

# **Determining the Structures and Properties of Biologically-relevant Ions**

by

**Jarrold Psutka**

A thesis  
presented to the University of Waterloo  
in fulfillment of the  
thesis requirement for the degree of  
Master of Science  
in  
Chemistry

Waterloo, Ontario, Canada, 2018

© Jarrod Psutka 2018

## **Author's Declaration**

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

## Statement of Contributions

Chapter 4 investigates the properties of protonated nucleobases and their methylated derivatives. The unmethylated version is part of a published manuscript with the following co-authors: Ahdia Anwar, Stephen W.C. Walker, Thorsten Dieckmann, John S. Janiszewski, J. Larry Campbell, and W. Scott Hopkins. The methylated work was carried out with Stephanie Ryall and a future publication is in the preparatory phase.

Chapter 6 investigates transformation products of trimethoprim and contains data from a recently published manuscript in *Analytical Chemistry*. The co-authors, along with me, are: Annick Dion-Fortier, Thorsten Dieckmann, J. Larry Campbell, Pedro A. Segura, and W. Scott Hopkins.

## **Abstract**

Gas phase studies of biologically relevant ions are increasing in popularity due to the possibility of high throughput analysis requiring minimum sample concentrations. This thesis explores the potential of differential mobility spectrometry-mass spectrometry (DMS-MS) in combination with quantum chemical calculation methods to probe the structures, energetics, and dynamics of three distinct classes of biomolecules. The first project outlines the use of DMS-MS to separate and identify protonated forms of methylated and unmethylated nucleobases to gain a fundamental understanding of their gas phase properties in relation to their role in nucleic acids. Next, DMS-MS and calculations were conducted for a large RNA system, the Varkud Satellite ribozyme active site loop VI, to study differences between its active and inactive conformations, especially through the use of negative mode hydrogen-deuterium exchange. Finally, DMS-MS was used to identify transformation products of trimethoprim, an antibiotic often found in environmental wastewaters in a reliable and efficient method. Ultimately, DMS-MS and quantum calculations have been shown to be a powerful analytical tool to investigate structures and properties of biomolecules. The methodologies described herein can have an impact in a wide variety of industries, from drug discovery to environmental wastewater cleanup.

## **Acknowledgements**

I would like to thank my supervisors, Dr. Scott Hopkins and Dr. Thorsten Dieckmann, for allowing me to conduct research projects under their expertise and offering advice, assistance, and scientific knowledge throughout my degree. I really enjoyed my time within these labs and could not have succeeded without them.

I would also like to thank Dr. Terry McMahon and Dr. John Honek for agreeing to be on my committee and providing valuable insight during my proposal. Having such a diverse and knowledgeable committee of renowned scientists made my experience here great and I am excited to disseminate my results.

Additionally, I would like to thank the members of the Hopkins lab and Dieckmann lab for providing assistance throughout my research and always providing a welcoming, enjoyable, and fun environment filled with many learning opportunities. Specifically, I would like to thank Christian Ieritano for his valuable insight involving calculations and DMS work, Ahdia Anwar and Stephanie Ryall for their collaborations involving nucleobases, Dr. Steve Walker for his collaborative work and availability for questions, as well as Mike Lecours, Ce Zhou, and Josh Featherstone for their computing knowledge and creation of many time saving scripts. Dr. Pedro Segura and Annick-Dion Fortier reached out to collaborate with us and I am truly grateful for their contribution towards the trimethoprim work.

Finally I would like to thank my family and all my friends for their support over my academic journey. I am grateful to have a great support team in my life and look forward to continue to be challenged in the future.

## Table of Contents

Author's Declaration.....	ii
Statement of Contributions .....	iii
Abstract.....	iv
Acknowledgements.....	v
List of Figures.....	viii
List of Tables .....	ix
List of Abbreviations .....	x
1.0 Introduction.....	1
2.0 Methodology.....	4
2.1 Introduction to Computational Methods.....	4
2.1.1 Density Functional Theory .....	4
2.1.2 Basin Hopping .....	5
2.2 Introduction to Experimental Methods.....	7
2.2.1 Differential Mobility Spectrometry .....	7
2.2.2 Collision Induced Dissociation.....	11
2.2.3 Hydrogen-Deuterium Exchange.....	12
2.2.4 Fenton Reaction.....	14
2.2.5 RNA Synthesis and Purification.....	15
3.0 Separation and Identification of Unmethylated and Methylated Nucleobase Tautomers .	17
3.1 Introduction.....	17
3.2 Methods.....	20
3.2.1 Experimental Methods.....	20
3.2.2 Computational Methods .....	22
3.3 Results and Discussion .....	22
3.3.1 Protonated (C + H) <sup>+</sup> Dispersion Plots.....	22
3.3.2 Computational (C + H) <sup>+</sup> Results.....	24
3.3.3 Collision-Induced Dissociation Results for (C + H) <sup>+</sup> .....	25
3.3.4 Hydrogen-Deuterium Exchange Results for (C + H) <sup>+</sup> .....	25
3.3.5 Declustering Potential Variation for <i>m/z</i> 112 .....	28
3.3.6 (C + H) <sup>+</sup> Ionogram Peak Assignments and Analysis .....	29
3.3.7 Results for (A + H) <sup>+</sup> , (T + H) <sup>+</sup> , (U + H) <sup>+</sup> , and (G + H) <sup>+</sup> .....	30
3.3.8 Methylated Nucleobase Results.....	34

3.4 Conclusions.....	40
4.0 Probing the Structure, Properties, and Dynamics of RNA in the Gas Phase Using the Varkud Satellite Ribozyme .....	42
4.1 Introduction.....	42
4.2 Methods.....	46
4.2.1 Experimental Methods.....	46
4.2.2 Computational Methods .....	49
4.3 Results and Discussion .....	50
4.3.1 Dispersion Plots of VSVI .....	50
4.3.2 Sugar Puckering Dynamics .....	52
4.3.3 Substrate Stem Loop I Accommodation in VSVI .....	53
4.3.4 HDX of VSVI.....	55
4.4 Conclusions.....	58
5.0 Separation and Identification of Trimethoprim Transformation Products .....	60
5.1 Introduction.....	60
5.2 Methods.....	63
5.2.1 Experimental Methods.....	63
5.2.2 Computational Methods .....	65
5.3 Results and Discussion .....	66
5.3.1 Fenton-Reaction Mixture LC-QqTOFMS Analysis .....	66
5.3.2 Protonated TP306 Dispersion Plots .....	66
5.3.3 Declustering Potential Study of Protonated TP306.....	68
5.3.4 Quantum Chemical Calculation Results for (TP306 + H) <sup>+</sup> .....	69
5.3.5 Collision Induced Dissociation Results for (TP306+H) <sup>+</sup> Species .....	70
5.3.6 DMS in Combination with Additional Analytical Techniques .....	74
5.4 Conclusions.....	75
6.0 Thesis Conclusions .....	76
References.....	80
APPENDIX A.....	87
APPENDIX B .....	117
APPENDIX C .....	123

## List of Figures

Figure 2.1 General Scheme of a Typical Basin Hop Routine .....	6
Figure 2.2 Schematic Diagram of DMS-MS and Within the DMS Plates.....	9
Figure 2.3 Sample Dispersion Plot With Three Major Clustering Behaviours.....	11
Figure 2.4 Literature Proposed Negative Mode HDX Mechanism of Nucleotides .....	13
Figure 3.1 Dispersion Plots and Ionogram for $(C + H)^+$ .....	23
Figure 3.2 The breakdown curves obtained for $(C + H)^+$ ( $m/z$ 112) at each isolated ionogram peak .....	26
Figure 3.3 HDX Results for $(C+H)^+$ ( $m/z$ 112) saturated partial pressure of $D_2O$ .....	27
Figure 3.4 Declustering Potential Variation of $(C + H)^+$ .....	29
Figure 3.5 The lowest energy proton-bound cytosine-methanol clusters .....	30
Figure 3.6 Dispersion Plots for Protonated Nucleobases in $N_2$ .....	31
Figure 3.7 Lowest Energy Structures of each Protonated Nucleobase.. .....	33
Figure 3.8 Dispersion Plots for $(MeG + H)^+$ .....	35
Figure 3.9 Breakdown Curves for $(MeG + H)^+$ .....	36
Figure 3.11 Dispersion Plots of Methylated Nucleobases in Dry Nitrogen.....	39
Figure 3.12 Lowest Energy Structures of the Methylated Nucleobases.. .....	40
Figure 4.1 The overall structure of the Varkud Satellite Ribozyme.. .....	44
Figure 4.2 Truncated form of VS loop VI. ....	46
Figure 4.3 Dispersion Plots of VSVI .....	51
Figure 4.4 The 6-mer of VSVI in its inactive C3'-endo conformation. ....	53
Figure 4.5 The 6-mer of VSVI in its active C2'-endo conformation .....	54
Figure 4.6 HDX profiles and difference spectra for the -6 charge state of VSVI in varying modifiers. ....	57
Figure 5.1 DMS obtained plots and structures for $(TP306 + H)^+$ .....	67
Figure 5.2 Declustering potential ramping of suspected $(TP306 + H)^+$ isomers.. .....	69
Figure 5.3 Fragmentation spectra and breakdown curves of the lowest in energy $(TP306 + H)^+$ isomers..	72
Figure 5.4 The fragmentation channels observed via ionogram Peak I.....	72
Figure 5.5 The fragmentation channels observed via ionogram Peak II.....	73



## List of Tables

Table 3.1 Population Distributions of Each Nucleobase Tautomer .....	34
Table 3.2 Population Distributions of Each Methylated Nucleobase Tautomer.....	38
Table 4.1 Transcription Reaction Mixture for VSVI.....	47

## List of Abbreviations

ACE	Acetone
ACN	Acetonitrile
AMBER	Assisted model building with energy refinement
B3LYP	Becke, three parameter, Lee-Yang-Parr
CCSD(T)	Coupled cluster single/double/triple excitations
CE	Collision energy
CID	Collision induced dissociation
CV	Compensation voltage
DFT	Density functional theory
DMS	Differential mobility spectrometry
DNA	Deoxyribonucleic acid
DP	Declustering potential
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
EPI	Enhanced product ion
ESI	Electrospray ionization
EtAc	Ethyl acetate
FT-ICR	Fourier-transform ion cyclotron resonance
HDX	Hydrogen-deuterium exchange
IMS	Ion mobility spectrometry
IPA	Isopropyl alcohol
IRMPD	Infrared multiple photon dissociation
LC	Liquid chromatography
<i>m/z</i>	Mass-to-charge ratio
MeOH	Methanol
MS	Mass spectrometry

NMR	Nuclear magnetic resonance
RNA	Ribonucleic acid
SPME	Solid-phase microextraction
SV	Separation voltage
TPOMs	Transformation products of organic micropollutants
Tris	Tris(hydroxymethyl)aminomethane
UV	Ultraviolet
VS	Varkud satellite
VSVI	Varkud satellite ribozyme active site loop VI

## 1.0 Introduction

Biomolecules are an essential set of compounds, important to the survival of all living organisms. Being such a fundamental part of nature, much work has been done by scientists to characterize and learn the structures and functions of a diverse set of biomolecules. New methods to probe the characteristics of biomolecules are always sought after due to the often complex nature of obtaining biomolecular samples or the time-consuming and difficult manner in which their structure is determined, such as solution phase nuclear magnetic resonance (NMR) or X-ray crystallography. Recently, research has focused on probing the structure of biomolecules in the gas phase to investigate whether the gas phase structure accurately reflects the biologically relevant solution phase form. Results show that care must be placed in sample preparation and electrospray conditions and that the higher order structures of biomolecules are observable in the gas phase, such as mass spectrometry studies of G-quadruplexes.<sup>1,2</sup> However as larger systems are sprayed in the gas phase and solvent is removed, intra- and intermolecular interactions become more important compared to hydrophobic interactions, which may cause different folding motifs in certain proteins.<sup>3,4</sup>

This thesis seeks to bridge the gap between the gas phase properties of biomolecules and biomolecular clusters and how they behave in the condensed phase. To accomplish this, it is important to select relevant model systems and then subsequently increase the complexity of the system once an established research protocol has proven accurate. To begin the study, an increasingly powerful analytical technique known as differential mobility spectrometry (DMS) was used to probe and identify tautomeric forms of nucleobases and their methylated counterparts.<sup>5</sup> Nucleobases are vital to all living things as they are the nitrogenous base found in

deoxyribonucleic acid (DNA) and ribonucleic acid (RNA), which chemically stores all relevant information necessary for cellular function and life. An understanding of the energetics and population distribution of protonated nucleobases in the gas phase acts as a proof of concept of the method as DMS-MS techniques are combined with computational calculations and previously gathered experimental data to determine method effectiveness. Because nucleobases are not found bare in nucleotides, the study was furthered by introducing a methyl group in the place of normal sugar attachment to mimic the state of nucleobases within DNA or RNA. We again determined the energetically favourable tautomeric forms of each methylated nucleobase and became confident in extending the gas phase research to larger systems.

A logical next step is to use the same gas phase techniques to study the interactions between nucleobases and nucleotides, particularly within a nucleic acid structure found in nature. Ribozymes are RNA structures with catalytic activity whose function directly depends on its higher order structure.<sup>6</sup> Due to the phosphate backbone within the nucleotides of the ribozyme, an overall negative charge is present which requires mass spectrometry techniques on the ribozyme to be done with a negative polarity. There is a lack of ion mobility based studies on nucleic acids in the negative mode in the literature, so not only does the research provide ribozyme dynamics information, but it also provides fundamental information regarding solvent-anion interactions with nucleic acids. Traditional methods of RNA probing including solution-phase NMR and X-ray crystallography, often require a relatively large amount of sample and long analysis time. Gas phase research provides an attractive alternative due to low concentration requirements (~ 100 ng/mL) and rapid analysis time.

A driving force behind this research is to push the limitations of DMS and how we can fully utilize its emerging potential as an analytical technique for biomolecules. Perhaps one of

the most important areas relevant to research today is any biomolecule that has an impact on the environment, particularly the increasing threat of antibiotic resistance. Antibiotics act to eliminate harmful microorganisms that cause illnesses in humans and animals. However, the overprescription and widespread agricultural use of antibiotics has led to microorganisms rapidly developing resistance mechanisms to the antibiotic, rendering them ineffective.<sup>7</sup> Many studies have been done to attempt to screen drug candidates, such as antibiotics, in the gas phase and correlate their gas phase behaviour to effectiveness against bacteria in an attempt to find new medicine.<sup>8-11</sup> The emergence of specific antibiotics, such as trimethoprim, in environmental wastewaters is an increasing concern as exposure to low doses can lead to development of resistance so a method to treat the wastewater and identify any metabolites and reaction products is essential. Common methods used to treat wastewaters include reacting it with hydroxyl radicals<sup>12,13</sup>; however the structure of metabolite transformation products are unknown so it remains difficult to assess the effectiveness of the treatment method and whether products have been properly mineralized for removal from water. DMS has been successfully employed to determine the transformation products of trimethoprim following a typical form of water treatment, and opens the door to potential environmental impacts of DMS, including analyzing complex sample analytes.

An extensive literature search has revealed an opportune spot for mass spectrometry based techniques for analyzing the structure, energetics, dynamics, and behaviour of relevant biomolecule clusters. In combination with quantum chemical calculations through a variety of different methods, DMS-MS gas phase probing of biomolecules has been shown to be a valuable method. This thesis outlines distinct methodologies to assess multiple classes of biomolecules, from nucleic acids to antibiotics, in order to gain better fundamental understandings of gas phase

chemistry as well as important diagnostic techniques useful for their medical, environmental, and physical impacts. Additional side-projects following similar themes were also investigated, such as exploring conformations of gold binding peptides and using thermometer ion species to determine the heat of ions within the DMS. However, for brevity and conciseness of this thesis, they will not be discussed.

## **2.0 Methodology**

### **2.1 Introduction to Computational Methods**

#### **2.1.1 Density Functional Theory**

Density functional theory (DFT) is an *ab initio* quantum chemistry method used to investigate the electronic structure of matter. It is currently perceived as an attractive technique to determine ground state properties of systems because of its compromise between a high degree of accuracy and relatively low computational cost. As its name suggests, density functional theory involves calculating the charge density from single electron orbitals and compares it to the charge density in the many-electron system to determine a ground state electronic charge density.<sup>14</sup> This ground state electronic configuration allows the predicted relative energies of systems to be computed and the densities can also be used for crystallographic structure predictions and comparison to IR, RAMAN, and NMR spectra.<sup>15-17</sup>

DFT has been shown to be highly accurate for a large number of systems, including the chemical species investigated in this thesis.<sup>18,19</sup> There are some molecules for which DFT may not be the ideal method for energetics calculations unless correction factors are introduced, so research needs to be done to determine what basis set is best for the particular system. For

example, sometimes DFT will spread out the charge density as electrons are over-delocalized or nearly degenerate energy levels may be neglected so calculations should always be checked to ensure rationality.<sup>20,21</sup>

### **2.1.2 Basin Hopping**

Basin hopping is a highly effective method that employs molecular mechanics calculations to exhaustively explore the potential energy surface of a particular system. The basin hopping routine is essentially a modified Monte Carlo simulation where the user guides the system towards the global minimum structure by imposing restrictions on accepted energies, geometries, and temperatures during random distortions to the molecule or system.<sup>22</sup> Rotations, dihedral angle modifications, and translations are all varied by random amounts within a specified range to generate a new structure which will be accepted or rejected based on its given energy. The basin hopping algorithm used in this work was created by Mike Lecours and a general flow chart following its logic is highlighted in Figure 2.1.

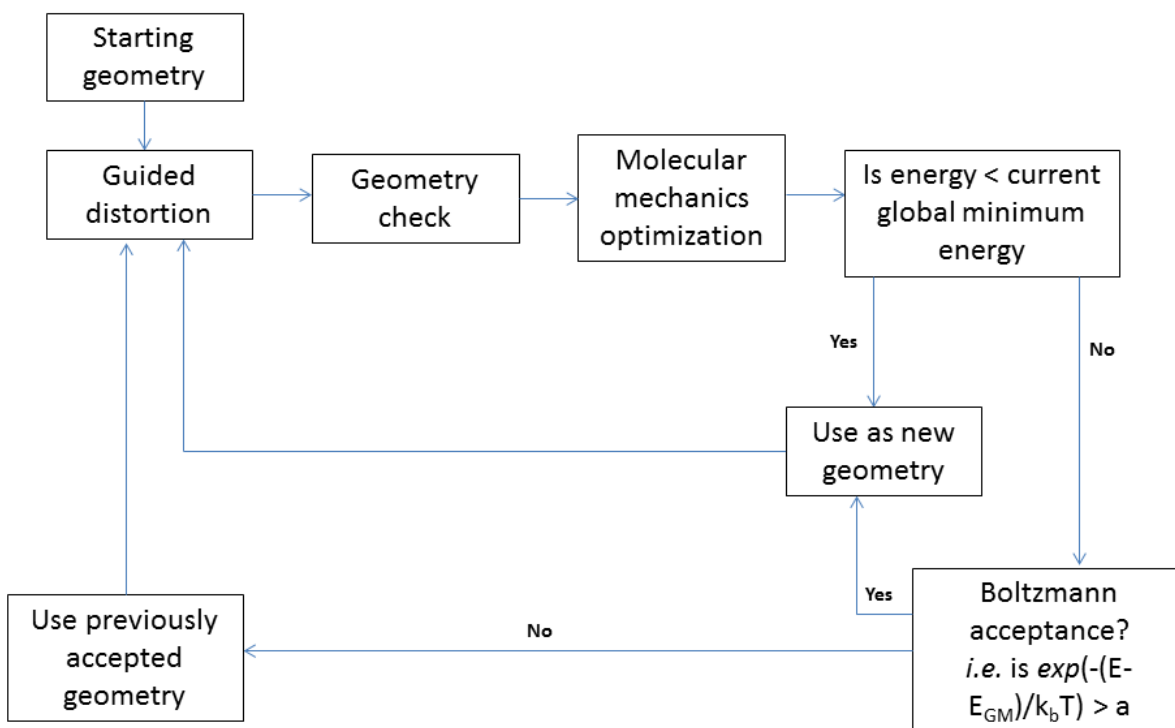
The basin hopping routine continues in the depicted manner until a user-specified number of unique geometries are accepted. The determined energies are rough but give a good prediction of low energy geometries that can be carried forward to more refined levels of theory, such as DFT, to more confidently identify the global minimum.

Molecular mechanics calculations vary in usefulness based on the user-defined force field. For biomolecules, especially proteins and nucleic acids, the AMBER force field has been shown to be effective in determining relative energies based on geometry and is defined as<sup>23</sup>:



$$V(r^N) = \sum_{bonds} k_b(l - l_0)^2 + \sum_{angles} k_a(\theta - \theta_0)^2 + \sum_{torsions} \sum_n \frac{1}{2} V_n[1 + \cos(n\omega - \gamma)]$$

$$+ \sum_{j=1}^{N-1} \sum_{i=j+1}^N f_{ij} \left\{ \epsilon_{ij} \left[ \left( \frac{r_{0ij}}{r_{ij}} \right)^{12} - 2 \left( \frac{r_{0ij}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \right\}$$



**Figure 2.1 General Scheme of a Typical Basin Hop Routine** The initial geometry is subject to user inputted random distortions before the routine decides if it is an acceptably low in energy structure based on its thermal Boltzmann distribution acceptance.  $E_{GM}$  represents the current global minimum energy and  $E$  is the energy of the current step.  $a$  is a random integer that is defined by the user (based on strictness of Boltzmann acceptance criteria), that falls between 0 and 1 to ultimately determine whether a structure should be accepted as a unique geometry.

The equation defines the potential energy of the system and is broken into four main summations. The first summations all assume the system is a harmonic oscillator and is split into three main parts. The first term involves energy due to bond stretching away from the equilibrium bond length. The second term involves energy due to bond bending and different distortion angles deviated from the equilibrium bond angle. The final harmonic oscillator term involves twisting and distortion energies, or torsion, of a bond based on its total degrees of freedom and is represented by a Fourier series. The final terms compute energies due to electrostatic interactions and van der Waals interactions.<sup>23</sup>

## **2.2 Introduction to Experimental Methods**

### **2.2.1 Differential Mobility Spectrometry**

A recently developed and optimized variant of ion mobility spectrometry (IMS), differential mobility spectrometry (DMS), and subsequent post-DMS MS techniques are extremely powerful experimental tools to probe, characterize, and study systems in the gas phase of varying size and complexity. It is an intriguing method to utilize in my Master's work to help bridge the gap between the study of the physicochemical properties of small, molecule systems and large biomolecules. DMS is principally used for the separation of isobaric species in the gas phase based on subtle differences in their differential mobility under atmospheric pressure. Following electrospray ionization, gas phase analyte ions travel through the instrument under the flow of a carrier gas between two parallel planar electrodes.<sup>24</sup> An asymmetric waveform separation voltage (SV) perpendicular to the flow of the carrier gas is applied which acts to tilt ions off of their linear trajectory towards either of the electrodes. A compensation voltage (CV) orthogonal to the gas flow axis is applied to return the ions back onto the cell axis such that they exit the DMS cell and enter Q1.<sup>24,25</sup>

Ions are separated based on mobility differences between the high and low-field conditions. The SV consists of a high-field portion and a low-field portion whose area under the curve is equal over time. The introduction of the high-field initially causes a zig-zag motion towards one of the electrodes and then as the SV waveform is changed to a low-field in the second half, the ion returns to its original distance from the electrode, in a slower manner.<sup>26</sup> The collision gas allows for differences between the ion's differential mobility in these alternating conditions to be amplified and described by its  $\alpha$  parameter (described below).

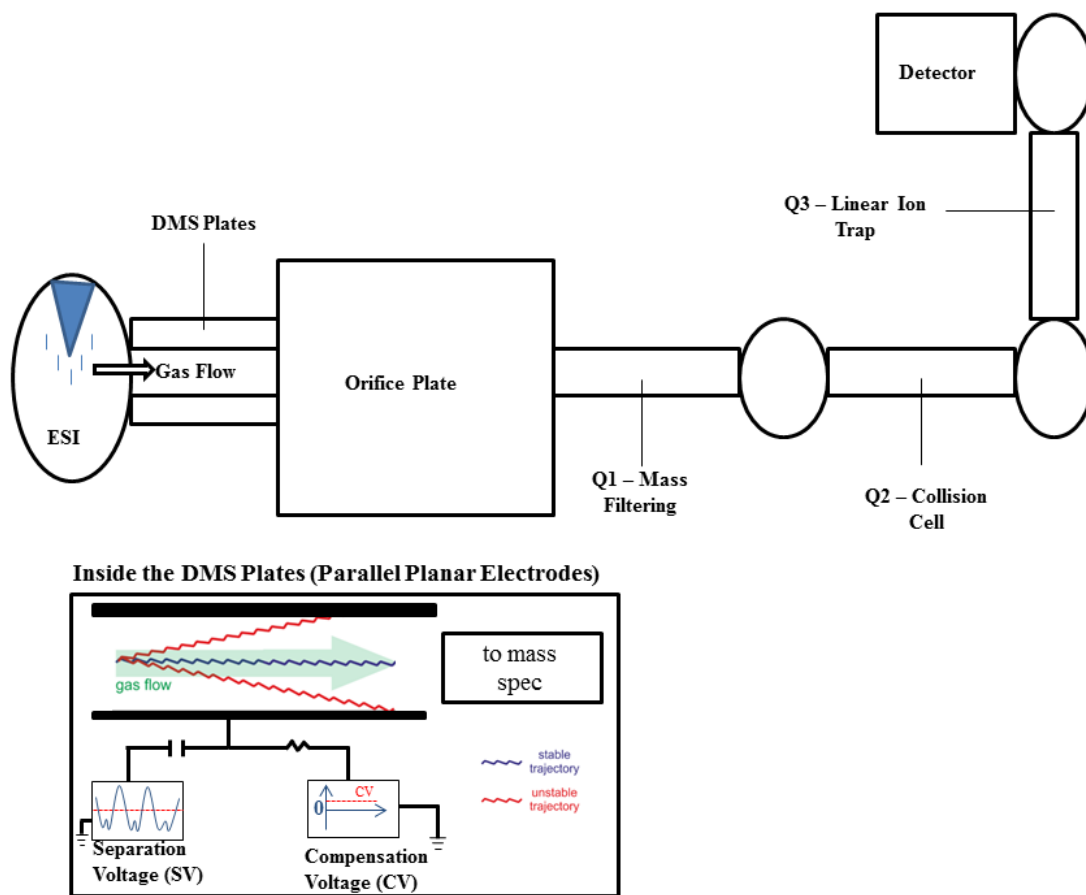
A schematic diagram of the process is shown in Figure 2.2. Therefore we can generate dispersion plots which characterize a particular ion at the CV value they are transmitted at a particular SV to provide information on the dynamic clustering environment between analyte and solvent.

Ion mobility through DMS has non-linear dependence on the asymmetric AC electric field and can be defined as  $K(E)$ . Mathematically, this can be expressed as<sup>26</sup>:

$$K\left(\frac{E}{N}\right) = K(0)\left[1 + \alpha\left(\frac{E}{N}\right)\right]$$

Where  $K(0)$  is mobility coefficient under the low-field  $\alpha(E/N)$  is a normalized function that demonstrates mobility changes via the following formula<sup>26</sup>:

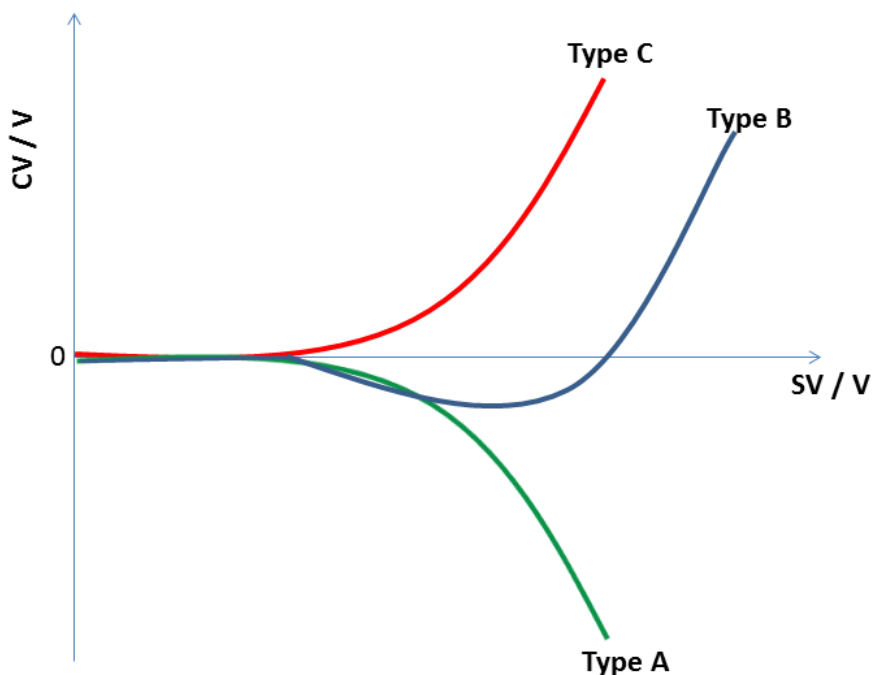
$$\alpha\left(\frac{E}{N}\right) = \frac{K\left(\frac{E}{N}\right) - K(0)}{K(0)}$$



**Figure 2.2 Schematic Diagram of DMS-MS and Within the DMS Plates.** The DMS-MS setup (above) involves an ESI source providing ions to the DMS plates which are carried *via* a carrier gas. Q1 acts as a mass filter, Q2 acts as a collision cell, while Q3 can further select species as a linear ion trap before reaching the final mass spectrometer detector. Within the parallel DMS plates (below), the asymmetric waveform of separation voltage changes the trajectory of ions towards each plate while compensation voltage returns the ion to a linear trajectory in order to reach the mass spectrometer.

The addition of chemical modifiers in the form of volatile solvents to the flow of the carrier gas has been shown to dramatically affect gas phase clustering interactions<sup>26</sup>, which can be visualized in dispersion plot trends due to altered transmitted CV values, as shown in Figure 2.3. Polar protic modifiers such as methanol or isopropyl alcohol, and aprotic modifiers such as acetonitrile or acetone are commonly used to study changes to dynamic clustering and

declustering within the DMS cell. Ion solvation occurs during the low-field portion of the separation voltage waveform, while spontaneous declustering occurs in the high-field portion. The variable geometries and differential mobilities are able to be better separated due to their clustering/declustering interactions. The first type of behaviour visualized in the dispersion plot, denoted Type A, is the result of strong clustering interactions between analyte and solvent, resulting in a sharp downward trend to negative CV values with increasing SV.<sup>26</sup> The clusters are formed during the low-field portion of the waveform and an increase to a high-field results in greater energy collisions which can cause ions to become released from the cluster or a reduction in cluster size, and thus an increase of the overall mobility of the analyte. In contrast, Type C behaviour is commonly referred to as “hard-sphere” behaviour due a lack of clustering interactions and only collisions with the carrier gas. This leads to an increasingly positive CV trend upon increasing SV values and is often observed when no modifier is present. The positive shift is due to the high energy collisions in the high field portion causing the analyte to lose momentum while this behavior is not observed in the low-field portion due to much less energy. The final displayed type on the dispersion plot is an intermediate behaviour, Type B, where initially clustering is observed and the CV shift is negative until a turnaround point is reached and declustering leads to positively shifted CV values at the highest SV values.<sup>26</sup> It is this turnaround point that has recently become a very interesting value to measure, as it has been implicated to a stronger correlation between DMS behaviour and molecular properties through the machine learning process, particularly in the case of studied pharmaceuticals in unpublished data from the Hopkins lab.



**Figure 2.3 Sample Dispersion Plot With Three Major Clustering Behaviours.** A sample dispersion plot which characterizes clustering behaviour of an ion in DMS by plotting the ion's optimal CV value at a specific SV. The three major types of clustering are shown and are Type A (strong clustering, green), Type B (intermediate, blue), and Type C (hard sphere collisions, red).

### 2.2.2 Collision Induced Dissociation

Following separation via DMS, a powerful MS technique used to aid in the identification of analyte ion structures is to fragment the selected ion through collision induced dissociation (CID) and monitor the formation of fragmented product ions.<sup>27</sup> The first quadrupole mass filter (Q1) selects for the desired  $m/z$  of the analyte of interest before passing through to Q2 and colliding with a relatively inert gas, such as argon or nitrogen. The resulting charged chemical species are detected in Q3 and are monitored to determine analytically meaningful products

along with increased collision energy. CID can provide useful information not only through the  $m/z$  of the fragment ions, but since it reflects relative bond energies in the structure which have been shown to be a useful method in *de novo* natural product structure elucidation.<sup>28</sup>

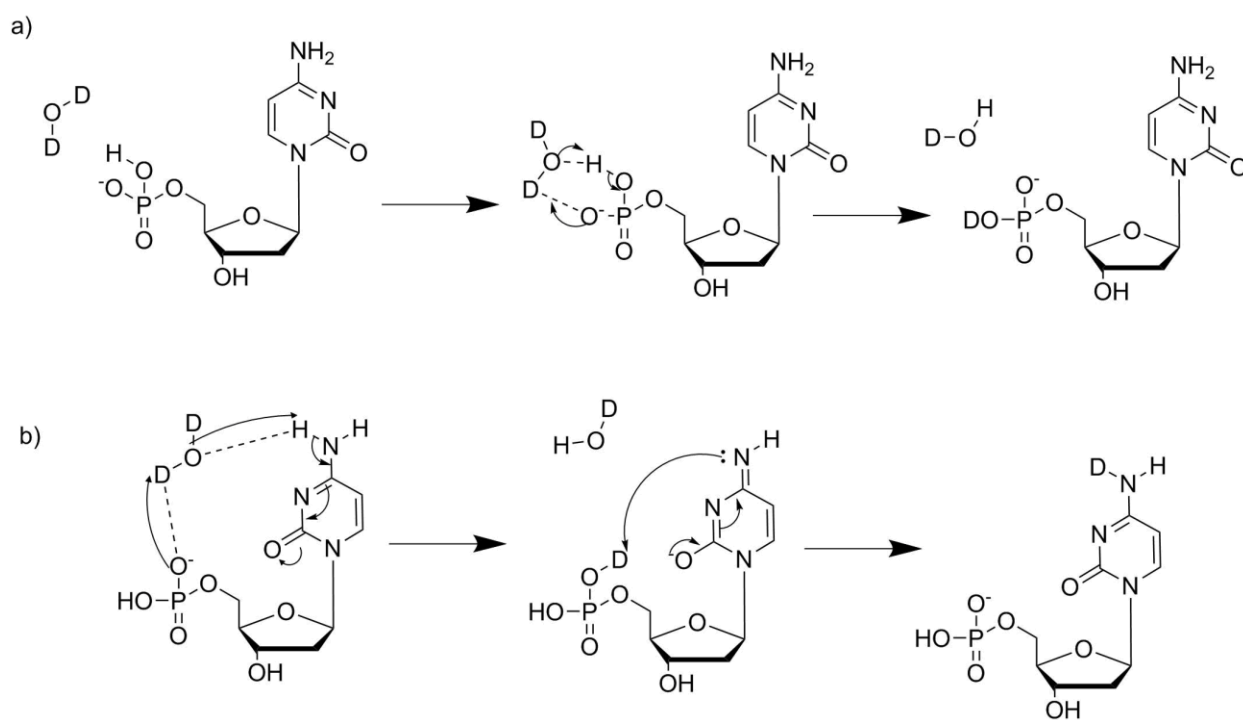
### 2.2.3 Hydrogen-Deuterium Exchange

One of the most important techniques in probing the structures and conformations of biomolecules is that of hydrogen-deuterium exchange (HDX). In gas-phase HDX, labile analyte hydrogen atoms are exchanged with deuterium from a deuterated solvent such as D<sub>2</sub>O or MeOD.<sup>29</sup> The number of deuterium exchanged can be directly counted from a mass spectrum as deuterium has a mass of 2 amu compared to 1 for a hydrogen atom. Exchange rates are directly correlated to gas phase basicity/acidity of the analyte proton in addition to how easily the solvent can access the exchange site due to bulk or steric effects.<sup>29</sup>

In the case of larger biomolecules, such as nucleic acids, much work on HDX has been carried out in the positive mode and research is needed to determine the HDX behavior of nucleic acids in the negative mode.<sup>29-31</sup> Robinson *et al.*, used Fourier-transform ion cyclotron resonance (FT-ICR) combined with HDX to determine that the negatively charged phosphate is essential in initiating the HDX mechanism of nucleotides in the negative mode and is illustrated in Figure 2.4, as no exchange was observed in cyclic nucleotides.<sup>32</sup> Within nucleic acids, there are multiple sites which can participate in HDX, including protonation sites on the nucleobase, O-H groups on the phosphate backbone, the 2' hydroxyl group of RNA, and any terminal hydroxyl group. Ligand-induced conformational changes to oligonucleotides, or other biomolecules such as proteins, have been observed using HDX experiments. If a ligand has a destabilizing effect on the analyte, it is likely an increase in the rate of deuterium uptake will occur due to exchange sites being more exposed to the deuterated solvent, while the opposite is

true for a decreased rate of deuterium uptake.<sup>29</sup> It has also been shown that hydrogen bonding can protect nucleobases from HDX in the positive mode, which infers that chemical modifier vapors in the DMS will also play a role in the HDX mechanism.<sup>30</sup>

Gas phase HDX has been used to study glycine oligomers<sup>33</sup>, peptides<sup>34</sup>, and proteins<sup>35</sup> and exchange is visible on a millisecond time scale, advantageous to conventional solution phase HDX which often has high degrees of back exchange.

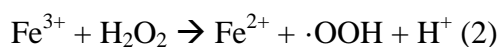
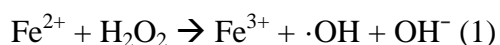


**Figure 2.4 Literature Proposed Negative Mode HDX Mechanism of Nucleotides.** (a) a proposed mechanism which solely relies on the negatively charged phosphate group for exchange and (b) a mechanism which involves interaction of the deuterated solvent with both the phosphate and nucleobase, with nucleophilic attack of nucleobase functional groups to abstract the deuterium from the phosphate. Adapted from reference.<sup>32</sup>



## 2.2.4 Fenton Reaction

Methods of organic reactions that can mimic biological processes are extremely important for being able to study environmental transformation processes in the lab. The Fenton reaction, discovered in 1894, has been shown to be a powerful method to transform many contaminants through the production of hydroxyl radicals and is used in organic pollutant cleanup, biodegradability improvement, and treatment of wastewaters.<sup>36</sup> Ultimately, the addition of iron and hydrogen peroxide generates hydroxyl radicals in the following reactions, which can react with organic pollutants.



As hydroxyl radicals become increasingly present in solution, there are four major behaviours which can occur<sup>37</sup>:

a) addition of the hydroxyl to the system (e.g.  $\cdot\text{OH} + \text{C}_6\text{H}_6 \rightarrow (\text{OH})\text{C}_6\text{H}_6$ )

b) hydrogen abstraction (e.g.  $\cdot\text{OH} + \text{CH}_3\text{OH} \rightarrow \text{CH}_2\text{OH} + \text{H}_2\text{O}$ )

c) electron transfer (e.g.  $\cdot\text{OH} + [\text{Fe}(\text{CN})_6]^{4-} \rightarrow [\text{Fe}(\text{CN})_6]^{3-} + \text{OH}^-$ )

d) radical interaction (e.g.  $\cdot\text{OH} + \cdot\text{OH} \rightarrow \text{H}_2\text{O}_2$ )

It is behaviours a) and b) which are desired in the transformation of organic pollutants, so parameters during the reaction must be strictly controlled to ensure the desirable outcome. The major parameter to control is the pH of the solution. An ideal pH is between 3-5 as higher pH values will precipitate the iron and decompose  $\text{H}_2\text{O}_2$  to  $\text{H}_2\text{O}$  and not produce the reactive hydroxyl species.<sup>38</sup> Several different strategies to drive the Fenton reaction and remove organic pollutants in industrial wastewaters have been shown and include photo-Fenton oxidation and electro-assisted Fenton oxidation, and ultimately can have a removal efficiency of over 95%.<sup>36</sup>

### 2.2.5 RNA Synthesis and Purification

Due to the expensive nature of commercial RNA synthesis, RNA sequences can be transcribed in the lab following the T7 polymerase transcription protocol from a corresponding DNA template. T7 RNA polymerase is an enzyme from the T7 bacteriophage that when given a DNA template with a promoter region, transcription cofactors, and free nucleotides, is able to carry out the synthesis of RNA from DNA in the 5' to 3' direction.<sup>39</sup> The enzyme requires  $Mg^{2+}$  as a cofactor for synthesis and dithiothreitol (DTT) is added to a transcription mixture in order to maintain the active oxidation state of the enzyme.<sup>40</sup> The optimal pH of the T7 transcription reaction is ~7.8, so the reaction is often carried out in a tris(hydroxymethyl)aminomethane (Tris) buffer in a 37 °C water bath for optimal enzyme temperature.

