The Dissertation Committee for Carolyn Leigh Mazzitelli Certifies that this is the approved version of the following dissertation:

Electrospray Ionization Tandem Mass Spectrometric Techniques for the Analysis of Drug/DNA Complexes

Committee:

Jennifer S. Brodbelt, Supervisor

Allen J. Bard

Brent L. Iverson

Sean M. Kerwin

Keith J. Stevenson

Electrospray Ionization Tandem Mass Spectrometric Techniques for the Analysis of Drug/DNA Complexes

by

Carolyn Leigh Mazzitelli, B.S.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

May 2007

For my parents, Rick and Diana.

Acknowledgements

This work would not be possible without the guidance and support of countless colleagues, friends and family members. I would first like to thank my adviser Dr. Jennifer Brodbelt for being an excellent teacher and mentor. Her wisdom, patience, and dependability has made my time as her student an enjoyable learning experience.

I am also indebted to my collaborators for providing me with their compounds and increasing my understanding and appreciation of medical and organic chemistry. I'd like to thank Dr. Sean Kerwin, Mireya Rodriquez, Dr. Brent Iverson and Yongjun Chu. I also wish to thank Dr. Allen Bard and Dr. Fu-Ren Fan for providing samples and helpful discussions for a study not included in this dissertation.

The Brodbelt group is also deserving of many thanks, not only for their helpful comments and questions throughout the years, but for providing many valuable diversions from research. From the post-it notes to the chopsticks, the owls to the unicorns, Madam Mam's to the ASMS hospitality suits, they made working fun.

I would also like to thank Mom, Dad, Andi, and Jennie for their invaluable moral support during my time in Austin. They deserve extra thanks for venturing from the mild Pacific Northwest into the searing heat of Texas for many enjoyable visits.

And finally, I thank Richard. Whether from near or far, his constant encouragement, sense of humor, and love has sustained me.

Electrospray Ionization Tandem Mass Spectrometric Techniques for the Analysis of Drug/DNA Complexes

Publication No._____

Carolyn Leigh Mazzitelli, Ph.D.

The University of Texas at Austin, 2007

Supervisor: Jennifer S. Brodbelt

Many anticancer and antibacterial therapies are based on the interaction of small molecule drugs with DNA. Increasing interest in the development of DNA-interactive agents has fostered the need for sensitive and versatile analytical techniques that are capable of characterizing the DNA/ligand interactions and are compatible with library-based screening methods. Electrospray ionization mass spectrometry (ESI-MS) has emerged as a useful technique for the analysis of non-covalent complexes formed between DNA and small molecules due to its low sample consumption and fast analysis time. The work presented in this dissertation is aimed at exploring, optimizing, and validating ESI-MS methods for characterizing DNA-ligand interactions.

ESI-MS is first used to assess the binding of threading bis-intercalators to duplexes containing different sequences to determine high affinity binding sites of the ligands. Preliminary DNAse footprinting experiments identified possible specific binding sites of the ligands and ESI-MS experiments revealed that the ligands bound to DNA duplexes containing the respective specific binding sequences. The metal-mediated

binding of benzoxazole ligands with different side chains to duplex DNA is also examined. Cu^{2+} and Ni^{2+} were found to promote the most dramatic increase in ligand binding, and ligands exhibiting the most dramatic metal-mediated or metal-enhanced binding were also determined to be the most cytotoxic.

The quadruplex DNA binding selectivity of perylene diimides is evaluated by screening the binding of the ligands to quadruplex, duplex and single strand DNA by ESI-MS. Three ligands, one containing basic side chains, one containing anionic side-chains, and one benzannulated compound were determined to be the most-quadruplex selective. The ESI-MS results correlated well with spectroscopic experiments. The relative gas-phase stabilities of different quadruplex DNA structures were investigated using molecular dynamics simulations and ESI-MS. The stabilities from the $E_{1/2}$ values generally paralleled the RMSD and relative free energies of the quadruplexes based on MD energy analysis. Finally an ESI-MS technique employing the $KMnO_4$ reaction with DNA to determine conformational changes to the duplex structure upon ligand binding is detailed. Thymines in most intercalator/duplex complexes are more susceptible to oxidation by $KMnO_4$ than those in duplex DNA. CAD and IRMPD experiments are used to identify the site of oxidation.

Table of Contents

Chapter 1: Introduction

The therapeutic basis of many anticancer and antibacterial drugs is the non-covalent binding of small molecules to nucleic acid substrates.[1,2] Increased interest in the development and characterization of DNA-interactive agents has fostered the need for sensitive and versatile analytical techniques that are capable of characterizing the binding of these compounds. ESI-MS has emerged as a useful tool for the analysis of non-covalent drug/DNA complexes due to its high sensitivity and fast analysis time making it adaptable to high throughput screening techniques. In gently transferring non-covalent complexes in solution to the gas-phase, ESI-MS allows for the characterization of many aspects of the drug/DNA binding interaction including complex binding stoichiometries, ligand sequence selectivities and relative binding affinities.

The development of ESI-MS as an analytical tool for the analysis of non-covalent drug/DNA complexes is the focus of this dissertation. A review of the different types of DNA-interactive compounds and the development of ESI-MS as a technique for the evaluation of drug/DNA complexes is presented in this introduction. This chapter concludes with an overview of the chapters to follow.

1.1 DNA-Interactive Compounds

The development of ligands that bind non-covalently to DNA structures represents an important area of medicinal chemistry and pharmaceutical research.[1,2] DNA-interactive compounds are used therapeutically to treat wide range of diseases including malaria,[3] hepatitis,[4] HIV,[5] as well as numerous bacterial infections[6,7] and cancers.[8-10] The primary mode of action of many of these drugs is to cause conformational changes or damage to the DNA targets, which ultimately interferes with DNA transcription or hinders replication, thereby interrupting cell growth and proliferation.[1]

While the clinical importance of these types of compounds has been demonstrated, the study of drug/DNA interactions remains an important area of medicinal chemistry research. The desire to more fully understand the mode of action of many current drugs, and to develop new compounds that are capable of selectively targeting cells directly affected by the disease, thereby reducing many negative side effects of the treatments, drive much of the current research.

1.1.1 Duplex DNA-Interactive Drugs

Until recently, the primary focus of drug-DNA research was on compounds that targeted duplex DNA. The two primary modes of binding by small molecules to duplex DNA are by intercalation and minor groove binding.[11] Intercalators comprise a class of compounds that are characterized as containing one or more planar, aromatic moieties that are inserted between adjacent base pairs upon DNA binding.[12] To accommodate the binding of a ligand in this manner, the duplex undergoes elongation due to a separation of base pairs, and unwinding around the site of intercalation.[12,13] Intercalators typically bind to G/C rich sequences, with a general preference for CpG sites due to the lower energy of unstacking at this sequence. Once bound, the complex is stabilized by van der Waals interactions between the chromophore of the intercalator and the adjacent nucleobases.[11] The intercalator-induced distortion of the DNA duplex causes an interference in DNA-protein interactions, ultimately disrupting transcription and replication.[12]

Figure 1.1 shows the structures of acridine, daunomycin, and actinomycin-D which are common intercalating compounds. Some intercalators, like daunomycin and actinomycin-D, contain side-chains that enhance the interaction by associating with the minor grove of duplex after intercalation of the planar chromophore.[14,15] Another way of increasing the binding constants of intercalators is by synthesizing polyintercalators, which contain multiple aromatic groups connected by chemical chains or scaffolds.[12]

When bound to DNA, the aromatic groups intercalate at multiple sites along the duplex, while the chemical linker interacts with the minor and/or major groove. The structure of echinomycin, one of the few commercially available bis-intercalators, is shown in Figure 1.1. Echinomycin is characterized by two quinoxaline chromophores connected by a central bicyclic peptide that interacts with the DNA minor groove upon binding.[16, 17] The potential advantages of polyintercalating ligands are increased binding constants, slower dissociation rates, increased sequence specificity, and the ability to target longer stretches of DNA.[18]

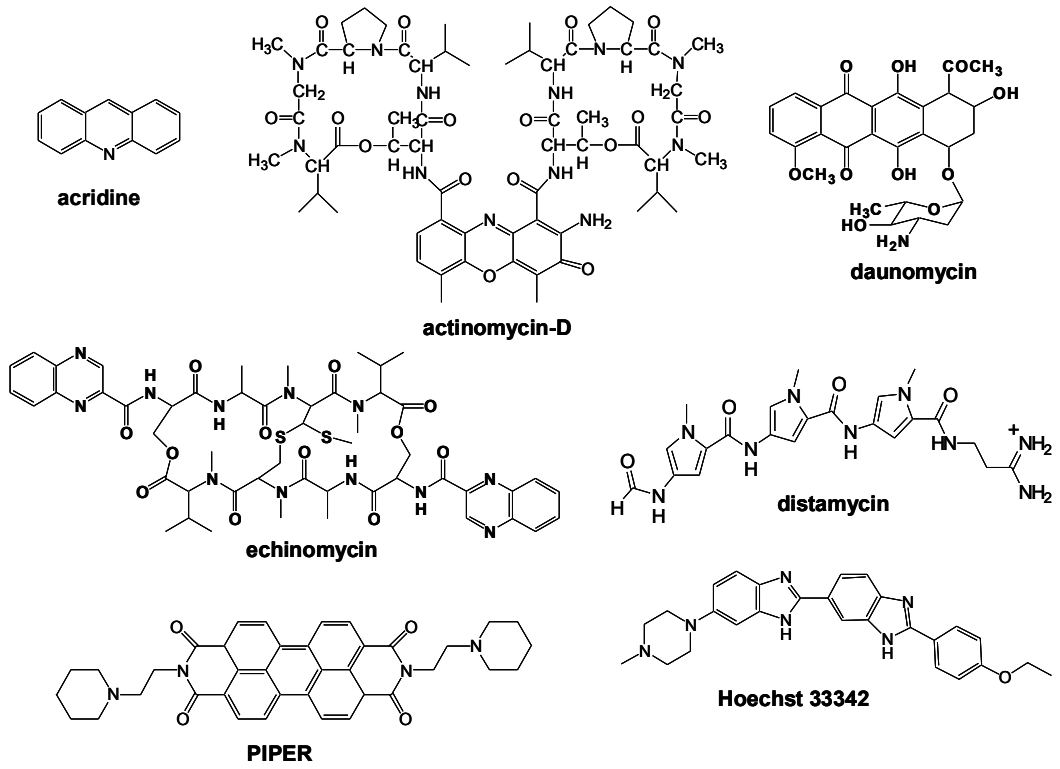


Figure 1.1: DNA-interactive drug molecules.

In addition to the intercalators, minor groove binders comprise the other main class of duplex-interactive drugs. As demonstrated by the structures of Hoechst 33342 and distamycin in Figure 1.1, minor groove binders are characterized by a crescent shape

which is similar to the curvature of the DNA minor groove.[13,19,20] DNA complexes containing minor groove binders are stabilized by van der Waals and hydrophophic interactions, as well as hydrogen bonds formed between the ligand and DNA. In contrast to the intercalators, minor groove binders typically exhibit selectivity for A/T rich sequences, likely due to increased hydrogen bonding between the ligands and adenines and thymines, as well as the smaller width of the minor groove of A/T tracts which allows for better binding interactions.[21] Upon complex formation, minor groove binders do not cause the significant conformational changes to the duplex that intercalators do, however they are effective at inhibiting protein binding usually by preventing the binding of transcription factors.[13]

1.1.2 Quadruplex-DNA Interactive Drugs

Traditionally, DNA-interactive compounds have targeted double stranded DNA, however there is increasing interest in the development of quadruplex DNA-interactive compounds.[22-24] Hydrogen bonds can form between a planar arrangement of four guanine nucleobases to create a G-tetrad as shown in Figure 1.2. Quadruplex DNA forms when two or more G-quartets are stacked on top of one another and stabilized by monovalent cations such as Na^+, K^+, and NH_4^+ coordinated to the central cavity of the quadruplex.[22-24] While a variety of sequences containing two or more tandem repeats of G-rich sequences are capable of forming quadruplex DNA, there is significant interest in examining the intramolecular quadruplex formation in 3' overhangs of the human telomeric sequence, 5-TTAGGG-3'.[25]

Telomeres are the DNA structures found on the terminal ends of chromosomes and are believed to play a key role in protecting DNA from degradation and fusion, ensuring the complete replication of DNA sequences and signaling cell senescence and death.[26] In normal, somatic cells, the telomere length is shortened with each DNA

replication and subsequent cell division process. As the length of the telomeres is shortened, regions of coding DNA are not replicated, eventually leading to cell death.[27] However, in most tumor cells, the length of the telomeres is maintained by telomerase, a reverse transcriptase enzyme.[28] The discovery that telomerase is over-expressed in 80-90% of all human tumor cells established a promising link between cell immortality characteristic of malignant tumors and telomerase activity.[29]

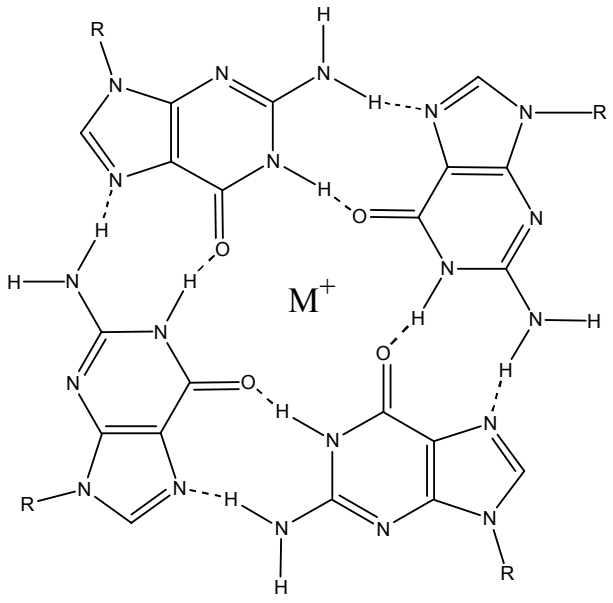


Figure 1.2: The structure of a G-tetrad. M^+ represents a monovalent cation and the dashed lines show hydrogen bonds.

Small molecules that bind to and stabilize the quadruplex structure in telomeric DNA have been found to inhibit telomerase activity and represent a promising area of anticancer drug research.[22-24,30,31] Ethidium bromide was one of the first compounds that was discovered to bind to quadruplex DNA.[32] Molecular models showed that this molecule binds by intercalation and undergoes pi-stacking interactions between the chromophore and the G-tetrad. Subsequent studies of related anthraquinone analogs found that adding flexible side chains to the molecule allowed for interaction with the

grooves of the quadruplex, and these analogs exhibited inhibition of telomerase activity.[33,34]

Another mode of binding by small molecules to quadruplexes is by stacking on the ends of the tetrads.[22,35] Several cationic porphyrin compounds, specifically *para*- and *ortho*- 5,10,15,20-tetra(*N*-methyl-4-pyridyl)porphine (H_2TMPy), bind to quadruplexes and have a suspected end-stack binding mode, with -overlap between the aromatic moieties and the G-tetrads.[36] The *ortho*-H_2TMPy compound was also found to create "sandwich" complexes which are formed when one porphyrin molecule stacks on the ends of two quadruplexes. The perylene diimides comprise another promising class of quadruplex-interactive compounds with a suspected end-stacking and/or sandwich binding mode.[22,23,30] As demonstrated by the structures of an early analog PIPER[37] (structure shown in Figure 1.1), these compounds are characterized by the perylene chromophore functionalized with the tetracarboxylic diimide groups with different side-chains. Early studies found that PIPER exhibited telomerase inhibition,[37] and more recent studies have focused on developing analogs that selectively bind to quadruplex DNA over duplex DNA in an attempt to reduce the cytotoxic side-effects of the compounds.[38-41]

1.2 Analysis of Drug/DNA complexes by ESI-MS

The therapeutic importance and structural diversity amongst DNA-interactive compound underscores the need for new analytical technologies capable of characterizing non-covalent drug/DNA complexes. ESI-MS has been established as a useful technique for this area of application due to its versatility, sensitivity and fast analysis time.[42-44] The development of electrospray ionization greatly expanded the scope of mass spectrometry by allowing for the analysis of large biomolecules[45] and the subsequent discovery that non-covalent complexes involving biomolecules, such as proteins and nucleic acids, were also preserved.[46-49] ESI is well-suited for the analysis of these types of complexes because

it creates highly charged ions, allowing analytes with large molecular weights to be analyzed by robust and inexpensive mass analyzers with limited mass ranges, such as the quadrupole ion trap which has a typical *m/z* range of 2000. ESI is also a gentle ionization technique allowing non-covalent complexes present in solution to be transferred to the gas-phase while maintaining many of the native binding interactions.

Upon the discovery that ions of DNA duplexes[50] and later, non-covalent complexes containing duplexes with minor groove binders,[51] could be observed by ESI-MS, interest in further examining the technique was stimulated. ESI-MS offers several advantages as a screening tool for the analysis of non-covalent drug/DNA complexes compared to traditional bioanalytical techniques such as surface plasmon resonance, fluorescence, circular dichroism, UV-Vis spectroscopy, NMR and crystallography. Compared to high resolution structural techniques such as NMR and crystallography, mass spectrometry is more sensitive, requiring micrograms of sample, compared to the milligrams typically required for those techniques. Another advantage is that ESI-MS does not require that ligands or DNA be labeled or tethered to a surface, which could affect ligand binding. This also reduces the time and complexity of an experiment. Finally, ESI-MS is well-suited for high throughput screening techniques due to its rapid analysis time and ability to be automated.[44]

1.2.1 ESI-MS Analysis of Duplex DNA/Drug Complexes

Early studies of drug/DNA complexes by ESI-MS focused on examining the binding of commercially available drugs to duplex DNA to determine if the binding behavior observed by mass spectrometry mirrored known solution activity. In seminal studies by Gale et al., complexes formed between minor groove binders including distamycin, pentamidine, and Hoechst 33258 and a self-complementary 12-mer duplex were examined.[51,52] The binding stoichiometries of the DNA complexes containing the

minor groove binders observed in the mass spectra were consistent with previous NMR reports. For example, ESI-MS experiments found that at low distamycin/DNA molar ratios, 1:1 complexes were present, while 2:1 complexes emerged at higher molar ratios. This study also established experimental conditions required for the analysis of drug/DNA complexes by ESI-MS, including using low ESI voltages and heated capillary temperatures, as well as an ammonium acetate buffer instead of conventional alkali metal buffers. Ammonium acetate is very amenable to ESI-MS because of the greater lability of the ammonium ion over Na^+ and K^+, thus reducing the extent of counterion adduction to the DNA ions.

ESI-MS studies of the binding of intercalators daunomycin and nogalamycin to duplexes of different sizes and sequence composition were also undertaken,[53,54] with results suggesting that mass spectrometry can be used to assess sequence selectivity of drugs, as the intercalators formed complexes with greater relative abundance with duplexes with the largest GC-base pair composition. The binding of commercially available intercalators (ethidium bromide, amsacrine, and ascididemin) and minor groove binders (Hoechst 33258, netropsin, distamycin, berenil, and DAPI) was further examined by Gabelica at al.[55] The DNA complexes containing the minor groove binders were found to have more well-defined 1:1 and 2:1 binding stoichiometries, while the complexes containing the intercalators had more variable binding stoichiometries that were suggestive of less specific binding compared to the groove binders. The binding of a series of commercially available minor groove binders and intercalators were also examined by Wan. et al.[56] Differences in sequences selectivities of intercalators and minor groove binders were compared, with the intercalators exhibiting high affinity for GC-rich sequences, while minor groove binders displayed a preference for AT-rich sequences. These previous ESI-MS studies of non-covalent complexes of duplex DNA

with commercially available compounds indicated a good correlation between the mass spectral results and known solution behavior.

1.2.2 ESI-MS Analysis of Quadruplex DNA/Drug Complexes

The use of mass spectrometry has recently been extended to examine non-covalent complexes containing quadruplex DNA after initial studies confirmed that quadruplexes stabilized by sodium[57] and ammonium[58] cations can be observed using ESI-MS. Complexes formed between quadruplexes and a variety of ligands have been examined including the perylene diimides,[59,60] ethidium bromide derivatives,[61] cyclo[n]pyrolles,[62] pyridoacridines,[63] cryptolepine and neocryptolepine,[64] telomestatin,[65] and the bis-intercalator ditercalinium.[66] In addition to confirming that ligand/quadruplex structures can be examined using ESI-MS, the binding selectivity of the ligands for quadruplexes over other DNA structures has been evaluated due to the desire to develop quadruplex-selective compounds. The results of these studies have shown good agreement between the ligand binding affinities for different DNA structures determined by ESI-MS with results from spectroscopic[60] and competition dialysis experiments.[61,63,64]

1.3 Overview of Chapters

The objective of this dissertation is to (1) extend to use of ESI-MS for the analysis of new types drug/DNA complexes, and (2) develop new techniques for the evaluation of drug binding sites in DNA complexes. Chapter 2 is a brief description of the primary experimental methods used throughout this dissertation. Chapters 3 and 4 explore the use of ESI-MS to examine complexes containing novel drugs with duplex DNA. The binding of a series of threading bis-intercalators to duplex DNA is evaluated in Chapter 3. Binding affinities of the ligands for duplexes with different sequences were evaluated to assess possible specific binding, while the CAD spectra of the bis-intercalator/duplex

complexes were found to be unique from complexes containing commercially available monointercalators.

Chapter 4 focuses on metal-mediated binding to duplex DNA. The binding of a series of benzoxazoles analogs with different amide- and ester-linked side chains to duplex DNA in the absence and presence of biologically relevant divalent metal cations is examined. The ligands with the shorter side chains only formed DNA complexes in the presence of metal cations, most notably with Cu^{2+}. The ligand exhibiting the most dramatic metal-enhanced DNA binding also demonstrated the greatest cytotoxic activity.

ESI-MS of quadruplex DNA is the focus of chapters 5 and 6. In chapter 5, the binding of a series of perylene diimides to quadruplex, duplex, and single strand DNA was examined to assess the quadruplex selectivity of each ligand. There was good correlation between solution-based spectroscopic binding studies and the ESI-MS analysis of G-quadruplex selectivity. The collisional activated dissociation (CAD) spectra of the perylene diimide/quadruplex complexes were consistent with previous studies of ligands with an end-stack binding mode. The gas-phase stabilities of different G-quadruplexes were examined in Chapter 6 using energy-variable collisional activated dissociation and molecular dynamics (MD) simulations. The MD simulations found that the quadruplex structures are maintained in the gas-phase, while the stabilities of the quadruplexes determined by energy-variable experiments generally paralleled the RMSD and relative free energies of the quadruplexes based on the MD energy analysis.

In the final chapter of this dissertation, Chapter 7, a technique that utilizes ESI-MS with the $KMnO_4$ oxidation of thymines to probe conformational changes to duplex DNA upon ligand binding is presented. In intercalator/DNA complexes, thymines around the ligand binding site are most susceptible to oxidation, allowing the binding sites to be

assessed. The site of oxidation is determined using CAD or infrared multiphoton dissociation (IRMPD).

1.4 REFERENCES

(1) Goodman, L. S.; Hardman, J. G.; Limbird, L. E.; Gilman, A. G. *Goodman & Gilman's the pharmacological basis of therapeutics*, 10th ed.; McGraw-Hill: New York, 2001.

(2) Propst, C. L.; Perun, T. J. *Nucleic acid targeted drug design*; M. Dekker: New York, 1992.

(3) Ginsburg, H.; Nissani, E.; Krugliak, M.; Williamson, D. H. *Mol. Biochem. Parasit.* **1993**, *58*, 7-15.

(4) Civitico, G.; Wang, Y. Y.; Luscombe, C.; Bishop, N.; Tachedjian, G.; Gust, I.; Locarnini, S. *J. Med. Virol.* **1990**, *31*, 90-97.

(5) Asante-Appiah, E.; Skalka, A. M. *Antivir. Res.* **1997**, *36*, 139-156.

(6) Cremieux, A.; Chevalier, J.; Sharples, D.; Berny, H.; Galy, A. M.; Brouant, P.; Galy, J. P.; Barbe, J. *Res. Microbiol.* **1995**, *146*, 73-83.

(7) Williams, R. A. D.; Lambert, P. A.; Singleton, P. *Antimicrobial Drug Action*; BIOS Scientific: Oxford, 1996.

(8) Galpin, A. J.; Evans, W. E. *Clin. Chem.* **1993**, *39*, 2419-2430.

(9) Allwood, M. *The Cytotoxics Handbook*; Radcliffe Medical Press: Oxford, 1993.

(10) Witt, K. L.; Bishop, J. B. *Mut. Res.-Fund. Mol. M.* **1996**, *355*, 209-234.

(11) Haq, I.; Ladbury, J. *J. Mol. Recognit.* **2000**, *13*, 188-197.

(12) Brana, M. F.; Cacho, M.; Gradillas, A.; de Pascual-Teresa, B.; Ramos, A. *Curr.Pharm. Des.* **2001**, *7*, 1745-1780.

(13) Bischoff, G.; Hoffmann, S. *Curr. Med. Chem.* **2002**, *9*, 321-348.

(14) Sobell, H. M. *Proc. Natl. Acad. Sci. U. S. A.* **1985**, *82*, 5328-5331.

(15) Moore, M. H.; Hunter, W. N.; Destaintot, B. L.; Kennard, O. *J. Mol. Biol.* **1989**, *206*, 693-705.

(16) Wakelin, L. P. G.; Waring, M. J. *Biochem. J.* **1976**, *157*, 721-740.

(17) Waring, M. J.; Wakelin, L. P. G. *Nature* **1974**, *252*, 653-657.

(18) Lee, J.; Guelev, V.; Sorey, S.; Hoffman, D. W.; Iverson, B. L. *J. Am. Chem. Soc.* **2004**, *126*, 14036-14042.

(19) Rao, K. E.; Dasgupta, D.; Sasisekharan, V. *Biochemistry* **1988**, *27*, 3018-3024.

(20) Timsit, Y.; Moras, D. *Embo J.* **1994**, *13*, 2737-2746.

(21) Geierstanger, B. H.; Wemmer, D. E. *Annu. Rev. Biophys. Biomol. Struct.* **1995**, *24*, 463-493.

(22) Kerwin, S. M. *Curr. Pharm. Des.* **2000**, *6*, 441-471.

(23) Han, H. Y.; Hurley, L. H. *Trends Pharmacol.l Sci.* **2000**, *21*, 136-142.

(24) Hurley, L. H.; Wheelhouse, R. T.; Sun, D.; Kerwin, S. M.; Salazar, M.; Fedoroff, O. Y.; Han, F. X.; Han, H. Y.; Izbicka, E.; Von Hoff, D. D. *Pharmacol. Ther.* **2000**, *85*, 141-158.

(25) Moyzis, R. K.; Buckingham, J. M.; Cram, L. S.; Dani, M.; Deaven, L. L.; Jones, M. D.; Meyne, J.; Ratliff, R. L.; Wu, J. R. *Proc. Natl. Acad. Sci. U. S. A.* **1988**, *85*, 6622-6626.

(26) Wellinger, R. J.; Sen, D. *Eur. J. Cancer* **1997**, *33*, 735-749.

(27) Harley, C. B.; Futcher, A. B.; Greider, C. W. *Nature* **1990**, *345*, 458-460.

(28) Nakamura, T. M.; Morin, G. B.; Chapman, K. B.; Weinrich, S. L.; Andrews, W. H.; Lingner, J.; Harley, C. B.; Cech, T. R. *Science* **1997**, *277*, 955-959.

(29) Counter, C. M.; Hahn, W. C.; Wei, W. Y.; Caddle, S. D.; Beijersbergen, R. L.; Lansdorp, P. M.; Sedivy, J. M.; Weinberg, R. A. *Proc. Natl. Acad. Sci. U. S. A.***1998**, *95*, 14723-14728.

(30) Rezler, E. M.; Bearss, D. J.; Hurley, L. H. *Curr. Opin. Pharmacol.* **2002**, *2*, 415-423.

(31) Rezler, E. M.; Bearss, D. J.; Hurley, L. H. *Annu. Rev. Pharmacol. Toxicol.* **2003**, *43*, 359-379.

(32) Guo, Q.; Lu, M.; Marky, L. A.; Kallenbach, N. R. *Biochemistry* **1992**, *31*, 2451-2455.

(33) Kelland, L. R.; Gowan, S. M.; Perry, P. J.; Neidle, S. *Br. J. Cancer* **1998**, *78*, 18-18.

(34) Perry, P. J.; Gowan, S. M.; Reszka, A. P.; Polucci, P.; Jenkins, T. C.; Kelland, L. R.; Neidle, S. *J. Med. Chem.* **1998**, *41*, 3253-3260.

(35) Jenkins, T. C. *Curr. Med. Chem.***2000**, *7*, 99-115.

(36) Izbicka, E.; Wheelhouse, R. T.; Raymond, E.; Davidson, K. K.; Lawrence, R. A.; Sun, D. Y.; Windle, B. E.; Hurley, L. H.; Von Hoff, D. D. *Cancer Res.* **1999**, *59*, 639-644.

(37) Fedoroff, O. Y.; Salazar, M.; Han, H. Y.; Chemeris, V. V.; Kerwin, S. M.; Hurley, L. H. *Biochemistry* **1998**, *37*, 12367-12374.

(38) Kern, J. T.; Kerwin, S. M. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 3395-3398.

(39) Kerwin, S. M.; Chen, G.; Kern, J. T.; Thomas, P. W. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 447-450.

(40) Kerwin, S. M.; Thomas, P. W.; Kern, J. T. *Abstracts of Papers of the American Chemical Society* **2001**, *222*, U663-U663.

(41) Kern, J. T.; Thomas, P. W.; Kerwin, S. M. *Biochemistry* **2002**, *41*, 12568-12568.

(42) Beck, J. L.; Colgrave, M. L.; Ralph, S. F.; Sheil, M. M. *Mass. Spectrom. Rev.* **2001**, *20*, 61-87.

(43) Hofstadler, S. A.; Griffey, R. H. *Chem. Rev.* **2001**, *101*, 377-390.

(44) Hofstadler, S. A.; Sannes-Lowery, K. A. *Nat. Rev. Drug Discov.* **2006**, *5*, 585-595.

(45) Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. *Science* **1989**, *246*, 64-71.

(46) Smith, R. D.; Lightwahl, K. J.; Winger, B. E.; Loo, J. A. *Org. Mass Spectrom.* **1992**, *27*, 811-821.

(47) Bruce, J. E.; Anderson, G. A.; Chen, R. D.; Cheng, X. H.; Gale, D. C.; Hofstadler, S. A.; Schwartz, B. L.; Smith, R. D. *Rapid Commun. Mass Spectrom.* **1995**, *9*, 644-650.

(48) Loo, J. A. *Bioconjugate Chem.* **1995**, *6*, 644-665.

(49) Loo, J. A. *Mass Spectrom. Rev.* **1997**, *16*, 1-23.

(50) Lightwahl, K. J.; Springer, D. L.; Winger, B. E.; Edmonds, C. G.; Camp, D. G.; Thrall, B. D.; Smith, R. D. *J. Am. Chem. Soc.* **1993**, *115*, 803-804.

(51) Gale, D. C.; Smith, R. D. *J. Am. Soc. Mass. Spectrom.* **1995**, *6*, 1154-1164.

(52) Gale, D. C.; Goodlett, D. R.; Lightwahl, K. J.; Smith, R. D. *J. Am. Chem. Soc.* **1994**, *116*, 6027-6028.

(53) Kapur, A.; Beck, J. L.; Sheil, M. M. *Rapid Commun. Mass Spectrom.* **1999**, *13*, 2489-2497.

(54) Colgrave, M. L.; Beck, J. L.; Sheil, M. M.; Searle, M. S. *Chem. Commun.* **2002**, 556-557.

(55) Gabelica, V.; De Pauw, E.; Rosu, F. *J. Mass Spectrom.* **1999**, *34*, 1328-1337.

(56) Wan, K. X.; Shibue, T.; Gross, M. L. *J. Am. Chem. Soc.* **2000**, *122*, 300-307.

(57) Goodlett, D. R.; Camp, D. G.; Hardin, C. C.; Corregan, M.; Smith, R. D. *Biol. Mass Spectrom.* **1993**, *22*, 181-183.

(58) Rosu, F.; Gabelica, V.; Houssier, C.; Colson, P.; De Pauw, E. *Rapid Commun. Mass Spectrom.* **2002**, *16*, 1729-1736.

(59) David, W. M.; Brodbelt, J.; Kerwin, S. M.; Thomas, P. W. *Anal. Chem.* **2002**, *74*, 2029-2033.

(60) Mazzitelli, C. L.; Brodbelt, J. S.; Kern, J. T.; Rodriguez, M.; Kerwin, S. M. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 593-604.

(61) Rosu, F.; De Pauw, E.; Guittat, L.; Alberti, P.; Lacroix, L.; Mailliet, P.; Riou, J. F.; Mergny, J. L. *Biochemistry* **2003**, *42*, 10361-10371.

(62) Baker, E. S.; Lee, J. T.; Sessler, J. L.; Bowers, M. T. *J. Am. Chem. Soc.* **2006**, *128*, 2641-2648.

(63) Guittat, L.; De Cian, A.; Rosu, F.; Gabelica, V.; De Pauw, E.; Delfourne, E.; Mergny, J. L. *Biochim. Biophys. Acta* **2005**, *1724*, 375-384.

(64) Guittat, L.; Alberti, P.; Rosu, F.; Van Miert, S.; Thetiot, E.; Pieters, L.; Gabelica, V.; De Pauw, E.; Ottaviani, A.; Riou, J. F.; Mergny, J. L. *Biochimie* **2003**, *85*, 535-547.

(65) Rosu, F.; Gabelica, V.; Shin-ya, K.; De Pauw, E. *Chem. Commun.* **2003**, 2702-2703.

(66) Carrasco, C.; Rosu, F.; Gabelica, V.; Houssier, C.; De Pauw, E.; Garbay-Jaureguiberry, C.; Roques, B.; Wilson, W. D.; Chaires, J. B.; Waring, M. J.; Bailly, C. *ChemBiochem* **2002**, *3*, 1235-1241.

Chapter 2: Experimental Methods

The experiments described in this dissertation involve the analysis of DNA by electrospray ionization mass spectrometry (ESI-MS). A review of DNA nomenclature as well as brief descriptions of the solution and instrument conditions used in the ESI-MS analysis of DNA and drug/DNA complexes are contained in this chapter.

2.1 DNA Nomenclature and Annealing Conditions

DNA is a biopolymer that is composed of repeating units called nucleotides. Each nucleotide contains one of four nucleobases, guanine (G), cytosine (C), thymine (T), and adenine (A), linked to pentose ring and phosphate group. Multiple nucleotides are linked together to form an oligonucleotide.[1] The ends of oligonucleotides are asymmetric and are labeled as the 5' and 3' end as demonstrated by the structure of 5'-GTAC-3' shown in Figure 2.1 In in vivo environments, DNA typically exists as a double helix or duplex, which forms when there is hydrogen bonding between nucleobases on two different single strands that are arranged in an anti-parallel fashion, meaning the 5' end of one strand is oriented to the 3' end of the second strand.[2] Hydrogen bonds are formed between complementary base pairs, with cytosines binding to guanines via three hydrogen bonds, and adenines binding to thymines by two hydrogen bonds as shown in Figure 2.2.[1]

For the ESI-MS studies described herein, single strand oligonucleotides were ordered from Integrated DNA Technologies (Coralville, IA) or TriLink Biotechnologies Inc. (San Diego, CA), custom synthesized on the 1.0 μmole scale with reverse-phase HPLC purification. Stock solutions of each oligonucleotide were prepared at 2 mM concentration in 250 mM ammonium acetate. The concentration of the DNA strands were verified spectroscopically using Beer's law and extinction coefficients provided by the DNA manufacturer. To prepare duplex DNA, solutions containing equimolar

concentrations of complementary strands were made in 250 mM ammonium acetate so that the final concentration of each oligonucleotide was 1 mM. The solution was then heated to 90 °C for 10 min., and then cooled to room temperature during a period of 7 hours. Quadruplex DNA was annealed using a similar procedure. Solutions containing quadruplex forming oligonucleotides were prepared at 1 mM concentration in 150 mM ammonium acetate. The quadruplex solutions were heated to 90 °C for 10 min., followed by cooling to room temperature during a period of 7 hours.

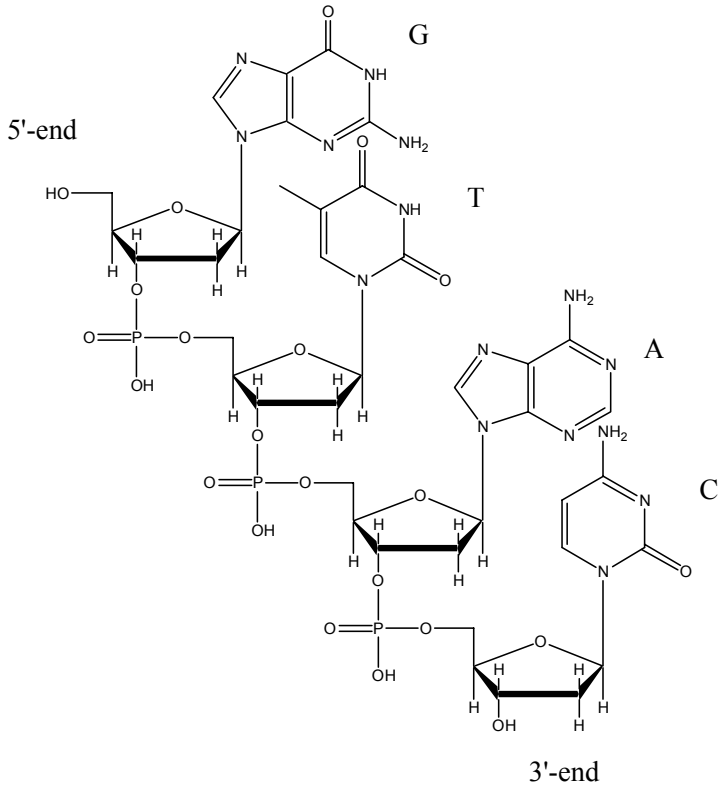


Figure 2.1: Structure of the 5’-GTAC-3’ oligonucleotide.

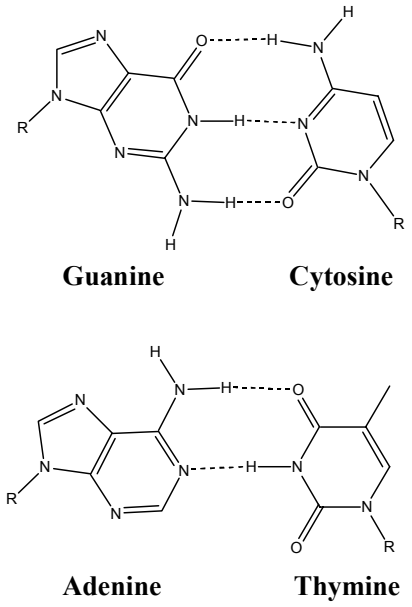


Figure 2.2: Hydrogen bonding between DNA nucleobases.

2.2 SOLUTION AND INSTRUMENT CONDITIONS FOR ESI-MS ANALYSIS OF DNA

In these studies, analytical solutions were typically contained the single strand, duplex or quadruplex DNA at a concentration of 10 μM in a buffer containing 50 mM ammonium acetate with 25% (v/v) methanol. Methanol was included in the analytical solutions to increase the volatility of the aqueous solutions. The pH of the solutions was nominally 6.8. The samples were directly infused into a Thermo Finnigan LCQ Duo ion trap mass spectrometer (San Jose, CA) at flow rate of 3 μL/min using a Harvard Apparatus syringe pump (Holliston, MA). The Xcalibur software package (Thermo Finnigan, San Jose, CA) was used to control the instrument. The mass range of this instrument is 2000 mass units and a typical lower limit of concentration of DNA oligonucleotides using the LCQ is 2.5 μM.

Ions were generated in the negative ion mode using an ESI voltage of 3.5 kV, and a heated capillary temperature ranging from 90 to 120 °C. Due to the large aqueous

composition of the analytical solutions, nitrogen sheath and auxiliary gas flows of 40 and 20 arbitrary units, respectively, were used to assist the desolvation. During operation of the instrument, the base pressure of the ion trap was 1 x 10^{-5} Torr with helium. Ion accumulation times of 50 to 100 ms were used and spectra were acquired by summing 300 scans.

During collisional activated dissociation (CAD) experiments, the desired precursor ion was isolated in the trap using resonance ejection. The collision energy applied to the trap (typically reported as a percentage of 5 $V_{0\text{-}p}$) was increased, causing the trapped ions to undergo energetic collisions with background bath gas molecules (typically He). The default activation time of 30 ms and q_z value of 0.25 was used. Typically, the CAD energy was increased until the abundance of the precursor ion was reduced to nominally 10% of its original abundance.

2.3 REFERENCES

(1) Watson, J. D.; Crick, F. H. C. *Nature* **1953**, *171*, 964-967.

(2) Arnott, S.; Chandrasekaran, R.; Selsing, R. *Structure and Conformation of Nucleic Acids and Protein-Nucleic Acid Interactions*; Univeristy Park Press: Baltimore, 1975.

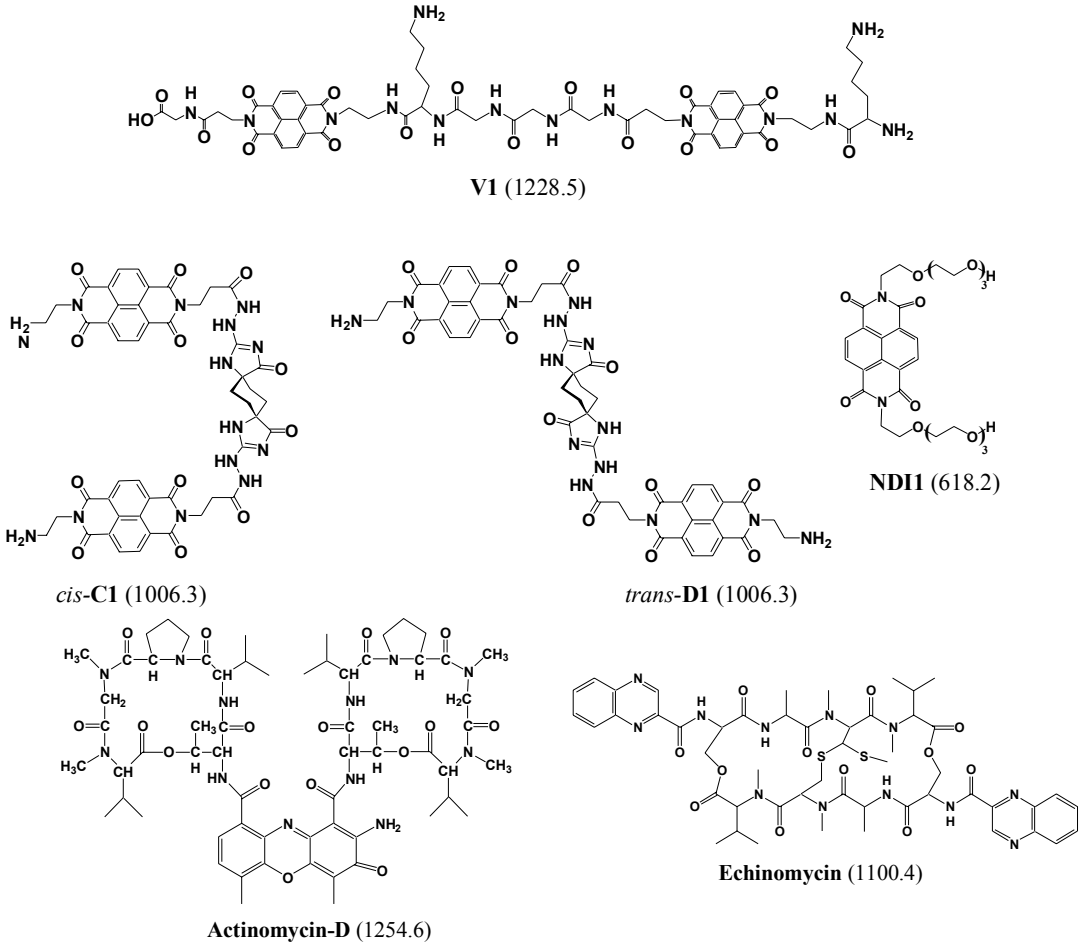
Chapter 3: Screening of Threading Bis-Intercalators Binding to Duplex DNA by Electrospray Ionization Tandem Mass Spectrometry

3.1 INTRODUCTION

Many anticancer, antitumor, and antibacterial therapies are based on the interaction of small molecules with DNA,[1, 2] fostering the need for sensitive and versatile analytical techniques that are both capable of characterizing the ligand/DNA interactions and compatible with library-based screening methods. Electrospray ionization mass spectrometry (ESI-MS) shows promise as a screening tool for the evaluation drug/DNA complexes due to its low sample consumption and rapid analysis time.[3,4] During the electrospray process, non-covalent complexes are transferred to the gas-phase with minimal internal energy, allowing many of the binding interactions to be maintained. The preservation of these non-covalent complexes allows information about binding stoichiometry and selectivity to be elucidated from the mass spectra, while tandem mass spectrometry techniques, such as collisional activation dissociation (CAD), can be used to examine the binding mode.

One important class of DNA-interactive drugs are ligands that bind via intercalation of one or more aromatic groups between base pairs of duplex DNA.[1, 2, 5] There have been numerous ESI-MS studies that have examined the interaction between duplex DNA and well-studied, commercially available monointercalators such as the anthracyclines,[6-11] porphyrins,[12,13] ruthenium compounds,[11-15] ethidium bromide,[16-18] actinomycin-D,[12,13] and aureolic acids.[19] Characteristics such as ligand binding stoichiometry,[7,8,11-14,16] sequence selectivity,[7,11,13,14,16,19] binding mode,[12,13,15] and complex stability[9,12,16] have been examined with promising results correlating the binding trends

observed in the mass spectra to known solution behavior. While the binding of many commercial monointercalators has been well-studied by ESI-MS, there has been only one study[20] that has focused on a bisintercalator, ditercalinium. Based on ESI-MS measurements of the complexation of ditercalinium to a series of DNA sequences, it was found that ditercalinum bound better to quadruplex structures than to duplexes.[20]



Scheme 3.1: Structures of intercalator ligands. Molecular weights of compounds in Da are given in parenthesis.

The development of novel polyintercalating ligands is of interest because of the potential improvements in antitumor activity and sequence specificity. A novel class of

DNA polyintercalators that shows great promise for binding to long stretches of DNA with sequence specificity and high affinity contain 1,4,5,8-tetracarboxylic naphthalene diimide (NDI) units connected in a head-to-tail arrangement by flexible scaffolds[21-24] (see Scheme 3.1). These compounds are known as threading polyintercalators because, upon intercalation, one of the functional groups attached to the diimide nitrogen resides in the DNA major groove, while the other is in the minor groove.[21] The potential advantages of developing a compound that binds to duplex DNA by threading polyintercalation include enhanced sequence specificity due to ligand interactions with the major and minor groove, the ability to bind to longer DNA sequences with a relatively low molecular weight compound, disruption of protein-ligand interactions that occur in both DNA grooves, and lower binding off-rates,[22]

A series of bis-intercalators, compounds that contain two NDI units, have been synthesized as precursors to longer compounds with more than two intercalation groups (Scheme 3.1). The primary objective of this study is to evaluate the binding behavior of *trans*-**D1** and *cis*-**C1** which are new ligands containing a rigid spiro-tricyclic scaffold in the *trans*- and *cis*- orientations, respectively. While few results have been reported for the newer ligands trans-D1 and *cis*-**C1**,[25] the binding behavior of **V1**, the ligand containing the peptide scaffold, has been examined in extensive footprinting and NMR-based studies.[22,23] In the present study, the binding behavior of the well-characterized compound **V1** will be examined to establish that the binding behavior observed by ESI-MS can be correlated to the results of traditional solution-based experiments. After developing a framework with **V1**, the binding of the new compounds, *trans*-**D1** and *cis*-**C1**, will be reported with an emphasis on comparing differences in binding behavior of the compounds that result from the trans- versus cis-orientations of the scaffold. Here we examine binding stoichiometries, sequence selectivities, and concentration dependent

binding of the bis-intercalators, evaluate the CAD fragmentation patterns of the observed DNA/drug complexes, and compare the results to ones obtained by conventional footprinting techniques. The binding of the bis-intercalator, echinomycin (Scheme 3.1) was also assessed by ESI-MS to serve as a comparison to the results of **V1**, *cis*-**C1**, and *trans*-**D1**. We also will compare the results of our study to other ESI-MS based studies involving monointercalators.

Another objective of this study is to demonstrate that mass spectrometry can be used as a screening tool for bis-intercalator ligands. To accomplish this aim, we demonstrate that the ESI-MS results mirror those established by traditional techniques such as NMR and DNAse I footprinting experiments to lend legitimacy to ESI-MS as an analysis tool for future studies involving novel polyintercalating compounds. It is not anticipated that ESI-MS will replace traditional techniques such as NMR and DNAse I footprinting, but instead that it be used as an initial screening tool. As drug discovery efforts shift to combinatorial synthesis of compound libraries, ESI-MS is suited to narrow down the pool of promising ligands that would then be examined in more detail using traditional methods.

3.2 EXPERIMENTAL

3.2.1 Chemicals

Single strand oligodeoxynucleotides (ODNs), custom synthesized as ammonium salts on the 1.0 μmole or 250 nmole scale with purification by HPLC, were obtained from Integrated DNA Technologies (Coralville, IA) and used without further purification. Stock solutions of each ODN were prepared at 2 mM concentration in deionized water. A portion of the stock solutions were diluted to 1 mM and set aside for experiments involving single strand ODNs. Duplex DNA was annealed by preparing solutions containing two complementary single strand ODNs, each at 0.7 mM concentration in 250

mM ammonium acetate. The annealing solutions were heated to 90 °C and then slowly cooled over a period of 4 hours. Table 3.1 shows the sequences used in this study. The synthesis of ligands **V1,**[26] *trans*-**D1,**[25] *cis*-**C1**,[25] and **NDI1**[27] have been previously reported. Concentrations were determined spectroscopically using Beer's law. The extinction coefficients for the DNA strands were provided by the manufacturer and those of the ligands are 26 300 M^{-1} cm^{-1} (385 nm) for *cis*-**C1**, 36,000 M^{-1} cm^{-1} (383 nm) for *trans*-**D1**, and 26 300 M^{-1} cm^{-1} (385 nm) for **V1**.[25]

3.2.2 DNAse I Footprinting.

Plasmid pBR322 (New England BioLabs, Ipswich, MA) was digested with NheI, dephosphorylated with CIAP, 5'-32P-end labeled with [-32P]-ATP and T4 kinase, digested with EcoRI (all enzymes were purchased from New England Biolabs) and purified by native polyacrylamide gel electrophoresis (PAGE), following standard protocols.[28] The 92 bp synthetic fragment (PAGE grade) was purchased from Midland Certified (Midland, TX) and labeled (32P) similarly. The DNAse I (Amersham Biosciences, Piscataway, NJ) footprinting was carried out according to the procedure described previously.[29] The DNA fragments were separated on an 8% (231 bp) or 12% (92 bp) denaturing polyacrylamide gel. The gels were exposed on a phosphor screen and analyzed with Quantity One 4.5 software from Bio-Rad (Hercules, CA).

