

**Structures and Energetics of Guanine Tetrads and
Quadruplexes Isolated in the Gas Phase**

By

Mohammad Azargun

A thesis submitted to the School of Graduate Studies
in partial fulfillment of the requirement for the degree of
Ph.D. of Science in Chemistry

Department of Chemistry
Memorial University of Newfoundland
September 2017

St. John's

Newfoundland and Labrador

Abstract

Structures, thermochemistries, reactivities, and kinetics of chemical systems in the gas phase can be determined by a variety of ion activation techniques. The internal energy of ions can be increased by colliding the accelerated ions with a target gas, collision-induced dissociation (CID). This technique can lead to a full or partial structural identification, and also to the relative stabilities of the ion clusters. Ions in the gas phase may reach a dissociation threshold by absorbing multiple photons from the laser radiation if the laser frequency is resonant with a vibrational mode which is called infrared multiphoton dissociation (IRMPD). This technique turns into IRMPD spectroscopy by varying the laser frequency and recording the resultant fragments as a function of laser frequency. Blackbody infrared radiative dissociation (BIRD) is another photodissociation technique used in this work that is well suited to slowly activate weakly-bound non-covalent interactions in order to determine their thermochemistries and also gain kinetic insights. Theoretical techniques have been used along with these activation techniques as a powerful complementary tool.

This work provides physical chemistry insights about guanine tetrads and quadruplexes in the gas phase. The first phase of this research focuses on the alkali metal G-tetrads, $M(9\text{-eG})_4^+$ ($M=\text{Li, Na, K, Rb, Cs}$; $9\text{eG} = 9\text{-ethylguanine}$). After their gas phase formation, their structures, stabilities, and energetics were examined by a combination of SORI-CID, IRMPD spectroscopy, and computational chemistry. The role of hydrogen bonding, the size of cation, and electrostatic interactions in the stability and structures were also examined. The formation of alkali metal G-quadruplexes, $(M(9\text{eG}))_8^+$ ($M = \text{Na,}$

K, Rb, Cs), was pursued in the next phase. The BIRD technique, in conjunction with master equation modeling, was applied to determine the binding energies and transition state nature of these G-quadruplexes. IRMPD spectroscopy, along with theoretical techniques, was used for structural elucidation. In the final step, gas phase formation of larger G-quadruplexes was assessed. The $K_2(9\text{-ethylguanine})_{12}^+$ quadruplex was successfully generated and isolated in the gas phase. BIRD and SORI-CID techniques were used to determine the quadruplex binding energies and gas phase stability, respectively. The structure of this quadruplex was determined by IRMPD spectroscopy and computational chemistry.

Acknowledgments

Foremost, I would like to express my sincere gratitude to my supervisor Prof. Travis D. Fridgen for the continuous support of my Ph.D study and related research, for his patience, guidance, motivation, and immense knowledge. It was a great fortune to have him as supervisor and mentor for my Ph.D study.

I would like to thank my supervisory committee Dr. Chris Flinn and Dr. Sunil Pansare for their great help and insightful comments.

My special thanks also goes to all the staff at chemistry department for their help and support during my Ph.D at the chemistry department.

Finally, I must express my very profound gratitude to my dearest parents, Sedighe (RIP) and Ahmad, for their endless support and love. My thanks also go to my brothers and sisters who love me unconditionally. Last but not least, I wish to express my thanks and love to my wife, Haniyeh ♥, for providing me with continuous support and encouragement throughout these years.

Table of Contents

Structures and Energetics of Guanine Tetrads and Quadruplexes Isolated in the Gas Phase.....	I
Abstract.....	II
Acknowledgments	IV
Table of Contents	V
List of Figures.....	X
List of Tables	XXIII
List of Schemes.....	XXIV
List of Appendices.....	XXV
List of Abbreviation and Terms	XXVI
Chapter 1. Introduction	1
1.1. Physical Chemistry Analysis of Gas Phase Ions via Mass Spectrometry	1
1.2. CID, IRMPD and BIRD for Biological Systems.....	13
1.2.1 CID Applications for Biological Systems.....	13
1.2.2 IRMPD Spectroscopy Applications for Biological Systems	18
1.2.3 BIRD Applications for Biological Systems	33
1.3 Objectives of Thesis	39

References	41
Chapter 2. Experimental and Theory	52
2.1. Instrumentation.....	52
2.1.1 Electrospray Ionization	54
2.1.2. Fourier Transform Ion Cyclotron Resonance Mass Spectrometry (FT-ICR MS).....	56
2.2. Tandem Mass Spectrometry Techniques.....	62
2.2.1. Collision Induced Dissociation	62
2.2.2. Sustained Off-Resonance Irradiation-Collision Induced Dissociation (SORI-CID).....	65
2.2.3 Blackbody Infrared Radiative Dissociation	66
2.2.4. Infrared Multiple Photon Dissociation Spectroscopy	69
2.2.4.1 IRMPD Mechanism	71
2.2.4.2 Optical Parametric Oscillators (OPO).....	74
2.2.4.3 Free Electron Laser (FEL)	75
2.3. Computational Chemistry.....	76
2.3.1. <i>ab initio</i> Methods	78
2.3.2 Density Functional Theory.....	79
2.3.3. Basis Sets	81
2.3.4. Calculation Approach.....	82

2.3.5. Master Equation Modeling.....	83
References	86
Chapter 3. Guanine Tetrads: An IRMPD Spectroscopy, Energy Resolved SORI-CID, and Computational Study of M(9-ethylguanine)₄⁺ (M=Li, Na, K, Rb, Cs) in the Gas Phase.....	96
3.1. Introduction	97
3.2. Methods	101
3.2.1. Experimental	101
3.2.2. Computational	103
3.3. Results and Discussion	103
3.3.1 Computed Structures of Metalated G-Tetrads, M(9eG) ₄ ⁺	105
3.3.2 IRMPD Spectroscopy of M(9eG) ₄ ⁺	107
3.3.3 Relative Gas-Phase Stabilities of the M(9eG) ₄ ⁺ Tetrads by Energy-Resolved SORI-CID	110
3.4. Conclusions	114
References	116
Chapter 4. The Intrinsic Stabilities and Structures of Metal Cationized Guanine Quadruplexes.....	121
4.1. Introduction	122

4.2. Methods	126
4.2.1. Experimental	126
4.2.2. Computational	127
4.2.2.1. Geometries and Binding Energies	127
4.2.2.2. Master Equation Modeling	127
4.3. Results and Discussion	129
4.3.1. Structures of G-Quadruplexes: IRMPD Spectroscopy	129
4.3.2. BIRD Experiments and Activation Energies of G-quadruplexes	132
4.3.3. Theoretical Dissociation Enthalpies and Gibbs Energies	139
4.4. Conclusions	140
References	142
Chapter 5. Intrinsic properties of $K_2(9\text{-ethylguanine})_{12}^{2+}$ by BIRD, IRMPD Spectroscopy, and Energy Resolved SORI-CID	148
5.1. Introduction	149
5.2. Methods	153
5.2.1. Experimental	153
5.3. Results and Discussion	156
5.3.1. Unimolecular Dissociation	157
5.3.2. IRMPD Spectroscopy $K_2(9eG)_{12}^{2+}$	158

5.3.3. Relative Gas-Phase Stabilities of Potassium Cationized Quadruplexes by Energy-Resolved SORI-CID	160
5.3.4. BIRD Experiments	162
5.4. Conclusion	165
References	166
Chapter 6. Summary and Outlook	172
6.1. Summary	172
6.2. Future Work	177
References:	180
Appendix 1	182
Appendix 2	187

List of Figures

- Figure 1.1.** ESI mass spectrum of myoglobin protein (10 μM in 10 mM ammonium acetate, pH 6.9) featuring two protein charge states (8^+ and 9^+). This shows how ESI can detect very large biomolecules by generating multi-charged species. It must be noted that due to the intact precursor ionization, structural and sequencing information is not extractable, and this requires additional mass stages, tandem mass spectrometry, to form fragment ions. Reprinted with permission from *Int. J. Mass Spectrom* **2000**, 200, 175-186. Copyright (2000) Elsevier Ltd.¹⁰ 3
- Figure 1.2.** A schematic of a tandem in space mass spectrometer, triple quadrupole. Selected precursor ions in MS1 are transferred to the collision cell, Quad 2, to undergo dissociation. The MS2 mass scans the fragment ions. 4
- Figure 1.3.** CID (a) and SID (b) mass spectra of D^{11+} (the 11^+ cytochrome C dimer) under the same conditions (laboratory collision energy CE and voltage difference ΔV). Tandem mass spectra depict different monomer products and dissociation pathways. Reprinted with permission from *J. Am. Chem. Soc* **2006**, 128, 15044-15045. Copyright (2006) American Chemical Society.³⁶ 7
- Figure 1.4.** SID ion activation mechanism might favor a dissociation pathway. In case of complex proteins, stepwise ion activation, CID, enables unfolding of subunits (A pathway) while a sudden activation, SID, can lead to a different transition state and products. Reprinted with permission from *J. Am. Soc. Mass Spectrom.* **2008**, 19, 190-208. Copyright (2008) American Society for Mass Spectrometry.³⁷ 7

Figure 1.5. Experimental IRMPD of $[\text{Pb}(\text{dCMP})\text{-H}]^+$ compared with B3-LYP/6-31+G(d,p) calculated spectra of three isomers (left panel). DFT calculated structures of some lowest energy isomers are depicted at right panel. The experimental spectrum is in good agreement with the lowest energy structure, dCytS6, in which the metal interacts with both the phosphate group and the carbonyl group of the nucleobase moiety. Reprinted and adopted with permission from *Int. J. Mass spectrom* **2011**, 304, 154-164. Copyright (2010) Elsevier B.V.⁴¹ 11

Figure 1.6. A general schematic of structural analysis of polypeptides by (a) CID and (b) ECD techniques. While removal of labile groups largely occurs in CID, ECD yields complete sequencing by keeping these vulnerable bonds. Reprinted with permission from *Curr. Opin. Biotechnol.* **2004**, 15, 12-16. Copyright (2003) Published by Elsevier Ltd.¹⁹ 13

Figure 1.7. Electrospray mass spectrum of an aqueous solution of uracil and CuCl_2 Reprinted with permission from *ChemPhysChem* **2012**, 13, 588-596. Copyright (2012) John Wiley and Sons.⁵⁶ 15

Figure 1.8. Gibbs energy diagram (298 K) of the $[\text{Cu}(\text{Ura-H})(\text{Ura})]^+$ fragmentation scheme attained with coupling SORI-CID and computational chemistry studies. It is shown that the fragmentation process initiates with an isomerization of the lowest energy structure. Reprinted with permission from *ChemPhysChem* **2012**, 13, 588-596. Copyright (2012) John Wiley and Sons.⁵⁶ 15

- Figure 1.9.** Notation of the possible dissociation pathways and ion products of a protonated peptide by Roepstorff and Fohlman nomenclature. a, b and c are indicative of fragment ions with the positive charge retained on the N-terminus while z, x, y ion fragments have the positive charge on the carboxy termination. Reprinted with permission from *Biochim. Biophys-Proteins and Proteomics*. **2006**, 1764, 1811-1822. Copyright (2006) Elsevier Ltd.⁶¹ 17
- Figure 1.10.** A comparison between a) CID and b) ETD tandem mass spectrum of O-GlcNAc containing peptides (1 pmol/μL). The CID spectra only includes b/y ion pairs resulted from peptide bond cleavages while a perfect set of c and z ions are obtained by ETD. Reprinted with permission from Reprinted with permission from *Biochim. Biophys-Proteins and Proteomics*. **2006**, 1764, 1811-1822. Copyright (2006) Elsevier Ltd.⁶¹ 17
- Figure 1.11.** Comparison of IRMPD spectra of [Pb(Aa-H)]⁺ complexes. Except Lys which has a basic side chain, other five amino acids exhibit similar bands. Reprinted with permission from *J. Phys. Chem. B* **2011**, 115, 11506-11518. Copyright (2011) American Chemical Society.⁷⁵ 19
- Figure 1.12.** Three main possible attachment of Pb²⁺ in to amino acids in [Pb(Aa-H)]⁺ complexes denoted as type A, B and C.. Reprinted with permission from *J. Phys. Chem. B* **2011**, 115, 11506-11518. Copyright (2011) American Chemical Society.⁷⁵ .20
- Figure 1.13.** IRMPD spectra of three amino acid complexes and their comparisons to the computed IR spectra of type A, B and C complexes shown in Figure 1.12.

Reprinted with permission from <i>J. Phys. Chem. B</i> 2011 , <i>115</i> , 11506-11518. Copyright (2011) American Chemical Society. ⁷⁵	20
Figure 1.14. IRMPD spectra of [Pb(Lys-H)] ⁺ complex compared to the computed IR spectra of type A, B and C complexes. Lys prefers a type C structure opposite to others shown in 1.13 Figure that preferred a type A configuration. Reprinted with permission from <i>J. Phys. Chem. B</i> 2011 , <i>115</i> , 11506-11518. Copyright (2011) American Chemical Society. ⁷⁵	21
Figure 1.15. An illustration of the IRMPD spectra of (Gly) ₂ H ⁺ compared to the computed IR spectra of the three lowest energy isomers. The GG-CS01 isomer fits better to the experimental IRMPD spectrum while coexistence of the other two isomers can be observed. Reprinted with permission from <i>J. Am. Chem. Soc.</i> 2007 , <i>129</i> , 4864-4865. Copyright (2007) American Chemical Society. ⁷⁶	22
Figure 1.16. IRMPD spectra of (Pro) ₂ H ⁺ together with the computed IR spectra of the three lowest energy isomers. The isomer containing a zwitterionic proline is the best match to the experimental IRMPD spectrum. The symmetric stretching vibration of the carboxylate group at 1396 cm ⁻¹ is considered the diagnostic band for zwitterionic structures. Reprinted with permission from <i>J. Am. Chem. Soc.</i> 2007 , <i>129</i> , 4864-4865. Copyright (2007) American Chemical Society. ⁷⁶	23
Figure 1.17. IRMPD spectra of the bovine ubiquitin in different charge states. The position of CO stretch at 1650 cm ⁻¹ is indicative of an alpha-helical structure which undergoes a blue shift upon increase in charge state. While the mid band at 1530 cm ⁻¹	

remains unchanged, the increase in charge states causes a dramatic increase in intensity for the unknown peak at around 1490 cm ⁻¹ . Reprinted with permission from <i>Phys. Chem. Chem. Phys.</i> 2005 , 7, 1345-1348. Copyright (2005) Royal Society of Chemistry. ⁸⁰	25
Figure 1.18. Left panel) IRMPD spectrum of protonated adenine compared to its three lowest energy tautomers. Right panel) the chemical structure of adenine tautomers and 9-methyladenine. Reprinted with permission from <i>ChemPhysChem</i> 2011 , 12, 1921-1927. Copyright (2011) John Wiley and Sons. ⁸¹	27
Figure 1.19. Comparisons of IRMPD spectra and computed IR spectra of two lowest energy isomers for three nucleobases (T, C, U). For all bases, the experimental spectrum is reproduced by the computed lowest energy isomer (the enol form). Reprinted with permission from <i>ChemPhysChem</i> 2007 , 8, 2235-2244. Copyright (2007) John Wiley and Sons. ⁷²	28
Figure 1.20. The Computed lowest energy structures of <i>cis</i> -[Pt(NH ₃) ₂ (G)Cl] ⁺ with relative free Gibbs energies (kJ mol ⁻¹ , italic numbers in parentheses). N7 is the preferred site for platination. Reprinted with permission from <i>J. Am. Chem. Soc.</i> 2013 , 135, 1445-1455. Copyright © 2012 American Chemical Society. ⁸⁵	30
Figure 1.21. The IRMPD spectrum of <i>cis</i> -[Pt(NH ₃) ₂ (G)Cl] ⁺ compared to four lowest energy isomers. N7 is the preferred site for platination. Reprinted with permission from <i>J. Am. Chem. Soc.</i> 2013 , 135, 1445-1455. Copyright © 2012 American Chemical Society. ⁸⁵	30

- Figure 1.22.** A comparison of the IRMPD spectrum of *cis*-[Pt(NH₃)₂(5'-dGMP-H)]⁺ (left structure with blue spectrum) and *cis*-[PtCl(NH₃)₂(5'-dGMP)]⁺ (right structure with green spectrum). The deprotonated ion shows a simultaneous attachment to the N7 of guanine and an O atom of the phosphate group with a characteristic hydrogen bond (NH₃---O=C) with a band at 3000cm⁻¹. In *cis*-[PtCl(NH₃)₂(5'-dGMP)]⁺, the metal is exclusively linked to N7 of the guanine residue with two hydrogen bonds (NH₃---O=C and NH₃---OP) appearing on IRMPD spectrum at ≈ 3200 cm⁻¹. Reprinted with permission from *Inorg Chem* **2015**, *54*, 3513-3522. Copyright 2015 American Chemical Society.⁸⁶ 31
- Figure 1.23.** Illustration of a guanine molecule (left) and a guanine tetrad (right). The black sphere represents a sodium cation. Reprinted with permission from *Chem. Commun.* **2014**, *50*, 14767-14770. Copyright 2014, Royal Society of Chemistry.⁸⁷ 32
- Figure 1.24.** The IRMPD spectrum of sodium G-tetrad (in two regions) compared to calculated infrared spectrum of the lowest energy isomer. A good match is reproduced by the computational spectrum except two bands at 3235 and 3322 cm⁻¹ which are absent in the computed spectra. These two bands can be related to overtones of bands at 1608 and 1680 cm⁻¹. Reprinted with permission from *Chem. Commun.* **2014**, *50*, 14767-14770. Copyright 2014, Royal Society of Chemistry.⁸⁷ 33
- Figure 1.25.** Different configuration resulted from complexation of lead to amino acids along with the corresponding hydrated clusters. Reprinted with permission from *Phys. Chem. Chem. Phys.* **2012**, *14*, 15118-15126. Copyright 2012 Royal Society of Chemistry.⁹¹ 34

Figure 1.26. Arrhenius plots for a series of 7-mers oligonucleotides. From top to bottom, plot for loss of adenine from $d(A_7)_2^-$ (filled square), $d(AATTAAT)_2^-$ (filled circle), and $d(TTAATTA)_2^-$ (open circle). The open square plot represents the dissociation of $d(CCGGCCG)_2^-$ (both C and G dissociation). Plots for loss of neutral cytosine (filled triangle), and the loss of neutral guanine (plus), from $d(CCGGCCG)_2^-$, are also established. Reprinted with permission from *J. Am. Soc. Mass Spectrom* **1998**, *9*, 1117-1124. Copyright 1998, American Society for Mass Spectrometry.⁹⁵36

Figure 1.27. The observed activation energies (bottom with open circles) and threshold dissociation energies modeled by master equation modeling (upper curve with filled circles). These results suggest that the inner solvation shell is composed of six uracil molecules. Reprinted with permission from *Phys. Chem. Chem. Phys.* **2012**, *14*, 3304-3315. Copyright 2011 Royal Society of Chemistry.⁹⁶38

Figure 1.28. An illustration of $U_{14}Ca^{2+}$ cluster. The inner shell consists of 6 uracil molecules. Four pairs of dimeric uracil molecules, in the outer shell, are bound to the inner shell through hydrogen bonding. Reprinted with permission from *Phys. Chem. Chem. Phys.* **2012**, *14*, 3304-3315. Copyright 2011 Royal Society of Chemistry.⁹⁶39

Figure 2.1. Bruker Apex-Qe7.0 T Fourier transform ion cyclotron mass spectrometer FT-ICR MS coupled to OPO laser located at Memorial University.53

Figure 2.2. Bruker Apex Qe 70 FTICR-MS located at Memorial University. The schematic describes the Apollo II ion source, Qh region, hexapole collision cell, ion transfer optics and ICR cell.54

Figure 2.3. A schematic of the electrospray ionization mechanism.	55
Figure 2.4. Cyclotron motion of ions in the presence of a spatially uniform magnetic field. The Lorentz force acts on ions perpendicular to the direction of the magnetic field. It is shown that the anion's cyclotron motion trajectory is opposite to that of positive ions.	57
Figure 2.5. A diagram of a Penning trap showing the configuration of plates compared to the magnetic field.	59
Figure 2.6. A diagram showing how a packet of undetectable ions at the center of an ICR cell can be excited by using an RF potential to a coherent and detectable motion.	60
Figure 2.7. A diagram illustrating how a raw time domain transient is Fourier transformed to a frequency domain which can be finally calibrated to a mass spectrum. Reprinted with permission from <i>Analyst</i> 2005 , 130, 18-28. Copyright (2004) Royal Society of Chemistry. ²²	61
Figure 2.8. Representation of vibrational energy levels of (a) an imaginary dissociation channel without anharmonic effect and (b) a species with anharmonic effect, showing convergence in vibrational spacings. Reprinted with permission from Fridgen, T. D.; McMahon, T. B. "IRMPD" in " <i>Encyclopedia of Mass Spectrometry</i> , 2005 , Vol. 4. Copyright (2005) Elsevier ⁴¹	72
Figure 2.9. A general schematic of an IRMPD process representing photon absorption in a specific vibrational mode followed by an IVR process. Right after every photon absorption, fast intramolecular vibrational relaxation stores energy gained over other	

vibrational modes, which is much easier in large polyatomic molecules. This cycle keeps repeating until the internal energy deposited surmounts the lowest dissociation threshold making ions undergo unimolecular dissociation. Reprinted with permission from Fridgen, T. D.	73
Figure 2.10. A representation of acquiring an IRMPD band. A scan of trapped ions $K_2(9eG)_{12}^{2+}$, 9eG denotes 9-ethylguanine, with the laser in the range of 3470 to 3426 cm^{-1} results in ion fragments that can be subsequently used to build the corresponding IRMPD bands in that wavelength range. Note that a full IRMPD spectrum can be obtained by scanning the desirable range of IR area.....	74
Figure 2.11. A diagram of free electron laser components. The undulator magnet plays a crucial role in which very-high-speed electrons become accelerated radially emitting electromagnet radiation. Reprinted with permission from <i>Infrared Phys. Tech.</i> 1995 , 36, 297-308. Copyright (1995) Elsevier. ⁴⁶	76
Figure 3.1. Electrospray mass spectrum of 10 mL of 0.1 mM solution of 9-ethylguanine to which two drops of 0.1 mM KCl was added.....	105
Figure 3.2. Computed structures of the metal cationized G-tetrads indicating the structural parameters listed in Table 3.1.	106
Figure 3.3. IRMPD spectra of the $M(9eG)_4^+$ G-tetrads in the 900 – 1850 cm^{-1} and the 2700 – 3800 cm^{-1} regions (black traces). The underlying grey traces are the computed spectra (B3LYP/6-31+G(d,p), Def2SVPD on metals) for the lowest energy G-tetrad structures scaled by 0.97 in both regions. The insets for the fingerprint regions of the	

Na ⁺ and K ⁺ tetramers are without an attenuating element, roughly 10x the laser intensity.....	108
Figure 3.4. Tetramer intensity vs the center of mass energy decay curve for the M(9eG) ₄ ⁺ G-tetrads from the energy-resolved SORI-CID spectra experiments. Reservoir pressure was 10 mbar with Ar. The higher energy required to dissociate Na(9eG) ₄ ⁺ indicates a structure more stable to loss of 9-ethylguanine.....	111
Figure 3.5. The relative center of mass collision energies (kJ mol ⁻¹) to affect 50 % dissociation for the five alkali metal cationized G-tetrads. The experimental values were done at three different reservoir pressures as indicated in the legend. The white bars show the computed binding energies for loss of a 9-ethylguanine from M(9eG) ₄ ⁺ . The computed values are from the B3LYP/6-311+G(3df,3pd) calculations listed in Table 3.2.....	112
Figure 4.1. Experimental vibrational (IRMPD) spectra of the M(9eG) ₈ ⁺ quadruplexes (black trace) and computed infrared spectra (grey trace).....	131
Figure 4.2. Kinetics plots for Na(9eG) ₈ ⁺ and K(9eG) ₈ ⁺ at 341 K and 399 K, respectively.	134
Figure 4.3. Arrhenius plots for the M(9eG) ₈ ⁺ quadruplexes and the Na(9eG) ₇ ⁺ complex.	135
Figure 4.4. Summary of Arrhenius activation energies (black bars), master equation model dissociation thresholds fitted to the experimental Arrhenius plots (grey bars), and computed dissociation enthalpies (white bars), all in kJ mol ⁻¹	140

Figure 5.1. Electrospray mass spectrum of an aqueous solution containing 10 mL of 0.1 mM solution of 9-ethylguanine with two drops of 0.1 mM KCl. Note that $K_2(9eG)_{12}^{2+}$ is the largest potassiated G-quadruplex detected.	157
Figure 5.2. Dissociation products resulted from unimolecular dissociation of $K_2(9eG)_{12}^{2+}$ by SORI-CID, IRMPD and BIRD activation techniques.....	158
Figure 5.3. Experimental vibrational (IRMPD) spectra of the $K_2(9eG)_{12}^{2+}$, $K(9eG)_8^+$ and $K(9eG)_4^+$ (black trace). The corresponding computed infrared spectra (grey trace) are also depicted.	160
Figure 5.4. Energy resolved intensity profiles for SORI-CID of $K_2(9eG)_{12}^{2+}$	161
Figure 5.5. The quadruplex intensity vs. collision energy plot resulted from the energy-resolved SORI-CID experiments at reservoir pressure of 10 mbar with Ar.	162
Figure 5.6. BIRD kinetics plot for $K_2(9eG)_{12}^{2+}$ at 385 K. Only octamer, tetramer, monomer are the fragment products.	163
Figure 5.7. Arrhenius plots for the $K_2(9eG)_{12}^{2+}$, $K(9eG)_8^+$, and $Na(9eG)_8^+$ quadruplexes. It was previously shown that among G-quadruples composed of two G-tetrads, $K(9eG)_8^+$ was the most stable cluster while $Na(9eG)_8^+$ was the weakest. ⁴⁸ ...	164
Figure S3.1. Profiles of the computed structures of the $M(9eG)_4^+$ complexes computed using B3LYP/6-31+G(d,p) with Def2SVPD on the metal centers. The calculations are also empirically corrected for dispersion.	182
Figure S3.2 a) Comparison of computed IR spectra for three isomers of $Na(9eG)_4^+$ with the experimental IRMPD spectra.	183

Figure S3.2 b) Comparison of computed IR spectra for three isomers of $\text{Li}(9\text{eG})_4^+$ with the experimental IRMPD spectra.	184
Figure S3.2 c) Comparison of computed IR spectra for three isomers of $\text{K}(9\text{eG})_4^+$ with the experimental IRMPD spectra.	185
Figure S3.3. Energy resolved intensity profiles for SORI-CID of the $\text{M}(9\text{-eG})_4^+$ complexes.....	186
Figure S3.4. Energy-resolved G-tetrad intensity profiles for the SORI-CID of $\text{M}(9\text{eG})_4^+$ with a) a reservoir pressure of 5 mbar and b) a reservoir pressure of 15 mbar with Ar.	187
Figure S4.1. Electrospray mass spectrum of an aqueous solution of KCl/9-ethylguanidine.	187
Figure S4.2. IRMPD spectrum of potassium quadruplex compared to potassium G-tetrad and computed infrared spectrum. Lowest energy structures are also depicted.	188
Figure S4.3. A diagram of molecular degrees of freedom as a function of reaction rate constants. This diagram illustrates that sodium G-quadruplex lies within the small molecule regime and its observation Arrhenius parameters are required to be modeled by master equation modeling. Larger cations, K, Rb and Cs, are categorized in the large regime.....	189
Figure S4.4. BIRD rete constants plots of alkali metal quadruplexes ($\text{M}(9\text{eG})_8^+$) at different temperatures	191
Figure S4.5. Kinetics plots for $\text{M}(9\text{eG})_8^+$ quadruplexes at different temperatures	205

Figure S4.6. Master equation modeling results for $\text{K}(9\text{eG})_8^+$ and $\text{Na}(9\text{eG})_8^+$ quadruplex²⁰⁸

List of Tables

Table 1.1. Arrhenius activation energies (E_0) and master equation modeled threshold dissociation energies	35
Table 1.2. Arrhenius parameters for loss from doubly deprotonated oligomers	37
Table 2.1 Thermochemical parameters accessible by BIRD experiments	68
Table 3.1. B3LYP/6-31+G(d,p) structural parameters for the $M(9eG)_4^+$ guanine tetrads, $M=Li, Na, K, Rb,$ and $Cs,$ and $(9eG)_4$	107
Table 3.2. 298 K Dissociation enthalpies of neutral and alkali metal cationized G-tetrads, $(9eG)_4$ and $M(9eG)_4^+$	114
Table 4.1. Summary of the observed Arrhenius parameters for the dissociation of $M(9eG)_8^+$ G-quadruplexes, and those obtained from master equation modeling of the temperature dependent rate constants for $Na(9eG)_8^+$ and $K(9eG)_8^+$	136
Table 5.1. Observed Arrhenius parameters for the dissociation of $K_2(9eG)_{12}^{2+}$ and the $M(9eG)_8^+$ G-quadruplexes ⁴⁸	164
Table S4.1. BIRD rate constants of alkali metal G-quadruplexes at different temperatures	192

List of Schemes

Scheme 3.1. The guanine nucleobase (a), and the structure of a guanine tetrad complex (b).....	98
Scheme 4.1. The structure of a G-tetrad complex (a), and a G-quadruplex cluster templated by potassium cations (b).....	122

List of Appendices

Appendix 1: Contains figures S3.1, S3.2, S3.3, and S3.4 related to Chapter 3, *guanine tetrads: An IRMPD spectroscopy, energy resolved SORI-CID, and computational study of $M(9\text{-ethylguanine})_4^+$ ($M=\text{Li, Na, K, Rb, Cs}$) in the gas phase*

Appendix 2: Contains figures S4.1, S4.2, S4.3, S4.4, S4.5, S4.6 and Table S4,1 related to Chapter 4, *the intrinsic stabilities and structures of metal cationized guanine quadruplexes*

List of Abbreviation and Terms

ab initio	a Latin term for “from the beginning”
B3LYP	Becke-(3 parameter)-Lee-Yang-Parr
CRM	charged residue model
CID	collision-induced dissociation
DFT	density functional theory
ECD	electron capture dissociation
ESI	electrospray ionization
ETD	electron transfer dissociation
CLIO	Centre Laser Infrarouge d’Orsay
FEL	free electron laser
FT	Fourier transform
FTICR	Fourier transform ion cyclotron resonance
ICR	ion cyclotron resonance
IEM	ion evaporation model
IR	infrared
IRMPD	infrared multiple-photon dissociation
IVR	intramolecular vibrational redistribution
Laser	light amplification by stimulated emission of radiation
MALDI	matrix-assisted laser desorption ionization
MS	mass spectrometry

MS/MS or MSn	tandem mass spectrometry
Nd:YAG	neodymium-doped yttrium aluminum garnet
OPO	optical parametric oscillator
RF	radio frequency
SORI-CID	sustained off-resonance collision-induced
SVP	split valence polarization
ECP	effective core potential
DNA	deoxyribonucleic acid
RNA	ribonucleic acid
HF	Hartree-Fock
E_a	activation energy
WC	Watson-Crick

Chapter 1. Introduction

1.1. Physical Chemistry Analysis of Gas Phase Ions via Mass Spectrometry

A “complete chemical laboratory” is a term that the mass spectrometer is often described as. The emergence and development of a number of important scientific aspects have been attained by the contributions of mass spectrometry. Mass spectrometry emerged with the discovery of the electron by J.J. Thompson and has progressively grown to a tool that today is widely used to perform complex tasks: answering questions from quantum theory to the determination of genome sequences,^{1,2} proteomics,³⁻⁵ and fine isotopic analysis.^{6,7}

In mass spectrometry the mass of an ionized analyte, whether from a complex or simple sample, is determined. Traditional ionization techniques such as electron impact (EI) and chemical ionization (CI) methods were the main obstacle for the introduction of mass spectrometry to the analysis of biomolecules, massive fragile macromolecules with

non-covalent interactions such as DNA and proteins. These classic ionization techniques, EI and CI, ionize only volatile molecules that limits the feasibility of mass spectrometry research within a very narrow dynamic range, up to about 1000 Da. The emergence of soft ionization methods and tandem mass spectrometry techniques was the starting point of the dominance of mass spectrometry as a versatile instrument to study the physical chemistry of macromolecules in the gas phase.

Electrospray ionization⁸ (ESI) and matrix-assisted laser desorption ionization⁹ (MALDI) are soft ionization techniques and opened up the possibility to volatilize biological molecules in the MDa regime. These soft ionization techniques not only extended the mass range by intact ionization of macromolecules but also allows the analysis of very large non-volatile biomolecules by producing multiply charged ions, moving them into the detectable mass region (Figure 1.1).¹⁰ ESI and MALDI have both proved to yield similar ionization qualities, in term of softness (low fragmentation), however ESI is arguably the most versatile soft ionization technique due to the fact that MALDI instrumentation is more complex, requiring a laser and complex sample preparation. In addition, MALDI generates only small amounts of multiply charged ions.

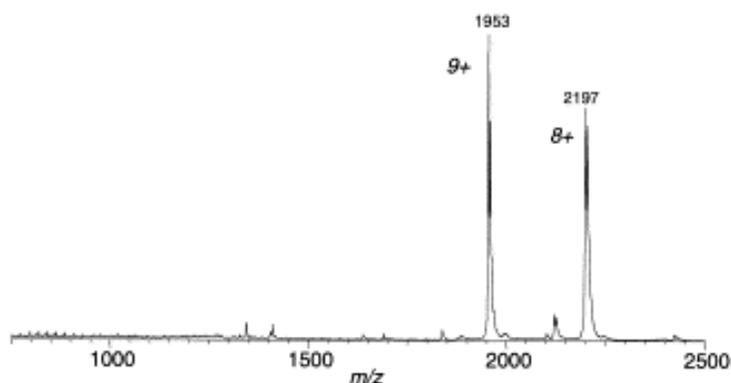


Figure 1.1. ESI mass spectrum of myoglobin protein (10 μ M in 10 mM ammonium acetate, pH 6.9) featuring two protein charge states (8^+ and 9^+). This shows how ESI can detect very large biomolecules by generating multi-charged species. It must be noted that due to the intact precursor ionization, structural and sequencing information is not extractable, and this requires additional mass stages, tandem mass spectrometry, to form fragment ions. Reprinted with permission from *Int. J. Mass Spectrom* **2000**, *200*, 175-186. Copyright (2000) Elsevier Ltd.¹⁰

The second important development that significantly diversified and broadened the application of mass spectrometry was tandem mass spectrometry, the ability to perform multiple steps of mass isolation and fragmentation, followed by mass analysis.¹¹⁻¹³ Tandem mass spectrometry became a requirement since soft ionization methods minimize the in-source dissociation and fragmentation, resulting in a decline in the potential for structure elucidation and also biomolecule sequencing (Figure 1.1). As a result, incorporation of at least one other mass spectrometry step is a necessity to study the physical chemistry of the analyte ions.

Tandem mass spectrometry instruments were first developed by coupling different mass analyzers such as quadrupoles mass filters, sectors, and time of flight (TOF) mass analyzers, in a way that each mass analyzer, depending on its position, acted as a mass separator, collision chamber, or mass analyzer (Figure 1.2).^{14,15} The evolution

of tandem mass spectrometry has provided tandem in time mass spectrometers¹⁶ which incorporate all the stages in a single mass analyzer. These instruments such as Fourier transform ion cyclotron resonance mass spectrometers (FT-ICR), quadrupole ion trap, and orbitrap enable scientists to isolate ions and trap them for an extended period of time during which they can be fragmented and the fragment ions remain isolated in the mass spectrometer allowing for mass detection. This feature has enabled the development of many physical/analytical chemistry techniques to understand the detailed physical chemistry of noncovalent complex ions. This includes the determination of thermochemistry, structural elucidation, and the study of ion-molecule reactions.

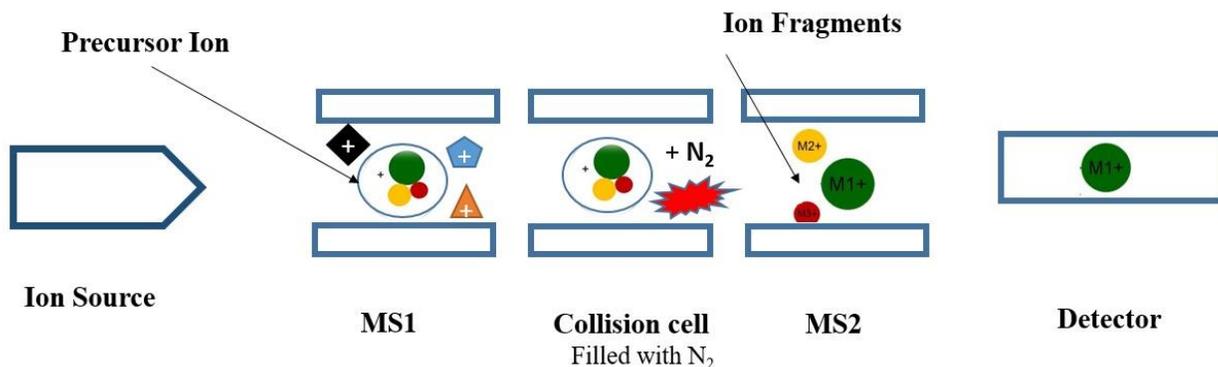


Figure 1.2. A schematic of a tandem in space mass spectrometer, triple quadrupole. Selected precursor ions in MS1 are transferred to the collision cell, Quad 2, to undergo dissociation. The MS2 mass scans the fragment ions.

Ion activation methods in tandem mass spectrometry impart packets of energy, in the eV to keV range, to induce fragmentation. This provides a powerful tool to obtain a deep understanding of ion structure, reactivity, and thermochemistry. Various ion activation techniques have emerged in terms of scientific purposes and amount of energy required for dissociation.¹⁷⁻²⁰

Collision-induced dissociation (CID) is the simplest and most common fragmentation technique which can be easily employed in both tandem-in-space and tandem-in-time mass spectrometers. In CID, fragmentation of molecular ions is induced by the collision of accelerated ions into relatively motionless neutral gases, called targets, most commonly Ar and N₂. This turns a portion of the ion's translational energy into internal energy followed by dissociation. Surface-induced dissociation^{20,21} (SID) is an analogous version of CID in which ions are accelerated into a surface rather than bumping into neutral gases. The amount of transferable energy during a head to head collision depends on the mass of the target gas; the higher the target mass, the greater the energy transferred to the ion. By utilizing a solid surface instead of a target gas, the mass jumps to a theoretically infinite value followed by a maximum in the center of mass energy transfer. This allows one to reach collisional energies up to keV which is a virtue to access the cleavage of strong covalent interactions such as aromatic cycles.

SID has been mated to several tandem configurations including quadrupole hybrids,^{22,23} quadrupole ion traps,²⁴ and TOF/TOF.²⁵⁻²⁷ In addition, some research groups have employed FTICR for SID experiments.²⁸⁻³¹ In this case, a surface is placed at one of the FTICR into which ions are accelerated. It is believed that this arrangement improves the mass resolution since the need to inject a target gas in CID is eliminated.¹⁷

SID can yield better structural information than CID when the ions are very large, massive biologically relevant molecules such as proteins and peptides, since it induces extensive fragmentation.³² Another advantage of SID over CID stems from the fact that in SID, ions are excited with a narrow internal energy distribution due to a single

collision with the surface. In contrast, the CID excitation mechanism involves a variable number and orientation of collisions with a target gas, resulting in a broad internal distribution. This is very worthwhile when identifying isomers that interconvert easily, and SID can yield distinguishable mass spectra for isomers.³³⁻³⁵ SID has been shown to be promising to study dissociation pathways of protein complexes.^{36,37} CID and SID spectra of a multi-charged dimer of cytochrome C (D^{11+}) are illustrated in Figure 1.3. Except for the ion activation techniques, all other experimental conditions are identical (quadrupole-TOF with a nanospray ion source). As it can be seen in Figure 1.3, the dimer protein dissociates through different pathways. CID unsymmetrically forms M^{3+} and M^{8+} monomers whereas the most symmetrical dissociation pathway (M^{5+}/M^{6+}) occurs when SID takes place. It is proposed that owing to the step-by-step energy pumping in CID, large-scale subunit unfolding (M^{8+}) occurs leading to an unsymmetrical dissociation pathway (Figure 1.4). Due to the nature of SID, a one collision activation, leading to a close-packed transition state is the source of the relatively symmetric pathway. However, SID and CID results can be used complementarily since they produce different fragment ions, specifically when the goal is to obtain sequencing data.

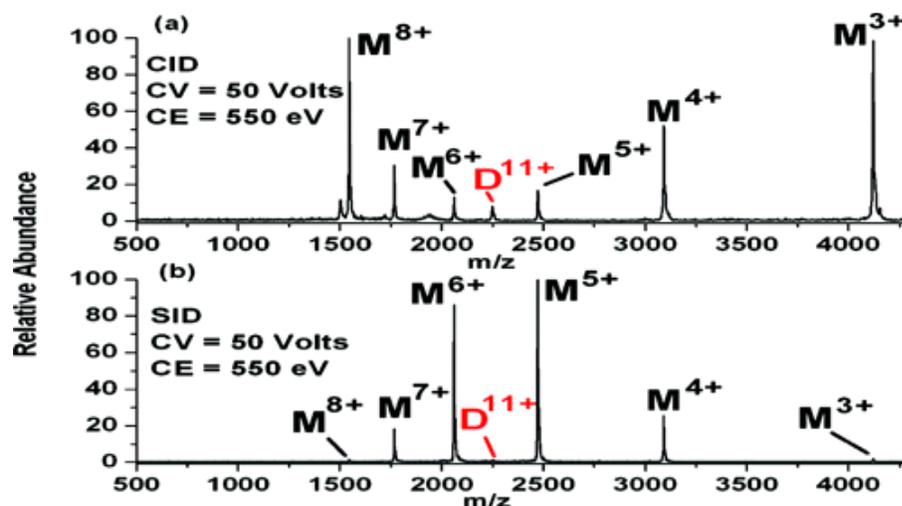


Figure 1.3. CID (a) and SID (b) mass spectra of D¹¹⁺ (the 11⁺ cytochrome C dimer) under the same conditions (laboratory collision energy CE and voltage difference ΔV). Tandem mass spectra depict different monomer products and dissociation pathways. Reprinted with permission from *J. Am. Chem. Soc.* **2006**, *128*, 15044-15045. Copyright (2006) American Chemical Society.³⁶

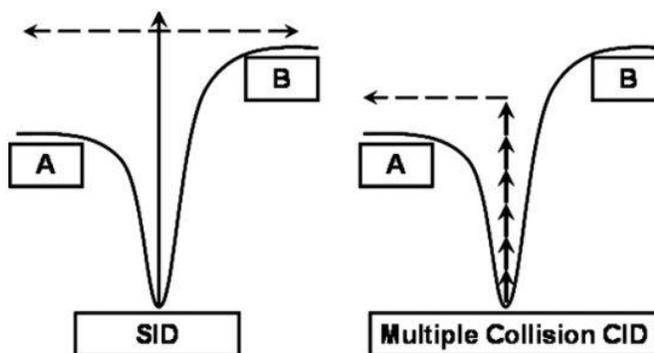


Figure 1.4. SID ion activation mechanism might favor a dissociation pathway. In case of complex proteins, stepwise ion activation, CID, enables unfolding of subunits (A pathway) while a sudden activation, SID, can lead to a different transition state and products. Reprinted with permission from *J. Am. Soc. Mass Spectrom.* **2008**, *19*, 190-208. Copyright (2008) American Society for Mass Spectrometry.³⁷

Ion activation can also occur through irradiating ions with an intense beam of infrared radiation. Infrared multiphoton dissociation (IRMPD) is an ion activation

technique in which ion fragments form via a multiphoton absorption process. Common chemical bond breakage requires 100-400 kJ mol⁻¹ and less when dealing with non-covalent interactions. To this regard, several tens of infrared photons are required to reach a dissociation threshold. IRMPD takes place through a two-step process which is fully discussed in the Chapter 2. Briefly, a vibrational normal mode which is in resonance with an intense laser beam absorbs one photon at a time ($\nu = 0 \rightarrow 1$) and under low pressure conditions, in the absence of collisions, the energy of the absorbed photon is randomized throughout the ion by a process called intramolecular vibrational energy redistribution (IVR). Subsequent photon absorption at the same fundamental transition followed by IVR is repeated until one of the dissociation channels obtains enough energy to undergo a bond cleavage.

The efficiency and performance of CID varies greatly with instrument design. For example, the time required to perform an FTICR-MS experiment increases dramatically by adding a CID step due to the added time to pump down the collision gas prior to the detection step, decreasing the instrument duty cycle. Loss of ions due to scattering is another associated issue with CID, constraining the sensitivity and the chance to effectively transfer the ions to the next MS steps due to missing some ions in every CID experiment. Neither of the above issues are associated with IRMPD, but IRMPD requires additional equipment, a tunable laser to irradiate the ions, which makes it a more expensive method than CID. The amount of energy deposited in CID dramatically varies for different instruments. With ion trap instruments and low collision energies CID is affected through multiple collisions during which slowly activation can interrupt non-

covalent interactions and lead to rearrangement reactions.¹⁷ In contrast, high collision energies in the keV range are attainable by sector and time of flight (TOF) mass spectrometers. In these instruments, an electronic excitation accelerates a beam of ions to kinetic energies up to tens of keV such that only a single collision is necessary to cleave a chemical bond. These two CID energy regimes usually show different dissociation products that allow access to more ion fragments, which may carry important structural information.^{38,39} It is worth noting that the amount of internal energy transferred during CID can also be influenced by the nature and pressure of collision gas. The higher pressure of a gas can lead to higher number of collisions, and a larger collisional gas can yield more inelastic collisions, so more energy is deposited. Ion activation mechanisms of CID and IRMPD are discussed in a greater detail in Chapter 2.

Although in most cases both IRMPD and CID give the same fragmentation patterns and produce the lowest-energy product ions, some comparative studies have clearly shown that these activation methods present different losses for complexes.⁴⁰ CID can be a less discriminative technique against higher energy fragmentation pathways compared to IRMPD. In CID, the energy deposited per collision was about 50 kJ mol⁻¹ whereas a photon (1000 cm⁻¹ for CO₂ laser) can deposit a maximum of 10 kJ mol⁻¹ energy into the ion. Therefore, CID cannot act so energy selective when some competitive fragmentation pathways with close energy are present.

During IRMPD spectroscopy, an intense and tunable laser irradiates ions allowing them to gain enough internal energy so that the deposited energy surpasses at least the lowest activation energy barrier followed by dissociation. By recording ion fragmentation

as a function of laser wavelength an IRMPD spectrum can be constructed that is somewhat similar to an IR spectrum. The IRMPD spectrum can be compared to infrared spectrum computed for different isomers and conformers using techniques such as DFT (density functional theory). IRMPD spectroscopy enables one to obtain more detailed structural outcomes, but it is customary that complementary computational analysis towards structural elucidation is used. Figure 1.5⁴¹ features an example how IRMPD along with computational chemistry can be used to determine the lowest energy structure out of possible isomers. The IRMPD spectrum of Pb^{2+} /deoxycytidine-5'-monophosphate (dCMP) complex is shown in the left panel (Figure 1.5) along with the computed IR spectra of three isomers. It is apparent that the IR spectrum of the lowest energy isomer, dCytS6, is the best match to the experimental spectrum.

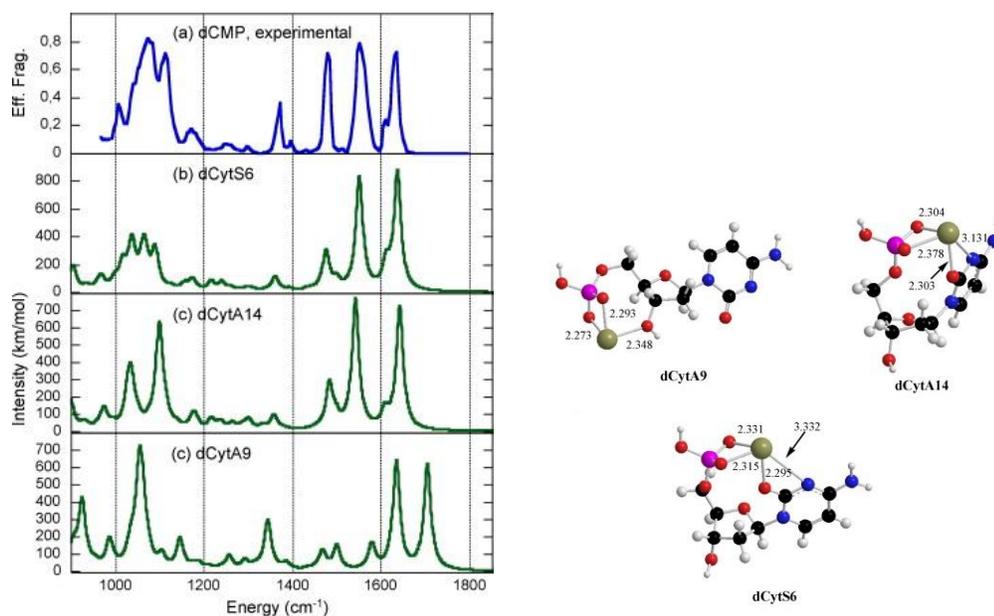


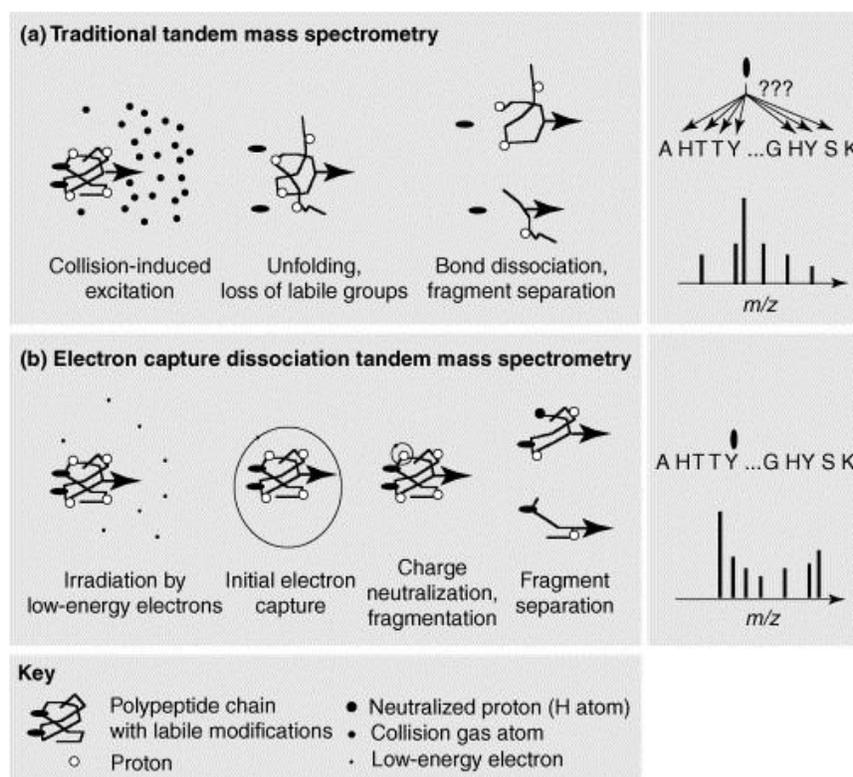
Figure 1.5. Experimental IRMPD of $[\text{Pb}(\text{dCMP})\text{-H}]^+$ compared with B3-LYP/6-31+G(d,p) calculated spectra of three isomers (left panel). DFT calculated structures of some lowest energy isomers are depicted at right panel. The experimental spectrum is in good agreement with the lowest energy structure, dCytS6, in which the metal interacts with both the phosphate group and the carbonyl group of the nucleobase moiety. Reprinted and adopted with permission from *Int. J. Mass spectrom* **2011**, 304, 154-164. Copyright (2010) Elsevier B.V.⁴¹

Blackbody infrared radiative dissociation (BIRD) is a very slow-heating activation technique well suited to the study of non-covalent interactions including hydrogen bonding, dipole-dipole, and even van der Waals interactions.⁴² The BIRD technique is exclusive to ion-trap devices due to a need for long trapping times. In BIRD, the heated walls of the ICR cell radiate a very weak flux of IR photons that are absorbed by the isolated ions. For weakly-bound ions, the elevated internal energy may surpass the threshold for dissociation. With BIRD a temperature dependence on the rate constant can be used to extract the thermochemistry and transition state properties.⁴³⁻⁴⁵ IRMPD, CID,

and blackbody infrared dissociation (BIRD) are the main tools used in the present thesis and will be explained in more detail in coming chapters.

Apart from having distinct mechanisms and instrumentation, all the ion activation methods explained above share a common characteristic. These methods, so-called “ergodic” techniques,¹⁷ activate ions through vibrational excitation followed by a fast energy redistribution over the whole molecule leading to a preferential dissociation at the weakest bonds. In this regard, a technique with the capability to cleave inaccessible barriers became a requirement for structure elucidation and sequencing of large biomolecules was the main goal. Based on the distinct mechanism of dissociation, electron capture dissociation, ECD, dissociates biomolecules up to 45 kDa at various locations which these locations remain intact using IRMPD and CID methods.

In the ECD process, a beam of low energy electrons, less than 0.5 eV, are captured by ions in an ion trapping mass spectrometer. Since the final fragments stem from a radical intermediate with varied bond lengths, strength, and potential energy surface, the ion fragments differ significantly compared to the ones produced by other fragmentation techniques. Figure 1.6¹⁹ compares the analysis of modified polypeptides by using a collisional ion activation technique (CID) and a method based on ECD. This shows that labile groups and secondary structures are lost during CID while ECD tends to preserve them. Due to better preservation of post-translational modifications (PMTs) and labile groups a more informative sequencing of proteins can be reached.



Current Opinion in Biotechnology

Figure 1.6. A general schematic of structural analysis of polypeptides by (a) CID and (b) ECD techniques. While removal of labile groups largely occurs in CID, ECD yields complete sequencing by keeping these vulnerable bonds. Reprinted with permission from *Curr. Opin. Biotechnol.* **2004**, *15*, 12-16. Copyright (2003) Published by Elsevier Ltd.¹⁹

In the three next sections, some applications of CID, IRMPD and BIRD, the main ion activation methods used in this thesis, will be explained. Further mechanistic and instrumentation details will be discussed in Chapter 2.

1.2. CID, IRMPD and BIRD for Biological Systems

1.2.1 CID Applications for Biological Systems

Collision induced dissociation is the most implemented ion activation technique due to both simplicity of incorporation in almost all mass spectrometers and the ability to

be used as a universal method (i.e. all molecular ions undergo collisions). CID was the first ion activation technique invented, as early as the observation of the first mass spectrum by Sir J. J. Thomson.

