Characterization of Nucleic Acid Non-Covalent Interactions by Fourier Transform Ion Cyclotron Resonance Tandem Mass Spectrometry and Gas-Phase Hydrogen/Deuterium Exchange

by

Jingjie Mo

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Chemistry)
in The University of Michigan
2008

Doctoral Committee:

Assistant Professor Kristina I. Håkansson, Chair
Professor Robert T. Kennedy
Associate Professor Nils G. Walter
Assistant Professor Hashim M. Al-Hashimi
Assistant Professor Jason E. Gestwicki

2008

To My Dear Family

Acknowledgements

After all these years, I've got quite a list of people who gave me the opportunity to complete my thesis research. I would like to take this opportunity to express my sincere gratitude to all of them.

First, I am deeply indebted to my supervisor, Dr. Kristina Håkansson, for her detailed and constructive comments and for her continuous support throughout this work. I am very lucky to have been one of the earliest students to join her group. Her wide knowledge and logical way of thinking have been of great value to me. This present work is possible due to her understanding, encouraging and personal guidance. I sincerely thank Kristina for all her stimulating conversations, suggestions, encouragement, guidance and research support throughout my Ph.D. research.

Second, I would like to thank my Ph.D. committee, who monitored the progress of my research work and provided valuable comments on my research proposal and this thesis: Prof. Hashim Al-Hashimi, Prof. Jason Gestwicki, Prof. Robert T. Kennedy, Prof. Nils G. Walter and Prof. Rowena G. Mathews. This dissertation could not have been completed without your critique, guidance and help.

Third, I would like to thank my labmates: Julie, Natasa, Jason, Yibing, Chris, Hangtian, Bo, Hyun Ju, Wen, and former labmates: Hye Kyong, Jiong, Haichuan, Gabriella, and

Liping, for providing a stimulating and fun environment in which to learn and grow. I have truly enjoyed being your colleague!

My research has been funded by a starter grant from the Petroleum Research Fund, a research award from the American Society for Mass Spectrometry (to Kristina) and a CAREER award from the National Science Foundation (CHE-05-47699). I would like to also thank the Department of Chemistry at the University of Michigan for providing me with financial support through a teaching assistantship.

Last, but not least, I would like to thank my parents for their continuous support. They bore me, raised me, supported me, taught me, and loved me. To them I dedicate this thesis. Finally, I have to say 'thank-you' to all my friends, wherever they are, for their valuable friendship.

Jingjie Mo

Mar. 20, 2008

Ann Arbor, MI

Table of Contents

List of Figures

Figure

List of Schemes

Scheme

List of Tables

Table

List of Abbreviations

Abbreviations

3D	three-dimensional
A	Adenine
A	Adenosine
A	Ampere
AI EDD	Activated ion electron capture dissociation
B	Base
C	Cytidine
C	Cytosine
CD	Circular dichroism
CID/CAD	Collision induced dissociation/Collision activated dissociation
CHEF	Correlated harmonic excitation fields
Da	Dalton
DIP	diphenylphenanthroline
DNA	Deoxyribonucleic Acid
e	Electron
ESI	Electrospray ionization
ECD	Electron capture dissociation
EDD	Electron detachment dissociation
EDTA	Ethylenediaminetetraacetic acid

EID	Electron induced dissociation
FA	Formic acid
FRET	Fluorescence resonance energy transfer
FT-ICR	Fourier transform ion cyclotron resonance
G	Guanine
G	Guanosine
HDX	Hydrogen/deuterium exchange
HPLC	High performance liquid chromatography
ICR	Ion cyclotron resonance
IRMPD	Infrared multiphoton dissociation
k	Thousand
m/z	mass-to-charge ratio
mRNA	Messenger RNA
μ	Micro
M	Million
M	Molarity
M	Precursor ion
MALDI	Matrix assisted laser desorption/ionization
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
NH_4OAc	Ammonium acetate
OMe	Methoxy
PAGE	Polyacrylamide gel electrophoresis
Phen	phenanthroline
ppm	part per million

PTMs	Posttranslational modifications
Q	Quadrupole (mass filter)
rRNA	Ribosomal RNA
RNA	Ribonucleic Acid
S	Serine (Ser in three letter code)
S/N	Signal-to-noise
SORI	Sustained off-resonance irradiation
tRNA	Transfer RNA
T	Tesla, a unit of magnetic field
T	Thymidine
T	Thymine
U	Uracil
U	Uridine
V	Volume
Y	Tyrosine (Tyr in three letter code)
W-C	Watson-Crick

Abstract

Characterization of Nucleic Acid Non-Covalent Interactions by Fourier Transform Ion Cyclotron Resonance Tandem Mass Spectrometry and Gas-Phase Hydrogen/Deuterium Exchange

by

Jingjie Mo

Chair: Kristina I. Håkansson

One of the long-term goals of biological chemistry is to understand the relationship between the 3-dimensional structures of biomolecules and their biological functions. Nucleic acids play important roles in a variety of fundamental biological processes and understanding of their structures has a profound impact on the progress of biological chemistry research. Common biophysical methods used to probe biomolecular higher order structures include X-ray crystallography, nuclear magnetic resonance spectroscopy, fluorescence, circular dichroism, calorimetry, infrared spectroscopy, Raman spectroscopy, and electron spin resonance spectroscopy. Mass spectrometry (MS) has a major advantage over these methods due to its capability of identifying and determining the abundances of different complexes from direct observation because the mass of every

molecule serves as the intrinsic detection “label”.

The utility of electrospray ionization (ESI) in MS to characterize non-covalent interactions of nearly every type of biomolecule has been described in numerous publications. In this thesis work, based on the “soft” characteristic of ESI, non-covalent interactions involved in higher order structures of nucleic acids as well as their complexes with small molecules were analyzed. Two strategies have been developed: first, a combination of infrared multiphoton dissociation and electron detachment dissociation and, second, gas-phase hydrogen/deuterium exchange (HDX) with MS detection.

With the first strategy, different cleavage patterns were observed for three isomeric DNA 15-mers, indicating that this approach allows probing of higher order structures of nucleic acids. However, the cleavage efficiency is low due to the intrinsic instrument configuration. The second strategy was applied to compare a series of nucleic acid hairpins and was shown to be able to correlate their solution-phase stabilities with their gas-phase HDX rates. This method was also used to compare the binding affinities of A-site RNAs with different aminoglycoside antibiotics and the results were found to be similar to those in previous experiments. Furthermore, we propose a relay mechanism for oligonucleotide gas-phase HDX in both positive and negative ion mode, and show that the HDX rates of nucleic acids are dependent on the gas-phase basicity (positive ion mode) and acidity (negative ion mode) of the nucleobases as well as their structural flexibility.

CHAPTER 1

INTRODUCTION

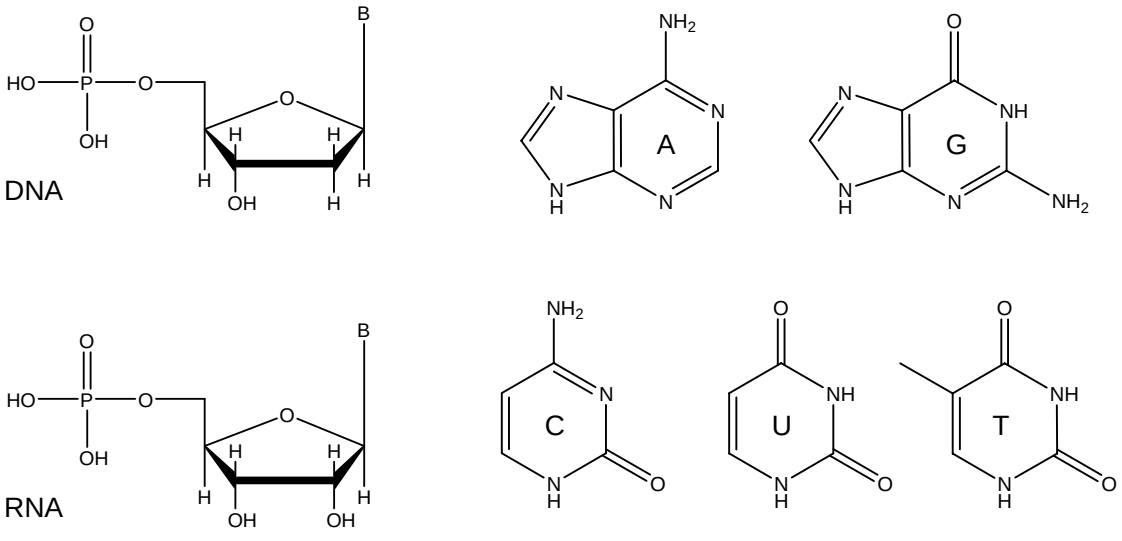
1.1. Non-Covalent Interactions in Nucleic Acids

Nucleic acids are macromolecules composed of nucleotide chains. Each nucleotide consists of three components (Scheme 1.1): a nitrogenous heterocyclic base (nucleobase), which is either a purine or a pyrimidine; a pentose sugar; and a phosphate group. Nucleic acid types differ in the structure of the sugar in their nucleotides: DNA contains 2-deoxyribose units while RNA contains ribose units. In addition, the nucleobases found in these two nucleic acid types are different: adenine (A), cytosine (C), and guanine (G) are found in both RNA and DNA, while thymine (T) mainly occurs in DNA and uracil (U) mainly occurs in RNA. Other rare nucleic acid bases can occur, for example inosine in strands of mature transfer RNA (tRNA). The order in which nucleobases appear in a nucleic acid often codes for the information carried by the molecule, i.e., nucleobases can serve as a genetic alphabet from which the amino acid sequence of each protein in our bodies is synthesized. The 'flow' of information from an organism's genome to the synthesis of its encoded proteins is referred to as the central dogma, where DNA is the principal inherited genetic material. In some

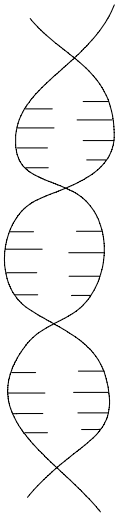
organisms, the inherited genetic material is RNA instead of DNA. For example, almost 60% of all characterized viruses have RNA genomes and these are more common in plant viruses than in animal viruses. Another important role that RNA plays in biological systems is as an "enzyme", so called "ribozyme". As early as in 1967, RNA was suggested to have catalytic function due to its various secondary structures. The catalytic domain of a ribozyme can break and re-form the phosphodiester bonds between nucleotides, lowering the activation energy for these reactions similar to protein-catalyzed reactions. The first ribozyme was discovered in the 1980s by Cech and Altman, which won them a Nobel prize.[1, 2] After that, the number of confirmed naturally occurring ribozymes continuously increases. Some ribozymes are self-splicing, i.e., they catalyze cleavage and resealing of their own nucleotide chain, whereas others catalyze these reactions on separate RNA molecules. Recently discovered ribozymes include the ribosomal RNAs of the protein biosynthetic machinery[3] and the small nuclear RNAs of the eukaryotic splicing machinery.[4]

Non-covalent interactions have a constitutive role in biological systems as well as in artificial supramolecular structures.[5, 6] In contrast to covalent interactions within molecules, non-covalent interactions are weak interactions that bind together different kinds of building blocks into supramolecular entities[7, 8] and are responsible for the three-dimensional conformations that biomolecules adopt, therefore playing a very significant role in the flexibility of these biomolecules, their interactions with each other, and with other molecules inside a cell. For example, non-covalent interactions enable the formation of DNA double helices (Scheme 1.2) and RNA secondary structures (Scheme 1.3), such as the hairpin, through hydrogen bonding; they enable enzymes to bind their

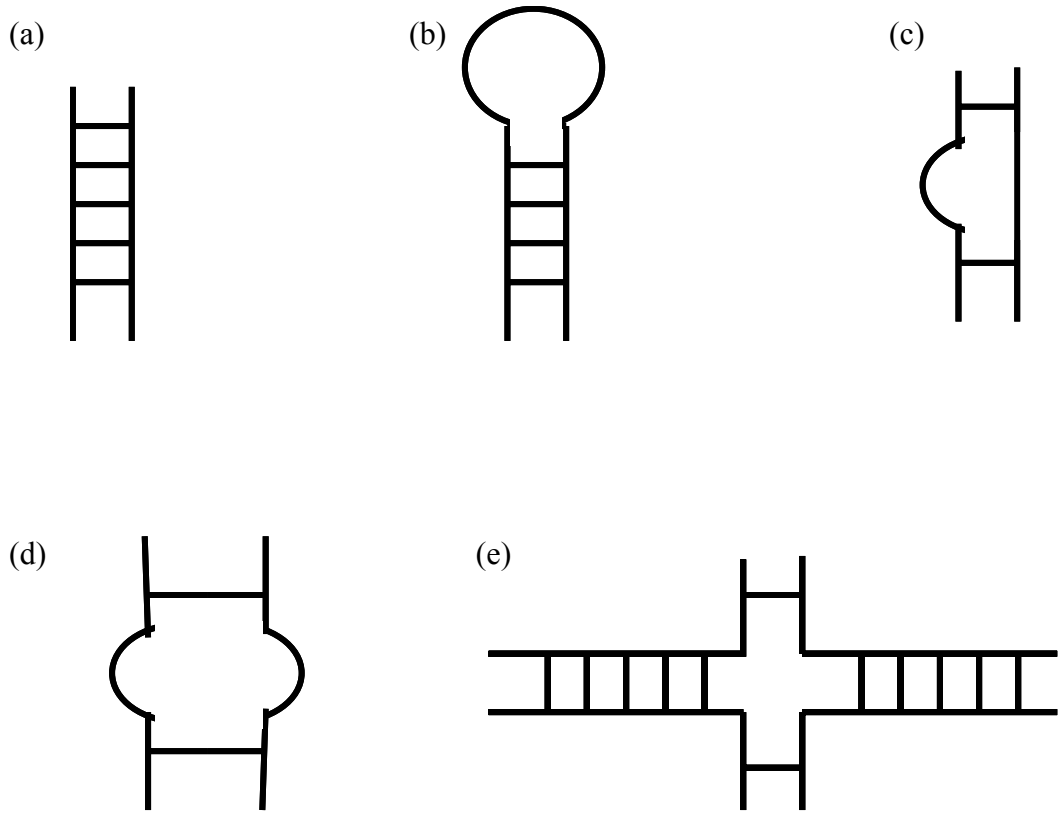
substrates, antibodies to bind their antigens, transcription factors to bind each other, and proteins to bind to their receptors. Non-covalent interactions also permit the assembly of such macromolecular machineries as ribosomes, actin filaments, microtubules, and many more.[9] Covalent bonds are generally shorter than 2 Å and require overlapping of partially occupied orbitals of interacting atoms that share a pair of electrons, whereas non-covalent interactions can function within a range of several angstroms, and thus no orbital overlapping is necessary because the attraction arises from the electronic properties of the building blocks.[9] The various non-covalent interactions can be classified as: ionic interactions, van der Waals interactions, hydrogen bonds, and hydrophobic interactions.[9]



Scheme 1.1. Structures of nucleotides and nucleobases (B = nucleobase). Purines include A and G and pyrimidines include C, T, and U.



Scheme 1.2. DNA double strand helix.



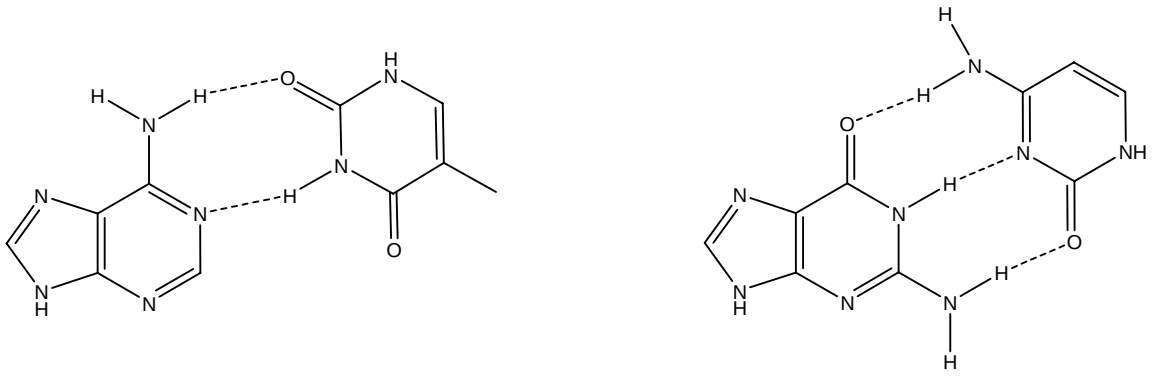
Scheme 1.3. RNA secondary structures: (a) double strand region, (b) hairpin (or loop), (c) base-mismatches, (d) symmetric bulges, (e) a four-way junction.

1.1.1. Non-Covalent Interactions in Nucleic Acid Higher Order Structure

Polynucleotide chains are intrinsically flexible molecules and have the potential to form many different higher-order structures. Their flexibility derives from rotation around bonds in the sugar-phosphate backbone. The chemical properties of nucleic acid components are primary determinants in structure formation, i.e., the driving forces for the formation of higher order nucleic acid structures are the properties of the constituent nucleotides and the non-covalent interactions between them in the form of base pairing and base stacking. In nucleic acids, the highly directional nature of the hydrogen bonding between nucleobase pairs is key to their secondary structure.

Watson and Crick showed that not only is the DNA molecule double-stranded through G-C and A-T base pairing (Scheme 1.4), but these two strands also wrap around each other to form a coil, or helix.[10] The duplex structure is the major existing form of DNA in living organisms and provides a simple mechanism for DNA replication. Similarly, in RNA, the nucleobases form G-C and A-U base pairs. The specific pairing between these complementary nucleobases is called Watson-Crick (W-C) base pairing and hydrogen bonding is the main interaction between them, e.g., each G-C base pair contains three hydrogen bonds and each A-T/U base pair contains two. As a general rule, nucleic acid chains tend to maximize base pairing, resulting in two common secondary structures in vivo, the double helix in DNA (Scheme 1.2), and the hairpin in RNA (Scheme 1.3).[11] However, W-C base pairing is not the only possible motif that nucleobases adopt, e.g., G-G pairing is found in telomeric DNA[12] and G-U pairing is a common 'mispair' in the codon-anticodon interaction in the ribosome.[13] These non-W-C base pairs usually produce characteristic three-dimensional shapes that stabilize certain

nucleic acid structures for proteins to recognize. Other poorly matched base pairs (e.g., A-C and G-T), in which the hydrogen-bonded atoms are imperfectly aligned, can insert extra flexibility in polynucleotide chains.



Scheme 1.4. Hydrogen bonds in Watson-Crick (W-C) base pairs.

In addition, interactions between nucleobases, interactions between nucleic acids and water (hydration) as well as metal cations in aqueous solution are also of critical importance in determining their physical properties and biological function. For example, the B-form DNA double helix (with a wide and shallow major groove) needs about 30%, by weight, water to maintain its native conformation in the crystalline state, and partial dehydration converts it to A-form DNA (with a narrow and deep major groove).[14] Nucleic acids have a number of functional groups that can hydrogen bond to water, with RNA having a greater extent of hydration than DNA due to its extra ribose 2' hydroxyl groups. The negatively charged backbone of nucleic acids is normally electrostatically compensated by cations (counterions, e.g., Na^+, K^+, or Mg^{2+}) condensed on their surface to form a thermally fluctuating sheath commonly referred to as the ion atmosphere.[15]

While monovalent cations mostly bind in a non-specific manner, divalent metal cations are more likely to occupy structurally well-defined binding positions in RNA and DNA architectures.[16] The presence of these counterions can affect nucleic acid structures as well as their ability to bind with drugs because they can screen and shield the negative backbone surface.

The diverse functional activities of biomolecules are attributed to their characteristic secondary and tertiary structures. For example, the 3D conformation of an RNA molecule determines many of its biological properties, including its involvement in gene expression, regulatory processes, mRNA splicing, transport and translation.[17] One of the most important forms of RNA are the ribozymes, which act as molecular scissors that cut RNAs and thereby provide a useful way of regulating gene function by effectively turning off a gene. A number of laboratories around the world are now using ribozymes to study gene function in, e.g., HIV, the AIDS virus, and in cancer research.[18-22] However, ribozyme assays are not straightforward because, inside a cell, there is a large number of RNAs and when a ribozyme is introduced, it is possible to cut several of them and thereby cause a large number of genes to turn off. This specificity problem might cause complicated results and therefore needs to be taken care of, e.g., by using trans-splicing ribozyme[23] or by controlling the salt concentration in solution.[24]

Another important RNA group is ribosomal RNA (rRNA) in the ribosome, which is the protein synthetic factory in the cell and contains two subunits (40S and 60S for eukaryotes and 30S and 50S for prokaryotes). Ribosomes are extremely abundant in cells, e.g., they make up at least 80% of the RNA molecules found in a typical eukaryotic cell. Eukaryotic ribosomes contain four different rRNA molecules: 18S, 5.8S, 28S, and 5S,

while prokaryotic ribosomes contain three: 16S, 5S and 23S. There are also proteins in the ribosome, however, the "active site" of the ribosome, where the chemical reactions occur to pass genetic information onto proteins (translation), only contains rRNA. Increasing evidence indicates that rRNAs play crucial roles in ribosome function.[25, 26]

1.1.2. Non-Covalent Interactions in Nucleic Acid Complexes

The interactions between DNA and drugs are of great importance in molecular biology and medicinal chemistry and they have wide applications in nucleic acid recognition, regulation of biological processes, and the development of therapeutic agents against cancers and virus-related diseases. After the first DNA duplex-netropsin complex structure was solved in 1985,[27] crystal structures of DNA duplexes bound to a series of drugs have been reported including distamycin and synthetic drugs.[28, 29] These studies demonstrated that netropsin and distamycin molecules interact with both DNA strands through hydrogen bonding and van der Waals interactions with sequence specificity for A-T base pairs that are located in the DNA duplex minor groove.[27, 30] They also identified specific molecular features that affect the relative binding affinities of these complexes. This established binding mode opened up avenues for designing new drugs with higher affinity[31] and different specificity.[32]

The ability of RNA to interact with small molecules has long been recognized.[33] Such interactions offer several strategic advantages, e.g., presenting a drug target lacking a known cellular repair mechanism, containing a large repertoire of structural diversity, and having the potential to achieve regulation that would be difficult to obtain by other approaches. Manipulation of RNA interactions has the potential not only to prevent, but also to enhance, gene expression. Antibiotics are a chemically and structurally diverse

collection of molecules that can interact with rRNA to exert profound effects on the translation process. For example, aminoglycoside antibiotics are known to bind to the A-site of the 16S rRNA and inhibit protein synthesis and RNA splicing, resulting in cell death.[34] As the first class of compounds known to bind specifically to subdomains of larger RNA sequences, aminoglycoside antibiotics are useful for understanding the design principles required to produce new classes of therapeutic agents to bind RNAs.[35, 36] Functional insights from structural studies of antibiotics bound to ribosomal subunits have revealed that rRNA-small molecule recognition is based on a combination of shape recognition, electrostatic, and hydrogen-bonding interactions.[37, 38] Additionally, RNA systematic evolution of ligands by exponential enrichment (SELEX) has enabled the selection of minimal nucleic acid recognition motifs for ligand binding, demonstrating that RNA three-dimensional structures can form a large number of highly specific ligand binding sites.[39]

In addition to small molecules, nucleic acids can also bind each other, or bind other macromolecules such as proteins. Formation of such complexes plays important roles in all aspects of genetic activity within an organism, such as transcription, translation, packaging, rearrangement, replication, and repair.[40] One example is the triple helix formed by an oligonucleotide binding specifically to the major groove of a DNA duplex through Hoogsteen or reverse Hoogsteen hydrogen bonds.[41] This binding specificity enables various applications in genome mapping[42] and therapeutic strategies[43] because transcription can be efficiently prevented. Another example is the telomeres, which are protein-DNA structures that make up the ends of linear chromosomes in eukaryotes.[44] Without proper telomere structure, chromosomes become unstable and

cells die. Telomere DNA is composed of tandemly repeated short sequences with one strand rich in G and T that extends past the duplex portion of the telomere to form a single stranded 3' end, which the telomere protein recognizes and binds to form a unique capping complex.

1.2. Traditional Methods for Characterizing Nucleic Acid Structure

1.2.1. Biophysical Methods

In many cases, nucleic acid structures are examined in vitro, under non-physiological conditions, such as after denaturation or chemical synthesis. High resolution techniques include X-ray crystallography and NMR spectroscopy.[45] The former provides a three-dimensional picture of the electron density within a crystal, from which the mean atomic positions, their chemical bonds, their disorder, and various other information can be derived. NMR spectroscopy complements X-ray crystallography by providing information about the dynamic structure of a molecule,[46, 47] however, its application is generally limited by the size of the molecule because large molecules generate highly complex spectra. One weakness shared by these two methods is the requirement for relatively large amounts of time and sample. In addition, although the structure of DNA has been successfully resolved by X-ray crystallography, RNAs frequently show poorly resolved regions in X-ray electron density maps, mainly because a large number of RNA molecules exist inside cells as mixtures of various secondary and tertiary structures that carry out different functions. Thus, it is challenging to isolate RNA species at sufficient quantities in their native forms for successful crystallization and analysis.

Fluorescence resonance energy transfer (FRET) is a technique used to detect conformational changes within molecules, or the interaction between two molecules, by revealing the distance between two chromophores in the molecules.[48] The distance range accessible by FRET is ideal for many biomolecules and fills the gap between other techniques, such as NMR and cryoelectron microscopy.[49] In addition, distance changes during a biopolymer folding process can be monitored in real time by FRET under a wide variety of buffer conditions and therefore provide dynamic structural information. However, this technique frequently requires labeling with fluorophores, which may require complicated sample preparation, and may change the local structure of the molecule.

Circular dichroism (CD) and UV-melting studies are also used to characterize nucleic acid structure and stability, but with relatively low resolution. The former technique is able to determine the form (A, B, or Z) of DNA double helices and to investigate global structural changes, however, the location of the differences between two structures cannot be deduced. The latter technique is useful for measuring the stability of inter- and intra-stand interactions, as well as for calculating thermodynamic parameters for nucleic acid higher order structure. Both of these techniques have some applications in characterizing binding properties of nucleic acid-ligand complexes.

1.2.2. Enzymatic and Chemical Approaches

In addition to the biophysical methods described in the previous section, enzymatic and chemical probing approaches can also be applied to nucleic acid structural analysis. These strategies employ either nucleases or small molecules that cause strand scission at specific residues or structural motifs such that nucleobase or sugar

accessibility can be elucidated and correlated to the folded structure of the molecule.[50] Other information such as structural and chemical differences between nucleic acids containing either unmodified or naturally modified bases at specific locations can also be assessed.[51, 52] Numerous types of DNases and RNases have been isolated and characterized. They differ among properties such as substrate specificity, cofactor requirements, location of cleavage (endonucleases cleave internally and exonucleases chew in from the ends).[9] The most widely used nucleases are DNase I and RNase A, both of which are purified from bovine pancreas: DNase I preferentially cleaves the 5' side of pyrimidine (C or T) residues in either double-stranded or single-stranded DNAs and therefore it is an endonuclease; RNase A is an endoribonuclease that cleaves single-stranded RNA on the 3' side of pyrimidine residues.[9] Transition metals can function as widely used chemical probes when they form complexes with a variety of ligands, e.g., $Fe(EDTA)^{2-}$ is used for studying solvent exposed backbone residues,[53] $Rh(DIP)_3^{3+}$ can probe single stranded, solvent-accessible Gs, pseudouridines (Ψs), and G•U wobble pairs,[54, 55] $Cu(phen)^{+}$ can probe for single-stranded, solvent-accessible regions,[56] and $CoCl_2$ can probe for single-stranded, solvent-accessible Gs.[57]

Enzymatic and chemical probing reactions are generally analyzed by polyacrylamide gel electrophoresis (PAGE), which not only provides information about the length of a polynucleotide chain but also provides some information about its shape, because the speed at which a polynucleotide moves through a gel matrix is related to the degree of compactness of the molecule. After electrophoresis, polynucleotides are detected by 'staining' with ethidium bromide, a dye that intercalates between stacked bases and, in doing so, becomes fluorescent when exposed to UV irradiation. This

method is more widely used for detection of DNA than for RNA because there are two important differences in the physical characteristics of their electrophoresis: the first is that RNA must be pretreated to disrupt any internal base pairing, i.e., secondary structure needs to be effectively destroyed. This pretreatment is necessary because the extensive base pairing in RNA molecules means that they have very diverse conformations, which can affect the relative mobility of the molecules. RNA samples are therefore pretreated by heating or by addition of agents such as formamide, which disrupts hydrogen bonds and denatures the RNA. The second important difference between electrophoresis of DNA and that of RNA is that the latter must be performed under conditions that buffer against alkalinity, since RNA is vulnerable to hydrolysis at alkaline conditions. Although PAGE is generally useful, there are some disadvantages associated with this technique. The first is that, for visualization purposes, the DNA fragment is generally labeled with a radioactive, fluorescent, or biotin label, because standard ethidium bromide staining lacks the sensitivity to detect small amounts of DNA (detection limits are 15 ng for single stranded DNA and 0.25 μg for double stranded DNA[58]) and this labeling may cause structural changes, or pose a problem with waste disposal. The second disadvantage is the inherent limited resolution of PAGE. A third one is that low molecular weight species will often run off the gel in the time it takes to separate larger molecular weight species, causing loss of information. Finally, due to the inherent nature of the probing experiments, only those binding sites that lead to cleavage can be observed.

1.3. Other Solution-Phase Methods for Characterizing Nucleic Acid Complexes

There are several additional solution-phase methods for characterizing binding properties of nucleic acid complexes, such as radioimmunoassay,[59] surface plasmon resonance (SPR) assay,[60] and gel mobility shift assay.[61, 62]

Radioimmunoassay is a highly sensitive and specific method that uses the competition between radiolabeled and unlabeled substances for a target, such as in the antigen-antibody reaction, to determine the concentration of the unlabeled substance. For a limited amount of target, adding unlabeled substance will have an inhibitory effect on the binding of the radiolabeled substance, whose concentration can be monitored and the increasing concentration of free radiolabeled substance can then be correlated to the concentration of the unlabeled substance that is bound to the target. The radioimmunoassay principle has found wide application in the measurement of a large and diverse group of substances, such as human serum albumin,[63] insulin,[64] and thyrotropin,[65] that are of clinical and biological interest.

Surface plasmons are surface electromagnetic waves that propagate parallel to a metal/dielectric (or metal/vacuum) interface. When total internal reflection occurs at an interface between two transparent media of different refractive index, the electromagnetic field component of the incident light penetrates a short (tens of nanometers) distance into the medium of lower refractive index, creating an evanescent wave. If the interface is coated with a thin layer of metal (gold), resonance energy transfer can happen between the surface plasmons and the evanescent wave, which causes a decrease of the reflected light intensity at a specific incidence angle, producing a sharp shadow (SPR signal). This

assay can be used to determine the binding constant (in a chemical reaction, it is the association rate divided by the dissociation rate) because the resonance conditions are influenced by the material adsorbed onto the thin metal film. Satisfactory linear relationship is found between resonance energy and mass concentration of biomolecules such as proteins, sugars, and DNA. The SPR signal is therefore a measure of mass concentration at the surface, which means that the analyte and ligand association and dissociation can be observed and ultimately their rate constants as well as equilibrium constants can be calculated.

Typically, radioimmunoassay and SPR assays measure only the equilibrium concentration of either the free ligand or the complex and provide little information about the binding stoichiometry.[66] Gel mobility shift assay is a common technique used to characterize protein-DNA/RNA interactions and sometimes stoichiometry can be obtained. This method generally involves electrophoretic separation of such complex mixtures on a polyacrylamide or agarose gel. For best results, four gel lanes are compared: the first one contains only the nucleic acid and the second one contains both the protein and the nucleic acid. If binding occurs, there will be additional bands compared to the first lane at shorter moving distances ("shifted" up), which corresponds to the larger and less mobile complexes. From the ratio of the amount of bound to unbound nucleic acid, the binding affinity of the protein may be estimated. The third lane contains a competitor oligonucleotide and the same protein as in the second lane to determine the most favorable binding sequence for the protein. The use of different oligonucleotides of defined sequences allows the identification of the precise binding site. The fourth lane contains an additional antibody compared to the second lane, which

recognizes the protein and creates an even larger complex with larger "shift". This method is referred to as a supershift assay, and is used to unambiguously identify a protein present in a protein-nucleic acid complex. One problem of the gel mobility shift assay is that the determined binding stoichiometry is not always correct (mass of the complex does not correspond to the determined stoichiometry) and mass spectrometry may be needed to obtain a true result.[67-69]

1.4. Mass Spectrometric Approaches

Mass spectrometry (MS) can directly determine the binding stoichiometry of a complex by measuring its mass as well as elucidating a variety of information, such as dissociation constants, discrimination of binding amongst multiple ligands, and identification of ligand binding sites. DNA- and RNA-small molecule complexes are of considerable interest because they lend insight into the mechanism of action of many therapeutic agents that can be used as structural probes,[70, 71] synthetic restriction enzymes,[72] and DNA repair agents.[73]

Gale *et al*.[74] have applied MS to characterize complexes of duplex DNAs with distamycin A, a minor groove binder and showed that low concentrations of ligand produced a 1:1 complex, whereas high ligand concentrations produced a 1:2 complex, consistent with NMR studies. Gabelica *et al*.[75, 76] studied duplex DNAs with both intercalators and minor groove binders and found that the minor groove binders appear to interact with the DNA in a specific manner and with discrete stoichiometries, whereas the intercalators exhibited a range of stoichiometries. These reports generated much interest for follow-up studies, where the interactions of many types of minor groove binders with

DNAs of various sizes were investigated.[77, 78]

The study of RNA-small molecule interactions has received attention over the years. Many of these studies involve the investigation of aminoglycoside binding to the A-site of rRNA. Hofstadler *et al*.[79] demonstrated that binding of five aminoglycosides to two rRNAs could be measured in parallel using MS. Griffey *et al.*[80] showed the possibility to estimate the binding affinities of these different complexes simultaneously by comparing their relative peak abundances observed in the mass spectrum because the binding of the ligands to the rRNA target does not appear to alter the ionization efficiency of the complex.[81-83] They also varied the concentration of aminoglycosides to generate different complex/free RNA abundance ratios to calculate dissociation constants (K_d) for different complexes. Sannes-Lowery *et al.*[84] have examined the effect of solution conditions such as organic solvent and buffer on K_d values measured by MS and found that experiments should be carried out at as high as possible salt concentration to minimize nonspecific binding while organic solvent does not interfere with ligand binding to RNA. They also compared two methods for estimating the K_d value: maintaining the concentration of the ligand fixed, and maintaining the concentration of RNA fixed. The calculated K_d values differed two-fold between these two methods as did K_d values determined at high and low concentrations of RNA.

With the development of soft ionization methods, e. g., matrix assisted laser desorption ionization (MALDI)[85, 86] and electrospray ionization (ESI),[87, 88] which are able to preserve solution-phase structure when molecules transfer into the gas phase, the use of MS for structural analysis of biological molecules has become increasingly routine. ESI MS, an atmospheric pressure interface which can be easily coupled to liquid

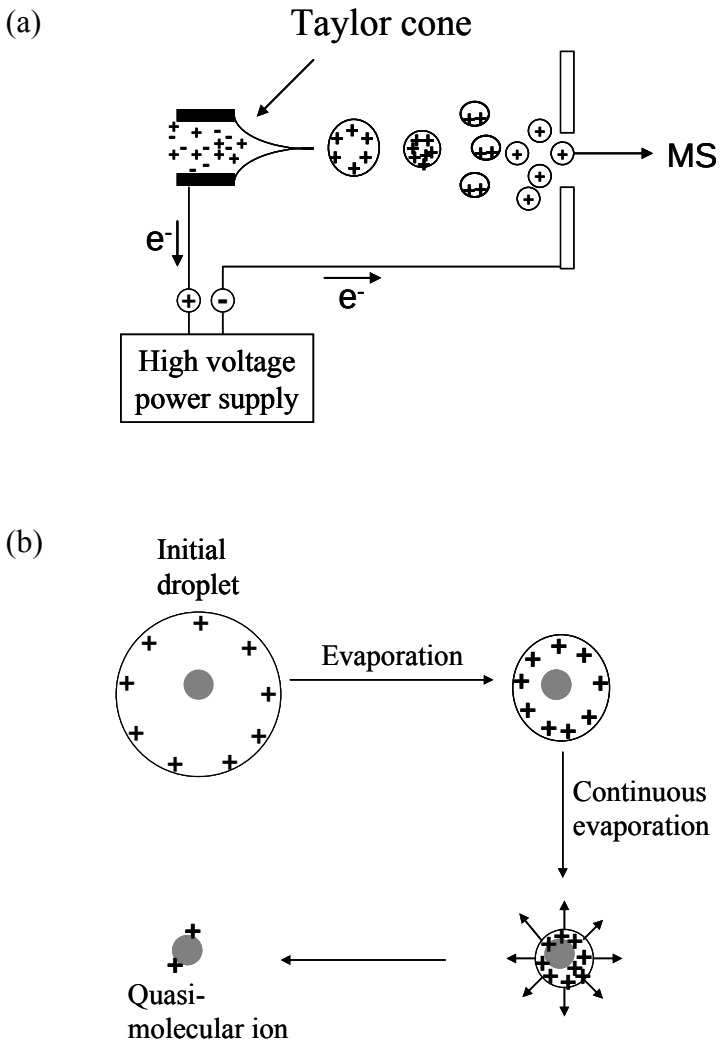
chromatography and auto-sampling methods, is becoming an important technique for various applications such as drug discovery and analysis. ESI is a "soft" ionization process that can transfer intact biomolecular ions from large and complex species containing non-covalent interactions from solution phase into the gas phase where they can be characterized by mass spectrometric analysis.[87-91] This property of ESI is particularly useful for MS analysis of nucleic acids because their higher order structures are mainly formed by hydrogen bonds, which are largely enhanced in the gas phase due to the removal of polar solvent. ESI has been shown to successfully ionize oligonucleotides as large as 8,000 base pairs, corresponding to a molecular mass of 5 M Da.[92]

The major reason for ESI being a “soft” ionization is that there is no “sudden” increase in energy. During ESI, an analyte solution is infused through a needle to which a high electric field is applied. This high electric field causes partial separation of positive and negative electrolyte ions in the solution, for example, in the positive ion mode, positive ions are enriched at the surface of the liquid emerging from the needle tip whereas negative ions are driven towards the inside of the needle. The repulsion between positive ions at the surface and the pull of the electric field (due to the large drop in potential at the needle exit) on the positive ions overcome the surface tension of the liquid and expand the liquid into a cone (Taylor cone) that emits a spray of small positively charged droplets (i.e., nebulization, Scheme 1.5 a and b, adopted from[93]). Both organic solvent (which decreases surface tension) and nebulizing gas[94] can be added to assist nebulization. During this process, analytes are solvated inside droplets without being imparted with high energy. To be analyzed by mass spectrometry, the charged droplets need to evolve into free ions by evaporating solvent. This process is achieved by

both a pressure drop inside the mass analyzer, and addition of hot drying gas. Solvent evaporation reduces the volume of the droplets at constant charge, i.e., increases droplet charge density, and leads to Coulomb fission of the droplets (Scheme 1.5c, adopted from Fenn *et al*.[87]). The actual process of ion formation is debated,[88, 93, 95, 96] but the final result is that the solvent in the charged droplets is completely removed and analyte ions are formed with either single or multiple charges thereby being ready for MS analysis. A major advantage of ESI is that it produces multiply charged ions. The importance of multiple charging was demonstrated by Fenn *et al*.[87] for proteins with multiple charges up to 64+ and molecular weights of 76 kDa. Multiple charging allows high mass ions to be analyzed with mass analyzers of limited mass-to-charge (m/z) range. Another advantage of forming multiply charged ions from large molecules is that it facilitates tandem mass spectrometric analysis.[97] The number of charges on analyte ions varies, depending on several parameters such as the presence of acidic/basic groups in the analyte, solvent, pH, and temperature. For negative ion analysis of nucleic acids and their complexes, charges normally reside on backbone phosphate groups due to their high acidities.[98, 99] One of the key factors of sample preparation for successful ESI MS analysis is the removal of all non-volatile cations (i.e. Na^+, K^+, Mg^{2+}) because the presence of these ions can lead to signal suppression and more complex spectra.

DNA duplexes were the first reported nucleic acid non-covalent complexes observed by ESI MS.[100, 101] Ding *et al.*[102] reported that the relative abundance of DNA duplex ions in ESI mass spectra qualitatively reflected the abundance of these duplexes in solution, e.g., the observed abundance of non-complementary DNA dimers was significantly lower than that of a complementary duplex. Ganem *et al*.[101] reported that

the relative abundance of duplex versus single-stranded DNA was higher for a G-C rich 8-mer as compared to an A-T rich 8-mer. This result is consistent with solution-phase observations and theoretical predictions that G-C base pairs are more stable than A-T base pairs.[9] Similar results have been obtained by our group (Figure 1.1). In addition, ESI is also able to preserve non-covalent interactions in other systems, e.g., complexes of small drug molecules bound to single stranded,[103] hairpin,[104] or duplex DNA,[105] as well as complexes between proteins and oligonucleotides.[82, 106] The binding properties and stoichiometry of these non-covalent complexes characterized by ESI MS are in good agreement with those derived from more conventional solution-phase techniques.[90]



Scheme 1.5. Description of the ESI process. Due to the high electric field applied at the ESI needle, a Taylor cone forms at the end of the needle and generates charged droplets (a). Before and during entrance into the mass analyzer, solvent in the droplets evaporates with the assistance of drying gas, which reduces droplet size and causes a build up of charge density on the droplet surfaces. The droplets will break up when Columbic forces overcome the surface tension. This process repeats until complete desolvation and quasi-molecular ions are generated. (a) is adopted from Kebarle *et al.* and (b) is adopted from Fenn *et al*.

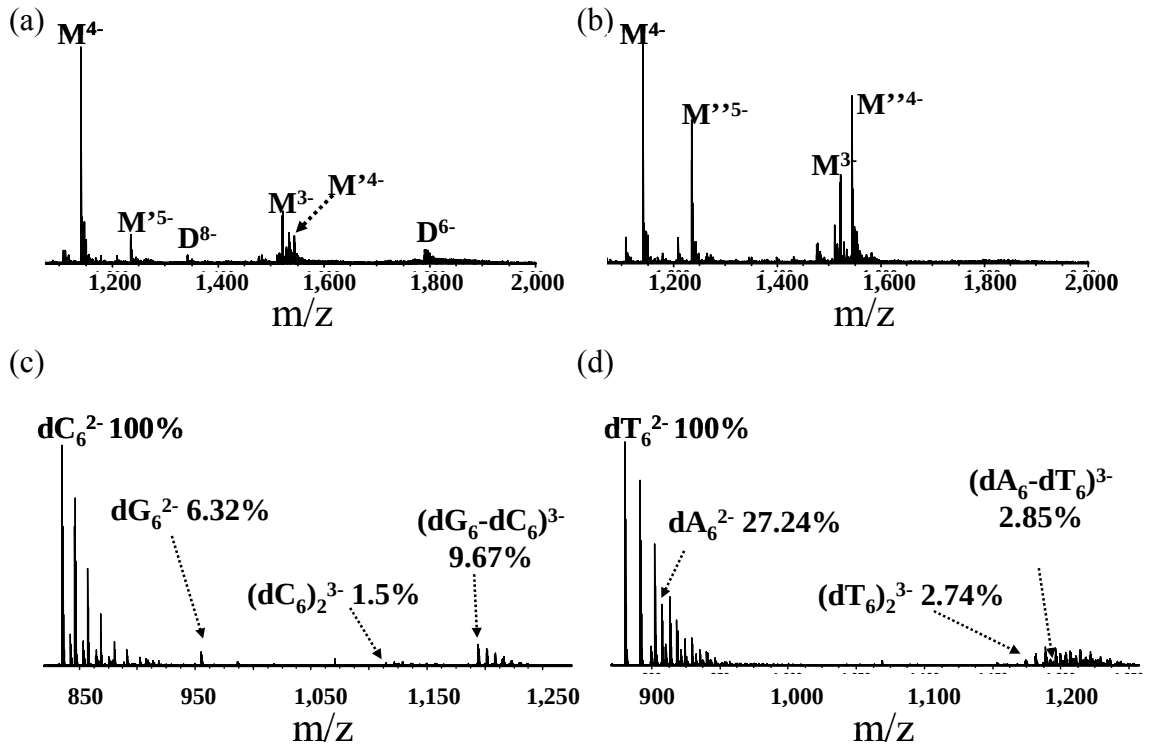


Figure 1.1. Comparison of DNA duplex formed between complementary strands (a) and non-complementary strands (b); between relatively strong duplex (dG_6-dC_6) (c) and relatively weak duplex (dA_6-dT_6) (d). M = d(GGTACCTATCAGTGATAGAG); M' = d(CTATCACTGATAGGT); M'' = d(TCTAACCTGATGATG); D = duplex. Duplexes were prepared by mixing two DNAs at the same concentration and annealing at 95 ^{o}C for 10 minutes. The concentrations of M, M' and M'' in (a) and (b) are the same as are the concentrations of dG_6 and dC_6 in (c) and the concentrations of dA_6 and dT_6 in (d). Duplexes were only observed between complementary strands M and M', dG_6 and dC_6, dA_6 and dT_6. No duplex was observed between non-complementary strand M and M''. Duplex dG_6-dC_6 is more efficiently formed than duplex dA_6-dT_6 because it has three hydrogen bonds within each G-C base pair while A-T base pairs only have two hydrogen bonds.