Table 3.1: DNA sequences used in this study.

Nam	Sequence	Molecular Weight
ds1	d(GGGCGGTACCGCGG/CCGCGGTACCGCCC)	8531.5
ds2	d(GCGGGGATGGGGCG/CGCCCCATCCCCGC)	8531.6
ds3	d(GCGGGAATTGGGCG/CGCCCAATTCCCGC)	8529.6
ds4	d(GCGGAAATTTGGCG/CGCCAAATTTCCGC)	8527.7
ds5	d(GGGACAGTGAGGGG/CCCCTCACTGTCCC)	8529.6
ds6	d(GGTTGGGCCCAAGG/CCTTGGGCCCAACC)	8529.6
ds7	d(GGGGTCGCCGGGGG/CCCCCGGCGACCCC)	8532.6

3.2.3 Mass Spectrometry

Stock solutions of **V1**, *trans*-**D1** and *cis*-**C1** were prepared in deionized water at 1 mM. Analytical solutions containing duplex or single strand DNA and one ligand were prepared at equimolar 10 μM concentration (unless noted otherwise) in 50 mM ammonium acetate with 25% methanol to enhance the volatility of the solution. After allowing the solutions to equilibrate for 30 min., they were directly infused into a ThermoFinnigan LCQ Duo mass spectrometer (San Jose, CA) using a Harvard syringe pump (Holliston, MA) at 3 μL/min. Ions were generated in the negative ion mode with an electrospray voltage of 3.5 kV. The temperature of the heated capillary was set at 90 to 110 °C and nitrogen sheath and auxiliary gas flows of 10 and 40 arbitrary units, respectively, were used to aid in desolvation. The base pressure in the ion trap region was nominally ~1 x 10^{-5} torr. Instrument conditions were optimized for each complex using the automatic tuning function of the Xcalibur software package (Finnigan, San Jose, CA). Spectra were acquired by summing 300 scans with an ion accumulation time of 100 ms.

Tandem mass spectrometry experiments were performed using collisional activated dissociation (CAD). The desired precursor ion was isolated in the trap using resonance ejection, followed by fragmentation induced by increasing the resonance voltage applied to the trap. An activation time of 30 ms was used for all experiments. The CAD energy was increased until the abundance of the precursor ion was reduced to ~10% relative abundance. These experiments required CAD energies of ~12 – 14%.

3.3 RESULTS AND DISCUSSION

3.3.1 Complexes with V1

The DNA binding of ligand **V1** has been previously examined by NMR and DNAse I footprinting techniques that identified the specific binding sites of this compound.[23] **V1** was found to have a binding preference for d(GGTACC)$_2$ sequences with NMR results confirming that the –Gly$_3$-Lys- peptide scaffold was located in the major groove.[23] The **V1**-d(GGTACC)$_2$ complex was formed with a dominant 1:1 binding stoichiometry.

The binding of **V1** to a DNA duplex containing the preferred binding sequence was evaluated by ESI-MS to confirm that results revealed in the mass spectra can be correlated to the solution binding behavior. The 14-mer d(GGGCGGTACCGCGG/CCGCGGTACCGCCC) (ds1) was used for this study because it contains the specific binding sequence of **V1** (GGTACC). All of the duplex DNA sequences selected for this study were non-self-complementary to allow the duplex and single strand ions to be unambiguously distinguished in the mass spectra. Duplexes with 14 base-pairs were selected for this study because previous ion mobility/molecular dynamics studies have reported that DNA duplexes greater than 12 base pairs maintain the helical conformations in the gas-phase better than smaller duplexes.[30,31] Even larger duplexes were not chosen to ensure that there was only one high affinity binding site per

duplex, allowing different binding sequences to be assessed individually. Each duplex was designed to contain the proposed high affinity ligand binding site in the center of the sequence. The terminal ends of the duplexes were selected to be G/C rich to enable good annealing of the sequences, and the sequences directly adjacent to the proposed specific binding site were selected based on the flanking sequences identified in DNAse I footprinting experiments to maintain consistency.

ESI mass spectra of solutions containing 10 μM of **V1** and 10 μM of ds1 in an ammonium acetate/methanol buffer were evaluated first. Both the 5- and 6- charge states are prominent for the duplex/V1 complexes. For a typical solution containing **V1** and ds1, the only complexes present in the spectrum possess a ligand/DNA binding stoichiometry of 1:1 with 2:1 complexes being at less than 5% of the relative abundance of the 1:1 complexes.

Concentration-dependent binding studies were undertaken to examine the extent of complexation changes as a result of varying the ligand/DNA ratios. Solutions containing ds1 at 10 μM and either 2.5, 5.0, 10, or 20 μM of **V1** were analyzed by ESI-MS (spectra not shown). As the ligand/DNA molar ratio was increased, the relative ion abundance of the 1:1 complexes in the 5- and 6- charge states increased, while that of the free duplex decreased. However, no 2:1 complexes emerged, even with excess ligand in solution. While it is possible that there are multiple sites on the duplexes for which **V1** could bind to form 2:1 complexes, binding of **V1** at a second site is not anticipated to be as strong because the ligand bound at the high affinity site, GGTACC, would cover six base pairs and thus presumably hinder the bis-intercalative binding of a second ligand. These results are consistent with solution-based studies of **V1**[23] and suggest ESI-MS is a promising tool for the analysis of DNA complexes containing bis-intercalators.

3.3.2 Binding selectivities of *trans*-D1 and *cis*-C1

To explore the effect of the scaffold on the binding specificity of the polyintercalators, two new bis-intercalators, *trans*-**D1** and *cis*-**C1,** were synthesized containing a rigid spiro-tricyclic scaffold (Scheme 3.1). As shown in Scheme 3.1, *trans*-**D1** is a *trans*-oriented ligand, while *cis*-**C1** is *cis*-oriented. While the design and synthesis of these compounds has been previously reported,[25] less is known about the DNA duplex binding of these compounds. We aimed to compare the binding of these compounds in a mass spectrometry based study.

To begin an ESI-MS evaluation of the binding behavior of *trans*-**D1** and *cis*-**C1**, the complexation of each ligand with DNA duplexes containing varying A-T and G-C base pair composition was evaluated. The NDI intercalator unit has exhibited a preference for G-G steps,[23] so it was of interest to assess the binding selectivities of the new compounds**.** ESI mass spectra were obtained for solutions containing *trans*-**D1** or *cis*-**C1** with a series of three duplexes with varying amounts of G/C and A/T base pair content: d(GCGGGGATGGGGCG/CGCCCCATCCCCGC) (ds2), d(GCGGGAATTGGGCG/CGCCCAATTCCCGC) (ds3), and d(GCGGAAATTTGGCG/CGCCAAATTTCCGC) (ds4).

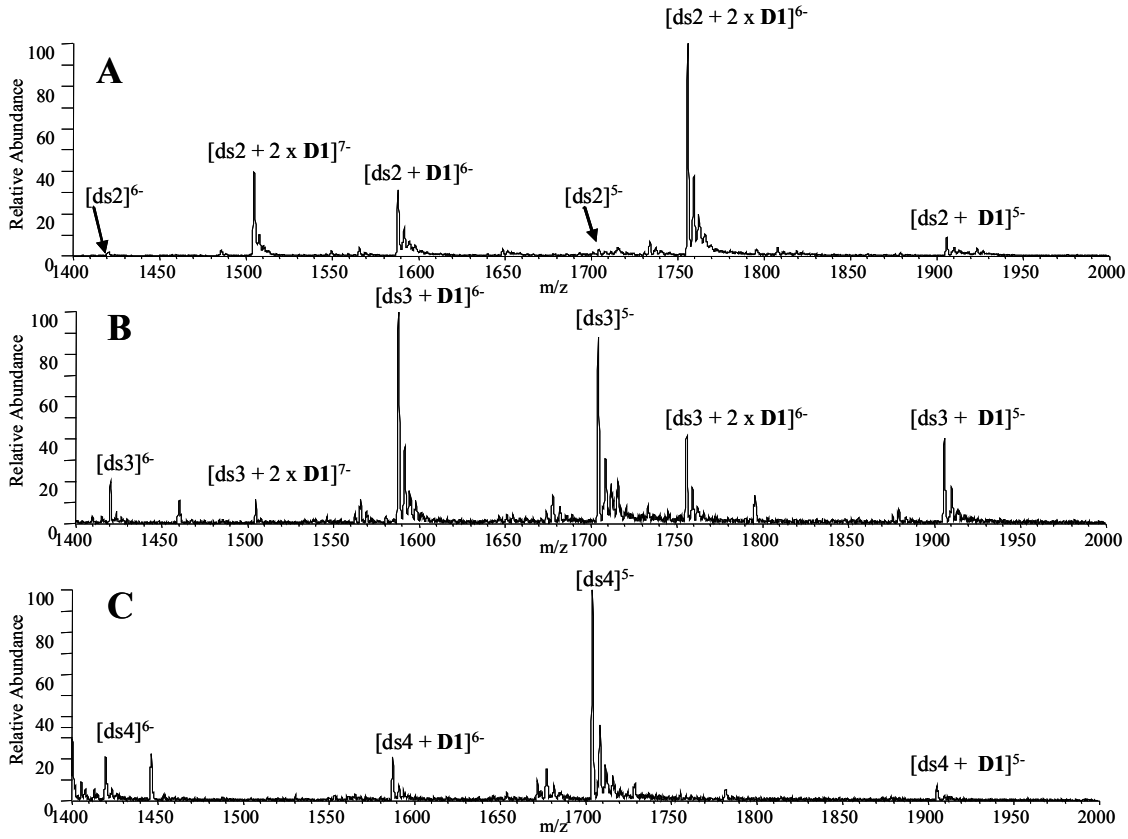


Figure 3.1 ESI mass spectra for complexes containing *trans*-**D1** and equimolar (10 μM) amounts of (A) ds2, d(GCGGGGATGGGGCG/CGCCCCATCCCCGC) (B) ds3, d(GCGGGAATTGGGCG/CGCCCAATTCCCGC and (C) ds4, d(GCGGAAATTTGGCG/CGCCAAATTTCCGC).

While both ligands exhibited GC base pair selectivity, the preference was more pronounced for *trans*-**D1**. As shown in Figure 3.1A, *trans*-**D1** readily forms abundant complexes with ds2, with ligand/DNA binding stoichiometries of 2:1 and 1:1, and little unbound DNA is present in the spectrum. The mass spectrum of *trans*-**D1** with ds3 (Figure 3.1B) shows that the relative abundance of the 2:1 complexes is significantly decreased compared to Figure 3.1A, and the abundance of the unbound DNA has increased. The spectrum of *trans*-**D1** with ds4, the duplex containing the most AT base

pairs, reveals that only 1:1 complexes are present and with significantly lower abundances relative to the duplex ion present at m/z 1705 (Figure 3.1C). As the AT-content increases, both the binding stoichiometry and relative abundance of the complexes formed between *trans*-**D1** and the duplex decrease dramatically.

The extent of complexation was calculated by expressing the sum of the abundances of ions from DNA/ligand complexes as a fraction of the total abundances of all ions from DNA as has been previously reported for ligand/DNA complexes.[32] Ions in the 5- and 6- charge states were used in this calculation. The binding results for *trans*-**D1** (discussed above) and those for *cis*-**C1** (spectra not shown) with ds2, ds3, and ds4 are summarized in Table 3.2.

Our results indicate that like *trans*-**D1**, *cis*-**C1** also forms more abundant complexes with the GC-rich DNA duplexes, as demonstrated by a value of 0.52 for the fraction of bound DNA with ds2, compared to 0.39 with ds3 and 0.32 with ds4. These results are consistent with solution dissociation kinetics experiments in which *cis*-**C1** and *trans*-**D1** demonstrated a strong preference for binding to poly (dGdC) over poly (dAdT) sequences.[25] This is a trend that has been previously reported for NDI intercalation.[21, 33] The observed poly d(GdC) preference of the ligands may be the result of a variety of binding interactions. Increased hydrogen bonding between the intercalators and functional groups in the major and minor grooves of GC-rich sequences could account for this preference. The imide carbonyls on the NDI units may also undergo a favorable electrostatic interaction with the N^2 amino group on G:C base pairs.[33] Steric and hydrophobic interactions could also play roles.[21-23] While the ligands demonstrate a general preference for GC-rich sequences, likely because of the NDI units, the functional linker imparts specific binding preferences.

Table 3.2: Fraction of bound DNA[a] for intercalator ligands and DNA duplexes.[d,f]

	ds1[b,c]	ds2	ds3	ds4	ds5[d]	ds6[c]	ds7[d]	ss1	ss5
V1	0.66	n/a[e]	n/a[e]	n/a[e]	0.52	n/a[e]	n/a[e]	0.00	n/a[e]
cis-**C1**	0.86	0.52	0.39	0.32	0.38	0.51	n/a[e]	0.00	n/a[e]
trans-**D1**	0.40	0.94	0.64	0.18	0.78	n/a[e]	0.85	n/a[e]	0.17
NDI1	0.53	n/a[e]	n/a[e]	n/a[e]	0.41	n/a[e]	n/a[e]	n/a[e]	n/a[e]

[a]All values +/- 0.05. This value was calculated to be the greatest standard deviation for the results of three experiments done with the samples. [b]Sequence contains proposed binding site of **V1**. [c]Sequence contains proposed binding site of *cis*-**C1.** [d]Sequence contains proposed binding site of *trans*-**D1.** [e]The abundances for all of the sodium adducts ions associated with a complex were included in the relative abundance calculations. [e]"n/a" indicates data was not collected because the results were not relevant to the study. [f]Solutions contained equimolar (10 μM) concentrations of ligand and DNA.

3.3.3 DNAse I Footprinting

In addition to evaluating the GC versus AT sequence selectivity of *trans*-**D1** and *cis*-**C1**, the binding of the compounds to duplexes containing potential specific binding sites was also evaluated by DNAse I footprinting and ESI-MS. It has been established that the naphthalenetetracarboxylic diimide based-intercalator units possess a preference for binding G-G steps.[23] In order to investigate the specificity of binding, DNAse I footprinting studies with *trans*-**D1**, *cis*-**C1** and reference ligand **V1** were carried out using a synthetic 92 bp DNA fragment containing 5 5'-GGNNCC-3' sites (Figure 3.2). As

reported previously, **V1** has a distinct binding site at 5'-GGTACC-3' with a K_d ~100 nM.[23]

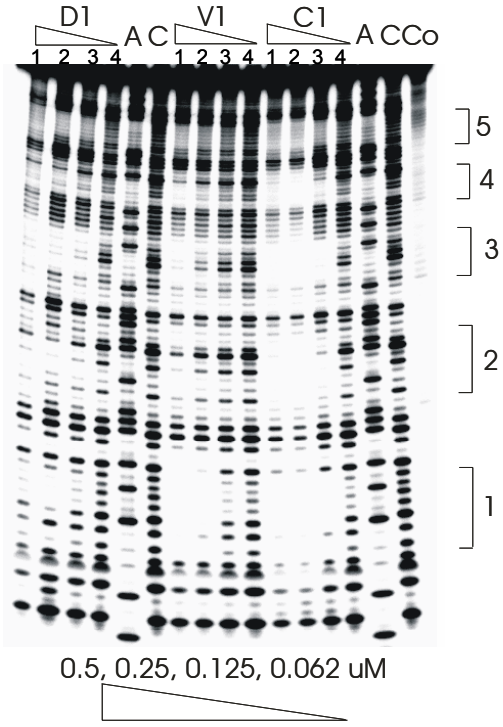


Figure 3.2: Footprinting of compounds *cis*-**C1**, *trans*-**D1** and **V1** on 92mer DNA with the (-) strand labeled on its 5'-end. Lane A represents Adenine-specific sequencing reaction. Lane Co contains DNA without DNAse I. For **D1**, **V1**, and **C1**, lanes 1-4 contain 0.5, 0.25, 0.125, and 0.062 μM ligand, respectively. Lane C contains DNA with DNAse but no compound. Sequences at 1: 5'-GGTACC; 2: 5'-GGATCC; 3: 5'-GGGCCC; 4: 5'-GGCGCC; 5: 5'-GGGGCC.

trans-**D1** showed some non-specific binding behavior, from targeting less than six-base pairs to only binding at high concentration (> 0.25 μM) (Figure 3.2). *cis*-**C1**, however, shows very different binding characteristics from the others. It not only binds to GGTACC with an even higher affinity than **V1**, but it also binds to other sequences with good affinity, such as GGGCCC and GGATCC. The variety of binding sites with

good binding affinity (K_d ~100 nM) demonstrates the potential of spiro-tricyclic scaffold as a "universal" scaffold for polyintercalators.

To further investigate the binding specificity of *trans*-**D1,** a 231 bp *Eco*RI-*Nhe*I restriction fragment of plasmid pBR322 was chosen for a second round of footprinting experiments. This sequence was previously used to screen the binding specificity of NDI ligands.[23] **V1** was used as a reference compound. As expected, one binding site for **V1** was clearly seen with this DNA fragment at the GGTACC sequence (results not shown). For *trans*-**D1**, two potential binding sites were identified: CAGTGA and GGCGAC.

3.3.4 ESI-MS Evaluation of *trans*-D1 and *cis*-C1 Binding Sequences

After identifying some possible specific binding sequences of *cis*-**C1** and *trans*-**D1** using DNAse I footprinting experiments, the binding of these ligands to DNA duplexes containing the most promising specific sequences were further evaluated by ESI-MS. Duplex ds1 contains the GGTACC sequence which is a potential specific binding site for *cis*-**C1** and is the same sequence identified for **V1,** and duplex ds5 contains a possible specific binding sequence CAGTGA for *trans*-**D1**. The full sequences of these duplexes are shown in Table 3.1. Additional experiments were done with ds6, which contains a second possible binding site for *cis*-**C1**, GGGCCC, and ds7, containing the second potential site for *trans*-**D1**, GGCGAT.

The mass spectrum acquired for a solution containing *cis*-**C1** and ds1 at equimolar concentrations in 50 mM ammonium acetate with 25% methanol demonstrates that *cis*-**C1** forms very abundant 1:1 complexes with this duplex (Figure 3.3A). The abundance of the unbound DNA ions are very low, suggesting *cis*-**C1** undergoes extensive complexation with this duplex. In addition, there are no complexes with a 2:1 binding stoichiometry, which is expected since the duplex contains a specific binding site in the center of this sequence. Based on the abundances of the ions in Figure 3.3A, the fraction

of bound DNA for the solution of ds1 with *cis*-**C1** was calculated to be 0.86 (Table 3.2). These results suggest that the binding of *cis*-**C1** with ds1 is more extensive than the binding of **V1** with the same duplex, as the fraction of bound DNA for **V1** with the duplex (shown in Figure 3.1B) was found to be 0.66 (Table 3.2). The higher binding affinity between *cis*-**C1** and the GGTACC sequence was also demonstrated in the footprinting experiments discussed above. *cis*-**C1** was also found to form complexes with ds6, which contained the GGGCCC binding site (spectra not shown). However, the fraction of bound DNA for *cis*-**C1** and ds6 was 0.51, which suggests less extensive complexation between the *cis*-**C1** and ds6 compared to ds1 and indicates a preference for the GGTACC binding sequence by the ligand.

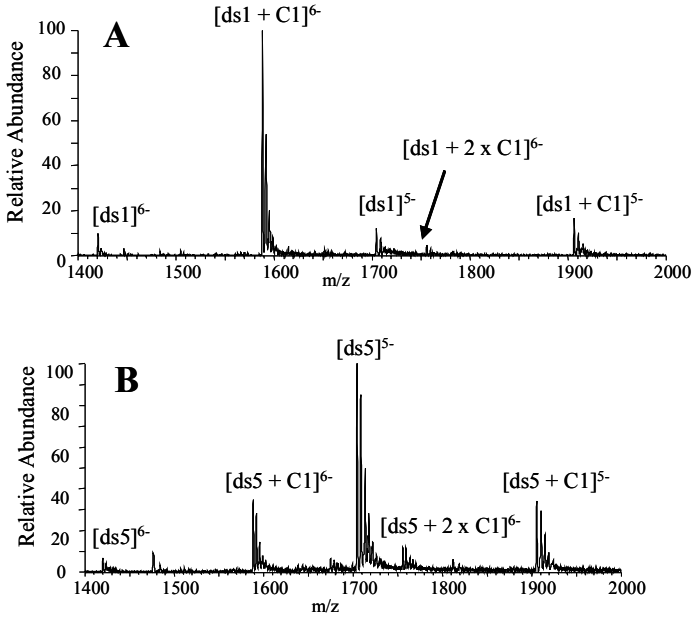


Figure 3.3: ESI mass spectra for complexes containing *cis*-**C1** and equimolar (10 μM) amounts of (A) ds1 d(GGGCGGTACCGCGG/CCGCGGTACCGCCC) and (B) ds5, d(GGGACAGTGAGGGG/CCCCTCACTGTCCC).

Experiments aimed at evaluating the binding between *cis*-**C1** and a DNA duplex that does not contain a target binding sequences were also undertaken. Duplex ds5 was selected for this experiment because the target binding sequence in the duplex, CAGTGA, was identified as a possible specific binding site for *trans*-**D1** but not *cis*-**C1**. The ESI mass spectrum of a solution containing equimolar (10 μM) concentrations of *cis*-**C1** and ds5 is shown in Figure 3.3B. The extent of complexation between *cis*-**C1** and the duplex is lower than what was observed in the spectra of solutions containing *cis*-**C1** with ds1 (Figure 3.3A) and ds6 (spectra not shown). In Figure 3.3B the abundances of the unbound DNA ions are considerably greater than the abundance of the 1:1 complexes. The fraction of bound DNA was calculated to be 0.38 which is significantly lower than the fraction of bound DNA for *cis*-**C1** with ds1 (0.86) and moderately lower than that of *cis*-**C1** with ds6 (0.51). These results indicate *cis*-**C1** forms more abundant complexes with ds1 and ds6 which is consistent with the specific binding site identified by DNAse I footprinting experiments.

Similar experiments were undertaken involving *trans*-**D1** and ds5 and ds7, each which contain a possible specific binding site of the ligand (CAGTGA and GGCGAC, respectively) and ds1, which was identified as a specific binding site for *cis*-**C1** but not *trans*-**D1**. The results of these experiments (spectra not shown) are summarized in Table 3.2, and they suggest that while the structures of *trans*-**D1** and *cis*-**C1** are similar, they exhibit different binding behavior with the duplexes. The fraction of bound DNA in the spectrum of *trans*-**D1** with ds5 was calculated to be 0.78, which is much greater than the fraction of bound DNA of *cis*-**C1** with the same duplex (0.38). *trans*-**D1** also formed abundant complexes with ds7, as indicated by a fraction of bound DNA of 0.85 for the ligand with the duplex. Conversely, the fraction bound for ds1 with *trans*-**D1** was only 0.41 compared to 0.86 observed with *cis*-**C1**.

As a further comparison, the fraction of bound DNA of **V1** with ds1 is also summarized in Table 3.2 since **V1** and *cis*-**C1** have the same possible specific binding sequence. The 0.66 value for fraction bound is not as great as that observed with *cis*-**C1**, suggesting *cis*-**C1** might be an improvement over **V1** in terms of forming abundant complexes with the GGTACC sequence. The fraction of bound DNA in a spectrum of **V1** with ds5, a duplex that does not contain a specific binding site for **V1**, is 0.52 which is lower than that with ds1. However, the difference in the fraction of bound DNA for **V1** with ds1 (containing the specific binding sequence) and ds5 (no specific binding sequence) is not as great as that observed with *cis*-**C1**. These results indicate that *cis*-**C1** shows the most promising selectivity for binding to its target sequences over other sequences and demonstrate that the relative binding behavior of the bis-intercalators observed by ESI-MS correlates with DNAase I footprinting results.

Because **NDI1** contains only one intercalating unit and lacks the scaffold designed to interact with the groove of duplex DNA, this compound functions as a monointercalator and was used as a reference ligand to compare its complexation with the same duplexes as used in the experiments above for *cis*-**C1** and *trans*-**D1**. The ESI-mass spectra indicate the formation of 1:1 and 2:1 **NDI1**:duplex complexes, with the fractions bound for ds1 and ds5 summarized in Table 3.2. The greater abundances of the 2:1 **NDI1**:duplex complexes is consistent with a lower specificity of **NDI1**.

3.3.5 Concentration Dependent Binding of *trans*-D1 and *cis*-C1

To further explore how the complexation of *trans*-**D1** and *cis*-**C1** changes with ligand/DNA molar ratios, concentration dependent binding was assessed. A series of solutions containing *trans*-**D1** or *cis*-**C1** with a duplex containing a proposed specific binding site (ds5 for *trans*-**D1** and ds1 for *cis*-**C1**) were prepared with a DNA concentration of 10 μM and a variable ligand concentration of 2.5, 5.0, 10, or 20 μM. The

mass spectra of *trans*-**D1** with ds5 at these molar ratios demonstrate that when the ligand concentration is increased relative to the DNA concentration, changes in the extent of ligand complexation are observed in the mass spectra. At a *trans*-**D1**/ds5 molar ratio of 0.25, complexes with 1:1 binding stoichiometry are present, but with low relative abundances compared to the unbound DNA ions. As the molar ratio is increased to 0.5, the relative abundance of the 1:1 complexes increases, and low abundance 2:1 complexes emerge while the abundance of the free DNA decreases. When the molar ratio is increased to 1:1 or even 2:1, the abundances of the 2:1 complexes increase further, while the abundance of the unbound DNA ions diminish.

The appearance of the 2:1 complexes at higher *trans*-**D1**/ds5 molar ratios is notable since ds5 contains only one relatively high affinity binding site for *trans*-**D1**. The results imply that *trans*-**D1** is able to bind to the DNA duplex in a non-specific manner when the ligand concentration is increased relative to the DNA. At this point, it is unknown how the 2:1 complexes are formed, but some possible scenarios include that the ligand may bind to the DNA via the intercalation of only one of the NDI units, two *trans*-**D1** ligands could be aggregating in solution and then binding to the DNA duplex, the second ligand may non-specifically aggregate to the DNA, or there could be two binding sites on the duplex that are mutually exclusive. This last scenario is the least likely since the duplex contains one proposed specific binding site at the center of the sequence and the ligand bound at the higher affinity site would hinder the binding of the second molecule. The presence of *trans*-**D1** footprints that are less than six bases long in Figure 3.2 suggest the ligand is able to partially bind to the duplex via one NDI unit. Future NMR modeling work will shed light on the structure of these higher binding stoichiometry complexes.

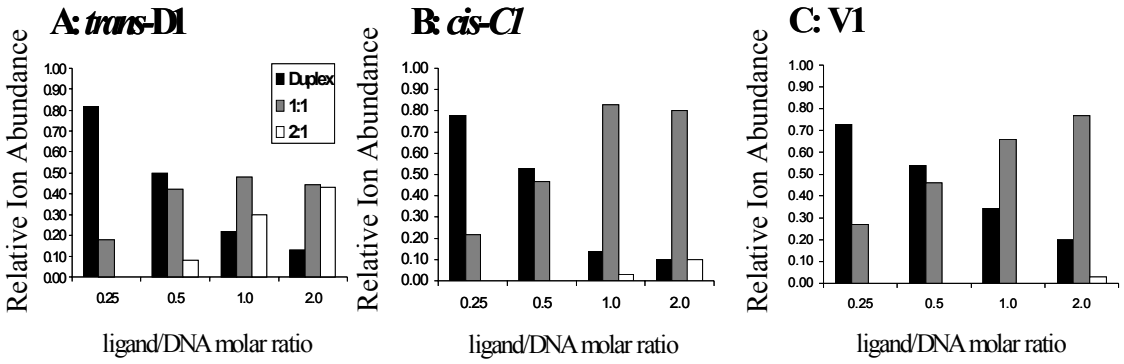


Figure 3.4: Summary plots of concentration dependent binding studies of (A) *trans*-D1 with ds5, (B) *cis*-C1 with ds1, and (C) V1 with ds1, indicating distribution of free duplexes, 1:1 ligand:duplex complexes, and 2:1 ligand:duplex complexes. Solutions contained the specified duplex DNA at 10 μM and ligand at, 2.5 μM, 5.0 μM, 10 μM, and 20 μM.

The results of the concentration dependent binding studies for *trans*-**D1** with ds5, *cis*-**C1** with ds1 (spectra not shown), and **V1** with ds1 (discussed earlier) are summarized by the graphs in Figure 3.4, which reflect the relative ion abundance of the unbound duplex ions (black bars), 1:1 complexes (grey bars), and 2:1 complexes (light bars), grouped by ligand/DNA ratio. The results for *cis*-**C1** with ds1 are significantly different from those of *trans*-**D1** with ds5. As the ligand molar ratio is increased, the relative ion abundance of the free DNA duplex decreases and the abundances of the 1:1 complexes increase. However, when there is excess *cis*-**C1** in solution, the 1:1 binding stoichiometry is dominant and only very low abundance 2:1 complexes are observed. The results of *cis*-**C1** with ds1 are similar to **V1** with ds1, which are also summarized in Figure 3.4C and suggest that like **V1**, *cis*-**C1** is binding specifically to the DNA duplex that contains a target site.

3.3.6 ESI-MS/MS Studies of Complexes Containing V1, *trans*-D1 and *cis*-C1.

An additional goal of this study was to determine if a bis-intercalative binding mode could be distinguished from a mono-intercalative binding mode using CAD since full scan mass spectra provide little insight into binding interactions between ligands and DNA. Until now, no MS/MS studies have been done on complexes containing bis-intercalators. Our group[34] and others[12, 35] have used collisional activated dissociation experiments to examine the fragmentation patterns of intercalator/DNA complexes, including complexes of actinomycin-D,[12,35] daunomycin,[34] and nogalamycin.[34] For complexes containing actinomycin-D, the predominant dissociation route is loss of the drug, but some strand separation (with retention of actinomycin D by one strand) and some nucleobase loss are also observed to a lesser extent.[12, 35] Our own CAD results for duplex/actinomycin-D complexes confirm that the primary fragmentation route is the disruption of the non-covalent interactions between actinomycin-D and the duplex, resulting in loss of actinomycin-D. Our earlier studies of complexes containing daunomycin or nogalamycin in a quadrupole ion trap mass spectrometer indicated that the drug/DNA complexes followed charge state dependent fragmentation patterns.[34] At lower charge states, the dominant fragmentation pathway was ejection of the drug leaving the intact duplex, while at higher charge states, separation of the individual single strand components of the duplex occurred, leaving the drug bound to one of the single strands. Dissociation by cleavage of a nucleobase was insignificant or not observed for complexes containing daunomycin or nogalamycin.[34] CAD of complexes containing reference ligand **NDI1,**[27] a monointercalator, resulted in fragmentation patterns consistent with those obtained for the daunomycin and nogalamycin complexes described above (spectra not shown). Complexes in the 5- charge state dissociated via loss of a neutral **NDI1** ligand,

while complexes in the 6- charge state produced fragment ions resulting from ligand ejection and predominant strand scission (spectra not shown).

CAD experiments were undertaken in the present study to examine what, if any, differences exist in the fragmentation pathways of complexes containing **V1**, *trans*-**D1**, and *cis*-**C1** compared to the monointercalators. The low charge state $[ds + L]^{5-}$ complexes, where L represents either **V1**, *cis*-**C1**, or *trans*-**D1,** produced the same fragmentation pattern**,** characterized by the guanine nucleobase loss (Figure 3.5A). This fragmentation pattern is different from what is commonly observed for the complexes containing commercial monointercalators, which predominantly dissociate via ejection of the ligand.[12, 35] For **V1**, *trans*-**D1** and *cis*-**C1**, all 1:1 complexes in the 6- charge state dissociated via dominant guanine nucleobase loss in addition to strand scission, with the ligand remaining bound to a single strand as demonstrated by the CAD spectrum of $[ds1 + \mathbf{C1}]^{6-}$ shown in Figure 3.5B. There were also very low abundance ions resulting from ejection of the negatively charged ligand, leaving the intact duplex. The predominance of nucleobase loss rather than ligand ejection of the 1:1 complexes in the 5- and 6- charge states is indicative of stronger binding interactions between the bis-intercalators and duplex DNA compared to traditional monointercalators, resulting from two intercalation sites and specific hydrogen bonding interactions between the scaffold and the DNA grooves. It is interesting that the strand scission pathway occurs in such a way that the ligand remains bound exclusively to only one of the ODNs (i.e. ss1a in Figure 3.5B, but not to ss1b). Although ss1a is more G-rich than ss1b, the underlying reason for this ODN selectivity is not clear.

The 2:1 complexes in the 6- charge state produced a different CAD pattern than those described above for the 1:1 complexes. *Trans*-**D1** was the only compound to form 2:1 complexes with sufficient abundance for CAD experiments, and the resulting CAD

mass spectrum of [ds5 + 2 x **D1**]$^{6-}$ is shown in Figure 3.5C. The most abundant product ion results from ejection of the negatively charged ligand, leaving the [ds5 + **D1**]$^{5-}$ complex. Ions resulting from strand scission and guanine base loss ions from the precursor complex are present but with significantly lower abundances. This result suggests that the second molecule is bound differently (and likely more weakly) than the first ligand in the 2:1 complexes which is similar to what has been previously observed with complexes containing nogalamycin [34]. Furthermore, the duplex DNA used in these experiments contain only one specific binding site for the ligands, so the presence of the 2:1 complexes indicates some non-specific binding by the second ligand, which is expected to be a weaker binding interaction.

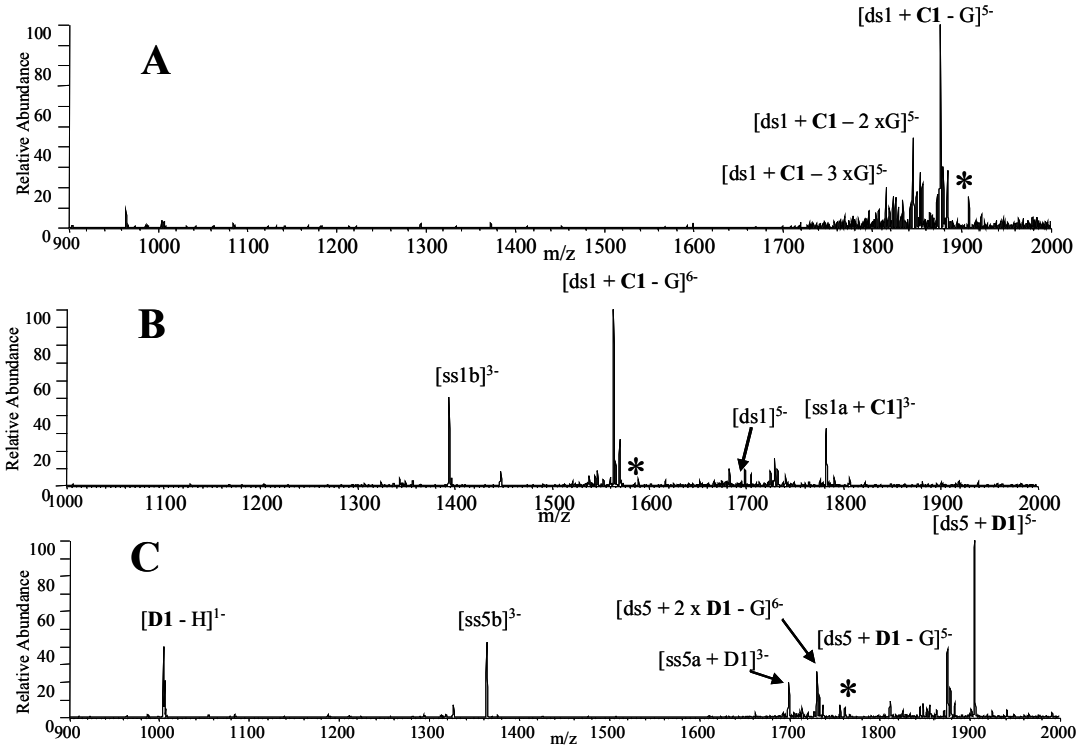


Figure 3.5: CAD mass spectra of (A) [ds1 + *cis*-**C1**]$^{5-}$, (B) [ds1 + *cis*-**C1**]$^{6-}$ and (C) [ds5 + 2 x *trans*-**D1**]$^{6-}$. The precursor ion is indicated by the asterisk. Solutions contained equimolar concentrations (10 μM) of the ligand and DNA duplex.

All complexes in the 7- charge state, regardless of binding stoichiometry, dissociated by strand scission leaving the ligands bound to a single strand (spectra not shown). This fragmentation pattern is consistent with past CAD studies involving intercalators[34] and is believed to result from coulombic repulsion of the more highly charged phosphate backbones.

The CAD fragmentation patterns of complexes containing echinomycin (Scheme 3.1) with duplex DNA were also evaluated in this study to serve as a comparison to the results of the new bis-intercalators. Echinomycin has been found to bind to duplex DNA via bis-intercalation, with the two quinoxaline rings preferably intercalating at CpG sites, and the bicyclic peptide scaffold oriented toward the DNA minor groove where hydrogen bonds are formed between the peptide and the nucleobases. In general, little ligand ejection was observed in the CAD spectra. Complexes in the 5- charge state undergo guanine base loss upon collisional activation. Guanine nucleobase loss is also observed in the CAD spectra of all complexes in the 6- charge state, in addition to ions resulting from strand scission. Complexes in the 7- charge state dissociated by strand separation. The CAD spectra of complexes containing enchinomycin are similar to those containing V1, trans-D1, or cis-C1, and are markedly different than those of the complexes containing monointercalators.

3.3.7 Single Strand Binding of V1, *trans*-D1 and *cis*-C1

The bis-intercalators evaluated in this study were designed to engage in two primary types of binding interactions with duplex DNA: intercalation interactions between the NDI units and the nucleobases, and hydrogen bonding interactions between the peptide scaffold and the minor or major groove of the DNA duplex. To determine if the ligands bind selectively to duplex DNA via these interactions over other DNA

structures, ESI-MS was used to analyze solutions of *trans*-**D1** or *cis*-**C1** with single strand DNA. The single strand ODNs used for these experiments, d(GGGCGGTACCGCGG) (ss1) and d(GGGACAGTGAGGGG) (ss5), were one of the two complementary single strand ODNs used to anneal duplexes ds1 and ds5, respectively. Based on the results of the duplex DNA binding studies of the ligands discussed above, *cis*-**C1** formed the most abundant complexes with ds1 so the binding of *cis*-**C1** to ss1 was evaluated, while *trans*-**D1** formed more abundant complexes with ds5, so ss5 was used. Using these single strand ODNs ensures that the same target binding sequences present in the duplexes are also found in the single strand sequences.

Solutions containing one single strand ODN and either **V1,** *trans*-**D1** or *cis*-**C1** at equimolar 10 μM concentration in ammonium acetate/methanol buffer were prepared and analyzed using the same instrument conditions used in experiments involving the duplex DNA discussed above. While *cis*-**C1** and **V1** did not form any complexes with ss1, *trans*-**D1** formed low abundance ligand/DNA complexes with ss5 (spectra not shown). The results of the single strand binding study are summarized by the fraction of bound DNA values shown in Table 3.2. In general, the complexes that formed between the *trans*-**D1** and the single strand DNA were relatively low in abundance compared to ligand/duplex DNA complexes (Table 3.2), suggesting that the ligand prefers binding to duplex DNA. It is also interesting to note that *trans*-**D1** is less selective for duplex DNA compared to *cis*-**C1,** which echoes the earlier observation that *trans*-**D1** exhibited considerable concentration dependent binding behavior with the duplex DNA and formed non-specific 2:1 complexes.

3.4 CONCLUSIONS

The utility of ESI-MS as a tool for screening non-covalent complexes formed between threading bis-intercalators and DNA is demonstrated in this study. Binding

stoichiometries and ligand sequence selectivity can be quickly assessed and qualitatively compared using the ESI-mass spectra, while CAD experiments provide information about ligand binding interactions. Our results demonstrated that **V1** forms abundant 1:1 complexes with ds1, the duplex containing its specific binding sequence 5'-GGTACC-3', and forms more abundant complexes with ds1 over ds5, which does not contain the specific binding sequence of the ligand. These results correlate well with previous DNAse I footprinting and NMR studies. Experiments involving **trans**-*D1* indicate the ligand extensively binds to duplex ds5 containing the target sequence 5'-CAGTGA-3' and forms significantly less abundant complexes with ds1, which does not contain a specific binding site identified by footprinting experiments. However, at higher ligand/DNA molar ratios, *trans*-**D1** forms 2:1 complexes which is indicative of non-specific binding by the ligand. *Cis*-**C1** exhibited the most promising specific binding behavior to ds1 containing the 5'-GGTACC-3' target sequence as evidenced by the formation of significantly more abundant complexes with this duplex over ds5 which did not contain a specific binding site. Unlike *trans*-**D1**, *cis*-**C1** did not form extensive 2:1 complexes at ligand/DNA molar ratios greater than one, conveying that the *cis* structure of **C1** is favorable for more specific binding.

In general the CAD spectra of the bis-intercalators are characterized by dominant guanine base loss in the 5- and 6- charge states, with increasing degrees of strand scission as the charge state increases. The different CAD fragmentation patterns exhibited by the bis-intercalator/duplex complexes compared to complexes containing known monointercalators mirror the shift in binding interaction of the ligands, characterized by intercalation at two sites and hydrogen bonding interactions with the DNA major or minor groove.

3.5 REFERENCES

(1) Goodman, L. S., Hardman, J. G., Limbird, L. E., and Gilman, A. G. *Goodman & Gilman's the Pharmacological Basis of Therapeutics*, 10th ed.; McGraw-Hill: New York, 2001 (1381-1460).

(2) Propst, C. L., and Perun, T. J. *Nucleic Acid Targeted Drug Design*; M. Dekker: New York, 1992, (1-12).

(3) Hofstadler, S. A.; Griffey, R. H. *Chem. Rev.* **2001**, *101*, 377-390.

(4) Beck, J. L.; Colgrave, M. L.; Ralph, S. F.; Sheil, M. M. E. *Mass Spectrom. Rev.* **2001,** *20*, 61-87.

(5) Brana, M. F.; Cacho, M.; Gradillas, A.; de Pascual-Teresa, B.; Ramos, A. *Curr. Pharm. Des.* **2001,** *7*, 1745-1780.

(6) Triolo, A.; Arcamone, F. M.; Raffaelli, A.; Salvadori, P. *J. Mass Spectrom.* **1997,** *32*, 1186-1194.

(7) Kapur, A.; Beck, J. L.; Sheil, M. M. *Rapid Commun. Mass Spectrom.* **1999,** *13*, 2489-2497.

(8) Gupta, R.; Kapur, A.; Beck, J. L.; Sheil, M. M. *Rapid Commun. Mass Spectrom.* **2001,** *15*, 2472-2480.

(9) Colgrave, M. L.; Beck, J. L.; Sheil, M. M.; Searle, M. S. *Chem. Commun*. **2002,** *6*, 556-557.

(10) Furlan, R. L. A.; Watt, S. J.; Garrido, L. M.; Amarante-Mendes, G. P.; Nur-E-Alam, M.; Rohr, J.; Brana, A.; Mendez, C.; Salas, J. A.; Sheil, M. M.; Beck, J. L.; Padilla, G. *J. Antibiot.* **2004,** *57*, 647-654.

(11) Urathamakul, T.; Beck, J. L.; Sheil, M. M.; Aldrich-Wright, J. R.; Ralph, S. F. *Dalton Trans.* **2004,** *17*, 2683-2690.

(12) Wan, K. X.; Gross, M. L.; Shibue, T. *J. Am. Soc. Mass Spectrom.* **2000,** *11*, 450-457.

(13) Wan, K. X.; Shibue, T.; Gross, M. L. *J. Am. Chem. Soc.* **2000,** *122*, 300-307.

(14) Beck, J. L.; Gupta, R.; Urathamakul, T.; Williamson, N. L.; Sheil, M. M.; Aldrich-Wright, J. R.; Ralph, S. F. *Chem. Commun.* **2003,** *5*, 626-627.

(15) Gupta, R.; Beck, J. L.; Ralph, S. F.; Sheil, M. M.; Aldrich-Wright, J. R. *J. Am. Soc. Mass Spectrom.* **2004,** *15*, 1382-1391.

(16) Gabelica, V.; De Pauw, E.; Rosu, F. *J. Mass Spectrom.* **1999,** *34*, 1328-1337.

(17) Greig, M. J.; Robinson, J. M. *Journal of Biomolecular Screening* **2000,** *5*, 441-454.

(18) Rosu, F.; De Pauw, E.; Guittat, L.; Alberti, P.; Lacroix, L.; Mailliet, P.; Riou, J. F.; Mergny, J. L. *Biochemistry* **2003,** *42*, 10361-10371.

(19) Reyzer, M. L.; Brodbelt, J. S.; Kerwin, S. M.; Kumar, D. *Nucleic Acids Res.* **2001,** *29*, e103.

(20) Carrasco, C.; Rosu, F.; Gabelica, V.; Houssier, C.; De Pauw, E.; Garbay-Jaureguiberry, C.; Roques, B.; Wilson, W. D.; Chaires, J. B.; Waring, M. J.; Bailly, C. *ChemBiohem.* **2002,** *3*, 1235-1241.

(21) Lokey, R. S.; Kwok, Y.; Guelev, V.; Pursell, C. J.; Hurley, L. H.; Iverson, B. L. *J. Am. Chem. Soc.* **1997,** *119*, 7202-7210.

(22) Lee, J.; Guelev, V.; Sorey, S.; Hoffman, D. W.; Iverson, B. L. *J. Am. Chem. Soc.* **2004,** *126*, 14036-14042.

(23) Guelev, V.; Lee, J.; Ward, J.; Sorey, S.; Hoffman, D. W.; Iverson, B. L. *Chem. Biol.* **2001,** *8*, 415-425.

(24) Guelev, V.; Sorey, S.; Hoffman, D. W.; Iverson, B. L. *J. Am. Chem Soc.* **2002,** *124*, 2864-2865.

(25) Chu, Y.; Lynch, V.; Iverson, B. L. *Tetrahedron,* **2006,** *62*, 5536-5548.

(26) Guelev, V. M.; Cubberley, M. S.; Murr, M. M.; Lokey, R. S.; Iverson, B. L. *Method. Enzymol.* **2001,** *340*, 556-570.

(27) Cubberley, M. S.; Iverson, B. L.. *J. Am. Chem. Soc.* **2001,** *123*, 7560-7563.

(28) Sambrook, J.; Russell, D. W. *Molecular cloning : a laboratory manual*; 3rd ed.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, N.Y., **2001**, 5.4-5.86**.**

(29) Guelev, V. Ph.D. Thesis, University of Texas at Austin, Austin, TX, **2002**.

(30) Rueda, M.; Luque, F. J.; Orozco, M. *J. Am. Chem. Soc.* **2005,** *127*, 11690-11698.

(31) Gidden, J.; Ferzoco, A.; Baker, E. S.; Bowers, M. T. *J. Am. Chem. Soc.* **2004,** *126*, 15132-15140.

(32) Mazzitelli, C. L.; Kern, J. T.; Rodriguez, M.; Brodbelt, J. S.; Kerwin, S. M. *J. Am. Soc. Mass Spectrom.* **2006,** *17*, 593-604.

(33) Liu, Z.-R.; Hecker, K. H.; Rill, R. L. *J. Biomol. Struct. Dynam.* **1996,** *14*, 331-339

(34) Keller, K. M.; Zhang, J. M.; Oehlers, L.; Brodbelt, J. S. *J. Mass Spectrom.* **2005,** *40*, 1362-1371.

(35) Gabelica, V.; De Pauw, E. *J. Am. Soc. Mass Spectrom.* **2002,** *13*, 91-98.

Chapter 4: The Metal-Mediated Binding by Benzoxazoles to Duplex DNA Evaluated by ESI-MS

4.1 INTRODUCTION

The discovery of the selective cytotoxic activity of UK-1,[1-3] a bis(benzoxazole) isolated as a secondary metabolite from *Streptomyces,* has stimulated the development of other benzoxazole and benzimidazole compounds with similar anticancer activities.[4-6] As a topoisomerase II inhibitor, one of the unique properties of UK-1 is its ability to bind biologically important divalent metal cations[3] and its metal-mediated DNA binding.[3,7] While UK-1 has demonstrated cytotoxicty against a number of cell lines, it does not inhibit the growth of bacteria, yeast or fungi,[1,8] making the mechanism of this selective cytotoxic activity and the metal binding properties of UK-1 and new analogs of great interest.

In a recent study, we used electrospray ionization mass spectrometry (ESI-MS) in conjunction with cytotoxicity assays to examine several simple analogs of UK-1 to explore the metal ion binding requirements of these compounds, assess metal-mediated DNA binding and evaluate anticancer and antibacterial activity.[6] Interestingly, the only ligand that exhibited anticancer activity, **2** (Figure 4.1), was also the only metal-mediated DNA binder with a preference for Ni^{2+} as determined by ESI-MS experiments. Two other compounds, **4** and **5,** formed complexes with DNA in a non-specific, non-metal mediated manner and showed antibacterial but not cytotoxic behavior. These results suggested a correlation between metal-mediated binding and anticancer activity of the compounds.