The utility of CID ranges from very small molecules to macromolecules up to MDa in size. CID has been employed for nucleotides and oligonucleotides⁴⁶⁻⁴⁸, peptides and proteins^{49,50}, pharmaceutical drugs⁵¹, organometallic molecules^{52,53}, and polymers.^{54,55}

Sustained-off resonance irradiation collision induced dissociation (SORI-CID) is the version of CID in ICR mass spectrometers which is explained in detail in the second chapter. SORI-CID has helped shed light on the structures and fragmentation pathways of nucleobase/metal dication complexes. An electrospray mass spectrum⁵⁶ (Figure 1.7) of an aqueous solution comprised of CuCl₂ and uracil has shown that the [Cu(Ura-H)(Ura)]⁺ complex forms as the most abundant ion. This complex possesses many isomeric possibilities. Computational chemistry with the aid of IRMPD spectroscopy has determined the lowest energy structures which have close interconversion barriers. Surprisingly, SORI-CID experiments showed HNCO as the major dissociation pathway, rather than neutral uracil loss. Extensive isotopic labelling, 2-¹³C, 3-¹⁵N, 1,3-¹⁵N₂, 5-d, and 6-d, followed by SORI-CID experiments revealed that the 3- and 2- positions are the source of the nitrogen and carbon atom. Further MSⁿ experiments along with theoretical results concluded a concise fragmentation pattern, starting from the lowest energy structure (Figure 1.8)

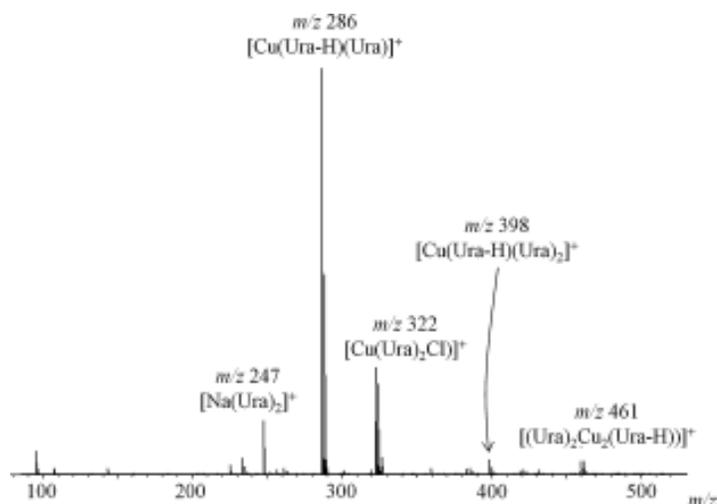


Figure 1.7. Electrospray mass spectrum of an aqueous solution of uracil and CuCl_2 Reprinted with permission from *ChemPhysChem* **2012**, *13*, 588-596. Copyright (2012) John Wiley and Sons.⁵⁶

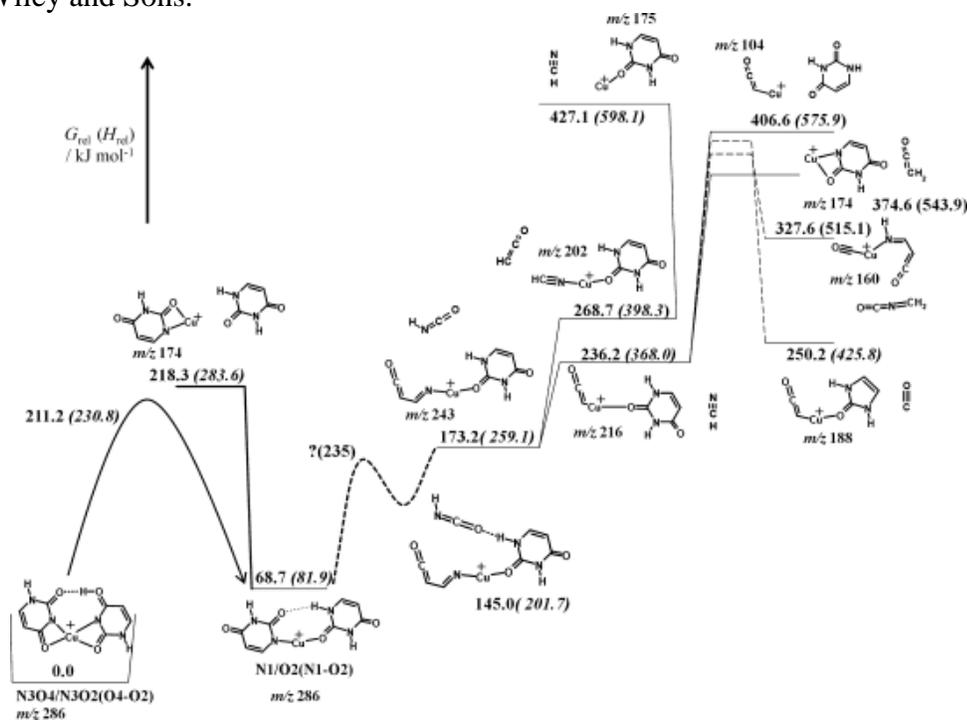


Figure 1.8. Gibbs energy diagram (298 K) of the $[\text{Cu}(\text{Ura-H})(\text{Ura})]^+$ fragmentation scheme attained with coupling SORI-CID and computational chemistry studies. It is shown that the fragmentation process initiates with an isomerization of the lowest energy structure. Reprinted with permission from *ChemPhysChem* **2012**, *13*, 588-596. Copyright (2012) John Wiley and Sons.⁵⁶

Amino acid sequences of peptides can be patterned and diagnosed through collisional activation. It was shown⁵⁷ long ago that CID, is able to produce distinctive mass spectra leading to a complete sequencing of certain soluble peptides.⁵⁸ Later, some research groups developed computer programs to enable scientists to compare the tandem mass spectra of a sequentially-unknown peptides to a library of CID spectra. This arguably appeared to be unsuccessful due to variability of instrumental and experimental factors.^{59,60} Although CID has been a valuable and inseparable tool for peptide sequencing and proteomics experiments, it is believed that methods based on collisional-activated dissociation results in incomplete sequencing.

The possible cleavages and dissociation products for protonated peptides are shown in Figure 1.9⁶¹ based on the notation system developed by Roepstorff⁶² to name peptide fragment ions. The nomenclature used in identifying the fragments represents both the position of fragmentation and the location of the charge. Satisfactory sequencing occurs when an activation method can produce a complete set, or most of the possible ion fragments by cleaving the backbone bonds of a peptide. CID methods tend mainly to form b and y fragment ions as the peptide bonds are the most labile, while a/x and c/z pairs are energetically inaccessible (Figure 1.10 a). A full or near-full sequencing coverage is hard to obtain unless a combination of two or more activation methods are used. For instance, electron excitation activation methods, ECD and electron transfer dissociation (ETD), form a/x and c/z pairs (Figure 1.10 b).⁶¹ As a result, CID and ECD/ETD can be used as complementary methods to CID to achieve better sequencing results.

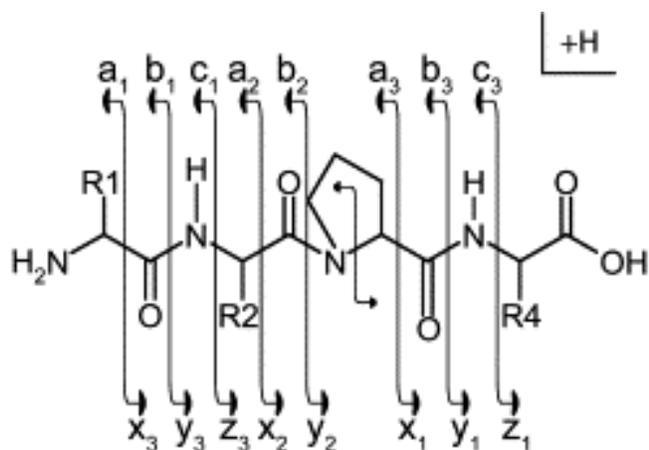


Figure 1.9. Notation of the possible dissociation pathways and ion products of a protonated peptide by Roepstorff and Fohlman nomenclature. a, b and c are indicative of fragment ions with the positive charge retained on the N-terminus while z, x, y ion fragments have the positive charge on the carboxy termination. Reprinted with permission from *Biochim. Biophys-Proteins and Proteomics*. **2006**, 1764, 1811-1822. Copyright (2006) Elsevier Ltd.⁶¹

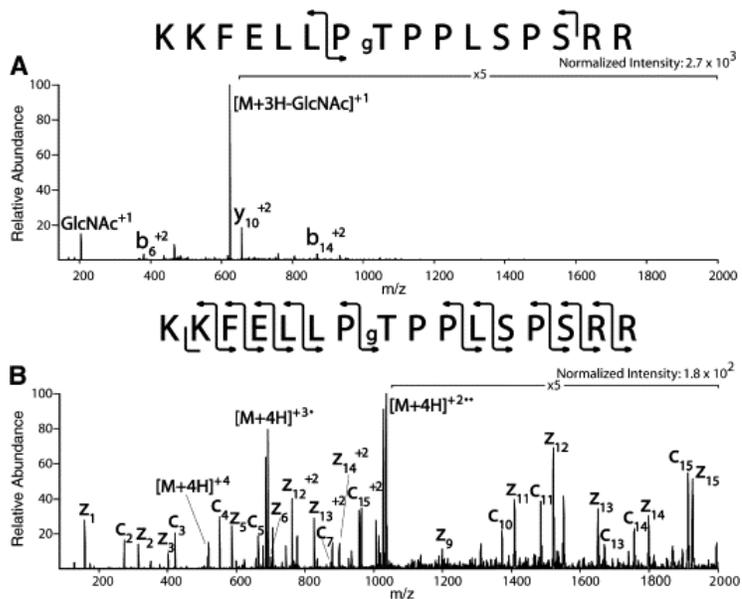


Figure 1.10. A comparison between a) CID and b) ETD tandem mass spectrum of O-GlcNAc containing peptides (1 pmol/ μ L). The CID spectra only includes b/y ion pairs resulted from peptide bond cleavages while a perfect set of c and z ions are obtained by ETD. Reprinted with permission from Reprinted with permission from *Biochim. Biophys-Proteins and Proteomics*. **2006**, 1764, 1811-1822. Copyright (2006) Elsevier Ltd.⁶¹

1.2.2 IRMPD Spectroscopy Applications for Biological Systems

Infrared spectroscopy has long been understood to contain important structural information that can lead to structural elucidation.^{63,64} However, the low density of ions in an ion trap requires a modified version of infrared spectroscopy. IRMPD spectroscopy was introduced in the 2000s to allow the IR spectroscopic investigation of gas phase ions. Instead of attenuation of the infrared light due to absorption by the ions, the consequence of photon absorption—dissociation—is monitored.^{17,65} IRMPD spectroscopy has been a helpful tool to probe the effect of metalation and protonation on the configuration and structure of a wide range of biomolecules such as amino acids,⁶⁶⁻⁶⁹ peptides,⁶⁹⁻⁷¹ nucleobases.⁷²⁻⁷⁴ The IRMPD mechanism is explained in great details in Chapter 2 and here some applications for biological systems are discussed.

Amino acids are of great interest as being one of the most significant biological building blocks, precursors to proteins. The cationization and hydration of amino acid has been studied by IRMPD spectroscopy.^{66,68} A comprehensive IRMPD spectroscopy work by Burt⁷⁵ and co-workers has studied the effect of lead (Pb^{2+}) on a number of amino acids (Aa) including proline (Pro), alanine (Ala), valine (Val), leucine (Leu), isoleucine (Ile), and lysine (Lys). The electrospray mass spectrum of a solution containing $\text{Pb}(\text{NO}_3)_2$ and amino acids revealed the ion cluster to be deprotonated, $[\text{Pb}(\text{Aa-H})]^+$. The IRMPD spectroscopy in the region of $3200\text{--}3800\text{ cm}^{-1}$ was conducted on six amino acid complexes and the spectra are illustrated in Figure 1.11. It is evident that Ala, Val, Leu, and Ile spectra feature almost identical patterns with two bands at 3440 cm^{-1} (very weak) and 3550 cm^{-1} (strong) which are indicative of N-H stretch and carboxylic acid O-H

stretch, respectively. On the other hand, lysine (Lys) exhibits a completely different spectrum with the absence of an O-H stretch at 3550 cm^{-1} and has three weaker bands in the N-H stretching region. The lack of O-H stretch can be ascribed to the fact that the deprotonation takes place at the carboxylic acid while for the other five amino acids the amine moiety is subject to deprotonation. Coordination of Pb^{2+} to the amino acids can lead to three main configurations illustrated in Figure 1.12.

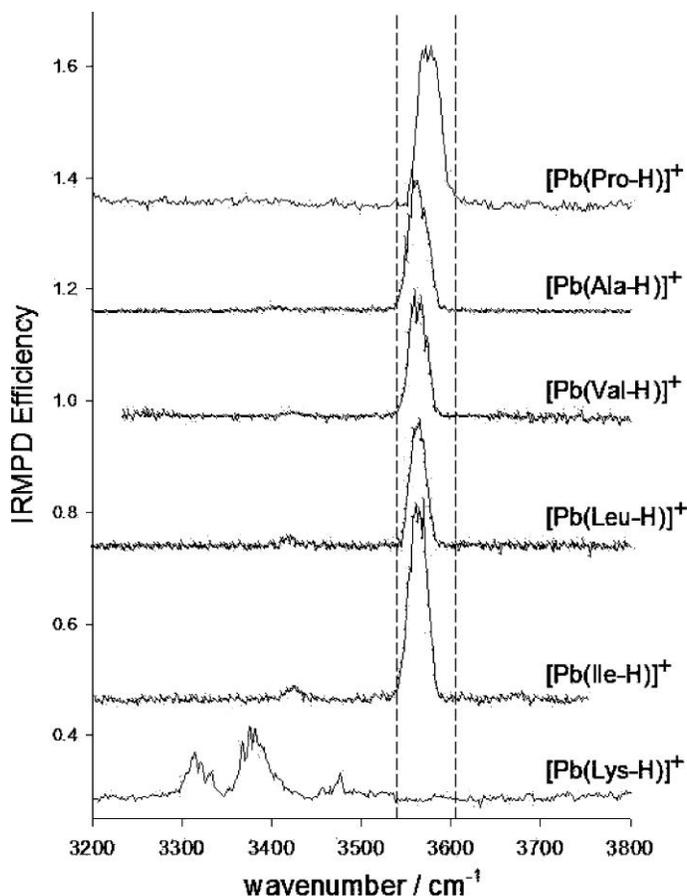


Figure 1.11. Comparison of IRMPD spectra of $[\text{Pb}(\text{Aa-H})]^+$ complexes. Except Lys which has a basic side chain, other five amino acids exhibit similar bands. Reprinted with permission from *J. Phys. Chem. B* **2011**, *115*, 11506-11518. Copyright (2011) American Chemical Society.⁷⁵

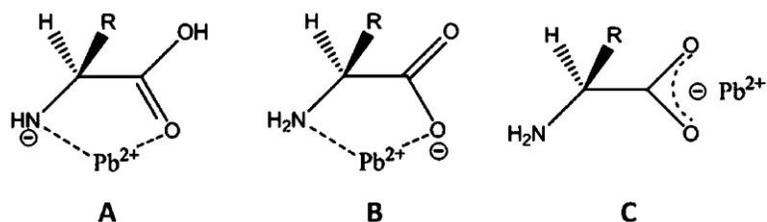


Figure 1.12. Three main possible attachment of Pb^{2+} in to amino acids in $[\text{Pb}(\text{Aa-H})]^+$ complexes denoted as type A, B and C.. Reprinted with permission from *J. Phys. Chem. B* **2011**, *115*, 11506-11518. Copyright (2011) American Chemical Society.⁷⁵

To gain insights into the configuration for each amino acid complex, computational studies were performed on each individual complex which some results are summarized in Figure 1.13.

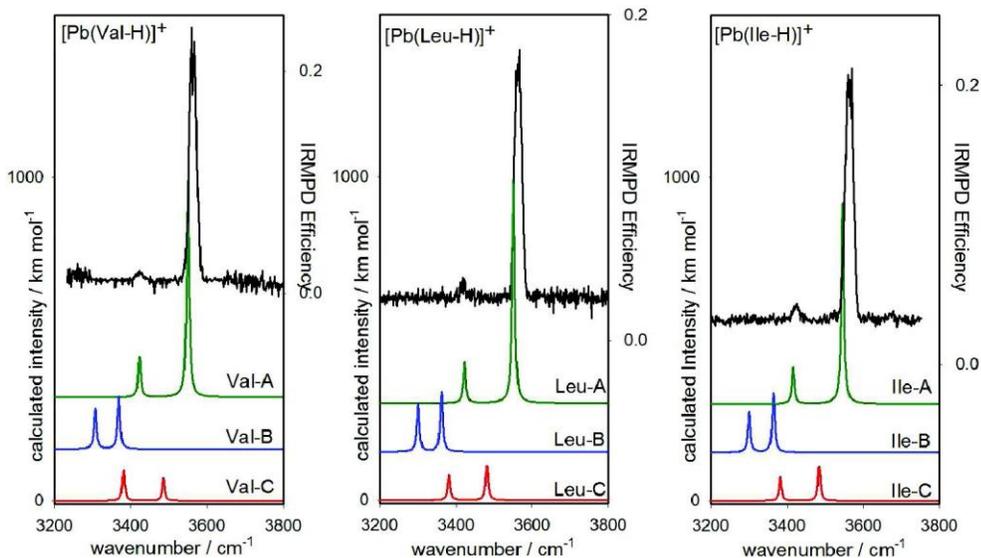


Figure 1.13. IRMPD spectra of three amino acid complexes and their comparisons to the computed IR spectra of type A, B and C complexes shown in Figure 1.12. Reprinted with permission from *J. Phys. Chem. B* **2011**, *115*, 11506-11518. Copyright (2011) American Chemical Society.⁷⁵

Comparison of the IRMPD spectra of Val, Leu, and Ile complexes with the calculated IR spectra of A, B and C type complexes (shown in Figure 1.12) clearly illustrates the formation of A type complexes. All these three amino acids possess side chains containing only C and H which tends not to interact with lead due to lack of polarity. As a consequence, Pb^{2+} as a soft Lewis acid favours softer base, amine groups, over the carboxylate anion which is considered a hard base. This stabilization energy can cope with the higher energy required to deprotonate an amine group relative to a carboxylic one. In contrast, the IRMPD comparison for $[\text{Pb}(\text{Lys-H})]^+$, Figure 1.14, indicates that the IRMPD spectrum is best match with the computed spectrum corresponding to the type C configuration. This different preference stems from the fact that the basic side chain in lysine introduces extra stabilization favoring a C-type structure with a deprotonation at its carboxylic moiety.

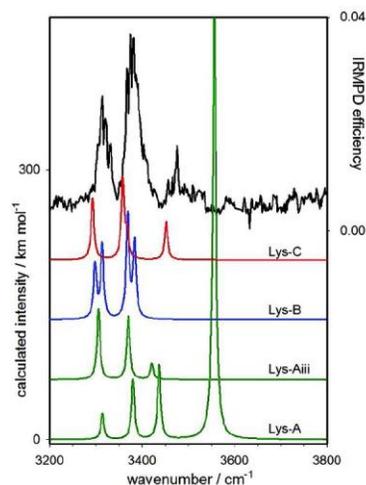


Figure 1.14. IRMPD spectra of $[\text{Pb}(\text{Lys-H})]^+$ complex compared to the computed IR spectra of type A, B and C complexes. Lys prefers a type C structure opposite to others shown in 1.13 Figure that preferred a type A configuration. Reprinted with permission from *J. Phys. Chem. B* **2011**, *115*, 11506-11518. Copyright (2011) American Chemical Society.⁷⁵

The structure and reactivity of larger clusters of amino acids can be studied by IRMPD. An IRMPD work by McMahon⁷⁶ and coworkers on amino acid proton-bound homodimers has provided evidence for the co-existence of zwitterionic isomers which is an isomer preferred in condensed phases. A lower region IRMPD spectrum, 1000 cm^{-1} -1800 cm^{-1} , of glycine proton-bound dimer, $(\text{Gly})_2\text{H}^+$, along with the computed spectra for three lowest energy isomers are indicated in Figure 1.15. According to the computational work, the GG-CS01 isomer is the most stable isomer followed by GG-CS02, and GG-ZW01. GG-CS01 and GG-ZW01 have similar structures except that neutral glycine has turned to a zwitterionic form in GG-ZW01.

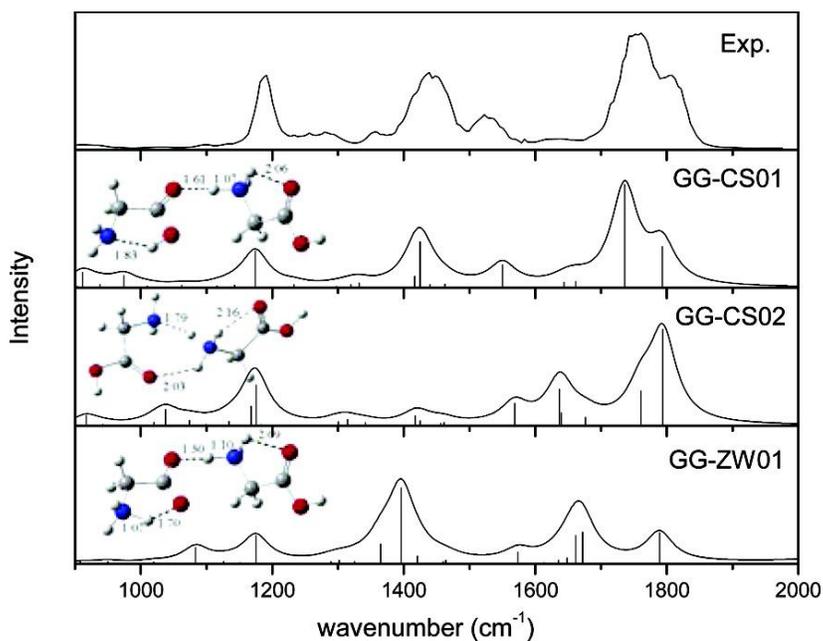


Figure 1.15. An illustration of the IRMPD spectra of $(\text{Gly})_2\text{H}^+$ compared to the computed IR spectra of the three lowest energy isomers. The GG-CS01 isomer fits better to the experimental IRMPD spectrum while coexistence of the other two isomers can be observed. Reprinted with permission from *J. Am. Chem. Soc.* **2007**, *129*, 4864-4865. Copyright (2007) American Chemical Society.⁷⁶

Comparison of experimental bands to the computed ones can lead to assigning GG-CS01 as the best match to the experimental spectrum resulting in being the largest contributor. The absence of a strong band centered at 1439 cm^{-1} , hydrogen-bonded O-H bending, can be considered as a strong evidence to rule out the GG-CS02 isomer as the main isomer which it was believed to be based on previous research⁷⁷ in the N-H stretching region, $2800\text{-}3700\text{ cm}^{-1}$. However, it is not possible to entirely rule out the existence of GG-CS02, and GG-ZW01. Based on further analyses, contributions of 5 and 1% has been assigned, respectively. Performing similar experiments on $(\text{Pro})_2\text{H}^+$ which more easily forms a zwitterionic tautomer produced some very interesting results. Figure 1.16 features the IRMPD spectra of $(\text{Pro})_2\text{H}^+$ which is compared to the computed lowest energy structures.

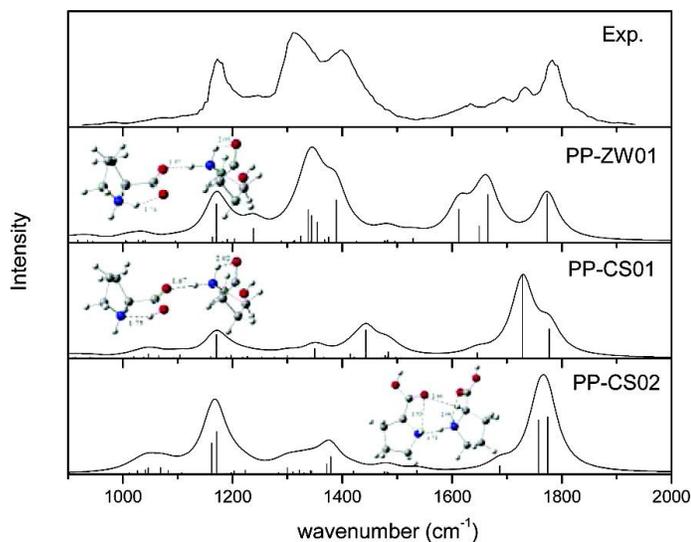


Figure 1.16. IRMPD spectra of $(\text{Pro})_2\text{H}^+$ together with the computed IR spectra of the three lowest energy isomers. The isomer containing a zwitterionic proline is the best match to the experimental IRMPD spectrum. The symmetric stretching vibration of the carboxylate group at 1396 cm^{-1} is considered the diagnostic band for zwitterionic structures. Reprinted with permission from *J. Am. Chem. Soc.* **2007**, *129*, 4864-4865. Copyright (2007) American Chemical Society.⁷⁶

Unlike $(\text{Gly})_2\text{H}^+$, the proton-bound proline IR spectrum better fits the experimental spectra and also it is 8 kJ mol^{-1} (298 K) more stable than the next stable isomer, PP-CS01. PP-ZW01 theoretical IR bands are greatly consistent with the experimental band and most importantly, the symmetric stretching vibration of the carboxylate group at 1396 cm^{-1} is the clearest evidence. This band, 1396 cm^{-1} , is considered the characteristic band used to diagnose the presence of a zwitterionic amino acid.^{66,78}

While IRMPD has proven itself as a powerful tool for structural studies of amino acids and small peptides, the extension of its capability to larger biomolecules such as proteins is limited. Proteins and enzymes possess a plethora of degrees of freedom leading to broad and very complex IRMPD spectra which may be difficult to interpret. One of earliest IR photodissociation studies on proteins was conducted by McLafferty and coworkers.⁷⁹ The odd-electron ions of bovine ubiquitin ions were photodissociated by IR photons resulting in a broad spectrum at 3350 cm^{-1} which was diagnostically unsuccessful. Later, a relatively resolved IRMPD spectrum of bovine cytochrome *c* was obtained by von Helden *et al.*⁸⁰ in the mid-IR region, $1400\text{-}1800\text{cm}^{-1}$, that exhibited some characteristic features for different charge states. Due to lack of any photodissociation products for the protonated cytochrome *c*, tagging with K^+ was performed that turned the protein into IRMPD detectable species losing K^+ during dissociation. The IRMPD spectra of different charge states of cytochrome *c* are depicted and compared in Figure 1.17. Three distinct bands are resolved in present spectra. The band around 1650 cm^{-1} , CO stretch, is routinely known as the diagnostic infrared signature of the secondary structure

of proteins in the solution phase. The position of this band is in the range of the CO stretch of alpha-helical structures in solution phase (1648-1658 cm^{-1}) which indicates the presence of an identical structure in gas phase. Upon an increase in charge state, a distinguishable blue-shift occurs which is attributed to presence of less hydrogen bonding.

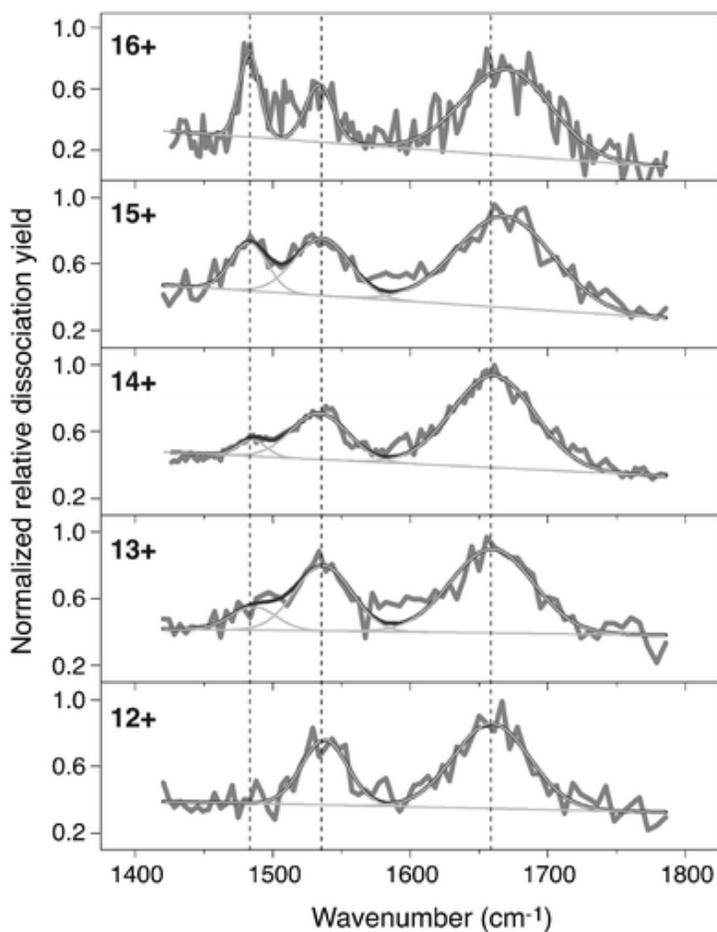


Figure 1.17. IRMPD spectra of the bovine ubiquitin in different charge states. The position of CO stretch at 1650 cm^{-1} is indicative of an alpha-helical structure which undergoes a blue shift upon increase in charge state. While the mid band at 1530 cm^{-1} remains unchanged, the increase in charge states causes a dramatic increase in intensity for the unknown peak at around 1490 cm^{-1} . Reprinted with permission from *Phys. Chem. Chem. Phys.* **2005**, 7, 1345-1348. Copyright (2005) Royal Society of Chemistry.⁸⁰

The NH bending band around 1530 cm^{-1} seems insensitive to the change in number of charges as it remains relatively constant in terms of intensity and position. Oppositely, the band centered at 1483 cm^{-1} dramatically increases in intensity as the charge state decreases. This band has not yet been assigned to any vibrational source due to the absence of any analogous band in the solution phase.

The physical chemistry of nucleobases and their larger clusters such as G-quadruplexes has been studied by IRMPD spectroscopy. Nucleobases can exist in several tautomeric forms with relatively similar thermochemical stabilities suggesting they may coexist. Determination of the most stable adenine tautomer was subject of an IRMPD spectroscopic study by Oomens and coworkers.⁸¹ Adenine can form a variety of tautomers so computational work was used to limit the possible structures considered by only choosing tautomers with relative Gibbs energies of less than 10 kJ mol^{-1} . Figure 1.18 shows the IRMPD spectrum of protonated adenine (AH^+) along with the IR spectra for the three lowest energy tautomers

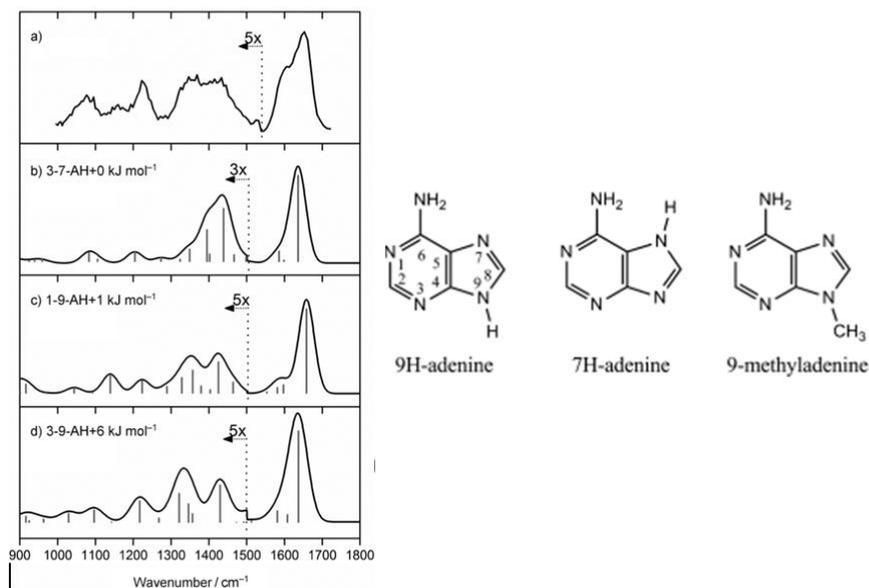


Figure 1.18. Left panel) IRMPD spectrum of protonated adenine compared to its three lowest energy tautomers. Right panel) the chemical structure of adenine tautomers and 9-methyladenine. Reprinted with permission from *ChemPhysChem* **2011**, *12*, 1921-1927. Copyright (2011) John Wiley and Sons.⁸¹

The IRMPD spectra exhibits a strong band centered at 1650 cm^{-1} with a left shoulder at 1605 cm^{-1} which all originate from the scissor mods of the NH_2 group. Few bands appearing below 1500 cm^{-1} can be assigned to a variety of in-plane bending modes. Comparison of the experimental spectrum to the computed bands, specifically the shoulder band at 1605 cm^{-1} and bands between $1300\text{-}1500\text{ cm}^{-1}$, reveals that 1-9-AH⁺ tautomer most likely the largest contributor to the IRMPD spectrum. However, considering the energy differences of tautomers (Figure 1.18) the coexistence of all three forms cannot be ruled out. Moreover, a complementary study by IR-UV ion-dip spectroscopy in the $525\text{ to }1750\text{ cm}^{-1}$ showed that the 9H tautomer is most likely the dominant tautomer in the neutral state.

In a separate study⁷², the tautomeric forms of protonated uracil (U), thymine (T) and cytosine (C) were probed at the Free Electron Laser facility of the Centre Laser Infrarouge d'Orsay (CLIO), France. Protonated bases generated by electrospray ionization were isolated in a quadruple ion trap and irradiated by a free electron laser in the 1000-2000 cm^{-1} range. The experimental IRMPD spectrum (Figure 1.19) for all three protonated bases matches the computed IR spectra of the enolic form. The IRMPD spectrum shows a very small band at 1800 cm^{-1} which no corresponding reproduction can be traced in the enolic form spectra. This band can be interpreted as the slight coexistence of the second lowest energy isomer called the oxo form.

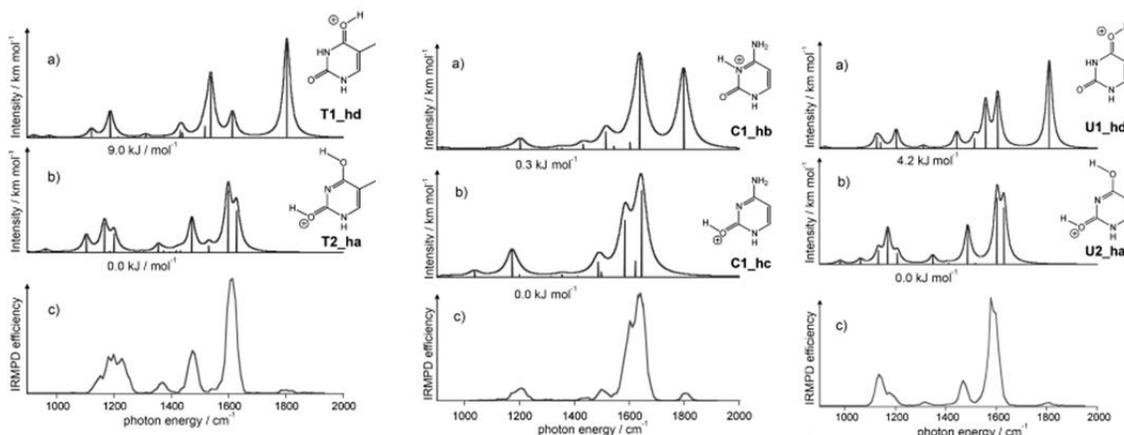


Figure 1.19. Comparisons of IRMPD spectra and computed IR spectra of two lowest energy isomers for three nucleobases (T, C, U). For all bases, the experimental spectrum is reproduced by the computed lowest energy isomer (the enol form). Reprinted with permission from *ChemPhysChem* **2007**, 8, 2235-2244. Copyright (2007) John Wiley and Sons.⁷²

Cisplatin (*cis*-[Pt(NH₃)₂Cl₂]) has been an effective drug to cure cancer and tumors where its function is established based on an interaction with guanine of genomic DNA.⁸²⁻⁸⁴ The coordination of cisplatin to guanine and the resultant complex can be probed by IRMPD spectroscopy.⁸⁵ Figure 1.20 illustrates the possible lowest energy structures for *cis*-[Pt(NH₃)₂(G)Cl]⁺ complex in which the structure (PtGN7-1) with guanine bound to platinum through N7 appears as the lowest energy isomer. This structure also establishes a hydrogen bond between the carbonyl oxygen at C6 and a hydrogen atom of one of NH₃ ligands, which is the only difference with the second lowest energy structure (PtGN7-2). The (PtGN7-2) isomer can be considered as an intermediate structure that quickly converts to (PtGN7-1) isomer. The IRMPD spectrum of *cis*-[Pt(NH₃)₂(G)Cl]⁺ in the 2900 to 3700 cm⁻¹ region along with the computed IR spectra of lowest energy structures are depicted in Figure 1.21. Except for an intensity mismatch below 3000 cm⁻¹, the experimental spectrum is reproduced by the lowest energy isomer (PtGN7-1) which proves that like the condensed phase, gas phase guanine platination occurs through N7.

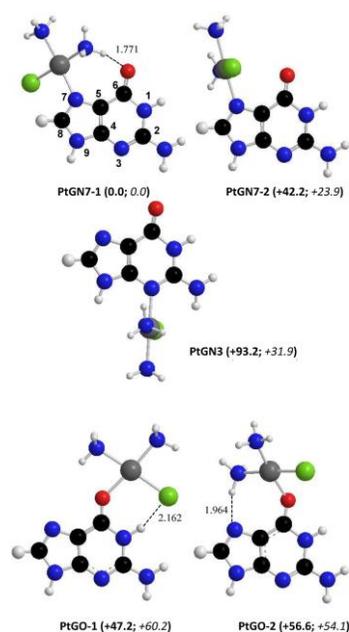


Figure 1.20. The Computed lowest energy structures of $cis\text{-}[\text{Pt}(\text{NH}_3)_2(\text{G})\text{Cl}]^+$ with relative free Gibbs energies (kJ mol^{-1} , italic numbers in parentheses). N7 is the preferred site for platination. Reprinted with permission from *J. Am. Chem. Soc.* **2013**, *135*, 1445-1455. Copyright © 2012 American Chemical Society.⁸⁵

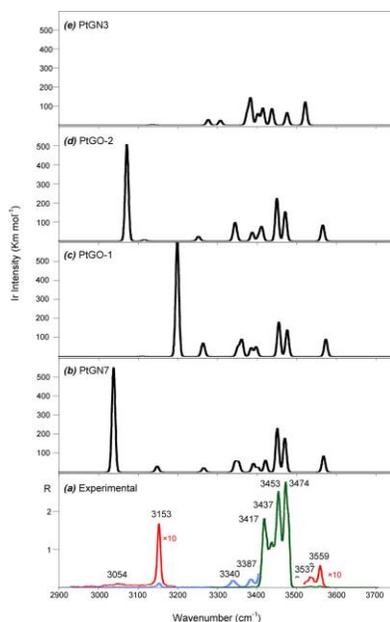


Figure 1.21. The IRMPD spectrum of $cis\text{-}[\text{Pt}(\text{NH}_3)_2(\text{G})\text{Cl}]^+$ compared to four lowest energy isomers. N7 is the preferred site for platination. Reprinted with permission from *J. Am. Chem. Soc.* **2013**, *135*, 1445-1455. Copyright © 2012 American Chemical Society.⁸⁵

A more realistic interaction of cisplatin to DNA has been examined by an IRMPD study of $cis\text{-[Pt(NH}_3)_2(5'\text{-dGMP-H})]^+$ and $cis\text{-[PtCl(NH}_3)_2(5'\text{-dGMP})]^+$ where 5-dGMP denotes the 2-deoxyguanosine-5-monophosphate molecule.⁸⁶ Computational work along with the spectral study of the 700-1900 cm^{-1} and 2800-3800 cm^{-1} region proved that $cis\text{-[Pt(NH}_3)_2(5'\text{dGMP-H})]^+$ forms a structure in which the metal simultaneously coordinates to the N7 of the guanine residue and an O atom on the phosphate group (Figure 1.22). On the other hand, $cis\text{-[PtCl(NH}_3)_2(5'\text{-dGMP})]^+$ ion structure forms a monodentate complex with a metal coordination to the N7 of the guanine residue. Both observations are in agreement with the NMR results obtained in the solution phase.

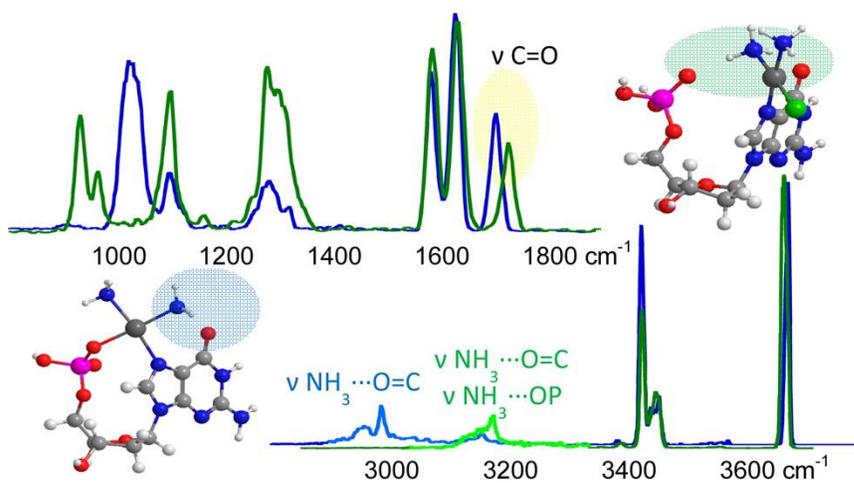


Figure 1.22. A comparison of the IRMPD spectrum of $cis\text{-[Pt(NH}_3)_2(5'\text{-dGMP-H})]^+$ (left structure with blue spectrum) and $cis\text{-[PtCl(NH}_3)_2(5'\text{-dGMP})]^+$ (right structure with green spectrum). The deprotonated ion shows a simultaneous attachment to the N7 of guanine and an O atom of the phosphate group with a characteristic hydrogen bond ($\text{NH}_3\cdots\text{O}=\text{C}$) with a band at 3000cm^{-1} . In $cis\text{-[PtCl(NH}_3)_2(5'\text{-dGMP})]^+$, the metal is exclusively linked to N7 of the guanine residue with two hydrogen bonds ($\text{NH}_3\cdots\text{O}=\text{C}$ and $\text{NH}_3\cdots\text{OP}$) appearing on IRMPD spectrum at $\approx 3200\text{ cm}^{-1}$. Reprinted with permission from *Inorg Chem* **2015**, 54, 3513-3522. Copyright 2015 American Chemical Society.⁸⁶

IRMPD spectroscopy is able to structurally probe larger adducts formed by nucleobases such as guanine tetrads (Figure 1.23).⁸⁷ G-tetrads are of great importance as they form at the guanine rich sections of DNA and are shown to perform very crucial biological functions.⁸⁸⁻⁹⁰ Four guanine molecules can configure differently around the metal forming various possible structures. The lowest energy structure(s) can be studied by IRMPD spectroscopy.

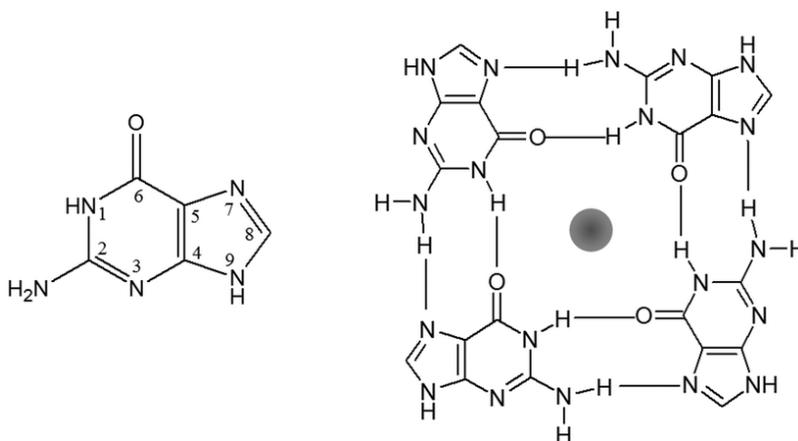


Figure 1.23. Illustration of a guanine molecule (left) and a guanine tetrad (right). The black sphere represents a sodium cation. Reprinted with permission from *Chem. Commun.* **2014**, 50, 14767-14770. Copyright 2014, Royal Society of Chemistry.⁸⁷

Fraschetti and coworkers⁸⁷ studied sodium G-tetrad conformers and concluded that Hoogsteen and Watson–Crick-type adducts are the most stable since establishing variety of hydrogen bonds. G-tetrad Hoogsteen type structures can be totally planar (D_{4h}) or slightly twisted (S_4) though their computed infrared spectra have been found to be almost identical. The IRMPD spectrum (in two regions) and the computed spectrum of the planar conformer are compared in Figure 1.24. The experimental spectrum is closely reproduced by the computed bands indicating that the G-tetrad conformer exhibits a

planar structure. Moreover, since the energy difference between the totally planar conformer and the slightly twisted structure is less than 2 kJ mol^{-1} , a possible coexistence has been offered.

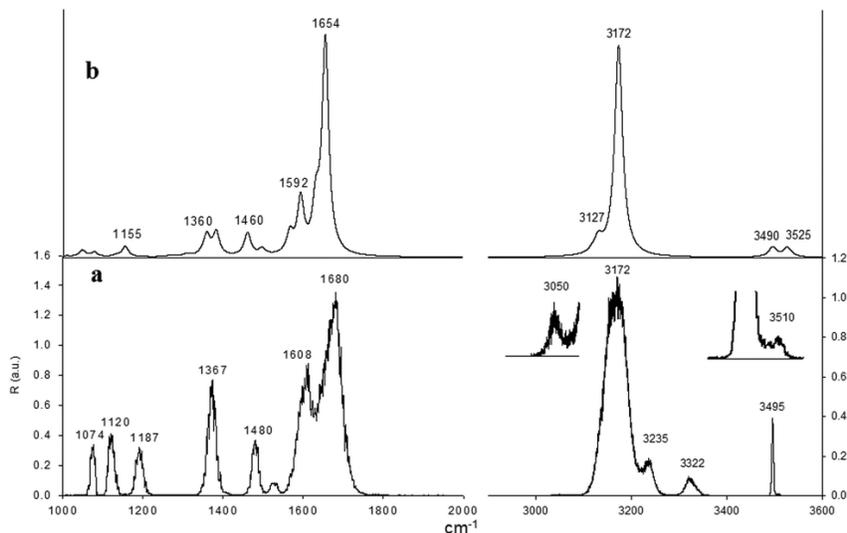


Figure 1.24. The IRMPD spectrum of sodium G-tetrad (in two regions) compared to calculated infrared spectrum of the lowest energy isomer. A good match is reproduced by the computational spectrum except two bands at 3235 and 3322 cm^{-1} which are absent in the computed spectra. These two bands can be related to overtones of bands at 1608 and 1680 cm^{-1} . Reprinted with permission from *Chem. Commun.* **2014**, *50*, 14767-14770. Copyright 2014, Royal Society of Chemistry.⁸⁷

1.2.3 BIRD Applications for Biological Systems

Blackbody infrared radiative dissociation was invented as a very slow heating activation technique that relies on the absorption of ambient blackbody photons by the ion of interest. A low pressure environment (10^{-10} mbar) guarantees lack of any translational activation which allows one to determine the thermal unimolecular dissociation rate constants (k_{uni}). This can be performed by a record of normalized intensity of parent ion and dissociation products as a function of time. A plot of $\text{Ln}(k_{\text{uni}})$

versus inverse temperature (T^{-1}) can result in determination of Arrhenius activation energies (E_0) which might lead to finding binding energies (discussed in Chapter 2).

BIRD is well suited to study macromolecules with weak noncovalent interactions since large molecules dissociate at a rate that is lower than the rate of energy exchange with the blackbody source.¹⁸ However, although the Arrhenius activation energy might meaningfully underestimate the small molecules binding energies, master equation modeling¹⁸ can be employed to correct the experimental discrepancy. Fridgen and coworkers⁹¹ hired BIRD to probe the binding energy and structure of small amino acid clusters, $[\text{Pb}(\text{Aa-H})\text{H}_2\text{O}]^+$ (Aa = Gly, Ala, Val, Leu, Ile, Phe, Glu, and Lys). Lead amino acid complexes can adopt different structures resulting in various monohydration configurations shown in Figure 1.25. Owing to this structural variation, the water binding energy is expected to show different values. As a result, BIRD kinetics can result in getting insight into the structure of the aforementioned clusters.

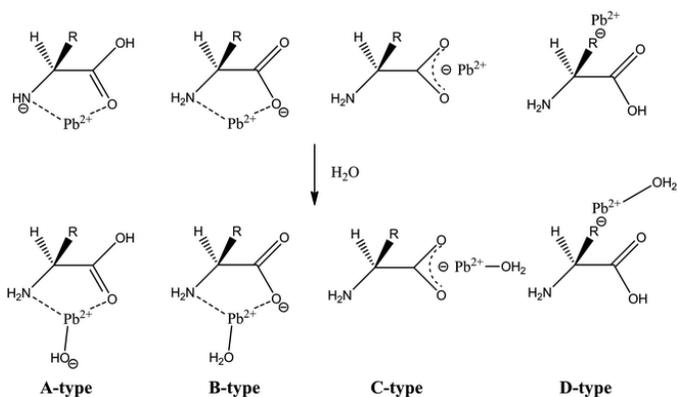


Figure 1.25. Different configuration resulted from complexation of lead to amino acids along with the corresponding hydrated clusters. Reprinted with permission from *Phys. Chem. Chem. Phys.* **2012**, *14*, 15118-15126. Copyright 2012 Royal Society of Chemistry.⁹¹

BIRD experiments on monohydrated amino acids yield activation energies ranging from 40-89 kJ mol⁻¹ which are summarized in Table 1.1. As explained before, experimental BIRD outputs of small molecules must be modeled by master equation modeling. Water dissociation energies outputted from master equation modeling (Table 1.1) present a meaningful underestimation up to 100%.

Table 1.1. Arrhenius activation energies (E_0) and master equation modeled threshold dissociation energies

[Pb(Aa-H)H ₂ O] ⁺ complex	E_a (kJ mol ⁻¹)	Master equation (kJ mol ⁻¹)
Gly	88.7 ± 12.1	113.6 ± 12.1
Ala	74.8 ± 6.9	105.3 ± 8.5
Val	68.9 ± 1.4	102.9 ± 1.4
Leu	68.1 ± 4.5	101.7 ± 4.8
Ile	65.4 ± 1.5	102.9 ± 1.5
Phe	36.6 ± 5.6	76.6 ± 5.6
Lys	58.0 ± 3.6	95.7 ± 3.6
Glu	42.3 ± 1.2	76.6 ± 1.2

Gly, Ala, Val, Leu, and Ile, as nonpolar amino acids, have been shown by IRMPD spectroscopy to adopt A type structures.^{75,92} This is also reflected in the threshold dissociation energies where all six amino acids exhibit dissociation energies greater than 100 kJ mol⁻¹. It is worth noting that A type isomers form an NH covalent bond (by transfer of a proton from a water molecule to the deprotonated amino group) upon hydration which demands a significant amount of energy for water removal.

BIRD has also been used to determine the stability of oligonucleotides.^{93,94} Oligonucleotides dissociate through a cleavage at the glycosidic bond resulting in the loss of a neutral nucleobase. The dissociation kinetics of several oligomers with seven nucleobases units (d(A)₇²⁻, d(AATTAAT)²⁻, d(TTAATTA)²⁻, d(T)₇²⁻, and

d(CCGGCCG)²⁻ have been determined by Williams and coworkers.⁹⁵ The zero-pressure limit activation parameters (Figure 1.26 and Table 1.2) show that d(A)₇²⁻, d(AATTAAT)²⁻, and d(TTAATTA)²⁻ dissociate at an average activation energy (E_a) of ~1.0 eV which arises from the loss of adenine in all cases. The d(CCGGCCG)²⁻ oligomer, on the other hand, undergoes detachment of both guanine and cytosine with Arrhenius parameters valued at $E_a = 1.32 \pm 0.03$ eV and $A = 10^{13.3 \pm 0.3} \text{ s}^{-1}$. Surprisingly, d(AATTAAT)²⁻, and d(TTAATTA)²⁻ show no thymine loss. Furthermore, d(T)₇²⁻ is observed to remain intact even at the highest temperature allowed by the instrument, 207°C, and a 600 s reaction delay. It can be inferred from the Arrhenius parameters that the identity of the nucleobase influences both the energetics and the entropy change of the process.

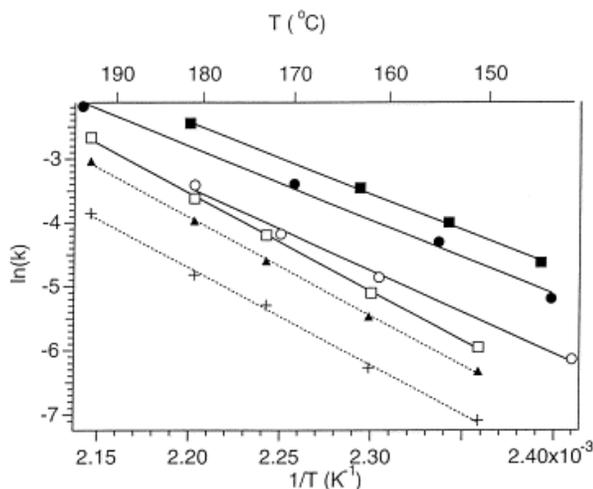


Figure 1.26. Arrhenius plots for a series of 7-mers oligonucleotides. From top to bottom, plot for loss of adenine from d(A₇)₂⁻ (filled square), d(AATTAAT)₂⁻ (filled circle), and d(TTAATTA)₂⁻ (open circle). The open square plot represents the dissociation of d(CCGGCCG)₂⁻ (both C and G dissociation). Plots for loss of neutral cytosine (filled triangle), and the loss of neutral guanine (plus), from d(CCGGCCG)₂⁻, are also established. Reprinted with permission from *J. Am. Soc. Mass Spectrom* **1998**, *9*, 1117-1124. Copyright 1998, American Society for Mass Spectrometry.⁹⁵

Table 1.2. Arrhenius parameters for loss from doubly deprotonated oligomers

Oligonucleotide	E_a (eV)	Log A	Primary Loss
d(A7) ₂ ⁻	1.03±0.05	10.3±0.6	-AH
d(AATTAAT) ₂ ⁻	1.00±0.07	9.8±0.8	-AH
d(TTAATTA) ₂ ⁻	1.12±0.06	10.9±0.8	-AH
d(T) ₇ ²⁻	-----	-----	-----
d(CCGGCCG) ₂ ⁻	1.32±0.03	13.3±0.3	-GH, -CH
→ -CH	1.35±0.02	13.1±0.2	-CH
→ -GH	1.31±0.04	12.5±0.3	-GH

BIRD analysis of complementary and noncomplementary DNA duplexes such as (TGCA)₂³⁻, (CCGG)₂³⁻, (AATTAAT)₂³⁻, (CCGGCCG)₂³⁻, A₇.T₇³⁻, A₇.A₇³⁻, T₇.T₇³⁻, and A₇.C₇³⁻ have evidenced the retention of Watson-Crick base pairing (WC) in complementary duplexes in the gas phase.⁹⁴ A₇.A₇³⁻ and T₇.T₇³⁻ with no WC base pairing undergo an extensive loss of neutral adenine while this process is impeded in A₇.T₇³⁻ due to presence of WC pairing. Additionally, the activation energy (E_0) for dissociation into single strands for A₇.A₇³⁻ and T₇.T₇³⁻ yields the same values (≈ 1.4 eV) whereas it shows a significantly higher E_0 for A₇.T₇³⁻ (1.68 eV). This extra stability in A₇.T₇³⁻ shows the increased interaction strength between strands due to existence of WC base pairing.