Only a small fraction of the ions generated by ESI at atmospheric pressure are subsequently transmitted through the mass spectrometer and ultimately detected. From ion mobility studies, it is known that more than 50% of the ESI current, generally ~ 50 to 200 nA, is potentially useful ion current.[95] However, the ion current entering the mass

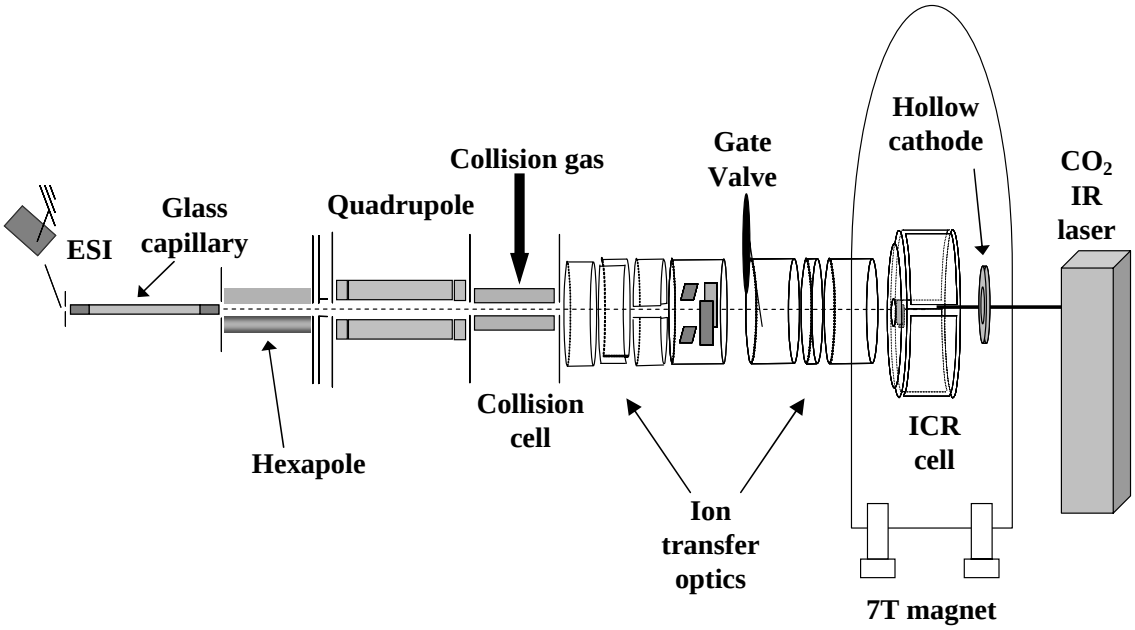
analyzer is typically three orders of magnitude or more smaller.[107] Most ion loss occurs when ions transfer from a high pressure region to a low pressure region, typically at the interface between the capillary exit and further transmission regions (e.g., skimmers) into the mass analyzer. Improvement of ion transmission efficiency in the first interface has been achieved by using a muticapillary inlet,[108, 109] while in the second interface it has been achieved by using ion funnels.[107, 110-112] Our group has installed an Apollo II ion funnel source (Bruker Daltonics, Billerica, MA) to improve the transmission efficiency in the second interface for nucleic acids with higher order structures and non-covalent complexes between nucleic acids and small molecules. Higher sensitivity, and softer transmission conditions were obtained with the ion funnel. More details are discussed in the Appendix.

1.5. Fourier Transform Ion Cyclotron Resonance Mass Spectrometry (FT-ICR MS)

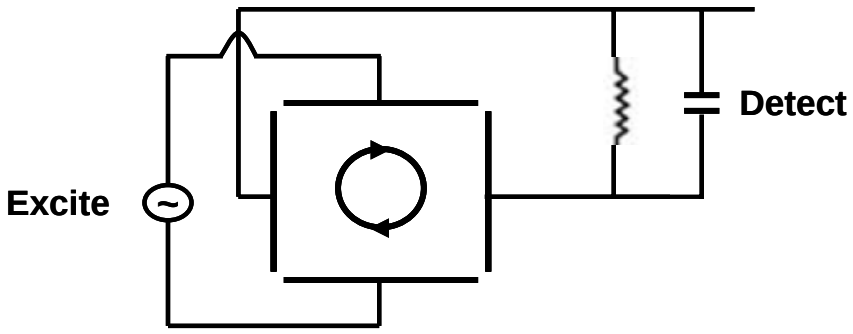
1.5.1. FT-ICR Mass Spectrometer

FT-ICR MS was first introduced by Comisarow and Marshall in 1974[113, 114] and its first coupling with an ESI source was reported in 1989 by Henry *et al.*[115] The major components of our FT-ICR instrument (shown in Scheme 1.6) are the ion source, the external quadrupole (Q) for precursor ion selection, the ion transfer optics, and the ICR cell that is located in the center of the magnetic field. Additional components, including an IR laser and an indirectly heated hollow cathode, serve to perform tandem MS. Ions are transferred from the ESI source and trapped in the ICR cell radially by the Lorenz force and axially by electrostatic trapping electrodes. Following trapping, ions are

excited into coherent motion (ions with same m/z ratio form an ion cloud) by an RF voltage, which is applied differentially between two opposite ICR cell plates and contains a band of frequencies covering a range of m/z values. Another pair of plates detects image current induced by ion clouds when they orbit close to the plates (Scheme 1.7). The resulting time domain signal is converted to the frequency domain by Fourier transformation, and ion cyclotron frequency (ω) is further converted to *m*/*z* values by using the cyclotron equation: $\omega = zeB/m$ in which z is the number of elementary charges, e is the elementary charge, and B is the magnetic field strength.[116] Because ions with different m/z values are separated only by their cyclotron frequency and the frequency measurement is inherently accurate, unparalleled mass accuracy (sub-ppm) can be obtained.[117] FT-ICR MS also offers higher resolving power than other mass analyzers such as time-of-flight or sector instruments. One reason is that the ion motion can last for a long time (several seconds) in the cell due to the ultrahigh vacuum environment, resulting in a large number of cycles for detection, which is proportional to the resolving power.[118] Another reason is that the cyclotron frequency is independent of ion velocity, therefore the performance is not limited by the spread in the ion initial position, direction, and speed.[119] Resolving power exceeding 10^6 can be obtained for large proteins. For example, Kelleher *et al.*[120] have detected a 112 kDa protein with unit mass resolution. One advantage of isotopic resolution is that it allows direct determination of the charge state of an ion from its isotopic distribution.[121] The high resolving power also makes FT-ICR analysis feasible for highly complex mixtures (e.g., large protein digests) without the need for chromatographic separation.[122]



Scheme 1.6. Configuration of our 7 T Q-FT-ICR mass spectrometer.



Scheme 1.7. Excitation and detection of an ion trapped inside the ICR cell. An RF voltage waveform containing the resonance frequency of the ion is applied differentially between one pair of cell plates. The other pair of plates detects an image current of the orbiting ion cloud.

One existing technical limit of FT-ICR MS is the requirement of about 100 charges of a given m/z ratio to induce a detectable image current. In addition, ESI is a

concentration-sensitive technique[93] and limited in analyzing dilute samples. One solution is to couple ESI FT-ICR MS with capillary electrophoresis (CE), where subattomole detection limits have been achieved.[123] However, CE also requires the analyte solution to be relatively concentrated. Nano-LC has been shown to constitute an attractive alternative for analyte solutions too dilute for CE analysis.[124] The coupling of nano-LC to FT-ICR MS also helps the on-line desalting of samples. Although FT-ICR MS does not present the lowest detection limit of all the mass analyzers, the unique combination of accurate mass measurement, ultrahigh resolution, and multistage tandem MS makes it extremely attractive for biological analysis.

1.5.2. Tandem Mass Spectrometry

The coupling of an ESI source to high resolution FT-ICR mass spectrometers allows for the development of several different tandem MS experiments. Tandem MS involves as least two stages of mass analysis[125] and in our instrument configuration, the first stage is normally the quadrupole isolation of precursor ions and the second stage is the detection of fragment ions produced by different fragmentation techniques. For even more tandem MS stages, the fragments from the second stage, instead of being detected, are subjected to further fragmentation. Most tandem MS fragmentation strategies can be divided into two categories, vibrational excitation mode and ion-electron interaction mode. The former category includes collision-activated dissociation (CAD),[126] sustained off-resonance irradiation (SORI) CAD,[127] infrared multiphoton dissociation (IRMPD),[128] blackbody infrared dissociation (BIRD),[129] and nozzle-skimmer dissociation,[130] and the latter involves radical chemistry and mainly includes electron capture dissociation

(ECD),[131] electron detachment dissociation (EDD),[132] and electron induced dissociation (EID).[133, 134]

Vibrational excitation tandem MS is a well-established technique for oligonucleotide sequencing and a maximum length of 108 nucleotides has been reported by using IRMPD.[135] The main advantages of applying tandem MS for sequence analysis compared to other traditional methods are the speed of analysis, sensitivity, high resolution, high mass accuracy and no need for enzyme.[119] The introduction of internal calibrants results in even higher mass accuracy, making it possible to directly distinguish very small mass differences between fragments.[136] However, there are still disadvantages associated with these vibrational excitation fragmentation methods: first, secondary fragmentation such as water and nucleobase loss complicates spectral interpretation and reduces sensitivity. Second, inherent lowest-energy fragmentation pathways result in non-random dissociation and possibly incomplete sequence coverage, e.g., cleavage on the 3' side of T is limited in IRMPD.[137] Ion-electron interaction tandem MS mode, on the other hand, has been proposed to be a non-ergodic fragmentation technique[131] which precludes site-specific dissociation and therefore provides complementary information to the vibrational excitation mode. Another benefit from this pathway is that neutral nucleobase loss is extensively limited.[138] One limitation of the ion-electron interaction mode is the low fragmentation efficiency, i.e., only a small fraction (5~20%) of the precursor ions can be converted into product ions.[119] For nucleic acids, EDD may provide higher efficiency than ECD because the acidic sugar-phosphate backbone is easily deprotonated to form anions. Another characteristic of ion-electron reactions is that they present "soft" fragmentation pathways, e.g., the backbone of peptides and

proteins can be cleaved without losing labile posttranslational modifications[139-142], thereby allowing their localization. This feature has also been applied in the investigation of protein higher order structure[143] as well as in kinetic studies of gas-phase protein folding and unfolding,[144, 145] because backbone covalent bonds can be ruptured without breaking non-covalent interactions. The application of ECD and EDD for characterization of nucleic acid higher order structure is part of this thesis work and will be discussed in Chapter 4.

In addition to the ability to sequence single stranded oligonucleotides, tandem MS can also be applied to probe complex structures, e.g., to sequence DNA duplexes[146] and to localize the binding sites of RNA-ligand complexes.[80] McLafferty *et al.*[147] have applied IRMPD to completely sequence a 64 base-pair DNA duplex by accurately assigning the fragments produced from cleavages at the termini of the respective single strands as well as from internal sites. Griffey *et al.*[148] showed that for RNA-DNA duplexes, more extensive fragmentation occurs for duplexes possessing mismatches compared to those with solely W-C base pairs and most of the cleavage sites are centered on the mismatch site and surrounding residues of the DNA strand. These results led them to use tandem MS (CAD) to interrogate the binding site of a 27-mer RNA bound to different aminoglycoside antibiotics.[149] They modified the A-site (an A-A mismatch and a bulge A) of the RNA to a chimeric 2'-*O*-methylribonucleotide model, where the three As were replaced with dAs. Without the bound aminoglycoside antibiotics, extensive fragmentation at the modified A-site was observed; however, upon binding of the ligands, the A-site was protected, indicating that it is the binding site, which is consistent with results from NMR and X-ray studies.[150, 151] These authors also tested the model RNA

against a 216 member combinatorial library of ligands and found that some of them could bind to the model RNA although they did not protect the A-site from fragmentation, indicating that these ligands bind at some other locations away from the A-site. One reason for successfully assigning the binding sites of these RNA-aminoglycoside antibiotics complexes is that they have relatively high binding affinities so that the ligands do not fall off in CAD. However, for weakly bound complexes, vibrational excitation tandem MS may not be suitable for localization of the binding sites because the loss of ligands would be preferential. ECD and EDD could constitue alternative methods due to their "soft" characteristic.[152, 153]

1.6. Gas-Phase Hydrogen/Deuterium Exchange

The three-dimensional shape of large molecules is of considerable interest to biologists. Hydrogen/deuterium exchange (HDX) is used as a probe of surface accessibility of molecules and ESI FT ICR MS is an ideal means for monitoring such reactions in solution[154] and in the gas phase.[155] A comparison of gas- and solution-phase conformations can provide insight into the nature of the ionization process, particularly the role of solvent. For nucleic acids, gas-phase HDX is advantageous over solution-phase HDX because their base pairs open and close too fast in solution, typically within several milliseconds. MS is especially well suited for gas-phase HDX as deuterium incorporation can be monitored directly from mass spectral peak shifts with reaction time, which is easily controlled by data acquisition programs. Initial gas-phase HDX experiments were performed inside an ICR cell, which provides extended trapping periods and allows for kinetic studies.[156, 157] Such experiments are facilitated by high

magnetic field due to several factors:[158] first, improved isolation of monoisotopic species leading to simpler data analysis; second, longer ion trapping periods leading to extended range of accessible rate constants; third, higher dynamic range permitting simultaneous analysis of different species; fourth, enhanced reproducibility; and finally the decrease in required data acquisition period. Enhanced gas-phase HDX performance has been achieved by performing the reaction in the external hexapole ion reservoir of an FT ICR mass spectrometer.[159] This experimental setup results in more than 100 times faster HDX compared to in-cell exchange, which requires lower gas pressures and additional pump-down periods. The short experimental times facilitate the quantitation of the number of labile hydrogens for less reactive species. There are several candidates of deuterated reagents, such as D_2O, D_2S and ND_3, among which D_2S is most suitable for nucleic acids because its p*K*a value is closer to that of the nucleic acid backbone phosphate groups compared with the other two reagents.[157, 160-163] Furthermore, D_2S can very easily be introduced into a mass spectrometer because it is a gas at room temperature. There are several studies attempting to elucidate the nucleic acid gas-phase HDX mechanism, but still no ultimate agreement has been reached. For example, it has been reported that, for mono-nucleotides exchanging with D_2O, cooperative interaction between a nucleobase and a phosphate group is required for exchange to happen,[156] which is dependent on the identity and orientation of the nucleobase as well as the position and flexibility of the phosphate moiety.[164] However, when exchanging with D_2S, only the phosphate group and the terminal hydroxyl group but not the nucleobase or the 2' hydroxyl group have been proposed to be involved.[160] Studies on gas-phase HDX of polynucleotide chains suggest that the reaction kinetics is dependent on factors such as charge state, chain

length, properties of the exchange reactants, and the salt content in the buffer.[165] The enhanced reactivity observed for polynucleotides compared to mononucleotides suggests that the flexibility of polyphosphate chains plays a critical role in forming accessible conformations for remote H(D)-donor sites.

Although the mechanism is not yet clear, gas-phase HDX of nucleic acids has been quite extensively applied. Kamel *et al.*[166] have applied gas-phase HDX to verify the CAD fragmentation pathways for both protonated and deprotonated pyrimidine bases, which are the principal fragment ions of pyrimidine antiviral agents. Their results also determined the specific influence of substitution in the pyrimidine ring on the site selectivity of the fragmentation pathways. Griffey *et al.*[167] have used gas-phase HDX to probe conformations of two phosphorothioate oligonucleotides, which are used as antisense therapeutics. Results showed that each molecule adopts more than one gas-phase conformation, which is represented by populations with different HDX rates. The slow interconversion between these populations is consistent with a model where initial ion conformations containing internal hydrogen bonds between the bases and charged residues are disrupted over time. Gabelica *et al.*[168] found that the $[(TGGGGT)^4 + 3NH_4^+]$ quadruplex, which is known to be a very rigid structure in the gas phase by molecular dynamic simulations, underwent very fast exchange in both positive and negative ion modes compared to the corresponding single-stranded TGGGGT. Although the hydrogens of the ammonium ions are not exchanged, the presence of inner cations is still essential for the fast exchange to be possible. Therefore, they suggested that the concept that compact DNA structures exchange H for D slower than unfolded ones is a misconception.

1.7. Dissertation Overview

This dissertation research is focused on mass spectrometric methodology development for characterization of nucleic acid non-covalent interactions. Two strategies have been utilized: tandem MS (up to MS^3) and gas-phase HDX combined with MS. An introduction to the first strategy is provided earlier in this Chapter while the mechanism of the second strategy is discussed in detail in Chapters 2 and 3 for negative and positive ion mode, respectively. Chapter 4 describes the application of the first strategy, including EDD, IRMPD, activated ion EDD (AI EDD) and EDD/IRMPD MS^3 tandem MS techniques, for distinguishing three isomeric DNA 15-mers. Chapter 5 describes the application of the second strategy for correlating the solution-phase stabilities of hairpin structured nucleic acids with their gas-phase HDX rates with D_2S. Chapter 6 demonstrates another potential application of the second strategy in comparing complex binding affinities by their HDX rates. The last chapter summarizes all the results obtained during this thesis work and also suggests some future directions. Due to the lability of non-covalent interactions, the ionization and ion transfer conditions of the mass spectrometer need to be optimized to be as 'soft' as possible. Such optimization was crucial for the work described in Chapters 4-6. Details of this process are discussed in the Appendix.

1.8. Bibliography

(1) Cech, T. R. *J. Am. Med. Assoc.* **1988**, *260*, 3030-3034.

(2) Zaug, J.; Been, M. D.; Cech, T. R. *Nature* **1986**, *324*, 429-433.

(3) Nissen, P.; Hansen, J.; Ban, N.; Moore, P. B.; Steitz, T. A. *Science* **2000**, *289*, 920-930.

(4) Valadkhan, S.; Manley, J. L. *Nature* **2001**, *413*, 701-707.

(5) Lehn, J. M. *Science* **1993**, *260*, 1762-1763.

(6) Lehn, J. M. *Supramolecular Chemistry*; Wiley-VCH, **1995**.

(7) Müller-Dethlefs, K.; Hobza, P. *Chem. Rev.* **2000**, *100*, 143-167.

(8) Schneider, H.-J.; Yatsimirsky, A. *Principles and methods in supramolecular chemistry*; John Wiley & Sons Ltd.: Chichester, England, **2000**.

(9) Voet, D.; Voet, J. G. *Biochemistry*; John Wiley & Sons, Inc.: New York, **1994**.

(10) Watson, J. D.; CRICK, F. H. C. *Nature* **1953**, *171*, 737-738.

(11) Tuerk, C.; Gauss, P.; Thermes, C.; Groebe, D.; Gayles, M.; Guild , N.; Stormo, G.; d'Aubenton-Carafa, Y.; Uhlenbeck, O.; Tinoco, I.; Brody, E.; Gold, L. *Proc. Nat. Acad. Sci. U.S.A.* **1988**, *85*, 1364-1368.

(12) Krafft, C.; Benevides, J. M.; Thomas, J., G. J. *Nucleic Acids Res.* **2002**, *30*, 3981-3991.

(13) Gautheret, D.; Konings, D.; Gutell, R. R. *RNA* **1995**, *1*, 807-814.

(14) Brovchenko, I.; Krukau, A.; Oleinikova, A.; Mazur, A. *Water Percolation Governs Polymorphic Transition and Conductivity of DNA, from Computational Biophysics to Systems Biology*; John von Neumann Institute for Computing, **2007**, 3.

(15) Manning, G. S. *Biophys. Chem.* **1977**, *7*, 95-102.

(16) Šponer, J. E.; Burda, J. V.; Leszczynki, J.; Šponer, J. *Computational Studies of RNA and DNA*; Springer Netherlands, **2006**, 22.

(17) Noller, H. F. *Science* **2005**, *309*, 1508-1514.

(18) Smythe, J. A.; Symonds, G. *Inflammation Res.* **1995**, *44*, 11-15.

(19) Qureshi, A.; Zheng, R.; Parlett, T.; Shi, X.; Balaraman, P.; Cheloufi, S.; Murphy, B.; Guntermann, C.; Eagles, P. *Biochem. J,* **2006**, *394*, 511-518.

(20) Kashani-Sabet, M.; Liu, Y.; Fong, S.; Desprez, P.-Y.; Liu, S.; Tu, G.; Nosrati, M.; Handumrongkul, C.; Liggitt, D.; Thor, A. D.; Debs, R. J. *Proc. Natl Acad. Sci. U.S.A.* **2002**, *99*, 3878-3883.

(21) Choi, K. S.; Lee, T. H.; Jung, M. H. *Cancer Gene Therapy* **2003**, *10*, 87-95.

(22) Norris, J. S.; Hoel, B.; Voeks, D.; Maggouta, F.; Dahm, M.; Pan, W.; Clawson, G. *Cancer Gene Therapy* Springer Netherlands, **2006**,

(23) Ayre, B. G.; Köhler, U.; Goodman, H. M.; Haseloff, J. *Proc. Natl Acad. Sci. U.S.A.* **1999**, *96*, 3507-3512.

(24) Ando, T.; Tanaka, T.; Kikuchi, Y. *J. Biochem.* **2003**, *133*, 445-451.

(25) Dahlberg, A. E. *Cell* **1989**, *57*, 525-529.

(26) Noller, H. F. *Annu. Rev. Biochem.* **1991**, *60*, 191-227.

(27) Kopka, M. L.; Yoon, C.; Goodsell, D.; Pjura, P.; Dickerson, R. E. *Proc. Natl Acad. Sci. U.S.A.* **1985**, *82*, 1376-1380.

(28) Wang, A. H.-J.; Teng, M. *Crystallographic and Modeling Methods in Molecular Design*; Springer-Verlag: New York, **1990**, 123-150.

(29) Kopka, M. L.; Larsen, T. A. *Nucleic Acid Targeted Drug Design*; Marcel Dekker: New York, **1992**, 303-374.

(30) Coll, M.; Frederick, C. A.; Wang, A. H.-J.; Rich, A. *Proc. Natl Acad. Sci. U.S.A.* **1987**, *84*, 8385-8389.

(31) Mrksich, M.; Dervan, P. B. *J. Am. Chem. Soc.* **1993**, *115*, 9892-9897.

(32) Geierstanger, B. H., Mrksich, M., Dervan, P. B. & Wemmer, D. E. *Nature Struct. Biol.* **1996**, *3*, 321-324.

(33) Westhof, E.; Masquida, B.; Jaeger, L. *Folding Des.* **1996**, *1*, R78-R89.

(34) Moazed, D.; Noller, H. F. *Nature* **1987**, *327*, 389-394.

(35) Park, W. K. C.; Auer, M.; Jaksche, H.; Wong, C. *J. Am. Chem. Soc.* **1996**, *118*, 10150-10155.

(36) Wong, C.-H.; Hendrix, M.; Manning, D. D.; Rosenbohm, C.; Greenberg, W. A. *J. Am. Chem. Soc.* **1998**, *120*, 8319-8327.

(37) Brodersen, D. E.; Clemons Jr., W. M.; Carter, A. P.; Morgan-Warren, R. J.; Wimberly, B. T.; Ramakrishnan, V. *Cell* **2000**, *103*, 1143-1154.

(38) Carter, A. P.; Clemons, W. M.; Brodersen, D. E.; Morgan-Warren, R. J.; Wimberly, B. T.; Ramakrishnan, V. *Nature* **2000**, *407*, 340-348.

(39) Gold, L.; B., P.; Uhlenbeck, O.; Yarus, M. *Annu. Rev. Biochem.* **1995**, *64*, 763-797.

(40) Luscombe, N. M.; Austin, S. E.; Berman, H. M.; Thornton, J. M. *Genome Biol.* **2000**, *1*, 1-37.

(41) Sun, J. S.; Hélène, C. *Curr. Opin. Struct. Biol.* **1993**, *3*, 345-356.

(42) Strobel, S. A.; Doucette-Stamm, L. A.; Riba, L.; Housman, D. E.; Dervan, P. B. *Science* **1991**, *254*, 1639-1642

(43) Cooney, M.; Czernuszewicz, G.; Postel, E. H.; Flint, S. J.; Hogan, M. E. *Science* **1988**, *241*, 456-459.

(44) Horvath, M. P.; Schweiker, V. L.; Bevilacqua, J. M.; Ruggles, J. A.; Schultz, S. C. *Cell* **1998**, *95*, 963-974.

(45) Saenger, W. *Principles of Nucleic Acid Structure*; Springer-Verlag: New York, **1984**, 116-158.

(46) Al-Hashimi, H. M. *ChemBioChem* **2005**, *6*, 1472-1473.

(47) Latham, M. P.; Brown, D. J.; McCallum, S. A.; Pardi, A. *ChemBioChem* **2005**, *6*, 1492-1505.

(48) Stryer, L.; Haugland, R. P. *Proc. Natl Acad. Sci. U.S.A.* **1967**, *58*, 719-726.

(49) Walter, N. G.; Harris, D. A.; Pereira, M. J. B.; Rueda, D. *Biopolymers* **2001**, *61*, 224-242.

(50) Ehresmann, C.; Baudin, F.; Mougel, M.; Romby, P.; Ebel, J.; Ehresmann, B. *Nucleic Acids Res.* **1987**, *15*, 9109-9127.

(51) Agris, P. F. *Prog. Nucleic Acid Res. Mol. Biol.* **1996**, *53*, 79-129.

(52) Chow, C. S.; Bogdan, F. M. *Chem. Rev.* **1997**, *79*, 1489-1514.

(53) Latham, J. A.; Cech, T. R. *Science* **1989**, *245*, 276-282.

(54) Chow, C. S.; Barton, J. K. *Biochemistry* **1992**, *31*, 5423-5429.

(55) Chow, C. S.; Cunningham, P. R.; Lee, K.; Meroueh, M.; SantaLucia, J.; Varma, S. *Biochimie* **2002**, *84*, 859-868.

(56) Murakawa, G. J.; Chen, C. H.; Kuwabara, M. D.; Nierlich, D. P.; Sigman, D. S. *Nucleic Acids Res.* **1989**, *17*, 5361-5375.

(57) Chen, X.; Woodson, S. A.; Burrows, C. J.; Rokita, S. E. *Biochemistry* **1993**, *32*, 7610-7616.

(58) Simpson, L. *Gel Electrophoresis and Photography. An Application Note*; UVP, Inc., 1998.

(59) Berson, S. A.; Yalow, R. S. *Gastroenterology* **1972**, *62*, 1061-1084.

(60) Homola, J.; Yee, S. S.; Gauglitz, G. *Sensors and Actuators B: Chemical* **1999**, *54*, 3-15.

(61) Garner, M. M.; Revzin, A. *Nucleic Acids Res.* **1981**, *9*, 3047-3060.

(62) Fried, M.; Crothers, D. M. *Nucleic Acids Res.* **1981**, *9*, 505-6525.

(63) Werner, S. C.; Acebedo, G.; Radichevich, I. *J. Clin. Endocrinol. Metab.* **1974**, *38*, 493-495.

(64) Yalow, R. S.; Berson, S. A. *J. Clin. Invest.* **1996**, *39*, 1157-1175.

(65) Acebedo, G.; Hayek, A.; Klegerman, M.; Crolla, L.; Bermes, E.; Brooks, M. *Biochem. Biophys. Res. Commun.* **1975**, 449-456.

(66) Winzor, D. J.; Sawyer, W. H. *Quantitative Characterization of Ligand Binding*; Wiley, **1995**, 176.

(67) Mei, H.-Y.; Mack, D. P.; Galan, A. A.; Halim, N. S.; Heldsinger, A.; Loo, J. A.; Moreland, D. W.; Sannes-Lowery, K. A.; Sharmeen, L.; Truong, H. N.; Czarnik, A. W. *Bioorg. Med. Chem.* **1997**, *5*, 1173-1184.

(68) Mei, H. Y.; Cui, M.; Heldsinger, A.; Lemrow, S. M.; Loo, J. A.; Sannes-Lowery, K. A.; Sharmeen, L.; Czarnik, A. W. *Biochemistry* **1998**, *37*, 14204-14212.

(69) Sannes-Lowery, K. A.; Mei, H.-Y.; Loo, J. A. *Int. J. Mass Spectrom.* **1999**, *193*, 115-122.

(70) Chow, C. S.; Barton, J. K. *J. Am. Chem. Soc.* **1990**, *112*, 2839-2841.

(71) Sitlani, A.; Dupureur, C. M.; Barton, J. K. *J. Am. Chem. Soc.* **1993**, *115*, 12589-12590.

(72) Fitzsimons, M. P.; Barton, J. K. *J. Am. Chem. Soc.* **1997**, *119*, 3379-3380.

(73) Dandliker, P. J.; Holmlin, R. E.; Barton, J. K. *Science* **1997**, *275*, 1465-1468

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

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

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

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

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

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

(80) Griffey, R. H.; Hofstadler, S. A.; Sannes-Lowery, K. A.; Ecker, D. J.; Crooke, S. T. *Proc. Nat. Acad. Sci. U.S.A.* **1999**, *96*, 10129-10133.

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

(82) Cheng, X.; Chen, R.; Bruce, J. E.; Schwartz, B. L.; Anderson, G. A.; Hofstadler, S. A.; Gale, D. C.; Smith, R. D.; Gao, J.; Sigal, G. B.; Mammen, M.; Whitesides, G. M. *J. Am. Chem. Soc.* **1995**, *117*, 8859-8860.

(83) Smith, R. D.; Bruce, J. E.; Wu, Q.; Lei, Q. P. *Chem. Soc. Rev.* **1997**, *26*, 191-202.

(84) Sanner-Lowery, K. A.; Griffey, R. H.; Hofstadler, S. A. *Anal. Biochem.* **2000**, *280*, 264-270.

(85) Tanaka, K.; Waki, H.; Ido, Y.; Akita, S.; Yoshida, Y.; Yoshida, T. *Rapid Commun. Mass Spectrom.* **1988**, *2*, 151-153.

(86) Karas, M.; Hillenkamp, F. *Anal. Chem.* **1988**, *60*, 2299-2301.

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

(88) Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F. *Mass Spectrom. Rev.* **1990**, *9*, 37-70.

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

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

(91) Mo, J.; Hakansson, K. *Anal. Bioanal. Chem.* **2006**, *386*, 675-681.

(92) Schultz, J. C.; Hack, A. C.; Benner, W. H. *J. Am. Soc. Mass Spectrom.* **1998**, *9*, 305-313.

(93) Kebarle, P.; Tang, L. *Anal. Chem.* **1993**, *65*, 972A-986A.

(94) Bruins, A. P.; Covey, T. R.; Henion, J. D. *Anal. Chem.* **1987**, *59*, 2642-2646.

(95) Smith, R. D.; Loo, J. A.; Ogorzalek Loo, R. R.; Busman, M.; Udseth, H. R. *Mass Spectrom. Rev.* **1991**, *10*, 359-451.

(96) Enke, C. G. *Anal. Chem.* **1997**, *69*, 4885-4893.

(97) Jockusch, R. A.; Schnier, P. D.; Price, W. D.; Strittmatter, E. F.; Demirev, P. A.; Williams, E. R. *Anal. Chem.* **1997**, *69*, 1119-1126.

(98) Lum, R. C.; Grabowski, J. J. *J. Am. Chem. Soc.* **1992**, *114*, 8619-8627.

(99) Rodgers, M. T.; Campbell, S.; Marzluff, E. M.; Beauchamp, J. L. *Int. J. Mass Spectrom. Ion Proc.* **1994**, *137*, 121-149.

(100) Light-Wahl, 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.

(101) Ganem, B.; Li, Y.-T.; Henion, J. D. *Tetrahedron Lett.* **1993**, *34*, 1445-1448.

(102) Ding, J.; Anderegg, R. J. *J. Am. Soc. Mass Spectrom.* **1995**, *6*, 159-164.

(103) Pócsfalvi, G.; Landa, G. D.; Ferranti, P.; Ritieni, A.; Randazzo, G.; Malorni, A. *Rapid Commun. Mass Spectrom.* **1997**, *11*, 256-272.

(104) Gao, Q. Y.; Cheng, X. H.; Smith, R. D.; Yang, C. F.; Goldberg, I. H. *J. Mass. Spectrom.* **1996**, *31*, 31-36.

(105) Gale, D. C.; Smith, R. D. *J. Am. Chem. Soc.* **1995**, *6*, 1154-1164.

(106) Greig, M. J.; Hans Gaus, H.; Cummins, L. L.; Sasmor, H.; Griffey, R. H. *J. Am. Chem. Soc.* **1995**, *117*, 10765-10766.

(107) Shaffer, S. A.; Tang, K.; Anderson, G. A.; Prior, D. C.; Udseth, H. R.; Smith, R. D. *Rapid Commun. Mass Spectrom.* **1997**, *11*, 1813-1817.

(108) Kim, T.; Udseth, H. R.; Smith, R. D. *Anal. Chem.* **2000**, *72*, 5014-5019.

(109) Kim, T.; Tang, K.; Udseth, H. R.; Smith, R. D. *Anal. Chem.* **2001**, *73*, 4162-4170.

(110) Shaffer, S. A.; Prior, D. C.; Anderson, G. A.; Udseth, H. R.; Smith, R. D. *Anal. Chem.* **1998**, *70*, 4111-4119.

(111) Shaffer, S. A.; Tolmachev, A.; Prior, D. C.; Anderson, G. A.; Udseth, H. R.; Smith, R. D. *Anal. Chem.* **1999**, *71*, 2957-2964.

(112) Kim, T.; Tolmachev, A. V.; Harkewicz, R.; Prior, D. C.; Anderson, R.; Udseth, H. R.; Smith, R. D.; Bailey, T. H.; Rakov, S.; Futrell, J. H. *Anal. Chem.* **2000**, *72*, 2247-2255.

(113) Comisarow, M. B.; Marshall, A. G. *Chem. Phys. Lett.* **1974**, *25*, 282-283.

(114) Comisarow, M. B.; Marshall, A. G. *Chem. Phys. Lett.* **1974**, *26*, 489-490.

(115) Henry, K. D.; Williams, E. R.; Wang, B.-H.; McLafferty, F. W.; Shabanowitz, J.; Hunt, D. F. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 9075-9078.

(116) Amster, I. J. *J. Mass Spectrom.* **1996**, *31*, 1325-1337.

(117) Conrads, T. P.; Anderson, G. A.; Veenstra, T. D.; Pasa-Tolic, L.; Smith, R. D. *Anal. Chem.* **2000**, *72*, 3349-3354.

(118) Marshall, A. G.; Comisarow, M. B.; Parisod, G. *J. Chem. Phys.* **1979**, *71*, 4434-4444.

(119) Hakansson, K.; Cooper, H. J.; Hudgins, R. R.; Nilsson, C. L. *Curr. Org. Chem.* **2003**, *7*, 1503-1525.

(120) Kelleher, N. L.; Senko, M. W.; Siegel, M. M.; McLafferty, F. W. *J. Am. Soc. Mass Spectrom.* **1997**, *8*, 380-383.

(121) Henry, K. D.; McLafferty, F. W. *Org. Mass Spectrom.* **1990**, *25*, 490-492.

(122) Hendrickson, C. L.; Emmett, M. R. *Annu. Rev. Phys. Chem.* **1999**, *50*, 517-536.

(123) Valaskovic, G. A.; Kelleher, N. L.; McLafferty, F. W. *Science* **1996**, *273*, 1199-1202.

(124) Emmett, M. R.; Caprioli, R. M. *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 605-613.

(125) McLafferty, F. W. *Tandem Mass Spectrometry*; Wiley: New York, **1983**,

(126) Hayes, R. N.; Gross, M. L. *Methods Enzymol.* **1990**, *193*, 237-263.

(127) Gauthier, J. W.; Trautman, T. R.; Jacobson, D. B. *Anal. Chim. Acta* **1991**, *246*, 211-225.

(128) Little, D. P.; Speir, J. P.; Senko, M. W.; O'Connor, P. B.; McLafferty, F. W. *Anal. Chem.* **1994**, *66*, 2809-2815.

(129) Dunbar, R. C.; McMahon, T. B. *Science* **1998**, *279*, 194-197.

(130) Little, D. P.; Chorush, R. A.; Speir, J. P.; Senko, M. W.; Kelleher, N. L.; McLafferty, F. W. *J. Am. Chem. Soc.* **1994**, *116*, 4893-4897.

(131) Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. *J. Am. Chem. Soc.* **1998**, *120*, 3265-3266.

(132) Budnik, B. A.; Haselmann, K. F.; Zubarev, R. A. *Chem. Phys. Lett.* **2000**, *342*, 299-302.

(133) Budnik, B. A.; Haselmann, K. F.; Elkin, Y. N.; Gorbach, V. I.; Zubarev, R. A. *Anal. Chem.* **2003**, *75*, 5994-6001.

(134) Cody, R. B.; Freiser, B. S. *Anal. Chem.* **1979**, *51*, 541-551.

(135) Little, D. P.; Aaserud, D. J.; Valaskovic, G. A.; McLafferty, F. W. *J. Am. Chem. Soc.* **1996**, *118*, 9352-9359.

(136) Hofstadler, S. A.; Griffey, R. H.; Pasa-Tolic, L.; Smith, R. D. *Rapid Commun. Mass Spectrom.* **1998**, *12*, 1400-1404.

(137) Hakansson, K.; Chalmers, M. J.; Quinn, J. P.; McFarland, M. A.; Hendrickson, C. L.; Marshall, A. G. *Anal. Chem.* **2003**, *75*, 3256-3262.

(138) Hakansson, K.; Cooper, H. J.; Emmett, M. R.; Costello, C. E.; Marshall, A. G.; Nilsson, C. L. *Anal. Chem.* **2001**, *73*, 4530-4536.

(139) Kelleher, N. L.; Zubarev, R. A.; Bush, K.; Furie, B.; Furie, B. C.; McLafferty, F. W.; Walsh, C. T. *Anal. Chem.* **1999**, *71*, 4250-4253.

(140) Mirgorodskaya, E.; Roepstorff, P.; Zubarev, R. A. *Anal. Chem.* **1999**, *71*, 4431-4436.

(141) Stensballe, A.; Norregaard-Jensen, O.; Olsen, J. V.; Haselmann, K. F.; Zubarev, R. A. *Rapid Commun. Mass Spectrom.* **2000**, *14*, 1793-1800.

(142) Guan, Z. Q. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 1443-1447.

(143) Horn, D. M.; Ge, Y.; McLafferty, F. W. *Anal. Chem.* **2000**, *72*, 4778-4784.

(144) Horn, D. M.; Breuker, K.; Frank, A. J.; McLafferty, F. W. *J. Am. Chem. Soc.* **2001**, *123*, 9792-9799.

(145) Breuker, K.; Oh, H.; Horn, D. M.; Cerda, B.; McLafferty, F. W. *J. Am. Chem. Soc.* **2002**, *124*, 6407 - 6420.

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

(147) McLafferty, F. W.; Aaserud, D. J.; Guan, Z.; Little, D. P.; Kelleher, N. L. *Int. J. Mass Spectrom. Ion Processes* **1997**, *165/166*, 457-466.

(148) Griffey, R. H.; Greig, M. J. *Detection of Base Pair Mismatches in Duplex DNA and RNA Oligonucleotides using Electrospray Mass Spectrometry*; Proc. SPIE: Bellingham, **1997**.

(149) Griffey, R. H.; Greig, M. J.; An, H.; Sasmor, H.; Manalili, S. *J. Am. Chem. Soc.* **1999**, *121*, 474-475.

(150) Fourmy, D.; Recht, M. I.; Puglisi, J. D. *J. Mol. Biol.* **1998**, *277*, 347-362.

(151) Fourmy, D.; Yoshizawa, S.; Puglisi, J. D. *J. Mol. Biol.* **1998**, *277*, 333-345.

(152) Schultz, K. N.; Hakansson, K. *Int. J. Mass Spectrom.* **2004**, *234*, 123-130.

(153) Yang, J.; Mo, J.; Adamson, J. T.; Hakansson, K. *Anal. Chem.* **2005**, *77*, 1876-1822.

(154) Wang, F.; Li, W.; Emmett, M. R.; Hendrickson, C. L.; Marshall, A. G.; Zhang, Y.-L.; Wu, L.; Zhang, Z.-Y. *Biochemistry* **1998**, *37*, 15289 - 15299.

(155) Green, M. K.; Lebrilla, C. B. *Mass Spectrom. Rev.* **1997**, *16*, 53-71.

(156) Robinson, J. M.; Greig, M. J.; Griffey, R. H.; Mohan, V.; Laude, D. A. *Anal. Chem.* **1998**, *70*, 3566-3571.

(157) Freitas, M. A.; Shi, S. D. H.; Hendrickson, C. L.; Marshall, A. G. *J. Am. Chem. Soc.* **1998**, *120*, 10187-10193.

(158) Freitas, M. A.; Hendrickson, C. L.; Emmett, M. R.; Marshall, A. G. *J. Am. Soc. Mass Spectrom.* **1998**, *9*, 1012-1019.

(159) Hofstadler, S. A.; Sannes-Lowery, K. A.; Griffey, R. H. *J. Mass Spectrom.* **2000**, *35*, 62-70.

(160) Freitas, M. A.; Marshall, A. G. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 780-785.

(161) Ausloos, P.; Lias, S. G. *J. Am. Chem. Soc.* **1981**, *103*, 3641-3647.

(162) Green-Church, K. B.; Limbach, P. A.; Freitas, M. A.; Marshall, A. G. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 268-277.

(163) DePuy, C. H.; Bierbaum, V. M. *Acc. Chem. Res.* **1981**, *14*, 146-153.

(164) Chipuk, J. E.; Brodbelt, J. S. *J. Am. Soc. Mass Spectrom.* **2007**, *18*, 724-736.

(165) Crestoni, M. E.; Fornarini, S. *J. Mass Spectrom.* **2003**, *38*, 854-861.

(166) Kamel, A. M.; Munson, B. *J. Am. Soc. Mass Spectrom.* **2007**, *18*, 1477-1492.

(167) Griffey, R. H.; Greig, M. J.; Robinson, J. M.; Laude, D. A. *Rapid Commun. Mass Spectrom.* **1999**, *13*, 113-117.

(168) Gabelica, V.; Rosu, F.; Witt, M.; Baykut, G.; de Pauw, E. *Rapid Commun. Mass Spectrom.* **2005**, *19*, 201-208.