To improve upon the solubility of the benzoxazoles in aqueous solutions and to allow further examination of the metal-mediated DNA binding and anticancer activity, a

series of analogs of **2** have been synthesized with different ester and amide-linked side-chains (Figure 4.1).[9]

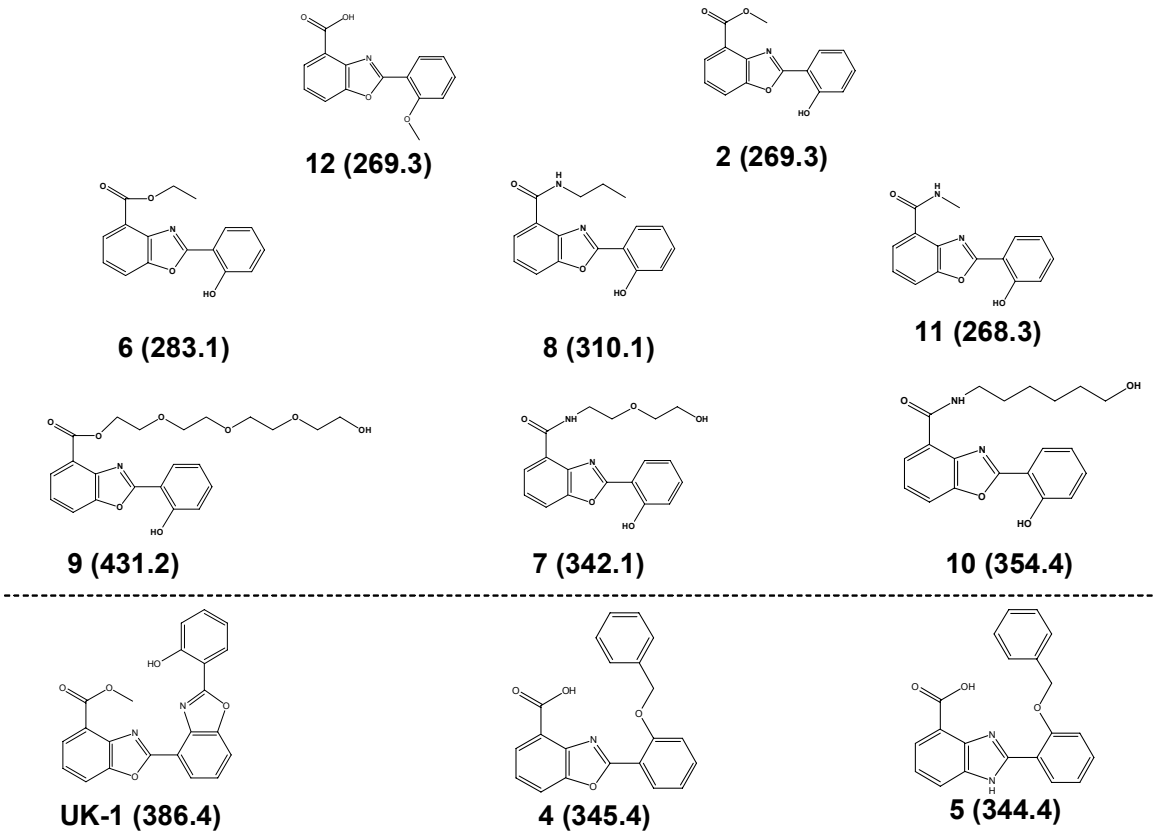


Figure 4.1: Structures of the benzoxazole and benzimidazole ligands. Molecular weights are given (in Da) in the parentheses. UK-1, 2, 4, and 5 were examined in a previous study.[6]

In the present study, ESI-MS was used to screen the binding of the new ligands to a series of divalent metals including Mg^{2+}, Ni^{2+}, Cu^{2+}, and Zn^{2+}. Ligand binding to duplex DNA in the presence and absence of metal ions was also assessed. ESI-MS has been shown to be a useful tool for the analysis of non-covalent ligand/DNA complexes due to

its sensitivity, low sample consumption, and fast analysis time that make it amenable to high throughput screening.[10-12] Early studies focused on examining well-studied, commercially available duplex DNA-binding compounds including both minor groove binders[13-16] and intercalators[17-19] to establish that the binding stoichiometries, selectivities, and specificities observed by ESI-MS correlate with known solution behavior. More recent studies have extended the use of the technique to novel compounds, including studies by our group examining the metal-dependent binding of UK-1 and related benzoxazoles and benzimidazole analogs.[6,7] Our previous studies have demonstrated that one of the main advantages of a mass spectrometry-based analysis technique for metal-mediated DNA binding ligands over traditional spectroscopic techniques is that information about ligand/metal and ligand/metal/DNA binding stoichiometries is obtained.

For the present study, anticancer cytotoxicity assays were also done so that the activity of these analogs could be correlated to the metal binding behavior elucidated in the mass spectrometry studies and compared to other previously reported UK-1 analogs. Previous studies had identified **2** as a simplified analog of UK-1 that retains the selective cytotoxicity of the natural product.[4] The analogs of **2** examined here were designed to increase the potential for metal ion complexation and metal mediated DNA binding by modifying the ester side chain of **2**, especially through the addition of addition sites for metal ion coordination.

4.2 EXPERIMENTAL

4.2.1 Chemicals

The synthesis of **2** and **12** has been previously reported.[4,20] The synthesis of the benzoxazole analogs **6**, **7**, **8**, **9**, **10**, and **11** will be reported separately.[9] The oligodeoxynucleotides (ODNs) using in this study, d(GCGGGGATGGGGCG),

d(CGCCCCATCCCCGC), d(GCGGGAATTGGGCG), d(CGCCCAATTCCCGC), d(GCGGAAATTTGGCG), and d(CGCCAAATTTCCGC) were purchased from IDT Technologies (Coralville, IA) as ammonium salts. Duplex DNA was annealed by preparing equimolar (1 mM) concentrations of the non-self-complementary ODNs in 250 mM ammonium acetate. Concentrations were verified spectroscopically using Beer's law and the extinction coefficients for the DNA strands provided by the manufacturer. The solutions were heated to 90 °C and allowed to cool to room temperature over 7 hours.

4.2.2 Mass Spectrometry

Analytical solutions containing a ligand and metal salt, a ligand and DNA duplex, or a ligand, metal salt and DNA duplex were prepared at equimolar (10 μM) concentrations unless noted otherwise, in 50 mM ammonium acetate solutions with 50% methanol. The samples were directly infused at 3 μL/min into a Thermo Electron (San Jose, CA) LCQ mass spectrometer. For the DNA binding experiments, the instrument was operated in the negative ion mode with an electrospray voltage of 3.5 kV and a heated capillary temperature of 90 - 110 °C with sheath and auxiliary gas flows of 40 and 10 arbitrary units, respectively. Ligand/metal ion solutions were examined in the positive ion mode using an electrospray voltage of 4.5 kV and the same heated capillary and gas flow rates used for the solutions containing DNA.

4.3.3 Cytotoxicity Assays

Cytotoxicity was determined using the AlamarBlue cell viability assay as described previously.[4] Briefly, aliquots of 100 μl cell suspension (1-3 x 10^3 cells) were placed in microtiter plates in an atmosphere of 5% CO_2 at 37 °C. After 24 h, 100 μL of culture media and 2 μl of the compound in DMSO were added to each well in duplicate, and the plates incubated an additional 72 h at 37 °C. Compounds, along with Mitomycin-C as a positive control were evaluated at final concentrations ranging from 0.001 to 50

μM. Cell viability was determined by removing the culture media from each well, and adding 200 μL of fresh media and 20 μL of AlamarBlue reagent (Biosource) , followed by an additional 6 h incubation and fluorescence measurement Beckman Coulter DTX880 plate reader with excitation at 530 nm and emission at 590 nm. The precent growth was calcualted from the fluorescence data using the equation:

$$\% \text{ Growth} = 100*(F_t - F_o)/ (F_c - F_o)$$

Where:

F_o = the averaged measured fluorescent intensities of AlamarBlue reagent at the time just before the exposure of the cells to the test substance.

F_t = the averaged measured fluorescent intensities of AlamarBlue reagent after 72 h exposure of the cells to the test substance at a particular concentration.

F_c = the averaged measured fluorescent intensities of AlamarBlue reagent after 72 h exposure of the cells to the vehicle without the test substance.

The concentration of compound required to inhibit growth by 50% (IC_{50}) was determined by nonlinear regression fitting the %Growth data to the equation:

$$y = \text{Min} + (\text{Max}–\text{Min})/(1+10\text{^}[(x-\log IC_{50})*H)]$$

Where:

x = compound concentration.

y = % Growth.

Min= the minimum response plateau (0%Growth).

Max= the maximum response plateau (100% Growth).

H = the Hill slope co-efficient.

4.3 RESULTS AND DISCUSSION

4.3.1 Ligand Binding to Metals

The binding of the benzoxazole ligands with metals including Mg^{2+}, Ca^{2+}, Ni^{2+}, Cu^{2+}, and Zn^{2+} was screened by ESI-MS to determine the binding preferences of the analogs and identify the metal ions that are most promising for enhancing ligand binding. Solutions containing a metal ion and a ligand were prepared at equimolar concentrations in the same ammonium acetate buffer used for subsequent DNA binding experiments and analyzed in the positive ion mode (Figure 4.2). No complexes were formed between any of the ligands and either Mg^{2+} and Ca^{2+}. Figure 4.2 show mass spectra for solutions containing Mg^{2+} with **11** (Figure 4.2a) and **9** (Figure 4.2b), demonstrating the absence of ligand/metal complexes. To ensure that the binding between the ligands and these metals was not concentration dependant, the metal ion concentration was increased 10 fold to 100 μM, however no complexation was observed.

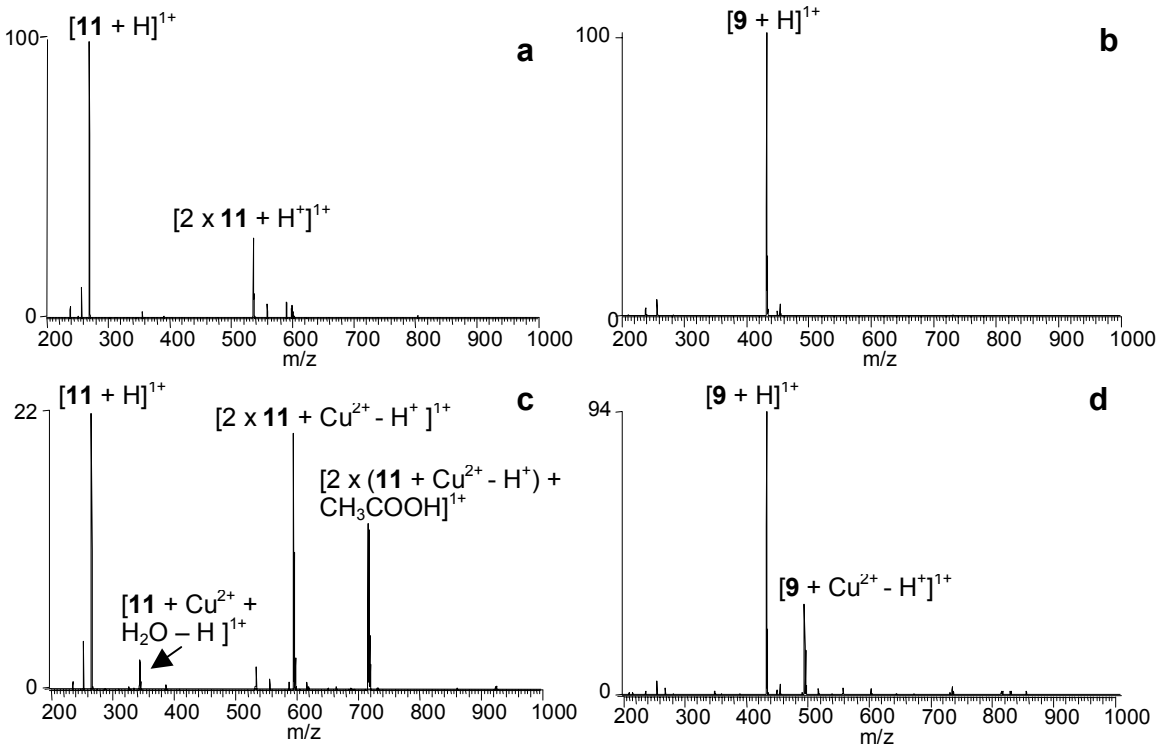


Figure 4.2: ESI-mass spectra for solutions containing benzoxazole ligands and metal ions: (a) 11 and Cu^{2+}, (b) 9 and Cu^{2+}, (c) 11 and Mg^{2+}, and (d) 9 and Mg^{2+}.

These results appear at first to contradict the results of previous studies in which the analogs of UK-1, including **2**, were found to form complexes with Mg^{2+}.[4] However the prior binding studies were undertaken in a methanol solvent due to the low solubility of the early analogs in aqueous buffers.[4] When methanol was used as a solvent in the current ESI-MS experiments, **6**, **7**, **8**, **9**, **10**, **11** and **2** formed abundant complexes with Mg^{2+} (spectra not shown). These results demonstrate the importance of the solvent in the metal ion binding of the benzoxazoles, presumably due to the differences in metal ion solvation in different solvents. To maintain consistency with the buffer used for the

DNA binding experiments, the ammonium acetate buffer was used for the remaining metal ion binding experiments.

6, **7**, **8**, **9**, **10**, **11** and **2** formed complexes with the divalent transition metals Ni^{2+}, Cu^{2+}, and to a lesser degree, Zn^{2+}. Typical ligand/metal ion binding stoichiometries ranged from 1:1 to 2:2, with the 2:2 complexes most likely being dimers of the 1:1 complexes. Ligands with the shorter side chains, **6**, **8**, and **11** formed abundant 2:1 complexes, as shown in Figure 4.2c for a solution containing **11** with Cu^{2+}. The short side-chain ligands were also capable of forming 1:1 complexes, however a water adduct was always bound to the complex, conceivably to fill the coordination shell of the metal. **7**, **9**, and **10**, the long side-chain compounds, were able to form abundant 1:1 complexes without a solvent adduct as demonstrated by the spectrum of **9** with Cu^{2+} shown in Figure 4.2d. It is likely that the long side chain is able to wrap partially around the metal ion upon binding, thereby filling the coordination shell without addition of solvent molecules.

The only benzoxazole ligand that did not form complexes with Ni^{2+}, Cu^{2+}, and Zn^{2+} was **12** (spectra not shown). This compound is the only analog examined in this study that contains a methyl ester group rather than a hydroxyl group on the phenyl moiety of the compound. The 2-(2'-hydroxylphenyl)benzoxazole moiety is also present in synthetic metal ion chelators,[21, 22] and is believed to play a key role in the metal ion chelation of the compounds.[3] The lack of metal ion binding observed by **12** in the ESI-MS experiments further implicates the role of the hydroxyl phenyl group in the metal ion binding of the benzoxazoles and makes **12** a good negative control compound for ligand/metal binding examined by ESI-MS.

The collisionally activated dissociation (CAD) mass spectra of the ligand/metal complexes were also examined by ESI-MS. Upon collisional activation, the 2:1 complexes containing **6**, **8** and **11** dissociate via the ejection of a neutral ligand with the

rapid adduction of a water molecule (always present in the trap in trace amounts) to the resulting 1:1 complex (spectra not shown). This is the same fragmentation pattern reported for complexes containing the anti-cancer analog **2** with Cu^{2+} and Ni^{2+}.[6] When the resulting 1:1 complexes were subjected to a second stage of CAD (MS^3), the complexes did not undergo further observable fragmentation. It is likely that the attached water molecule is dislodged during collisional activation, but then rapidly re-attaches prior to ion detection. This type of solvent adduction process has been commonly observed for transition metal complexes in a quadrupole ion trap.[23-27] This was the same fragmentation pattern observed for the dissociation of the water-solvated 1:1 complexes containing **6**, **8** and **11** and a transition metal.

Collisional activation of the 2:1 benzoxazole:metal complexes that contained the ligands with longer side-chains, **7**, **9**, and **10**, showed dissociation via the loss of one ligand, leaving 1:1 complexes (spectra not shown). Upon MS^3, these 1:1 complexes produced different fragmentation pathways that were dependent on the ligand. Complexes containing **7** and **9**, the ligands with a polyethylene glycol side chain, dissociated via the loss of small portions of the side chain, such as C_2H_4O, while the metal ion remained bound to the remainder of the ligand. These losses are not observed for **10**, which contains a six carbon alkyl chain with a terminal hydroxyl group. Instead, the 1:1 complexes containing **10** dissociate via the loss of the metal ion with the spontaneous adduction of methanol or water. In general, these initial ESI-MS experiments not only revealed that Ni^{2+}, Cu^{2+}, and Zn^{2+} are apparently the favored metals for ligand binding, but also showed differences in the preferred binding stoichiometries and fragmentation patterns for complexes containing ligands with long versus short side-chains.

4.3.2 DNA Binding of the Ligands Without Metals

In previous ESI-MS studies, benzoxazole and benzimidazole ligands that exhibited antibacterial activity were found to bind to DNA in the absence of metal cations, while compounds exhibiting anticancer activity, notably UK-1 and **2**, only formed complexes with duplex DNA in a metal-mediated manner. Therefore, the DNA binding of the new analogs of **2** (Figure 4.1) in the absence of metal cations was of great interest and was examined by ESI-MS in the present study. Three non-self-complementary DNA duplexes, d(GCGGGGATGGGGCG/CGCCCCATCCCCGC), d(GCGGGAATTGGGCG/CGCCCAATTCCCGC), and d(GCGGAAATTTGGCG/CGCCAAATTTCCGC), were used for the initial screening study. The duplex sequences were selected to have different degrees of G/C and A/T base pair composition to account for possible sequence selectivities of the ligands. Solutions containing one ligand and one duplex at equimolar (10 μM) concentration were prepared in 50 mM ammonium acetate with 50% methanol. The high methanol composition was necessary to ensure the ligands remained soluble in the analytical solutions. To allow for comparison of the DNA binding of the new analogs with **2**, experiments involving **2** were included in the present study to maintain consistency with the earlier study.[6]

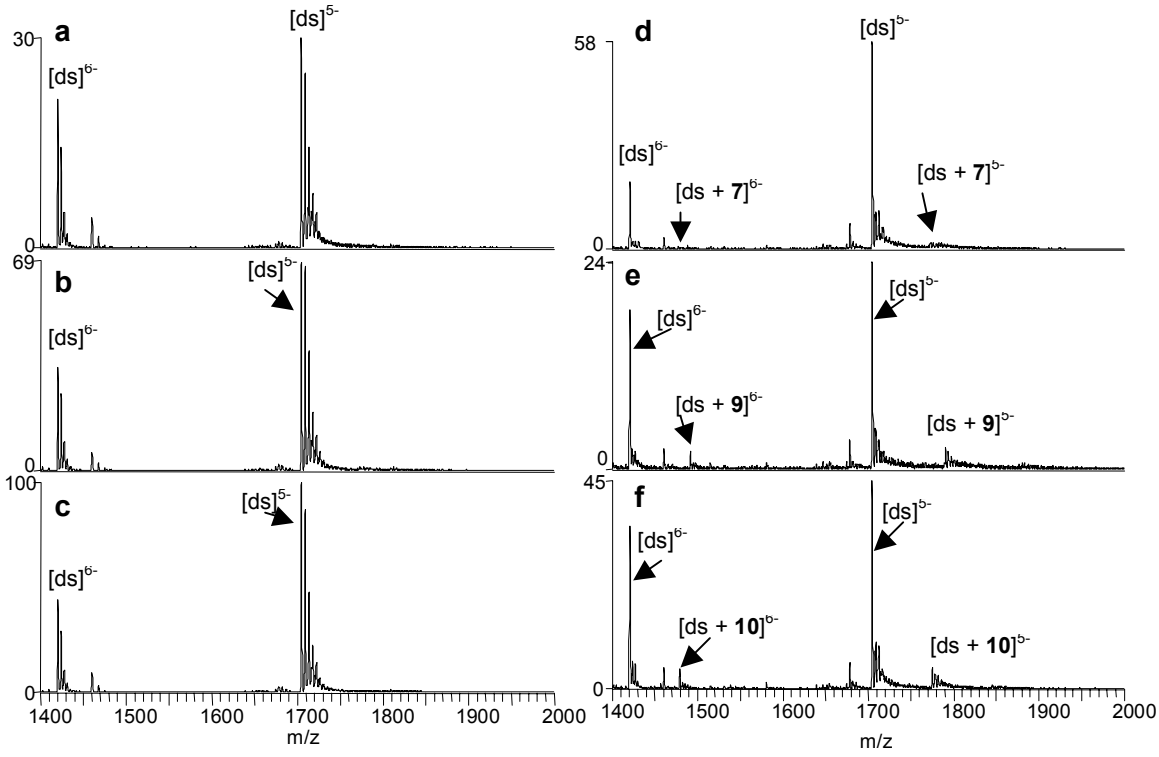


Figure 4.3: ESI-mass spectra of solutions containing duplex 3/4 and (a) **6**, (b) **8**, (c) **11**, (d) **7**, (e) **9**, and (f) **10**.

As demonstrated by the series of spectra shown in Figure 4.3 for solutions containing one of six ligands with d(GCGGGAATTGGGCG/CGCCCAATTCCCGC), three of the ligands, **7** (Figure 4.3d), **9** (Figure 4.3e), and **10** (Figure 4.3f), were found to form low abundance complexes with 1:1 ligand/DNA binding stoichiometries, while **6** (Figure 4.3a), **8** (Figure 4.3b) and **11** (Figure 4.3c) did not form any complexes detectable by ESI-MS. As expected, **2** likewise did not bind to DNA in the absence of metal ions (spectra not shown). The binding results of the ligands with the other two duplexes, d(GCGGGGATGGGGCG/CGCCCCATCCCCGC) and d(GCGGAAATTTGGCG/CGCCAAATTTCCGC), were similar to those shown in Figure 4.3, suggesting that these ligands do not have significant sequence selectivities.

The ligands that do bind to the duplex all contain ester- or amide-linked side chains that are longer than those compounds that did not bind to duplex DNA, suggesting the side chains could play a role in promoting non-metal mediated DNA binding of the benzoxazoles.

The non-metal mediated binding of **7**, **9**, and **10** differs from that of the antibacterial ligands **4** and **5** examined in our previous study. The complexes formed by **4** and **5** had greater relative abundances, and the binding stoichiometries ranged from 1:1 to 3:1 and were highly dependent on ligand concentration.[6] The binding behavior of ligands **4** and **5** was similar to that of commercial intercalators examined in previous ESI-MS studies.[17, 8] While **7**, **9**, and **10** formed complexes with duplexes without metals, the binding stoichiometries never exceed 1:1, and their relative abundances are low.

4.3.3 DNA binding of the Ligands with Metals

Metal cations are documented to play a key role in the DNA binding of UK-1 and **2,** and this metal-mediated binding behavior is thought to be related to the anti-cancer activity of these ligands over other non-metal-mediated benzoxazole and benzimidazole compounds. Previous studies of the solution metal ion binding ability of UK-1 and analogs had indicated that Mg^{2+} and Zn^{2+} were implicated in the metal-mediated DNA binding by these compounds,[4] while in ESI-MS studies Ni^{2+}, Co^{2+}, and Zn^{2+} promoted the greatest duplex binding for UK-1,[7] and the duplex binding of **2** was found to be mediated by Ni^{2+}.[6]

To determine if the binding of the new benzoxazole ligands are also metal-mediated, ESI-MS experiments were undertaken for solutions containing equimolar (10 μM) concentrations of a benzoxazole ligand, duplex d(GCGGGAATTGGGCG/CGCCCAATTCCCGC), and a metal salt. Based on the first section of results described above, only the metals that were found to form complexes

with the ligands, Ni^{2+}, Cu^{2+}, and Zn^{2+}, were used in this phase of the study. The most significant enhancement in duplex/ligand binding was observed with Cu^{2+} as demonstrated by the spectra shown in Figure 4.4, which correspond to solutions containing Cu^{2+}, the duplex and either **6** (Figure 4.4a), **8** (Figure 4.4b), and **11** (Figure 4.4c). While none of these three ligands bound DNA in the absence of metal ions, complexation is observed for all three in the presence of copper, with the most significant degree of binding observed for **8** and **11**. Complexes containing ligand/metal/DNA binding stoichiometries of 1:1:1 and 2:2:1 were observed for **8** and **11**, while only low abundance 1:1:1 were formed with **6**. Similar results were obtained for **2** (spectra not shown). The ions that are not labeled in Figure 4.4 correspond to sodium adducts of the DNA and ligand/Cu^{2+}/DNA complexes, with the source of sodium contamination likely being the benzoxazole solutions.

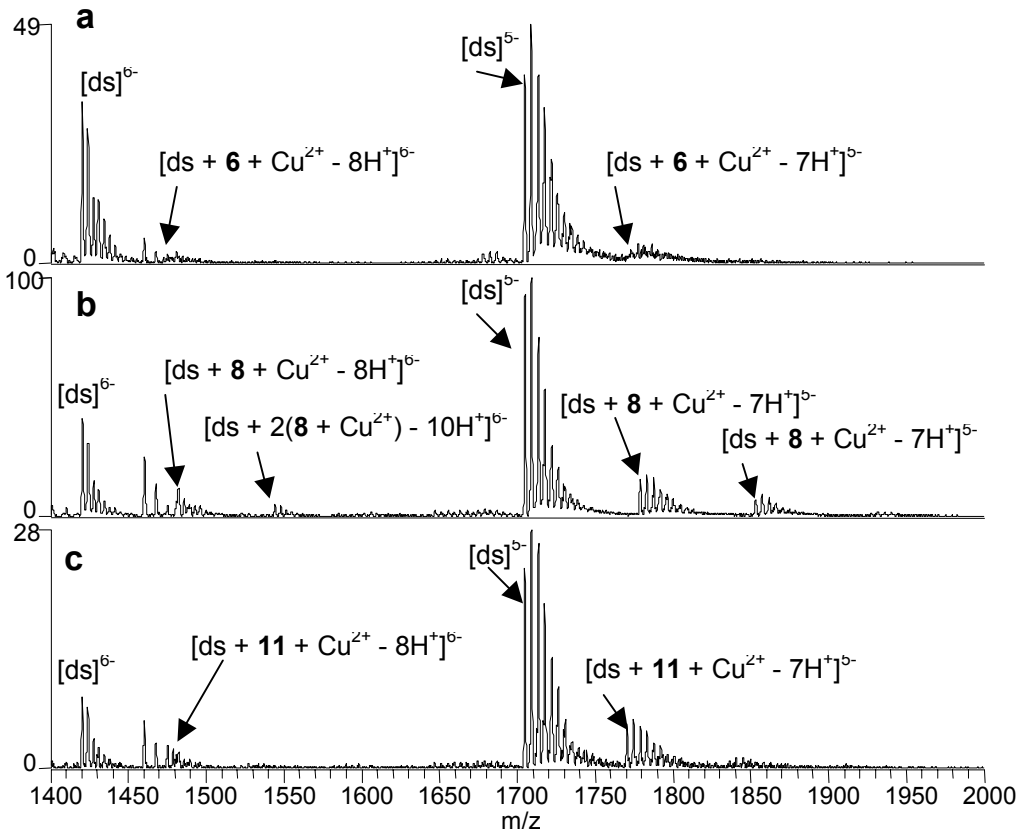


Figure 4.4: ESI-mass spectra of solutions containing duplex 3/4 and Cu^{2+} with (a) **6**, (b) **8**, and (c) **11**.

While **7**, **9**, and **10** were found to bind to duplex DNA in the absence of metal cations, their binding was enhanced in the presence of Cu^{2+}, most significantly for **9** and **10** (spectra not shown). To compare the changes in the degree of ligand binding upon addition of the metal salt, the fraction of bound DNA values were calculated by expressing the sum of the abundances of all ions attributed to DNA/ligand complexes as a fraction of the total abundances of all DNA-containing ions (both free DNA and DNA/ligand complexes) as has been previously reported for other ligand/DNA complexes.[29] The abundances for all of the sodium adducts associated with each complex

were included in the calculation, and only ions in the 5- charge state were used since this was the dominant charge state in the mass spectra. The increase in the extent of binding by **7**, **9** and **10** is reflected in the fraction of bound DNA values summarized in Table 4.1. The results here and summarized in Table 4.1 demonstrate that Cu^{2+} has the greatest impact on ligand binding, enhancing the binding of **7**, **9**, and **10**, and promoting binding by the other ligands with the most dramatic metal-mediated behavior seen for **8** and **11**, and to a lesser degree, **6**.

In general, Ni^{2+} had a moderate impact on ligand binding as demonstrated by the spectra shown in Figure 4.5. Extensive binding of Ni^{2+} to the duplex DNA is observed in Figure 4.5, but this does not translate into an enhancement of ligand binding. **11** was the only new analog that exhibited binding in the presence of Ni^{2+} yet did not bind to DNA without metal ions, as shown in Figure 4.5c. A low abundance 1:1:1 complex containing **11** is formed (Figure 4.5c), but with lower abundance relative to the complexes formed by **11** in the presence of Cu^{2+}. The duplex binding of **9** was also enhanced in the presence of Ni^{2+} (spectra not shown). For **9**, the fraction of bound DNA increased from 0.17 in the absence of metal ions to 0.34 in the presence of Ni^{2+} (Table 4.1). Experiments with **2** confirmed the results of our previous study that found the binding of the ligand to be Ni^{2+}-mediated. The fraction of bound DNA values based on the **2**/Ni^{2+}/DNA complexes was 0.18 which is similar to that of the **11**/Ni^{2+}/DNA complexes, 0.14.

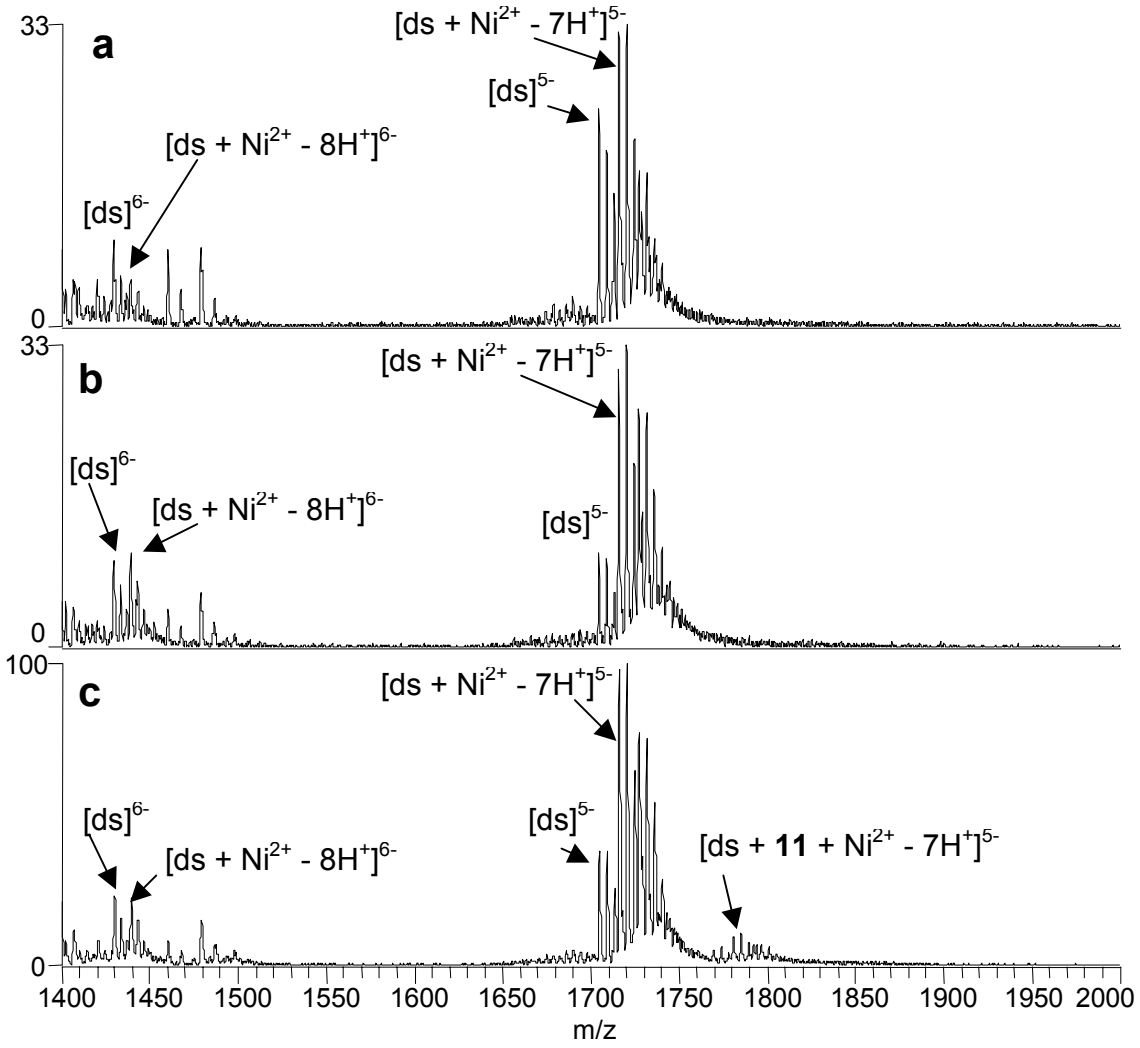


Figure 4.5: ESI-mass spectra of solutions containing duplex 3/4 and Ni^{2+} with (a) **6**, (b) **8**, and (c) **11**.

Figures 4.5A and 4.5B show that neither **6** nor **8** bound to the DNA with Ni^{2+}, nor did Ni^{2+} enhance the binding of **7** or **10** (spectra not shown). As summarized in Table 4.1, the fraction of bound DNA values for the **7** and **10** in the absence metal ions were 0.09 and 0.20, respectively (Figures 4.3d and 4.3f). While both ligands formed 1:1:1

ligand/Ni^{2+}/DNA complexes, the relative abundances of these complexes were either the same as or lower than the abundances of the complexes formed in the absence of metal ions. This is reflected in the fraction of bound DNA values for solutions containing the ligands with Ni^{2+} and DNA which were calculated to be 0.08 for **7** and 0.13 for **10**.

Table 4.1: Fraction of bound DNA values[a] for the benzoxazole ligands and d(GCGGGAATTGGGCG/CGCCCAATTCCCGC) with Ni2+, Cu2+, or Zn2+, and in the absence of metals. Solutions contained equimolar (10 μM) concentration of ligand, DNA, and, where appropriate, metal ion.

Ligand	No metal	Ni^{2+}	Cu^{2+}	Zn^{2+}
2	0	0.18	0.17	0
6	0	0	0.14	0
8	0	0	0.40	0
11	0	0.14	0.27	0
7	0.09	0.08	0.15	0.17
9	0.17	0.34	0.27	0.27
10	0.2	0.13	0.32	0.17

[a] All values $\pm$ 0.05.

Zn^{2+} did not have a significant impact on the benzoxazole ligand binding to DNA. Zn^{2+} did not promote DNA binding by **6**, **8**, **11**, or **2**, the ligands that likewise did not bind to DNA in the absence of metals (spectra not shown). **10** formed 1:1:1 ligand/Zn^{2+}/DNA complexes, however the fraction of bound DNA value for ligand binding in the presence of zinc was 0.17 which is not significantly different than it was in the absence of metal, 0.20 (spectra not shown). The DNA binding of **7** and **9** was enhanced by Zn^{2+}, as

indicated by an increase in their fraction of bond DNA values (spectra not shown). For **7**, the fraction of bound DNA without metals was 0.08 and in the presence of Zn^{2+} it increased to 0.17, while the values increased from 0.17 to 0.27 for **9**.

Aside from **2**, the ligands displaying the most pronounced metal-mediated DNA binding based on the ESI-MS results are **8** and **11**. The binding of **6** was also mediated by Cu^{2+}, however the complexes formed by the ligand are lower in relative abundance than those formed by **8** and **11**. The structures of **8** and **11** are similar to **2**. This result is somewhat unsurprising due to the similarities in their structures. The ester linkage of **2** is changed to an amide-linkage in **11**. While both **11** and **2** exhibited metal mediated binding, copper had a more positive impact on **11** binding than it did on **2**. Likewise, Ni^{2+} promoted greater binding for **2** than it did for **11**. This difference may result from the preferential coordination of Cu(II) to nitrogen atoms over those oxygen atoms.[30] Similar differences can be seen for the binding of **8**, which also has an amide-linked side chain that is extended by two carbon atoms compared to **11** and **2.** Compared to **2**, **8** also demonstrates greater DNA binding with copper and no binding with Ni^{2+}.

While the longer side-chains of **7**, **9** and **10** enhance non-metal-mediated DNA binding, they tend to preclude the dramatic metal-mediated DNA binding behavior observed for the ligands with the shorter side-chains. The longer side-chains of the ligands are believed to enhance the metal binding of the ligands by wrapping around and coordinating to the metals. The ester-linked polyethylene glycol side chain of **9** produced the best enhancement of metal mediated binding, as all three metal ions, Ni^{2+}, Cu^{2+}, and Zn^{2+}, increased the binding of the ligand to DNA relative to the binding in the absence of the metal. The results for **7**, with the shorter, amide-linked polyethylene glycol side chain were not as favorable. While the binding was enhanced by Cu^{2+} and Zn^{2+}, the overall fraction of bound DNA values were generally lower than the values for the other ligands

in the presence of metals. With an amide-linked side chain that did not contain ethylene glycol groups, the binding of **10** was only enhanced by Cu^{2+}. The binding results of the ligands with the longer side chains suggests that the longer, polyethylene glycol side chain of **9** is the most favorable for overall metal-mediated binding.

4.3.4 CAD Spectra of the DNA Complexes

The fragmentation pathways of the DNA complexes containing the benzoxazole ligands were also evaluated. The ligand/DNA complexes that were formed in the absence of metal ions dissociated by the neutral ejection of ligand, leaving the bare duplex ion. An example of this fragmentation route is shown in Figure 4.6A for the CAD spectrum of $[ds + \mathbf{9}]^{5-}$. This pathway is similar to what has been observed in past CAD studies of DNA complexes containing traditional DNA intercalators.[31-33] Complexes with ligand and metal bound to DNA dissociated via the ejection of the neutral ligand, with the metal ion being retained by the DNA duplex as shown in Figure 4.6B for $[ds + \mathbf{11} + Cu^{2+}]^{5-}$. Past ESI-MS studies of duplex DNA complexes containing UK-1 with a metal ion reported the same fragmentation pathway.[7] While it is difficult to draw conclusions about the binding mode of the benzoxazole/metal ion/DNA complexes based on CAD results, they do suggest that the binding interactions involved in complexes containing the new benzoxazole ligands are similar to that of UK-1.

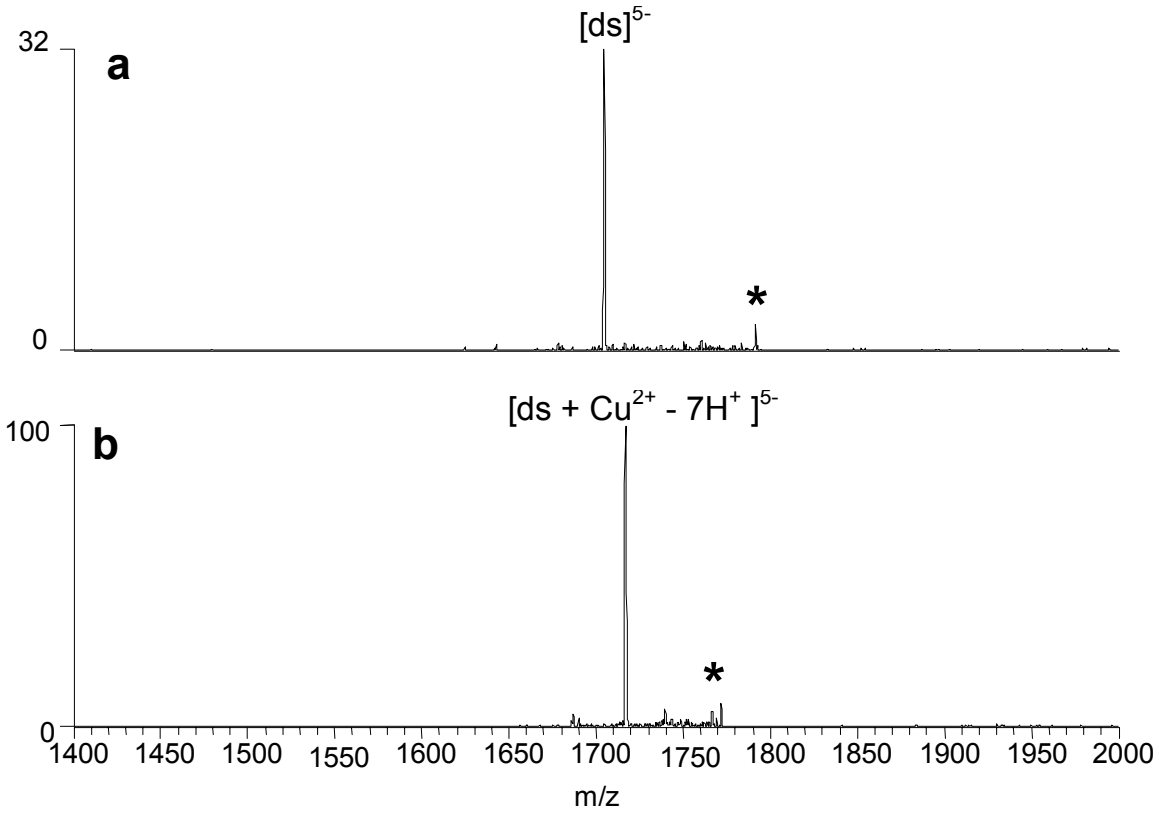


Figure 4.6: CAD spectra of (A) [ds + **9**]$^{5-}$ and (B) [ds + **11** + Cu^{2+}]$^{5-}$.

4.3.5 Cytotoxicity Assays of Benzoxazole Ligands

There are some interesting correlations between the degree of metal-mediated binding behavior of the benzoxazole ligands determined by ESI-MS and their cytotoxicity to two cancer cell lines. As summarized by the IC_{50} values shown in Table 4.2, the new ligands with the greatest cytotoxicity against the A549 lung cancer cell line were **8**, **9**, and **11**, all having IC_{50} values that were similar to **2**. Both **8** and **11** exhibited the most dramatic degree of metal-mediated DNA binding as their fraction of bound DNA values were increased from 0 in the absence of metal ions to 0.40 and 0.27,

respectively, in the presence of Cu^{2+}. **11** was also found to form complexes in the presence of Ni^{2+}. While **9** was not a truly metal-mediated DNA binding ligand, it exhibit the most consistent and dramatic enhancement in DNA binding in the presence of metal ions of the ligands that bound to DNA in the absence of metal ions.

Table 4.2: IC50 values(in μM) for the benzoxazole compounds against the A549 (lung cancer) and MCF7 (breast cancer) cell lines.

Ligand	A549	MCF7
2	12 ± 3	4 ± 2
6	40 ± 10	15 ± 5
8	14 ± 4	30 ± 10
11	11 ± 1	10 ± 8
7	41 ± 17	> 50
9	11 ± 1	13 ± 2
10	39 ± 11	31 ± 7
12	>50	>50

9 and **11** also showed cytotoxicity against the MCF7 breast cancer cell lines with IC_{50} values of 13 ± 2 μM and 10 ± 8 μM, respectively. Both of these values are on par with the IC_{50} value of **2**, 4 ± 2 μM9 ± 7. Interestingly, **8**, which showed activity against A549 cells comparable to that of **2** was relatively inactive against MCF7 cells (IC_{50} = 30 ± 10 μM), and **6**, which was relatively inactive against A549 cells, retained some activity against MCF7 cells (IC_{50} = 15 ± 5 μM). At this point the reasons for these differences are unclear, but we note that of the two analogs of **2** containing the most conservative structural changes, **6** and **11**, only **11** demonstrates enhanced metal-mediated DNA

binding relative to **2** and this analog also displays the most comparable cytoxicity against both cancer cell lines, relative to **2**. **9** also demonstrated good cytotoxic activity, and while it was not a truly metal-mediated DNA binder, its DNA complexation was consistently enhanced by the presence of Cu^{2+}, Ni^{2+}, and Zn^{2+}. In contrast, compound **12**, which does not bind metal ions, was the only compound in this series that did not display cytotoxicity towards either cell line.

4.4 CONCLUSIONS

The correlation between significant metal-mediated or metal-enhanced binding determined by ESI-MS and anticancer activity of the benzoxazoles ligands assessed by cytotoxicity assays is demonstrated in this study. ESI-MS experiments reveal that Cu^{2+} and Ni^{2+} form the most abundant complexes with **6**, **7**, **8**, **9**, **10**, and **11**, while less abundant complexes are formed with Zn^{2+}. For the complexes containing short side-chains, **6**, **8**, and **11**, 2:1 ligand/metal ion binding stoichiometries were predominant, which the compounds with longer side-chains, **8**, **9**, and **10**, formed abundant 1:1 complexes. DNA binding experiments reveal that the analogs with longer side-chains formed complexes with duplex DNA in the absence of metal ions, while those with shorter side-chains did not.

Of the ligands that did not bind to duplex DNA in the absence of metal ions, the DNA binding by **8** and **11** was enhanced most dramatically by Cu^{2+}. Ni^{2+} influenced duplex binding for **11** and **2**, and enhanced the binding by **9**. The metal ion with the least substantial effect was Zn^{2+} which only enhanced the binding of **7** and **9**, two ligands that formed complexes with DNA regardless of the presence of metals.

Of the compounds examined in this study, both **8** and **11** were also found to be the most cytotoxic against the A549 lung cancer cell line and **11** demonstrated moderate cytotoxicity against MCF7 breast cancer cells. Metal ions also enhanced the DNA

binding of the ligands with the long side-chains, most notably for **9**, which also exhibited the highest level of cytotoxicity of the long side-chain compounds.

4.5 REFERENCES

(1) Ueki, M.; Ueno, K.; Miyadoh, S.; Abe, K.; Shibata, K.; Taniguchi, M.; Oi, S. M. *J. Antibiot.* **1993**, *46*, 1089-1094.

(2) Shibata, K.; Kashiwada, M.; Ueki, M.; Taniguchi, M. UK-1, *J. Antibiot.* **1993**, *46*, 1095-1100.

(3) Reynolds, M. B.; DeLuca, M. R.; Kerwin, S. M. *Bioorg. Chem.* **1999**, *27*, 326-337.

(4) Kumar, D.; Jacob, M. R.; Reynolds, M. B.; Kerwin, S. M. *Bioorg. Med. Chem.* **2002**, *10*, 3997-4004.

(5) Wang, B. B.; Maghami, N.; Goodlin, V. L.; Smith, P. J. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 3221-3226.

(6) Oehlers, L.; Mazzitelli, C. L.; Brodbelt, J. S.; Rodriguez, M.; Kerwin, S. *J. Am. Soc. Mass Spectrom.* **2004**, *15*, 1593-1603.

(7) Reyzer, M. L.; Brodbelt, J. S.; Kerwin, S. M.; Kumar, D. *Nucleic Acids Res.* **2001**, *29*, art. no.-e103.

(8) Sato, S.; Kajiura, T.; Noguchi, M.; Takehana, K.; Kobayasho, T.; Tsuji, T. *J. Antibiot.* **1997**, *54*, 102.

(9) Rodriguez, M.; Kerwin, S. *in preparaion* **2007**.

(10) Hofstadler, S. A.; Griffey, R. H. *Chem. Rev.* **2001**, *101*, 377-390.

(11) Beck, J. L.; Colgrave, M. L.; Ralph, S. F.; Sheil, M. M. *Mass Spectrom. Rev.* **2001**, *20*, 61-87.

(12) Hofstadler, S. A.; Sannes-Lowery, K. A. *Nat. Rev. Drug Dis.* **2006**, *5*, 585-595.

(13) Gale, D. C.; Smith, R. D. *J. Am. Soc. Mass Spectrom.* **1995**, *6*, 1154-1164.

(14) Gabelica, V.; Galic, N.; Rosu, F.; Houssier, C.; De Pauw, E. *J. Mass Spectrom.* **2003**, *38*, 491-501.

(15) Gabelica, V.; Rosu, F.; Houssier, C.; De Pauw, E. *Rapid Commun. Mass Spectrom.* **2000**, *14*, 464-467.

(16) Rosu, F.; Gabelica, V.; Houssier, C.; De Pauw, E. *Nucleic Acids Res.* **2002**, *30*, e82.

(17) Gabelica, V.; De Pauw, E.; Rosu, F. *J. Mass Spectrom.* **1999**, *34*, 1328-1337.

(18) Beck, J. L.; Gupta, R.; Urathamakul, T.; Williamson, N. L.; Sheil, M. M.; Aldrich-Wright, J. R.; Ralph, S. F. *Chem. Commun.* **2003**, *5*, 626-627.

(19) Wan, K. X.; Shibue, T.; Gross, M. L. *J. Am. Chem. Soc.* **2000**, *122*, 300-307.

(20) DeLuca, M. R.; Kerwin, S. M. *Tetrahedron Lett.* **1997**, *38*, 199-202.

(21) Hoveyda, H. R.; Rettig, S. J.; Orvig, C. *Inorg. Chem.* **1993**, *32*, 4909-4913.

(22) Tanaka, K.; Kumagai, T.; Aoki, H.; Deguchi, M.; Iwata, S. *J. Org. Chem.* **2001**, *66*, 7328-7333.

(23) Vachet, R. W.; Callahan, J. H. *J. Mass Spectrom.* **2000**, *35*, 311-320.

(24) Williams, S. M.; Brodbelt, J. S. *J. Am. Soc. Mass Spectrom.* **2004**, *15*, 1039-1054.

(25) Satterfield, M.; Brodbelt, J. S. *J. Am. Soc. Mass. Spectrom.* **2001**, *12*, 537-549.

(26) Perera, B. A.; Ince, M. P.; Talaty, E. R.; Van Stipdonk, M. J. *Rapid Commun. Mass Spectrom.* **2001**, *15*, 615-622.

(27) Vachet, R. W.; Hartman, J. A. R.; Callahan, J. H. *J. Mass Spectrom.* **1998**, *33*, 1209-1225.

(28) Kapur, A.; Beck, J. L.; Sheil, M. M. *Rapid Commun. Mass Spectrom.* **1999**, *13*, 2489-2497.

(29) Mazzitelli, C. L.; Kern, J. T.; Rodriguez, M.; Brodbelt, J. S.; Kerwin, S. M. *J. Am. Soc. Mass Spectrom.* **2006,** *17*, 593-604.

(30) Douglas, B.; McDaniel, D.; Alexander, J. *Concepts and Models of Inorganic Chemistry*, Third ed.; John Wiley & Sons, Inc.: New York, NY, 1994.

(31) Wan, K. X.; Gross, M. L.; Shibue, T. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 450-457.

(32) Keller, K. M.; Zhang, J. M.; Oehlers, L.; Brodbelt, J. S. *J. Mass Spectrom.* **2005**, *40*, 1362-1371.

(33) Gabelica, V.; De Pauw, E. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 91-98.

Chapter 5: Evaluation of Binding of Perylene Diimide and Benzannulated Perylene Diimide Ligands to DNA by Electrospray Ionization Mass Spectrometry

5.1 INTRODUCTION

Small molecules that interact non-covalently with nucleic acid structures comprise an important class of anticancer, antitumor, and antibacterial therapies.[1, 2] Increased interest in the development and evaluation of DNA-interactive agents has stimulated the need for sensitive analytical techniques that can not only characterize drug/DNA interactions but are also compatible with library-based screening. Electrospray ionization-mass spectrometry (ESI-MS) has emerged as a useful tool for examining non-covalent drug/DNA complexes because its low sample consumption and fast analysis time makes it well-suited for high throughput screening techniques.[3, 4] The full-scan mass spectra can be used to evaluate binding stoichiometries and selectivity, while binding mode and structural information can be examined via tandem mass spectrometry techniques such as collisional activated dissociation (CAD).[5-7]

Much of the past work done in this area has focused on analyzing well-characterized drug/duplex DNA complexes, with promising results that indicate behavior in the gas-phase can be correlated to solution.[5-25] For example, Gabelica and co-workers demonstrated that binding stoichiometries and relative ion abundances observed in the mass spectra reflect known solution binding behavior.[6] Minor groove and intercalation binding modes of well-studied duplex-interactive drugs were distinguished by Wan and co-workers using collisional activated dissociation (CAD) experiments.[7]

Recent work by our group and others has extended the use of ESI-MS to evaluate non-covalent interactions of small molecules with G-quadruplex DNA,[18, 20, 23, 24, 26-28] a higher order nucleic acid structure being investigated as an anticancer drug target.[29, 30] G-

quadruplex DNA is formed from hydrogen bonding between a planar arrangement of four guanine nucleobases with the central cavity serving as a binding site for monovalent metal ions, the presence of which is generally required for the formation of the quadruplex. A variety of G-quadruplex structures can be formed from hydrogen bonding between guanines on up to four separate strands of oligodeoxynucleotides (ODNs).[29]

There is interest in studying G-quadruplex DNA as an anticancer target because of its role in telomere maintenance in cancer cells.[31, 32] Telomeres are the extremities of linear chromosomes that play an important role in protecting the chromosome ends from fusion and degradation and ensuring the complete replication of chromosomal DNA. Composed of G-rich sequences (TTAGGG is the human telomeric sequence), the single strand 3' overhangs of telomeres have the ability to form G-quadruplex structures. Ligands that selectively bind and stabilize the quadruplex structure of telomeric DNA may lead to the inhibition of telomerase, the enzyme responsible for the synthesis of telomeric DNA.[33] While telomerase is not active in most somatic cells, high levels of telomerase activity have been associated with most cancer cells.[34] Developing a drug that exhibits binding selectivity for G-quadruplex DNA over other DNA structures is also a primary focus of G-quadruplex-interactive drug design given the implications for reducing cytotoxic side-effects of anticancer agents.