Following completion of the transcription, it is required to add urea to denature the T7 enzyme and EDTA to chelate free  $Mg^{2+}$ , so that further reaction or degradation does not occur.<sup>39</sup> To separate the synthesized RNA product from the reaction mixture and ensure the synthesis went to completion, gel electrophoresis is used. The method operates under the principle that RNA, which possesses a net negative charge, migrates towards a positive charge at the end of the gel following the application of an electric field.<sup>41</sup> As the RNA traverses the polyacrylamide gel, fragments are separated on the basis of size as small molecules, such as unreacted free nucleotides, will reach the end of the gel much faster than the sequence of interest. Based on the percent concentration of the polyacrylamide gel, it can be calculated how far the synthesized sequence should travel in relation to reference markers, such as blue dyes or molecular marker standards which qualitatively allow the user to know when to stop the process.

The ring systems of the nitrogenous bases in RNA absorb light at 260 nm. When illuminated with a UV light on a fluorescent plate, through a process known as UV shadowing,

RNA fragments appear as dark bands in the gel where they can easily be extracted using a razor blade.<sup>42</sup> To use synthesized RNA in mass spectrometric analysis, it is necessary that they are further purified to minimize potential contaminants. The RNA containing gel pieces can be added to a NaCl solution and rotated to allow for diffusion of nucleic acids into solution prior to the common procedure of ethanol precipitation. Cold ethanol effectively lowers the dielectric constant of the RNA containing solution as RNA precipitates out of solution and centrifugation can aggregate the now solid pelleted RNA from the supernatant aqueous ethanol solution.<sup>42,43</sup>

As a final purification step, it is desirable to remove any residual moisture and reduce sample volume to concentrate the RNA sample. This leads to a much more stable sample that can be stored for a longer period of time. A key method to accomplish this is through the vacuum drying technique of lyophilisation where the frozen sample is placed under vacuum conditions and cold temperatures to induce sublimation of moisture.<sup>44</sup> Moreover, ion exchange chromatography is often performed to remove impurities such as salt cations or leftover T7 RNA polymerase. This involves the use of an anion exchange column that has a positively charged resin which can interact with the negatively charged RNA, thus having a longer elution time that can be measured with UV/VIS spectrophotometry.<sup>45</sup>

### **3.0 Separation and Identification of Unmethylated and Methylated Nucleobase Tautomers**

This chapter is comprised of results featured in a published manuscript in the *International Journal of Mass Spectrometry* and additional work which has been conducted and prepared to be published. The study investigates the analytical separation power of differential mobility spectrometry on the nucleotide tautomers of cytosine, guanine, adenine, thymine, and uracil. Quantum chemical calculations in tandem with DMS and MS-MS techniques were used to rapidly and effectively probe and identify gas phase ensembles of protonated nucleobase tautomers. We show that the relative populations of tautomers can be tuned through variation of declustering potential and the ability of declustering potential to act as a verification method for interfering clusters. The study provides fundamental DMS information of small but biologically relevant molecules, that has been extended to investigate methylated versions of the nucleobases with the methyl group in place of (deoxy)ribose attachment to mimic biological relevance. Ultimately, the methodology and information obtained here will be used to improve fundamental understanding of nucleic acids and provide insight into similar studies for larger nucleic molecules, such as ribozymes. A special thanks to Stephanie Ryall for her work on the methylated nucleobases and figure preparation.

#### **3.1 Introduction**

Nucleobases are the nitrogenous bases found in deoxyribonucleic acid (DNA) and ribonucleic acid (RNA) and consist of the of the single-ring pyrimidines of cytosine (C), thymine (T), and uracil (U), as well as the double-ring system purines of guanine (G), and adenine (A). These nucleobases play the vital biological role of encoding and expressing genetic information as well as maintaining key nucleic acid structural motifs relevant to the function of nucleic acids,

discussed in Chapter 4. Much work has been done to theoretically<sup>46-49</sup> and experimentally<sup>50-53</sup> investigate physicochemical properties of nucleobases, including gas-phase acidity and basicity determination.<sup>54-60</sup> The determination of protonation site and tautomeric forms of the nucleobases is very relevant in modern research due to their implication in mutagenic processes and stabilization of higher order structures.<sup>61</sup> Based on previous experimental and theoretical work, including infrared multiple photon dissociation (IRMPD), spectroscopic studies have shown the existence of multiple tautomeric forms of nucleobases existing in the gas phase.<sup>62,63</sup> For example, Salpin *et al.* observed at least two tautomeric forms of (C + H)<sup>+</sup>, (T + H)<sup>+</sup>, and (U + H)<sup>+</sup> through IRMPD work,<sup>64</sup> while Bakker *et al.* used vibrational spectroscopy of monohydrated protonated uracil and cytosine to infer the presence of multiple tautomeric forms.<sup>65</sup> Corresponding quantum chemical calculations at the B3LYP/6-311++G(d,p) level of theory suggested that the species corresponded to keto-enol tautomeric forms with a small subset of oxo tautomers. This leads to the postulation of keto-enol tautomerisation in the gas phase as a possible mechanism of isomeric interconversion which may be accessible in biologically relevant conditions.<sup>66</sup> Therefore it is likely that molecular properties of protonated nucleobases in the gas phase are due to the contribution from multiple tautomeric structures present at varying populations depending on experimental parameters leading to a need for structural separation prior to gas phase spectrometric or spectroscopic analysis.

The separation of tautomeric forms of ions principally utilizes ion mobility spectrometry, such as travelling wave ion mobility spectrometry.<sup>67-70</sup> In this work, we demonstrate that the solution phase population distribution of an ion is not consistently reflected in ESI-MS results and in fact varies as a function of several experimental parameters and source conditions. We use DMS to probe and separate tautomeric nucleobase in the same manner that has previously been

reported in the Hopkins lab study on N- and O-protonated tautomers of *para*-aminobenzoic acid.<sup>71</sup> An extended post-DMS MS-MS characterization of each nucleobase tautomer was conducted to observe their physicochemical properties and determine their chemical structure, in conjunction with quantum chemical calculations. In addition to CID, HDX behaviours have been shown to be different between tautomeric forms and relative to solvent flow rate. As described in the experimental section, chemical modifiers were used to enhance tautomeric separation and it was observed that caution must be exercised, particularly when using strong polar protic modifiers or even high vapor pressures of HDX reagent, to avoid possible *in situ* tautomerization *via* proton-transfer relay networks established upon ion-solvent clustering.<sup>71</sup> This leads to the important realization that ESI solvent effects are not the only criteria which determine relative isomer distributions in the gas phase and that it is also affected by instrumental conditions that can tune sub-populations within a gas phase ensemble.

Following successful separation and identification of bare protonated nucleobase tautomers, the logical next step was to further extend the analysis to larger systems with a greater degree of biological relevance. Within the structure of nucleic acids, isolated nucleobases are only found in synthetic and salvage pathways<sup>72</sup> while higher order structures are composed of nucleotide polymers where nucleobases are joined to a sugar group and negatively charged phosphate backbones. In place of ribose or deoxyribose, we placed a methyl group at the site of sugar attachment. This eliminates one of the bare nucleobase protonation sites and adds a flexible electronic group to mimic the steric effects of the sugar. Nucleosides can be methylated at either the nucleobase or ribose, and methylated nucleobases have been shown to have distinct roles in nature, particularly within RNA. For example, the degree of methylation of rRNA subunits has been linked to antibiotic resistance.<sup>73</sup> The addition of methyl groups to nucleotides

also have biological implications and can be necessary to distinct processes. Perhaps the most well-known example is the addition of methyl groups to the polyA tail of processed mRNA (7-methylguanine and N6-methyladenine) implying a possible regulatory role.<sup>74</sup>

The addition of methyl groups in the site of sugar attachment has been used in previous nucleobase studies, again for simplicity and biological relevance. Chin *et. al*, used 9-methylguanine to block the possibility of 7/9NH tautomerization (so only keto-enol tautomerization was possible), before using optical spectroscopy methods and quantum mechanical calculations to determine tautomer and rotamer populations.<sup>75</sup> They observed inefficiency of UV spectroscopy to detect all relevant expected rotamers of 9-methylguanine and its monohydrate due to species with very similar energies, which provides a need for methods such as DMS to supplement theoretical tautomer population calculations.<sup>75</sup>

## 3.2 Methods

### 3.2.1 Experimental Methods

**Reagents and Chemicals.** Dried, solid nucleobases (cytosine, guanine, adenine, thymine, and uracil) as well as their methylated derivatives (9-methyladenine, 9-methylguanine, 1-methyluracil, and 1-methylthymine) were purchased from Sigma-Aldrich. 1-methylcytosine was not available at the time of this research.

**Differential Mobility – Mass Spectrometry.** DMS experiments were carried out using a SelexION differential mobility spectrometer paired with a QTRAP 5500 mass spectrometer (SCIEX). Positive mode ESI was used with an electrospray voltage of 5.5 kV, an ambient source temperature (~32 °C), nebulizing gas pressure of 20 psi, and auxiliary gas pressure of 0 psi. Separations were conducted at a low DMS temperature setting corresponding to 150 °C. Nitrogen was used as the curtain gas (20 psi) and collision gas for MS-MS experiments (~9

mTorr). Analyte solutions of each protonated nucleobase were made to concentrations of 10 ng/mL or 100 ng/mL in a 1:1 mixture of milli-Q water and methanol with an added 0.1% (by volume) formic acid to aid in protonation.

DMS experiments involved running each sample at separation voltages (SV) from 0 V to 4000 V in 500 V increments. At each SV, the compensation voltage (CV) was ramped from -80 V to 15 V with a 0.1 V step size to generate ionograms. Dispersion plots, as discussed in Chapter 2, were generated to observe ion transmission patterns as a function of SV and CV. To determine the effects of solvent modifiers, each set of ionograms for each nucleobase was run in (a) a dry N<sub>2</sub> environment, (b) a pure N<sub>2</sub> environment seeded with 1.5% (mole ratio) methanol, and (c) a pure N<sub>2</sub> environment seeded with 1.5% (mole ratio) isopropyl alcohol (IPA). To increase confidence in isomeric structure assignment, collision induced dissociation (CID) was performed on each ionogram peak at an SV of 3500 V and its corresponding CV. Breakdown curves were generated, which plot fragment ion intensities as a function of collision energy (CE), through the ramping of CE of the Q2 ion trap from 0 V to 60 V in 0.25 V increments. Mass spectra were recorded at each CE interval through the use of enhanced product ion (EPI) scans. The EPI scan specifically selects for the *m/z* of the protonated nucleobase to be fragmented and records all fragment ion *m/z* intensities which came from the parent ion.

Hydrogen-deuterium exchange (HDX) involved the infusion of deuterium oxide into the throttle gas at varying rates of saturation corresponding to different rates of D<sub>2</sub>O vapor production. The first set of experiments involved bubbling of D<sub>2</sub>O to saturate the N<sub>2</sub> with deuterated reagent while the second involved sampling of the headspace above the liquid D<sub>2</sub>O reagent leading to slower rates and lower D<sub>2</sub>O partial pressure. Through these experiments, mass



spectra of each suspected isomer peak corresponding to a unique ionogram peak were generated both prior to and post deuterium exchange.

### 3.2.2 Computational Methods

Using chemical intuition, all possible tautomeric forms and protonation sites of each nucleobase were explored computationally. All were geometrically optimized with frequency calculations carried out using DFT at the B3LYP level of theory and 6-311++G(d,p) basis set. Thermochemical corrections and energetic parameters were generated through the harmonic frequency calculations which also produced vibrational spectra in the cases of (C + H)<sup>+</sup>, (T + H)<sup>+</sup>, and (U + H)<sup>+</sup> to compare with IRMPD data. The four tautomers of each nucleobase with the lowest standard Gibbs energies were then computationally treated with the coupled cluster single, double, and perturbative triple excitations method CCSD(T)/6-311++G(d,p) level of theory to further improve electronic energies and are reported here.

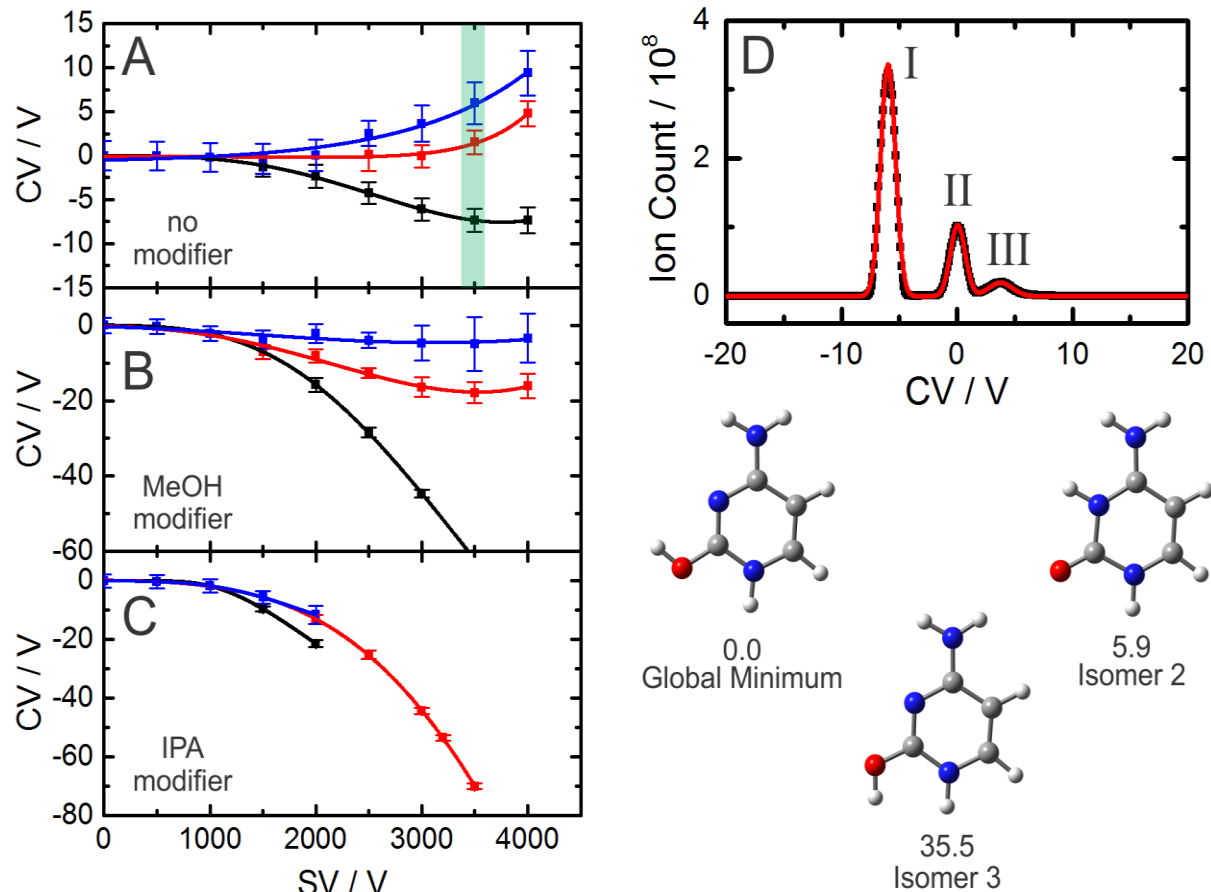
## 3.3 Results and Discussion

Although all experiments were run with all protonated nucleobases, the results of protonated cytosine will be discussed at length as a representative case while the data for all other unmethylated nucleobases will be provided.

### 3.3.1 Protonated (C + H)<sup>+</sup> Dispersion Plots

Figure 3.1A-C displays the dispersion plots recorded for (C + H)<sup>+</sup> (*m/z* 112) in a pure N<sub>2</sub> environment, N<sub>2</sub> seeded with 1.5% MeOH vapor, and N<sub>2</sub> seeded with 1.5% IPA vapor, respectively. The presence of three distinct peaks suggests the presence of at least three different tautomers, with optimal separation achieved in the unmodified pure N<sub>2</sub> environment. Type C and

weak Type B clustering were observed in the unmodified environment while MeOH- and IPA-modified environments saw ion trajectories shifted to lower CV values<sup>76–78</sup> and stronger Type A behaviour due to ion-solvent clustering interactions involving (C + H)<sup>+</sup> tautomers.<sup>79</sup> The optimized tautomer separation in pure N<sub>2</sub> seen at an SV = 3500 V is shown in Figure 3.1D and three distinct peaks are clearly resolved. Interesting to note that weaker intensity features diminish at high SV (greater than 2500 V) in IPA, which is likely due to *in situ* tautomerization induced by IPA<sup>78</sup> or proton scavenging from (C + H)<sup>+</sup> to IPA due to IPA having a higher gas phase basicity than the probed tautomers under the experimental parameters.



**Figure 3.1 Dispersion Plots and Ionogram for (C + H)<sup>+</sup>.** The dispersion plots obtained for (C + H)<sup>+</sup> (*m/z* 112) in (A) a pure N<sub>2</sub> environment, (B) a N<sub>2</sub> environment seeded with 1.5% (mole ratio) methanol vapor, and (C) a N<sub>2</sub> environment seeded with 1.5% (mole ratio) isopropyl alcohol vapor. Error bars are 2σ obtained from Gaussian fits to the ionogram peaks. (D) The ionogram of (C + H)<sup>+</sup> recorded in a pure N<sub>2</sub> environment at SV = 3500 V (highlighted in green in A). Inset:

The three lowest energy tautomers of  $(C + H)^+$  as calculated at the CCSD(T)/6-311++G(d,p)//B3LYP/6-311++G(d,p) level of theory. Reported energies are standard Gibbs energies in kJ/mol relative to the global minimum.

### 3.3.2 Computational $(C + H)^+$ Results

To correlate the ionogram peaks to structures, further experiments needed to be conducted in addition to their DMS clustering behaviour. Following high level electronic energy calculations using CCSD(T), three low energy  $(C + H)^+$  tautomers were identified and are displayed along with their relative energies in Figure 3.1 (inset). Assuming ionization efficiencies proportional to relative energies in Figure 3.1D, it was initially inferred that the global minimum O-protonated form corresponded to the largest intensity peak (I), the N-protonated form labelled Isomer 2 corresponded to peak II, and the alternate conformation of O-protonated  $(C + H)^+$  corresponded to peak III. A 5.9 kJ/mol difference in energy between the global minimum and isomer 2 structure would be seen as relatively close ionization efficiencies in an ionogram, while isomer 3, at an energy 35.5 kJ/mol greater than the global minimum, would be much less likely to be observed in the gas phase without the help of kinetic trapping. The presence of three peaks for  $m/z$  112 does not agree with IRMPD results from Salpin *et al.* who observed the presence of only two tautomers, which can be due to a variety of reasons.<sup>64</sup> It is possible that the different methods used to generate gas phase ions of protonated cytosine resulted in different relative tautomeric populations or that similar conformational isomers (*i.e.*, the global minimum and Isomer 3) display similar vibrational spectra in the 1000 – 2000  $\text{cm}^{-1}$  region, which was supported by our frequency calculations. This could potentially lead to the IRMPD studies to dismiss the overlap in this region as only one isomer when multiple tautomers

could be present. To solve this problem, it was decided that additional post-DMS spectrometric methods should be used to further verify the identity of the resolved ionogram peaks.

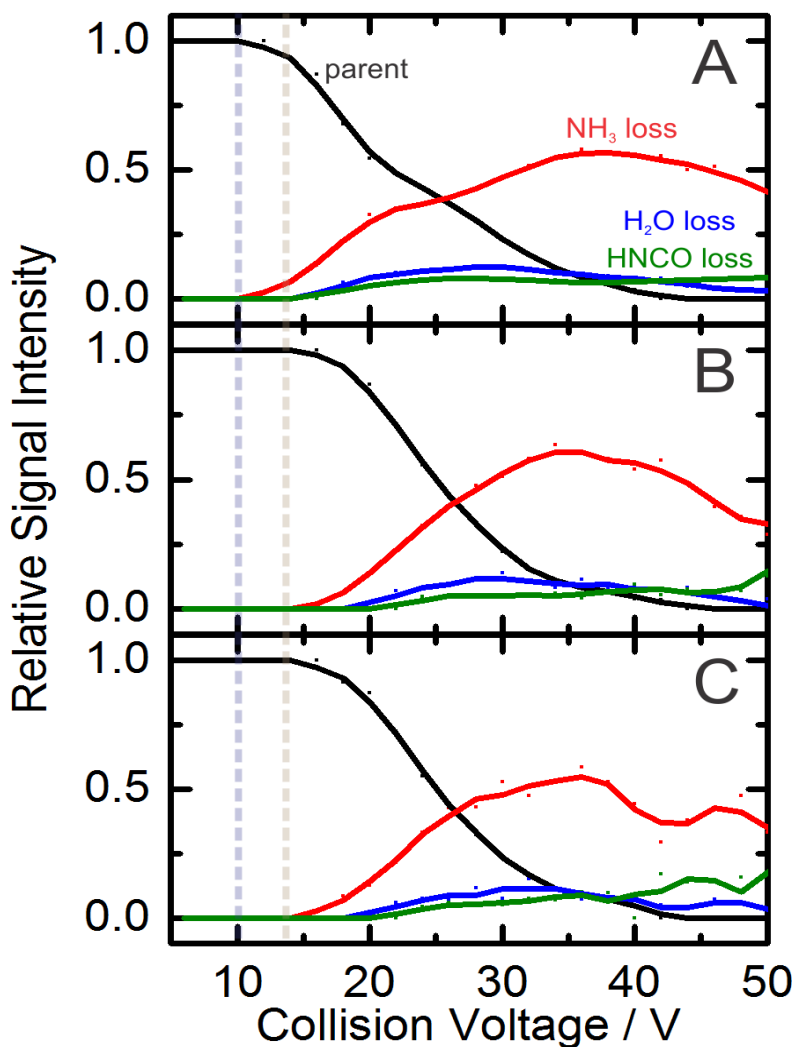
### 3.3.3 Collision-Induced Dissociation Results for $(C + H)^+$

The first method of post-DMS isomer structure identification was through the employment of CID, as explained in the experimental section. A breakdown curve displaying the onset of fragmentation patterns for each ionogram peak is displayed in Figure 3.2 and is shown for (A) peak I, (B) peak II, and (C) peak III. The expected fragments of protonated cytosine have been discussed in literature and agree with the loss of  $NH_3$ ,  $H_2O$  and  $HNCO$  ions we observed here.<sup>80-82</sup> The major difference between the breakdown curves is the onset of when each fragment loss is first observed.  $NH_3$  and  $H_2O$  loss from peak I are seen at a CE  $\sim 5$  V lower than when the fragments are first seen in peaks II and III, which supports evidence of at least two tautomeric forms while peaks II and III still cannot be differentiated as unique tautomers from each other.

### 3.3.4 Hydrogen-Deuterium Exchange Results for $(C + H)^+$

An additional tool in our DMS-MS setup is the ability to perform HDX on differential mobility separated species to note differences in deuterium exchange and uptake. Looking at the structure of  $(C + H)^+$ , there are a total of four protons available for exchange, but there is little evidence of exchange of the third and fourth nuclei in both levels of  $D_2O$  saturation. Additionally, all three peaks exhibit clearly different HDX behaviour which is depicted in Figure

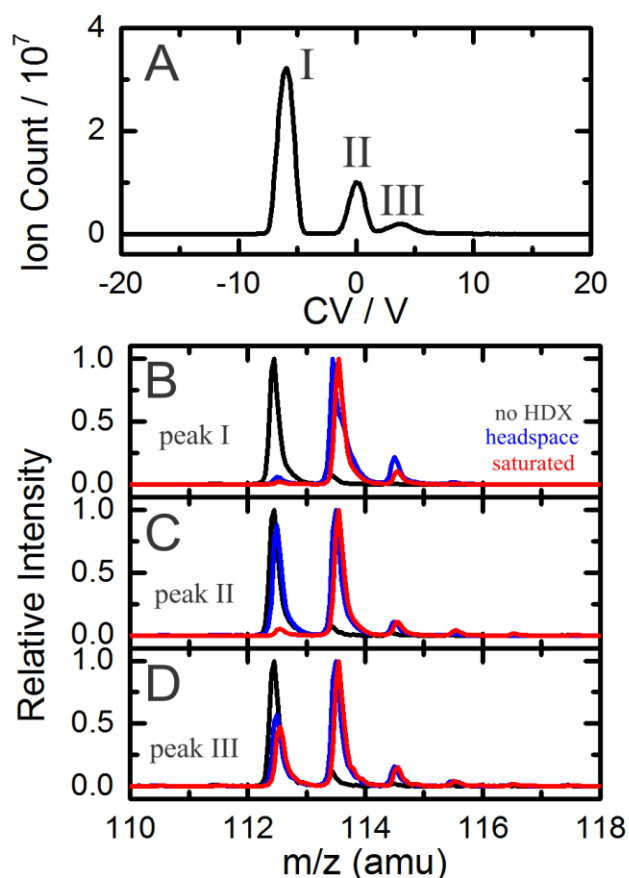
3.3. Peak I displays rapid exchange of one H atom which would be expected from the postulated structure due to the acidity of the protonated carbonyl within the global minimum structure.



**Figure 3.2** The breakdown curves obtained for  $(C + H)^+$  ( $m/z$  112) at each isolated ionogram peak displayed in Figure 3.1D. (A) Peak I at a CV = -6 V, (B) Peak II at a CV = 1 V, and (C) Peak III at a CV = 4 V. Collision energy was increased in 2.5 V increments from 0 to 50 V and mass spectra were obtained at each increment.

In contrast, peak II shows relatively low first H exchange under low  $D_2O$  partial pressure while going to completion under high  $D_2O$  partial pressure. The differences due to partial pressure may be indicative of HDX-induced tautomerization or saturation of the environment under high

partial pressure. However, a lower rate of HDX in comparison to peak I would correspond to the lower relative acidity of the protonated nitrogen atoms in the postulated isomer 2.



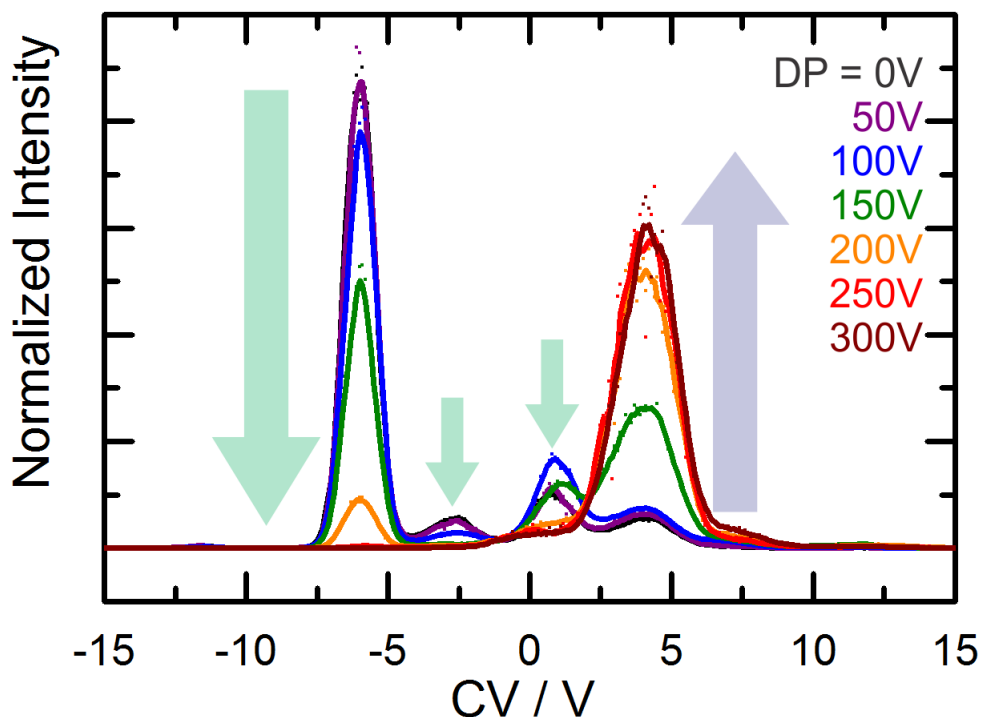
**Figure 3.3 HDX Results for  $(C+H)^+$  ( $m/z$  112)** (A) The ionogram recorded for  $(C + H)^+$  ( $m/z$  112) at SV = 3500 V. The results of HDX when gated on CV corresponding to (B) peak I, (C) peak II, and (D) peak III. The black traces show ionograms in the absence of  $D_2O$  while the blue trace shows a low partial pressure introduction of  $D_2O$  and the red displays HDX in  $N_2$  at 18 °C with a saturated partial pressure of  $D_2O$

The HDX profile obtained for peak III is visibly different to those obtained for peaks I and II and at first glance is difficult to interpret. When saturated with  $D_2O$  at a high partial pressure, the first proton is not completely exchanged for the structure corresponding to peak III indicating a lower rate of HDX compared to peaks I and II. However, in the low partial pressure environment, the rate of HDX of peak III surpasses peak II while the overall HDX profile for peak III appears identical in low and high  $D_2O$  partial pressure environments. The inconsistency

and unexpected results obtained for peak III lead to a hypothesis that source conditions or alternate DMS-MS parameter variations were responsible for the unexplained effect.

### 3.3.5 Declustering Potential Variation for $m/z$ 112

The variations of source conditions including flow rate, ESI voltage, and solution concentrations contributed no observable differences in subsequent ionogram recordings of  $(C + H)^+$ . It was not until the declustering potential (DP) was varied during the experimental setup that different results were obtained and insight could be shed on the tautomeric peak assignments. The ionogram for  $(C + H)^+$  was recorded at an  $SV = 3500$  V and CV value corresponding to the location of each suspected peak while each ionogram was run at declustering potentials of 0 to 300 V in 50 V increments and the results are shown in Figure 3.4. An increase in declustering potential allows for larger mass clusters to pass through Q1 before exposed to the increased energy provided by the DP (*i.e.*, comparable to CID) which would lead to a decreased ion intensity for a bare, unique tautomer. As expected, peaks I and II both depleted proportionally to an increasing DP, while peak III surprisingly increased in intensity. The unsuspected behaviour of peak III implies that it is not a unique tautomer of  $(C + H)^+$ . We propose that given the increased ion intensity at high DP values, peak III corresponds to  $(C + H)^+$  clustering with solvent with the solvent dissociating after Q1 when the energy available is sufficient and releasing a bare species of  $m/z$  112. This is further supported by mass spectra obtained at low DP showing masses corresponding to  $(C + H)^+$  clustered with protons of the solvent, specifically water and formic acid. Consequently, we have determined that monitoring the ion intensity changes of ionogram peaks as a function of declustering potential is an adequate way of determining the presence of ion-solvent clusters which had previously not been addressed in DMS studies.



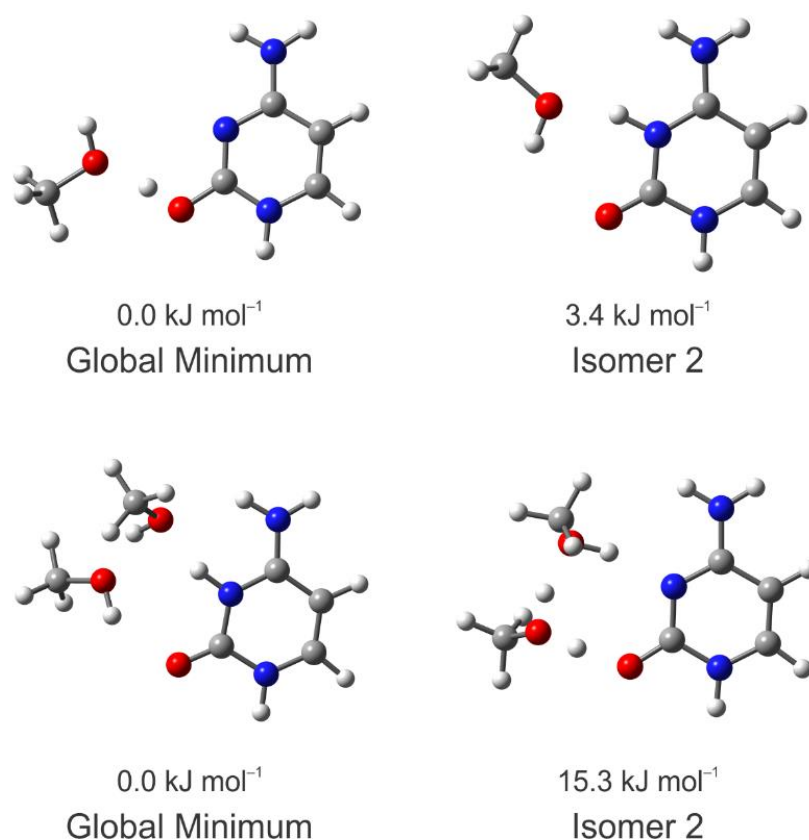
**Figure 3.4 Declustering Potential Variation of  $(C + H)^+$**  The ionogram obtained for  $(C + H)^+$  ( $m/z$  112) at  $SV = 3500$  V in a pure  $N_2$  environment as declustering potential is varied in 50 V increments from 0 V (black) to 300 V (brown).

### 3.3.6 $(C + H)^+$ Ionogram Peak Assignments and Analysis

Based on the above analysis, we have confidently separated and identified two unique tautomers of  $(C + H)^+$ . Peaks I and II are assigned to the global minimum structure and isomer 2, respectively, while peak III is assigned to a ion-solvent cluster that fragments to produce  $m/z$  112 with sufficient energy. The two identified structures accord with IRMPD spectroscopic data and their structural similarity lead to a proposed mechanism of interconversion via proton transfer.<sup>64</sup> A network of polar protic solvents was shown to facilitate proton transfer through a hydrogen bond network to readily interconvert  $(C + H)^+$  tautomers as demonstrated in Figure 3.5 with methanol. Therefore interconversion between isomers prior to MS characterization is likely in



methanol-modified and high partial pressure HDX environments, and is supported by a loss in signal intensity for weak features in alcohol-modified environments.

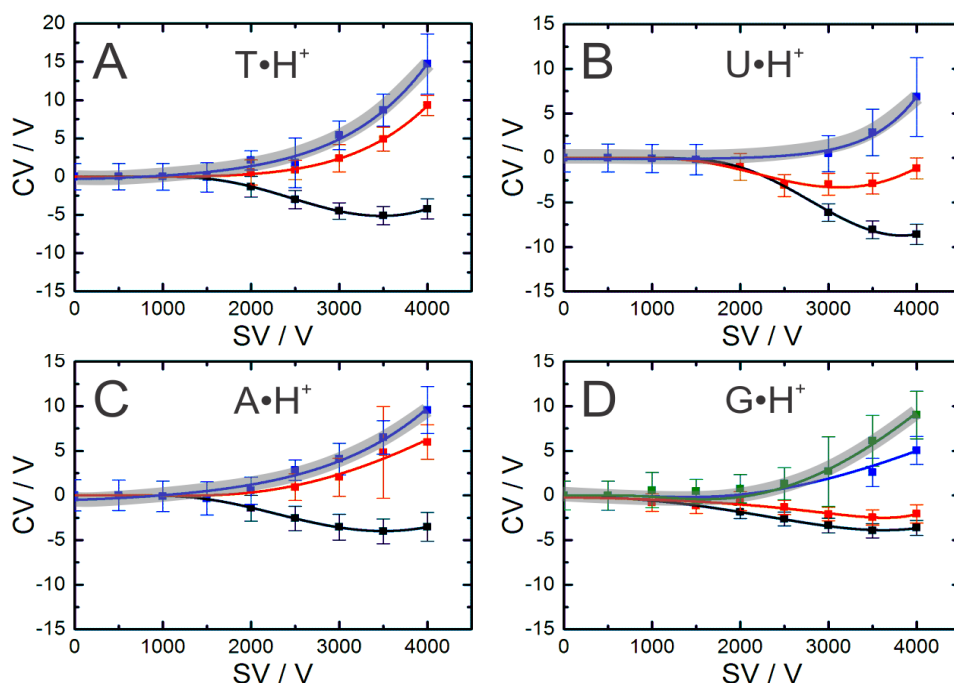


**Figure 3.5** The lowest energy proton-bound cytosine-methanol clusters, with one and two methanol molecules placed near the site of protonation and optimized. Proton transfer was observed along the intermolecular hydrogen-bond network between protonation sites which could facilitate interconversion between the global minimum and isomer 2. Standard Gibbs energies are reported at the B3LYP/6-311++G(d,p) level of theory.

### 3.3.7 Results for (A + H)<sup>+</sup>, (T + H)<sup>+</sup>, (U + H)<sup>+</sup>, and (G + H)<sup>+</sup>

The same methodology discussed to separate, probe, and identify tautomers of (C + H)<sup>+</sup> was carried out with the rest of the bare protonated nucleobases. The dispersion plots recorded for each in a dry N<sub>2</sub> environment are shown in Figure 3.6 while the dispersion plots in MeOH- and IPA-modified environments are shown in Appendix A. The breakdown curves and HDX for each protonated tautomer can also be found in Appendix A. Within each dispersion plot, a trace

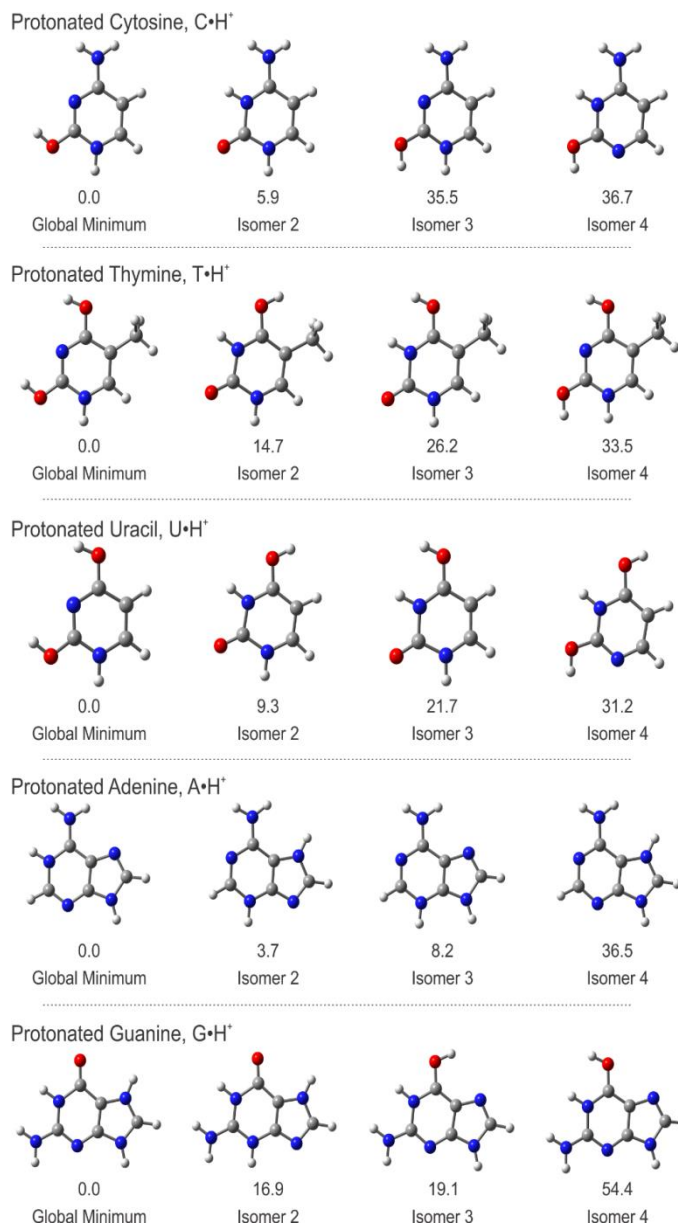
highlighted in grey indicates that the specific species increased in intensity at high DP values and can be attributed to a higher order cluster as opposed to an observed unique tautomer. The number of tautomers observed for each protonated nucleobases accords with the number of isomers observed in IRMPD data as we see two tautomers for  $(A + H)^+$ ,  $(T + H)^+$ , and  $(U + H)^+$ , and three unique tautomers for  $(G + H)^+$ . In the case of  $(G + H)^+$ , one of the three signals accounts for  $\sim 80\%$  of the total ion signal indicating a global minimum structure significantly more stable and low in energy than the next lowest in energy structures.



**Figure 3.6 Dispersion Plots for Protonated Nucleobases in  $N_2$ .** The dispersion plots for  $(T + H)^+$ ,  $(U + H)^+$ ,  $(A + H)^+$ , and  $(G + H)^+$  obtained in a dry  $N_2$  environment. A greyed out trace indicates the peak was attributed to a larger cluster that fragmented to produce the ion of interest post-DMS.