BIRD kinetics can be used as a tool to probe the nature of ligand coordination to metals. In work by Gillis *et al.*⁹⁶ BIRD was used along with computational chemistry to

examine calculated uracil clusters, U_nCa^{2+} ($n = 14-4$). BIRD activation energies and dissociation energies determined by master equation modeling exhibit an interesting size dependency. A summary of observed activation energies and dissociation energies as a function of cluster size are indicated in Figure 1.27. The results reveal that the threshold dissociation energy for U_6Ca^{2+} has a value significantly greater than energies obtained for larger clusters ($n = 7-14$). It can be inferred that the sixth uracil is the first uracil removed from the inner shell, first solvent shell, or in other words, Ca^{2+} coordinates directly to only six uracil molecules in the first solvation shell. Threshold dissociation energies for clusters with more than 6 uracil molecules showed an alternating low-high odd-even pattern. This finding was attributed to the fact that uracil molecules in the outer shell exist as dimers (Figure 1.28)

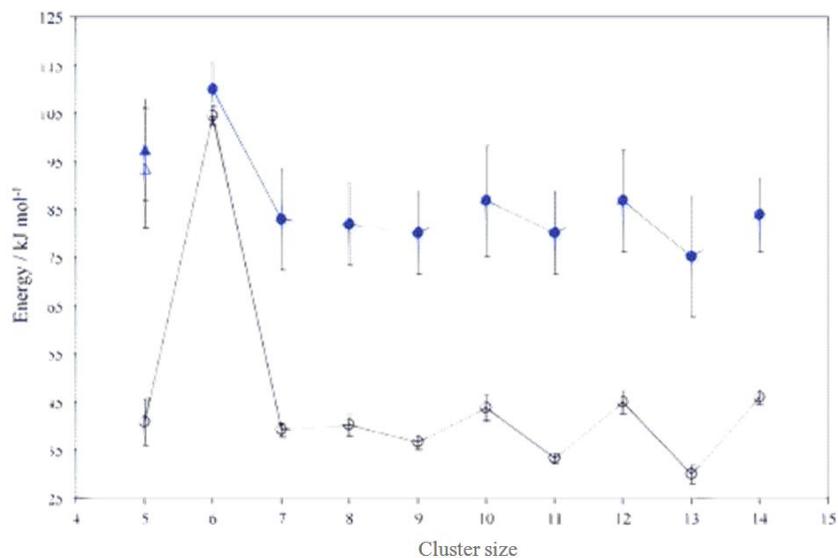


Figure 1.27. The observed activation energies (bottom with open circles) and threshold dissociation energies modeled by master equation modeling (upper curve with filled circles). These results suggest that the inner solvation shell is composed of six uracil molecules. Reprinted with permission from *Phys. Chem. Chem. Phys.* **2012**, *14*, 3304-3315. Copyright 2011 Royal Society of Chemistry.⁹⁶

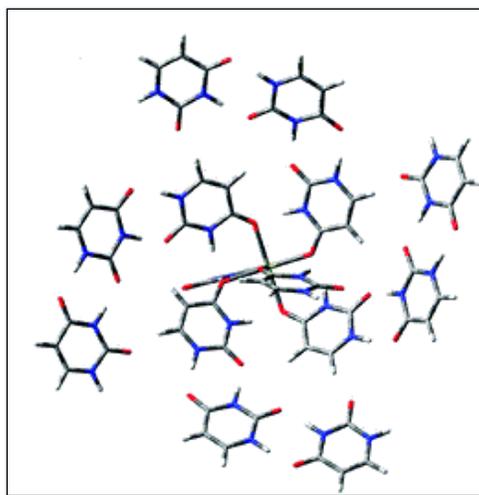


Figure 1.28. An illustration of $U_{14}Ca^{2+}$ cluster. The inner shell consists of 6 uracil molecules. Four pairs of dimeric uracil molecules, in the outer shell, are bound to the inner shell through hydrogen bonding. Reprinted with permission from *Phys. Chem. Chem. Phys.* **2012**, *14*, 3304-3315. Copyright 2011 Royal Society of Chemistry.⁹⁶

1.3 Objectives of Thesis

The main scope of this thesis is to study the intrinsic property of guanine adducts (alkali metal G-tetrads and G-quadruplexes) in the gas phase by using tandem mass spectrometry along with computational chemistry. The motivation behind experiments stems from the key role that guanine adducts play in biological systems.

In Chapter 3, the structure and energetics of the alkali metal cationized G-tetrads, $M(9-eG)_4^+$ ($M = Li, Na, K, Rb, Cs$; $9eG = 9\text{-ethylguanine}$), were probed through tandem mass spectrometry and computational chemistry. To study the gas phase structures, IRMPD experiments were performed in both the fingerprint region ($900\text{--}1850\text{ cm}^{-1}$) and the N-H/C-H stretching region ($2700\text{--}3800\text{ cm}^{-1}$). It was concluded that all five complexes form the metal-centered G-tetrad structure which is a square structure with the metal at the center. The gas phase stabilities of G-tetrads were studied by energy resolved CID showing that the totally

planar $\text{Na}(\text{9eG})_4^+$ was the most stable complex followed by the Li^+ , K^+ , Rb^+ , and Cs^+ G-tetrads in order. This stability order was reproduced by the computational results.

In Chapter 4, the intrinsic properties of self-assembled guanine quadruplexes, $\text{M}(\text{9eG})_8^+$ ($\text{M} = \text{Na}, \text{K}, \text{Rb}, \text{Cs}$) were studied by BIRD and computational chemistry. BIRD kinetics can shed light on the effect of metals on the stability and configuration of G-quadruplexes. In biological systems, K^+ is the ion of choice in guanine quadruplexes and the Arrhenius activation energy obtained from BIRD experiments showed potassium G-quadruplex to be the most stable adduct followed by $\text{K}^+ > \text{Rb}^+ > \text{Cs}^+ \gg \text{Na}^+$. The structural elucidation of alkali metal G-quadruplexes by IRMPD revealed spectroscopic features very similar to G-tetrads' which proves the G-quadruplexes to be formed from two G-tetrad subunits with a metal between in, a sandwich structure.

In Chapter 5, In Chapter 5, the structure and energetics of a larger G-quadruplex composed of three G-tetrads and two K^+ , $\text{K}_2(\text{9eG})_{12}^{2+}$, were studied. IRMPD spectroscopy in the N–H/C–H stretching region ($2700\text{--}3800\text{ cm}^{-1}$) showed that $\text{K}_2(\text{9eG})_{12}^{2+}$ configures in a sandwich-like structure with two G-tetrads sandwiching each of the two K^+ , similar to smaller G-quadruplexes. Energy-resolved SORI-CID and BIRD experiments were conducted on $\text{K}_2(\text{9eG})_{12}^{2+}$ to examine its binding energy and stability in the gas phase. SORI-CID experiments showed that $\text{K}_2(\text{9eG})_{12}^{2+}$ dissociates at a higher collisional energy compared to $\text{K}(\text{9eG})_8^+$. Results from BIRD experiments were consistent with SORI-CID experiments showing a greater activation energy for $\text{K}_2(\text{9eG})_{12}^{2+}$ relative to $\text{K}(\text{9eG})_8^+$.

References

- (1) Koster, H.; Tang, K.; Fu, D. J.; Braun, A.; van den Boom, D.; Smith, C. L.; Cotter, R. J.; Cantor, C. R.: A strategy for rapid and efficient DNA sequencing by mass spectrometry. *Nat biotechnol* **1996**, *14*, 1123-8.
- (2) Booth, M. J.; Branco, M. R.; Ficiz, G.; Oxley, D.; Krueger, F.; Reik, W.; Balasubramanian, S.: Quantitative sequencing of 5-methylcytosine and 5-hydroxymethylcytosine at single-base resolution. *Science* **2012**, *336*, 934-937.
- (3) Aebersold, R.; Mann, M.: Mass spectrometry-based proteomics. *Nature* **2003**, *422*, 198-207.
- (4) Wilhelm, M.; Schlegl, J.; Hahne, H.; Gholami, A. M.; Lieberenz, M.; Savitski, M. M.; Ziegler, E.; Butzmann, L.; Gessulat, S.; Marx, H.; Mathieson, T.; Lemeer, S.; Schnatbaum, K.; Reimer, U.; Wenschuh, H.; Mollenhauer, M.; Slotta-Huspenina, J.; Boese, J.-H.; Bantscheff, M.; Gerstmair, A.; Faerber, F.; Kuster, B.: Mass-spectrometry-based draft of the human proteome. *Nature* **2014**, *509*, 582-587.
- (5) Chambers, M. C.; Maclean, B.; Burke, R.; Amodei, D.; Ruderman, D. L.; Neumann, S.; Gatto, L.; Fischer, B.; Pratt, B.; Egertson, J.; Hoff, K.; Kessner, D.; Tasman, N.; Shulman, N.; Frewen, B.; Baker, T. A.; Brusniak, M.-Y.; Paulse, C.; Creasy, D.; Flashner, L.; Kani, K.; Moulding, C.; Seymour, S. L.; Nuwaysir, L. M.; Lefebvre, B.; Kuhlmann, F.; Roark, J.; Rainer, P.; Detlev, S.; Hemenway, T.; Huhmer, A.; Langridge, J.; Connolly, B.; Chadick, T.; Holly, K.; Eckels, J.; Deutsch, E. W.; Moritz, R. L.; Katz, J. E.; Agus, D. B.; MacCoss, M.; Tabb, D. L.; Mallick, P.: A cross-platform toolkit for mass spectrometry and proteomics. *Nat Biotech* **2012**, *30*, 918-920.
- (6) Miladinović, S. M.; Kozhinov, A. N.; Gorshkov, M. V.; Tsybin, Y. O.: On the utility of isotopic fine structure mass spectrometry in protein identification. *Anal.Chem.* **2012**, *84*, 4042-4051.
- (7) Nikolaev, E. N.; Jertz, R.; Grigoryev, A.; Baykut, G.: Fine structure in isotopic peak distributions measured using a dynamically harmonized Fourier Transform ion cyclotron resonance Cell at 7 T. *Anal. Chem* **2012**, *84*, 2275-2283.

- (8) Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M.: Electrospray ionization for mass spectrometry of large biomolecules. *Science* **1989**, *246*, 64-71.
- (9) Karas, M.; Bachmann, D.; Bahr, U.; Hillenkamp, F.: Matrix-assisted ultraviolet laser desorption of non-volatile compounds. *Int. J. Mass Spectrom* **1987**, *78*, 53-68.
- (10) Loo, J. A.: Electrospray ionization mass spectrometry: a technology for studying noncovalent macromolecular complexes. *Int. J. Mass Spectrom* **2000**, *200*, 175-186.
- (11) McLafferty, F.: Tandem mass spectrometry. *Science* **1981**, *214*, 280-287.
- (12) Nesvizhskii, A. I.; Keller, A.; Kolker, E.; Aebersold, R.: A statistical model for identifying proteins by tandem mass spectrometry. *Anal. Chem.* **2003**, *75*, 4646-4658.
- (13) Taylor, P. J.: Matrix effects: the Achilles heel of quantitative high-performance liquid chromatography–electrospray–tandem mass spectrometry. *Clin. Biochem.* **2005**, *38*, 328-334.
- (14) Yost, R. A.; Fetterolf, D. D.: Tandem mass spectrometry (MS/MS) instrumentation. *Mass Spectrom. Rev.* **1983**, *2*, 1-45.
- (15) Cowan, T. M.: Neonatal screening by tandem mass spectrometry. *NeoReviews* **2005**, *6*, e539-e548.
- (16) Marshall, A. G.; Hendrickson, C. L.; Jackson, G. S.: Fourier transform ion cyclotron resonance mass spectrometry: A primer. *Mass Spectrom. Rev.* **1998**, *17*, 1-35.
- (17) Sleno, L.; Volmer, D. A.: Ion activation methods for tandem mass spectrometry. *J. Mass Spectrom.* **2004**, *39*, 1091-1112.
- (18) Dunbar, R. C.: BIRD (blackbody infrared radiative dissociation): Evolution, principles, and applications. *Mass Spectrom. Rev.* **2004**, *23*, 127-158.
- (19) Zubarev, R. A.: Electron-capture dissociation tandem mass spectrometry. *Curr. Opin. in Biotechnol.* **2004**, *15*, 12-16.

- (20) Grill, V.; Shen, J.; Evans, C.; Cooks, R. G.: Collisions of ions with surfaces at chemically relevant energies: Instrumentation and phenomena. *Rev. Sci. Instrum.* **2001**, *72*, 3149-3179.
- (21) Mabud, M. A.; Dekrey, M. J.; Cooks, R. G.: Surface-induced dissociation of molecular ions. *Int. J. Mass Spectrom* **1985**, *67*, 285-294.
- (22) Bier, M.; Amy, J.; Cooks, R.; Syka, J.; Ceja, P.; Stafford, G.: A tandem quadrupole mass spectrometer for the study of surface-induced dissociation. *Int. J. Mass Spectrom* **1987**, *77*, 31-47.
- (23) Wysocki, V. H.; Ding, J.-M.; Jones, J. L.; Callahan, J. H.; King, F. L.: Surface-induced dissociation in tandem quadrupole mass spectrometers: a comparison of three designs. *J. Am. Soc. Mass Spectrom.* **1992**, *3*, 27-32.
- (24) Lammert, S.; Cooks, R.: Surface-induced dissociation of molecular ions in a quadrupole ion trap mass spectrometer. *J. Am. Soc. Mass Spectrom.* **1991**, *2*, 487-491.
- (25) Williams, E. R.; Fang, L.; Zare, R. N.: Surface induced dissociation for tandem time-of-flight mass spectrometry. *Int. J. Mass Spectrom* **1993**, *123*, 233-241.
- (26) Stone, E.; Gillig, K. J.; Ruotolo, B.; Fuhrer, K.; Gonin, M.; Schultz, A.; Russell, D. H.: Surface-induced dissociation on a MALDI-ion mobility-orthogonal time-of-flight mass spectrometer: sequencing peptides from an “in-solution” protein digest. *Anal. Chem.* **2001**, *73*, 2233-2238.
- (27) Williams, E. R.; Zare, R. N.: Surface induced dissociation with reflectron time-of-flight mass spectrometry. Google Patents, 1992.
- (28) Williams, E. R.; Henry, K. D.; McLafferty, F. W.; Shabanowitz, J.; Hunt, D. F.: Surface-induced dissociation of peptide ions in Fourier-transform mass spectrometry. *J. Am. Soc. Mass Spectrom.* **1990**, *1*, 413-416.
- (29) Fernandez, F. M.; Wysocki, V. H.; Futrell, J. H.; Laskin, J.: Protein identification via surface-induced dissociation in an FT-ICR mass spectrometer and a patchwork sequencing approach. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 700-709.

(30) Laskin, J.; Denisov, E.; Futrell, J.: Comparative study of collision-induced and surface-induced dissociation. 2. Fragmentation of small alanine-containing peptides in FT-ICR MS. *J. Phys. Chem. B* **2001**, *105*, 1895-1900.

(31) Chorush, R. A.; Little, D. P.; Beu, S. C.; Wood, T. D.; McLafferty, F. W.: Surface-induced dissociation of multiply-protonated proteins. *Anal. Chem.* **1995**, *67*, 1042-1046.

(32) Wysocki, V. H.; Jones, C. M.; Galhena, A. S.; Blackwell, A. E.: Surface-induced dissociation shows potential to be more informative than collision-induced dissociation for structural studies of large systems. *J. Am. Soc. Mass Spectrom.* **2008**, *19*, 903-913.

(33) Hayakawa, S.; Feng, B.; Cooks, R. G.: Differentiation of isomeric $C_3H_4^+$ ions and T \rightarrow V energy partitioning in surface-induced dissociation. *Int. J. Mass Spectrom* **1997**, *167*, 525-539.

(34) Hayward, M. J.; Mabud, M. A.; Cooks, R. G.: Ion/surface collisions for distinction of isomeric $[C_6H_6].\text{cndot.}^+$ and $[C_6H_6]^{2+}$ ions. *J. Am. Chem. Soc.* **1988**, *110*, 1343-1346.

(35) Schaaf, T.; Qu, Y.; Farrell, N.; Wysocki, V. H.: Investigation of the trans effect in the fragmentation of dinuclear platinum complexes by electrospray ionization surface-induced dissociation tandem mass spectrometry. *J. Mass Spectrom.* **1998**, *33*, 436-443.

(36) Jones, C. M.; Beardsley, R. L.; Galhena, A. S.; Dagan, S.; Cheng, G.; Wysocki, V. H.: Symmetrical gas-phase dissociation of noncovalent protein complexes via surface collisions. *J. Am. Chem. Soc.* **2006**, *128*, 15044-15045.

(37) Wysocki, V. H.; Joyce, K. E.; Jones, C. M.; Beardsley, R. L.: Surface-induced dissociation of small molecules, peptides, and non-covalent protein complexes. *J. Am. Soc. Mass Spectrom.* **2008**, *19*, 190-208.

(38) McLafferty, F. W.; Tureček, F.: *Interpretation of mass spectra*; University science books, 1993.

- (39) Wysocki, V. H.; Kenttämaa, H. I.; Cooks, R. G.: Internal energy distributions of isolated ions after activation by various methods. *Int. J. Mass Spectrom* **1987**, *75*, 181-208.
- (40) Ali, O. Y.; Randell, N. M.; Fridgen, T. D.: Primary fragmentation pathways of gas phase $[M(\text{Uracil-H})(\text{Uracil})]^+$ Complexes (M=Zn, Cu, Ni, Co, Fe, Mn, Cd, Pd, Mg, Ca, Sr, Ba, and Pb): loss of uracil versus HNCO. *ChemPhysChem* **2012**, *13*, 1507-1513.
- (41) Salpin, J.-Y.; Gamiette, L.; Tortajada, J.; Besson, T.; Maître, P.: Structure of $\text{Pb}^{2+}/\text{dCMP}$ and $\text{Pb}^{2+}/\text{CMP}$ complexes as characterized by tandem mass spectrometry and IRMPD spectroscopy. *Int. J. Mass Spectrom* **2011**, *304*, 154-164.
- (42) Dunbar, R. C.; McMahon, T. B.; Thdlmann, D.; Tanner, D. S.: zero-pressure thermal-radiation-induced dissociation of gas-phase cluster Ions: comparison of theory and experiment. *J. Am. Chem. Soc* **1995**, *117*, 12819-12825.
- (43) Price, W. D.; Schnier, P. D.; Jockusch, R. A.; Strittmatter, E. F.; Williams, E. R.: Unimolecular reaction kinetics in the high-pressure limit without collisions. *J. Am. Chem. Soc.* **1996**, *118*, 10640-10644.
- (44) Wong, R. L.; Paech, K.; Williams, E. R.: Blackbody infrared radiative dissociation at low temperature: Hydration of $\text{X}_2^+(\text{H}_2\text{O})_n$, for X= Mg, Ca. *Int. J. Mass Spectrom* **2004**, *232*, 59-66.
- (45) Gross, D. S.; Zhao, Y.; Williams, E. R.: Dissociation of heme-globin complexes by blackbody infrared radiative dissociation: Molecular specificity in the gas phase? *J. Am. Soc. Mass Spectrom.* **1997**, *8*, 519-524.
- (46) Harper, B.; Neumann, E. K.; Solouki, T.: DNA oligonucleotide fragment ion rearrangements upon collision-induced dissociation. *J. Am. Soc. Mass Spectrom.* **2015**, *26*, 1404-1413.
- (47) Gao, Y.; McLuckey, S. A.: Electron transfer followed by collision-induced dissociation (NET-CID) for generating sequence information from backbone-modified oligonucleotide anions. *Rapid Commun. Mass Spectrom.* **2013**, *27*, 249-257.

(48) Gao, Y.; Yang, J.; Cancilla, M. T.; Meng, F.; McLuckey, S. A.: Top-down interrogation of chemically modified oligonucleotides by negative electron transfer and collision induced dissociation. *Anal. Chem.* **2013**, *85*, 4713-4720.

(49) Ortiz, D.; Martin-Gago, P.; Riera, A.; Song, K.; Salpin, J.-Y.; Spezia, R.: Gas-phase collision induced dissociation mechanisms of peptides: Theoretical and experimental study of N-formylalanyl amide fragmentation. *Int. J. Mass Spectrom* **2013**, *335*, 33-44.

(50) Williams, J. P.; Brown, J. M.; Campuzano, I.; Sadler, P. J.: Identifying drug metallation sites on peptides using electron transfer dissociation (ETD), collision induced dissociation (CID) and ion mobility-mass spectrometry (IM-MS). *Chem. Commun.* **2010**, *46*, 5458-5460.

(51) Kleinova, M.; Belgacem, O.; Pock, K.; Rizzi, A.; Buchacher, A.; Allmaier, G.: Characterization of cysteinylolation of pharmaceutical-grade human serum albumin by electrospray ionization mass spectrometry and low-energy collision-induced dissociation tandem mass spectrometry. *Rapid Commun. Mass Spectrom.* **2005**, *19*, 2965-2973.

(52) Pike, S. D.; Pernik, I.; Theron, R.; McIndoe, J. S.; Weller, A. S.: Relative binding affinities of fluorobenzene ligands in cationic rhodium bisphosphine η^6 -fluorobenzene complexes probed using collision-induced dissociation. *J. Organomet. Chem.* **2015**, *784*, 75-83.

(53) Altuntas, E.; Weber, C.; Schubert, U. S.: Detailed characterization of poly(2-ethyl-2oxazoline)s by energy variable collision-induced dissociation study. *Rapid Commun. Mass Spectrom.* **2013**, *27*, 1095-1100.

(54) Fouquet, T.; Chendo, C.; Toniazzo, V.; Ruch, D.; Charles, L.: Collision-induced dissociation of synthetic polymers containing hydride groups: the case of poly(methylhydrosiloxane) homopolymers and poly(methylhydrosiloxane)-co-(dimethylsiloxane) copolymers. *Rapid Commun. Mass Spectrom.* **2013**, *27*, 88-96.

- (55) Zhang, W.; Quernheim, M.; Räder, H. J.; Müllen, K.: Collision-induced dissociation ion mobility mass spectrometry for the elucidation of unknown structures in strained polycyclic aromatic hydrocarbon macrocycles. *Anal. Chem.* **2016**, *88*, 952-959.
- (56) Ali, O. Y.; Fridgen, T. D.: Structures and fragmentation of [Cu(Uracil-H)(Uracil)]⁺ in the gas phase. *ChemPhysChem* **2012**, *13*, 588-596.
- (57) Hunt, D. F.; Yates, J. R.; Shabanowitz, J.; Winston, S.; Hauer, C. R.: Protein sequencing by tandem mass spectrometry. *Proc. Natl. Acad. Sci.* **1986**, *83*, 6233-6237.
- (58) Biemann, K.: Mass spectrometry of peptides and proteins. *Annu. Rev. Biochem.* **1992**, *61*, 977-1010.
- (59) Yates, J. R.; Morgan, S. F.; Gatlin, C. L.; Griffin, P. R.; Eng, J. K.: Method to compare collision-induced dissociation spectra of peptides: potential for library searching and subtractive analysis. *Anal. Chem.* **1998**, *70*, 3557-3565.
- (60) Yates, J. R.; Eng, J. K.; McCormack, A. L.; Schieltz, D.: Method to correlate tandem mass spectra of modified peptides to amino acid sequences in the protein database. *Anal. Chem.* **1995**, *67*, 1426-1436.
- (61) Mikesch, L. M.; Ueberheide, B.; Chi, A.; Coon, J. J.; Syka, J. E. P.; Shabanowitz, J.; Hunt, D. F.: The utility of ETD mass spectrometry in proteomic analysis. *Biochim. Biophys. Acta* **2006**, *1764*, 1811-1822.
- (62) Roepstorff, P.; Fohlman, J.: Letter to the editors. *Biol. Mass Spectrom.* **1984**, *11*, 601-601.
- (63) Smith, B. C.: Fourier transform infrared spectroscopy. *CRC, Boca Raton, FL* **1996**.
- (64) Rao, C. N. R.: Chemical applications of infrared spectroscopy. **1963**.
- (65) Polfer, N. C.: Infrared multiple photon dissociation spectroscopy of trapped ions. *Chem. Soc. Rev.* **2011**, *40*, 2211-2221.
- (66) Kapota, C.; Lemaire, J.; Maître, P.; Ohanessian, G.: Vibrational signature of charge solvation vs salt bridge isomers of sodiated amino acids in the gas phase. *J. Am. Chem. Soc.* **2004**, *126*, 1836-1842.

- (67) Bush, M. F.; Oomens, J.; Saykally, R. J.; Williams, E. R.: Effects of alkaline earth metal ion complexation on amino acid zwitterion stability: results from infrared action spectroscopy. *J. Am. Chem. Soc.* **2008**, *130*, 6463-6471.
- (68) Wu, R.; McMahon, T. B.: An investigation of protonation sites and conformations of protonated amino acids by IRMPD spectroscopy. *ChemPhysChem* **2008**, *9*, 2826-2835.
- (69) Prell, J. S.; Flick, T. G.; Oomens, J.; Berden, G.; Williams, E. R.: Coordination of trivalent metal cations to peptides: results from IRMPD spectroscopy and theory. *J Phys Chem A* **2009**, *114*, 854-860.
- (70) Wu, R.; McMahon, T. B.: Protonation sites and conformations of peptides of glycine (Gly1- 5H⁺) by IRMPD spectroscopy. *J. Phys. Chem. B* **2009**, *113*, 8767-8775.
- (71) Polfer, N. C.; Oomens, J.; Dunbar, R. C.: Alkali metal complexes of the dipeptides PheAla and AlaPhe: IRMPD spectroscopy. *ChemPhysChem* **2008**, *9*, 579-589.
- (72) Salpin, J. Y.; Guillaumont, S.; Tortajada, J.; MacAleese, L.; Lemaire, J.; Maitre, P.: Infrared spectra of protonated uracil, thymine and cytosine. *ChemPhysChem* **2007**, *8*, 2235-2244.
- (73) Oomens, J.; Moehlig, A. R.; Morton, T. H.: Infrared multiple photon dissociation (IRMPD) spectroscopy of the proton-bound dimer of 1-methylcytosine in the gas phase. *J. Phys. Chem. Lett.* **2010**, *1*, 2891-2897.
- (74) Gillis, E. A.; Rajabi, K.; Fridgen, T. D.: Structures of hydrated Li⁺-Thymine and Li⁺-uracil complexes by IRMPD spectroscopy in the N-H/O-H stretching region. *J. Phys. Chem. A* **2008**, *113*, 824-832.
- (75) Burt, M. B.; Decker, S. G. A.; Atkins, C. G.; Rowsell, M.; Peremans, A.; Fridgen, T. D.: Structures of Bare and Hydrated [Pb(AminoAcid-H)]⁺ Complexes using infrared multiple photon dissociation spectroscopy. *J. Phys. Chem. B* **2011**, *115*, 11506-11518.

- (76) Wu, R.; McMahon, T. B.: Infrared multiple photon dissociation spectra of proline and glycine proton-bound homodimers. Evidence for zwitterionic structure. *J. Am. Chem. Soc.* **2007**, *129*, 4864-4865.
- (77) Oh, H.-B.; Lin, C.; Hwang, H. Y.; Zhai, H.; Breuker, K.; Zabrouskov, V.; Carpenter, B. K.; McLafferty, F. W.: Infrared photodissociation spectroscopy of electrosprayed ions in a Fourier Transform mass spectrometer. *J. Am. Chem. Soc.* **2005**, *127*, 4076-4083.
- (78) Chen, X. G.; Li, P.; Holtz, J. S. W.; Chi, Z.; Pajcini, V.; Asher, S. A.; Kelly, L. A.: Resonance raman examination of the electronic excited states of glycyglycine and other dipeptides: observation of a carboxylate→amide charge transfer transition. *J. Am. Chem. Soc.* **1996**, *118*, 9705-9715.
- (79) Oh, H.; Breuker, K.; Sze, S. K.; Ge, Y.; Carpenter, B. K.; McLafferty, F. W.: Secondary and tertiary structures of gaseous protein ions characterized by electron capture dissociation mass spectrometry and photofragment spectroscopy. *Proc. Natl. Acad. Sci.* **2002**, *99*, 15863-15868.
- (80) Oomens, J.; Polfer, N.; Moore, D. T.; van der Meer, L.; Marshall, A. G.; Eyler, J. R.; Meijer, G.; von Helden, G.: Charge-state resolved mid-infrared spectroscopy of a gas-phase protein. *Phys. Chem. Chem. Phys.* **2005**, *7*, 1345-1348.
- (81) van Zundert, G. C.; Jaqx, S.; Berden, G.; Bakker, J. M.; Kleinermanns, K.; Oomens, J.; Rijs, A. M.: IR Spectroscopy of isolated neutral and protonated adenine and 9-methyladenine. *ChemPhysChem* **2011**, *12*, 1921-1927.
- (82) Group, I. A. L. C. T. C.: Cisplatin-based adjuvant chemotherapy in patients with completely resected non-small-cell lung cancer. *N Engl J Med* **2004**, *2004*, 351-360.
- (83) Khuri, F. R.; Nemunaitis, J.; Ganly, I.; Arseneau, J.; Tannock, I. F.; Romel, L.; Gore, M.; Ironside, J.; MacDougall, R.; Heise, C.: A controlled trial of intratumoral ONYX-015, a selectively-replicating adenovirus, in combination with cisplatin and 5-fluorouracil in patients with recurrent head and neck cancer. *Nat. Med.* **2000**, *6*, 879-885.

- (84) Scagliotti, G. V.; Parikh, P.; Von Pawel, J.; Biesma, B.; Vansteenkiste, J.; Manegold, C.; Serwatowski, P.; Gatzemeier, U.; Digumarti, R.; Zukin, M.: Phase III study comparing cisplatin plus gemcitabine with cisplatin plus pemetrexed in chemotherapy-naive patients with advanced-stage non-small-cell lung cancer. *J. Clin. Oncol.* **2008**, *26*, 3543-3551.
- (85) Chiavarino, B.; Crestoni, M. E.; Fornarini, S.; Scuderi, D.; Salpin, J.-Y.: Interaction of cisplatin with adenine and guanine: a combined IRMPD, MS/MS, and theoretical study. *J. Am. Chem. Soc.* **2013**, *135*, 1445-1455.
- (86) Chiavarino, B.; Crestoni, M. E.; Fornarini, S.; Scuderi, D.; Salpin, J.-Y.: Interaction of cisplatin with 5'-dGMP: a combined IRMPD and theoretical study. *Inorg Chem* **2015**, *54*, 3513-3522.
- (87) Frascchetti, C.; Montagna, M.; Guarcini, L.; Guidoni, L.; Filippi, A.: Spectroscopic evidence for a gas-phase librating G-quartet- Na^+ complex. *Chem. Commun.* **2014**, *50*, 14767-14770.
- (88) Hsu, S.-T. D.; Varnai, P.; Bugaut, A.; Reszka, A. P.; Neidle, S.; Balasubramanian, S.: A G-rich sequence within the c-kit oncogene promoter forms a parallel G-quadruplex having asymmetric G-tetrad dynamics. *J. Am. Chem. Soc.* **2009**, *131*, 13399-13409.
- (89) Benimetskaya, L.; Berton, M.; Kolbanovsky, A.; Benimetsky, S.; Stein, C. A.: Formation of a G-tetrad and higher order structures correlates with biological activity of the RelA (NF- κ B p65) 'antisense' oligodeoxynucleotide. *Nucleic Acids Res.* **1997**, *25*, 2648-2656.
- (90) Jing, N.; Marchand, C.; Liu, J.; Mitra, R.; Hogan, M. E.; Pommier, Y.: Mechanism of Inhibition of HIV-1 Integrase by G-tetrad-forming Oligonucleotides in Vitro. *J. Biol. Chem.* **2000**, *275*, 21460-21467.
- (91) Burt, M. B.; Decker, S. G.; Fridgen, T. D.: Water binding energies of $[\text{Pb}(\text{amino acid-H})\text{H}_2\text{O}]^+$ complexes determined by blackbody infrared radiative dissociation. *Phys. Chem. Chem. Phys.* **2012**, *14*, 15118-15126.

(92) Atkins, C. G.; Banu, L.; Rowsell, M.; Blagojevic, V.; Bohme, D. K.; Fridgen, T. D.: Structure of [Pb(Gly-H)]⁺ and the monosolvated water and methanol solvated species by infrared multiple-photon dissociation spectroscopy, energy-resolved collision-induced dissociation, and electronic structure calculations. *J. Phys. Chem. B* **2009**, *113*, 14457-14464.

(93) Daneshfar, R.; Klassen, J. S.: Arrhenius activation parameters for the loss of neutral nucleobases from deprotonated oligonucleotide anions in the gas phase. *J. Am. Soc. Mass Spectrom.* **2004**, *15*, 55-64.

(94) Schnier, P. D.; Klassen, J. S.; Strittmatter, E. F.; Williams, E. R.: Activation energies for dissociation of double strand oligonucleotide anions: evidence for Watson–Crick base pairing in Vacuo. *J. Am. Chem. Soc.* **1998**, *120*, 9605-9613.

(95) Klassen, J. S.; Schnier, P. D.; Williams, E. R.: Blackbody infrared radiative dissociation of oligonucleotide anions. *J. Am. Soc. Mass Spectrom.* **1998**, *9*, 1117-1124.

(96) Gillis, E. A. L.; Demireva, M.; Nanda, K.; Beran, G.; Williams, E. R.; Fridgen, T. D.: Structures and energetics of electrosprayed uraciln Ca²⁺ clusters (n = 14-4) in the gas phase. *Phys. Chem. Chem. Phys.* **2012**, *14*, 3304-3315.

Chapter 2. Experimental and Theory

Mass spectrometry has opened up the possibility to shed light on the intrinsic properties—in the absence of solvent—of biologically relevant ions by their detailed study in the gas phase. These studies can result in precious information about biomolecule structures, energetics, and reactivities.¹⁻³ Due to its versatility and capability to incorporate a variety of ion activation techniques, mass spectrometry is an ideal technique for the study of gaseous ions. Ion trap mass spectrometers are particularly useful since ions can be isolated for an extended period which enables one to perform further experiments. Most mass spectrometry experiments involve at least four steps including ionization, trapping, ion activation, and finally, data analysis. This chapter aims to explain and discuss the instrumentation and methods used to get insights into the physical chemistries of guanine cluster ions.

2.1. Instrumentation

The present work was performed using a Bruker Apex-Qe 7.0 T FT-ICR mass spectrometer in the Laboratory for the Study of the Energetics, Structures, and Reactivities of Gas Phase Ions at Memorial University which is depicted in Figure 2.1.

A more detailed schematic of the instrument is illustrated in Figure 2.2. The instrument is composed of four distinct parts including electrospray ionization (ESI) source, quadruple/hexapole region, ion transfer optics, and the ICR cell which is the heart of the instrument.



Figure 2.1. Bruker Apex-Qe7.0 T Fourier transform ion cyclotron mass spectrometer FT-ICR MS coupled to OPO laser located at Memorial University.

Ions are generated in the ESI source which is held at a relatively high pressure (3 mbar) compared to the other subunits of the instrument. The ions formed are then guided into the quadrupole mass filters or ion guide (10^{-6} mbar) where the ion beam can be mass selected and then transferred into the hexapole collision/storage cell where collision induced dissociation or microsolvation of ions occurs at pressures of about 10^{-2} mbar. Ion optics then direct the ions to the ICR cell at 10^{-10} mbar where the most important gas phase experiments take place. In the ICR cell it is possible to trap and isolate ions of

interest, following which activation techniques can be used to make ions dissociate from which structural, kinetic, and thermochemical information can be derived.

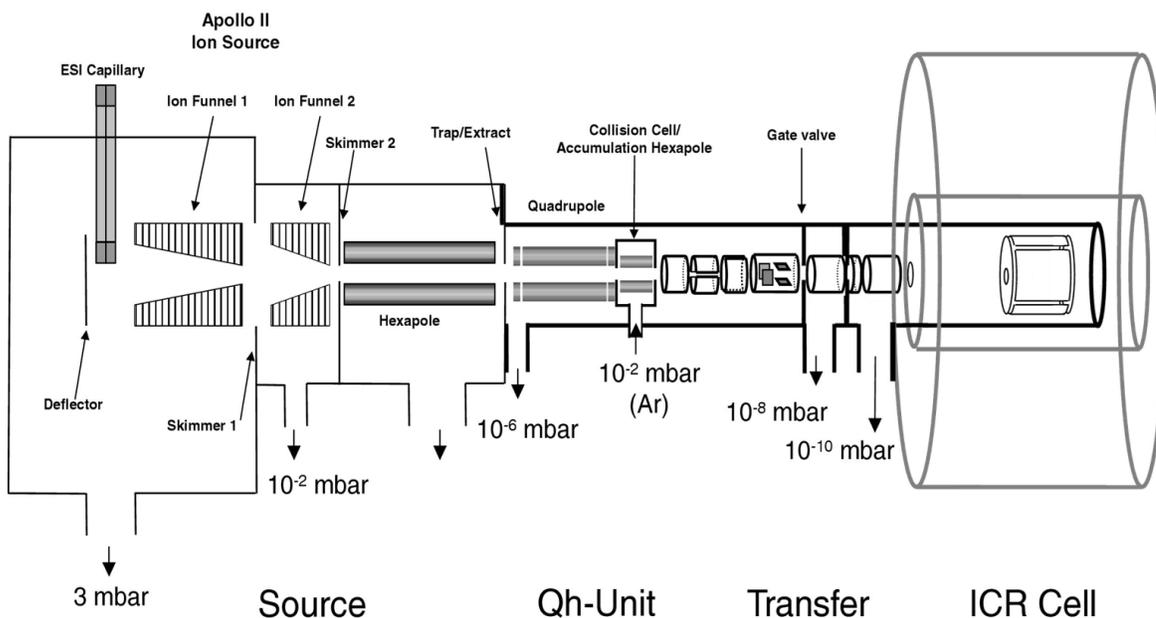


Figure 2.2. Bruker Apex Qe 70 FTICR-MS located at Memorial University. The schematic describes the Apollo II ion source, Qh region, hexapole collision cell, ion transfer optics and ICR cell. Reprinted with permission from *J Am Soc Mass Spectrom* **2009**, *20*, 411-418. Copyright (2009) American Society for Mass Spectrometry.⁴

2.1.1 Electrospray Ionization

The formation of ions is a critical and inseparable part of every mass spectrometer. Ionization techniques have been evolving from the very first ionization techniques such as electron impact, to the many desorption electrospray ionization techniques. The conventional ionization techniques were found to be inefficient at ionizing large nonvolatile biomolecules. Even small volatile compounds often underwent severe fragmentation during the ionization process which limited the possibility of analyzing intact molecular ions. The emergence of electrospray ionization revolutionized

mass spectrometry in the field of biochemistry which deals with biomacromolecules whose three dimensional structures are determined by weak non-covalent interactions.⁵

Figure 2.3 illustrates the two general mechanisms for ion formation during an electrospray process.⁶ The process initiates with dispersion of the solution, composed of analyte and solvent, into small droplets when exiting the thin capillary, which is held at a very high voltage (2-4 kV). A large electric field at the tip of the capillary dominates the surface tension and distorts the shape of the liquid flow into a cone with a rounded tip called a Taylor cone.⁷ At a critical voltage the rounded tip turns into a jet and due to an excessive number of similar charges, Columbic repulsion breaks the jet into smaller droplets. These droplets carry hundreds and thousands of charges depending on the size of the droplets and undergo desolvation influenced by the heated counter flow of drying gas.

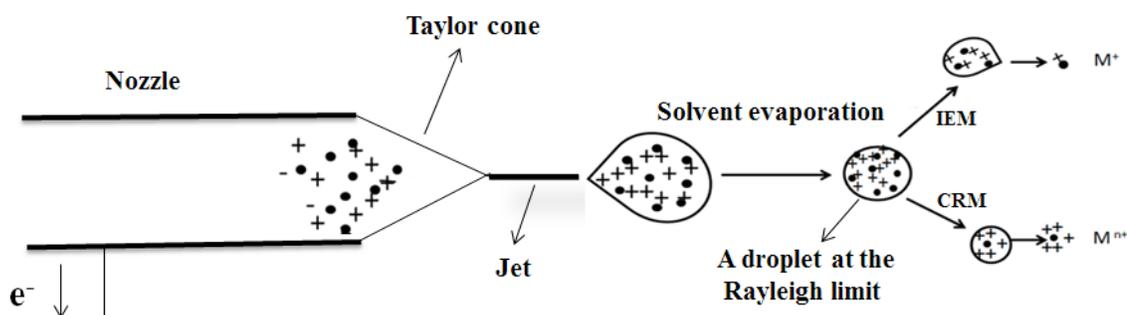


Figure 2.3. A schematic of the electrospray ionization mechanism.

Reduction of the droplet size through solvent evaporation destabilizes it until Columbic repulsion surpasses the forces of attraction that keeps the droplets together. At this point, the Rayleigh limit, the unstable droplets explode, forming smaller charged

droplets. The repetition of this process, desolvation and Coulomb fission, generates very small droplets leading to the final production of gas phase ions, either by the ion evaporation model (IEM) or the charged residue model (CRM) of gas phase ion formation from electrospray.

The IEM refers to the direct emission of analyte ion from the droplet surface when the size of charged droplet is less than 20 nm at the Rayleigh limit.^{8,9} The IEM mechanism seems to mostly occur when dealing with smaller analyte ions such as the small organic metal cation complexes studied here.^{8,10}

On the other hand, large biological molecules are transferred into gas phase through the CRM process that suggests cascades of desolvation and fission until a multi-charged droplet containing one analyte molecule is formed.¹¹ Evaporation of the remaining solvent eventually liberates a multi-charged gas phase ion that can be transferred to the mass spectrometer for further analysis. The dominance of the CRM mechanism in the electrospray ionization of macromolecules with masses above 20 kDa can be inferred and backed up by the observation of multi-charged ions during electrospray ionization.¹¹

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

Since its invention in 1974 by Marshall and Comisarow,¹² FT-ICR has become an indispensable tool for the study of the physical chemistry of ion-molecule complexes including those composed of DNA, RNA, proteins, and polymers. The FT-ICR is

characterized by incomparable resolving power, excellent mass accuracy, and large dynamic range; but above all, its ion trapping ability. The ion trapping capability is augmented by powerful ion activation techniques such as SORI-CID, BIRD, and IRMPD; these benefits have made the FT-ICR well suited for scientific research including the spectroscopy of gas phase ions,^{13,14} ion decomposition or ion-molecule reaction kinetics,¹⁵⁻¹⁷ and complicated isotopic studies.^{18,19}

It was long understood that a magnetic field can induce charged species to revolve in a circular path perpendicular to the magnetic field due to the centripetal (Lorentz) force as illustrated in Figure 2.4. The Lorentz force is given by Equation 2.1.²⁰

$$F = q\mathbf{v} \times \mathbf{B} \quad 2.1$$

in which q , \mathbf{v} , and \mathbf{B} are the ion charge, velocity, and magnetic field, respectively.

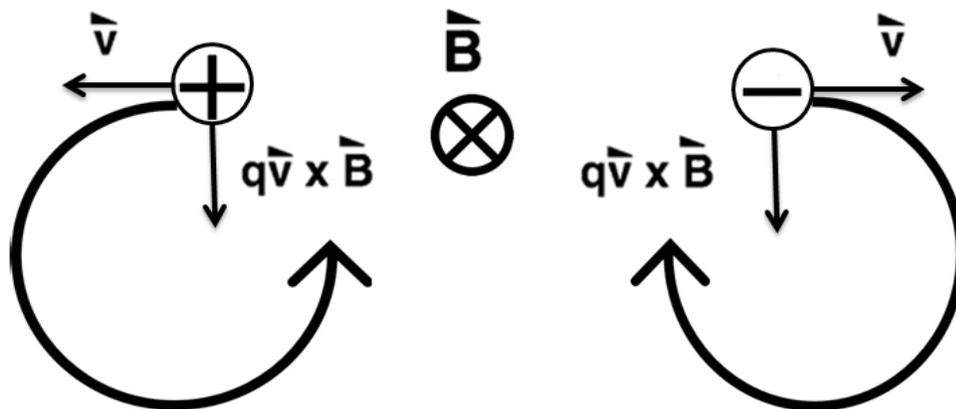


Figure 2.4. Cyclotron motion of ions in the presence of a spatially uniform magnetic field. The Lorentz force acts on ions perpendicular to the direction of the magnetic field. It is shown that the anion's cyclotron motion trajectory is opposite to that of positive ions.

To maintain a stable circular motion, the Lorentz force must be equal to the centrifugal force of the ions $F = \frac{mv^2}{r}$, so equating to 2.1 Equation yields

$$\frac{mv^2}{r} = qvB \quad 2.2$$

Defining the cyclotron frequency as $\omega = v/r$ ^{20,21} and that $q = ze$, Equation 2.2 can be rewritten as

$$\frac{m}{z} = \frac{eB}{\omega} \quad 2.3$$

Equation 2.3 presents two important facts. First, in a fixed magnetic field the cyclotron frequency is dependent only on the m/z of the ions. Since all ions with the same m/z have identical cyclotron frequencies, determination of their cyclotron frequencies can result in the determination of their m/z . Second, the cyclotron frequency is independent of the kinetic energy of the ion.^{12,20}

Figure 2.5 illustrates the configuration of a closed cylindrical ICR, also called a Penning trap.²⁰ The cell consists of three sets of components; the detection (D) and excitation electrodes (E) forming the sides of the cylinder, and orthogonal to the magnetic field are electrostatic trapping plates (T) forming the top and bottom of the cylinder.

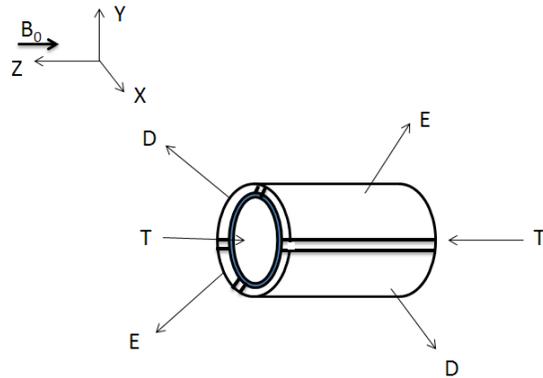


Figure 2.5. A diagram of a Penning trap showing the configuration of plates compared to the magnetic field.

Ions revolve in the xy -plane when entering the ICR cell due to the presence of the magnetic field along the z axis. The process of trapping is not still complete since ions have the option to escape the cell along the z axis. To block the escape of ions along the z -axis and keep them in the center of ICR cell, the trapping plates are held at a potential of between 0.5 to 1.5 V with the same polarity as the trapped ions.

To detect the ions, an RF potential is applied to the excitation plates which causes the ions with a cyclotron frequency equal to the RF frequency to absorb energy. During irradiation ion packets composed of ions with identical masses (and identical cyclotron frequencies) gain kinetic energy and rotate to higher cyclotron radii (see Figure 2.6 with one ion packet for simplicity). At higher orbit radii, the coherent ion packets are now close enough to the detection plates to produce a signal by drawing electrons to a plate as positively charged ions approach it. As the ions pass each detection plate, a corresponding amount of charge is dragged to the plate. The oscillation of charge between the two plates produces a sinusoidal current signal. This induced image current,

also called the time domain ICR signal, includes the information required to produce a mass spectrum. A complex time domain ICR signal, resulting from several ions with different mass to charge ratios, can be converted to a frequency domain signal using a mathematical treatment called “Fourier transform” and ultimately to a mass spectrum (see Figure 2.7).²²

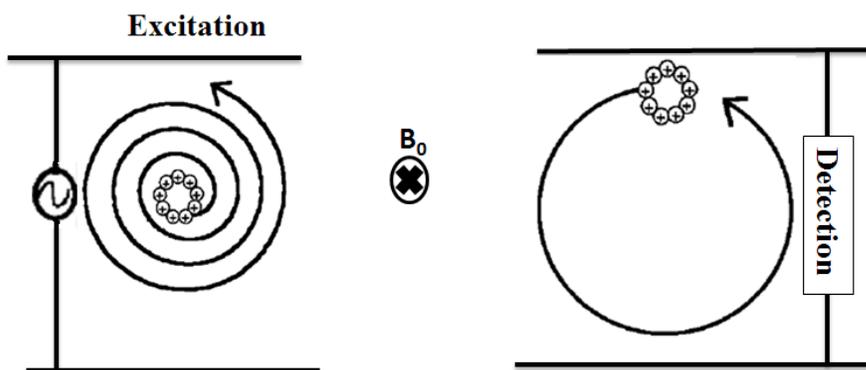


Figure 2.6. A diagram showing how a packet of undetectable ions at the center of an ICR cell can be excited by using an RF potential to a coherent and detectable motion.

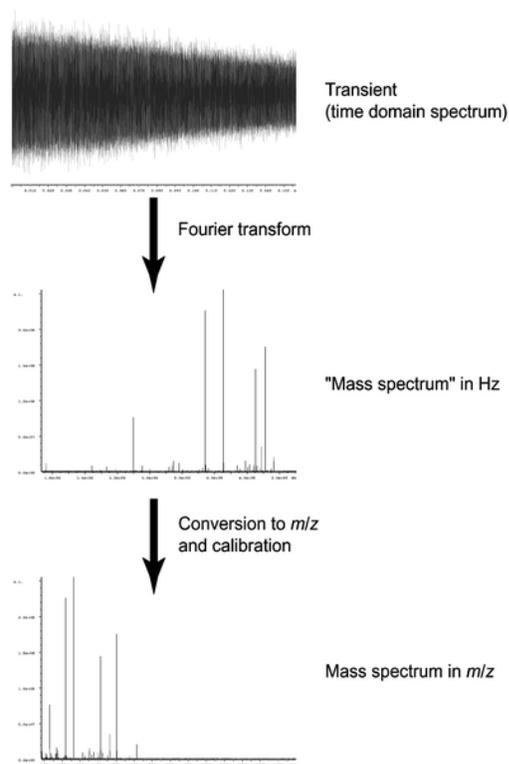


Figure 2.7. A diagram illustrating how a raw time domain transient is Fourier transformed to a frequency domain which can be finally calibrated to a mass spectrum. Reprinted with permission from *Analyst* **2005**, *130*, 18-28. Copyright (2004) Royal Society of Chemistry.²²

It must be noted that by applying a single RF potential, only ions with the same cyclotron frequency move to a detectable radius, so other ions stay unaffected. In order to detect all the masses, the RF frequency must be scanned in a proper range that allows one to excite and detect all ions that may be present in the ICR cell. However, it is also possible to excite all ions except those with a range of masses (frequencies), resulting in ion isolation, which becomes very important when performing tandem mass spectrometry experiments.

2.2. Tandem Mass Spectrometry Techniques

Tandem mass spectrometry (MS/MS) is among the most frequently used and well suited techniques for the study of the physical chemistry of gaseous ions. Structural and thermochemical studies conducted by MS/MS generally rely on detection of fragmentation of mass-selected ions in the gas phase. These types of experiments all involve an excitation step allowing the ion of interest to gain enough internal energy so that the deposited energy surpasses at least the lowest energy dissociation pathway. To date, various configurations have been applied to conduct MS/MS experiments offering different levels of performance and advantages.

Ion trapping MS/MS configurations, including ion cyclotron resonance mass spectrometer (ICR-MS), are among the most versatile instruments to perform these sorts of studies. In the present work, infrared multiple photon dissociation (IRMPD) spectroscopy, collision induced dissociation (CID-specifically sustained off-resonance irradiation, SORI-CID), and blackbody infrared radiative dissociation (BIRD) have been used in order to obtain physical chemistry information of ionized guanine adducts. These ion activation techniques will be explained in the following three sections.

2.2.1. Collision Induced Dissociation

The energetic collisions of biologically relevant ions with a neutral inert gas is the most commonly used activation method, namely collision-induced dissociation (CID). In the CID process, ions can undergo multiple collisions with a target gas, most commonly N_2 or Ar, and the translational energy is converted to and stored as internal energy of the

ion, followed by fragmentation. The amount of converted energy depends on the experimental parameters applied in the CID process including the kinetic energy of the ion of interest, the pressure of the CID cell, the nature of the target gas, the mass of ions, and the time scale of the CID process.^{23,24}

The term *collision endothermicity* (q) is used to define the amount of kinetic energy converted into internal energy upon collision. Theoretically, ions can undergo a totally elastic collision, $q = 0$, when no conversion of energy to internal energy takes place. In an inelastic collision, colliding pairs can stick together and in this case the collision endothermicity gets its maximum value, q_{max} , as the whole of the translational energy of the system is converted into internal energy, leaving the colliding partners motionless. The mechanism of CID lies somewhere between the two aforementioned possibilities. It means the relative translational energy of colliding species can change to some extent, $q > 0$, that is highly dependent on the relative kinetic energy of particles and their respective masses, and according to the law of conservation energy, this amount of energy is converted into internal energy. Briefly, in CID, the ion and target which is an inert gas such as argon, collide and stick together for a period of time in which their translational energies are transferred to internal energy, but after a period of time the ion and neutral separate, taking with them some of their original translational energy.

Equation 2.4^{24,25} formulates the maximum available translational energy of ions, the center of mass collision energy (E_{com}), that can appear as internal energy while the target gas particles are static. N and m_p represent the masses of the target gas and the parent ions, respectively, and E_{lab} is the lab frame kinetic energy of the ions.

$$E_{com} = \left(\frac{N}{m_p + N}\right)E_{lab} \quad 2.4$$

It is evident that the E_{com} is dependent on the masses of both colliding particles; the larger the neutral target, the higher the value of E_{com} . Further, the larger the ions the less lab frame energy is converted to internal energy.

Numerous tandem mass spectrometry instruments have been used to conduct CID experiments for ions either trapped or in continuous linear motion in beam-type instruments such as sectors and time of flight mass spectrometers. Based on the instrument used, CID experiments can be classified into two categories including high-energy activation (fast activation) and low-energy activation (slow activation).^{24,25}

High-energy CID experiments are mostly performed by means of beam-type instruments like time of flight (TOF) or sector-based tandem mass spectrometers in which the kinetic energy of the precursor ions is on the order of few hundreds or thousands of eV ($\sim 100 \text{ kJ mol}^{-1}$ per eV). As a result, one collision is enough to dissociate the ion and the time scale for dissociation is on the order of a few microseconds. High-energy CID experiments are well-suited for those types of studies aiming to reach high-energy dissociation channels or extensive fragmentation. Low-energy CID refers to experiments where the kinetic energy of the precursor ion lies typically between 1-100 eV range giving rise to a slow activation process in which ions are exposed to tens to thousands of collisions with the inert gas molecules over the course of a few hundred ms. Slow activation CID experiments are typically carried out in ion trapping instruments or with linear quadrupole instruments. In low-energy CID, only the lowest energy dissociation pathways are accessed.