One promising class of quadruplex selective-ligands are perylene diimides (PDIs).[35-43] Recent solution studies of N,N'-bis[2-(1-piperidino)-ethyl]-3,4,9,10-perylenetetracarboxylic acid diimide (PIPER)[35] and N,N'-bis(4-morpholinylpropyl)-3,4,9,10-perylenetetracarboxylic acid diimide (Tel01)[38] indicate that these molecules exhibit G-quadruplex affinity by stacking on the faces of the terminal G-tetrads, thereby stabilizing the G-quadruplex structure. We previously demonstrated the selectivity of Tel01 for quadruplex over duplex DNA based on ESI-MS studies in a quadrupole ion

trap mass spectrometer.[26] A more comprehensive series of PDI ligands, including those in which the perylene diimide chromophore is modified through substitution or benzannulation, (Figure 5.1) have been designed to further investigate the structural features most important for G-quadruplex DNA binding. In this study we demonstrate the utility of ESI-MS as a screening tool to characterize binding stoichiometry, distinguish between selective and non-selective binding, and determine the binding mode of these novel perylene diimide ligands.

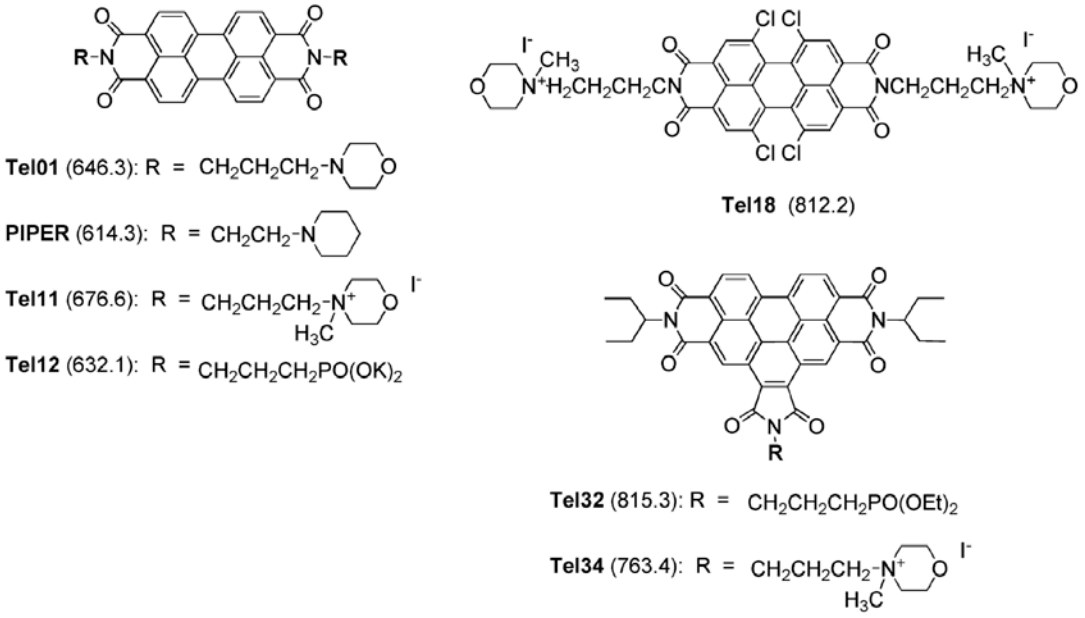


Figure 5.1: Structures of perylene diimide analogs. Molecular weights of compounds (Tel01, PIPER, Tel32) or organic ions (Tel11, Tel12, Tel18, Tel34) in Da are given in parenthesis.

5.2 EXPERIMENTAL

5.2.1 Chemicals

PIPER, Tel01, Tel11, Tel12, were prepared as described previously.[35, 37, 38] The manuscript detailing the synthesis of **Tel18, Tel32** and **Tel34** is in preparation. Ammonium salts of oligodeoxynucleotides (ODNs) custom synthesized on the 1.0 μmole scale with purification by RP-HPLC were obtained from TriLink Biotechnologies Inc. (San Diego, CA) and Integrated DNA Technologies (Coralville, IA) and used without further purification. The oligonucleotide d(TTTTTTTT) was synthesized on a 10 μM scale using a DNA synthesizer and purified by RP-HPLC. Collected fractions were combined and extensively dialyzed against deionized water before being lyophilized completely. A 2 mM stock solution of the G-quadruplex-forming ODN (d(TTGGGGGT)) was prepared in deionized water. A portion of the initial stock solution was diluted to 500 μM in deionized water and set aside for single strand experiments. The remaining solution was annealed by diluting in 150 mM ammonium acetate, heating to 90° C and slowly cooling to room temperature over a period of 7 hours. The self-complementary duplex-forming ODNs (d(GCGGGGATGGGGCG/CGCCCCATCCCCGC), d(GCGGGAATTGGGCG/CGCCCAATTCCCGC), and d(GCGGAAATTTGGCG/CGCCAAATTTCCGC)) were annealed in 250 mM ammonium acetate by heating to 90° C and slowly cooling to room temperature over a period of 2-3 hours.

5.2.2 Absorption Spectroscopy

Spectra were recorded on a UNICO model 2102 UV-spectrophotometer. Experiments were carried out in polystyrene cuvettes to minimize nonspecific binding of the ligands to the surface of the cuvettes. For DNA binding experiments, the absorption

spectra were obtained under high and low salt conditions. For the high salt conditions, the compound (20 µM) in 70 mM potassium phosphate, 100 mM potassium chloride, 1 mM EDTA buffer (pH 7) was analyzed alone or in the presence 20 µM structure of G4-DNA $[d(TAGGGTTA)]_4$, G4'-DNA $[d(TTAGGG)_4]$ double-strand DNA $[d(CGCGCGATATCGCGCG)]_2$, or single-stranded DNA d(TTTTTTTT). For the low salt conditions, the compound (10 µM) in 3:1 25 mM ammonium acetate/methanol was analyzed alone or in the presence of 10 µM structure of G4-DNA $[d(TTGGGGGT)]_4$, double-stranded DNA $[d(GCAAATTTCG)]_2$ or single-stranded DNA d(TTTTTTTT). Samples were monitored until equilibrium was achieved, as evidenced by constant absorbance readings.

5.2.3 Fluorescence Spectroscopy and Resonance Light Scattering

Spectra were recorded on a Hitachi model F-2000 spectrofluorometer. Quartz cuvettes were treated with SigmaCote for 1 hour followed by extensive washing with water to minimize nonspecific binding of the ligands to the surface of the cuvette. Fluorescence and resonance light scattering (RLS) measurements were carried out on the ligands under both high- and low-salt conditions. The high salt conditions employed solutions of compound in 70 mM potassium phosphate, 100 mM potassium chloride, 1 mM EDTA buffer at the indicated pH. Low-salt conditions employed solutions of compound in 3:1 25 mM ammonium acetate/methanol at the indicated pH. Solutions were allowed to equilibrate for the specified time before scans were performed at 25 °C using the noted excitation and emission wavelengths.

5.2.4 Mass Spectrometry

Stock solutions of the ligands were prepared in either 0.1% aqueous trifluoroacetic acid (PIPER and Tel01), deionized water (Tel11, Tel12, and Tel18), acetonitrile (Tel32) or dimethylsulfoxide (Tel34) and stored at room temperature.

Analytical solutions were prepared containing G-quadruplex, duplex, or single strand DNA and one ligand each at equimolar 10 μM concentrations in 3:1 25 mM ammonium acetate/methanol solvent. A Harvard syringe pump (Holliston, MA) set at a flow rate of 3 μL/min was used to directly infuse the sample solutions into a ThermoFinnigan LCQ Duo mass spectrometer (San Jose, CA). The ESI source was operated in the negative ion mode with an electrospray voltage of 3.5 kV and a heated capillary temperature of 90 to 120° C. To aid in desolvation, nitrogen sheath and auxiliary gas were set at 40 and 10 arbitrary units, respectively. The base pressure of the trap was ~1 x 10^{-5} Torr. Spectra were acquired by summing 300 scans, with an ionization time ranging from 100 to 250 ms.

Collisional activated dissociation (CAD) experiments were performed on selected complexes by isolating the desired precursor ion in the ion trap using resonance ejection, followed by fragmentation promoted by increasing the collisional energy applied to the trap (reported as a percentage of 5 $V_{0\text{-}p}$) until the intensity of the precursor ion was reduced to approximated 10 % of its original intensity. An activation time of 30 ms was used in all CAD experiments.

5.3 RESULTS AND DISCUSSION

5.3.1 Fluorescence and Resonance Light Scattering Studies of PDI Aggregation

PDIs like Tel01 and PIPER are known to undergo pH-dependent self-association to form aggregates in solution.[37-39] The selectivity of PDIs towards G-quadruplex DNA depends on the aggregation state of these ligands, which, in addition to pH, may also be affected by ligand concentration, buffer ionic strength, and the presence of non-aqueous co-solvents.[36, 38] Previous solution studies of the aggregation state and G-quadruplex DNA binding selectivity of PIPER,[36,38] Tel01,[36,38] Tel11,[37] and Tel12[37] were carried out under relatively high-salt (170 mM KCl) phosphate buffer. In order to determine the

aggregation state of these and the other PDI ligands under the conditions of the ESI-MS analysis discussed later, the fluorescence and resonance light scattering spectra were determined in 3:1 25 mM ammonium acetate/methanol.

As reported previously, the intense fluorescence of the monomeric PDIs is quenched upon ligand self-association in solution.[36-38] Solutions of PIPER (1 μM) in methanolic ammonium acetate buffer display a pronounced fluorescence emission at 550 nm whereas solutions of Tel01 do not (spectra not shown). For aggregates in which there is good overlap between adjacent monomer chromophores, a large increase in the intensity of scattered light is observed when the incident light is the same wavelength as the aggregate absorbance.[44] This resonance light scattering (RLS) signal has been used as a qualitative measure of the presence of PDI aggregates in solution.[36-38] Solutions of Tel01 display a large RLS signal; however, the RLS signal for PIPER is absent (data not shown). The absence of strong fluorescence and the presence of a strong RLS signal for solutions of Tel01 indicate that this PDI is aggregated under these conditions. In contrast, the lack of an RLS signal and the presence of strong fluorescence for solutions of PIPER indicate that this PDI exists in the monomeric state. Similar observations were made for solution of Tel01 and PIPER in high-salt buffer conditions, indicating that the aggregation state of these two PDIs is largely unaffected by the methanolic buffer conditions employed in the ESI-MS analysis.

The fluorescence and RLS spectra of the other PDIs were also examined under the methanolic buffer conditions. As shown in Table 5.1, PIPER, Tel11, and Tel18 are not aggregated in solution, as judged by the strong fluorescence and lack of RLS signals of these solutions. The PDI Tel12 displays only weak fluorescence and no RLS spectra. The weak fluorescence of Tel12 may be due to ligand dimerization, as has been noted for this PDI in 150 mM KCl phosphate buffer;[37] however, the fluorescence intensity of Tel12

in methanolic buffer is higher than that in the high-salt buffer (data not shown), indicating that dimerization is less extensive in the methanolic buffer. Because Tel12 has been found to dimerize, but does not produce a RLS spectrum, this ligand self-associates. Self-association refers to the formation of dimers, as is the case for Tel12, as well as high-order species and aggregates, which we functionally define as PDI species that can be observed by resonance-light scattering.

Table 5.1: Resonance light scattering and fluorescence data for PDIs.

Ligand	Relative Intrinsic Fluorescence[a]	Relative RLS Signal[b]
TEL01	0.649	36
PIPER	25.31	ns[c]
TEL11	25.83	ns[c]
TEL12	12.74	ns[c]
TEL18	41.82	ns[c]
TEL32	0.189	6
TEL34	0.524	5

[a] Relative intensity of fluorescence emission (550 nm) from a 1 μM solution of PDI in 3:1 25 mM ammonium acetate /methanol buffer, pH 7. Excitation was at 520 nm for all ligands except Tel32 and Tel24, which were excited at 495 nm. [b] Relative intensity of the resonance light scattering peak at 470 nm for solutions of PDI (1 μM) in 3:1 25 mM ammonium acetate /methanol buffer, pH 7. [c] No significant RLS signal observed.

Like Tel01, solutions of the benzannulated PDI Tel34 in methanolic buffer do not fluoresce; however, unlike Tel01, Tel34 solutions do not display a strong RLS signal,

indicating that if Tel34 is aggregated in solution, the aggregates are RLS-silent. Attempts to compare the aggregation state of Tel34 in methanolic ammonium acetate and 150 mM KCl phosphate buffer were unsuccessful due to the very low solubility of this PDI in the high-salt buffer. Similarly, the low solubility of Tel32, even in methanolic ammonium acetate, made characterization of this PDI difficult. Immediately prepared solutions of Tel32 showed very low fluorescence and RLS signals; however, over a period of minutes the ligand began to precipitate from these solutions.

5.3.2 DNA Binding Studies by UV-Vis Absorption Spectroscopy

The UV-Vis absorption spectra of the PDIs (10 μM in 3:1 25 mM ammonium acetate/methanol) were determined in the absence and presence of 10 μM of G-quadruplex DNA, duplex DNA, or single-stranded DNA. The parallel stranded G-quadruplex $[d(T_2G_5T)]_4$ (G5), which is similar to the well-characterized $[d(T_2G_4T)]_4$, was employed for both these solution binding studies as well as the ESI-MS analysis. An advantage of this parallel-stranded G-quadruplex over intramolecular G-quadruplexes is the ease with which the G-quadruplex form can be distinguished from single-stranded DNA in the ESI-MS spectra (see below). The duplex DNA formed by the self-complementary ODN d(GCAAATTTCG) and single-stranded $d(T_8)$ were also used in these solution binding studies.

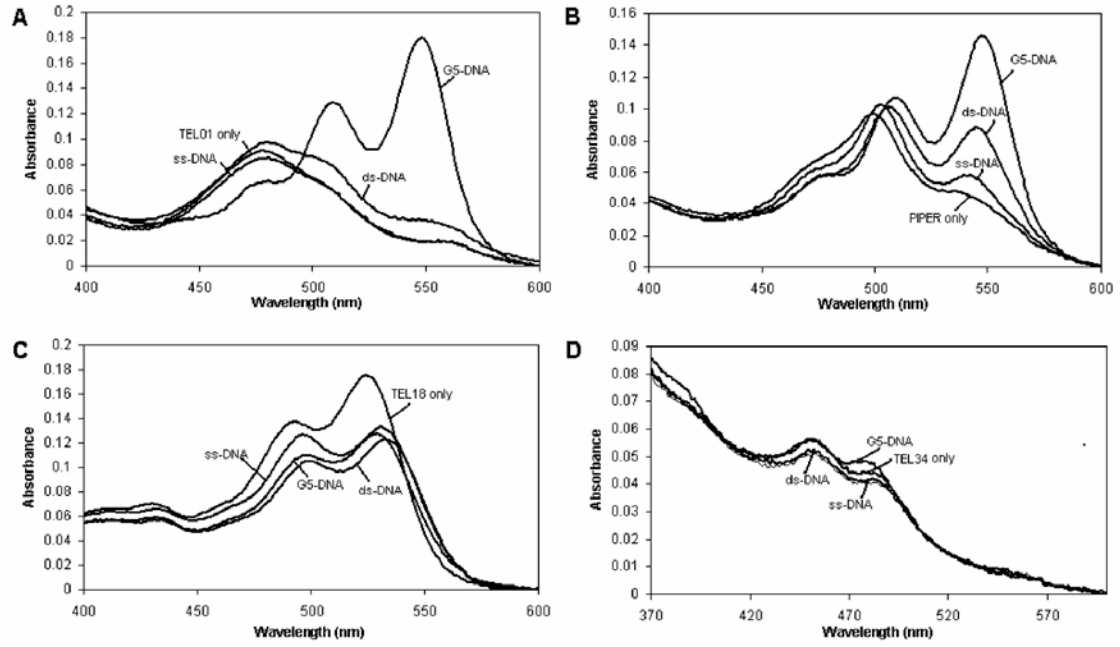


Figure 5.2: UV-Vis absorbance spectra of 10 µM Tel01(A), PIPER (B), Tel18 (C), and Tel34 (D) alone and in the presence of 10 µM G5-DNA $[d(T_2GT)]_4$, ds-DNA $[d(GCAAATTTCG)]_2$, or ss-DNA $[d(T_8)]$ in 3:1 25 mM ammonium acetate/methanol, pH 7.

As shown in Figure 5.2A, the UV-Vis spectrum of Tel01 shows an absorbance peak at 470 nm. Similar absorbance spectra are recorded for solutions of Tel01 in the presence of equimolar duplex or single-stranded DNA; however, in the presence of equimolar G-quadruplex DNA, the absorbance spectrum of Tel01 is remarkably different, consisting of two large peaks at 550 and 510 nm. The UV-Vis spectrum of PIPER displays an absorbance peak at 500 nm, which shifts to 510 nm in the presence of equimolar single-stranded, double-stranded, or G-quadruplex DNA (Figure 5.2B). Additionally, in the presence of each of these DNA structures, there is a new absorbance from the PIPER chromophore at 550 nm. The intensity of this long-wavelength band

increases in the order single-stranded DNA < double-stranded DNA < G-quadruplex DNA.

These UV-Vis spectral changes for Tel01 and PIPER indicate a difference in the G-quadruplex DNA binding selectivity of these two ligands.[37,38] Tel01, whose absorbance spectrum in the absence of DNA is blue-shifted due to ligand aggregation, forms a complex with G-quadruplex DNA that absorbs in the 550 nm region. PIPER, whose spectrum in the absence of DNA does not indicate ligand aggregation, forms similar long-wavelength absorbing complexes with G-quadruplex, duplex, and even single-stranded DNA (Figure 5.2B).

Other PDIs shown in Figure 5.1 also undergo UV-Vis spectral changes in the presence of G-quadruplex DNA. Tel11 and Tel12, whose absorbance spectra in the absence of DNA is similar to that of PIPER, also form complexes with G-quadruplex DNA characterized by absorbance peaks at 550 nm (data not shown). Tel11, but not Tel12, also gives rise to a peak at 550 nm in the presence of double-stranded DNA. As expected, the UV-Vis absorbance spectrum of the chromophore-modified PDI Tel18 is different from that of PIPER and the other PDIs. The Tel18 spectrum, consisting of a peak at 525 nm with a shoulder at 492 nm, undergoes hypochromic and bathochromic shifts in the presence of G-quadruplex DNA (Figure 5.2C). Similar changes are also observed in the presence of duplex DNA, and to a lesser extent, single-stranded DNA. The absorbance spectrum of Tel34 includes a long wavelength peak at 450 nm with a shoulder at 490 nm (Figure 5.2D). In the presence of G-quadruplex DNA, there is a shift of the 490 nm absorbance to 480 nm, accompanied by slight hyperchromism. Immediately after the addition of double-stranded or single-stranded DNA, there is a slight hypochromism in the Tel34 absorbance spectrum, but no appreciable changes in the position of the absorbance peaks (Figure 5.2D). Over a period of an hour, solutions

of Tel34 containing duplex or single-stranded DNA exhibit a pronounced decrease in UV-Vis absorbance that is accompanied by the formation of insoluble material. This behavior is not observed in the presence of G5 DNA, presumably because the complex formed between this benzannulated PDI and the G-quadruplex DNA is more soluble in the buffer than the ligand itself. The benzannulated and uncharged PDI Tel32 is also relatively insoluble, and at the concentration employed for these absorbance studies (10 μM), the ligand precipitated from solution over the course of a few minutes.

5.3.3 Fluorescence Quenching Studies

It has previously been shown that PDI fluorescence is quenched upon ligand binding to DNA structures.[36-38] Fluorescence quenching experiments of the PDIs with quadruplex, duplex, and single-stranded DNA allow better quantification and insight into the affinity and selectivity differences of these ligands for G-quadruplex DNA. Solutions of 1 μM PDI in 3:1 25 mM ammonium acetate/methanol buffer, pH 7 were titrated with stock solutions of each DNA structure in the same buffer. Fluorescence spectra were recorded at appropriate emission and excitation wavelengths for each class of chromophore: unmodified PDIs Tel01, Tel12, Tel12, Tel34, and PIPER (emission at 545 nm, excitation at 495 nm); chlorinated PDI Tel18 (emission at 555 nm, excitation at 495 nm); and benzannulated PDI Tel32 (emission at 565 nm, excitation at 469 nm (data not shown). While all three DNA samples quench the fluorescence of Tel18, there is clearly a difference between the efficiency of fluorescence quenching, with G5-DNA causing the most extensive quenching at low concentrations, single-stranded DNA producing the least quenching, and duplex DNA leading to intermediate quenching. In the case of G-quadruplex DNA, there is nearly 80% fluorescence quenching after the addition of just one-half equivalent of DNA, indicating that multiple Tel18 ligands are binding to the DNA.

The results of the fluorescence quenching studies of the PDIs are summarized in Table 5.2, which shows the percent PDI fluorescence quenching after the addition of one equivalent of the different DNA structures. In accord with the UV-Vis absorbance studies (Figure 5.2A), Tel01 displays a high degree of selectivity for G-quadruplex DNA binding, as shown by the minimal fluorescence quenching in the presence of either duplex or single-stranded DNA compared to G-quadruplex DNA.

Other PDIs, such as PIPER and Tel11, appear to interact more strongly with G-quadruplex DNA when compared to Tel01, but show very little selectivity for G-quadruplex DNA versus duplex DNA. Both PIPER and Tel18 also show significant interactions with single-stranded DNA; the addition of one equivalent of single-stranded DNA to these PDIs quenches their fluorescence by more than 50%. Tel18 and Tel34 appear to be of intermediate selectivity for G-quadruplex DNA versus duplex DNA and have moderate affinity for single-stranded DNA. Tel12 does not interact strongly with G-quadruplex DNA, as evidenced by only 29% fluorescence quenching in the presence of one equivalent of G5-DNA; however, this PDI is relatively selective for G-quadruplex DNA and does not bind to single-stranded DNA. There was insignificant fluorescence quenching of Tel32, indicating that this PDI does not interact with any of these DNA structures.

Table 5.2: Fluorescence quenching of PDIs by quadruplex, duplex, and single-stranded DNA.

Ligand	Percent Fluorescence Quenching in the presence of equimolar DNA[a]		
	G5-DNA[b]	ds-DNA[c]	ss-DNA[d]
TEL01[e]	64 ± 4%	6 ± 2 %	2 ± 1%
PIPER[e]	100 ± 1 %	94 ± 2 %	57 ± 4 %
TEL11[e]	99.8 ± 0.1 %	96 ± 1 %	67 ± 5 %
TEL12[e]	29 ± 6 %	16 ± 8%	13 ± 6 %
TEL18[f]	92.9 ± 0.1 %	66 ± 3%	12 ± 9 %
TEL32[e]	4 ± 2 %	1 ± 2%	7 ± 2 %
TEL34[g]	90 ± 2 %	62 ± 4%	40 ± 6 %

[a]Fluorescence emission intensity of solutions of 1 µM PDI and 1µM DNA in 3:1 25mM ammonium acetate /methanol buffer, pH 7, expressed as a percent of emission intensity for solutions of the PDI alone. [b] $[d(T_2G_5T)]_4$. [c] $[d(GCAAATTTCG)]_2$. [d] $[d(T_8)]$. [e]Excitation = 495 nm, emission = 545 nm. [f]Excitation = 495 nm, emission = 555 nm. [g]Excitation = 469 nm, emission = 565 nm.

5.3.4 Binding Stoichiometry of Perylene Diimides by Electrospray Ionization Mass Spectrometry

To begin an ESI-MS investigation of the binding selectivity of the perylene diimide ligands, the complexation of each ligand with G-quadruplex DNA was evaluated. The intermolecular, parallel stranded G-quadruplex $[d(T_2G_5T)]_4$ (G5) was used in this study. Rosu *et al*. were the first to demonstrate that G-quadruplex DNA can be annealed in an ammonium acetate buffer and analyzed by ESI-MS.[27] When ammonium acetate is used, the G-quadruplex structure is stabilized by NH_4^+ counterions, rather than by Na^+ or

K^+, resulting in cleaner spectra as a result of the greater lability of the associated counterions during the ESI process. The ESI mass spectrum of the G-quadruplex without any added ligands (Figure 5.3A) demonstrates that the G-quadruplex species can readily be detected in the -6 charge state, and to a lesser degree in the -7 charge state with two to four ammonium adducts associated with each quadruplex. It is likely that the ammonium ions are bound in the central cavity of the G-tetrad, stabilizing the quadruplex. For visual simplicity, the NH_4^+ ions are not labeled in the remaining figures.

Solutions containing the G-quadruplex and one of each of the five non-benzannulated ligands (Tel01, Tel11, Tel12, Tel18, and PIPER) at equimolar concentrations of 10 μM in a 3:1 25 mM ammonium acetate/methanol solution exhibited a variety of binding stoichiometries as demonstrated by the mass spectra shown in Figure 5.3. For the solution containing Tel18 (Figure 5.3B), complexes with Tel18/G5 binding stoichiometries of 1:1 and 2:1 are present with greater abundances than the free G-quadruplex ions. Other ligands formed less abundant complexes with the G-quadruplex. This is apparent in the spectrum of G5 with Tel12 (Figure 5.3C) in which only 1:1 complexes are present at lower abundances than the unbound G-quadruplex ions.

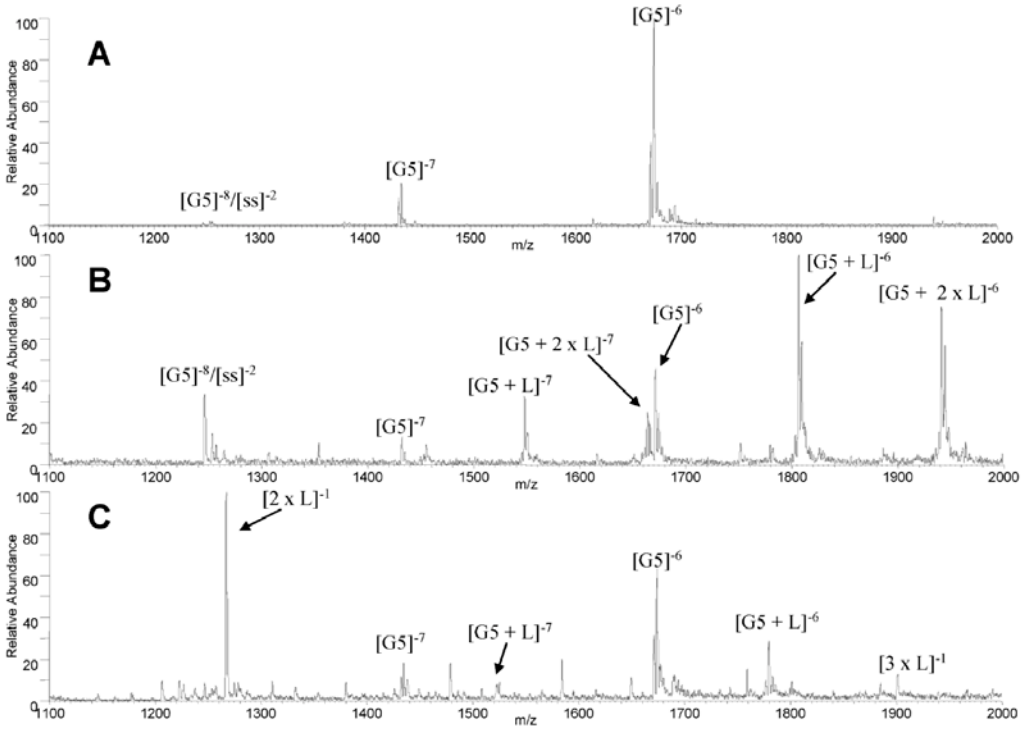


Figure 5.3: ESI mass spectra for complexes containing $[d(T_2G_5T)]_4$ (G5) (A) alone, and with (B) Tel18, and (C) Tel12. Complexes containing one or more ligands are labeled with a "L".

With the exception of Tel01 and Tel12, 1:1 and 2:1 binding stoichiometries were consistently observed. This result is in agreement with a previous NMR-based study of the perylene diimide analog PIPER that found 1:1 and 2:1 complexes were typically formed in solution by end-stacking on the faces of the terminal G-tetrads.[35] The 3:1 complexes present in the ESI mass spectrum of Tel01 with G5 may be the result of ligand aggregation, as discussed above. Recent x-ray crystallographic studies of an anthraquinone bound to a parallel-stranded G-quadruplex DNA show three ligands bound to a single G-tetrad face.[45] Presumably, a similar sort of self-association of multiple PDI ligands on G-tetrad faces could occur for Tel01, which undergoes aggregation in solution even in the absence of G-quadruplex DNA. Ligand aggregation is not observed in

solution for the analogs with charged side chains (Tel11, Tel12, and Tel18). Aggregation is less expected for these ligands due to charge repulsion, which is especially true in the gas-phase, thus explaining why aggregation was not observed with those ligands. Charge repulsion may also explain why only 1:1 complexes were observed for Tel12. As the only ligand with negatively charged side chains, Tel12 may experience coulombic repulsion with the anionic backbone of the DNA in solution and the gas-phase, preventing the binding of multiple ligands. In general, the predominant 1:1 and 2:1 binding stoichiometries observed for the ligands with G5 DNA is expected based on an end-stacking binding mode.

5.3.5 Binding Stoichiometry of Benzannulated Ligands by ESI-MS

The interactions of G-quadruplex DNA with a series of benzannulated perylene diimides (Figure 5.1) with an extended chromophore, Tel32 and Tel34, were also examined by ESI-MS. The benzannulated analogs may engage in increased end-stacking interactions with the G-tetrad as a result of the larger conjugated ring system. The ESI-MS results for this set of analogs (spectra not shown) indicate that Tel34 forms complexes with the quadruplex DNA, but only with binding stoichiometries of 1:1. Interestingly, Tel11 and Tel18 have the same side chain as Tel34, but both formed complexes with higher binding stoichiometries than Tel34, indicating that the benzannulated analogs may interact differently with the G-quadruplex. The larger chromophore of the benzannulated analogs may prevent multiple ligands from binding to the G5 structure due to steric hindrance. While Tel34 formed complexes with G-quadruplex DNA, the other benzannulated analog, Tel32, was the only perylene diimide ligand that did not form any complexes with G5 DNA.

5.3.6 Concentration Effects

To examine how changes in the solution conditions affect the binding stoichiometries of the ligand/G5 DNA complexes, solutions containing 10 μM of G5 DNA and either 2.5, 5.0, 10, and 20 μM of each PDI ligand were analyzed by ESI-MS (spectra not shown). The binding stoichiometries of the analogs containing charged side chains (Tel11, Tel12, Tel18, and Tel34), did not change from those observed at a 1:1 ligand/G5 ratio with either higher or lower ligand/G5 molar ratios. For example, when the Tel12/G5 molar ratio was 1:4, only complexes with a 1:1 binding stoichiometry were present in the ESI mass spectra. When the Tel12/G5 molar ratio was increased from 1:4 to 2:1, the relative abundances of the 1:1 peaks remained the same and no 2:1 complexes appeared.

Ligands containing basic side chains (Tel01 and PIPER) do show concentration dependent binding. At ligand/DNA molar ratios of 1:4 to 1:1, the complexes with binding stoichiometries of 1:1 are the most abundant, with 2:1 complexes present at relative abundances that are approximately three times less than that observed for the 1:1 complexes. When the ligand/DNA ratio is increased above 1:1, the 2:1 complexes become nearly twice as abundant as the 1:1 complexes, and very low abundance 3:1 complexes emerge that were not present at lower ligand/DNA ratios. The concentration dependent binding of Tel01 was demonstrated in a past study by David et al.[26] and is confirmed by experiments in this study. These results indicate that while all of the perylene diimides bind G-quadruplex DNA at primarily 1:1 and 2:1 binding stoichiometries, ligands with side chains that are basic, Tel01 and PIPER, demonstrate the ability to bind with higher stoichiometries at increased ligand/G5 molar ratios. Both Tel01 and PIPER can undergo aggregation in solution in a pH-dependent fashion.[37, 38] The observation of higher binding stoichiometries for these two ligands may be due to ligand

self-association during ESI, as supported by the RLS and fluorescence data discussed above and summarized in Table 5.1.

5.3.7 ESI-MS/MS Studies of G5 DNA/Perylene Diimide Complexes

ESI-MS/MS experiments were undertaken to examine the fragmentation patterns of ligand/G-quadruplex complexes via collisional activated dissociation. While full-scan ESI mass spectra are useful for comparing the binding stoichiometries of the various ligands with DNA, ligand binding modes are generally not distinguishable. Past ESI-MS/MS studies of drug/DNA complexes suggest that complexes with different binding modes produce distinct fragmentation patterns upon collisional activation.[6, 7, 26] In this study, all complexes, including $[G5 + L]^{-6}$, $[G5 + 2 \times L]^{-6}$, $[G5 + 3 \times L]^{-6}$, $[G5 + L]^{-7}$, and $[G5 + 2 \times L]^{-7}$, where L represents a perylene diimide ligand, were subjected to CAD experiments. All complexes, regardless of the charge of the ligand, the charge state of the complex, or the binding stoichiometry, dissociated via guanine base losses from the precursor ions in addition to the ejection of one anionic or neutral ligand, leaving the intact quadruplex. Triplex and single strand oligonucleotide ions without bound ligand are also present in the spectra as a result of further dissociation of the quadruplex. Ligands Tel12, Tel01, and PIPER were ejected as anions, while Tel11, Tel18, and Tel34 were ejected as neutral species. An example of this fragmentation pattern is shown in Figure 5.4 for the CAD spectrum of the $[G5 + Tel012]^{-6}$ complex. In addition to ions resulting from guanine base loss, the -5 charge state of the intact quadruplex and the -2 charge state of the single strand species are present. Presumably the -3 charge state of the triplex is also formed, however the *m/z* of this ion lies outside the detectable mass range of our instrument.

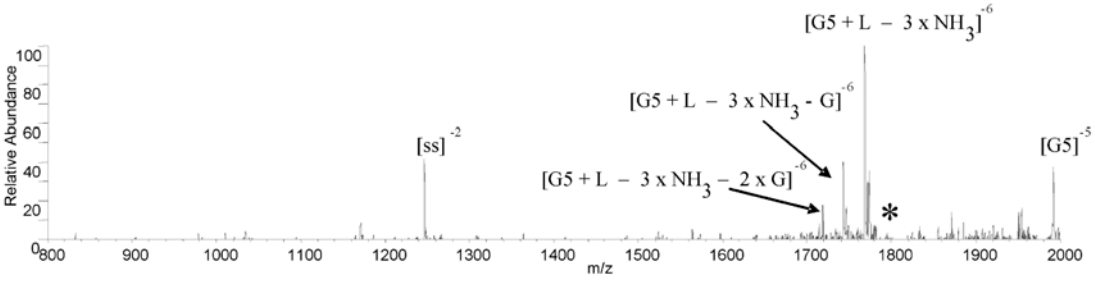


Figure 5.4: CAD product ion spectrum $[G5 + Tel12]^{-6}$. An activation energy of 1.98 V was used. Complexes containing one or more ligands are labeled with a "L". Parent ions are labeled with an astrisk.

The facile loss of ligand observed in the CAD spectra of the DNA/perylene diimide complexes is consistent with past studies of ligands that bind to G-quadruplex DNA by end-stacking[26,27] suggesting the perylene diimides bind to quadruplex DNA in a similar manner. Previously published results by our group demonstrated that complexes containing distamycin A and diethyloxadicarbocyanine (DTC), ligands that presumably bind to G-quadruplex DNA via groove binding,[46,47] produce single stranded, double stranded, and triplex ODNs with and without bound ligand upon CAD.[26] It is important to note that both distamycin A and DTC are cationic, while the postulated quadruplex end-stacking ligands involved in a past ESI-MS/MS study[26] were neutral. It has been difficult to determine whether the differences between the fragmentation patterns of the end-stacking and the groove-binding quadruplex-interactive ligands were the result of different binding modes and/or ligand charge. However, in this study, all analogs, including the cationic Tel11, Tel18, and Tel34 ligands dissociated via the ejection of ligand from the intact G-quadruplex, suggesting that the charge of the ligand has less of an impact on the dissociation pathway than binding mode.

5.3.8 Duplex and single strand DNA binding of ligands by ESI-MS

The evaluation of duplex versus quadruplex binding selectivity is a critical issue in the design of new anti-cancer drugs. To evaluate selectivity, ESI mass spectra of solutions containing the perylene diimide ligands with duplex DNA (abbreviated as ds) have also been obtained to allow comparison of the selectivity of the ligands for quadruplex versus duplex versus single strand ODNs. To account for potential sequence selectivity, a series of three non-self-complementary duplex ODNs with varying G-C and A-T base pair composition were used (d(GCGGGGATGGGGCG/CGCCCCATCCCCGC), d(GCGGGAATTGGGCG/CGCCCAATTCCCGC), and d(GCGGAAATTTGGCG/CGCCAAATTTCCGC)). As shown in Figure 5.5, there is considerable variation in duplex versus quadruplex selectivity. Some ligands, such as Tel18, formed numerous complexes with the duplex DNA (Figure 5.5B), while others formed no complexes, as demonstrated by the spectrum containing Tel34 (Figure 5.5C). The increased number of adducts, especially for the single strand ions, present in Figure 5.5C are due to higher levels of salts present in the Tel034 ligand sample. The ligands that did bind to the duplex DNA, namely PIPER, Tel11, and Tel18, formed complexes with binding stoichiometries of 4:1, 3:1, 2:1, and 1:1, and did not demonstrate any notable base-pair selectivity.

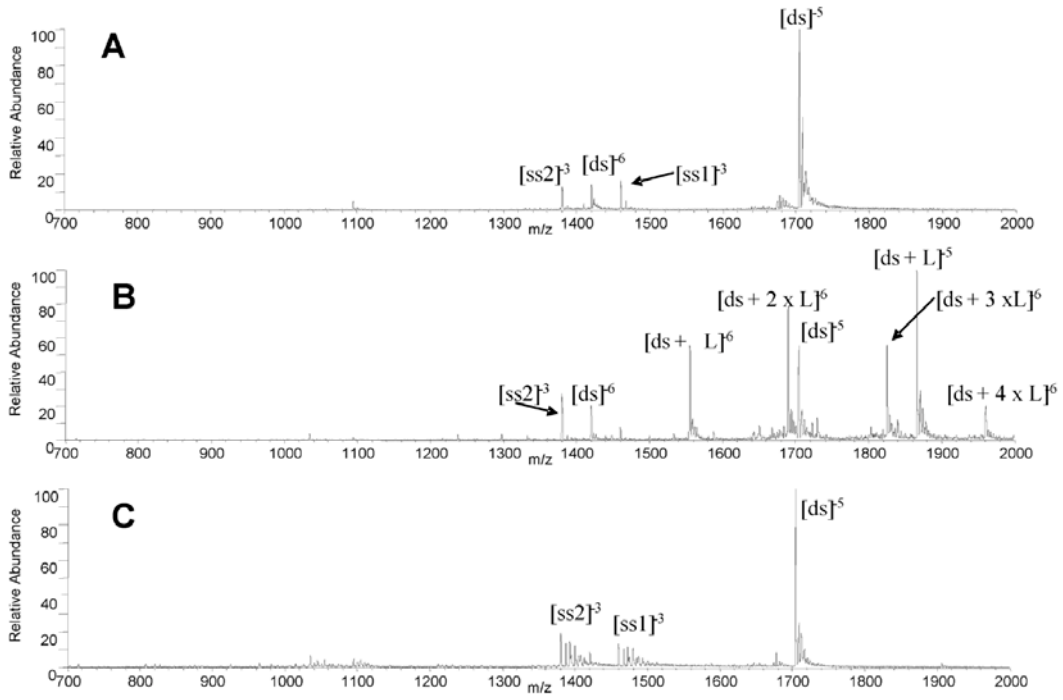


Figure 5.5: ESI mass spectra for complexes containing d(GCGGGAATTGGGCG/CGCCCAATTCCCGC): (A) alone, and with (B) Tel18 (extensive binding) and (C) Tel34 (no binding). The d(GCGGAATTGGGCG) single strand is labeled "ss1" and d(CGCCCAATTCCCGC) is labeled "ss2."

To further examine the interaction between the ligands and duplex DNA, CAD experiments were performed on all ligand/DNA complexes present in the mass spectra. The complexes subjected to CAD included $[ds + L]^{-5}$, $[ds + L]^{-6}$, $[ds + 2 \times L]^{-6}$, $[ds + 3 \times L]^{-6}$, $[ds + 4 \times L]^{-6}$, where "L" represents either PIPER, Tel11, or Tel18. All complexes produced the same charge-state dependent CAD pattern, regardless of the ligand or binding stoichiometry. At lower charge states, such as -5, the complexes dissociated by loss of nucleobases (specifically A, G, and C) from the intact complex, while at higher

charge states, strand scission was observed, with one or more ligands remaining bound to the single strand ODN.

The complexes containing PIPER, the only analog without cationic side chains that bound to duplex DNA, exhibited an additional fragmentation route. At low charge states, neutral ligand ejection was observed (a new route), leaving the duplex ion, in addition to base losses from the complex. These two processes are competitive since both types of fragments are formed from the same parent ion. The fragmentation pattern produced by the PDI/duplex DNA complexes, characterized by both non-covalent strand scission and ligand ejection, in addition to covalent cleavages of high proton affinity nucleobases from the complex, is similar to the results of past MS/MS studies of intercalator/duplex DNA complexes.[5,7] In quadrupole ion trap instruments, complexes containing minor groove binding ligands also undergo strand scission and cleavage of nucleobases, and covalent cleavages of the terminal portions of the duplex resulting in losses of a_n and (d_n + B) ions are also observed upon CAD.[7] The loss of these small sequence ions is never observed in the CAD spectra of the PDI/duplex complexes. Intercalation has been suggested for perylene diimide molecules that bind to duplex DNA,[37, 38] and would be expected with a duplex DNA-interactive molecule containing a planar chromophore. Future work is necessary to examine the correlation between the fragmentation patterns of the perylene diimide/duplex complexes and ligand binding mode.

ESI mass spectra of solutions containing one ligand and a single strand ODN were also acquired. The single strand sequence used to anneal quadruplex DNA, $d(T_2G_5T)$, was used for these experiments to determine whether the analogs that selectively bound to the G-quadruplex DNA were selective for the G-quadruplex structure or for the sequence of DNA. To ensure that there were no Hoogsteen base pair

interactions between the guanine residues of the single strand ODN, ESI mass spectra were first obtained for the ODN alone (spectra not shown). The mass spectrum shows that the most abundant ions are found at *m/z* 831 and 1247, corresponding to the $[ss]^{-3}$ and $[ss]^{-2}$ species respectively. The appearance of this mass spectrum is significantly different than the one obtained for the annealed G-quadruplex (Figure 5.3A).

Solutions containing the single strand ODN and each ligand were prepared at equimolar 10 μM concentration in 3:1 water/methanol. The ESI mass spectra show that the two ligands with cationic side chains that formed complexes with duplex DNA and quadruplex DNA, Tel11 and Tel18, also formed complexes with the single strand ODN, while the remaining ligands exhibited no single strand binding (spectra not shown). These results suggest that Tel11 and Tel18 indiscriminately bind to DNA, regardless of the higher order structure. This lack of selectivity may be the result of electrostatic interactions between the cationic side chains of the analogs and the phosphate backbone of DNA. However, Tel34 also has a cationic side chain but was only found to bind to G-quadruplex DNA, suggesting that the ligand/DNA complexes observed in the ESI mass spectra are not merely the result of non-specific electrostatic interactions in the gas phase.

5.3.9 ESI-MS evaluation of binding selectivity of ligands

After evaluating the binding of the ligands to G-quadruplex, duplex, and single strand DNA by ESI-MS/MS, the overall selectivity of the analogs is summarized in Table 5.3. The values in the table represent the fraction of bound DNA calculated using the following equation:

$$\text{Fraction of bound DNA} = \frac{I_{(1:1)} + I_{(2:1)} + I_{(3:1)} + I_{(4:1)}}{I_{(DNA)} + I_{(1:1)} + I_{(2:1)} + I_{(3:1)} + I_{(4:1)}}$$

where I_{DNA} is the relative abundance of the free (unbound DNA) and $I_{(n:1)}$ are the relative abundances of the ligand/DNA complexes. This equation is a modified version of an equation developed by Rosu et. al.[17] that was used to estimate the concentration of all individual species at equilibrium based on the relative abundances of DNA ions in the mass spectrum, with the assumption that the abundances of the free and bound DNA ions are proportional to their relative concentrations in solution. We are not convinced that the ESI response factors of the unbound DNA ions and the DNA/perylene diimide complexes are identical, so instead we calculate the fraction of bound DNA to obtain a relative comparison of the extent of ligand binding to the different DNA structures.

Table 5.3: Summary of ligand binding selectivity[a]

Ligand	PIPER	Tel01	Tel11	Tel12	Tel18	Tel34	Tel32
G5 Binding	0.95	0.65	0.95	0.30	0.75	0.30	–
Duplex Binding	0.80	0.20	0.95	--	0.90	--	--
Single Strand Binding	--	--	0.10	--	0.10	--	--
Selectivity for G-Quadruplex	poor	good	poor	good	poor	good	poor

[a] Values in the chart represent the fraction of bound DNA. "--" indicates no complexes were observed. All values +/-0.05.

While PIPER, Tel11, and Tel18 readily formed complexes with the quadruplex DNA, these three analogs also bound to duplex DNA which is undesirable for target selectivity. Both Tel11 and Tel18 were also found to bind to single strand DNA,

suggesting that the cationic side chains of the ligands may promote indiscriminant binding to DNA due to electrostatic interactions with the anionic phosphate backbone. Interestingly, Tel34 has a cationic side chain but did not bind to duplex or single strand DNA, thus exhibiting significant selectivity. This suggests the side-chains of the benzannulated analogs have a different role in DNA binding, which may be a key issue for G-quadruplex selectivity.

The only analog with anionic side chains, Tel12, formed 1:1 complexes with the G5 DNA, but did not bind to duplex or single strand DNA. Whether the anionic side chains play a role in the selectivity of the analog is still unclear, but the binding of the ligand to the G-quadruplex suggests that the complexes observed in the ESI mass spectra are not solely the result of electrostatic interactions. Finally, PIPER was found to bind to G-quadruplex and duplex DNA, while the structurally-similar ligand, Tel01, demonstrated selectivity for the G-quadruplex.

There is good correlation between these ESI-MS results and the spectroscopic studies of G-quadruplex DNA binding by these PDIs (see above). Both the UV-Vis spectroscopic studies and ESI-MS analysis demonstrate that all of the PDIs examined here, with the exception of Tel32, bind to parallel-stranded G-quadruplex DNA. Spectroscopic and ESI-MS studies also both reveal the selectivity of certain PDIs (e.g., Tel01) for G-quadruplex versus duplex DNA and the lack of selectivity of other PDIs (e.g., Tel11) for any of the DNA structures studied here. Additional insights into the origin of the G-quadruplex binding selectivity observed for certain PDIs is provided by ESI-MS. Of the PDIs studied, three were found to undergo extensive self-association (Tel12, Tel34) or aggregation (Tel01) in solution. All three of these ligands display selectivity for G-quadruplex DNA. However, ESI-MS reveals that the stoichiometries of G-quadruplex DNA binding by these ligands are different. Tel01 binds to G-quadruplex

DNA to form complexes with as many as three ligands associated with a single quadruplex. Tel12 and Tel34 only form 1:1 complexes with G-quadruplex DNA. Non-selective PDIs such as PIPER and Tel11, which are not aggregated or highly self-associated in solution, also form 1:1 and 2:1 ligand/DNA complexes with G-quadruplex DNA. Clearly, the selectivity of PDIs for G-quadruplex DNA binding is correlated with the self-association state of these ligands in solution and not their self-association on G-quadruplex DNA.

5.4 CONCLUSIONS

The results of this study demonstrate the use of ESI-MS to evaluate non-covalent interactions between perylene diimide ligands and G-quadruplex DNA and to screen the ligands for selectivity. There is good correlation between spectroscopic binding studies and ESI-MS analysis of G-quadruplex DNA binding. Qualitative comparisons can be made to quickly screen ligands with minimal consumption of analyte, making this technique compatible with high throughput screening techniques. Our results found, with the exception of Tel32 which exhibited no DNA binding, all ligands evaluated in this study formed complexes with G-quadruplex DNA. Binding stoichiometries of 1:1 and 2:1 were observed for PIPER, Tel11, and Tel18; this is a binding stoichiometry that is consistent with telomerase inhibitors that bind to G-quadruplex DNA by stacking on the faces of the terminal G-tetrads [35, 40]. Interesting, the three most quadruplex-selective ligands formed complexes with quadruplex DNA with a different binding stoichiometry. Tel01 formed 3:1 complexes, and Tel12 and Tel18 only formed 1:1 complexes. This unusual behavior will be explored in future studies. CAD experiments also suggested an end-stack mode of binding since all complexes dissociated by ejecting the ligand, leaving the intact G-quadruplex, triplex, or single strand ODN. Based on binding to G-quadruplex, duplex, and single stranded DNA, the analogs Tel01, Tel12, and Tel34, are

most promising in terms of G-quadruplex selectivity. Our future work will focus on obtaining more quantitative information about ligand/G-quadruplex complexes, such as binding constants, and identification of specific binding sites.

5.5 REFERENCES

(1) Goodman, L. S., Hardman, J. G., Limbird, L. E., and Gilman, A. G. *Goodman & Gilman's the Pharmacological Basis of Therapeutics*, 10th ed.; McGraw-Hill: New York, 2001.

(2) Propst, C. L., and Perun, T. J. *Nucleic Acid Targeted Drug Design*; M. Dekker: New York, 1992.

(3) Hofstadler, S. A., and Griffey, R. H. *Chem. Rev.* **2001** *101*, 377-390.

(4) Beck, J. L.;, Colgrave, M. L., Ralph, S. F., and Sheil, M. M. *Mass Spectrom. Rev.* **2001**, *20*, 61-87.

(5) Gabelica, V. and De Pauw, E. *J. Am. Soc. Mass Spectrom.***2001**, *13*, 91-98.