CHAPTER 2

OLIGONUCLEOTIDE GAS-PHASE HYDROGEN/DEUTERIUM EXCHANGE WITH D_2S IN THE COLLISION CELL OF A Q-FT-ICR MASS SPECTROMETER

In this Chapter, we present the implementation of gas-phase HDX in the collision cell of a hybrid Q-FT-ICR mass spectrometer and investigate the mechanism of oligonucleotide gas-phase HDX with D_2S. In this configuration, multiply charged oligonucleotide anions undergo significant exchange with D_2S at reaction intervals ranging from 0.11 to 60.1 s. For DNA homohexamers, relative exchange rates are $dC_6 \sim dA_6 > dG_6 > dT_6$, correlating with the gas-phase acidities of nucleobases (C > A > T > G), except for guanine. Our results are consistent with a relay mechanism in which D_2S interacts with both a backbone phosphate group and a neutral nucleobase through hydrogen bonding. We propose that the faster exchange of polyguanosine compared to polythymidine is due to the larger size of guanine and the orientation of its labile hydrogens, which may result in gas-phase conformations more favorable for forming complexes with D_2S. Similar trends were observed for RNA homohexamers, although their HDX rates were faster than those for DNA, suggesting they can also exchange via another relay process involving the 2' hydroxyl group. HDX of DNA duplexes further supports the involvement of nucleobase hydrogens because duplexes exchanged slower

than their corresponding single strands, presumably due to the intermolecular hydrogen bonds between nucleobases. This work constitutes the first investigation of the mechanisms of nucleic acid gas-phase HDX. Our results on duplexes show promise for applying this strategy to characterize structured nucleic acids, which will be discussed in Chapter 5.

2.1 Introduction

HDX is a well established technique for probing the solution-phase structure of biological molecules based on the accessibility of various potentially exchangeable hydrogens.[1-4] However, due to the high opening rates of nucleobase pairs, HDX of structured nucleic acids in solution is very fast (~ms), rendering it difficult to track reaction details. Gas-phase HDX[5-9] in which rate constants are much smaller than those in solution has been proposed as an alternative method for structural analysis of nucleic acids.[8, 10-16] This approach has become more feasible for larger nucleic acids due to the introduction of soft ionization methods, such as MALDI[17, 18] and ESI.[19, 20] Comparison of gas- and solution-phase conformations can provide insight into the nature of these ionization processes, particularly the role of solvent. Mass spectrometers are well suited for performing gas-phase HDX experiments because deuterium incorporation can be monitored directly from mass spectral peak shifts with reaction time, which is easily controlled by data acquisition programs. In particular, ion trapping techniques, such as ICR and quadrupole ion traps, can provide gas-phase HDX kinetics because ions can be trapped for extended time periods, allowing for observation of the relatively slow exchange processes.[8]

DePuy *et al.*[21] suggested that HDX can be observed between anions and exchanging reagent that is as much as 20 kcal/mol less acidic. Ausloos *et al.*[22] also stated that HDX was not observed for protonated compounds if the gas-phase basicities of the exchanging reagents differed by more than 20 kcal/mol. However, HDX can still occur between reagents with large acidity differences, presumably due to complex formation.[5, 12, 23, 24] Among the deuterated reagents explored for gas-phase HDX of nucleic acids, D_2S is most suitable[16, 23] because its p*K*a is closer to the deprotonated backbone phosphate groups than those of other reagents, such as ND_3.[22, 24] Consequently, HDX rate constants were found to be much higher for D_2S than for D_2O (minutes vs. hours at 5 x 10^{-8} Torr pressure) due to the smaller gas-phase acidity difference between D_2S and nucleic acids (> 20 kcal/mol vs. > 40 kcal/mol for D_2O).[23] Furthermore, D_2S can very easily be introduced into a mass spectrometer because it is a gas at room temperature. Most previous HDX studies of nucleic acids involved mono-(deoxy)-nucleotides with a few exceptions in which five- and 20-mer phosphorothioate oligodeoxynucleotides were characterized.[13, 14, 25] However, the exchange mechanism has only been discussed for mononucleotide phosphates. A relay process involving a 5' phosphate group and the 3' hydroxyl group of the sugar but not the nucleobase or the 2' hydroxyl group has been proposed for exchange of 5'-monophosphate anions with D_2S.[16] By contrast, the 2' hydroxyl group has been proposed to play an important role in gas-phase HDX of mononucleotide cations.[24] In this chapter, we attempt to elucidate the role of nucleobases in the gas-phase HDX process of nucleic acid anions. Furthermore, in order to explore the analytical utility of this approach, we applied gas-phase HDX with D_2S to DNA duplexes to determine whether their exchangeable nucleobase hydrogens involved in base

pairing would be protected from exchange and thereby display slower exchange rates. In a previous study, a quadruplex was shown to undergo more rapid HDX than its constituent monomer,[15] a behavior that was not expected. However, quadruplexes do not contain W-C base pairs.

2.2. Experimental Section

2.2.1. Sample Preparation

DNA and RNA hexamers, including dT_6, dC_6, dA_6, d(GCATAC), d(GCATGC), d(TGGGGT), A_6, U_6 and 2'-methylated (2'-OMe) RNA were purchased from either TriLink Biotechnologies (San Diego, CA) or the Yale Keck Facility (New Haven, CT). All reagents were used without further purification. Oligonucleotides were diluted into 5% v/v isopropanol (Fisher, Fair Lawn, NJ) with 10 mM ammonium acetate (Fisher) to a concentration of 10-100 μM (adjusted to yield a signal-to-noise ratio of 10 in 8 scans). Duplex formation was promoted by annealing DNA at 95 °C for 5 minutes, cooling down to 70 °C for 30 minutes, and storing overnight at 4 °C.

2.2.2. Mass Spectrometry

All experiments were conducted on an actively shielded 7 T Q-FT-ICR mass spectrometer (Bruker Daltonics, Billerica, MA, described in Chapter 1) in negative ion mode. Oligonucleotide solutions were infused via an external Apollo II ion funnel electrospray ion source at a flow rate of 50 μL/h with the assistance of N_2 nebulizing gas. The inlet capillary was set to 2.8 kV for generation of oligonucleotide anions. N_2 drying gas (120 °C, 3.9 L/s) was applied to assist ESI droplet desolvation. HDX data were generated from doubly deprotonated precursor ions for DNA or RNA hexamers and triply

deprotonated ions for DNA duplexes. The entire precursor isotopic envelope was quadrupole selected and stored in a hexapole collision cell in the presence of D_2S for 0.11 to 60.1 s. Multiple ICR cell fills were not used to avoid generation of mixtures of ion populations with different exposure time to D_2S (there is residual D_2S outside the storage hexapole). D_2S (Cambridge Isotope Laboratories, Andover, MA) was leaked into the collision cell to a pressure of ~ $5x10^{-6}$ mbar (gauge factory calibrated for nitrogen; no additional calibration was performed). Although different gas-phase conformations of nucleic acids may exist, hydrogens residing on different positions may exchange at different rates. Percentage (%) of HDX was calculated based on the average m/z of the entire oligonucleotide isotopic distribution, i.e., by considering the relative abundance of each isotopic peak, according to the following equation (adapted from Zhang and Smith[26])

$$\% \text{ deuterium incorporation} = [(m/z)_{obs.} - (m/z)_0]/[(m/z)_{max.} - (m/z)_0]*100\%$$

in which $(m/z)_{obs.}$ = observed average m/z following a particular exchange time, $(m/z)_0$ = average m/z prior to HDX, and $(m/z)_{max.}$ = expected average m/z at full deuteration of all exchangeable hydrogens. Error bars (higher errors are expected with more deuterium incorporated) were generated from data (three repeats) acquired on the same day due to difficulties with reproducing the D_2S pressure in the collision cell. Number (#) of HDX was also calculated for comparison according to the following equation

$$\# \text{ deuterium incorporation} = [(m/z)_{obs.} - (m/z)_0] * z$$

and results showed similar HDX rate trends as using % of HDX (one example is shown in Chapter 6), therefore, this Chapter only shows data calculated using % of HDX (the same data handling is used in later Chapters).

2.3. RESULTS AND DISCUSSION

2.3.1. Number of Exchangeable Hydrogens

The HDX process of oligonucleotides requires a deuterium acceptor group and a hydrogen donor group.[12, 23] For both DNA and RNA, the most acidic sites are the backbone phosphate groups,[23, 27] which are deprotonated in negative ion mode and therefore constitute possible deuterium acceptors during HDX. Each phosphate group (excluding the deprotonated ones) has one exchangeable hydrogen while nucleobases have varied numbers: adenine and cytosine both have two exchangeable hydrogens, guanine has three, and thymine and uracil both have one. In addition, the two end hydroxyl groups in DNA and RNA each have one exchangeable hydrogen. Therefore, the total numbers of exchangeable hydrogens for the doubly deprotonated DNA hexamer anions are seventeen for dA_6, dC_6 and d(GCATAC), nineteen for d(TGGGGT), eleven for dT_6, and eighteen for d(GCATGC). For RNA, the 2' hydroxyl group of each sugar adds additional exchangeable hydrogens. Therefore, the doubly deprotonated RNA hexamers have six more exchangeable hydrogens than their corresponding DNA hexamers.

2.3.2. Gas-phase Acidity of Nucleic Acids

It is generally believed that HDX is more favorable if the gas-phase acidities of the analyte and the deuterated reagent are close in value. Calculated acidities of the deuterium donor (D_2S) and acceptor (phosphate diester bridge in nucleic acids) are 351.3 kcal/mol and 329.0 kcal/mol, respectively.[23, 27, 28] Their difference is more than 20 kcal/mol, which is energetically unfavorable for HDX. However, previous experiments showed that exchange can still be observed between reagents with large acidity

differences, which was attributed to complex formation between reagents, followed by a relay exchange process.[7, 16, 23, 24] Prior experimental determination of the gas-phase acidities of nucleobases yielded the following trend: C > A > T > G, all with lower values than that of D_2S, thus, acidity differences between nucleobases and D_2S follow the trend C < A < T < G. Chan and Enke found that reaction efficiency increased as the gas-phase acidity difference between anions and neutral reagent decreased.[29] Therefore, the order of DNA HDX rates should hypothetically follow the acidity trend; C > A > T > G.

2.3.3. HDX of DNA

Upon HDX, oligonucleotide m/z values (z = 2) shift to higher values by 0.5 for each deuterium incorporated. Figure 2.1 displays the mass spectra of dA_6 (left) and dT_6 (right) doubly deprotonated anions following different HDX times. Deuterium incorporation for dA_6 is fast during the first 10 s and then slows down up to 20 s after which it remains close to constant for the remainder of the exchange time (Figure 2.1a). Similar behavior was observed for dC_6, d(GCATGC), and for d(GCATAC) (see Figure 2.2), but to a different extent (i. e., 75% incorporation for dA_6, 74% for dC_6, 67% for d(GCATGC), and 86% for d(GCATAC)). dT_6 (Figure 2.1b) and d(TGGGGT) (Figure 2.2) displayed different behavior with continuous exchange throughout the entire reaction period (up to one minute).

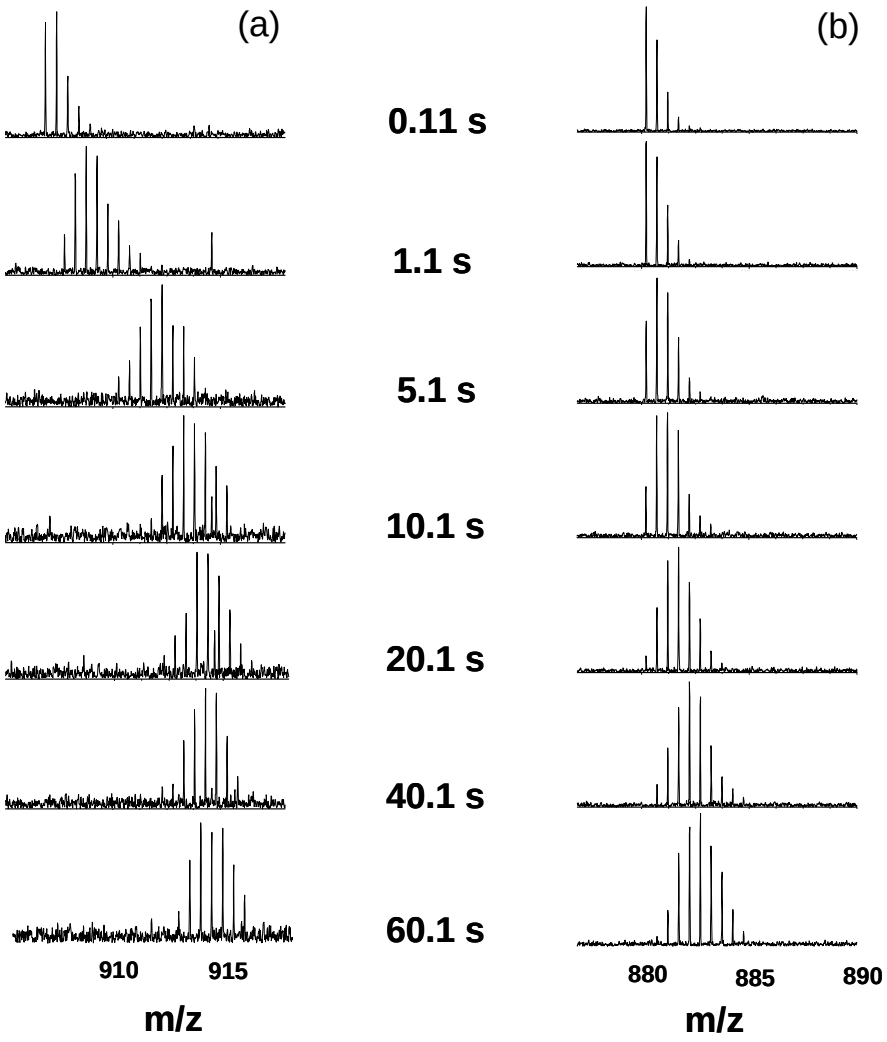


Figure 2.1. m/z shift of doubly deprotonated DNA following different HDX time periods. dA_6 exchanged rapidly during the first 10 s, then slower up to ~20 s after which the exchange plateaued (a). dT_6 exchanged at a slower rate but kept exchanging over the entire reaction period (b).

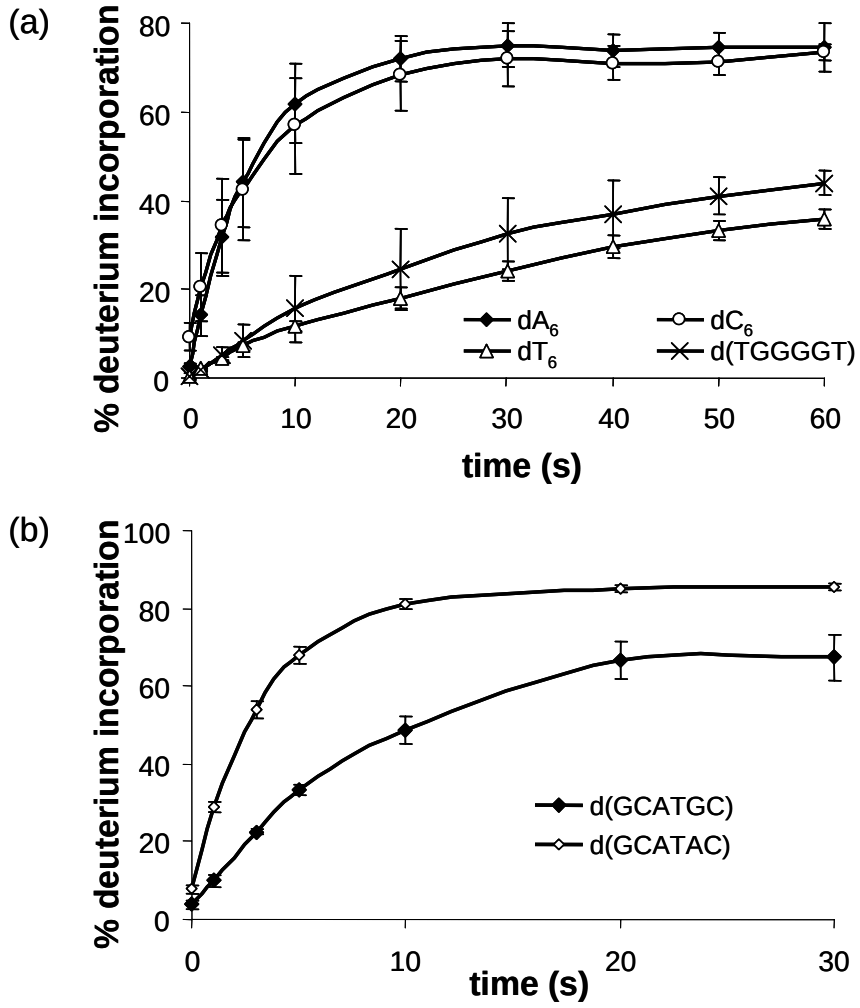
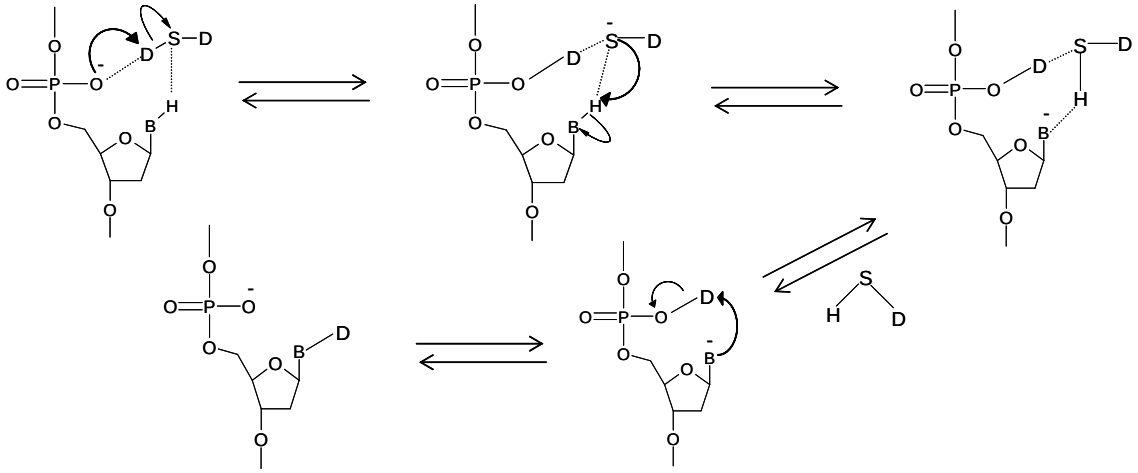


Figure 2.2. HDX of DNA hexamers in negative ion mode. Observed rates are dC_6 ~ dA_6 > d(TGGGGT) > dT_6 (a) and d(GCATAC) > d(GCATGC) (b), i.e., C ~ A > G > T.

Comparisons of HDX rates for all six DNA hexamers with D_2S are shown in Figure 2.2. Based on DNA oligomer structure, only the 5'and 3' hydroxyl groups can undergo HDX via the mechanism proposed for nucleoside-5'-monophosphates.[16] However, the number of experimentally observed exchanges for hexamer DNAs is much higher than two; e.g., doubly deprotonated dA_6 exchanged about twelve hydrogens with D_2S in one minute although it only contains five backbone exchangeable hydrogens.

This behavior indicates that the remaining exchanges must occur at nucleobases. Furthermore, if a deprotonated phosphate is located in the middle of the oligonucleotide rather than at the ends, nucleobases may be more accessible for forming a complex with D_2S than the end hydroxyl groups. Therefore, contrary to nucleoside monophosphates whose initial HDX step was proposed not to involve the nucleobases, these DNA hexamers are likely having their nucelobases participate in the initial complexation step. Our experimentally observed rates were dC_6 ~ dA_6 > d(TGGGGT) > dT_6, and d(GCATAC) > d(GCATGC), providing the overall trend C ~ A > G > T. This behavior correlates with the expected trend C > A > T > G, except for guanosine, which has lower gas-phase acidity than T but exchanged faster. A relay mechanism that may explain this behavior is proposed in Scheme 2.1. Here, D_2S forms a complex with a deprotonated backbone phosphate group and a neutral DNA nucleobase through hydrogen bonding. The phosphate group is deuterated by D_2S while the resulting DS^- anion is protonated by the nucleobase, followed by transfer of a deuteron from the phosphate group to the nucleobase. In this proposed mechanism, the faster than expected exchange observed for dG_6, may be due to the large size of guanine and the preferred orientation of its labile hydrogens (primary amine),[10, 12, 23] resulting in a gas-phase conformation that favors formation of an oligonucleotide-D_2S complex. On the other hand, both Freitas *et al.*[23] and Robinson *et al.*[12] found that 5' dGMP exchanged slower than other nucleoside monophosphates, which was attributed to hydrogen bonding between guanine and the 5' phosphate (geometry optimization of 5' dGMP obtained from a semiempirical PM3 calculation suggests that the 5' phosphate can form a hydrogen bond with the nucelobase

when the glycosidic bond adopts a syn orientation[12]). However, several gas-phase conformations likely exist for the oligonucleotide dG_6.



Scheme 2.1. Proposed HDX relay mechanism for DNA oligonucleotides in negative ion mode. Exchange of nucleobases is influenced by their gas-phase acidities. The data in Figure 2.2 correlate well with this hypothesis, except for dG_6 (gas-phase acidity of G is lower than for T). One possible explanation is that, due to the larger size of G compared to T, a conformation favoring D_2S complexation with the phosphate and a nucleobase is more readily formed.

2.3.4. HDX of RNA Compared to DNA

Because RNA has an additional hydroxyl group at the 2' position of each sugar, it may undergo gas-phase HDX through a different mechanism than DNA. When comparing A_6 to dA_6 (Figure 2.3a), faster initial HDX rate was observed for the RNA (A_6 > dA_6). Similar behavior was seen when comparing dT_6 and U_6 (Figure 2.3b). This experimentally observed behavior indicates that the 2' OH group of RNA may be involved in HDX because it constitutes the only structural difference between RNA and DNA (excluding uracil/thymine). The doubly deprotonated A_6 anion has six

exchangeable hydrogens at 2' sugar positions, and five exchangeable hydrogens at backbone phosphate groups and terminal hydroxyl groups. However, it exchanged on average thirteen hydrogens with D_2S within one minute, which is more than the sum of these two groups on the backbone (eleven), suggesting that at least some of the exchanged hydrogens must reside on the nucleobases. HDX of RNA nucleobases may occur via the same mechanism as for DNA (Scheme 2.1). However, an alternative mechanism (which could occur simultaneously as the one in Scheme 1) involving the 2' OH group is also possible. In such a mechanism, D_2S forms a complex with a deprotonated phosphate group and a 2' OH group instead of a nucleobase, followed by intermolecular relay HDX between D_2S and the RNA backbone, and intramolecular deuterium rearrangement between the phosphate group and the 2' hydroxyl group. The faster HDX of RNA compared to DNA could be either due to higher acidity of the 2'OH compared to that of nucleobases (higher acidity of the hydrogen donor group should facilitate the second step in relay HDX), or due to steric effects.

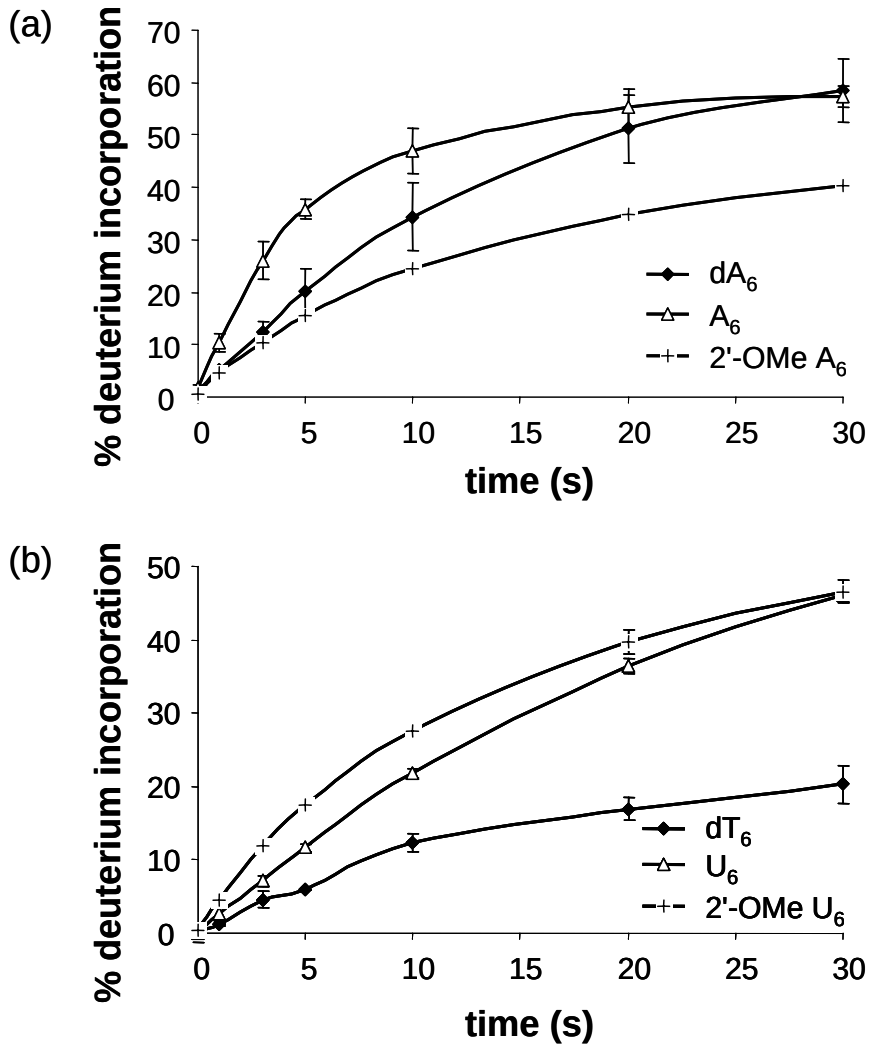


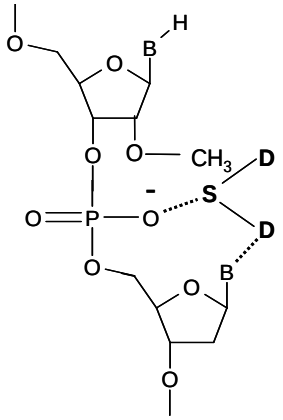
Figure 2.3. HDX of DNA compared to RNA and 2'-OMe RNA. Observed rates are RNA > DNA > 2'-OMe RNA except 2'-OMe U_6 which exchanges faster than dT_6.

2.3.5. HDX of 2'-OMe RNA Compared to DNA

HDX experiments on a modified form of RNA in which the 2' hydroxyl groups are methylated (2'-OMe A_6, Scheme 2.2) were performed to further investigate the role of the 2' hydroxyl group in RNA HDX. Due to this methylation, there is no exchangeable hydrogen on the sugar ring, suggesting that similar exchange behavior may

be observed for 2'-OMe RNA compared to DNA. However, slower exchange was observed for 2'-OMe A_6 compared to its corresponding DNA (Figure 2.3a). This behavior may be understood from steric hindrance by the 2' methyl group compared to 2' hydrogen, disfavoring formation of a complex with D_2S (Scheme 2.2). By contrast, the modified RNA 2'-OMe U_6 displayed faster HDX than its corresponding DNA (dT_6, Figure 2.3b). However, the thymine methyl group may impose a similar steric effect compared to uracil, thereby counteracting the effect of the 2' chemical group. Compared to U_6, 2'-OMe U_6 had a very similar HDX behavior (Figure 2.3b), suggesting that the Scheme 2.1 mechanism dominates for these oligonucleotides.

2'-OMe RNA



Scheme 2.2. Structures of 2'-OMe RNA. The methyl group at the 2' position of 2'-OMe A_6 may induce steric hindrance, thereby impeding its complexation with D_2S, resulting in a slower HDX process compared to dA_6.

2.3.6. HDX of DNA Duplexes

After establishing that (deoxy-)oligonucleotide gas-phase HDX rates depend on nucleobase acidities, the presence of a 2' hydroxyl group, and structural and steric effects, we set out to explore whether exchange rates can still reflect structural changes. As an example, we performed HDX of two DNA duplexes; one formed from the self-complementary oligonucleotide d(GCATGC), which can form a W-C duplex, and one formed from d(TGGGGT), which can form a duplex through Hoogsteen hydrogen bonding. Both these duplexes displayed slower exchange rates than their corresponding single strand DNAs, as shown in Figure 2.4. Hydrogen bonding clearly protects nucleobase hydrogens from HDX, thereby revealing intriguing possibilities for characterizing higher order structure of other folded nucleic acids. Furthermore, the reduced HDX rates observed for duplexes further support our claim that nucleobases are involved in HDX of oligonucleotides.

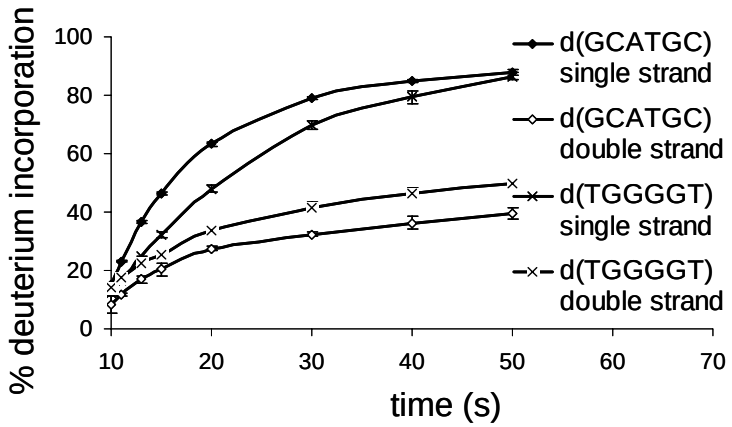


Figure 2.4. HDX of DNA duplexes compared to their corresponding single strands. Due to the low efficiency of duplex formation, a longer accumulation time (10 s) was used compared to the hexamers (Figures 2.2 and 2.3). Both W-C and Hoogsteen duplexes exchanged slower than their corresponding single strands because hydrogen bonds protect nucleobases from exchanging.

2.4. CONCLUSION

In this Chapter, we show that gas-phase HDX can be easily and rapidly performed in the external collision cell of a Q-FT-ICR mass spectrometer. Observed exchange behavior for DNA, RNA and 2'- methylated RNA is consistent with relay mechanisms originating by complexation of the biomolecules with the deuterated reagent and followed by inter- and intramolecular hydrogen/deuterium transfer. We propose that negatively charged phosphate groups and neutral nucleobases both participate in the HDX process of DNA and RNA while 2' OH groups are also involved in RNA HDX, resulting in faster exchange rates compared to DNA. The order of DNA HDX rates follows the gas-phase acidities of nucleobases, except for guanine, presumably due to its larger size and the orientation of its labile hydrogens, facilitating a gas-phase conformation that favors the formation of a complex with D_2S. Despite this complicated HDX behavior, two different kinds of DNA duplexes displayed slower exchange rates than their corresponding single strands, which also further supports the suggestion that nucleobases are involved in HDX of oligonucleotides. Thus, there is still precedent for the use of gas-phase HDX combined with mass spectrometry as a valuable technique for investigating gas-phase structures of nucleic acids. Chapter 5 will discuss utilization of this method for determining gas-phase higher order structural stability of larger nucleic acids.

2.5. BIBLIOGRAPHY

(1) Englander, S. W.; Sosnick, T. R.; Englander, J. J.; Mayne, L. *Curr. Opin. Struct. Biol.* **1996**, *6*, 18.

(2) Wales, T. E.; Engen, J. R. *Mass Spectrom. Rev.* **2006**, *25*, 158-170.

(3) Hoofnagle, A. N.; Resing, K. A.; Ahn, N. G. *Annu. Rev. Biophys. Biomol. Struct.* **2003**, *32*, 1-25.

(4) Smith, D. L.; Zhang, Z. Q. *Mass Spectrom. Rev.* **1994**, *13*, 411-429.

(5) Freitas, M. A.; Hendrickson, C. L.; Emmett, M. R.; Marshall, A. G. *J. Am. Soc. Mass Spectrom.* **1998**, *9*, 1012-1019.

(6) Campbell, S.; Rodgers, M. T.; Marzluff, E. M.; Beauchamp, J. L. *J. Am. Chem. Soc.* **1994**, *116*, 9765.

(7) Gard, E.; Green, M. K.; Bregar, J.; Lebrilla, C. B. *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 623-631.

(8) Green, M. K.; Lebrilla, C. B. *Mass Spectrom. Rev.* **1997**, *16*, 53-71.

(9) Wyttenbach, T.; Bowers, M. T. *J. Am. Soc. Mass Spectrom.* **1999**, *10*, 9-14.

(10) Chipuk, J. E.; Brodbelt, J. S. *J. Am. Soc. Mass Spectrom.* **2007**, *18*, 724-736.

(11) Crestoni, M. E.; Fornarini, S. *J. Mass Spectrom.* **2003**, *38*, 854-861.

(12) Robinson, J. M.; Greig, M. J.; Griffey, R. H.; Mohan, V.; Laude, D. A. *Anal. Chem.* **1998**, *70*, 3566-3571.

(13) Griffey, R. H.; Greig, M. J.; Robinson, J. M.; Laude, D. A. *Rapid Commun. Mass Spectrom.* **1999**, *13*, 113-117.

(14) Hofstadler, S. A.; Sannes-Lowery, K. A.; Griffey, R. H. *J. Mass Spectrom.* **2000**, *35*, 62-70.

(15) Gabelica, V.; Rosu, F.; Witt, M.; Baykut, G.; de Pauw, E. *Rapid Commun. Mass Spectrom.* **2005**, *19*, 201-208.

(16) Freitas, M. A.; Marshall, A. G. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 780-785.

(17) Tanaka, K.; Waki, H.; Ido, Y.; Akita, S.; Yoshida, Y.; Yoshida, T. *Rapid Commun. Mass Spectrom.* **1988**, *2*, 151-153.

(18) Karas, M.; Hillenkamp, F. *Anal. Chem.* **1988**, *60*, 2299-2301.

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

(20) Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F. *Mass Spectrom. Rev.* **1990**, *9*, 37-70.

(21) DePuy, C. H.; Bierbaum, V. M. *Acc. Chem. Res.* **1981**, *14*, 146-153.

(22) Ausloos, P.; Lias, S. G. *J. Am. Chem. Soc.* **1981**, *103*, 3641-3647.

(23) Freitas, M. A.; Shi, S. D. H.; Hendrickson, C. L.; Marshall, A. G. *J. Am. Chem. Soc.* **1998**, *120*, 10187-10193.

(24) Green-Church, K. B.; Limbach, P. A.; Freitas, M. A.; Marshall, A. G. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 268-277.

(25) Hofstadler, S. A.; Sannes-Lowery, K. A.; Griffey, R. H. *Rapid Commun. Mass Spectrom.* **1999**, *13*, 1971-1979.

(26) Zhang, Z.; Smith, D. L. *Protein Sci.* **1993**, *2*, 522-531.

(27) Rodgers, M. T.; Campbell, S.; Marzluff, E. M.; Beauchamp, J. L. *Int. J. Mass Spectrom. Ion Proc.* **1994**, *137*, 121-149.

(28) Lum, R. C.; Grabowski, J. J. *J. Am. Chem. Soc.* **1992**, *114*, 8619-8627.

(29) Chan, S.; Enke, C. G. *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 282-291.

CHAPTER 3

ROLES OF PHOSPHATE GROUPS AND NUCLEOBASES IN POSITIVE ION MODE DNA GAS-PHASE HYDROGEN/DEUTERIUM EXCHANGE WITH D_2S

A relay mechanism involving both backbone phosphate groups and nucleobases has been proposed in Chapter 2 for negative ion mode nucleic acid gas-phase HDX with D_2S. However, details about how these individual groups participate in HDX have not been established. To investigate DNA HDX behavior in more detail, we performed gas-phase HDX experiments on modified DNAs in which these functional groups are removed, including methyl-phosphonate DNA (Met-DNA, phosphate hydroxyl group replaced by methyl group) and abasic DNA (one nucleobase is replaced by hydrogen), and compared their behavior to those of natural DNAs. Because Met-DNAs are difficult to deprotonate due to the lack of acidic sites, experiments were conducted in positive ion mode. Similar to negative ion mode, both the phosphate groups and nucleobases are involved in positive ion mode DNA HDX and the process appears to be initiated by complexation of D_2S with a hydrogen donor group and a deuterium acceptor group in DNA, followed by inter- and intra-molecular hydrogen/deuterium transfer. For example, phosphate hydroxyl groups can act as hydrogen donors with the nucleobases as deuterium acceptors, rendering HDX rates dependent on nucleobase proton affinities. Our results

show a correlation between HDX rates for the four DNA nucleobases and their proton affinities. Results also suggest that adenine can form intra-molecular hydrogen bonds with phosphate groups, thereby hampering the HDX process.

3.1. Introduction

Elucidation of biomolecular gas-phase structure would be very useful if it could mimic solution-phase structure because most biological activities occur in solution. With the development of soft ionization methods that can preserve solution-phase structure when molecules transfer into the gas phase, e.g., MALDI[1, 2] and ESI,[3-6] MS is becoming a more and more routine technique for characterizing solution-phase structures.[7-10] Gas-phase HDX is one method that can be coupled with MS for probing structures based on the accessibility of various potentially exchangeable hydrogens. This approach has been used extensively to distinguish between isomeric species,[11, 12] deduce reaction mechanisms,[13-16] and infer structural features of complexes,[17, 18] and it is becoming a practical way to investigate biomolecular structures. Although this combined method has been more broadly applied for structural characterization of peptides and proteins,[19-25] it is highly suited for characterization of nucleic acid structure because the high opening rates of nucleobase pairs makes solution-phase HDX challenging.

Several groups have examined gas-phase HDX of nucleic acids, including investigation of the HDX mechanism, verification of fragmentation pathways, and characterization of complex conformation.[26-34] Understanding of the gas-phase HDX mechanism is critical for its application and correct interpretation of the corresponding data. Chapter 2 suggests that besides the phosphate and terminal hydroxyl groups, the

nucleobases also participate in oligonucleotide HDX[35] and observed rates were found to be related to the gas-phase acidities of nucleobases as well as their conformational accessibility to form a complex with the exchanging reagent (D_2S). In this Chapter, we aim to investigate more details about the HDX mechanism, i.e., the role of each individual DNA chemical group, including the phosphate groups and the nucleobases, by comparing the HDX behavior of modified DNAs with that of natural DNAs. In the modified DNAs, either all the backbone phosphate groups are methylated (Met-DNA), or the nucleobases are removed and replaced with hydrogens such that the complexation step in the previously proposed mechanism should not be possible. Experiments were conducted in positive ion mode because the Met-DNA does not contain acidic sites and therefore could not be observed in negative ion mode. D_2S was still used as the exchanging reagent to avoid the corrosiveness of ND_3 (the latter reagent has been reported to be more effective in positive ion mode because it has higher proton affinity than D_2S[24, 34]).

3.2. Experimental Section

3.2.1. Sample Preparation

DNAs (dA_6, dC_6, d(GCATGC), d(GCATAC), 12mer-G d(GGGGATATGGGG) and 12mer-C d(CCCCATATCCCC)) as well as their modified forms (abasic variants and Met-DNA) were purchased from either TriLink Biotechnologies (San Diego, CA) or the Yale Keck Facility (New Haven, CT). Anion exchange C_{18} Ziptips (Millipore Corporation, Billerica, Massachusetts) were used to desalt 12mer-G while all the other samples were used without further purification. Samples were diluted into 20% methanol

(Fisher, Fair Lawn, NJ) with 0.1% formic acid (Fisher) to a concentration of 10-100 μM (adjusted to yield a signal-to-noise ratio of more than 10 in 8 scans).

3.2.2. Mass Spectrometry

All experiments were performed in positive ion mode with a 7 Tesla Q-FT-ICR mass spectrometer (Bruker Daltonics, Billerica, Massachusetts) described Chapter 1.[35-37] Analyte solutions were infused at a flow rate of 50 μL/h with assistance of N_2 nebulizing gas. Apollo II ion funnel electrospray ion source (Bruker Daltonics) was used to generate multiply charged precursor cations (doubly charged for all oligonucleotides except 12mer-G and 12mer-C which were mainly triply charged). Protonated precursor ions were mass selected by the quadrupole with the isolation window large enough to include the entire isotopic distribution but no salt adduct peaks. Selected precursor ions were externally accumulated in a hexapole for a short time (0.1 s) in presence of the deuterated reagent D_2S (Cambridge Isotope Laboratories) at a pressure of ~ $5x10^{-6}$ mbar. The HDX reaction and the calculation of percentage of deuterium incorporation were performed as described in Chapter 2.

3.3. RESULTS AND DISCUSSION

3.3.1. Number of Exchangeable Hydrogens

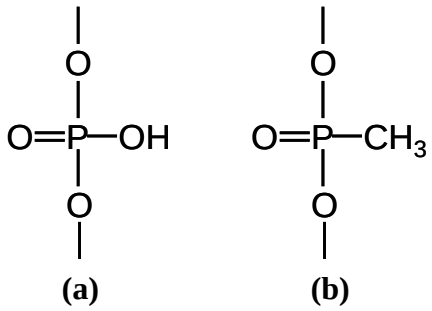
Protonated oligonucleotides have several groups that contain exchangeable hydrogens, e.g., phosphate hydroxyl groups, nucleobases, terminal hydroxyl groups, and the protonation sites. The latter are most likely to be on the nucleobases, except for thymine which has low proton affinity. Consequently, the total numbers of exchangeable hydrogens for the protonated DNA cations used here are 21 for dA_6, dC_6 and

d(GCATAC), 22 for d(GCATGC), 46 for 12mer-G, and 38 for 12mer-C. Due to the large difference of proton affinities between the DNA and the D_2S (> 20 kcal/mole), the HDX process is likely to undergo a relay mechanism.[34] D_2S would first form a complex with a hydrogen donor group, which can be the terminal hydroxyl groups, a nucleobase, or the protonation site and a deuterium acceptor group, which can be the phosphate double-bonded oxygen (P=O), or a nucleobase, and then undergo inter- and intra-molecular hydrogen/deuterium transfer. Our goal was to investigate the role of these individual groups on oligonucleotide gas-phase HDX behavior with D_2S as exchange reagent.

3.3.2. Role of Phosphate Groups in Positive Ion Mode DNA HDX

The role of phosphate groups in positive ion mode DNA HDX was evaluated by comparing the HDX behavior of natural DNA hexamers with their modified derivatives (Met-DNA), in which all the five backbone phosphate groups were methylated (structure shown in Scheme 3.1) and therefore providing no exchangeable hydrogens. The protonation sites of Met-DNA are most likely to be on the nucleobases except for Met-dT_6, which may be protonated on the methylated backbone phosphate group ($P=OH^+$, proton affinity of the double-bonded oxygen might decrease due to the methylation of the phosphate group), or the terminal hydroxyl groups because thymine has very low proton affinity.[34] Figure 3.1 compares the gas-phase HDX behavior of three Met-DNAs with their corresponding natural DNAs. Both Met-d(GCATAC) and Met-dC_6 (Figure 3.1 a and b) show lower HDX rates and exchange to smaller extents than their natural DNAs, indicating that the exchangeable hydrogens on the phosphate hydroxyl groups are involved in natural DNA HDX. The corresponding mechanisms of natural DNA and

Met-DNA are suggested in Scheme 3.2 a and b respectively: for natural DNA, the phosphate hydroxyl group is the hydrogen donor and a neutral nucleobase is the deuterium acceptor; for Met-DNA, a protonated nucleobase is the hydrogen donor and the phosphate group is the deuterium acceptor. Alternatively, natural DNA may undergo the same pathway as Met-DNA if HDX occurs at a protonated nucleobase. However, if the latter pathway is dominant, it would not explain the observed drastically different HDX behavior for natural and methylated DNA. Figure 3.1c shows a comparison of Met-dA_6 with dA_6, which display a different trend compared to the previous two examples: here, Met-dA_6 actually exchanges faster than A_6. This behavior may be understood from the larger size of adenine that makes it possible to form hydrogen bonds with backbone phosphate groups in dA_6 (see inset of Figure 3.1c). Prevention of such hydrogen bonds in Met-dA_6 could result in the observed increase in both HDX rate and extent. Overall, phosphate groups appear to contribute to DNA HDX in two ways, i.e., contribute exchangeable hydrogens to facilitate HDX as well as form hydrogen bonds with purines to prevent HDX.



(a) **(b)**

Scheme 3.1. Structure of natural DNA (a) and DNA in methyl-phosphonate form (Met-DNA) (b).