2.2.2. Sustained Off-Resonance Irradiation-Collision Induced Dissociation (SORI-CID)

SORI-CID is the FT-ICR version of CID which is considered to be a very slow activation technique (very low-energy CID) with collision energies in the low 10's to about a 1000 kJ mol⁻¹, (0.1 – 10 eV) useful to induce dissociation of non-covalent interactions as well as covalent interactions after many hundreds of collisions. It was discussed in section 2.1.2 that ions trapped in the ICR cell can become excited by using an RF potential whose frequency matches the natural cyclotron frequency of the ions. This excitation, on-resonance excitation, cause ions to gain kinetic energy quickly such that they are ejected from the ICR cell. To overcome this issue, scientists developed off-resonance activation which applies an RF frequency slightly higher or lower than the cyclotron frequency of the ions of interest (typically by about 500 Hz). This causes the ions' kinetic energy to fluctuate sinusoidally over the radiation time.^{24,26,27} The acceleration and deceleration cycle prolongs the activation process to 100's of ms with upwards of thousands of collisions. SORI-CID activates the very lowest energy dissociation pathways and can make macromolecules undergo isomerizations and dissociations through a slow heating process.

The maximum translational energy that ions can obtain during a SORI-CID process can be measured by following equation (Equation 2.5) where β is a geometrical factor (dependent on the shape of ICR cell, here 0.92), d is the diameter of the ICR cell (6×10^{-2} m), V_{p-p} is the peak to peak voltage, q is the ion charge m is the ion mass and $\Delta\nu$ is the frequency difference between natural cyclotron frequency and RF frequency (typically 500 Hz).

$$E_{lab} = \frac{\beta^2 q^2 V_{p-p}}{\pi^2 m d^2 \Delta v^2} \quad 2.5$$

Experimentally, the ion's kinetic energy can be changed by varying the V_{p-p} value while the other factors remain unchanged.

2.2.3 Blackbody Infrared Radiative Dissociation

It was discussed in section 2.1.2 that FT-ICR can enable one to effectively trap and isolate ions for an extended periods; hours, even days. The trapping capability becomes even more beneficial when taking into account that ion isolation occurs in an extremely low pressure environment, 10^{-10} mbar, which guarantees very few collisions (~1 collision every 10 s). However, weakly-bound ions can be observed to dissociate under this low-pressure condition by absorbing infrared photons emitted from ambient blackbody radiation field. This ion activation process in “zero”-pressure environments refers to blackbody infrared radiative activation from which unimolecular dissociation occurs and can lead to important mechanistic and thermochemical data.²⁸⁻³⁰

BIRD, as a very slow heating process, is well suited to investigate binding energies of non-covalent adducts with binding energies less than about 150 kJ mol^{-1} . To perform a successful BIRD experiment leading to binding energy insights, the ions are only activated by accumulation of energy through energy exchange with the ambient blackbody radiation field. Binding energy data in BIRD are determined from the temperature dependent dissociation rate constants. The BIRD mechanism can be described as:



in which k_{abs} and k_{em} are photon absorption and photon emission rate constants, respectively, and k_{dis} is the dissociation rate constant.

By setting the temperature of ICR cell to a specific temperature (T), the ion under study undergoes unimolecular dissociation and the rate of this dissociation, k_{uni} , can be determined. This can be performed by a record of the ion's normalized intensity, I , as a function of time, t , and a fit of these measurable factors to the integrated first-order rate law, Equation 2.7,

$$[I] = [I]_0 e^{-k_{uni}t} \quad 2.7$$

A repeat of BIRD experiments at multiple temperatures yields temperature dependent rate constants, k_{uni} , from which an Arrhenius plot, Equation 2.8, is prepared.

$$k_{uni} = A e^{-E_a/k_B T} \quad 2.8$$

An Arrhenius plot enables one to extract ion activation energies, E_a , and pre exponential factors, A . Table 2.1 summarizes some thermochemical quantities which can be determined by BIRD technique.

Table 2.1 Thermochemical parameters accessible by BIRD experiments

Quantity	Description
E_a^{obs} Activation energy of dissociation (observed)	$E_a^{obs} = \frac{\partial \ln k_{uni}}{\partial (1/k_B T)}$
E_a^∞ High pressure/REX limiting activation energy of dissociation	When systems are within REX, E_a^{obs} and E_a^∞ are equal
A^{obs} Pre-exponential factor (observed)	$k_{uni} = A^{obs} e^{-E_a^{obs}/k_B T}$
A^∞ High pressure/REX limiting pre-exponential factor	$k_{uni}^\infty = A^\infty e^{-E_a^\infty/k_B T}$
ΔS^\ddagger Entropy of activation	$A^\infty = \frac{e k_B T}{h} e^{-\Delta S^\ddagger/k_B}$

The validity of the observed Arrhenius values (E_a^{obs} and A^{obs}) is dependent on the relative rates of energy exchange and the size of the molecule. In this regard, molecules are classified into two regimes, based on size, including small molecule and large molecule regimes. Considering the BIRD mechanism (Equation 2-6), the collisionless unimolecular rate constant can be given by Equation 2.9 in which the steady state approximation is incorporated.³¹

$$k_{uni} = k_d \left(\frac{k_{1,rad}}{k_{-1,rad} + k_d} \right) \quad 2.9$$

where $k_{1,rad}$ and $k_{-1,rad}$ are the temperature dependent blackbody photon absorption and emission rate constants, respectively, and k_d is the dissociation rate constant.

When dealing with the small molecule regime, thermal equilibrium of dissociating ions with the surroundings cannot be established since the dissociation rate constant greatly surpasses the emission rate constant ($k_d \gg k_{-1,rad}$). Briefly, ions with activation energy above E_0 are subject to removal due to quick dissociation which distorts the Boltzmann distribution of internal energy of the ion population. As a consequence, the unimolecular rate constant is merely governed by the energy exchange rate instead of the reaction rate. The Arrhenius parameters resulting from this regime are underestimated compared to the parameters obtained from rapid exchange limit (i.e. Boltzmann distribution of ions).²⁹ The observed parameters for small molecules must be corrected by using of master equation modeling^{31,32} which will be explained in section 2.3.5.

In the large molecule regime, the observed Arrhenius values (E_a^{obs} and A^{obs}) can be considered equal to the limiting so-called infinite-pressure values (E_a^∞ and A^∞). The rate of energy exchange between large ions with many degrees of freedom and the blackbody radiation field (the body of ICR cell) is higher than the dissociation rate ($k_d \ll k_{-1,rad}$). At this rapid exchange limit the ion population is thermally equilibrated with its surroundings, resulting in a Boltzmann distribution of internal energies of the ion population. At thermal equilibrium, one can readily extract threshold dissociation energies, E_0 , directly from the Arrhenius plot.³²

2.2.4. Infrared Multiple Photon Dissociation Spectroscopy

Exposure of molecules to the infrared region of light has been long used to identify molecules through detection of functional groups and structural elucidation. Due

to insensitivity at low ion density, conventional infrared absorption spectroscopy is limited to the condensed phases and concentrated gas phase. In order to record infrared spectra of gaseous ions, the consequence of resonant infrared absorption is observed. There are a number of different consequence spectroscopies that provide vibrational (structural) information. For example, under certain conditions it is possible to tag ions with an inert gas atom and the consequence of resonant absorption of an infrared photon is loss of the tag, which is detected by mass spectrometry.³³⁻³⁵ In the present case, IRMPD spectroscopy is used, by which the ions absorb multiple photons, increasing the ion's internal energy beyond the threshold for dissociation and dissociation of the ion is detected, again, by mass spectrometry.

It was demonstrated in 1978 that a low power infrared laser can dissociate trapped ions,³⁶ and that was the starting point of IRMPD spectroscopy's evolution as a consequence infrared spectroscopy. Since then, the applications of IRMPD spectroscopy in the gas phase to structural and thermochemical studies has undergone tremendous changes and growth, including the studying relatively large biomolecule ions ionized by ESI.

In IRMPD spectroscopy, ions are mass-selected and trapped for a short period of time (on the order of seconds) during which they are irradiated by an intense tunable infrared laser. Early IRMPD studies utilized tunable CO₂ lasers,³⁶⁻³⁸ and although these lasers were intense, the range of infrared region provided was limited and discontinuous, constraining their spectroscopic usefulness. However, the entire chemically valuable region of the infrared spectrum became accessible to ions in mid 2000s when widely

tunable lasers such as free electron lasers (FELs)³⁹ were mated with FTICR instruments. Soon after, optical parametric oscillators/amplifiers (OPO/A) (another tunable infrared laser) began to be used in university laboratories.⁴⁰ The operation of these lasers is different from conventional lasers and will be discussed in the following sections, 2.2.4.2 and 2.2.4.3.

2.2.4.1 IRMPD Mechanism

IRMPD is considered a consequence spectroscopy as the absorption or emission is not detected to obtain the infrared spectrum, but the consequence of multiphoton absorption—dissociation—is detected. For an ion/molecule complex to undergo chemical bond breakage or isomerization, the required energy is in the range of 100-400 kJ mol⁻¹, but it can be less for very weakly-bound complexes. To deposit this amount of energy into the ion of interest, several tens of infrared photons are required. The first requirement for the IRMPD process is the existence of a vibrational normal mode that is resonant with the laser frequency so that a photon is absorbed. If the vibrational mode was considered a harmonic oscillator (Figure 2.8 a),⁴¹ the vibrational energy levels would have identical energy differences. Therefore, the threshold for dissociation would be easily reached by absorbing sequential photons. However, real molecules are anharmonic making the aforementioned mechanism impossible since anharmonicity distorts the vibrational energy spacings to convergence as seen in Figure 2.8 b. As a result, soon after absorbing only one or a few photons, the laser light will run out of resonance making the absorption process stop.

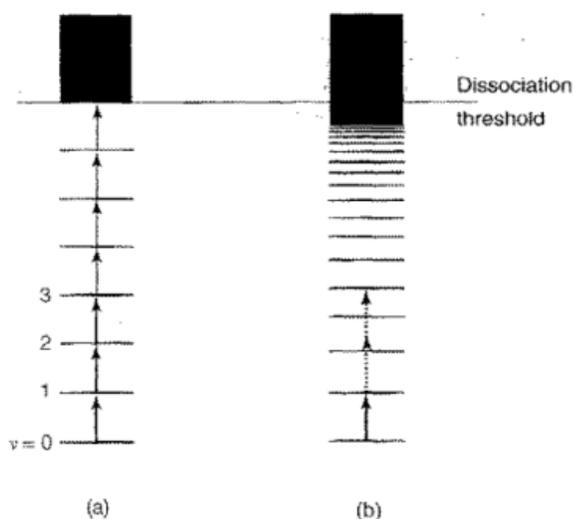


Figure 2.8. Representation of vibrational energy levels of (a) an imaginary dissociation channel without anharmonic effect and (b) a species with anharmonic effect, showing convergence in vibrational spacings. Reprinted with permission from Fridgen, T. D.; McMahon, T. B. "IRMPD" in *Encyclopedia of Mass Spectrometry*, **2005**, Vol. 4. Copyright (2005) Elsevier⁴¹

The mechanism of IRMPD is depicted in Figure 2.9.⁴² After the absorption of the first photon ($v = 0 \rightarrow 1$), and also due to anharmonicity and coupling of vibrational modes, the energy deposited can be redistributed over the other available vibrational degrees of freedom making available the fundamental mode for subsequent photon absorption. The depopulation is attributed to fast ($\approx 10^{-13}$ s) intramolecular vibrational energy redistribution (IVR) in which the energy of a localized vibrational excitation randomizes rapidly throughout other vibrational modes when the ion is in a collisionless environment. Subsequent photon absorption at the same fundamental transition followed by IVR is repeated until one of the dissociation energy thresholds is surpassed and dissociation is observed.

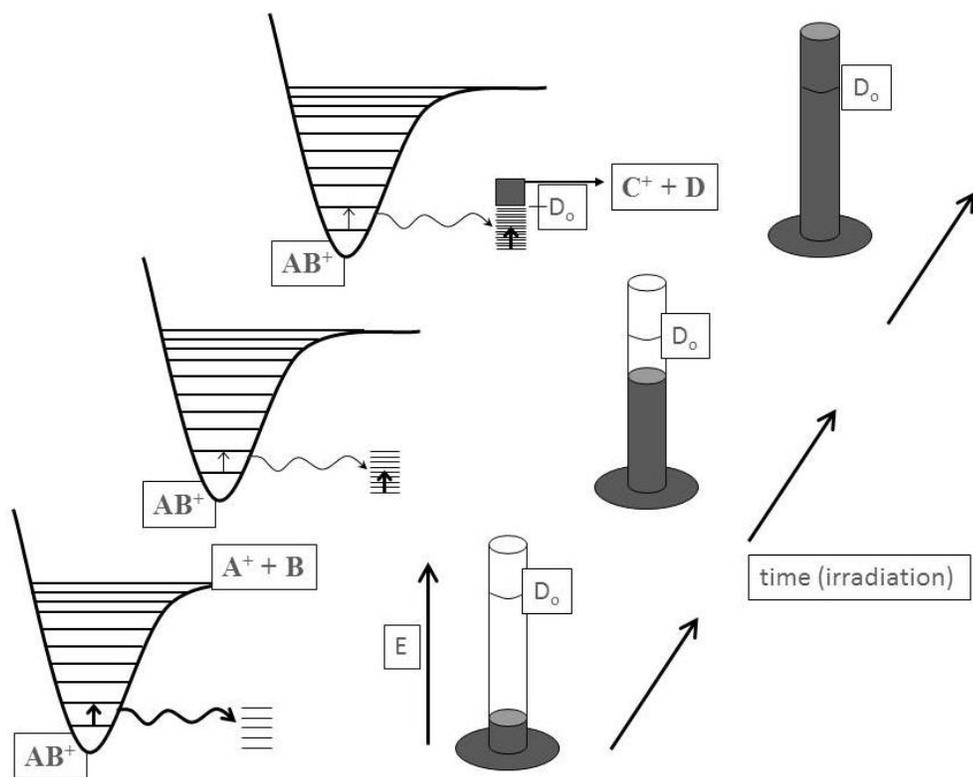


Figure 2.9. A general schematic of an IRMPD process representing photon absorption in a specific vibrational mode followed by an IVR process. Right after every photon absorption, fast intramolecular vibrational relaxation stores energy gained over other vibrational modes, which is much easier in large polyatomic molecules. This cycle keeps repeating until the internal energy deposited surmounts the lowest dissociation threshold making ions undergo unimolecular dissociation. The figure is drawn by Dr. Fridgen for theses and presentations.

IRMPD spectroscopy, then, can be conducted by irradiating trapped ions with a tunable infrared laser over a range of wavelengths (Figure 2.10). The IRMPD spectrum is produced by plotting the IRMPD yield (Equation 2.10) as a function of wavelength

$$IRMPD\ Yield = -\log_{10}\left(\frac{I_{parent}}{I_{parent} + \sum_i I_{fragment(i)}}\right) \quad 2.10$$

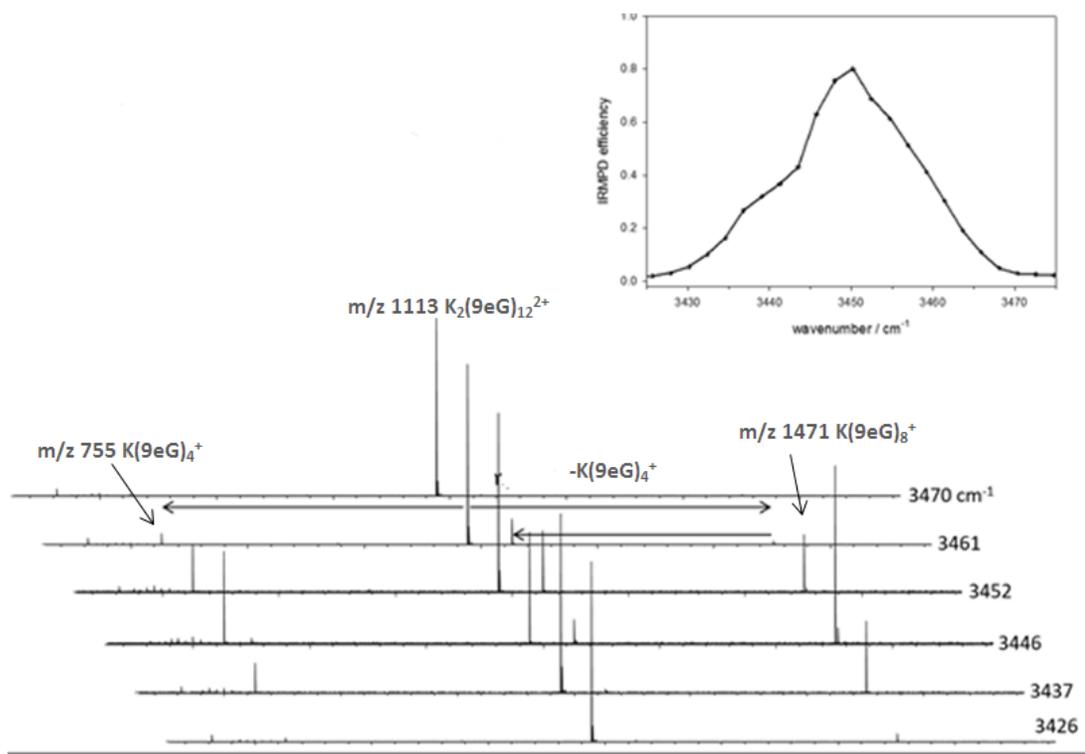


Figure 2.10. A representation of acquiring an IRMPD band. A scan of trapped ions $K_2(9eG)_{12}^{2+}$, 9eG denotes 9-ethylguanine, with the laser in the range of 3470 to 3426 cm^{-1} results in ion fragments that can be subsequently used to build the corresponding IRMPD bands in that wavelength range. Note that a full IRMPD spectrum can be obtained by scanning the desirable range of IR area.

2.2.4.2 Optical Parametric Oscillators (OPO)

Lasers emerged as unique source of radiation well-suited for spectroscopy applications, but being single-color (very narrow wavelength) required scientists to build technologies with ability of tuning the laser output to a wider range of wavelengths. Optical parametric oscillators were invented five years after the emergence of the first laser, and were demonstrated to successfully scan an input source of light.^{43,44} Every OPO consists of a non-linear crystal such as potassium titanyl phosphate (KTP) or lithium niobate ($LiNbO_3$) which can split the radiation from an input or pump laser, into

two beams with lower frequencies shown in Equation 2.11 in which ω_p is the pump laser frequency,

$$\omega_s + \omega_i = \omega_p \tag{2.11}$$

ω_s and ω_i are the output frequencies which are called the signal and idler beam, respectively. The output with higher frequency is conventionally named signal and the one with lower frequency is called the idler output. The frequency of the output depends on the angle of the non-linear crystal with respect to the pump input, and by altering the angle a wide range of mid-infrared radiation can be obtained.

IRMPD spectra in the higher energy region, 2700-3800 cm^{-1} , were acquired using an infrared (IR) optical parametric oscillator (OPO) located at Memorial University, manufactured by LaserSpec, the idler output which is tunable from 2700 to 4000 cm^{-1} , with a bandwidth of 2 cm^{-1} . The OPO is built around a periodically poled lithium niobate crystal and is pumped using a diode-pumped, solid state, Nd:YAG laser. The OPO operates at 20 kHz, with a pulse length of a few nanoseconds and can generate an output power of about 3 W at 3 μm . The output power is limited to 1 W and was attenuated using infrared filters for the present work.

2.2.4.3 Free Electron Laser (FEL)

Unlike conventional lasers that use excited atoms or molecules as the active lasing medium, the free electron laser employs very fast electrons to produce a laser beam. FELs operate based on the concept that charged particles that are accelerated emit electromagnetic radiation—synchrotron radiation.⁴⁵ Figure 2.11 depicts a free electron laser. A gigaampere current of electrons produced by an electron gun is accelerated to

MeV of kinetic energy. This electron beam is directed into a set of alternating, undulator, magnets, which causes the electrons to wiggle as the electron beam travels through the axis of the undulator. The wiggling electrons lose some of their energy in the form of radiated monochromatic photons. Mirrors at either end amplify the laser beam. The free electron laser is currently the only laser which is tunable from microwave to x-ray frequencies which makes it very practical for spectroscopy purposes, but the cost and foot print are the current obstacle to equip labs with this unique instrument. The laser frequency can be tuned by changing the electron velocity, the magnetic strength, or the undulator magnet position.^{46,47}

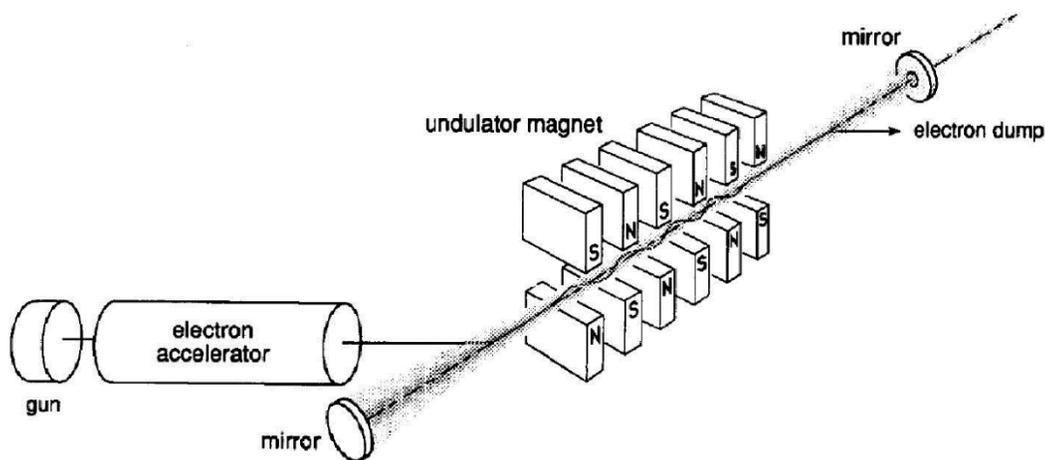


Figure 2.11. A diagram of free electron laser components. The undulator magnet plays a crucial role in which very-high-speed electrons become accelerated radially emitting electromagnetic radiation. Reprinted with permission from *Infrared Phys. Tech.* **1995**, 36, 297-308. Copyright (1995) Elsevier.⁴⁶

2.3. Computational Chemistry

Emergence of powerful computers with the ability to perform millions of calculations and data processing in a reasonable period of time has enabled computer

programmers to model and simulate the chemical properties of molecules without conducting any experiments.⁴⁸ Computational techniques are capable of predicting and computing molecular structures, reaction mechanisms, energetics, spectroscopic data, etc.⁴⁸⁻⁵¹ Although computational chemistry is a trusted method to examine the chemical properties of molecules, these techniques have been implemented as complementary methods to chemical experiments. For instance, IRMPD spectroscopy can probe the existence of a variety of functional groups, but a comparison of computed IR spectra to the experimental IRMPD spectra can help lead to a full structural elucidation.^{42,52-54} *Ab initio* methods and density functional theory (DFT) were employed in this work to compute electronic structures, binding energies, stabilities and energetics, and spectroscopic values such as IR band positions and intensities which will be discussed in greater detail in the following sections. Besides differences in the approaches used, all computational techniques try to determine the lowest energy structure through probing the relative potential energy surface of a sets of possible isomers.^{55,56} Computational techniques differ in terms of accuracy, cost, and time consumption, and they can be combined in order to obtain satisfactory results. For instance, geometry optimization, which has been proved to be less dependent on the selected method and basis set, can be performed by density functional theory methods which are both time and cost efficient. The DFT-optimized structures then can be subjected to more accurate techniques such as *ab initio* to determine single point energies.

2.3.1. *ab initio* Methods

Ab initio refers to methods entirely based on first principles quantum mechanics in which the simplification of the Schrödinger equation for multi-electron systems is acquired by some approximations.⁵⁷ The Born-Oppenheimer approximation⁵⁸ (BO approximation) is the underlying and indispensable approximation that makes the electronic wavefunction tractable and is incorporated in all *ab initio* methods. The BO approximation considers the nuclear coordinates fixed, leading to decoupling and separation of the motion of nuclei and electrons. This approximation allows the separation of the wavefunction into an electronic and nuclear component in which the resultant equation (Equation 2.12) can be solved in two simpler steps. It can be inferred as a reasonable approximation considering the size difference of fast moving electrons and heavy nuclei.

$$\Psi_{total} = \psi_{electronic} + \psi_{nuclear} \tag{2.12}$$

It must be noted that this approximation still includes the Columbic attractions between the nuclei and electrons but with a constant distance over time.

Hartree–Fock⁵⁹ (HF) is the simplest *ab initio* technique that has been long used to determine wavefunctions and energies of atoms, molecules, and nanostructures. Beside the BO approximation, HF makes other simplifications including a non-relativistic momentum operator, and a mean field approximation. The mean field approximation defines the electrons' motions entirely independent of each other while still each electron is influenced by the average electric field of the other electrons, instead of instantaneous electron-electron repulsion, arising from all electrons. Although the electron repulsion of

same-spin electrons is accounted for, this assumption totally neglects the interaction between electrons with opposite spins when being in close proximity, leading to an incorrect estimation of correlation energy. In other words, the Hartree-Fock technique overestimates the electron repulsion resulting in energies higher than the exact energies that can be obtained by solving Schrödinger equation. The HF technique is a simple method that has been shown to be effective for geometry optimization, especially for preparing partly optimized structures for higher level theories. Post-Hartree-Fock⁶⁰⁻⁶² *ab initio* methods have been developed to improve the results accuracy by inclusion of electron correlation. Møller-Plesset⁶³ (MP) theory is one of several post-Hartree-Fock computational methods that relies on Rayleigh–Schrödinger perturbation theory (RS-PT) in order to have a better estimation of electron correlation energies. MP_n, where n denotes the order of perturbation applied, has shown great potential for very accurate geometry optimization, energy calculations, and spectroscopy data at the price of more cost and lengthy computations. As a result, MP methods can be employed to improve the calculated results from lower level methods.

2.3.2 Density Functional Theory

Density functional theory refers to a computational technique which uses the electron density as the underlying factor, unlike the HF techniques that solve a complicated wavefunction.⁶⁴⁻⁶⁶ The complexity of HF techniques stems from the high number of variables included in electronic wavefunction (the coordinates of all electrons) which increases calculation time, whereas in DFT, the number of variables is reduced to x, y and z coordinates of the electron density.⁶⁵ This reduction in the number of variables

dramatically shortens the time required for electronic calculations. The Hohenberg-Kohn theorem⁶⁷ is the centerpiece of the DFT techniques that asserts that all ground state properties of molecules and atoms can be extracted by using functionals, functions of another function. In this theorem, a functional is defined as the spatially dependent electron density, ρ , which is a function of position $\rho(r)$. As a result, since the electronic energy of a system is a functional of the electron density, knowing the electron density functional can lead to the desired energy calculation. The energy functional for a system with interacting electrons can be solved by the Kohn-Sham theorem⁶⁸. The Kohn-Sham theorem simplifies the system into a tractable problem with non-interacting electrons moving in a specific potential. This potential is a summation of the external potential and the Coulombic interactions between the electrons, the correlation and exchange interactions.

There have been several methods developed based on the DFT concept.^{69,70} Semi-empirical hybrid DFT techniques, such as B3LYP,^{71,72} are the most common DFT methods that use HF exchange energy and empirical sources in order to improve the results in terms of accuracy and allocated time. Although there are more accurate functionals developed, B3LYP is the most common functional owing to its reliable and rapid geometry optimizations, and infrared spectrum determinations for ion clusters. The main drawback of B3LYP appears when it comes to thermodynamic and binding energy determinations which results in highly deviated results. To overcome this problem, the ion molecules are first optimized by B3LYP functional and then the energetics are determined by treating the optimized structures with higher level techniques such as MP₂.

2.3.3. Basis Sets

Basis sets in quantum chemical calculations are a set of mathematical functions, basis functions, which are combined linearly in order to form molecular orbitals.^{50,73,74} Basis sets differ in terms of the nature and number of functions which lead to different calculation time, cost, and accuracy. In this thesis, 6-31+G(d,p) and 6-311+G(3df,3pd) were used as basis sets during B3LYP calculations. These basis sets are classified as Pople basis sets, split-valence basis sets, in which each valence orbital is defined by more than one basis function.⁷⁵⁻⁷⁸ Pople basis sets have a typical format of C-VvG; each letter contains function(s) concerning the core or valence electrons. In this formula, C and Vv are indicative of the Gaussian functions that deal with the core and valence electrons, respectively. While the core electrons are described by a single basis function comprised of C Gaussian functions, the valence electrons are treated with two basis functions of V and v Gaussians. The valence electrons, in this case composed of two basis function refers to the term split-valence double-zeta. Beside the functions describing the core and valence electrons, other functions can also be incorporated into the basis sets to include polarization and diffusion effects. In the basis set, the diffusion function appears as + or ++ which is necessary for systems such as anions and excited states with more spread electron interactions. The polarization functions are shown in basis sets as a parentheses with the orbitals name, (d,p) or (2d,2f), giving more flexibility to the orbital shapes. For instance, in the (d,p) functions, d refers to the addition of d-type functions to the p-orbitals while p corresponds to the addition of p-type functions to the light atoms such as hydrogen and helium.

Effective core potential⁷⁹⁻⁸¹ (ECP) basis sets were implicated in this work for electronic calculations of heavy metals, and also larger molecules such as G-quadruplexes. Evidently, an increase in the number of electrons increases calculation time and cost dramatically. ECP methods remove the calculation part that concerns the core electrons and instead, it considers a collective effect that core electrons can have on valence electrons. This approximation speeds up calculations by eliminating many variables from the problem.

2.3.4. Calculation Approach

Throughout this work, Gaussian 09⁸² was used to perform all quantum chemical calculations and the procedure was as follows. Predicted isomers of a molecule were subject to the B3LYP density functional geometry optimization resulting in optimized structures with corresponding infrared spectra. B3LYP frequency calculations also provide some thermodynamic data which are used as complementary data to higher level B3LYP calculations, i.e. B3LYP calculations with larger basis sets, in order to determine more accurate thermodynamics. The internal energy of a system can be defined by Equation 2.13 which can be then used to determine enthalpy, H , and Gibbs energy, G , by means of Equations 2.14 and 2.15, respectively.

$$E = E_{elec} + E_{vib} + E_{trans} + E_{rot} + E_{zp} \quad 2.13$$

$$H = E + k_B T \quad 2.14$$

$$G = H - TS \quad 2.15$$

Internal energy in Equation 2.13 is a summation of electronic, vibrational, translational, rotational, and zero point energy. k_B and S in Equations 2.14 and 2.15 represent

Boltzmann constant and total entropy, respectively. In this work, for G-tetrads, B3LYP density functional theory with the 6-31+G(d,p) basis set on C, H, N, and O and the Def2SVPD basis on metals were used for optimizations and calculation of the infrared spectra. In order to improve the accuracy of thermochemistries, electronic energies were refined using the 6-311+G(3df,3pd) basis on C, H, N, and O and the Def2TZVPP basis on metals. Relative thermochemistries at 298 K (enthalpies and Gibbs energies) were then calculated from B3LYP/6-311++G(3df,3pd) electron energies and thermal corrections from the B3LYP/6-31+G(d,p) (Equations 2.16 and 2.17).

$$\Delta H_{LB} = \Delta E_{LB} + H_{corr(SB)} \quad 2.16$$

$$\Delta G_{LB} = \Delta E_{LB} + G_{corr(SB)} \quad 2.17$$

in which LB and SB stands for larger, 6-311++G(3df,3pd), and smaller, 6-31+G(d,p), basis sets, respectively.

This two-step process allows one to improve the accuracy meanwhile saving a great deal of calculation time and cost by avoiding optimization and frequency calculations using a larger basis set. Calculation of relative enthalpy ($\Delta_{rel}H$) and Gibbs energy ($\Delta_{rel}G$) of the optimized isomers at 298 K allows the relative stability of a series of isomers to be determined and most importantly the lowest energy isomer.

2.3.5. Master Equation Modeling

When interpreting BIRD thermodynamic results, an attention to the size regime is essential, which was discussed in detail in section 2.2.3.^{29,31,83} The observed activation energies can greatly deviate from E_0 , threshold dissociation energy, if the systems are not

large enough to exchange energy at a rate equal or greater than the rate they dissociate. For slow dissociating systems ($k_d \ll k_{-1,rad}$) like the K, Rb and Cs G-quadruplexes in Chapter 4, the ion cluster is dissociating at the rapid exchange rate and the observed Arrhenius factors can be interpreted as the reliable values toward determining thermodynamic information. However, the unimolecular dissociation kinetics of molecules in the non-REX limit ($k_d \ll k_{-1,rad}$) must be determined by assistance of master equation modeling.

In this work, a matrix form of the master equation (Equation 2.18) was employed to obtain E_0 for potassiated and sodiated G-quadruplexes from the BIRD Arrhenius plots (Chapter 4); this process has been explained elsewhere^{31,84,85} and brief details are provided here.

$$\frac{dN_i(t)}{dt} = \sum_{j \neq i} k_{i,j} N_j(t) - k_d N_i(t) \quad 2.18$$

In this equation, $N_i(t)$ and $N_j(t)$ account for the population in the corresponding energy levels of i and j , respectively, and $k_{i,j}$ and k_d , refer to the internal state to the state transitions and dissociation processes, respectively. The master equation has two components: the transport (or J) matrix includes the rate constants for emission, absorption, and dissociation; and a second matrix that contains the initial population distribution, $[N_i(0)]$. Solution of this matrix equation results in the probability of energy transfer from one energy state to another possible state. To model the BIRD temperature dependant rate constant, the Boltzmann population at a desired temperature was allowed to redistribute until the internal energy distribution reaches a steady-state in which linear

BIRD kinetics are achieved. The modelled BIRD rate constant is obtained from the slope of the $\ln([M]_t)$ vs. t plot, where $[M]_t$ is the fractional population left at reaction time t .

Finally, an Arrhenius plot was graphed by only using the modeled rate constants for the lowest and highest temperatures. In the modeling process, a range of pre-exponential factors (tightness or looseness of the transition state), A , were examined to obtain reliable fitting. The IR intensities were scaled to achieve the desired pre-exponential factors. Calculations were followed by increasing E_0 in the range of 0.95 – 1.75 eV in 0.05 eV increments. For each set of E_0 and $\ln(A)$ values, an Arrhenius plot was constructed which was compared to the experimental plot, and the best fit to the plot yielded the threshold activation energy for the G-quadruplex.

References

- (1) Gillis, E. A. L.; Demireva, M.; Nanda, K.; Beran, G.; Williams, E. R.; Fridgen, T. D.: Structures and energetics of electrosprayed uracil-Ca²⁺ clusters (n = 14-4) in the gas phase. *Phys. Chem. Chem. Phys.* **2012**, *14*, 3304-3315.
- (2) Fernandez, F. M.; Wysocki, V. H.; Futrell, J. H.; Laskin, J.: Protein identification via surface-induced dissociation in an FT-ICR mass spectrometer and a patchwork sequencing approach. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 700-709.
- (3) Wilhelm, M.; Schlegl, J.; Hahne, H.; Gholami, A. M.; Lieberenz, M.; Savitski, M. M.; Ziegler, E.; Butzmann, L.; Gessulat, S.; Marx, H.: Mass-spectrometry-based draft of the human proteome. *Nature* **2014**, *509*, 582-587.
- (4) Rajabi, K.; Easterling, M. L.; Fridgen, T. D.: Solvation of electrosprayed ions in the accumulation/collision hexapole of a hybrid Q-FTMS. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 411-418.
- (5) Whitehouse, C. M.; Dreyer, R.; Yamashita, M.; Fenn, J.: Electrospray ionization for mass-spectrometry of large biomolecules. *Science* **1989**, *246*, 64-71.
- (6) Kebarle, P.; Tang, L.: From ions in solution to ions in the gas phase-the mechanism of electrospray mass spectrometry. *Anal. Chem.* **1993**, *65*, 972A-986A.
- (7) Taylor, G.: Disintegration of water drops in an electric field. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*; The Royal Society, 1964; Vol. 280; pp 383-397.

- (8) Iribarne, J.; Thomson, B.: On the evaporation of small ions from charged droplets. *J. Chem. Phys.* **1976**, *64*, 2287-2294.
- (9) Nguyen, S.; Fenn, J. B.: Gas-phase ions of solute species from charged droplets of solutions. *Proc. Natl. Acad. Sci.* **2007**, *104*, 1111-1117.
- (10) Gamero-Castano, M.; De La Mora, J. F., J.: Direct measurement of ion evaporation kinetics from electrified liquid surfaces. *J. Chem. Phys.* **2000**, *113*, 815-832.
- (11) De La Mora, J. F.: Electrospray ionization of large multiply charged species proceeds via Dole's charged residue mechanism. *Anal. Chim. Acta* **2000**, *406*, 93-104.
- (12) Comisarow, M. B.; Marshall, A. G.: Selective-phase ion cyclotron resonance spectroscopy. *Can. J. Chem.* **1974**, *52*, 1997-1999.
- (13) Nieckarz, R. J.; Oomens, J.; Berden, G.; Sagulenko, P.; Zenobi, R.: Infrared multiple photon dissociation (IRMPD) spectroscopy of oxazine dyes. *Phys. Chem. Chem. Phys.* **2013**, *15*, 5049-5056.
- (14) van Stipdonk, M. J.; Patterson, K.; Gibson, J. K.; Berden, G.; Oomens, J.: IRMPD spectroscopy reveals a novel rearrangement reaction for modified peptides that involves elimination of the N-terminal amino acid. *Int. J. Mass Spectrom* **2015**, *379*, 165-178.
- (15) Watkins, M. A.; Winger, B. E.; Shea, R. C.; Kenttämä, H. I.: Ion-molecule reactions for the characterization of polyols and polyol mixtures by ESI/FT-ICR mass spectrometry. *Anal. Chem.* **2005**, *77*, 1385-1392.
- (16) Nibbering, N. M.: Gas-phase ion/molecule reactions as studied by Fourier transform ion cyclotron resonance. *Acc. Chem. Res.* **1990**, *23*, 279-285.

- (17) Stenson, A. C.; Ruddy, B. M.; Bythell, B. J.: Ion molecule reaction H/D exchange as a probe for isomeric fractionation in chromatographically separated natural organic matter. *Int. J. Mass Spectrom* **2014**, *360*, 45-53.
- (18) Angthararuk, D.; Harir, M.; Schmitt-Kopplin, P.; Sutthivaiyakit, S.; Kettrup, A.; Sutthivaiyakit, P.: Degradation products of profenofos as identified by high-field FTICR mass spectrometry: Isotopic fine structure approach. *J Environ Sci Health B* **2017**, *52*, 10-22.
- (19) Ruddy, B. M.; Blakney, G. T.; Rodgers, R. P.; Hendrickson, C. L.; Marshall, A. G.: Elemental composition validation from stored waveform inverse Fourier transform (SWIFT) isolation FT-ICR MS isotopic fine structure. *J. Am. Soc. Mass Spectrom.* **2013**, *24*, 1608-1611.
- (20) Marshall, A. G.; Hendrickson, C. L.; Jackson, G. S.: Fourier transform ion cyclotron resonance mass spectrometry: A primer. *Mass Spectrom. Rev.* **1998**, *17*, 1-35.
- (21) Marshall, A. G.; Hendrickson, C. L.: Fourier transform ion cyclotron resonance detection: principles and experimental configurations. *Int. J. Mass Spectrom* **2002**, *215*, 59-75.
- (22) Barrow, M. P.; Burkitt, W. I.; Derrick, P. J.: Principles of Fourier transform ion cyclotron resonance mass spectrometry and its application in structural biology. *Analyst* **2005**, *130*, 18-28.
- (23) Sleno, L.; Volmer, D. A.: Ion activation methods for tandem mass spectrometry. *J. Mass Spectrom.* **2004**, *39*, 1091-1112.
- (24) Cooks, R. G.: Special feature: Historical. Collision-induced dissociation: Readings and commentary. *J. Mass Spectrom.* **1995**, *30*, 1215-1221.
- (25) Hayes, R. N.; Gross, M. L.: Collision-induced dissociation. *Methods Enzymol.* **1990**, *193*, 237-263.

- (26) Herrmann, K. A.; Somogyi, Á.; Wysocki, V. H.; Drahos, L.; Vékey, K.: Combination of sustained off-resonance irradiation and on-resonance excitation in FT-ICR. *Anal. Chem.* **2005**, *77*, 7626-7638.
- (27) Senko, M. W.; Speir, J. P.; McLafferty, F. W.: Collisional activation of large multiply charged ions using Fourier transform mass spectrometry. *Anal. Chem.* **1994**, *66*, 2801-2808.
- (28) Price, W. D.; Schnier, P. D.; Williams, E. R.: Tandem mass spectrometry of large biomolecule ions by blackbody infrared radiative dissociation. *Anal. Chem.* **1996**, *68*, 859-866.
- (29) Dunbar, R. C.: BIRD (blackbody infrared radiative dissociation): Evolution, principles, and applications. *Mass Spectrom. Rev.* **2004**, *23*, 127-158.
- (30) Dunbar, R. C.; McMahon, T. B.; Thoelmann, D.; Tonner, D. S.; Salahub, D. R.; Wei, D.: Zero-pressure thermal-radiation-induced dissociation of gas-phase cluster ions: comparison of theory and experiment for $(\text{H}_2\text{O})_2\text{Cl}^-$ and $(\text{H}_2\text{O})_3\text{Cl}^-$. *J. Am. Chem. Soc.* **1995**, *117*, 12819-12825.
- (31) Price, W. D.; Williams, E. R.: Activation of Peptide ions by blackbody radiation: factors that lead to dissociation kinetics in the rapid energy exchange limit. *J. Phys Chem A* **1997**, *101*, 8844-8852.
- (32) Dunbar, R. C.; McMahon, T. B.; Thdlmann, D.; Tanner, D. S.: Zero-pressure thermal-radiation-induced dissociation of gas-phase cluster ions: comparison of theory and experiment. *J. Am. Chem. Soc.* **1995**, *117*, 12819-12825.
- (33) Roithová, J.; Gray, A.; Andris, E.; Jašík, J.; Gerlich, D.: Helium tagging infrared photodissociation spectroscopy of reactive ions. *Acc. Chem. Res.* **2016**, *49*, 223-230.

- (34) Kamrath, M. Z.; Relph, R. A.; Guasco, T. L.; Leavitt, C. M.; Johnson, M. A.: Vibrational predissociation spectroscopy of the H₂-tagged mono- and dicarboxylate anions of dodecanedioic acid. *Int. J. Mass Spectrom* **2011**, *300*, 91-98.
- (35) Altinay, G.; Metz, R. B.: Comparison of IRMPD, Ar-tagging and IRLAPS for vibrational spectroscopy of Ag⁺(CH₃OH). *Int. J. Mass Spectrom* **2010**, *297*, 41-45.
- (36) Bomse, D.; Woodin, R.; Beauchamp, J.: Molecular activation with low-intensity CW infrared laser radiation. Multiphoton dissociation of ions derived from diethyl ether. *J. Am. Chem. Soc.* **1979**, *101*, 5503-5512.
- (37) Isenor, N.; Merchant, V.; Hallsworth, R.; Richardson, M.: CO₂ laser-induced dissociation of SiF₄ molecules into electronically excited fragments. *Can. J. Phys.* **1973**, *51*, 1281-1287.
- (38) Woodin, R.; Bomse, D.; Beauchamp, J.: Multiphoton dissociation of molecules with low power continuous wave infrared laser radiation. *J. Am. Chem. Soc.* **1978**, *100*, 3248-3250.
- (39) Prazeres, R.; Glotin, F.; Insa, C.; Jaroszynski, D.; Ortega, J.: Two-colour operation and applications of the CLIO FEL in the mid-infrared range. *Nucl. Instr. Meth. Phys. Res.* **1998**, *407*, 464-469.
- (40) Oh, H.; Breuker, K.; Sze, S. K.; Ge, Y.; Carpenter, B. K.; McLafferty, F. W.: Secondary and tertiary structures of gaseous protein ions characterized by electron capture dissociation mass spectrometry and photofragment spectroscopy. *Proc. Natl. Acad. Sci.* **2002**, *99*, 15863-15868.
- (41) Gross, M. L.; Caprioli, R.; Nibbering, N. M. M.: *The encyclopedia of mass spectrometry. Volume 4, Volume 4*; Elsevier: Amsterdam [etc.], 2004.
- (42) Polfer, N. C.; Oomens, J.: Vibrational spectroscopy of bare and solvated ionic complexes of biological relevance. *Mass Spectrom. Rev.* **2009**, *28*, 468-494.

- (43) Armstrong, J.; Bloembergen, N.; Ducuing, J.; Pershan, P.: Interactions between light waves in a nonlinear dielectric. *Phys. Rev.* **1962**, *127*, 1918.
- (44) Giordmaine, J.; Miller, R. C.: Tunable coherent parametric oscillation in LiNbO₃ at optical frequencies. *Phys. Rev. Lett.* **1965**, *14*, 973.
- (45) Elder, F. R.; Gurewitsch, A. M.; Langmuir, R. V.; Pollock, H. C.: Radiation from electrons in a synchrotron. *Phys. Rev.* **1947**, *71*, 829-830.
- (46) Oepts, D.; van der Meer, A. F. G.; van Amersfoort, P. W.: The free-electron-laser user facility FELIX. *Infrared Phys Techn* **1995**, *36*, 297-308.
- (47) Silfvast, W. T.: *Laser fundamentals*; Cambridge University Press: Cambridge, 1996.
- (48) Bader, R. F. W.: Atoms in molecules. In *Encyclopedia of Computational Chemistry*; John Wiley & Sons, Ltd, 2002.
- (49) Cramer, C. J.: *Essentials of computational chemistry: theories and models*; John Wiley & Sons, 2013.
- (50) Jensen, F.: *Introduction to computational chemistry*; John Wiley & Sons, 2016.
- (51) von Ragué Schleyer, P.: Encyclopedia of computational chemistry. *J. Am. Chem. Soc* **1990**, *112*, 6127-6129.
- (52) Polfer, N. C.; Oomens, J.; Dunbar, R. C.: IRMPD spectroscopy of metal-ion/tryptophan complexes. *Phys. Chem. Chem. Phys.* **2006**, *8*, 2744-2751.
- (53) Wu, R.; McMahon, T. B.: An investigation of protonation sites and conformations of protonated amino acids by IRMPD spectroscopy. *ChemPhysChem* **2008**, *9*, 2826-2835.

- (54) Prell, J. S.; Flick, T. G.; Oomens, J.; Berden, G.; Williams, E. R.: Coordination of trivalent metal cations to peptides: results from IRMPD spectroscopy and theory. *J Phys Chem A* **2009**, *114*, 854-860.
- (55) Lewars, E. G.: *Computational chemistry: introduction to the theory and applications of molecular and quantum mechanics*; Springer, 2016.
- (56) Rivail, J.; Rinaldi, D.; Leszczynski, J.: Computational chemistry: Review of current trends. *World Scientific Publishing* **1995**.
- (57) Hehre, W. J.: *Ab initio molecular orbital theory*; Wiley-Interscience, 1986.
- (58) Combes, J. M.; Duclos, P.; Seiler, R.: The Born-Oppenheimer approximation. In *Rigorous Atomic and Molecular Physics*; Velo, G., Wightman, A. S., Eds.; Springer US: Boston, MA, 1981; pp 185-213.
- (59) Fischer, C. F.: Hartree--Fock method for atoms. A numerical approach. **1977**.
- (60) Johnson, E. R.; Becke, A. D.: A post-Hartree-Fock model of intermolecular interactions: Inclusion of higher-order corrections. *J. Chem. Phys.* **2006**, *124*, 174104.
- (61) Bartlett, R. J.; Stanton, J. F.: Applications of Post-Hartree-Fock methods: A tutorial. *Rev. Comput. Chem.* **2007**, 65-169.
- (62) Magnasco, V.: Post-Hartree-Fock Methods. *Methods of molecular quantum mechanics: An Introduction to Electronic Molecular Structure*, 133-139.
- (63) Møller, C.; Plesset, M. S.: Note on an approximation treatment for many-electron systems. *Phys. Rev.* **1934**, *46*, 618-622.
- (64) Parr, R. G.: Density functional theory of atoms and molecules. In *Horizons of Quantum Chemistry*; Springer, 1980; pp 5-15.

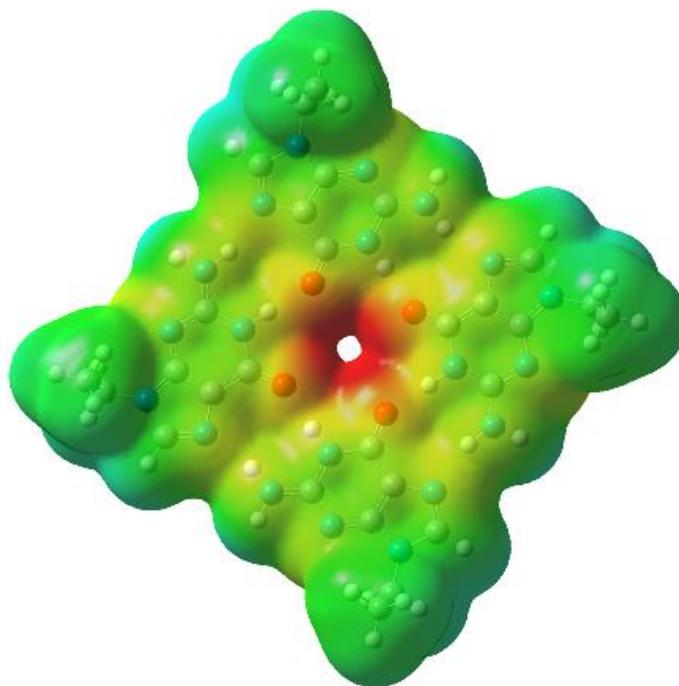
- (65) Schlüter, M.; Sham, L. J.: Density functional theory. *Phys Today* **1982**, *35*, 36-43.
- (66) Gross, E. K.; Dreizler, R. M.: *Density functional theory*; Springer Science & Business Media, 2013; Vol. 337.
- (67) Hohenberg, P.; Kohn, W.: Inhomogeneous electron gas. *Phys. Rev.* **1964**, *136*, B864-B871.
- (68) Kohn, W.; Sham, L. J.: Self-consistent equations including exchange and correlation effects. *Phys. Rev.* **1965**, *140*, A1133-A1138.
- (69) Chong, D. P.: *Recent Advances in Density Functional Methods:(Part I)*; World Scientific, 1995; Vol. 1.
- (70) Labanowski, J. K.; Andzelm, J. W.: *Density functional methods in chemistry*; Springer Science & Business Media, 2012.
- (71) Becke, A. D.: Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A* **1988**, *38*, 3098-3100.
- (72) Lee, C.; Yang, W.; Parr, R. G.: Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **1988**, *37*, 785-789.
- (73) van Lenthe, E.; Baerends, E. J.: Optimized Slater-type basis sets for the elements 1–118. *J. Comput. Chem.* **2003**, *24*, 1142-1156.
- (74) de la Vega, J. G.; Miguel, B.: Basis sets for computational chemistry. *Introduction to Advanced Topics of Computational Chemistry*, edited by LA Montero, LA Diaz, and R. Bader, 41-80.
- (75) Rassolov, V. A.; Pople, J. A.; Ratner, M. A.; Windus, T. L.: 6-31G* basis set for atoms K through Zn. *J. Chem. Phys.* **1998**, *109*, 1223-1229.

- (76) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A.: Self-consistent molecular orbital methods. XX. A basis set for correlated wave functions. *J. Chem. Phys.* **1980**, *72*, 650-654.
- (77) Francl, M. M.; Pietro, W. J.; Hehre, W. J.; Binkley, J. S.; Gordon, M. S.; DeFrees, D. J.; Pople, J. A.: Self-consistent molecular orbital methods. XXIII. A polarization-type basis set for second-row elements. *J. Chem. Phys.* **1982**, *77*, 3654-3665.
- (78) Hariharan, P.; Pople, J.: Theory of splitting basis set calculation. *Chim. Acta* **1973**, *28*, 213.
- (79) Cundari, T. R.; Benson, M. T.; Lutz, M. L.; Sommerer, S. O.: Effective core potential approaches to the chemistry of the heavier elements. *Rev. Comput. Chem.* **1996**, 145-202.
- (80) Hay, P. J.; Wadt, W. R.: Ab initio effective core potentials for molecular calculations. Potentials for the transition metal atoms Sc to Hg. *J. Chem. Phys.* **1985**, *82*, 270-283.
- (81) Cundari, T. R.; Stevens, W. J.: Effective core potential methods for the lanthanides. *J. Chem. Phys.* **1993**, *98*, 5555-5565.
- (82) Frisch, M.; Trucks, G.; Schlegel, H.; Scuseria, G.; Robb, M.; Cheeseman, J.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G.: Gaussian 09, revision D. 01. Gaussian, Inc., Wallingford CT, 2009.
- (83) Dunbar, R. C.; McMahon, T. B.: Activation of unimolecular reactions by ambient blackbody radiation. *Science* **1998**, *279*, 194-197.
- (84) Price, W. D.; Schnier, P. D.; Williams, E. R.: Binding energies of the proton-bound amino acid dimers Gly·Gly, Ala·Ala, Gly·Ala, and Lys·Lys measured by blackbody infrared radiative dissociation. *J. Phys. Chem. B* **1997**, *101*, 664-673.

(85) Price, W. D.; Schnier, P. D.; Jockusch, R. A.; Strittmatter, E. F.; Williams, E. R.: Unimolecular reaction kinetics in the high-pressure limit without collisions. *J. Am. Chem. Soc.* **1996**, *118*, 10640-10644.

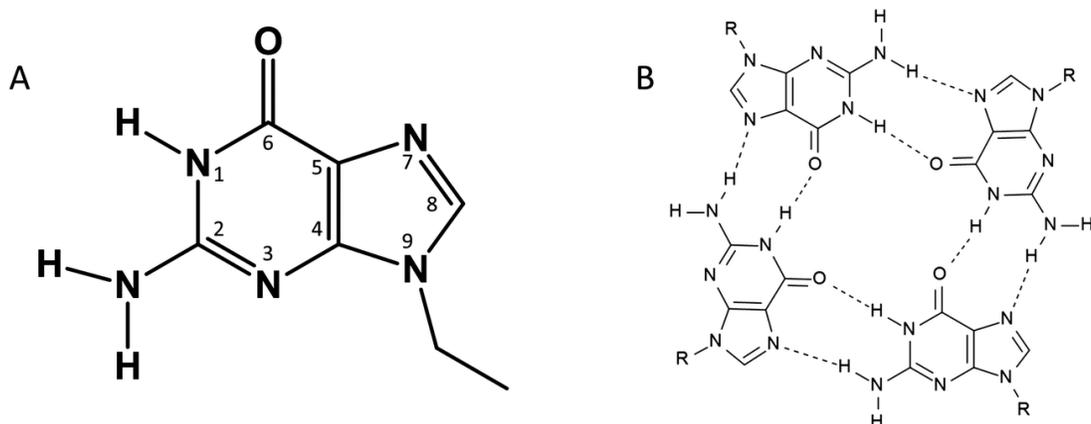
Chapter 3. Guanine Tetrads: An IRMPD Spectroscopy, Energy Resolved SORI-CID, and Computational Study of M(9-ethylguanine)₄⁺ (M=Li, Na, K, Rb, Cs) in the Gas Phase

This work has been published in part as *M. Azargun and T.D. Fridgen.*, Guanine tetrads: an IRMPD spectroscopy, energy resolved SORI-CID, and computational study of M(9-ethylguanine)₄⁺ (M = Li, Na, K, Rb, Cs) in the gas phase., *Phys. Chem. Chem. Phys.*, 2015, 17, 25778-25785. This work has been conducted and written by me and under supervision of Dr. Travis D. Fridgen.