(6) Gabelica, V., De Pauw, E., and Rosu, F. *J. Mass Spectrom.* **1999**, *34*, 1328-1337.

(7) Wan, K. X., Gross, M. L., and Shibue, T. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 450-457.

(8) Gale, D. C., and Smith, R. D. *J. Am. Soc. Mass Spectrom.* **1995**, *6*, 1154-1164.

(9) Triolo, A., Arcamone, F. M., Raffaelli, A., and Salvadori, P. *J. Mass Spectrom.*, **1997**, *32*, 1186-1194.

(10) Hofstadler, S. A., Sannes-Lowery, K. A., Crooke, S. T., Ecker, D. J., Sasmor, H., Manalili, S., and Griffey, R. H. *Anal. Chem.* **1999**, *71*, 3436-3440.

(11) Kapur, A., Beck, J. L., and Sheil, M. M. *Rapid Commun. Mass Spectrom.* **1999**, *13*, 2489-2497.

(12) Greig, M. J., and Robinson, J. M. *J. Biomol. Screen.***2000**, *5*, 441-454.

(13) Wan, K. X., Shibue, T., and Gross, M. L. *J. Am. Chem. Soc.***2000**, *122*, 300-307.

(14) Gabelica, V., Rosu, F., Houssier, C., and De Pauw, E. *Rapid Commun. Mass Spectrom.* **2000**, *14*, 464-467.

(15) Iannitti-Tito, P., Weimann, A., Wickham, G., and Sheil, M. M. *Analyst* **2000**, *125*, 627-633.

(16) Gupta, R., Kapur, A., Beck, J. L., and Sheil, M. M. *Rapid Commun. Mass Spectrom.* **2001**, *15*, 2472-2480.

(17) Rosu, F., Gabelica, V., Houssier, C., and De Pauw, E. *Nucleic Acids Res.* **2002**, *30*, e82.

(18) Carrasco, C., Rosu, F., Gabelica, V., Houssier, C., De Pauw, E., Garbay-Jaureguiberry, C., Roques, B., Wilson, W. D., Chaires, J. B., Waring, M. J., and Bailly, C. *ChemBiochem* **2002**, *3*, 1235-1241.

(19) Colgrave, M. L., Beck, J. L., Sheil, M. M., and Searle, M. S. *Chem. Comm.* **2002**, *6*, 556-557.

(20) Guittat, L., Alberti, P., Rosu, F., Van Miert, S., Thetiot, E., Pieters, L., Gabelica, V., De Pauw, E., Ottaviani, A., Riou, J. F., and Mergny, J. L. *Biochimie* **2003**, *85*, 535-547.

(21) Gabelica, V., Galic, N., Rosu, F., Houssier, C., and De Pauw, E. *J. Mass Spectrom.* **2003**, *38*, 491-501.

(22) Colgrave, M. L., Iannitti-Tito, P., Wickham, G., and Sheil, M. M. *Aust. J. Chem.* **2003**, *56*, 401-413.

(23) Rosu, F., De Pauw, E., Guittat, L., Alberti, P., Lacroix, L., Mailliet, P., Riou, J. F., and Mergny, J. L. *Biochemistry* **2003**, *42*, 10361-10371.

(24) Rosu, F., Gabelica, V., Shin-ya, K., and De Pauw, E. C*hem. Commun.* **2003**, *21*, 2702-2703.

(25) Beck, J. L., Gupta, R., Urathamakul, T.. Williamson, N. L., Sheil, M. M., Aldrich-Wright, J. R., and Ralph, S. F. *Chem Commun.* **2003**, *9*, 626-627.

(26) David, W. M., Brodbelt, J., Kerwin, S. M., and Thomas, P. W. *Anal. Chem.***2002**, *74*, 2029-2033.

(27) Rosu, F., Gabelica, V., Houssier, C., Colson, P., and Pauw, E. D. *Rapid Commun. Mass Spectrom.***2002**, *16*, 1729-1736.

(28) Vairamani, M., and Gross, M. L. *J. Am. Chem. Soc.* **2003**, *125*, 42-43.

(29) Kerwin, S. M. *Curr. Pharm. Des.* **2000**, *6*, 441-471.

(30) Hurley, L. H., Wheelhouse, R. T., Sun, D., Kerwin, S. M., Salazar, M., Fedoroff, O. Y., Han, F. X., Han, H. Y., Izbicka, E., and Von Hoff, D. D. *Pharmacol. Ther.* **2000**, *85*, 141-158.

(31) Wellinger, R. J., and Sen, D. *Eur. J. Cancer* **1997**, *33*, 735-749.

(32) Bearss, D. J., Hurley, L. H., and Von Hoff, D. D. *Oncogene* **2000**, *19*, 6632-6641.

(33) Mergny, J. L., Mailliet, P., Lavelle, F., Riou, J. F., Laoui, A., and Helene, C. *Anti-Cancer Drug Des.* **1999**, *14*, 327-339.

(34) Counter, C. M., Hahn, W. C., Wei, W. Y., Caddle, S. D., Beijersbergen, R. L., Lansdorp, P. M., Sedivy, J. M., and Weinberg, R. A. *Proc. Nal. Acad. Sci. U.S.A.* **1998**, *95*, 14723-14728.

(35) Fedoroff, O. Y., Salazar, M., Han, H. Y., Chemeris, V. V., Kerwin, S. M., and Hurley, L. H. *Biochemistry* **1998**, *37*, 12367-12374.

(36) Kern, J. T., and Kerwin, S. M. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 3395-3398.

(37) Kern, J. T., Thomas, P. W., and Kerwin, S. M. *Biochemistry* **2002** *41*, 11379-11389.

(38) Kerwin, S. M., Chen, G., Kern, J. T., and Thomas, P. W. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 447-450.

(39) Rossetti, L., Franceschin, M., Bianco, A., Ortaggi, G., and Savino, M. *Bioorg. Med. Chem. Lett.* **2002,** *12*, 2527-2533.

(40) Read, M. A., and Neidle, S. *Biochemistry* **2000**, *39*, 13422-13432.

(41) Han, H. Y., Bennett, R. J., and Hurley, L. H. *Biochemistry* **2000**, *39*, 9311-9316.

(42) Han ,H. Y., Cliff, C. L., and Hurley, L. H. *Biochemistry* **1999**, *38*, 6981-6986.

(43) Read, M. A., Wood, A. A., Harrison, J. R., Gowan, S. M., Kelland, L. R., Dosanjh, H. S., and Neidle, S. *J. Med. Chem.* **1999**, *42*, 4538-4546.

(44) Pasternack, R. F., and Collings, P. J. *Science* **1995**, *269*, 935-939.

(45) Clark, G. R., Pytel, P. D., Squire, C. J., and Neile, S. *J. Am. Chem. Soc.* **2003**, *125*, 4066-4067.

(46) Randazzo, A., Galeone, A., and Mayol, L. *Chem. Comm.* **2001,** *11*, 1030-1031

(47) Kerwin, S. M., Sun, D., Kern, J. T., Rangan, A., and Thomas, P. W. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 2411-2414.

Chapter 6: Gas-Phase Stability of G-quadruplex DNA Determined by Electrospray Ionization Tandem Mass Spectrometry and Molecular Dynamics Simulations

6.1 INTRODUCTION

The basis of many anticancer and antitumor therapies is the interaction between small molecule drugs and nucleic acid structures. While most current DNA-interactive therapies target duplex DNA, G-quadruplex DNA has attracted recent interest as a potential anti-cancer drug target because of its role in telomere maintenance.[1-3] Telomeres are the regions of non-coding DNA found on the extremities of chromosomes.[4] In addition to protecting the chromosomes from fusion and degradation, telomeres allow for the complete replication of the chromosomal DNA. With each subsequent cell division process, the length of the telomeres is shortened until a critical length is reached, leading to cell senescence and death.[5] Telomerase is a reverse transcriptase enzyme that is responsible for maintaining the length of telomeric DNA. While this enzyme is inactive in most human somatic cells, high levels of telomerase activity is found in 80-90% of human cancer cells. [6, 7]

Telomeric DNA is composed of tandem repeats of G-rich sequences, such as the $d(T_2AG_3)_n$ sequence in mammals,[8] $d(T_2G_4)_n$ in *Tetrahymena*[9] and $d(T_4G_4)_n$ in *Oxytricha.*[10] These and other G-rich sequences have been shown to form G-quadruplex structures *in vitro.* G-quadruplex DNA forms via Hoogsteen hydrogen bonding between a planar arrangement of four guanine nucleobases, termed a G-quartet. In the quadruplex structure, the G-quartets stack on top of one another and the overall structure is stabilized by monovalent cations such as Na^+, K^+, or NH_4^+ coordinated in the central cavity of the tetrad [11]. Structural polymorphism resulting from different quadruplex sequences and strand orientations has been demonstrated in vitro.[12] A DNA sequence containing a single

G-rich repeat can form a four-stranded parallel G-quadruplex. Strands containing two or more G-rich regions can form G-G hairpins and dimerize in different orientations to form a two-stranded quadruplex, while a sequence with four G-rich repeats can fold upon itself to form a intramolecular quadruplex.[12]

The stabilization of the G-quadruplex structure of the 3' single strand overhang of telomeric DNA by ligands and metal cations has been demonstrated to inhibit telomerase activity *in vivo*, thus stimulating interest in the study of telomeric quadruplex formation and the development of G-quadruplex selective ligands and anticancer compounds.[13-15] Electrospray ionization mass spectrometry (ESI-MS) has emerged as a useful tool for the analysis of quadruplex DNA and its non-covalent complexes with ligands due to its low sample consumption and fast analysis time. During the ESI process, non-covalent complexes in solution are transferred to the gas-phase with minimal internal energy, allowing many of the binding interactions to be maintained. The preservation of these non-covalent complexes allows information about binding stoichiometry and selectivity to be elucidated from the mass spectra Our group[16,17] and others[18-23] have focused on developing ESI-MS based techniques to assess whether the binding of quadruplex interactive ligands observed in the gas-phase can be correlated to known solution behavior with the ultimate goal of developing ESI-MS as a screening tool for drug/DNA complexes.

There have been several ESI-MS studies aimed at examining structural features of duplex[24-26] and hairpin[27,28] DNA in solution and the gas-phase, primarily based on correlation of relative abundances of ions observed in ESI mass spectra as a function of solution conditions or based on collisional activated dissociation (CAD) patterns. CAD is a versatile technique that has been used to characterize gas-phase stabilities and fragmentation pathways of DNA duplexes.[22,25,26,29,30] For example, the dissociation profiles

of a series of 6-, 8- and 12-base pair self-complementary duplexes have been shown to correspond to the known solution melting behavior of DNA, demonstrating the utility of CAD to provide relative comparisons of the gas-phase stabilities of DNA duplexes.[29] Few studies have been undertaken to elucidate the gas-phase structures or stabilities of quadruplex DNA. One recent study focused on using ion mobility mass spectrometry and molecular dynamics simulations to examine gas-phase conformations of quadruplexes with different lengths of the telomeric repeat $d(TTAGGG)_n$ where n = 1, 2, 4, and 6.[31] The experimental cross sections of the quadruplex structures determined by ion mobility matched those determined by molecular modeling, with the gas-phase structures also corresponding to solution structures unveiled by circular dichroism. Only one study has been devoted to the CAD fragmentation of DNA quadruplexes, and it found that $[d(TGGGT)]_4$ ions dissociated via the loss of a single strand ion, producing a triplex species.[32]

Molecular modeling techniques comprise another promising tool for the study of nucleic acid structures in a solvent free environment. Molecular dynamics (MD) simulations are suitable for the detailed study of the structural, energetic, and dynamic properties of macromolecules, including proteins and nucleic acids. Recently, many research groups have studied G-quadruplexes with MD simulations in a solvent-free environment.[18,31,33-35] An interesting conclusion of those studies was that, in the proper charge state, G-quadruplex structures could be maintained with some distortion during the solvent-free MD simulations. The additional finding that the collision cross sections of MD snapshots were in good agreement with those measured by ion mobility experiments[18, 31] indicates that both are viable techniques for the determination of strand orientation patterns for G-quadruplexes. Baker et al. also found this approach to be useful for determining the proper binding patterns of ligands with G-quadruplexes.[18] However, it

should be noted that similar collision cross sections can be associated with extremely different conformations; therefore this method may not be used to measure how well a G-quadruplex maintains its G-quartets in the solvent-free environment of ESI-MS experiments.

In the present study, we use the capabilities of both ESI-MS/MS and molecular dynamics simulations to assess the relative stabilities of different quadruplex structures in the gas phase. First, tandem mass spectrometry, specifically collisional activation dissociation, was used to evaluate the fragmentation patterns of gas-phase quadruplex ions and create energy-variable dissociation curves which provide a way to compare the relative stabilities of different quadruplex structures. Second, molecular dynamics simulations were performed to examine the structures and trajectories of the quadruplexes in a solvent free environment to not only confirm that the quadruplexes are maintained, but also to assess their relative stabilities using free energy measurements. To the best of our knowledge, this is the first attempt to predict the relative stabilities for a set of G-quadruplexes through free energy analysis.

6.2 EXPERIMENTAL

6.2.1 Chemicals

Single strand oligodeoxynucleotides (ODNs), custom synthesized as ammonium salts on the 1.0 μmole scale with purification by HPLC, were obtained from Integrated DNA Technologies (Coralville, IA) and used without further purification. The sequences of the quadruplex-forming ODNs are shown in Table 6.1. Stock solutions of $d(G_4T_4G_4)$ (G2), $d(T_2AG_3)_2$ (G3), $d(T_2AG_3)_4$ (G4), $d(T_2G_4)_4$ (G5), and $d(G_2T_4)_4$ (G6) were prepared at 0.8 mM concentration in 150 mM ammonium acetate buffer while $d(TG_4T)$ (G1) and $d(T_2G_3T)$ (G7) were prepared at 1 mM in 150 mM ammonium acetate buffer. Each quadruplex structure was annealed by heating the stock solutions to 90 °C for 10 min.

followed by cooling to 25 ° C during a period of 7 h. The ammonium acetate buffer was chosen for the annealing solutions since quadruplexes form in the presence of the NH_4^+ ion, and NH_4^+ is more suitable than Na^+ or K^+ for ESI-MS studies due to its greater lability.[32] Stock solutions of the duplex forming ODNs, d($GCG_3A_2T_2G_3CG$) and d($CGC_3A_2T_2C_3GC$) were prepared at 2 mM in 250 mM ammonium acetate. Equal volumes of each single strand were combined and the duplex was annealed using the same procedure described above for the quadruplex DNA. A 500 μM stock solution of the single strand ODN d(TTGGGGGT) was prepared in deionized water. *N,N'*-bis(2-morpholinylpropyl)-3,4,9,10-perylenetetracarobxylic acid diimide (Tel01) was prepared as previously reported.[36] A 1 mM stock solution of the ligand was prepared in 0.1% aqueous trifluoroacetic acid.

Table 6.1: Quadruplex sequences used in this study.

Quadruplex	Sequence	Type of Quadruplex	Number of Tetrads	Molecular Weight (Da)
G1	$[d(TG_4T)]_4$	parallel 4-stranded[31]	4	7453.2
G2	$[d(G_4T_4G_4)]_2$	antiparallel 2-stranded[32]	4	7577.0
G3	$[d(T_2AG_3)_2]_2$	parallel or antiparallel 2-stranded[31]	3	7513.0
G4	$d(T_2AG_3)_4$	antiparallel intramolecular[31]	3	7575.0
G5	$d(T_2G_4)_4$	intramolecular	4	7639.0
G6	$d(G_2T_4)_4$	intramolecular	4	7438.9
G7	$[d(T_2G_3T)]_4$	4-stranded	3	7353.0

6.2.2 Mass Spectrometry

ESI-MS experiments were undertaken on a ThermoFinnigan LCQ Duo mass spectrometer (San Jose, CA). Analytical solutions of the quadruplex were prepared at 10

μM concentration in 50 mM ammonium acetate buffer with 25% (v/v) methanol to aid in the desolvation of the DNA. Equimolar (10 μM) concentrations of Tel01 with each quadruplex, duplex or single strand ODN were prepared in the ammonium acetate buffer for the ligand binding study, The samples were directly infused into the mass spectrometer at 3 μL/min using a Harvard Syringe Pump (Holliston, MA). Ions were generated in the negative ion mode using a needle voltage of 3.5 kV, a heated capillary temperature of 110 °C and sheath and auxiliary gas flows of 40 and 10 arbitrary units, respectively. An ion accumulation time of 100 ms was used and spectra were acquired by summing 300 scans.

For collisional activated dissociation (CAD) experiments, the desired precursor ion was isolated in the trap using resonance ejection, followed by fragmentation induced by increasing the resonance voltage applied to the trap. An activation time of 30 ms was used for all experiments. To obtain the energy-variable dissociation curves, the applied CAD voltage, a parameter which influences the collision energy, was increased from 0.70 to 1.70 V in increments of 0.11 V. The data presented represent an average of three experiments done on three different days to account for possible instrument instability. Aside from the energy variable experiments, CAD was performed by increasing the collision voltage to the amplitude that resulted in dissociation of the precursor ion to 10% of its initial abundance.

6.2.3 Molecular Modeling

Gas phase MD simulations were performed for quadruplexes G1-G6. The initial structure of G1 was taken from a parallel-stranded crystal structure resolved at 0.95 Å (PDB: 1S45)[37] and the starting structure of G2 was taken from an antiparallel-stranded crystal structure resolved at 1.86 Å (PDB: 1JB7).[38] For G3-G6, a parallel-stranded crystal structure resolved at 2.1 Å (PDB 1KFI)[39] and an antiparallel-stranded NMR structure

(PDB 143D)[40] were applied as templates to construct the initial structures of both strand orientations for each G-quadruplex. An additional starting structure for G6 was modeled by mutating G to T and T to G on the model structure of antiparallel G5. It should be noted that for a particular quadruplex our decision on which strand orientations to model was based on both experimental and theoretical findings. To investigate the role of cations in stabilizing the G-quartet, two models were built for G1, one with and one without ammoniums residing in the central channel. In total, twelve model structures, namely, G1-ammonium, G1-no-cation, G2-antiparallel, G3-antiparallel, G3-parallel, G4-antiparallel, G4-parallel, G5-antiparallel, G5-parallel, G6-antiparallel, G6-parallel, G6-T-tetrad, were used for the MD simulations.

Although the most abundant charge state observed in ESI-MS experiments can be readily identified from the mass spectra, the exact locations of the deprotonation sites are not clear. Rueda et al. proposed two neutralization protocols to model the desired charge state for a G-quadruplex: a distributed model and a more realistic localized model.[33-35] In this study, we adopted the a modified version of localized model. Instead of maximizing the sum of the distances between the charged phosphates, we calculate electrostatic energies of all possible ways of assigning deprotonation sites and the assignment that gave the most favorable electrostatic energy was chosen for the subsequent molecular mechanical studies. If a G-quadruplex has ammonium adducts, a set of 3-dimensional fine grids (0.2 Å) that enclosed the quadruplex was constructed. Grid points that were within 2.0 Å of any atom of the quadruplex were dropped off. A +1 charge probe was placed at each grid point to calculate the electrostatic energy between the point charge and the quadruplex. Finally, an ammonium ion was placed at the grid point that gave the optimum electrostatic energy. This procedure was repeated *n* times to incorporate *n* ammonium molecules.

All of the minimizations, MD simulations and subsequent structural and energetic analyses were carried out with the AMBER 8 suite of programs[41] using the Parm99 force field.[42] A delicate minimization protocol was designed to minimize the initial structures step by step: (1) Only hydrogen atoms were optimized for 1000 steps. (2) A 1000-step minimization was carried out with the G-quartet guanine nucleosides restrained by a 50 kcal/mol$Å^2$ harmonic potential. (3) The G-quartet guanine bases were restrained by a 50, 10, 5 and 1 kcal/mol$Å^2$ harmonic potentials in a set of subsequent 1000-step minimizations. (4) 1000-step full minimization was carried out for the complete molecules.

The fully minimized structures were then subjected to a set of MD simulations. Similar to the restrained minimizations, the G-quartet guanine bases were restrained by a 10, 5, 1, 0.5, 0.1 kcal/mol$Å^2$ harmonic potential in a set of subsequent 100 picosecond restrained MD simulations. In each step, the last snapshot was taken as the starting structure for the next simulation run. Finally, a completely unrestrained 50 nanosecond MD simulation was performed. All the simulations were carried out at 300 K with a time step of 1 femtosecond. No non-bonded cutoff was applied and the SHAKE algorithm was applied for all the bonds involving hydrogen atoms.[43]

A large variety of geometrical analyses including root-mean-square displacements (RMSD), hydrogen bonding, stacking between two guanines, and free energy analysis, were performed for 50 snapshots evenly selected from the last 30 nanosecond simulations. Hydrogen bonds were formed when the heteroatom-heteroatom distance was less than 4.0 Å and the donor-hydrogen/donor-acceptor angle was deviated less than 60° from linearity. A stacking interaction was recognized when the distance between two base rings was less than 5.5 Å, the angle between the two normals to the base planes was less than 30°, and the angle between the normal of any base plane and

the vector between the centers of the two bases was less than 50°. The numbers of stacking pairs were 4, 8 and 12 for an ideal G-quadruplex with two, three and four tetrads, respectively.

The free energy of a molecule was composed of two parts: the molecular mechanical energy (E_{MM}) and the conformational entropy (TS). The E_{MM} energies were calculated with the Anal program of AMBER 8 on the minimized structure of each snapshot,[41] while the conformational entropies were estimated by normal model analysis on a thoroughly minimized structure with the Nmode program of AMBER 8.[41] The principles behind the free energy calculations were the thermodynamics and the widely used MM-PBSA free energy theories[44-46] (Solvation free energies were omitted because all the molecules were studied in the gas phase). As discussed above, free energy analysis was essential in interpreting the CAD experimental data.

6.3 RESULTS AND DISCSUSSION

6.3.1 Mass Spectra of the G-quadruplexes

The quadruplex structures selected for this study represent a variety of quadruplex-forming sequences and strand orientations. The variation allows one to evaluate which structures are most stable in the gas-phase and thus most suitable for future ESI-MS studies of quadruplex DNA. Sequences G1 and G7 were selected to form four-stranded quadruplex structures with four and three quartets, respectively. Previous circular dichroism (CD) studies performed on G1 in an ammonium acetate buffer indicate that it forms a parallel, four stranded quadruplex structure.[31] Both G2 and G3 contain two G-rich regions per strand and form two-stranded quadruplex structures. CD studies have found that G2 forms an antiparallel structure[32] while the CD spectrum for G3 has characteristics of both parallel and antiparallel orientations.[31] The intramolecular quadruplex-forming sequences selected for this study, G4, G5, and G6, contain four

repeats each of G-rich sequences capable of forming three, four, and two quartets, respectively. The major species in the CD spectrum of G4 in an ammonium acetate buffer is the antiparallel quadruplex, with evidence that the parallel quadruplex is also present to a lesser degree.[31] All quadruplexes in the present study contain 24 nucleobases to ensure that the molecular weights and degrees of freedom of the DNA molecules were comparable for the energy-variable dissociation studies.

Full scan ESI mass spectra of solutions containing each quadruplex were obtained to evaluate the annealing efficiencies. Representative spectra are shown in Figure 6.1 for the four-stranded quadruplex G1 (Figure 6.1a), the two-stranded structure G2 (Figure 6.1b), and the intramolecular quadruplex G4 (Figure 6.1c). The 5- charge state proved to be the most abundant for all of the quadruplex structures, and thus was used for all subsequent CAD and variable-energy dissociation experiments. The 5- ions of intermolecular quadruplexes G1 (Figure 6.1a) and G7 (spectra not shown) can also unambiguously assigned as four stranded structures, while G2 (Figure 6.1b) and G3 can be defined as two-stranded structures for G2 and G3 based on the m/z values which are distinctive for species containing one, two, three or four strands.

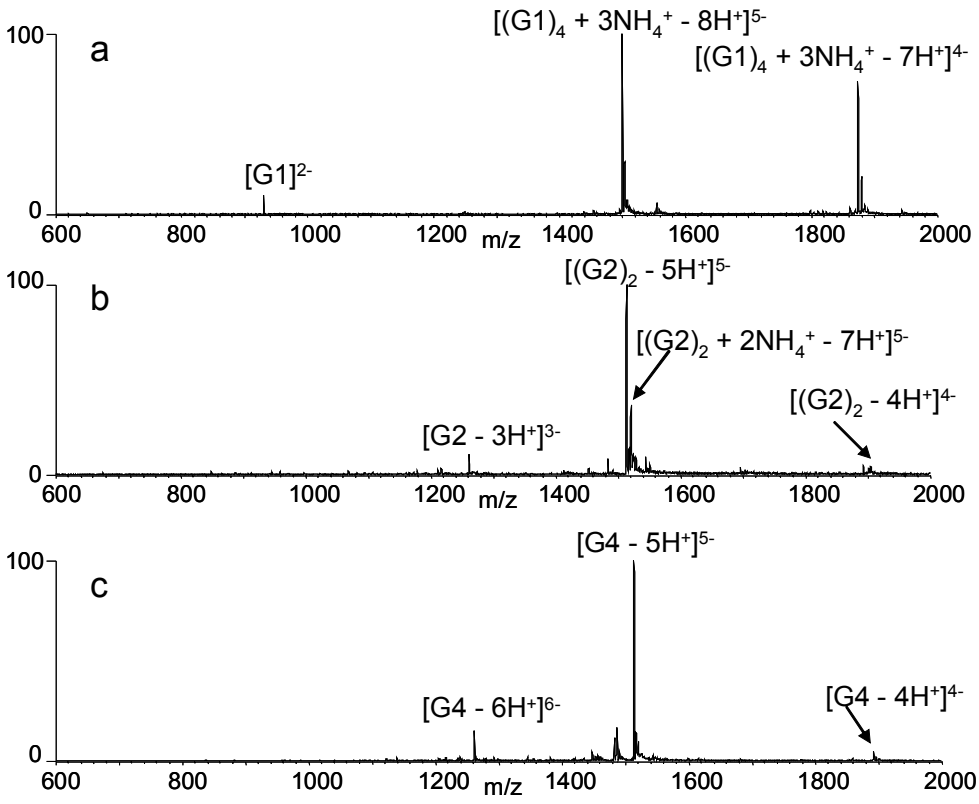


Figure 6.1: ESI mass spectra of solutions containing 10 μM of (a) G1, (b) G2, and (c) G4.

Ammonium adducts are also associated with the four-stranded quadruplexes G1 and G7. The dominant 5- ion of G1 contained three ammoniums (Figure 6.1a) while the corresponding ion of G7 contained two adducted ammoniums (spectra not shown). These results are consistent with the previous observation that *n*-1 ammonium cations are incorporated in quadruplex structures containing *n* tetrads.[32] As shown in Figure 6.1b and 6.1c, no ammoniums are associated with the most abundant ions for the two-stranded G2 and single stranded G4 quadruplex. Similar results were obtained for other two-stranded quadruplex G3 and the single stranded G5 and G6 (spectra not shown). In addition, an ion corresponding to the two-stranded G2 structure with two ammoniums is present in the mass spectra but it is approximately one third the abundance of the dominant ion without

adducted ammonium (Figure 6.1b). These results are consistent with previous studies of quadruplexes by ESI-MS that found adducted ammoniums are only retained by the dominant quadruplex species for the four-stranded quadruplex G1, suggesting tighter binding of the ammonium ions by this structure.[31, 32]

While Figure 6.1 confirms that ions containing strand stoichiometries consistent with quadruplex structures for G1 and G2 are present, it does not verify that the quadruplex structures are maintained in the gas-phase nor does it indicate whether G4 is in the folded quadruplex conformation or a denatured state. ESI-MS studies of quadruplexes have been largely limited to using multiple stranded quadruplexes because the *m/z* of an intramolecular quadruplex is the same as a denatured single strand structure, thereby precluding direct confirmation that the quadruplex is maintained in the gas-phase. While recent studies have confirmed that quadruplex structures are conserved in the gas-phase, they did not provide information about the relative stabilities of the structures.[31] Here, molecular dynamics studies were carried out to demonstrate that quadruplex structures can be maintained in the gas-phase, while further calculations and CAD experiments were used to assess quadruplex stability.

6.3.2 Ligand Binding of the Quadruplexes

The binding of the quadruplexes to one quadruplex-selective ligand was used to demonstrate that the quadruplexes were present in the annealed solutions. In a previous study we demonstrated that that the perylene diimide *N,N'*-bis(2-morpholinylpropyl)-3,4,9,10-perylenetetracarobxylic acid diimide (Tel01)[36] demonstrated binding selectivity for quadruplexes over duplex and single strand DNA by forming significantly lower abundance complexes with duplex DNA and exhibited no binding to single strand DNA.[17] Here, we examined the binding of Tel01 to the quadruplexes, one duplex, and one single strand oligonucleotide by ESI-MS. Solutions containing equimolar concentrations of the

DNA and ligand were prepared and analyzed by ESI-MS. The fraction of bound DNA values were determined from the mass spectra as described previously,[17] and the results are summarized in Table 6.2. Tel01 formed abundant complexes with each quadruplex, with fraction of bound DNA values ranging from 0.80 for G5 to 0.95 for G6. These values are on par with the results of our previous study.[17] The binding of Tel01 to a duplex d(GCGGGAATTGGGCG/CGCCCAATTCCCGC) (DS1) was indicated by the presence of Tel01/DS1 complexes in the mass spectrum, but they are significantly lower in abundance than the complexes formed between the ligand and the quadruplexes. Tel01 also did not form any complexes with the single strand d(TTGGGGGT) (SS1). This sequence was selected since it is G-rich like many of the quadruplexes however it was not annealed and ESI-MS experiments were done to confirm that only the single strand species was present in solution (spectra not shown). Had the quadruplexes not properly annealed and remained in the denatured single strand conformation, little or no Tel01 binding would be expected. These results confirm that the quadruplexes are present in solution and that the ESI-MS results reflect the anticipated quadruplex selectivity.

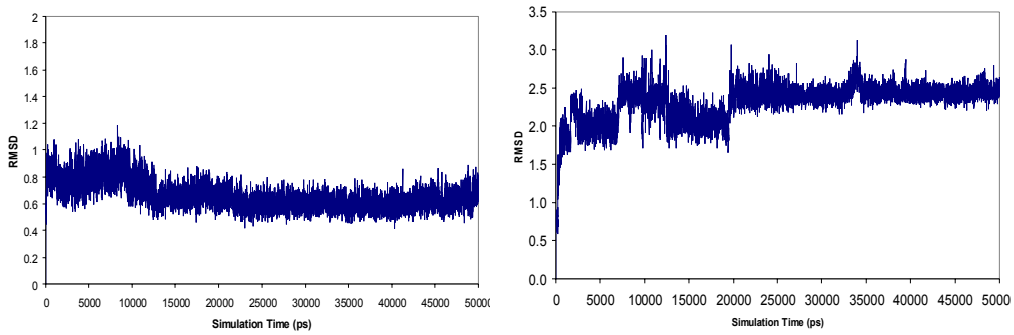
Table 6.2: Fraction of bound DNA values for equimolar concentrations (10 μM) of each DNA with Tel01. [a]All values +/- 0.05.

DNA Sequence	Fraction DNA Bound to Tel01[a]
G1	0.95
G2	0.84
G3	0.93
G4	0.93
G5	0.80
G6	0.95
DS1	0.20
SS1	0.00

6.3.3 Molecular Dynamics Trajectories of the Quadruplexes in the Gas-Phase

Molecular dynamics simulations were undertaken to demonstrate that the quadruplex conformations are maintained under conditions used for the ESI-MS/MS experiments. A net charge of 5- was assigned to the twelve model quadruplexes since this was the most abundant charge state observed in the ESI-MS experiments. Plots of the RMSD values of the tetrad forming guanines versus simulations time were constructed using the final minimized structure as a reference point. In most cases, the trajectories fluctuated vigorously during the first few nanoseconds and then became quite stable after 20 nanoseconds as demonstrated by the plots for G1 shown in Figure 6.2. The reference structures for the comparisons are the last structures of a series of minimizations which have their G-quartets well maintained. This behavior is consistent with what Rueda et al. found in their simulations of DNA in the gas phase.[33-35]

The final snapshots of the G1 simulations with and without the ammonium ions are depicted in Figure 6.3. From Figures 6.3a and 6.3b, it is clear that the ammonium cations (blue) reside in the centers of two adjacent tetrad planes. Furthermore, a comparison of Figures 6.3b and 6.3d shows the impact of the ammonium ions on the G-tetrad stabilization as the deviations in the tetrad position are more pronounced when the ammoniums are absent. The final snapshots of the structures with the most energetically favorable strand orientations of G2-G6 are presented in Figure 6.4. The antiparallel structures of G2, G3, G4, and G6 and the parallel structure of G5 were identified as the most favorable strand orientations of the quadruplexes in the gas-phase. The results of the MD simulations demonstrate that the G-quartets of the quadruplexes are maintained in the solvent-free environment.



(a) (b)

Figure 6.2: Root-mean-square displacements (Å) of G-tetrad guanines over 50 nanosecond MD simulations (compared to minimized reference structures) for two G1 models: (a) G1 with three ammoniums residing in the central cavity between the stacked tetrads, and (b) G1 in the absence of ammoniums in the central cavity).

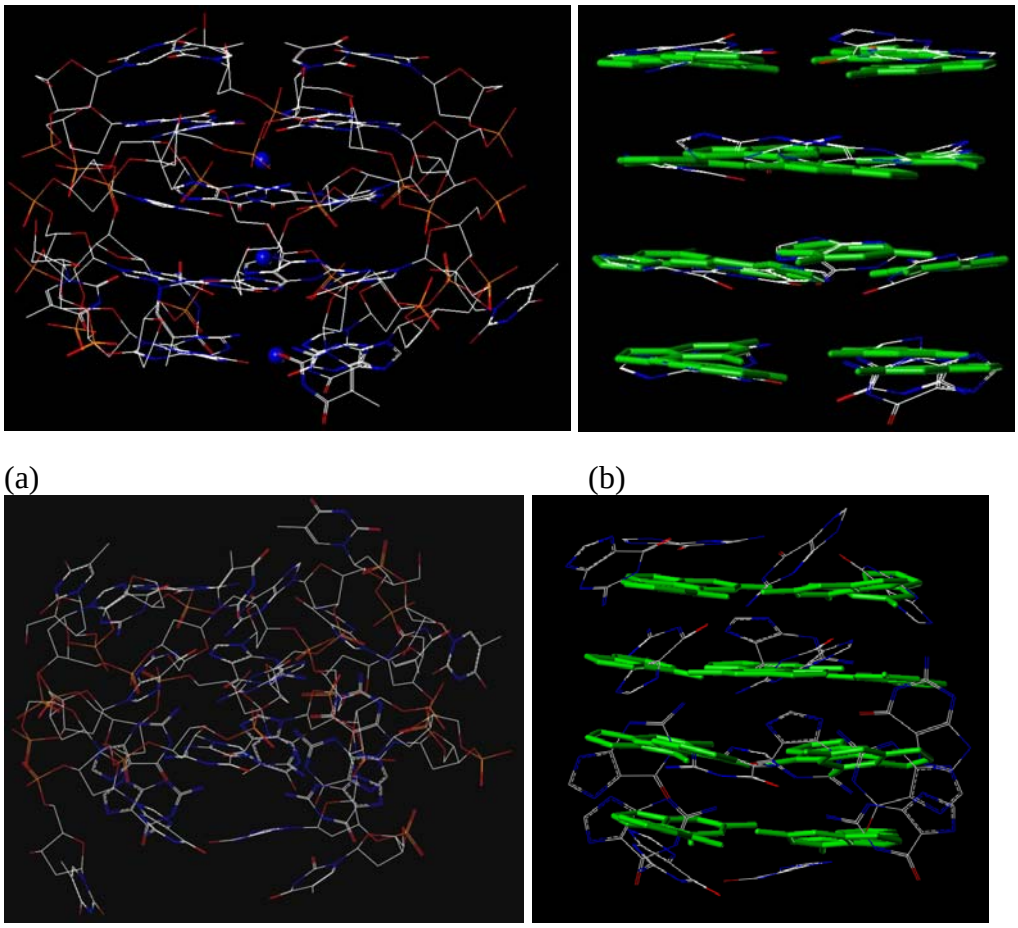


Figure 6.3: The final snapshot of a 50 nanosecond MD simulations for G1: (a) an all heavy atom model with three ammoniums residing in the central channel, (b) G-quartets with three ammonium ions (the green capped sticks of the G-quartets in the reference structure are presented for comparison purposes) (c) an all heavy atom model in the absence of ammonium ions, and (d) G-quartets in the absence of ammonium ions (the green capped sticks of the G-quartets in the reference structure are presented for comparison purposes).

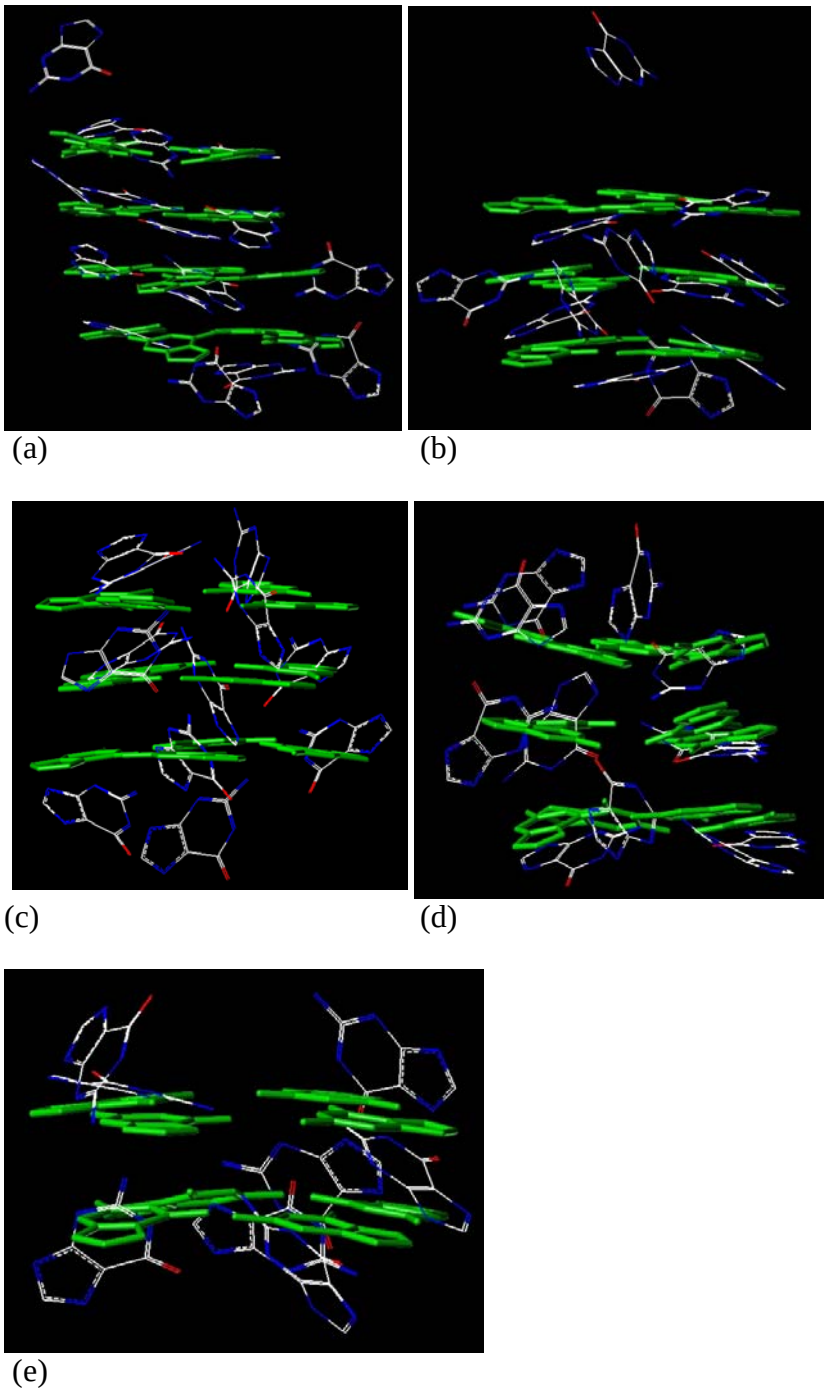


Figure 6.4: The G-quartets of G2-G6 in the final snapshots of 50 nanosecond MD simulations. Only the energetically more favorable strand orientations are presented: (a) G2-antiparallel, (b) G3-antiparallel, (c) G4-antiparallel, (d) G5-parallel, and (e) G6-antiparallel. The green capped sticks of the G-quartets in the reference structure are presented for comparison purposes.

6.3.4 CAD Fragmentation Pathways of the Quadruplexes

While significant attention has been devoted to examining the fragmentation pathways of DNA duplexes of different length, sequence and base pair composition,[29, 47] fewer studies have focused on quadruplexes. Thus, CAD spectra of G1, G2, G3, G4, G5, G6, and G7 (5- charge state) were obtained to determine how different quadruplex strand stoichiometries and sequences affect the fragmentation patterns. The quadruplex ions dissociate via the loss of a nucleobase, strand separation, or a combination of the two (Figure 6.5). The primary fragmentation route of the four-stranded quadruplexes G1 and G7 was strand separation to produce a triplex ion and single strand species in the 3- and 2- charge states, respectively. Guanine base loss from the resulting triplex species was also observed as demonstrated by the CAD spectrum of $[(G1)_4]^{5-}$ shown in Figure 6.5a.

In contrast to the four-stranded quadruplexes, the two-stranded quadruplexes dissociate primarily through guanine base loss from the precursor ion, although some strand separation products are also present. Interestingly, different degrees of strand separation are observed in the CAD spectra of G2 (Figure 6.5b) and G3 (Figure 6.5c) even though an applied collision voltage of 1.25 V was used for each ion. The relative abundances of the single strand ions in the CAD spectrum of G2 are significantly lower than those of the single strand ions in the spectrum of G3, suggesting the different sequences or orientations of the quadruplexes impact the fragmentation pathways. G2 has two repeats of four guanine nucleobases, allowing it to form a quadruplex containing four G-tetrads, while the sequence of G3 allows it to form only three tetrads. With a greater number of tetrads, there are more Hoogsteen hydrogen bonds holding the single strands of G2 together, making covalent cleavage of guanines favored over non-covalent strand separation. This result is reminiscent of an increase in nucleobase loss observed in the CAD spectra of DNA duplexes that parallels the total number of interstrand hydrogen

bonds, either by increasing the GC content of the sequence and/or lengthening the duplex.[29,30,47,48]

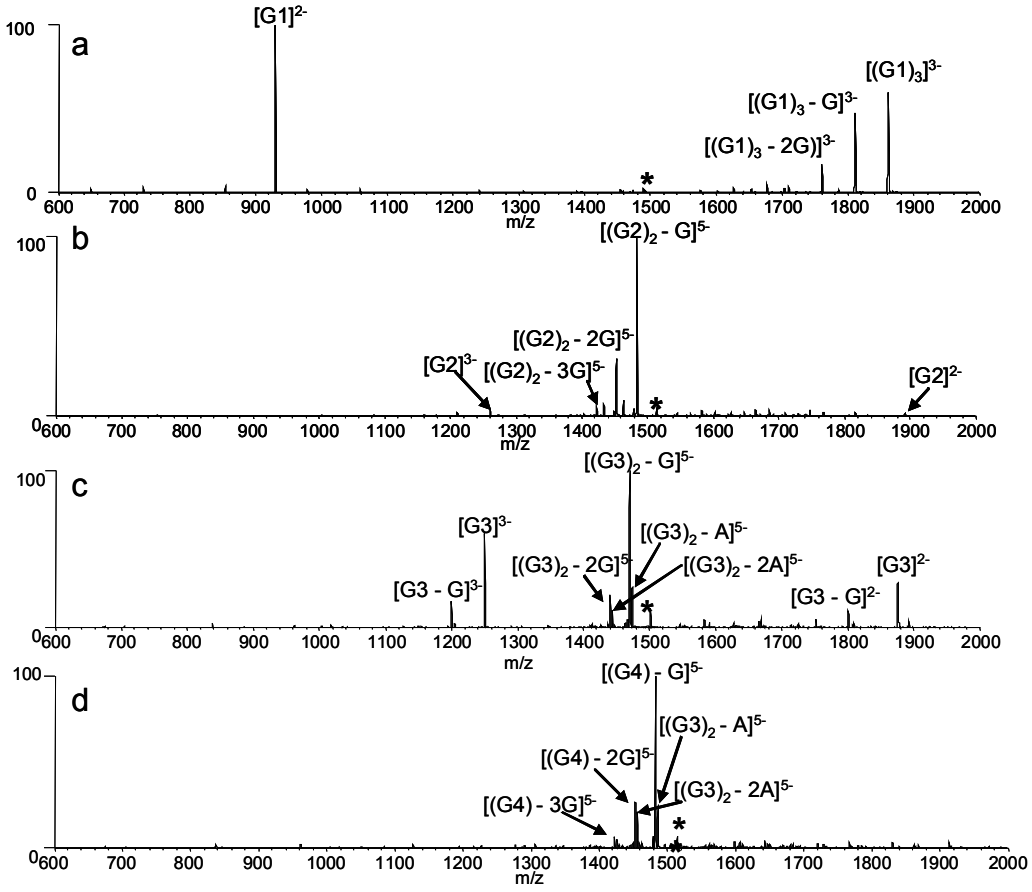


Figure 6.5: CAD spectra of the 5- charge state of: (a) $[(G1)_4]$, (b) $[(G2)_2]$, (c) $[(G3)_2]$, and (d) [G4].

Guanine nucleobase loss is also the primary fragmentation pathway of the intramolecular quadruplexes, as demonstrated by the CAD spectrum of G4 in Figure 6.5d. Similar spectra were acquired for G5 and G6 (data not shown). This result was expected since the quadruplex is single stranded, and thus cannot undergo strand separation. The CAD spectra of the intramolecular quadruplexes resemble those of single strand ions that have not formed a quadruplex, and thus give no diagnostic information about differences in the binding interactions between the quadruplexes.

6.3.5 Energy-Variable Dissociation Studies

While there are some general differences in the CAD spectra of the four, two and single stranded quadruplexes, the fragmentation pathways alone do not provide insight into the relative stabilities of the quadruplexes, nor can they distinguish the intramolecular structures (G4, G5, and G6) or the four-stranded structures (G1 and G7). To probe the relative stabilities of the quadruplexes in the gas-phase, energy-variable dissociation experiments were undertaken. The quadruplexes that dissociate via the same fragmentation pathway may require different amounts of energy to induce their dissociation as a result of different strand orientations and/or binding interactions. For this part of the study, the quadruplex ions in the 5- charge state were isolated and subjected to increasing collision energy based on the applied collisional activation voltage. The abundance of the precursor ion relative to the abundance of the product ions was monitored as a fraction of 1.0 and plotted as a function of CAD voltage. The reported $E_{1/2}$ values are calculated by determining the CAD voltage at which the relative abundance of the precursor ion is reduced to 50%.[26,29,49] A greater $E_{1/2}$ value implies a more stable ion, one that requires greater energization to induce dissociation. A similar approach has been used for in-source CAD and thermal denaturation experiments of DNA duplexes.[30]

The four-stranded quadruplexes G1 and G7 both dissociate via strand separation to produce single strand ions in the 2- charge state and triplex species with a 3- charge. However, the CAD voltages ($E_{1/2}$ values) required to produce these ions are significantly different as demonstrated by the dissociation profiles shown in Figure 6.6. The $E_{1/2}$ value for G1 was calculated to be 1.33 ± 0.05 V, while that of G7 was 1.14 ± 0.05 V, indicating G1 is substantially more stable in the gas phase than G7. This result is unsurprising since the sequence of G1 contains a single repeat of four guanines, allowing it to form four

tetrads, while G7 can only form three tetrads. With four tetrads, there are more hydrogen bonds between the four strands (16 for G1 versus 12 for G7) and greater stacking interactions between the tetrads. The differences in binding interactions between the single strands in G1 versus G7 are reflected in the $E_{1/2}$ values of the quadruplexes.

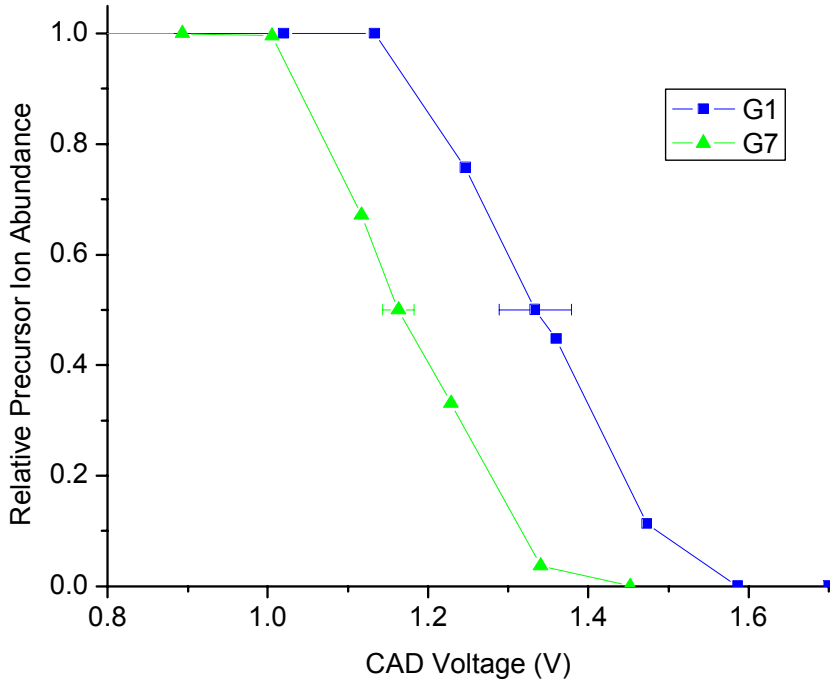


Figure 6.6: Energy-variable dissociation curves for $[(G1)_4]^{5-}$ and $[(G7)_4]^{5-}$.

The gas-phase stabilities of the intramolecular quadruplexes G4, G5, and G6 are also distinguished based on different $E_{1/2}$ values. Unlike G1 and G7, the strand separation pathway is not possible for the intramolecular quadruplexes. Therefore, the emergence of base loss ions and the formation of a_n-*B* and *w* ions is diagnostic for the disruption of the quadruplex structures. In order to cleave the covalent bond between the guanine and the ribose and disassemble the quadruplex structure, both Hoogsteen hydrogen bonding as well as tetrad stacking interactions must be broken. The dissociation curves shown in Figure 6.7 demonstrate that G4 and G5 have similar gas-phase stabilities, with $E_{1/2}$ values

of 1.20 ± 0.02 V and 1.18 ± 0.01 V, respectively, while G6 is significantly less stable with an $E_{1/2}$ value of 1.10 ± 0.02. The differences in the gas-phase stabilities may be partially explained by the number of tetrads per structure. The least stable intramolecular quadruplex, G6, contains four repeats of two guanines, allowing for only two tetrads in the quadruplex structure, compared to three tetrads in G4 or four tetrads in G6. However the $E_{1/2}$ values of G4 and G5 are approximately the same, even though G5 can form four G-tetrads and G4 can only form three. These results indicate there are also other factors governing the stabilities of the quadruplexes. As shown below, the structural and free energy analysis on the MD trajectories shed light on the trend of $E_{1/2}$ values of the G-quadruplexs.