Again, electronic structure calculations at the CCSD(T)/6-311++G(d,p)//B3LYP/6-311++G(d,p) level of theory were conducted to estimate relative standard Gibbs energies of the possible protonated nucleobase tautomers. The four lowest energy tautomers for all five of the protonated nucleobases are shown in Figure 3.7. Based on the calculated relative energies of

each protonated nucleobase tautomer, there is a direct correlation between the energies and number of tautomers experimentally resolved through DMS. It appears that for  $(C + H)^+$ ,  $(T + H)^+$ , and  $(U + H)^+$  there are two relatively low tautomeric forms that can interconvert via solvent-mediated proton transfer between adjacent basic sites, while the third lowest energy tautomer is at least 20 kJ/mol higher in energy than the global minimum and are kinetically trapped during the ESI process. Therefore the isomers >20 kJ/mol higher in energy than the global minimum will have negligible contributions to the ensemble population in a stochastic Boltzmann distribution under the tested experimental conditions. In the case of  $(A + H)^+$ , the calculated energies suggest that three tautomers should be resolved *via* DMS while we only see two resolved structures. It should be noted that at low DP values, a slight peak asymmetry to the low-CV side of the most intense ion peak is observed which may be attributed to this third tautomer but unable to be fully resolved under the current conditions and is shown in the Supporting Information. For  $(G + H)^+$ , we observe one dominant ion which correlates to the much lower in energy of the global minimum to the second and third lowest energy structures at 16.9 kJ/mol and 19.1 kJ/mol, respectively. Table 3.1 outlines the observed percentage distribution of each structure at low and medium DP values in comparison to expected calculated values. There appears to be significant differences between observed and calculated percentage distributions in the case of all protonated nucleobases except for  $(U + H)^+$ . Based on our observations with varying DP values, we propose that the gas phase ensembles of protonated nucleobases are generated and kinetically trapped in non-equilibrium conditions, which can be tuned by varying the discussed instrument conditions.



**Figure 3.7 Lowest Energy Structures of each Protonated Nucleobase.** The four lowest in energy tautomers for each protonated nucleobase as calculated at the CCSD(T)/6-311++G(d,p) level of theory with B3LYP/6-311++G(d,p) thermochemical corrections. Standard Gibbs energies are reported in kJ/mol.

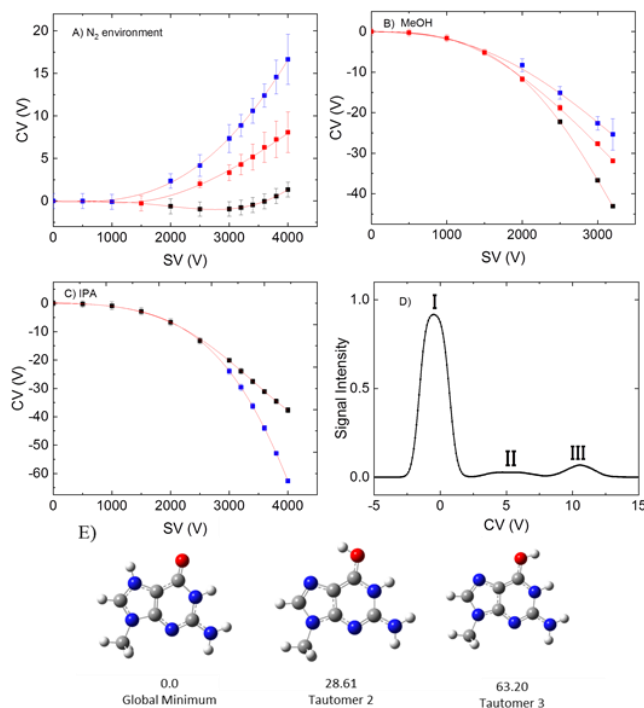
**Table 3.1 Population Distributions of Each Nucleobase Tautomer.** The percentage of the ensemble population of the lowest energy tautomer forms of each protonated nucleobase at DP = 0V, DP = 150 V, and theoretical calculated values.  $1\sigma$  error is reported in parentheses. Calculations were carried out at the CCSD(T)//B3LYP level of theory employing a 6-311++G(d,p) basis set.

Species	Percentage of Population		
	DP = 0 V	DP = 150 V	Calculated
<u>(C + H)<sup>+</sup></u>			
Global Minimum	79.5 (0.4)	73.5 (0.5)	91.5
Tautomer 2	8.1 (0.5)	26.5 (0.9)	8.5
Tautomer 3	12.4 (0.5)	0.0	0.0
<u>(A + H)<sup>+</sup></u>			
Global Minimum	73.6 (4.2)	100.0	79.3
Tautomer 2	17.7 (1.2)	0.0	17.8
Tautomer 3	8.7 (4.1)	0.0	2.9
<u>(G + H)<sup>+</sup></u>			
Global Minimum	74.7 (0.4)	79.5 (0.4)	99.8
Tautomer 2	10.8 (4.5)	20.5 (1.3)	0.1
Tautomer 3	14.6 (1.8)	trace	0.0
<u>(T + H)<sup>+</sup></u>			
Global Minimum	30.9 (0.5)	86.6 (0.9)	99.7
Tautomer 2	39.4 (1.4)	13.4 (1.0)	0.3
Tautomer 3	29.7 (1.5)	0.0	0.0
<u>(U + H)<sup>+</sup></u>			
Global Minimum	91.1 (0.3)	94.6 (0.3)	97.7
Tautomer 2	6.2 (0.3)	5.4 (0.4)	2.3
Tautomer 3	2.7 (0.5)	0.0	0.0

### 3.3.8 Methylated Nucleobase Results

Once again, the outlined methodology was used to separate and classify tautomers of the methylated nucleobases of 9-methylguanine (MeG), 9-methyladenine (MeA), 1-methylthymine (MeT), and 1-methyluracil (MeU). The lowest energy protonated structures corresponded exactly to the energetic ordering found in their complementary unmethylated version with the exception of the sugar attachment site nitrogen as it could no longer be protonated due to the presence of the methyl group. The dispersion plots for (MeG + H)<sup>+</sup> along with the ionogram at SV = 3400 V is shown in Figure 3.8 while the remaining protonated methylated nucleobases information is

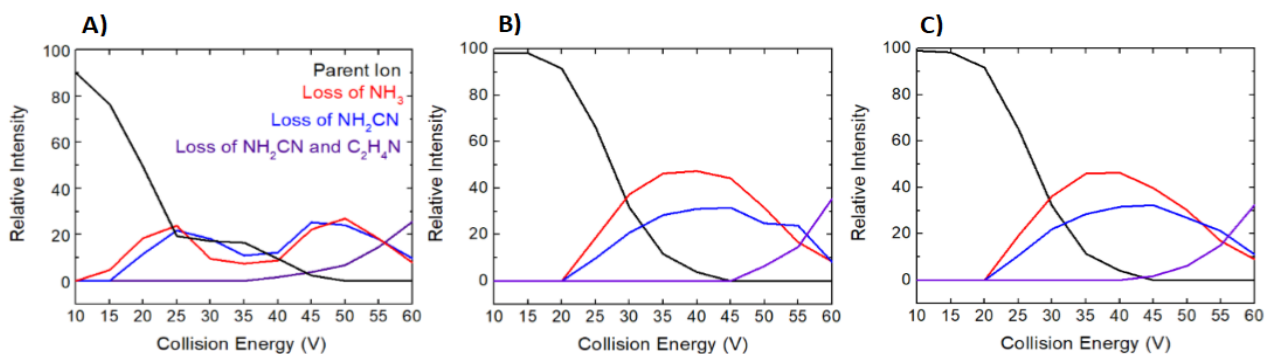
found in the APPENDIX A. The lowest in energy structures, analogous to its unmethylated counterpart, with relative energies are shown in Figure 3.8E. Breakdown curves, HDX profiles, and DP scans were performed once again to verify the identities of unique tautomers and are shown for  $(\text{MeG} + \text{H})^+$  in Figure 3.9 and Figure 3.10. As expected, there is slight difference in the fragmentation patterns due to the presence of the extra methyl group changing the nominal mass of each fragmented ion, but the overall trends discussed above remain consistent.



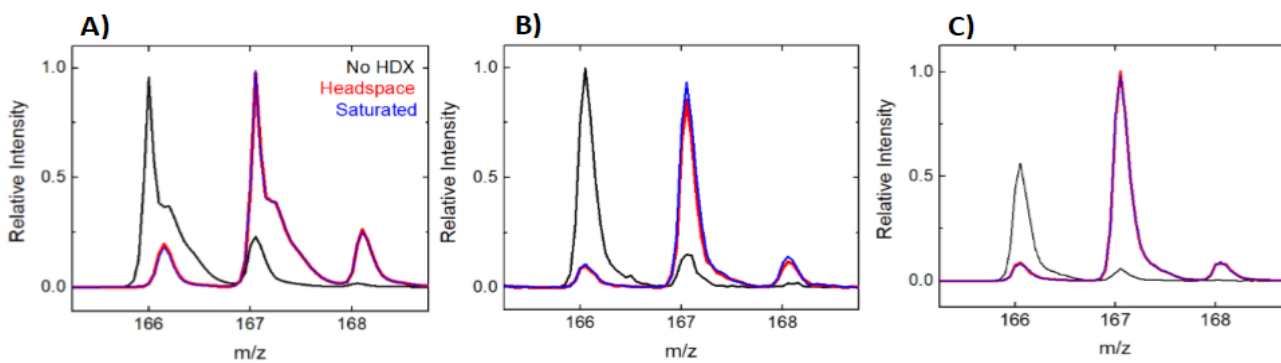
**Figure 3.8 Dispersion Plots for  $(\text{MeG} + \text{H})^+$**  in (A) a pure  $\text{N}_2$  environment, and a nitrogen environment seeded with 1.5 % (v/v) (B) MeOH vapor and (C) IPA vapor. (D) The ionogram of the  $m/z = 165$  peak in a pure nitrogen environment at an SV = 3400 V. (E) The three lowest energy calculated tautomers of  $(\text{MeG} + \text{H})^+$  and their relative standard Gibbs' free energies (in  $\text{kJ mol}^{-1}$ ) using CCSD(T)/6-311++G(d,p)

In the case of  $(\text{MeG} + \text{H})^+$ , three distinct peaks are visible in an  $\text{N}_2$  as well as  $\text{N}_2$  seeded with 1.5 % (v/v) MeOH vapor environment. In dry nitrogen, two of the suspected tautomers displayed Type C behaviour while one species displayed Type B behaviour. The chemically

modified environments yielded Type A and weak Type B behaviours. Evidently, the methylated nucleobases cluster with solvent in a similar manner to their unmethylated counterparts to produce similar clustering trends. Based on the energetics calculations, it can be assumed that the N-protonated global minimum structure is assigned to ionogram peak I, while tautomers 2 and 3 (O-protonated species) are assigned to peaks III and II, respectively. The peak identities were verified in the same manner as the bare protonated nucleobases.



**Figure 3.9 Breakdown Curves for  $(\text{MeG} + \text{H})^+$  in a pure  $\text{N}_2$  environment** at  $\text{SV} = 3400 \text{ V}$  at a CV of (A)  $-0.5 \text{ V}$ , (B)  $5.1 \text{ V}$ , and (C)  $10.6 \text{ V}$ . Collision energy was ramped from  $10 \text{ V}$  to  $60 \text{ V}$  in  $0.25$  increments.



**Figure 3.10 HDX Curves for  $(\text{MeG} + \text{H})^+$**  at an  $\text{SV} = 3400 \text{ V}$  and CV gated at (A)  $-0.4 \text{ V}$ , (B)  $5.1 \text{ V}$ , and (C)  $10.6 \text{ V}$ . The blue trace was saturated with  $\text{D}_2\text{O}$  while the red trace involved sampling the headspace of the  $\text{D}_2\text{O}$  reagent.

CID spectra of the studied methylated nucleobases are not present in literature, however it can be expected that similar patterns to unmethylated nucleobases will be observed. In the case of  $(\text{MeG} + \text{H})^+$ , the appearance of suspected product ions are once again seen with  $\text{NH}_3$  loss,  $\text{NH}_2\text{CN}$  loss and subsequent  $\text{C}_2\text{H}_4\text{N}$  loss dominating the spectra. Similar to  $(\text{G} + \text{H})^+$ , the three peaks assigned to tautomers of  $(\text{MeG} + \text{H})^+$  all show the same product formation, however the onset of fragmentation as a function of collision energy is what allows for differentiation with peak I showing a significantly different onset to peaks II and III. The N-protonated global minimum assigned tautomer displays an earlier loss of  $\text{NH}_3$  and  $\text{NH}_2\text{CN}$  at a CE 5 V lower than observed in the two O-protonated tautomers, again comparable to the breakdown curves of  $(\text{G} + \text{H})^+$ .

With the methylation on N9,  $(\text{MeG} + \text{H})^+$  has four exchangeable protons for HDX, however, exchange is only seen with one or two hydrogens, as was the case in  $(\text{G} + \text{H})^+$ . Exchange rates are again very similar and it is difficult to be used in this case for differentiation, so DP scans were done to aid in tautomer assignment. The percent populations of each suspected tautomer of  $(\text{MeG} + \text{H})^+$  at increments of DP = 0 V, 150 V, and 300 V are displayed in Table 3.2. Unlike in the case of  $(\text{G} + \text{H})^+$ , no weak features are observed at low DP values so high energy tautomers are not visible. At DP values greater than 150 V, all three peaks decrease in intensity until they are completely depleted at DP = 300 V indicating the presence of unique tautomers and no higher order clusters dissociating to produce the parent ion, which was observed in the bare nucleobases. Therefore, we can confidently assign peak I to the N7-protonated species, and II and III to the two different orientations of O-protonated  $(\text{MeG} + \text{H})^+$ . The global minimum structure was found to be  $28.61 \text{ kJ mol}^{-1}$  lower in energy than tautomer 2



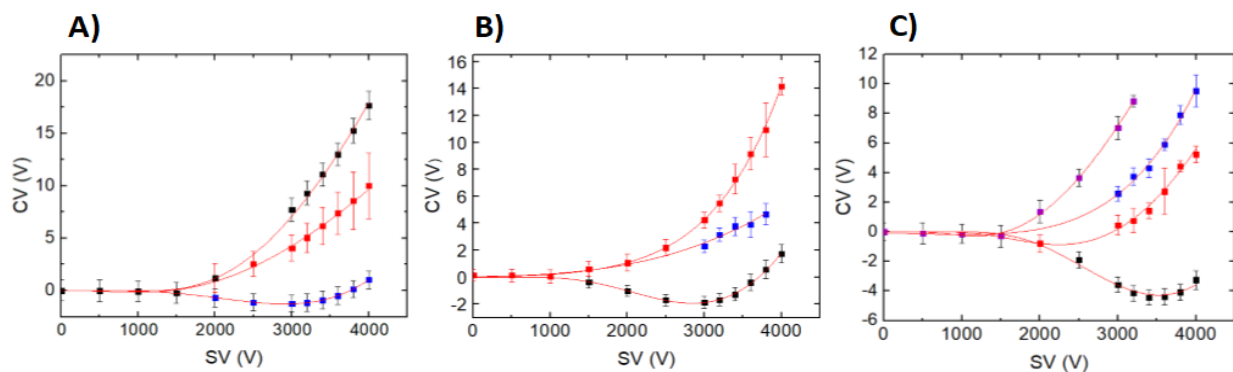
which explains why it dominates the population distribution in the gas phase. Tautomer 3, at an energy 63.2 kJ mol<sup>-1</sup> higher than the determined global minimum was observed in the gas phase despite its high energy. Kinetic trapping during the ESI process may explain its emergence within the DMS despite its small contribution to the overall ion population.

**Table 3.2 Population Distributions of Each Methylated Nucleobase Tautomer.** The percentage of the ensemble population of the lowest energy tautomer forms of each protonated methylated nucleobase at DP = 0V, DP = 150 V, and theoretical calculated energy values. Calculations were carried out at the CCSD(T)//B3LYP level of theory employing a 6-311++G(d,p) basis set.

Species	Percentage of Population		Calculated Energy (kJ mol <sup>-1</sup> )
	DP = 0 V	DP = 150 V	
<u>(MeG + H)<sup>+</sup></u>			
Global Minimum	89.75	58.84	0
Tautomer 2	0	16.26	28.6
Tautomer 3	4.21	24.91	63.2
<u>(MeA + H)<sup>+</sup></u>			
Global Minimum	96.49	81.29	0
Tautomer 2	0	14.13	7.7
Tautomer 3	0	4.174	23.6
<u>(MeT + H)<sup>+</sup></u>			
Global Minimum	97.33	85.73	0
Tautomer 2	0	10.40	13.1
Tautomer 3	0	3.87	25.1
<u>(MeU + H)<sup>+</sup></u>			
Global Minimum	96.49	83.04	0
Tautomer 2	0	8.26	6.9
Tautomer 3	0	5.14	20.1

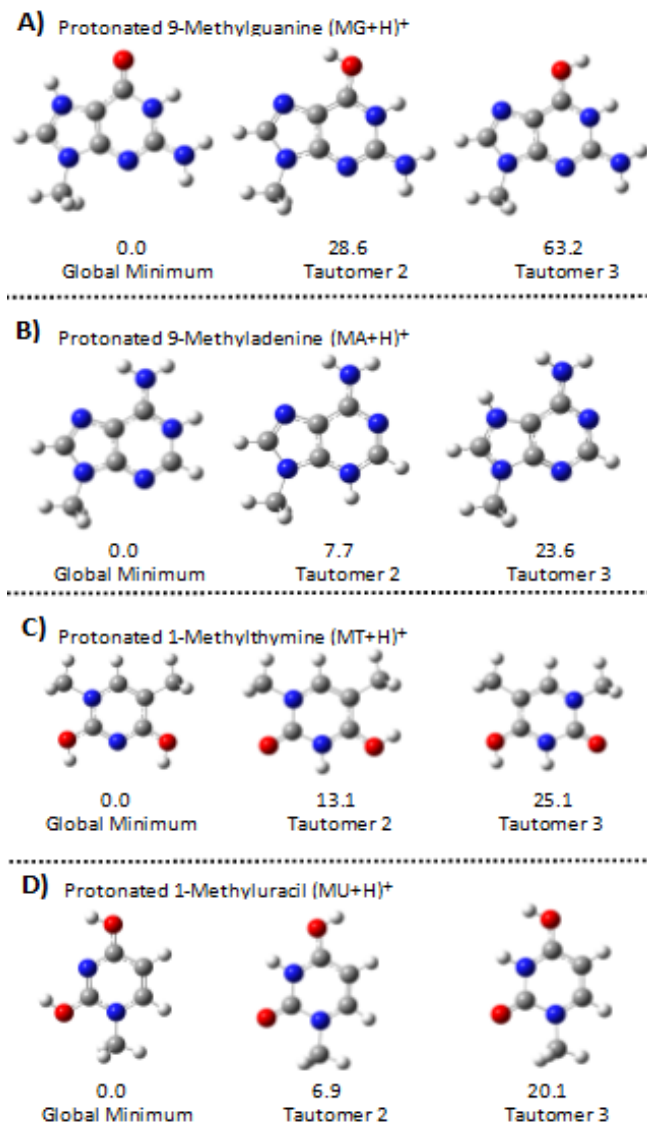
The same analysis was carried out on the protonated versions of 9-methyladenine, 1-methylthymine, and 1-methyluracil. Their dispersion plots in dry nitrogen are displayed in Figure 3.11 while the modified dispersion plots as well as breakdown curves and HDX data can be found in Appendix A. Dry N<sub>2</sub> without any chemical modifier was the ideal environment for separation in the case of each protonated methylated nucleobase. Three distinct tautomers are

visible for  $(\text{MeA} + \text{H})^+$  and  $(\text{MeT} + \text{H})^+$  while four tautomers (as dictated by separate curves on the dispersion plot) are visible for  $(\text{MeU} + \text{H})^+$ . Once again, all signals depleted upon increasing DP values, ensuring that each signal was in fact a unique tautomer and not a higher order cluster and the mass spectra are shown in Appendix B. The three lowest in energy calculated tautomers for each species are shown in Figure 3.12. Based on the relative energies it made sense why one tautomer (the global minimum structure) dominated the population distribution.



**Figure 3.11 Dispersion Plots of Methylated Nucleobases in Dry Nitrogen.** (A)  $(\text{MeA} + \text{H})^+$  at an  $m/z = 150$ , (B)  $(\text{MeT} + \text{H})^+$  at an  $m/z = 141$ , and (C)  $(\text{MeU} + \text{H})^+$  at an  $m/z = 127$ .

Interesting to note, and not fully observed in the analysis of bare nucleobases, was depletion of specific protonated methylated nucleobase signal intensity of  $(\text{MeU} + \text{H})^+$  and  $(\text{MeT} + \text{H})^+$  at increasing SV in the presence of polar protic modifiers such as MeOH or IPA. This is likely due to proton scavenging from the methylated nucleobase to the solvent if the solvent has a higher gas phase basicity than the nucleobase ions. Calculations will be done to determine these values and support the gas phase basicity assumption.



**Figure 3.12 Lowest Energy Structures of the Methylated Nucleobases.** Shown are the three lowest in energy tautomers of (A) (MeG + H)<sup>+</sup>, (B) (MeA + H)<sup>+</sup>, (C) (MeT + H)<sup>+</sup>, and (D) (MeU + H)<sup>+</sup> calculated at the CCSD(T)/6-311++G(d,p) level of theory and thermochemically corrected with B3LYP/6-311++G(d,p).

### 3.4 Conclusions

DMS-MS experiments were used as a powerful tool in conjunction with quantum chemical calculations to separate and identify tautomers of different forms of protonated nucleobases (*i.e.*, bare and methylated). In all cases, a dominant energetically favourable

tautomer correlated to calculated protonation sites was observed experimentally along with smaller ionogram features corresponding to higher-energy metastable species as confirmed by CID, HDX, and DP ramping. For the studies of  $(C + H)^+$ ,  $(T + H)^+$ , and  $(U + H)^+$ , the presence of two or more protonated forms from DMS separation were also found in IRMPD studies and the same low energy structures were determined spectroscopically. This is an impactful discovery as it leads to the possibility of DMS being used as a pre-filter to separate tautomeric species prior to laser interrogation and subsequently the deconvolution of obtained spectra, as has been observed in studies on lipid and saccharide isomers.<sup>83,84</sup>

Increasing the complexity of the system through the introduction of methyl groups at the sugar binding site of nucleobases was also successful and followed the trends of bare protonated nucleobases. It was found that the solvent environment can have stronger interactions with the gas phase ions as previously experienced and lead to ion depletion through proton abstraction when polar protic solvents are used. Additionally, declustering has been shown to be a valuable tool in determining whether or not a peak corresponds to a unique tautomer or is part of a larger solvent-ion cluster. We have shown that modifying these parameters was able to change the gas phase relative ensemble population distributions of nucleobase tautomers and should be optimized on an experiment-to-experiment case to ensure accurate and efficient results.

## **4.0 Probing the Structure, Properties, and Dynamics of RNA in the Gas Phase Using the Varkud Satellite Ribozyme**

This chapter describes an investigation of the Varkud Satellite ribozyme active site loop VI (VSVI) as a model system of RNA. It is an extension from previous chapters and tests the outlined methodology and analysis methods with respect to increased size, complexity, and lack of knowledge surrounding the chosen system. The standard methods for quantum chemical calculations that proved useful for the small systems described in previous chapters were deemed ineffective when dealing with the large, electron rich ribozyme, so alternate methods needed to be explored. We show that key residue protonation alters the electrostatic environment of the ribozyme and modelling through semi-empirical and DFT-B methods indicates important differences in relative structural stabilities to support current theories of inactive and active conformations. Experimentally, DMS and HDX were used in an attempt to separate equilibrium conformations of VSVI and provide fundamental gas phase knowledge of an RNA system. We found that acetone was able to separate individual charge states of VSVI and, through HDX experiments, that the negatively charged phosphate groups are important to the exchange mechanism, which can be impeded through the formation of strong hydrogen-bonding networks with polar protic solvents. Modifying the chemical exchange environment for RNA in the gas phase with solvent modifiers in a negative polarity had not previously been investigated, so results are fundamental. The knowledge gained about the behaviour of RNA in the gas phase can be extended in the future to additional large, biological systems.

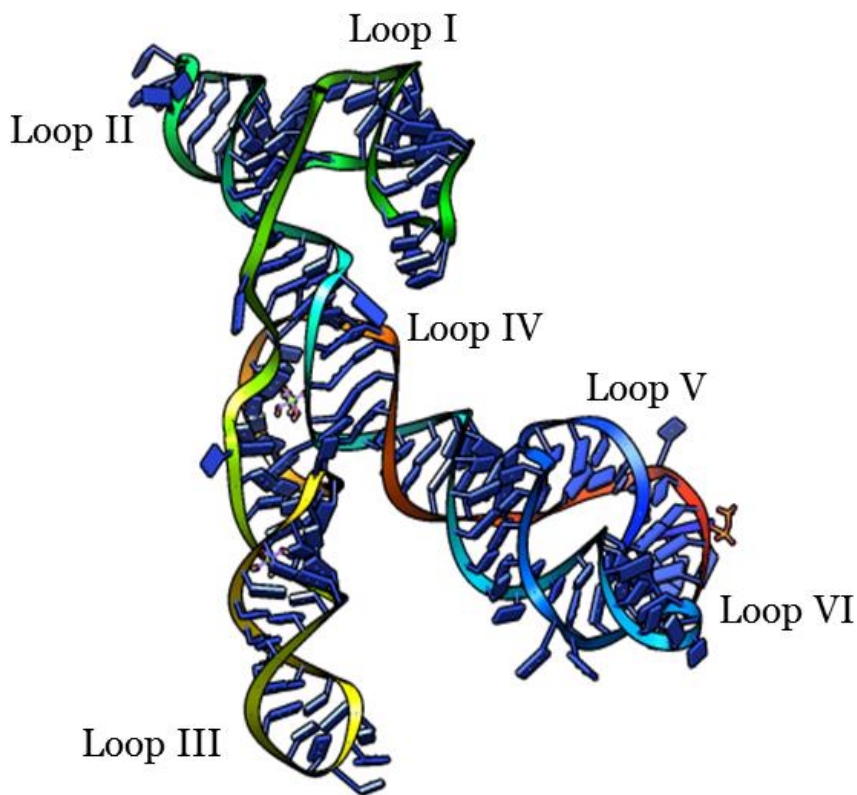
### **4.1 Introduction**

Over the past 40 years, nucleic acid research has become increasingly interesting following the discovery of catalytic RNA molecules known as ribozymes, which play important

roles in various cellular mechanisms. Biological catalysis was once thought to be conducive only to protein enzymes. With the discovery of ribozymes, a need for their structural characterization and how the overall three-dimensional structure facilitates catalytic reactions became a grand challenge for the RNA research community. The specific ribozyme studied in this research is the Varkud satellite (VS) ribozyme, which is an 881 nucleotide species isolated from *Neurospora* mitochondria.<sup>85</sup> Much work has been put into characterizing the overall structure of the VS ribozyme including NMR,<sup>86-88</sup> X-ray crystallography,<sup>89</sup> and mutagenesis and chemical studies<sup>90,91</sup> which provided primary and secondary structural information. Because of the very large size of the system and lack of knowledge present around ribozyme dynamics, the overall tertiary structure of the VS ribozyme was elucidated using a divide-and-conquer approach where each subdomain was structurally characterized individually and compared exhaustively to biochemical data to construct an in depth image of possible VS ribozyme conformations.<sup>88</sup> Recently, the X-ray crystallographic structure of the VS ribozyme was solved by Suslov *et al.*, in the presence of both Mg<sup>2+</sup> and K<sup>+</sup>.<sup>89</sup> Overall, the sequence of the VS ribozyme necessary for self-cleavage and ligation consists of 160 nucleotides divided into six major subdomains designated loops I through VI inclusively and is shown in Figure 4.1. The ribozyme is responsible for a self-cleavage and ligation reaction essential for replication of the RNA and generates products with 2',3'-cyclic phosphate and 5'-OH termini as found in other small ribozymes.<sup>88</sup>

Though all loops have been shown to play a role in stabilizing the structure and mechanism of the ribozyme, loop I undergoes a conformational change to become an activated substrate while loop VI acts as the active site to cleave loop I. Loops III, IV, and V aid in the formation of the active site and maintain a structure that allows the loop I substrate to possess the flexibility needed to bind the active site loop VI and initiate self-cleavage.<sup>88</sup> The focus of this

chapter involves the structure and dynamics within the possible conformations of the active site loop VI (VSVI). Two key characteristics within VSVI that allow for an accessible active site and efficient catalysis include a non-canonical G-A base pair and a sequence of nucleotides that generate a S-turn pattern.



**Figure 4.1 The overall structure of the Varkud Satellite Ribozyme.** The 160 nucleotides are divided into six major loops (I through VI, inclusively), with loop I acting as the substrate and loop VI acting as the active site for the overall self-cleavage reaction. Figure from PDB code 4R4P.

The presence of a protonated adenine ( $A_{730}$ ) adjacent to the G-A base pair has been postulated to have a favourable energetic effect in the active site, however the adenine in the G-A base pair

(A<sub>756</sub>) has been shown to play the role of a general acid in conjunction with a general base in loop I for cleavage to occur.<sup>92</sup> Additionally, NMR studies have found that the guanine in the G-A base pair (G<sub>729</sub>) has a ribose that undergoes sugar puckering in an exchange between C2'-endo and C3'-endo conformations on a millisecond timescale in solution.<sup>92</sup> Stacking interactions have been shown to be another important contribution to the overall folding process of the ribozyme as certain key nucleobases stack to either protrude from the active site cleft or stabilize the site of loop I binding. For example, C<sub>755</sub> and A<sub>756</sub> from VSVI protrude to the opposite side of the helix while the active site remodels and instead stacks between A<sub>621</sub> and A<sub>639</sub> from the substrate loop I to ensure successful binding.<sup>93</sup> It is also this protrusion and stacking interaction that allows the supposed catalytic nucleotides (G<sub>638</sub> and A<sub>756</sub>) to stack and position reactive groups close to the phosphate which is cleaved in the proposed mechanism.<sup>94</sup> It is advantageous to further probe these unique structural features to gain a better understanding of ribozyme dynamics.

The overall proposed mechanism of cleavage involves an initial proton transfer from a 2'-OH to the N1 position of G<sub>638</sub> before a nucleophilic attack on the scissile phosphate's 2' oxygen. A<sub>756</sub> then acts as a general acid as it stabilizes the transition state and protonates the 5'-O leaving group as the reaction progresses.<sup>94</sup>

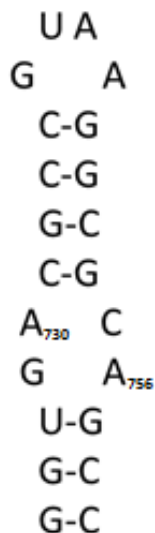
Computational and mass spectrometry-based investigations involving the VS ribozyme might provide additional information and support traditional NMR and X-ray crystallography-based techniques. Here we demonstrate that DFT calculations and DMS-HDX experiments are able to distinguish two structural motifs, which are hypothesized to be the inactive and active conformations of VSVI.



## 4.2 Methods

### 4.2.1 Experimental Methods

**RNA Synthesis.** The truncated form of VSVI, as synthesized and characterized by Flinders *et al.*, was used for all experiments. The truncated form consists of 22 nucleotides and has been shown to retain function which is important as it is biologically relevant and the smaller size is easier to work with both computationally and experimentally. Its structure is shown in Figure 4.2 and was synthesized following the RNA synthesis workflow outlined in **Section 2.2.5**. The VSVI DNA template was ordered from Integrated DNA Technologies and used for transcription in a reaction mixture shown in Table 4.1. The starting concentration of the template mixture was 20  $\mu\text{M}$  and a total volume of 500  $\mu\text{L}$ . A tris(hydroxymethyl)aminomethane (Tris) buffer was used to maintain an optimal reaction pH of  $\sim 7.8$ . Transcription was carried out in a 37  $^{\circ}\text{C}$  water bath for a total duration of 6 hours before 1X 10 M urea and 500 mM EDTA was added to terminate the reaction.



**Figure 4.2 Truncated form of VS loop VI.** The 22 nucleotide sequence and secondary structure of the truncated form of VSVI, which retains catalytic activity. Working with a smaller sequence that still maintains biochemical activity allows for a significantly more efficient analysis.

A 15% polyacrylamide gel was used for gel electrophoresis. Upon completion, the gel was visualized over a fluorescent plate and illuminated with UV light. The gel pieces containing the VSVI RNA were cut into small pieces using a razor and placed in a conical vial with equal volume of 300 mM NaCl to be spun overnight on a rotating bar. For ethanol precipitation, cold ethanol was added to the aqueous phase after NaCl extraction at 2.5 times the volume. The mixture was centrifuged for 45 minutes at 0 °C and 13200 x g, with g being the relative centrifugal force. The supernatant of the centrifuged sample was discarded and the pellet was resuspended in 1 mL of milli-Q water and placed in a -70 °C freezer for 1 hour.

**Table 4.1 Transcription Reaction Mixture for VSVI.** The table lists the components of the transcription reaction to synthesize the VSVI RNA to be purified along with the volume required.

<b>Name of Reagent</b>	<b>Volume Added (μL)</b>
Milli-Q Water	165
Transcription Buffer	50
Magnesium Chloride (MgCl <sub>2</sub> ) [45 mM]	25
Dithiothreitol (DTT) [100 mM]	25
Nucleotide Triphosphate Mastermix (equal parts CTP, ATP, GTP, UTP) [100 mM each]	150
VSVI Template [20 μM]	35
T7 RNA Polymerase	50

Lyophilization removed moisture from the sample before once again being suspended in 1 mL of milli-Q water. The sample was then subjected to anion exchange chromatography with a diethylaminoethane (DEAE) containing resin. A 1.5 M NaCl solution was used to elute the target RNA sample as it was detected spectrophotometrically based on its absorbance before undergoing a second lyophilisation step. As a final step to desalt the sample for mass spectrometric use, size exclusion chromatography was used and ammonium acetate solution

eluted the sample prior to storage in a  $-40\text{ }^{\circ}\text{C}$  freezer. Additional RNA with the VSVI sequence was purchased from Sigma Aldrich and 2.612 mg of RNA was delivered and added to the previously synthesized samples for a sufficient amount for all outlined experiments.

**Differential Mobility Spectrometry – Mass Spectrometry.** Due to the limited mass range of the QTRAP 5500 linear ion mass spectrometer at the University of Waterloo (accurate measurements only up to  $m/z$  1250) and a large  $m/z$  of the most stable charge states of VSVI ( $m/z > 1000$ ), all experiments were run using instrumentation located at SCIEX (Concord, ON). The instrument used was a modified TripleTOF® 5600 that had a DMS built in and an ExD cell (which was not active for these experiments). A time-of-flight (ToF) detection system was used which allowed for a much larger sampled  $m/z$  range than the Waterloo instrument. Contrary to other work outlined in this thesis, analysis of VSVI required a negative instrument polarity and an electrospray voltage of  $-4500\text{ V}$ . Low temperature DMS conditions ( $\sim 150\text{ }^{\circ}\text{C}$ ) were used with a 10 psi nebulizing gas pressure and ambient source pressure. The resolving gas flow was set to a value of 2.0 (arbitrary units) which has been shown to increase analyte separation in DMS experiments in comparison to the previous nucleobase experiments, which were done without the addition of a resolving gas flow. Separation voltages (SV) were stepped from  $SV = 0$  to  $4000\text{ V}$  in  $1000\text{ V}$  increments to  $SV = 2000\text{ V}$ , and  $500\text{ V}$  increments to  $2500\text{--}4000\text{ V}$ , while the compensation voltage (CV) was ramped from  $-20\text{ V}$  to  $+20\text{ V}$  in  $0.1$  step increments. Dispersion plots were generated in dry nitrogen, as well as with a series of solvent modifiers introduced at a  $1.5\%$  mole ratio with varying polarity and proticity. The tested modifiers included: methanol, isopropyl alcohol, acetonitrile, acetone, and ethyl acetate.

**Hydrogen Deuterium Exchange.** Following initial dispersion plot generation, the experiments were repeated, but with deuterium oxide ( $\text{D}_2\text{O}$ ) infused directly into the throttle gas

*via* a secondary syringe. D<sub>2</sub>O was t-infused into the throttle gas at a rate of 20 μL/min administered through an Oxford syringe pump. Mass spectra were obtained before and after deuterium exchange and relative rates of exchange were compared throughout the different charge states. The use of deuterated ammonia (Sigma-Aldrich) as the exchange reagent was also tested, but it resulted in rapid and complete exchange for all species present.

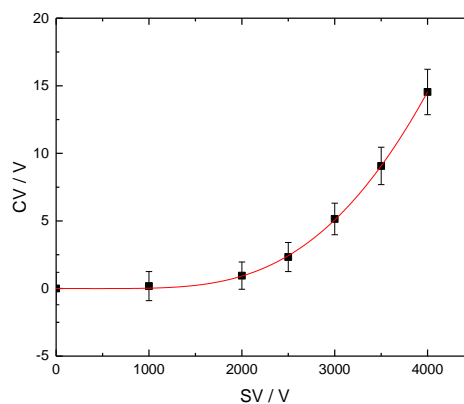
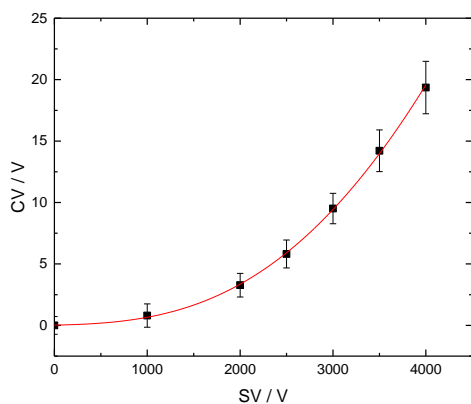
#### **4.2.2 Computational Methods**

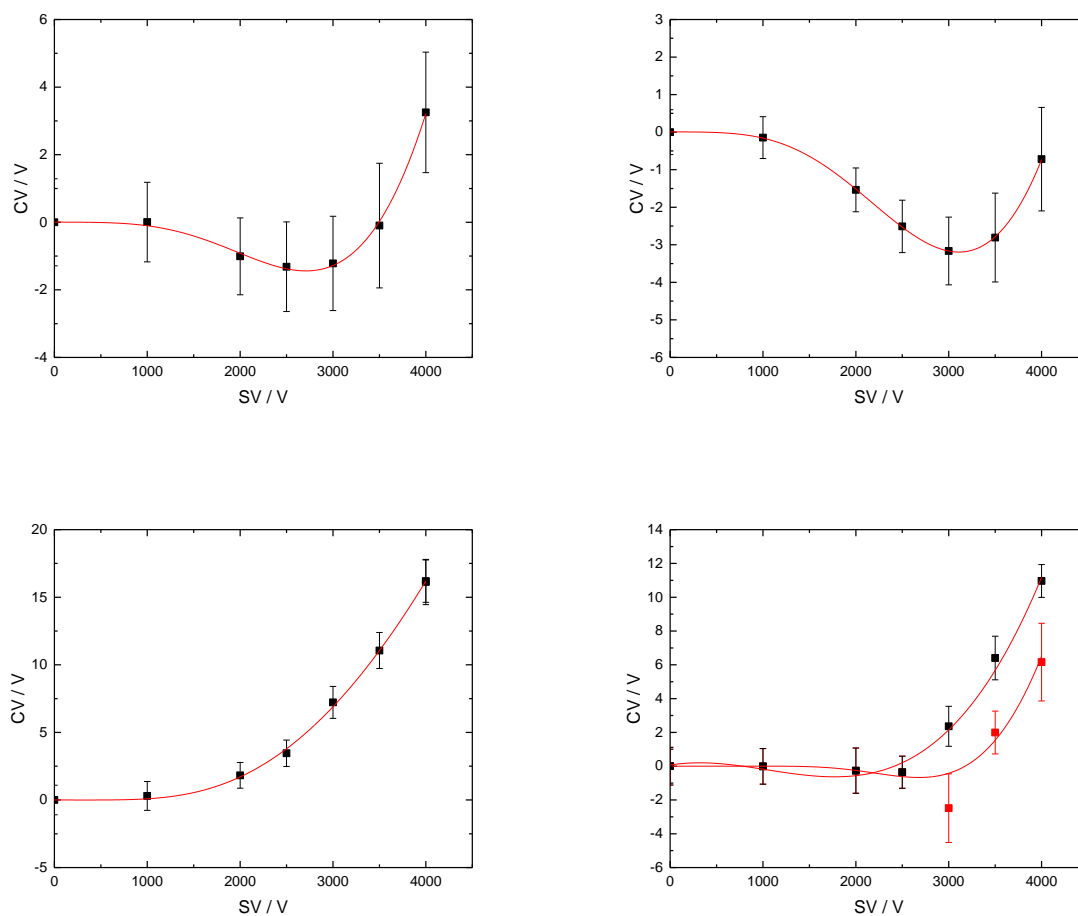
Due to the large electronic size of the system, care must be used when selecting a computational method to ensure meaningful results. Due to time constraints over the course of this research, high level calculations were not a practical avenue of assessment, so a variety of methods to study the VSVI structures and dynamics were used. The starting structure of VSVI was obtained from solution-phase NMR work located on the protein data bank (PDB code 1TJZ), and a major assumption was made that conformational changes observed in the gas phase reflect those in solution, or at least highlight the major intramolecular interactions necessary for function. Following initial molecular mechanics attempts at optimization of the entire structure, a subset of VSVI was used for further calculations, specifically a 6-mer of three adjacent base pairs that contains suspected key nucleobases necessary for catalytic activity (A<sup>+</sup><sub>730</sub>, G<sub>729</sub>, U<sub>728</sub>, C<sub>755</sub>, A<sub>756</sub>, and G<sub>757</sub>). Optimizations to study sugar puckering were performed with this 6-mer, initially using the semi-empirical PM7 method, then using DFT with the B3LYP functional and 6-311++G(d,p) basis set. Negative charges were assumed to be localized along the phosphate backbone, and the protonation states of each nucleobase were assumed to be consistent with their solution phase states at pH 7 with the exception of the known A<sup>+</sup><sub>730</sub>.