3.1. Introduction

The guanine nucleobase (Scheme 3.1A) can form hydrogen-bonded self-assemblies^{1,2,3,4} (Scheme 3.1B) due to its unique arrangement of both a pair of hydrogen donating (N1, NH₂) and accepting (N7, O6) groups. Stable tetrameric clusters, called G-tetrads (or G-quartets), are known to be formed in guanine rich parts of nucleic acids such as DNA.³ Metal cations are known to facilitate G-tetrad formation.⁶ Monovalent metallic cations can be accommodated in the cavity at the center of the tetrad and bound by the carbonyl oxygen atoms of all four guanines. These metal cations neutralize and stabilize the negative electrostatic potential produced by four oxygen atoms in close proximity.^{2,7} The carbohydrate and phosphate backbone make G-tetrad complexes more stackable forming G-quadruplexes.^{8,9}



Scheme 3.1

G-tetrad and G-quadruplex clusters have received much attention as they contribute to important biological processes.^{10,11} At one time, G-quadruplex clusters were believed to be exclusively formed in human telomeric DNA^{12,13,14,15} protecting the chromatid from fusing with neighboring peers and preventing loss of genetic information during replication. Recently, however, research has shown, that G-quadruplexes exist in various locations of human DNA performing numerous functions.^{16,17,18,19,20} For instance, Beaudoin and coworkers²⁰ have studied human mRNA and clearly showed that G-quadruplex can also act as translational suppressor, stopping RNA transcription. In fact, very recently guanine-rich tracks within DNA have been linked to amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD).^{21, 22, 23}

Not only do metallic monovalent cations like Li^+ , Na^+ , and K^+ lower the negative charge density created by the oxygen atoms, but they also can induce formation of these clusters.⁴ Further, although it has been commonly accepted that G-tetrads are templated by cations, work done in the solution phase using an (N,N-dimethylaniline)guanosine

derivative, indicates that clusters can also form in the absence of metal cations.²⁴ The alkali metal ion affinity of guanine and other nucleobases has been examined in the gas phase by Cerda and coworkers²⁵ using the kinetic method in which a metallated heterodimer, $[\text{nucleobase} + \text{B}]\text{M}^+$ where B is a reference base with known cation affinity, is dissociated at different internal energies. The dissociation rates of the heterodimers into both metallated monomers are compared to obtain the difference in metal cation affinities of monomers. Results show that Li^+ is the most strongly bound alkali metal to guanine, followed by Na^+ and K^+ which shows that these complexes are ion-dipole complexes where the smaller ion has a strong electrostatic interaction. These kinetic method values agree closely with those determined by the threshold collision-induced dissociation (TCID) technique conducted by Rodgers and Armentrout.²⁶

Experimental studies in the gas phase along with computational techniques show that the stability of the complexes, G-tetrads and G-quadruplexes, and also their structures correlate with the size of metal cations. The nucleobases in a tetrad rearrange from the planar geometry when a metal cation is present, presumably to maximize ion-dipole interactions while maintaining the strong hydrogen bonding network.⁷ Fukushima and Iwahashi² concluded that Na^+ is more strongly bound to 9-ethylguanine tetrads than Li^+ and K^+ in the solution phase by comparing peak intensities of the corresponding metal-bound tetrads when a mixture of alkali metal cations and 9-ethylguanine solution was electrosprayed into a mass spectrometer.

Divalent and trivalent cations behave similarly to monovalent cations, forming huge multi-quartet clusters or quadruplexes when exposed to guanine and its derivatives.

Cd^{2+} -bound guanosine tetrads were observed following electrospray of a solution of Cd^{2+} and guanosine.²⁷ In addition, an electrospray mass spectrometry study performed by Kwan *et al.*²⁸ revealed that trivalent lanthanide cations like Tb^{3+} can form G-tetrad adducts as well as octameric complexes.

Ion mobility work²⁹ demonstrated the potency of the ammonium cation to initiate the formation of self-assembled clusters. Sandwich-shaped clusters were observed and their cross-sections were measured and compared to those determined by theoretical methods. Results verify the formation of quadruplexes and G-tetrads in presence of non-metallic cations and more importantly these complexes are stable in solution and keep their structures.

Since the advent of ESI³⁰, mass spectrometry techniques have become widely used to study non-volatile biological molecules such as nucleobases and their clusters. ESI-MS coupled to a plethora of activation techniques^{31,32,33,34,35} have become very useful for studying the intrinsic physical properties of gaseous biological ions. Collisional activation is one of the most common activation techniques.^{36,37} Sustained off-resonance irradiation collision-induced dissociation (SORI-CID) is the Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR) version of the CID experiment. It has unique advantages including providing more control on the kinetic energy of parent ions and the ability to increase the activation time resulting in more collisions.³⁷

To our knowledge, the intrinsic physical chemistries of alkali metal cation-bound G-tetrad clusters have not been examined or compared experimentally in terms of structures and stability. In this work, we have used the energy-resolved SORI-CID

activation technique along with computational chemistry to gain more insight into the effect of different alkali metal cations on the structure and energetics of the G-tetrad. The IRMPD spectrum of $\text{Na}(\text{Gua})_4^+$ has been recently published, and is consistent with a planar structure in which Na^+ is at the center of the G-tetrad.³⁸ Here, we also present and compare IRMPD spectra in both the fingerprint and N-H/C-H stretching region of G-tetrad complexes with all the alkali metals.

3.2. Methods

3.2.1. Experimental

A Bruker ApexQe 7.0 hybrid Fourier transform ion cyclotron resonance (FTICR) mass spectrometer was used to carry out all experiments. 9-ethylguanine was purchased from Sigma-Aldrich and used without further purification like all other chemicals. To prepare solutions, a few drops of 1 mM salt solution (LiCl, NaCl, KCl, RbCl, and CsCl) were added to 10 mL of 0.1 mM 9-ethylguanine solution in 18 M Ω water. The final solution was electrosprayed by an Apollo II ion source coupled to FTICR mass spectrometer at 70-150 $\mu\text{L h}^{-1}$. SORI-CID experiments were done by isolating the tetramers under study inside the ICR cell ($P = 10^{-10}$ mbar) and exposing them to Ar inside the ICR cell at higher pressures ($P \approx 10^{-5} - 10^{-6}$ mbar). At these pressures, in the 250 ms excitation time, there are on the order of 10's to 100's of collisions. Since the activation is done by SORI, the kinetic energies vary over the period of excitation and we quote the maximum center of mass kinetic energies. The average kinetic energies are expected to

be 2/3 the maximum kinetic energies. The maximum lab frame kinetic energies are computed using the following equation:³⁹

$$E_{\text{lab}}^{\text{max}} = \frac{\beta^2 q^2 V_{p-p}}{32\pi^2 m d^2 \Delta\nu^2}$$

where β is a geometrical factor of the ICR cell (0.9 in the present case), q is the charge on the ion, V_{p-p} is the peak to peak excitation voltage, m is the mass of the ion, d is the diameter of the ICR cell (6 cm) and $\Delta\nu$ is the frequency offset (500 Hz). The lab frame energies were multiplied by $m_{\text{Ar}} / (m_{\text{Ar}} + m_{\text{M}(\text{9eG})_i^+})$ to obtain the center of mass kinetic energies whose maxima ranged from 0.1 to 2.3 eV.

IRMPD experiments were performed by using two setups. IRMPD spectra in the 2700 – 4000 cm^{-1} region were obtained in the Laboratory for the Study of the Energetics, Structures, and Reactions of Gaseous Ions at Memorial University using an IR OPO, manufactured by LaserSpec, tuneable from 1.4 to 4.5 μm , with bandwidth of 2 cm^{-1} . The OPO, built around a periodically poled lithium niobate crystal, is pumped by a diode pump solid state Nd:YAG laser. The OPO operates at 20 kHz, with pulse duration of few nanoseconds and generates output power near 3 W at 3 μm . The power was limited to 1 watt in the present experiments. In the fingerprint region, all experiments were performed using a Fourier-transform ion cyclotron resonance mass spectrometer (FT-ICR-MS) coupled to a mid-infrared free electron laser (FEL) (5 cm^{-1} bandwidth) at the Centre Laser Infrarouge d'Orsay (CLIO).

3.2.2. Computational

In this research, the Gaussian 09⁴⁰ software package was used to perform all calculations. Hybrid density functional (DFT) techniques have shown their potential and reliability to structurally optimize and also determine the lowest energy structures of nucleobase clusters. B3LYP density functional theory with the 6-31+G(d,p) basis set on C, H, N, and O and the Def2SVPD basis set and ECP's were used on all metals were used for optimizations and calculation of the infrared spectra. Electronic energies were refined using the 6-311+G(3df,3pd) basis on C, H, N, and O and the Def2TZVPP basis and ECP on all metals. 298 K relative thermochemistries (enthalpies and Gibbs energies) were calculated from B3LYP/6-311++G(3df,3pd) electron energies and thermal corrections from the B3LYP/6-31+G(d,p) (note the different basis for metals). These thermochemical values are denoted as B3LYP/6-311++G(3df,3pd)//B3LYP 6-31+G(d,p). All dissociation energies were corrected for basis set superposition error using the counterpoise correction method as implemented in G09. All calculations were corrected for dispersion using GD3 empirical corrections.⁴¹

3.3. Results and Discussion

In this research, 9-ethylguanine was chosen since it is significantly more soluble compared to guanine resulting in mass spectra with improved intensities. Moreover, N9 is blocked by an ethyl group eliminating the possibility for interaction of N9 with the metal cation, blocking it from hydrogen bonding, and preventing it from participating in

tautomerization. It is also a better model for biological systems as N9 is the site of glycosylation in nucleic acids.

An electrospray mass spectrum of an aqueous solution prepared by adding only 9-ethylguanine and KCl is shown in Figure 3.1. The mass spectrum reveals the desired potassiated tetrad, $K(9eG)_4^+$, at m/z 755.3 as well as a potassiated octamer, $K(9eG)_8^+$, at m/z 1471 and the doubly charged ion $K_2(9eG)_{12}^{2+}$ at m/z 1113. Clearly, complexes composed of multiples of four 9-ethylguanines are magic number complexes, with other complexes (i.e. dimeric or trimeric) observed only in minute quantities. To our knowledge, the $K(9-eG)_8^+$, and $K_2(9-eG)_{12}^+$, are the largest guanine quadruplexes observed in the gas phase. However, the most abundant ion observed is the sodiated tetrad, $Na(9-eG)_4^+$, at m/z 739. The observation of $Na(9-eG)_4^+$ when Na^+ is only present as an impurity is consistent with previous work concluding the exceptional solution phase stability of the $Na(9-eG)_4^+$ over tetrads associated with other metal cations.² It should also be noted that despite the large intensity of $Na(9-eG)_4^+$, no larger clusters are observed. This is the topic of a future manuscript on guanine quadruplexes.

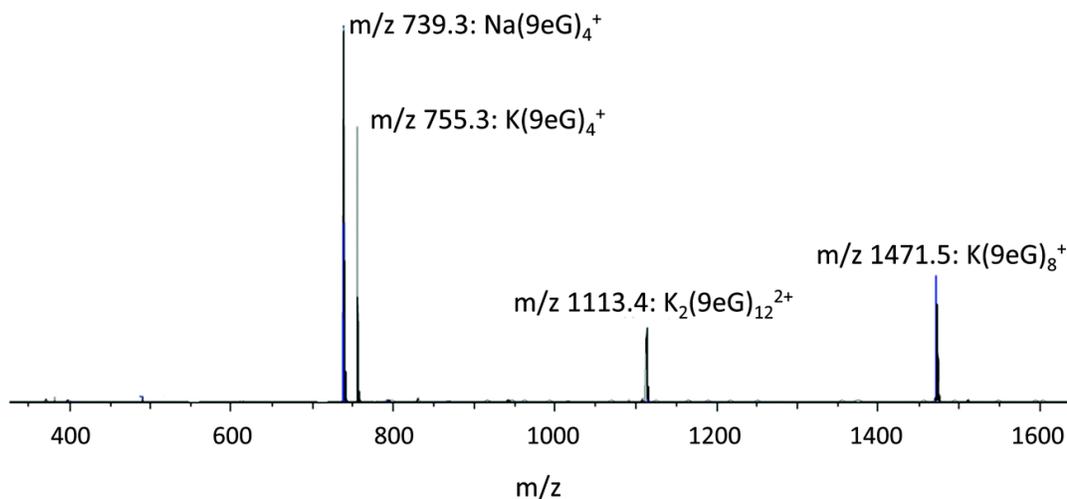


Figure 3.1. Electrospray mass spectrum of 10 mL of 0.1 mM solution of 9-ethylguanine to which two drops of 0.1 mM KCl was added.

3.3.1 Computed Structures of Metalated G-Tetrads, $\text{M}(9\text{eG})_4^+$

The computed structures of the G-tetrads are summarized in Figure 3.2 with structural parameters provided in Table 3.1. Individual snapshots of each $\text{M}(9\text{eG})_4^+$ structure are given in Figure S3.1. In all cases the O—O distances are significantly shorter compared to the neutral G-tetrad. The addition of Li^+ to the neutral G-tetrad shrinks the O—O distance by almost 1 Å due to strong ion-dipole interactions. In $\text{Na}(9\text{eG})_4^+$ the O—O distance is shortened by about 0.5 Å and for the largest ion, Cs^+ , the distance is only shorter by a few hundredths of an angstrom compared to the neutral tetramer. The greater contraction of the tetrad with the smaller sized alkali metal cation is consistent with ion-dipole interactions occurring between the metal cation and the 9-ethylguanines. The NH—N hydrogen bond is also shorter for all the metal cation/G-tetrad complexes due to the increasing ionic nature of the hydrogen bond by addition of the charged metal. Interestingly, the NH—N hydrogen bond is the shortest in $\text{Na}(9\text{eG})_4^+$,

of all the complexes. The NH—O hydrogen bond increases slightly upon the addition of a metal cation which is expected since the metal cation is also bound to the oxygen atom in each of the 9-ethylguanines. $\text{Na}(\text{9eG})_4^+$ is the only planar G-tetrad other than the neutral complex. $\text{Li}(\text{9eG})_4^+$ is very different from the other structures as the Li^+ sits inside a rather twisted G-tetrad in order to maximize the ion-dipole interactions but also to relieve any repulsion that would be present by not allowing the hydrogen bonds to get too short, yet maintaining the strong hydrogen bonds. For K^+ , Rb^+ , and Cs^+ , the metal cation sits on top of a non-planar, slightly distorted G-tetrad. The larger the metal cation, the higher above the tetrad the metal cation sits as indicated by the OOM angle. There are also two values for each of the structural parameters listed for the K^+ , Rb^+ , and Cs^+ complexes. That is due to the complexes resembling, somewhat, two 9eG dimers in those tetraplexes.

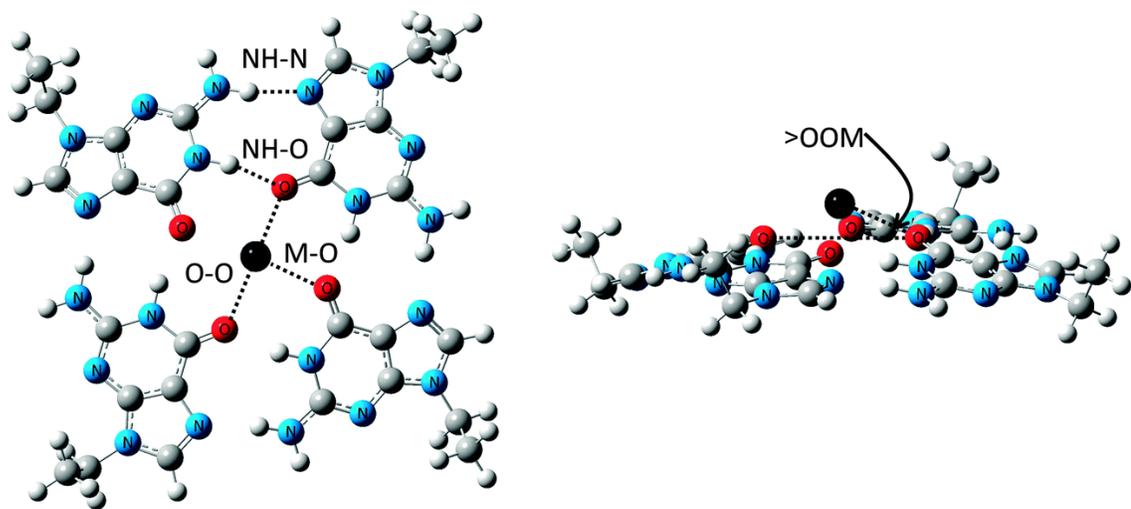


Figure 3.2. Computed structures of the metal cationized G-tetrads indicating the structural parameters listed in Table 3.1.

Table 3.1. B3LYP/6-31+G(d,p) structural parameters for the $M(9eG)_4^+$ guanine tetrads, $M = Li, Na, K, Rb,$ and $Cs,$ and $(9eG)_4.$

Species	O-O	NH—N	NH—O	M-O	π OOM
$(9eG)_4$	4.991	1.970	1.755		
$Li(9eG)_4^+$	3.947	1.895	1.776	2.031	13.9
$Na(9eG)_4^+$	4.582	1.866	1.839	2.291	0.0
$K(9eG)_4^+$	4.781/4.847	1.888/1.897	1.855/1.864	2.655/2.660	24.1/26.0
$Rb(9eG)_4^+$	4.797/4.898	1.886/1.901	1.844/1.856	2.838/2.846	30.4/32.6
$Cs(9eG)_4^+$	4.841/4.946	1.889/1.911	1.829/1.844	3.028/3.046	35.2/38.5

3.3.2 IRMPD Spectroscopy of $M(9eG)_4^+$

The IRMPD spectra for the $M(9eG)_4^+$ tetrads are compared to each other and to the computed spectra for the lowest-energy structures in both the 2700 – 3800 cm^{-1} and the 900 – 1850 cm^{-1} regions in Figure 3.3. The experimental spectra for all of the $M(9eG)_4^+$ are quite similar to one another in both regions, indicating similarities in their structures. In fact, there really is no way to distinguish between the complexes composed of a different alkali metal cations based on the IRMPD spectra. More importantly, the IRMPD spectra and the computed IR spectra for the lowest energy structures are in very good agreement although there is one observed band in the high energy region that is not accounted for by the computed IR spectra and will be discussed below. We also note that the present experimental spectrum of $Na(9eG)_4^+$ is in good agreement with the those of $Na(guanine)_4^+$ which was concluded to belong to a planar complex similar to the lowest energy complex shown in Figure 3.2 for $Na(9eG)_4^+.$ ³⁸

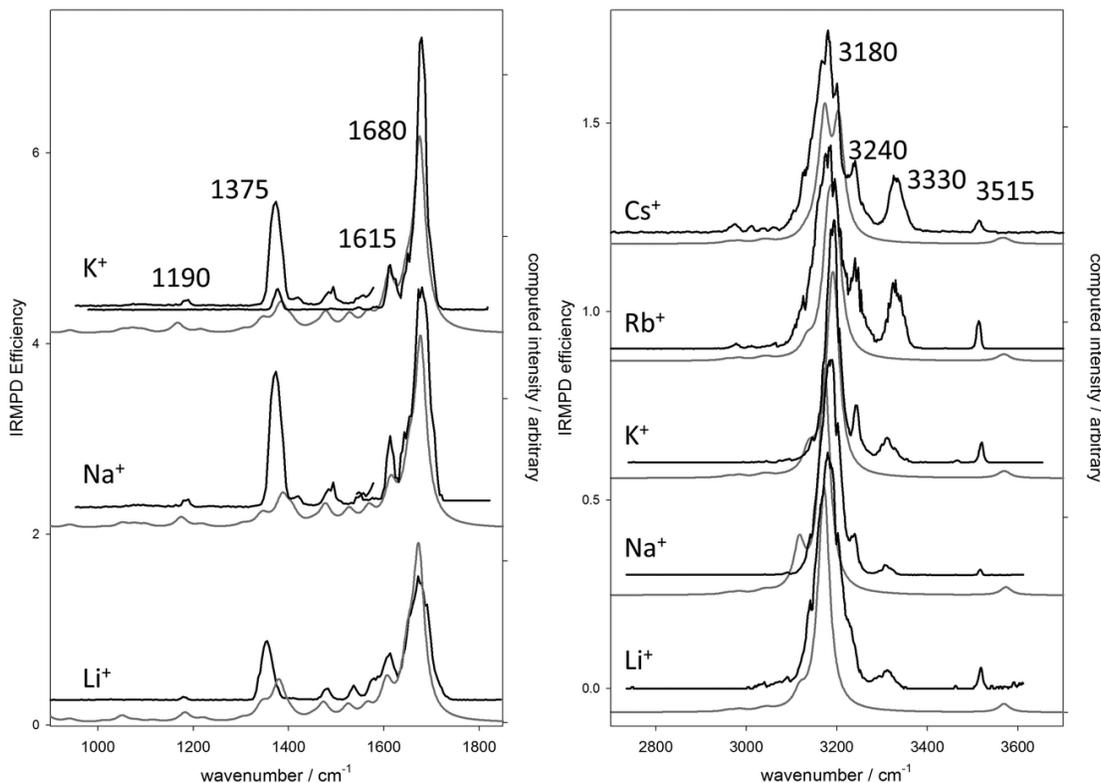


Figure 3.3. IRMPD spectra of the $M(9eG)_4^+$ G-tetrads in the $900 - 1850 \text{ cm}^{-1}$ and the $2700 - 3800 \text{ cm}^{-1}$ regions (black traces). The underlying grey traces are the computed spectra (B3LYP/6-31+G(d,p), Def2SVPD on metals) for the lowest energy G-tetrad structures scaled by 0.97 in both regions. The insets for the fingerprint regions of the Na^+ and K^+ tetramers are without an attenuating element, roughly 10x the laser intensity.

In the lower energy region, the most prominent band is at 1680 cm^{-1} due to C=O stretching and C-N-H bending and a resolved shoulder at about 1615 cm^{-1} due to C-N-H bending. There is also a band at 1375 cm^{-1} due to C-NH₂ stretching and CH₂ wagging. Between these main absorptions are three weak features (at ~ 1480 , 1535 , and 1575 cm^{-1}) due mainly to C-C and C-N stretching. Another weak band is observed at 1190 cm^{-1} due to deformation and C-N stretching. In the higher energy region, the main absorption occurs at about 3180 cm^{-1} and is due to the hydrogen bonded N-H stretching vibrations.

The highest energy band occurs at 3515 cm^{-1} is due to the free N-H stretch (or antisymmetric stretch of the NH_2 group). These calculations predict this band to be at a slightly higher wavenumber position, about 3570 cm^{-1} . Harmonic calculations have been shown to overestimate the wavenumber position of the antisymmetric NH_2 stretch of adenine, a similar purine base,^{42,43,44} as well as cytosine.⁴⁵ Anharmonic calculations were shown to correctly reproduce the position of the NH_2 antisymmetric stretch and our own anharmonic frequency calculations on $\text{Na}(\text{9eG})_4^+$ predict the NH_2 stretch to occur at 3502 cm^{-1} , in better agreement with the experimentally observed band. There is a strong band observed at 3330 cm^{-1} as well as a shoulder to the 3180 cm^{-1} band, at 3240 cm^{-1} that are resolved for the Na^+ , K^+ , Rb^+ and Cs^+ clusters and have no fundamental band predicted in these regions. An obvious possibility for these bands are overtones of the very strong C=O stretching absorption observed at 1680 cm^{-1} . Fraschetti *et al.* used *ab initio* molecular dynamics simulations³⁸ to show that these two absorptions could also belong to the hydrogen bonded N-H stretches of a rapidly fluctuating structure due to a very flat potential energy surface for distortion in and out of planarity of this non-covalently bonded complex.

In Figure S3.2 the experimental spectra for $\text{Na}(\text{9eG})_4^+$, $\text{Li}(\text{9eG})_4^+$, and $\text{K}(\text{9eG})_4^+$ are compared to computed spectra for the lowest energy isomer and two higher energy isomers that we were able to find and optimize. One of the isomers has Hoogsteen-type pairing, but with only three 9-ethylguanines, the fourth is bound to the metal through O6 and N7. These isomers are 56, 85, and 61 kJ mol^{-1} higher in Gibbs energy than the lowest energy isomer for the Li^+ , Na^+ , and K^+ tetrads. An even higher energy isomer has

the metal cation bound to all four 9-ethylguanines through N7 and interactions between O6 and H8, and is more than 140 kJ mol⁻¹ higher in Gibbs energy relative to the lowest energy structure. The best matches to the experimental spectra are the lowest energy G-tetrad structures. We conclude therefore, that the IRMPD spectra point to very similar G-tetrad-like structures with Hoogsteen base pairing in the gas phase for all five M(9eG)₄⁺ complexes, those shown in Fig. 3.2/S3.1.

3.3.3 Relative Gas-Phase Stabilities of the M(9eG)₄⁺ Tetrads by Energy-Resolved SORI-CID

Energy-resolved SORI-CID activation was used to determine the relative stabilities of the alkali metal cationized G-tetrads. After isolation, the G-tetrads were excited to collision energies of between 0 and 0.7 eV and exposed to Ar gas with a reservoir pressure of argon of 10 mbar. In Figure S3.3 are the breakdown diagrams for all five complexes. This Figure shows that in all cases, the complexes dissociate, losing sequential 9-ethyl guanine molecules as the trimer intensity occurs at lower energy than the dimer intensity followed by a rise in the monomer intensity. Only in the case of the Cs⁺ cluster is the bare ion observed and only at the highest collision energies. It would not be possible to observe a signal for the bare metal cations of Li⁺ or Na⁺ due to the low mass cut-off of the ICR due to the frequency generator. However, we note that in the K⁺ or Rb⁺ experiments over the energy range studied, no bare metal cations were observed. We cannot completely rule out the direct loss of a neutral dimer from the M(9eG)₄⁺ complexes. For example, for Na(9eG)₄⁺ there is only a small accumulation of Na(9eG)₃⁺

and the most abundant product is $\text{Na}(9\text{eG})_2^+$. This means that on a relative basis, it may be more energetically feasible to lose a neutral dimer than a neutral monomer.

In Figure 3.4, the $\text{M}(9\text{eG})_4^+$ intensities for all five G-tetrads are plotted against the collision energy. A line is drawn parallel to the energy axis at 50% dissociation of the G-tetrads. It is apparent that $\text{Na}(9\text{eG})_4^+$ requires the greatest energy to affect 50% dissociation, followed by $\text{Li}(9\text{eG})_4^+$ and then the K^+ , Rb^+ and Cs^+ G-tetrads. On the basis of ion-dipole interactions between the metal cations and guanine, the gas-phase stabilities of the tetrads would be expected to follow the order $\text{Li}^+ > \text{Na}^+ > \text{K}^+ > \text{Rb}^+ > \text{Cs}^+$ due to the decreasing charge density of the metal cation.^{25,26}

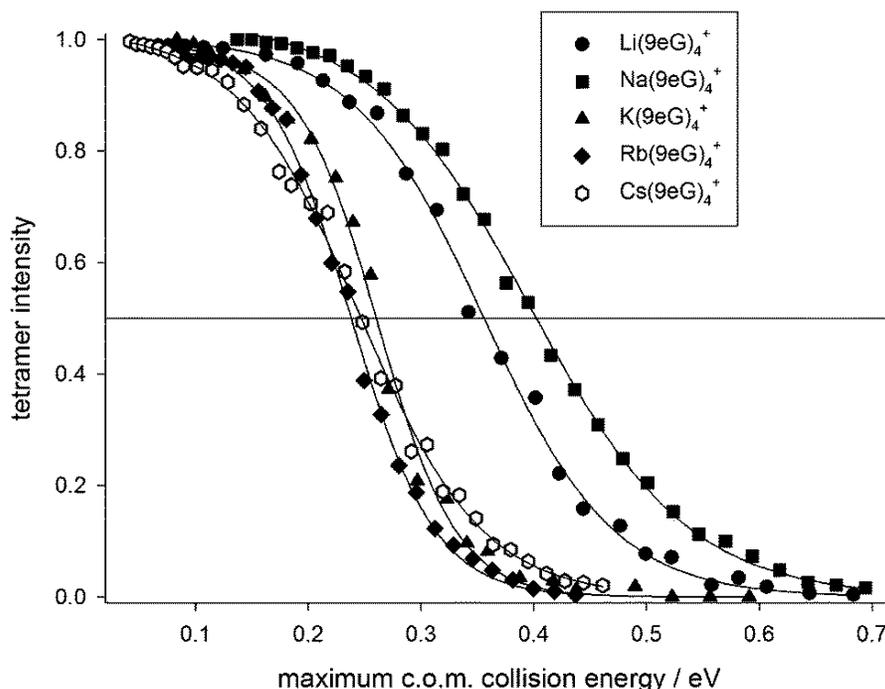


Figure 3.4. Tetramer intensity vs the center of mass energy decay curve for the $\text{M}(9\text{eG})_4^+$ G-tetrads from the energy-resolved SORI-CID spectra experiments. Reservoir pressure was 10 mbar with Ar. The higher energy required to dissociate $\text{Na}(9\text{eG})_4^+$ indicates a structure more stable to loss of 9-ethylguanine.

The experiments were repeated with an Ar reservoir pressure of 5 and 15 mbar (Figure S3.4), effectively decreasing and increasing the pressure inside the ICR cell. The energies required to affect dissociation were observed to increase at the lower pressure and decrease at the higher pressure due to a decreased and increased, respectively, collision frequency, as expected. The ordering of the energy required to dissociate $M(9eG)_4^+$ was the same at all three energies, $Na^+ > Li^+ > K^+ > Rb^+ > Cs^+$. In Figure 3.5, the relative center of mass energies to affect 50% dissociation are plotted along with the computed dissociation enthalpies for the following reaction:

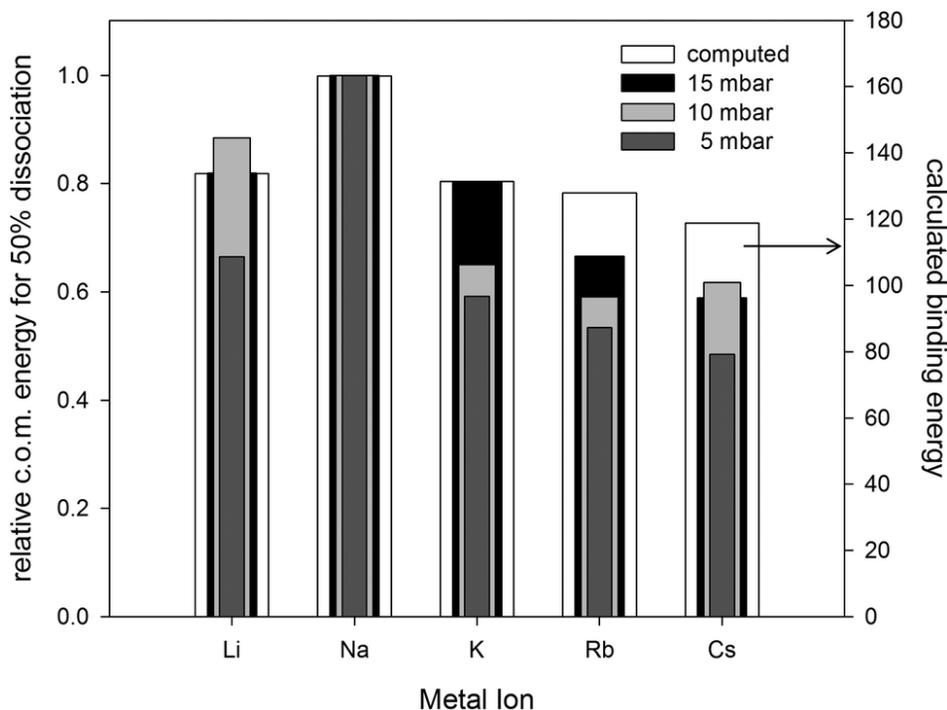


Figure 3.5. The relative center of mass collision energies (kJ mol^{-1}) to affect 50 % dissociation for the five alkali metal cationized G-tetrads. The experimental values were done at three different reservoir pressures as indicated in the legend. The white bars show the computed binding energies for loss of a 9-ethylganine from $M(9eG)_4^+$. The computed values are from the B3LYP/6-311+G(3df,3pd) calculations listed in Table 3.2.

It is clear from Figure 3.5 that the energy resolved SORI-CID experiments and theory agree that the Na⁺ tetrad is the most stable to dissociation. This agrees with conclusions from solution phase studies by simply electrospraying solutions containing 9-ethylguanine and all the metal cations at the same concentration.²

The computed 298 K dissociation enthalpies for both levels of theory are summarized in Table 3.2. As well, single-point calculations were done on the complexes, but after removing the metal cation. These energies (all computed with the larger basis) were compared to the energies of the neutral complex and were found to be 63.3, 30.9, 23.9, 21.9, and 20.2 kJ mol⁻¹ for the Li⁺, Na⁺, K⁺, Rb⁺, and Cs⁺ complexes, respectively. These energies are destabilization energies, the effect of the metal cation on the energy of the neutral G-tetrad by distorting it from its normal geometry. Clearly, the twisted and compact structure of Li(9eG)₄⁺ has the largest effect, distorting the tetrad, and raising its energy, by the greatest amount. Na⁺ distorts the tetrad by the next greatest amount followed by K⁺, Rb⁺, and Cs⁺. If these energies are added to the dissociation energies of the M(9eG)₄⁺ complexes and the dissociation energy of the neutral tetrad itself (37.1 kJ mol⁻¹) is removed, the resulting value (the last column of Table 3.2) represents the amount (at least relatively) by which the metal cation stabilizes the G-tetrad to loss of a neutral 9-eG. The trend for these values is a monotonic decrease as the size of the metal cation increases or the charge density decreases as would be expected for a purely ion-dipole complex. These calculations show that the secondary effect of the metal cation, distorting the G-tetrad to maximize ion-dipole interactions while maintaining hydrogen

bonding interactions, seems to be the reason for $\text{Li}(9\text{eG})_4^+$ being less stable than $\text{Na}(9\text{eG})_4^+$ in the gas phase.

Table 3.2. 298 K Dissociation enthalpies (kJ mol^{-1}) of neutral and alkali metal cationized G-tetrads, $(9\text{eG})_4$ and $\text{M}(9\text{eG})_4^+$.

Species	B3LYP/ 6-31+G(d,p) ^a	B3LYP/ 6-311+G(3df,3pd) ^b	Remove G-tetrad distortion energy ^c
$(9\text{eG})_4$	37.5	37.1	
$\text{Li}(9\text{eG})_4^+$	106.8	133.8	160.0
$\text{Na}(9\text{eG})_4^+$	136.3	163.4	157.2
$\text{K}(9\text{eG})_4^+$	104.2	131.4	118.2
$\text{Rb}(9\text{eG})_4^+$	98.6	127.8	112.6
$\text{Cs}(9\text{eG})_4^+$	88.1	118.8	101.9

a: Def2SVPD on metal cations and empirical dispersion

b: Def2TZVPP on metal cations and empirical dispersion

c: calculated using B3LYP/6-311(3df,3pd) and the calculations in column 3.

3.4. Conclusions

We have investigated gas phase structures and relative stabilities of the alkali metal cationized G-tetrads, $\text{M}(9\text{eG})_4^+$ ($\text{M} = \text{Li}, \text{Na}, \text{K}, \text{Rb}, \text{Cs}$). These intrinsic properties were studied by a combination of mass spectrometric techniques. IRMPD spectroscopy in both the fingerprint region ($900 - 1850 \text{ cm}^{-1}$) and the N-H/C-H stretching region ($2700 - 3800 \text{ cm}^{-1}$) showed that the gas phase structures for all five complexes are most consistent with the metal-centered G-tetrad structure and that they all give very similar IRMPD spectra. Calculations also predict very similar IR spectra in both regions studied.

Energy-resolved SORI-CID was used to compare the gas phase stabilities of the G-tetrads and showed that $\text{Na}(9\text{eG})_4^+$ is the most stable in agreement with solution phase studies. The stability of the other four complexes were found to decrease in the order $\text{Li}(9\text{eG})_4^+ > \text{K}(9\text{eG})_4^+ > \text{Rb}(9\text{eG})_4^+ > \text{Cs}(9\text{eG})_4^+$. This experimental energy ordering was reproduced by electronic structure calculations of the energies. Based upon the charge density of the ions it was expected that the Li^+ G-tetrad might be more stable than the Na^+ G-tetrad. Calculations were used to show that the lower stability to loss of 9-ethylguanine for the Li^+ complex could be due to a strong distortion and therefore destabilization of the neutral G-tetrad due to maximizing ion-dipole interactions while also maintaining hydrogen bonding interactions.

The present study does not explain why it is the potassium ion that is most commonly associated with G-quadruplexes in DNA, but we hope to answer that question in a forthcoming article.

References

- (1) Pinnavaia, T. J.; Miles, H. T.; Becker, E. D.: Self-assembled 5'-guanosine monophosphate, nuclear magnetic resonance evidence for a regular, ordered structure and slow chemical exchange. *J. Am. Chem. Soc.* **1975**, *97*, 7198-7200.
- (2) Fukushima, K.; Iwahashi, H.: 1:1 Complex of guanine quartet with alkali metal cations detected by electrospray ionization mass spectrometry. *Chem. Commun.* **2000**, 895-896.
- (3) Davis, J. T.; Spada, G. P.: Supramolecular architectures generated by self-assembly of guanosine derivatives. *Chem. Soc. Rev.* **2007**, *36*, 296-313.
- (4) Koch, K. J.; Aggerholm, T.; Nanita, S. C.; Graham Cooks, R.: Clustering of nucleobases with alkali metals studied by electrospray ionization tandem mass spectrometry: implications for mechanisms of multistrand DNA stabilization. *J. Mass Spectrom* **2002**, *37*, 676-686.
- (5) Meyer, M.; Brandl, M.; Sühnel, J.: Are guanine tetrads stabilized by bifurcated hydrogen bonds? *J Phys Chem A* **2001**, *105*, 8223-8225.
- (6) Campbell, N.; Neidle, S.: G-Quadruplexes and Metal Ions. In *Interplay between Metal Ions and Nucleic Acids*; Sigel, A., Sigel, H., Sigel, R. K. O., Eds.; Springer Netherlands, 2012; Vol. 10; pp 119-134.
- (7) Mezzache, S.; Alves, S.; Paumard, J.-P.; Pepe, C.; Tabet, J.-C.: Theoretical and gas-phase studies of specific cationized purine base quartet. *Rapid Commun. Mass Spectrom.* **2007**, *21*, 1075-1082.
- (8) Davis, J. T.: G-Quartets 40 years later: from 5'-GMP to molecular biology and supramolecular chemistry. *Angew. Chem. Int. Ed.* **2004**, *43*, 668-698.
- (9) Burge, S.; Parkinson, G. N.; Hazel, P.; Todd, A. K.; Neidle, S.: Quadruplex DNA: sequence, topology and structure. *Nucleic Acids Res.* **2006**, *34*, 5402-5415.
- (10) Métifiot, M.; Amrane, S.; Litvak, S.; Andreola, M.-L.: G-quadruplexes in viruses: function and potential therapeutic applications. *Nucleic Acids Res.* **2014**.

- (11) Wu, Y.; Brosh, R. M.: G-quadruplex nucleic acids and human disease. *FEBS J.* **2010**, *277*, 3470-3488.
- (12) Neidle, S.; Parkinson, G. N.: The structure of telomeric DNA. *Curr. Opin. Struct. Biol.* **2003**, *13*, 275-283.
- (13) Pandita, T. K.: Telomeres and Telomerase. In *Encyclopedia of Cancer (Second Edition)*; Editor-in-Chief: Joseph, R. B., Ed.; Academic Press: New York, 2002; pp 353-361.
- (14) Smith, F. W.; Feigon, J.: Quadruplex structure of *Oxytricha* telomeric DNA oligonucleotides. *Nature* **1992**, *356*, 164-8.
- (15) Balagurumoorthy, P.; Brahmachari, S. K.: Structure and stability of human telomeric sequence. *J. Biol. Chem.* **1994**, *269*, 21858-69.
- (16) Siddiqui-Jain, A.; Grand, C. L.; Bearss, D. J.; Hurley, L. H.: Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 11593-11598.
- (17) De Armond, R.; Wood, S.; Sun, D.; Hurley, L. H.; Ebbinghaus, S. W.: Evidence for the presence of a guanine quadruplex forming region within a polypurine tract of the hypoxia inducible factor 1 α promoter. *Biochemistry* **2005**, *44*, 16341-16350.
- (18) Simonsson, T.; Kubista, M.; Pecinka, P.: DNA tetraplex formation in the control region of c-myc. *Nucleic Acids Res.* **1998**, *26*, 1167-1172.
- (19) Huppert, J. L.; Bugaut, A.; Kumari, S.; Balasubramanian, S.: G-quadruplexes: the beginning and end of UTRs. *Nucleic Acids Res.* **2008**, *36*, 6260-6268.
- (20) Beaudoin, J.-D.; Perreault, J.-P.: 5'-UTR G-quadruplex structures acting as translational repressors. *Nucleic Acids Res.* **2010**, *38*, 7022-7036.
- (21) Haeusler, A. R.; Donnelly, C. J.; Periz, G.; Simko, E. A. J.; Shaw, P. G.; Kim, M.-S.; Maragakis, N. J.; Troncoso, J. C.; Pandey, A.; Sattler, R.; Rothstein, J. D.; Wang, J.: C9orf72 nucleotide repeat structures initiate molecular cascades of disease. *Nature* **2014**, *507*, 195-200.
- (22) Renton, A. E.; Majounie, E.; Waite, A.; Simón-Sánchez, J.; Rollinson, S.; Gibbs, J. R.; Schymick, J. C.; Laaksovirta, H.; van Swieten, J. C.; Myllykangas, L.;

Kalimo, H.; Paetau, A.; Abramzon, Y.; Remes, A. M.; Kaganovich, A.; Scholz, S. W.; Duckworth, J.; Ding, J.; Harmer, D. W.; Hernandez, D. G.; Johnson, J. O.; Mok, K.; Ryten, M.; Trabzuni, D.; Guerreiro, R. J.; Orrell, R. W.; Neal, J.; Murray, A.; Pearson, J.; Jansen, I. E.; Sondervan, D.; Seelaar, H.; Blake, D.; Young, K.; Halliwell, N.; Callister, J.; Toulson, G.; Richardson, A.; Gerhard, A.; Snowden, J.; Mann, D.; Neary, D.; Nalls, M. A.; Peuralinna, T.; Jansson, L.; Isoviita, V.-M.; Kaivorinne, A.-L.; Hölttä-Vuori, M.; Ikonen, E.; Sulkava, R.; Benatar, M.; Wu, J.; Chiò, A.; Restagno, G.; Borghero, G.; Sabatelli, M.; The, I. C.; Heckerman, D.; Rogaeva, E.; Zinman, L.; Rothstein, J.; Sendtner, M.; Drepper, C.; Eichler, E. E.; Alkan, C.; Abdullaev, Z.; Pack, S. D.; Dutra, A.; Pak, E.; Hardy, J.; Singleton, A.; Williams, N. M.; Heutink, P.; Pickering-Brown, S.; Morris, H. R.; Tienari, P. J.; Traynor, B. J.: A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. *Neuron* **2011**, *72*, 257-268.

(23) DeJesus-Hernandez, M.; Mackenzie, I. R.; Boeve, B. F.; Boxer, A. L.; Baker, M.; Rutherford, N. J.; Nicholson, A. M.; Finch, N. A.; Gilmer, H. F.; Adamson, J.; Kouri, N.; Wojtas, A.; Sengdy, P.; Hsiung, G.-Y. R.; Karydas, A.; Seeley, W. W.; Josephs, K. A.; Coppola, G.; Geschwind, D. H.; Wszolek, Z. K.; Feldman, H.; Knopman, D.; Petersen, R.; Miller, B. L.; Dickson, D.; Boylan, K.; Graff-Radford, N.; Rademakers, R.: Expanded GGGGCC hexanucleotide repeat in non-coding region of C9ORF72 causes chromosome 9p-linked frontotemporal dementia and amyotrophic lateral sclerosis. *Neuron* **2011**, *72*, 245-256.

(24) Sessler, J. L.; Sathiosatham, M.; Doerr, K.; Lynch, V.; Abboud, K. A.: A G-quartet formed in the absence of a templating metal cation: A new 8-(N,N-dimethylaniline)guanosine derivative. *Angew. Chem. Int. Ed.* **2000**, *39*, 1300-1303.

(25) Cerda, B. A.; Wesdemiotis, C.: Li⁺, Na⁺, and K⁺ Binding to the DNA and RNA nucleobases. bond energies and attachment sites from the dissociation of metal ion-bound heterodimers. *J. Am. Chem. Soc.* **1996**, *118*, 11884-11892.

- (26) Rodgers, M. T.; Armentrout, P. B.: Noncovalent interactions of nucleic acid bases (Uracil, Thymine, and Adenine) with alkali metal ions. Threshold collision-induced dissociation and theoretical studies. *J. Am. Chem. Soc.* **2000**, *122*, 8548-8558.
- (27) Moriwaki, H.: Complexes of cadmium ion with guanine bases detected by electrospray ionization mass spectrometry. *J. Mass Spectrom.* **2003**, *38*, 321-327.
- (28) Kwan, I. C. M.; She, Y.-M.; Wu, G.: Trivalent lanthanide metal ions promote formation of stacking G-quartets. *Chem. Commun.* **2007**, 4286-4288.
- (29) Baker, E. S.; Bernstein, S. L.; Bowers, M. T.: Structural characterization of G-quadruplexes in deoxyguanosine clusters using ion mobility mass spectrometry. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 989-997.
- (30) Fenn, J.; Mann, M.; Meng, C.; Wong, S.; Whitehouse, C.: Electrospray ionization for mass spectrometry of large biomolecules. *Science* **1989**, *246*, 64-71.
- (31) Fridgen, T. D.: Infrared consequence spectroscopy of gaseous protonated and metal ion cationized complexes. *Mass Spectrom. Rev.* **2009**, *28*, 586-607.
- (32) Polfer, N. C.; Oomens, J.: Vibrational spectroscopy of bare and solvated ionic complexes of biological relevance. *Mass Spectrom. Rev.* **2009**, *28*, 468-494.
- (33) Dunbar, R. C.: BIRD (blackbody infrared radiative dissociation): Evolution, principles, and applications. *Mass Spectrom. Rev.* **2004**, *23*, 127-158.
- (34) Gillis, E. A. L.; Demireva, M.; Nanda, K.; Beran, G.; Williams, E. R.; Fridgen, T. D.: Structures and energetics of electrosprayed uracilnCa²⁺ clusters (n = 14-4) in the gas phase. *Phys. Chem. Chem. Phys.* **2012**, *14*, 3304-3315.
- (35) Price, W. D.; Schnier, P. D.; Williams, E. R.: Tandem mass spectrometry of large biomolecule ions by blackbody infrared radiative dissociation. *Anal. Chem.* **1996**, *68*, 859-866.
- (36) Sleno, L.; Volmer, D. A.: Ion activation methods for tandem mass spectrometry. *J. Mass Spectrom.* **2004**, *39*, 1091-1112.
- (37) Laskin, J.; Futrell, J. H.: Collisional activation of peptide ions in FT-ICR mass spectrometry. *Mass Spectrom. Rev.* **2003**, *22*, 158-181.

- (38) Fraschetti, C.; Montagna, M.; Guarcini, L.; Guidoni, L.; Filippi, A.: Spectroscopic evidence for a gas-phase librating G-quartet- Na^+ complex. *Chem. Commun.* **2014**, 50, 14767-14770.
- (39) Peltz, C.; Drahos, L.; Vékey, K.: SORI excitation: Collisional and radiative processes. *J. Am. Soc. Mass Spectrom.* **2007**, 18, 2119-2126.
- (40) Frisch, M. J. F. e.: GAUSSIAN 09 (Revision A.01). *Gaussian, Inc., Wallingford, CT* **2009**.
- (41) Grimme, S.; Ehrlich, S.; Goerigk, L.: Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* **2011**, 32, 1456-1465.
- (42) Mao, C.; LaBean, T. H.; Reif, J. H.; Seeman, N. C.: Logical computation using algorithmic self-assembly of DNA triple-crossover molecules. *Nature* **2000**, 407, 493-496.
- (43) Rajabi, K.; Gillis, E. A. L.; Fridgen, T. D.: Structures of alkali metal ion–adenine complexes and hydrated complexes by IRMPD spectroscopy and electronic structure calculations. *J Phys Chem A* **2010**, 114, 3449-3456.
- (44) Gillis, E. A. L.; Fridgen, T. D.: The hydrated Li^+ –adenine–thymine complex by IRMPD spectroscopy in the N–H/O–H stretching region. *Int. J. Mass spectrom.* **2010**, 297, 2-8.
- (45) Wang, F.; Downton, M. T.; Kidwani, N.: Adenine tautomer electronic structural signatures studied using dual space analysis. *J. Chem. Theory Comput.* **2005**, 04, 247-264.

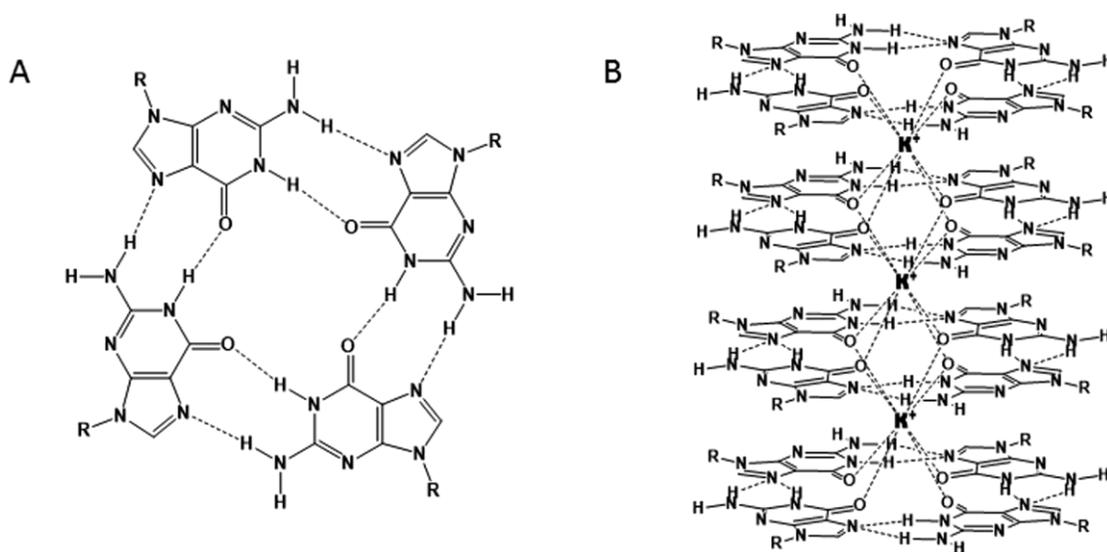
Chapter 4. The Intrinsic Stabilities and Structures of Metal Cationized Guanine Quadruplexes

This work has been published in part as *M. Azargun, Y. Jami-Alahmadi and T.D.*

Fridgen., The intrinsic stabilities and structures of alkali metal cationized guanine quadruplexes, *Phys. Chem. Chem. Phys.*, 2017, 19, 1281-1287. This work has been conducted by me and Y. Jami-Alahmadi and written by me and under supervision of Dr. Travis D. Fridgen.

4.1. Introduction

Guanine-rich sequences of DNA have been long known to form G-tetrads¹ which are square planar hydrogen bonded complexes (Scheme 4.1a). These clusters can be readily formed due to unique locations of hydrogen donating and accepting groups in the guanine nucleobase. The unique symmetric arrangement of the four guanine molecules causes the four carbonyl oxygens to form a cavity with a highly negative potential that can accommodate cations. G-quadruplexes are formed by two or more G-tetrads where G-tetrads stack on top of each other with metal cations such as K^+ between each (Scheme 4.1b).



Scheme 4.1

G-quadruplexes were first discovered to be at the end of each chromatid—the telomeric region—and function to preserve genetic information during replication and obstruct the fusion of individual chromatids. As part of the human aging process, the telomere region shortens as cells replicate until it is entirely removed from the DNA,

genetic information is lost, resulting in the death of cells.^{2,3} Fast replicating cells with a mechanism for preventing telomeres from getting too short, are implicated in many types of cancers and much research has been dedicated to understanding their structures, functions, and effects of chemical environment on them.^{4,5,6} More recently, research has revealed that G-quadruplexes can exist at other guanine-rich locations of DNA and RNA⁷ such as the promotor region,^{8,9} replication initiation sites,^{10,11} and 5' and 3'-untranslated (UTR) regions^{12,13,14} performing different functions.

De Armond and coworkers¹⁵ demonstrated that the promoter region of HIF-1 α gene, hypoxia inducible factor 1 alpha, may generate a potassium cation stabilized G-quadruplex structure that administers the gene expression process, and mutagenesis of this structure results in stoppage of gene expression. The transcriptional activity of the human RET (rearranged during transfection) proto-oncogene is found to be suppressed due to formation of a specific G-quadruplex complex.⁹ It was found that the promoter region of human RET gene contains a polypurine/polypyrimidine tract where a guanine-rich sequence exists. This guanine rich sequence may form specific G-quadruplexes resulting in the suppression of RET human gene transcription. Furthermore, G-quadruplexes have the potential to cure various diseases including cancers,^{16,6,17} tumors,^{18,19} and AIDS.^{20,21} For instance, the G-quadruplex structure in the promoter region of vegfr-2 gene can be selectively targeted by specific drugs to terminate the growth of the tumor by inhibition of tumor angiogenesis.¹⁸ Besides the biological importance of guanine complexes, they have received much attention for their potential

applications in materials science and in nanotechnology due to their affinity for self-assembly into supramolecular scaffolds.^{22,23}

The affinity of the guanine nucleobase to biologically relevant cations such as Na^+ and K^+ has been investigated and it follows $\text{Li}^+ > \text{Na}^+ > \text{K}^+ > \text{Rb}^+ > \text{Cs}^+$ order in the gas phase, in line with the magnitude of electrostatic interaction.^{24,25} In recent work,²⁶ we comprehensively assessed the structure and effect of alkali metal cations on the binding of gas phase G-tetrads using infrared multiple photon dissociation (IRMPD) spectroscopy and energy-resolved collision induced dissociation (CID) of $\text{M}(\text{9eG})_4^+$ complexes (9eG = 9-ethylguanine, M = alkali metal). It was found that $\text{Na}(\text{9eG})_4^+$ dissociates at higher energies compared to other alkali metal G-tetrads following this order $\text{Na}^+ \gg \text{Li}^+ > \text{K}^+ > \text{Rb}^+ > \text{Cs}^+$. These results contradict the gas phase binding energies that might be expected since Li^+ is a smaller, more densely-charged ion than Na^+ . This was attributed to the distortion energy brought into the system by Li^+ which strongly distorts the planar hydrogen bonded structure.^{26,27} Na^+ , on the other hand, perfectly fits in the cavity of the tetrad and preserves the planar structure.