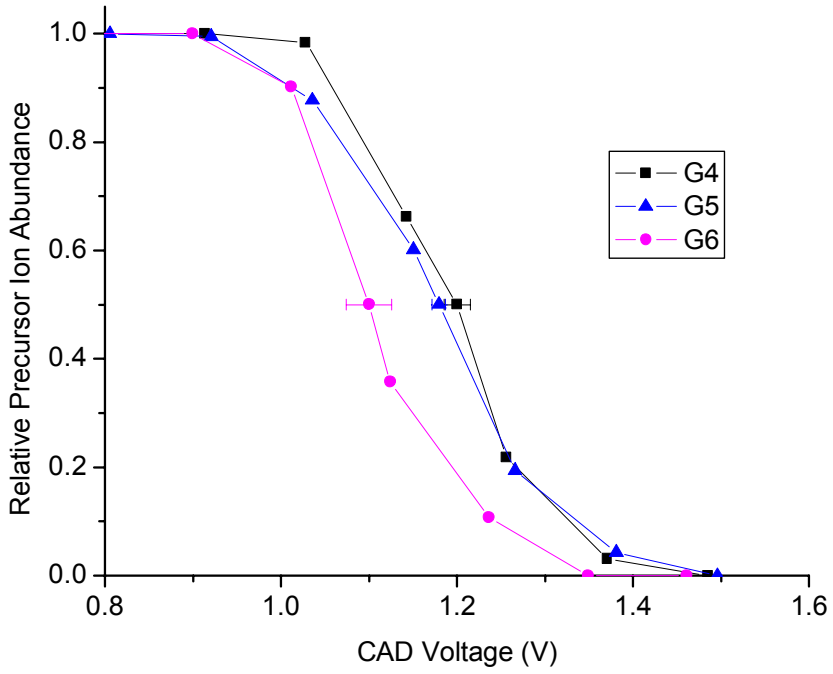


Figure 6.7: Energy-variable dissociation curves for $[G4]^{5-}$, $[G5]^{5-}$, and $[G6]^{5-}$.

Energy-variable dissociation curves were also constructed for the two-stranded quadruplexes G2 and G3 (Figure 6.8), both of which dissociate via strand separation and guanine base loss from the parent complex. Their $E_{1/2}$ values are virtually the same within the margin of errors of the measurements. While both quadruplexes are two-stranded, G2 is able to form four G-tetrads while G3 can only form three, so one would expect a greater $E_{1/2}$ value for G2. These results again suggest that other factors influence the apparent gas-phase stabilities of the quadruplexes, possibly strand orientation because G2 forms an antiparallel quadruplex, while G3 likely forms a mixture of parallel and antiparallel structures.

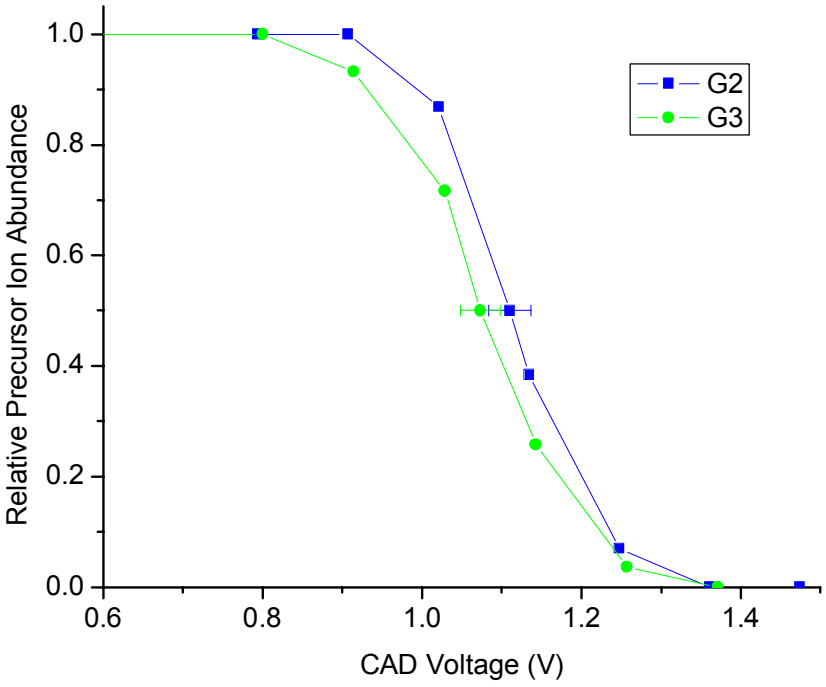


Figure 6.8: Energy-variable dissociation curves for $[(G2)_2]^{5-}$ and $[(G3)_2]^{5-}$.

In addition to comparing $E_{1/2}$ values amongst quadruplexes that have the same strand stoichiometries and thus fragmentation patterns, the $E_{1/2}$ values of all quadruplexes were also compared relative to each other (Table 6.3). The $E_{1/2}$ values for the

quadruplexes fall into three main groups. G1, the parallel four-stranded quadruplex with four tetrads and three retained ammonium ions, has by far the largest $E_{1/2}$. The quadruplexes with the next highest $E_{1/2}$ values are two of the intramolecular quadruplexes with three and four tetrads, G4 and G5, respectively. The next grouping based on $E_{1/2}$ values include the multistranded quadruplexes with three tetrads, G7 with four strands, and G3 with two strands, followed by the intramolecular quadruplex with two tetrads G6, and finally the two stranded quadruplex with four tetrads.

Table 6.3: Summary of $E_{1/2}$ values for the quadruplexes.

Quadruplex	$E_{1/2}$ Value (V)	Number of tetrads
G1	1.33 +/- 0.05	4
G2	1.07 +/- 0.03	4
G3	1.11 +/- 0.03	3
G4	1.20 +/- 0.02	3
G5	1.18 +/- 0.01	4
G6	1.10 +/- 0.02	2
G7	1.16 +/- 0.02	3

It is apparent that the $E_{1/2}$ values are not solely artifacts of the DNA sequence. For example, G3 and G4 both contain four d(T_2AG_3) repeats, with all four on one strand for G4 and two on two different strands for G3. However, their $E_{1/2}$ values are significantly different, with G4 having a greater $E_{1/2}$ value. Base content alone also does not account for the trend apparent in Table 6.3, as both G1 and G5 both contain 16 guanines and 8 thymines in the quadruplex, but their $E_{1/2}$ values are significantly different. It has also been demonstrated in the discussion above that quadruplexes with the same strand stoichiometries have different $E_{1/2}$ values. For example, G4 and G6 could be distinguished

by their $E_{1/2}$ values, even though both form intramolecular quadruplexes. Therefore, we ascribe the differences in $E_{1/2}$ values of the quadruplexes to differences in gas-phase stabilities of different quadruplex structures.

6.3.6 Molecular Dynamics Structure and Energy Analysis

To provide insight into the relative stabilities of the quadruplexes determined by the energy-variable dissociation experiments, molecular dynamics energy analyses were performed on the quadruplexes. The root mean square displacement (RMSD) is one of the key parameters used to evaluate the stability of a MD trajectory. The smaller the RMSD values, the more similar the MD snapshots are to the reference structure. In this work, we were interested in the relative stabilities of the G-tetrads in the gas phase MD simulations. Therefore, only the G-tetrad atoms were used in least-square fittings to calculate the RMSD values, with the final structures of a series of minimizations used as references. The RMSD of G-tetrads during the MD simulations are listed in Table 6.4. It is shown that G1-ammonium has the most stable MD trajectory and its mean RMSD of G-quartets is only 0.66 Å which is significantly smaller than all other model molecules including G1-no-cation.

The molecular mechanical energies and the entropies, as well as the free energies of the quadruplexes, are listed in Table 6.4. There is more than one strand orientation for G3, G4, G5, and G6. According to the calculated free energies, the antiparallel strand orientation is more favorable for G4 and G6, and only slightly more favorable for G3. On the other hand, the free energy of parallel G5 is virtually the same as the antiparallel orientation, indicating that G5 likely forms a mixture of parallel and antiparallel structures.

Table 6.4: Root mean square displacements (RMSD) of guanines in G-tetrads and the free energies and their components of the final minimized structures and 50 snapshots collected from the MD simulations. For a quadruplex, the energetically more favorable stoichiometry or strand orientation according to the free MD free energies is in bold.

Quadruplex	Mean RMSD (Å)	Energy (kcal/mol)		Entropy (cal/mol/K)		Free Energy (kcal/mol)	
		Minimization	MD	Minimization	MD	Minimization	MD
G1-ammonium	0.66	-1840.9	-1962.6	2132.7	2044.9	-2480.7	-2576.0
G1-no-cation	2.32	-1680.6	-1795.6	2112.4	2038.9	-2314.3	-2407.3
G2-antiparallel	3.22	-1779.3	-1868.3	2091.2	2049.5	-2406.7	-2483.2
G3-antiparallel	3.82	-1593.8	-1782.1	2210.2	2041.7	-2256.9	-2394.6
G3-parallel	2.96	-1625.7	-1782.5	2098.6	2016.8	-2256.3	-2387.5
G4-antiparallel	3.09	-1645.2	-1818.0	2212.0	2034.5	-2308.8	-2428.4
G4-parallel	2.23	-1668.4	-1803.1	2113.2	2038.0	-2302.4	-2414.5
G5-antiparallel	3.96	-1719.8	-1917.8	2194.2	2053.1	-2378.1	-2533.4
G5-parallel	2.56	-1773.4	-1921.9	2125.9	2041.8	-2411.1	-2534.4
G6-antiparallel	3.58	-1174.6	-1534.1	2259.3	2042.1	-1852.4	-2146.7
G6-parallel	2.70	-1225.2	-1514.6	2313.0	2034.9	-1919.1	-2125.1
G6- T-tetrad	4.44	-1332.7	-1515.8	2083.9	2026.7	-1957.9	-2123.8

The relative free energies between the minimized and the MD structures are listed in Table 6.5. The hydrogen bonding and stacking interaction analysis on the MD trajectories were performed (data not shown). The overall conclusion is that the number of stacking interactions and the number of hydrogen bonds among guanine tetrads slightly decrease in comparison to the final minimized structures, while the number of hydrogen bonds formed between guanine tetrads and other atoms increases significantly.

To estimate the relative stability of the quadruplexes, one needs to consider many factors including RMSD values and relative free energies between the minimized and the

MD structures. It can be argued that the smaller the distortion of the G-quartets, the more stable the G-quadruplex and the higher the $E_{1/2}$ value. This translates to the smaller the relative free energy between the distorted and the undistorted structures, the more stable the G-quadruplex

If only the mean RMSD are considered, the rank order of stability is: G1 > G5 > G4 > G2 > G6 > G3, while the rank order becomes G2 > G1 > G4 > G5 > G3 > G6 when the relative free energies are taken into account. Thus, according to structural and energetic analysis, G1 and G2 are more stable than G4 and G5, and G4 and G5 are more stable than G3 and G6 according to the structural and energetic analysis. It should be noted that the above comparisons were made only for the energetically more favorable stoichiometry or strand orientation of each quadruplex, and the rank order may be changed if one considers the contribution of other strand orientations.

Table 6.5: The relative free energies between the last minimized and MD structures. For each quadruplex, only the energetically more favorable stoichiometry or strand orientation is considered.

Quadruplex	Relative Free Energy (kcal/mol)
G1-ammonium	95.3
G2-antiparallel	93
G3-antiparallel	137.7
G4-antiparallel	119.6
G5-parallel	123.3
G6-antiparallel	165.9

The conclusions about the relative quadruplex stabilities determined by RMSD and relative free energy results correlate well with the rankings determined by the energy variable CAD experiments. One exception to the general agreement is the result for G2.

This quadruplex had the lowest $E_{1/2}$ value, however it was determined to be the most stable quadruplex according to the free energy analysis and ranked fourth based on RMSD comparison. This discrepancy could be attributed to the fragmentation pathways of this structure in the gas phase. Previous studies of the fragmentation mechanisms of single strand and duplex oligonucleotides have found that base loss and subsequent back bone fragmentation is initiated by protonation of the nucleobase, followed by covalent cleavage.[50] In duplex DNA, the terminal ends of the duplex must unzip to accommodate this nucleobase loss. When high-proton affinity bases such as guanine are located at the terminal positions of a duplex, more extensive nucleobase loss and further unzipping will occur, leading to decreased stability of the quadrupelxes.[29] If a similar fragmentation mechanism is assumed for quadruplex DNA, G2 can undergo more facile nucleobase loss and thus has decreased gas-phase stability since it is the only quadruplex containing G-tracts on both the 5' and 3' ends of the strands forming the quadruplex. The gas-phase stabilities of the other quadruplexes do not appear to be affected by this anomaly.

Aside from G2, the results of the molecular dynamics simulation support the general conclusions of the energy-variable dissociation study. Of the quadruplexes assessed, the parallel four stranded structure G1 was the most stable, likely due to the four-tetrads and stabilizing NH_4^+ adducts that did not remain bound to the other structures. The next most stable structures were the intramolecular quadruplexes G4 and G5 which have approximately equal gas-phase stabilities. While G5 has more tetrads than G4 with four compared to three, the quadruplexes have different strand orientations as determined by the molecular dynamics simulations. The parallel strand orientation of G5 may negate the stabilizing effects of the additional tetrad compared to the antiparallel orientation of G4. The least stable quadruplexes are the double stranded antiparallel G3 with three tetrads and the intramolecular quadruplex G6 with two tetrads. G6 has fewer tetrads than

both G4 and G5 so it is unsurprisingly less stable. G3 and G4 have the same number of tetrads, however the antiparallel intramolecular orientation of G4 apparently increases its gas-phase stability relative to G3.

6.4 CONCLUSIONS

The CAD fragmentation patterns and $E_{1/2}$ values for a series of quadruplexes have been acquired to assess their relative gas-phase stabilities. State of the art molecular dynamics simulations have also been applied to investigate the structural and energetic properties of G-quadruplexes in solvent-free environment. It is concluded that G-quadruplexes can maintain their structures in the gas phase although the G-quartets are distorted to some degree. The ammonium ions were also shown to play an important rule in stabilizing the G-quartets for a four-stranded parallel quadruplex. Free energy analysis is useful not only in identifying the most favorable stoichiometry or strand orientation, but also for ranking the relative stability of a set of quadruplexes. The relative gas-phase stabilities of the quadruplexes determined by the molecular dynamics simulations correlates well with the ranking determined from $E_{1/2}$ values of the quadruplexes, suggesting energy-variable dissociation curves are also useful as a means for rapidly screening the stability of quadruplexes.

6.5 REFERENCES

(1) Rezler, E. M.; Bearss, D. J.; Hurley, L. H. *Curr. Opin. Pharmacol.* **2002**, *2*, 415-423.

(2) Bearss, D. J.; Hurley, L. H.; Von Hoff, D. D. *Oncogene* **2000**, *19*, 6632-6641.

(3) Rezler, E. M.; Bearss, D. J.; Hurley, L. H. *Annu. Rev. Pharmacol. Toxicol.* **2003**, *43*, 359-379.

(4) Wellinger, R. J.; Sen, D. *Eur. J. Cancer* **1997**, *33*, 735-749.

(5) Harley, C. B.; Futcher, A. B.; Greider, C. W. *Nature* **1990**, *345*, 458-460.

(6) Kim, N. W.; Piatyszek, M. A.; Prowse, K. R.; Harley, C. B.; West, M. D.; Ho, P. L. C.; Coviello, G. M.; Wright, W. E.; Weinrich, S. L.; Shay, J. W. *Science* **1994**, *266*, 2011-2015.

(7) Nakamura, T. M.; Morin, G. B.; Chapman, K. B.; Weinrich, S. L.; Andrews, W. H.; Lingner, J.; Harley, C. B.; Cech, T. R. *Science* **1997**, *277*, 955-959.

(8) Moyzis, R. K.; Buckingham, J. M.; Cram, L. S.; Dani, M.; Deaven, L. L.; Jones, M. D.; Meyne, J.; Ratliff, R. L.; Wu, J. R. *Pro.Natl. Acad. Sci. U. S. A.* **1988**, *85*, 6622-6626.

(9) Blackburn, E. H.; Gall, J. G. *J. Mol. Biol.***1978**, *120*, 33-53.

(10) Haider, S.; Parkinson, G. N.; Neidle, S. *J. Mol. Biol.* **2002**, *320*, 189-200.

(11) Tohl, J.; Eimer, W. *Biophys. Chem.* **1997**, *67*, 177-186.

(12) Kerwin, S. M. *Curr. Pharm. Des.* **2000**, *6*, 441-471.

(13) Mergny, J. L.; Mailliet, P.; Lavelle, F.; Riou, J. F.; Laoui, A.; Helene, C. *Anti-Cancer Drug Des.* **1999**, *14*, 327-339.

(14) Hurley, L. H.; Wheelhouse, R. T.; Sun, D.; Kerwin, S. M.; Salazar, M.; Fedoroff, O. Y.; Han, F. X.; Han, H. Y.; Izbicka, E.; Von Hoff, D. D. *Pharmacol. Ther.* **2000**, *85*, 141-158.

(15) Han, H. Y.; Hurley, L. H. *Trends Pharmacol. Sci.* **2000**, *21*, 136-142.

(16) David, W. M.; Brodbelt, J.; Kerwin, S. M.; Thomas, P. W. *Anal. Chem.* **2002**, *74*, 2029-2033.

(17) Mazzitelli, C. L.; Kern, J. T.; Rodriguez, M.; Brodbelt, J. S.; Kerwin, S. M. *J. Am. Soc. Mass Spectrom.* **2006**.

(18) Baker, E. S.; Lee, J. T.; Sessler, J. L.; Bowers, M. T. *J. Am. Chem. Soc.* **2006**, *128*, 2641-2648.

(19) Rosu, F.; De Pauw, E.; Guittat, L.; Alberti, P.; Lacroix, L.; Mailliet, P.; Riou, J. F.; Mergny, J. L. *Biochemistry* **2003**, *42*, 10361-10371.

(20) Guittat, L.; Alberti, P.; Rosu, F.; Van Miert, S.; Thetiot, E.; Pieters, L.; Gabelica, V.; De Pauw, E.; Ottaviani, A.; Riou, J. F.; Mergny, J. L. *Biochimie* **2003**, *85*, 535-547.

(21) Guittat, L.; De Cian, A.; Rosu, F.; Gabelica, V.; De Pauw, E.; Delfourne, E.; Mergny, J. L. *Biophys. Acta* **2005**, *1724*, 375-384.

(22) Carrasco, C.; Rosu, F.; Gabelica, V.; Houssier, C.; De Pauw, E.; Garbay-Jaureguiberry, C.; Roques, B.; Wilson, W. D.; Chaires, J. B.; Waring, M. J.; Bailly, C. *ChemBioChem* **2002**, *3*, 1235-1241.

(23) Rosu, F.; Gabelica, V.; Shin-ya, K.; De Pauw, E. *Chem. Commun.* **2003**, 2702-2703.

(24) Gidden, J.; Ferzoco, A.; Baker, E. S.; Bowers, M. T. *J. Am. Chem. So.c* **2004**, *126*, 15132-15140.

(25) Gabelica, V.; De Pauw, E. *Int. J. Mass Spectrom.* **2002**, *219*, 151-159.

(26) Pan, S.; Sun, X. J.; Lee, J. K. *Int. J. Mass Spectrom.* **2006**, *253*, 238-248.

(27) Guo, X. H.; Bruist, M. F.; Davis, D. L.; Bentzley, C. M. *Nucleic Acids Res.* **2005**, *33*, 3659-3666.

(28) Guo, X. H.; Liu, Z. Q.; Liu, S. Y.; Bentzley, C. M.; Bruist, M. F. *Anal. Chem.* **2006**, *78*, 7259-7266.

(29) Wan, K. X.; Gross, M. L.; Shibue, T. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 450-457.

(30) Gabelica, V.; De Pauw, E. *J.Mass Spectrom.* **2001**, *36*, 397-402.

(31) Baker, E. S.; Bernstein, S. L.; Gabelica, V.; De Pauw, E.; Bowers, M. T. *Int. J. Mass Spectrom.* **2006**, *253*, 225-237.

(32) Rosu, F.; Gabelica, V.; Houssier, C.; Colson, P.; De Pauw, E. *Rapid Commun. Mass Spectrom.y* **2002**, *16*, 1729-1736.

(33) Rueda, M.; Kalko, S. G.; Luque, F. J.; Orozco, M. *J. Am. Chem. Soc.* **2003**, *125*, 8007-8014.

(34) Rueda, M.; Luque, F. J.; Orozco, M. *J. Am. Chem. Soc.* **2005**, *127*, 11690-11698.

(35) Rueda, M.; Luque, F. J.; Orozco, M. *J. Am. Chem. Soc.* **2006**, *128*, 3608-3619.

(36) Kerwin, S. M.; Chen, G.; Kern, J. T.; Thomas, P. W. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 447-450.

(37) Caceres, C.; Wright, G.; Gouyette, C.; Parkinson, G.; Subirana, J. A. *Nucleic Acids Res.* **2004**, *32*, 1097-1102.

(38) Horvath, M. P.; Schultz, S. C. *J. Mol. Biol.* **2001**, *310*, 367-377.

(39) Muller, S.; Diederichs, K.; Breed, J.; Kissmehl, R.; Hauser, K.; Plattner, H.; Welte, W. *J. Mol. Biol.* **2002**, *315*, 141-153.

(40) Wang, Y.; Patel, D. J. *Structure* **1993**, *1*, 263-282.

(41) Case, D. A.; Darden, T. A.; Cheatham, T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Merz, K. M.; Wang, B.; Pearlman, D. A.; Crowley, M.; Brozell, S.; Tsui, V.; Gohlke, H.; Mongan, J.; Hornack, V.; Cui, G.; Berzoa, P.; Schafmeister, C.; Caldwell, J. W.; Ross, W. S.; Kollman, P. A. **2005**, AMBER 8, University of California, San Francisco.

(42) Wang, J. M.; Cieplak, P.; Kollman, P. A. *Journal of Comput. Chem.* **2000**, *21*, 1049-1074.

(43) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *Journal of Comput. Phys.* **1977**, *23*, 327-341.

(44) Srinivasan, J.; Cheatham, T. E.; Cieplak, P.; Kollman, P. A.; Case, D. A. *J.Am. Chem. Soc.* **1998**, *120*, 9401-9409.

(45) Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S. H.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.; Srinivasan, J.; Case, D. A.; Cheatham, T. E. *Acc. Chem. Res.* **2000**, *33*, 889-897.

(46) Wang, J.; Hou, T.; Xu, X. *Current Computer-Aided Drug Des.* **2006**, *2*, 287-306.

(47) Keller, K. M.; Zhang, J. M.; Oehlers, L.; Brodbelt, J. S. *J. Mass Spectrom.* **2005**, *40*, 1362-1371.

(48) Gabelica, V.; De Pauw, E. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 91-98.

(49) Pan, S.; Verhoeven, K.; Lee, J. K. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 1853-1865.

(50) Gale, D. C.; Smith, R. D. *J. Am. Soc. Mass Spectrom.* **1995**, *6*, 1154-1164.

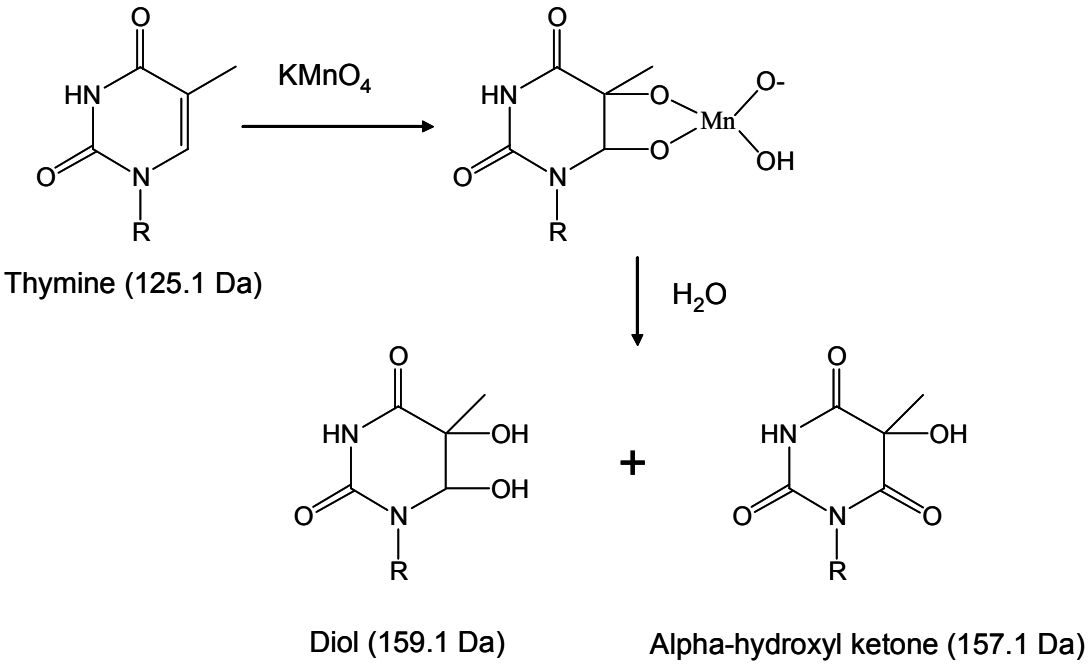
Chapter 7: Probing Ligand Binding to Duplex DNA using $KMnO_4$ Reactions and Electrospray Ionization Tandem Mass spectrometry

7.1 INTRODUCTION

Electrospray ionization mass spectrometry (ESI-MS) has been established as a useful tool for the analysis of non-covalent complexes formed between small molecule drugs and DNA.[1-3] The advantages of using ESI-MS in this capacity are low sample consumption coupled with fast analysis time, which make the technique adaptable to high throughput screening techniques. In early reports, the binding of well-studied, commercially available drugs to duplex DNA was examined, with results indicating that the binding trends observed by ESI-MS such as binding stoichiometries, sequence selectivities, binding affinities and complex stabilities, can be correlated to known solution binding behavior.[4-10] Recent accounts have extended the use of the technique to look at more novel types of ligand/DNA complexes such as those containing bis-intercalators,[11] ligands that are metal-mediated,[12, 13] and quadruplex DNA,[14-21] further demonstrating the capabilities of the method.

While the success of using ESI-MS to screen binding affinities and stoichiometries has been established, one of the drawbacks of the technique is the lack of structural information that can be directly obtained; drug/DNA complexes are observed, but information about the drug binding site or alterations in the DNA structure due to drug binding can not be easily determined. Chemical probes that modify nucleic acid substrates in a structure-dependant manner[22, 23] coupled with ESI-MS/MS analysis represent a promising technique for the structural analysis of DNA and RNA. Solvent accessibility probes (such as dimethylsufate (DMS), kethoxal, and 1-cyclohexyl-3-(2-morpholinoethyl)carbodiimide metho-*p*-toluenesulfonate (CMCT)) that modify

nucleotides that are not base-paired or involved in binding interactions with proteins or ligands have been used with Fourier transform mass spectrometry (FTMS) to examine structural features of RNA[24-26] and RNA/protein[27] complexes. In these previous studies, tandem mass spectrometry or chemical digestion techniques were used to sequence the RNA and identify the site of chemical modification, thereby eliminating the need for polyacrylamide gel electrophoresis (PAGE) analysis of the products, resulting in a fast, powerful analytical method for RNA and RNA/ligand complexes.[24-27]



Scheme 7.1: Mechanism of oxidation of thymine nucleobases by permanganate. The mass of thymine and the resulting diol and alpha-hydroxyl ketone products are given in parenthesis.

In the present study, we have exploited the use of chemical probes with ESI-MS analysis by developing a technique that utilizes $KMnO_4$ to detect conformational changes in duplex DNA and identify the ligand binding site in non-covalent drug/DNA

complexes. Under neutral conditions, $KMnO_4$ oxidizes thymine and, to a much lesser extent, cytosine by attacking the double bond of the nucleobase to produce a mixture of the diol and an alpha-hydroxylketone products as demonstrated in Scheme 1.[22, 23, 28] Each oxidized base results in a mass shifts of +34 for the diol and +32 Da for the alpha-hydroxylketone, which can easily be detected by mass spectrometry. Thymines in unstacked, single strand-like DNA are susceptible to reaction with $KMnO_4$ while double stranded DNA is resistant to the oxidation.[29] As a result, $KMnO_4$ has been used to detect conformational changes upon ligand[30, 31] and protein binding,[32-35] determine the base pair composition,[36, 37] uncover the specific structure of DNA (Z-DNA, hairpins, curvatures, parallel stranded helices, etc.),[38-41] and ascertain conformational changes of mismatched DNA, in conjunction with gel electrophoresis analysis.[42, 43]

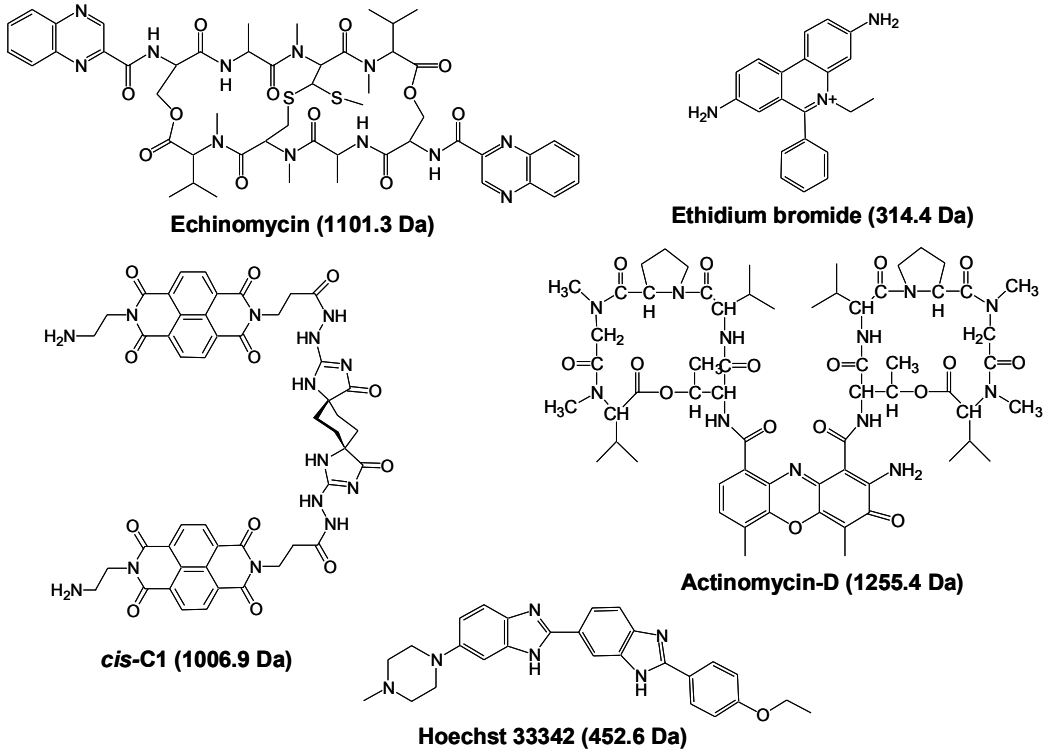


Figure 7.1: Structures of ligands used in this study with molecular weights given in parenthesis.

In past studies using the $KMnO_4$ reaction to examine conformational changes upon small molecule binding, it was determined that the permanganate ion reacts with thymines around the ligand binding site as a result of the unwinding and extension of the duplex upon intercalator binding.[30, 31] Therefore, ligand binding sites can be elucidated after determining the positions of the oxidized thymines. The site of oxidation is determined using a piperidine heat treatment to cleave the DNA at the site of the oxidation, followed by polyacrylamide gel electrophoresis (PAGE) analysis to sequence the fragments.[30, 31] We have developed a streamlined technique to oxidize drug/DNA complexes in solution, analyze the extent of oxidation using ESI-MS and determine the site of oxidation using two tandem mass spectrometry techniques: collisional activated dissociation (CAD)[44] and infrared multiphoton dissociation (IRMPD).[45, 46] Our technique simplifies the experiment by eliminating the need for time consuming and labor intensive PAGE analysis and use of radiolabeled DNA, while also offering detailed binding site and sequence information. We establish the technique by examining the oxidation of the duplex DNA of different sizes and sequences as well as drug/DNA complexes containing echinomycin, ethidium bromide, actinomycin-D, Hoechst 33342, and *cis*-C1 (structures shown in Figure 7.1).

7.2 EXPERIMENTAL

7.2.1 Materials

Single strand oligonucleotides were purchased from Integrated DNA Technologies (Coralville, IA) and synthesized as ammonium salts on the 1 μmole scale and purified by reverse phase HPLC. Actinomycin-D, ethidium bromide, Hoechst 33342, and $KMnO_4$ were purchased from Thermo Fisher Scientific (Waltham, MA) and used without further purification. Echinomycin was purchased from Sigma-Aldrich (St. Louis,

MO) and used without further purification and *cis*-C1 was prepared as previously described.[47] All solvents were of HPLC grade purity.

7.2.2 $KMnO_4$ Reaction

Prior to the $KMnO_4$ oxidation reactions, stock solutions of each single strand oligo were prepared at 2 mM in 250 mM ammonium acetate buffer. Solutions containing complementary single strands, each at 1 mM concentration were prepared in 250 mM ammonium acetate and annealed by heating the solution to 90 °C, followed by cooling to room temperature over a period of 7 hours. Stock solutions of the drug molecules were prepared at 1 mM concentration in water for actinomycin-D, Hoechst 33342 and *cis*-C1, and in methanol for echinomycin and ethidium bromide.

Solutions containing a DNA duplex or single strand and were prepared at 40 µM in 50 µL of 90 mM ammonium acetate. Where indicated, a drug was added at 120 µM and the solution was allowed to equilibrate for 30 min. to allow for binding. To initiate oxidation 5 µL of a 5 mM $KMnO_4$ solution prepared in water was added, and the reaction mixture was incubated for 4-30 min. at room temperature. After the desired incubation time, $KMnO_4$ was immediately removed from the solution using a Pierce PepClean C_{18} spin column (Rockford, IL). The DNA was eluted from the column using 40 µL of 50% acetonitrile solution and then diluted to 100 µL so that the final solution contained 50 mM ammonium acetate.

7.2.3 Mass Spectrometry

Analytical solutions were directly infused into a ThermoFinnigan LCQ Duo mass spectrometer (San Jose, CA) at 3 µL/min using a Harvard Apparatus PHD 2000 syringe pump (Holliston, MA). Negative ions were produced using an ESI voltage of 3.5 kV and a heated capillary temperature of 110 °C. Nitrogen sheath and auxiliary gas flows of 40 and 20 arbitrary units, respectively were used to aide the desolvation of the ions. Spectra

were acquired using an ionization time of 50 -100 ms and by summing 300 scans. For CAD experiments the desired precursor ion was isolated in the trap using default activation time of 30 ms and a q_z value of 0.25. The collision energy was increased until the abundance of the precursor ion was reduced to ~10% relative abundance. The base pressure of the trap was nominally 1 x 10^{-5} torr.

IRMPD experiments were undertaken on a modified ThermoFinnigan LCQ Deca XP mass spectrometer equipped with model 48-5 Synrad 50 W continuous wave laser (Mukilteo, WA) that has been described previously.[46] Briefly, q_z values of 0.2 to 0.1 were used and the laser was triggered during the activation portion of the scan function. Activation times of 1 to 3 ms at 50 W laser power were used to achieve dissociation. The base pressure of the trap was nominally 2.8 x 10^{-5} torr which corresponds to a helium pressure of 1 mTorr.

7.3 RESULTS AND DISCUSSION

The capabilities of the new ESI-MS method involving the the $KMnO_4$ reaction will be evaluated by examining the oxidation of single strand and then duplex DNA. The oxidation of drug/duplex DNA complexes with varying duplex sequences, drug molecules, and duplex length will also be presented. The use of CAD and IRMPD to identify the position of the oxidized thymines is then demonstrated.

7.3.1 Oxidation of Single Strand DNA

To confirm that the oxidation protocol, subsequent work-up, and mass spectrometric analysis allows effective monitoring and detection of the oxidation of thymines in DNA, the $KMnO_4$ reaction was first carried out on solutions containing the single strand d(GCGGATATATGGCG), which contains three possible sites for thymine oxidation. Figure 7.2a shows the ESI mass spectrum of a solution containing the single strand prior to reaction with $KMnO_4$. To obtain this spectrum, the reaction and sample

purification procedures were followed as described in the Experimental section, except the $KMnO_4$ was not added to the solution. The 3- and 4- charge states of the single strand are present with very low abundance sodium adducts. The spectral region around $[ss]^{4-}$ (*m/z* 1070 to 1115) is shown in the inset of Figure 7.2a, and the sodium adduct ions are labeled with asterisks. The corresponding spectrum of the single strand after a 5 min. reaction with $KMnO_4$ is shown in Figure 7.2b. Compared to Figure 7.2a, new ions are present in Figure 7.2b and the *m/z* values of the new ions are shifted by multiples of +34 Da compared to the *m/z* of the single strand ions. The *m/z* values of the new ions are consistent with the formation of up to three oxidized thymines on the single strand ions via the reaction shown in Scheme 1. The ions in Figure 7.2 that correspond to single strand DNA containing the oxidized thymine(s) are labeled with "a black triangle", and the number of oxidation adducts is indicated by the number in parenthesis. The formation of up to three adducts is observed and is expected since three thymines are present in the single stand. The spectral enlargements of the region around $[ss]^{4-}$ also demonstrate that peaks pertaining to sodium adducts (again labeled with asterisks) can be distinguished from those containing oxidation adducts (labeled with black triangles).

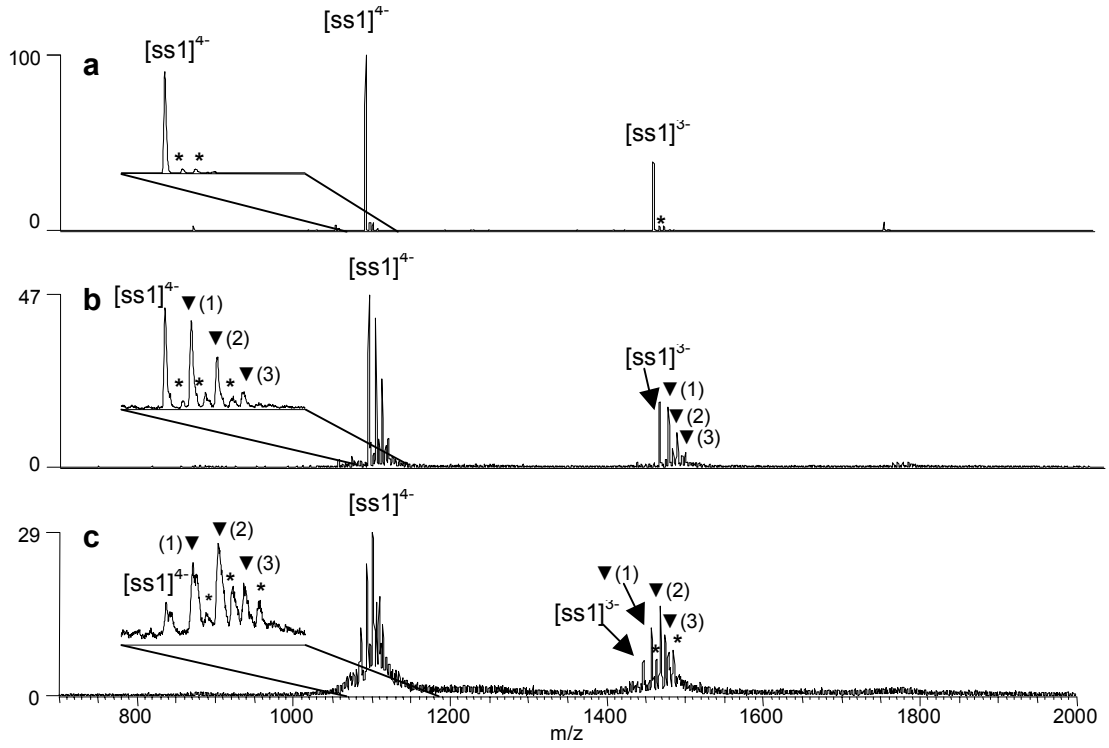


Figure 7.2: ESI-mass spectra showing solutions containing the single strand d(GCGGATATATGGCG) (ss1) (a) before reaction with KMnO4, (b) after 4 min.reaction with KMnO4, and (c) after 30 min. reaction with KMnO4. Spectral elargements of the region around m/z 1070 to 1115 are shown in the inset. Sodium adducts are labeled with an asterisk. Ions containing oxidized thymines are labeled with a black triangle, with the number in parenthesis indicating the number of oxidation adducts.

After 30 min. reaction with $KMnO_4$, the oxidation of the single strand is more extensive, as demonstrated by the spectrum shown in Figure 7.2c. Up to three oxidation adducts are still observed, and the abundance of the ions containing the oxidation adducts relative to the single strand ions containing no adducts is significantly increased in Figure 7.2c compared to Figure 7.2b. These results indicate that at longer incubation times, more extensive oxidation of the single strand DNA occurs. These initial results confirm that the

oxidation process, as well as the extent of oxidation as a function of time, can be monitored by ESI-MS.

7.3.2 Oxidation of Duplex DNA

After observing significant oxidation of single strand DNA, a similar series of experiments were performed with a duplex to compare the degree of thymine oxidation. The duplex d(GCGGATATATGGCG/CGCCATATATCCGC) (ds1) was used and contains the same single strand used in the oxidation experiments described above, then annealed with its complementary strand. The spectrum of the duplex prior to oxidation is shown in Figure 7.3a. The dominant charge state of the duplex is 5-, and the region around this ion (*m/z* 1700-1730) is enlarged and shown in the inset. After 4 min. reaction with $KMnO_4$, two low abundance ions are detected with mass shifts that are consistent with the formation of duplexes containing one or two oxidized thymines (Figure 7.3b). The abundance of these adducts relative to the unoxidized $[ds]^{5-}$ ion is significantly lower than the abundance of the oxidation adducts formed after the $KMnO_4$ reaction with the single strand oligonucleotide d(GCGGATATATGGCG) (Figure 7.2b). After 30 min. reaction with the permanganate ion, the relative abundances of the oxidation adducts of the duplex do not significantly change, as demonstrated by the spectrum shown in Figure 7.3c. Furthermore, no more than two oxidation adducts are formed even though there are six thymines in the duplex sequence. It is also interesting to note that increasing the reaction time from 4 to 30 min. resulted in more extensive oxidation of the single strand (Figure 7.2b versus 7.2c), while the degree of oxidation of the duplex did not change. The striking differences between the spectra obtained for the oxidation of single strand (Figure 7.2) and duplex DNA (Figure 7.3) demonstrates the resistance of duplex DNA to $KMnO_4$ oxidation. This resistance stems from the stacking of the nucleobases which sterically hinders the access and attack of the permanganate ion on the double bond of

thymine.[28] Based on the results shown in Figures 7.1 and 7.2, a reaction period of 20 min. was selected for all subsequent $KMnO_4$ reactions since this time allowed sufficient reaction of thymines on single strand DNA, while the extent of oxidation of the duplex did not change substantially at even longer times.

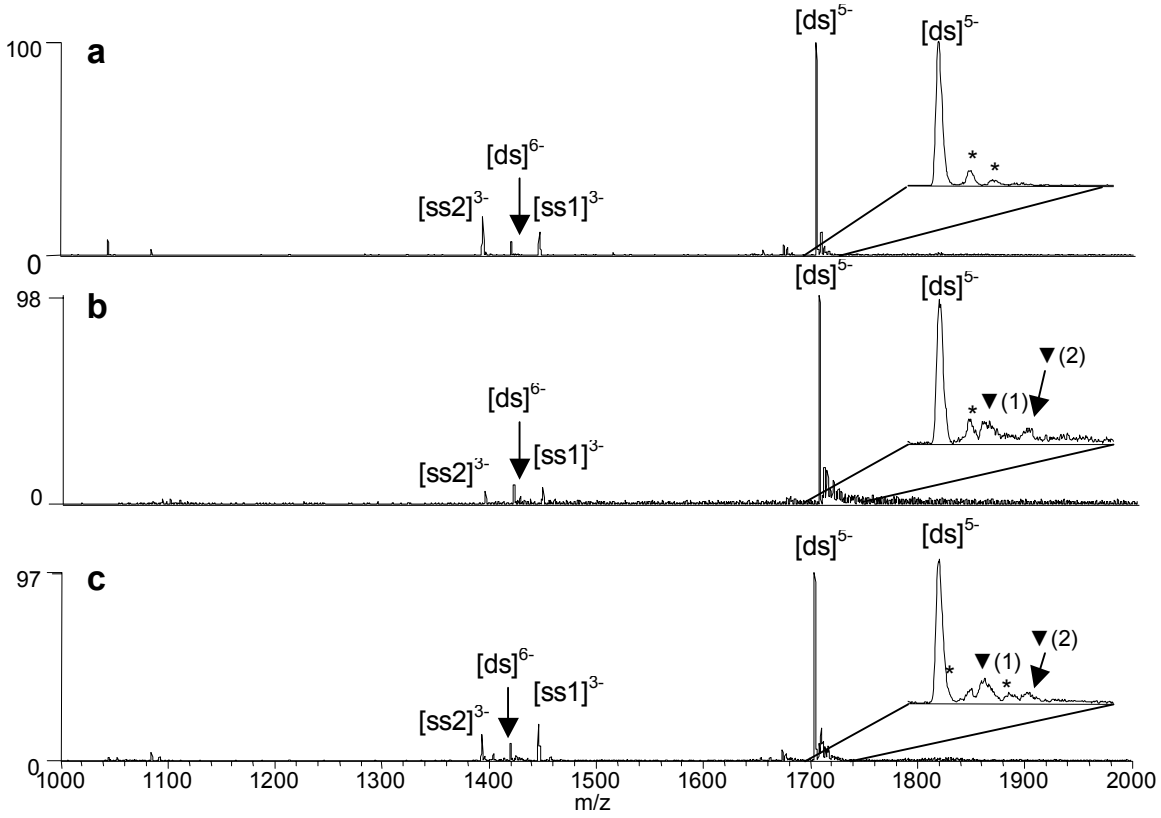


Figure 7.3. ESI-mass spectra showing solutions containing the duplex dGCGGATATATGGCG/CGCCATATATCCGC) (a) before reaction with KMnO4, (b) after 4 min.reaction with KMnO4, and (c) after 30 min. reaction with KMnO4. Spectral elargements of the region around m/z 1700 - 1730 are shown in the inset. Sodium adducts are labeled with an asterisk. Ions containing oxidized thymines are labeled with a black triangle, with the number in parenthesis indicating the number of oxidation adducts.

7.3.3 Oxidation of DNA Complexes Containing Echinomycin

The $KMnO_4$ oxidation of drug/DNA duplex complexes was also examined by ESI-MS to assess the ability to identify or predict structural changes of DNA upon ligand binding based on the observed changes in the oxidation patterns. Drug binding sites also can be determined from the oxidation pattern since the thymines that encompass the drug binding site are most susceptible to oxidation.[30, 31] Previous PAGE-based studies have found that intercalators induce changes to the DNA conformation, and in doing so, greatly increasing the reactivity of thymines to the permanganate ions.[30, 31] This effect is especially pronounced for echinomycin (Figure 7.1), a bis-intercalator antibiotic that binds to duplex DNA via the intercalation of two quinoxaline chromophores at CpG sites, with the central bicyclic peptide residing in the minor groove. Intercalation of echinomycin causes considerable unwinding and distortion of the DNA duplex.[48-51] In the present study, the duplex d(GCGGATATATGGCG/CGCCATATATCCGC) (ds1) was selected for experiments with echinomycin since this sequence contains multiple CpG ligand binding sites. Prior to examining the oxidation of the echinomycin/duplex complexes, the oxidation of the duplex in the absence of the ligand was assessed as a control. The results are shown in Figure 7.4a, and they are virtually the same as those shown in Figure 7.3b and 7.3c for the same duplex, respectively.

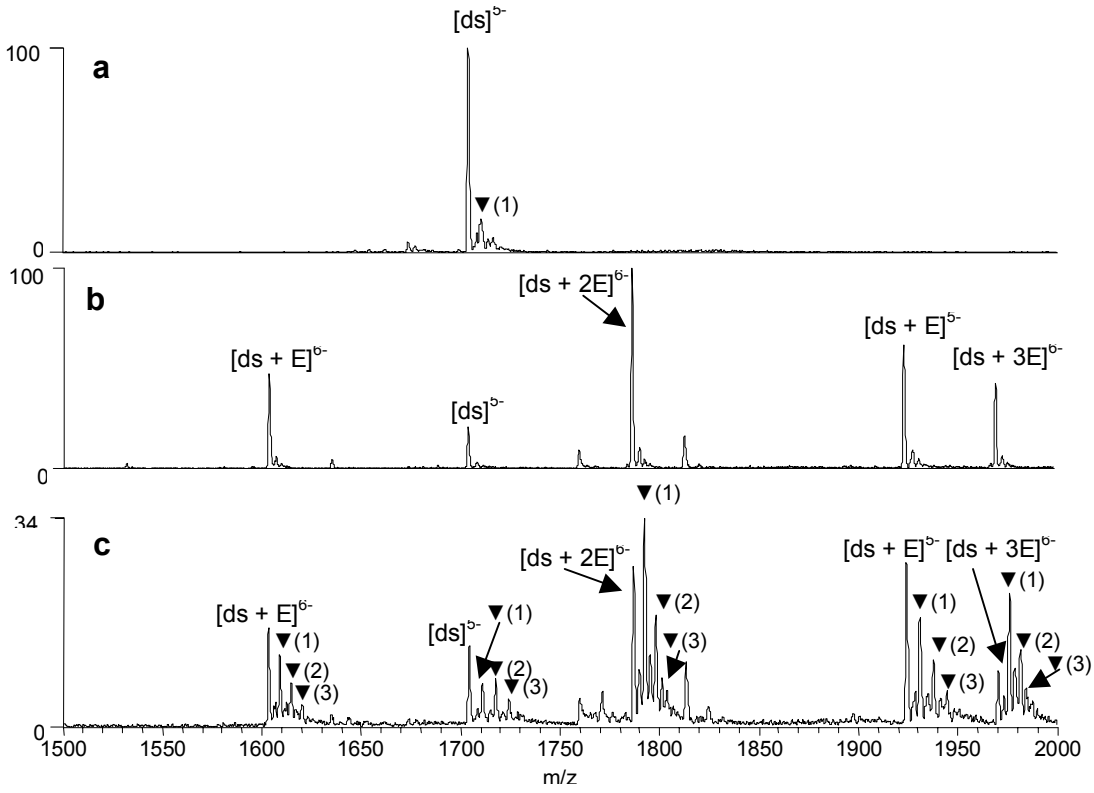


Figure 7.4: ESI-mass spectra showing solutions containing the duplex dGCGGATATATGGCG/CGCCATATATCCGC) (a) after 20 min, reaction with KMnO4, (b) with echinomycin (E), prior to reaction with KMnO4, and (c) with echinomycin (E), after 20 min. reaction with KMnO4. Ions containing oxidized thymines are labeled with a black triangle, with the number in parenthesis indicating the number of oxidation adducts.