## 4.3 Results and Discussion

### 4.3.1 Dispersion Plots of VSVI

Following electrospray into the DMS-MS, a variety of charge states were observed. Relative charge state intensity varied depending on which solvent modifier was used (*vide infra*). Owing to its larger mass range, the ToF instrument facilitated measurement of the DMS behaviour of all charge states simultaneously. Figure 4.3 shows the dispersion plots obtained for the  $-6$  charge states in pure  $N_2$ ,  $N_2$  seeded with 1.5 % mole ratio methanol (MeOH),  $N_2$  seeded with 1.5 % mole ratio isopropyl alcohol (IPA),  $N_2$  seeded with 1.5 % mole ratio acetonitrile (ACN),  $N_2$  seeded with 1.5 % mole ratio ethyl acetate (EtAc), and  $N_2$  seeded with 1.5 % mole ratio acetone (ACE), while the dispersion plots for the other charge states are shown in Appendix B. Interestingly, separation was only observed with an ACE solvent vapor modifier and not with any of the other polar protic or aprotic solvents, an unexpected result not observed in prior DMS studies. In the positive mode DMS study of protonated nucleobases and their methylated counterparts, separation was easily visible in dry  $N_2$  as well as in MeOH, and to a lesser extent, IPA.<sup>5</sup>





**Figure 4.3 Dispersion Plots of VSVI** in (left to right) dry nitrogen, nitrogen seeded with 1.5 % MeOH vapor, nitrogen seeded with 1.5 % IPA vapor, nitrogen seeded with 1.5% ACN vapor, nitrogen seeded with 1.5% ethyl acetate vapor, and nitrogen seeded with 1.5% acetone.

In terms of clustering interactions, Type C behaviour was observed in dry nitrogen, methanol modifier, ethyl acetate modifier, and for one of the acetone modifier traces. Type B behaviour was observed in IPA modifier, acetonitrile modifier, and the second trace observed when employing an acetone modifier. The Type B behaviour observed in acetone appears slightly different than conventional Type B dispersion plot fits as rather than having a smooth and steady decline in CV before increasing at a well-defined SV at  $CV_{\min}$ , the fit sharply decreases between  $SV = 2500$  V and  $SV = 3000$  V before abruptly increasing after  $SV = 3000$  V.

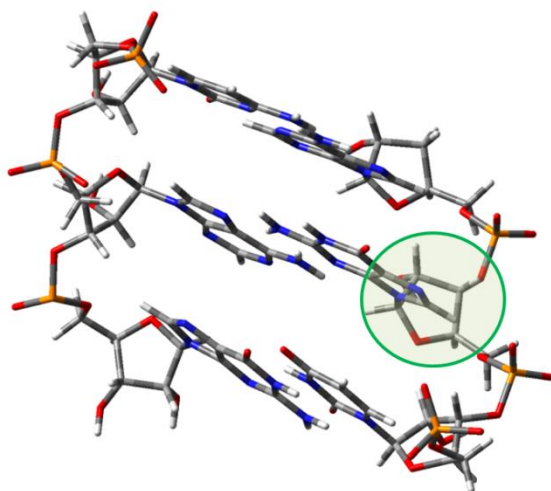
To confirm this behaviour, it would have been ideal to perform the experiment in smaller SV increments around this minimum point.

### 4.3.2 Sugar Puckering Dynamics

The three central base pairs in VSVI were optimized and energy values were obtained for multiple different starting conformations. NMR experiments have observed that the ribose on G<sub>729</sub> is in exchange between a C2'-endo and C3'-endo conformation, and is possibly a driving force for active site activation.<sup>92</sup> The computationally pre-optimized 6-mer was subject to a manual restructuring of the G<sub>729</sub> to a C2'-endo and a C3'-endo conformation and A<sub>730</sub> was studied in both the protonated and neutral states to yield multiple possible conformations with either a -5 or -6 overall charge. The energy values obtained indicate that the most favourable C2'-endo state (hypothesized active conformation) occurred when A<sub>730</sub> was protonated and ~5 kJ/mol lower in energy than when A<sub>730</sub> is neutral and the structures are shown in Figure 4.4 and Figure 4.5. This suggests that the protonation of A<sub>730</sub> imparts an electrostatic effect to drive the conformational change to an active C2'-endo conformation. It was also postulated that the energy barrier between the inactive and active forms of VSVI was low, so the additional energy available with A<sup>+</sup><sub>730</sub> and C2'-endo conformation on guanine may facilitate aid in activation. However, it is important to investigate two other important dynamics factors: (1) the accessibility of base pairs for the substrate stem loop I and (2) the potential catalytic activity of A<sub>756</sub>.

The conformational change on the ribose had an effect on the overall local structure of the 6-mer, with the AGU segment flipping in order to bring A<sup>+</sup><sub>730</sub> close to G<sub>757</sub>. It appears that the rearrangements with A<sup>+</sup><sub>730</sub> and a C2'-endo form leave space in the active site able to accommodate the substrate loop I due to protruding nucleobases. More calculations and

molecular dynamics simulations involving the substrate loop I in combination with active site loop VI should be done in the future to further confirm the proposed mechanism.



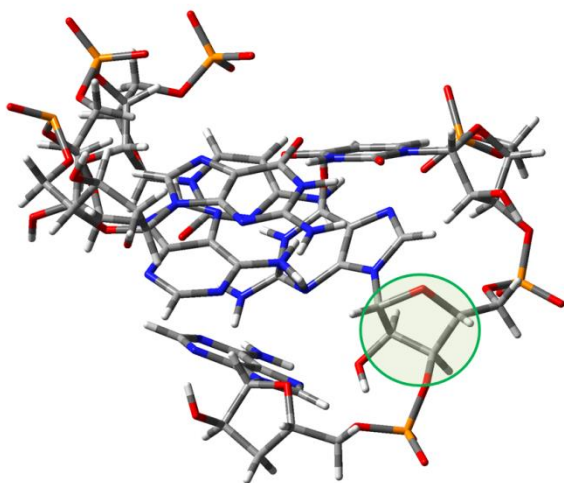
**Figure 4.4 The 6-mer of VSVI in its inactive C3'-endo conformation.** The main interactions between nucleobases when the circled ribose is in a C3'-endo conformation appear to be hydrogen bonding between paired bases. The structure maintained rigidity in the phosphate backbone and energetics were calculated using DFT at the B3LYP/6-311G++(d,p) basis set.

#### 4.3.3 Substrate Stem Loop I Accommodation in VSVI

Chemical intuition and nucleic acid knowledge suggests there to be multiple sources of instability in VSVI that would allow for conformational change upon substrate binding. For example, the supposed catalytic A<sub>756</sub> is located in a sheared base pair with the guanine that



contains the repuckered ribose. The potential electrostatic driving force of  $A^+_{730}$  is also in a sheared base pair with  $C_{755}$ .



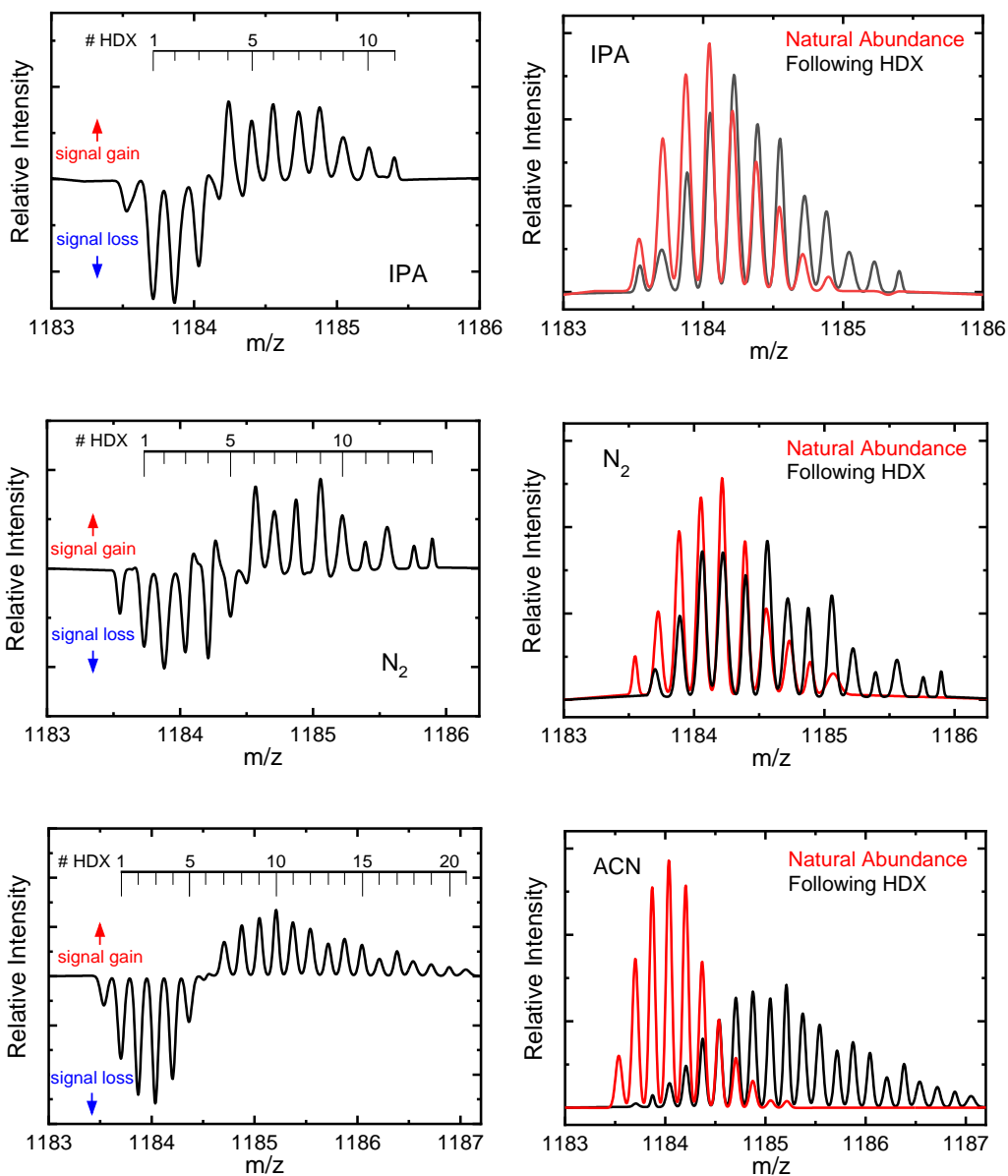
**Figure 4.5 The 6-mer of VSVI in its active C2'-endo conformation.** The main interactions between nucleobases when the circled ribose is in a C2'-endo conformation appear to be stacking interactions among distorted bases. As the sugar pucker the backbone reoriented to protrude specific nucleobases, and can be visualized to accommodate a substrate loop.

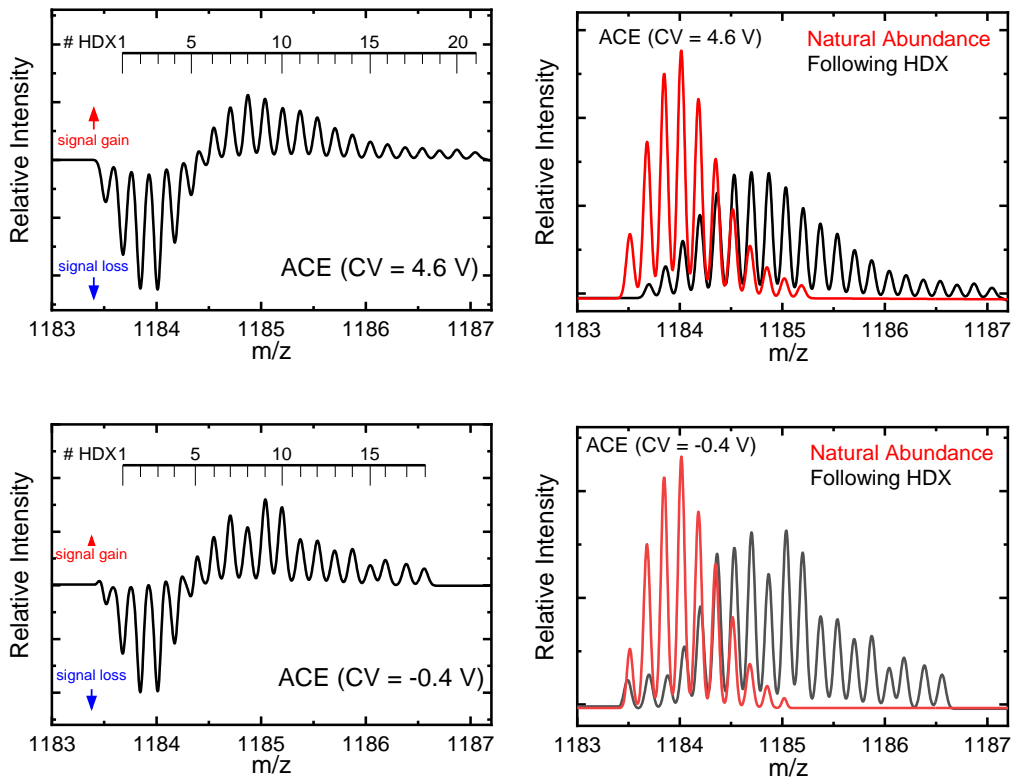
Both the sheared base pairs act to distort the helix and subsequently narrow the minor groove. In fact, all three base pairs in the studied 6-mer are non-canonical.  $U_{728}$  and  $G_{757}$  are involved in a mismatch (wobble), but is commonly found in A-form RNA helices along with the canonical C-G and A-U base pairing. It is possible for loops I and VI to interact in space following the above distortions, however future calculations should involve both species together to ensure accurate reaction locations and processes.

#### 4.3.4 HDX of VSVI

The mechanism for gas phase HDX of negatively charged nucleic acids has been studied previously.<sup>31</sup> A few mechanisms have been proposed, most of which involve the backbone phosphates interacting with the deuterated solvent.<sup>31</sup> Following HDX, the mass spectra for each charge state was recorded and the relative rates of deuterium uptake were observed, as illustrated in Figure 4.6. Only the  $-6$  charge state for each chemical modifier is depicted in Figure 4.6, while the HDX spectra for rest of the investigated charge states can be found in Appendix B. Figure 4.6 also depicts the difference spectrum for each modifier which is the resultant spectrum when the non-HDX data is subtracted from HDX data and indicates where signal is lost and gained during HDX relative to deuterium uptake. It appears that solvent modifiers had a major effect on deuterium exchange with acetone enhancing uptake and IPA impeding uptake. The overall order for gaseous environments and deuterium exchange from most exchange to least was: acetone  $\approx$  acetonitrile  $>$  nitrogen  $>$  isopropyl alcohol. Based on the chemical properties of the solvent modifiers, the polarity did not seem to play a role in exchange but the protic solvent modifiers displayed less exchange while aprotic solvent modifiers displayed the most exchange. This can be explained by envisioning the solvent environment creating a network around the analyte, potentially competing with any deuterated solvent based on gas phase basicity. One idea is that acetone displaces any clustered water or methanol that is interacting with the RNA following electrospray, and then is displaced with  $D_2O$  post-DMS. This would increase the exposure of RNA to  $D_2O$ , thus enhancing the likelihood of exchange to occur. From the perspective of IPA, a polar protic solvent, there are hydrogen atoms bonded to oxygen which have been linked to the formation of hydrogen bonding networks.<sup>95</sup> It is possible that an extended hydrogen bonding network of IPA molecules was formed surrounding VSVI and

competing with VSVI for exchangeable deuterium atoms and decreasing the overall rate of exchange. According to the proposed negative mode HDX mechanism involving nucleosides, it is possible that IPA interacts with the negatively charged phosphate that has been implicated in the mechanism which impedes the overall exchange rate.





**Figure 4.6 HDX profiles and difference spectra for the  $-6$  charge state of VSVI in varying modifiers.** The  $-6$  charge state was t-infused with  $D_2O$  and mass spectra were recorded in (top to bottom): isopropyl alcohol, dry nitrogen, acetonitrile, and acetone at CV = 4.6 V and  $-0.4$  V respectively. The difference spectrum is the resultant spectrum when the non-HDX spectrum is subtracted from the HDX spectrum to indicate signal loss or signal gain and number of deuterium atoms exchanged. Raw mass spectra were first fit with a Gaussian smooth before the Origin multi-peak fit analysis was performed and baseline was corrected. Peaks were extrapolated in  $0.005$   $m/z$  increments from  $m/z = 1183$  to  $m/z = 1187$  before they were normalized by sum to ensure consistency. The fitting procedure and low signal is the reason of the abrupt drop off at high  $m/z$  in several of the difference spectra.

The two separated species in the 1.5 % (v/v) acetone modified environment displayed subtle differences in hydrogen-deuterium exchange patterns. The differences are also present in the other charge states, as illustrated in Appendix B. However, it is quite possible that the subtle differences in exchange are due to spectral noise. Based on the postulated differences between active and inactive conformations, it would make sense that the HDX rates are quite similar. A slight increase in deuterium uptake may be related to the protrusion of nucleobases in the active conformation to accommodate substrate loop I, which exposes additional exchangeable

hydrogens. However this may be balanced by newly formed stacking interactions as seen in the C2'-endo conformation involving the protruded nucleobases, such as A<sub>730</sub><sup>+</sup> in an interaction with G<sub>757</sub> which makes exchange more difficult.

#### 4.4 Conclusions

The above method outlines a potential alternative to traditional NMR and X-ray crystallography techniques to probe the structure and dynamics of large RNA molecules using DMS and quantum chemical calculations. We found that the protonation of a key adenine residue in the active site loop of the Varkud Satellite ribozyme allowed for conformational changes between speculated active and inactive conformations due to electrostatics. Additionally, sugar repuckering observed on the timescale of NMR experiments was seen in guided computational calculations which allowed for minor groove narrowing and increased stacking interaction between nucleobases to accommodate the substrate loop.

DMS and HDX studies provided fundamental results on how large RNA molecules behave in the gas phase during these experiments in a negative polarity. Two species of VSVI, postulated to be active and inactive conformations, were able to be separated in DMS with the use of an acetone modifier. The rest of the DMS behaviours showed typical clustering behaviours with each modifier, however separation only in acetone was difficult to explain. The use of solvent modifier also had an effect on HDX results, with protic solvents such as IPA impeding the overall exchange rate while aprotic solvents such as acetonitrile or acetone appeared to enhance the exchange rate relative to an unmodified nitrogen environment. We rationalize this by noting that D<sub>2</sub>O was t-infused prior to Q1 so solvent modifier, such as IPA, could engage in a hydrogen bond network with the ribozyme and impede the exchange rate while simultaneously competing with phosphate groups for deuterium exchange.

VSVI has a unique dynamic and energetic profile and further study is necessary to fully understand its mechanism. Through the use of DFT calculations, MM simulations, and DMS-MS-HDX experiments, several postulated statements about VSVI function, including nucleobase rearrangement, from NMR experiments were confirmed. Based on the work here, it can be thought that the VS ribozyme is able to catalyze its reaction by key catalytic acidic and basic nucleobases, sugar puckering dynamics, and specific nucleobase protonations to allow for differed electrostatics as highlighted both experimentally and computationally.

## 5.0 Separation and Identification of Trimethoprim Transformation Products

This chapter is comprised of results also featured in a recently published manuscript in *Analytical Chemistry*. The study investigates and outlines a method to separate and identify transformation products of common environmental contaminants. We used differential mobility spectrometry in conjunction with MS-MS techniques and quantum chemical calculations to outline a new method to rapidly identify hydroxylated products of trimethoprim, an antibiotic of increasing concern due to its emergence in wastewaters. The products were generated using an electro-assisted Fenton reaction to mimic environmental transformation processes, and results found an interfering contaminant that impeded DMS results illustrating the importance of proper analytical clean-up procedures prior to analysis. The study was able to find that the methodology is able to accurately determine transformation products of small molecules at low, environmental concentrations (<100 ng/mL) in a time-efficient, high-throughput fashion.

### 5.1 Introduction

Organic micropollutants have been extensively studied in recent years due to their prevailing emergence in bodies of water and effects they may have on aquatic ecosystems.<sup>96-98</sup> Wastewaters, such as sewage treatment facilities, have been shown to contain organic micropollutants at levels conducive to biological activity which demonstrates the importance of determining how the compounds can be treated to minimize environmental harm.<sup>99</sup> A promising method to transform micropollutants is to use advanced oxidation techniques to produce H<sub>2</sub>O, CO<sub>2</sub>, and NH<sub>3</sub>; however some species are unable to be completely mineralized and undergo transformations to products with unknown structures and properties.<sup>11,100,101</sup> The determination of these transformation products of organic micropollutants (TPOMs) is essential to evaluate the

effectiveness of the oxidation process and its potential success in minimizing negative effects on the environment. However, conventional techniques of structural determination such as nuclear magnetic resonance, infrared spectroscopy, and X-ray diffraction are unreliable due to the very low environmental concentration of the TPOMs and the complex matrices in which the contaminants exist. Consequently, mass spectrometry-based methods have taken over as the primary technique for TPOM identification, which has led to the implementation of a classification scheme to evaluate the applicability and accuracy a mass spectrometry-based method has in characterizing small molecule structures.<sup>102</sup> Within this system the lowest level is level 5, which corresponds to accurate mass measurements. To improve the level of the method, additional information from analysis must be obtained such as obtaining a match from a spectral library or other diagnostic evidence to reach level 2. The use of a reference standard is required to reach the highest level of certainty, level 1. Current high-resolution tandem mass spectrometry studies have limitations when spectral libraries of the TPOM are absent and multiple positional isomers are possible, which can potentially conserve or modify biological activity of the compound.<sup>103</sup> In order to mitigate this problem, additional methods of providing detailed TPOM structural information are required and ion mobility spectrometry (IMS) has been proposed as a potential solution.<sup>104,105</sup> IMS employs a constant electric field and a drift gas opposing ion flow to separate gas phase ions based on differences in charge, size, shape, and mass. However, IMS has had difficulties in the classification of TPOM structure and demonstrates a need to improve experimental workflow to optimize efficiency and accuracy.

In this research, we employed the emerging technique of differential mobility spectrometry-mass spectrometry (DMS-MS) in an attempt to resolve current methodology limitations and provide accurate structural information quickly and efficiently. As explained in



the experimental method section, DMS differs from IMS by operating under atmospheric conditions as opposed to the low pressure environment in IMS. Additionally, DMS applies a high-frequency asymmetric waveform that is absent in IMS and these differences allow for DMS to have higher sensitivity and better selectivity, which is advantageous for environmental sampling.<sup>24,106</sup> There have been several studies that have highlighted the relevance and potential of DMS as a method for characterizing the structure and properties of biologically relevant species, including the ability to detect and separate isomers<sup>107</sup>, conformers<sup>108</sup>, and tautomers<sup>5</sup> that are often absent from analysis using conventional analytical techniques. Understanding and identifying the relevant isomeric forms of TPOMs is essential to ensure the biological impact they pose and optimal removal from any contaminated environment. To enhance separation within the DMS, chemical modifiers such as methanol, isopropyl alcohol, and acetonitrile can be used to induce clustering effects and evaporation of ion-solvent clusters.<sup>24</sup> Some knowledge of the chemical structure of interest is advantageous to postulate on whether a polar protic or aprotic solvent modifier will result in optimal species separation.

Owing to the complex matrices in which TPOMs are naturally found, traditional LCMS-MS techniques often require the sample to be treated extensively to produce usable data. DMS has increasingly been used to analyze and separate analytes within complex samples<sup>76,109,110</sup>, including Porta *et al.* who utilized DMS-MS to separate and identify many drugs from human kidney and tissue cross sections.<sup>111</sup> They also saw a 3-6 fold improvement in analysis time over the traditional LCMS methods to demonstrate the optimization power of differential mobility spectrometry. Research has also been done to employ DMS-MS in the rapid characterization of small organic molecules in environmental sources such as ground water. Noestheden *et al.* were able to separate and characterize naphthenic acid derivatives from ground water affected by oil

sands processing through DMS highlighting its potential for environmental analysis.<sup>112</sup> Here, we demonstrate the potential of DMS-MS as a rapid and efficient tool to separate and identify TPOMs. We focus on the transformation products of trimethoprim, an antibiotic that has widely been reported in environmental waters.<sup>113–115</sup>

## 5.2 Methods

### 5.2.1 Experimental Methods

**Reagents and Chemicals.** Trimethoprim was purchased from Santa-Cruz Biotechnology. Reaction reagents with a purity >98% were purchased from Sigma-Aldrich and included iron (II) sulfate ( $\text{FeSO}_4$ ) and sodium sulfate ( $\text{Na}_2\text{SO}_4$ ). Sulfuric acid ( $\text{H}_2\text{SO}_4$ ) and 0.1 M sodium hydroxide ( $\text{NaOH}$ ) were purchased from Fisher Scientific. Solvents and reagents used in LC-MS and solid-phase extraction procedures were purchased from Fisher Scientific and included acetonitrile (Optima LC-MS grade), methanol (Optima LC-MS grade), water (Optima LC-MS grade), formic acid (Optima LC-MS grade), and ethylenediaminetetraacetic acid disodium salt ( $\text{Na}_2\text{EDTA}$ , ACS grade).

**Preparation and Purification of Trimethoprim Transformation Product.** To generate the trimethoprim transformation product, the electro-assisted Fenton reaction<sup>116,117</sup> was conducted using a potentiostat/galvanostat EG&G model 273A (Princeton Applied Research) in a cell with two electrodes and one compartment in the galvanostatic mode. A platinum reference electrode was used with a glassy carbon working electrode. To agitate the reaction, a bubble generator was inserted into the cell solution which consisted of 50 mM of  $\text{Na}_2\text{SO}_4$  as electrolyte, 0.1 mM of  $\text{FeSO}_4$  as an iron source for the reaction, and 0.2 mM of trimethoprim.  $\text{H}_2\text{SO}_4$  was

used to adjust the solution pH to 2.7. The current was adjusted to 1 mA for a total reaction time of 60 min. The solution was transferred to an amber vial upon completion of the reaction.

Initial mass spectrometry analysis of the reaction mixture revealed a potential isobaric contaminant which led to the development of a solid-phase extraction (SPE) cleanup procedure. 0.1 M NaOH was added to the solution to adjust the pH to 9 and 200 mg/L Na<sub>2</sub>EDTA was added to complex residual iron suspected to contribute to the contaminant. The sample was introduced in Strata-X polymeric reversed phase SPE cartridges (200 mg, 6 mL, 33 μm) from Phenomenex previously conditioned with 5 mL ACN/MeOH 1:1 followed by H<sub>2</sub>O at pH 9. The cartridge was then rinsed with 2x5 mL of H<sub>2</sub>O to remove the hydrophilic impurity and 2x2.5 mL of ACN/MeOH 1:1 was used to elute it from the cartridge. Subsequent sample analysis did not show the presence of the interfering compound. Following characterization of the *m/z* 307.1412 species (*vide infra*), it was no longer necessary to conduct the SPE cleanup procedure since the *m/z* 307.1412 species could be identified using diagnostic multiple reaction monitoring (MRM) transitions. The wash was also analyzed using similar methods and showed a single peak of *m/z* 307 eluting at 0.33 min. The analysis of the eluate showed the presence of three peaks of *m/z* 307 eluting between 4 and 10 min with more than 5 % relative abundance.

**Differential Mobility Spectrometry – Mass spectrometry.** The DMS experiments were carried out on a SelexION differential mobility spectrometer paired with a QTRAP 5500 linear ion trap mass spectrometer (SCIEX). Positive mode ESI conditions were used with a 5.5 kV electrospray voltage, 20 psi nebulizing gas pressure, and ambient source temperature (~33 °C). The DMS temperature was set to low (~150 °C) and N<sub>2</sub> curtain gas pressure was set at 20 psi. Separation voltages (SV) were varied from SV = 0 – 4000 V in 500 V increments, and at each SV the compensation voltage (CV) was ramped from –50 V to +20 V in 0.1 V step increments.

The resulting series of ionograms were used to generate dispersion plots that indicate how the observed ions behave in the dynamic DMS environment.<sup>25</sup> To vary the clustering and collision environment of the DMS cell, polar protic modifiers, including methanol or isopropyl alcohol (1.5 % mole ratio, Sigma-Aldrich), were introduced into the curtain gas. A 1.5 % mole ratio was used as it is the lowest setting available on the SCIEX software and has been shown to enhance separation through clustering interactions while limiting potential solvent-induced tautomerization effects that may be visible at higher mole ratios.<sup>5,71</sup> Electrospray solutions of TP306 were diluted to a concentration of 5 – 100 ng/mL in a 1:1 solution of milli-Q water and methanol with 0.1% formic acid to aid in protonation of analytes.

Collision-induced dissociation (CID or MS/MS) experiments utilized ~9 mTorr of nitrogen as a collision gas to induce fragmentation of selected species. The selected species corresponded to distinct ionogram peaks and thus SV and CV values were fixed to isolate these isomers (*vide infra*). Collision energy values were ramped from +5 V to +120 V in 0.25 V step increments.

### 5.2.2 Computational Methods

Based on chemical intuition, all possible hydroxylation sites of neutral trimethoprim (methylene linker, methoxy-containing ring, amine-containing ring, and N-oxides) were manually generated and all possible protonated forms of these species were optimized with density functional theory (DFT) at the B3LYP level of theory using a 6-311G++(d,p) basis set as implemented in the Gaussian 16 software package.<sup>118</sup> Upon completion of geometry optimizations, harmonic frequencies were calculated to estimate thermochemical corrections and ensure that each structure was a local minimum (*i.e.*, isomer) on the potential energy surface. To investigate the possibility of solvent-induced tautomerization effects, the lowest energy

structures were re-optimized in the presence of an isopropyl alcohol molecule placed at the known charge site (site of protonation). All resulting structures were sorted based on their standard Gibbs' energies to determine the structures that are energetically favourable and likely to exist in the gas phase, and thus be probed by DMS.

## 5.3 Results and Discussion

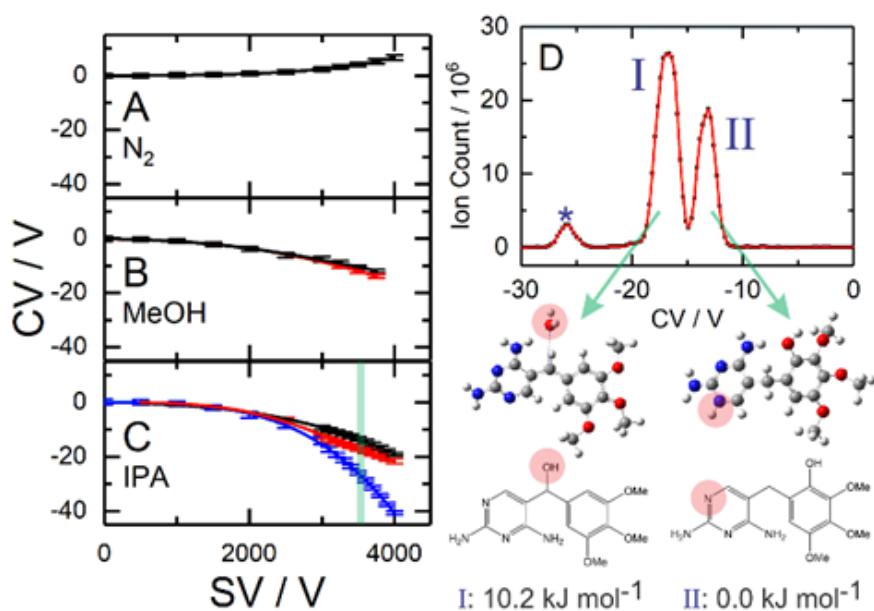
### 5.3.1 Fenton-Reaction Mixture LC-QqTOFMS Analysis

Liquid chromatography-quadrupole-time-of-flight mass spectrometry (LC-QqTOFMS; Shimadzu Nexera / Bruker Maxis) analysis of the trimethoprim solution following Fenton reaction showed the presence of a major peak eluting at 2.3 min, which corresponded to TP306. The compound, or unresolved mixture of compounds, appeared as a protonated species with  $m/z$  307.1412 [*i.e.*, (TP306 + H)<sup>+</sup>]. An additional compound with a similar nominal  $m/z$  of 306.8527 eluted at 0.3 min and was observed in preliminary DMS-MS experiments. This second compound interfered with DMS-MS analysis as both species were simultaneously transmitted by the first quadrupole mass filter. Therefore the SPE cleanup procedure discussed in the experimental methods section was required to remove the hydrophobic  $m/z$  306.8527 species for preliminary experiments. DMS-MS experiments of the initial untreated Fenton-reacted solution displayed peaks with irregular splitting patterns and unexpected CID products. A possible origin of the contaminant could be from leftover unreacted Fe<sup>+</sup> ions complexing with a compound present in the reaction mixture and was not investigated further.

### 5.3.2 Protonated TP306 Dispersion Plots

The resulted dispersion plots for protonated TP306 ( $m/z$  307.1) are shown in Figure 5.1A-C and represent resulting peaks in pure N<sub>2</sub>, N<sub>2</sub> seeded with methanol (MeOH), and N<sub>2</sub> seeded

with isopropyl alcohol (IPA), respectively. A pure nitrogen environment did not allow for isomeric separation so the introduction of polar protic solvents as modifiers was required. Ideal separation was achieved using the 1.5% IPA in N<sub>2</sub>, and the resulting ionogram at SV = 3500 V for (TP306+H)<sup>+</sup> is shown in Figure 5.1D. The observation of three distinct resolved peaks in an ionogram infers the presence of at least three isomeric structures; however caution must be exercised when drawing conclusions from separation in IPA and polar protic solvent environments due to the previously discussed possibility of solvent induced tautomerization which can act the skew relative isomer distributions.<sup>5,71</sup> An additional possibility that must be considered, as mentioned in the nucleobase section, is that a peak may be due to a solvent cluster dissociating and releasing a species with the same desired nominal *m/z* as the species of interest as opposed to a unique isomer. To further probe this possibility and the possibility of isobaric interference, additional characterization of the ions associated with each ionogram peak is necessary.

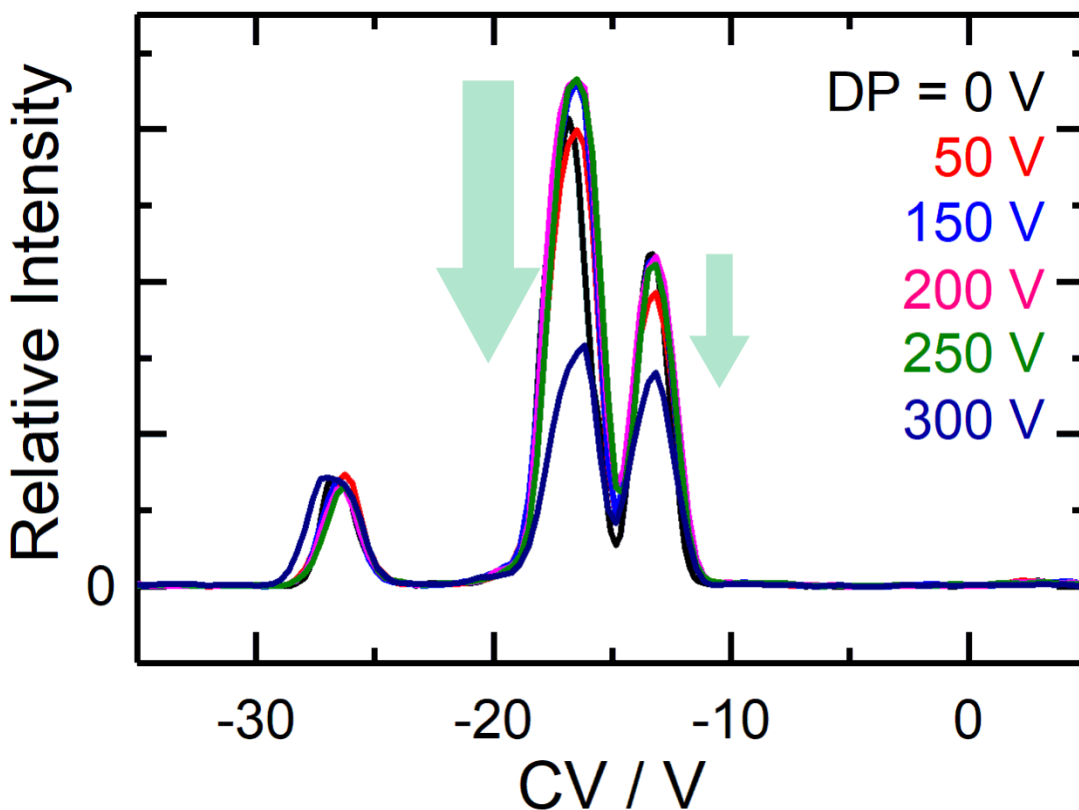


**Figure 5.1** DMS obtained plots and structures for (TP306 + H)<sup>+</sup>. Dispersion plots recorded for (TP306 + H)<sup>+</sup> in (A) a pure N<sub>2</sub> environment, and N<sub>2</sub> seeded with 1.5 % (v/v) (B) methanol vapor and (C) isopropanol vapor. (D) The ionogram observed at SV = 3500 V in the IPA-

modified environment (highlighted in green in XC). Inset: the two lowest energy isomers of (TP306 + H)<sup>+</sup> as calculated at the B3LYP/6-311G++(d,p) level of theory. Standard Gibbs' energies are reported.

### 5.3.3 Declustering Potential Study of Protonated TP306

Variation of declustering potential (DP) has been shown to infer whether or not an ionogram peak is a unique isomer or simply a fragment from a higher-order solvent-containing cluster.<sup>5,119</sup> The DP was ramped at the input of the mass spectrometer and in-source fragmentation was monitored in order to verify whether the observed ion signals were from (TP306 + H)<sup>+</sup>. Figure 5.2 displays the behaviour of each ionogram peak as a function of DP voltage at a fixed SV of 3500 V. The two most intense features in the ionogram depleted as DP increased to 300 V, as expected from bare ions of (TP306 + H)<sup>+</sup> due to onset fragmentation. In contrast, the weakest feature (indicated with an asterisk in Figure 5.1D) remained at a relatively constant intensity over the entire range of sampled DP values. This suggests that in-source fragmentation of the species producing this peak is balanced by production of *m/z* 307.1 from dissociation of larger clusters. To further characterize each peak in an attempt to determine structural molecular arrangement, collision induced dissociation (CID) was performed at an SV of 3500 V and a CV value which corresponded to when the isomer was transmitted at SV = 3500 V, and was guided by our computational study.



**Figure 5.2 Declustering potential ramping of suspected  $(\text{TP306} + \text{H})^+$  isomers.** The ionogram observed at  $\text{SV} = 3500 \text{ V}$  in the IPA-modified environment as a function of declustering potential (DP). The ion signals at  $\text{CV} = -13 \text{ V}$  and  $-16 \text{ V}$  deplete with increasing DP voltage due to in-source CID, thus indicating that these peaks are associated with  $(\text{TP306} + \text{H})^+$  and not other noncovalent adduct forms.