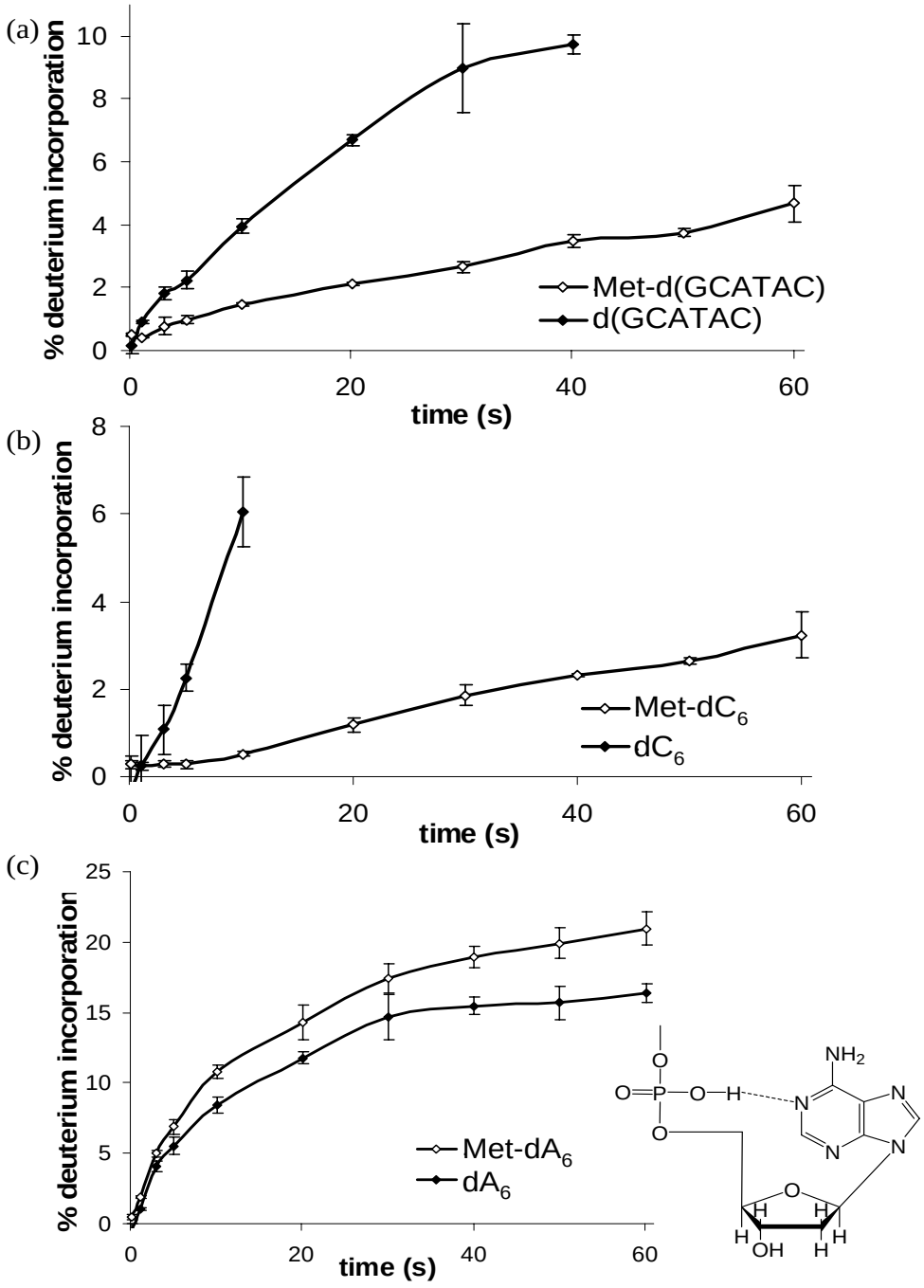
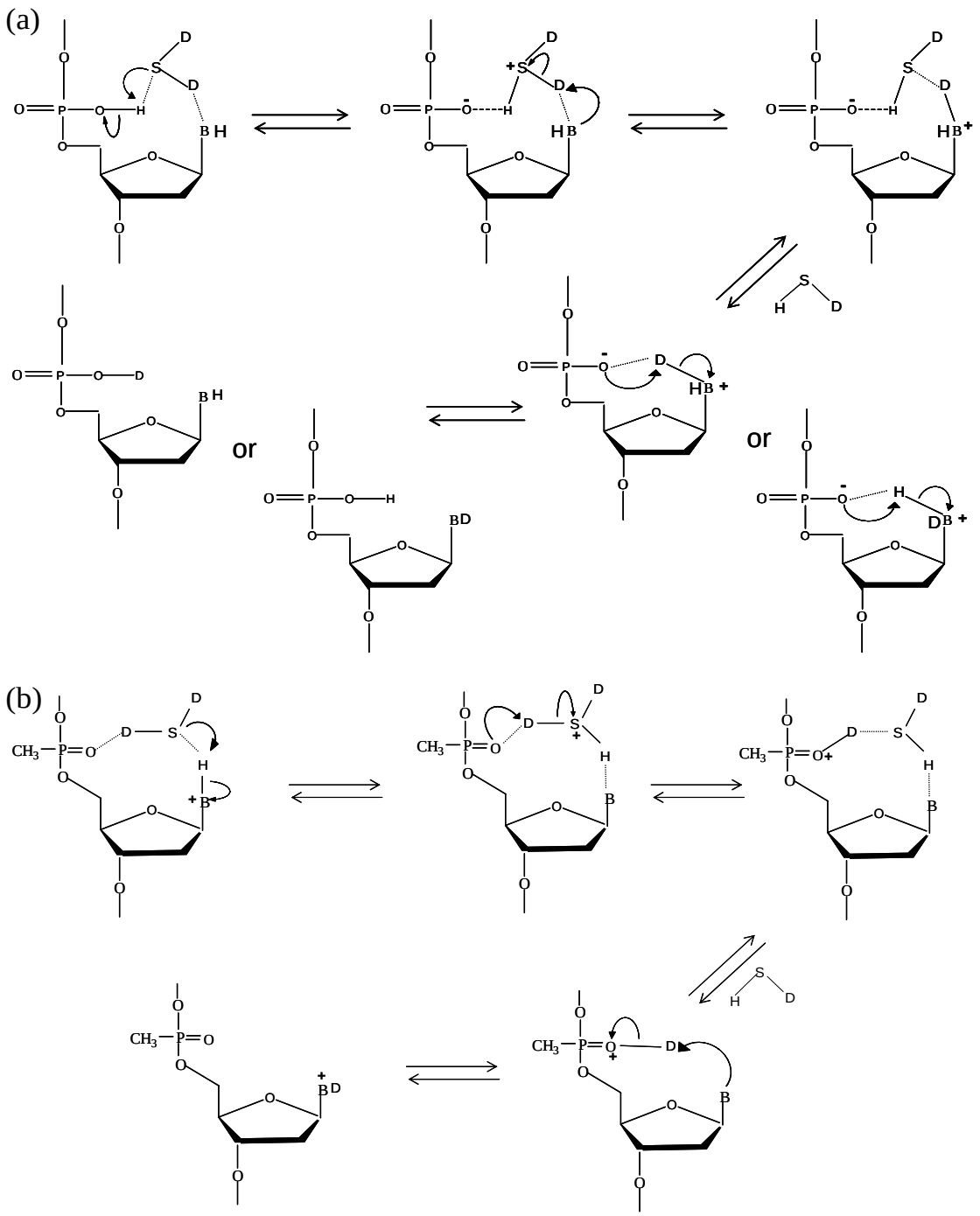


Figure 3.1. HDX of Met-DNAs compared to natural DNAs. Observed rates are Met-d(GCATAC) < d(GCATAC) and Met-dC_6 < dC_6, indicating that phosphate hydroxyl groups are involved in DNA HDX and that removing them results in lower HDX rate and extent. In contrast, Met-dA_6 exchanges faster than dA_6, perhaps due to the large size of adenine, which makes it possible to form hydrogen bonds with backbone phosphate groups in dA_6 (inset), therefore impending HDX.



Scheme 3.2. Proposed nucleic acid gas-phase HDX mechanisms in positive ion mode. In mechanism (a), DNA forms a complex with D_2S, a phosphate hydroxyl group acts as hydrogen donor and a nucleobase is the deuterium acceptor. In mechanism (b), a protonated nucleobase is hydrogen donor and a phosphate P=O group is deuterium acceptor. Natural DNA can undergo both of these two pathways while Met-DNA can only undergo pathway (b), resulting in different HDX behavior.

3.3.3. Role of Nucleobases in Positive Ion Mode DNA HDX

Gas-phase HDX rates of six DNAs with different sequences were compared to characterize the nucleobase effect (Figure 3.2). Observed rates are dA_6 > dC_6 > d(GCATAC) > d(GCATGC) (data for dT_6 could not be acquired due to its low proton affinity and resulting difficulty in generating protonated precursor ions) and 12mer-G close to 12mer-C (these two 12mers were used in order to compare the HDX behavior of cytosine and guanine because dG_6 was difficult to observe, perhaps due to quadruplex formation), suggesting the order of HDX rates for the four DNA nucleobases is A > G ~ C > T, which is consistent with the order of their proton affinities (A > G > C >> T)[38]. HDX of Met-dA_6, Met-dC_6, Met-dG_6, and Met-dT_6 (see Figure 3.3) resulted in a similar trend as that of natural DNAs, i.e., T > A > G ~ C, except for Met-dT_6. The particularly fast HDX process of Met-dT_6 may be due to its different sites of protonation, and an HDX mechanism for Met-dT_6 is proposed in Scheme 3.3. Here, the protonated phosphate group is the hydrogen donor and the nucleobase is the deuterium acceptor.

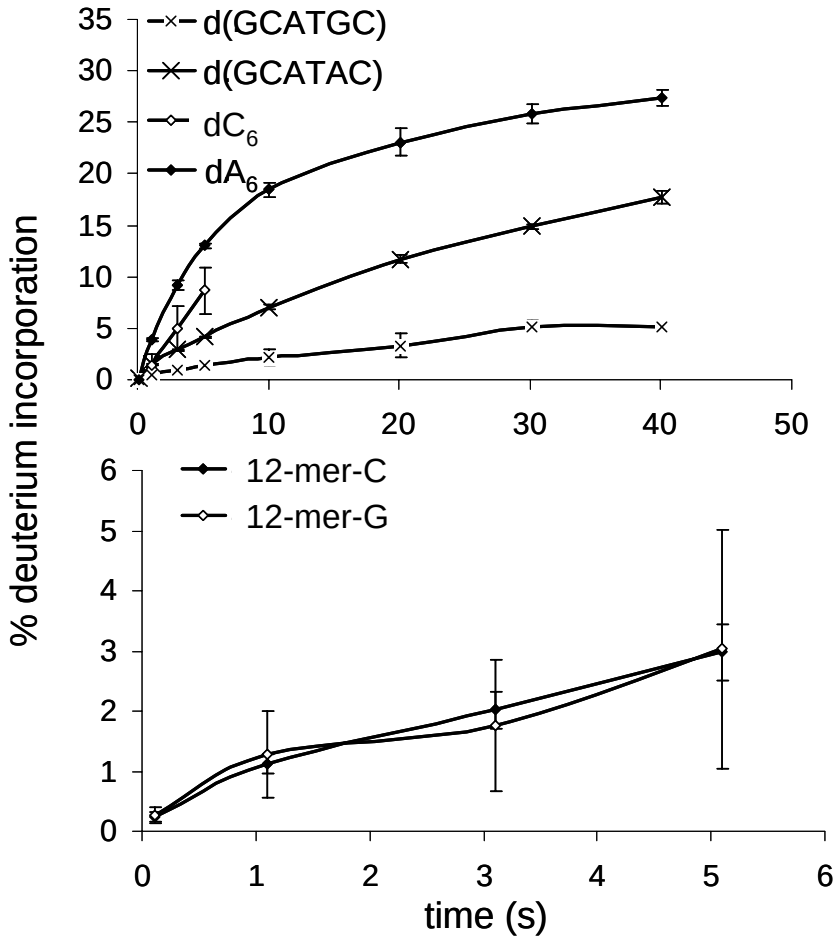


Figure 3.2. Comparison of HDX rates of six DNAs. Observed rates are dA_6 > dC_6 > d(GCATAC) > d(GCATGC) & 12mer-G (d(GGGGATATGGGG))~ 12mer-C(d(CCCCATATCCCC)), i. e., A > G ~ C > T, which is consistent with the proton affinities of the nucleobases.

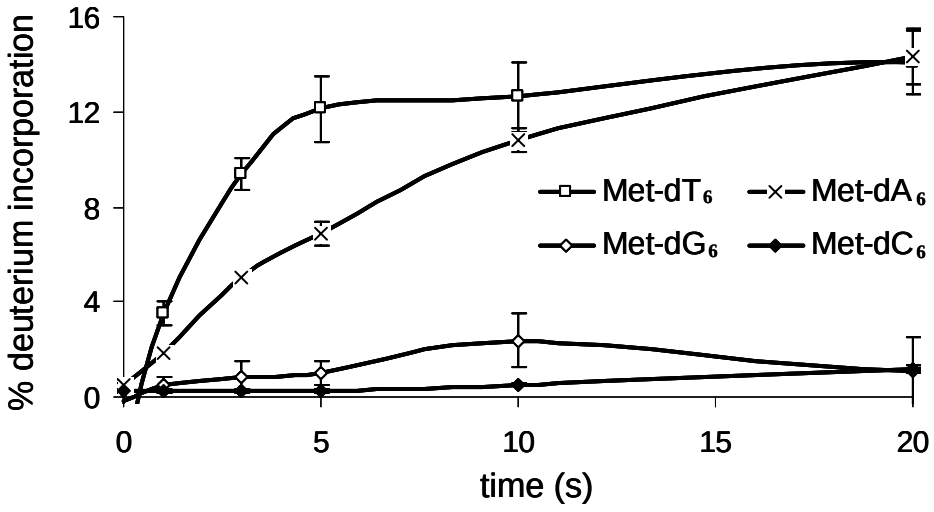


Figure 3.3. Comparison of HDX rates of four Met-DNA hexamers. Observed rates are Met-dT_6 > Met-dA_6 > Met-dG_6 ~ Met-dC_6, i. e., T > A > G ~ C. The unexpectedly fast HDX rate of Met-dT_6 may be due to the different protonation sites compared to other Met-DNAs, resulting in a different HDX pathway.

The correlation between the HDX rates of different nucleobases and their proton affinities indicates that nucleobases are involved in DNA HDX; however, it is not clear yet if they directly participate in HDX, where their exchangeable hydrogens exchange with deuteriums, or just have some structural effect on the complexation of DNA with D_2S. From the percentages of deuterium incorporation of these DNAs, e.g., dA_6 exchanges about five hydrogens after 40 s, which is less than the number of exchangeable hydrogens on the backbone (seven); it appears that the nucleobase exchangeable hydrogens do not necessarily exchange with deuteriums. In order to further investigate the nucleobase behavior, HDX rates of d(GCATAC) with abasic sites was determined (Figure 3.4). Under the same experimental conditions, observed HDX rates follow the trend d(GCATXC) >> d(GCXTAC) ~ d(GCATAC) > d(GXATAC) (X is the abasic site,

i.e., the nucleobase is replaced by hydrogen). The remarkably higher HDX rate of d(GCATXC) compared to the other abasic isomeric oligonucleotides suggests that the presence of adenine close to the 3' terminus of d(GCATAC) has a significant effect in slowing down the HDX process, probably due to a similar effect as mentioned above, i.e., the large size of adenine facilitates formation of hydrogen bonds with the backbone phosphate groups and therefore prevents HDX. However, if the adenine is located in the middle of the DNA (d(GCXTAC) compared with d(GCATAC)), it seems to have a very small effect on the HDX rate. When removing the cytosine close to the 5' terminus of d(GCATAC), both the HDX rate and the exchange extent decrease (d(GXATAC) compared with d(GCATAC)), indicating that the exchangeable hydrogen of this cytosine is directly involved in HDX. This behavior is consistent with the mechanisms suggested in Scheme 3.2 that propose deuterium can eventually reside on nucleobases. Overall, HDX of DNAs is a sequence dependent process and the nucleobases are involved in two ways: directly participating in HDX to increase the rate, which is related to their proton affinities, and by forming hydrogen bonds with the backbone phosphate groups to decrease the HDX rate, which is related to their structures.

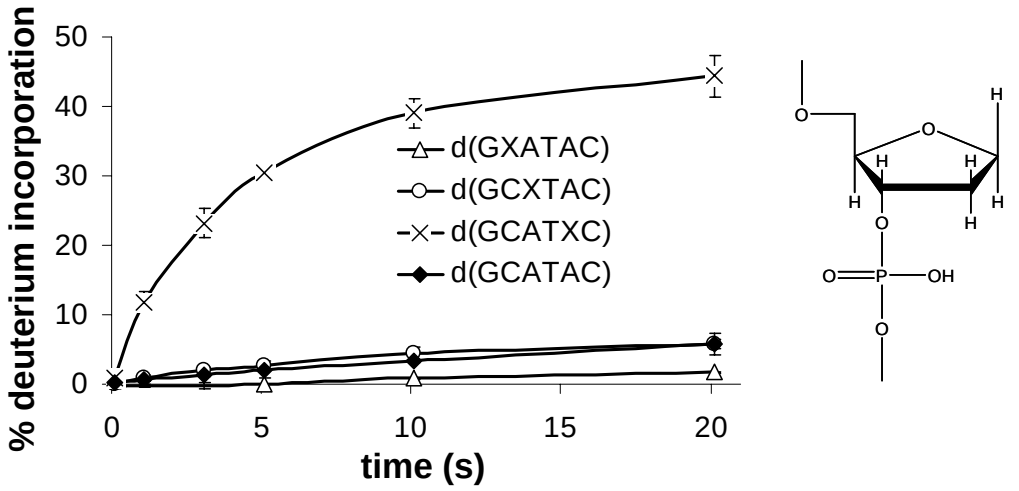
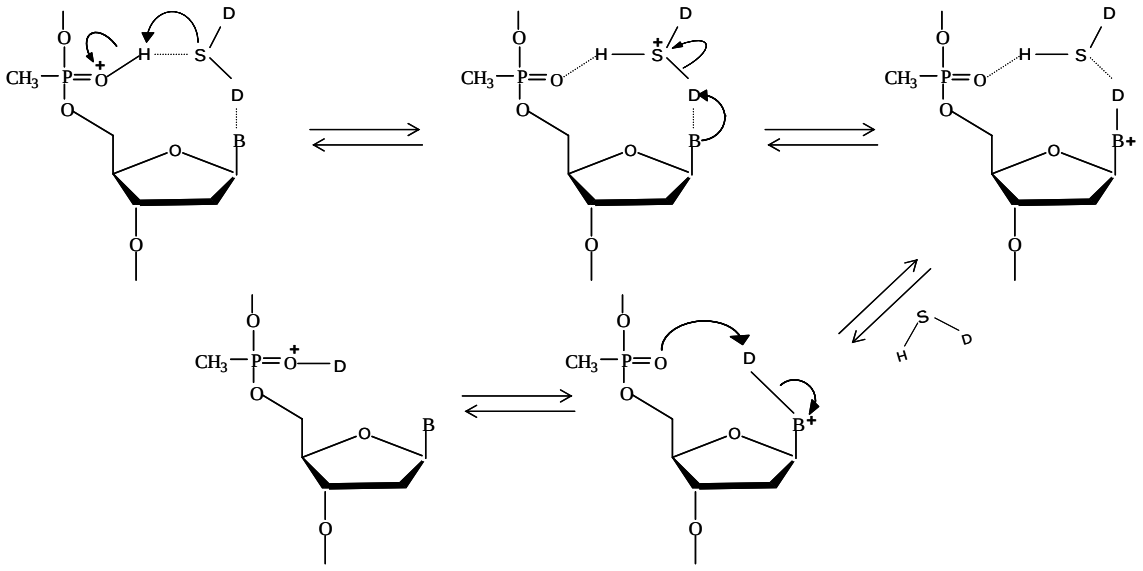


Figure 3.4. Comparison of HDX rates of d(GCATAC) with its abasic variants, i.e., one nucleobase is replaced with hydrogen, see inset. Observed rates are d(GCATXC) >> d(GCXTAC) ~ d(GCATAC) > d(GXCTAC) (X is the abasic site), indicating that both the structure and position of nucleobases have influence on HDX rates.



Scheme 3.3. Proposed relay HDX mechanism for Met-dT_6 in positive ion mode. A protonated backbone phosphate group acts as hydrogen donor when DNA forms complexes with D_2S and a nucleobase is the deuterium acceptor.

3.4. CONCLUSION

Similar to negative ion mode, DNA HDX in positive ion mode appears to be initiated by a complexation step between D_2S and DNA, involving both the nucleobases and the backbone phosphate groups. Both of these two groups influence HDX in two ways: by contributing exchangeable hydrogens for HDX (i.e., having deuterium reside on them after HDX), and by forming hydrogen bonds with each other to prevent HDX. For nucleobases, the former effect is related to their proton affinities and the latter effect is more likely for purines due to their large size. In addition, it seems more significant when purines are located close to the DNA termini.

3.5. BIBLIOGRAPHY

(1) Tanaka, K.; Waki, H.; Ido, Y.; Akita, S.; Yoshida, Y.; Yoshida, T. *Rapid Commun. Mass Spectrom.* **1988**, *2*, 151-153.

(2) Karas, M.; Hillenkamp, F. *Anal. Chem.* **1988**, *60*, 2299-2301.

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

(4) Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F. *Mass Spectrom. Rev.* **1990**, *9*, 37-70.

(5) Kebarle, P. *J. Mass Spectrom.* **2000**, *35*, 804-817.

(6) Tang, K. Q.; Page, J. S.; Smith, R. D. *J. Am. Soc, Mass Spectrom.* **2004**, *15*, 1416–1423.

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

(8) Loo, J. A.; Sannes-Lowery, K. A.; Hu, P.; Mack, D. P.; Mei, H.-Y. *NATO ASI Ser. C* **1998**, *510*, 83-99.

(9) Geels, R. B. J.; van der Vies, S. M.; Heck, A. J. R.; Heeren, M. A. *Anal. Chem.* **2006**, *20*, 7191-7196.

(10) Spelbrink, R. E. J.; Kolkman, A.; Slijper, M.; Killian, J. A.; Kruijff, B. d. *J. Biol. Chem.* **2005**, *280*, 28742-28748.

(11) Freiser, B. S.; Woodin, R. L.; Beauchamp, J. L. *J. Am. Chem. Soc.* **1975**, *97*, 6893 - 6894.

(12) Ranasinghe, A.; Cooks, R. G.; Sethi, S. K. *Org. Mass Spectrom.* **1992**, *27*, 77-88.

(13) Squires, R. R.; Bierbaum, V. M.; Grabowski, J. J.; DePuy, C. H. *J. Am. Chem. Soc.* **1983**, *105*, 5185 - 5192.

(14) Ausloos, P.; Lias, S. G. *J. Am. Chem. Soc.* **1981**, *103*, 3641-3647.

(15) Lias, S. G. *J. Phys. Chem.* **1984**, *88*, 4401 - 4407.

(16) Hunt, D. F.; Sethi, S. K. *J. Am. Chem. Soc.* **1980**, *102*, 6953 - 6963.

(17) Winger, B. E.; Light-Wahl, K. J.; Rockwood, A. L.; Smith, R. D. *J. Am. Chem. Soc.* **1992**, *114*, 5897 - 5898.

(18) Suckau, D.; Shi, Y.; Beu, S. C.; Senko, M. W.; Quinn, J. P.; Wampler FM, I.; McLafferty, F. W. *Proc. Nat. Acad. Sci. U.S.A.* **1993**, *90*, 790-793.

(19) Green, M. K.; Lebrilla, C. B. *Mass Spectrom. Rev.* **1997**, *16*, 53-71.

(20) Freitas, M. A.; Marshall, A. G. *Int. J. Mass Spectrom.* **1999**, *183*, 221-231.

(21) Freitas, M. A.; Hendrickson, C. L.; Emmett, M. R.; Marshall, A. G. *Int. J. Mass Spectrom.* **1999**, *185/186/187*, 565-575.

(22) Brodbelt, J. S. *Mass Spectrom. Rev.* **1997**, *16*, 91-110.

(23) Freitas, M. A.; Hendrickson, C. L.; Emmett, M. R.; Marshall, A. G. *J. Am. Soc. Mass Spectrom.* **1998**, *9*, 1012-1019.

(24) Campbell, S.; Rodgers, M. T.; Marzluff, E. M.; Beauchamp, J. L. *J. Am. Chem. Soc.* **1995**, *117*, 12840 - 12854.

(25) Wyttenbach, T.; Bowers, M. T. *J. Am. Soc. Mass Spectrom.* **1999**, *10*, 9-14.

(26) Robinson, J. M.; Greig, M. J.; Griffey, R. H.; Mohan, V.; Laude, D. A. *Anal. Chem.* **1998**, *70*, 3566-3571.

(27) Griffey, R. H.; Greig, M. J.; Robinson, J. M.; Laude, D. A. *Rapid Commun. Mass Spectrom.* **1999**, *13*, 113-117.

(28) Hofstadler, S. A.; Sannes-Lowery, K. A.; Griffey, R. H. *J. Mass Spectrom.* **2000**, *35*, 62-70.

(29) Freitas, M. A.; Marshall, A. G. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 780-785.

(30) Crestoni, M. E.; Fornarini, S. *J. Mass Spectrom.* **2003**, *38*, 854-861.

(31) Chipuk, J. E.; Brodbelt, J. S. *J. Am. Soc. Mass Spectrom.* **2007**, *18*, 724-736.

(32) Gross, J.; Leisner, A.; Hillenkamp, F.; Hahner, S.; Karas, M.; Schafer, J.; Lutzenkirchen, F.; Nordhoff, E. *J. Am. Soc. Mass Spectrom.* **1998**, *9*, 866-878.

(33) Gabelica, V.; Rosu, F.; Witt, M.; Baykut, G.; de Pauw, E. *Rapid Commun. Mass Spectrom.* **2005**, *19*, 201-208.

(34) Green-Church, K. B.; Limbach, P. A.; Freitas, M. A.; Marshall, A. G. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 268-277.

(35) Mo, J.; Hakansson, K. *Anal. Chem.* **2007**, *79*, 7893-7898.

(36) Mo, J.; Hakansson, K. *Anal. Bioanal. Chem.* **2006**, *386*, 675-681.

(37) Yang, J.; Mo, J.; Adamson, J. T.; Hakansson, K. *Anal. Chem.* **2005**, *77*, 1876-1822.

(38) Green-Church, K. B.; Limbach, P. A. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 24-32.

CHAPTER 4

CHARACTERIZATION OF NUCLEIC ACID HIGHER ORDER STRUCTURE BY HIGH RESOLUTION TANDEM MS

MS is extensively used for the identification and sequencing of nucleic acids but has so far seen limited use for characterization of their higher order structure. In this Chapter, we have applied a range of different tandem MS techniques, including EDD, IRMPD, activated ion (AI) EDD, and EDD/IRMPD MS^3, in an FT-ICR mass spectrometer to the characterization of three isomeric 15-mer DNAs with different sequences and predicted solution-phase structures. Our goal was to explore whether their structural differences could be directly probed with these techniques. We found that all three 15-mers had higher order structures in the gas phase although preferred structures were only predicted for two of them in solution. Nevertheless, EDD, AI EDD, and EDD/IRMPD MS^3 experiments yielded different cleavage patterns with less backbone fragmentation for the more stable solution-phase structure as compared to the other two 15-mers. By contrast, no major differences were observed in IRMPD although the extent of backbone cleavage was higher with that technique for all three 15-mers. Thus, experiments utilizing the radical ion chemistry of EDD can provide complementary structural information compared to traditional slow heating methods, such as IRMPD, for structured nucleic acids.

4.1. Introduction

Hairpins represent the dominant secondary structure element in RNA. Stable RNA hairpins define nucleation sites for folding,[1] determine tertiary interactions in RNA enzymes,[2, 3] protect mRNAs from degradation,[4, 5] are recognized by RNA-binding proteins, and are involved in mRNA localization,[6] retroviral encapsidation and packaging.[7, 8] Because only a few RNA structures have been solved, determination of RNA hairpin motifs can provide valuable information for RNA folding and structural prediction.[9] Conformational changes involving hairpins can be characterized by NMR spectroscopy or FRET techniques employing fluorescent labels[10-12]. However, the former technique requires large sample quantities and has an upper mass limit whereas the latter only provides distance constraints in a limited region of the molecule and, thus, does not provide information on the overall structure.

In this Chapter, we attempt to apply mass spectrometric approaches for characterizing nucleic acid higher order structure. Tandem MS techniques, accomplished via CAD[13] or IRMPD,[14] have been employed for probing the extent of hydrogen bonding in DNA duplexes[15] and the binding sites of high affinity RNA ligands.[16] However, to our knowledge, it has not been employed to probe the higher order structure of RNA or DNA hairpins. ECD[17-20] and EDD[21, 22] are two rather recent fragmentation techniques that involve a radical ion intermediate, produced via electron attachment to multiply charged cations (ECD) or electron removal from multiply charged anions (EDD). These two techniques have been proposed to be able to cleave backbone covalent bonds in gas-phase biomolecular ions while preserving the non-covalent interactions that determine their higher order structure.[22, 23] The combination of ECD with gas-phase unfolding

through IR heating (AI ECD) has been employed to probe the gas-phase structure of protein cations.[24] In this Chapter, we sought to apply the combination of EDD and IRMPD to differentiate between three 15-mer DNA sequences with or without a preferred solution-phase structure. Because EDD operates in negative ion mode, it is better suited than ECD for analysis of nucleic acids, which undergo facile deprotonation at the phosphate backbone and therefore yields higher molecular ion signal in negative mode ESI.[25]

4.2. Experimental Section

4.2.1. Sample Preparation

The oligodeoxynucleotides d(TCTAACCTGATGATG) (15-mer 0), d(CTATCACTGATAGGT) (15-mer 1), and d(CTATCACTGGATAGT) (15-mer 2) were purchased from TriLink BioTechnologies (San Diego, CA) as their crude ammonium salts and used without further purification. These oligonucleotides have the same base composition (i.e., identical molecular weight) but differ in sequence and stability. Their ΔGs as calculated by the Mfold Web Server (37° C, 1 M Na^+)[26] are: >0 (15-mer 0), -1.0 kcal/mole (15-mer 1), and -1.8 kcal/mole (15-mer 2). The DNAs were dissolved in 95:5 HPLC grade water:isopropanol (Fisher, Fair Lawn, NJ) with 10 mM ammonium acetate (Fisher) to a concentration of 10 μM.

4.2.2. Mass Spectrometry

All experiments were conducted on an actively shielded 7 T Q-FT-ICR mass spectrometer (Bruker Daltonics, Billerica, MA) described in Chapter 1. The oligonucleotide solutions were infused via an external Apollo electrospray ion source at a

flow rate of 60-80 μL/h with the assistance of N_2 nebulizing gas. Ionization was performed in negative mode with an ESI voltage of 4200 V. Care was taken to ensure that the source transfer conditions were as soft as possible (i.e. the source voltages were kept as low as possible while still allowing efficient ion transfer. More details are provided in the Appendix). Specifically, the capillary exit was kept at - 50 V, skimmer 1 at - 40 V, and skimmer 2 at - 6 V. Following ESI, ions were mass-selectively externally accumulated (4-8 m/z isolation window) for 1-4 s, transferred through high-voltage ion optics, and captured by gated trapping in an infinity ICR cell. This accumulation/transfer/trapping sequence was looped twice (six times for MS^3) for improved signal-to-noise ratio. All mass spectra were acquired with XMASS (version 6.1, Bruker Daltonics) in broadband mode from m/z 200 to 2,500 with 512k data points. Data processing, including Hanning apodization and one zero fill, was performed with the MIDAS analysis software.[27]

For EDD experiments, electrons from an indirectly heated hollow dispenser cathode[28] (inner and outer diameters of 3.5 and 7.6 mm, respectively, located 88 mm from the ICR cell) were introduced into the ICR cell for 1 s. A heating current of 1.8 A was applied to a heater element located behind the cathode and the bias voltage was - 18 V. A lens electrode (6 mm inner diameter) located immediately in front of the cathode was kept at - 19 V. IRMPD was performed with a vertically mounted 25 W, 10.6 μm, CO_2 laser (Synrad, Mukilteo, WA). The laser beam is deflected by two mirrors for alignment through the hollow dispenser cathode to the center of the ICR cell. The beam enters the vacuum system through a BaF_2 window. Photon irradiation was performed at 10% laser power. The irradiation time (30-45 ms) was adjusted to yield as many product

ions as possible but still maintain the precursor ion species as the most abundant peaks in the spectra. For AI EDD, precursor ions were heated by the IR laser to unfold their higher order structure but without generating any backbone product ions (10% laser power, 18-22 ms irradiation time) and electron irradiation (- 18 V, 1 s) was applied immediately following the laser pulse. For EDD/IRMPD MS^3 experiments, EDD (- 18 V, 1 s) was performed to generate charge reduced radical ions, which were isolated in the cell by correlated harmonic excitation fields (CHEF)[29]) and then subjected to IRMPD (10% laser power, 100-200 ms irradiation time). Tandem mass spectra were summed over 16 scans for EDD and IRMPD, 32 scans for AI EDD and 64 scans for MS^3. All experiments were repeated three to five times on different days to allow error analysis. Product ion abundances were normalized to their charge states, i.e. divided by their charge, because FT-ICR signal is proportional to charge.

4.3. Results and Discussion

Solution-phase structures of the three isomeric 15-mer DNAs were predicted by the Mfold web server (37° C, 1 M Na^+). 15-mer 1 and 15-mer 2 both have a major solution-phase structure corresponding to a hairpin with a three- or four-nucleotide loop and a two- or one-nucleotide dangling 3' end, respectively, whereas no favored solution-phase structure was found for 15-mer 0 (see Schemes 4.1-4.3). DNA hairpin loops of four to five nucleotides usually have higher stability than smaller or larger loops.[30-32] Consistently, ΔG of 15-mer 2 is more negative (- 1.8 kcal/mol) than that of 15-mer 1 (- 1.0 kcal/mol), whereas 15-mer 0 has a positive ΔG.

4.3.1. EDD of Three 15-mer DNAs

Figure 4.1 shows the EDD spectra of the 4^- charge state of the three 15-mer DNAs. This charge state was selected for fragmentation because it was the most abundant following ESI. For 15-mer 1 and 2, which have solution-phase hairpin structures, only charge reduction to form a triply charged radical ion ($[M - 4H]^{3-\bullet}$) and a doubly charged biradical ($[M - 4H]^{2-\bullet\bullet}$, this species can recombine to form an even-electron ion[33]) was observed (Figure 4.1 a and b). This behavior is consistent with the proposed ability of EDD to preserve non-covalent interactions.[22] If we assume that some higher order structure survives into the gas phase (as previously suggested[15]), the W-C hydrogen bonds involved in the hairpin structures of 15-mer 1 and 2 are not disrupted and, therefore, no product ions can be observed even though backbone covalent bonds may have been cleaved. By contrast, two product ions (w_{14} and d_{14} (McLuckey nomenclature[34])) formed from cleavage of the phosphodiester backbone between the two first and two last nucleotides, respectively, were observed in EDD of 15-mer 0 (Figure 4.1c and inset). Such fragments were anticipated because 15-mer 0 does not have a preferred solution-phase structure and was therefore not expected to be prevented from dissociating in EDD. Traditional MS/MS techniques, including CAD and IRMPD, generally only provide (*a* - B) (B = nucleobase) and *w*-type backbone product ions.[35, 36] However, we have demonstrated that *d*-type ions are commonly observed in EDD.[22, 37] Nevertheless, the number of product ions (only two) was much lower than expected in EDD of 15-mer 0. Our previous work on hexamer DNAs[22] and RNAs[38] did not reveal any significant oligonucleotide sequence-dependence in EDD. Thus, the observed low degree of fragmentation is not likely attributable to the particular sequence of 15-mer 0.

This behavior suggests that, although unstructured in solution, 15-mer 0 is folded in the gas phase. Such intramolecular charge solvation has been observed previously, both in gas-phase hydrogen exchange experiments and theoretical studies.[39-41] Thus, EDD may not provide data directly related to solution-phase behavior but still shows promise for interrogating the gas-phase structure of nucleic acids.

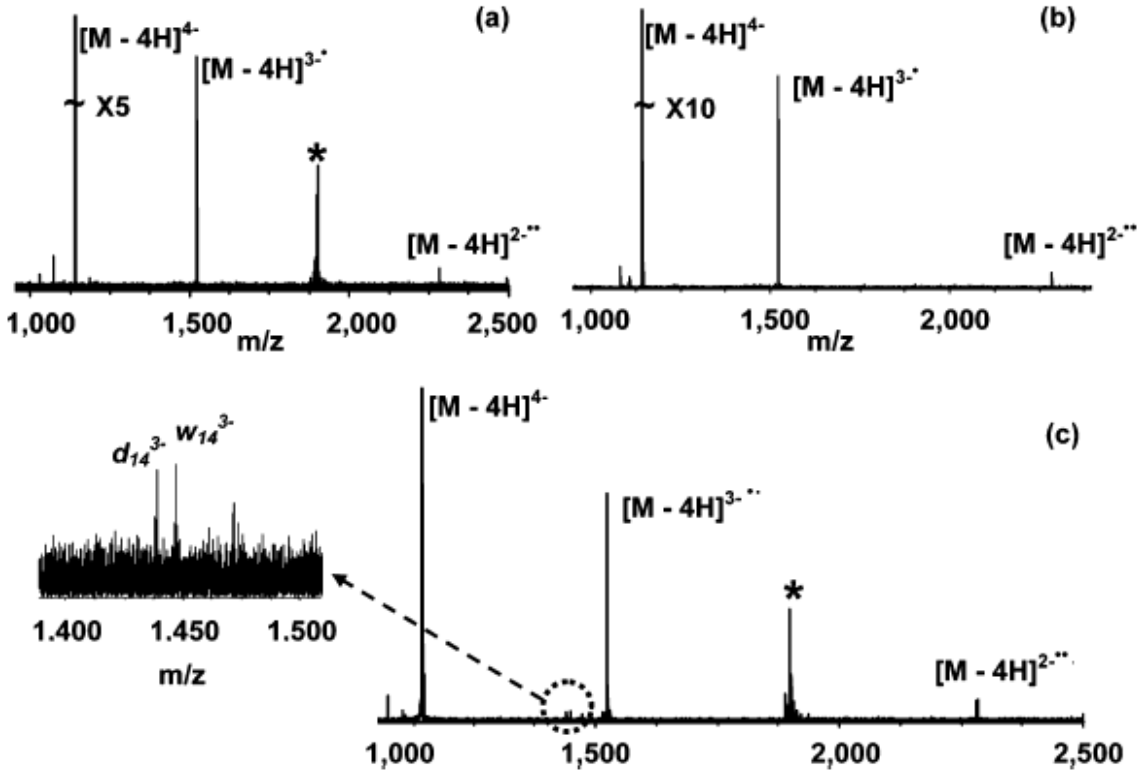
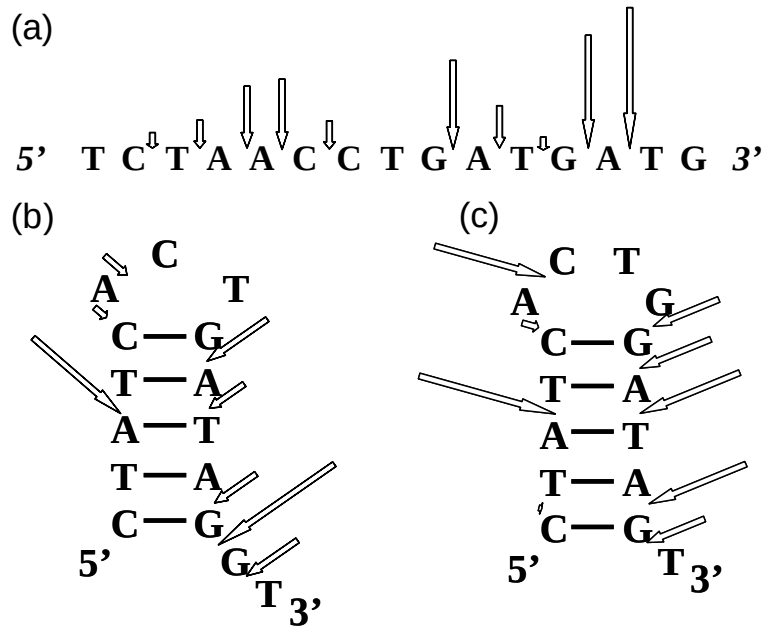


Figure 4.1. EDD of three 15-mer isomeric DNAs. Only charge reduction was observed for 15mer-1 (a) and 2 (b), which have predicted preferred solution-phase structures. Two backbone product ions (inset) were observed for 15mer-0 (c), which does not have a preferred solution-phase structure. Asterisks denote noise peaks.

4.3.2. IRMPD of Three 15-mer DNAs

Gabelica *et al.*[42] have shown that the amount of collisional energy required to dissociate a DNA duplex-drug anionic complex to 50% of its original abundance can be related to the binding strength of the complex. Because IRMPD dissociates nucleic acids through a similar mechanism as CAD (vibrational excitation followed by nucleobase loss and subsequent backbone cleavage[43, 44]), it should theoretically be possible to relate the laser fluence needed to deplete the 15mers to 50% of their original value to the number of hydrogen bonds and thereby, indirectly, to their gas-phase structure. However, in our set-up, the laser beam diameter is limited due to the necessity of passing through the hollow cathode. Thus, it is almost certain that there is only partial overlap between the laser beam and ion cloud, which results in inconclusive data. Figure 4.2 shows the backbone (including mainly *w* and (*a* - B) ions but also some (*c* - B), (*x* - B), (*w* - B) and (*z* - B) ions) and nucleobase loss cleavages observed in IRMPD of the three 15-mers. These data have been normalized to the charge state of the product ions (as discussed in the experimental section). Very limited cleavage is seen 5' to thymidine residues (which have very low proton affinity[45]) for all three 15-mers, consistent with a cleavage mechanism involving proton transfer to the nucleobase as the initiating step.[43, 44] For 15-mer 1 and 2, the cleavages 5' to cytidine within the hairpin loops are not observed. However, such a cleavage is also absent from the IRMPD spectrum of 15-mer 0. Thus, the IRMPD data (summarized in Scheme 4.1 for all three 15-mers) do not reveal any significant differences in higher order structure. However, that result is expected because IRMPD should first unfold any higher order structure prior to causing backbone cleavage due to the slow heating involved.



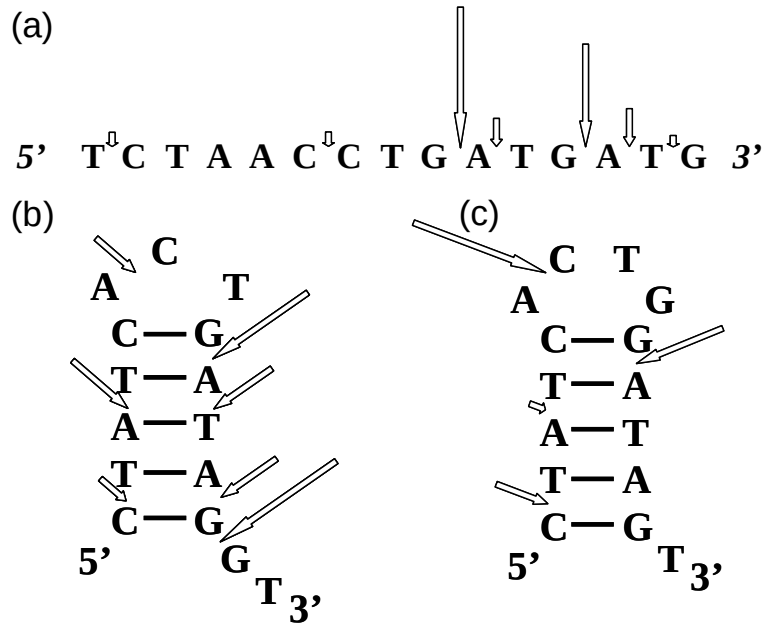
Scheme 4.1. Backbone cleavages (including mainly *w* and (*a* - B) ions but also some (*c* - B), (*x* - B), (*w* - B) and (*z* - B) ions) observed in IRMPD of 15mer-0 (a), 15-mer-1 (b), and 15mer-2 (c). The lengths of the arrows are scaled to the relative product ion abundance, normalized to charge.