The intrinsic chemical properties of G-quadruplexes have not been investigated due to the low solubility of guanine. An electrospray ionization (ESI) study of nucleobases clustering with alkali metal cations showed that the guanine only clusters with Na^+ and the tetramer cluster was the largest guanine cluster observed, while the other nucleobases were observed to form much larger clusters with up to 15mers.²⁸ Fukushima²⁹ and coworkers attributed the lack of larger guanine clusters to the absence of phosphate groups, resulting in the reduced stackability of individual tetrads. Using a

more soluble guanosine, Cooks and coworkers³⁰ obtained clusters larger than tetramers (although no quadruplexes) and demonstrated that the presence of sugar appears not to influence the clustering process. G-quadruplexes with cadmium(II)³¹ and terbium(III)³² have been observed as minor constituents in ESI-MS but no further analyses were performed.

A comprehensive computational work³³ in both aqueous and biological environments demonstrated that the sugar–phosphate backbone does not affect the affinity order of G-quadruplexes to alkali metals. It was also concluded that even the parallel and the antiparallel double layers bind to alkali metals with the same affinity. In addition, a solid-state experimental work showed that it is the affinity of monovalent cations for the cavity channel that determines the stability of G-quadruplexes while the binding of metals to the backbone part has an insignificant role in stability.³⁴ In the present work, guanine quadruplexes are simplified by eliminating the sugar–phosphate backbone. The ethyl group in 9-ethylguanine substitutes the sugar–phosphate backbone and blocks the 9-position from interaction with the metal cation

In the present work, for the very first time, the gas-phase stabilities of G-quadruplexes, $M(9\text{-eG})_8^+$, of alkali metal cations were studied by blackbody infrared radiative dissociation (BIRD, first given the acronym ZTRID³⁵ for zero-pressure thermal radiation induced dissociation). Vibrational spectroscopy was used to probe their structures and computational chemistry techniques were used to compute structures, infrared spectra, and theoretical binding energies.

4.2. Methods

4.2.1. Experimental

An ApexQe7 Bruker Fourier transform ion cyclotron resonance (FTICR) mass spectrometer (MS) was used to conduct all experiments in the Laboratory for the Study of the Energetics, Structures, and Reactions of Gaseous Ions at Memorial University. 10 mL of 0.1 mM 9-ethylguanine solution in 18 M Ω cm water, to which 2-3 drops of 1 mM salt solutions (NaCl, KCl, RbCl, and CsCl) were added, were introduced in the FTICR using an Apollo II ESI source at flow rates of 100-120 μ L h⁻¹. A heating jacket positioned around the flight tube, from the ICR cell to the gate valve, heats the ICR cell to the desired temperature.³⁶ To carry out BIRD experiments, a single G-quadruplex was isolated in the ICR cell ($P = 10^{-10}$ mbar), and unimolecular dissociation (precursor and product ion intensities) were monitored as a function of time. These experiments were repeated for a number of temperatures over the 322 – 405 K range. BIRD rate constants (k) were obtained by fitting the experimental precursor intensity *vs.* time plots experimental results to the first order integrated rate law:

$$[I]_t = e^{-kt} \quad 4.1$$

where $[I]_t$ is the normalized intensity of the cluster at reaction time, t . The BIRD rate constants were plotted against the inverse temperature for each G-quadruplex. The Arrhenius parameters, the observed pre-exponential factor (A_{obs}) and activation energy ($E_{a,obs}$), were extracted by fitting the experimental rate constants and temperature (T) data to the following equation:

$$\ln k = \ln A_{obs} - \frac{E_{a,obs}}{k_B T} \quad 4.2$$

IRMPD spectra were acquired using an infrared (IR) optical parametric oscillator (OPO), manufactured by LaserSpec, which is tuneable from 2700 – 4000 cm^{-1} , with a bandwidth of 2 cm^{-1} . The OPO is built around a periodically poled lithium niobate crystal and is pumped by a diode-pumped, solid state, Nd:YAG laser. The OPO operates at 20 kHz, with a pulse length of a few nanoseconds and generates output power near 3 W at 3 μm . The output power was limited to 1 W and was attenuated 10 times for the present experiments.

4.2.2. Computational

4.2.2.1. Geometries and Binding Energies

Geometry optimizations and frequency calculations were done using B3LYP/SVP computations. For comparison with the experimental spectra, the computed frequencies were scaled by 0.97. These structures were then optimized with B3LYP and the 6-31+G(d,p) basis set on C, H, N, and O and the Def2-SVPD basis set and effective core potentials (ECP's) on Na and K. The effect of dispersion was taken into account using GD3 empirical corrections.³⁷ All reported thermochemistries are 298 K values using the electronic energies with the 6-31+G(d,p) basis set and the SVP thermal corrections.

4.2.2.2. Master Equation Modeling

The observed Arrhenius values (E_a^{obs} and A^{obs}) require careful treatment and interpretation since they might not be necessarily equal to the limiting infinite-pressure

values (E_a^∞ and A^∞). These values can only be considered equal when the rate of energy exchange between ions and the blackbody radiation field (the body of ICR cell) is higher than the dissociation rate. At this rapid exchange limit, which is the case for large ions with many degrees of freedom or slowly dissociating ions, the ion population is thermally equilibrated with its surroundings, resulting in a Boltzmann distribution of internal energies of the ion population. At thermal equilibrium, one can readily extract threshold dissociation energies, E_0 , directly from the Arrhenius plot.^{35,38,39,40} Smaller, or more weakly bound ions, in contrast, dissociate at a rate that is comparable to or higher than the rate that ions absorb and emit infrared radiation in the ICR cell. The observed E_a^{obs} and A^{obs} for quickly dissociating ions may deviate significantly from the limiting rapid exchange E_a^∞ and A^∞ values. Consequently, these factors are not solely useful to interpret the transition state,^{38,41} and master equation modeling must be used to extract E_0 values from the BIRD Arrhenius plots.^{38,42,40}

A matrix form of the master equation was employed to obtain E_0 for $\text{Na}(9eG)_8^+$ and $\text{K}(9eG)_8^+$ from the BIRD Arrhenius plots; this process has been explained elsewhere^{41,42} but brief details are provided here. The master equation has two components: the transport (or J) matrix includes the rate constants for emission, absorption, and dissociation; and a second matrix that contains the initial population distribution, $[N_i(0)]$. Solution of the master equation results in the probability of energy transfer between energy states. To model the BIRD temperature dependent rate constant, the Boltzmann population at a desired temperature was allowed to redistribute until the internal energy distribution reaches a steady-state in which linear BIRD kinetics is

achieved. In the modeling process, a range of pre-exponential factors (tightness or looseness of the transition state), A , were examined to obtain reliable fitting. In this work, the IR intensities were scaled to achieve the desired pre-exponential factors. Calculations were done by increasing E_0 in the range of 0.95 – 1.75 eV in 0.05 eV increments. For each set of E_0 and $\ln(A)$ values, an Arrhenius plot was constructed which was compared to the experimental plot, and the best fit to the plot yielded the threshold activation energy for the G-quadruplex.

4.3. Results and Discussion

In a previous work, an ESI mass spectrum of a solution containing only KCl and 9-ethylguanine (reproduced in Figure S4.1) showed three potassiated clusters including the potassiated tetrad ($\text{K}(9\text{eG})_4^+$) at m/z 755; the potassiated quadruplex, $\text{K}(9\text{eG})_8^+$, at m/z 1471 and the topic of this work; and $\text{K}_2(9\text{eG})_{12}^{2+}$ at m/z 1113. Interestingly, complexes with multiples of four 9-ethylguanine form dominantly and other complexes such as dimers and trimers exist in almost unobservable amounts. The most abundant peak despite not adding Na^+ was m/z 739, the very stable $\text{Na}(9\text{-eG})_4^+$ discussed in a previous communication.²⁶ It is interesting that despite the intensity of the sodiated tetrads, no sodiated quadruplexes were observed.

4.3.1. Structures of G-Quadruplexes: IRMPD Spectroscopy

A comparison of the IRMPD spectra in the 2700 – 3800 cm^{-1} region for the $\text{M}(9\text{eG})_8^+$ G-quadruplexes is provided in Figure 4.1. All four quadruplexes display very similar features in the infrared spectra. The strongest broad band at about 3180 cm^{-1} is

attributed to the hydrogen bonded N-H stretches and except for some broadening, it seems that the metals do not shift this band significantly. It can be inferred that hydrogen bonds play an important role in G-quadruplexes, and from metal to metal their strength remains preserved. According to the structural analysis, $\text{Na}(9\text{-eG})_8^+$ and $\text{K}(9\text{-eG})_8^+$ possess eight almost identical hydrogen bonds with the same bond length while in the other two larger metal complexes, these hydrogen bonds vary in length to a small extent. We believe that this might be one of the reasons for the broadening. A free N-H stretching band is observed at 3525 cm^{-1} . The band just above 3300 cm^{-1} was attributed to the C=O stretching overtone^{26,27} and it seems reasonable to make the same assignment for these quadruplexes. The grey spectra in Figure 4.1 are the computed absorption spectra for each of the quadruplex species and they predict the main hydrogen bonding absorption and the free N-H stretching.

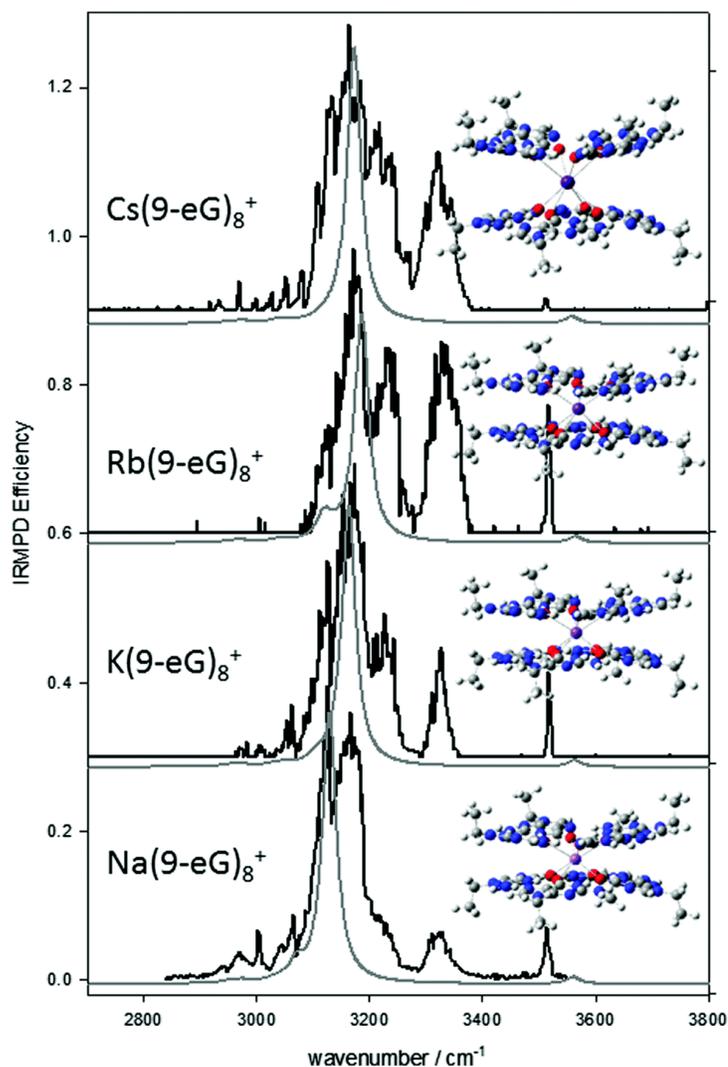


Figure 4.1. Experimental vibrational (IRMPD) spectra of the $M(9eG)_8^+$ quadruplexes (black trace) and computed infrared spectra (grey trace).

It is clear that these spectra show quite similar features, and, therefore, the G-quadruplexes have similar structures. In Figure S4.2, the IRMPD spectra of $K(9eG)_4^+$ and $K(9eG)_8^+$ are compared and it is clear that they too have the same features including some of the structures associated with the hydrogen bonded N–H stretching. The $M(9eG)_8^+$ spectra are in general more complex in the hydrogen bonded N–H stretching

region than $M(9eG)_4^+$, probably due the existence of more N–H bonds. It is also true that $M(9eG)_8^+$ is more difficult to acquire than $M(9eG)_4^+$ since the former absorbs more strongly the intense infrared laser and dissociates much more easily, resulting in some saturation of the IRMPD efficiency signal. The IRMPD spectra for $M(9eG)_4^+$ ($M = Li, Na, K, Rb, Cs$) G-tetrads^{26,27} were found to be virtually indistinguishable and agreed very well with the computed spectra for G-tetrad structures with Hoogsteen base pairing where the metal occupies the central negative hole bound to the carbonyl oxygen of all four 9eG's, as in Scheme 1a. Based on the similarity of the $M(9eG)_4^+$ and $M(9eG)_8^+$ spectra and the agreement of the experimental and computed spectra for $K(9eG)_8^+$ it is concluded that the complexes under investigation are sandwich-like quadruplex structures composed of two G-tetrads bound by the metal cation. The possibility of other conformer contributions were considered via computational chemistry. It is concluded that non-sandwich structures are at least 65 kJ mol⁻¹ higher in Gibbs energy and their computed infrared spectra do not compare well with the experimental IRMPD spectrum. Consequently, the possibility of non-sandwich structures is ruled out while the existence of sandwich conformers remains possible. Furthermore, the BIRD kinetics plots display quite linear first order kinetics, so the possibility of more than one structure with very different binding energies, which would result in non-linear kinetic plots is very unlikely.

4.3.2. BIRD Experiments and Activation Energies of G-quadruplexes

The G-quadruplexes undergo blackbody radiation induced dissociation in a timeframe such that temperature dependent rate constants can be obtained over a range of temperatures. There are two important differences noticed in these plots. Figure 4.2

shows BIRD breakdown plots for both $\text{Na}(9\text{eG})_8^+$ and $\text{K}(9\text{eG})_8^+$ at 341 K and 399 K, respectively. There are two important differences between the decomposition of the sodiated and potassiated quadruplexes highlighted in this plot. First, $\text{Na}(9\text{eG})_8^+$ loses 9eG sequentially to form $\text{Na}(9\text{eG})_7^+$, then $\text{Na}(9\text{eG})_6^+$ *etc.* $\text{K}(9\text{eG})_8^+$, on the other hand, seems to lose a neutral guanine tetrad, $(9\text{eG})_4$, to form $\text{K}(9\text{eG})_4^+$. $\text{Rb}(9\text{eG})_8^+$ and $\text{Cs}(9\text{eG})_8^+$ also show loss of neutral tetrad (see breakdown curves in the supplementary information). The dissociation products are of importance since they may provide some insights into the transition states for decomposition (*vide infra*). The second difference is that the decomposition of both $\text{Na}(9\text{eG})_8^+$ and $\text{K}(9\text{eG})_8^+$ shown in Figure 4.2 occurs over the same time frame—but a significantly higher temperature (some 50 K) is required for $\text{K}(9\text{eG})_8^+$ than for $\text{Na}(9\text{eG})_8^+$, hinting at a higher stability for the former.

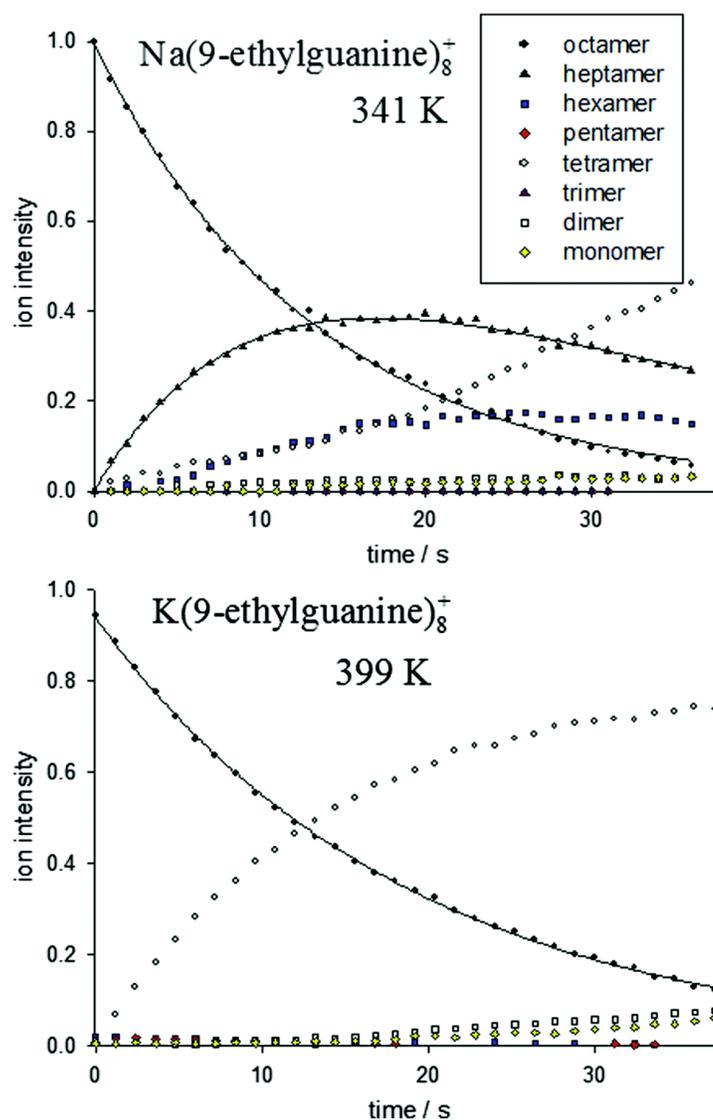


Figure 4.2. Kinetics plots for Na(9eG)₈⁺ and K(9eG)₈⁺ at 341 K and 399 K, respectively.

Arrhenius plots for each of the quadruplexes are depicted in Figure 4.3, and activation energies and experimental entropies of activation are shown in Table 4.1. The observed activation energy for dissociation in the gas phase of Na(9eG)₈⁺ has the lowest value among quadruplexes, about 100 kJ mol⁻¹ lower than that observed for K(9eG)₈⁺. This finding is analogous to the results from the solid-state NMR work by Wong³⁴ where the relative affinity of alkali metals to G-quadruplexes formed by guanosine 5'-

monophosphate followed the order $\text{K}^+ > \text{NH}_4^+ > \text{Rb}^+ > \text{Na}^+ > \text{Cs}^+ > \text{Li}^+$. The activation energies follow the trend expected for electrostatic ion–dipole interactions except for the low value observed for $\text{Na}(\text{9eG})_8^+$. The temperature dependent kinetic plots, from which the rate constants are determined, for each of the four $\text{M}(\text{9eG})_8^+$ are provided in the supplementary information. The BIRD rate constants for $\text{Na}(\text{9eG})_8^+$ are between two and three orders of magnitude larger than those for the other quadruplexes. Due to these large rate constants, $\text{Na}(\text{9eG})_8^+$ displays small-to-medium size kinetics (Fig. S4.3)⁴¹ and is in the slow energy exchange regime, and master equation modeling is likely required to extract E_0 values; for comparison master equation modeling was performed on both $\text{Na}(\text{9eG})_8^+$ and $\text{K}(\text{9eG})_8^+$.

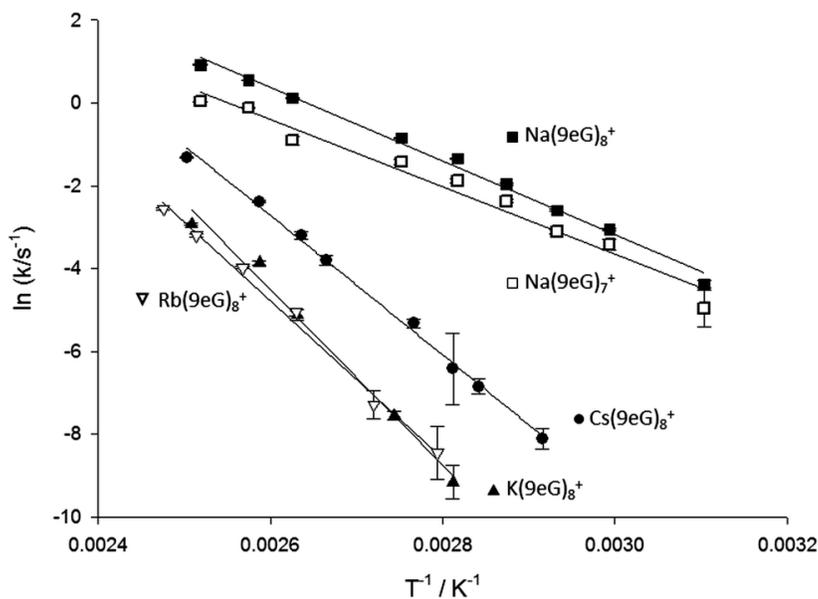


Figure 4.3. Arrhenius plots for the $\text{M}(\text{9eG})_8^+$ quadruplexes and the $\text{Na}(\text{9eG})_7^+$ complex.

Table 4.1. Summary of the observed Arrhenius parameters for the dissociation of $M(9eG)_8^+$ G-quadruplexes, and those obtained from master equation modeling of the temperature dependent rate constants for $Na(9eG)_8^+$ and $K(9eG)_8^+$

	$E_a^{\text{obs}}/\text{kJ mol}^{-1}$	$\log A_{\text{obs}}$	$\Delta S^\ddagger/\text{J K}^{-1} \text{ mol}^{-1}$	$E_0/\text{kJ mol}^{-1}$
$Na(9eG)_8^+$	74 ± 3	10.4	-58 ± 4	100 ± 7
$K(9eG)_8^+$	176 ± 11	21.9	166 ± 11	170 ± 7
$Rb(9eG)_8^+$	158 ± 7	19.4	119 ± 6	
$Cs(9eG)_8^+$	140 ± 4	17.5	88 ± 5	

As seen in Figure 4.2, $K(9eG)_8^+$ loses a neutral tetrad a process that requires little rearrangement. One can envisage a loose (late) transition state without a classical barrier and this is reflected in the experimental $\log A^{\text{obs}}$ and entropy of activation (Table 4.1). On the contrary some rearrangement is likely necessary for $Na(9eG)_8^+$ to form $Na(9eG)_7^+$ by losing neutral 9eG, resulting in a tight (early) transition state; again this is reflected in the experimental $\log A^{\text{obs}}$ and entropy consistent with other biomolecule dissociations⁴¹ with $\log A^{\text{obs}}$ of 12.6 or less ($\Delta S^\ddagger \leq -22 \text{ J K}^{-1} \text{ mol}^{-1}$). The observed activation energy for these molecules appears to be less than the E_0 , threshold activation energy. Like $K(9eG)_8^+$, cesium and rubidium quadruplexes fell apart by losing a neutral tetramer which was predictable since these two G-quadruplexes have analogous G-tetrad structures very similar to $K(9eG)_4^+$. Transition state entropies and $\log A^{\text{obs}}$ for $Rb(9eG)_8^+$ and $Cs(9eG)_8^+$ are relatively similar to those obtained for $K(9eG)_8^+$, indicating that these two metal clusters dissociate through a loose transition state similar to $K(9eG)_8^+$.

To fit the experimental rate constants for $Na(9eG)_8^+$, E_0 was increased in 0.05 eV increments in the range of 0.95 to 1.15 eV, and the fitting process was repeated for frequency factors, $\log A^\infty$, of 14, 15, 16, and 17. The best fit to the experimental BIRD

rate constants was obtained by using $\log A^\infty$ of 15, 16, or 17 and E_0 values of 1.00, 1.05, 1.10 eV, respectively. The threshold dissociation energy of about $100 \pm 7 \text{ kJ mol}^{-1}$ for Na(9eG)_8^+ is larger by about 26 kJ mol^{-1} from the observed activation energy ($E_a^{\text{obs}} = 74 \pm 3 \text{ kJ mol}^{-1}$). This difference reveals that the dissociation rate of Na(9eG)_8^+ is higher than the rate of energy exchange. For K(9eG)_8^+ , an E_0 range of 1.55 to 1.85 eV was examined for variety of $\log A^\infty$ (19 to 23). The best fit to the experimental rates was acquired with the E_0 of 1.65, 1.70, and 1.79 eV for $\log A^\infty$ values of 21, 22 and 23. The threshold dissociation energy for K(9eG)_8^+ then is about $170 \pm 7 \text{ kJ mol}^{-1}$ which lays within the standard deviation for the observed activation energy ($176 \pm 11 \text{ kJ mol}^{-1}$) obtained from the BIRD temperature dependent rate constants. The modeling shows that the potassium quadruplex dissociates at a rate close to or less than the energy exchange rate since the observed activation energy closely matches the limiting rapid exchange activation energy. The observed activation energies for the Rb and Cs quadruplexes, like K, are expected to reflect the true E_0 values. The stability for these four G-quadruplexes follows this order $\text{K}^+ > \text{Rb}^+ > \text{Cs}^+ \gg \text{Na}^+$ and thermochemistry explains why K^+ is the ion of choice for G-quadruplexes in the DNA of biological systems^{43,44,45,46} despite the greater abundance of Na^+ .

In a previous paper,²⁶ it was demonstrated that the intrinsic stability of the G-tetrads follow the order $\text{Na}^+ > \text{Li}^+ \gg \text{K}^+ > \text{Rb}^+ > \text{Cs}^+$ and it was the ability of Na^+ to sit in the plane of the tetrad and the stronger ion-dipole interaction between Na^+ and the tetrad accounted for this stability. The larger alkali cations, K^+ , Rb^+ , and Cs^+ , sit above the plane in the tetrad structure, favoring their ability to form a sandwich type, G-

quadruplex structures. In order to form G-quadruplexes, Na^+ in $\text{Na}(\text{9eG})_4^+$ needs to be pulled from the stable position in the center of the tetrad. It is interesting to note that in the solution phase, K^+ was also found to form the most thermodynamically stable quadruplex, although it was solvation and deformation effects that were determined to be responsible for the difference in these relative stabilities.³³ In the present work—in the gas-phase—only intrinsic properties are responsible for the relative stabilities of the quadruplexes and solvation has no effect on the stability ordering of quadruplexes. There are also very strong differences between the decomposition processes studied in the present work compared to Zaccaria et al.³³ In the latter, the theoretical thermochemistry for the loss of metal cation from the quadruplex was determined, while in the present work the quadruplexes lose either a G-tetrad (K^+ , Rb^+ , and Cs^+) or a monomer of 9-ethylguanine (Na^+).

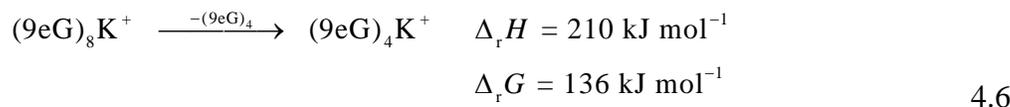
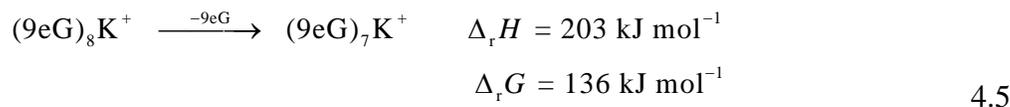
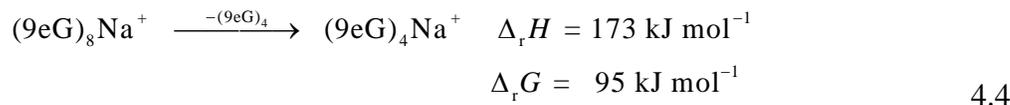
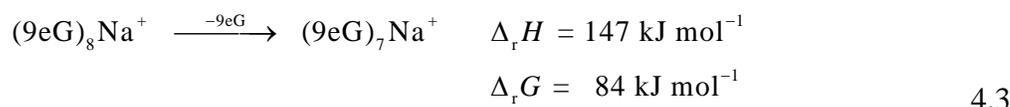
In an equilibrium titration study,⁴⁷ the process of formation of metal cation stabilized guanine quadruplexes has been suggested to be through intermediates consisting of G-tetrads stabilized by metal cations followed by further strand rearrangements to form the quadruplex. It was also determined that the formation of Na^+ stabilized quadruplexes was more complex than the formation of K^+ stabilized quadruplexes, in agreement with the gas-phase decomposition processes discovered here, where the Na^+ stabilized quadruplex loses one guanine at a time and the quadruplexes stabilized by the larger cations lose a G-tetrad all at once.

Na^+ was the smallest alkali metal which formed a G-quadruplex; attempts to produce $\text{Li}(\text{9eG})_8^+$ were unsuccessful. This is consistent with the $\text{Li}(\text{9eG})_4^+$ tetrad being

strongly distorted from planarity, much closer to a tetrahedral coordination of the 9eG about Li⁺.²⁶ That Li⁺ is not exposed in the tetrad, means it is unable to form quadruplexes. This is in agreement with the ability of lithium to denature and prohibit the formation of G-quadruplexes in nucleic acids and with the calculations of Zaccaria³³ in the aqueous phase, which showed that Li⁺ avoids forming sandwich clusters.

4.3.3. Theoretical Dissociation Enthalpies and Gibbs Energies

The computed 298 K dissociation enthalpies for two possible dissociation paths, loss of neutral tetramer and loss of monomer, are summarized below:



The computed thermochemistries for dissociation of Na(9eG)₈⁺ are consistent with the observation of the loss of monomer since the loss of tetramer is significantly higher in energy for this quadruplex. For K(9eG)₈⁺, the computations show the loss of monomer or the loss of tetrad to be virtually isoenergetic. The calculations overstabilize the complexes by between 40 and 50 kJ mol⁻¹ which may be reflective of the minimal basis set used and the omission of basis set superposition error. However, the calculations

do agree with the experimental conclusions that the binding of the potassiated quadruplex is significantly higher than the sodiated quadruplex. A summary of the dissociation thresholds for all complexes as well as the relative computed values for the Na^+ and K^+ complexes is provided in Figure 4.4.

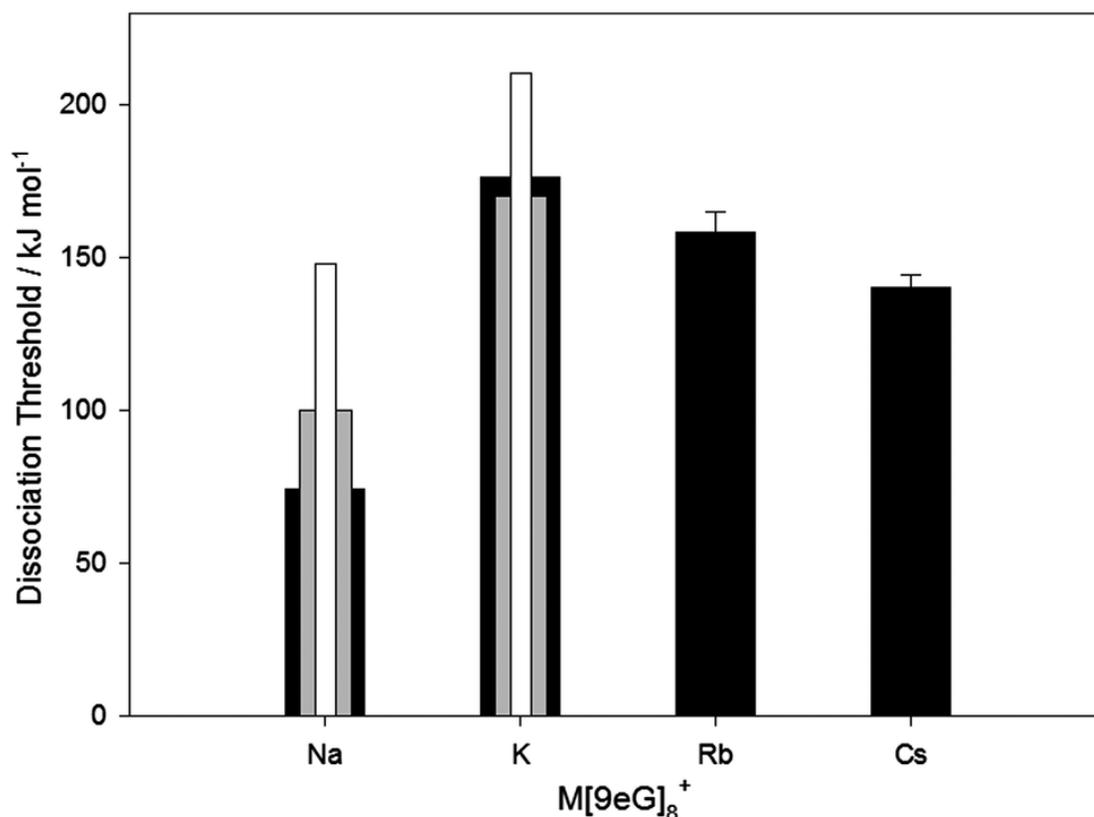


Figure 4.4. Summary of Arrhenius activation energies (black bars), master equation model dissociation thresholds fitted to the experimental Arrhenius plots (grey bars), and computed dissociation enthalpies (white bars), all in kJ mol^{-1} .

4.4. Conclusions

The intrinsic stabilities for self-assembled guanine quadruplexes was determined to be in the order $\text{K}(9\text{eG})_8^+ > \text{Rb}(9\text{eG})_8^+ > \text{Cs}(9\text{eG})_8^+ \gg \text{Na}(9\text{eG})_8^+$, which is consistent with the observation of K^+ being the ion of choice in guanine quadruplexes in nucleic

acids in biological systems. In the gas-phase, the sodiated quadruplex was found to lose one 9eG, whereas the quadruplexes of the heavier cations lose a neutral guanine tetrad. Vibrational spectroscopy on the gas-phase quadruplex ions was consistent with the structures in which the metal cations were sandwiched between two guanine tetrads for all complexes. The stability order for the quadruplexes is quite different than that for the metal cationized tetrads: $\text{Na}(9\text{eG})_4^+ > \text{Li}(9\text{eG})_4^+ \gg \text{K}(9\text{eG})_4^+ > \text{Rb}(9\text{eG})_4^+ > \text{Cs}(9\text{eG})_4^+$. The difference, mainly in the position of the sodiated complex, is attributed to the ability of Na^+ to sit in the center of a planar hydrogen-bonded tetrad complex while the heavier cations sit above the plane of the tetrad, readily able to bind to a second tetrad. To form a quadruplex, Na^+ must be pulled from its stable position in the center of the tetrad. It was not possible to form a lithiated quadruplex, which is consistent with the tetrad structure, where Li^+ is not exposed and with Li^+ being used as a means to degrade guanine quadruplexes. The calculations presented here are consistent with the observations of the potassiated quadruplex being thermochemically more stable than the sodiated quadruplex. Similarly, they are consistent with the observation of the loss of monomeric guanine from the sodiated quadruplex and a neutral tetrad from the potassiated quadruplex as the lowest energy dissociation pathways.

References

- (1) Gellert, M.; Lipsett, M. N.; Davies, D. R.: Helix formation by guanylic acid. *Proc. Natl. Acad. Sci. U.S.A.* **1962**, *48*, 2013-2018.
- (2) Shammass, M. A.: Telomeres, lifestyle, cancer, and aging. *Curr. Opin. Clin. Nutr. Metab. Care* **2011**, *14*, 28-34.
- (3) Aubert, G.; Lansdorp, P. M.: Telomeres and aging. *Physiol Rev* **2008**, *88*, 557-579.
- (4) Burge, S.; Parkinson, G. N.; Hazel, P.; Todd, A. K.; Neidle, S.: Quadruplex DNA: sequence, topology and structure. *Nucleic Acids Res.* **2006**, *34*, 5402-5415.
- (5) Huppert, J. L.: Four-stranded nucleic acids: structure, function and targeting of G-quadruplexes. *Chem. Soc. Rev.* **2008**, *37*, 1375-1384.
- (6) Murat, P.; Singh, Y.; Defrancq, E.: Methods for investigating G-quadruplex DNA/ligand interactions. *Chem. Soc. Rev.* **2011**, *40*, 5293-5307.
- (7) Huppert, J. L.: Hunting G-quadruplexes. *Biochimie* **2008**, *90*, 1140-1148.
- (8) You, H.; Wu, J.; Shao, F.; Yan, J.: Stability and kinetics of c-MYC promoter G-quadruplexes studied by single-molecule manipulation. *J. Am. Chem. Soc.* **2015**, *137*, 2424-2427.
- (9) Shin, Y. J.; Kumarasamy, V.; Camacho, D.; Sun, D.: Involvement of G-quadruplex structures in regulation of human RET gene expression by small molecules in human medullary thyroid carcinoma TT cells. *Oncogene* **2015**, *34*, 1292-1299.
- (10) Paeschke, K.; Capra, John A.; Zakian, Virginia A.: DNA replication through G-quadruplex motifs is promoted by the *S. cerevisiae* Pif1 DNA helicase. *Cell*, *145*, 678-691.

- (11) Zheng, K.W.; Wu, R.Y.; He, Y.D.; Xiao, S.; Zhang, J.Y.; Liu, J.Q.; Hao, Y.H.; Tan, Z.: A competitive formation of DNA:RNA hybrid G-quadruplex is responsible to the mitochondrial transcription termination at the DNA replication priming site. *Nucl. Acids Res.* **2014**, *42*, 10832-10844.
- (12) Kumari, S.; Bugaut, A.; Huppert, J. L.; Balasubramanian, S.: An RNA G-quadruplex in the 5[prime] UTR of the NRAS proto-oncogene modulates translation. *Nat Chem Biol* **2007**, *3*, 218-221.
- (13) Morris, M. J.; Basu, S.: An unusually stable G-quadruplex within the 5'-UTR of the MT3 matrix metalloproteinase mRNA represses translation in eukaryotic cells. *Biochemistry* **2009**, *48*, 5313-5319.
- (14) Agarwala, P.; Pandey, S.; Mapa, K.; Maiti, S.: The G-quadruplex augments translation in the 5' untranslated region of transforming growth factor β 2. *Biochemistry* **2013**, *52*, 1528-1538.
- (15) De Armond, R.; Wood, S.; Sun, D.; Hurley, L. H.; Ebbinghaus, S. W.: Evidence for the presence of a guanine quadruplex forming region within a polypurine tract of the hypoxia inducible factor 1 α promoter. *Biochemistry* **2005**, *44*, 16341-16350.
- (16) Drygin, D.; Siddiqui-Jain, A.; O'Brien, S.; Schwaebe, M.; Lin, A.; Bliesath, J.; Ho, C. B.; Proffitt, C.; Trent, K.; Whitten, J. P.; Lim, J. K. C.; Von Hoff, D.; Anderes, K.; Rice, W. G.: Anticancer activity of CX-3543: A direct inhibitor of rRNA biogenesis. *Cancer Res.* **2009**, *69*, 7653-7661.
- (17) Wang, J.-M.; Huang, F.-C.; Kuo, M. H.-J.; Wang, Z.-F.; Tseng, T.-Y.; Chang, L.-C.; Yen, S.-J.; Chang, T.-C.; Lin, J.-J.: Inhibition of cancer cell migration and invasion through suppressing the Wnt1-mediating signal pathway by G-quadruplex structure stabilizers. *J. Biol. Chem.* **2014**, *289*, 14612-14623.

- (18) Salvati, E.; Zizza, P.; Rizzo, A.; Iachettini, S.; Cingolani, C.; D'Angelo, C.; Porru, M.; Randazzo, A.; Pagano, B.; Novellino, E.; Pisanu, M. E.; Stoppacciaro, A.; Spinella, F.; Bagnato, A.; Gilson, E.; Leonetti, C.; Biroccio, A.: Evidence for G-quadruplex in the promoter of vegfr-2 and its targeting to inhibit tumor angiogenesis. *Nucleic Acids Res.* **2014**, *42*, 2945-2957.
- (19) Gowan, S. M.; Harrison, J. R.; Patterson, L.; Valenti, M.; Read, M. A.; Neidle, S.; Kelland, L. R.: A G-quadruplex-interactive potent small-molecule inhibitor of telomerase exhibiting in vitro and in vivo antitumor activity. *Mol Pharmacol* **2002**, *61*, 1154-1162.
- (20) Perrone, R.; Butovskaya, E.; Daelemans, D.; Palù, G.; Pannecouque, C.; Richter, S. N.: Anti-HIV-1 activity of the G-quadruplex ligand BRACO-19. *J. Antimicrob. Chemother.* **2014**, *69*, 3248-3258.
- (21) Métifiot, M.; Amrane, S.; Litvak, S.; Andreola, M.-L.: G-quadruplexes in viruses: function and potential therapeutic applications. *Nucl. Acids Res.* **2014**.
- (22) Davis, J. T.: G-Quartets 40 Years Later: From 5'-GMP to molecular biology and supramolecular chemistry. *Angew. Chem. Int. Ed.* **2004**, *43*, 668-698.
- (23) Davis, J. T.; Spada, G. P.: Supramolecular architectures generated by self-assembly of guanosine derivatives. *Chem. Soc. Rev.* **2007**, *36*, 296-313.
- (24) Cerda, B. A.; Wesdemiotis, C.: Li⁺, Na⁺, and K⁺ binding to the DNA and RNA nucleobases. Bond energies and attachment sites from the dissociation of metal ion-bound heterodimers. *J. Am. Chem. Soc.* **1996**, *118*, 11884-11892.
- (25) Rodgers, M. T.; Armentrout, P. B.: Noncovalent interactions of nucleic acid bases (uracil, thymine, and adenine) with alkali metal ions. Threshold collision-induced dissociation and theoretical studies. *J. Am. Chem. Soc.* **2000**, *122*, 8548-8558.

(26) Azargun, M.; Fridgen, T. D.: Guanine tetrads: an IRMPD spectroscopy, energy resolved SORI-CID, and computational study of M(9-ethylguanine)₄⁺ (M = Li, Na, K, Rb, Cs) in the gas phase. *Phys. Chem. Chem. Phys.* **2015**, *17*, 25778-25785.

(27) Fraschetti, C.; Montagna, M.; Guarcini, L.; Guidoni, L.; Filippi, A.: Spectroscopic evidence for a gas-phase librating G-quartet-Na⁺ complex. *Chem. Commun.* **2014**, *50*, 14767-14770.

(28) Koch, K. J.; Aggerholm, T.; Nanita, S. C.; Graham Cooks, R.: Clustering of nucleobases with alkali metals studied by electrospray ionization tandem mass spectrometry: implications for mechanisms of multistrand DNA stabilization. *J. Mass Spectrom.* **2002**, *37*, 676-686.

(29) Fukushima, K.; Iwahashi, H.: 1[Na^+] Complex of guanine quartet with alkali metal cations detected by electrospray ionization mass spectrometry. *Chem. Commun.* **2000**, 895-896.

(30) Aggerholm, T.; Nanita, S. C.; Koch, K. J.; Cooks, R. G.: Clustering of nucleosides in the presence of alkali metals: Biologically relevant quartets of guanosine, deoxyguanosine and uridine observed by ESI-MS/MS. *J. Mass Spectrom.* **2003**, *38*, 87-97.

(31) Moriwaki, H.: Complexes of cadmium ion with guanine bases detected by electrospray ionization mass spectrometry. *J. Mass Spectrom.* **2003**, *38*, 321-327.

(32) Kwan, I. C. M.; She, Y.-M.; Wu, G.: Trivalent lanthanide metal ions promote formation of stacking G-quartets. *Chem. Commun.* **2007**, 4286-4288.

(33) Zaccaria, F.; Paragi, G.; Fonseca Guerra, C.: The role of alkali metal cations in the stabilization of guanine quadruplexes: why K⁺ is the best. *Phys. Chem. Chem. Phys.* **2016**, *18*, 20895-20904.

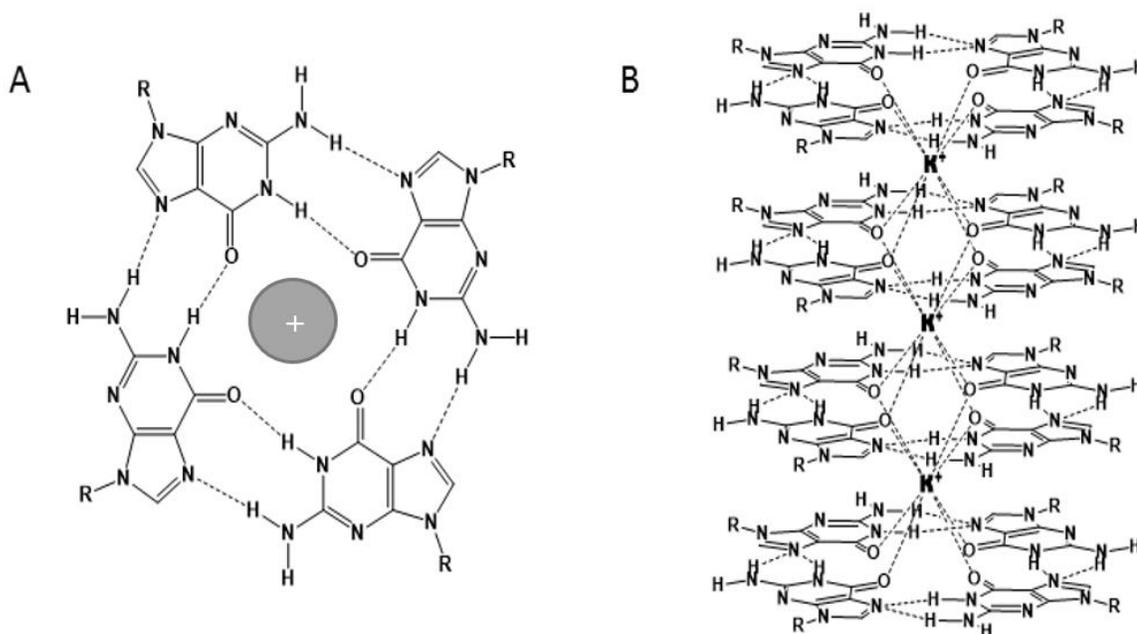
- (34) Wong, A.; Wu, G.: Selective binding of monovalent cations to the stacking G-quartet structure formed by guanosine 5'-monophosphate: A solid-state NMR study. *J. Am. Chem. Soc.* **2003**, *125*, 13895-13905.
- (35) Dunbar, R. C.; McMahon, T. B.; Thoelmann, D.; Tonner, D. S.; Salahub, D. R.; Wei, D. Comparison of theory and experiment for (H₂O)₂Cl⁻ and (H₂O)₃Cl. *J. Am. Chem. Soc.* **1995**, *117*, 12819-12825.
- (36) Gillis, E. A. L.; Demireva, M.; Nanda, K.; Beran, G.; Williams, E. R.; Fridgen, T. D.: Structures and energetics of electrosprayed uracil Ca²⁺ clusters (n = 14-4) in the gas phase. *Phys. Chem. Chem. Phys.* **2012**, *14*, 3304-3315.
- (37) Grimme, S.; Ehrlich, S.; Goerigk, L.: Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* **2011**, *32*, 1456-1465.
- (38) Dunbar, R. C.; McMahon, T. B.; Thoelmann, D.; Tonner.: Zero-pressure thermal-radiation-induced dissociation of gas-phase cluster ions. *J. Am. Chem. Soc.* **1995**, *117*, 12819-12825.
- (39) Price, W. D.; Schnier, P. D.; Williams, E. R.: Binding energies of the proton-bound amino acid dimers Gly·Gly, Ala·Ala, Gly·Ala, and Lys·Lys measured by blackbody infrared radiative dissociation. *J Phys Chem B.* **1997**, *101*, 664-673.
- (40) Dunbar, R. C.: Kinetics of thermal unimolecular dissociation by ambient infrared radiation. *J. Phys. Chem.* **1994**, *98*, 8705-8712.
- (41) Dunbar, R. C.: BIRD (blackbody infrared radiative dissociation): Evolution, principles, and applications. *Mass Spectrom. Rev.* **2004**, *23*, 127-158.
- (42) Price, W. D.; Williams, E. R.: Activation of peptide ions by blackbody radiation: factors that lead to dissociation kinetics in the rapid energy exchange limit. *J Phys Chem A* **1997**, *101*, 8844-8852.

- (43) Guschlbauer, W.; Chantot, J.-F.; Thiele, D.: Four-stranded nucleic acid structures 25 years later: from guanosine gels to telomer DNA. *J. Biomol. Struct. Dyn.* **1990**, *8*, 491-511.
- (44) Parkinson, G. N.; Lee, M. P. H.; Neidle, S.: Crystal structure of parallel quadruplexes from human telomeric DNA. *Nature* **2002**, *417*, 876-880.
- (45) Sun, D.; Guo, K.; Rusche, J. J.; Hurley, L. H.: Facilitation of a structural transition in the polypurine/polypyrimidine tract within the proximal promoter region of the human VEGF gene by the presence of potassium and G-quadruplex-interactive agents. *Nucleic Acids Res.* **2005**, *33*, 6070-6080.
- (46) Sen, D.; Gilbert, W.: A sodium-potassium switch in the formation of four-stranded G4-DNA. *Nature* **1990**, *344*, 410-414.
- (47) Gray, R. D.; Chaires, J. B.: Kinetics and mechanism of K⁺- and Na⁺-induced folding of models of human telomeric DNA into G-quadruplex structures. *Nucleic Acids Res.* **2008**, *36*, 4191-4203.

**Chapter 5. Intrinsic properties of $\text{K}_2(9\text{-ethylguanine})_{12}^{2+}$
by BIRD, IRMPD Spectroscopy, and Energy Resolved
SORI-CID**

5.1. Introduction

G-tetrads (G_4) and G-quadruplexes (G_{4S}) (Scheme 5.1A and 5.1B) are secondary structures of DNA that can be adopted in guanine-rich sequences of DNA and RNA.¹⁻⁴ Formation of thermodynamically stable G-tetrads is kinetically fast and assisted by the unique locations of hydrogen bond donating and accepting groups in the guanine nucleobase. This unique configuration eases the formation of eight hydrogen bonds. Metal cations are shown to stabilize and facilitate the formation of G-tetrads by partially neutralizing the excess negative potential from the four carbonyl oxygens of guanine molecules.^{5,6} Four-stranded G-quadruplexes can readily form by self-stacking of planar G-tetrads with a metal cation between each layer (see Scheme 1 B).⁷⁻¹¹



Scheme 5.1

G-quadruplexes were first observed in the telomere region—the ends of chromosomes which cap and preserve DNA from degradation and reaction with other

DNA molecules.^{12,13} DNA loses some of its telomeric region during mitotic cycles. This shortening of the telomere region has been understood to play an important role in the aging process¹⁴⁻¹⁶; once the telomeric length reaches its critical length, the cell dies. Telomeric G-quadruplexes not only contribute to the aging process, but they have also been shown to have potential for anti-cancer strategies.¹⁷⁻¹⁹ Telomeric G4s have been shown to induce cellular senescence in cancer cells by inhibiting telomerase activity, responsible for elongation of the telomeric region.²⁰ As a result, stabilization of telomeric G-quadruplexes by small ligands has presented promising results toward treatment of a variety of cancers.²¹⁻²³

The biological importance of G-quadruplexes has gained recognition by their observation in various regions of DNA other than the telomeric region.²⁴⁻²⁹ G4 sequences have been traced in the first intron, first exon, promoter, and 5'-untranslated regions (5'-UTRs). Work conducted by Hurley and coworkers³⁰ has proposed the presence of a G-quadruplex at the promoter region of c-Myc gene, an essential gene regulator activating the expression of many genes. It was shown that this G-quadruplex can function as a transcriptional suppressor in order to control the transcription of c-Myc gene. The inhibitory role of G4s has been incorporated in methods to induce formation of G4s in the promoter region of oncogenes, genes with potential to cause cancer.³¹ G-quadruplexes have found numerous therapeutic applications mainly as targets to land drugs and manipulate gene activity of diseased and cancerous cells. Work performed by Leonetti *et al.*³² selectively targeted DNA G-quadruplexes by an organic ligand, EMICRON; C₅₂H₅₉N₆O₄·4HCl. The resulting complex was shown to inflict significant damage to

telomeres in cancer cells, and cell proliferation in tumour cells was observed to be greatly suppressed. Due to their affinity for self-assembly into supramolecular scaffolds, G4s have also received much attention in material science, nanotechnology, and sensor designing.^{3,33-40}

The effect of biologically relevant ions including Na^+ and K^+ on the formation and stability of guanine tetrads has also been investigated in the gas phase.⁴¹ Infrared multiphoton dissociation (IRMPD) spectroscopy and energy-resolved collision induced dissociation (CID) studies of $\text{M}(9\text{eG})_4^+$ complexes (9eG = 9-ethylguanine, M = alkali metal) have revealed that Na^+ forms the strongest G-tetrads followed by Li^+ , K^+ , Rb^+ and Cs^+ . This conclusion was in agreement with solution phase experiments conducted by Fukushima and coworkers.⁴² It has been also shown spectroscopically and computationally that $\text{Na}(9\text{eG})_4^+$ adopts a square planar structure^{41,43} with Na^+ at the center while other larger metals such as K^+ , Rb^+ and Cs^+ sit above the plane formed by four guanine molecules.⁴¹ The observed stability trend, $\text{Na}^+ \gg \text{Li}^+ > \text{K}^+ > \text{Rb}^+ > \text{Cs}^+$, is not consistent with what is expected based on purely electrostatic grounds which predicts Li^+ to have the strongest interaction with guanine molecules.^{44,45} Calculations showed that Li^+ strongly distorts the planar hydrogen bonded structure while Na^+ preserves the planar structure.⁴¹

The chemistry of G-quadruplex clusters composed of two or more G-tetrads has been the subject of both solution and gas phase studies. Proton NMR chemical shift studies⁴⁶ examined the competition between K^+ and Na^+ for coordination to G-tetrads. In

this regard, the oligonucleotide d(G₃T₄G₃) was exposed to metal chloride (NaCl and KCl) solutions. It was reported that two Na⁺ cations in the G-quadruplex composed of three G-tetrads are displaced by two K⁺, a thermodynamically favourable process, ($\Delta_r G^\circ$) = -7.11 ± 0.15 kJ mol⁻¹. It was also suggested that the replacement of Na⁺ is not the result of an optimal fit of K⁺ in the G-tetrad cavities, but it is due to the higher hydration energy of Na⁺. This conclusion has been also supported by a computational work⁴⁷ proposing that although the cage formed by the G4-G4 structure acts as a better host for K⁺, still the process is governed by relative hydration energies. However, it is important to point out in gas phase studies—in the absence of solvent—K⁺ quadruplexes composed of two 9-ethylguanine tetrads were determined to be significantly more thermodynamically stable than those containing Na⁺.⁴⁸

The process of formation of metal cation stabilized guanine quadruplexes was the subject of an equilibrium titration study.⁴⁹ It was proposed that G4s form through intermediates composed of G-tetrads stabilized by metal cations with further strand rearrangements to yield the quadruplex. It was also suggested that the formation of G4s stabilized by potassium cation is less complex than processes resulting in Na⁺ stabilized G-quadruplexes. This observation is consistent with our previous gas phase studies⁴⁸ showing that decomposition of the Na⁺ stabilized quadruplex occurred through a one guanine at a time mechanism while K⁺ stabilized quadruplexes lose a neutral G-tetrad in a single step.

Although the sugar-phosphate backbone undoubtedly plays an important role in defining different G-quadruplex topologies, both computational and experimental works have shown that the affinity order of G-quadruplexes to alkali metals to be independent of the backbone moieties⁵⁰⁻⁵³. A recent computational work by Zaccaria⁵² in both the aqueous and biological environments has suggested that the affinity order of alkali metals to the G4s is not influenced by the sugar-phosphate backbone. The effect of different configurations of backbone moiety on metal binding was also examined, demonstrating that the parallel and the antiparallel double layers do not influence the affinity order. A solid-state study⁵³ on G-quadruplexes has also resulted in the same conclusion. It was determined that the affinity of monovalent cations to the G4-G4 cavity dominates the stability order of the resulting cluster. The binding of metals to the backbone units was shown to be insignificant in both G4s formation and complex stability.