### 5.3.4 Quantum Chemical Calculation Results for $(\text{TP306} + \text{H})^+$

In total, 38 different stable isomers of  $(\text{TP306} + \text{H})^+$  were identified in the potential energy surface search of trimethoprim transformation products. The two lowest in energy structures are shown in the inset of Figure 5.1. The second lowest in energy structure (assigned to peak I) optimized to a carbocation/water association complex at the B3LYP/6-311G++(d,p) level of theory, which is not unexpected as it is common for the protonation of an alcohol to lead to water loss and produce an intense  $(\text{M} - \text{H}_2\text{O})^+$  fragment peak as the base peak in the mass spectra.<sup>120,121</sup> However, this is not the case in the experimental spectra recorded for  $(\text{TP306} + \text{H})^+$



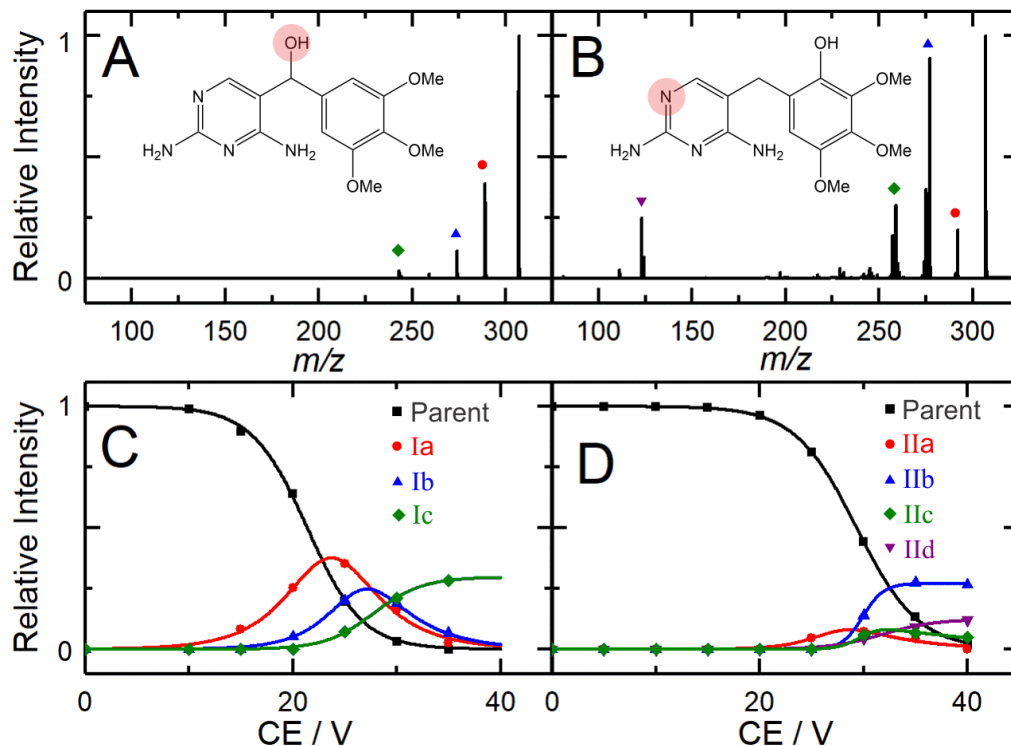
as an overall  $m/z$  307.1 peak is still observed as the parent. The optimized structures of the two lowest in energy species differed by only 10.2 kJ/mol which can both favourably exist within the gas phase and be detected in the mass spectrometer. Among the 38 studied isomers were N-oxides of trimethoprim that have been experimentally observed in metabolic pathways in pig liver microsomes.<sup>122</sup> The thermochemically corrected energy values following geometry optimizations of the N-oxides yielded high energy isomers with the lowest being 187.64 kJ/mol higher than the aromatic ring hydroxylation site which indicates N-oxides are not likely products in the electro assisted Fenton treatment of trimethoprim. Based on postulated Boltzmann distributions based on the energy calculations, we can pre-emptively assign peak I to the protonated methylene-hydroxylated product and peak II to the protonated hydroxylation product on the amine containing ring. The peak marked with an asterisk in Figure 5.1D remains difficult to determine, so to increase the confidence level in peak assignment, collision-induced dissociation was performed on each of the isolated species corresponding to a unique ionogram peak.

### 5.3.5 Collision Induced Dissociation Results for (TP306+H)<sup>+</sup> Species

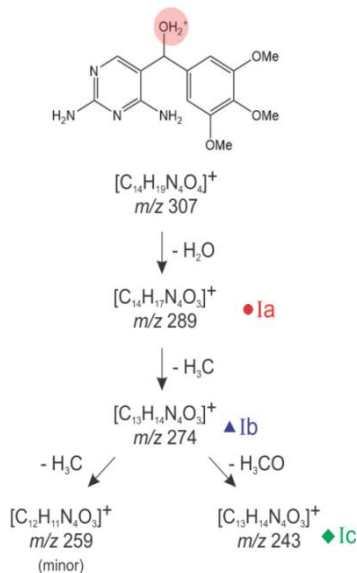
The assignment of isomers to ionogram peaks were reinforced through CID data. Through the monitoring of the fragmentation products of each species at a fixed SV and CV as a function of collision energy (CE), one can generate breakdown curves for the ion of interest to illustrate fragmentation patterns. The assignment of peak I to the second lowest energy isomer is supported by the observed fragmentation spectrum following CID (shown in Figure 5.3A), where a loss of  $m/z$  18 is observed corresponding to water loss from (TP306 + H)<sup>+</sup>. The breakdown curve for the second lowest energy isomer of (TP306 + H)<sup>+</sup> is shown in Figure 5.3C and three major product channels are observed for isomer 2 (Peak I). A proposed assignment of

fragmentation patterns is illustrated in Scheme I. Note that channel Ib is associated with demethylation of the exposed methyl ester groups bound to the resonance stabilized aromatic system, which has been observed in previous CID studies of trimethoprim.<sup>123</sup>

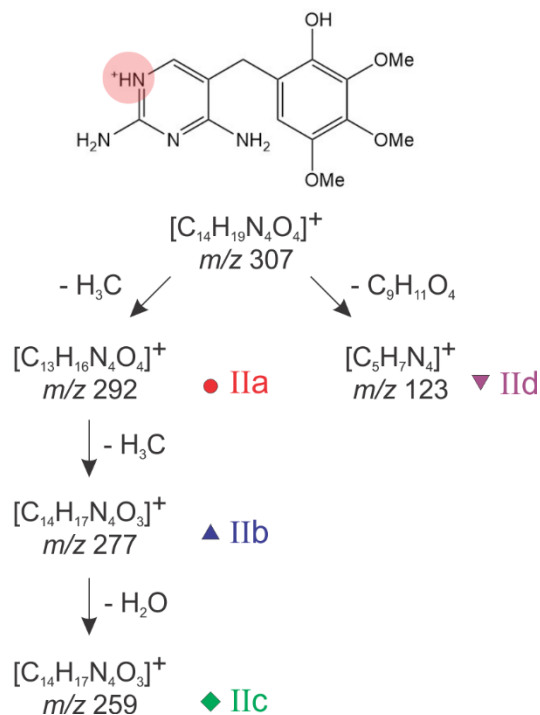
The global minimum isomer (Peak II) of (TP306 + H)<sup>+</sup> displays a distinctly different fragmentation spectrum and breakdown curve from the ones obtained from Peak I and is shown in Figure 5.3B and Figure 5.3D respectively. Four major breakdown products are observed, as opposed to the three observed from Peak I. The proposed assignment of each peak is shown in Scheme II. Successive demethylations of methyl ether groups are shown in channels IIa and IIb, while channel IIc corresponds to loss of H<sub>2</sub>O from the cationic fragment remaining following the double demethylation. The final channel, channel IId, corresponds to cleavage of the methylene linkage between the two aromatic ring systems. The assignment of peak II to the global minimum structure is further supported by the fact water loss would be a high energy process compared to the second lowest energy isomer, and that positive charge remains on the nitrogen-bearing fragment following cleavage at the bridging methylene group.



**Figure 5.3** Fragmentation spectra and breakdown curves of the lowest in energy (TP306 + H)<sup>+</sup> isomers. (A) Fragmentation spectrum observed for isomer 2 (peak I) following CID at CE = 20 V (est. lab frame). (B) Fragmentation spectrum observed for isomer 1 (peak II) following CID at CE = 35 V. Breakdown curves for (C) isomer 2 and (D) isomer 1. Product channels are described in the text and shown in Schemes I and II.



**Figure 5.4** The fragmentation channels observed via ionogram Peak I (see Figure 5.1D), attributed to isomer 2. The site of protonation is highlighted in red.



**Figure 5.5** The fragmentation channels observed via ionogram Peak II (see Figure 5.1D), attributed to isomer 1 (*i.e.*, global minimum). The site of protonation is highlighted in red.

The ambiguity of assigning the lowest intensity peak persisted through CID experiments. Again, based on high energy isomer structures from calculations, it was unlikely this peak belonged to protonated forms of the two lowest in energy hydroxylated sites, or alternative calculated structures such as N-oxides or the postulated hydroxylamine. The fragmentation pattern of the isolated species, as shown in Appendix C, did not correlate to any of these species and displayed consistent losses of  $m/z$  60, likely corresponding to IPA solvent clusters with the IPA evaporating upon heating. Based on the charge distribution, it is hypothesized that the contaminating species is a chloride containing ion core with IPA clusters and unrelated to any transformation products of trimethoprim.

### 5.3.6 DMS in Combination with Additional Analytical Techniques

While we have outlined a rapid, accurate, and resource-limited method for determining and identifying transformation products of trimethoprim, additional analytical techniques can be combined with DMS to enhance the desired results. If increased analytical separation is required, liquid chromatography (LC) can be combined with DMS to maximize the time and resource saving attributes of both orthogonal techniques as shown in the measurement of steroids by Ray *et al.*<sup>124</sup> Our method required an additional specific method and clean-up procedure in addition to the information provided by DMS in order to properly characterize the transformation products of trimethoprim, while LC alone would simply be able to separate the positional isomers. A major limiting factor of traditional one-dimensional LC is diminished peak capacities, which are increased in DMS leading to additional separation. For complex samples containing more than 50 compounds, only about 50% will be resolved using one-dimensional LC which highlights a major argument for the inclusion of DMS in these analyses.<sup>125</sup> In order to expand the methodology to the characterization of more complex samples and complex matrices, additional MS and pre-ESI analytical techniques can be advantageously used. Multiple reaction monitoring (MRM) analysis can be used to specifically monitor known fragments of already separated isomeric species when more than one set of transformation products is present in a sample. Additionally, techniques such as solid-phase microextraction (SPME) may be used prior to DMS to minimize matrix effects and maximize the intensities of the ions of interest.<sup>126</sup>

## 5.4 Conclusions

Transformation products of trimethoprim, an antibiotic commonly found in environmental waters, were generated using the electro-assisted Fenton reaction and analyzed to determine their chemical structure. Two major isomeric transformation products with masses 16 Da higher than trimethoprim were observed, separated, and isolated using differential mobility spectrometry and then probed with additional mass spectrometric analysis. IPA vapor at a 1.5% mole ratio was seeded into the N<sub>2</sub> gas flow in the DMS to enhance separation of the hydroxylated species. CID experiments were then carried out on each of the two suspected (TP306 + H)<sup>+</sup> isomers and revealed distinct fragmentation channels ideal for unique identification through multiple reaction monitoring. Quantum chemical calculations initially determined the structures of the transformation products and further supported our experimental observations. Following an electro-assisted Fenton-reaction, trimethoprim becomes hydroxylated on either the bridging methylene group or the methoxy-containing phenyl ring to generate a pair of regioisomers that are visible in the gas phase and have been seen in additional trimethoprim transformation studies. Assuming consistent ionization efficiencies, the two isomers are produced with relative populations of 64 % and 36 %, respectively. It is interesting to note that the electro-assisted Fenton treatment of trimethoprim did not yield any N-oxide isomers, which were hypothesized following *in vitro* treatment by pig liver microsomes and illustrates differences between biological processing and metabolic mechanisms.

We have demonstrated a promising methodology for identifying transformation products of small organic molecules in an efficient and timely manner. Based on signal-to-noise extrapolation, the limit of detection of the tested sample was 1.9 ng/mL, which indicates that environmental samples with a concentration < 1 ng/mL may need to be concentrated prior to

analysis.<sup>127</sup> DMS-MS/MS has been shown to be a high-throughput technique requiring relatively low solution concentrations which advantageously produce results complementary to other compound identifying techniques (*e.g.*, high resolution mass spectrometry and nuclear magnetic resonance). Furthermore, DMS analysis allows for a level of specificity ideal for complex samples and lead to a decreased LC run time or even a replacement of the traditional and more resource intensive LC workflow.<sup>128</sup> The methodology can also be used to derive quantitative measurements as demonstrated by work on steroids by Ray *et al.*<sup>124</sup> Moreover, given the capacity of the present technique to separate positional isomers, DMS-MS/MS expands the toolbox of environmental analysts to improve the identification of organic contaminants to higher levels of confidence.

## 6.0 Thesis Conclusions

The gas phase chemistry of biomolecules allows for a better understanding of the important energetic, dynamic, and structural properties that are present in their biologically relevant solution phase. The projects discussed in detail in this thesis explored the expansive usefulness of differential mobility spectrometry as a technique for probing and characterizing biomolecules and combined additional mass spectrometry techniques and quantum chemical calculations to reinforce the obtained results.

Chapter 3 discussed a fundamental study looking at the structures of tautomers of protonated nucleobases as well as their methylated counterparts. DMS was able to separate suspected tautomers of each protonated nucleobase which were then subject to identification *via* collision induced dissociation and hydrogen-deuterium exchange to supplement the calculated lowest in energy geometries. It was found that varying the declustering potential setting after

obtaining an ionogram is vital in validating whether a peak is a unique isomer or produced as an artefact following solvent evaporation yielding a species with the same suspected nominal  $m/z$ . Without investigating different declustering potentials, it is possible a peak is misidentified as an isomer when it can easily be a higher-order solvent cluster. Calculated structures agreed with the observed results and matched the fragmentation patterns seen in CID. Harmonic frequency calculations also matched precisely obtained IRMPD data to allow for confident assignment of gas phase tautomeric population distributions. The introduction of a methyl group to each protonated nucleobase at the site of typical sugar attachment mimicked the increased sterics as well as blocked off a previously available site of protonation to increase biological relevance. Following the same methodology as the bare protonated nucleobases, the methylated nucleobases showed the same trends in DMS separation and there was no change in protonation sites of low energy tautomers. It was found that care needs to be placed when using solvent modifiers in DMS, particularly isopropyl alcohol, as solvent induced tautomerization may artificially inflate relative isomer population distributions and skew their equilibrium away from how they exist in solution phase. Ultimately this project allowed for a fundamental understanding of the structures of protonated nucleobases and sets up a methodology and workflow to confidently assess larger biomolecular systems.

Chapter 4 continued with the theme of nucleic acids and vastly increased the electronic size of the system to study the dynamics of the VS ribozyme active site loop VI. The protonated form of a key nucleobase allowed for a sugar repuckering change to occur and a lower energetic barrier between inactive and active conformations *via* electrostatics. Additionally, the protrusion of key nucleotides agrees with current literature about the active site loop containing a residue that acts as a catalytic acid and able to accommodate the substrate loop I. DMS studies provided



a fundamental look at the behaviour and interaction of RNA and solvent modifiers within the negative mode in the gas phase. Acetone was found to be the only tested solvent modifier to allow for separation of VSVI conformers at multiple charge states. The study also reinforced the current proposed mechanism for gas phase-HDX of RNA in the negative mode which requires the phosphate for exchange. It was found that hydrogen-deuterium exchange rates occurred on the order of acetone > acetonitrile > nitrogen > methanol > isopropyl alcohol. Clearly the proticity of the solvent is an important factor and it can be envisioned that the highly protic isopropyl alcohol can establish hydrogen bond networks with the negatively charged phosphate groups and compete with deuterated solvent for exchange, thus impeding the overall rate.

The final chapter of the thesis expanded on the methodologies used to study nucleic acids and highlight the usefulness of DMS-MS techniques as a tool for probing and characterizing transformation products of antibiotics. The electro-assisted Fenton reaction was able to generate highly reactive hydroxyl radicals which in turn reacted with the common wastewater antibiotic, trimethoprim, to generate an unknown set of products to be characterized *via* DMS-MS. The ionograms originally displayed three distinct peaks, however subsequent analysis led to the discovery that one peak was a contaminant caused by a chloride-bound cluster with IPA, thus a clean-up procedure based on SPME was developed. DMS, collision induced dissociation, and quantum chemical calculations successfully identified two regioisomers of hydroxytrimethoprim, the hydroxylated transformation product of trimethoprim following Fenton reaction. The two isomers varied on site of hydroxylation, either on the bridging methylene group or methoxy-containing phenyl ring. The outlined protocol allows for a high throughput analysis requiring minimal concentrations, ideal for environmental samples, and a fast sampling time. It is hopeful that the technique will be extended in the future to analyze additional antibiotics as well as more

complex sample matrices, such as wastewater, to help remedy their chemical or biodegradation. DMS-MS in combination with quantum chemical calculations are powerful techniques in bridging the physicochemical gas phase properties of biomolecules and their solution phase behaviour. A better understanding of the gas phase properties of biomolecules can act as a powerful tool in high demand areas such as drug discovery, environmental cleanup, and energy sustainability, thus highlighting the importance to continue research in the area and closely associate physical chemistry and biochemistry.

## References

- (1) Chambers, V. S.; Marsico, G.; Boutell, J. M.; Di Antonio, M.; Smith, G. P.; Balasubramanian, S. *Nat. Biotechnol.* **2015**, *33*, 1–7.
- (2) Gidden, J.; Baker, E. S.; Ferzoco, A.; Bowers, M. T. *Int. J. Mass Spectrom.* **2005**, *240*, 183–193.
- (3) Warnke, S.; Von Helden, G.; Pagel, K. *J. Am. Chem. Soc.* **2013**, *135* (4), 1177–1180.
- (4) Wyttenbach, T.; Bowers, M. T. *J. Phys. Chem. B* **2011**, *115* (42), 12266–12275.
- (5) Anwar, A.; Psutka, J.; Walker, S. W. C.; Dieckmann, T.; Janizewski, J. S.; Larry Campbell, J.; Scott Hopkins, W. *Int. J. Mass Spectrom.* **2017**.
- (6) Serganov, A.; Patel, D. J. *Nat Rev Genet* **2007**, *8* (10), 776–790.
- (7) Llor, C.; Bjerrum, L. *Ther. Adv. Drug Saf.* **2014**, *5* (6), 229–241.
- (8) Liu, C.; Yves Le Blanc, J. C.; Schneider, B. B.; Shields, J.; Federico, J. J.; Zhang, H.; Stroh, J. G.; Kauffman, G. W.; Kung, D. W.; Ieritano, C.; Shepherdson, E.; Verbuyst, M.; Melo, L.; Hasan, M.; Naser, D.; Janiszewski, J. S.; Hopkins, W. S.; Campbell, J. L. *ACS Cent. Sci.* **2017**, *3* (2), 101–109.
- (9) Heck, A. J. R.; Jørgensen, T. J. D.; O’Sullivan, M.; von Raumer, M.; Derrick, P. J. *J. Am. Soc. Mass Spectrom.* **1998**, *9* (12), 1255–1266.
- (10) Lee, J. G.; Sagui, C.; Roland, C. *J. Phys. Chem. B* **2005**, *109* (43), 20588–20596.
- (11) Boxall, A. B. A.; Rudd, M. A.; Brooks, B. W.; Caldwell, D. J.; Choi, K.; Hickmann, S.; Innes, E.; Ostapyk, K.; Staveley, J. P.; Verslycke, T.; Ankley, G. T.; Beazley, K. F.; Belanger, S. E.; Berninger, J. P.; Carriquiriborde, P.; Coors, A.; DeLeo, P. C.; Dyer, S. D.; Ericson, J. F.; Gagné, F.; Giesy, J. P.; Guin, T.; Hallstrom, L.; Karlsson, M. V.; Larsson, D. G. J.; Lazorchak, J. M.; Mastrocco, F.; McLaughlin, A.; McMaster, M. E.; Meyerhoff, R. D.; Moore, R.; Parrott, J. L.; Snape, J. R.; Murray-Smith, R.; Servos, M. R.; Sibley, P. K.; Straub, J. O.; Szabo, N. D.; Topp, E.; Tetreault, G. R.; Trudeau, V. L.; Van Der Kraak, G. *Environ. Heal. Perspect.* **2012**, *120* (9), 1221–1229.
- (12) Katoh, M.; Kurata, T.; Watanabe, Y.; Sawaki, Y. *J. Am. Chem. Soc.* **1988**, *110* (22), 7472–7478.
- (13) Thomas, C.; Mackey, M. M.; Diaz, A. A.; Cox, D. P. *Redox Rep.* **2009**, *14* (3), 102–108.
- (14) Becke, A. D. *J. Chem. Phys.* **2014**, *140* (18), 0–18.
- (15) Katari, M.; Nicol, E.; Steinmetz, V.; van der Rest, G.; Carmichael, D.; Frison, G. *Chem. - A Eur. J.* **2017**, *23* (35), 8414–8423.

- (16) Miwa, K. *Phys. Rev. B* **2018**, *97* (7), 1–7.
- (17) Bagno, A.; Rastrelli, F.; Saielli, G. *J. Phys. Chem. A* **2003**, *107* (46), 9964–9973.
- (18) van Mourik, T.; Bühl, M.; Gaigeot, M.-P. *Philos. Trans. A. Math. Phys. Eng. Sci.* **2014**, *372* (2011), 20120488.
- (19) Kaschner, R.; Hohl, D. *J. Phys. Chem. A* **1998**, *102* (26), 5111–5116.
- (20) Mirzaei, M. *Zeitschrift fur Naturforsch. - Sect. A J. Phys. Sci.* **2010**, *65* (10), 844–848.
- (21) Kulik, H. J. *J. Chem. Phys.* **2015**, *142* (24).
- (22) Wales, D. J.; Doye, J. P. K. *J. Phys. Chem. A* **1997**, *101* (28), 5111–5116.
- (23) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117* (19), 5179–5197.
- (24) Campbell, Larry., Le Blanc, Y. *Bioanalysis* **2015**, *7* (7), 853–856.
- (25) Campbell, J. L.; Zhu, M.; Hopkins, W. S. *J. Am. Soc. Mass Spectrom.* **2014**, *25* (9), 1583–1591.
- (26) Schneider, B. B.; Nazarov, E. G.; Londry, F.; Vouros, P.; Covey, T. R. *Mass Spectrom. Rev.* **2016**, *35*, 687–737.
- (27) Johnson, A. R.; Carlson, E. E. *Anal. Chem.* **2015**, *87* (21), 10668–10678.
- (28) Kind, T.; Fiehn, O. *Bioanal. Rev.* **2010**, *2* (1), 23–60.
- (29) Skinner, J. J.; Lim, W. K.; Bédard, S.; Black, B. E.; Englander, S. W. *Protein Sci.* **2012**, *21* (7), 996–1005.
- (30) Keppel, T. R.; Howard, B. A.; Weis, D. D. *Biochemistry* **2011**, *50* (40), 8722–8732.
- (31) Mo, J.; Todd, G. C.; Hakansson, K. *Biopolymers* **2009**, *91* (4), 256–264.
- (32) Robinson, J. M.; Greig, M. J.; Griffey, R. H.; Mohan, V.; Laude, D. A. *Anal. Chem.* **1998**, *70* (17), 3566–3571.
- (33) Campbell, S.; Rodgers, M. T.; Marzluff, E. M.; Beauchamp, J. L. *J. Am. Chem. Soc.* **1995**, *117* (51), 12840–12854.
- (34) Rand, K. D.; Pringle, S. D.; Morris, M.; Brown, J. M. *Anal. Chem.* **2012**, *84* (4), 1931–1940.
- (35) Cryar, A.; Groves, K.; Quaglia, M. *J. Am. Soc. Mass Spectrom.* **2017**, *28* (6), 1192–1202.
- (36) Ebrahiem, E. E.; Al-Maghrabi, M. N.; Mobarki, A. R. *Arab. J. Chem.* **2017**, *10*, S1674–S1679.
- (37) Wink, D. A.; Nims, R. W.; Saavedra, J. E.; Utermahlen, W. E.; Ford, P. C. *Proc. Natl. Acad. Sci.* **1994**, *91* (14), 6604–6608.
- (38) Chang, C. Y.; Hsieh, Y. H.; Cheng, K. Y.; Hsieh, L. L.; Cheng, T. C.; Yao, K. S. *Water Sci. Technol.* **2008**, *58* (4), 873–879.
- (39) Cazenave, C.; Uhlenbeck, O. C. *Proc. Natl. Acad. Sci.* **1994**, *91* (15), 6972–

- 6976.
- (40) Sissi, C.; Palumbo, M. *Nucleic Acids Res.* **2009**, *37* (3), 702–711.
  - (41) Rio, D. C.; Ares, M.; Hannon, G. J.; Nilsen, T. W. *Cold Spring Harb. Protoc.* **2010**, *5* (6), 1–7.
  - (42) Petrov, A.; Wu, T.; Puglisi, E. V.; Puglisi, J. D. *RNA purification by preparative polyacrylamide gel electrophoresis*, 1st ed.; Elsevier Inc., 2013; Vol. 530.
  - (43) Piškur, J.; Rupprecht, A. *FEBS Lett.* **1995**, *375* (3), 174–178.
  - (44) Damsteegt, E. L.; McHugh, N.; Lokman, P. M. *Anal. Biochem.* **2016**, *511*, 92–96.
  - (45) Easton, L. E.; Shibata, Y.; Lukavsky, P. J. *RNA.* **2010**, 647–653.
  - (46) Fornaro, T.; Biczysko, M.; Monti, S.; Barone, V. *Phys. Chem. Chem. Phys.* **2014**, *16* (21), 10112–10128.
  - (47) Panigrahi, S.; Bhattacharya, A.; Banerjee, S.; Bhattacharyya, D. *J. Phys. Chem. C* **2012**, *116* (7), 4374–4379.
  - (48) Dong, C.; Yang, J.; Ning, H.; Li, C. *J. Mol. Struct. THEOCHEM* **2010**, *950*, 64–71.
  - (49) Baei, M. T.; Taghartapeh, M. R.; Lemeski, E. T.; Soltani, A. *Phys. B Condens. Matter* **2014**, *444*, 6–13.
  - (50) Kubota, M.; Kobayashi, T. *J. Electron Spectros. Relat. Phenomena* **1996**, *82*, 61–70.
  - (51) Sowers, L. C.; Fazakerley, G. V.; Eritja, R.; Kaplan, B. E.; Goodman, M. F. *Proc. Natl. Acad. Sci.* **1986**, *83* (15), 5434–5438.
  - (52) Buyl, F.; Smets, J.; Maes, G.; Adamowicz, L. *J. Phys. Chem.* **1995**, *99* (41), 14967–14979.
  - (53) Kwiatkowski, J. S.; Leszczynski, J. *J. Mol. Struct.* **1996**, *376* (1–3), 325–342.
  - (54) Russo, N.; Toscano, M.; Grand, A.; Jolibois, F. *J. Comput. Chem.* **1998**, *19* (9), 989–1000.
  - (55) Kryachko, E. S.; Nguyen, M. T.; Zeegers-Huyskens, T. *J. Phys. Chem. A* **2001**, *105* (8), 1288–1295.
  - (56) Chandra, A. K.; Michalska, D.; Wysokiński, R.; Zeegers-Huyskens, T. *J. Phys. Chem. A* **2004**, *108* (44), 9593–9600.
  - (57) Tho Nguyen, M.; Chandra, A. K.; Zeegers-Huyskens, T. *J. Chem. Soc. Faraday Trans.* **1998**, *94* (9), 1277–1280.
  - (58) Moustafa, H.; El-Taher, S.; Shibl, M. F.; Hilal, R. *Int. J. Quantum Chem.* **2002**, *87* (6), 378–388.
  - (59) Di Donna, L.; Napoli, A.; Sindona, G.; Athanassopoulos, C. *J. Am. Soc. Mass Spectrom.* **2004**, *15* (7), 1080–1086.
  - (60) Kurinovich, M. A.; Phillips, L. M.; Sharma, S.; Lee, J. K. *Chem. Commun. (Camb)*. **2002**, No. 20, 2354–2355.

- (61) Frank-Kamenetskii, M. D.; Mirkin, S. M. *Annu. Rev. Biochem.* **1995**, *64*, 65–95.
- (62) Colominas, C.; Luque, F. J.; Orozco, M. *J. Am. Chem. Soc.* **1996**, *118* (29), 6811–6821.
- (63) Podolyan, Y.; Gorb, L.; Leszczynski, J. *J. Phys. Chem. A* **2000**, *104* (31), 7346–7352.
- (64) Salpin, J. Y.; Guillaumont, S.; Tortajada, J.; MacAleese, L.; Lemaire, J.; Maitre, P. *ChemPhysChem* **2007**, *8* (15), 2235–2244.
- (65) Bakker, J. M.; Sinha, R. K.; Besson, T.; Brugnara, M.; Tosi, P.; Salpin, J. Y.; Maître, P. *J. Phys. Chem. A* **2008**, *112* (48), 12393–12400.
- (66) Singh, V.; Fedeles, B. I.; Essigmann, J. M. *Rna* **2015**, *21* (1), 1–13.
- (67) Lalli, P. M.; Iglesias, B. A.; Toma, H. E.; De Sa, G. F.; Daroda, R. J.; Silva Filho, J. C.; Szulejko, J. E.; Araki, K.; Eberlin, M. N. *J. Mass Spectrom.* **2012**, *47* (6), 712–719.
- (68) Karpas, Z.; Berant, Z.; Stimac, R. M. *Struct. Chem.* **1990**, *1*, 201–204.
- (69) Schröder, D.; Buděšínský, M.; Roithová, J. *J. Am. Chem. Soc.* **2012**, *134* (38), 15897–15905.
- (70) Xia, H.; Attygalle, A. B. *Anal. Chem.* **2016**, *88* (11), 6035–6043.
- (71) Campbell, J. L.; Yang, A. M. C.; Melo, L. R.; Hopkins, W. S. *J. Am. Soc. Mass Spectrom.* **2016**, *27* (7), 1277–1284.
- (72) Girke, C.; Daumann, M.; Niopek-Witz, S.; Mählmann, T. *Front. Plant Sci.* **2014**, *5* (September), 1–12.
- (73) Stojković, V.; Noda-Garcia, L.; Tawfik, D. S.; Fujimori, D. G. *Nucleic Acids Res.* **2016**, *44* (18), 8897–8907.
- (74) N’guyen Cong, H.; Bertaux, O.; Valencia, R.; Becue, T.; Fournier, T.; Biou, D.; Porquet, D. *J. Chromatogr.* **1994**, *661*, 193–204.
- (75) Chin, W.; Mons, M.; PiuZZi, F.; Tardivel, B.; Dimicoli, I.; Gorb, L.; Leszczynski, J. *J. Phys. Chem. A* **2004**, *108* (40), 8237–8243.
- (76) Schneider, B. B.; Covey, T. R.; Coy, S. L.; Krylov, E. V.; Nazarov, E. G. *Int. J. Mass Spectrom.* **2011**, *298*, 45–54.
- (77) Schneider, B. B.; Nazarov, E. G.; Covey, T. R. *Int. J. Ion Mobil. Spectrom.* **2012**, *15* (3), 141–150.
- (78) Schneider, B. B.; Covey, T. R.; Nazarov, E. G. *Int. J. Ion Mobil. Spectrom.* **2013**, *16* (3), 207–216.
- (79) Schneider, B. B.; Nazarov, E. G.; Londry, F.; Covey, T. R. *Int. J. Ion Mobil. Spectrom.* **2015**, *18* (3–4), 159–170.
- (80) Nelson, C. C.; McCloskey, J. A. *J. Am. Soc. Mass Spectrom.* **1994**, 339–349
- (81) Nelson, C. C.; McCloskey, J. A. *J. Am. Chem. Soc.* **1992**, *114* (10), 3661–3668.
- (82) Cheng, P.; Li, Y.; Li, S.; Zhang, M.; Zhou, Z. *Phys. Chem. Chem. Phys.*

- 2010**, *12* (18), 4667.
- (83) Lintonen, T. P. I.; Baker, P. R. S.; Suoniemi, M.; Ubhi, B. K.; Koistinen, K. M.; Duchoslav, E.; Campbell, J. L.; Ekroos, K. *Anal. Chem.* **2014**, *86* (19), 9662–9669.
- (84) Levin, D. S.; Vouros, P.; Miller, R. A.; Nazarov, E. G. *J. Am. Soc. Mass Spectrom.* **2007**, *18* (3), 502–511.
- (85) Saville, B. J.; Collins, R. a. *Proc. Natl. Acad. Sci. U. S. A.* **1991**, *88* (19), 8826–8830.
- (86) Michiels, P. J. A.; Schouten, C. H. J.; Hilbers, C. W.; Heus, H. A. *RNA* **2000**, *6* (12), 1821–1832.
- (87) Bonneau, E.; Girard, N.; Lemieux, S.; Legault, P. *Rna* **2015**, *21* (9), 1621–1632.
- (88) Dagenais, P.; Girard, N.; Bonneau, E.; Legault, P. *Wiley Interdiscip. Rev. RNA* **2017**, *8* (5), 1–18.
- (89) Suslov, N. B.; Dasgupta, S.; Huang, H.; Fuller, J. R.; Lilley, D. M. J.; Rice, P. A.; Piccirilli, J. A. **2016**, *11* (11), 840–846.
- (90) Collins, R. A. *Mol. Genet.* **2002**, *30*, 1122–1126.
- (91) Lilley, D. M. J. *Rna* **2004**, *10* (2), 151–158.
- (92) Flinders, J.; Dieckmann, T. *J. Mol. Biol.* **2004**, *341* (4), 935–949.
- (93) Andersen, A. A.; Collins, R. A. *Mol. Cell* **2000**, *5* (3), 469–478.
- (94) Wilson, T. J.; Li, N.-S.; Lu, J.; Frederiksen, J. K.; Piccirilli, J. A.; Lilley, D. M. J. *Proc. Natl. Acad. Sci.* **2010**, *107* (26), 11751–11756.
- (95) Kuznetsova, N.; Rau, D. C.; Parsegian, V. A.; Leikin, S. *Biophys. J.* **1997**, *72* (1), 353–362.
- (96) Monteiro, S.; Boxall, A. B. A. *Rev. Environ. Contam. Toxicol.* **2010**, *202*, 53–154.
- (97) Hughes, S. R.; Kay, P.; Brown, L. E. *Environ. Sci. Technol.* **2013**, *47* (2), 661–677.
- (98) Bradley, P. M.; Journey, C. A.; Romanok, K. M.; Barber, L. B.; Buxton, H. T.; Foreman, W. T.; Furlong, E. T.; Glassmeyer, S. T.; Hladik, M. L.; Iwanowicz, L. R.; Jones, D. K.; Kolpin, D. W.; Kuivila, K. M.; Loftin, K. A.; Mills, M. A.; Meyer, M. T.; Orlando, J. L.; Reilly, T. J.; Smalling, K. L.; Villeneuve, D. L. *Environ. Sci. Technol.* **2017**, *51* (9), 4792–4802.
- (99) Segura, P. A.; François, M.; Gagnon, C.; Sauv e, S. *Environ. Health Perspect.* **2009**, *117* (5), 675–684.
- (100) Brausch, J. M.; Connors, K. A.; Brooks, B. W.; Rand, G. M. *Rev. Environ. Contam. Toxicol.* **2012**, *218*, 1–99.
- (101) Vasquez, M. I.; Lambrianides, A.; Schneider, M.; K ummerer, K.; Fatta-Kassinos, D. *J. Hazard. Mater.* **2014**, *279*, 169–189.
- (102) Schymanski, E. L.; Jeon, J.; Gulde, R.; Fenner, K.; Ruff, M.; Singer, H. P.;

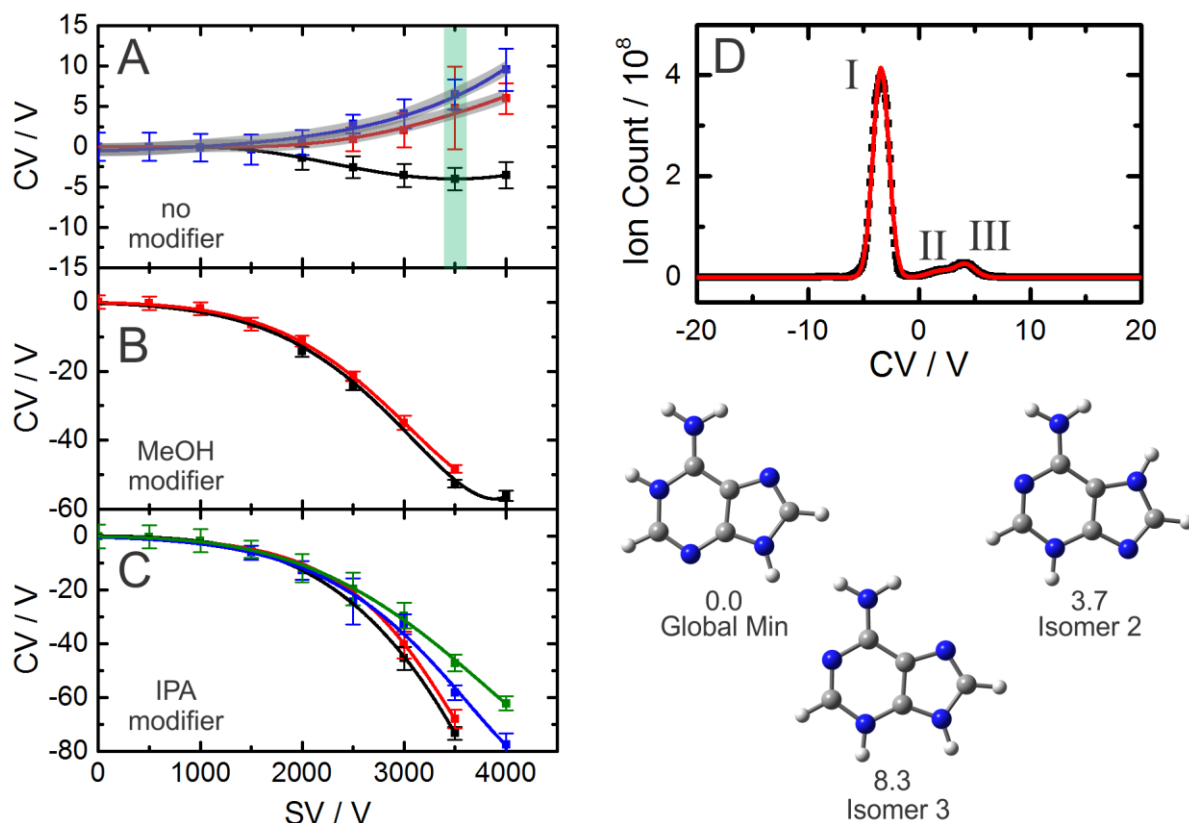
- Hollender, J. *Environ. Sci. Technol.* **2014**, *48* (4), 2097–2098.
- (103) Boxall, A. B. A.; Sinclair, C. J.; Fenner, K.; Kolpin, D.; Maund, S. J. *Environ. Sci. Technol.* **2004**, *38* (19), 368A–375A.
- (104) Jasak, J.; Le Blanc, Y.; Speer, K.; Billian, P.; Schoening, R. M. *J. AOAC Int.* **2012**, *95* (6), 1786–76.
- (105) Li, S.; Jia, J.; Gao, X.; He, X.; Li, J. *Anal. Chim. Acta* **2012**, *720*, 97–103.
- (106) Sinatra, F. L.; Wu, T.; Manolakos, S.; Wang, J.; Evans-Nguyen, T. G. *Anal. Chem.* **2015**, *87* (3), 1685–1693.
- (107) Campbell, J. L.; Baba, T.; Liu, C.; Lane, C. S.; Le Blanc, J. C. Y.; Hager, J. *W. J. Am. Soc. Mass Spectrom.* **2017**, *28* (7), 1374–1381.
- (108) Shvartsburg, A. A.; Smith, R. D. *Anal. Chem.* **2013**, *85* (14), 6967–6973.
- (109) Davis, C. E.; Bogan, M. J.; Sankaran, S.; Molina, M. A.; Loyola, B. R.; Zhao, W.; Benner, W. H.; Schivo, M.; Farquar, G. R.; Kenyon, N. J.; Frank, M. *IEEE Sens. J.* **2010**, *10* (1), 114–122.
- (110) Eiceman, G. A.; Krylov, E. V.; Tadjikov, B.; Ewing, R. G.; Nazarov, E. G.; Miller, R. A. *Analyst* **2004**, *129* (4), 297–304.
- (111) Porta, T.; Varesio, E.; Hopfgartner, G. *Anal. Chem.* **2013**, *85* (24), 11771–11779.
- (112) Noestheden, M. R.; Headley, J. V.; Peru, K. M.; Barrow, M. P.; Burton, L. L.; Sakuma, T.; Winkler, P.; Campbell, J. L. *Environ. Sci. Technol.* **2014**, *48* (17), 10264–10272.
- (113) Jewell, K. S.; Castronovo, S.; Wick, A.; Falås, P.; Joss, A.; Ternes, T. A. *Water Res.* **2016**, *88*, 550–557.
- (114) Managaki, S.; Murata, A.; Takada, H.; Bui, C. T.; Chiem, N. H. *Environ. Sci. Technol.* **2007**, *41* (23), 8004–8010.
- (115) Chang, H.; Hu, J.; Asami, M.; Kunikane, S. *J. Chromatogr. A* **2008**, *1190* (1–2), 390–393.
- (116) Jurva, U.; Wikström, H. V.; Bruins, A. P. *Rapid Commun. Mass Spectrom.* **2002**, *16* (20), 1934–1940.
- (117) Subramanyam, R.; Mishra, I. *J. Hazardous, Toxic, ...* **2012**, *19* (JULY), 218–229.
- (118) Becke, A. D. *J. Chem. Phys.* **1993**, *98* (7), 5648.
- (119) Campbell, M. T.; Glish, G. L. *Int. J. Mass Spectrom.* **2018**.
- (120) Schwarz, H.; Stahl, D. *Int. J. Mass Spectrom. Ion Phys.* **1980**, *36*, 285–289.
- (121) Wilfried, M. A. N.; Ricardo, A.; Correa, C. *Interpretation of MS-MS Mass Spectra of Drugs and Pesticides*; 2017.
- (122) Liu, Z. Y.; Wu, Y.; Sun, Z. L.; Wan, L. *Biomed. Chromatogr.* **2012**, *26* (9), 1101–1108.
- (123) Eckers, C.; Monaghan, J. J.; Wolff, J.-C. *Eur. J. Mass Spectrom. (Chichester, Eng).* **2005**, *11* (1), 73–82.