Figure 7.4b shows the ESI mass spectrum of a solution containing the duplex with echinomycin (abbreviated as "E") prior to $KMnO_4$ reaction. Echinomycin/duplex complexes are observed with drug/DNA binding stoichiometries ranging from 1:1 to 3:1 in the 6- charge state, confirming that echinomycin/duplex complexes are readily formed in solution and maintained upon transport to the gas phase. Then the $KMnO_4$ reaction was undertaken for twenty minutes on a solution containing echinomycin/duplex complexes, and the resulting mass spectrum is shown in Figure 7.4c. Compared to the oxidation of

duplex alone (Figure 7.4a), the degree of oxidation is significantly increased for the echinomycin/duplex complexes. Interestingly, the degree of oxidation changes with the echinomycin/duplex binding stoichiometry; as the number of bound ligands increases, the abundance of the ions containing oxidation adducts increases relative to the unoxidized ions. This is demonstrated by comparing the relative abundances of the oxidation adducts associated with the 1:1, 2:1, and 3:1 echinomycin/duplex complexes in the 6- charge state. It is expected that the unwinding, elongation and hence the distortion of the duplex should increase with the number of intercalator ligands bound to the duplex, thus making thymines in the duplex more susceptible to reaction with $KMnO_4$.[51]

To allow the semi-quantitative comparison of the degree of oxidation of DNA complexes, we calculate the percent oxidation of a given DNA complex, M, using the following equation:

$$\text{Percent Oxidation} = \frac{A_{[M+O]} + A_{[M+2xO]} + \ldots A_{[M+nxO]}}{A_{[M]} + A_{[M+O]} + A_{[M+2xO]} + \ldots A_{[M+nxO]}} \times 100\%$$

where $A_{[M]}$ is the ion abundance of the single strand, duplex or drug/DNA complex designated in the subscript bracket, $A_{[M+O]}$ corresponds to each complex containing one or more oxidation adducts, and n is the maximum number of oxidation adducts associated with the particular complex [M]. For example, for the spectra shown in Figure 7.4, the percent oxidation of the duplex in the absence of echinomycin is 15%, while the oxidation values of the corresponding echinomycin/DNA complexes are 57% for the 1:1 complex, 69% for the 2:1 complex, and 82% for the oxidation value obtained for the 3:1 complex (all in the 6- charge state). The increase in the extent of oxidation with each additional echinomycin molecule bound to the duplex is reflected in the percent oxidation values. It is interesting to note that the percent oxidation of the single strand

d(GCGGATATATGGCG) shown in Figure 7.2b was calculated to be 85%, which is very similar to the 3:1 echinomycin/duplex DNA complex. Clearly the multiple intercalation of echinomycin causes a dramatic distortion of double stranded DNA.

7.3.3.1 Effect of Duplex DNA Sequence

The $KMnO_4$ oxidation of complexes containing echinomycin with another duplex, d(GTAGAGTCGACCTG/CAGGTCGACTCTAC) (ds2) was also examined (spectra not shown). The sequence of this duplex was selected from a fragment of a 265-mer DNA sequence used in a past study by Bailly et al. which reported the binding of echinomycin by DNase I, methidium propyl EDTA (MPE)·Fe^{II} , diethyl pyrocarbonate (DEPC), and $KMnO_4$.[31] This former study identified multiple binding sites along the DNA fragment including one strong site at the sequence 5'-AGTCGACCT-3' which is thus incorporated in our sequence, ds2. The $KMnO_4$ reaction with this duplex in the absence of echinomycin was undertaken, and the percent oxidation for the duplex was 46% (spectra not shown). This is a higher degree of oxidation than the value determined for ds1 in the absence of echinomycin (15%). While both duplexes contain an equal number of thymines, we speculate that the higher degree of oxidation of ds2 results from the specific location of the thymines. All of the thymines in ds1 are located in the center of the duplex and are flanked by G/C base pairs, while in ds2, there are three thymines closer to the termini of the duplex. Therefore, a slight unwinding of duplex ds2 in solution, a process more likely to occur near the ends of the duplex, may make the thymines more susceptible to permanganate oxidation even in the absence of an intercalating ligand.

While the variation in the degree of thymine oxidation in different DNA sequences is interesting, the change in the percent oxidation upon ligand binding is more important to the goals of this study. Echinomycin was found to form abundant complexes with ds2 with 1:1 and 2:1 binding stoichiometries (spectra not shown). After reaction

with $KMnO_4$, the formation of one and, to a lesser degree, two oxidation adducts were formed for each complex. As summarized in Table 7.1, the percent oxidation values of the 1:1, 2:1, and 3:1 echinomycin/duplex complexes are 62%, 66%, and 70%, respectively. While the overall increase in the percent oxidation values for the complexes containing echinomycin with ds2 was not as great as the increase for the echinomycin/ds1 complexes described in the previous section, the increase was still significant. These results confirm that the oxidation of DNA upon echinomycin binding[31] can be determined by the ESI-MS technique. As will be discussed in the sections below, the specific sites of thymine oxidation on the duplex can be determined by CAD, and they correlate with the findings of the past gel electrophoresis study.[31]

Table 7.1: Percent oxidation values of the duplexes and duplex/ligand complexes.[a]

	[ds][b]	[ds + L]	[ds + 2L]	[ds + 3L]
ds1/echinomycin[c]	15	57	69	82
ds2/echinomycin[d]	46	62	66	70
ds3/echinomycin[e]	56	100	100	nd
ds4/echinomycin[f]	32	53	67	nd
ds5/echinomycin[g]	18	51	67	69
ds5/actinomycin-D[g]	18	19	nd	nd
ds1/ethidium bromide[c]	15	40[h]	nd	nd
ds5/*cis*-C1[g]	18	40	nd	nd
ds1/Hoechst 33342[c]	15	20	nd	nd

[a] All values +/- 5.
[b] Values for the duplex alone were determined from the oxidation of solutions containing the duplex without ligand.
[c] ds1 = d(GCGGATATATGGCG/CGCCATATATCCGC)
[d] ds2 = d(GTAGAGTCGACCTG/CAGGTCGACTCTAC)
[e] ds3= d(GCAGTGA/TCACTGC)
[f] ds4= d(GGACAGTGAGGGCAGTGAGGG/CCCTCACTGCCCTCACTGTCC)
[g] ds5 =(GCGGGGATGGGGCG/CGCCCCATCCCCGC) was used in the experiment.
[h] Because the $KMnO_4$ oxidation caused a loss of ethidium bromide from the duplex, the values for ethidium bromide were determined based oxidation adducts formed on the free duplex.

7.3.3.2 Effect of Duplex DNA Length

The effect of duplex length on the $KMnO_4$ oxidation reaction was also examined. Our previous ESI-MS studies of drug/DNA complexes typically employed duplexes with 14 base-pairs because ion mobility/molecular dynamics studies have reported that DNA duplexes greater than 12 base pairs maintain the helical conformation in the gas-phase better than shorter duplexes.[52, 53] It may, however, be desirable to examine the $KMnO_4$ oxidation of smaller duplexes for high throughput applications and targeted screening strategies. Since the oxidation reactions are undertaken in solution and mass spectrometry is simply used to detect the products in the present methodology, concerns about gas-phase conformations are not a primary issue. Experiments involving a seven base pair duplex, d(GCAGTGA/TCACTGC) (ds3) were undertaken to assess the lower size limit of the duplexes that can be examined by the ESI-MS/oxidation technique. For this shorter duplex, the dominant DNA ions are detected in the 3- charge state. The percent oxidation value for the $[ds]^{3-}$ ion after $KMnO_4$ reaction in the absence of any DNA-interactive ligand was determined to be 56% (spectrum not shown). This value is greater than what was observed for the 14 base pair duplexes in the absence of ligand, and suggests more extensive distortion of the shorter duplex in solution. This result is consistent with past studies that found that the B-form of duplex DNA is more easily distorted in shorter duplexes due to the reduced base stacking and hydrophobic interactions.[54] For the present study, when the oxidation reaction is carried out on a solution containing echinomycin and d(GCAGTGA/TCACTGC), up to two oxidation adducts are formed for each 1:1 and 2:1 echinomycin/duplex complex. The percent oxidation for these complexes is 100% as no unoxidized form of the duplex remains in the ESI mass spectrum (data not shown). These results reflect the greater distortion of smaller duplexes upon ligand binding

compared to longer ones dues to the smaller number of hydrogen bonds and stacking interactions, as well as the substantial degree of distortion caused by the intercalator ligand.

Experiments with a longer 21 base pair duplex were also undertaken. The oxidation of the duplex, d(GGACAGTGAGGGCAGTGAGGG/CCCTCACTGCCCTCACTGTCC) (ds4) was first evaluated in the absence of echinomycin (spectrum not shown), and the percent oxidation was determined to be 32%. After echinomycin binding, the percent oxidation value of the 1:1 echinomycin:duplex complexes was determined to be 53% and the value for the 2:1 complex was 67% (spectrum not shown). While the increase in the degree of oxidation for ds4 compared to the echinomycin/ds4 complex is not as dramatic as was observed for the complexes containing the shorter duplexes (7-mers or 14-mers), it is reproducible and consistent with the expected duplex distortion caused by intercalation.

One of the main challenges of using longer DNA duplexes in an ESI-MS/$KMnO_4$ experiment is that at higher charge states, it is more difficult to distinguish the formation of an oxidized thymine from a sodium adduct which are always present in ESI mass spectra of nucleic acids. The dominant charge state for drug/DNA complexes containing 14 base pair duplexes is 6-. The observed mass shift upon formation of an oxidized thymine in the 6- charge state is 5.7 Da, while the shift for a sodium adduct is 3.8 Da. The *m/z* values of the sodium and oxidation adducts in this charge state can be easily distinguished using a quadrupole ion trap mass spectrometer. However, complexes containing the 21 base pair duplex are observed in the 9- and 8- charge states, and at these high charge states, the *m/z* differences between the sodium and oxidation adducts are decreased. In the 8- charge state, the ions corresponding to sodium adducts are shifted by 2.9 Da, while those of the oxidation adducts are shifted by 4.3 Da. These peaks can

generally distinguished in the mass spectrum; however, for the 9- charge state, the peaks for the sodium and oxidation adducts overlap since the mass shift for a sodium adduct is 2.5 Da, while the shift for an oxidation adduct is only 3.7 Da. Analyses of the $KMnO_4$ reaction with DNA present in the 9- charge state or greater require a mass analyzer with greater mass accuracy and resolving power.

7.3.4 Oxidation of Complexes Containing Other Drugs

$KMnO_4$ reactions with complexes containing other types of DNA interactive ligands, including three other intercalators and a minor groove binding agent, were also examined by ESI-MS. Actinomycin-D (Figure 7.1) binds to DNA duplexes via the intercalation of the phenoxazone chromophore, preferably at 5'-GC-3' sites, while the two cyclic peptides bind to the minor groove.[55, 56] As a monointercalator, the unwinding angle and elongation of the duplex upon actinomycin-D binding is less substantial compared to the bis-intercalator echinomycin,[51, 57, 58] therefore the $KMnO_4$ reactivity of an actinomycin-D/duplex complex is expected to be reduced compared to complexes containing echinomycin. To enhance the duplex binding by actinomycin-D, the duplex d(GCGGGGATGGGGCG/CGCCCCATCCCCGC) (ds5) was used due to its G/C rich sequence. The ESI mass spectrum of the solution containing the duplex with actinomycin-D prior to the oxidation reaction reveals abundant complexes with binding stoichiometries of 1:1 in the 6- and 5- charge state (spectrum not shown). As summarized in Table 7.1, the percent oxidation value for this duplex in the absence of ligand was determined to be 18% (spectrum not shown). After 20 min. reaction with $KMnO_4$, one low abundance oxidation adduct was formed for the actinomycin-D/duplex complexes (spectrum not shown). The percent oxidation for the 1:1 complexes was calculated to be 19%, which is not significantly different than the percent oxidation of the duplex in the absence of the ligand (Table 7.1).

For comparison, the $KMnO_4$ reaction with complexes containing echinomycin with ds5 were also examined. Echinomycin/duplex binding stoichiometries of 1:1, 2:1, and 3:1 were observed in the mass spectrum, and the corresponding percent oxidation values calculated for these complexes were 51%, 67%, and 69%, respectively. The smaller duplex distortion imparted by actinomycin-D compared to that of echinomycin is reflected in the lower percent oxidation values for these actinomycin-D/duplex complexes and is likely the result of the monointercalative binding mode of actinomycin-D. It is also possible that the cyclic peptides of the molecule hinder the access of the permanganate ions to the thymines.

Ethidium bromide is another monointercalator that has been shown to induce $KMnO_4$ reactivity in previous studies.[30] Like actinomycin-D, the DNA unwinding caused by ethidium bromide is not as significant as that of echinomycin,[51] resulting in decreased thymine hyperactivity. The conformational changes induced by ethidium bromide binding were thus examined by our oxidation/ESI-MS protocol. Prior to the $KMnO_4$ reaction, a solution containing ethidium bromide with ds1 was incubated, followed by C_{18} spin column purification. In the mass spectrum of the resulting solutions, abundant ethidium/duplexes complexes with binding stoichiometries of 1:1 and 2:1 were present in the mass spectra (spectrum not shown). The spectrum of the solution containing ethidium bromide with ds1 after the 20 min. reaction with $KMnO_4$ showed no ethdium bromide/ds1 complexes (spectra not shown). However, the abundances of the oxidation adducts of the free duplex in the 5- charge state were significantly greater than those of the oxidation adducts of ds1 in the absence of ligand. The percent oxidation of ds1 in the absence of ligand was 15%, while after the addition of the ethidium bromide, the percent oxidation was increased to 40% despite the apparent lack of retention of the ethidium bromide by the duplexes after $KMnO_4$ oxidation

These results suggest that ethidium bromide reacts with $KMnO_4$, causing it to dislodge from the duplex, either during the oxidation reaction. In the absence of $KMnO_4$, ethidium bromide remains bound to the duplex after the spin column clean-up step. The increase in the degree of oxidation of the drug-free $[ds]^{5-}$ ion in the spectrum obtained for the solution that contained both the duplex and ethidium bromide indicates that the ethidium bromide induced the thymine reactivity of the duplex *prior* to its dislocation. Thus, the $KMnO_4$ reactivity induced by ethidium can be assessed by examining the relative increase in the percent oxidation of the free duplex. The percent oxidation results summarized in Table 7.2 indicate that the duplex distortion induced by ethidium bromide was greater than that caused by actinomycin-D, but not as dramatic as that caused by echinomycin.

In a recent study, we examined the binding affinities of a series of new threading bis-intercalators[47] for different DNA duplexes by ESI-MS/MS.[11] One compound, *cis*-C1, demonstrated a high degree of binding specificity based on our ESI-MS screening results, and the bis-intercalative binding mode was confirmed by subsequent NMR studies.[59] We examined the $KMnO_4$ reactivity of DNA duplex complexes containing *cis*-C1 to determine how the degree of oxidation induced by the new bis-intercalator compares to that of echinomycin. An abundant complex with 1:1 binding stoichiometry in the 6- charge state is present in the ESI mass spectrum of a solution containing *cis*-C1 with ds5 prior to the $KMnO_4$ reaction (spectra not shown). After the permanganate reaction, one oxidation adduct was formed for the *cis*-C1/ds5 complex, and the percent oxidation was calculated to be 40% (Table 7.1). Compared to the percent oxidation of the duplex in the absence of the bis-intercalator ligand, 18%, a modest increase in the degree of oxidation is observed upon *cis*-C1 binding. The duplex distortion induced by *cis*-C1 is considerably greater than that caused by actinomycin with the same duplex (only 19%) which is

expected since actinomycin is a monointercalator whereas *cis*-C1 is a bis-intercalator. However, the percent oxidation suggests that the distortion of the duplex by *cis*-C1 is not as great as that caused by echinomycin, as evidenced by the moderately greater percent oxidation value for 1:1 complexes containing echinomycin with ds5 (51%).

Hoechst 33342 (structure shown in Figure 7.1) is a well-studied minor-groove binder known to target A/T rich regions of a duplex. Compounds that bind to the minor groove of DNA typically do not distort the duplex as significantly as do intercalators[60, 61] and as such, do not induce $KMnO_4$ reactivity. The ESI mass spectrum of a solution containing Hoechst 33342 (abbreviated as "H") and d(GCGGATATATGGCG/CGCCATATATCCGC) (ds1) prior to the permanganate reaction is shown in Figure 7.5a. An abundant 1:1 complex is present in the 5- charge state at *m/z* 1795 and a much less abundant 2:1 complex is present at *m/z* 1885. The corresponding mass spectrum of the solution after the $KMnO_4$ reaction is shown in Figure 7.5b. The relative abundance of the Hoechst 33342/duplex complex at *m/z* 1795 has decreased, and a new peak is present at *m/z* 1791 that is consistent with a dehydrated Hoescht 33342/duplex species, not a typical oxidation adduct. To identify the new species at m/z 1791, CAD experiments were undertaken on the ion (spectra not shown). The ion dissociated to form an ion at m/z 1705, which is consistent with the m/z of the original duplex. Upon a second stage of collisional activation, the resulting ion of m/z 1705 dissociated in a manner consistent with the fragmentation of native duplex ions in the 5- charge state (spectra not shown) These results suggest an unusual reaction between the bound Hoechst 33342 and $KMnO_4$ but not the oxidation of the duplex The Hoechst ligand apparently reacts with $KMnO_4$, dehydrates, and yet remains bound to the duplex to form the observed product of m/z 1791. This reaction also caused the decrease in the relative abundance of the original 1:1 Hoechst 33342/duplex complex.

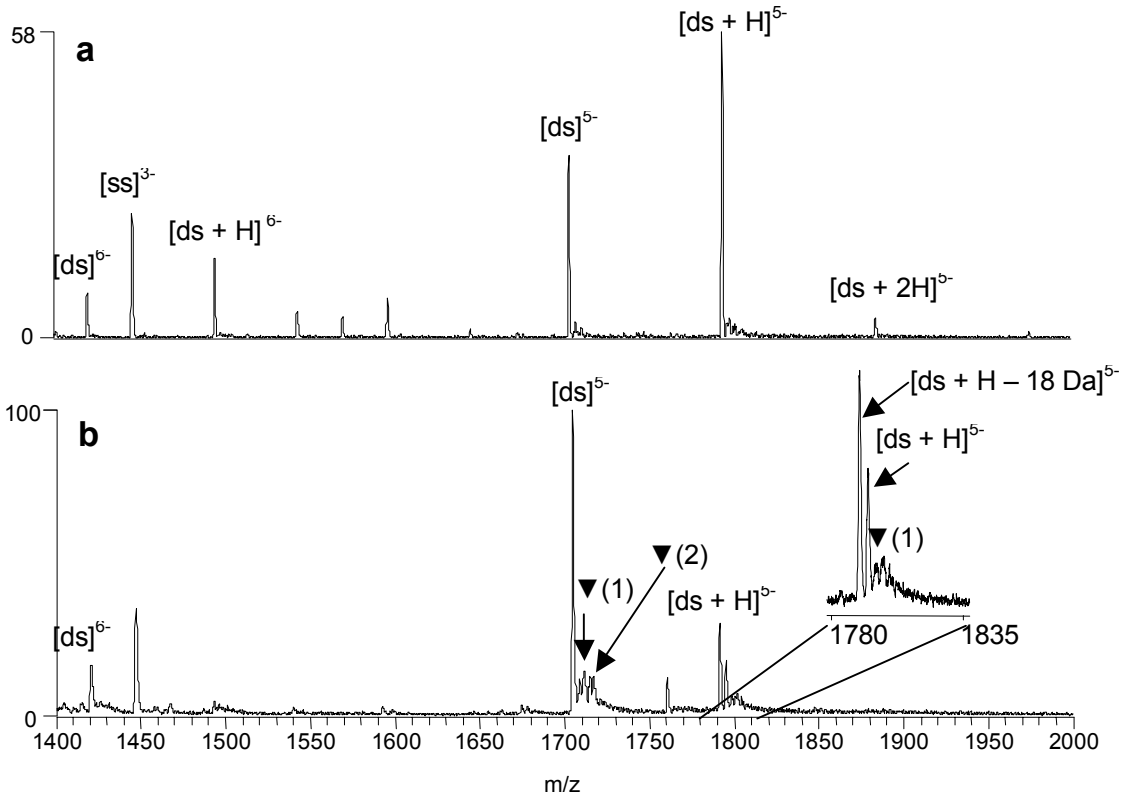


Figure 7.5: ESI-mass spectra showing solutions containing the duplex dGCGGATATATGGCG/CGCCATATATCCGC) and Hoechst 33342 (H) (a) before reaction with KMnO4 and (b) after 20 min. reaction with KMnO4. Ions containing oxidized thymines are labeled with a black triangle, with the number in parenthesis indicating the number of adducts each peak contains.

Low abundance oxidation adducts associated with the complex at m/z 1791 and 1795 are present as shown in the spectral enlargement in the inset of Figure 7.5b; however, the percent oxidation for these complexes does not exceed 20% and in fact is no greater than the oxidation value of the drug-free duplex The significantly lower percent

oxidation value of the Hoechst 33342/duplex complexes compared to the echinomycin/duplex complexes is consistent with the low degree of distortion of the duplex by minor groove binder ligands.

7.3.5 CAD of Oxidized DNA

In addition to observing oxidation adducts induced by ligand binding to duplex DNA, another important aspect of the $KMnO_4$ experiment is determining the site of oxidation. In traditional gel-based experiments, oxidation sites are mapped via a multi-step procedure in which the oxidized DNA is cleaved adjacent to the modification sites by a piperidine heat treatment, followed by DNA precipitation, and then subsequent gel-analysis of the products.[30, 31] Rather than employing the piperidine treatment, our strategy entails direct analysis of the oxidized DNA adducts based on CAD to determine the site of oxidation, thus simplifying the analytical scheme.

The CAD spectra of intercalator/DNA complexes have been previously evaluated.[8, 62, 63] These complexes typically dissociate via the ejection of the ligand, leaving the drug-free duplex ion, or by separation of the single strands of the duplex with the ligand remaining bound to one of the strands. The MS/MS dissociation pathways of oligonucleotides have also been extensively examined,[44, 64] and systematic interpretation of the fragmentation patterns has been used to determine the site of covalent modifications of nucleic acids based on diagnostic *a –B* and *w* sequence ions that result from the backbone fragmentation of single strand oligonucleotides.[24-27, 65]

In the present study, the most abundant oxidation adducts were isolated in the trap and subjected to CAD. For example, Figure 7.6a shows the CAD spectrum of the $[ds + 2 \times E + O]^{6-}$ ion containing two molecules of echinomycin and duplex ds1 that was originally observed in the ESI mass spectrum shown in Figure 7.4c at m/z 1789. After the first stage of CAD, the complex dissociates by strand separation with one echinomycin

molecule remaining bound to each single strand. The mass of the ions containing the d(CGCCATATATCCGC) single strand (labeled ss2 in Figure 7.6a) is shifted by 34 Da, indicating that the oxidation occurred exclusively on ss2 rather than ss1. The $[ss2 + O]^{3-}$ product ion was subjected to a second stage of collisional activation (MS^3) to gain information about the specific position of the oxidized base. As shown in Figure 7.6b, the MS^3 spectrum contains many informative *a-B* and *w* ions, some of which contain the mass shift characteristic of the oxidized base. The sequence overlaid on Figure 7.6b summarizes the diagnostic cleavages (labeled with slashmarks) in the spectrum, with those containing the mass shift labeled with a "black triangle".

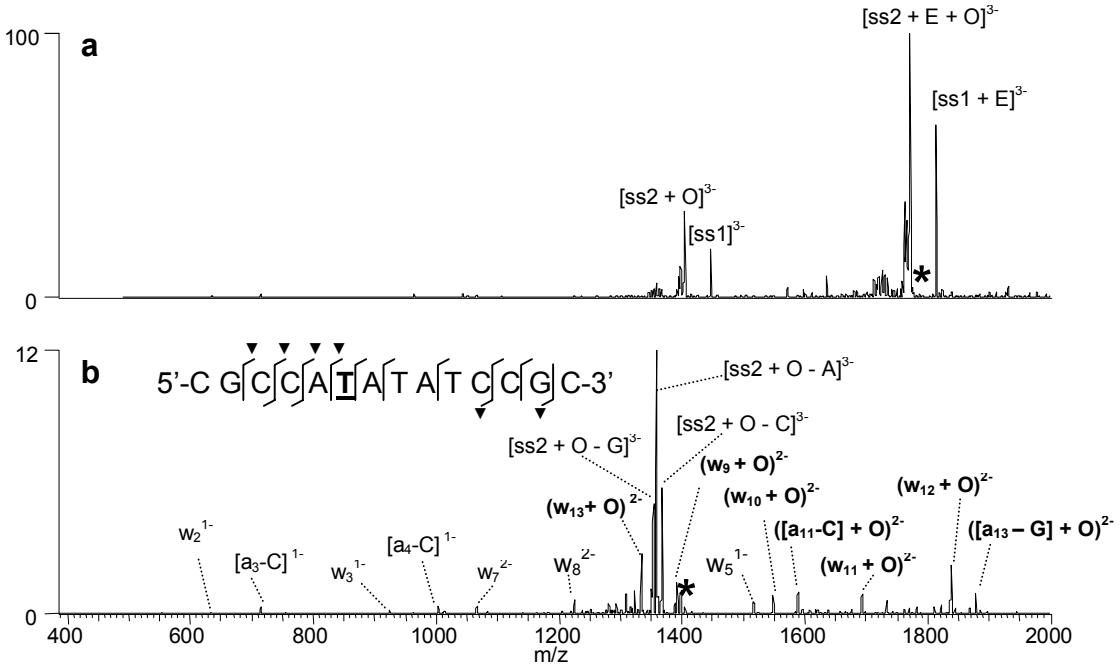


Figure 7.6: ESI-MS^3 experiments for $[ds + 2E + O]^{6-}$ containing d(GCGGATATATGGCG/CGCCATATATCCGC) and echinomycin (E): (a) CAD spectrum of the initial $[ds + 2E + O]^{6-}$ complex and (b) MS^3 spectrum of the $[ss2 + O]^{3-}$ product ion. "O" is indicative of an oxidation adduct. The sequence structure in Figure 7.6b summarizes the sequence coverage. The fragments containing an oxidation adduct are labeled with a black triangle. The thymine that was determined to be oxidized is underlined in the sequence shown in Figure 7.6b.

Based on the cleavage pattern, T6 can be identified as the sole oxidized residue. The key fragment ions for making this assignment are the w_8 ion that does not contain the mass shift and the (w_9 + O) ion that does contain the oxidation mass shift. Because the echinomycin molecule binds to the G/C rich regions of the duplex, it follows that the thymines closest to this region, like T6, will be most easily oxidized. The clarity of the fragmentation pattern is remarkable, as well as the fact that the oxidation occurs exclusively at one specific thymine site. The abundant oxidation of T6 over T10, the other thymine in the sequence that is close to the G/C rich region of the sequence may occur because this thymine is flanked by two adenines and previous studies have found thymines adjacent to short A-tracts are more susceptible to the permanganate reaction.[66] The preference for thymine reaction on the C-rich single strand is not full understood at this time and merits further study.

The sites of thymine oxidation determined by CAD experiments for the other drug/DNA complexes examined in this study are summarized in Table 7.2. In most cases, the CAD results confirm a single thymine oxidation, not a distribution or ensemble of oxidation sites. Thus, the oxidation process displays striking selectivity that makes mass spectrometric analysis a natural fit for pinpointing the oxidation sites. The thymines identified in Table 7.2 are all with in 4-5 bases of the proposed 5'-CG-3' intercalator binding sites or adjacent to G/C rich regions of the sequence that are likely intercalator binding sites. Future studies will be aimed at examining trends in the sites of oxidation. One result of note involves the complexes containing echinomycin and d(GTAGAGTCGACCTG/CAGGTCGACTCTAC) (ds2). As discussed above, the oxidation of complexes containing this duplex were examined to maintain consistency with a previous gel electrophoresis study of the binding of echinomycin to duplex DNA

assessed by chemical probes, including $KMnO_4$.[31] The high affinity binding site of echinomycin on the duplex was determined to be the central 5'-TCGA-3' site.[31] Based on the strand separation products present in the CAD spectrum of the $[ds + E + O]^{6-}$ complex in our present study, the site of oxidation was determined to be located on the first single strand, d(GTAGAGTCGACCTG) (spectra not shown). However, the MS^3 spectrum of the $[ss + O]^{3-}$ ion resulting from the first stage of CAD was inconclusive since there were product ions suggesting that the oxidized thymine could be the T2 or T13. It is likely that the $[ss + O]^{3-}$ species consisted of two population of ions in which the thymine at position 2 was oxidized in one population, while the thymine at position 13 was oxidized in the other, suggesting that the duplex can unravel at both ends. Unlike the other DNA sequences used in this study, d(GTAGAGTCGACCTG) contains two thymines (T2 and T13) that are located at equivalent positions from the ends of the duplex, and thus may have roughly equal susceptibility to thymine oxidation upon echinomycin binding.

Table 7.2: Site of thymine oxidation for DNA/ligand complexes.

Ligand	Duplex	Sequence and Thymine Oxidation Site(s)[a]
echinomycin	ds1	d(GCGGATATATGGCG/CGCCA**T**ATATCCGC)
echinomycin	ds2	d(G**T**AGAGTCGACC**T**G/CAGGTCGACTC**T**AC)
echinomycin	ds5	d(GCGGGGATGGGGCG/CGCCCCA**T**CCCCGC)
echinomycin	ds3	d(GCAGTGA/**T**CACTGC)
echinomycin	ds4	d(GGACAGTGAGGGCAGTGAGGG/CCCTCAC**T**GCCCTCACTGTCC)
ethidium bromide	ds1	d(GCGGATATATGGCG/CGCCA**T**ATATCCGC)
cis-C1	ds5	d(GCGGGGATGGGGCG/CGCCCCA**T**CCCCGC)

[a]The site(s) of thymine oxidation are shown in bold.

To further ascertain the regions of DNA unwinding for this duplex, CAD studies were performed on the $[ds2 + E + 2 \times O]^{6-}$ complex, the complex containing two sites of oxidation. Upon the first stage of CAD, the resulting strand separation products revealed that the second site of thymine oxidation was located on the second strand, d(CAGGTCGACTCTAC). Subsequent MS^3 experiments indicated that T12 is the position of the second oxidized thymine (spectra not shown). This is the same thymine that was found to be oxidized in the gel-based study by Bailly et al. using a longer duplex.[31] When the structure of the duplex is considered, the thymine at the twelth position on the second single strand d(CAGGTCGACTCTAC) is adjacent to the second thymine on the complementary first strand, d(GTAGAGTCGACCTG), which confirms that this region of the DNA is especially distorted upon binding echinomycin. The sites of thymine oxidation correlate with the results of the past gel experiments[31] and support that echinomycin is bound to the 5'-TCGA-3' region, thus promoting oxidation around this site.

7.3.6 IRMPD of Oxidized DNA

IRMPD has also proven useful for gaining sequence information of DNA molecules and sites of modification.[45, 46] During IRMPD, ion activation is independent of the RF voltage applied to the trap which eliminates the low mass cut-off problem characteristic of CAD experiments in a quadrupole ion trap.[45] Ion activation by IR photoabsorption is a non-resonant process, resulting in increased secondary fragmentation of DNA and subsequently more *a* - *B* and *w* ions without the need for sequential stages of ion isolation, activation and dissociation that are common for CAD strategies.[45] To explore the potential merits of IRMPD for the characterization of DNA oxidation adducts and determination of oxidation sites, several pilot IRMPD experiments

were undertaken in comparison to the CAD results described above. Figure 7.7a shows the IRMPD spectrum of the $[ds + 2 \times E + O]^{6-}$ complex containing echinomycin and duplex ds5 using 50 W laser power and an activation time of only 1.7 ms. The types and relative abundances of the fragment ions present in the IRMPD mass spectrum are quite different from those in the CAD spectrum of the same complex shown in Figure 7.6a. While strand separation ions dominated the CAD spectrum, they are essentially absent in the IRMPD spectrum because these ions have undergone efficient IR absorption and been converted into informative sequence *a* - *B* and *w* ions. The site of the oxidation can be determined readily from these sequence ions, as shown by the cleavages marked on the duplex sequence structure overlaid in Figure 7.6a, thus eliminating the need for a second stage of activation which is typically necessary for more complete sequence coverage in the analogous CAD experiments.

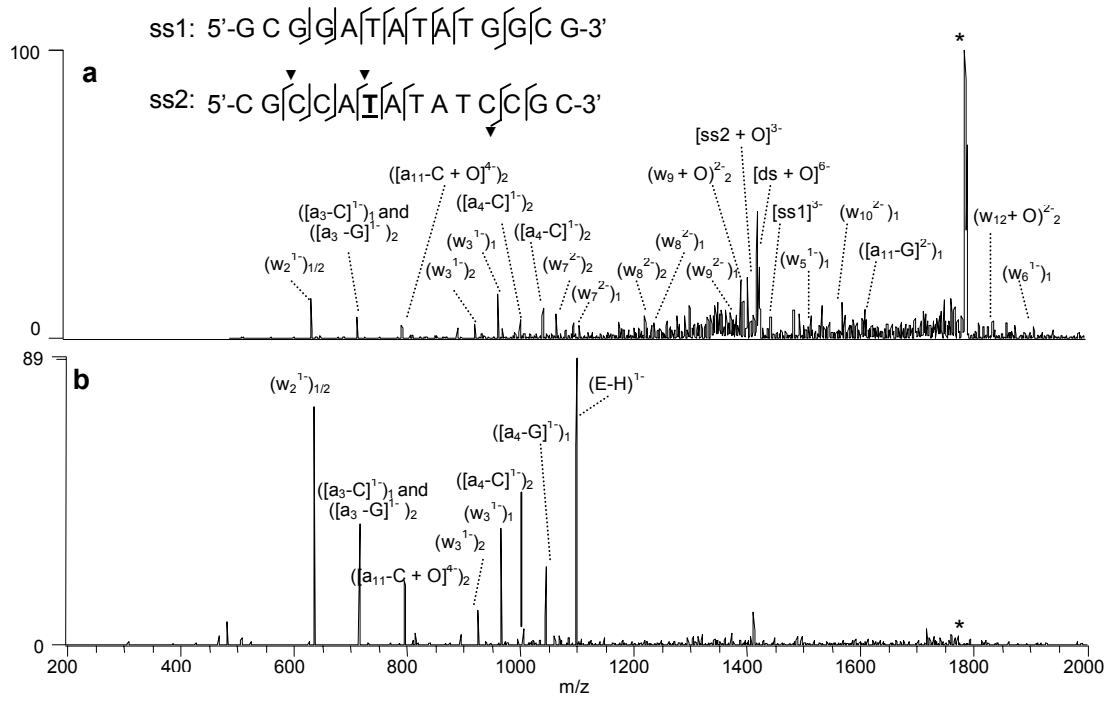


Figure 7.7: IRMPD spectra of $[ds + 2 \times E + O]^{6-}$ using (a) 99% laser power and 1.7 ms irradiation time and (b) 99% laser power and 2.5 ms irradiation time.

The relative abundance of the sequence ions produced upon IRMPD can be controlled by varying the irradiation time. This is demonstrated in Figure 7.7b which shows the IRMPD spectrum of the same complex that was dissociated in Figure 7.7a, $[ds + 2 \times E + O]^{6-}$, with an irraidation time of 2.5 ms. The precursor ion has been totally converted to fragment ions, and the relative abundances of the lower mass *a* - *B* and *w* ions are significantly increased compared to those in Figure 7.7a. The ability to effectively "tune" the relative abundances of the sequence ions using different irradiation times during an IRMPD experiment is particularly beneficial for identifying the site of thymine oxidation without requiring more elaborate MS^3 experiments.

7.4 CONCLUSIONS

An ESI-MS technique employing $KMnO_4$ has been developed to assess the distortion of DNA duplexes upon ligand binding. Adducts consistent with the oxidation of thymine nucleobases are detected in the ESI mass spectrum, while CAD and IRMPD can be used to identify the specific sites of oxidation. Ligand binding sites can be determined based on the thymine oxidation pattern. The technique presented here represents an alternative to traditional gel-based $KMnO_4$ experiments and is attractive due to the elimination of the need to use both radiolabeled DNA and the piperidine heat treatment required to identify the site of oxidation, while also offering excellent sensitivity and facile adaptation to high throughput screening applications.

Significant differences in the extent of oxidation of thymines in single strand DNA compared to duplex DNA were observed, and the susceptibility of thymines in drug/duplex complexes were significantly enhanced, presumably due to conformational changes of the duplex upon drug binding. The bis-intercalator echinomycin caused the most extensive thymine oxidation, followed by the threading bis-intercalator *cis*-C1, and the monointercalator ethidium bromide. Actinomycin-D, another monointercalator, and the minor groove binder Hoechst 33342 did not substantially increase the reactivity of thymines with $KMnO_4$. In addition to the type of drug bound to the duplex, other factors that influenced the extent of thymine oxidation were the size of the duplex, the number of bound ligands, and the proximity of the thymines to the termini of the duplex. The ability to assess the $KMnO_4$ reactivity induced by ligands that become unbound from the duplex, either by reacting with the permanganate ion or as a result of the post-oxidation C_{18} clean-up procedure, is also demonstrated. Multi-stage CAD experiments allowed confident determination of the positions of the oxidized thymines. As an alternative to CAD and

MS^3 approaches, IRMPD offers a promising option that results in more abundant sequence ions due to enhanced secondary fragmentation.

7.5 REFERENCES

(1) Hofstadler, S. A.; Sannes-Lowery, K. A. *Nat. Rev. Drug. Discov.* **2006**, *5*, 585-595.

(2) Hofstadler, S. A.; Griffey, R. H. *Chem. Rev.* **2001**, *101*, 377-390.

(3) Beck, J. L.; Colgrave, M. L.; Ralph, S. F.; Sheil, M. M. *Mass. Spectrom. Rev.* **2001**, *20*, 61-87.

(4) Gale, D. C.; Smith, R. D. *J. Am. Soc. Mass. Spectrom.* **1995**, *6*, 1154-1164.

(5) Gabelica, V.; De Pauw, E.; Rosu, F. *J. Mass. Spectrom.* **1999**, *34*, 1328-1337.

(6) Kapur, A.; Beck, J. L.; Sheil, M. M. *Rapid Commun. Mass Spectrom.* **1999**, *13*, 2489-2497.

(7) Gabelica, V.; Rosu, F.; Houssier, C.; De Pauw, E. *Rapid Commun. Mass Spectrom.* **2000**, *14*, 464-467.

(8) Wan, K. X.; Gross, M. L.; Shibue, T. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 450-457.

(9) Wan, K. X.; Shibue, T.; Gross, M. L. *J. Am. Chem. Soc.* **2000**, *122*, 300-307.

(10) Rosu, F.; Gabelica, V.; Houssier, C.; De Pauw, E. *Nucleic Acids Res.* **2002**, *30*, -.

(11) Mazzitelli, C. L.; Chu, Y.; Reczek, J. J.; Iverson, B. L.; Brodbelt, J. S. *J. Am. Soc. Mass. Spectrom.* **2006**, *in press*.

(12) Reyzer, M. L.; Brodbelt, J. S.; Kerwin, S. M.; Kumar, D. *Nucleic Acids Res.* **2001**, *29*, art. no.-e103.

(13) Oehlers, L.; Mazzitelli, C. L.; Brodbelt, J. S.; Rodriguez, M.; Kerwin, S. *J. Am. Soc. Mass Spectrom.* **2004**, *15*, 1593-1603.

(14) Rosu, F.; Gabelica, V.; Houssier, C.; Colson, P.; De Pauw, E. *Rapid Commun. Mass Spectrom.* **2002**, *16*, 1729-1736.

(15) David, W. M.; Brodbelt, J.; Kerwin, S. M.; Thomas, P. W. *Anal. Chem.* **2002**, *74*, 2029-2033.

(16) Guittat, L.; Alberti, P.; Rosu, F.; Van Miert, S.; Thetiot, E.; Pieters, L.; Gabelica, V.; De Pauw, E.; Ottaviani, A.; Riou, J. F.; Mergny, J. L. *Biochimie* **2003**, *85*, 535-547.

(17) Rosu, F.; De Pauw, E.; Guittat, L.; Alberti, P.; Lacroix, L.; Mailliet, P.; Riou, J. F.; Mergny, J. L. *Biochemistry* **2003**, *42*, 10361-10371.

(18) Rosu, F.; Gabelica, V.; Shin-ya, K.; De Pauw, E. *Chem. Commun.* **2003**, 2702-2703.

(19) Guittat, L.; De Cian, A.; Rosu, F.; Gabelica, V.; De Pauw, E.; Delfourne, E.; Mergny, J. L. *Biochim. Biophys. Acta* **2005**, *1724*, 375-384.

(20) Baker, E. S.; Lee, J. T.; Sessler, J. L.; Bowers, M. T. *J. Am. Chem. Soc.* **2006**, *128*, 2641-2648.

(21) Mazzitelli, C. L.; Kern, J. T.; Rodriguez, M.; Brodbelt, J. S.; Kerwin, S. M. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 593-604.

(22) Nielsen, P. E. *Journal of Molec. Recognit.* **1990**, *3*, 1-25.

(23) Bui, C. T.; Rees, K.; Cotton, R. G. H. *Curr. Pharmacogenomics* **2004**, *2*, 325-332.

(24) Kellersberger, K. A.; Yu, E.; Kruppa, G. H.; Young, M. M.; Fabris, D. *Anal. Chem.* **2004**, *76*, 2438-2445.

(25) Yu, E. T.; Zhang, Q. G.; Fabris, D. *J. Mol. Biol.* **2005**, *345*, 69-80.

(26) Yu, E.; Fabris, D. *Anal. Biochem.* **2004**, *334*, 356-366.

(27) Yu, E.; Fabris, D. *J. Mol. Biol.* **2003**, *330*, 211-223.

(28) Bui, C. T.; Rees, K.; Cotton, R. G. H. *Nucleos. Nucleot. Nucl.* **2003**, *22*, 1835-1855.

(29) Sasse-Dwight, S.; Gralla, J. D. *J. Biol. Chem.* **1989**, *264*, 8074-8081.

(30) Jeppesen, C.; Nielsen, P. E. *Febs Lett.* **1988**, *231*, 172-176.

(31) Bailly, C.; Gentle, D.; Hamy, F.; Purcell, M.; Waring, M. J. *Biochem. J.* **1994**, *300*, 165-173.

(32) Gille, H.; Messer, W. *Embo J.* **1991**, *10*, 1579-1584.

(33) Hsieh, D. J.; Camiolo, S. M.; Yates, J. L. *Embo J.* **1993**, *12*, 4933-4944.

(34) Farah, J. A.; Smith, G. R. *J. Mol. Biol.* **1997**, *272*, 699-715.

(35) Schaubach, O. L.; Dombroski, A. J. *J. Biol. Chem.* **1999**, *274*, 8757-8763.

(36) Rubin, C. M.; Schmid, C. W. *Nucleic Acids Res.* **1980**, *8*, 4613-4619.

(37) Jones, A. S.; Walker, R. T. *Nature* **1964**, *202*.

(38) Jiang, H.; Zacharias, W.; Amirhaeri, S. *Nucleic Acids Res.* **1991**, *19*, 6943-6948.

(39) Desantis, P.; Palleschi, A.; Savino, M.; Scipioni, A. *Biochemistry* **1990**, *29*, 9269-9273.

(40) Mitas, M.; Yu, A.; Dill, J.; Kamp, T. J.; Chambers, E. J.; Haworth, I. S. *Nucleic Acids Res.* **1995**, *23*, 1050-1059.

(41) Klysik, J.; Rippe, K.; Jovin, T. M. *Biochemistry* **1990**, *29*, 9831-9839.

(42) Bui, C. T.; Lambrinakos, A.; Cotton, R. G. H. *Biopolymers* **2003**, *70*, 628-636.

(43) Cotton, R. G. H. *Biochem. J.* **1989**, *263*, 1-10.

(44) Mcluckey, S. A.; Vanberkel, G. J.; Glish, G. L. *J. Am. Soc. Mass Spectrom.* **1992**, *3*, 60-70.

(45) Keller, K. M.; Brodbelt, J. S. *Anal. Biochem.* **2004**, *326*, 200-210.

(46) Wilson, J. J.; Brodbelt, J. S. *Anal. Chem.* **2006**, *78*, 6855-6862.

(47) Chu, Y. J.; Lynch, V.; Iverson, B. L. *Tetrahedron* **2006**, *62*, 5536-5548.

(48) Vandyke, M. M.; Dervan, P. B. *Science* **1984**, *225*, 1122-1127.

(49) Low, C. M. L.; Drew, H. R.; Waring, M. J. *Nucleic Acids Res.* **1984**, *12*, 4865-4879.

(50) Gao, X.; Patel, D. J. *Q. Rev. Biophysics* **1989**, *22*, 93-138.

(51) Waring, M. J.; Wakelin, L. P. G. *Nature* **1974**, *252*, 653-657.

(52) Rueda, M.; Luque, F. J.; Orozco, M. *J. Am. Chem. Soc.* **2005**, *127*, 11690-11698.

(53) Gidden, J.; Ferzoco, A.; Baker, E. S.; Bowers, M. T. *J. Am. Chem. Soc.* **2004**, *126*, 15132-15140.

(54) Sinden, R. R. *DNA Structure and Function*; Academic Press: San Diego, 1994.

(55) Sobell, H. M. *Proc. Natl. Acad. Sci. USA* **1985**, *82*, 5328-5331.

(56) Sobell, H. M.; Jain, S. C.; Sakore, T. D.; Nordman, C. E. *Nature new Biol.* **1972**, *231*, 200-205.

(57) Wakelin, L. P. G.; Waring, M. J. *Biochem. J.* **1976**, *157*, 721-740.

(58) Wang, J. C. *Biochim. Biophys. Acta* **1971**, *232*, 246-251.

(59) Chu, Y.; Sorrey, S.; Hoffman, D. W.; Iverson, B. L. *J. Am. Chem. Soc.* **2006**, *in press*.

(60) Neidle, S. *Biopolymers* **1997**, *44*, 105-121.

(61) Neidle, S.; Pearl, L. H.; Skelly, J. V. *Biochem J.* **1987**, *243*, 1-&.

(62) Keller, K. M.; Zhang, J. M.; Oehlers, L.; Brodbelt, J. S. *J. Mass Spectrom.* **2005**, *40*, 1362-1371.

(63) Gabelica, V.; De Pauw, E. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 91-98.

(64) Mcluckey, S. A.; Habibigoudarzi, S. *J. Am. Chem. Soc.* **1993**, *115*, 12085-12095.

(65) Wang, Y. S.; Zhang, Q. C.; Wang, Y. S. *J. Am. Soc. Mass Spectrom.* **2004**, *15*, 1565-1571.

(66) Mccarthy, J. G.; Williams, L. D.; Rich, A. *Biochemistry* **1990**, *29*, 6071-6081

Chapter 8: Conclusions

The results of this dissertation have contributed to the field of bioanalytical mass spectrometry by extending the use of ESI-MS to evaluate novel types of drug/DNA complexes, and by demonstrating how ESI-MS can be used to gain structural information about DNA and drug/DNA complexes. The clinical importance and structural diversity of DNA-interactive compounds is presented in chapter 1. In recent years ESI-MS has been developed as a tool for the analysis of drug/DNA complexes because it offers the advantages of improved sensitivity and faster analysis times over traditional, solution-based techniques. As reviewed in chapter 1, most of the work done in this area has focused on examining DNA complexes containing well-studied, commercially available duplex-interactive drugs. These past studies were important because they demonstrated that not only are non-covalent drug/DNA complexes retained in the gas-phase, but the sequence selectivities, binding affinities and complex stoichiometries observed by ESI-MS mirror known solution behavior. These reports laid the groundwork for the studies presented in this dissertation.

In chapter 3, the binding of threading bis-intercalators with *cis*- and *trans*-oriented structures to DNA duplexes with different sequences is presented. The compounds used in this study were designed to undergo more specific binding interactions with DNA duplexes. ESI-MS was used to quickly screen the binding of the compounds to a series of duplexes containing different potential high affinity binding sites identified in DNase I footprinting experiments. The binding affinities of the compounds for different sequences as determined by ESI-MS correlated with the results of the footprinting experiments, further establishing ESI-MS as a screening tool. Specific binding was assessed in ESI-MS titration binding experiments that determined the *cis*-oriented ligand exhibited more

promising specific binding behavior by forming complexes with well-defined stoichiometries, even at high ligand/DNA molar ratios. The *trans*-oriented compound formed ligand/DNA complexes with 2:1 and 3:1 binding stoichiometries even when each duplex contained only one high affinity binding site, which is suggestive of non-specific binding.

The evaluation of metal-mediated DNA binding by ESI-MS is demonstrated in chapter 4. The binding of a series of benzoxazole analogs with different amide- and ester-linking side-chains with duplex DNA is examined. These compounds are unique in that some of them, notably the ligands with shorter side chains, will only complexes with duplex DNA in the presence of divalent metal cations. The ligands form complexes with the metals and then bind to duplex DNA forming ligand/metal/DNA complexes with 1:1:1 and 2:2:1 binding stoichiometries. Cu^{2+} was the most promising for promoting DNA binding, following by Ni^{2+}. Another interesting result of the study was that the analogs with the most dramatic metal-mediated or metal-enhanced binding behavior were also found to be the most cytotoxic against lung and breast cancer cell lines.

ESI-MS analysis of quadruplex DNA was explored in chapters 5 and 6. The quadruplex-selectivity of perylene diimides with different side chains was determined by ESI-MS in chapter 5 by evaluating the binding of the ligands to quadruplex, duplex and single strand DNA. Three ligands, one containing basic side chains, one containing anionic side-chains, and one benzannulated compound were determined to be the most-quadruplex selective. The ESI-MS results correlated well with spectroscopic experiments, demonstrating that ESI-MS can be used to investigate the selectivity of ligands for different DNA structures which is increasingly important for developing compounds reducing the cytotoxic side-effects.

The success of ESI-MS for examining ligand binding to quadruplex DNA spurred interest in further examining the gas-phase stabilities of different quadruplex structures as detailed in chapter 6. Energy-variable CAD experiments and molecular dynamics simulations were performed on series of four-, two- and single stranded-quadruplexes with different sequences and numbers of G-tetrads. The molecular dynamics simulations confirmed that the quadruplex structure is maintained with some distortion in the gas-phase. Free energy analyses were also performed on the different structures and the stability ranking of the quadrupelxes correlated well with the relative stabilities determined by the energy-resolved CAD experiments. The results of this study suggest that energy-resolved CAD curves could be used to rapidly screen the gas-phase stabilities of different quadruplex structures.