In the present work, gas phase $K_2(9\text{-ethylguanine})_{12}^{2+}$ was studied for the first time using ion-activation techniques including blackbody infrared radiative dissociation (BIRD), sustained off-resonance irradiation collision induced dissociation (SORI-CID). Infrared multiphoton dissociation (IRMPD) spectroscopy along with computational chemistry was used to shed light on the cluster structure.

5.2. Methods

5.2.1. Experimental

All experiments were conducted using a Bruker ApexQe 7.0 hybrid Fourier transform ion cyclotron resonance (FTICR) mass spectrometer in the Laboratory for the

Study of the Energetics, Structures, and Reactions of Gaseous Ions at Memorial University.^{54,55} 9-Ethylguanine (9eG) was purchased from Sigma-Aldrich and used without further purification. The solution was prepared by the addition of two drops of 1 mM KCl solution to 10 mL of 0.2 mM 9eG solution in 18 MΩ cm water. An Apollo II ion source coupled to the FTICR mass spectrometer was used to electrospray the solution at $\approx 100 \mu\text{L h}^{-1}$. SORI-CID experiments were carried out by isolating the G-quadruplex cluster inside the ICR cell ($P = 10^{-10}$ mbar) exciting, the ion by standard off-resonance irradiation techniques, while exposing the cluster ion to Ar inside the ICR cell at pressures of about 10^{-5} - 10^{-6} mbar. This pressure ensures 10's to 100's collisions in a 250 ms excitation window. The average kinetic energies are expected to be 2/3 the maximum lab frame kinetic energies which are determined from Equation 5.1⁵⁶

$$E_{\text{lab}}^{\text{max}} = \frac{\beta^2 q^2 V_{p-p}}{32\pi^2 m d^2 \Delta\nu^2} \quad 5.1$$

where β is a constant factor known as geometrical factor of the ICR cell (0.9 in the present case), q is the charge on the ion, V_{p-p} is the peak to peak excitation voltage, m is the mass of the ion, d is the diameter of the ICR cell (6 cm), and $\Delta\nu$ is the frequency offset (500 Hz). The center of mass ion kinetic energies can be obtained by a multiplication of the lab frame energies by $m_{\text{Ar}} / (m_{\text{Ar}} + m_{\text{quadruplex}})$.

An infrared (IR) optical parametric oscillator (OPO), manufactured by LaserSpec and is tuneable from $2700 - 4000 \text{ cm}^{-1}$ with a bandwidth of 2 cm^{-1} , was used to obtain IRMPD spectra. The OPO is built around a periodically poled lithium niobate crystal and

is pumped by a diode-pumped, solid state, Nd:YAG laser. The OPO operates at 20 kHz, with a pulse duration of a few nanoseconds and generates output power near 3 W at 3 μm . The output power from the laser was limited to 1 W for the present experiments and a broad-band filter decreasing the maximum laser intensity to ~ 300 mW was also used. Following isolation of $\text{K}_2(9\text{eG})_{12}^{2+}$, the laser was scanned at 2 cm^{-1} intervals irradiated for 1 s to produce the IRMPD spectrum. The IRMPD efficiency is the negative logarithm of the ratio of the precursor ion intensity divided by the total ion intensity.

For BIRD experiments, $\text{K}_2(9\text{eG})_{12}^{2+}$ was isolated in the ICR cell with a background pressure of 10^{-10} mbar, and the precursor and product ion intensities were recorded after a varying period of reaction time at a specific temperature. BIRD kinetics were observed between 373 – 405K. The BIRD rate constant (k) for each temperature was determined by a fitting the experimental quadruplex intensity vs time plots to the integrated first order rate law, Equation 5.2

$$[I]_t = e^{-kt} \quad 5.2$$

in which $[I]_t$ is the normalized intensity of the quadruplex at reaction time, t . A plot of the BIRD rate constants as a function of the inverse temperature results in the Arrhenius parameters, pre-exponential factor (A^{obs}) and activation energy (E_a^{obs}). These parameters were obtained by fitting the experimental rate constants and temperature (T) data to the Equation 5.3.

$$\ln k = \ln A_{obs} - \frac{E_{a,obs}}{k_B T} \quad 5.3$$

5.3. Results and Discussion

In this present work, 9eG was used instead of guanine due to better solubility resulting in improved intensity of the quadruplex mass spectrum. Owing to the fact that N9 is blocked by an ethyl group, the possibility of interaction between K^+ and 9eG through N9 is also eliminated. This N9 blockage makes the 9-ethylguanine a better model for biological systems since N9 is the site of attachment to the phosphate backbone of DNA and RNA. The sugar-phosphate backbone has been shown to play an insignificant role in the affinity of metal cations to the quadruplexes and also their stabilities.⁵² Here we have introduced a simplified G-quadruplex system, in the absence of the sugar-phosphate backbone, composed of three G-tetrads to probe the intrinsic physical chemistries of the gas-phase quadruplex.

Figure 5.1 depicts an electrospray mass spectrum of 9-ethylguanine/KCl aqueous solution. $K(9eG)_4^+$, $K(9eG)_8^+$, and $K_2(9eG)_{12}^{2+}$ show up, each being made of multiples of four 9-ethylguanines, the G-tetrad building block. The structure and stability of $K(9eG)_4^+$ and $K(9eG)_8^+$ to loss of 9eG were discussed in previous communications and it was shown that among the alkali metals, $K(9eG)_8^+$ is the most stable G-quadruplex while the planar $Na(9eG)_4^+$ was the strongest G-tetrad.^{41,48} Interestingly, $K_2(9eG)_{12}^{2+}$ was the largest G-quadruplex generated which is consistent with most G-quadruplexes found in the biological systems which are formed by three G4s and 2 potassium cations.⁵⁷⁻⁵⁹

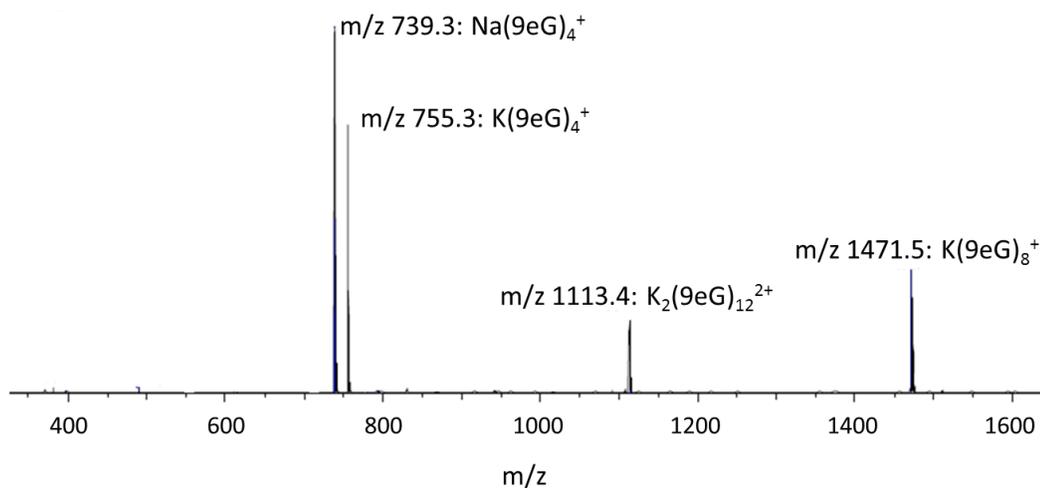


Figure 5.1. Electrospray mass spectrum of an aqueous solution containing 10 mL of 0.1 mM solution of 9-ethylguanine with two drops of 0.1 mM KCl. Note that $\text{K}_2(9\text{eG})_{12}^{2+}$ is the largest potassiumated G-quadruplex detected.

5.3.1. Unimolecular Dissociation

Figure 5.2 shows the unimolecular dissociation of $\text{K}_2(9\text{eG})_{12}^{2+}$ conducted by using SORI-CID, IRMPD, and BIRD activation techniques. It is apparent that, in all cases, $\text{K}_2(9\text{eG})_{12}^{2+}$ dissociates as the main dissociation channel into two ionic products, $\text{K}(9\text{eG})_4^+$ and $\text{K}(9\text{eG})_8^+$. SORI-CID and IRMPD mass spectra look similar as some of other fragments such as the pentamer and trimer show up. These two activation techniques are known as harder techniques by which more energetic dissociation channels are also susceptible to activation. However, BIRD as the softer technique exclusively activated the lowest energy dissociation pathway which is the loss of a G-tetrad, $\text{K}(9\text{eG})_4^+$.

The intensity of $\text{K}(9\text{eG})_8^+$ in all three mass spectra is much less than the $\text{K}(9\text{eG})_4^+$ while both should show similar intensities. This inconsistency stems from

the fact that the $\text{K}(\text{9eG})_8^+$ cluster quickly dissociates upon formation by losing a $(\text{9eG})_4$ cluster and forms more $\text{K}(\text{9eG})_4^+$.

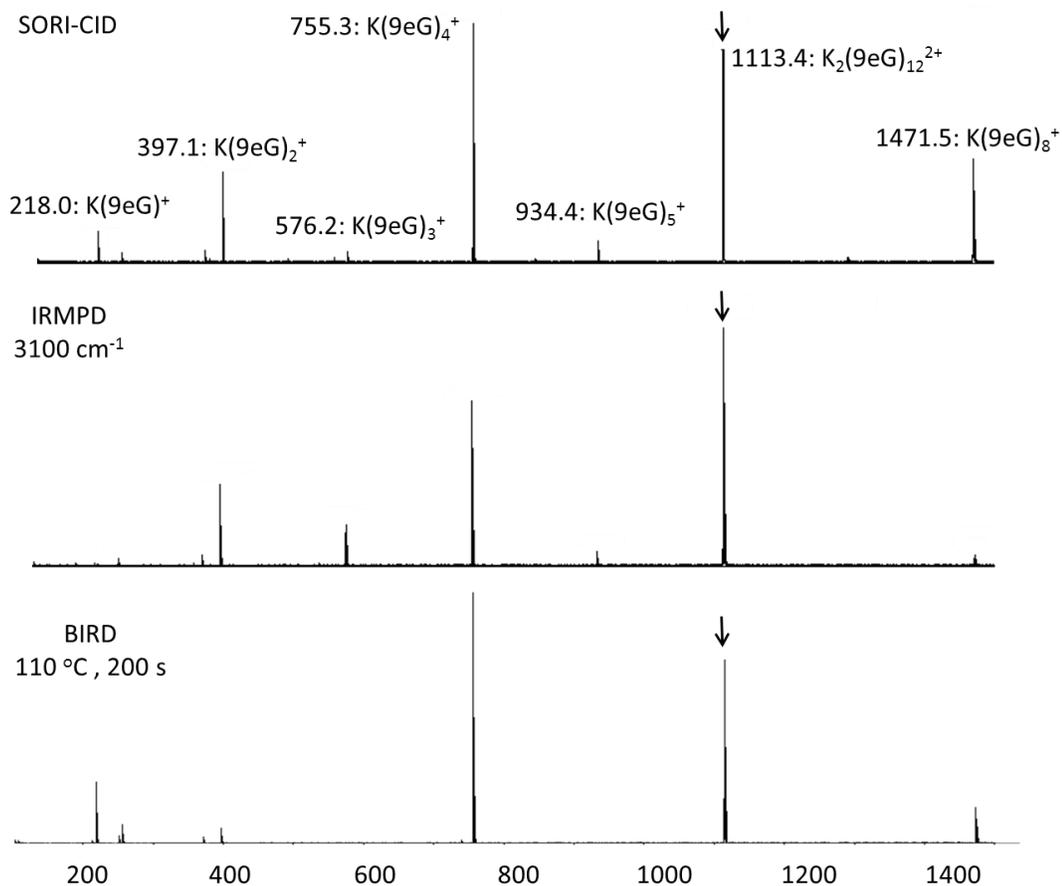


Figure 5.2. Dissociation products resulted from unimolecular dissociation of $\text{K}_2(\text{9eG})_{12}^{2+}$ by SORI-CID, IRMPD and BIRD activation techniques.

5.3.2. IRMPD Spectroscopy $\text{K}_2(\text{9eG})_{12}^{2+}$

Figure 5.3 features a comparison of the IRMPD spectrum of $\text{K}_2(\text{9eG})_{12}^{2+}$ as well as the IRMPD spectra of $\text{K}(\text{9eG})_8^+$ and $\text{K}(\text{9eG})_4^+$ clusters in the $2700 - 3800 \text{ cm}^{-1}$ region. It is apparent that all clusters show very similar spectral features containing three main

bands centered at about 3160, and 3500 cm^{-1} attributed to hydrogen bonded N-H stretching and free N-H stretching, respectively.^{41,43,48} The feature observed at about 3320 cm^{-1} has been assigned to the C=O stretching overtone.^{41,43}

In our previous work, it was shown by IRMPD spectroscopy and computational chemistry that $M(9eG)_4^+$ clusters ($M = \text{Li, Na, K, Rb and Cs}$) form similar structures with Hoogsteen base pairing in which metals occupy the cavity (Scheme 5.1). It was observed that $M(9eG)_4^+$ clusters were indistinguishable solely based on IRMPD spectra. In addition, IRMPD spectroscopy on $M(9eG)_8^+$ ($M = \text{Na, K, Rb, and Cs}$) clusters resulted in the same conclusion. The $M(9eG)_8^+$ experimental IRMPD spectra displayed very similar features which were interpreted to conclude that these clusters exist in a sandwich structure with a metal cation between two G-tetrads. Computational studies in conjunction with the vibrational spectra were used to rule out the co-existence of non-sandwich structures. Figure 5.3 shows that the larger G-quadruplex composed of three G-tetrads, $K_2(9eG)_{12}^{2+}$, presents a very similar spectrum compared to $K(9eG)_8^+$ and $K(9eG)_4^+$. It is concluded that $K_2(9eG)_{12}^{2+}$ also adopts a sandwich structure in which three G-tetrads stack with two potassium cations between layers.

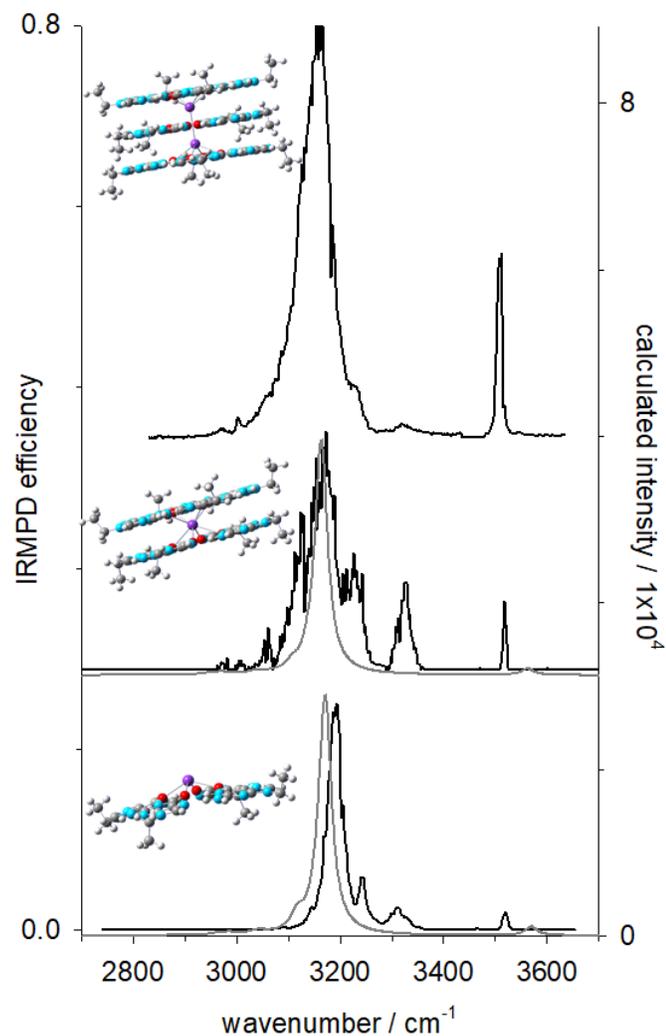


Figure 5.3. Experimental vibrational (IRMPD) spectra of the $K_2(9eG)_{12}^{2+}$, $K(9eG)_8^+$ and $K(9eG)_4^+$ (black trace). The corresponding computed infrared spectra (grey trace) are also depicted.

5.3.3. Relative Gas-Phase Stabilities of Potassium Cationized Quadruplexes by Energy-Resolved SORI-CID

Energy-resolved SORI-CID activation was used to probe the relative gas phase stabilities of potassium the quadruplexes, $K_2(9eG)_{12}^{2+}$ and $K(9eG)_8^+$. Clusters were isolated and then excited to collision energies of between 0 and 0.5 eV and then exposed

to Ar gas with a reservoir pressure of 10 mbar Ar. The energy resolved intensity profiles for $K_2(9eG)_{12}^{2+}$ are depicted in Figure 5.4.

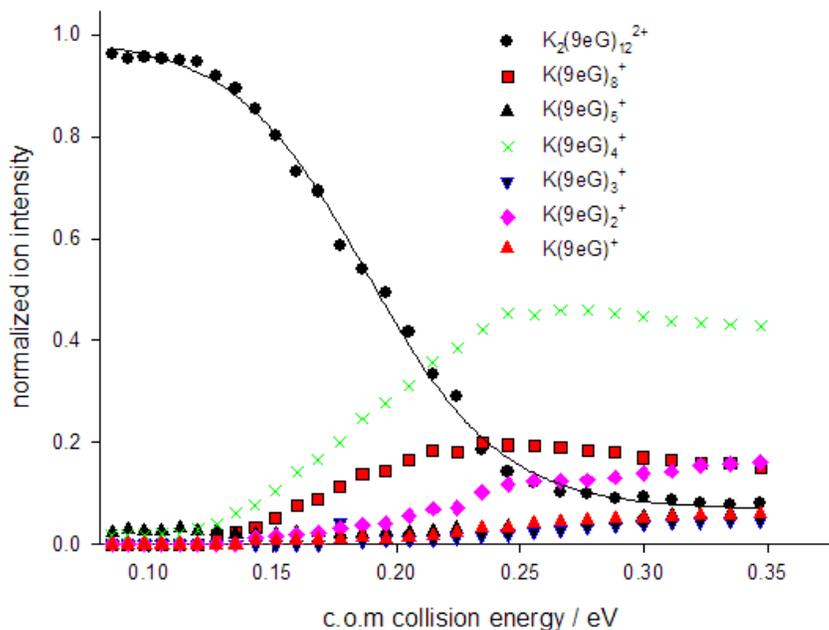


Figure 5.4. Energy resolved intensity profiles for SORI-CID of $K_2(9eG)_{12}^{2+}$.

It is apparent that $K_2(9eG)_{12}^{2+}$ quadruplex undergoes charge separation into two singly charged ions, $K(9eG)_4^+$ and $K(9eG)_8^+$. Furthermore, $K(9eG)_8^+$ from $K_2(9eG)_{12}^{2+}$ loses a neutral tetramer, as was previously observed,⁴⁸ to also form $K(9eG)_4^+$. Only minute amounts of other ionic fragments other than $K(9eG)_4^+$ and $K(9eG)_8^+$ were observed showing that these clusters primarily dissociate by losing a G-tetrad unit. This dissociation pattern was reproduced by IRMPD activation, discussed above, as well as BIRD activation that will be discussed in next section. Figure 5.5 depicts plots of the normalized intensities of precursor $K(9eG)_8^+$ and $K_2(9eG)_{12}^{2+}$ quadruplexes against the

center of mass collision energy from two different experiments beginning with isolated $K(9eG)_8^+$ and $K_{12}(9eG)_{12}^{2+}$. It is clear that $K(9eG)_8^+$ dissociates at lower energies than that required for $K_{12}(9eG)_{12}^{2+}$; a line is drawn to mark the 50% dissociation of clusters. $K_2(9eG)_{12}^{2+}$ clearly shows more stability to collision induced dissociation compared to the smaller quadruplex, $K(9eG)_8^+$. This greater stability of $K_2(9eG)_{12}^{2+}$ versus $K(9eG)_8^+$ is due to the ratio of cation to guanine tetrad, $2/3$ versus $1/2$, respectively; this results in stronger bonding between K^+ and the tetrad.

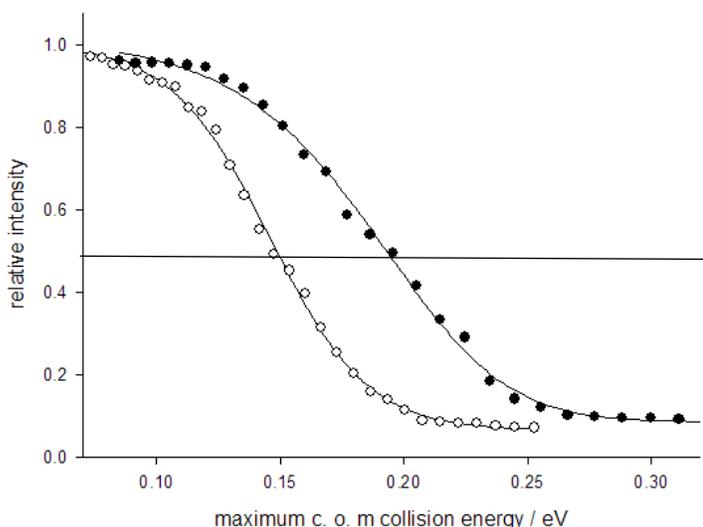


Figure 5.5. The quadruplex intensity vs. collision energy plot resulted from the energy-resolved SORI-CID experiments at reservoir pressure of 10 mbar with Ar.

5.3.4. BIRD Experiments

The temperature dependent rate constants of G-quadruplexes can be determined by conducting BIRD experiments over a range of temperatures.⁴⁸ $K_2(9eG)_{12}^{2+}$ undergoes BIRD and loses $K(9eG)_4^+$ leaving a $K(9eG)_8^+$ complex behind. $K(9eG)_8^+$ also quickly loses a neutral tetramer and produces a $K(9eG)_4^+$ cluster (see Figure 5.6). Figure 5.7

compares the BIRD Arrhenius plots for $K_2(9eG)_{12}^{2+}$ from this work and that of $K(9eG)_8^+$ from previous work.⁴⁸ The corresponding observed activation energies (E_0) and $\log A^{\text{obs}}$ are summarized in Table 5.1. Interestingly, the activation energy extracted from BIRD kinetics revealed that $K_2(9eG)_{12}^{2+}$ dissociate about 50 kJ mol^{-1} higher than the value observed for $K(9eG)_8^+$. This agrees with the results of the SORI-CID experiments discussed in section 5.3.3 which showed that higher energy was required to dissociate $K_2(9eG)_{12}^{2+}$. This observation is in line with most G-quadruplexes detected in the biological systems in which three G-tetrads along with two potassium ions are involved.⁵⁷⁻⁵⁹ It is not completely clear why G-quadruplexes composed of three G-tetrads show better stability compared to smaller G-quadruplexes. However, it can be related to the higher number of K cations that improves the electrostatic interactions and pi-pi stacking.⁶⁰

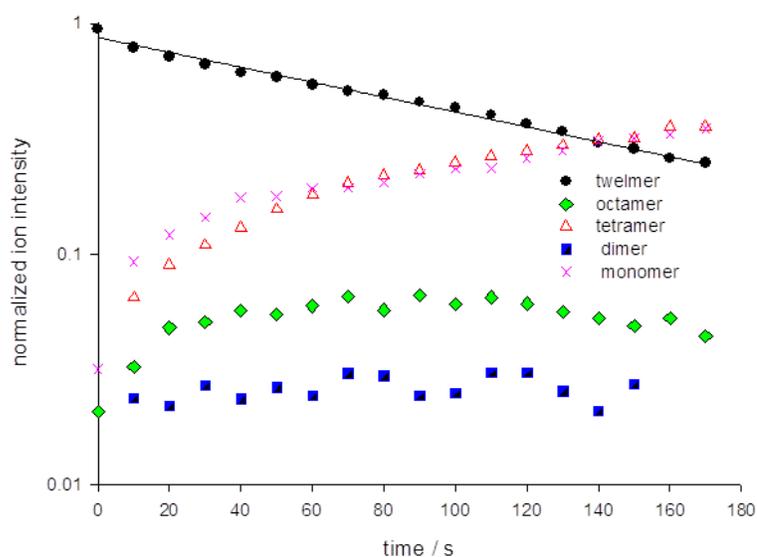


Figure 5.6. BIRD kinetics plot for $K_2(9eG)_{12}^{2+}$ at 385 K. Only octamer, tetramer, monomer are the fragment products.

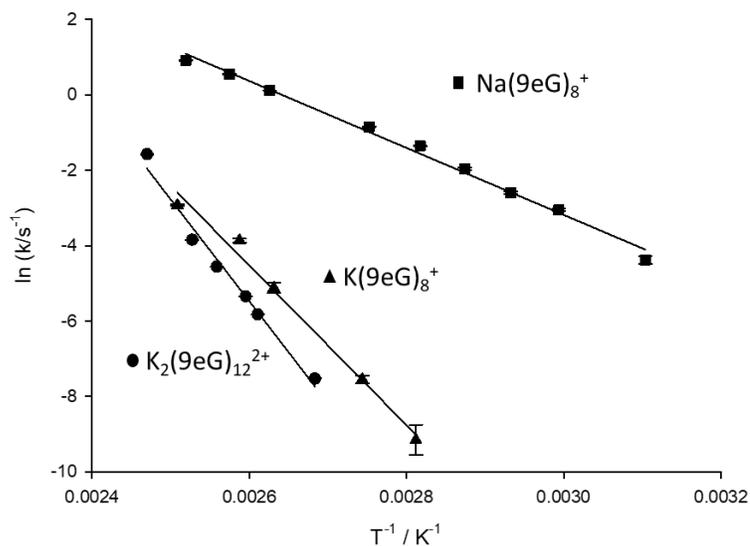


Figure 5.7. Arrhenius plots for the $\text{K}_2(9\text{eG})_{12}^{2+}$, $\text{K}(9\text{eG})_8^+$, and $\text{Na}(9\text{eG})_8^+$ quadruplexes. It was previously shown that among G-quadruples composed of two G-tetrads, $\text{K}(9\text{eG})_8^+$ was the most stable cluster while $\text{Na}(9\text{eG})_8^+$ was the weakest.⁴⁸

Table 5.1. Observed Arrhenius parameters for the dissociation of $\text{K}_2(9\text{eG})_{12}^{2+}$ and the $\text{M}(9\text{eG})_8^+$ G-quadruplexes⁴⁸

	$E_a^{\text{obs}} / \text{kJ mol}^{-1}$	$\log A^{\text{obs}}$	$\Delta S^\ddagger / \text{J K}^{-1} \text{mol}^{-1}$
$\text{K}_2(9\text{eG})_{12}^{2+}$	225 ± 15	28.2	286 ± 20
$\text{Na}(9\text{eG})_8^+$	74 ± 3	10.4	-58 ± 4
$\text{K}(9\text{eG})_8^+$	176 ± 11	21.9	166 ± 11
$\text{Rb}(9\text{eG})_8^+$	158 ± 7	19.4	119 ± 6
$\text{Cs}(9\text{eG})_8^+$	140 ± 4	17.5	88 ± 5

It can be inferred that potassium G-quadruplexes dissociate through a loose (late) transition state with lack of a major rearrangement (no classical barrier) since the fragments look similar when separated as when in the quadruplex. This observation is opposite to the $\text{Na}(\text{9eG})_8^+$ cluster that loses sequential 9-ethylguanine and the tetrad must rearrange during this dissociation. The transition state looseness is also reflected in the $\log A^{\text{obs}}$ and entropy of activations summarized in Table 5.1.

5.4. Conclusion

A combination of mass spectrometry and computational chemistry has been used in the present work to assess the intrinsic properties of the potassiumated 9-ethylguanine quadruplex, $\text{K}_2(\text{9eG})_{12}^{2+}$. IRMPD spectroscopy in the N-H/C-H stretching region ($2700 - 3800 \text{ cm}^{-1}$) revealed that this G-quadruplex configures in a sandwich-type structure with two G-tetrads sandwiching each of the two K^+ , very similar to the structure determined for the $\text{M}(\text{9eG})_8^+$ complexes. The stability of $\text{K}_2(\text{9eG})_{12}^{2+}$ toward dissociation and its binding energy was examined by means of energy-resolved SORI-CID and BIRD kinetics experiments. SORI-CID experiments showed that $\text{K}_2(\text{9eG})_{12}^{2+}$ is a more stable cluster toward dissociation to $\text{K}(\text{9eG})_8^+$ and $\text{K}(\text{9eG})_4^+$ compared to the stability of the $\text{K}(\text{9eG})_8^+$ cluster which loses $(\text{9eG})_4$. Temperature dependent BIRD kinetics for $\text{K}_2(\text{9eG})_{12}^{2+}$ were consistent with energy-resolved SORI-CID results yielding an activation energy valued at $225 \pm 15 \text{ kJ mol}^{-1}$ which was greater than that determined for $\text{K}(\text{9eG})_8^+$.

References

- (1) Burge, S.; Parkinson, G. N.; Hazel, P.; Todd, A. K.; Neidle, S.: Quadruplex DNA: sequence, topology and structure. *Nucleic Acids Res.* **2006**, *34*, 5402-5415.
- (2) Phan, A. T.; Kuryavyi, V.; Patel, D. J.: DNA architecture: from G to Z. *Curr. Opin. Struc. Biol.* **2006**, *16*, 288-298.
- (3) Davis, J. T.: G-quartets 40 years later: from 5'-GMP to molecular biology and supramolecular chemistry. *Angew. Chem. Int. Ed.* **2004**, *43*, 668-698.
- (4) Simonsson, T.: G-quadruplex DNA structures variations on a theme. *Biol. Chem.* **2001**, *382*, 621-628.
- (5) Campbell, N. H.; Neidle, S.: G-quadruplexes and metal ions. In *interplay between metal ions and nucleic acids*; Springer, 2012; pp 119-134.
- (6) Georgiades, S. N.; Abd Karim, N. H.; Suntharalingam, K.; Vilar, R.: Interaction of metal complexes with G-quadruplex DNA. *Angew. Chem. Int. Ed.* **2010**, *49*, 4020-4034.
- (7) Patel, D. J.; Phan, A. T.; Kuryavyi, V.: Human telomere, oncogenic promoter and 5'-UTR G-quadruplexes: diverse higher order DNA and RNA targets for cancer therapeutics. *Nucleic Acids Res.* **2007**, *35*, 7429-7455.
- (8) Neidle, S.; Balasubramanian, S.: Fundamentals of quadruplex structures. in *quadruplex nucleic acids*; The Royal Society of Chemistry, 2006; pp 1-30.
- (9) Dai, J.; Carver, M.; Yang, D.: Polymorphism of human telomeric quadruplex structures. *Biochimie* **2008**, *90*, 1172-1183.
- (10) Wu, G.; Wong, A.; Gan, Z.; Davis, J. T.: Direct detection of potassium cations bound to G-quadruplex structures by solid-state 39K NMR at 19.6 T. *J. Am. Chem. Soc.* **2003**, *125*, 7182-7183.
- (11) Nagatoishi, S.; Nojima, T.; Juskowiak, B.; Takenaka, S.: A pyrene-labeled G-quadruplex oligonucleotide as a fluorescent probe for potassium ion detection in biological applications. *Angew. Chem. Int. Ed.* **2005**, *117*, 5195-5198.

- (12) Parkinson, G. N.; Lee, M. P.; Neidle, S.: Crystal structure of parallel quadruplexes from human telomeric DNA. *Nature* **2002**, *417*, 876-880.
- (13) Wang, Y.; Patel, D. J.: Solution structure of the human telomeric repeat d[AG₃(T₂AG₃)₃] G-tetraplex. *Structure* **1993**, *1*, 263-282.
- (14) Harley, C. B.: 9 Telomeres and aging. *Cold Spring Harbor Monograph Archive* **1995**, *29*, 247-263.
- (15) Harley, C. B.; Vaziri, H.; Counter, C. M.; Allsopp, R. C.: The telomere hypothesis of cellular aging. *Exp. Gerontol.* **1992**, *27*, 375-382.
- (16) Aubert, G.; Lansdorp, P. M.: Telomeres and aging. *Physiol. Rev.* **2008**, *88*, 557-579.
- (17) Collie, G. W.; Promontorio, R.; Hampel, S. M.; Micco, M.; Neidle, S.; Parkinson, G. N.: Structural basis for telomeric G-quadruplex targeting by naphthalene diimide ligands. *J. Am. Chem. Soc.* **2012**, *134*, 2723-2731.
- (18) Oganessian, L.; Bryan, T. M.: Physiological relevance of telomeric G-quadruplex formation: a potential drug target. *BioEssays* **2007**, *29*, 155-165.
- (19) Neidle, S.: Human telomeric G-quadruplex: The current status of telomeric G-quadruplexes as therapeutic targets in human cancer. *FEBS J.* **2010**, *277*, 1118-1125.
- (20) Pennarun, G.; Granotier, C.; Gauthier, L. R.; Gomez, D.; Hoffschir, F.; Mandine, E.; Riou, J.-F.; Mergny, J.-L.; Mailliet, P.; Boussin, F. D.: Apoptosis related to telomere instability and cell cycle alterations in human glioma cells treated by new highly selective G-quadruplex ligands. *Oncogene* **2005**, *24*, 2917-2928.
- (21) Collie, G. W.; Haider, S. M.; Neidle, S.; Parkinson, G. N.: A crystallographic and modelling study of a human telomeric RNA (TERRA) quadruplex. *Nucleic Acids Res.* **2010**, *38*, 5569-5580.
- (22) Collie, G.; Reszka, A. P.; Haider, S. M.; Gabelica, V.; Parkinson, G. N.; Neidle, S.: Selectivity in small molecule binding to human telomeric RNA and DNA quadruplexes. *Chem. Commun.* **2009**, 7482-7484.

- (23) Drygin, D.; Lin, A.; Bliesath, J.; Ho, C. B.; O'Brien, S. E.; Proffitt, C.; Omori, M.; Haddach, M.; Schwaebe, M. K.; Siddiqui-Jain, A.: Targeting RNA polymerase I with an oral small molecule CX-5461 inhibits ribosomal RNA synthesis and solid tumor growth. *Cancer Res.* **2011**, *71*, 1418-1430.
- (24) Arora, A.; Dutkiewicz, M.; Scaria, V.; Hariharan, M.; Maiti, S.; Kurreck, J.: Inhibition of translation in living eukaryotic cells by an RNA G-quadruplex motif. *RNA* **2008**, *14*, 1290-1296.
- (25) Shahid, R.; Bugaut, A.; Balasubramanian, S.: The BCL-2 5' untranslated region contains an RNA G-quadruplex-forming motif that modulates protein expression. *Biochemistry* **2010**, *49*, 8300-8306.
- (26) Balkwill, G. D.; Derecka, K.; Garner, T. P.; Hodgman, C.; Flint, A. P. F.; Searle, M. S.: Repression of translation of human estrogen receptor α by G-quadruplex formation. *Biochemistry* **2009**, *48*, 11487-11495.
- (27) Huppert, J. L.; Balasubramanian, S.: G-quadruplexes in promoters throughout the human genome. *Nucleic Acids Res.* **2007**, *35*, 406-413.
- (28) Qin, Y.; Hurley, L. H.: Structures, folding patterns, and functions of intramolecular DNA G-quadruplexes found in eukaryotic promoter regions. *Biochimie* **2008**, *90*, 1149-1171.
- (29) Wu, Y.; Brosh, R. M.: G-quadruplex nucleic acids and human disease. *FEBS J.* **2010**, *277*, 3470-3488.
- (30) Siddiqui-Jain, A.; Grand, C. L.; Bearss, D. J.; Hurley, L. H.: Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. *Proc. Natl. Acad. Sci.* **2002**, *99*, 11593-11598.
- (31) Rangan, A.; Fedoroff, O. Y.; Hurley, L. H.: Induction of duplex to G-quadruplex transition in the c-myc promoter region by a small molecule. *J. Biol. Chem.* **2001**, *276*, 4640-4646.
- (32) Porru, M.; Artuso, S.; Salvati, E.; Bianco, A.; Franceschin, M.; Diodoro, M. G.; Passeri, D.; Orlandi, A.; Savorani, F.; D'Incalci, M.: Targeting G-quadruplex

DNA structures by EMICORON has a strong antitumor efficacy against advanced models of human colon cancer. *Mol. Cancer Ther.* **2015**, *14*, 2541-2551.

(33) German, I.; Buchanan, D. D.; Kennedy, R. T.: Aptamers as ligands in affinity probe capillary electrophoresis. *Anal. Chem.* **1998**, *70*, 4540-4545.

(34) Nutiu, R.; Li, Y.: Structure-switching signaling aptamers. *J. Am. Chem. Soc.* **2003**, *125*, 4771-4778.

(35) Ueyama, H.; Takagi, M.; Takenaka, S.: A novel potassium sensing in aqueous media with a synthetic oligonucleotide derivative. Fluorescence resonance energy transfer associated with guanine quartet–potassium ion complex formation. *J. Am. Chem. Soc.* **2002**, *124*, 14286-14287.

(36) Marsh, T. C.; Henderson, E.: G-Wires: Self-assembly of a telomeric oligonucleotide, d(GGGGTTGGGG), into large superstructures. *Biochemistry* **1994**, *33*, 10718-10724.

(37) Marsh, T. C.; Vesenka, J.; Henderson, E.: A new DNA nanostructure, the G-wire, imaged by scanning probe microscopy. *Nucleic Acids Res.* **1995**, *23*, 696-700.

(38) Dai, T.-Y.; Marotta, S. P.; Sheardy, R. D.: Self-assembly of DNA oligomers into high molecular weight species. *Biochemistry* **1995**, *34*, 3655-3662.

(39) Kotia, R. B.; Li, L.; McGown, L. B.: Separation of nontarget compounds by DNA aptamers. *Anal. Chem.* **2000**, *72*, 827-831.

(40) Davis, J. T.; Spada, G. P.: Supramolecular architectures generated by self-assembly of guanosine derivatives. *Chem. Soc. Rev.* **2007**, *36*, 296-313.

(41) Azargun, M.; Fridgen, T. D.: Guanine tetrads: an IRMPD spectroscopy, energy resolved SORI-CID, and computational study of M(9-ethylguanine)₄⁺ (M = Li, Na, K, Rb, Cs) in the gas phase. *Phys. Chem. Chem. Phys.* **2015**, *17*, 25778-25785.

(42) Fukushima, K.; Iwahashi, H.: 1:1 Complex of guanine quartet with alkali metal cations detected by electrospray ionization mass spectrometry. *Chem. Commun.* **2000**, 895-896.

- (43) Fraschetti, C.; Montagna, M.; Guarcini, L.; Guidoni, L.; Filippi, A.: Spectroscopic evidence for a gas-phase librating G-quartet- Na^+ complex. *Chem. Commun.* **2014**, *50*, 14767-14770.
- (44) Rodgers, M. T.; Armentrout, P. B.: Noncovalent interactions of nucleic acid bases (uracil, thymine, and adenine) with alkali metal ions. Threshold collision-induced dissociation and theoretical studies. *J. Am. Chem. Soc.* **2000**, *122*, 8548-8558.
- (45) Cerda, B. A.; Wesdemiotis, C.: Li^+ , Na^+ , and K^+ binding to the DNA and RNA nucleobases. Bond energies and attachment sites from the dissociation of metal Ion-bound heterodimers. *J. Am. Chem. Soc.* **1996**, *118*, 11884-11892.
- (46) Hud, N. V.; Smith, F. W.; Anet, F. A.; Feigon, J.: The selectivity for K^+ versus Na^+ in DNA quadruplexes is dominated by relative free energies of hydration: a thermodynamic analysis by ^1H NMR. *Biochemistry* **1996**, *35*, 15383-15390.
- (47) Elder, F. R.; Gurewitsch, A. M.; Langmuir, R. V.; Pollock, H. C.: Radiation from electrons in a synchrotron. *Phys. Rev.* **1947**, *71*, 829-830.
- (48) Azargun, M.; Jami-Alahmadi, Y.; Fridgen, T. D.: The intrinsic stabilities and structures of alkali metal cationized guanine quadruplexes. *Phys. Chem. Chem. Phys.* **2017**, *19*, 1281-1287.
- (49) Gray, R. D.; Chaires, J. B.: Kinetics and mechanism of K^+ - and Na^+ -induced folding of models of human telomeric DNA into G-quadruplex structures. *Nucleic Acids Res.* **2008**, *36*, 4191-4203.
- (50) Williamson, J. R.; Raghuraman, M. K.; Cech, T. R.: Monovalent cation-induced structure of telomeric DNA: The G-quartet model. *Cell*, *59*, 871-880.
- (51) Sundquist, W. I.; Klug, A.: Telomeric DNA dimerizes by formation of guanine tetrads between hairpin loops. *Nature* **1989**, *342*, 825-829.
- (52) Zaccaria, F.; Paragi, G.; Fonseca Guerra, C.: The role of alkali metal cations in the stabilization of guanine quadruplexes: why K^+ is the best. *Phys. Chem. Chem. Phys.* **2016**, *18*, 20895-20904.

- (53) Wong, A.; Wu, G.: Selective binding of monovalent cations to the stacking G-quartet structure formed by guanosine 5'-monophosphate: A solid-state NMR study. *J. Am. Chem. Soc.* **2003**, *125*, 13895-13905.
- (54) Burt, M. B.; Fridgen, T. D.: Structures and physical properties of gaseous metal cationized biological ions. *Eur. J. Mass Spectrom* **2012**, *18*, 235-250.
- (55) Fridgen, T. D.: Infrared consequence spectroscopy of gaseous protonated and metal ion cationized complexes. *Mass Spectrom. Rev.* **2009**, *28*, 586-607.
- (56) Peltz, C.; Drahos, L.; Vékey, K.: SORI excitation: Collisional and radiative processes. *J. Am. Soc. Mass Spectrom.* **2007**, *18*, 2119-2126.
- (57) Parkinson, G. N.; Lee, M. P. H.; Neidle, S.: Crystal structure of parallel quadruplexes from human telomeric DNA. *Nature.* **2002**, *417*, 876-880.
- (58) Xu, Y.: Chemistry in human telomere biology: structure, function and targeting of telomere DNA/RNA. *Chem. Soc.Rev.* **2011**, *40*, 2719-2740.
- (59) Wang, Y.; Patel, D. J.: Guanine residues in d(T₂AG₃) and d(T₂G₄) form parallel-stranded potassium cation stabilized G-quadruplexes with anti glycosidic torsion angles in solution. *Biochemistry.* **1992**, *31*, 8112-8119.
- (60) Mignon, P.; Loverix, S.; Steyaert, J.; Geerlings, P.: Influence of the π - π interaction on the hydrogen bonding capacity of stacked DNA/RNA bases. *Nucleic Acids Res.* **2005**, *33*, 1779-1789.

Chapter 6. Summary and Outlook

6.1. Summary

Ion trap mass spectrometers are capable of storing ions for an extended period of time in a low-pressure environment. This capability empowers one to determine ion geometries and thermochemistries when mating ion traps to ion activation techniques including photodissociation and collisional activation.¹ The Fourier transform ion cyclotron mass spectrometer (FTICR) has been the most widely used ion trap mass spectrometer to probe the physical chemistry of ions in the gas phase. By coupling an external light source (e.g. FEL, and OPO laser) to an FTICR, photodissociation of trapped ions can be monitored. IRMPD spectroscopy is well suited to photo-dissociate ions in the gas phase that, like typical infrared spectroscopy, can be used to identify functional groups leading to structure elucidation. Theoretical chemistry plays an important role in structural identification by comparing computed IR and experimental IRMPD spectra in order to assign the most likely structures.

BIRD, another photodissociation technique, has been conducted in an FTICR; BIRD is a very slow heating process that activates weakly bound clusters.² The heated walls of an ICR cell emit low energy infrared photons that make trapped ions engage in an energy exchange process with the heated walls leading to dissociation. Temperature dependent BIRD rate constants can then be obtained by repeating BIRD experiments at a range of temperatures. BIRD provides valuable physical chemistries including ion activation energies, binding energies, and mechanistic insights.

Energy resolved SORI-CID is a collisional activation technique that has also been used in this present work. This technique can activate ions through a slow collisional process in which ions undergo up to thousands of collisions. SORI-CID not only provides structural information but also is beneficial to examine the gas phase relative stability of ions under study.

The main objective of this thesis was to investigate the intrinsic properties of alkali metal guanine adducts in the gas phase. In Chapter 3, the structure and stability of alkali metal guanine tetrads were comprehensively assessed by a combination of mass spectrometric techniques and computational chemistry. Infrared multiple photon dissociation (IRMPD) spectroscopy was used to probe the structure of $M(9eG)_4^+$ ($M = \text{Li, Na, K, Rb, Cs}$) in both the fingerprint region ($900 - 1850 \text{ cm}^{-1}$) and the N-H/C-H stretching region ($2700 - 3800 \text{ cm}^{-1}$). IRMPD spectra of these five alkali metal tetrads showed very similar features in both regions. The computed infrared spectrum of the lowest energy structure for each G-tetrad was consistent with the experimental spectrum showing a metal-centered G-tetrad structure with Hoogsteen base pairing. The IRMPD

spectrum of $\text{Na}(9\text{eG})_4^+$ was also shown to be consistent with the $\text{Na}(\text{guanine})_4^+$ spectrum.³ This shows that the ethyl groups at N9 appear to have little impact on the structure of the G-tetrads. In order to study the gas phase stability of G-tetrads, energy resolved SORI-CID experiments were conducted. These experiments revealed that $\text{Na}(9\text{eG})_4^+$ is the most stable G-tetrad followed by the other metal cation centered tetrads in the order: $\text{Li}(9\text{eG})_4^+ > \text{K}(9\text{eG})_4^+ > \text{Rb}(9\text{eG})_4^+ > \text{Cs}(9\text{eG})_4^+$. This stability order was also obtained by computational methods. The lowest energy structure found for $\text{Na}(9\text{eG})_4^+$ was determined to be planar due to the optimal fit of the Na^+ cation for the cavity formed by the oxygen atoms of four carbonyl groups. As a result, Na^+ has the lowest distortion and therefore destabilization of the neutral G-tetrad. Our calculations showed that Li^+ distorts the neutral G-tetrad to a great extent, weakening hydrogen bonds. This finding explains why $\text{Li}(9\text{eG})_4^+$ is less stable than $\text{Na}(9\text{eG})_4^+$ in spite of it having the strongest electrostatic interactions with 9-ethylguanine molecules.

The work in Chapter 3 did not explain why potassium is the ion of choice in biological G-quadruplexes. In Chapter 4, the gas phase structure and energetics of alkali metal G-quadruplexes was pursued. The simplest G-quadruplexes, $\text{M}(9\text{eG})_8^+$, which were composed of two stacked G-tetrads with a metal cation between them were formed for Na^+ , K^+ , Rb^+ , Cs^+ . Na^+ was the smallest cation that formed a G-quadruplex as all attempts to generate $\text{Li}(9\text{eG})_8^+$ failed. This is in agreement with our observation in the third chapter that Li^+ strongly distorted the G-tetrad from planarity, much closer to a tetrahedral coordination of the 9eG about Li^+ . As a result, the tendency of Li^+ to be a non-planar G-tetrad means it cannot form a sandwich quadruplex structure. This is consistent

with the potential of lithium to degrade and prohibit the formation of G-quadruplexes in nucleic acids and with the calculations of Zaccaria⁴ in the aqueous phase, which concluded that Li^+ avoids forming sandwich clusters. In order to gain insight into the structure of alkali metal quadruplexes, vibrational spectroscopy experiments were performed in the N-H stretching region. The experimental spectra of the quadruplexes displayed very similar features in a way that the clusters were indistinguishable solely based on IRMPD spectra. Comparison of IRMPD spectra with computed infrared spectra showed consistency with structures in which the metal cations were sandwiched between two guanine tetrads for all complexes. The intrinsic stabilities and binding energies of the quadruplexes were determined by BIRD experiments which resulted in the following energetic ordering: $\text{K}(9\text{eG})_8^+ > \text{Rb}(9\text{eG})_8^+ > \text{Cs}(9\text{eG})_8^+ \gg \text{Na}(9\text{eG})_8^+$. This explains why K^+ is the ion of choice for G-quadruplexes in the DNA of biological systems despite the greater abundance of Na^+ .

This stability order for the quadruplexes was quite different than of the metal cationized tetrads: $\text{Na}(9\text{eG})_4^+ > \text{Li}(9\text{eG})_4^+ \gg \text{K}(9\text{eG})_4^+ > \text{Rb}(9\text{eG})_4^+ > \text{Cs}(9\text{eG})_4^+$. The position of the sodiated complex is attributed to the ability of the Na^+ to sit in the centre of a planar hydrogen-bonded tetrad complex while the heavier cations sit above the plane of the tetrad, readily able to bind to a second tetrad. To form a quadruplex, Na^+ must be pulled from its stable position in the centre of the tetrad. BIRD experiments also showed that $\text{Na}(9\text{eG})_8^+$ loses one 9eG whereas the quadruplexes of the heavier cations lose a neutral guanine tetrad. This was also reflected in the thermochemistry of the transition

states where the entropy of transition state for the heavier cations are significantly more negative than that for $\text{Na}(\text{9eG})_8^+$.

In Chapter 5, the structure and energetics of the K^+ associated quadruplex, $\text{K}_2(\text{9eG})_{12}^+$ were studied. So far, all attempts failed to form larger G-quadruplexes with Na^+ , Rb^+ , and Cs^+ , and K^+ was the only alkali metal that formed a quadruplex with three stacked G-tetrads (i.e. $\text{G}_4\text{-K-G}_4\text{-K-G}_4$). Some works conducted in the solution phase proposed that the affinity of the K^+ cation to form larger quadruplexes stems from relative hydration energies especially when compared to Na^+ .^{5,6} However, our intrinsic studies, in the absence of any solvent, showed that a significant part of the K^+ ability to form larger quadruplexes comes from the optimal fit of K^+ in the $\text{G}_4\text{-G}_4$ cage. The IRMPD spectrum of $\text{K}_2(\text{9eG})_{12}^{2+}$ was recorded and its comparison to the spectra of the smaller G-quadruplexes and G-tetrads showed very similar spectroscopic features. It was concluded that the $\text{K}_2(\text{9eG})_{12}^{2+}$ adopts a sandwich-like structure like the smaller quadruplexes. BIRD kinetics experiments and energy-resolved SORI-CID were used to examine the stability and binding energy of this quadruplex, $\text{K}_2(\text{9eG})_{12}^{2+}$. Results from SORI-CID experiments concluded that $\text{K}_2(\text{9eG})_{12}^{2+}$ is a more stable cluster toward dissociation to $\text{K}(\text{9eG})_8^+$ and $\text{K}(\text{9eG})_4^+$ compared to the stability of the $\text{K}(\text{9eG})_8^+$ cluster which loses $(\text{9eG})_4$. Temperature dependent BIRD kinetics for $\text{K}_2(\text{9eG})_{12}^{2+}$ yielded an activation energy valued at $225 \pm 15 \text{ kJ mol}^{-1}$ which was higher than that determined for $\text{K}(\text{9eG})_8^+$.

6.2. Future Work

For future work, it would be worthwhile to look at the formation of G-tetrads and G-quadruplexes with doubly charged cations. Structural studies and gas phase stabilities of doubly charged G-quadruplexes can reveal how metal cations with more charges can affect the intrinsic properties of these clusters. It was highlighted in this work that alkali metal cations template the formation of G-tetrads in which eight hydrogen bonds are involved. Alkali metal cations filled the cavity formed by the carbonyl groups, stabilizing the complex by neutralizing the negative potential of the cavity. Doubly charged cations may change the intrinsic properties of G-tetrads due to the different charge distribution. The possibility of formation of deprotonated G-tetrads can change the strength of hydrogen bonds and the structure. It is also interesting to examine if doubly charged cations can form sandwich structures, quadruplexes, and how their stability and structure vary compared to alkali metal G-quadruplexes. There have been reports on the formation of trivalent metal G-quadruplexes⁷ but the gas phase physical chemistry of these structures has not been studied. Activation techniques such as IRMPD, BIRD and CID along with IRMPD spectroscopy can be used to study the stability and structure of these quadruplexes as well.

It will be also interesting to compare the G-adducts formed by metal cations with similar sizes but different charge states. This could be very helpful in order to understand the role of size in formation of G-tetrads and G-quadruplexes. It was shown in Chapters 4 and 5 that the optimal fit of K^+ cation to the cavity is the main reason that favors this cation to be the ion of choice in biological G-quadruplexes. In this regard, it can be

possible to probe the formation of G-quadruplexes with a doubly charged cation (Ba^{2+}) that possesses a similar ionic size compared to the K^+ cation. Formation and study of larger G-quadruplexes, composed of three or more G4s, with a variety of metal cations can also be pursued in order to compare their physical chemistries.

Some works have shown⁸⁻¹³ that certain metal cations such as Ag^+ , Hg^{2+} , and Pb^{2+} can destroy or alter the structure of biological G-quadruplexes, which can cause serious health issues. It would be helpful to look into the gas phase properties of G-quadruplexes with these metals to assess their affinity toward guanine base, stability, and structure. In addition, gas phase experiments can be designed to examine the destabilizing effect of these metals on biological G-quadruplexes containing Na^+ and K^+ . On the other hand, studies have shown that certain metals can offset and override the destabilization effect of the aforementioned metals.¹³ For instance, Ba^{2+} (in the presence of specific ligands) has been observed to cancel the destructive effect of Ag^+ and Hg^{2+} . The interaction mechanism is not well understood. Hence it would be worthwhile to study the Ba^{2+} , Ag^+ , and Hg^{2+} associated G-quadruplexes to assess the structures and energetics.

G-quadruplexes can be used as anti-cancer targets since they are inhibitory to telomerase performance.¹⁴ The therapeutic effectiveness of G-quadruplexes for anti-cancer purposes can be guaranteed when this cluster is stabilized with certain ligands.¹⁵ It can be interesting to study the interaction of stabilizing ligands in the gas phase. Gas phase formation of ligand-quadruplexes followed by ion activation analysis can evaluate the amount of stability that ligands introduce into the system. The IRMPD spectroscopy

can be used to assess how the ligands interact and bind to the G-quadruplex and the structures which result.

There have been arguments about the role of sugar-phosphate backbones in the stability, structure, and metal coordination of G-quadruplexes. It should also be attempted to form G-adducts with guanosine instead of guanine (or 9-ethylguanine). It is useful to experimentally explain if the sugar-phosphate backbone impacts the structure, fragment products, and stability of these adducts in the gas phase.