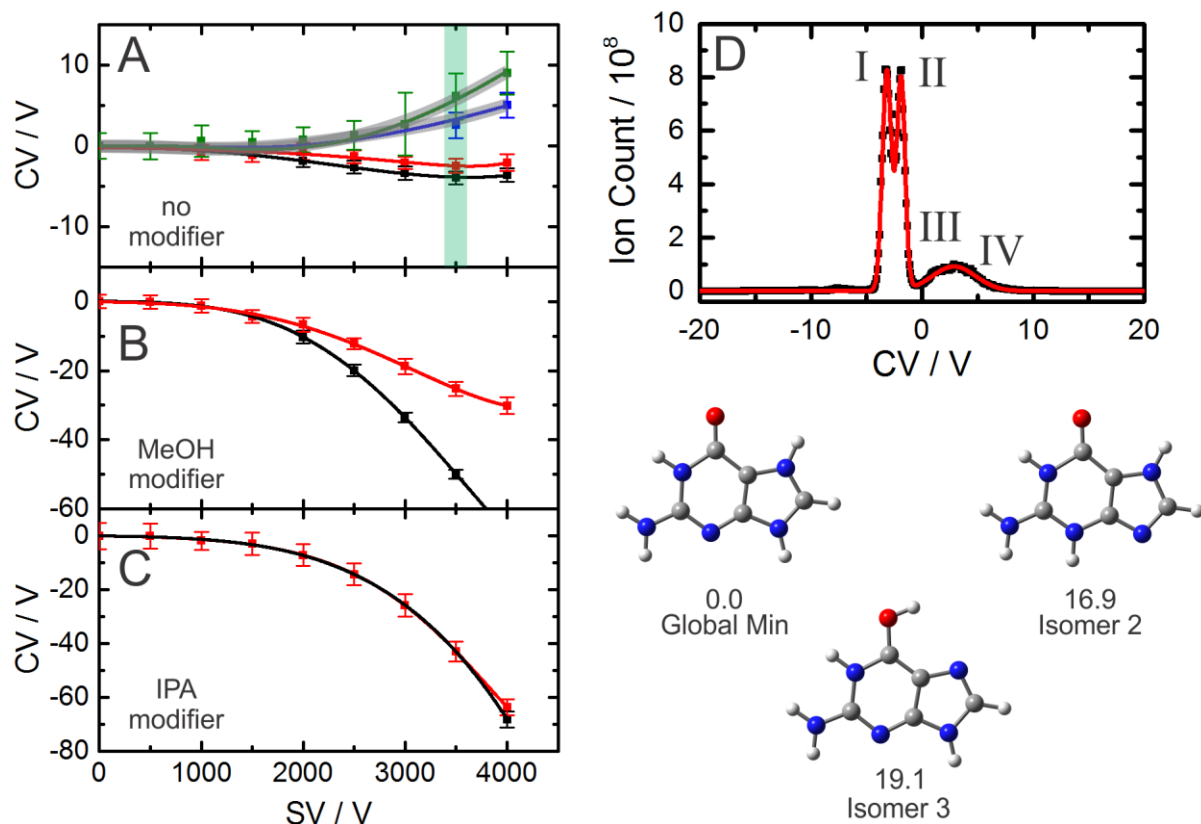
- (124) Ray, J. A.; Kushnir, M. M.; Yost, R. A.; Rockwood, A. L.; Wayne Meikle, A. *Clin. Chim. Acta* **2014**, *438*, 330–336.
- (125) Stoll, D. R.; Carr, P. W. *Anal. Chem.* **2017**, *89* (1), 519–531.
- (126) Gómez-Ríos, G. A.; Liu, C.; Tascon, M.; Reyes-Garcés, N.; Arnold, D. W.; Covey, T. R.; Pawliszyn, J. *Anal. Chem.* **2017**, *89* (7), 3805–3809.
- (127) Armbruster, D. A.; Pry, T. *Clin. Biochem. Rev.* **2008**, *29 Suppl 1* (August), S49-52.
- (128) Liu, C.; Gómez-Ríos, G. A.; Schneider, B. B.; Le Blanc, J. C. Y.; Reyes-Garcés, N.; Arnold, D. W.; Covey, T. R.; Pawliszyn, J. *Anal. Chim. Acta* **2017**, *991*, 89–94.

## APPENDIX A

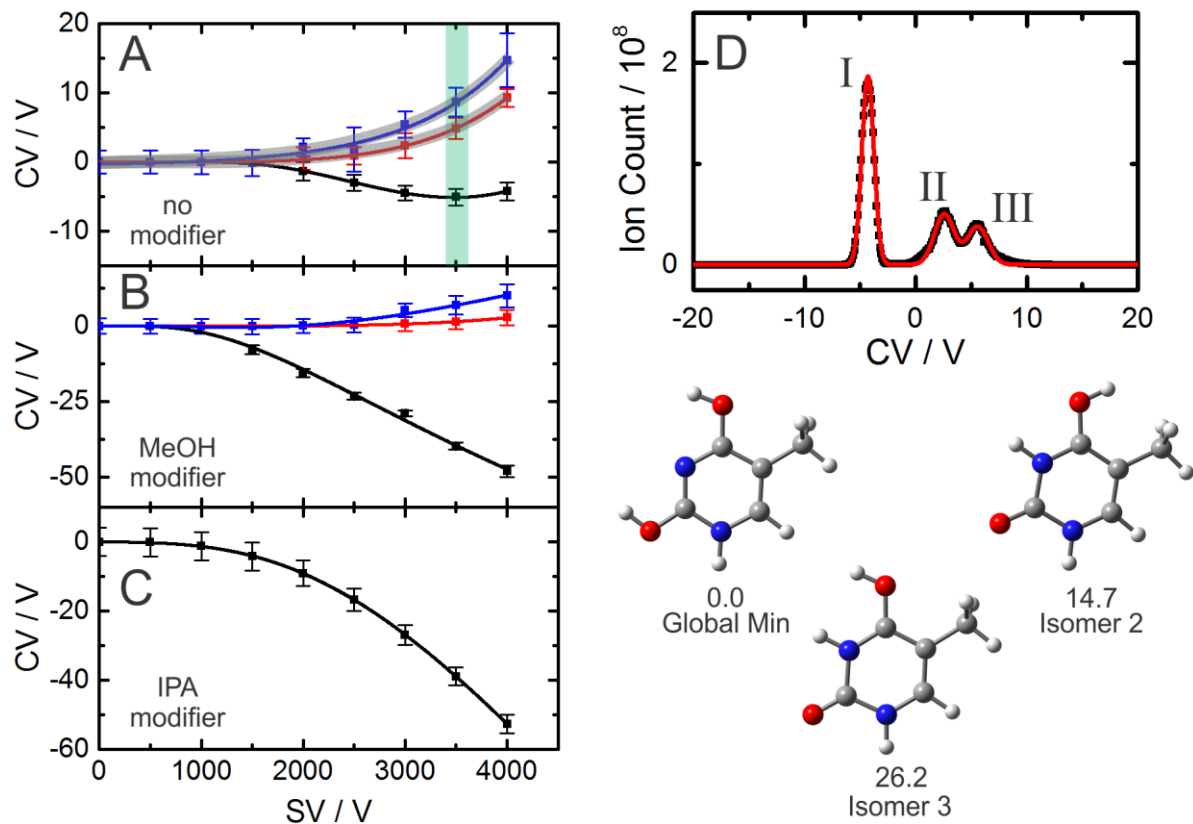
### Dispersion Plots



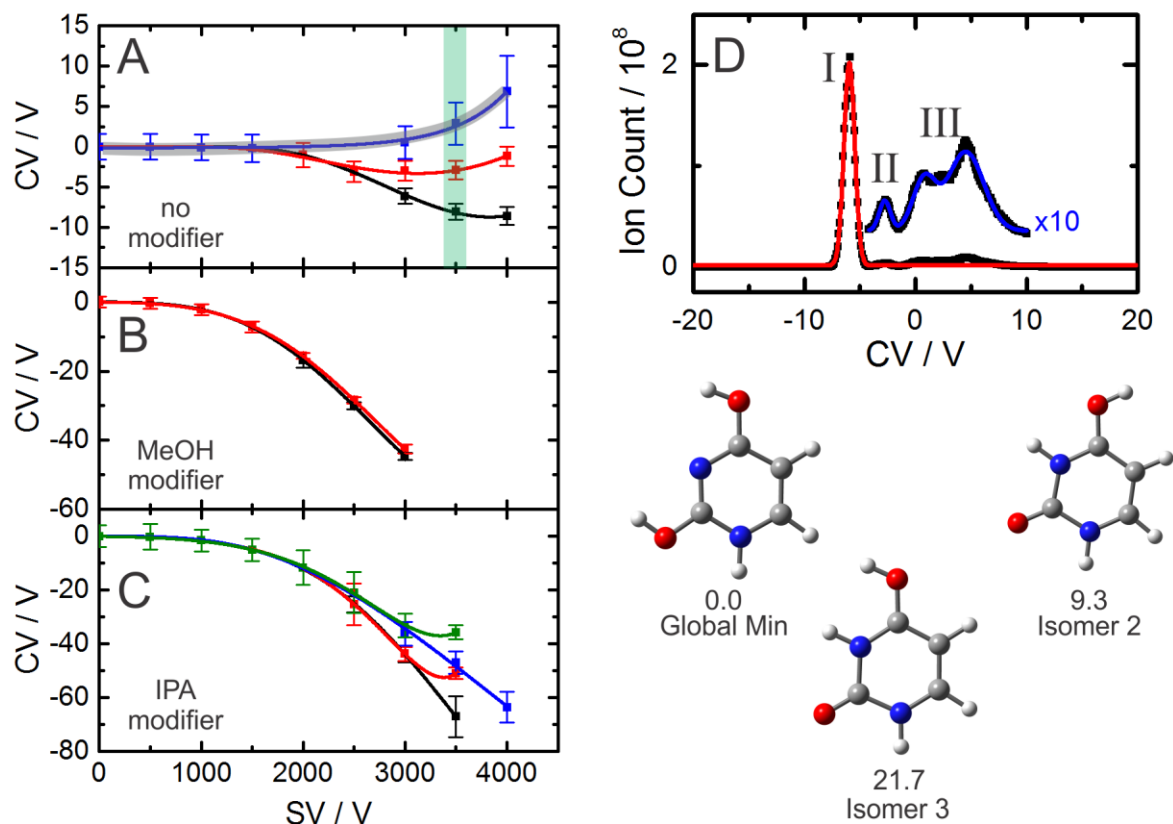
The dispersion plot obtained for  $A\cdot H^+$  ( $m/z = 136$  amu) with a DMS cell containing (A) a pure  $N_2$  environment, and a  $N_2$  environment seeded with 1.5% (v/v) (B) methanol vapor, and (C) isopropyl alcohol vapor. Error bars are  $2\sigma$  obtained from Gaussian fits to the ionogram peaks. (D) The ionogram recorded for the  $m/z = 136$  amu peak in a pure  $N_2$  environment with  $SV = 3500$  V (highlighted green in A). (Inset) The three lowest energy tautomers of  $A\cdot H^+$  as calculated at the CCSD(T)/6-311++G(d,p)//B3LYP/6-311++G(d,p) level of theory. Energies are reported as standard Gibbs' energies in  $\text{kJ mol}^{-1}$ .



The dispersion plot obtained for  $G\cdot H^+$  ( $m/z = 152$  amu) with a DMS cell containing (A) a pure  $N_2$  environment, and a  $N_2$  environment seeded with 1.5% (v/v) (B) methanol vapor, and (C) isopropyl alcohol vapor. Error bars are  $2\sigma$  obtained from Gaussian fits to the ionogram peaks. (D) The ionogram recorded for the  $m/z = 152$  amu peak in a pure  $N_2$  environment with  $SV = 3500$  V (highlighted green in A). (Inset) The three lowest energy tautomers of  $G\cdot H^+$  as calculated at the CCSD(T)/6-311++G(d,p)//B3LYP/6-311++G(d,p) level of theory. Energies are reported as standard Gibbs' energies in  $\text{kJ mol}^{-1}$ .

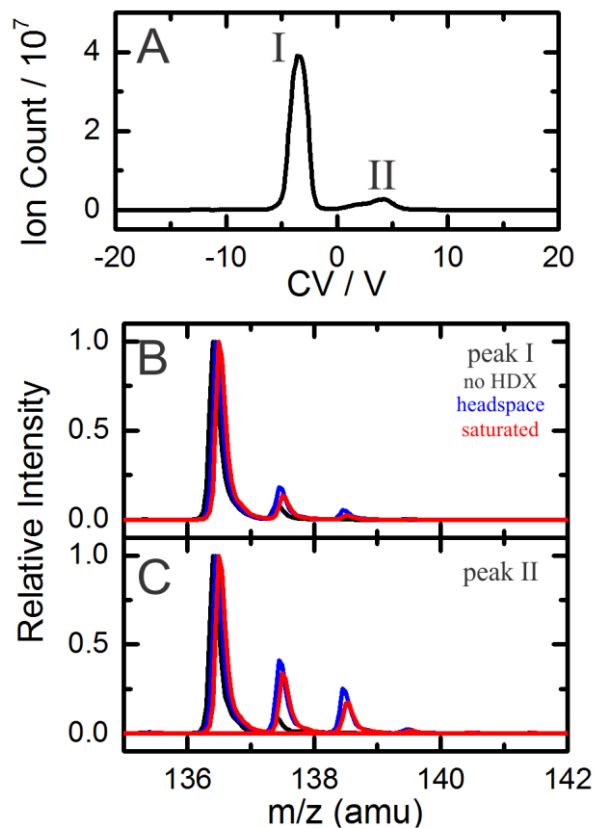


The dispersion plot obtained for  $T\cdot H^+$  ( $m/z = 127$  amu) with a DMS cell containing (A) a pure  $N_2$  environment, and a  $N_2$  environment seeded with 1.5% (v/v) (B) methanol vapor, and (C) isopropyl alcohol vapor. Error bars are  $2\sigma$  obtained from Gaussian fits to the ionogram peaks. (D) The ionogram recorded for the  $m/z = 127$  amu peak in a pure  $N_2$  environment with  $SV = 3500$  V (highlighted green in A). (Inset) The three lowest energy tautomers of  $T\cdot H^+$  as calculated at the CCSD(T)/6-311++G(d,p)//B3LYP/6-311++G(d,p) level of theory. Energies are reported as standard Gibbs' energies in  $\text{kJ mol}^{-1}$ .

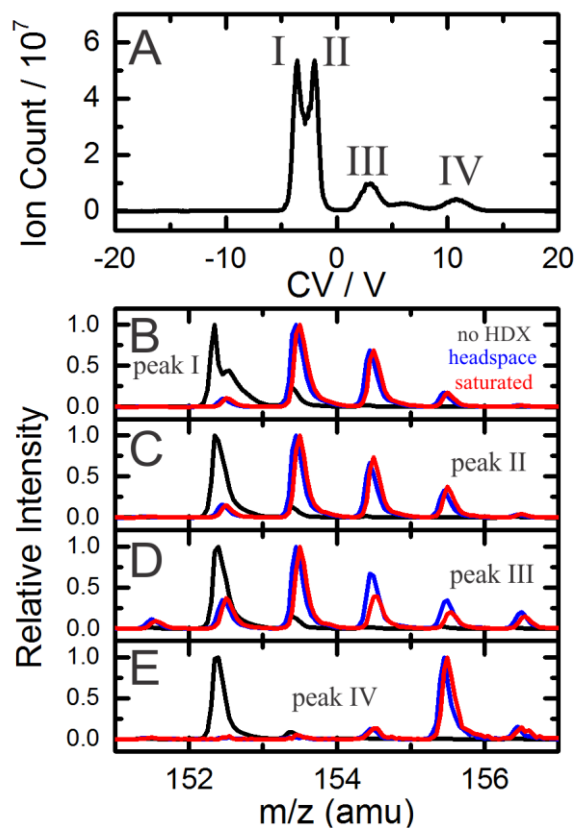


The dispersion plot obtained for  $U\cdot H^+$  ( $m/z = 113$  amu) with a DMS cell containing (A) a pure  $N_2$  environment, and a  $N_2$  environment seeded with 1.5% (v/v) (B) methanol vapor, and (C) isopropyl alcohol vapor. Error bars are  $2\sigma$  obtained from Gaussian fits to the ionogram peaks. (D) The ionogram recorded for the  $m/z = 113$  amu peak in a pure  $N_2$  environment with SV = 3500 V (highlighted green in A). (Inset) The three lowest energy tautomers of  $U\cdot H^+$  as calculated at the CCSD(T)/6-311++G(d,p)//B3LYP/6-311++G(d,p) level of theory. Energies are reported as standard Gibbs' energies in  $\text{kJ mol}^{-1}$ .

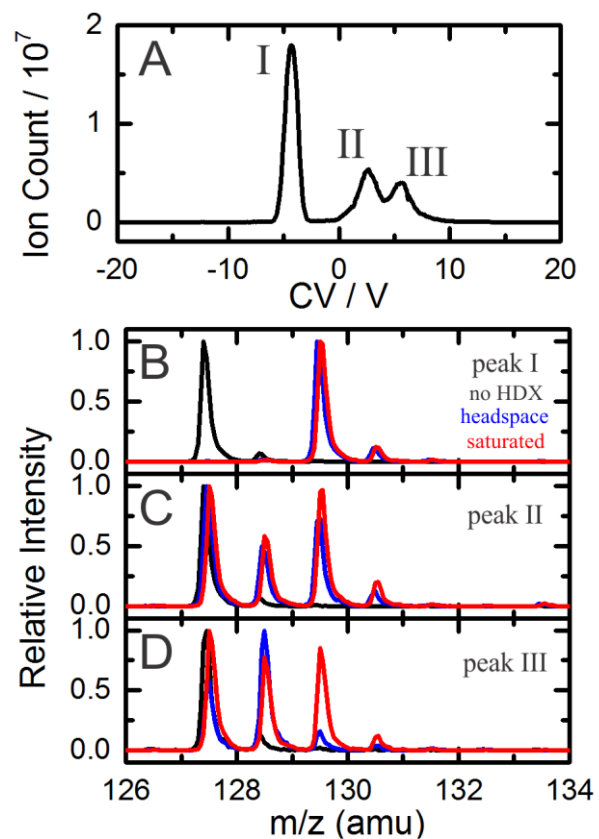
## HDX Profiles



(A) The ionogram recorded when gating on  $A \cdot H^+$  ( $m/z = 136$  amu). The results of HDX experiments when isolating on (B) peak I, and (C) peak II. The black traces show the observed mass distributions in the absence of HDX reagent. The blue traces show the effect of introducing a low vapor pressure of  $D_2O$ , and the red traces are observed following HDX in  $N_2$  at  $18^\circ C$  seeded with a saturated partial pressure of  $D_2O$ .



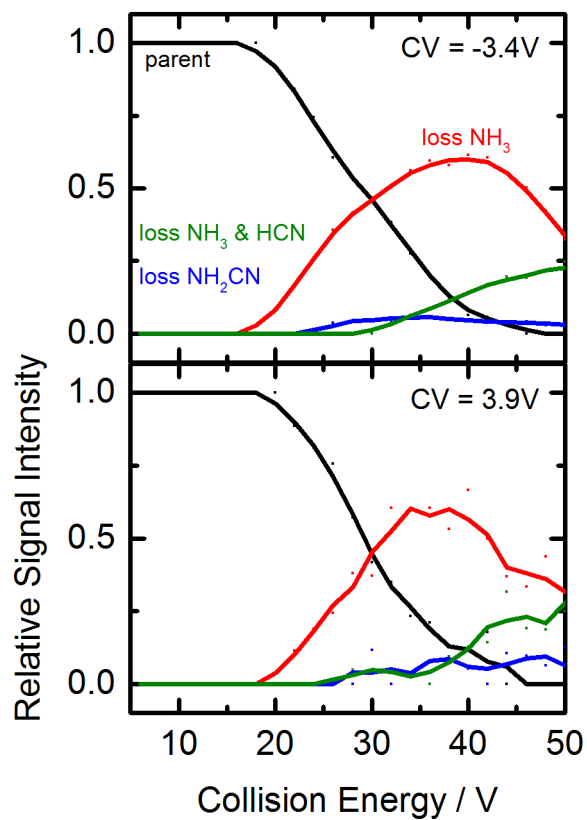
(A) The ionogram recorded when gating on  $G \cdot H^+$  ( $m/z = 152$  amu). The results of HDX experiments when isolating on (B) peak I, (C) peak II, (D) peak III, and (E) peak IV. The black traces show the observed mass distributions in the absence of HDX reagent. The blue traces show the effect of introducing a low vapor pressure of  $D_2O$ , and the red traces are observed following HDX in  $N_2$  at  $18^\circ C$  seeded with a saturated partial pressure of  $D_2O$ .



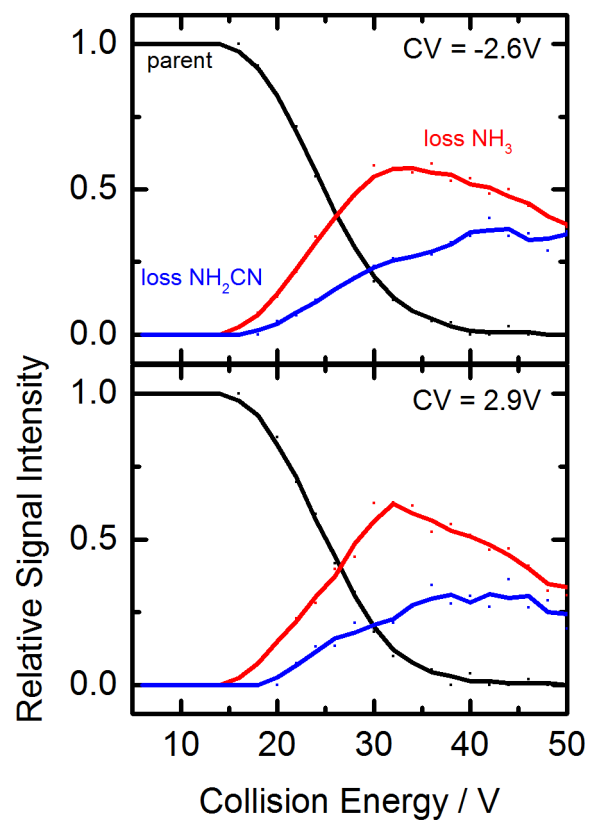
(A) The ionogram recorded when gating on  $T \cdot H^+$  ( $m/z = 127$  amu). The results of HDX experiments when isolating on (B) peak I, (C) peak II, and (D) peak III. The black traces show the observed mass distributions in the absence of HDX reagent. The blue traces show the effect of introducing a low vapor pressure of  $D_2O$ , and the red traces are observed following HDX in  $N_2$  at  $18^\circ C$  seeded with a saturated partial pressure of  $D_2O$ .



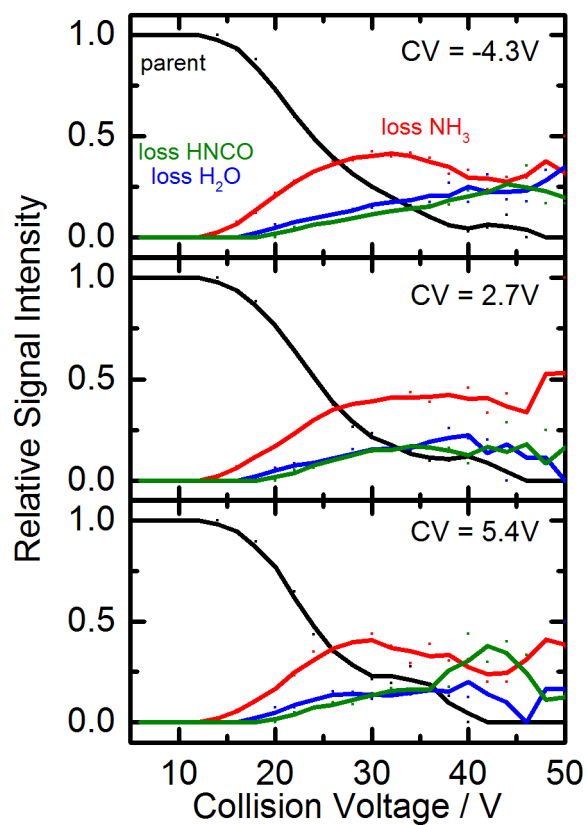
## Breakdown Curves



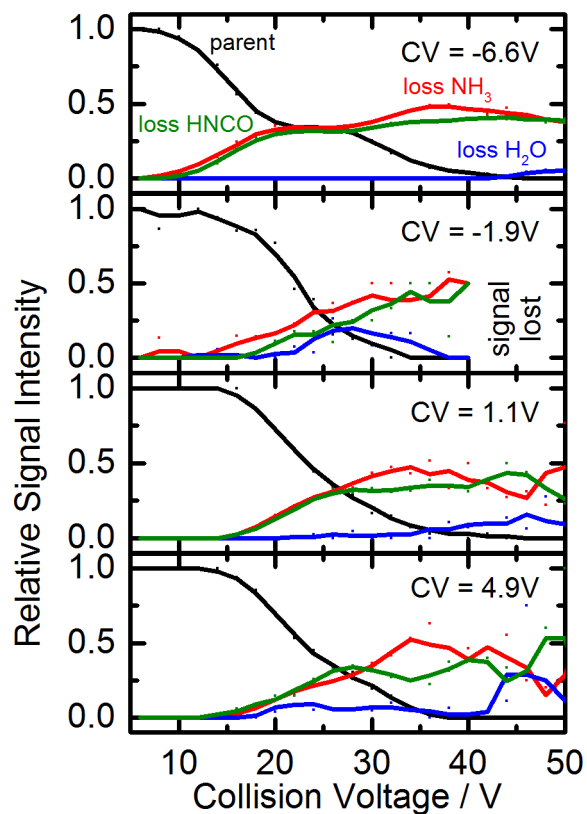
The breakdown curves obtained for  $A\cdot H^+$  ( $m/z = 136$  amu) when isolating on the ionogram peaks plotted in Figure S5A. **(A)** peak I (CV = -3.4 V), and **(B)** peak II (CV = 3.9 V). Collision energy was increased in 2.5 V increments from 0 to 50 V.



The breakdown curves obtained for  $G\cdot H^+$  ( $m/z = 152$  amu) when isolating on the ionogram peaks plotted in Figure S6A. **(A)** peak I ( $CV = -2.6$  V), and **(B)** peak II ( $CV = 2.9$  V). Collision energy was increased in 2.5 V increments from 0 to 50 V.

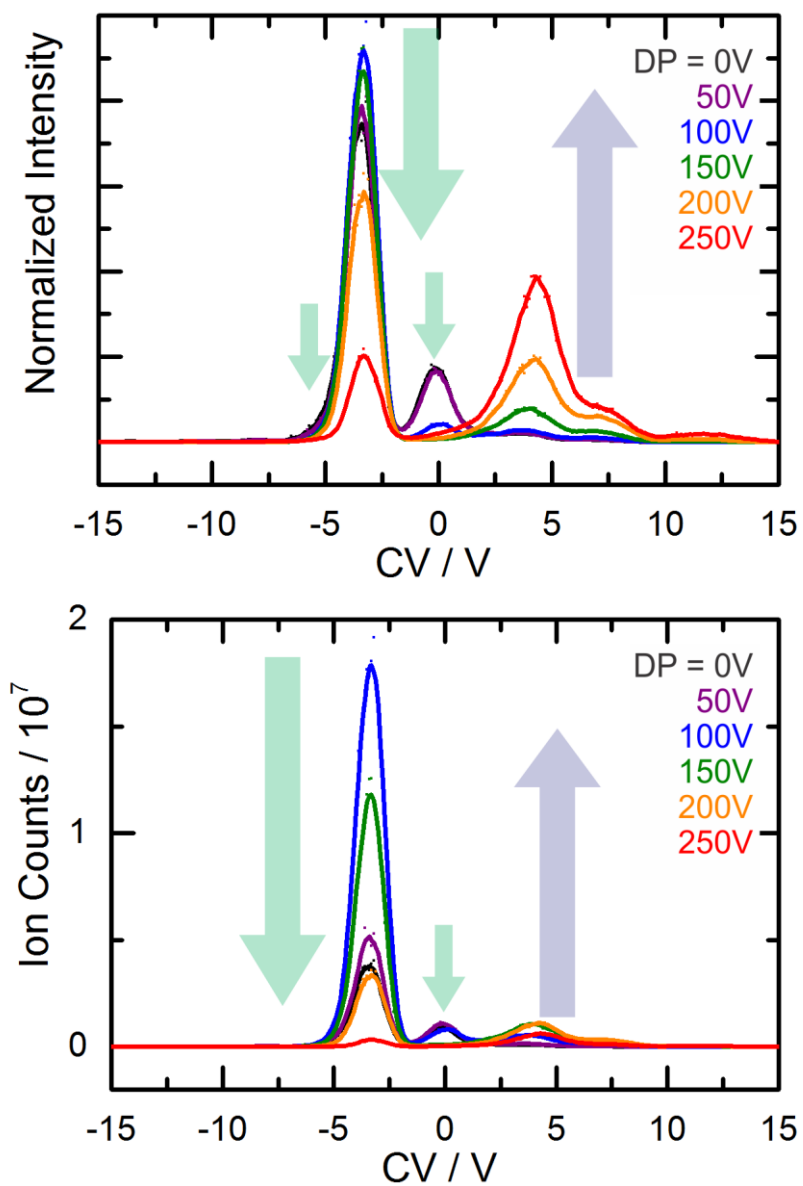


The breakdown curves obtained for  $T\cdot H^+$  ( $m/z = 127$  amu) when isolating on the ionogram peaks plotted in Figure S7A. **(A)** peak I ( $CV = -4.3$  V), **(B)** peak II ( $CV = 2.7$  V), and **(C)** peak III ( $CV = 5.4$  V). Collision energy was increased in 2.5 V increments from 0 to 50 V.

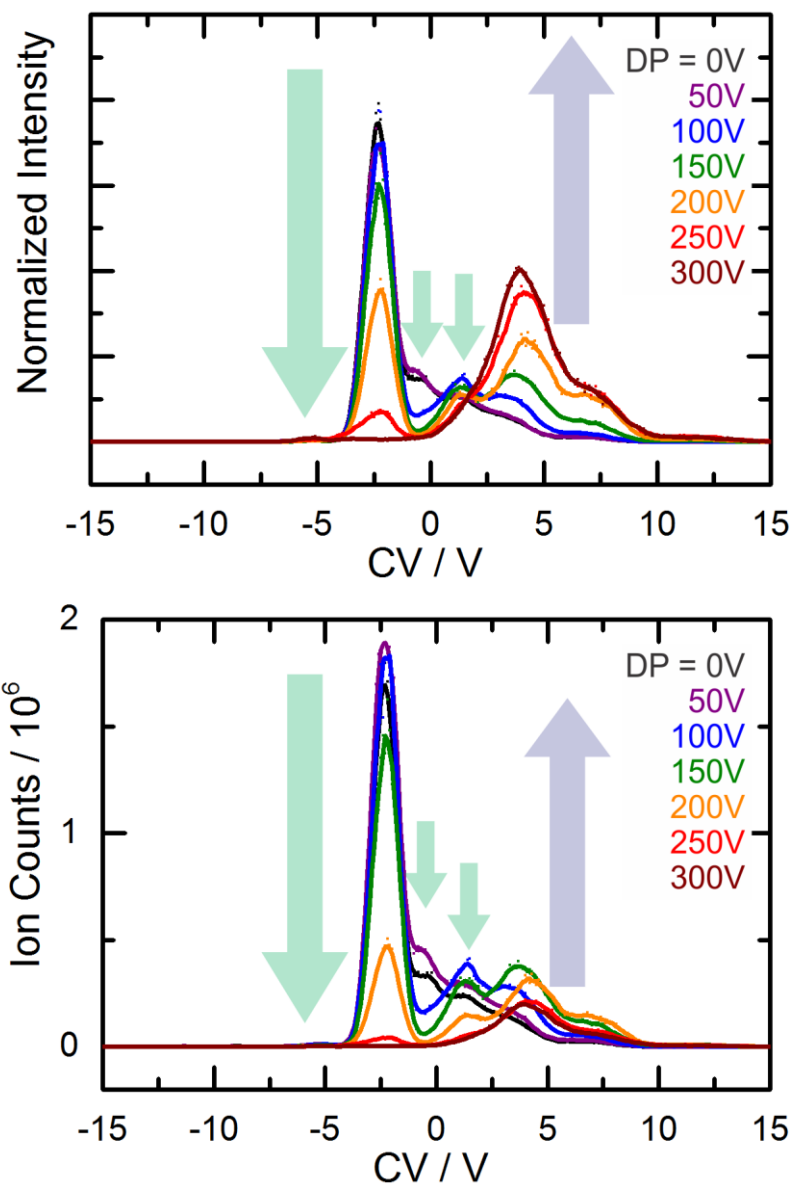


The breakdown curves obtained for  $U\cdot H^+$  ( $m/z = 113$  amu) when isolating on the ionogram peaks plotted in Figure S4D. **(A)** peak I (CV = -6.6 V), **(B)** peak II (CV = -1.9 V), **(C)** peak III (CV = 1.1 V), and **(D)** peak III (CV = 4.9 V). Collision energy was increased in 2.5 V increments from 0 to 50 V.

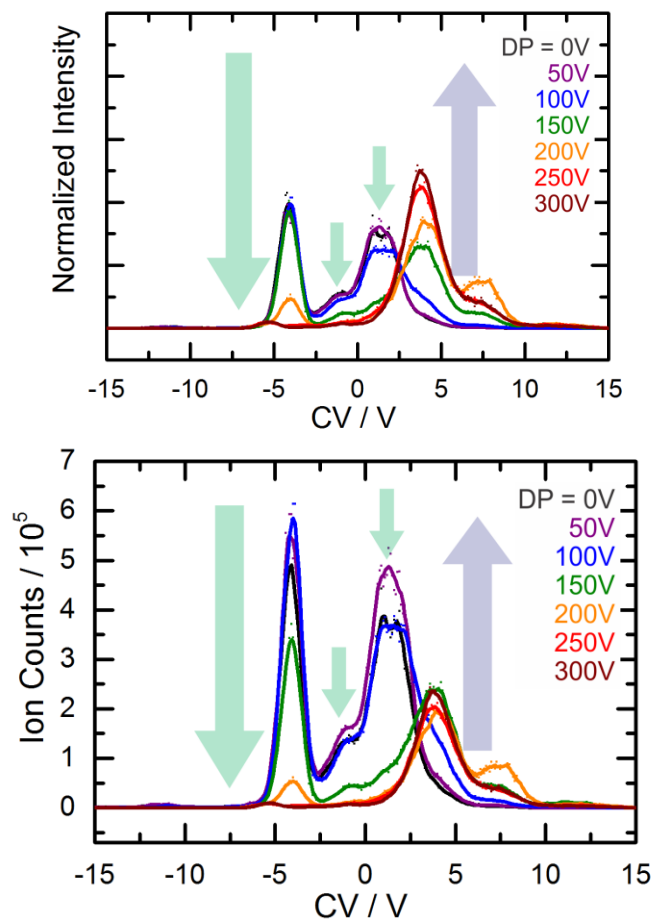
## Declustering Potential Scans



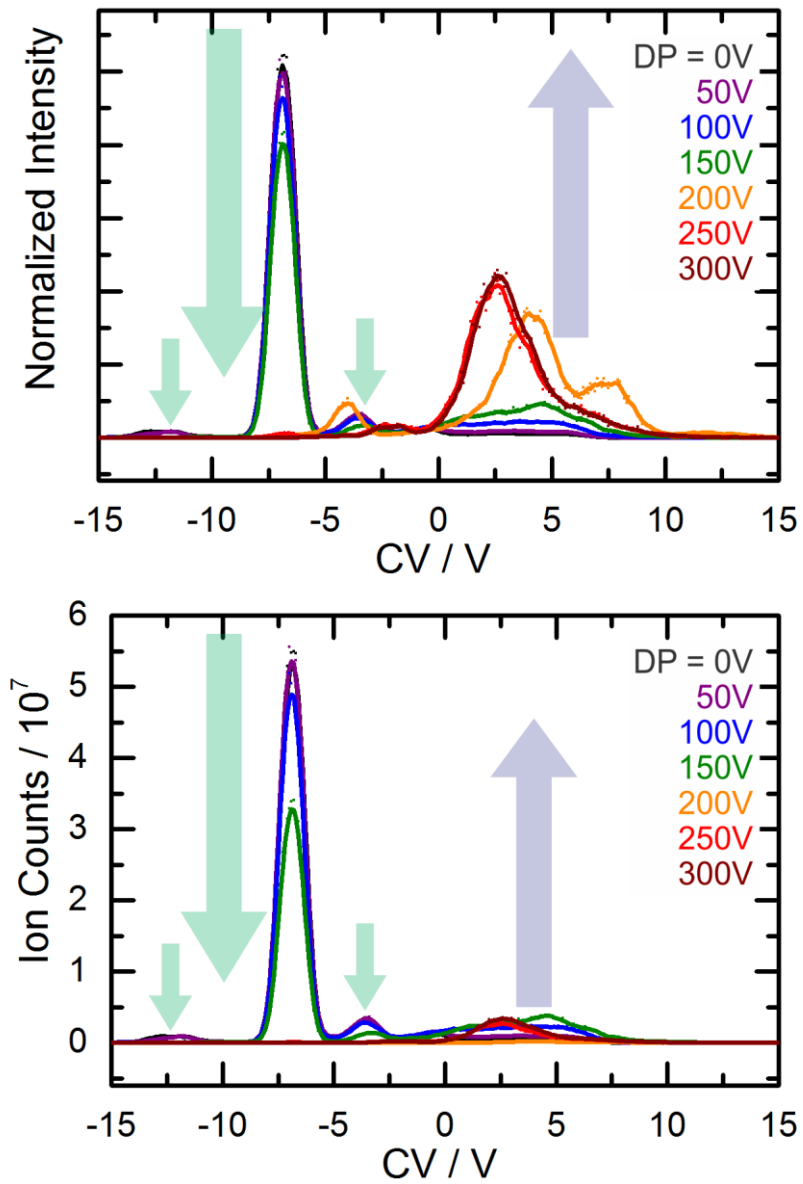
The ionogram recorded for  $A\cdot H^+$  ( $m/z = 136$  amu) in a pure  $N_2$  environment at  $SV = 3500$  V as the declustering potential, DP, is stepped from 0–300 V in 50 V increments.



The ionogram recorded for  $G\cdot H^+$  ( $m/z = 152$  amu) in a pure  $N_2$  environment at  $SV = 3500$  V as the declustering potential, DP, is stepped from 0–300 V in 50 V increments.



The ionogram recorded for  $T\cdot H^+$  ( $m/z = 127$  amu) in a pure  $N_2$  environment at  $SV = 3500$  V as the declustering potential, DP, is stepped from 0–300 V in 50 V increments.



The ionogram recorded for  $U\cdot H^+$  ( $m/z = 113$  amu) in a pure  $N_2$  environment at  $SV = 3500$  V as the declustering potential, DP, is stepped from 0–300 V in 50 V increments.