The seventh chapter of this dissertation describes the use of ESI-MS to gain information about conformational changes to DNA upon ligand binding, as well as determine ligand binding sites by utilizing the $KMnO_4$ reaction. The ESI-MS technique was developed by examining the oxidation of thymines on single strand and duplex DNA, with the duplexes exhibiting a significant resistance to the reaction. DNA complexes containing different drugs were also subjected to the oxidation reaction. Bis-intercalators were found to induce the most $KMnO_4$ reactivity, while minor groove binders cause little to no increase in the oxidation. CAD and IRMPD experiments were also used to determine the site of oxidation based on oligonucleotide fragmentation patterns.

The work presented in this dissertation demonstrates that ESI-MS is a useful tool for examining drug/DNA complexes. My work has shown that tandem mass spectrometry can be used to map the sequence selectivities and relative binding affinities of new DNA interactive ligands in a way that consumes very little sample and is easily

adaptable to high throughput screening applications. Ligands that are selective for quadruplex, duplex, and single strand structures can be pinpointed, binding stoichiometries can be determined, and non-specific interactions can be differentiated from specific binding modes. Moreover, ESI-MS/MS can be combined with chemical probe methods to yield a complementary method for determining ligand binding sites and identifying conformational changes of DNA upon ligand interaction.

Despite these inroads, there is considerable room for future work in this field. One major challenge that merits future work is developing ESI-MS techniques that are capable of analyzing longer DNA strands. The present method has proven to be successful for examining oligonucleotides containing 50 bases or less, however the ability to analyze larger DNA structures will make the technique more biologically relevant. The use of longer DNA strands will certainly necessitate more high performance mass analyzers with large mass ranges, or the use of chemical or enzymatic cleaving agents to cleave the DNA into smaller pieces more suitable for ESI-MS analysis.

The use of chemical and enzymatic probes to gain structural information about ligand binding sites and binding modes is another important area of future study. In addition to the $KMnO_4$ reaction present in chapter 7, there are numerous other footprinting and chemical probe techniques that could be used in conjunction with ESI-MS including DNase I, the hydroxyl radical, and MPE-Fe^{II}. The use of ESI-MS to examine drug/DNA complexes is growing field, the capabilities of which are far from being fully developed.

References

(1) Goodman, L. S.; Hardman, J. G.; Limbird, L. E.; Gilman, A. G. *Goodman & Gilman's the pharmacological basis of therapeutics*, 10th ed.; McGraw-Hill: New York, 2001.

(2) Propst, C. L.; Perun, T. J. *Nucleic acid targeted drug design*; M. Dekker: New York, 1992.

(3) Ginsburg, H.; Nissani, E.; Krugliak, M.; Williamson, D. H. *Mol. Biochem. Parasit.* **1993**, *58*, 7-15.

(4) Civitico, G.; Wang, Y. Y.; Luscombe, C.; Bishop, N.; Tachedjian, G.; Gust, I.; Locarnini, S. *J. Med. Virol.* **1990**, *31*, 90-97.

(5) Asante-Appiah, E.; Skalka, A. M. *Antivir. Res.* **1997**, *36*, 139-156.

(6) Cremieux, A.; Chevalier, J.; Sharples, D.; Berny, H.; Galy, A. M.; Brouant, P.; Galy, J. P.; Barbe, J. *Res. Microbiol.* **1995**, *146*, 73-83.

(7) Williams, R. A. D.; Lambert, P. A.; Singleton, P. *Antimicrobial Drug Action*; BIOS Scientific: Oxford, 1996.

(8) Galpin, A. J.; Evans, W. E. *Clin. Chem.* **1993**, *39*, 2419-2430.

(9) Allwood, M. *The Cytotoxics Handbook*; Radcliffe Medical Press: Oxford, 1993.

(10) Witt, K. L.; Bishop, J. B. *Mut. Res.-Fund. Mol. M.* **1996**, *355*, 209-234.

(11) Haq, I.; Ladbury, J. *J. Mol. Recognit.* **2000**, *13*, 188-197.

(12) Brana, M. F.; Cacho, M.; Gradillas, A.; de Pascual-Teresa, B.; Ramos, A. *Curr.Pharm. Des.* **2001**, *7*, 1745-1780.

(13) Bischoff, G.; Hoffmann, S. *Curr. Med. Chem.* **2002**, *9*, 321-348.

(14) Sobell, H. M. *Proc. Natl. Acad. Sci. U. S. A.* **1985**, *82*, 5328-5331.

(15) Moore, M. H.; Hunter, W. N.; Destaintot, B. L.; Kennard, O. *J. Mol. Biol.* **1989**, *206*, 693-705.

(16) Wakelin, L. P. G.; Waring, M. J. *Biochem. J.* **1976**, *157*, 721-740.

(17) Waring, M. J.; Wakelin, L. P. G. *Nature* **1974**, *252*, 653-657.

(18) Lee, J.; Guelev, V.; Sorey, S.; Hoffman, D. W.; Iverson, B. L. *J. Am. Chem. Soc.* **2004**, *126*, 14036-14042.

(19) Rao, K. E.; Dasgupta, D.; Sasisekharan, V. *Biochemistry* **1988**, *27*, 3018-3024.

(20) Timsit, Y.; Moras, D. *Embo J.* **1994**, *13*, 2737-2746.

(21) Geierstanger, B. H.; Wemmer, D. E. *Annu. Rev. Biophys. Biomol. Struct.* **1995**, *24*, 463-493.

(22) Kerwin, S. M. *Curr. Pharm. Des.* **2000**, *6*, 441-471.

(23) Han, H. Y.; Hurley, L. H. *Trends Pharmacol.l Sci.* **2000**, *21*, 136-142.

(24) Hurley, L. H.; Wheelhouse, R. T.; Sun, D.; Kerwin, S. M.; Salazar, M.; Fedoroff, O. Y.; Han, F. X.; Han, H. Y.; Izbicka, E.; Von Hoff, D. D. *Pharmacol. Ther.* **2000**, *85*, 141-158.

(25) Moyzis, R. K.; Buckingham, J. M.; Cram, L. S.; Dani, M.; Deaven, L. L.; Jones, M. D.; Meyne, J.; Ratliff, R. L.; Wu, J. R. *Proc. Natl. Acad. Sci. U. S. A.* **1988**, *85*, 6622-6626.

(26) Wellinger, R. J.; Sen, D. *Eur. J. Cancer* **1997**, *33*, 735-749.

(27) Harley, C. B.; Futcher, A. B.; Greider, C. W. *Nature* **1990**, *345*, 458-460.

(28) Nakamura, T. M.; Morin, G. B.; Chapman, K. B.; Weinrich, S. L.; Andrews, W. H.; Lingner, J.; Harley, C. B.; Cech, T. R. *Science* **1997**, *277*, 955-959.

(29) Counter, C. M.; Hahn, W. C.; Wei, W. Y.; Caddle, S. D.; Beijersbergen, R. L.; Lansdorp, P. M.; Sedivy, J. M.; Weinberg, R. A. *Proc. Natl. Acad. Sci. U. S. A.***1998**, *95*, 14723-14728.

(30) Rezler, E. M.; Bearss, D. J.; Hurley, L. H. *Curr. Opin. Pharmacol.* **2002**, *2*, 415-423.

(31) Rezler, E. M.; Bearss, D. J.; Hurley, L. H. *Annu. Rev. Pharmacol. Toxicol.* **2003**, *43*, 359-379.

(32) Guo, Q.; Lu, M.; Marky, L. A.; Kallenbach, N. R. *Biochemistry* **1992**, *31*, 2451-2455.

(33) Kelland, L. R.; Gowan, S. M.; Perry, P. J.; Neidle, S. *Br. J. Cancer* **1998**, *78*, 18-18.

(34) Perry, P. J.; Gowan, S. M.; Reszka, A. P.; Polucci, P.; Jenkins, T. C.; Kelland, L. R.; Neidle, S. *J. Med. Chem.* **1998**, *41*, 3253-3260.

(35) Jenkins, T. C. *Curr. Med. Chem.***2000**, *7*, 99-115.

(36) Izbicka, E.; Wheelhouse, R. T.; Raymond, E.; Davidson, K. K.; Lawrence, R. A.; Sun, D. Y.; Windle, B. E.; Hurley, L. H.; Von Hoff, D. D. *Cancer Res.* **1999**, *59*, 639-644.

(37) Fedoroff, O. Y.; Salazar, M.; Han, H. Y.; Chemeris, V. V.; Kerwin, S. M.; Hurley, L. H. *Biochemistry* **1998**, *37*, 12367-12374.

(38) Kern, J. T.; Kerwin, S. M. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 3395-3398.

(39) Kerwin, S. M.; Chen, G.; Kern, J. T.; Thomas, P. W. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 447-450.

(40) Kerwin, S. M.; Thomas, P. W.; Kern, J. T. *Abstracts of Papers of the American Chemical Society* **2001**, *222*, U663-U663.

(41) Kern, J. T.; Thomas, P. W.; Kerwin, S. M. *Biochemistry* **2002**, *41*, 12568-12568.

(42) Beck, J. L.; Colgrave, M. L.; Ralph, S. F.; Sheil, M. M. *Mass. Spectrom. Rev.* **2001**, *20*, 61-87.

(43) Hofstadler, S. A.; Griffey, R. H. *Chem. Rev.* **2001**, *101*, 377-390.

(44) Hofstadler, S. A.; Sannes-Lowery, K. A. *Nat. Rev. Drug Discov.* **2006**, *5*, 585-595.

(45) Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. *Science* **1989**, *246*, 64-71.

(46) Smith, R. D.; Lightwahl, K. J.; Winger, B. E.; Loo, J. A. *Org. Mass Spectrom.* **1992**, *27*, 811-821.

(47) Bruce, J. E.; Anderson, G. A.; Chen, R. D.; Cheng, X. H.; Gale, D. C.; Hofstadler, S. A.; Schwartz, B. L.; Smith, R. D. *Rapid Commun. Mass Spectrom.* **1995**, *9*, 644-650.

(48) Loo, J. A. *Bioconjugate Chem.* **1995**, *6*, 644-665.

(49) Loo, J. A. *Mass Spectrom. Rev.* **1997**, *16*, 1-23.

(50) Lightwahl, K. J.; Springer, D. L.; Winger, B. E.; Edmonds, C. G.; Camp, D. G.; Thrall, B. D.; Smith, R. D. *J. Am. Chem. Soc.* **1993**, *115*, 803-804.

(51) Gale, D. C.; Smith, R. D. *J. Am. Soc. Mass. Spectrom.* **1995**, *6*, 1154-1164.

(52) Gale, D. C.; Goodlett, D. R.; Lightwahl, K. J.; Smith, R. D. *J. Am. Chem. Soc.* **1994**, *116*, 6027-6028.

(53) Kapur, A.; Beck, J. L.; Sheil, M. M. *Rapid Commun. Mass Spectrom.* **1999**, *13*, 2489-2497.

(54) Colgrave, M. L.; Beck, J. L.; Sheil, M. M.; Searle, M. S. *Chem. Commun.* **2002**, 556-557.

(55) Gabelica, V.; De Pauw, E.; Rosu, F. *J. Mass Spectrom.* **1999**, *34*, 1328-1337.

(56) Wan, K. X.; Shibue, T.; Gross, M. L. *J. Am. Chem. Soc.* **2000**, *122*, 300-307.

(57) Goodlett, D. R.; Camp, D. G.; Hardin, C. C.; Corregan, M.; Smith, R. D. *Biol. Mass Spectrom.* **1993**, *22*, 181-183.

(58) Rosu, F.; Gabelica, V.; Houssier, C.; Colson, P.; De Pauw, E. *Rapid Commun. Mass Spectrom.* **2002**, *16*, 1729-1736.

(59) David, W. M.; Brodbelt, J.; Kerwin, S. M.; Thomas, P. W. *Anal. Chem.* **2002**, *74*, 2029-2033.

(60) Mazzitelli, C. L.; Brodbelt, J. S.; Kern, J. T.; Rodriguez, M.; Kerwin, S. M. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 593-604.

(61) Rosu, F.; De Pauw, E.; Guittat, L.; Alberti, P.; Lacroix, L.; Mailliet, P.; Riou, J. F.; Mergny, J. L. *Biochemistry* **2003**, *42*, 10361-10371.

(62) Baker, E. S.; Lee, J. T.; Sessler, J. L.; Bowers, M. T. *J. Am. Chem. Soc.* **2006**, *128*, 2641-2648.

(63) Guittat, L.; De Cian, A.; Rosu, F.; Gabelica, V.; De Pauw, E.; Delfourne, E.; Mergny, J. L. *Biochim. Biophys. Acta* **2005**, *1724*, 375-384.

(64) Guittat, L.; Alberti, P.; Rosu, F.; Van Miert, S.; Thetiot, E.; Pieters, L.; Gabelica, V.; De Pauw, E.; Ottaviani, A.; Riou, J. F.; Mergny, J. L. *Biochimie* **2003**, *85*, 535-547.

(65) Rosu, F.; Gabelica, V.; Shin-ya, K.; De Pauw, E. *Chem. Commun.* **2003**, 2702-2703.

(66) Carrasco, C.; Rosu, F.; Gabelica, V.; Houssier, C.; De Pauw, E.; Garbay-Jaureguiberry, C.; Roques, B.; Wilson, W. D.; Chaires, J. B.; Waring, M. J.; Bailly, C. *ChemBiochem* **2002**, *3*, 1235-1241.

(1) Watson, J. D.; Crick, F. H. C. *Nature* **1953**, *171*, 964-967.

(2) Arnott, S.; Chandrasekaran, R.; Selsing, R. *Structure and Conformation of Nucleic Acids and Protein-Nucleic Acid Interactions*; Univeristy Park Press: Baltimore, 1975.

(1) Goodman, L. S., Hardman, J. G., Limbird, L. E., and Gilman, A. G. *Goodman & Gilman's the Pharmacological Basis of Therapeutics*, 10th ed.; McGraw-Hill: New York, 2001 (1381-1460).

(2) Propst, C. L., and Perun, T. J. *Nucleic Acid Targeted Drug Design*; M. Dekker: New York, 1992, (1-12).

(3) Hofstadler, S. A.; Griffey, R. H. *Chem. Rev.* **2001**, *101*, 377-390.

(4) Beck, J. L.; Colgrave, M. L.; Ralph, S. F.; Sheil, M. M. E. *Mass Spectrom. Rev.* **2001,** *20*, 61-87.

(5) Brana, M. F.; Cacho, M.; Gradillas, A.; de Pascual-Teresa, B.; Ramos, A. *Curr. Pharm. Des.* **2001,** *7*, 1745-1780.

(6) Triolo, A.; Arcamone, F. M.; Raffaelli, A.; Salvadori, P. *J. Mass Spectrom.* **1997,** *32*, 1186-1194.

(7) Kapur, A.; Beck, J. L.; Sheil, M. M. *Rapid Commun. Mass Spectrom.* **1999,** *13*, 2489-2497.

(8) Gupta, R.; Kapur, A.; Beck, J. L.; Sheil, M. M. *Rapid Commun. Mass Spectrom.* **2001,** *15*, 2472-2480.

(9) Colgrave, M. L.; Beck, J. L.; Sheil, M. M.; Searle, M. S. *Chem. Commun*. **2002,** *6*, 556-557.

(10) Furlan, R. L. A.; Watt, S. J.; Garrido, L. M.; Amarante-Mendes, G. P.; Nur-E-Alam, M.; Rohr, J.; Brana, A.; Mendez, C.; Salas, J. A.; Sheil, M. M.; Beck, J. L.; Padilla, G. *J. Antibiot.* **2004,** *57*, 647-654.

(11) Urathamakul, T.; Beck, J. L.; Sheil, M. M.; Aldrich-Wright, J. R.; Ralph, S. F. *Dalton Trans.* **2004,** *17*, 2683-2690.

(12) Wan, K. X.; Gross, M. L.; Shibue, T. *J. Am. Soc. Mass Spectrom.* **2000,** *11*, 450-457.

(13) Wan, K. X.; Shibue, T.; Gross, M. L. *J. Am. Chem. Soc.* **2000,** *122*, 300-307.

(14) Beck, J. L.; Gupta, R.; Urathamakul, T.; Williamson, N. L.; Sheil, M. M.; Aldrich-Wright, J. R.; Ralph, S. F. *Chem. Commun.* **2003,** *5*, 626-627.

(15) Gupta, R.; Beck, J. L.; Ralph, S. F.; Sheil, M. M.; Aldrich-Wright, J. R. *J. Am. Soc. Mass Spectrom.* **2004,** *15*, 1382-1391.

(16) Gabelica, V.; De Pauw, E.; Rosu, F. *J. Mass Spectrom.* **1999,** *34*, 1328-1337.

(17) Greig, M. J.; Robinson, J. M. *Journal of Biomolecular Screening* **2000,** *5*, 441-454.

(18) Rosu, F.; De Pauw, E.; Guittat, L.; Alberti, P.; Lacroix, L.; Mailliet, P.; Riou, J. F.; Mergny, J. L. *Biochemistry* **2003,** *42*, 10361-10371.

(19) Reyzer, M. L.; Brodbelt, J. S.; Kerwin, S. M.; Kumar, D. *Nucleic Acids Res.* **2001,** *29*, e103.

(20) Carrasco, C.; Rosu, F.; Gabelica, V.; Houssier, C.; De Pauw, E.; Garbay-Jaureguiberry, C.; Roques, B.; Wilson, W. D.; Chaires, J. B.; Waring, M. J.; Bailly, C. *ChemBiohem.* **2002,** *3*, 1235-1241.

(21) Lokey, R. S.; Kwok, Y.; Guelev, V.; Pursell, C. J.; Hurley, L. H.; Iverson, B. L. *J. Am. Chem. Soc.* **1997,** *119*, 7202-7210.

(22) Lee, J.; Guelev, V.; Sorey, S.; Hoffman, D. W.; Iverson, B. L. *J. Am. Chem. Soc.* **2004,** *126*, 14036-14042.

(23) Guelev, V.; Lee, J.; Ward, J.; Sorey, S.; Hoffman, D. W.; Iverson, B. L. *Chem. Biol.* **2001,** *8*, 415-425.

(24) Guelev, V.; Sorey, S.; Hoffman, D. W.; Iverson, B. L. *J. Am. Chem Soc.* **2002,** *124*, 2864-2865.

(25) Chu, Y.; Lynch, V.; Iverson, B. L. *Tetrahedron*, **2006,** *62*, 5536-5548.

(26) Guelev, V. M.; Cubberley, M. S.; Murr, M. M.; Lokey, R. S.; Iverson, B. L. *Method. Enzymol.* **2001,** *340*, 556-570.

(27) Cubberley, M. S.; Iverson, B. L.. *J. Am. Chem. Soc.* **2001,** *123*, 7560-7563.

(28) Sambrook, J.; Russell, D. W. *Molecular cloning : a laboratory manual*; 3rd ed.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, N.Y., **2001**, 5.4-5.86**.**

(29) Guelev, V. Ph.D. Thesis, University of Texas at Austin, Austin, TX, **2002**.

(30) Rueda, M.; Luque, F. J.; Orozco, M. *J. Am. Chem. Soc.* **2005,** *127*, 11690-11698.

(31) Gidden, J.; Ferzoco, A.; Baker, E. S.; Bowers, M. T. *J. Am. Chem. Soc.* **2004,** *126*, 15132-15140.

.(32) Mazzitelli, C. L.; Kern, J. T.; Rodriguez, M.; Brodbelt, J. S.; Kerwin, S. M. *J. Am. Soc. Mass Spectrom.* **2006,** *17*, 593-604.

(33) Liu, Z.-R.; Hecker, K. H.; Rill, R. L. *J. Biomol. Struct. Dynam.* **1996,** *14*, 331-339

(34) Keller, K. M.; Zhang, J. M.; Oehlers, L.; Brodbelt, J. S. *J. Mass Spectrom.* **2005,** *40*, 1362-1371.

(35) Gabelica, V.; De Pauw, E. *J. Am. Soc. Mass Spectrom.* **2002,** *13*, 91-98.

(1) Ueki, M.; Ueno, K.; Miyadoh, S.; Abe, K.; Shibata, K.; Taniguchi, M.; Oi, S. M. *J. Antibiot.* **1993**, *46*, 1089-1094.

(2) Shibata, K.; Kashiwada, M.; Ueki, M.; Taniguchi, M. UK-1, *J. Antibiot.* **1993**, *46*, 1095-1100.

(3) Reynolds, M. B.; DeLuca, M. R.; Kerwin, S. M. *Bioorg. Chem.* **1999**, *27*, 326-337.

(4) Kumar, D.; Jacob, M. R.; Reynolds, M. B.; Kerwin, S. M. *Bioorg. Med. Chem.* **2002**, *10*, 3997-4004.

(5) Wang, B. B.; Maghami, N.; Goodlin, V. L.; Smith, P. J. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 3221-3226.

(6) Oehlers, L.; Mazzitelli, C. L.; Brodbelt, J. S.; Rodriguez, M.; Kerwin, S. *J. Am. Soc. Mass Spectrom.* **2004**, *15*, 1593-1603.

(7) Reyzer, M. L.; Brodbelt, J. S.; Kerwin, S. M.; Kumar, D. *Nucleic Acids Res.* **2001**, *29*, art. no.-e103.

(8) Sato, S.; Kajiura, T.; Noguchi, M.; Takehana, K.; Kobayasho, T.; Tsuji, T. *J. Antibiot.* **1997**, *54*, 102.

(9) Rodriguez, M.; Kerwin, S. *in preparaion* **2007**.

(10) Hofstadler, S. A.; Griffey, R. H. *Chem. Rev.* **2001**, *101*, 377-390.

(11) Beck, J. L.; Colgrave, M. L.; Ralph, S. F.; Sheil, M. M. *Mass Spectrom. Rev.* **2001**, *20*, 61-87.

(12) Hofstadler, S. A.; Sannes-Lowery, K. A. *Nat. Rev. Drug Dis.* **2006**, *5*, 585-595.

(13) Gale, D. C.; Smith, R. D. *J. Am. Soc. Mass Spectrom.* **1995**, *6*, 1154-1164.

(14) Gabelica, V.; Galic, N.; Rosu, F.; Houssier, C.; De Pauw, E. *J. Mass Spectrom.* **2003**, *38*, 491-501.

(15) Gabelica, V.; Rosu, F.; Houssier, C.; De Pauw, E. *Rapid Commun. Mass Spectrom.* **2000**, *14*, 464-467.

(16) Rosu, F.; Gabelica, V.; Houssier, C.; De Pauw, E. *Nucleic Acids Res.* **2002**, *30*, e82.

(17) Gabelica, V.; De Pauw, E.; Rosu, F. *J. Mass Spectrom.* **1999**, *34*, 1328-1337.

(18) Beck, J. L.; Gupta, R.; Urathamakul, T.; Williamson, N. L.; Sheil, M. M.; Aldrich-Wright, J. R.; Ralph, S. F. *Chem. Commun.* **2003**, *5*, 626-627.

(19) Wan, K. X.; Shibue, T.; Gross, M. L. *J. Am. Chem. Soc.* **2000**, *122*, 300-307.

(20) DeLuca, M. R.; Kerwin, S. M. *Tetrahedron Lett.* **1997**, *38*, 199-202.

(21) Hoveyda, H. R.; Rettig, S. J.; Orvig, C. *Inorg. Chem.* **1993**, *32*, 4909-4913.

(22) Tanaka, K.; Kumagai, T.; Aoki, H.; Deguchi, M.; Iwata, S. *J. Org. Chem.* **2001**, *66*, 7328-7333.

(23) Vachet, R. W.; Callahan, J. H. *J. Mass Spectrom.* **2000**, *35*, 311-320.

(24) Williams, S. M.; Brodbelt, J. S. *J. Am. Soc. Mass Spectrom.* **2004**, *15*, 1039-1054.

(25) Satterfield, M.; Brodbelt, J. S. *J. Am. Soc. Mass. Spectrom.* **2001**, *12*, 537-549.

(26) Perera, B. A.; Ince, M. P.; Talaty, E. R.; Van Stipdonk, M. J. *Rapid Commun. Mass Spectrom.* **2001**, *15*, 615-622.

(27) Vachet, R. W.; Hartman, J. A. R.; Callahan, J. H. *J. Mass Spectrom.* **1998**, *33*, 1209-1225.

(28) Kapur, A.; Beck, J. L.; Sheil, M. M. *Rapid Commun. Mass Spectrom.* **1999**, *13*, 2489-2497.

(29) Mazzitelli, C. L.; Kern, J. T.; Rodriguez, M.; Brodbelt, J. S.; Kerwin, S. M. *J. Am. Soc. Mass Spectrom.* **2006,** *17*, 593-604.

(30) Douglas, B.; McDaniel, D.; Alexander, J. *Concepts and Models of Inorganic Chemistry*, Third ed.; John Wiley & Sons, Inc.: New York, NY, 1994.

(31) Wan, K. X.; Gross, M. L.; Shibue, T. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 450-457.

(32) Keller, K. M.; Zhang, J. M.; Oehlers, L.; Brodbelt, J. S. *J. Mass Spectrom.* **2005**, *40*, 1362-1371.

(33) Gabelica, V.; De Pauw, E. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 91-98.

(1) Goodman, L. S., Hardman, J. G., Limbird, L. E., and Gilman, A. G. *Goodman & Gilman's the Pharmacological Basis of Therapeutics*, 10th ed.; McGraw-Hill: New York, 2001.

(2) Propst, C. L., and Perun, T. J. *Nucleic Acid Targeted Drug Design*; M. Dekker: New York, 1992.

(3) Hofstadler, S. A., and Griffey, R. H. *Chem. Rev.* **2001** *101*, 377-390.

(4) Beck, J. L.;, Colgrave, M. L., Ralph, S. F., and Sheil, M. M. *Mass Spectrom. Rev.* **2001**, *20*, 61-87.

(5) Gabelica, V. and De Pauw, E. *J. Am. Soc. Mass Spectrom.***2001**, *13*, 91-98.

(6) Gabelica, V., De Pauw, E., and Rosu, F. *J. Mass Spectrom.* **1999**, *34*, 1328-1337.

(7) Wan, K. X., Gross, M. L., and Shibue, T. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 450-457.

(8) Gale, D. C., and Smith, R. D. *J. Am. Soc. Mass Spectrom.* **1995**, *6*, 1154-1164.

(9) Triolo, A., Arcamone, F. M., Raffaelli, A., and Salvadori, P. *J. Mass Spectrom.*, **1997**, *32*, 1186-1194.

(10) Hofstadler, S. A., Sannes-Lowery, K. A., Crooke, S. T., Ecker, D. J., Sasmor, H., Manalili, S., and Griffey, R. H. *Anal. Chem.* **1999**, *71*, 3436-3440.

(11) Kapur, A., Beck, J. L., and Sheil, M. M. *Rapid Commun. Mass Spectrom.* **1999**, *13*, 2489-2497.

(12) Greig, M. J., and Robinson, J. M. *J. Biomol. Screen.***2000**, *5*, 441-454.

(13) Wan, K. X., Shibue, T., and Gross, M. L. *J. Am. Chem. Soc.***2000**, *122*, 300-307.

(14) Gabelica, V., Rosu, F., Houssier, C., and De Pauw, E. *Rapid Commun. Mass Spectrom.* **2000**, *14*, 464-467.

(15) Iannitti-Tito, P., Weimann, A., Wickham, G., and Sheil, M. M. *Analyst* **2000**, *125*, 627-633.

(16) Gupta, R., Kapur, A., Beck, J. L., and Sheil, M. M. *Rapid Commun. Mass Spectrom.* **2001**, *15*, 2472-2480.

(17) Rosu, F., Gabelica, V., Houssier, C., and De Pauw, E. *Nucleic Acids Res.* **2002**, *30*, e82.

(18) Carrasco, C., Rosu, F., Gabelica, V., Houssier, C., De Pauw, E., Garbay-Jaureguiberry, C., Roques, B., Wilson, W. D., Chaires, J. B., Waring, M. J., and Bailly, C. *ChemBiochem* **2002**, *3*, 1235-1241.

(19) Colgrave, M. L., Beck, J. L., Sheil, M. M., and Searle, M. S. *Chem. Comm.* **2002**, *6*, 556-557.

(20) Guittat, L., Alberti, P., Rosu, F., Van Miert, S., Thetiot, E., Pieters, L., Gabelica, V., De Pauw, E., Ottaviani, A., Riou, J. F., and Mergny, J. L. *Biochimie* **2003**, *85*, 535-547.

(21) Gabelica, V., Galic, N., Rosu, F., Houssier, C., and De Pauw, E. *J. Mass Spectrom.* **2003**, *38*, 491-501.

(22) Colgrave, M. L., Iannitti-Tito, P., Wickham, G., and Sheil, M. M. *Aust. J. Chem.* **2003**, *56*, 401-413.

(23) Rosu, F., De Pauw, E., Guittat, L., Alberti, P., Lacroix, L., Mailliet, P., Riou, J. F., and Mergny, J. L. *Biochemistry* **2003**, *42*, 10361-10371.

(24) Rosu, F., Gabelica, V., Shin-ya, K., and De Pauw, E. C*hem. Commun.* **2003**, *21*, 2702-2703.

(25) Beck, J. L., Gupta, R., Urathamakul, T.. Williamson, N. L., Sheil, M. M., Aldrich-Wright, J. R., and Ralph, S. F. *Chem Commun.* **2003**, *9*, 626-627.

(26) David, W. M., Brodbelt, J., Kerwin, S. M., and Thomas, P. W. *Anal. Chem.***2002**, *74*, 2029-2033.

(27) Rosu, F., Gabelica, V., Houssier, C., Colson, P., and Pauw, E. D. *Rapid Commun. Mass Spectrom.***2002**, *16*, 1729-1736.

(28) Vairamani, M., and Gross, M. L. *J. Am. Chem. Soc.* **2003**, *125*, 42-43.

(29) Kerwin, S. M. *Curr. Pharm. Des.* **2000**, *6*, 441-471.

(30) Hurley, L. H., Wheelhouse, R. T., Sun, D., Kerwin, S. M., Salazar, M., Fedoroff, O. Y., Han, F. X., Han, H. Y., Izbicka, E., and Von Hoff, D. D. *Pharmacol. Ther.* **2000**, *85*, 141-158.

(31) Wellinger, R. J., and Sen, D. *Eur. J. Cancer* **1997**, *33*, 735-749.

(32) Bearss, D. J., Hurley, L. H., and Von Hoff, D. D. *Oncogene* **2000**, *19*, 6632-6641.

(33) Mergny, J. L., Mailliet, P., Lavelle, F., Riou, J. F., Laoui, A., and Helene, C. *Anti-Cancer Drug Des.* **1999**, *14*, 327-339.

(34) Counter, C. M., Hahn, W. C., Wei, W. Y., Caddle, S. D., Beijersbergen, R. L., Lansdorp, P. M., Sedivy, J. M., and Weinberg, R. A. *Proc. Nal. Acad. Sci. U.S.A.* **1998**, *95*, 14723-14728.

(35) Fedoroff, O. Y., Salazar, M., Han, H. Y., Chemeris, V. V., Kerwin, S. M., and Hurley, L. H. *Biochemistry* **1998**, *37*, 12367-12374.

(36) Kern, J. T., and Kerwin, S. M. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 3395-3398.

(37) Kern, J. T., Thomas, P. W., and Kerwin, S. M. *Biochemistry* **2002** *41*, 11379-11389.

(38) Kerwin, S. M., Chen, G., Kern, J. T., and Thomas, P. W. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 447-450.

(39) Rossetti, L., Franceschin, M., Bianco, A., Ortaggi, G., and Savino, M. *Bioorg. Med. Chem. Lett.* **2002,** *12*, 2527-2533.

(40) Read, M. A., and Neidle, S. *Biochemistry* **2000**, *39*, 13422-13432.

(41) Han, H. Y., Bennett, R. J., and Hurley, L. H. *Biochemistry* **2000**, *39*, 9311-9316.

(42) Han ,H. Y., Cliff, C. L., and Hurley, L. H. *Biochemistry* **1999**, *38*, 6981-6986.

(43) Read, M. A., Wood, A. A., Harrison, J. R., Gowan, S. M., Kelland, L. R., Dosanjh, H. S., and Neidle, S. *J. Med. Chem.* **1999**, *42*, 4538-4546.

(44) Pasternack, R. F., and Collings, P. J. *Science* **1995**, *269*, 935-939.

(45) Clark, G. R., Pytel, P. D., Squire, C. J., and Neile, S. *J. Am. Chem. Soc.* **2003**, *125*, 4066-4067.

(46) Randazzo, A., Galeone, A., and Mayol, L. *Chem. Comm.* **2001,** *11*, 1030-1031

(47) Kerwin, S. M., Sun, D., Kern, J. T., Rangan, A., and Thomas, P. W. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 2411-2414.

(1) Rezler, E. M.; Bearss, D. J.; Hurley, L. H. *Curr. Opin. Pharmacol.* **2002**, *2*, 415-423.

(2) Bearss, D. J.; Hurley, L. H.; Von Hoff, D. D. *Oncogene* **2000**, *19*, 6632-6641.

(3) Rezler, E. M.; Bearss, D. J.; Hurley, L. H. *Annu. Rev. Pharmacol. Toxicol.* **2003**, *43*, 359-379.

(4) Wellinger, R. J.; Sen, D. *Eur. J. Cancer* **1997**, *33*, 735-749.

(5) Harley, C. B.; Futcher, A. B.; Greider, C. W. *Nature* **1990**, *345*, 458-460.

(6) Kim, N. W.; Piatyszek, M. A.; Prowse, K. R.; Harley, C. B.; West, M. D.; Ho, P. L. C.; Coviello, G. M.; Wright, W. E.; Weinrich, S. L.; Shay, J. W. *Science* **1994**, *266*, 2011-2015.

(7) Nakamura, T. M.; Morin, G. B.; Chapman, K. B.; Weinrich, S. L.; Andrews, W. H.; Lingner, J.; Harley, C. B.; Cech, T. R. *Science* **1997**, *277*, 955-959.

(8) Moyzis, R. K.; Buckingham, J. M.; Cram, L. S.; Dani, M.; Deaven, L. L.; Jones, M. D.; Meyne, J.; Ratliff, R. L.; Wu, J. R. *Pro.Natl. Acad. Sci. U. S. A.* **1988**, *85*, 6622-6626.

(9) Blackburn, E. H.; Gall, J. G. *J. Mol. Biol.***1978**, *120*, 33-53.

(10) Haider, S.; Parkinson, G. N.; Neidle, S. *J. Mol. Biol.* **2002**, *320*, 189-200.

(11) Tohl, J.; Eimer, W. *Biophys. Chem.* **1997**, *67*, 177-186.

(12) Kerwin, S. M. *Curr. Pharm. Des.* **2000**, *6*, 441-471.

(13) Mergny, J. L.; Mailliet, P.; Lavelle, F.; Riou, J. F.; Laoui, A.; Helene, C. *Anti-Cancer Drug Des.* **1999**, *14*, 327-339.

(14) Hurley, L. H.; Wheelhouse, R. T.; Sun, D.; Kerwin, S. M.; Salazar, M.; Fedoroff, O. Y.; Han, F. X.; Han, H. Y.; Izbicka, E.; Von Hoff, D. D. *Pharmacol. Ther.* **2000**, *85*, 141-158.

(15) Han, H. Y.; Hurley, L. H. *Trends Pharmacol. Sci.* **2000**, *21*, 136-142.

(16) David, W. M.; Brodbelt, J.; Kerwin, S. M.; Thomas, P. W. *Anal. Chem.* **2002**, *74*, 2029-2033.

(17) Mazzitelli, C. L.; Kern, J. T.; Rodriguez, M.; Brodbelt, J. S.; Kerwin, S. M. *J. Am. Soc. Mass Spectrom.* **2006**.

(18) Baker, E. S.; Lee, J. T.; Sessler, J. L.; Bowers, M. T. *J. Am. Chem. Soc.* **2006**, *128*, 2641-2648.

(19) Rosu, F.; De Pauw, E.; Guittat, L.; Alberti, P.; Lacroix, L.; Mailliet, P.; Riou, J. F.; Mergny, J. L. *Biochemistry* **2003**, *42*, 10361-10371.

(20) Guittat, L.; Alberti, P.; Rosu, F.; Van Miert, S.; Thetiot, E.; Pieters, L.; Gabelica, V.; De Pauw, E.; Ottaviani, A.; Riou, J. F.; Mergny, J. L. *Biochimie* **2003**, *85*, 535-547.

(21) Guittat, L.; De Cian, A.; Rosu, F.; Gabelica, V.; De Pauw, E.; Delfourne, E.; Mergny, J. L. *Biophys. Acta* **2005**, *1724*, 375-384.

(22) Carrasco, C.; Rosu, F.; Gabelica, V.; Houssier, C.; De Pauw, E.; Garbay-Jaureguiberry, C.; Roques, B.; Wilson, W. D.; Chaires, J. B.; Waring, M. J.; Bailly, C. *ChemBioChem* **2002**, *3*, 1235-1241.

(23) Rosu, F.; Gabelica, V.; Shin-ya, K.; De Pauw, E. *Chem. Commun.* **2003**, 2702-2703.

(24) Gidden, J.; Ferzoco, A.; Baker, E. S.; Bowers, M. T. *J. Am. Chem. So.c* **2004**, *126*, 15132-15140.

(25) Gabelica, V.; De Pauw, E. *Int. J. Mass Spectrom.* **2002**, *219*, 151-159.

(26) Pan, S.; Sun, X. J.; Lee, J. K. *Int. J. Mass Spectrom.* **2006**, *253*, 238-248.

(27) Guo, X. H.; Bruist, M. F.; Davis, D. L.; Bentzley, C. M. *Nucleic Acids Res.* **2005**, *33*, 3659-3666.

(28) Guo, X. H.; Liu, Z. Q.; Liu, S. Y.; Bentzley, C. M.; Bruist, M. F. *Anal. Chem.* **2006**, *78*, 7259-7266.

(29) Wan, K. X.; Gross, M. L.; Shibue, T. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 450-457.

(30) Gabelica, V.; De Pauw, E. *J.Mass Spectrom.* **2001**, *36*, 397-402.

(31) Baker, E. S.; Bernstein, S. L.; Gabelica, V.; De Pauw, E.; Bowers, M. T. *Int. J. Mass Spectrom.* **2006**, *253*, 225-237.

(32) Rosu, F.; Gabelica, V.; Houssier, C.; Colson, P.; De Pauw, E. *Rapid Commun. Mass Spectrom.y* **2002**, *16*, 1729-1736.

(33) Rueda, M.; Kalko, S. G.; Luque, F. J.; Orozco, M. *J. Am. Chem. Soc.* **2003**, *125*, 8007-8014.

(34) Rueda, M.; Luque, F. J.; Orozco, M. *J. Am. Chem. Soc.* **2005**, *127*, 11690-11698.

(35) Rueda, M.; Luque, F. J.; Orozco, M. *J. Am. Chem. Soc.* **2006**, *128*, 3608-3619.

(36) Kerwin, S. M.; Chen, G.; Kern, J. T.; Thomas, P. W. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 447-450.

(37) Caceres, C.; Wright, G.; Gouyette, C.; Parkinson, G.; Subirana, J. A. *Nucleic Acids Res.* **2004**, *32*, 1097-1102.

(38) Horvath, M. P.; Schultz, S. C. *J. Mol. Biol.* **2001**, *310*, 367-377.

(39) Muller, S.; Diederichs, K.; Breed, J.; Kissmehl, R.; Hauser, K.; Plattner, H.; Welte, W. *J. Mol. Biol.* **2002**, *315*, 141-153.

(40) Wang, Y.; Patel, D. J. *Structure* **1993**, *1*, 263-282.

(41) Case, D. A.; Darden, T. A.; Cheatham, T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Merz, K. M.; Wang, B.; Pearlman, D. A.; Crowley, M.; Brozell, S.; Tsui, V.; Gohlke, H.; Mongan, J.; Hornack, V.; Cui, G.; Berzoa, P.; Schafmeister, C.; Caldwell, J. W.; Ross, W. S.; Kollman, P. A. **2005**, AMBER 8, University of California, San Francisco.

(42) Wang, J. M.; Cieplak, P.; Kollman, P. A. *Journal of Comput. Chem.* **2000**, *21*, 1049-1074.

(43) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *Journal of Comput. Phys.* **1977**, *23*, 327-341.

(44) Srinivasan, J.; Cheatham, T. E.; Cieplak, P.; Kollman, P. A.; Case, D. A. *J.Am. Chem. Soc.* **1998**, *120*, 9401-9409.

(45) Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S. H.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.; Srinivasan, J.; Case, D. A.; Cheatham, T. E. *Acc. Chem. Res.* **2000**, *33*, 889-897.

(46) Wang, J.; Hou, T.; Xu, X. *Current Computer-Aided Drug Des.* **2006**, *2*, 287-306.

(47) Keller, K. M.; Zhang, J. M.; Oehlers, L.; Brodbelt, J. S. *J. Mass Spectrom.* **2005**, *40*, 1362-1371.

(48) Gabelica, V.; De Pauw, E. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 91-98.

(49) Pan, S.; Verhoeven, K.; Lee, J. K. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 1853-1865.

(50) Gale, D. C.; Smith, R. D. *J. Am. Soc. Mass Spectrom.* **1995**, *6*, 1154-1164.

(1) Hofstadler, S. A.; Sannes-Lowery, K. A. *Nat. Rev. Drug. Discov.* **2006**, *5*, 585-595.

(2) Hofstadler, S. A.; Griffey, R. H. *Chem. Rev.* **2001**, *101*, 377-390.

(3) Beck, J. L.; Colgrave, M. L.; Ralph, S. F.; Sheil, M. M. *Mass. Spectrom. Rev.* **2001**, *20*, 61-87.

(4) Gale, D. C.; Smith, R. D. *J. Am. Soc. Mass. Spectrom.* **1995**, *6*, 1154-1164.

(5) Gabelica, V.; De Pauw, E.; Rosu, F. *J. Mass. Spectrom.* **1999**, *34*, 1328-1337.

(6) Kapur, A.; Beck, J. L.; Sheil, M. M. *Rapid Commun. Mass Spectrom.* **1999**, *13*, 2489-2497.

(7) Gabelica, V.; Rosu, F.; Houssier, C.; De Pauw, E. *Rapid Commun. Mass Spectrom.* **2000**, *14*, 464-467.

(8) Wan, K. X.; Gross, M. L.; Shibue, T. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 450-457.

(9) Wan, K. X.; Shibue, T.; Gross, M. L. *J. Am. Chem. Soc.* **2000**, *122*, 300-307.

(10) Rosu, F.; Gabelica, V.; Houssier, C.; De Pauw, E. *Nucleic Acids Res.* **2002**, *30*, -.

(11) Mazzitelli, C. L.; Chu, Y.; Reczek, J. J.; Iverson, B. L.; Brodbelt, J. S. *J. Am. Soc. Mass. Spectrom.* **2006**, *in press*.

(12) Reyzer, M. L.; Brodbelt, J. S.; Kerwin, S. M.; Kumar, D. *Nucleic Acids Res.* **2001**, *29*, art. no.-e103.

(13) Oehlers, L.; Mazzitelli, C. L.; Brodbelt, J. S.; Rodriguez, M.; Kerwin, S. *J. Am. Soc. Mass Spectrom.* **2004**, *15*, 1593-1603.

(14) Rosu, F.; Gabelica, V.; Houssier, C.; Colson, P.; De Pauw, E. *Rapid Commun. Mass Spectrom.* **2002**, *16*, 1729-1736.

(15) David, W. M.; Brodbelt, J.; Kerwin, S. M.; Thomas, P. W. *Anal. Chem.* **2002**, *74*, 2029-2033.

(16) Guittat, L.; Alberti, P.; Rosu, F.; Van Miert, S.; Thetiot, E.; Pieters, L.; Gabelica, V.; De Pauw, E.; Ottaviani, A.; Riou, J. F.; Mergny, J. L. *Biochimie* **2003**, *85*, 535-547.

(17) Rosu, F.; De Pauw, E.; Guittat, L.; Alberti, P.; Lacroix, L.; Mailliet, P.; Riou, J. F.; Mergny, J. L. *Biochemistry* **2003**, *42*, 10361-10371.

(18) Rosu, F.; Gabelica, V.; Shin-ya, K.; De Pauw, E. *Chem. Commun.* **2003**, 2702-2703.

(19) Guittat, L.; De Cian, A.; Rosu, F.; Gabelica, V.; De Pauw, E.; Delfourne, E.; Mergny, J. L. *Biochim. Biophys. Acta* **2005**, *1724*, 375-384.

(20) Baker, E. S.; Lee, J. T.; Sessler, J. L.; Bowers, M. T. *J. Am. Chem. Soc.* **2006**, *128*, 2641-2648.

(21) Mazzitelli, C. L.; Kern, J. T.; Rodriguez, M.; Brodbelt, J. S.; Kerwin, S. M. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 593-604.

(22) Nielsen, P. E. *Journal of Molec. Recognit.* **1990**, *3*, 1-25.

(23) Bui, C. T.; Rees, K.; Cotton, R. G. H. *Curr. Pharmacogenomics* **2004**, *2*, 325-332.

(24) Kellersberger, K. A.; Yu, E.; Kruppa, G. H.; Young, M. M.; Fabris, D. *Anal. Chem.* **2004**, *76*, 2438-2445.

(25) Yu, E. T.; Zhang, Q. G.; Fabris, D. *J. Mol. Biol.* **2005**, *345*, 69-80.

(26) Yu, E.; Fabris, D. *Anal. Biochem.* **2004**, *334*, 356-366.

(27) Yu, E.; Fabris, D. *J. Mol. Biol.* **2003**, *330*, 211-223.

(28) Bui, C. T.; Rees, K.; Cotton, R. G. H. *Nucleos. Nucleot. Nucl.* **2003**, *22*, 1835-1855.

(29) Sasse-Dwight, S.; Gralla, J. D. *J. Biol. Chem.* **1989**, *264*, 8074-8081.

(30) Jeppesen, C.; Nielsen, P. E. *Febs Lett.* **1988**, *231*, 172-176.

(31) Bailly, C.; Gentle, D.; Hamy, F.; Purcell, M.; Waring, M. J. *Biochem. J.* **1994**, *300*, 165-173.

(32) Gille, H.; Messer, W. *Embo J.* **1991**, *10*, 1579-1584.

(33) Hsieh, D. J.; Camiolo, S. M.; Yates, J. L. *Embo J.* **1993**, *12*, 4933-4944.

(34) Farah, J. A.; Smith, G. R. *J. Mol. Biol.* **1997**, *272*, 699-715.

(35) Schaubach, O. L.; Dombroski, A. J. *J. Biol. Chem.* **1999**, *274*, 8757-8763.

(36) Rubin, C. M.; Schmid, C. W. *Nucleic Acids Res.* **1980**, *8*, 4613-4619.

(37) Jones, A. S.; Walker, R. T. *Nature* **1964**, *202*.

(38) Jiang, H.; Zacharias, W.; Amirhaeri, S. *Nucleic Acids Res.* **1991**, *19*, 6943-6948.

(39) Desantis, P.; Palleschi, A.; Savino, M.; Scipioni, A. *Biochemistry* **1990**, *29*, 9269-9273.

(40) Mitas, M.; Yu, A.; Dill, J.; Kamp, T. J.; Chambers, E. J.; Haworth, I. S. *Nucleic Acids Res.* **1995**, *23*, 1050-1059.

(41) Klysik, J.; Rippe, K.; Jovin, T. M. *Biochemistry* **1990**, *29*, 9831-9839.

(42) Bui, C. T.; Lambrinakos, A.; Cotton, R. G. H. *Biopolymers* **2003**, *70*, 628-636.

(43) Cotton, R. G. H. *Biochem. J.* **1989**, *263*, 1-10.

(44) Mcluckey, S. A.; Vanberkel, G. J.; Glish, G. L. *J. Am. Soc. Mass Spectrom.* **1992**, *3*, 60-70.

(45) Keller, K. M.; Brodbelt, J. S. *Anal. Biochem.* **2004**, *326*, 200-210.

(46) Wilson, J. J.; Brodbelt, J. S. *Anal. Chem.* **2006**, *78*, 6855-6862.

(47) Chu, Y. J.; Lynch, V.; Iverson, B. L. *Tetrahedron* **2006**, *62*, 5536-5548.

(48) Vandyke, M. M.; Dervan, P. B. *Science* **1984**, *225*, 1122-1127.

(49) Low, C. M. L.; Drew, H. R.; Waring, M. J. *Nucleic Acids Res.* **1984**, *12*, 4865-4879.

(50) Gao, X.; Patel, D. J. *Q. Rev. Biophysics* **1989**, *22*, 93-138.

(51) Waring, M. J.; Wakelin, L. P. G. *Nature* **1974**, *252*, 653-657.

(52) Rueda, M.; Luque, F. J.; Orozco, M. *J. Am. Chem. Soc.* **2005**, *127*, 11690-11698.

(53) Gidden, J.; Ferzoco, A.; Baker, E. S.; Bowers, M. T. *J. Am. Chem. Soc.* **2004**, *126*, 15132-15140.

(54) Sinden, R. R. *DNA Structure and Function*; Academic Press: San Diego, 1994.

(55) Sobell, H. M. *Proc. Natl. Acad. Sci. USA* **1985**, *82*, 5328-5331.

(56) Sobell, H. M.; Jain, S. C.; Sakore, T. D.; Nordman, C. E. *Nature new Biol.* **1972**, *231*, 200-205.

(57) Wakelin, L. P. G.; Waring, M. J. *Biochem. J.* **1976**, *157*, 721-740.

(58) Wang, J. C. *Biochim. Biophys. Acta* **1971**, *232*, 246-251.

(59) Chu, Y.; Sorrey, S.; Hoffman, D. W.; Iverson, B. L. *J. Am. Chem. Soc.* **2006**, *in press*.

(60) Neidle, S. *Biopolymers* **1997**, *44*, 105-121.

(61) Neidle, S.; Pearl, L. H.; Skelly, J. V. *Biochem J.* **1987**, *243*, 1-&.

(62) Keller, K. M.; Zhang, J. M.; Oehlers, L.; Brodbelt, J. S. *J. Mass Spectrom.* **2005**, *40*, 1362-1371.

(63) Gabelica, V.; De Pauw, E. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 91-98.

(64) Mcluckey, S. A.; Habibigoudarzi, S. *J. Am. Chem. Soc.* **1993**, *115*, 12085-12095.

(65) Wang, Y. S.; Zhang, Q. C.; Wang, Y. S. *J. Am. Soc. Mass Spectrom.* **2004**, *15*, 1565-1571.

(66) Mccarthy, J. G.; Williams, L. D.; Rich, A. *Biochemistry* **1990**, *29*, 6071-6081.

Vita

Carolyn Leigh Mazzitelli was born on December 13, 1979 in Renton, WA to Rick and Diana Mazzitelli. After graduating from Eastlake High School in Sammamish, WA in 1998, she enrolled at the University of North Carolina at Chapel Hill. While at UNC, she worked as an undergraduate research assistant in the laboratories of Dr. Royce Murray. She received her B. S. degree in chemistry from UNC in May of 2002 and went on to join the doctoral program in the Department of Chemistry and Biochemistry at the University of Texas at Austin in August of the same year.

Permanent address: 16 Holly Ave. #3, Cambridge, MA 02138

This dissertation was typed by the author.