4.3.3. Activated Ion EDD of Three 15-mer DNAs

For AI EDD, the 15-mers were first subjected to a short IR laser pulse, insufficient to cause backbone cleavage but likely causing some unfolding of the DNAs prior to the EDD event. Backbone (including *w*, (*a* - B) and a few *d* ions) and nucleobase cleavages (normalized to charge) observed from this approach are shown in Figure 4.3 for all three 15-mers. As expected, more product ions were observed in AI EDD compared to EDD alone (Figure 4.1). However, the AI EDD data were different from the IRMPD data in several respects. First, nucleobase loss was more dominant than in IRMPD alone (Figure 4.2), which can be explained either by the lower degree of IR heating, channeling the fragmentation into these particular pathways, or by the different

cleavage mechanism in EDD. Our previous work for smaller oligonucleotides[22] did not show abundant nucleobase loss. However, those oligonucleotides were smaller than the 15-mers characterized here and therefore were not likely to have significant higher order structure. Alternatively, the higher degree of base loss for 15-mers compared to hexamers could be due to the larger number of bases present, channeling into the same number of fragmentation pathways. Interestingly, the relative abundance of the various nucleobase losses (i.e. loss of adenine vs. loss of guanine vs. loss of cytosine) was different in the three cases although the base composition of the three 15-mers is identical. For 15-mer 0, guanine loss dominated over adenine loss although there are three guanosine residues and four adenosine residues. However, that behavior is consistent with previous EDD data, which showed more abundant nucleobase loss for dG_6 compared to dC_6, dA_6, and dT_6.[22] The dominance of adenine loss over guanine loss for 15-mer 1 and 2 may result from different gas-phase structures as compared to 15-mer 0, based on our previous observations that there is no significant sequence-dependence in EDD.[22, 38] Guanine loss is more abundant for 15-mer 1 than for 15-mer 2, which may be due to the dangling guanosine residue at the 3' end. The second major difference between the IRMPD and AI EDD data is the extent of backbone cleavage. In IRMPD, the cleavage pattern was very similar for all three 15-mers (due to the unfolding followed by backbone cleavage as discussed above). By contrast, 15-mer 2 stands out in AI EDD due to the rather limited extent of cleavage: only four 3' product ions were detected. This limited fragmentation may be due to insufficient laser fluence to efficiently unfold this 15-mer prior to EDD, consistent with its higher solution-phase stability. The cleavage patterns for 15-mer 0 and 1 were rather similar, indicating that their gas phase structures

have similar stability. This conclusion may be justified by 15-mer 1 having a three nucleotide loop, which is known to be a less stable structure than a hairpin with a four nucleotide loop, such as the one in 15-mer 2.[30-32] Thus, the amount of energy required to unfold 15-mer 0 and 1 may be similar although their gas phase structures likely are rather different. The backbone cleavages observed in AI EDD of the three 15-mers are summarized in Scheme 4.2.

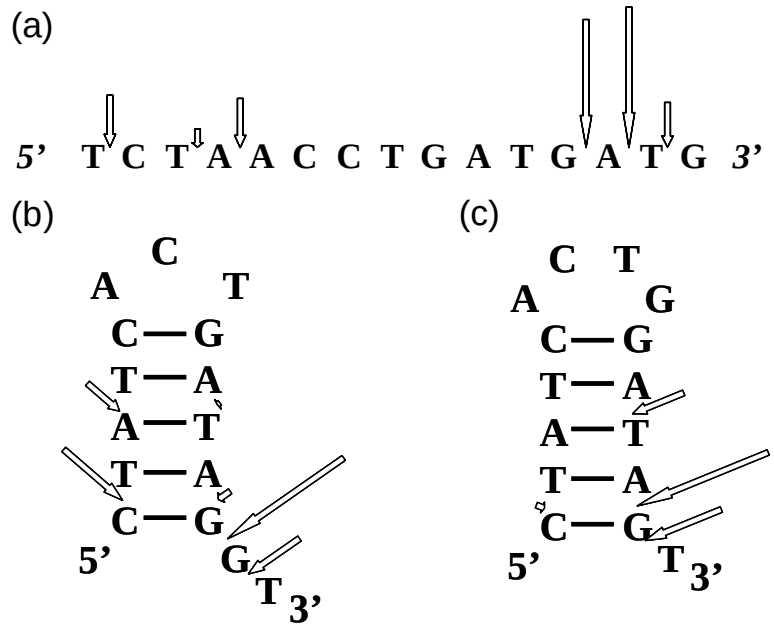


Scheme 4.2. Backbone cleavages (including *w*, (*a* - B) and a few *d* ions) observed in AI EDD of 15mer-0 (a), 15-mer-1 (b), and 15mer-2 (c). The lengths of the arrows are scaled to the relative product ion abundance, normalized to charge.

4.3.4. EDD/IRMPD MS^3 of Three 15-mer DNAs

In MS^3 experiments, EDD of the 4^- charge state of the three 15-mers was first performed to generate the charge reduced species $[M - 4H]^{3-\bullet}$ and $[M - 4H]^{2-\bullet\bullet}$ (see Figure 4.1). The radical species $[M - 4H]^{3-\bullet}$ was then isolated in the cell with a CHEF waveform

and fragmented by IRMPD. For this fragmentation event, longer IR photon irradiation time was generally required compared to direct IRMPD of the even-electron $[M - 4H]^{4-}$ precursor ions (Figure 4.2). This behavior can be explained by the summed EDD and CHEF events, which energize the $[M - 4H]^{3-\bullet}$ ions prior to IRMPD and therefore move them off-axis, out of the area covered by the centrally located IR laser beam. The backbone (including *w*, *d* and (*a* - B) ions and a few *a* radical ions) and nucleobase loss cleavages observed in EDD/IRMPD MS^3 (normalized to charge) are shown in Figure 4.4 for all 15-mers. These data are more similar to the IRMPD data (Figure 4.2) than to the AI EDD data (Figutre 4.3). The main difference compared to IRMPD is that backbone cleavages from the central portions of the 15-mers are lacking, indicating that the IR laser fluence was not sufficient to completely unfold the 15-mers, probably due to poor overlap with the laser beam as discussed above. However, the data confirms that EDD by itself does not disrupt hydrogen bonding within these 15mers. As for the AI EDD experiments, the data for 15-mer 1 and 2 are rather similar whereas 15-mer 2 once again stands out in that less backbone fragments are observed, confirming its more stable structure. The backbone cleavages observed in EDD/IRMPD MS^3 of the three 15-mers are summarized in Scheme 4.3.



Scheme 4.3. Backbone cleavages (including *w*, *d* and (*a* - B) ions and a few *a* radical ions) observed in EDD/IRMPD MS^3 of 15mer-0 (a), 15-mer-1 (b), and 15mer-2 (c). The lengths of the arrows are scaled to the relative product ion abundance, normalized to charge.

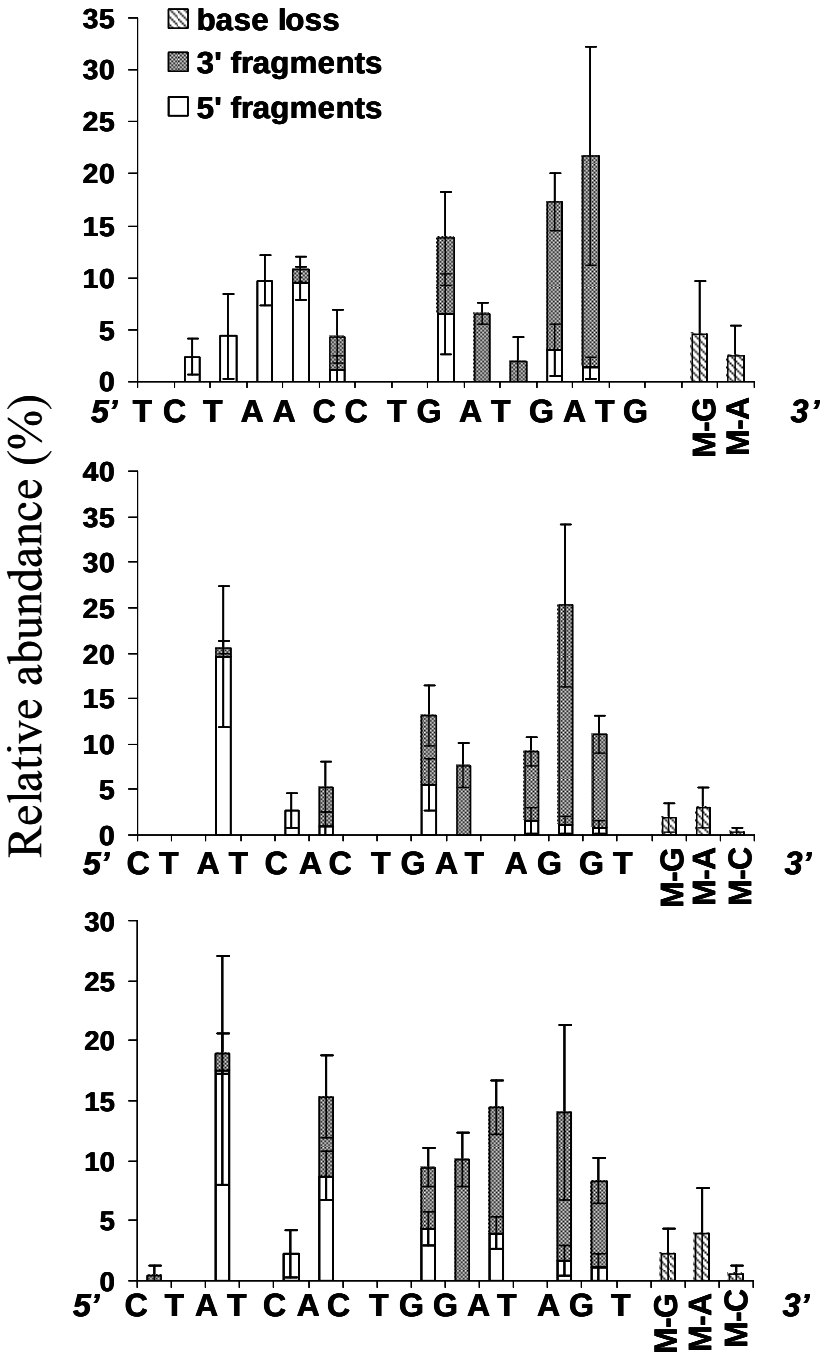


Figure 4.2. Backbone (including mainly *w* and (*a* - B) ions but also some (*c* - B), (*x* - B), (*w* - B) and (*z* - B) ions) and nucleobase cleavages (normalized to charge) observed following IRMPD of 15mer-0 (a), 15mer-1 (b), and 15mer-2 (c).

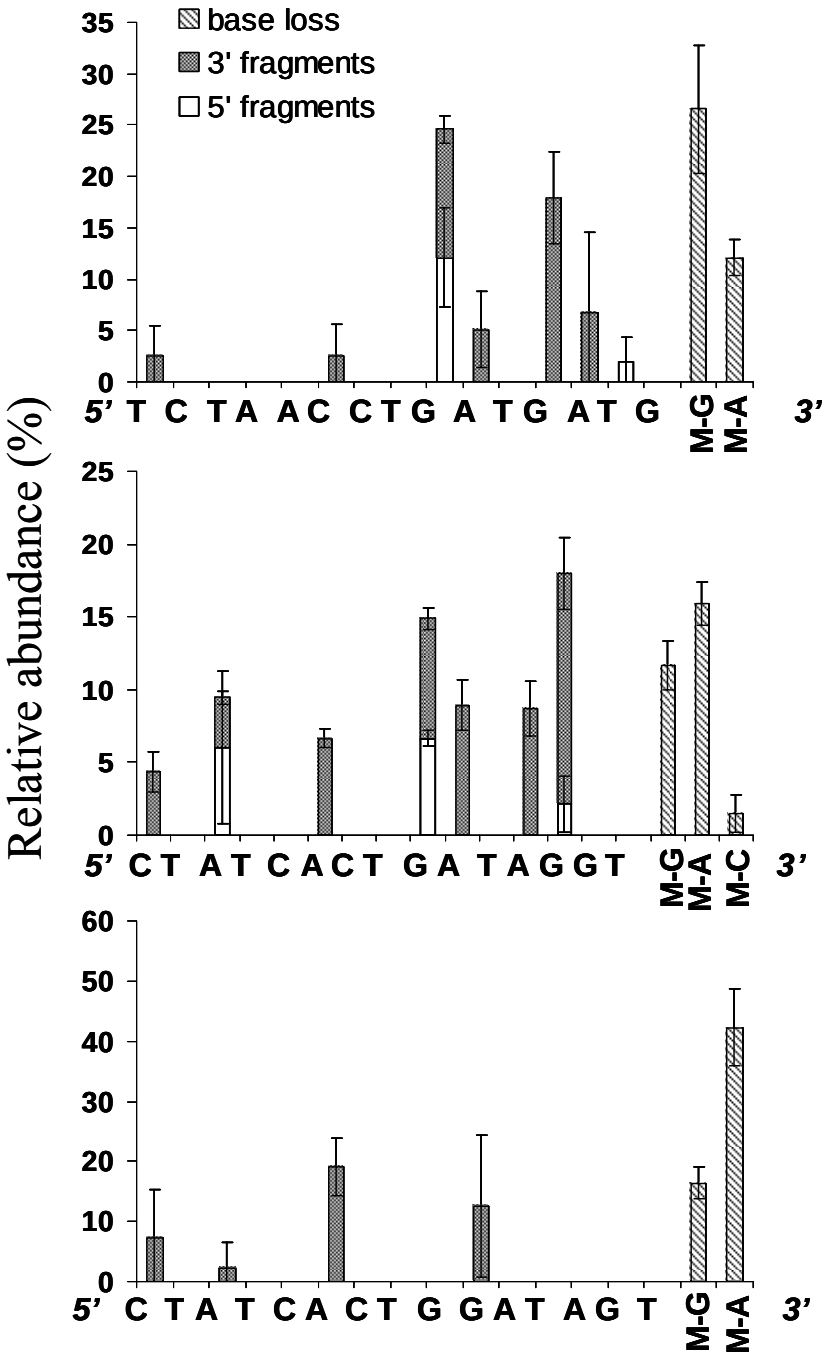


Figure 4.3. Backbone (including *w*, (*a* - B) and a few *d* ions) and nucleobase cleavages (normalized to charge) observed following AI EDD of 15mer-0 (a), 15mer-1 (b), and 15mer-2 (c).

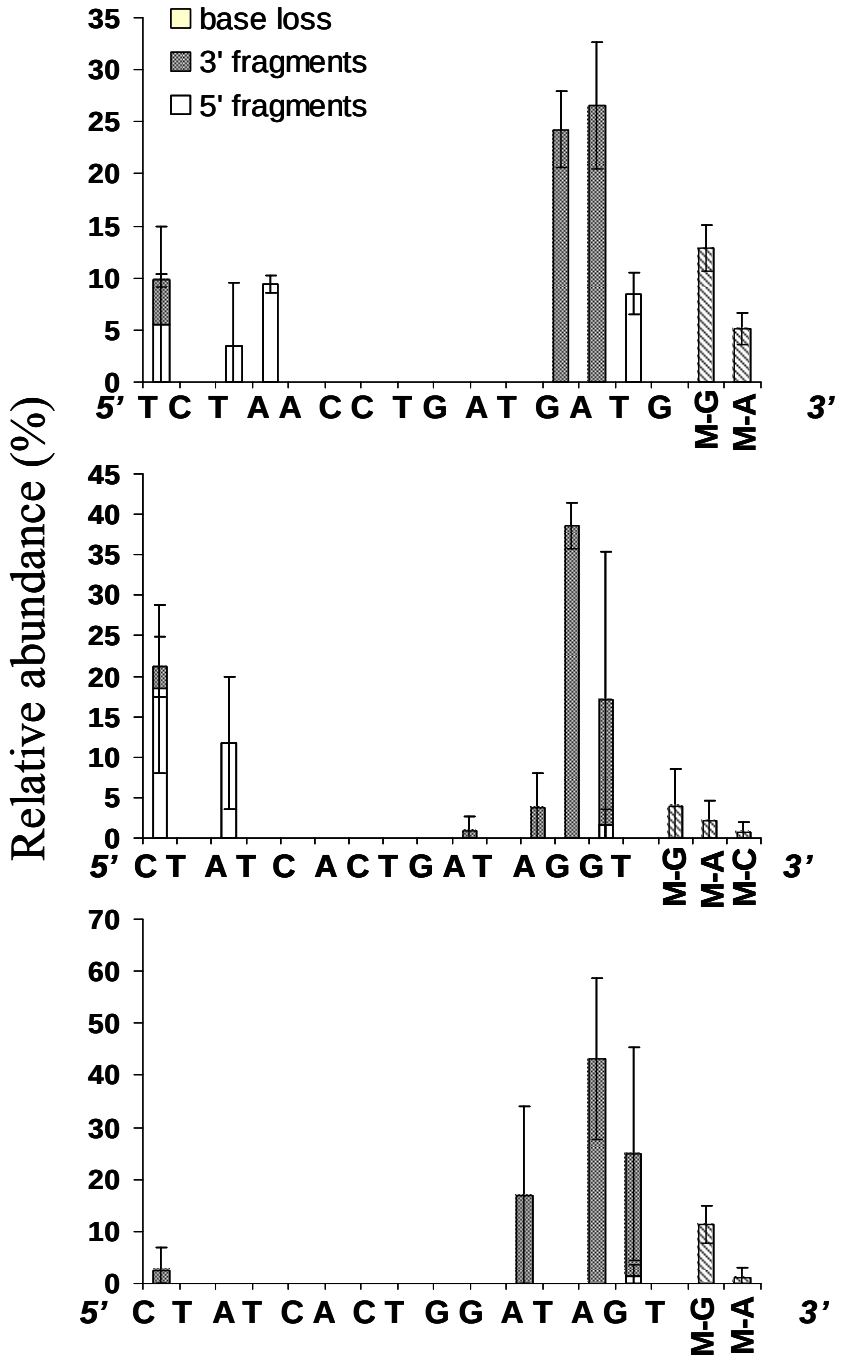


Figure 4.4. Backbone (including *w*, *d* and (*a* - B) ions and few *a* radical ions) and nucleobase cleavages (normalized to charge) observed following EDD/IRMPD MS^3 of 15mer-0 (a), 15mer-1 (b), and 15mer-2 (c).

4.4. Conclusion

Our results from EDD, AI EDD and EDD/IRMPD MS^3 experiments suggest that the three 15-mer DNAs investigated here all have higher order structures in the gas phase, contrary to their predicted solution-phase characteristics with higher order structure only predicted for 15-mer 1 and 2. However, the presented experiments demonstrate that the presence of higher order structures can be determined by these fragmentation techniques as well as some information on their stability. 15-mer 2 generated less product ions in both AI EDD and EDD/IRMPD MS^3, consistent with its more stable solution-phase structure. However, IRMPD did not differentiate between the three 15-mers although that technique would be preferred for sequencing purposes because backbone cleavage was observed to a much higher extent. Thus, as previously stated,[21, 22] EDD is a valuable complementary fragmentation method compared to traditional MS/MS techniques, including IRMPD, in that it can generate additional structural information.

4.5. BIBLIOGRAPHY

(1) Tuerk, C.; Gauss, P.; Thermes, C.; Groebe, D.; Gayles, M.; Guild , N.; Stormo, G.; d'Aubenton-Carafa, Y.; Uhlenbeck, O.; Tinoco, I.; Brody, E.; Gold, L. *Proc. Nat. Acad. Sci. U.S.A.* **1988**, *85*, 1364-1368.

(2) James, J.; I Tinoco, J. *Nucleic Acids Res.* **1993**, *21*, 3287-3293.

(3) Panayotatos, N.; Wells, R. *Nature* **1981**, *289*, 466-470.

(4) Adams, C. C.; Stern, D. B. *Nucleic Acids Res.* **1990**, *18*, 6003-6010.

(5) Klausner, R.; Rouault, T.; Harford, J. *Cell* **1993**, *72*, 19-28.

(6) Gottlieb, E. *Proc. Nat. Acad. Sci. U.S.A.* **1992**, *89*, 7164-7168.

(7) Skripkin, E.; Paillart, J.-C.; Marquet, R.; Ehresmann, B.; Ehresmann, C. *Proc. Nat. Acad. Sci. U.S.A.* **1994**, *91*, 4945-4949.

(8) Yang, S.; Temin, H. *EMBO J.* **1994**, *13*, 713-726.

(9) Gautheret, D.; Major, F. *J. Mol. Biol.* **1993**, *229*, 1049-1064.

(10) Al-Hashimi, H. M. *ChemBioChem* **2005**, *6*, 1472-1473.

(11) Klostermeier, D.; Millar, D. P. *Methods* **2001**, *23*, 240-254.

(12) Latham, M. P.; Brown, D. J.; McCallum, S. A.; Pardi, A. *ChemBioChem* **2005**, *6*, 1492-1505.

(13) Hayes, R. N.; Gross, M. L. *Methods Enzymol.* **1990**, *193*, 237-263.

(14) Little, D. P.; Speir, J. P.; Senko, M. W.; O'Connor, P. B.; McLafferty, F. W. *Anal. Chem.* **1994**, *66*, 2809-2815.

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

(16) Griffey, R. H.; Hofstadler, S. A.; Sannes-Lowery, K. A.; Ecker, D. J.; Crooke, S. T. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 10129-10133.

(17) Cooper, H. J.; Hakansson, K.; Marshall, A. G. *Mass Spectrom. Rev.* **2005**, *24*, 201-222.

(18) McLafferty, F. W.; Horn, D. M.; Breuker, K.; Ge, Y.; Lewis, M. A.; Cerda, B.; Zubarev, R. A.; Carpenter, B. K. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 245-249.

(19) Zubarev, R. A. *Mass Spectrom. Rev.* **2003**, *22*, 57-77.

(20) Zubarev, R. A. *Curr. Opin. Biotechnol.* **2004**, *15*, 12-16.

(21) Budnik, B. A.; Haselmann, K. F.; Zubarev, R. A. *Chem. Phys. Lett.* **2001**, *342*, 299-302.

(22) Yang, J.; Mo, J.; Adamson, J. T.; Hakansson, K. *Anal. Chem.* **2005**, *77*, 1876-1822.

(23) Horn, D. M.; Ge, Y.; McLafferty, F. W. *Anal. Chem.* **2000**, *72*, 4778-4784.

(24) Horn, D. M.; Breuker, K.; Frank, A. J.; McLafferty, F. W. *J. Am. Chem. Soc.* **2001**, *123*, 9792-9799.

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

(26) Zuker, M. *Nucleic Acids Res.* **2003**, *31*, 3406-3415.

(27) Senko, M. W.; Canterbury, J. D.; Guan, S.; Marshall, A. G. *Rapid Commun. Mass Spectrom.* **1996**, *10*, 1839-1844.

(28) Tsybin, Y. O.; Witt, M.; Baykut, G.; Kjeldsen, F.; Hakansson, P. *Rapid Commun. Mass Spectrom.* **2003**, *17*, 1759-1768.

(29) de Koning, L. J.; Nibbering, N. M. M.; van Orden, S. L.; Laukien, F. H. *Int. J. Mass Spectrom.* **1997**, *165*, 209-219.

(30) Blommers, M.; Walters, J.; Haashoot, C.; Aelen, J.; van der Marel, G.; Van Boom, J.; Hilbers, C. *Biochemistry* **1989**, *28*, 7491-7498.

(31) Haasnoot, C.; de Bruin, S.; Berendsen, R.; Janssen, H.; Binnendijk, T.; Hilbers, C.; van der Marel, G.; van Boom, J. *J Biomol Struct Dyn.* **1983**, *1*, 115-129.

(32) Haasnoot, C.; Hilbers, C.; van der Marel, G.; van Boom, J.; Singh, U.; Pattabiraman, N.; Kollman, P. *J Biomol Struct Dyn.* **1986**, *3*, 843-857.

(33) Kleinnijenhuis, A. J.; Duursma, M. C.; Breukink, E.; Heeren, R. M. A.; Heck, A. J. R. *Anal. Chem.* **2003**, *75*, 3219-3225.

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

(35) Hakansson, K.; Cooper, H. J.; Hudgins, R. R.; Nilsson, C. L. *Curr. Org. Chem.* **2003**, *7*, 1503-1525.

(36) Little, D. P.; Aaserud, D. J.; Valaskovic, G. A.; McLafferty, F. W. *J. Am. Chem. Soc.* **1996**, *118*, 9352-9359.

(37) Yang, J.; Hakansson, K. *53rd ASMS Conference on Mass Spectrometry and Allied Topics*, San Antonio, TX, June 5-9 2005; DVD-ROM.

(38) Yang, J.; Hakansson, K. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 1369-1375.

(39) Green-Church, K. B.; Limbach, P. A.; Freitas, M. A.; Marshall, A. G. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 268-277.

(40) Griffey, R. H.; Greig, M. J.; Robinson, J. M.; Laude, D. A. *Rapid Commun. Mass Spectrom.* **1999**, *13*, 113-117.

(41) Robinson, J. M.; Greig, M. J.; Griffey, R. H.; Mohan, V.; Laude, D. A. *Anal. Chem.* **1998**, *70*, 3566-3571.

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

(43) Wan, K. X.; Gross, J.; Hillenkamp, F.; Gross, M. L. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 193-205.

(44) Wan, K. X.; Gross, M. L. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 580-589.

(45) Green-Church, K. B.; Limbach, P. A. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 24-32.

CHAPTER 5

CHARACTERIZATIN OF NUCLEIC ACID HIGHER ORDER STRUCTURE BY GAS-PHASE HYDROGEN/DEUTERIUM EXCHANGE

In Chapter 2, gas-phase HDX experiments with MS detection were performed to show that anionic DNA duplexes have lower HDX rates than their constituent monomers, indicating that hydrogen bonding can shield hydrogens from exchanging with the bath gas D_2S. The same HDX assay is applied in this Chapter to investigate nucleic acid hairpin structure. Variations in hairpin solution-phase stabilities are achieved by changing their loop size, stem length, and stem composition (ratio of G/C and A/T(U) base pairs in the stem). These differences can be carried into the gas phase because ESI is a gentle ionization method that is able to preserve non-covalent interactions. The observed gas-phase HDX rates of these hairpins are consistent with their relative solution-phase stabilities as predicted by MFold web server, i.e., less stable nucleic acid hairpins exchange faster than more stable hairpins. To our knowledge, the presented experiments demonstrate for the first time that gas-phase HDX can be used to characterize nucleic acid higher order structure and the results suggest that the relative stabilities of nucleic acid hairpins in the gaseous phase are correlated with those in solution.

5.1. Introduction

The 3D configuration of an RNA molecule determines many of its biological properties, including its involvement in gene expression, regulatory processes, mRNA splicing, transport, and translation.[1-7] It is well established that single-stranded nucleic acid molecules possess the capability of condensing into compact structures by virtue of intrastrand interaction to form secondary structures.[8, 9] This type of interaction has been well characterized by X-ray crystallography,[10-12] NMR,[13, 14] FRET,[15] as well as relatively indirect approaches such as enzymatic digestion,[16-20] chemical probing,[17, 18, 21] electrophoresis,[22, 23] and phylogenetic conservation.[24, 25] Common mass spectrometric approaches for characterizing nucleic acid structure include ion mobility analysis,[26] and the use of MS in conjunction with solution-phase chemical footprinting.[27] Efficacy of these approaches is limited, however, as the first only provides information about the molecular cross section and the second strategy requires extensive sample manipulation prior to MS analysis. In Chapter 4, MS^n was applied to differentiate three isomeric DNA 15-mers, two of which were predicted to form hairpin structures in solution and the third without a preferred solution-phase conformation.[28] Results from EDD/IRMPD MS^3 and AI EDD in which precursor ions were briefly heated with an IR laser prior to EDD suggested a correlation between gas- and solution-phase structures for these 15-mers and provided some information on their relative stabilities as the more stable structure generated fewer product ions than the less stable ones. However, this approach is limited to relatively small molecules due to the low fragmentation efficiency of EDD.

Gas-phase HDX,[29-33] in which exchange rate constants are much smaller than those in solution has been proposed as an alternative method for structural analysis of

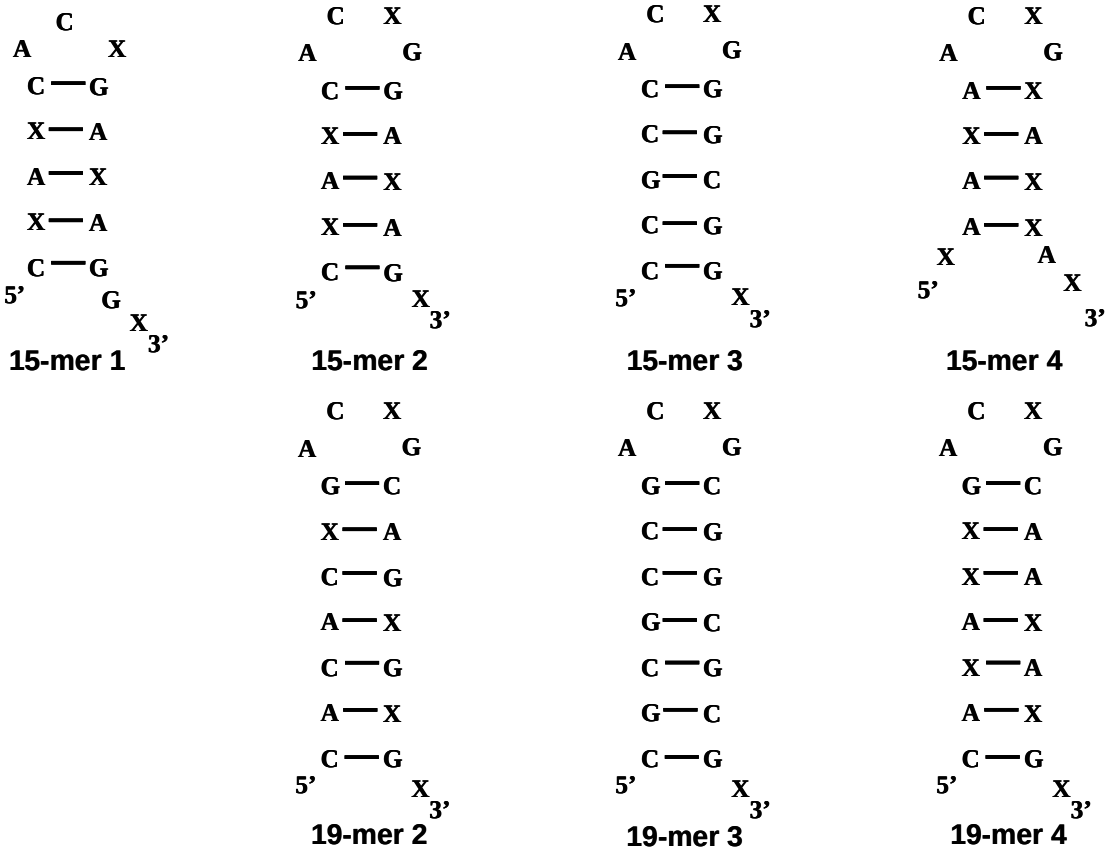
nucleic acids.[32, 34-40] Combining HDX with MS is a practical way to investigate gas-phase structures of biomolecules as well as their differences from solution-phase structures.[29-33, 36] FT-ICR MS is particularly well suited for gas-phase H/D exchange by virtue of its ultrahigh mass resolving power[32, 41-44] and the ability to trap and react gaseous ions for extended periods.[45] In Chapter 2, we showed that gas-phase HDX may be a useful technique for characterization of nucleic acid higher order structure because nucleic acid duplexes displayed lower HDX rates than their constituent monomers, due to hydrogens participating in base pairing being protected from exchange.[46] In this Chapter, we systematically investigate HDX of nucleic acid hairpins of varying stem length, loop size, and stem composition in an effort to seek a correlation between their solution-phase stabilities and gas-phase HDX rates.

5.2. Experimental Section

5.2.1. Sample Preparation

DNA (gel purified) and RNA (HPLC purified) were purchased from Yale Keck Facility (New Haven, CT) and Integrated DNA Technologies (Coralville, IA), respectively. Sequences of all the nucleic acids used are as follows (X = T for DNA and U for RNA) and their MFold[47, 48] predicted structures are shown in Scheme 5.1: 15-mer 1: CXAXCACXGAXAGGX, 15-mer 2: CXAXCACXGGAXAGX, 15-mer 3: CCGCCACXGGGCGGX, 15-mer 4: XAAXAACXGXAXXAX, 19-mer 2: CACACXGACXGCAGXGXGX, 19-mer 3, CGCGCCGACXGCGGCGCGX, 19-mer 4: CAXAXXGACXGCAAXAXGX. DNAs were used without further purification whereas RNAs were desalted by ethanol precipitation (protocol modified from Limbach *et al.*[49]:

1/3 volume of 10 M NH_4OAc (pH = 7) and 2.5 volume of 100% cold ethanol were added to the RNA aqueous solution, samples were vortexed and stored at -80 ºC for 3 hours, centrifuged for 15 min, and the supernatant was decanted. Then, 400 µL of cold 70% ethanol was added to the precipitate, samples were stored at -80 ºC for 2 hours, centrifuged for 15 min, the supernatant was decanted, the precipitate was dried down, and dissolved in electrospray solution).



Scheme 5.1. Nucleic acid structures predicted by Mfold Web Server[47, 48] in solution (X = T for DNA and U for RNA). DNA/RNA 15-mer 1 has the same base composition as DNA/RNA 15-mer 2 but their sequences differ, resulting in a three-nucleotide-loop hairpin structure for 15mer-1 and a four-nucleotide loop for 15-mer 2. All the other nucleic acids characterized in this study have the preferable four-nucleotide-loop hairpin structure. DNA/RNA 15- and 19-mer 2 represent mixed base pair stem hairpins; DNA/RNA 15- and 19-mer 3 represent G/C rich hairpins, and DNA/RNA 15- and 19-mer

4 represent A/T(U) rich hairpins. All the 19-mers have two more G/C base pairs in the hairpin stem than their corresponding 15-mers.

	Δ G of 15-mer 1	Δ G of 15-mer 2	Δ G of 15-mer 3	Δ G of 15-mer 4	Δ G of 19-mer 2	Δ G of 19-mer 3	Δ G of 19-mer 4
DNA	**- 1.0**	**- 1.8**	**- 5.8**	**- 0.6**	**- 6.2**	**- 10.5**	**- 3.9**
RNA	**- 2.5**	**- 3.3**	**- 8.8**	**+ 0.6**	**- 9.1**	**- 13.6**	**- 4.9**

Table 5.1. ΔG values (in kcal/mol, calculated by Mfold Web Server[47, 48]) of nucleic acid hairpins in aqueous solution of 1 M Na^+. More negative values represent more stable structures.

5.2.2. Mass Spectrometry

Gas-phase deprotonated precursor ions were generated by negative ion mode ESI at 50 μL/h through an external Apollo II source equipped with dual ion funnels (Bruker Daltonics, Billerica, MA) in a spray solution consisting of 25% (vol/vol) methanol (Fisher, Fair Lawn, NJ) and 50 mM NH_4OAc (Fisher). Ion source transfer conditions, including the ion funnel DC voltages and skimmer voltages, were adjusted to be as 'soft' as possible (higher charge states, which are not as stable as lower charge states, are observed under 'soft' conditions) to maximally preserve nucleic acid higher order structure in the solution-to-gas phase transition. Analyte concentrations varied between 10 to 100 μM in order to get a signal-to-noise ratio of 10 in 8 scans.

All mass spectra were collected with an actively shielded 7 Tesla FT-ICR mass spectrometer (Bruker Daltonics) with a quadrupole front end (APEX-Q, Bruker Daltonics) as previously described in Chapter 1.[28] Precursor ions in either 4- or 5- charge states

were selected by the quadrupole and accumulated externally in the collision cell for 0.5 s, in the presence of D_2S (Cambridge Isotope Laboratories, Andover, MA) at a pressure of ~ 5 x 10^{-6} mbar (gauge factory calibrated for nitrogen, no further calibration was done). Multiple ICR cell fills were not used to avoid generation of mixtures of ion populations with different exposure time to D_2S (there is residual D_2S outside the storage hexapole). Gas-phase HDX reactions were performed for 0.51 to 30.5 s and mass spectra indicating mass shifts at each time point were recorded by XMASS (version 6.1 Bruker Daltonics). Data were processed with the MIDAS analysis software.[50]

Percentage of HDX was calculated as described in Chapter 2. Error bars were generated from three individual experiments acquired on the same day due to difficulties with reproducing the D_2S pressure in the collision cell and the t-test was used for statistical analysis. Due to the fact that deuterium incorporation of nucleic acids reaches saturation after a short time period, typically within one minute, data were only collected for the first 30 s to decrease sample consumption and to save time.

5.3. Results and Discussion

5.3.1. Effects of Salt and Organic Solvent in Spray Solution on Gas-Phase HDX

Metal salts are known to play an important role in nucleic acid folding because they can shield negative charges on the backbone phosphate groups and thereby eliminate increased repulsive forces arising from compact secondary and tertiary structures. All nucleic acids investigated in this Chapter are predicted to have a hairpin structure by the Mfold web server[47, 48] in 1 M Na^+ aqueous solution (Scheme 5.1). However, such high Na^+ concentration in nucleic acid samples results in low negative mode ionization

efficiency and low precursor ion abundance due to the distribution of signal among multiple Na^+-bound adducts. To overcome this limitation, NH_4^+ ions from NH_4OAc instead of Na^+ have been used in previous studies. NH_4^+ ions neutralize backbone negative charges but can be stripped off from precursor ions during ESI. For example, Gabelica *et al.*[51] used 50 mM NH_4OAc for characterizing DNA duplexes and showed that DNA double helices were conserved. They also showed that with similar solvent (100 mM NH_4OAc) the structure of complexes between DNA duplexes and minor groove binders were preserved.[52] Baker *et al.*[53, 54] have employed CD, ion mobility and molecular dynamics simulations to compare the conformations of G-quadruplexes with different lengths of the telomeric repeat in the gas phase with those in the solution phase. Their results imply that ionic stabilization of G-quadruplexes is not required for the structure to survive in the gas phase because the solution strand orientations observed in the CD spectra were the same as the solvent-free strand orientations detected in the ion mobility measurements and molecular dynamics calculations strongly supported structural conservation upon ESI spraying and dehydration of these G-quadruplexes. The effect of different cations in our experiment was investigated by comparing the gas-phase HDX rates of DNA hairpins in 1 mM sodium acetate (NaOAc) to those in 50 mM NH_4OAc. Statistically similar HDX rates were obtained for both DNA 15-mer 1 and DNA 15-mer 2 (see Scheme 5.1 for structures) in these two different spray solutions, indicating similar effects of 1 mM Na^+ and 50 mM NH_4^+ on the gas-phase stabilities of nucleic acid hairpins (Figure 5.1). This small influence from different types of counter ions is probably due to the enhanced hydrogen bonding in the gas phase resulting from the removal of polar solvent.[55-59] Gabelica *et al.*[51, 52] have shown that thermal

denaturation curves of DNA duplexes as well as their complexes with minor groove binders in the gas phase are similar to those in solution phase. Higher Na^+ concentrations were not tested due to the difficulty of generating precursor ions of acceptable signal-to-noise ratios. Instead, NH_4OAc was varied from 25 mM to 500 mM to investigate the concentration influence. Our results show statistically similar behavior for DNA 15-mer 2 at all concentrations (except at 25 mM, where slightly faster HDX rate was observed, Figure 5.2). Thus, 50 mM NH_4OAc was arbitrarily used for all following experiments.

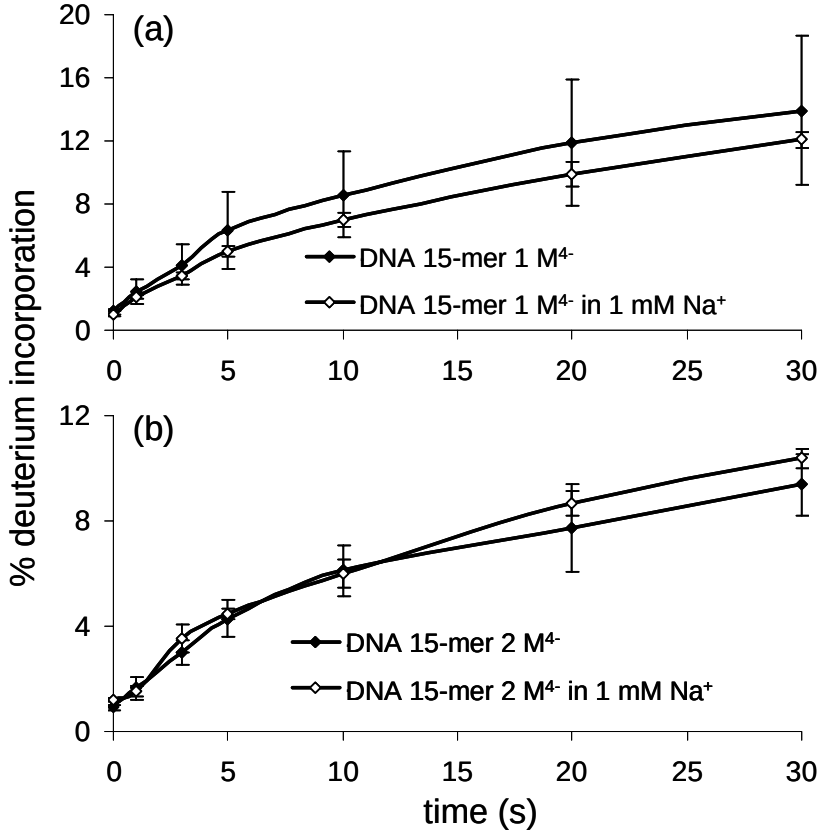


Figure 5.1. Gas-phase HDX of DNA 15-mer 1 (a) and 15-mer 2 (b) in 25% methanol and 1 mM Na^+ compared to those in 25% methanol and 50 mM NH_4OAc. Statistically similar results were observed for both DNAs in the two different spray solutions, indicating that 50 mM NH_4OAc has similar effect on nucleic acid stability as 1 mM Na^+.

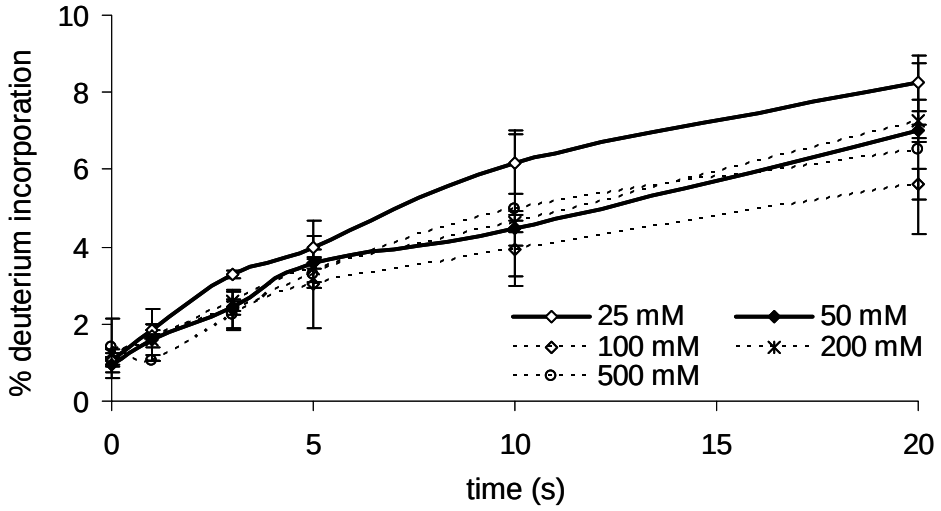


Figure 5.2. Gas-phase HDX of DNA 15-mer 2 electrosprayed at varying NH_4OAc concentrations. At 25 mM NH_4OAc (lowest concentration), the hairpin exchanges slightly faster than at higher concentrations. However, at 50 mM and above, there is no apparent influence of NH_4OAc concentration on the HDX rate. All further experiments were acquired at 50 mM NH_4OAc.