## Calculation Results

Thermochemical data for the protonated nucleobases A•H<sup>+</sup>, G•H<sup>+</sup>, C•H<sup>+</sup>, T•H<sup>+</sup>, and U•H<sup>+</sup> as calculated at the B3LYP/6-311++G(d,p) level of theory at a temperature of T = 298 K and a pressure of P = 1 atm.

### Protonated Adenine, A•H<sup>+</sup>

Isomer	Electronic Energy (hartree)	Zero Point Corrected Energy (hartree)	Gibbs Energy (hartrees)	Enthalpy (hartree)
A1_ha	-467.7862364	-467.6606034	-467.6521214	-467.6932854
A1_hb	-467.8223598	-467.6972598	-467.6890108	-467.7289478
A1_hc	-467.8200078	-467.6948738	-467.6866708	-467.7265808
A1_hd	-467.8088895	-467.6844655	-467.6759125	-467.7166165
A10_hb	-467.7930656	-467.6677786	-467.6596696	-467.6995776
A2_ha	-467.762282	-467.637553	-467.629042	-467.669811
A2_hb	-467.8044384	-467.6805774	-467.6718514	-467.7127504
A2_hc	-467.8223625	-467.6973225	-467.6890145	-467.7291495
A3_ha	-467.7661793	-467.6404263	-467.6320453	-467.6726923
A3_hb	-467.7833006	-467.6589916	-467.6504836	-467.6909626
A4_ha	-467.7284774	-467.6040174	-467.5954624	-467.6363064
A5_hb	-467.7985315	-467.6730855	-467.6650655	-467.7047595
A5_hc	-467.7771226	-467.6526716	-467.6442766	-467.6847426
A6_hb	-467.7706426	-467.6466176	-467.6382346	-467.6786806
A7_hb	-467.791943	-467.666887	-467.658745	-467.698689
A8_hb	-467.7935646	-467.6682706	-467.6602226	-467.7000166
A9_hb	-467.7824872	-467.6576602	-467.6494202	-467.6895352

### Protonated Guanine, G•H<sup>+</sup>

Isomer	Electronic Energy (hartree)	Zero Point Corrected Energy (hartree)	Gibbs Energy (hartrees)	Enthalpy (hartree)
G1_ha	-543.0793901	-542.9504341	-542.9408281	-542.9837781
G1_hb	-543.0651693	-542.9377213	-542.9273933	-542.9726193
G1_hc	-543.0264163	-542.8969523	-542.8873843	-542.9305603
G1_hd	-543.0615177	-542.9336937	-542.9235867	-542.9676407

<b>G1_hf</b>	-543.0895771	-542.9600751	-542.9505481	-542.9934471
<b>G10_ha</b>	-543.0140475	-542.8866475	-542.8767375	-542.9210575
<b>G10_hb</b>	-543.0180381	-542.8906291	-542.8808121	-542.9246611
<b>G11_ha</b>	-543.0162177	-542.8890177	-542.8797077	-542.9225627
<b>G11_hb</b>	-543.036913	-542.908075	-542.898645	-542.941694
<b>G2_ha</b>	-543.0611553	-542.9335563	-542.9234273	-542.9674943
<b>G2_hb</b>	-543.0585815	-542.9318215	-542.9220775	-542.9656035
<b>G2_hc</b>	-543.0329627	-542.9031547	-542.8936497	-542.9366957
<b>G2_hd</b>	-543.0814218	-542.9525218	-542.9427398	-542.9861178
<b>G3_ha</b>	-543.0389691	-542.9107321	-542.9008731	-542.9442981
<b>G3_hb</b>	-543.0207935	-542.8945295	-542.8845775	-542.9284855
<b>G3_hc</b>	-542.9454278	-542.8185278	-542.8082148	-542.8533818
<b>G4_ha</b>	-543.0483665	-542.9195605	-542.9101745	-542.9529225
<b>G4_hb</b>	-543.0377422	-542.9096662	-542.9000792	-542.9431852
<b>G4_hf</b>	-543.0500904	-542.9213074	-542.9116714	-542.9551574
<b>G5_ha</b>	-543.0433891	-542.9151531	-542.9054631	-542.9489331
<b>G5_hb</b>	-543.0455082	-542.9172012	-542.9075902	-542.9508272
<b>G6_ha</b>	-543.0243706	-542.8963896	-542.8866016	-542.9304636
<b>G6_hb</b>	-543.0302188	-542.9019288	-542.8923618	-542.9355698
<b>G7_ha</b>	-543.0035113	-542.8770513	-542.8675683	-542.9108943
<b>G7_hb</b>	-543.0256888	-542.8975158	-542.8879348	-542.9313398
<b>G8_ha</b>	-543.0486641	-542.9198271	-542.9104681	-542.9531851
<b>G8_hb</b>	-543.0362528	-542.9083958	-542.8987248	-542.9420218
<b>G8_hf</b>	-543.0536551	-542.9245501	-542.9150541	-542.9582401
<b>G9_ha</b>	-543.0415144	-542.9134444	-542.9037144	-542.9472754
<b>G9_hb</b>	-543.0417086	-542.9138816	-542.9040866	-542.9477166

Protonated Cytosine, C•H<sup>+</sup>

<b>Isomer</b>	<b>Electronic Energy (hartree)</b>	<b>Zero Point Corrected Energy (hartree)</b>	<b>Gibbs Energy (hartrees)</b>	<b>Enthalpy (hartree)</b>
<b>C1_ha</b>	-395.3771683	-395.2652673	-395.2574053	-395.2966393
<b>C1_hb</b>	-395.4289641	-395.3169051	-395.3093061	-395.3475201
<b>C1_hc</b>	-395.4284661	-395.3162371	-395.3087981	-395.3466311
<b>C1_hd</b>	-395.4145863	-395.3031403	-395.2954303	-395.3337793
<b>C2_ha</b>	-395.3735097	-395.2616727	-395.2537957	-395.2934007
<b>C2_hb</b>	-395.39698	-395.286607	-395.278544	-395.317539
<b>C2_hd</b>	-395.3518095	-395.2433295	-395.2355585	-395.2739515
<b>C3_ha</b>	-395.3781451	-395.2660971	-395.2583131	-395.2974671
<b>C3_hb</b>	-395.4143491	-395.3028441	-395.2952231	-395.3333321
<b>C4_ha</b>	-395.3403413	-395.2295123	-395.2214203	-395.2610903
<b>C5_hc</b>	-395.3751082	-395.2646842	-395.2568302	-395.2956512
<b>C5_hd</b>	-395.3746832	-395.2642972	-395.2564042	-395.2953142
<b>C6_hc</b>	-395.382475	-395.271502	-395.263792	-395.302303

<b>C6_hd</b>	-395.380275	-395.269558	-395.261706	-395.300517
--------------	-------------	-------------	-------------	-------------

Protonated Thymine, T•H<sup>+</sup>

<b>Isomer</b>	<b>Electronic Energy (hartree)</b>	<b>Zero Point Corrected Energy (hartree)</b>	<b>Gibbs Energy (hartrees)</b>	<b>Enthalpy (hartree)</b>
<b>T1_ha</b>	-454.6015387	-454.4755627	-454.4662457	-454.5083647
<b>T1_hb</b>	-454.6034927	-454.4773677	-454.4681067	-454.5101217
<b>T1_hc</b>	-454.6112872	-454.4844752	-454.4754072	-454.5171072
<b>T1_hd</b>	-454.6155442	-454.4886262	-454.4795912	-454.5212822
<b>T1_hf</b>	-454.570469	-454.444629	-454.435072	-454.478668
<b>T2_ha</b>	-454.6059178	-454.4792978	-454.4702568	-454.5117798
<b>T2_hb</b>	-454.6189561	-454.4916751	-454.4828511	-454.5239641
<b>T2_hd</b>	-454.5329734	-454.4086794	-454.3990744	-454.4418694
<b>T3_hb</b>	-454.5182132	-454.3944792	-454.3847772	-454.4276422
<b>T3_hc</b>	-454.5988146	-454.4727176	-454.4635536	-454.5053886
<b>T3_hd</b>	-454.6052203	-454.4788063	-454.4697503	-454.5114493
<b>T3_hf</b>	-454.5438213	-454.4190703	-454.4093883	-454.4532873
<b>T4_hb</b>	-454.5459386	-454.4210226	-454.4116826	-454.4537126
<b>T4_hd</b>	-454.5327438	-454.4083118	-454.3988068	-454.4414618

Protonated Uracil, U•H<sup>+</sup>

<b>Isomer</b>	<b>Electronic Energy (hartree)</b>	<b>Zero Point Corrected Energy (hartree)</b>	<b>Gibbs Energy (hartrees)</b>	<b>Enthalpy (hartree)</b>
<b>U1_ha</b>	-415.2656	-415.167138	-415.159512	-415.197905
<b>U1_hb</b>	-415.2678807	-415.1692447	-415.1617037	-415.1999247
<b>U1_hc</b>	-415.2787264	-415.1794584	-415.1721254	-415.2099414
<b>U1_hd</b>	-415.2834329	-415.1838719	-415.1766129	-415.2143049
<b>U1_hf</b>	-415.237362	-415.13921	-415.131311	-415.171113
<b>U2_ha</b>	-415.2711901	-415.1720571	-415.1647281	-415.2024621
<b>U2_hb</b>	-415.2845737	-415.1847597	-415.1776527	-415.2149727
<b>U2_hd</b>	-415.201573	-415.10474	-415.096899	-415.135735
<b>U3_hb</b>	-415.1825839	-415.0863849	-415.0783669	-415.1175119
<b>U3_hc</b>	-415.2651514	-415.1664374	-415.1590734	-415.1968124
<b>U3_hd</b>	-415.2721968	-415.1730008	-415.1657958	-415.2032518
<b>U3_hf</b>	-415.2099919	-415.1130699	-415.1058529	-415.1435279
<b>U4_hb</b>	-415.2118865	-415.1144365	-415.1068305	-415.1449925

<b>U4_hd</b>	-415.2000403	-415.1029353	-415.0952843	-415.1336043
--------------	--------------	--------------	--------------	--------------

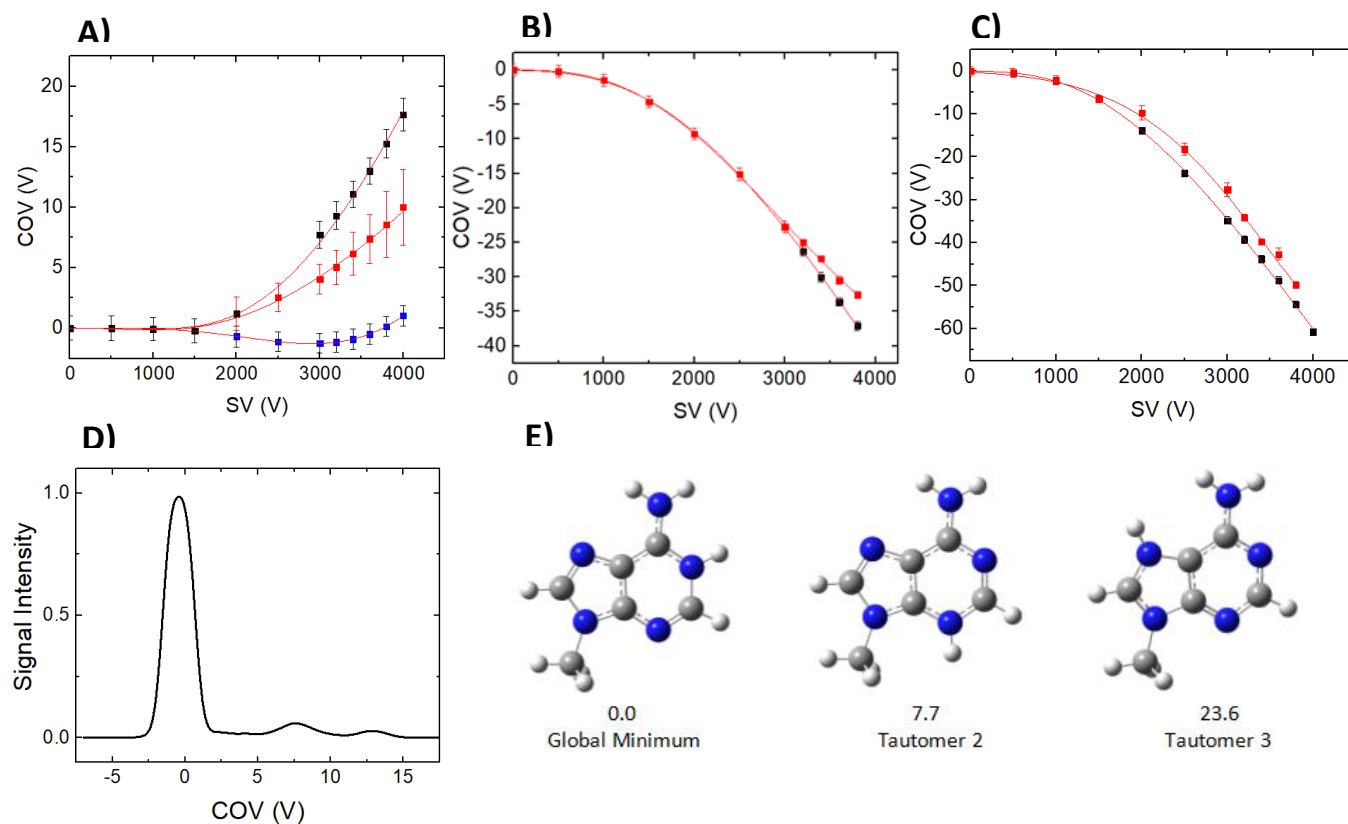
## CCSD(T)//B3LYP Results

Electronic single point energies calculated at the CCSD(T) level of theory, and thermochemical corrections calculated at the B3LYP level of theory using the 6-311++G(d,p) basis set. Calculations employed a temperature of T = 298 K and a pressure of P = 1 atm.

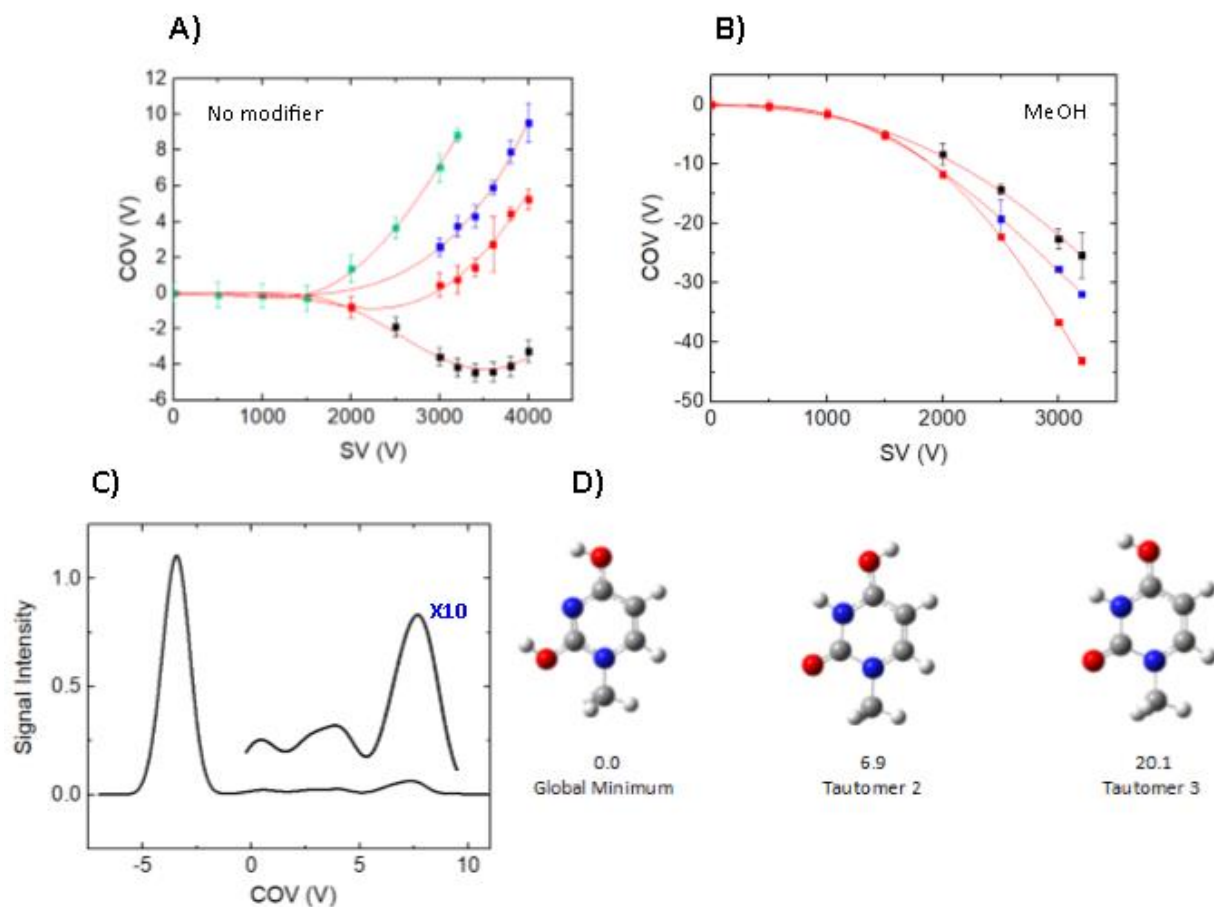
Tautomer	CCSD(T) Electronic Energy (hartree)	Relative ZPE- Corrected Energy (kJ/mol)	Relative Gibbs Energy (kJ/mol)	Relative Enthalpy (kJ/mol)	Entropy (Cal/mol Kelvin)
A1_hb.log	-466.636786	0	0	0	84.055
A1_hc.log	-466.6336084	8.35016	8.23056	8.30076	83.998
A1_hd.log	-466.6223724	35.71776	36.50816	34.51396	85.668
A2_hc.log	-466.6353456	3.58904	3.74244	3.22764	84.471
C1_hb.log	-394.4355744	5.37108	5.78708	4.79648	80.429
C1_hc.log	-394.4378102	0	0	0	79.625
C1_hd.log	-394.4236535	34.77162	35.47622	34.13462	80.712
C3_hb.log	-394.4231373	36.26714	36.74034	36.02274	80.207
G1_ha.log	-541.7317196	18.88146	19.08686	18.95426	90.395
G1_hb.log	-541.7173443	52.33644	54.41904	48.36884	95.186
G1_hf.log	-541.7395277	0	0	0	90.288
G2_hd.log	-541.7326968	16.19514	16.85814	15.61274	91.296
T1_hc.log	-453.4832962	25.5827	26.2171	24.6909	87.765
T1_hd.log	-453.4878143	14.11124	14.65984	13.15704	87.746

T2_ha.log	-453.4802763	32.93524	33.49944	32.43344	87.392
T2_hb.log	-453.4936047	0	0	0	86.528
U1_hc.log	-414.2611868	21.08106	21.66866	20.37906	79.59
U1_hd.log	-414.2661797	8.86132	9.25652	8.28932	79.329
U2_hb.log	-414.2698409	0	0	0	78.546
U3_hd.log	-414.2573088	30.97666	31.23146	30.87786	78.833

## Methylated Nucleobases

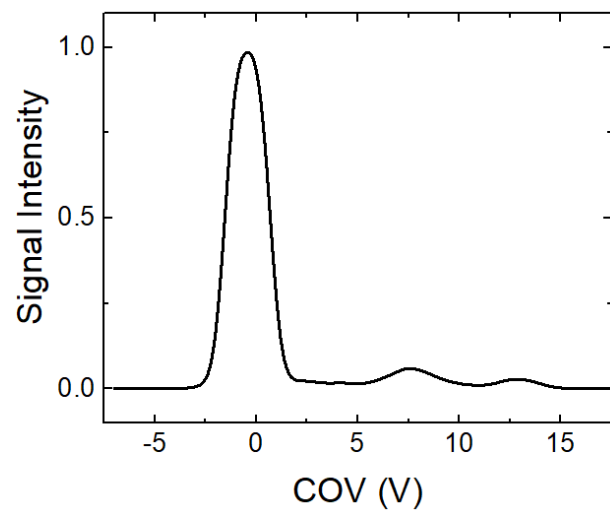


Dispersion plots of  $(\text{MeA} + \text{H})^+$  in (A) dry  $\text{N}_2$ , and an  $\text{N}_2$  environment seeded with 1.5 % (mole ratio) of (B) MeOH vapor, and (C), IPA vapor. (D) The ionogram of  $m/z = 150$  at  $\text{SV} = 3600 \text{ V}$  in dry  $\text{N}_2$ . (E) The three lowest in energy calculated structures of  $(\text{MeA} + \text{H})^+$  using CCSD(T)/6-311++G(d,p).

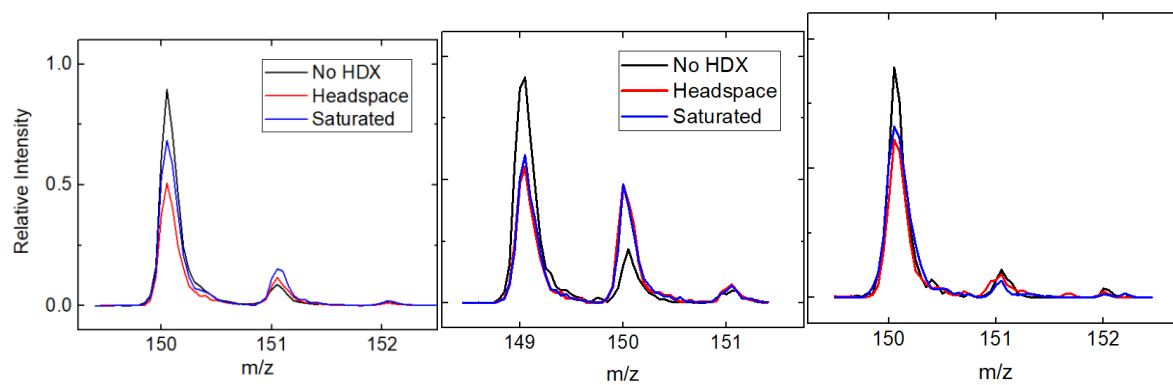


Dispersion plots of  $(\text{MeU} + \text{H})^+$  in (A) dry  $\text{N}_2$ , and an  $\text{N}_2$  environment seeded with 1.5 % (mole ratio) of (B) MeOH vapor (C) The ionogram of  $m/z = 127$  at  $\text{SV} = 3000 \text{ V}$  in dry  $\text{N}_2$ . (E) The three lowest in energy calculated structures of  $(\text{MeU} + \text{H})^+$  using CCSD(T)/6-311++G(d,p).

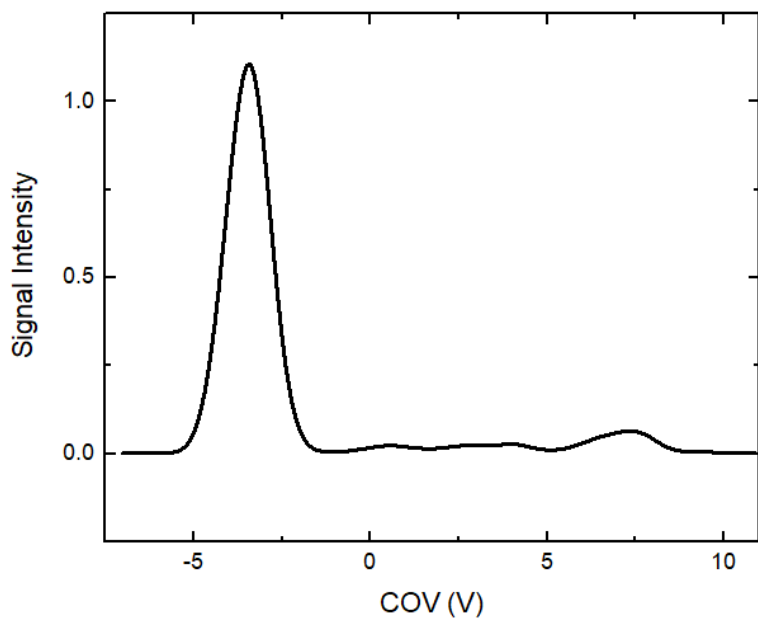
## HDX Profiles



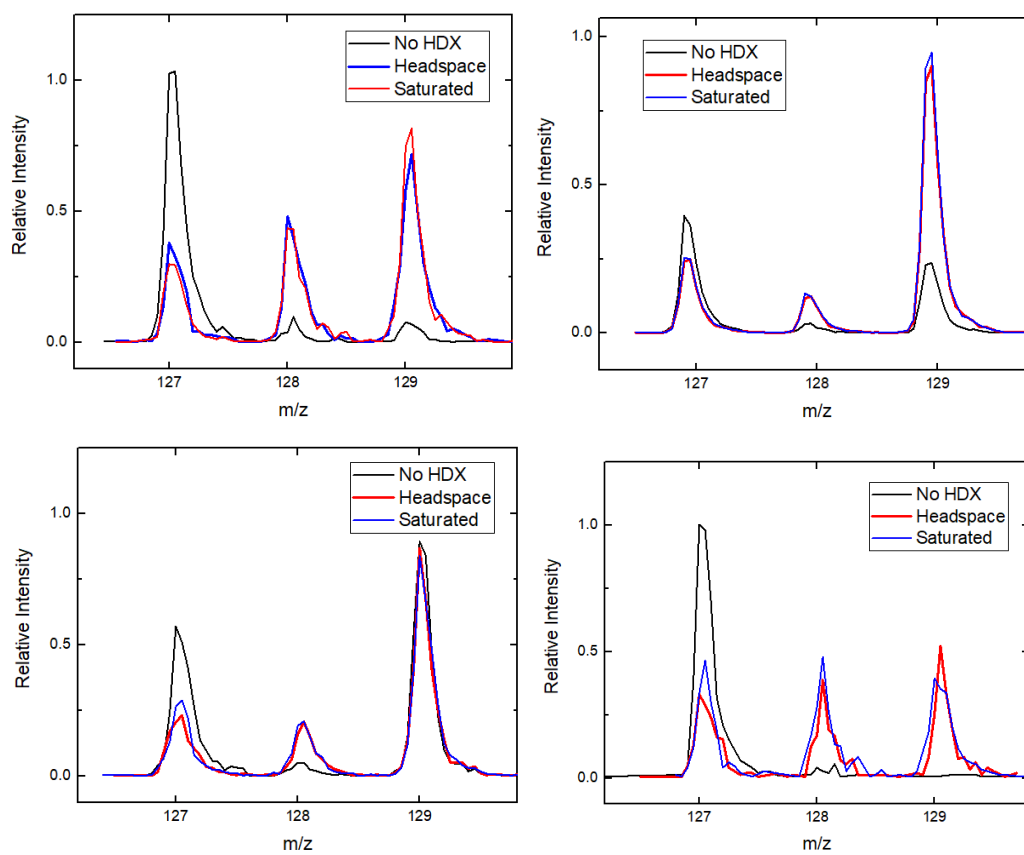
Ionogram of  $(\text{MeA} + \text{H})^+$  at  $\text{SV} = 3600 \text{ V}$  in dry  $\text{N}_2$ .



HDX traces corresponding to (left to right)  $\text{CV} = -0.9 \text{ V}$ ,  $\text{CV} = 7.2 \text{ V}$ , and  $\text{CV} = 13.0 \text{ V}$ . Both a saturated environment (blue) with  $\text{D}_2\text{O}$  and bubbling in the  $\text{D}_2\text{O}$  reagent through the headspace (red) are shown.



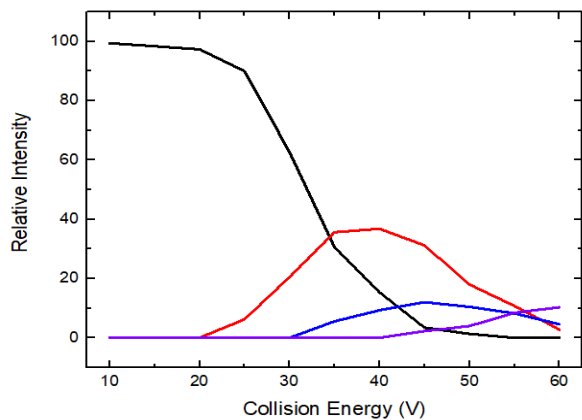
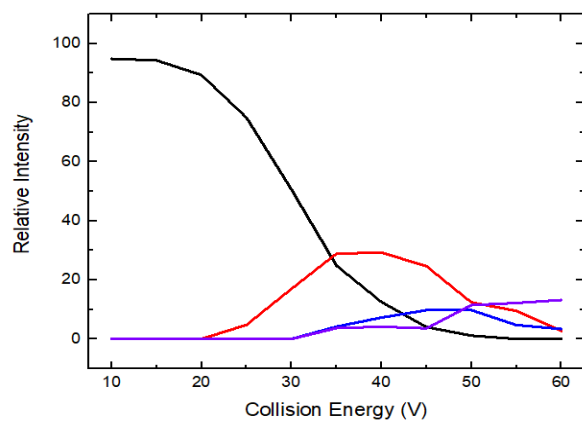
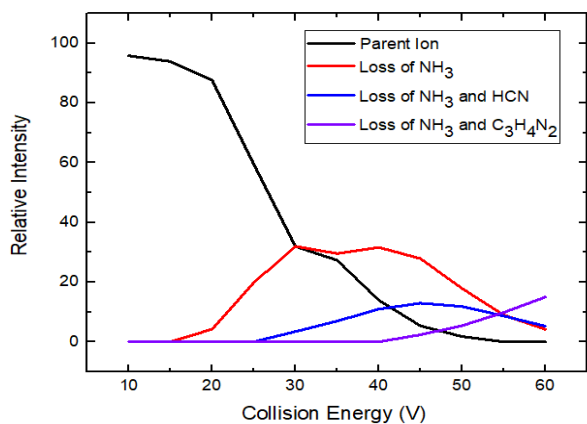
Ionogram of  $(\text{MeU} + \text{H})^+$  at  $\text{SV} = 3000 \text{ V}$  in dry  $\text{N}_2$ .



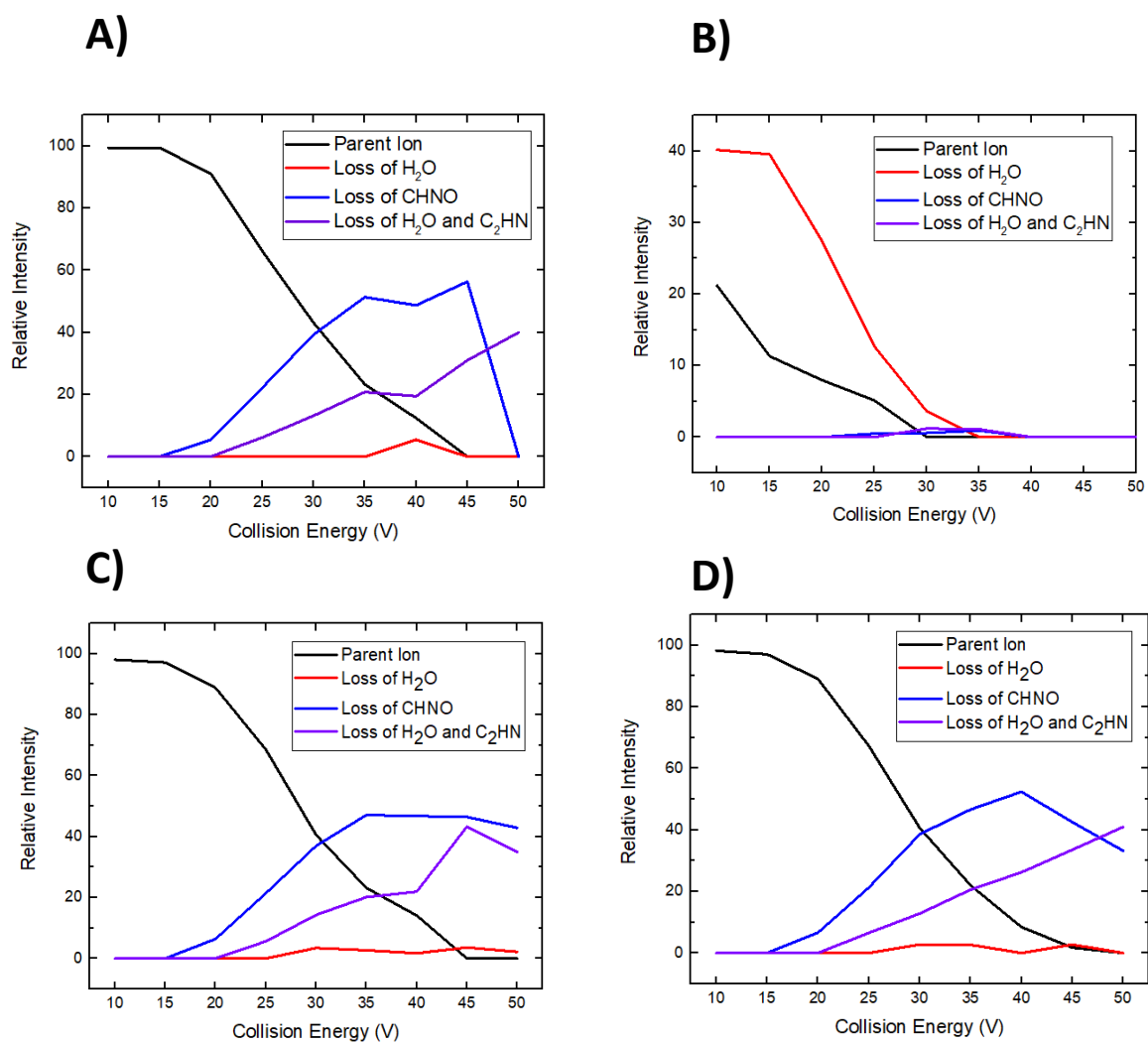
HDX traces corresponding to (top left to bottom right)  $\text{CV} = -3.5 \text{ V}$ ,  $\text{CV} = 0.5 \text{ V}$ ,  $\text{CV} = 3.8 \text{ V}$ , and  $\text{CV} = 7.5 \text{ V}$ . Both a saturated environment (blue) with  $\text{D}_2\text{O}$  and bubbling in the  $\text{D}_2\text{O}$  reagent through the headspace (red) are shown.



## Breakdown Curves

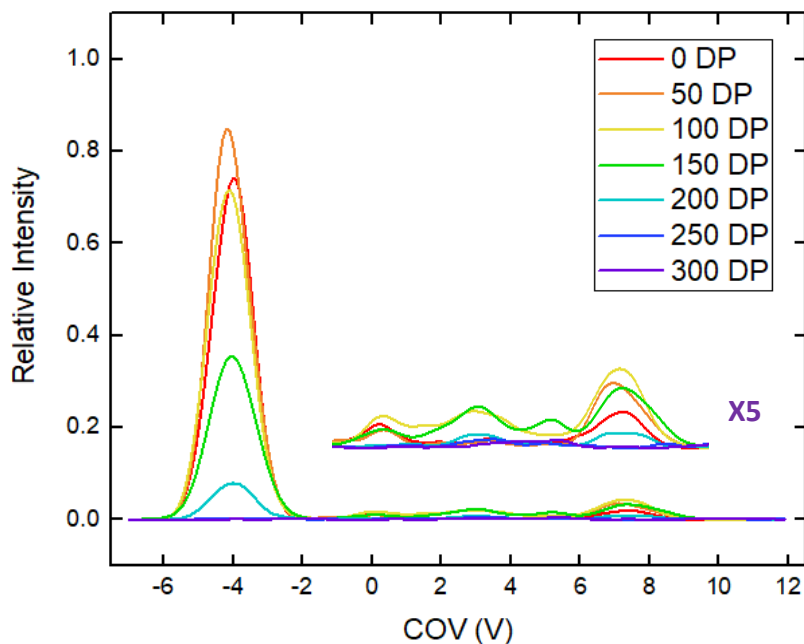
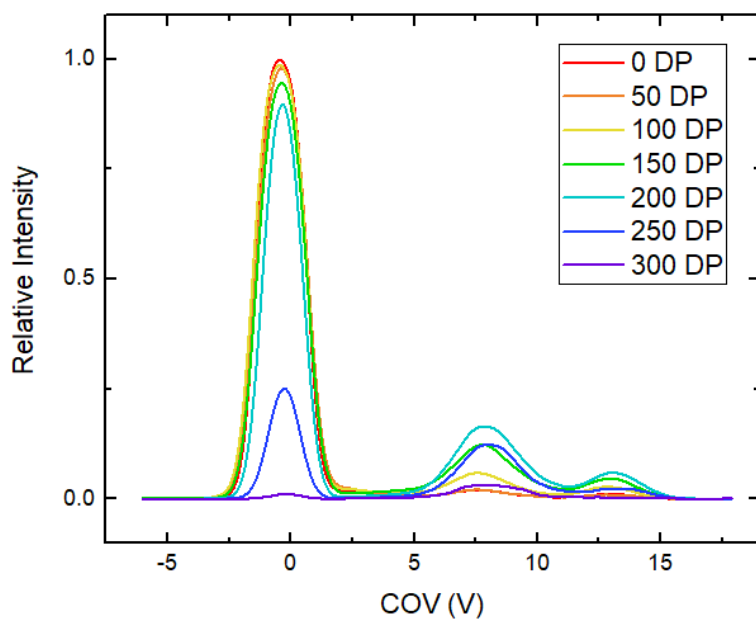


The obtained breakdown curves for (MeA + H)<sup>+</sup> at a CV equal to (top to bottom): -0.9 V, 7.2 V, and 13.0 V, as collision energy was increased from 10 to 60 V in 2.5 V increments.



Breakdown curves obtained for  $(\text{MeU} + \text{H})^+$  at CV equal to (A) -3.5 V, (B) 0.5 V, (C) 3.8 V, and (D) 7.5 V, while ramping the collision energy from 10 V to 50 V in 2.5 V increments.

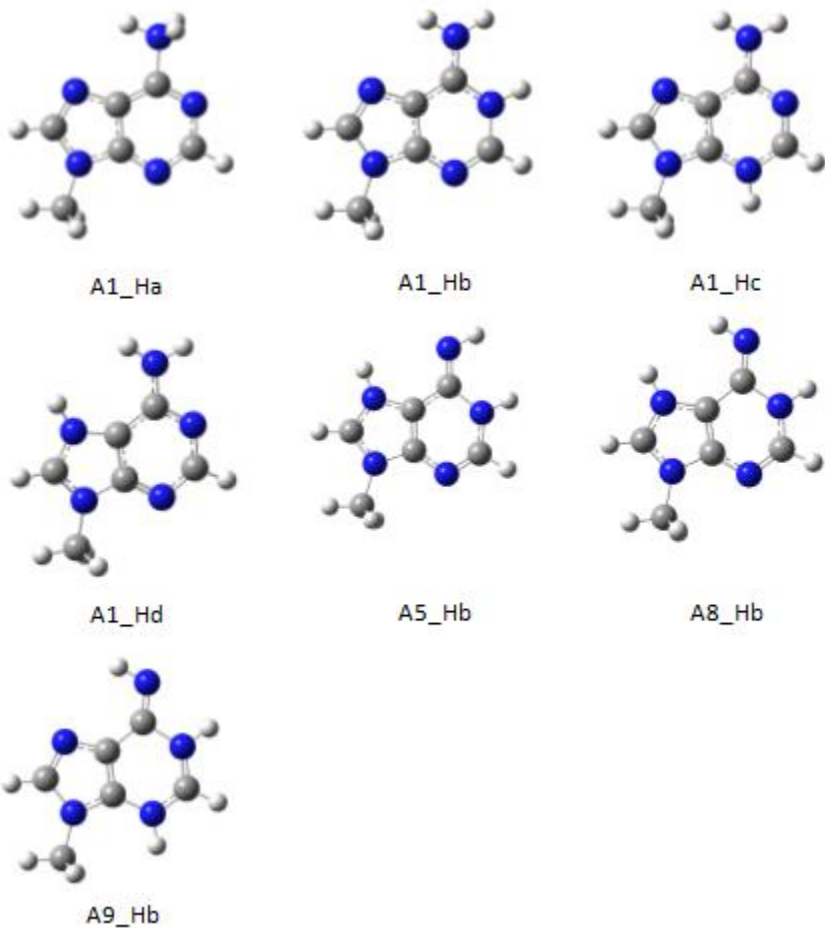
## Declustering Potential Scans



The ionograms obtained for (top) (MeA + H)<sup>+</sup> at SV = 3600 V in a pure N<sub>2</sub> environment and (bottom) (MeU + H)<sup>+</sup> at SV = 3600 V as the declustering potential was varied from 0 V to 300 V in 50 V increments. The CV = -1 V to 10 V is magnified for (MeU + H)<sup>+</sup> to aid in visualization.

## Calculation Results

Thermochemical data for the protonated methylated nucleobases was calculated at the B3LYP/6-311++G(d,p) level of theory at a temperature of  $T = 298$  K and a pressure of  $P = 1$  atm.



Tautomer	Electronic Structure (Hartree)	Gibbs Free Energy (Hartree)	Zero-point Corrected Energy (Hartree)	Enthalpy (Hartree)
A1_Hc	-507.144488	-507.025818	-506.99171	-506.981731
A1_Hd	-507.1370761	-507.0201751	-506.9851651	-506.9747221
A1_Hb	-507.1360223	-507.0166383	-506.9835143	-506.9742033
A5_Hb	-507.1260098	-507.0082148	-506.9731478	-506.9631928
A8_Hb	-507.121498	-507.003684	-506.968751	-506.958784
A1_Ha	-507.1111163	-506.9946143	-506.9579963	-506.9475563
A9_Hb	-507.1067056	-506.9884716	-506.9542066	-506.9442116



G1\_Ha

G1\_Hb

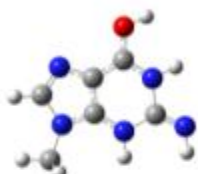
G1\_Hd



G1\_Hf



G4\_Ha



G4\_Hb



G4\_Hf



G8\_Ha

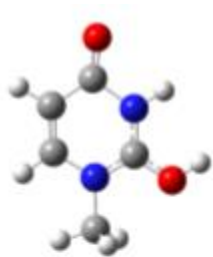


G8\_Hb

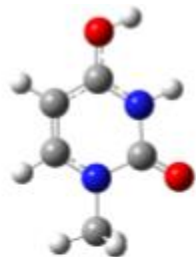


G8\_Hf

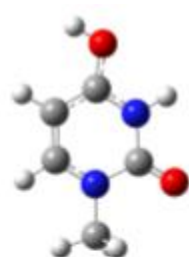
Tautomer	Electronic Structure (Hartree)	Gibbs Free Energy (Hartree)	Zero-point Corrected Energy (Hartree)	Enthalpy (Hartree)
G1_Hf	-582.4163774	-582.2955444	-582.2594054	-582.2479764
G1_Ha	-582.403897	-582.282438	-582.247581	-582.236872
G1_Hb	-582.3895262	-582.2696342	-582.2347962	-582.2242222
G1_Hd	-582.3853929	-582.2663109	-582.2299269	-582.2180379
G8_Hf	-582.3808731	-582.2602171	-582.2242011	-582.2129121
G4_Hf	-582.3771193	-582.2569873	-582.2207663	-582.2093293
G8_Ha	-582.3737458	-582.2535388	-582.2174368	-582.2061968
G4_Ha	-582.3732741	-582.2530291	-582.2169701	-582.2057201
G4_Hb	-582.3624978	-582.2431398	-582.2069158	-582.1954608
G8_Hb	-582.3611943	-582.2422593	-582.2058623	-582.1942973



U1\_Hb



U1\_Hc



U1\_Hd



U2\_Ha



U2\_Hb

Tautomer	Electronic Structure (Hartree)	Gibbs Free Energy (Hartree)	Zero-point Corrected Energy (Hartree)	Enthalpy (Hartree)
U1_Hd	-454.6092274	-454.5147864	-454.4821364	-454.4731914
U2_Hb	-454.6094877	-454.5145317	-454.4821027	-454.4733197
U1_Hc	-454.6043065	-454.5102595	-454.4775315	-454.4684925
U2_Ha	-454.5954996	-454.5013646	-454.4688476	-454.4598826
U1_Hb	-454.5914899	-454.4989139	-454.4652709	-454.4559509
U1_Hd	-454.6092274	-454.5147864	-454.4821364	-454.4731914

## CCSD(T)//B3LYP Results

Electronic single point energies calculated at the CCSD(T) level of theory with thermochemical corrections at the B3LYP level of theory and 6-311++G(d,p) basis set at a temperature of T = 298 K and pressure of P = 1 atm.

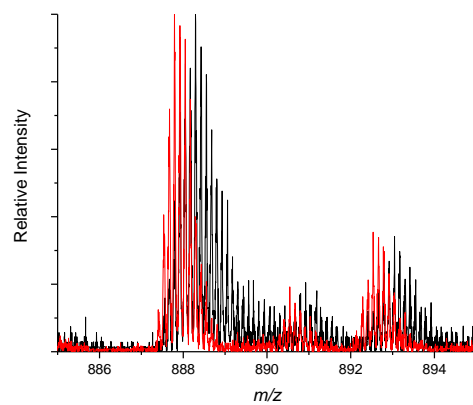
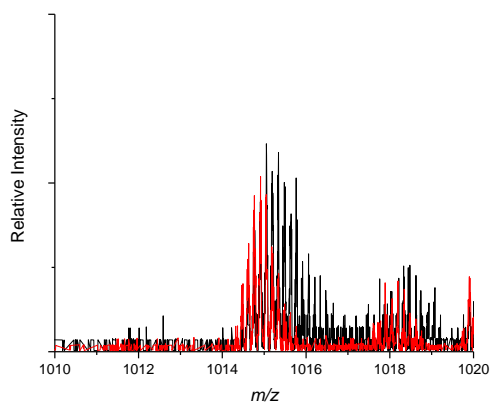
Tautomer	Electronic Energy (CCSD)	Relative Gibbs (kJ/mol)	Relative ZPE (kJ/mol)	Relative Enthalpy (kJ/mol)	Entropy (Cal/mol Kelvin)
A1_Hb	-505.7688293	0	0	0	89.313
A1_Hc	-505.7651855	7.692163645	10.27565565	12.02948965	92.789
A1_Hd	-505.7573324	23.66607321	28.61776621	31.58983222	95.665
G1_Hf	-580.8673	0	0	0	100.115
G1_Ha	-580.8570289	28.61049358	25.24460258	23.35424258	95.903
G1_Hb	-580.8422782	63.22429793	59.80852243	57.56371993	95.577
U2_Hb	-453.4213424	0	0	0	86.738
U1_Hd	-453.4181921	6.91898015	7.49921565	7.92454665	87.542
U1_Hc	-453.4127758	20.10516007	20.89018458	21.56231258	87.907

## APPENDIX B

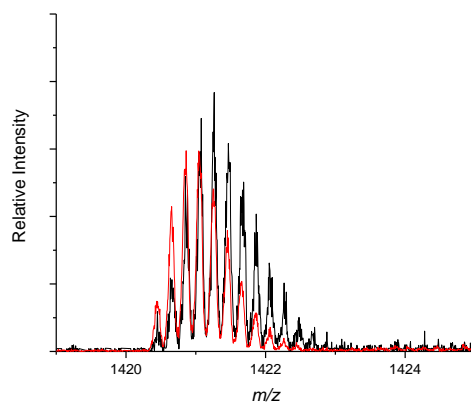
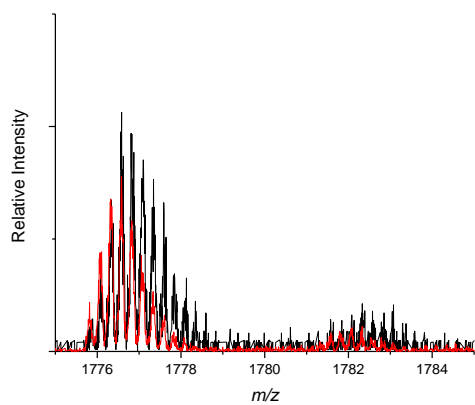
### HDX Spectra

All HDX spectra were ran using the method listed in the main text on the instrument at SCIEX. The red traces indicate no D<sub>2</sub>O present while black represents D<sub>2</sub>O exchange.

#### N<sub>2</sub> -7 state (left) and -8 state (right)

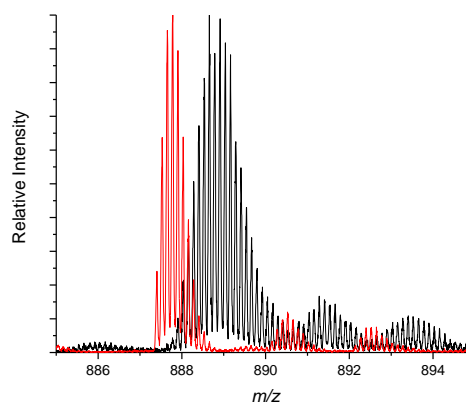
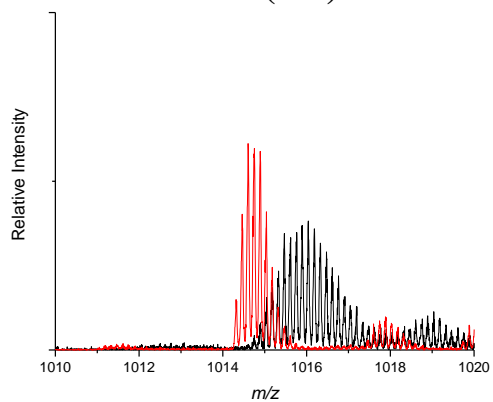


#### IPA -4 state (left) and -5 state (right)

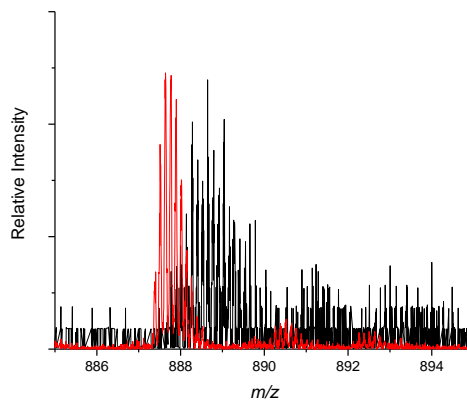
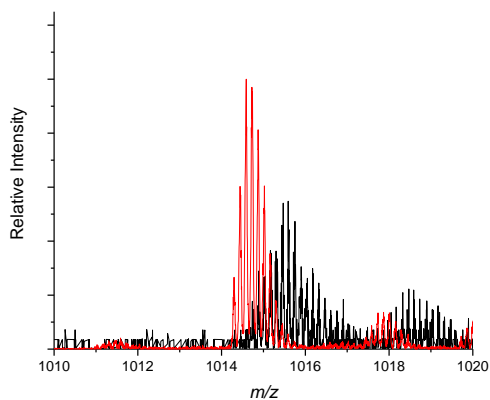




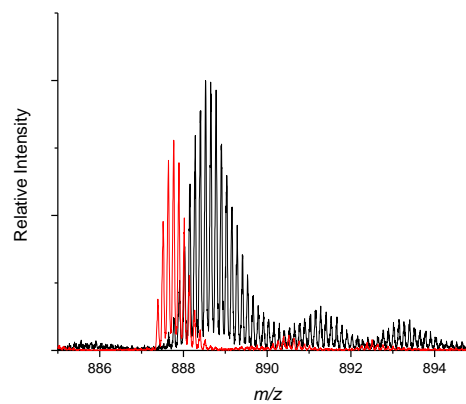
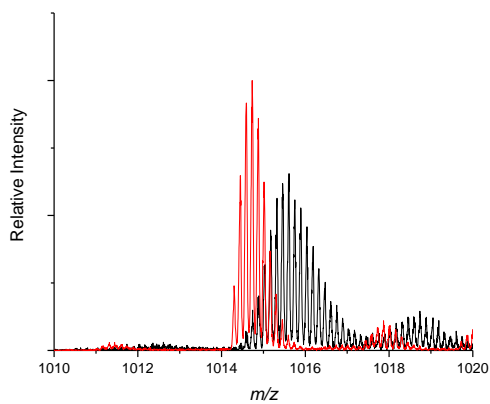
**ACN -7 state (left) and -8 state (right)**



**ACE CV = -0.4 V - 7 state (left) and - 8 state (right)**

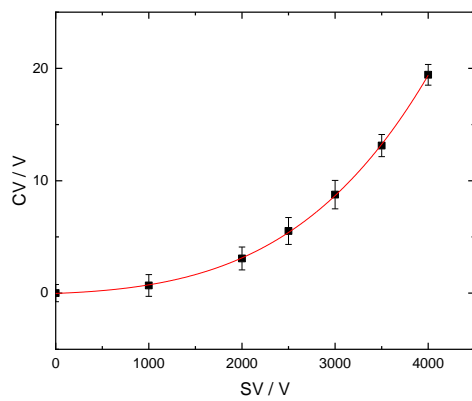
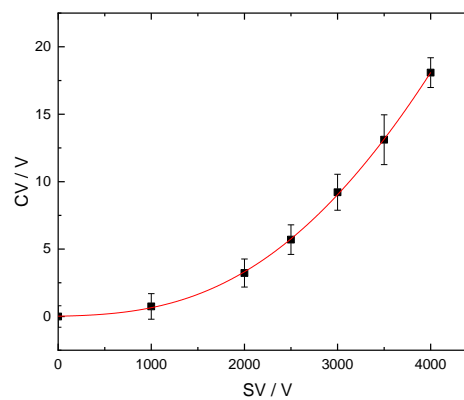
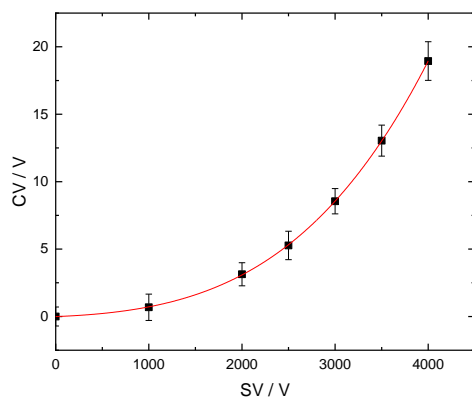


**ACE CE = 4.6 V -7 state (left) and -8 state (right)**

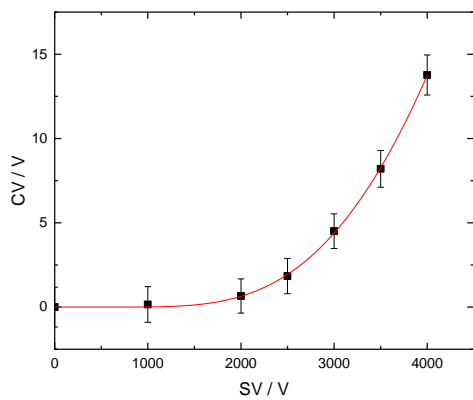
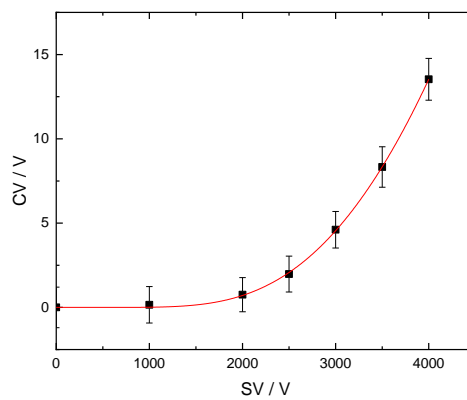
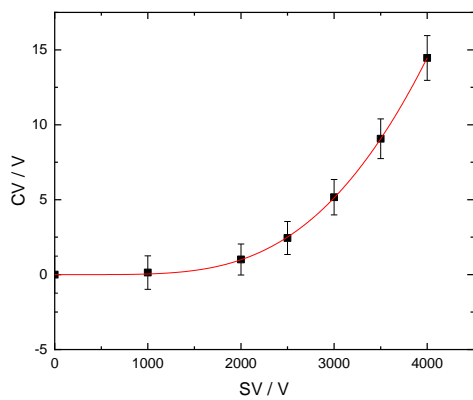


## Dispersion Plots

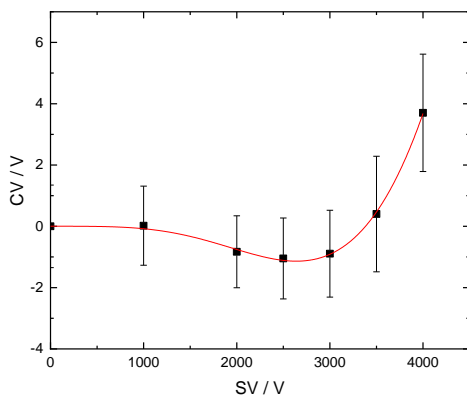
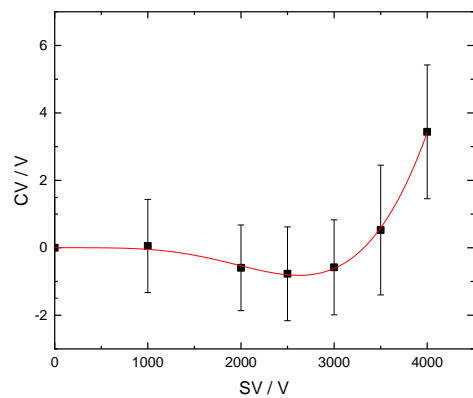
The following dispersion plots were obtained in dry nitrogen at the -5 (top left), -7 (top right), and -8 (bottom left) charge states.



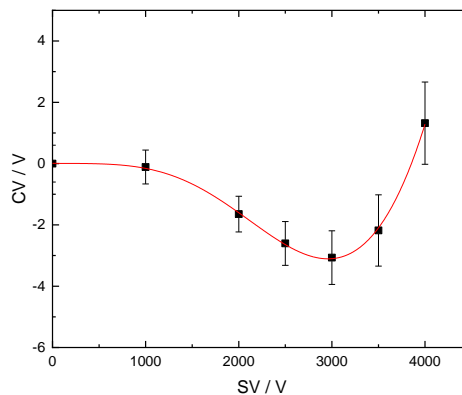
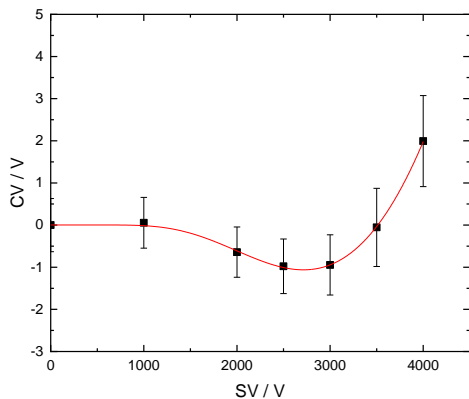
The following dispersion plots were obtained in dry nitrogen seeded with 1.5 % (mole ratio) MeOH at the -5 (top left), -7 (top right), and -8 (bottom left) charge states.



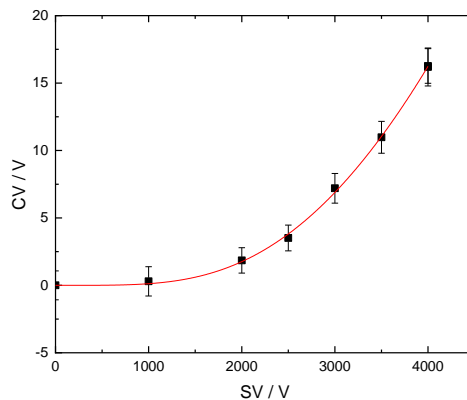
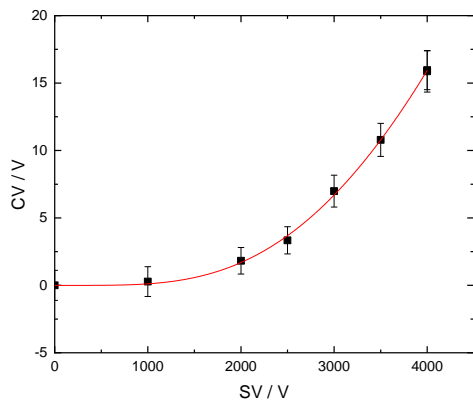
The following dispersion plots were obtained in dry nitrogen seeded with 1.5 % (mole ratio) IPA at the -4 (left), and -5 (right) charge states.



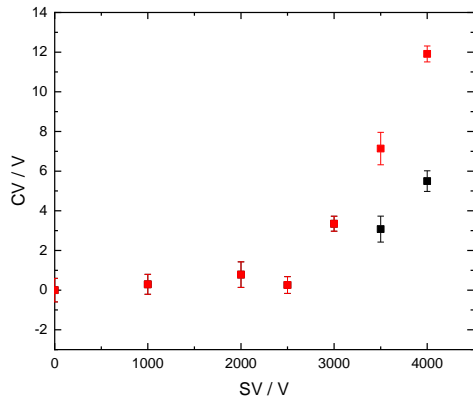
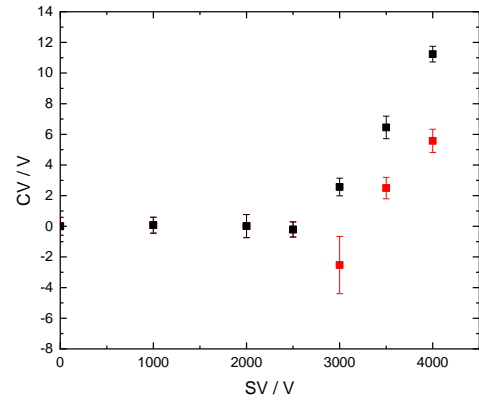
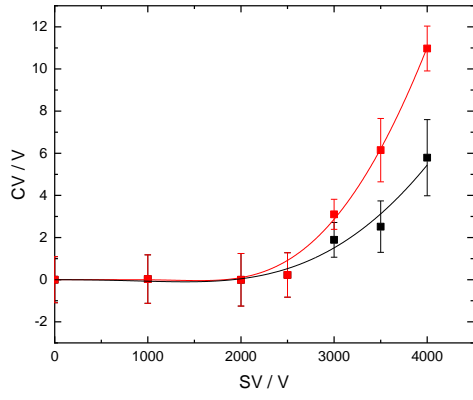
The following dispersion plots were obtained in dry nitrogen seeded with 1.5 % (mole ratio) ACN at the -5 (left) and -7 (right) charge states.



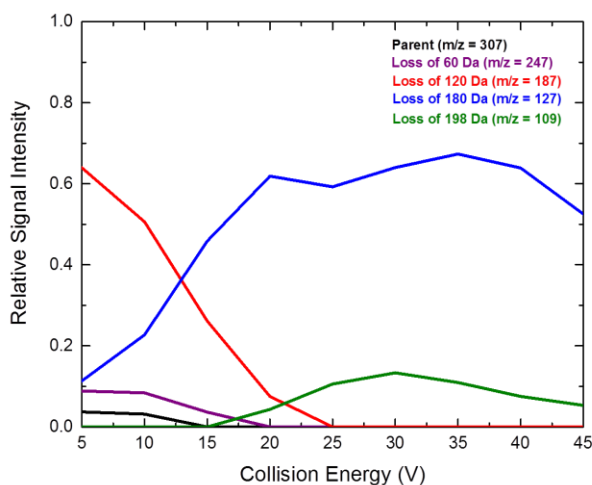
The following dispersion plots were obtained in dry nitrogen seeded with 1.5 % (mole ratio) ethyl acetate at the -5 (left) and -7 (right) charge states.



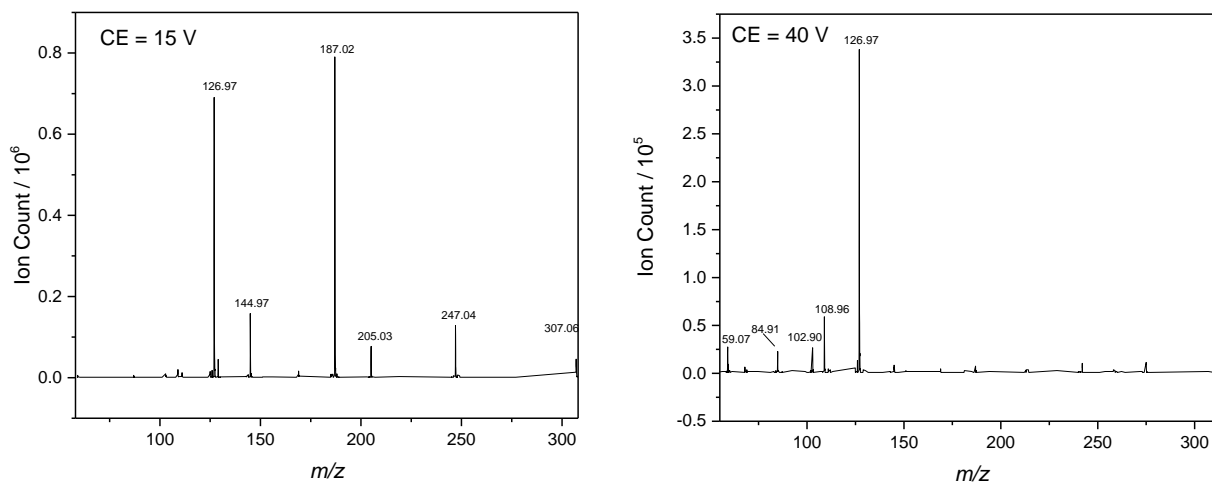
The following dispersion plots were obtained in dry nitrogen seeded with 1.5 % (mole ratio) ACE at the -5 (top left), -7 (top right), and -8 (bottom left) charge states.



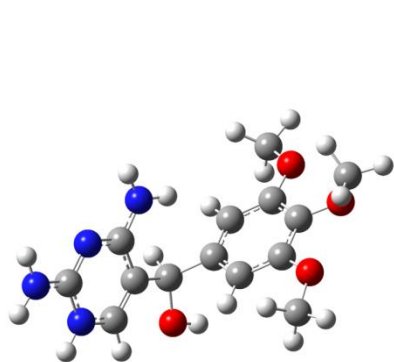
## APPENDIX C



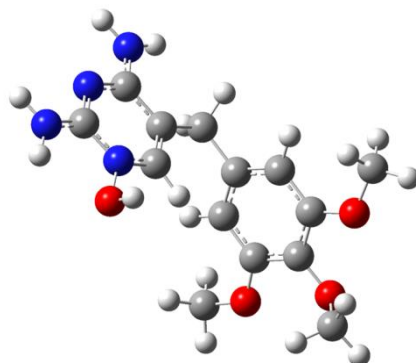
The breakdown curve for the isolated  $m/z$  306.8537 species at SV = 3500 V and CV = -26 V in  $N_2$  seeded with 1.5% (v/v) IPA vapor.



Fragmentation spectra recorded for the isolated  $m/z$  306.8537 species (peak \*) at SV = 3500 V and CV = -26 V in  $N_2$  seeded with 1.5% (v/v) IPA vapor. The fragment ions were recorded at collision energies of 15 V (left) and 40 V (right) and illustrate a suspected IPA cluster contaminant due to loss of 60 Da.

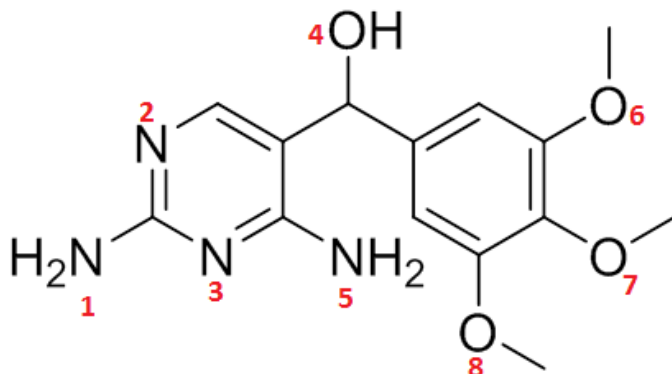


**31.30 kJ/mol**

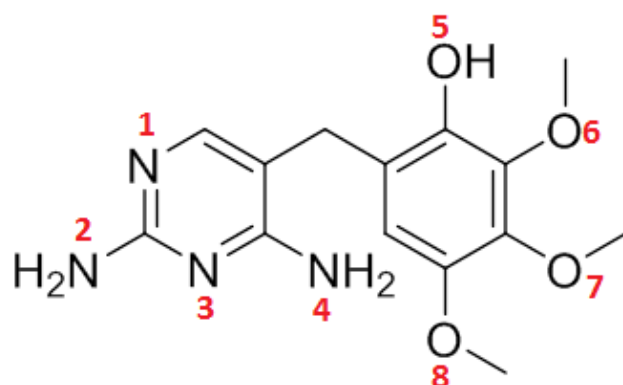


**187.64 kJ/mol**

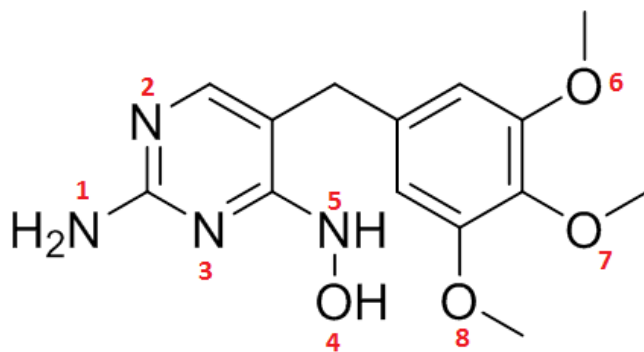
The third and fourth lowest in energy isomers of (TP306 + H)<sup>+</sup> as calculated at the B3LYP/6-311G++(d,p) level of theory relative to the reported global minimum structure. Standard Gibbs' energies are reported.



The methylene-hydroxylated isomer and protonation sites investigated computationally for this study. The isomer above is TPa and each protonated site is indicated numerically (*i.e.*, protonation on N<sub>2</sub> would be named TPa2).

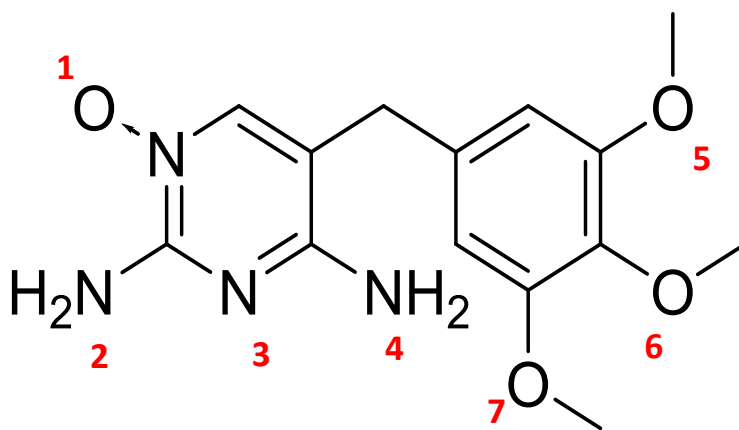


The methoxy-containing ring hydroxylated isomer and protonation sites investigated computationally for this study. The isomer above is TPb and each protonated site is indicated numerically (*i.e.*, protonation on N<sub>2</sub> would be named TPb2).

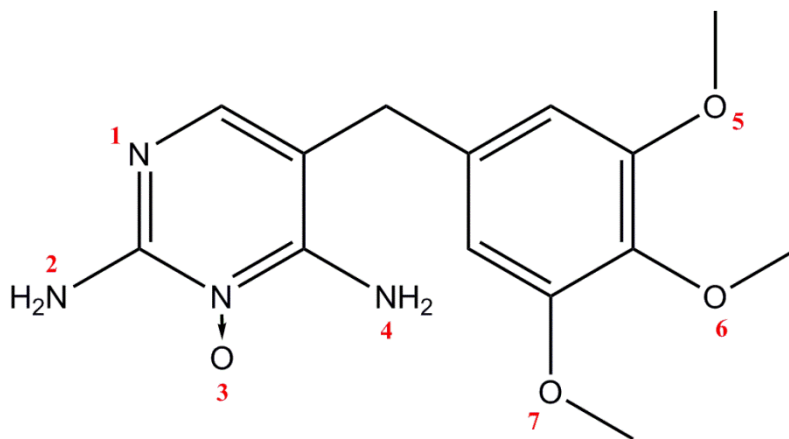


The NH-hydroxylated isomer and protonation sites investigated computationally for this study. The isomer above is TPc and each protonated site is indicated numerically (*i.e.*, protonation on N<sub>2</sub> would be named TPc2).





The first N-oxide isomer and protonation sites investigated computationally for this study. The isomer above is TPNa and each protonated site is indicated numerically (*i.e.*, protonation on N<sub>2</sub> would be named TPNa2).



The second N-oxide isomer and protonation sites investigated computationally for this study. The isomer above is TPNb and each protonated site is indicated numerically (*i.e.*, protonation on N<sub>2</sub> would be named TPNb2).

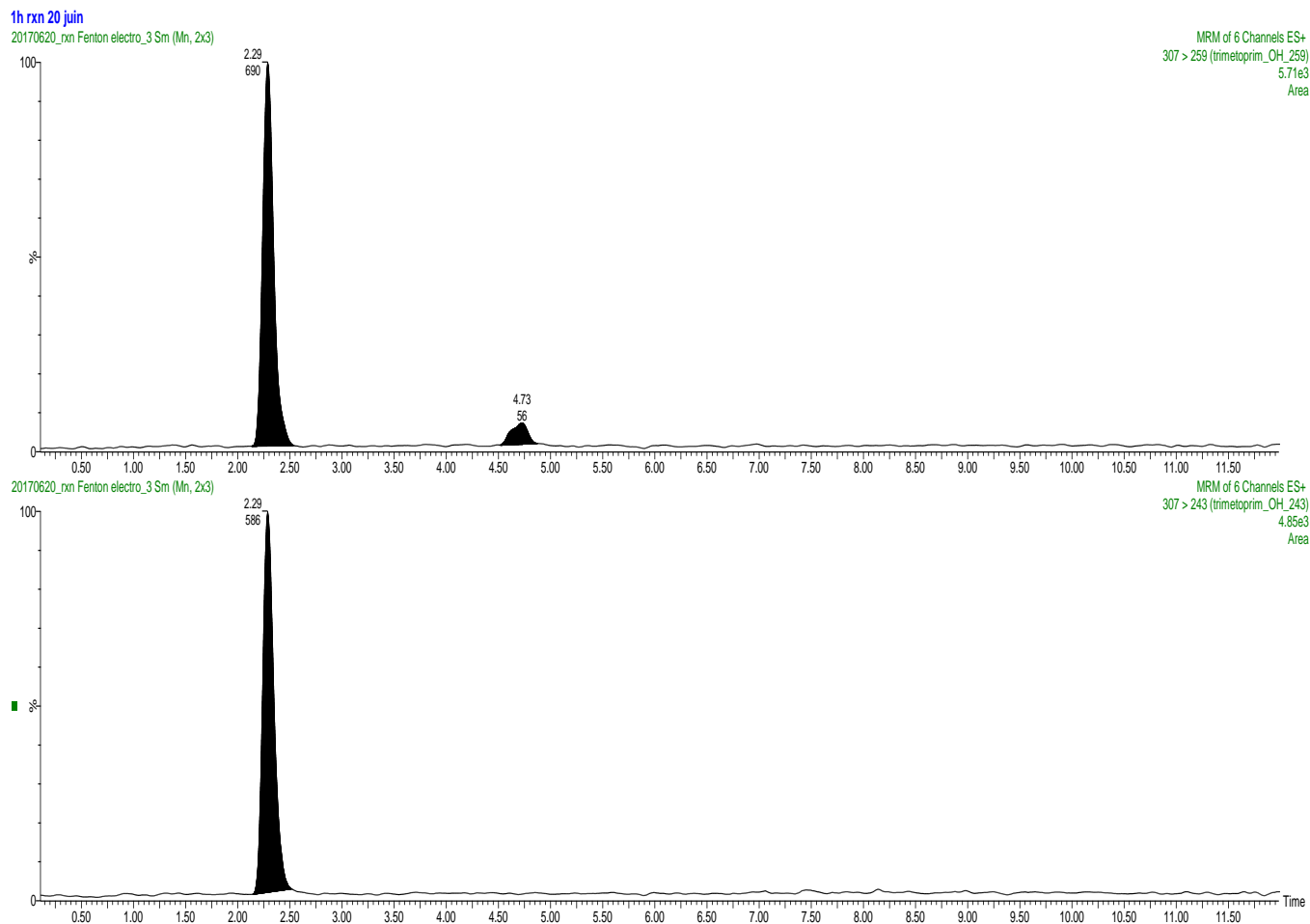
## Calculation Results

Thermochemical data for each (TP306 + H)<sup>+</sup> isomer calculated at B3LYP/6-311++G(d,p) level of theory at a temperature of T = 298 K and a pressure of P = 1 atm. For IPA calculations, a geometrically optimized IPA molecule was placed at the site of TP protonation and no proton transfer to the solvent was observed following optimization and energetic calculations. From the main text nomenclature, Isomer I corresponds to TPa4 and Isomer II corresponds to TPb1.

Isomer	Electronic Energy (hartree)	Zero Point Corrected Energy (hartree)	Gibbs Energy (hartrees)	Enthalpy (hartree)
TPb1	-1064.917814	-1064.583781	-1064.635408	-1064.560477
TPa4	-1064.907048	-1064.576631	-1064.631512	-1064.551646
TPa2	-1064.904385	-1064.571172	-1064.623488	-1064.54758
TPb3	-1064.904683	-1064.571483	-1064.623353	-1064.547973
TPa3	-1064.887042	-1064.554675	-1064.607098	-1064.530876
TPb4	-1064.88565	-1064.551257	-1064.602655	-1064.528285
TPb2	-1064.882462	-1064.547231	-1064.599105	-1064.524092
TPa1	-1064.866983	-1064.53277	-1064.585527	-1064.50925
TPa5	-1064.864171	-1064.53048	-1064.58355	-1064.506974
TPa7	-1064.853052	-1064.519738	-1064.57187	-1064.496184
TPb6	-1064.849533	-1064.517185	-1064.569676	-1064.493454
TPb8	-1064.849784	-1064.516749	-1064.568925	-1064.493319
TPb7	-1064.849614	-1064.516613	-1064.568831	-1064.493054
TPNa1	-1064.843675	-1064.511282	-1064.563941	-1064.487857
TPa8	-1064.840018	-1064.507872	-1064.560781	-1064.48404
TPa6	-1064.838078	-1064.505985	-1064.558983	-1064.482143
TPb5	-1064.833701	-1064.502715	-1064.556163	-1064.478869
TPc2	-1064.833583	-1064.500797	-1064.553984	-1064.477146
TPNb4	-1064.830441	-1064.498263	-1064.550428	-1064.475005
TPNb1	-1064.829541	-1064.497438	-1064.55001	-1064.474022
TPNb2	-1064.829541	-1064.497437	-1064.550006	-1064.474022
TPc4	-1064.811576	-1064.48429	-1064.540699	-1064.458716
TPc3	-1064.813732	-1064.481595	-1064.535249	-1064.45778
TPNa3	-1064.808584	-1064.478226	-1064.53189	-1064.454114
TPNb3	-1064.80174	-1064.469237	-1064.521483	-1064.446151
TPc1	-1064.797525	-1064.463607	-1064.517399	-1064.440065
TPNb6	-1064.793332	-1064.461007	-1064.513756	-1064.437592
TPNa4	-1064.792285	-1064.459658	-1064.51333	-1064.436175
TPNa6	-1064.79059	-1064.45853	-1064.511712	-1064.435004
TPc7	-1064.790015	-1064.456811	-1064.509931	-1064.433316
TPa2 (+IPA)	-1259.356684	-1258.913628	-1258.977385	-1258.883156
TPb1 (+IPA)	-1259.3695	-1258.925523	-1258.988322	-1258.895414
TPb3 (+IPA)	-1259.358125	-1258.914862	-1258.977667	-1258.884583
TPc2 (+IPA)	-1259.286612	-1258.843901	-1258.908376	-1258.81342

## Liquid Chromatography-Quadrupole-Time-of-Flight Mass Spectrometry (LC-QqTOF-MS)

These experiments used a liquid chromatograph manufactured by Shimadzu (Japan) composed of a Nexera LC-30AD pump module, a SIL-30AC autosampler and a CTO-30A as column oven module. This LC system was coupled to a high-resolution mass spectrometer, the Maxis quadrupole-time-of-flight mass spectrometer (QqTOFMS) made by Bruker (USA) equipped with an electrospray ionization (ESI) source operated in the positive mode. The chromatographic column was a Acquity HSS T3 (2.1×50mm, 1.8  $\mu$ m) from Waters (USA). The column temperature was 30 °C. The mobile phase was composed of solvent A (0.1% formic acid in water) and solvent B (0.1% formic acid in acetonitrile). The flow rate was 0.5 mL/min. The chromatographic separation used the following elution gradient described here as %B in the mobile phase as a function of time: 0 min (5%), 6 min (10%), 7 min (98%), 9 min (98%), 10 min (5%), 12 min (5%). Injection volume was 5  $\mu$ L. The QqTOFMS was calibrated with a sodium formate solution. The mass drift was monitored and all analyses were done within 4 h of the calibration. No lock mass solution was used. In these conditions, full width at half-maximum mass resolution (RFWHM) at  $m/z$  307 was about 20 000.



**The LC/MS MRM monitoring and elution of the  $m/z = 307$  species characterized in this experiment.**

**For a list of all XYZ coordinate files, please refer to the published Supplementary Information document found here : [10.1021/acs.analchem.8b00484](https://doi.org/10.1021/acs.analchem.8b00484)**