Organic solvent in the spray solution greatly facilitates high ionization efficiency because it assists the extraction of precursor ions from ESI droplets by decreasing surface tension.[60, 61] Although the presence of organic solvent may shift the folding equilibrium of nucleic acid hairpins in solution, when transferred into the gas phase, as mentioned above, the hydrogen bonds between base pairs should be enhanced due to the hydrophobic environment following removal of polar solvent.[55-59] Thus, we hypothesize the hairpin structure would still be favored in the gas phase. Nevertheless, the influence of organic solvent in the spray solution was investigated by comparing the gas-phase HDX rates of DNA 15-mer 2 in spray solutions with varying methanol percentages (5% to 95% with 5% increments). Results suggested that HDX behavior changed slightly at

different methanol percentage but no clear trend was observed (data not shown). Therefore, 25% methanol was used for all subsequent experiments to eliminate the variable effects of organic solvents on HDX (20% methanol was used in the work discussed above by Gabelica *et al*.[51, 52]).

5.3.2. Effect of Charge State on Gas-Phase HDX

ESI characteristically produces ions with multiple charge states,[62, 63] which would influence the biomolecular higher order structure because Coulomb repulsion increases with increasing number of charge sites in the molecule, thereby presenting more unfolded structures,[64-73] and rendering hydrogens more accessible to HDX. Gas-phase HDX experiments were used to estimate the stabilities of nucleic acid higher order structures at different charge states. Figure 5.3 shows the comparison of gas-phase HDX rates of both DNA (a, b) and RNA (c, d) hairpins 15-mer 1 and 2 at their 4- and 5- charge states. In all four cases, the 5- charge state exchanged faster than the 4- charge state of the same molecule, consistent with the fact that biomolecules at higher charge states are less structured.

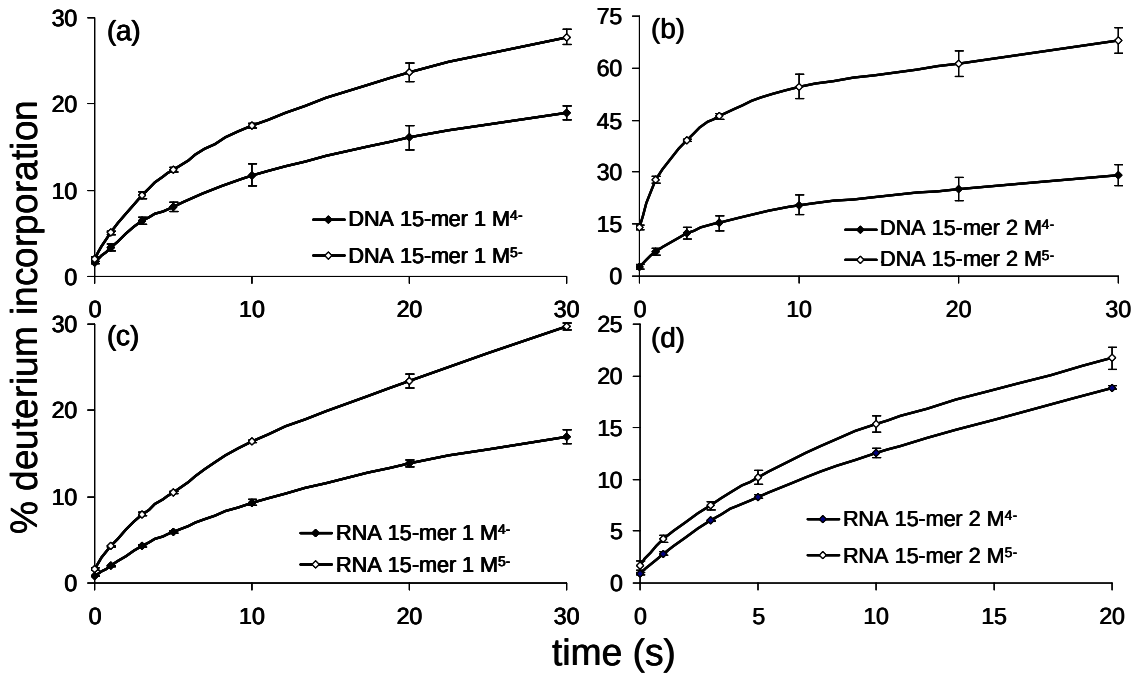


Figure 5.3. Gas-phase HDX of nucleic acid hairpins at different charge states. For all four nucleic acid hairpins (DNA 15-mer 1 (a), DNA 15-mer 2 (b), RNA 15-mer 1 (a) and RNA 15-mer 2 (d)), higher charge state (5-) exchanges faster than lower charge state (4-) of the same molecule, consistent with a more unfolded hairpin structure at higher charge state.

5.3.3. Effect of Hairpin Loop Size on Gas-Phase HDX

DNA/RNA 15-mers 1 and 2 have the same base composition but different sequences, resulting in a three-nucleotide and a four-nucleotide loop hairpin, respectively (Scheme 5.1). The stability of the former hairpin is predicted to be lower than the stability of the latter by MFold (see ΔG values in Table 5.1) and by thermal denaturation experiments.[74] Figure 5.4 displays gas-phase HDX of their 4- charge state precursor ions. Deuterium incorporation in 15-mer 1 is faster than that of 15-mer 2 for both DNA (Figure 5.4a, 30% vs. 20% in 20 s) and RNA (Figure 5.4b, 30% vs. 22% in 30s), indicating that the three-nucleotide loop hairpin is less stable than the four-nucleotide-loop hairpin,

consistent with the literature. The ratio of average deuterium incorporation into 15-mer 1 with respect to 15-mer 2 at each data point was 1.8 ± 0.2 for DNA and 1.5 ± 0.2 for RNA. When comparing these values to the ratio of MFold-predicted ΔG values for 15-mer 2 and 15-mer 1 (1.8 for DNA and 1.3 for RNA) in solution, they are numerically close, indicating that relative gas-phase HDX rates may reflect the relative solution-phase stability of nucleic acid hairpins with different loop sizes.

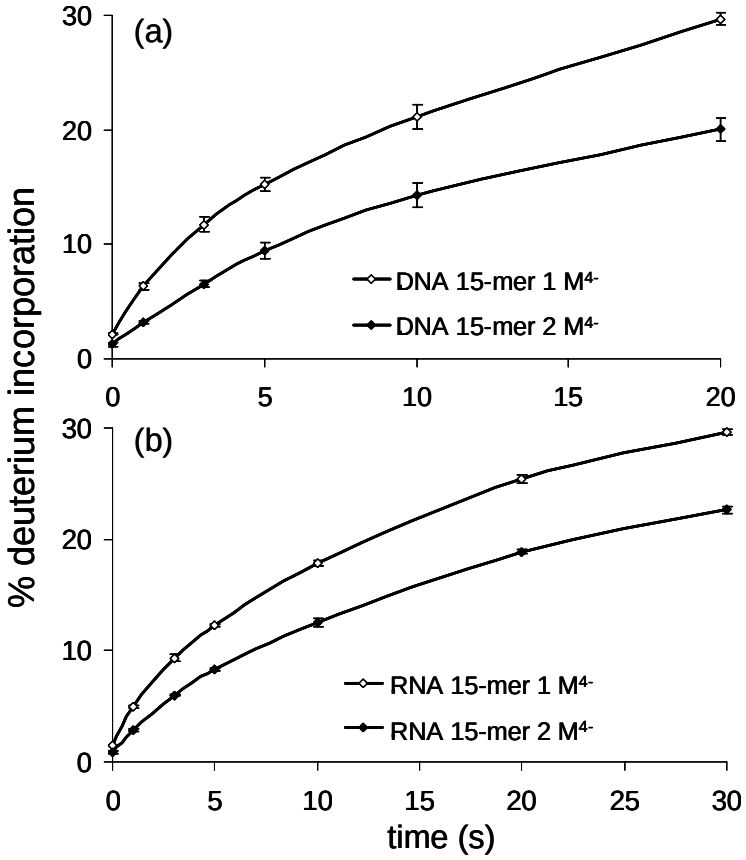


Figure 5.4. Gas-phase HDX of nucleic acid hairpins with a three-nucleotide loop compared to those with a four-nucleotide loop. The former undergoes exchange faster than the latter for both DNA (X = T) and RNA (X = U), consistent with their solution-phase stabilities as predicted by MFold.

5.3.4. Effect of Hairpin Stem Length on Gas-Phase HDX

Another factor that influences hairpin stability is its stem length; longer stems have more base pairs and should therefore be more stable and less susceptible to HDX. Six pairs of nucleic acid hairpins (one 15-mer and one 19-mer) were selected for investigating the effect of stability for different stem lengths on gas-phase HDX behavior (structures are shown in Scheme 5.1). The MFold predicted ΔG values of all six 19-mer hairpins (Table 5.1) predict that they are more stable than their corresponding 15-mer hairpins due to the additional two G/C base pairs in the stem. When comparing the same charge states (5-) of 19-mers and 15-mers, the relative exchange rates for both DNA and RNA correlate well with their relative solution-phase stabilities (Figure 5.5). However, when comparing species of the same charge density (5- charge state of 19-mer vs. 4- charge state of 15-mer), the results are not consistent with their relative solution-phase stabilities and no conclusive trend could be found (data not shown). Therefore, we conclude that gas-phase HDX is able to demonstrate relative stabilities of nucleic acid hairpins with different stem lengths of the same charge state but not of the same charge density.

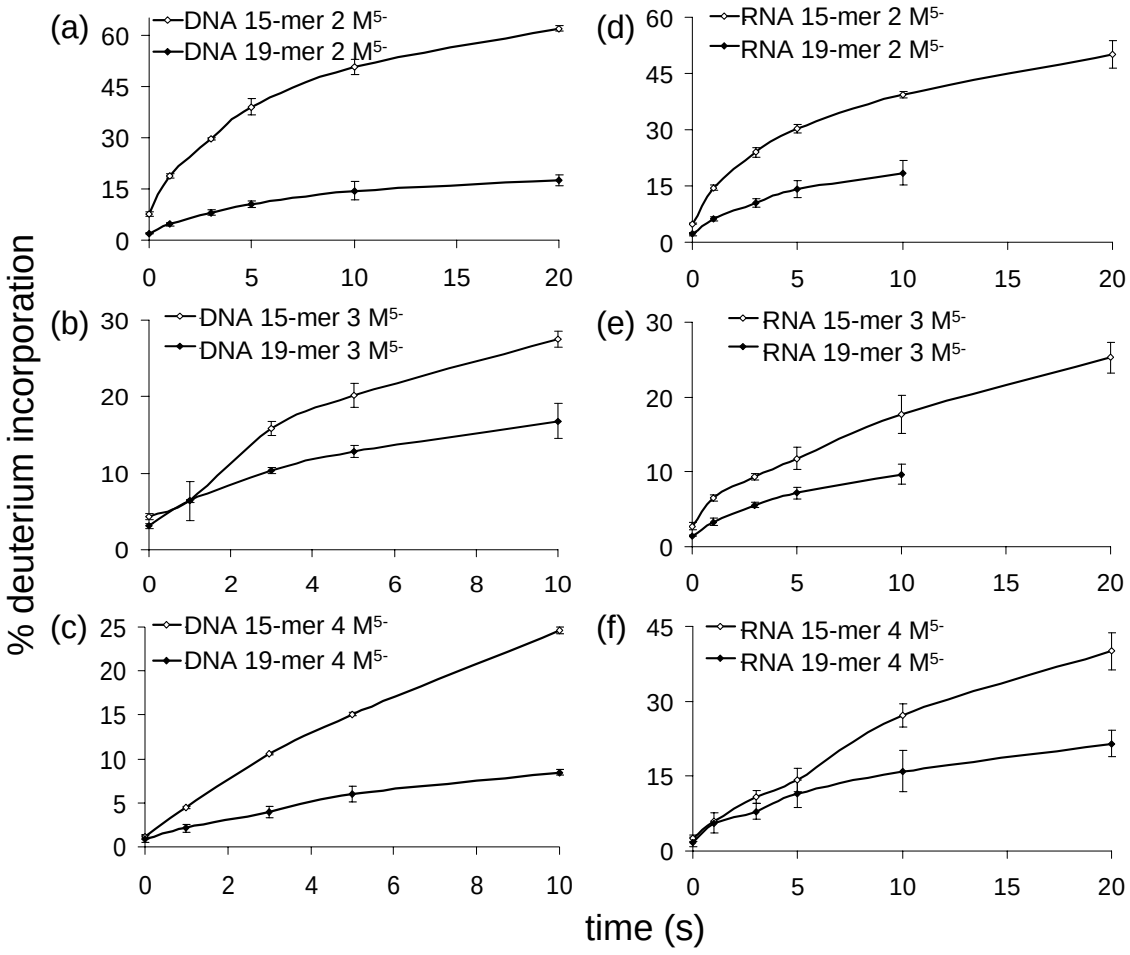


Figure 5.5. Gas-phase HDX of nucleic acid hairpins with different stem lengths. All of the three 19-mers contain two more G/C base pairs in the stem than their corresponding 15-mers, and, consequently, exchange slower than the 15-mers, which correlates with their predicted solution-phase stabilities.

5.3.5. Effect of Hairpin Stem Composition on Gas-Phase HDX

The last factor of hairpin stability investigated was stem composition. Due to the fact that a G/C base pair has one more hydrogen bond than an A/T(U) base pair, slower HDX behavior is expected for hairpins with G/C rich stems compared to those with A/T(U) rich or mixed base pair stems. To characterize this effect, gas-phase HDX

experiments were performed on DNA/RNA 15- and 19-mers 2, 3, and 4, which represent hairpins with mixed base pairs, G/C rich, and A/T(U) rich stems, respectively. The expected order of stabilities for these three groups is 3 > 2 > 4, and therefore the expected order of HDX rates is 3 < 2 < 4. Results are shown in Figure 5.6. In all cases, the mixed base pair stem hairpins exchange faster than their corresponding G/C rich hairpins, except for DNA 15-mers (Figure 5.6a) where these two groups exchange at similar rates, which is probably due to the relatively small difference in their ΔG values. A/T(U) rich hairpins always exchange faster than the other two hairpin types for DNAs as expected. For RNAs, however, they exchange faster than the G/C rich hairpins and slower than the mixed base pair stem hairpins. This unexpected behavior may be due to alternative gas-phase conformations of RNA A/U rich hairpins. For example, Mfold predicted a positive ΔG value for A/U rich RNA 15-mer 4, indicating hairpin formation is unfavorable in solution, thus difficult to compare its behavior with other hairpin structures. Overall, for nucleic acids capable of forming stable hairpin structures, their gas-phase HDX rates could be correlated to their relative solution-phase stabilities.

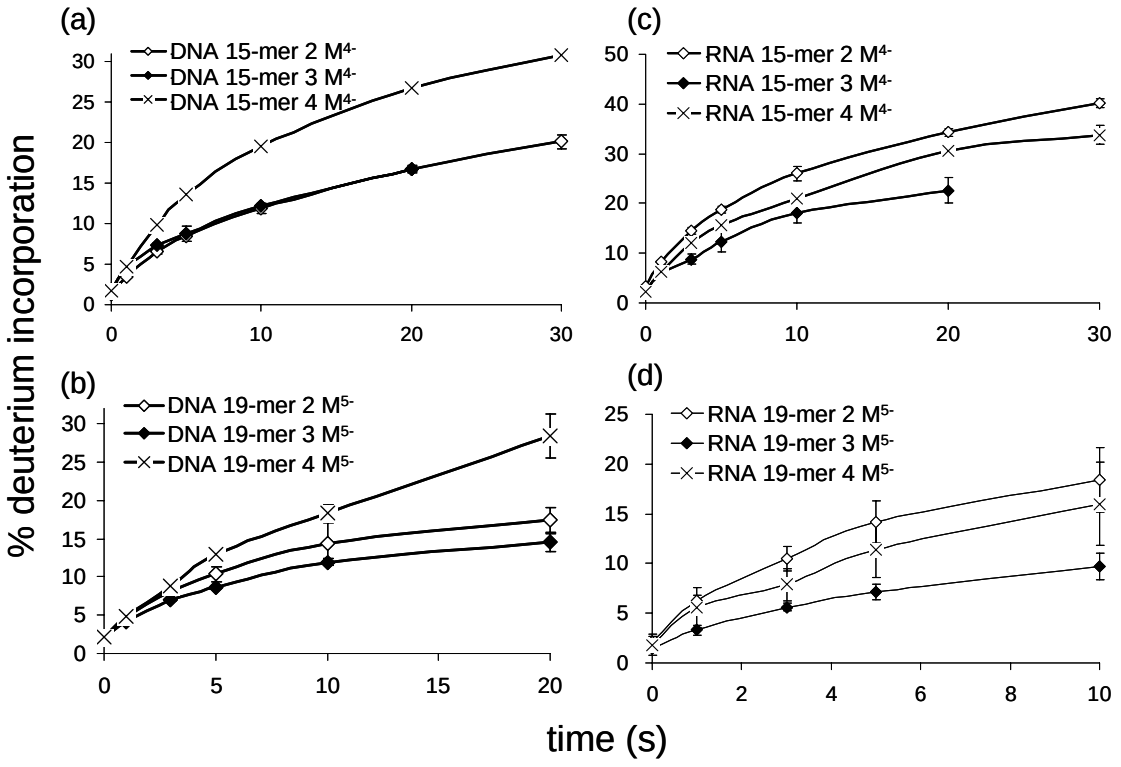


Figure 5.6. Gas-phase HDX of nucleic acid hairpins with different stem compositions (percentage of G/C base pairs in the stem). For correlation with ΔG values, HDX rates should follow the trend DNA/RNA 4 > DNA/RNA 2 > DNA/RNA 3. However, observed rates for RNA 15- and 19-mer 4 deviate from this behavior, possibly because they have too many A/T base pairs in the stem such that formation of a hairpin is not favorable. Instead, they may fold into other structures in the gas phase.

5.4. Conclusion

In this Chapter, we show that gas-phase HDX combined with MS is a useful technique for characterization of nucleic acid higher order structure. The stability of hairpins, which are the dominant secondary structure of nucleic acids, depends on the loop size, stem length, and stem composition. Results from gas-phase HDX experiments of a set of nucleic acid hairpins suggest that their relative solution-phase stabilities can be reflected by their gas-phase HDX rates, i.e., more stable hairpins exchange slower than

the less stable ones. For example, four-nucleotide-loop hairpins exchange slower than isomeric three-nucleotide-loop hairpins; long stem hairpins exchange slower than short stem hairpins, and G/C rich hairpins exchange slower than mixed stem hairpins, which exchange slower than A/T(U) rich hairpins (if they could form stable hairpin structure). The influence of salts in spray solution on nucleic acid higher order structure was also investigated and NH_4^+ was found to have similar effect for neutralizing the backbone negative charges as Na^+ and its concentration does not appear to have significant influence on HDX rates. However, organic solvent was found to have some influence on the gas-phase HDX rates but the trend is not clear. Overall, the presented gas-phase HDX experiments suggest that the relative structural stabilities of nucleic acid hairpins in the gaseous phase are consistent with those in solution.

5.5. Bibliography

(1) Tuerk, C.; Gauss, P.; Thermes, C.; Groebe, D.; Gayles, M.; Guild , N.; Stormo, G.; d'Aubenton-Carafa, Y.; Uhlenbeck, O.; Tinoco, I.; Brody, E.; Gold, L. *Proc. Nat. Acad. Sci. U.S.A.* **1988**, *85*, 1364-1368.

(2) Adams, C. C.; Stern, D. B. *Nucleic Acids Res.* **1990**, *18*, 6003-6010.

(3) Klausner, R.; Rouault, T.; Harford, J. *Cell* **1993**, *72*, 19-28.

(4) Gottlieb, E. *Proc. Nat. Acad. Sci. U.S.A.* **1992**, *89*, 7164-7168.

(5) Gesteland, R. F.; Cech, T. R.; Atkins, J. F. *The RNA World*; Cold Spring Harbor Laboratory Press: NY, **1999**.

(6) Skripkin, E.; Paillart, J.-C.; Marquet, R.; Ehresmann, B.; Ehresmann, C. *Proc. Nat. Acad. Sci. U.S.A.* **1994**, *91*, 4945-4949.

(7) Yang, S.; Temin, H. *EMBO J.* **1994**, *13*, 713-726.

(8) James, J.; I Tinoco, J. *Nucleic Acids Res.* **1993**, *21*, 3287-3293.

(9) Panayotatos, N.; Wells, R. *Nature* **1981**, *289*, 466-470.

(10) Ban, N.; Nissen, P.; Hansen, J.; Moore, P. B.; Steitz, T. A. *Science* **2000**, *289*, 905-920.

(11) Wimberly, B. T.; Brodersen, D. E.; Clemons, W. M. J.; Morgan-Warren, R. J.; Carter, A. P.; Vonrhein, C.; Hartsch, T.; Ramakrishnan, V. *Nature* **2000**, *407*.

(12) Carter, A. P.; Celmons, W. M.; Brodersen, D. E.; Morgan-Warren, R. J.; Wimberly, B. T.; Ramakrishnam, V. *Nature* **2000**, *407*, 340-348.

(13) Al-Hashimi, H. M. *ChemBioChem* **2005**, *6*, 1472-1473.

(14) Latham, M. P.; Brown, D. J.; McCallum, S. A.; Pardi, A. *ChemBioChem* **2005**, *6*, 1492-1505.

(15) Klostermeier, D.; Millar, D. P. *Methods* **2001**, *23*, 240-254.

(16) Noller, H. F.; Woese, C. R. *Science* **1981**, *212*, 403-411.

(17) Ehresmann, C.; Baudin, F.; Mougel, M.; Romby, P.; Ebel, J.; Ehresmann, B. *Nucleic Acids Res.* **1987**, *15*, 9109-9127.

(18) Parker, R. *Methods Enzymol.* **1989**, *180*, 510-517.

(19) Pavlakis, G. N.; Lockard, R. E.; Vamvakopoulos, N.; Lauren Rieser, L.; Rajbhandary, U. L.; Vournakis, J. N. *Cell* **1980**, *19*, 91-102.

(20) Yao, Y.; Wang, Q.; Hao, Y.; Tan, Z. *Nucleic Acids Res.* **2007**, *35*, e68.

(21) Peattie, D. A.; Gilbert, W. *Proc. Nat. Acad. Sci. U.S.A.* **1980**, *77*, 4679-4682.

(22) Ross, A.; Brimacombe, R. *Nature* **1979**, *281*, 271-276.

(23) Glotz, C.; Zwieb, C.; Brimacombe, R.; Edwards, K.; Kossel, H. *Nucleic Acids Res.* **1981**, *9*, 3287-3306.

(24) Gutell, R. R.; Gray, M. W.; Schnare, M. N. *Nucleic Acids Res.* **1993**, *21*, 3055-3074.

(25) Woese, C. R.; Pace, H. R. *The RNA world*; Cold Spring Harbor: New York, **1993**, 91-117.

(26) Gidden, J.; Baker, E. S.; Ferzoco, A.; Bowers, M. T. *Int. J. Mass Spectrom.* **2005**, *240*, 183-193.

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

(28) Mo, J.; Hakansson, K. *Anal. Bioanal. Chem.* **2006**, *386*, 675-681.

(29) Freitas, M. A.; Hendrickson, C. L.; Emmett, M. R.; Marshall, A. G. *J. Am. Soc. Mass Spectrom.* **1998**, *9*, 1012-1019.

(30) Campbell, S.; Rodgers, M. T.; Marzluff, E. M.; Beauchamp, J. L. *J. Am. Chem. Soc.* **1994**, *116*, 9765.

(31) Gard, E.; Green, M. K.; Bregar, J.; Lebrilla, C. B. *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 623-631.

(32) Green, M. K.; Lebrilla, C. B. *Mass Spectrom. Rev.* **1997**, *16*, 53-71.

(33) Wyttenbach, T.; Bowers, M. T. *J. Am. Soc. Mass Spectrom.* **1999**, *10*, 9-14.

(34) Chipuk, J. E.; Brodbelt, J. S. *J. Am. Soc. Mass Spectrom.* **2007**, *18*, 724-736.

(35) Crestoni, M. E.; Fornarini, S. *J. Mass Spectrom.* **2003**, *38*, 854-861.

(36) Robinson, J. M.; Greig, M. J.; Griffey, R. H.; Mohan, V.; Laude, D. A. *Anal. Chem.* **1998**, *70*, 3566-3571.

(37) Griffey, R. H.; Greig, M. J.; Robinson, J. M.; Laude, D. A. *Rapid Commun. Mass Spectrom.* **1999**, *13*, 113-117.

(38) Hofstadler, S. A.; Sannes-Lowery, K. A.; Griffey, R. H. *J. Mass Spectrom.* **2000**, *35*, 62-70.

(39) Gabelica, V.; Rosu, F.; Witt, M.; Baykut, G.; de Pauw, E. *Rapid Commun. Mass Spectrom.* **2005**, *19*, 201-208.

(40) Freitas, M. A.; Marshall, A. G. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 780-785.

(41) Amster, I. J. *J. Mass Spectrom.* **1996**, *31*, 1325-1337.

(42) Dienes, T.; Pastor, S. J.; Schch, S.; Scott, J. R.; Yao, J.; Cui, S.; Wilkins, C. L. *Mass Spectrom. Rev.* **1996**, *15*, 163-211.

(43) Laude, D. A., Jr.; Stevenson, E.; Robinson, J. M. In *Electrospray ionization mass spectrometry: fundamentals, instrumentation & applications*; Cole, R. B., Ed.; John Wiley & Sons, Inc.: New York, 1997, pp 291-319.

(44) Marshall, A. G.; Hendrickson, C. L.; Jackson, G. S. *Mass Spectrom. Rev.* **1998**, *17*, 1-35.

(45) Marshall, A. G.; Guan, S. *Rapid Commun. Mass Spectrom.* **1996**, *10*, 1819-1823.

(46) Mo, J.; Hakansson, K. *Anal. Chem.* **2007**, *79*, 7893-7898.

(47) Mathews, D. H.; Sabina, J.; Zuker, M.; Turner, D. H. *J. Mol. Biol.* **1999**, *288*, 911-940.

(48) Zuker, M. *Nucleic Acids Res.* **2003**, *31*, 3406-3415.

(49) Limbach, P. A.; Crain, P. F.; McCloskey, J. M. *J. Am. Soc. Mass Spectrom.* **1995**, *6*, 27-39.

(50) Senko, M. W.; Canterbury, J. D.; Guan, S.; Marshall, A. G. *Rapid Commun. Mass Spectrom.* **1996**, *10*, 1839-1844.

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

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

(53) Erin Shammel Baker, S. L. B., and Michael T. Bowers *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 989-997.

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

(55) Thomas, R. K. *Proc. Roy. Soc. Lond. A.* **1971**, 137-146.

(56) Gao, J.; Carbeck, J.; Q.; Smith, R. D.; Whitesides, G. M. *Biophys. J.* **1999**, *76*, 3253-3260.

(57) Gao, Q. Y.; Cheng, X. H.; Smith, R. D.; Yang, C. F.; Goldberg, I. H. *J. Mass. Spectrom.* **1996**, *31*, 31-36.

(58) Rogniaux, H.; Van Dorsselaer, A.; Barth, P.; Biellmann, J. F.; Barbanton, J.; van Zandt, M.; Chevrier, B.; Howard, E.; Mitschler, A.; Potier, N.; Urzhumtseva, L.; Moras, D.; Podjarny, A. *J. Am. Soc. Mass Spectrom.* **1999**, *10*, 635-647.

(59) Saenger, W. *Principles of Nucleic Acid Structure*; Springer-Verlag: New York, **1984**, 116-158.

(60) Ikonomou, M. G.; Blades, A. T.; Kebarle, P. *Anal. Chem.* **1991**, *63*, 1989-1998.

(61) Tang, L.; Kebarle, P. *Anal. Chem.* **1991**, *63*, 2709-2715.

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

(63) Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F. *Mass Spectrom. Rev.* **1990**, *9*, 37-70.

(64) Chowdhury, S. K.; Katta, V.; Chait, B. T. *J. Am. Chem. Soc.* **1990**, *112*, 9013-9015.

(65) Loo, J. A.; Edmonds, C. G.; Udseth, H. R.; Smith, R. D. *Anal. Chem.* **1990**, *62*, 693-698.

(66) Mirza, U. A.; Cohen, S. L.; Chait, B. T. *Anal. Chem.* **1993**, *65*, 1-6.

(67) Wagner, D. S.; Anderegg, R. J. *Anal. Chem.* **1994**, *66*, 706-711.

(68) Konermann, L.; Collings, B. A.; Douglas, D. J. *Biochemistry* **1997**, *36*, 5554-5559.

(69) Konermann, L.; I., R. F.; Mauk, A. G.; Douglas, D. J. *Biochemistry* **1997**, *36*, 6448-6454.

(70) Konermann, L.; Douglas, D. J. *Biochemistry* **1997**, *36*, 12296-12302.

(71) Konermann, L.; Douglas, D. J. *Rapid Commun. Mass Spectrom.* **1998**, *12*, 435-442.

(72) Konermann, L.; Douglas, D. J. *J. Am. Soc. Mass Spectrom.* **1998**, *9*, 1248-1254.

(73) Vincent W. S. Lee, V. W. S.; Chen, Y. L.; Konermann, L. *Anal. Chem.* **1999**, *71*, 4154-.

(74) Groebe, D. R.; Uhlenbeck, O. C. *Nucleic Acids Res.* **1988**, *16*, 11725-11735.

CHAPTER 6

CHARACTERIZATION OF RNA BINDING WITH AMINOGLYCOSIDE ANTIBIOTICS BY GAS-PHASE HYDROGEN/DEUTERIUM EXCHANGE

RNA molecules have well-defined tertiary structures that allow specific binding with small molecules. *E. coli* 16S rRNA is a component of the ribosome small subunit involved in ribosomal decoding and its interactions with mRNA and tRNA play important roles in protein translation. It has been defined that the decoding region in 16S rRNA, which contains two phylogenetically conserved sequences, is the binding site of aminoglycoside antibiotics. As the first class of compounds known to bind specifically to the subdomains of larger RNA sequences, aminoglycoside antibiotics are useful for understanding the design principles required to produce new classes of therapeutic agents to bind with RNAs. NMR studies have shown that conformational change is induced in RNA when it binds with aminoglycoside antibiotics. In this Chapter, we attempt to characterize the conformational changes induced by different aminoglycoside antibiotics using gas-phase HDX. Results show that all of the investigated RNA complexes have slower HDX processes than the corresponding free RNA, indicating that hydrogens are protected from exchange upon ligand binding. This effect is likely a combination of hydrogen bonding between the RNA and the ligand as well as formation of a more

compact structure. Additionally, the order of HDX rates of different complexes is consistent with the order of their solution-phase dissociation constants calculated from titration experiments.

6.1. Introduction

Ribosomes are complexes of rRNA and protein and consist of two subunits that fit together and work as one to translate mRNA into a polypeptide chain during protein synthesis by using amino acids delivered by tRNA. The function of rRNA is to provide a mechanism for decoding mRNA into amino acids and to interact with tRNAs during translation. A highly conserved rRNA sequence forms the site of interaction between codon and anticodon (A-site) in the small ribosomal subunit. Crystallographic work has shown that there are no ribosomal proteins close to the reaction site for polypeptide synthesis, suggesting the protein components of ribosomes act as a scaffold that may enhance the ability of rRNA to synthesize protein rather than directly participating in catalysis. Some groups of antibiotics can impair the decoding function of rRNA, in both prokaryotes and eukaryotes, by interacting with the A-site.[1-9]

The specificity of the interaction between aminoglycoside antibiotics and the A-site rRNA is governed by the 3D structure of the rRNA as well as electrostatic and nucleobase-specific interactions.[10] To characterize such complexes, solution-phase methods including radioimmunoassays, filter assays, and surface plasmon resonance assays have been used, which typically measure the equilibrium concentration of either the free ligand, or the complex but provide little or no information about the binding stoichiometry.[11] Gel mobility shift assay has been used in some cases to show that

multiple ligands can bind to RNA, however, the binding stoichiometry determined in those studies is not always correct (mass of the complex does not correspond to the determined stoichiometry).[12-14] In contrast, MS can directly determine the binding stoichiometry by measuring the mass of the complex formed as well as detecting all components of the equilibrium mixture. ESI[15, 16] is a very gentle ionization technique that transfers non-covalent complexes formed in solution into the gas phase where they can be characterized by mass spectrometric analysis.[17-19] ESI MS has been widely used to determine the stoichiometry and dissociation constants (K_d) for protein-protein, protein-ligand, and protein-oligonucleotide interactions[20-27] as well as RNA-ligand interactions.[28-31] It has been demonstrated that binding properties and stoichiometry characterized by MS are in good agreement with those derived from more conventional solution-phase techniques.[17, 20, 21, 23-27, 31-33]

Hofstadler *et al.*[34] demonstrated that binding of five aminoglycoside antibotics to 27 nucleotide models of the 16S (prokaryotic) and 18S (eukaryotic) rRNAs could be measured in parallel using ESI FT-ICR MS and, by adding a neutral mass tag to the 18S model, all complexes could be resolved in the same mass spectrum. Griffey *et al.*[35] were able to estimate the relative binding affinities of different complexes simultaneously by comparing their relative peak abundances in the mass spectrum because the binding of the ligand to the RNA target does not appear to alter the ionization efficiency.[26, 28, 31] This group also calculated the K_d values for different complexes by plotting the complex/free RNA abundance ratio vs. the varying ligand concentration (referred to as titration study in this Chapter). Sannes-Lowery *et al.*[10] further examined the effect of solution conditions such as organic solvent and buffer concentration on the measured K_d

values. They also compared two methods for measuring the K_d values: maintaining the concentration of the ligand fixed, or maintaining the concentration of the RNA fixed. The calculated K_d values differed two-fold between these two methods as did K_d values determined at high and low concentrations of RNA.

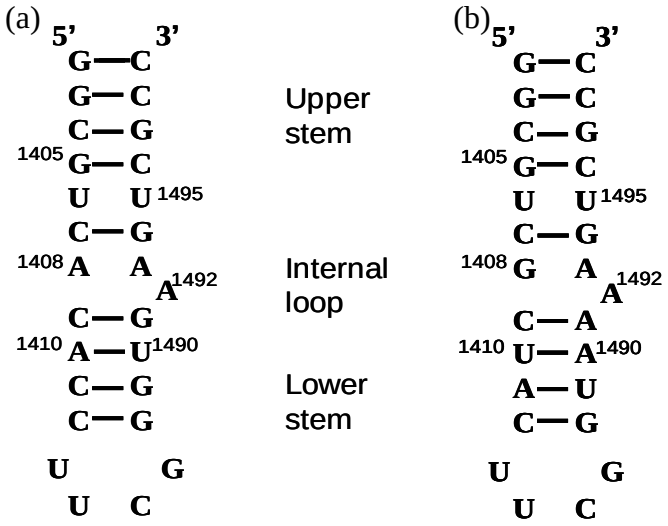
Gas-phase HDX combined with MS is a practical way to analyze nucleic acid structures because the exchange rate constants in the gas phase are much smaller than those in solution phase.[36-47] As shown in Chapter 2, it is a useful technique for characterization of hydrogen bonds because nucleic acid duplexes undergo less HDX extent than their constituent monomers.[48] Chapter 5 further demonstrated that gas-phase HDX rates of a series of nucleic acid hairpins are correlated with their solution-phase stabilities, i.e., hairpins with higher stability exchange at a slower rate. In this Chapter, we attempt to use gas-phase HDX to characterize the binding between rRNA and different aminoglycoside antibiotics.

6.2. Experimental Section

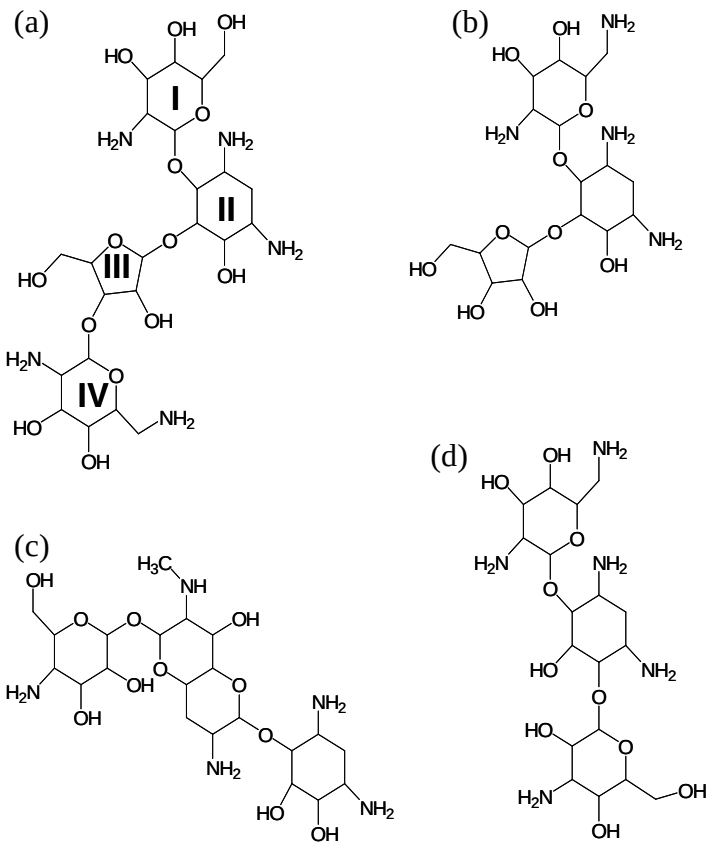
6.2.1. Sample Preparation

Two 27-mer RNAs, R1 and R2 (structures are shown in Scheme 6.1, adapted from[34, 35]), corresponding to the A-site structures of 16S and 18S respectively, were purchased from Integrated DNA Technologies (Coralville, IA) and desalted by ethanol precipitation (see detailed procedure described in Chapter 5 of this thesis). Aminoglycoside antibiotics, including paromomycin (PM), ribostamycin (RM), (Be)kanamycin (BK) and apramycin (AP) (structures are shown in Scheme 6.2) were purchased from Sigma-Aldrich (St. Louis, MO) and dissolved in water to generate 1 mM

stock solutions. Complexes were prepared by mixing RNA with aminoglycoside antibiotics at a molar ratio of 2-4 (different ratios were used in order to obtain the 1:1 stoichiometry species of the complex as the most abundant component). Concentration of RNA was adjusted to 10-100 μM in order to get a signal-to-noise ratio of 10 in 8 scans.



Scheme 6.1. Structures of R1 (a) and R2 (b).



Scheme 6.2. Structures of aminoglycoside antibiotics: (a) PM, (b) RM, (c) AP, (d) BK.

6.2.2. Mass Spectrometry

All mass spectra were collected with an actively shielded 7 Tesla Q-FT-ICR mass spectrometer with a quadrupole front end (APEX-Q, Bruker Daltonics, Billerica, MA) in negative ion mode as described in the previous Chapters. Complexes were electrosprayed at 50 µL/h through an external Apollo II ion funnel (Bruker Daltonics) in a spray solution consisting of 25% (vol/vol) methanol (Fisher, Fair Lawn, NJ) and 50 mM NH_4OAc (Fisher). Precursor ions of 6- charge state were selected by the quadrupole and accumulated externally in the collision cell for 4 s in the presence of D_2S (Cambridge

Isotope Laboratories, Andover, MA) at a pressure of ~ 5 x 10^{-6} mbar (gauge calibrated for nitrogen). Multiple ICR cell fills were not used to avoid generation of mixtures of ion populations with different exposure time to D_2S (there is residual D_2S outside the storage hexapole). Gas-phase HDX reactions were performed for 4 to 34 s and mass spectra indicating mass shifts at each time point were recorded by XMASS (version 1.0.6 Bruker Daltonics). Data were processed with the MIDAS analysis software.[49]

Number (#) and percentage (%) of deuterium incorporation were calculated based on the average m/z of the entire isotopic distribution, i.e., by considering the relative abundance of each isotopic peak, according to the following equations

$$\#\ \text{deuterium incorporation} = [(m/z)_{obs.} - (m/z)_0] * z$$

$$\%\ \text{deuterium incorporation} = [(m/z)_{obs.} - (m/z)_0]/[(m/z)_{max.} - (m/z)_0]*100\%$$

in which $(m/z)_{obs.}$ = observed average m/z following a particular exchange time, $(m/z)_0$ = average m/z prior to HDX, $(m/z)_{max.}$ = expected average m/z at full deuteration of all exchangeable hydrogen and z is the charge state. Error bars were generated from three individual experiments acquired on the same day due to difficulties with reproducing the D_2S pressure in the collision cell.

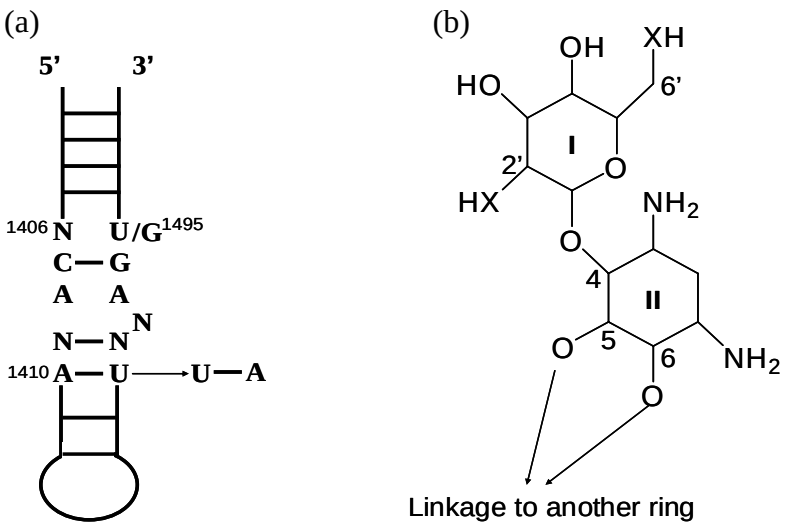
6.3. Results and Discussion

6.3.1. Conserved Structure Elements in A-site rRNA Complexes

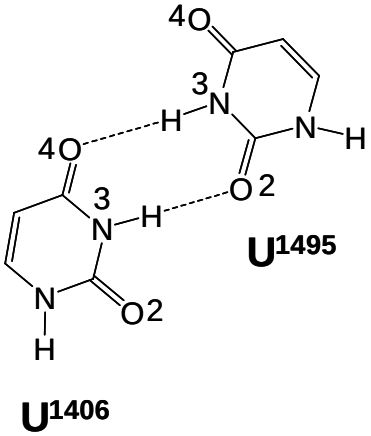
Purohit *et al.*[50] have shown by a chemical protection study that the small oligonucleotide model of rRNA (R1 and R2 in this case) can interact with mRNA, tRNA, and aminoglycoside antibiotics in a similar manner as that of the small subunit of the ribosome. NMR studies showed that the prokaryotic A-site rRNA is composed of two

helical stems with an asymmetric internal loop of three adenine residues.[51] In the absence of aminoglycoside antibiotics, this internal loop is closed by formation of a Hoogsteen base pair U^{1406}-U^{1495} and a W-C C^{1407}-G^{1494} base pair in the upper stem and another W-C base pair at positions 1409 and 1491 (specific sequence is not critical) in the lower stem.[52] A^{1408} is stacked within the helix, as are A^{1492} and A^{1493}. The structure of the internal loop is dynamic, as indicated by mixed sugar conformations.[52] A tetraloop UUCG (structure has been determined previously[53]) closes the lower stem.

PM belongs to the neomycin family (4,5-linked 2-deoxystreptamines) and is the most widely studied aminoglycoside antibiotic that binds in the major groove of the A-site rRNA. The universally conserved C^{1407}-G^{1494} base-pair, A^{1408}, A^{1493} and U/G^{1495} of the rRNA are required for specific binding (Scheme 6.3a). In addition, base pairing in the lower stem, and asymmetry of the internal loop resulting from the presence of A^{1492} are also required for specific binding.[51] Positions 1406 and 1495 in the upper stem are universally conserved uridines in the ribosome, and form the U-U pair[54], where the N3 and O4 of U^{1406} form hydrogen bonds with O2 and N3 of U^{1495} (Scheme 6.4). This U-U base pair can also be substituted by an A-U or U-G base pair. For the aminoglycoside antibiotic, NMR studies showed that the ring I and ring II of PM are essential for specific binding, particularly the two hydrogen bond donor groups at the 2' and 6' position of ring I and the two amino groups of ring II (Scheme 6.3b).[52] Ring III and ring IV contribute weakly to the interaction, suggesting RM, similar as PM but without ring IV, can bind to the A-site rRNA in a similar way as PM.



Scheme 6.3. (a) RNA (*E. coli*) sequence elements critical for specific aminoglycoside antibiotics binding to the A-site: N, any nucleotide; N—N, any W-C base pair. For position 1495, either a U or a G results in high binding affinity, and each of these two nucleotides presents a hydrogen bond acceptor in the major groove. (b) Conserved structural elements for aminoglycoside antibotics binding to the *E. coli* A-site RNA. X=NH or O.



Scheme 6.4. Schematic diagram of the U-U base pair conformation.

6.3.2. Gas-Phase HDX of R1 and its Complexes

NMR studies[52, 55] showed that there is a conformational change in the A-site RNA when it binds with aminoglycoside antibiotics and thermodynamic studies[56] suggested that this conformational change enhances the thermal stability of the A-site RNA because it is an exothermic process (favorable). In the previous Chapter, gas-phase HDX was shown to be able to correlate HDX rates of nucleic acid hairpins with their solution-phase stabilities. This same method is applied here to characterize the conformational change / stability of R1 and its complexes. Both the number (#) and the percentage (%) of deuterium incorporation (Figure 6.5 and 6.6) of the free and bound RNAs are measured for comparison, and the relative HDX trends are the same using these two algorithms. Results show that upon binding with either PM or RM, less # or % of deuterium is incorporated compared to the free R1 (Figures 6.5a and 6.6a), indicating that the bound RNA has a higher stability than the free form. This result also suggests that the conformational change caused by ligand binding induces a more compact structure of the RNA, which protects more hydrogens from exchanging with deuterium. Comparing PM and RM, the former appears to have a more stable complex structure with R1 than the latter (within 34 s, 10 hydrogens were exchanged for the R1-PM complex vs. 12 for the R1-RM complex), consistent with its lower dissociation constant calculated by titration method.[35] However, the absolute difference revealed by gas-phase HDX is different from that obtained by titration because the latter presents a >100-fold difference in their dissociation constants (16 μM for the R1-RM complex and 0.11 μM for the R1-PM complex).[35]

6.3.3. Gas-Phase HDX of R2 and its Complexes

R2 is the eukaryotic counterpart of R1 with a similar structure. The major difference between these two RNAs is the position 1408 in the A site: all eukaryotes have guanine at this position, whereas in prokaryotes it is adenine.[57] Additionally, R2 has a mismatched base pair in the lower stem (C^{1409}-A^{1491}). Therefore, R2 lacks the required structural elements for binding with the neomycin family. Griffey *et al*.[35] have shown that there is no complex formed between R2 and PM or RM, however, they showed that R2 can bind to AP, which belongs to the mono-substituted 2-deoxystreptamine class, with high affinity. Upon binding with AP, the residues in the internal loop (G^{1408}, A^{1492} and A^{1493}) are protected, similar to the protection at the A-site of PM bound R1.[35] BK, which belongs to the 4,6-linked 2-deoxystreptamine class of aminoglycoside antibiotics, can also bind to R2, but with a low affinity because the non-canonical C^{1409}-A^{1491} base pairing induces the shift of A^{1491} towards the major groove, which introduces unfavorable steric contacts between the RNA and the ring I of the ligand.[58] The dissociation constants of the R2-AP and R2-BK complexes are 0.5 and 1.6 μM, respectively[35] and their gas-phase HDX curves (Figures 6.5b and 6.6b) show a similar trend, i.e., the R2-AP complex exchanges slower than the R2-BK complex, indicating that a more compact structure is induced upon binding of AP compared to BK. The BK binding seems to have little effect on the R2 conformational change because the R2-BK complex exchanges at almost the same rate as the free R2.

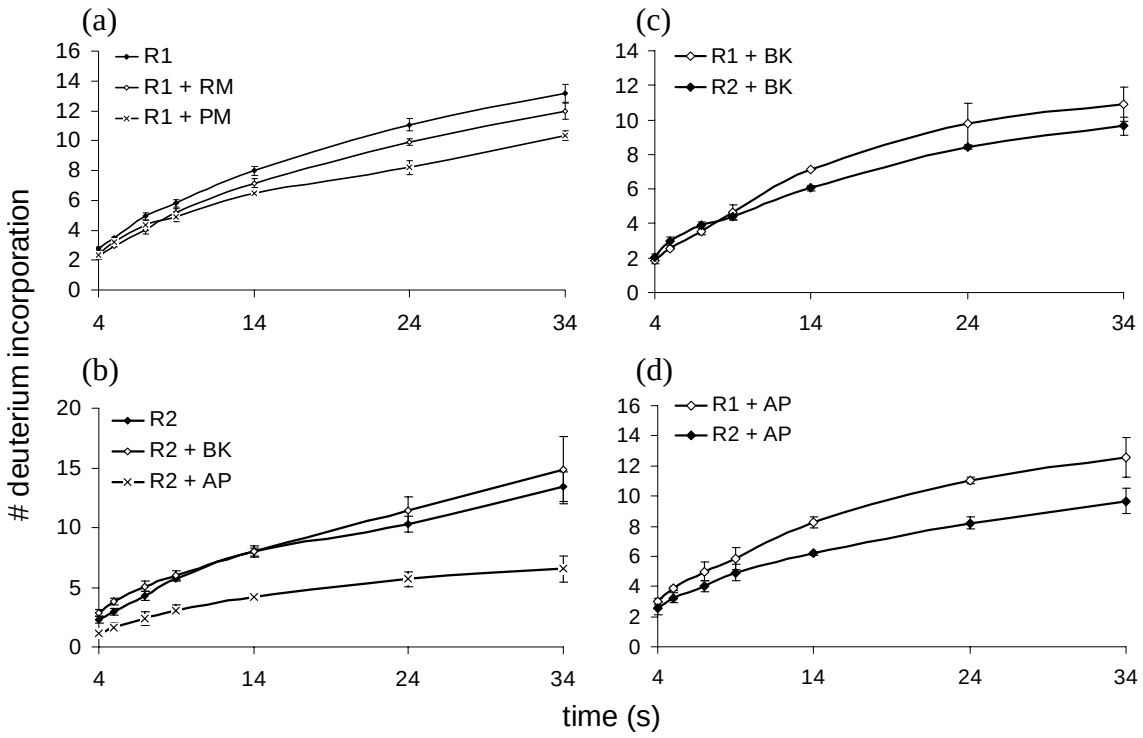


Figure 6.1. Comparison of gas-phase HDX of R1, R2 and their complexes in terms of # of deuterium incorporation.

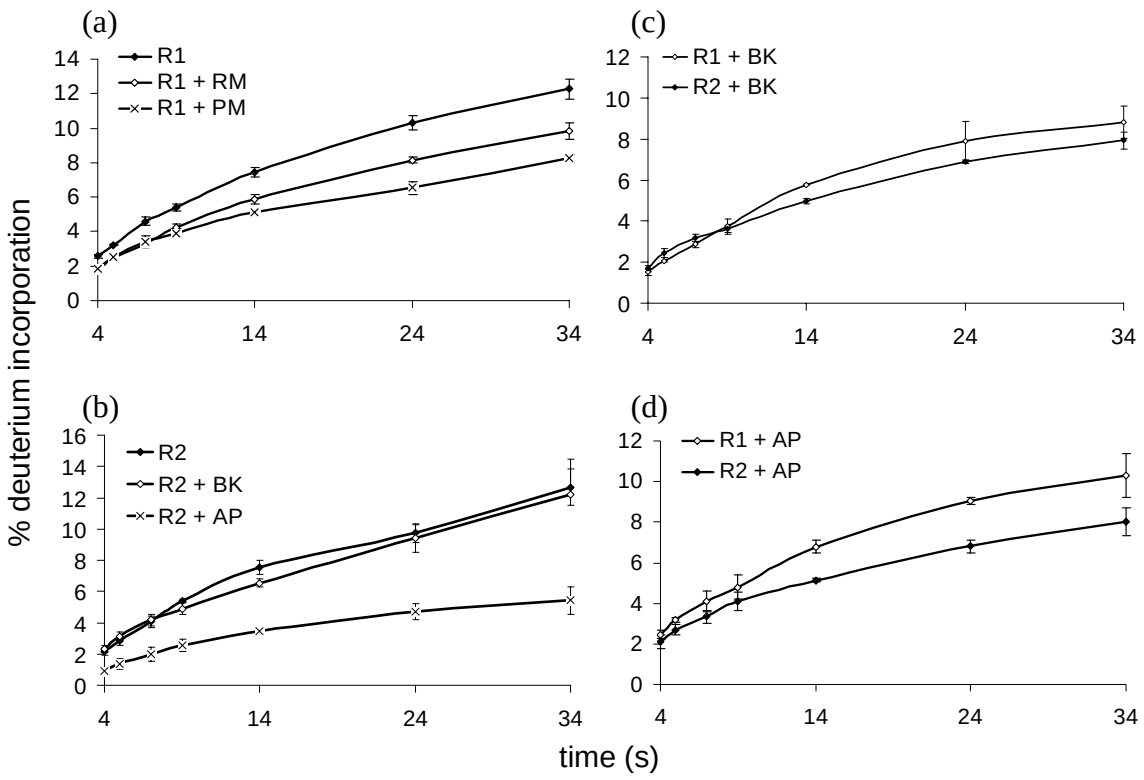


Figure 6.2. Comparison of gas-phase HDX of R1, R2 and their complexes in terms of % of deuterium incorporation.

6.3.4. Comparison of Gas-Phase HDX between R1 and R2 bound to the Same Ligand

In addition to comparing the gas-phase HDX of the same RNA with different ligands, we also compared gas-phase HDX of the same ligand binding with different RNA substrates. Figures 6.5c, 6.5d, 6.6c, and 6.6d show that both BK and AP can bind with R1 and R2 and a comparison between them demonstrates faster HDX rates of R1 complexes compared to the corresponding R2 complexes. This behavior is consistent with their relative dissociation constant values (2 μM for R1-AP complex, 0.5 μM for R2-AP complex, 2 μM for R1-BK complex, and 1.6 μM for R2-BK complex)[35] and

suggests less conformational change is induced in R1 compared to R2 when binding with the same aminoglycoside antibotic.

6.4. Conclusion

Conformational changes are observed when RNA binds with aminoglycoside antibiotics and gas-phase HDX is able to compare the conformational change / stability of different RNA-aminoglycoside antibiotic complexes. Faster HDX process correlates with smaller conformational change or lower stability of the complex. Results show that PM and RM, which belong to the 4,5-linked 2-deoxystreptamine class of aminoglycoside antibiotics, can bind prokaryotic A-site rRNA better than the eukaryotic rRNA, while BK, which belongs to the 4,6-linked 2-deoxystreptamine class, and AP, which belongs to the mono-substituted 2-deoxystreptamine class of aminoglycoside antibiotics, can bind the eukaryotic A-site rRNA better than the prokaryotic rRNA. All data show a trend of HDX rates consistent with that of their solution-phase dissociation constants calculated from titration experiments, although the absolute difference is different between the two techniques. Efforts to exponentially fit the observed HDX curves were inconclusive, likely because there are several (>10) groups of hydrogens with unique HDX behavior, thereby requiring a high number of exponential terms. Overall, gas-phase HDX provides a fast and direct estimation of the conformational change and stability of RNA complexes.

6.5. Bibliography

(1) Wallis, M. G.; Schroeder, R. *Prog. Biophys. Mol. Biol.* **1997**, *67*, 141-154.

(2) Wank, H.; Schroeder, R. *J. Mol. Biol.* **1996**, *258*, 53-61.

(3) Hoch, I.; Berens, C.; Westhof, E.; Schroeder, R. *J. Mol. Biol.* **1998**, *282*, 557-569.

(4) Tor, Y.; Hermann, Y.; Westhof, E. *Chem. Biol.* **1998**, *5*, R277-R283.

(5) Cundcliffee, E. *Recognition sites for antibiotics within rRNA. In the Ribosome: Structure, Function and Evolution*; Am. Soc. Microbiol.: Washington DC, **1990**, 479-490.

(6) Moazed, D.; Noller, H. F. *Nature* **1987**, *327*, 389-394.

(7) Davies, J.; Gorini, L.; Davis, B. D. *Mol. Pharmacol.* **1965**, *1*, 93-106.

(8) Palmer, E.; Wilhelm, J. M. *Cell* **1978**, *13*, 329-334.

(9) Wilhelm, J. M.; Pettitt, S. E.; Jessop, J. J. *Biochemistry* **1978**, *17*, 1143-1149.

(10) Sanner-Lowery, K. A.; Griffey, R. H.; Hofstadler, S. A. *Anal. Biochem.* **2000**, *280*, 264-270.

(11) Winzor, D. J.; Sawyer, W. H. *Quantitative Characterization of Ligand Binding*; Wiley, **1995**, 176.

(12) Mei, H.-Y.; Mack, D. P.; Galan, A. A.; Halim, N. S.; Heldsinger, A.; Loo, J. A.; Moreland, D. W.; Sannes-Lowery, K. A.; Sharmeen, L.; Truong, H. N.; Czarnik, A. W. *Bioorg. Med. Chem.* **1997**, *5*, 1173-1184.

(13) Mei, H. Y.; Cui, M.; Heldsinger, A.; Lemrow, S. M.; Loo, J. A.; Sannes-Lowery, K. A.; Sharmeen, L.; Czarnik, A. W. *Biochemistry* **1998**, *37*, 14204-14212.

(14) Sannes-Lowery, K. A.; Mei, H.-Y.; Loo, J. A. *Int. J. Mass Spectrom.* **1999**, *193*, 115-122.

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

(16) Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F. *Mass Spectrom. Rev.* **1990**, *9*, 37-70.

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

(18) Mo, J.; Hakansson, K. *Anal. Bioanal. Chem.* **2006**, *386*, 675-681.

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

(20) Greig, M. J.; Hans Gaus, H.; Cummins, L. L.; Sasmor, H.; Griffey, R. H. *J. Am. Chem. Soc.* **1995**, *117*, 10765-10766.

(21) Lim, H. K.; Hsieh, Y. L.; Ganem, B.; Henion, J. *J. Mass. Spectrom.* **1995**, *30*, 708-714.

(22) Gao, J., Cheng, X., Chen, R., Sigal, G. B., Bruce, J. E., Schwartz,; B. L., H., S. A., Anderson, G. A., Smith, R. D., and; Whitesides, G. M. *J. Med. Chem.* **1996**, *39*, 1949-1955.

(23) Wu, Q.; Cheng, X.; Hofstadler, S. A.; Smith, R. D. *J. Mass. Spectrom.* **1996**, *31*, 669-675.

(24) Jorgensen, T. J. D.; Roepstorff, P. *Anal. Chem.* **1998**, *70*, 4427-4432.

(25) Ayed, A.; Krutchinsky, A. N.; Ens, W.; Standing, K. G.; Duckworth, H. W. *Rapid Commun. Mass Spectrom.* **1998**, *12*, 339-344.

(26) Cheng, X.; Chen, R.; Bruce, J. E.; Schwartz, B. L.; Anderson, G. A.; Hofstadler, S. A.; Gale, D. C.; Smith, R. D.; Gao, J.; Sigal, G. B.; Mammen, M.; Whitesides, G. M. *J. Am. Chem. Soc.* **1995**, *117*, 8859-8860.

(27) Loo, J. A.; Hu, P.; McConnell, P.; Mueller, W. T. *J. Am. Soc. Mass Spectrom.* **1997**, *8*, 234-243.

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

(29) Loo, J. A.; Sannes-Lowery, K. A.; Hu, P.; Mack, D. P.; Mei, H.-Y. *NATO ASI Ser. C* **1998**, *510*, 83-99.

(30) Loo, J. A.; Sannes-Lowery, K. A. *Studying noncovalent interactions by electrospray ionization mass spectrometry.*, 2nd. ed.; CRC, **1998**, 469.

(31) Smith, R. D.; Bruce, J. E.; Wu, Q.; Lei, Q. P. *Chem. Soc. Rev.* **1997**, *26*, 191-202.

(32) Gao, Q. Y.; Cheng, X. H.; Smith, R. D.; Yang, C. F.; Goldberg, I. H. *J. Mass. Spectrom.* **1996**, *31*, 31-36.

(33) Wu, Q. Y.; Gao, J. M.; Josephmccarthy, D.; Sigal, G. B.; Bruce, J. E.; Whitesides, G. M.; Smith, R. D. *J. Am. Chem. Soc.* **1997**, *119*, 1157-1158.

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

(35) Griffey, R. H.; Hofstadler, S. A.; Sannes-Lowery, K. A.; Ecker, D. J.; Crooke, S. T. *Proc. Nat. Acad. Sci. U.S.A.* **1999**, *96*, 10129-10133.

(36) Chipuk, J. E.; Brodbelt, J. S. *J. Am. Soc. Mass Spectrom.* **2007**, *18*, 724-736.

(37) Crestoni, M. E.; Fornarini, S. *J. Mass Spectrom.* **2003**, *38*, 854-861.

(38) Robinson, J. M.; Greig, M. J.; Griffey, R. H.; Mohan, V.; Laude, D. A. *Anal. Chem.* **1998**, *70*, 3566-3571.

(39) Griffey, R. H.; Greig, M. J.; Robinson, J. M.; Laude, D. A. *Rapid Commun. Mass Spectrom.* **1999**, *13*, 113-117.

(40) Hofstadler, S. A.; Sannes-Lowery, K. A.; Griffey, R. H. *J. Mass Spectrom.* **2000**, *35*, 62-70.

(41) Gabelica, V.; Rosu, F.; Witt, M.; Baykut, G.; de Pauw, E. *Rapid Commun. Mass Spectrom.* **2005**, *19*, 201-208.

(42) Freitas, M. A.; Marshall, A. G. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 780-785.

(43) Green, M. K.; Lebrilla, C. B. *Mass Spectrom. Rev.* **1997**, *16*, 53-71.

(44) Freitas, M. A.; Hendrickson, C. L.; Emmett, M. R.; Marshall, A. G. *J. Am. Soc. Mass Spectrom.* **1998**, *9*, 1012-1019.

(45) Campbell, S.; Rodgers, M. T.; Marzluff, E. M.; Beauchamp, J. L. *J. Am. Chem. Soc.* **1994**, *116*, 9765.

(46) Gard, E.; Green, M. K.; Bregar, J.; Lebrilla, C. B. *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 623-631.

(47) Wyttenbach, T.; Bowers, M. T. *J. Am. Soc. Mass Spectrom.* **1999**, *10*, 9-14.

(48) Mo, J.; Hakansson, K. *Anal. Chem.* **2007**, *79*, 7893-7898.

(49) Senko, M. W.; Canterbury, J. D.; Guan, S.; Marshall, A. G. *Rapid Commun. Mass Spectrom.* **1996**, *10*, 1839-1844.

(50) Purohit, P.; Stern, S. *Nature* **1994**, *370*, 659-662.

(51) Recht, M. I.; Fourmy, D.; Blanchard, S. C.; Dahlquist, K. D.; Puglisi, J. D. *J. Mol. Biol.* **1996**, *262*, 421-436.

(52) Fourmy, D.; Recht, M. I.; Blanchard, S. C.; Puglisi, J. D. *Science* **1996**, *274*, 1367-1371.

(53) Allain, F. H.-T.; Varani, G. *J. Mol. Biol.* **1995**, *250*, 333-353.

(54) Baeyens, K. J.; De Bondt, H. L.; Holbrook, S. R. *Nature Struct. Biol.* **1995**, *2*, 56-62.

(55) Fourmy, D.; Yoshizawa, S.; Puglisi, J. D. *J. Mol. Biol.* **1998**, *277*, 333-345.

(56) Kaul, M.; Pilch, D. S. *Biochemistry* **2002**, *41*, 7695-7706.

(57) Van de Peer, Y.; Van den Broeck, I.; De Rijk, P.; De Wachter, R. *Nucleic Acids Res.* **1994**, *22*, 3488-3494.

(58) Zimmermann, G. R.; Shields, T. P.; Jenison, R. D.; Wick, C. L.; Pardi, A. *Biochemistry* **1998**, *37*, 9186-9192.

CHAPTER 7

CONCLUSION AND FUTURE OUTLOOK

MS has evolved and grown to become one of the most versatile and informative tools for investigating molecular structure, bonding, reactivity, equilibrium, and energetics in the gas phase.[1] It is increasingly perceived to be an essential tool in the drug discovery process for many of the key steps in the development of novel therapeutics, including the assessment of compound purity; quantitation of cellular uptake, distribution, metabolism, and excretion; and compound-specific pharmacokinetic analyses.[2-6] More recently, it has emerged as an effective technique for identifying lead compounds on the basis of characterizing non-covalent macromolecular-ligand interactions.[6] This approach offers several attractive properties for screening applications in drug discovery, including the small quantities of target and ligands required, and the capacity to study ligands or targets without having to label them. ESI MS is demonstrated to be a rapid screening method for identification of active compounds from crude mixtures.[7, 8] It allows the simultaneous analysis of mixtures of compounds based on their unique molecular masses. Active compounds can be identified directly from their non-covalent complexes with target molecules and control targets can be included in the screening mixture to provide a measure of binding

specificity. In addition, tandem MS measurements can be used to gain insight into the composition and structure of the binding species.

This dissertation focuses on method development using MS for characterization of nucleic acid non-covalent interactions. Two strategies have been utilized, one is tandem MS (up to MS^3) and the other is gas-phase HDX combined with MS.

7.1. Summary of Results

For the first time, we investigated the mechanism of oligonucleotide gas-phase HDX in both positive and negative ion mode. Relay mechanisms are proposed for both cases and the HDX rates of nucleic acids are dependent on the gas-phase basicity (for positive ion mode) and acidity (for negative ion mode) of the nucleobases as well as their structural flexibility. RNA shows a faster exchange process than the corresponding DNA due to the additional 2' hydroxyl group on the backbone sugar ring. Gas-phase HDX of several nucleic acid derivatives was compared with that of natural nucleic acids in order to verify the involvement of each individual functional group. Results demonstrate that the backbone phosphate groups, nucleobases, and 2' hydroxyl groups on the ribose rings of RNA are all involved in HDX.

After elucidating the mechanism of nucleic acid gas-phase HDX, we examined the ability of this technique to characterize higher order structures of nucleic acids as well as their complexes with small molecules. Gas-phase HDX of a DNA duplex underwent a slower HDX process than its constituent monomers, indicating that hydrogen bonds, which represent one of the key factors for forming nucleic acid higher order structure, can protect hydrogens from exchanging. Utilizing this characteristic of gas-phase HDX,

we compared a series of nucleic acid hairpins and found that the ones with higher stabilities in solution have lower HDX rates than those with lower stabilities. In addition, data for a three-nucleotide-loop hairpin and an isomeric four-nucleotide-loop hairpin suggest that this technique might be able to provide quantitative analysis of the structural stability.

Applying the gas-phase HDX technique to characterize the binding between RNA and different aminoglycoside antibiotics, we found that complexes with higher binding affinities (lower dissociation constants) in solution exchange slower compared to those with lower binding affinities. Our data show a trend of HDX rates for these complexes, which is consistent with their dissociation constants calculated from titration experiments, although the absolute values may be different. Overall, gas-phase HDX provides a fast and direct estimation of the stability of RNA complexes.

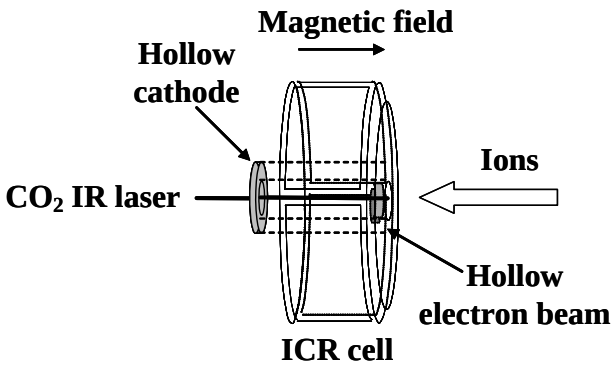
Tandem MS is another method investigated in this thesis work to characterize nucleic acid higher order structure. By applying a range of different tandem MS techniques, including EDD, IRMPD, AI EDD, and EDD/IRMPD MS^3, three isomeric 15-mer DNAs with different sequences and predicted solution-phase structures were characterized. We found that all three 15-mers had higher order structures in the gas phase although preferred structures were only predicted for two of them in solution. Nevertheless, experiments utilizing the radical ion chemistry of EDD (EDD, AI EDD, and EDD/IRMPD MS^3) yielded different cleavage patterns with less backbone fragmentation for the most stable solution-phase structure as compared to the other two. By contrast, no major differences were observed in IRMPD although the extent of backbone cleavage was higher for all the three 15-mers. This work demonstrates that the

presence of higher order structures can be determined by these fragmentation techniques as well as some information on their stability and that EDD is a valuable complementary fragmentation method compared to traditional MS/MS techniques, including IRMPD, in that it can generate additional structural information.

7.2. Prospects for Future Work

7.2.1. Improving the Efficiency of AI EDD and EDD/IRMPD MS^3

In Chapter 4, we applied different tandem MS techniques, including EDD, IRMPD, AI EDD, and EDD/IRMPD MS^3, to characterize the higher order structure of nucleic acids. Results suggest that by combining EDD, which involves radical ion chemistry, with traditional slow heating methods, such as IRMPD, complementary structural information can be obtained. However, limited numbers of fragments are observed in both the AI EDD and EDD/IRMPD MS^3 spectra, indicating low efficiencies of these two techniques. This low efficiency is probably due to poor overlap of the IR laser beam with the electron beam and/or with the analyte ion cloud because combined IRMPD/EDD is optimized only when the IR and electron beams have complete overlap with the ion cloud in the ICR cell. Scheme 7.1 shows our current instrumental setup for combined IRMPD and EDD events in the ICR cell.



Scheme 7.1. Schematic drawing of the setup for combined IRMPD and EDD FT-ICR MS (adopted from[9]).

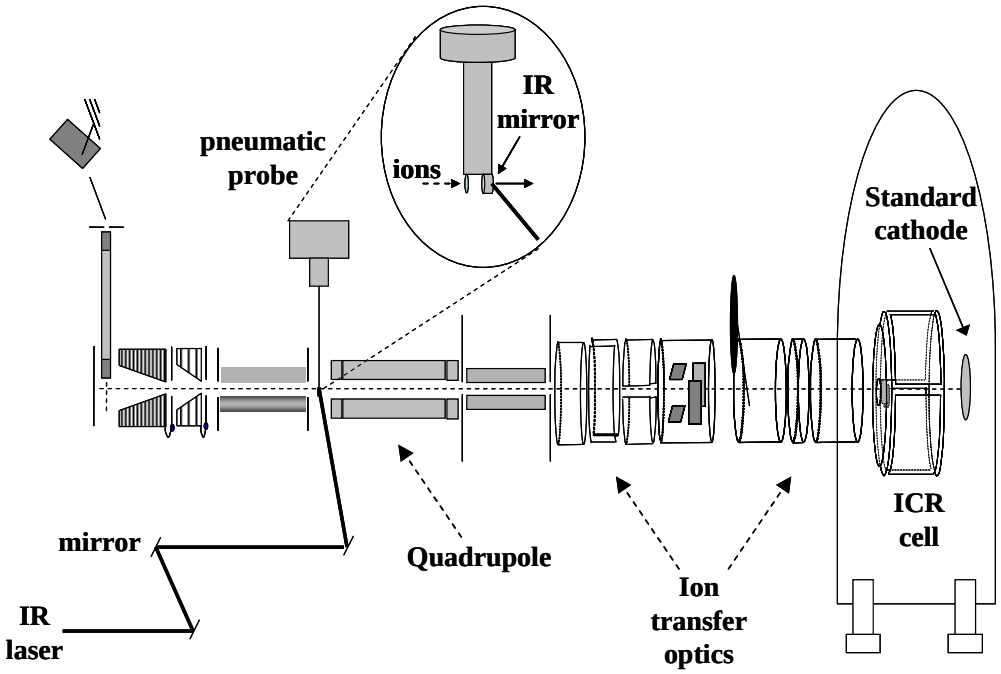
In this setup, an electron injection system based on an indirectly heated hollow dispenser cathode (donut-shaped) has been installed in the FT-ICR mass spectrometer. This cathode generates a hollow electron beam coaxial with the ICR cell axis. Additionally, an on-axis IR laser beam is positioned at the back of the cathode and passes through the hole at the centre of the cathode. This setup enables the consecutive or simultaneous use of the EDD and IRMPD methods within a single experimental sequence, however, it is difficult to have both events occur with high efficiency. For example, during EDD/IRMPD MS^3, if the ion cloud enters the ICR cell on-axis, it would have only a boundary interaction region with the electron beam because, in the strong magnetic field, the electron beam maintains the ring cross section of the cathode throughout the ICR cell. Therefore, the overlap of ions and electrons may be poor and EDD efficiency is therefore also low. It is possible to increase EDD efficiency by using sidekick trapping (which "kicks" the ions off-axis during entry to the cell), or ion off-resonance excitation (which also makes the ion cloud move off-axis),[10, 11] however, such excitation would

reduce the overlap of the on-axis IR beam with the ion cloud. In AI EDD, although the ion cloud is activated by the IR laser first, which may cause it to move off-axis, its overlap with the electron beam is still poor because it is difficult to control the precise excitation energy, which is dependent on analyte structure, to excite the ions exactly to the electron beam region.

Håkansson *et al.*[12] have combined the complementary capabilities of ECD (same instrumental setup as EDD) and IRMPD by utilizing an on-axis non-hollow dispenser cathode together with an angled IR laser beam. However, introduction of an angled beam requires an open ICR cell whereas our instrument has a closed cell. Thus, this combined implementation of ECD/EDD and IRMPD is not compatible with our instrument.

One solution to improve the overlap of both the electron and laser beams with the ion cloud on our instrument may be to collisionally cool the ion cloud to an on-axis position prior to fragmentation, which can be achieved by quadrupolar axilization,[13, 14] and have both the electron beam (not hollow shaped) and the IR laser beam positioned on-axis. The problem of this design is that both the IR and electron beams are introduced from the rear end of the ICR cell and, thus, they cannot both be on-axis if the cathode is not hollow. Mihalca *et al.*[15] demonstrated a novel FT-ICR MS setup, which allows simultaneous on-axis IRMPD and ECD in the ICR cell. A pneumatic probe was inserted into the ion-optical path, thereby enabling on-axis IR irradiation of the ion cloud from the front of the ICR cell simultaneously with electron irradiation, provided by a standard dispenser cathode (not hollow), from the rear of the ICR cell. Similar modifications could be done to our instrument. The key to this approach is the ability to direct the IR laser beam through the ion transfer optics to the ICR cell after the ions have been

transferred to the cell, specifically, the ability to align the IR laser beam along the axis, and the reproducibility of this optimum alignment. Such alignment can be achieved by inserting a pneumatically actuated IR mirror, which is attached to a guided pneumatic probe, between ion-optical elements (in front of the quadrupole) to reflect the IR beam along the axis of the instrument and into the ICR cell (Figure 7.2). This geometrical design would ensure good overlap between the electron beam, IR beam, and ion cloud. The flexibility to carry out both independent and combined EDD and IRMPD experiments with high efficiency will result in significant increase in product ions and thus more structural information.



Scheme 7.2. Schematic of an FT-ICR mass spectrometer with a pneumatic probe and the resulting path of an IR laser beam entering from the front of the ICR cell (adopted from[15]).

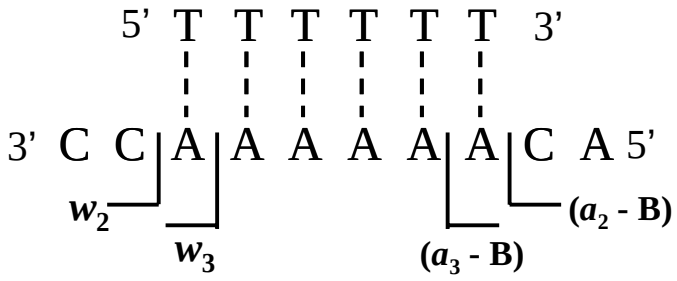
7.2.2. Investigating Hydrogen Scrambling during MS/MS

Chapters 5 and 6 showed that gas-phase HDX combined with MS provides valuable insights into nucleic acid higher order structure as well as their binding properties with small molecules. Both of these applications are based on HDX data of the entire molecule, which do not provide local information. It would be more desirable to be able to monitor deuterium exchange levels at each individual nucleotide. Previous attempts have been made to localize the HDX level for individual amino acids in proteins. For example, CAD has been used to fragment deuterated peptic peptides into shorter pieces and the observed N-terminal fragments yielded deuterium levels consistent with NMR results,[16, 17] which indicates no "hydrogen scrambling" in these fragments. However, the deuterium is apparently "scrambled" in most C-terminal fragments. Other work has shown that 100% scrambling appears to occur in both fragment types.[18] Kaltashov and Eyles[19-21] have discussed the use of ECD to eliminate the scrambling problem because it might be possible to reduce the degree of hydrogen migration due to the fast fragmentation pathway of ECD that may prevent migration of deuteriums (non-ergodic process).[22-24] Kweon *et al.*[25] have shown that the HDX rates of consecutive c-type ion pairs generated by ECD fragmentation of melittin have some correlation with previous NMR data, where the amide hydrogens of leucine 13 and alanine 15, located at an unstructured kink, appear as fast exchangers while the amide hydrogens of leucine 16 and lysine 23, buried within the helical regions, appear as slow exchangers. However, calculations based on c-type ions for other amide hydrogens do not correlate well with NMR data, and evidence for deuterium scrambling in ECD was obtained from z-type ions. In a recent paper, Rand *et al.*[26] stated that a low degree of intramolecular migration of

peptide amide hydrogens was observed in ECD, provided that vibrational excitation prior to ECD, including ESI and quadrupole isolation processes, is minimized.

Similar attempts can be made to elucidate whether hydrogen scrambling occurs in MS/MS of deuterated nucleic acids. Low degree of hydrogen scrambling may be expected because deprotonated analyte species have no mobile protons. A DNA duplex (Scheme 7.3) has been designed to investigate this issue. In this duplex, six A-T base pairs (this number of base pairs facilitates duplex formation as well as provides a sufficient number of fragments in MS/MS for data analysis) are used to form a duplex instead of G-C base pairs because oligonucleotides containing multiple G residues are not very stable in solution. In order to probe whether hydrogen scrambling occurs, additional nucleobases need to be added to one of the strands because all the T and A residues are presumably protected from HDX by hydrogen bonding. With the dangling residues, we can apply different fragmentation techniques (CAD, IRMPD and EDD) to cleave the strand at the backbone and large HDX extent would be expected for fragments with less A or T (involved in base pairing) if there is no hydrogen scrambling. The additional residues (including three C and one A) are added at the two ends of the poly-A strand due to the following four reasons: first, they eliminate other possible base pairing between these two strands and ensure the unique formation of the duplex; second, C and A have similar HDX rates, which simplifies data analysis based on nucleobase effects (as discussed in chapter 2);[27] third, an asymmetric sequence is used at the two ends in order to differentiate between 3' and 5' fragment ions; fourth, two nucleotides were added on each side because adding more nucleotides would result in a long strand that generates too many fragments for HDX data analysis and adding only one nucleotide would

generate mononucleotide (w_1 and (a_1 - B)) ions that are difficult to observe. Lower degree of hydrogen scrambling is expected in EDD than in slow heating MS/MS methods, such as CAD and IRMPD, because EDD is related to ECD.[28] The ionization conditions in ESI and the ion transfer conditions in the quadrupole and ion transfer optics need to be optimized to avoid excitation of the duplex prior to MS/MS (see discussion in the Appendix).



Scheme 7.3. A DNA duplex designed for investigating the degree of hydrogen scrambling in MS/MS of nucleic acids.

7.3. Bibliography

(1) Bowers, M. T.; Marshall, A. T.; McLafferty, F. W. *J. Phys. Chem.* **1996**, *100*, 12897-12910.

(2) Zhang, J.; McCombie, G.; Guenat, C.; Knochenmuss, R. *Drug Discov. Today* **2005**, *10*, 635-642.

(3) Siegel, M. M. *Curr. Topics Med. Chem.* **2002**, *2*, 13-33.

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

(5) Hofstadler, S. A.; Griffey, R. H. *Curr. Opin. Drug Discov. Dev.* **2000**, *3*, 423-431.

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

(7) Cummins, L. L.; Chen, S.; Blyn, L. B.; Sannes-Lowery, K. A.; Drader, J. J.; Griffey, R. H.; Hofstadler, S. A. *J. Nat. Prod* **2003**, *66*, 1186-1190.

(8) Gooding, K. B.; Higgsa, R.; Hodgea, B.; Stauffera, E.; Heinza, B.; McKnighta, K.; Phippsa, K.; Shapiroa, M.; Winklera, M.; Nga, W.-L.; Juliana, R. K. *J. Am. Soc. Mass Spectrom.* **2004**, *15*, 884-892.

(9) Tsybin, Y. O.; Witt, M.; Baykut, G.; Kjeldsen, F.; Hakansson, P. *Rapid Commun. Mass Spectrom.* **2003**, *17*, 1759-1768.

(10) Gorshkov, M. V.; Masselson, C. D.; Nikolaev, E. N.; Udseth, H. R.; Pasa-Tolic, L.; Smith, R. D. *Int. J. Mass Spectrom.* **2004**, *234*, 131-136.

(11) Mormann, M.; Peter-Katalinic, J. *Rapid Commun. Mass Spectrom.* **2003**, *17*, 2208-2214.

(12) Hakansson, K.; Chalmers, M. J.; Quinn, J. P.; McFarland, M. A.; Hendrickson, C. L.; Marshall, A. G. *Anal. Chem.* **2003**, *75*, 3256-3262.

(13) Guan, S.; Kim, H. S.; Marshall, A. G.; Wahl, M. C.; Wood, T. D.; Xiang, X. *Chem. Rev.* **1994**, *94*, 2161-2182.

(14) Hendrickson, C. L.; Drader, J. J.; Laude, D. A., Jr. *J. Am. Soc. Mass Spectrom.* **1995**, *6*, 448-452.

(15) Mihalca, R.; van der Burgt, Y. E. M.; McDonnell, L. A.; Duursma, M.; Cerjak, L.; Heck, A. J.; Heeren, R. M. A. *Rapid Commun. Mass Spectrom.* **2006**, *20*, 1838-1844.

(16) Deng, Y. Z.; Pan, H.; Smith, D. L. *J. Am. Chem. Soc.* **1999**, *121*, 1966-1967.

(17) Kim, M.-Y.; Maier, C. S.; Reed, D. J.; Deinzer, M. L. *J. Am. Chem. Soc.* **2001**, *123*, 9860-9866.

(18) Jorgensen, T. J.; Gardsvoll, H.; Ploug, M.; Roepstorff, P. *J. Am. Chem. Soc.* **2005**, *127*, 2785-2793.

(19) Kaltashov, I. A.; Eyles, S. J. *Mass Spectrom. Rev.* **2002**, *21*, 37-71.

(20) Kaltashov, I. A.; Eyles, S. J. *J. Mass Spectrom.* **2002**, *37*, 557-565.

(21) Eyles, S. J.; Kaltashov, I. A. *Methods* **2004**, *34*, 88-99.

(22) Turecek, F.; McLafferty, F. W. *J. Am. Chem. Soc.* **1984**, *106*, 2525-2528.

(23) Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. *J. Am. Chem. Soc.* **1998**, *120*, 3265-3266.

(24) Horn, D. M.; Ge, Y.; McLafferty, F. W. *Anal. Chem.* **2000**, *72*, 4778-4784.

(25) Kweon, H. K.; Håkansson, K. *Analyst* **2006**, *131*, 275-280.

(26) Rand, K. D.; Adams, C. M.; Zubarev, R. A.; Jorgensen, T. J. *J. Am. Chem. Soc.* **2008**, *130*, 1341-1349.

(27) Mo, J.; Hakansson, K. *Anal. Chem.* **2007**, *79*, 7893-7898.

(28) Budnik, B. A.; Haselmann, K. F.; Zubarev, R. A. *Chem. Phys. Lett.* **2001**, *342*, 299-302.

APPENDIX

IMPROVING SENSITIVITY AND ACHIEVING SOFTER TRANSMIMSSION CONDITIONS BY ION FUNNEL ESI

A.1. Introduction

The ability of producing intact molecular ions from high molecular weight compounds makes ESI an important tool for a wide variety of applications, including the analysis of biological molecules.[1-4] In ESI, ions are formed at atmospheric pressure and transferred to the lower pressure first region of the mass analyzer through a capillary (shown in Scheme 1.6 in Chapter 1). The presence of residual gas in this region of the mass spectrometer makes the pressure relatively high (about 1-10 Torr), resulting in radial expansion of the ion cloud due to gas collisions as well as Coulomb repulsion between ions. When passing through apertures (often in the form of skimmers) dividing the first pressure region from subsequent lower pressure regions, many ions are lost due to the small aperture size, which is required for maintaining differential pumping. This low ion transmission efficiency leads to low sensitivity of MS analysis. From ion mobility studies it is known that more than 50% of the ESI current, generally around 50 to 200 nA, is potentially useful ion current,[3] while the ion current measured entering the

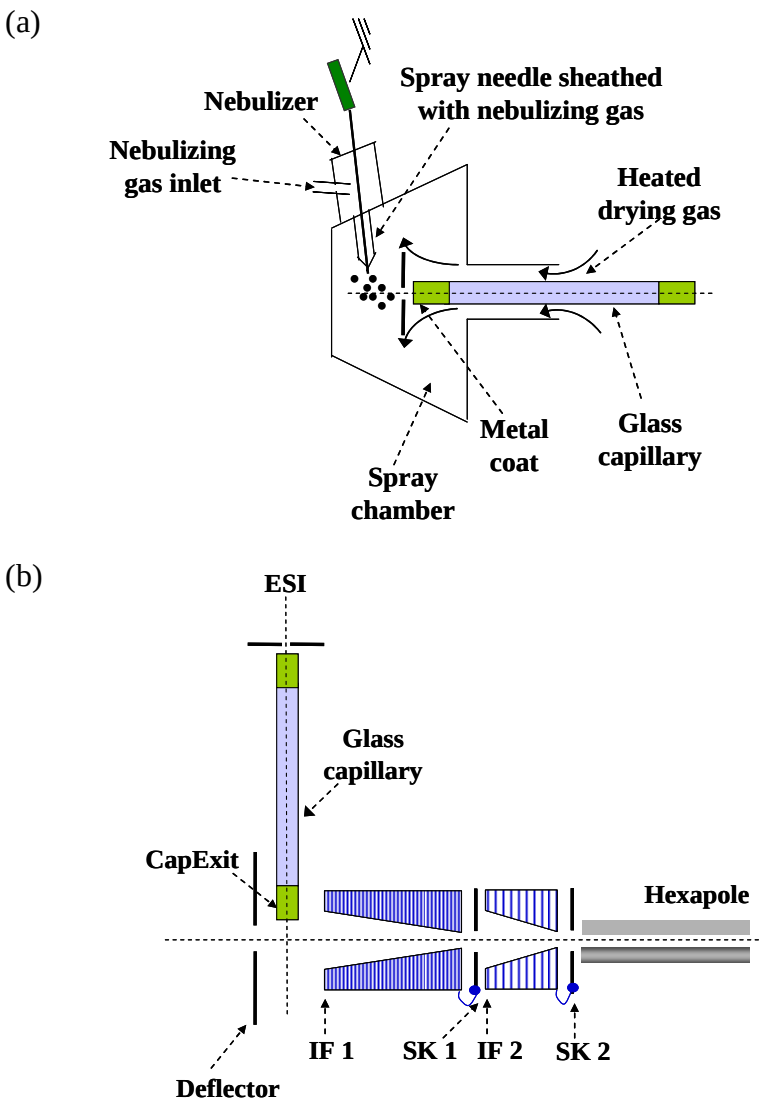
mass analyzer is typically about three orders of magnitude or more smaller.[5] The ability to effectively focus and transmit ions from relatively high pressure ion sources is the key factor that affects sensitivity and dynamic range in ESI MS. Conventional ion optics, which are based on electrostatic lenses, can effectively focus ions in a vacuum but they are ineffective at atmospheric pressure because of the extensive ion cloud expansion. One possible solution is to increase the inlet aperture size, however, this approach requires the use of pumps with larger pumping capacity. Even with this approach, a maximum possible aperture size is quickly reached because effective desolvation becomes more difficult and cooling effects, due to gas expansion, can be problematic.

It is well known that RF multipole devices can be used for ion focusing at low pressure (10^{-4} ~ 10^{-2} Torr),[6-10] however, they have limited use for more diffused ions at higher pressure (1-10 Torr) due to their relatively small acceptance aperture. Gerlich[11] demonstrated that ions can be effectively transmitted through a series of stacked cylindrical ring electrodes of fixed ring diameter using RF electric fields with opposite polarity on adjacent electrodes. In that arrangement, the electrodes can function as an ion 'pipe' to effectively confine the ion cloud because ion-neutral collisions lead to damping of the ion motion, thereby shrinking the ion cloud size. In 1997, Shaffer *et al.*[5, 12] introduced a novel ion funnel interface between the exit of the inlet capillary and the first differentially pumped region of the mass spectrometer, which improved ion transmission efficiency, and thus the ESI MS sensitivity, by over an order of magnitude compared to the normal configuration (capillary-skimmer inlet). Generally, an ion funnel is comprised of a series of ring electrodes with progressively smaller diameter (for the purpose of collisional focusing at elevated pressure) where opposite RF phases are

applied to adjacent and alternating electrode elements (Scheme A1b). At the same time, an axial DC field is co-applied to the device through resistive coupling to each ring electrode, resulting in effective transmission of ions to the low pressure region through a relatively small exit aperture, which is compatible with the acceptance aperture of the RF multipole in the later stage of the mass spectrometer.

More work has been done to improve the focusing and transmission efficiency of the initial ion funnel configuration. For example, Shaffer *et al.*[13] developed a new design of the ion funnel on a triple quadrupole mass spectrometer. This design comprises larger internal diameters of the ring electrodes that taper down to a relatively large exit aperture and provided significant improvement in low m/z ion transmission. These authors also showed that by ramping the RF amplitude of the ion funnel together with the m/z scan of the quadrupole mass analyzer, the range of efficient m/z ion transmission was greatly improved. Another way to enhance transmission of low m/z ions is to reduce the thickness of the ring electrodes and the spacing between them.[14]

In our instrument, a dual ion funnel device was installed and a sensitivity increase of more than an order of magnitude was achieved in positive ion mode. Kim *et al.*,[14] found that labile non-covalent complexes could be transmitted by using low RF amplitude, while undesired adducts could be removed by using high RF amplitude; however, sensitivity decreased with decreasing RF amplitude. On our instrument, similar "soft" transmission conditions for non-covalent complexes can be realized by decreasing the DC voltage fields at the interfaces between each two adjacent inlets in the ion funnel.



Scheme A1. Configurations of standard ESI source (a) and ion funnel ESI source (b). The parts before entering the capillary in (b) are the same as those in (a) and simplified as "ESI".

A.2. Experimental Section

All experiments described here were performed on a 7 T Q-FT-ICR mass spectrometer (Bruker Daltonics, Billerica, MA) with either an external standard (Apollo I) ESI source, or an Apollo II ion funnel ESI source. Sample solution is dispersed into

multiple small charged droplets during ESI as a result of a high electrostatic field. The standard ESI configuration is shown in Scheme A1a. To achieve reasonable sensitivity in mass analysis of liquid samples, the sample solution is first sprayed into very fine droplets, from which solvent can be easily evaporated prior to entering the vacuum system. Such nebulization is best achieved with the use of a pneumatic nebulizer consisting of two concentric tubes with a common central axis, where the inner tube transports the liquid sample and the outer tube transports the pressurized nebulizing gas (N_2). The two streams meet at the end of the tubes and the solution is dispersed into droplets with uniform size, small enough to facilitate the following desolvation process. Evaporation of the solvent results in the droplets reducing in size and causes an increase in charge density on their surfaces and finally results in Coulombic forces strong enough to break up the droplets. This process repeats until complete desolvation has occurred, generating sample ions. Heated drying gas (N_2), flowing in the opposite direction of the droplet stream, is used to aid volatilization and to carry away any uncharged material. The ions, e.g., positively charged, are attracted in the forward direction by an electrical field between the spray needle (ground potential) and the negatively biased metal-coated glass capillary.

In the ion funnel configuration, ions are generated in the same way as in the standard ESI source, however, after the capillary, there are additional parts to assist ion transmission (Scheme A1b). Another difference is that the inlet capillary in the ion funnel source is perpendicular, instead of on axis, to the transfer optics and MS analyzer for the purpose of reducing chemical noise and contaminants because, after exiting the capillary (voltage 'CapExit' ranges from 0 to 400 V), only charged ions can be deflected

to the MS analyzer by the 'Deflector', whose voltage ranges from 160 V to 320 V (Scheme A1b). The entrance of the first ion funnel is biased by the voltage Ion funnel 1 (IF 1, varying from 50 V to 240 V). The potential difference between IF 1 and the first skimmer (SK 1, varying from 0 V to 100 V) forms an electric field that pushes ions in the forward direction. The RF field applied to the first ion funnel together with collisions focuses the ions to efficiently pass through the aperture of skimmer 1. A typical pressure in the first pumping stage of the first ion funnel is 1-2 mbar, which is maintained by a roughing pump. After passing skimmer 1, ions enter a second ion funnel with a DC voltage IF 2, varying from 3 V to 35 V. The second funnel is followed by a second skimmer (SK 2), which is typically held at a voltage of 3-8 V. The pressure between the two skimmers is in the range of 0.1-0.3 mbar, maintained by a turbomolecular pump. After passing through the second skimmer, ions enter a source hexapole, which is biased at 0-6 V. The pressure in this range is $\sim 10^{-3}$ mbar, maintained by a second turbomolecular pump. For the ion funnels, the RF frequencies are fixed and the RF amplitudes and DC voltages are tunable. Each of these parameters influence both ion transmission efficiency and how 'hard' or 'soft' the ion transfer is. However, signal abundances are very sensitive to the RF amplitudes and they are therefore generally kept at optimized values. In our experiments, both the absolute and relative (with respect to the voltages on adjacent ion optical elements) ion funnel DC values influence ion transmission conditions and they were tuned to obtain optimum transmission efficiency for non-covalent complexes.

Model systems used for testing the ion funnel performance (both sensitivity and transmission efficiency) include apomyoglobin (Sigma-Aldrich, St. Louis, MO) for

positive ion mode and a hairpin structured DNA 15-mer (TriLink Biotechnologies, San Diego, CA) for negative ion mode. Additionally, a DNA duplex (dA_6-dT_6, single stranded DNA purchased from TriLink Biotechnologies) and a ternary complex formed between a DNA duplex (12mer-C: (d(CCCCATATCCCC) - 12mer-G: d(GGGGATATGGGG), single stranded DNA purchased from Yale Keck Facility) and a minor groove binder (distamycin A from Sigma-Aldrich) with a 1:2 stoichiometry were used to test the ion funnel transmission efficiency for non-covalent complexes in negative ion mode. In positive ion mode, samples were diluted into 20% v/v methanol (Fisher, Fair Lawn, NJ) with 0.1% formic acid (Fisher) while in negative ion mode samples were diluted into 5% v/v isopropanol (Fisher) with 10 mM ammonium acetate (Fisher). Duplex dA_6-dT_6 was prepared by mixing 50 μM dT_6 and 50 μM dA_6 and annealing at 95ºC for 10 minutes. The ternary complex was prepared by mixing 20 μM 12-mer C, 20 μM 12-mer G, and 50 μM distamycin A and annealing at 95ºC for 10 minutes. All samples were injected through a 100 μl Hamilton syringe at a flow rate of 50 μl/h. Detected peak abundances were adjusted to their charge states because higher charge state species generate more abundant signals compared to lower charge state species of same ion number.

A.3. Results and Discussion

Sensitivity of the standard ESI source was compared to that of the ion funnel ESI source by comparing the absolute signal abundances of the same sample at the same concentration. Apomyoglobin was used in positive ion mode and the 12-mer G-C duplex-distamycin A complex was used in negative ion mode, respectively (only data for

negative ion mode is shown here). An 80-fold increase in signal abundance was observed for apomyoglobin and ~2.5-fold improvement was seen for the complex. These improvements are probably due to the enhanced ion focusing and transmission conditions in the ion funnel ESI source. Figure A1 shows the ESI-FT-ICR mass spectra of the 5- charge state (deprotonated species as well as the corresponding salts adducts) of the ternary complex at a 1:2 stoichiometry (duplex:minor groove binder). The ion accumulation time with the standard ESI source was 5 s and the spectrum was summed over 8 scans (Figure A1a) whereas with the ion funnel ESI configuration, ions were accumulated for 4 s and the spectrum was summed over 4 scans (Figure A1b). Thus, the overall ion accumulation time for the standard ESI source (a) is about 2.5 times longer than that with the ion funnel ESI source (b) and their compared signal abundances therefore indicate a sensitivity increase of about 2.5-fold in negative ion mode following installation of the ion funnel ESI source.

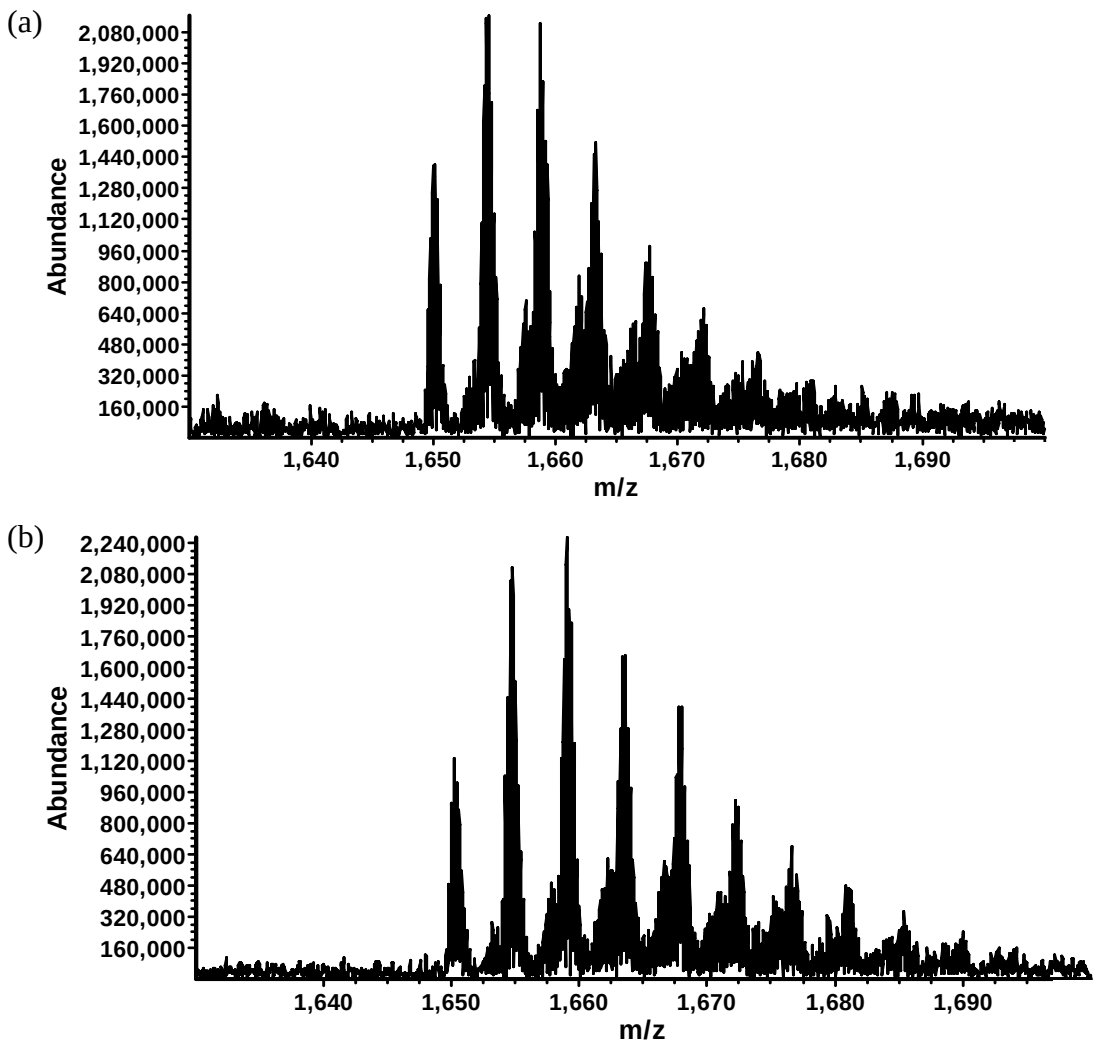


Figure A1. Sensitivity in standard ESI MS (a) compared to that with an ion funnel ESI source (b) in negative ion mode. Spectra show the 5- charge state (deprotonated species as well as the corresponding salts adducts) of a ternary complex formed between a DNA duplex (12-mer C-G) and distamycin A. Samples were prepared under identical conditions for these two cases. The accumulation time in (a) is 5 s and the spectrum is summed over 8 scans. The accumulation time in (b) is 4 s and the spectrum is summed over 4 scans. Therefore, the sensitivity is improved about 2.5-fold in the ion funnel ESI source compared to the standard ESI source.

Charge state distributions can be used as an indication of the softness of the ion source because higher charge state species are relatively less stable than lower charge

state species and, therefore, the survival of higher charge state species from the ionization process indicates that the source is soft. Apomyoglobin and a hairpin structured DNA 15-mer were used for comparing the softness of the standard ESI source with that of the ion funnel ESI source in positive and negative ion mode, respectively. Table A1 shows that at various voltages of the capillary exit (CapExit, see Scheme A1b), the most abundant isotopic clusters of apomyoglobin are centered around the charge states 17+ to 12+ in standard ESI, while with the ion funnel ESI source (Scheme A2b) they are centered around 22+ to 20+, indicating that the ion funnel is softer than the standard ESI source. Similar results were obtained in negative ion mode. Figure A2c shows the presence of 7- and 6- charge states for the DNA 15-mer (structure shown in the inset) in ion funnel ESI whereas in standard ESI mostly the 4- charge state is observed (Figure A2a). In addition to the generally softer transmission conditions compared to the standard ESI source, variation of the DC voltages at each interface within the ion source can also influence the softness of the ion transmission process. For example, Table A1 shows that changes in the CapExit voltage influence the charge state distribution of apomyoglobin. The spectra shown in Figure A2 b and c were both obtained with the ion funnel ESI source but with different DC voltages at each interface. Voltages in (b) are CapExit = - 400 V, Deflector = - 320 V, IF 1 = - 240V. SK 1 = - 100 V, IF 2 = - 20 V, SK 2 = - 8 V, Hexapole DC = - 2 V, and voltages in (c) are CapExit = - 180 V, Deflector = - 240 V, IF 1 = - 120V. SK 1 = - 60 V, IF 2 = - 6 V, SK 2 = - 5 V, Hexapole DC = - 4 V. It is clear that the conditions in A2c are softer than those in A2b because a higher charge state (7-) is present.

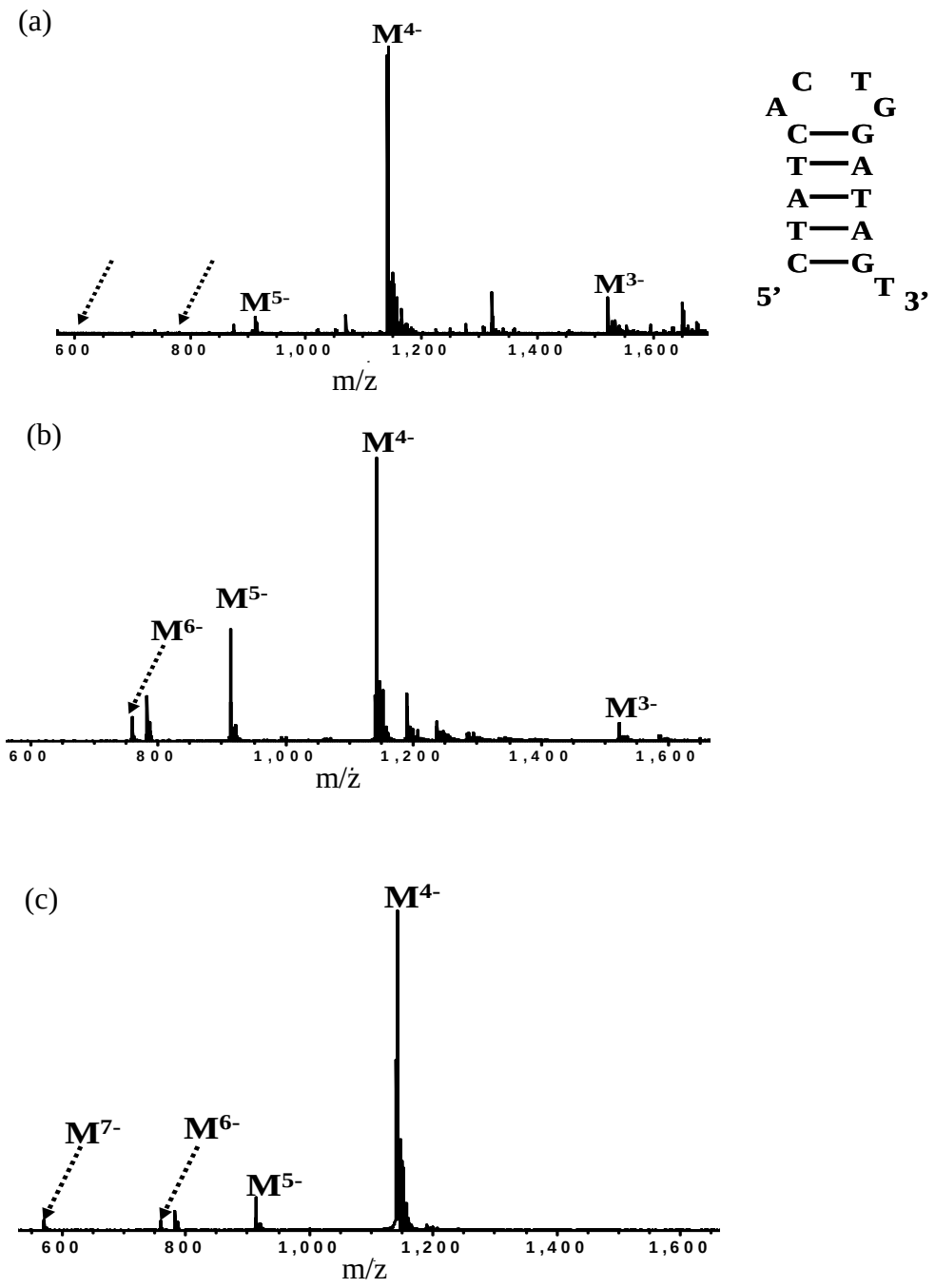


Figure A2. Charge state distribution of a hairpin DNA 15-mer (structure shown in the inset) with a standard ESI source (a) compared to that with an ion funnel ESI source (b and c) in negative ion mode. With the ion funnel source, there are higher charge states present in the spectra. Spectra in (b) and (c) are both obtained with the ion funnel ESI source but the transmission conditions in (c) are softer than those in (b).

To further investigate this DC voltage's effect on the softness of the ion funnel ESI source, additional experiments were performed. In positive ion mode, the charge state distribution of apomyoglobin was still used as an indication of the softness (Table A2). There are seven DC potentials within the ion funnel ESI source: the CapExit voltage, the Deflector voltage, the IF 1 voltage, the SK 1 voltage, the IF 2 voltage, the SK 2 voltage, and the Hexapole DC voltage. Table A2a shows that when the CapExit voltage gets closer to the Deflector voltage, the most abundant isotopic cluster shifts towards higher charge states, indicating softer transmission conditions. Table A2b shows that changes of the Deflector voltage does not have a significant influence on the charge state distribution. Table A2c shows the influence of the IF 1 voltage, which affects both the Deflector-IF 1 and the IF 1-SK 1 electric fields, on ion transmission. The former electric field appears to be most significant because when the potential difference between the Deflector and IF 1 is larger and the potential difference between IF 1 and SK 1 is smaller, lower charge state isotopic clusters become more abundant. Similarly, Table A2d shows that the SK 1 voltage is more important for achieving soft transmission conditions between skimmer 1 and the second ion funnel than for the ion funnel 1-skimmer 1 interface. Table A2e shows that the IF 2 voltage influences transmission conditions at both the skimmer 1-ion funnel 2, and the ion funnel 2-skimmer 2 interfaces. Finally, Tables A2 f and g show neither the SK 2 voltage, nor the Hexapole DC voltage have large effects on ion transmission. At voltages outside the ranges shown in the tables, there was no ion signal.

(a) Standard ESI source

CapExit	Relative abundance of each charge state															
	24+	23+	22+	21+	20+	19+	18+	17+	16+	15+	14+	13+	12+	11+	10+	9+
110	0.07	0.18	0.34	0.55	0.72	0.93	0.98	1.00	0.59	0.38	0.24	0.18	0.13			
130	0.09	0.15	0.26	0.52	0.77	0.87	0.88	1.00	0.56	0.30	0.10	0.22	0.10			
150		0.12	0.21	0.46	0.62	0.88	0.89	1.00	0.50	0.34	0.18	0.11	0.08			
170		0.08	0.18	0.33	0.71	0.79	0.80	1.00	0.52	0.32	0.18	0.21	0.08			
190		0.05	0.11	0.24	0.45	0.73	0.99	1.00	0.7	0.26	0.13	0.11	0.06			
210			0.05	0.12	0.32	0.58	0.73	1.00	0.72	0.41	0.15	0.13	0.08			
230				0.06	0.25	0.50	0.71	1.00	0.98	0.48	0.23	0.17	0.06			
250					0.08	0.25	0.56	0.98	1.00	0.67	0.36	0.14	0.09	0.05		
270						0.06	0.25	0.71	1.00	0.93	0.69	0.23	0.13	0.05		
290								0.25	0.74	1.00	0.92	0.58	0.11	0.04	0.05	
310									0.19	0.71	1.00	0.93	0.17	0.07	0.10	
330									0.03	0.13	0.49	1.00	0.27	0.10	0.14	0.12
350											0.24	0.99	1.00	0.22	0.63	0.59

(b) Ion funnel ESI source

CapExit	Relative abundance of each charge state													
	26+	25+	24+	23+	22+	21+	20+	19+	18+	17+	16+	15+	14+	13+
360	0.07	0.14	0.37	0.57	1.00	0.78	0.68	0.54	0.59	0.31	0.30	0.28	0.15	0.13
320	0.07	0.16	0.35	0.52	1.00	0.90	0.70	0.67	0.57	0.35	0.29	0.27	0.19	0.20
280	0.07	0.16	0.36	0.50	1.00	0.75	0.68	0.58	0.58	0.39	0.41	0.28	0.21	
240	0.06	0.16	0.33	0.59	1.00	0.93	0.75	0.72	0.58	0.47	0.34	0.33	0.14	0.17
200	0.10	0.20	0.45	0.52	1.00	0.85	0.85	0.74	0.81	0.44	0.42	0.31	0.23	0.20
160		0.14	0.36	0.51	0.87	1.00	0.89	0.85	0.86	0.66	0.74	0.51	0.26	
120			0.14	0.27	0.61	0.96	1.00	0.97	0.99	0.64	0.91	0.45	0.43	

Table A1. Apomyoglobin charge state distributions in standard ESI MS (a) compared to those from an ion funnel ESI source (b) in positive ion mode at different CapExit voltages. In standard ESI MS, the most abundant isotopic clusters are centered around charge states from 17+ to 12+ whereas with the ion funnel ESI source, they are centered around 22+ to 20+. The numbers shown in the tables are the relative peak abundances (adjusted to their charge states) normalized to the most abundant one.

(a) CapExit voltage change

CapExit	Deflector	Relative abundance of each charge state													
		26+	25+	24+	23+	22+	21+	20+	19+	18+	17+	16+	15+	14+	13+
400	260	0.05	0.12	0.30	0.41	0.92	0.98	1.00	0.73	0.67	0.37	0.30	0.27	0.12	0.11
360	260	0.07	0.14	0.37	0.57	1.00	0.78	0.68	0.54	0.59	0.31	0.30	0.28	0.15	0.13
320	260	0.07	0.16	0.35	0.52	1.00	0.90	0.70	0.67	0.57	0.35	0.29	0.27	0.19	0.20
280	260	0.07	0.16	0.36	0.50	1.00	0.75	0.68	0.58	0.58	0.39	0.41	0.28	0.21	0.00
240	260	0.06	0.16	0.33	0.59	1.00	0.93	0.75	0.72	0.58	0.47	0.34	0.33	0.14	0.17
200	260	0.10	0.20	0.45	0.52	1.00	0.85	0.85	0.74	0.81	0.44	0.42	0.31	0.23	0.20
160	260		0.14	0.36	0.51	0.87	1.00	0.89	0.85	0.86	0.66	0.74	0.51	0.26	
120	260			0.14	0.27	0.61	0.96	1.00	0.97	0.99	0.64	0.91	0.45	0.43	

(b) Deflector voltage change

CapExit	Deflector	IF 1	Relative abundance of each charge state													
			26+	25+	24+	23+	22+	21+	20+	19+	18+	17+	16+	15+	14+	13+
160	320	180	0.08	0.15	0.31	0.48	1.00	0.87	0.71	0.58	0.61	0.33	0.34	0.23		0.12
160	280	180		0.00	0.20	0.32	1.00	0.88	0.68	0.60	0.89	0.51	0.73	0.31	0.00	
160	260	180		0.00	0.15	0.41	0.78	1.00	0.86	0.79	0.75	0.61	0.75	0.47	0.23	
160	240	180		0.14	0.36	0.51	0.87	1.00	0.89	0.85	0.86	0.66	0.74	0.51	0.26	
160	220	180	0.09	0.17	0.38	0.63	0.98	1.00	0.95	0.77	0.81	0.54	0.55	0.43	0.28	
160	200	180	0.14	0.29	0.48	0.70	0.99	1.00	0.80	0.82	0.78	0.53	0.57	0.32	0.22	
160	180	180	0.08	0.17	0.32	0.53	1.00	0.85	0.63	0.65	0.60	0.52	0.40	0.27	0.27	

(c) IF 1 voltage change

Deflector	IF 1	SK 1	Relative abundance of each charge state												
			25+	24+	23+	22+	21+	20+	19+	18+	17+	16+	15+	14+	13+
260	240	50	0.15	0.32	0.50	0.93	1.00	0.65	0.35	0.32	0.19	0.18			
260	220	50	0.11	0.29	0.51	0.91	1.00	0.81	0.76	0.70	0.53	0.38	0.38		
260	200	50	0.15	0.42	0.65	1.00	1.00	0.97	0.77	0.87	0.52	0.63	0.35		
260	180	50	0.14	0.36	0.51	0.87	1.00	0.89	0.85	0.86	0.66	0.74	0.51	0.26	
260	160	50	0.10	0.38	0.60	0.93	0.94	0.97	0.88	1.00	0.65	0.80	0.65	0.31	
260	140	50		0.22	0.43	0.65	0.85	0.68	0.70	1.00	0.49	0.69	0.40	0.00	
260	120	50		0.26	0.38	0.83	0.93	0.77	0.75	1.00	0.70	0.89	0.49		0.50

(d) SK 1 voltage change

IF 1	SK 1	IF 2	Relative abundance of each charge state												
			25+	24+	23+	22+	21+	20+	19+	18+	17+	16+	15+	14+	13+
160	100	5.5										0.30	0.81	1.00	0.66
160	90	5.5									0.31	0.96	1.00	0.49	0.37
160	80	5.5							0.17	0.58	0.71	1.00	0.65	0.29	0.10
160	70	5.5					0.06	0.26	0.42	0.83	0.75	1.00	0.49	0.09	0.05
160	60	5.5			0.04	0.19	0.39	0.61	0.75	1.00	0.51	0.33	0.20	0.12	0.09
160	50	5.5	0.10	0.38	0.60	0.93	0.94	0.97	0.88	1.00	0.65	0.80	0.65	0.31	0.00
160	40	5.5	0.10	0.16	0.19	0.32	0.40	0.47	0.62	1.00	0.66	0.65	0.45	0.35	0.32
160	30	5.5	0.16	0.22	0.27	0.50	0.44	0.58	0.80	1.00	0.77	0.79	0.68	0.00	
160	20	5.5	0.18	0.18	0.28	0.48	0.65	0.74	0.80	0.98	1.00	1.00	0.97	0.00	
160	10	5.5	0.18	0.14	0.21	0.28	0.43	0.61	0.94	1.00	0.82	0.94	0.77	0.41	
160	0	5.5	0.10	0.17	0.20	0.26	0.34	0.49	0.52	1.00	0.62	0.85	0.48	0.34	

(e) IF 2 voltage change

SK 1	IF 2	SK 2	Relative abundance of each charge state											
			24+	23+	22+	21+	20+	19+	18+	17+	16+	15+	14+	13+
60	35	5			0.41	0.48	1.00	0.72	0.93	0.65	0.58			
60	30	5		0.28	0.49	0.44	0.67	1.00	0.90	0.51	0.80			
60	25	5		0.00	0.43	0.62	0.58	0.98	1.00	0.77	0.77	0.55		
60	20	5		0.00	0.13	0.62	0.74	1.00	0.82	0.50	0.37	0.31	0.12	
60	15	5		0.05	0.23	0.66	1.00	0.93	0.62	0.25	0.24	0.11	0.06	0.05
60	10	5	0.05	0.15	0.33	0.69	0.87	1.00	0.63	0.24	0.16	0.10		
60	5.5	5		0.08	0.18	0.39	0.72	1.00	1.00	0.52	0.44	0.20		
60	3	5			0.08	0.21	0.45	0.83	1.00	0.68	0.62	0.22	0.12	

(f) SK 2 voltage change

IF 2	SK 2	Hexapole DC	Relative abundance of each charge state										
			24+	23+	22+	21+	20+	19+	18+	17+	16+	15+	14+
10	15	4			0.14	0.27	0.54	0.94	1.00	0.95	0.61		
10	10	4			0.09	0.21	0.54	0.80	1.00	0.90	0.39	0.38	0.23	0.16
10	5	4		0.05	0.16	0.41	0.65	0.88	1.00	0.67	0.23	0.18	0.12	0.04

(g) Hexapole DC voltage change

SK 2	Hexapole DC	Relative abundance of each charge state											
		24+	23+	22+	21+	20+	19+	18+	17+	16+	15+	14+	13+
5	6	0.06	0.09	0.30	0.62	1.00	0.83	0.73	0.40	0.26	0.09	0.00	0.00
5	5	0.07	0.22	0.38	0.81	0.92	1.00	0.66	0.33	0.26	0.18	0.00	0.00
5	4	0.05	0.16	0.41	0.65	0.88	1.00	0.67	0.23	0.18	0.12	0.04	0.00
5	3	0.04	0.11	0.30	0.58	0.83	1.00	0.87	0.40	0.37	0.21	0.13	0.13
5	2	0.00	0.00	0.15	0.22	0.65	0.56	1.00	0.55	0.89	0.56	0.00	0.00

Table A2. Influences of DC voltages withitn the ion funnel ESI source on the charge state distribution of apomyoglobin in positive ion mode. The numbers shown in the tables are the relative peak abundances (adjusted to their charge states) normalized to the most abundant one.

In negative ion mode, the abundance of either the non-covalently bound dA_6-dT_6 duplex, or the 12-mer G-C duplex-distamycin A complex were used as an indication of the softness of the ion funnel ESI source. Table A3 summarizes observed changes in the ratio of duplex abundance relative to the sum of single strand abundances at different ESI source DC voltages. Among all the seven DC voltages, CapExit and IF 1 seemed to be more influential than the others. Lowering their values yielded a higher relative abundance of the duplex, however, the absolute duplex abundance decreased, which means that sensitivity is sacrificed. Table A4 summarizes the results for the ternary complex. In this case, CapExit and IF 2 seemed to be the most critical voltages: lowering the CapExit voltage resulted in an increase in the ratio of the complex abundance relative to the overall abundance of other species in the spectrum and a decrease in absolute signal abundance, whereas lowering IF 2 generated both a higher abundance ratio and a higher absolute signal abundance of the complex. For both the duplex and the ternary complex, SK 2 and Hexapole DC could not be varied much because the signal disappeared outside the voltage ranges shown in Tables A3 and A4 f and g.

(a) CapExit voltage change

CapExit	Deflector	Relative abundance of each charge state				Sum of abundances	duplex (3-)	
		dT_6^{3-}	dA_6^{3-}	dT_6^{2-}	dA_6^{2-}		Abundance	Ratio
-400	-240	60.48	80.34	5775.00	4720.00	10635.81		
-360	-240	816.20	883.33	5033.00	4165.00	10897.53	46.75	
-320	-240	1100.83	1143.33	4006.45	3755.00	10005.62	129.31	0.01
-280	-240	1103.03	1120.00	3806.80	3650.55	9680.38	205.10	0.02
-240	-240	863.07	886.67	3345.25	3518.00	8612.98	245.66	0.03
-200	-240	547.40	553.33	3401.50	3299.65	7801.88	234.88	0.03
-160	-240	93.52	79.61	1829.30	2088.40	4090.84	199.42	0.05
-120	-240			977.55	1057.80	2035.35	84.69	0.04

(b) Deflector voltage change

CapExit	Deflector	IF 1	Relative abundance of each charge state				Sum of abundances	duplex (3-)	
			dT_6^{3-}	dA_6^{3-}	dT_6^{2-}	dA_6^{2-}		Abundance	Ratio
-240	-320	-130					795.18	42.20	0.05
-240	-280	-130	117.83	109.30	1896.75	2063.65	4187.53	175.46	0.04
-240	-240	-130	863.07	886.67	3345.25	3518.00	8612.98	245.66	0.03
-240	-200	-130	1105.93	1123.33	3243.65	3300.15	8773.07	189.65	0.02
-240	-160	-130	600.00	570.40	1655.00	1851.20	4676.60	170.89	0.04

(c) IF 1 voltage change

Deflector	IF 1	SK 1	Relative abundance of each charge state				Sum of abundances	duplex (3-)	
			dT_6^{3-}	dA_6^{3-}	dT_6^{2-}	dA_6^{2-}		Abundance	Ratio
-240	-210	-40	765.367	863.33	3560.10	3286.90	8475.70	180.70	0.02
-240	-180	-40	1102.67	1142.87	3940.10	3563.85	9749.48	149.68	0.02
-240	-150	-40	1021.73	1083.33	3275.00	3349.35	8729.42	212.84	0.02
-240	-130	-40	863.07	886.67	3345.25	3518.00	8612.98	245.66	0.03
-240	-100	-40	416.67	416.67	2552.05	2763.20	6148.58	235.77	0.04
-240	-80	-40	193.86	183.94	2118.20	2418.30	4914.31	204.23	0.04
-240	-50	-40			1079.90	1019.30	2099.20	136.17	0.06

(d) SK 1 voltage change

IF1	SK1	IF2	Relative abundance of each charge state				Sum of abundances	duplex (3-)	
			dT_6^{3-}	dA_6^{3-}	dT_6^{2-}	dA_6^{2-}		Abundance	Ratio
-100	-60	-6	63.66	76.31	5772.50	4285.95	10198.42		
-100	-40	-6	416.67	416.67	2552.05	2763.20	6148.58	235.77	0.04
-100	-30	-6	520.00	493.47	2673.60	2957.00	6644.07	323.34	0.05
-100	-20	-6	489.80	530.00	2651.65	3064.95	6736.40	255.95	0.04
-100	-10	-6	516.67	563.33	2852.60	3005.00	6937.60	308.80	0.04
-100	0	-6	61.39	65.24	2263.10	2661.34	5051.08	276.24	0.05

(e) IF 2 voltage change

SK 1	IF 2	SK 2	Relative abundance of each charge state				Sum of abundances	duplex (3-)	
			dT_6^{3-}	dA_6^{3-}	dT_6^{2-}	dA_6^{2-}		Abundance	Ratio
-30	-20	-5					1632.20		
-30	-10	-5	1136.67	835.30	4094.70	4067.10	10133.77	155.78	0.02
-30	-6	-5	520.00	493.47	2673.60	2957.00	6644.07	323.34	0.05
-30	-3	-5	426.6.0	486.67	1818.95	2220.00	4952.22	76.68	0.02

(f) SK 2 voltage change

IF 2	SK 2	Hexapole DC	Relative abundance of each charge state				Sum of abundances	duplex (3-)	
			dT_6^{3-}	dA_6^{3-}	dT_6^{2-}	dA_6^{2-}		Abundance	Ratio
-6	-10	-2	278.02	110.56	428.72	228.34	1045.65		
-6	-8	-2	816.67	458.00	2570.75	2023.85	5869.27		
-6	-5	-2	520.00	493.47	2673.60	2957.00	6644.07	323.34	0.05
-6	-3	-2			141.95	263.68	405.62	60.87	0.15

(g) Hexapole DC voltage change

SK 2	Hexapole DC	Relative abundance of each charge state				Sum of abundances	duplex (3-)	
		dT_6^{3-}	dA_6^{3-}	dT_6^{2-}	dA_6^{2-}		Abundance	Ratio
-5	-4	543.33	448.37	2239.40	2286.65	5517.75	163.79	0.03
-5	-2	520.00	493.47	2673.60	2957.00	6644.07	323.34	0.05
-5	0	81.98	136.33	471.79	645.00	1335.10	118.47	0.09
-5	1	39.51	77.53	341.24	515.00	973.28	87.19	0.09
-5	3		41.84	198.52	308.64	548.99		

Table A3. Influences of DC voltages within the ion funnel ESI source on the relative abundance of a DNA duplex (dA_6-dT_6) in negative ion mode. Numbers shown in the Tables are the absolute peak abundances (adjusted to charge state) of the duplex or a single strand and the ratio is the duplex abundance relative to the sum of single strand abundances. Samples were accumulated for 5 s in the collision cell and spectra were summed over 4 scans.

(a) CapExit voltage change

CapExit	Deflector	Sum of abundancs	7-		6-		5-	
			Abundance	Ratio	Abundance	Ratio	Abundance	Ratio
-400	-240	2024.83					129.06	0.06
-360	-240	1785.80			16.78	0.01	99.93	0.06
-320	-240	988.41			50.08	0.05	77.47	0.08
-280	-240	492.08	20.43	0.04	56.07	0.11	41.76	0.08
-240	-240	170.99	18.77	0.11	37.34	0.22	28.46	0.17
-200	-240	68.28			26.57	0.39		

(b) Deflector voltage change

CapExit	Deflector	IF 1	Sum of abundances	7-		6-		5-	
				Abundance	Ratio	Abundance	Ratio	Abundance	Ratio
-360	-320	-100	70.12			18.05	0.26	22.36	0.32
-360	-280	-100	786.85			32.67	0.04	57.48	0.07
-360	-240	-100	1785.80			16.78	0.01	99.93	0.06
-360	-200	-100	1673.29			21.54	0.01	79.69	0.05
-360	-160	-100	889.03					63.59	0.07

(c) IF 1 voltage change

Deflector	IF 1	SK 1	Sum of abundances	7-		6-		5-	
				Abundance	Ratio	Abundance	Ratio	Abundance	Ratio
-240	-210	-40	459.70	16.74	0.04	54.19	0.12	28.56	0.06
-240	-180	-40	787.69			56.10	0.07	38.22	0.05
-240	-150	-40	1214.21			50.93	0.04	78.07	0.06
-240	-120	-40	1650.81			31.38	0.02	91.76	0.06
-240	-100	-40	1785.80			16.78	0.01	99.93	0.06
-240	-80	-40	1514.30					104.42	0.07
-240	-50	-40	1091.36					50.85	0.05

(d) SK 1 voltage change

IF 1	SK 1	IF 2	Sum of abundance	7-		6-		5-	
				Abundance	Ratio	Abundance	Ratio	Abundance	Ratio
-120	-100	-6							
-120	-80	-6	3919.86					84.72	0.02
-120	-60	-6	1921.29					119.64	0.06
-120	-40	-6	1703.60			34.29	0.02	104.21	0.06
-120	-20	-6	1591.74			33.50	0.02	85.72	0.05
-120	-10	-6	1552.55			33.47	0.02	78.60	0.05

(e) IF 2 voltage change

SK 1	IF 2	SK 2	Sum of abundances	7-		6-		5-	
				Abundance	Ratio	Abundance	Ratio	Abundance	Ratio
-60	-20	-5	1381.89					28.66	0.02
-60	-10	-5	1497.41					78.72	0.05
-60	-6	-5	1921.29					119.64	0.06

(f) SK 2 voltage change

IF 2	SK 2	Hexapole DC	Sum of abundances	7-		6-		5-	
				Abundance	Ratio	Abundance	Ratio	Abundance	Ratio
-6	-5	-4	1921.29					119.64	0.06
-6	-3	-4	1974.02					175.64	0.09

(g) Hexapole DC voltage change

SK 2	Hexapole DC	Sum of abundances	7-		6-		5-	
			Abundance	Ratio	Abundance	Ratio	Abundance	Ratio
-3	-4	1974.02					175.64	0.09
-3	-2	1236.69					224.67	0.18

Table A4. Influences of DC voltages within the ion funnel ESI source on the relative abundance of a ternary complex (12-mer G-C duplex-distamycin A) in negative ion mode. “Sum of abundance” is the sum of absolute peak abundances of all peaks (adjusted to their charge states) in the spectrum except the complex and the ratio for each charge state is the absolute peak abundance of the complex at that charge state relative to the “sum of abundance”. Samples were accumulated for 4 s in the collision cell and spectra were summed over 4 scans.

A.4. Conclusion

Our experiments comparing the performance of the standard ESI source with that of the ion funnel ESI source suggest that the latter has both higher sensitivity and softer ion transmission conditions than the former. Almost two orders of magnitude improvement in sensitivity was achieved in positive ion mode whereas about 2.5-fold improvement was seen in negative ion mode. The ion funnel ESI source is generally softer than the standard ESI source and, by tuning the DC voltages of the ion funnel source, even softer ion transmission conditions could be achieved. Among all the DC voltages, the inlet capillary exit voltage (CapExit) and the DC offsets of the two ion funnels (IF 1 and IF 2) seem to be the most influential ones. Lowering these voltages usually provided softer transmission conditions, however, the sensitivity is sometimes sacrificed.

A.5. Bibliography

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

(2) Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F. *Mass Spectrom. Rev.* **1990**, *9*, 37-70.

(3) Smith, R. D.; Loo, J. A.; Ogorzalek Loo, R. R.; Busman, M.; Udseth, H. R. *Mass Spectrom. Rev.* **1991**, *10*, 359-451.

(4) Smith, R. D.; Loo, J. A.; Edmonds, C. G.; Barinaga, C. J.; Udseth, H. R. *Anal. Chem.* **1990**, *62*, 882-899.

(5) Shaffer, S. A.; Tang, K.; Anderson, G. A.; Prior, D. C.; Udseth, H. R.; Smith, R. D. *Rapid Commun. Mass Spectrom.* **1997**, *11*, 1813-1817.

(6) Douglas, D. J.; French, J. B. *J. Am. Soc. Mass Spectrom.* **1992**, *3*, 398-408.

(7) Chen, Y.-L.; Collings, B. A.; Douglas, D. J. *J. Am. Soc. Mass Spectrom.* **1997**, *6*, 681-687.

(8) Javahery, G.; Thomson, B. *J. Am. Soc. Mass Spectrom.* **1997**, *6*, 697-702.

(9) Douglas, D. J. *J. Am. Soc. Mass Spectrom.* **1998**, *9*, 101-113.

(10) Dodonov, A.; Kozlovsky, V.; Loboda, A.; Raznikov, V.; Sulimenkov, I.; Tolmachev, A.; Kraft, A.; Wollnik, H. *Rapid Commun. Mass Spectrom.* **1997**, *11*, 1649-1656.

(11) Gerlich, D. *in State Selected and State-to-State Ion - Molecule Reaction Dynamics. Part 1. Experiment*; Wiley: New York, 1992.

(12) Shaffer, S. A.; Prior, D. C.; Anderson, G. A.; Udseth, H. R.; Smith, R. D. *Anal. Chem.* **1998**, *70*, 4111-4119.

(13) Shaffer, S. A.; Tolmachev, A.; Prior, D. C.; Anderson, G. A.; Udseth, H. R.; Smith, R. D. *Anal. Chem.* **1999**, *71*, 2957-2964.

(14) Kim, T.; Tolmachev, A. V.; Harkewicz, R.; Prior, D. C.; Anderson, R.; Udseth, H. R.; Smith, R. D.; Bailey, T. H.; Rakov, S.; Futrell, J. H. *Anal. Chem.* **2000**, *72*, 2247-2255.