References:

- (1) Sleno, L.; Volmer, D. A.: Ion activation methods for tandem mass spectrometry. *J. Mass Spectrom.* **2004**, *39*, 1091-1112.
- (2) Dunbar, R. C.: BIRD (blackbody infrared radiative dissociation): Evolution, principles, and applications. *Mass Spectrom. Rev.* **2004**, *23*, 127-158.
- (3) Fraschetti, C.; Montagna, M.; Guarcini, L.; Guidoni, L.; Filippi, A.: Spectroscopic evidence for a gas-phase librating G-quartet- Na^+ complex. *Chem. Commun.* **2014**, *50*, 14767-14770.
- (4) Zaccaria, F.; Paragi, G.; Fonseca Guerra, C.: The role of alkali metal cations in the stabilization of guanine quadruplexes: why K^+ is the best. *Phys. Chem. Chem. Phys.* **2016**, *18*, 20895-20904.
- (5) Hud, N. V.; Smith, F. W.; Anet, F. A.; Feigon, J.: The selectivity for K^+ versus Na^+ in DNA quadruplexes is dominated by relative free energies of hydration: a thermodynamic analysis by ^1H NMR. *Biochemistry* **1996**, *35*, 15383-15390.
- (6) Elder, F. R.; Gurewitsch, A. M.; Langmuir, R. V.; Pollock, H. C.: Radiation from Electrons in a Synchrotron. *Phys. Rev.* **1947**, *71*, 829-830.
- (7) Kwan, I. C.; She, Y.-M.; Wu, G.: Trivalent lanthanide metal ions promote formation of stacking G-quartets. *Chem. Commun.* **2007**, 4286-4288.
- (8) Zhao, J.; Chen, C.; Zhang, L.; Jiang, J.; Shen, G.; Yu, R.: A Hg^{2+} -mediated label-free fluorescent sensing strategy based on G-quadruplex formation for selective detection of glutathione and cysteine. *Analyst* **2013**, *138*, 1713-1718.
- (9) Zhou, X.-H.; Kong, D.-M.; Shen, H.-X.: Ag^+ and cysteine quantitation based on G-quadruplex-hemin DNazymes disruption by Ag^+ . *Anal. Chem.* **2009**, *82*, 789-793.
- (10) Karimata, H.; Miyoshi, D.; Fujimoto, T.; Koumoto, K.; Wang, Z.-M.; Sugimoto, N.: Conformational switch of a functional nanowire based on the DNA G-quadruplex. In *Nucleic Acids Symp. Ser.*; Oxford University Press, 2007; Vol. 51; pp 251-252.

- (11) Liu, L.; Liu, W.; Hong, T.; Weng, X.; Zhai, Q.; Zhou, X.: Ag⁺ and cysteine detection by Ag⁺–guanine interaction based on graphene oxide and G-quadruplex DNA. *Anal. Methods* **2012**, *4*, 1935-1939.
- (12) Liu, W.; Zhu, H.; Zheng, B.; Cheng, S.; Fu, Y.; Li, W.; Lau, T.-C.; Liang, H.: Kinetics and mechanism of G-quadruplex formation and conformational switch in a G-quadruplex of PS2.M induced by Pb²⁺. *Nucleic Acids Res.* **2012**, *40*, 4229-4236.
- (13) Lu, H.; Li, S.; Chen, J.; Xia, J.; Zhang, J.; Huang, Y.; Liu, X.; Wu, H.-c.; Zhao, Y.; Chai, Z.: Metal ions modulate the conformation and stability of a G-quadruplex with or without a small-molecule ligand. *Metallomics* **2015**, *7*, 1508-1514.
- (14) Wang, Q.; Liu, J.Q.; Chen, Z.; Zheng, K.W.; Chen, C.Y.; Hao, Y.H.; Tan, Z.: G-quadruplex formation at the 3' end of telomere DNA inhibits its extension by telomerase, polymerase and unwinding by helicase. *Nucleic Acids Res.* **2011**, *39*, 6229-6237.
- (15) Ilyinsky, N.; Varizhuk, A.; Beniaminov, A.; Puzanov, M.; Shchyolkina, A.; Kaluzhny, D.: G-quadruplex ligands: Mechanisms of anticancer action and target binding. *Molecular Biology* **2014**, *48*, 778-794.

Appendix 1

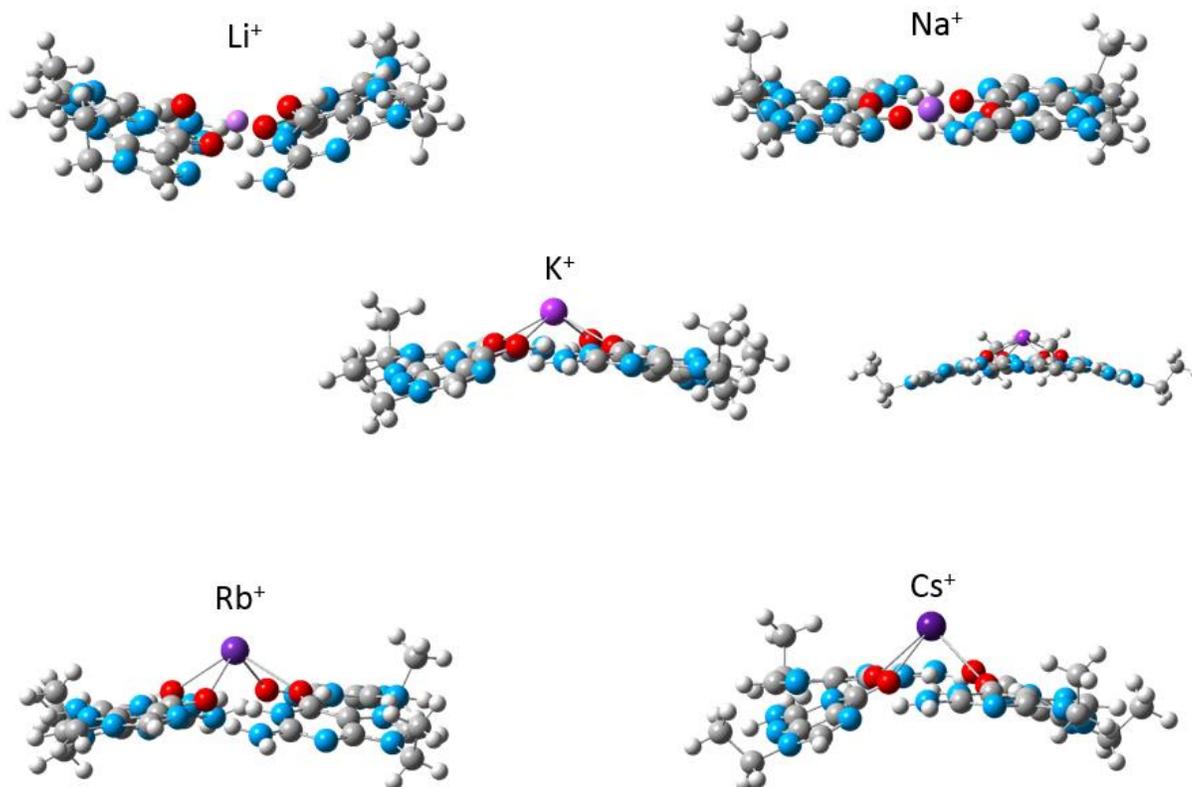


Figure S3.1. Profiles of the computed structures of the $M(9eG)_4^+$ complexes computed using B3LYP/6-31+G(d,p) with Def2SVPD on the metal centers. The calculations are also empirically corrected for dispersion.

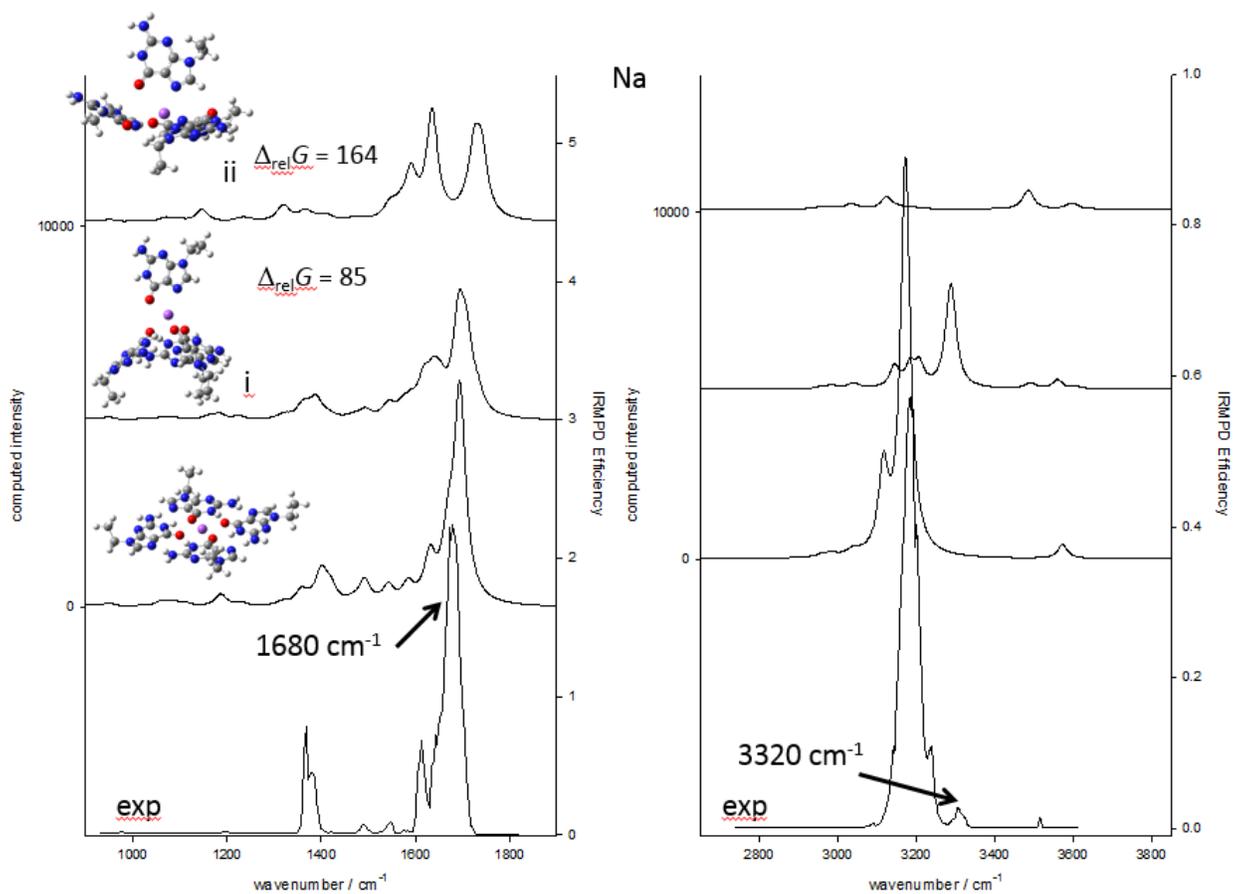


Figure S3.2 a) Comparison of computed IR spectra for three isomers of $\text{Na}(9\text{eG})_4^+$ with the experimental IRMPD spectra.

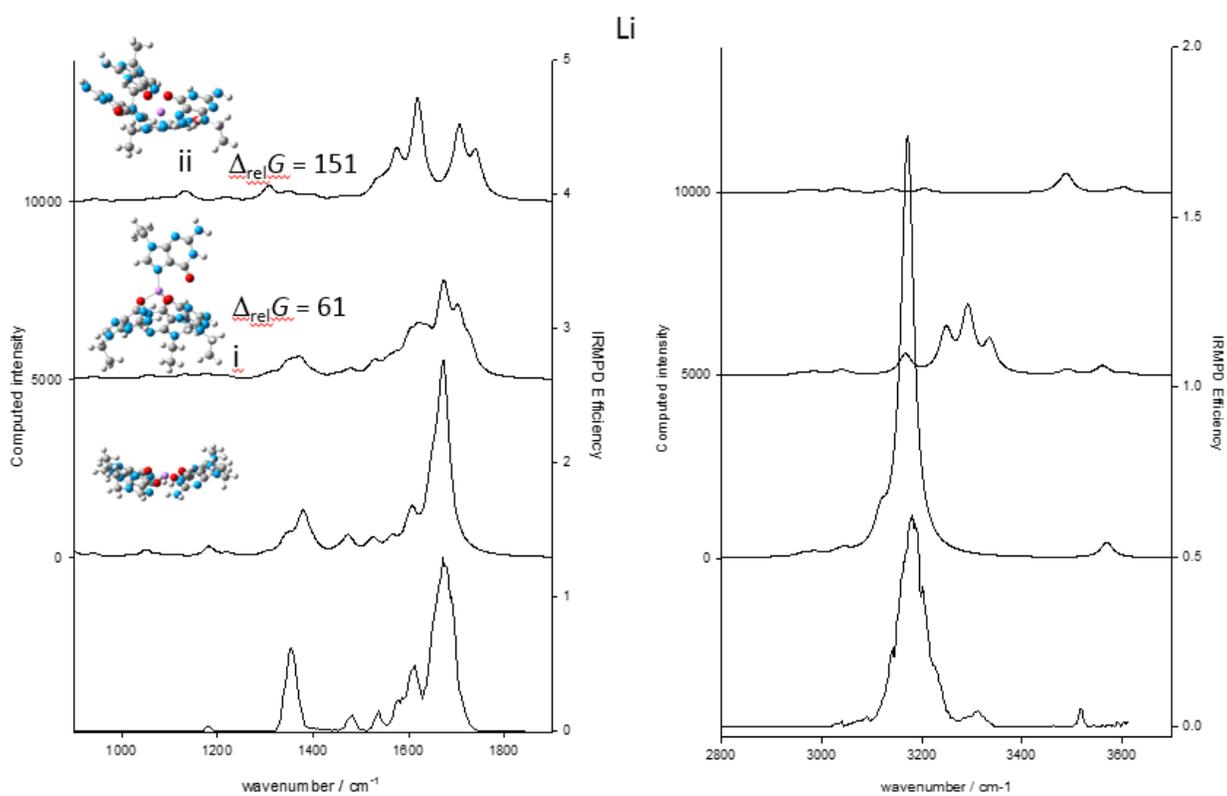


Figure S3.2 b) Comparison of computed IR spectra for three isomers of $\text{Li}(9eG)_4^+$ with the experimental IRMPD spectra.

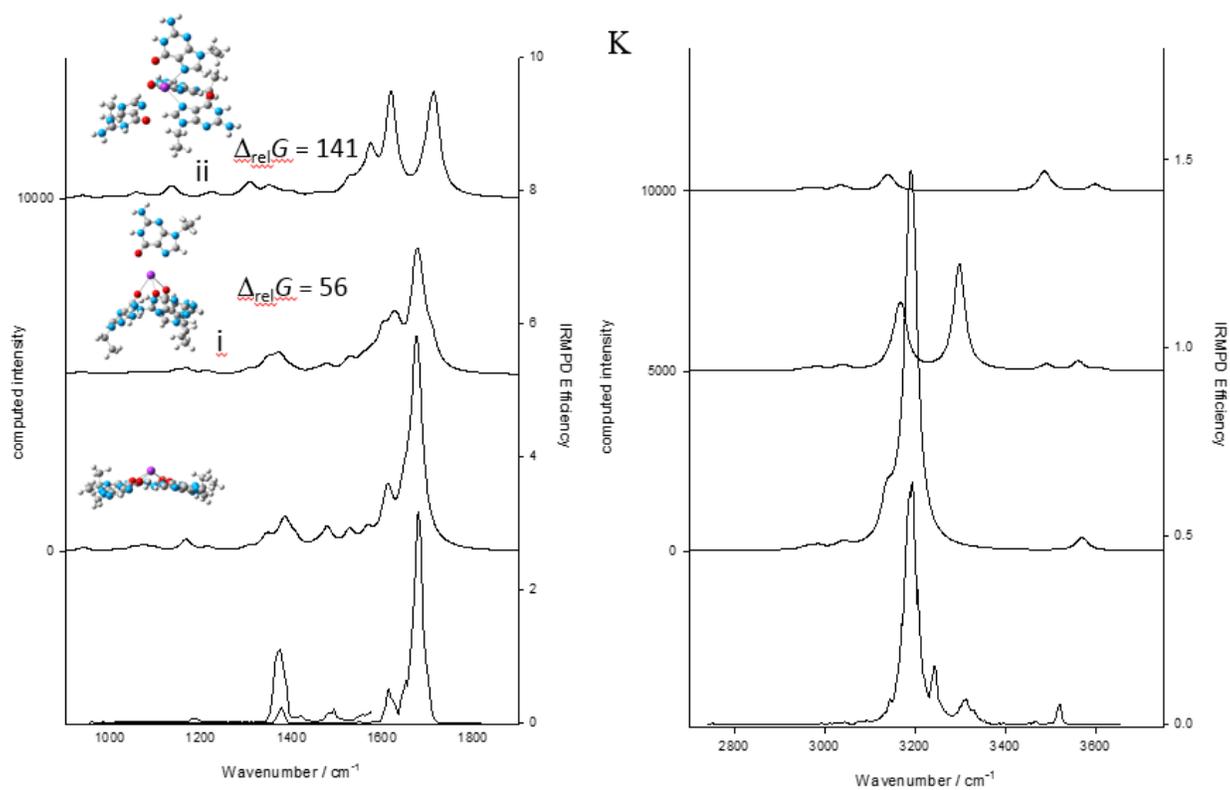


Figure S3.2 c) Comparison of computed IR spectra for three isomers of $K(9eG)_4^+$ with the experimental IRMPD spectra.

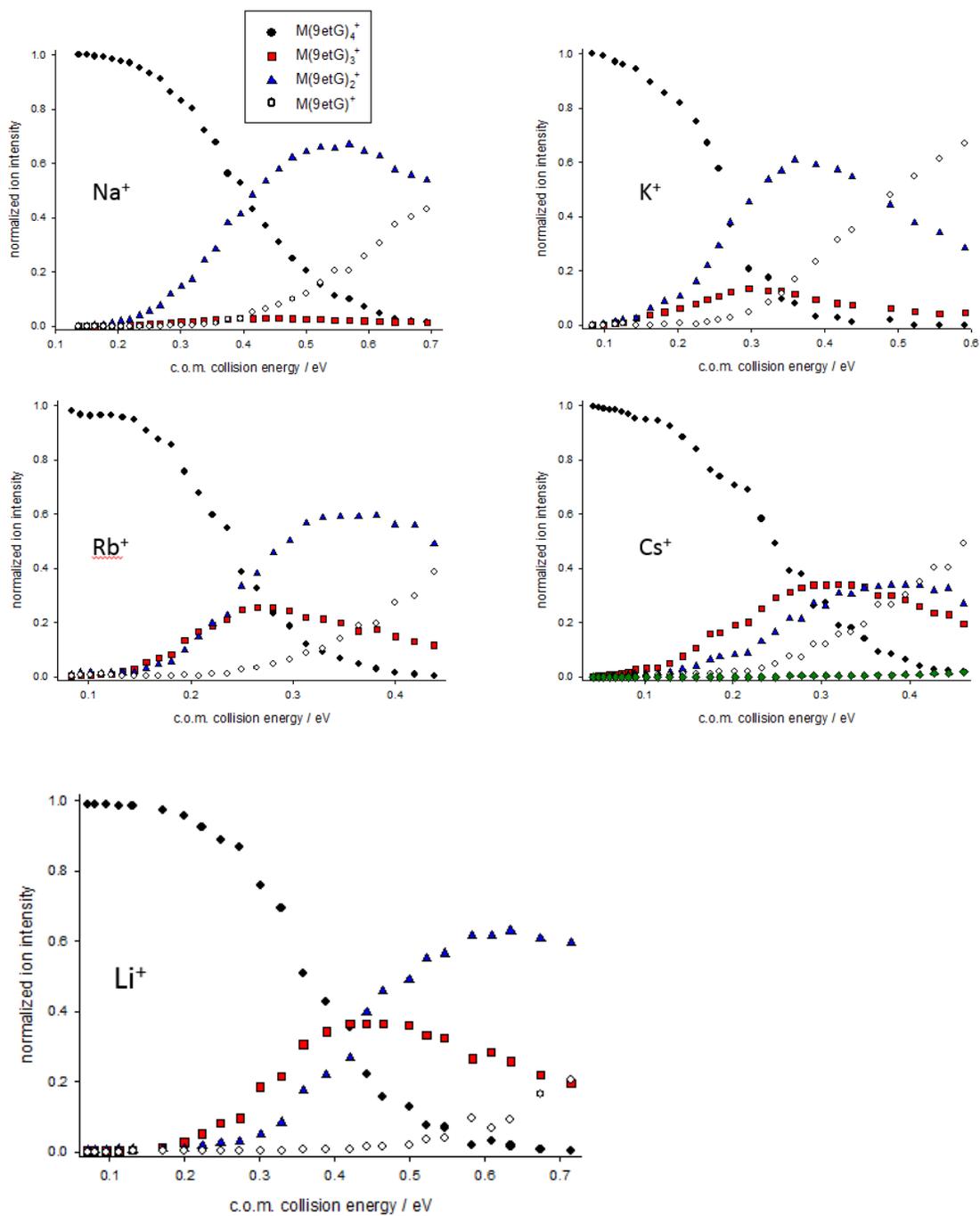


Figure S3.3. Energy resolved intensity profiles for SORI-CID of the M(9-eG)₄⁺ complexes.

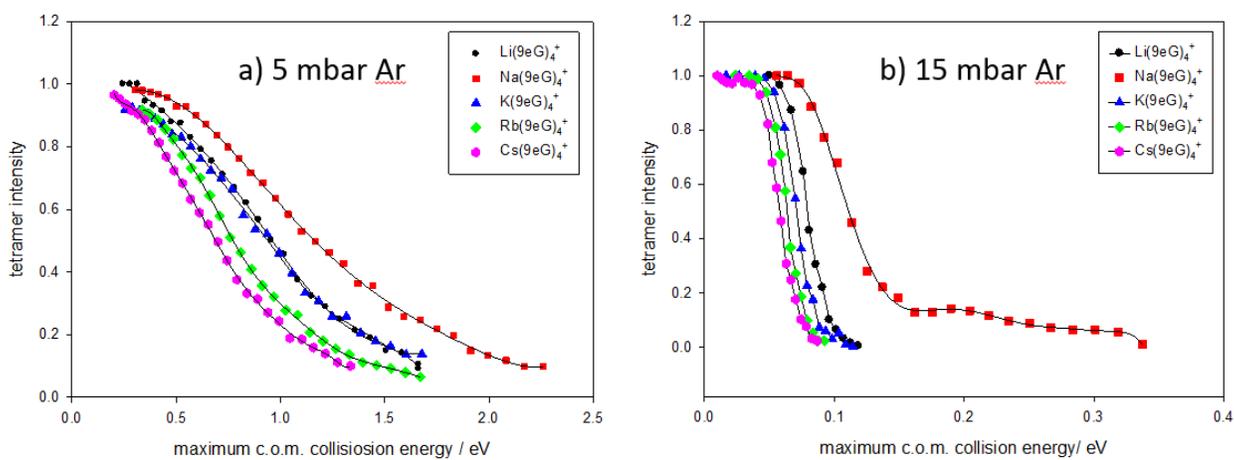


Figure S3.4. Energy-resolved G-tetrad intensity profiles for the SORI-CID of $M(9eG)_4^+$ with a) a reservoir pressure of 5 mbar and b) a reservoir pressure of 15 mbar with Ar.

Appendix 2

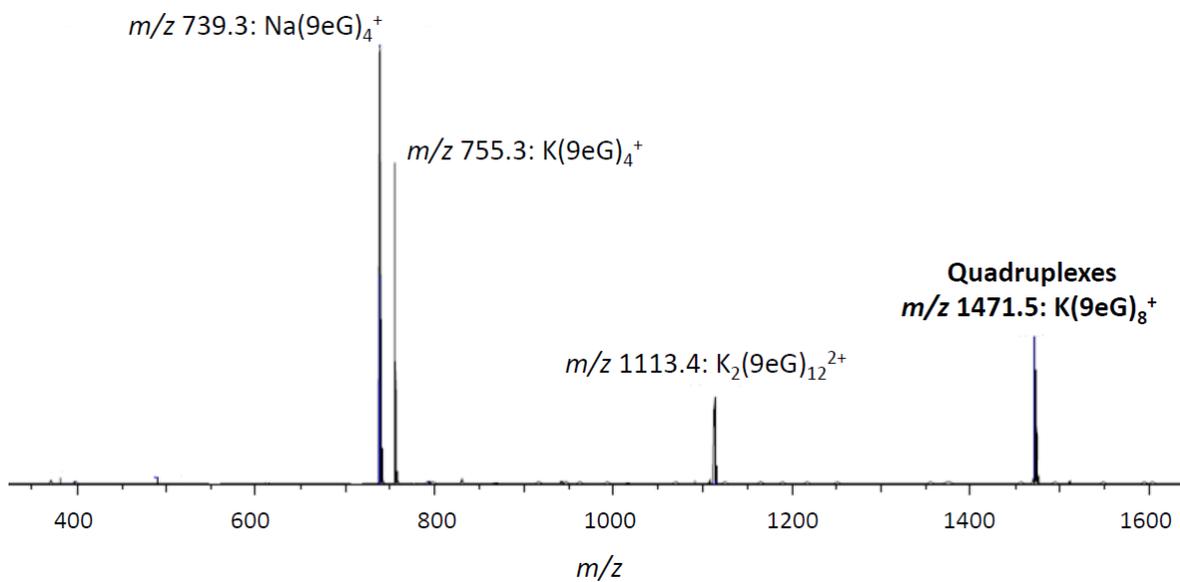


Figure S4.1. Electrospray mass spectrum of an aqueous solution of KCl/9-ethylguanine.

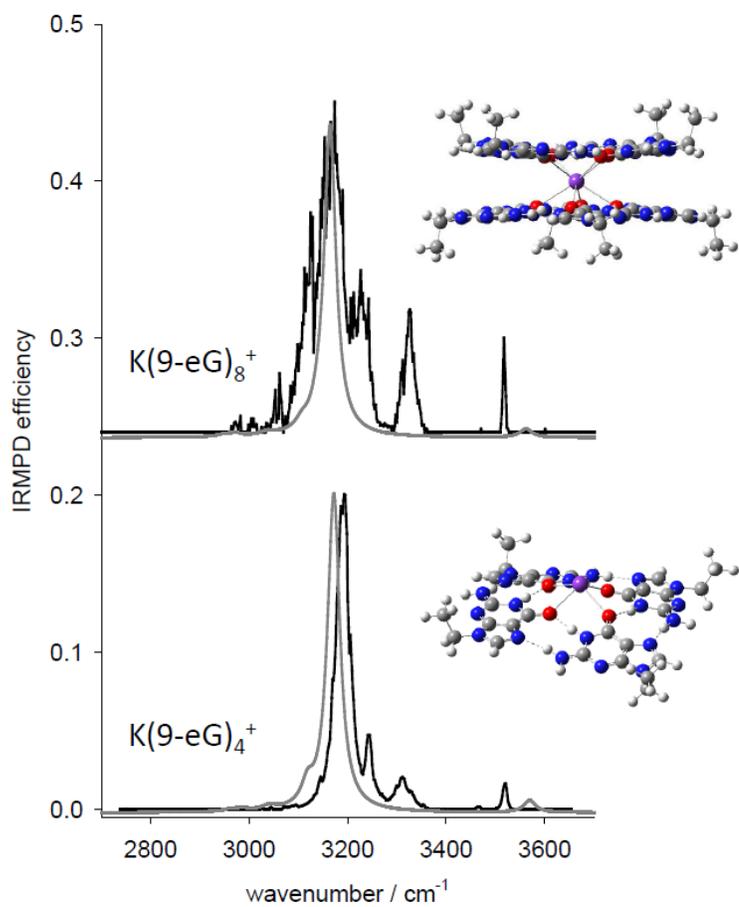
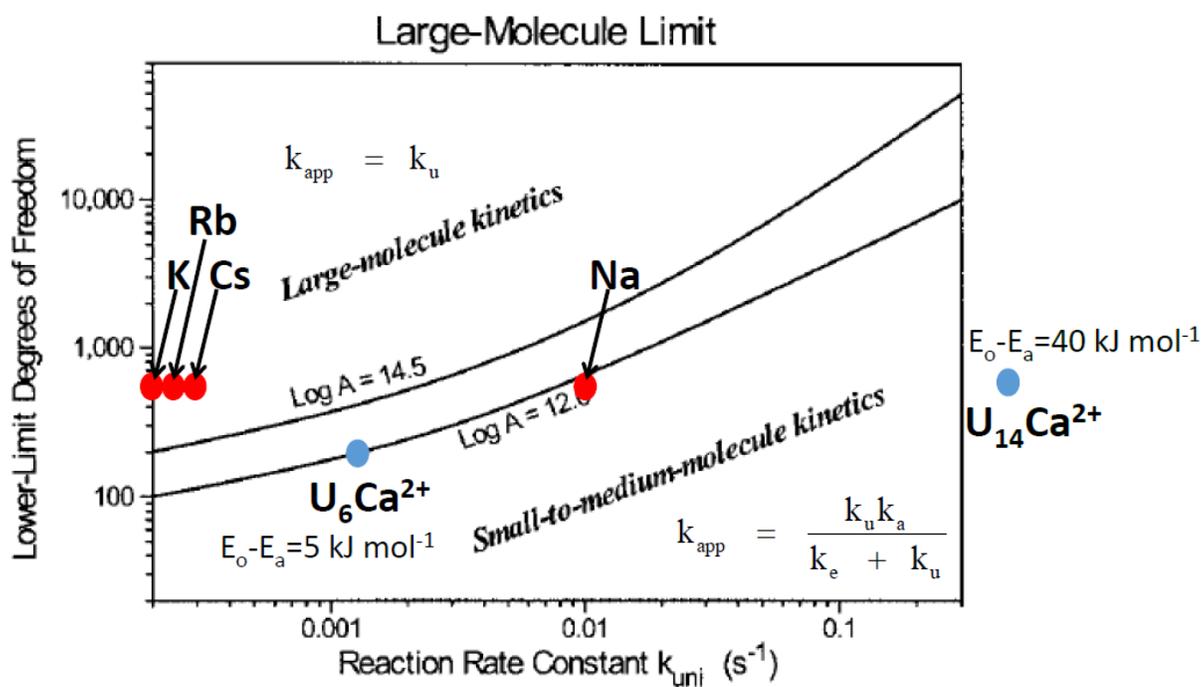


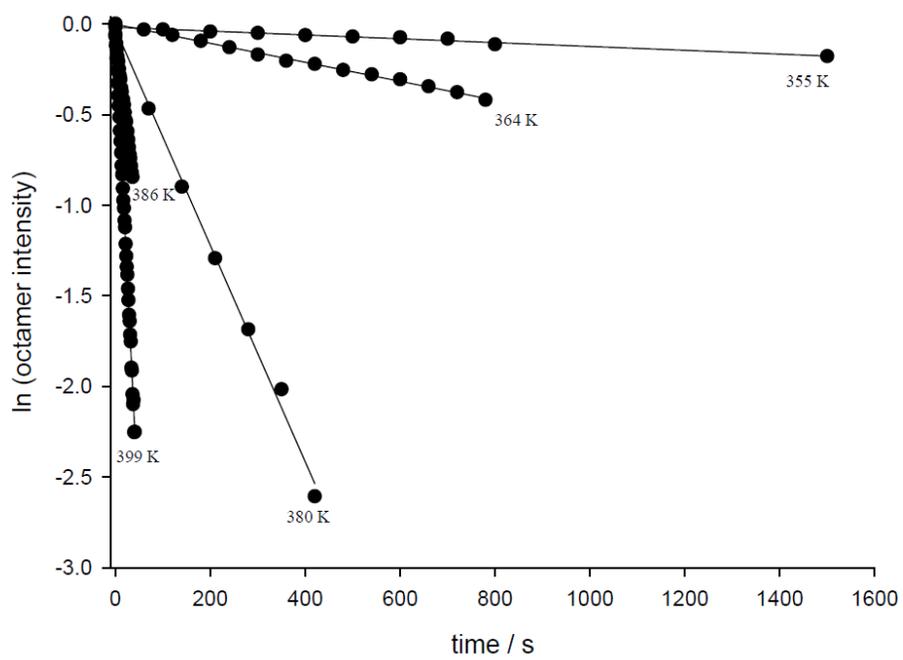
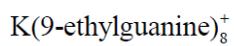
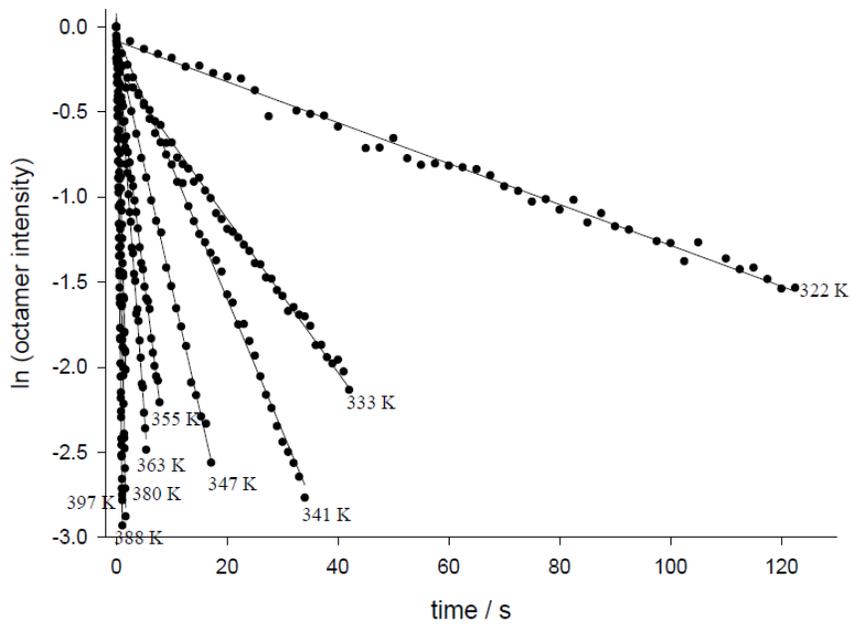
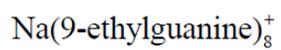
Figure S4.2. IRMPD spectrum of potassium quadruplex compared to potassium G-tetrad and computed infrared spectrum. Lowest energy structures are also depicted.



Dunbar, R.C. *Mass Spectrom. Rev.* **2004**, *23*, 127.

- E.A.L. Gillis, M. Demireva, K. Nanda, G.J.O. Beran, E. Williams, and T.D. Fridgen *Phys. Chem. Chem. Phys.* **2012**, *14*, 3304.

Figure S4.3. A diagram of molecular degrees of freedom as a function of reaction rate constants. This diagram illustrates that sodium G-quadruplex lies within the small molecule regime and its observation Arrhenius parameters are required to be modeled by master equation modeling. Larger cations, K, Rb and Cs, are categorized in the large regime.



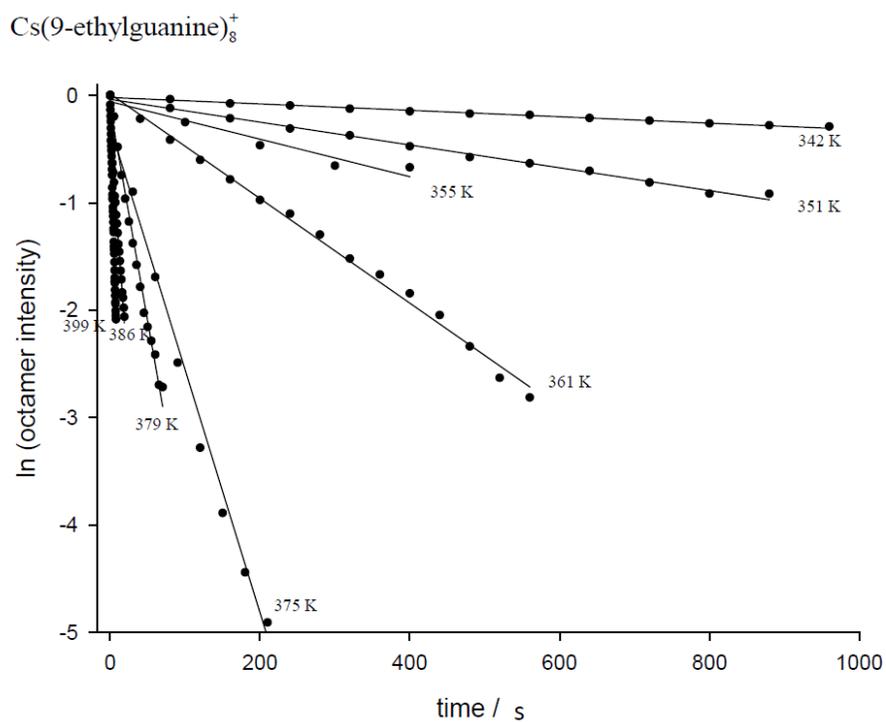
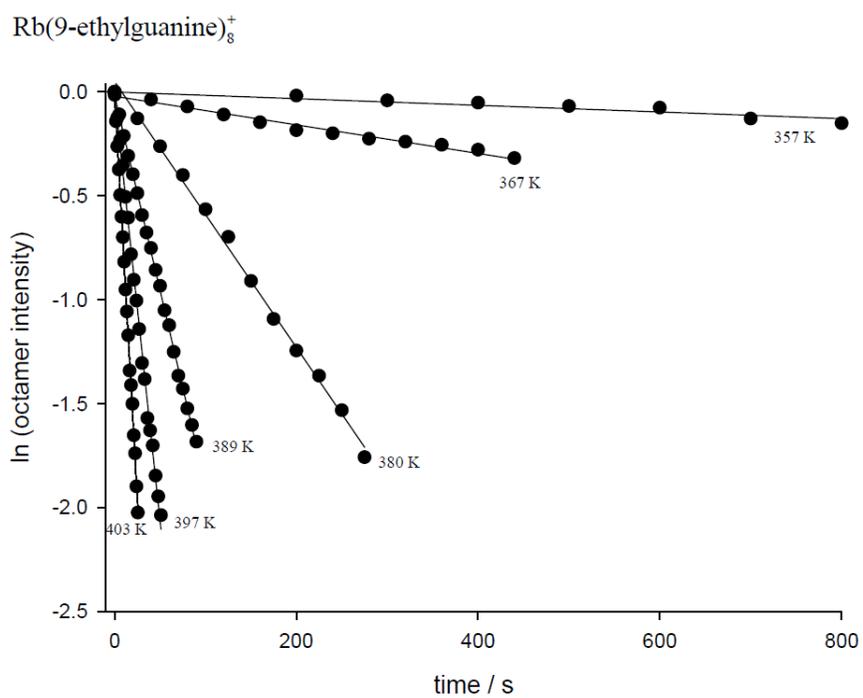
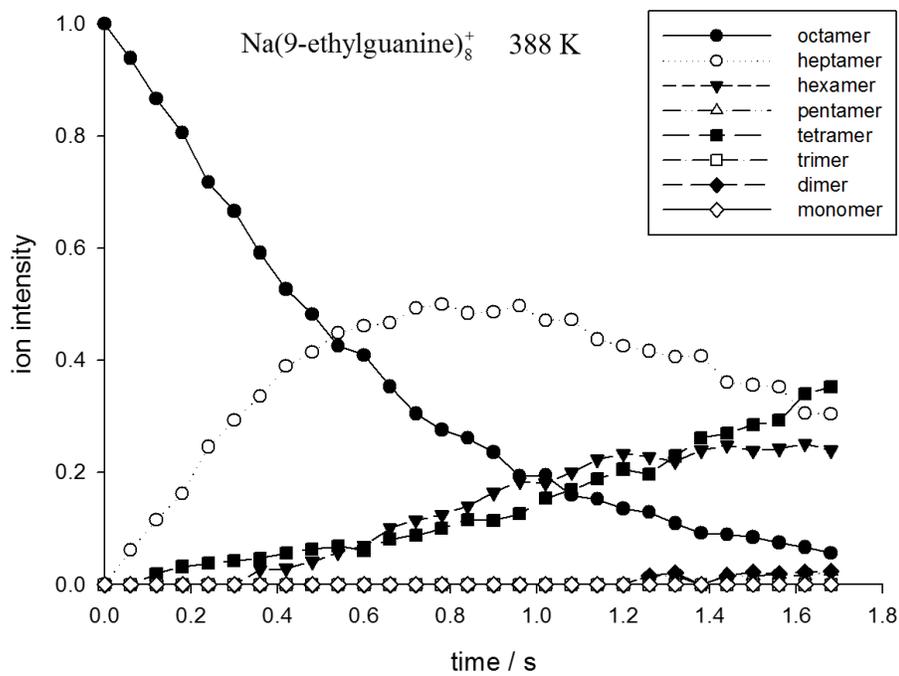
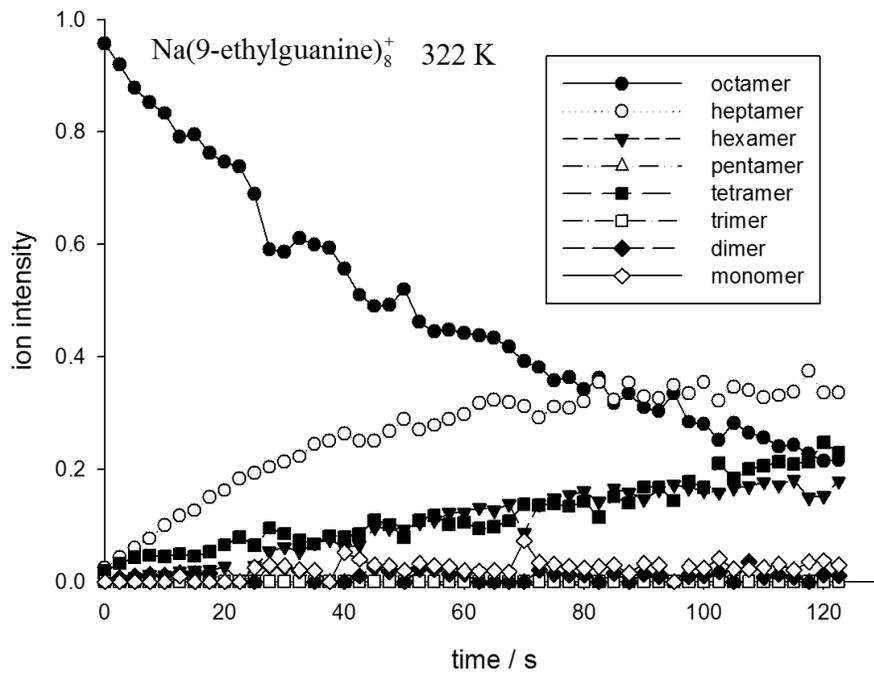
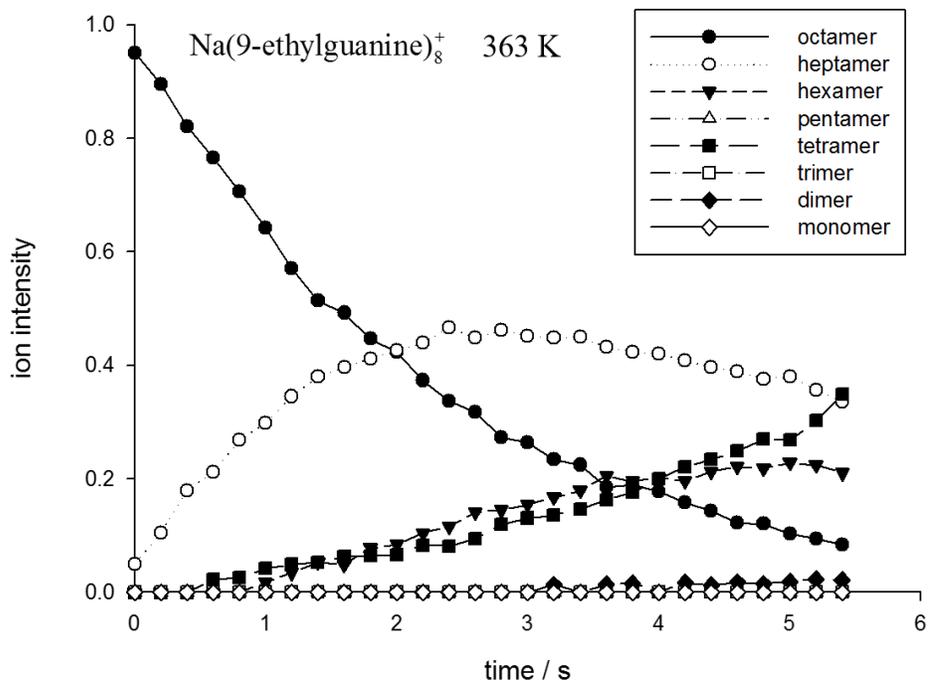
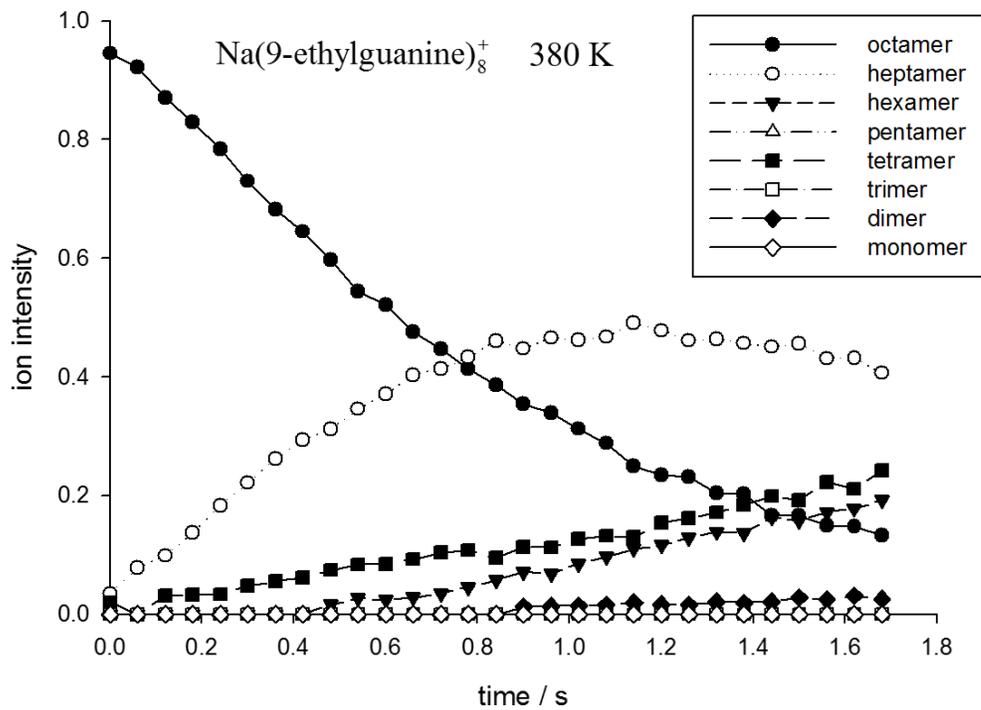


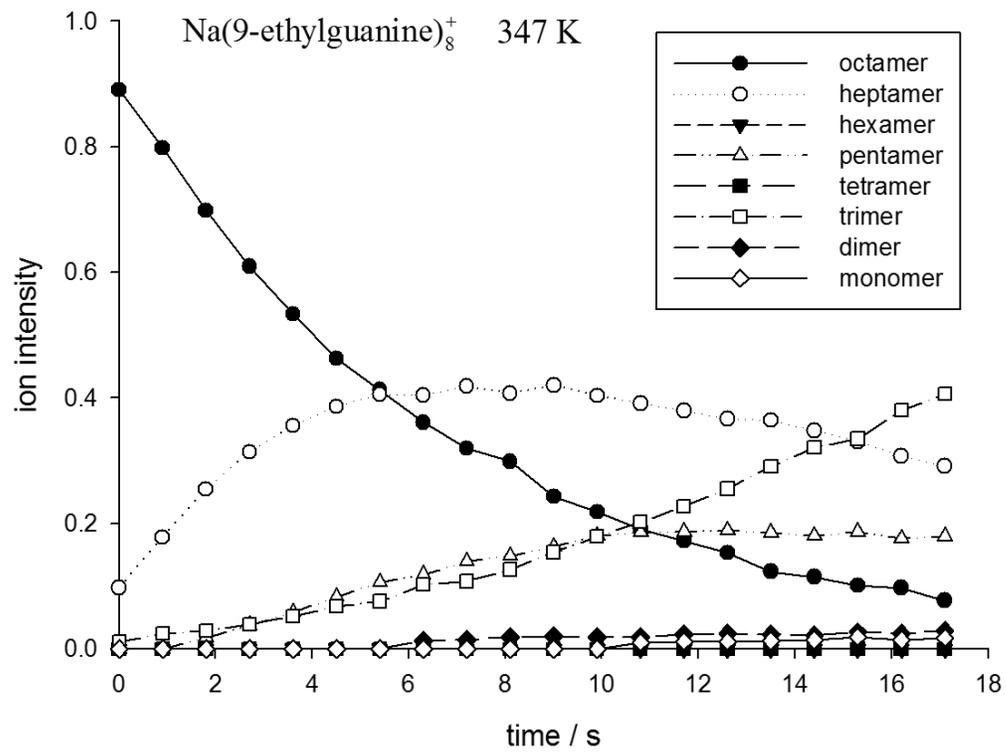
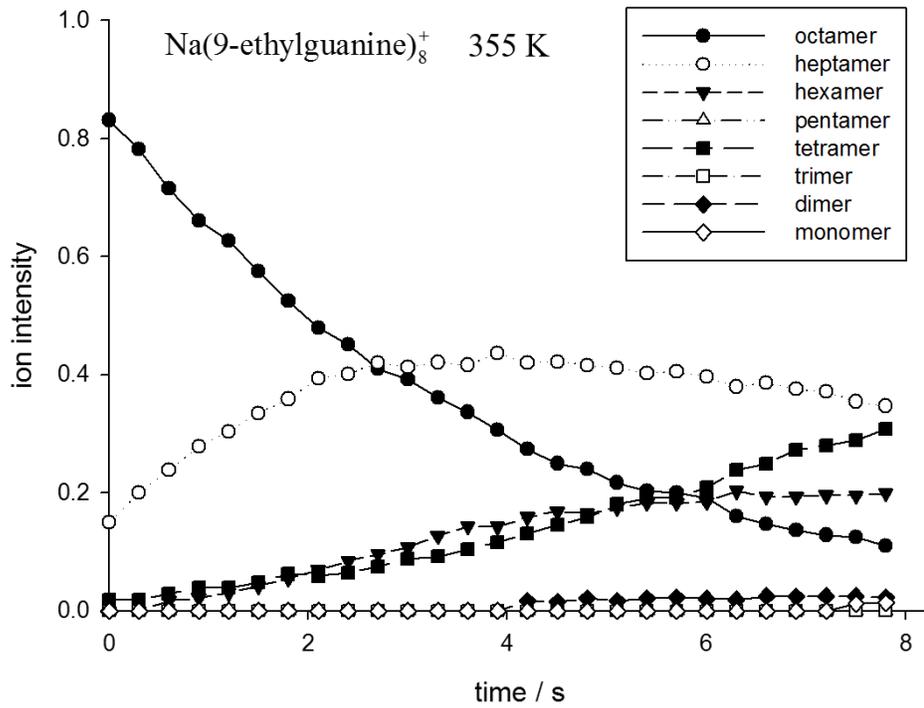
Figure S4.4. BIRD rete constants plots of alkali metal quadruplexes ($\text{M}(9\text{eG})_8^+$) at different temperatures

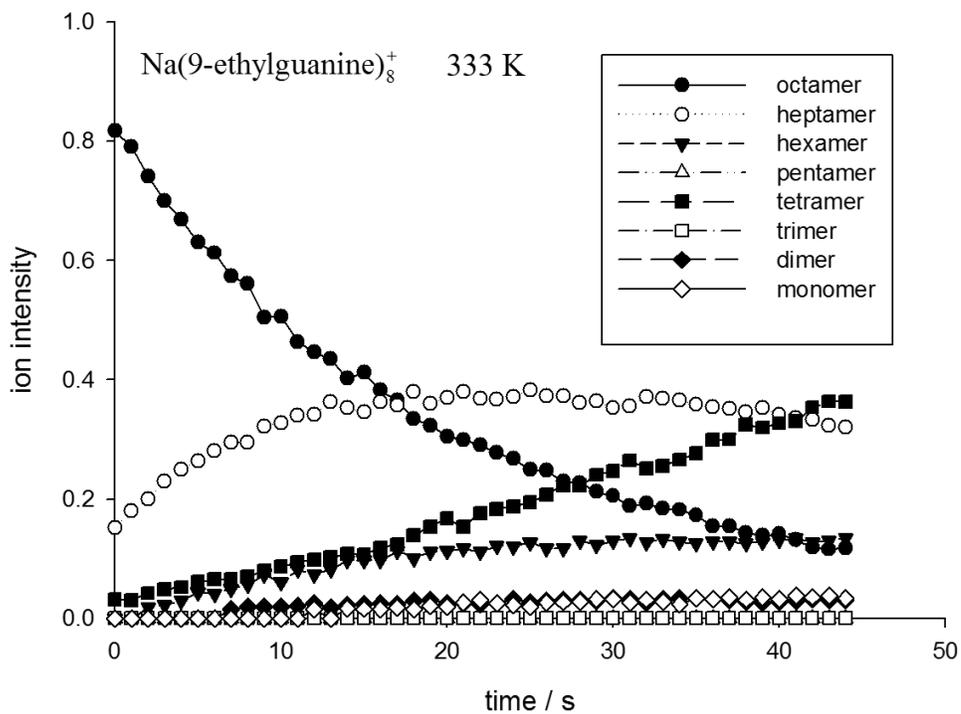
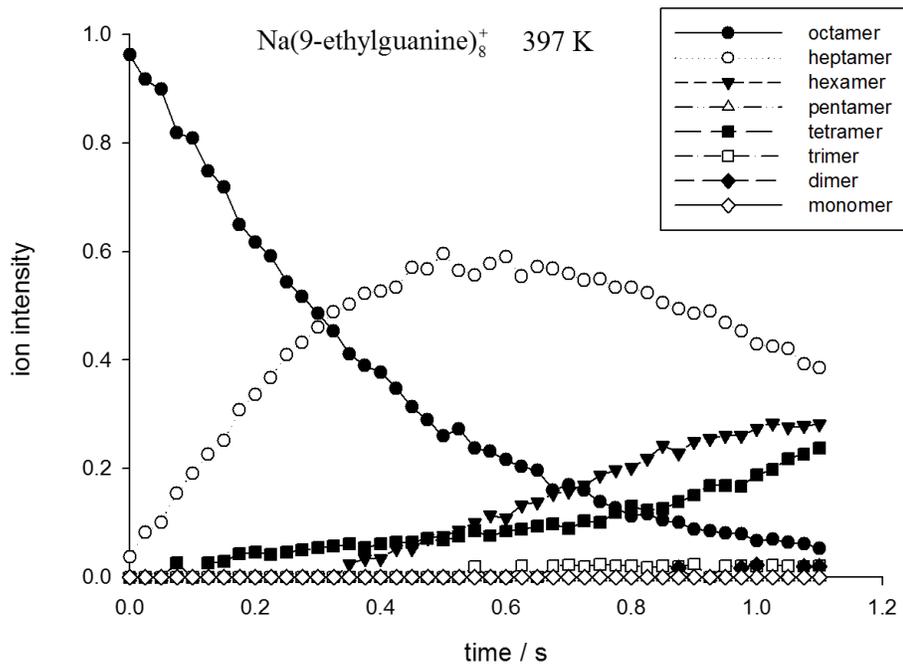
Table S4.1. BIRD rate constants of alkali metal G-quadruplexes at different temperatures

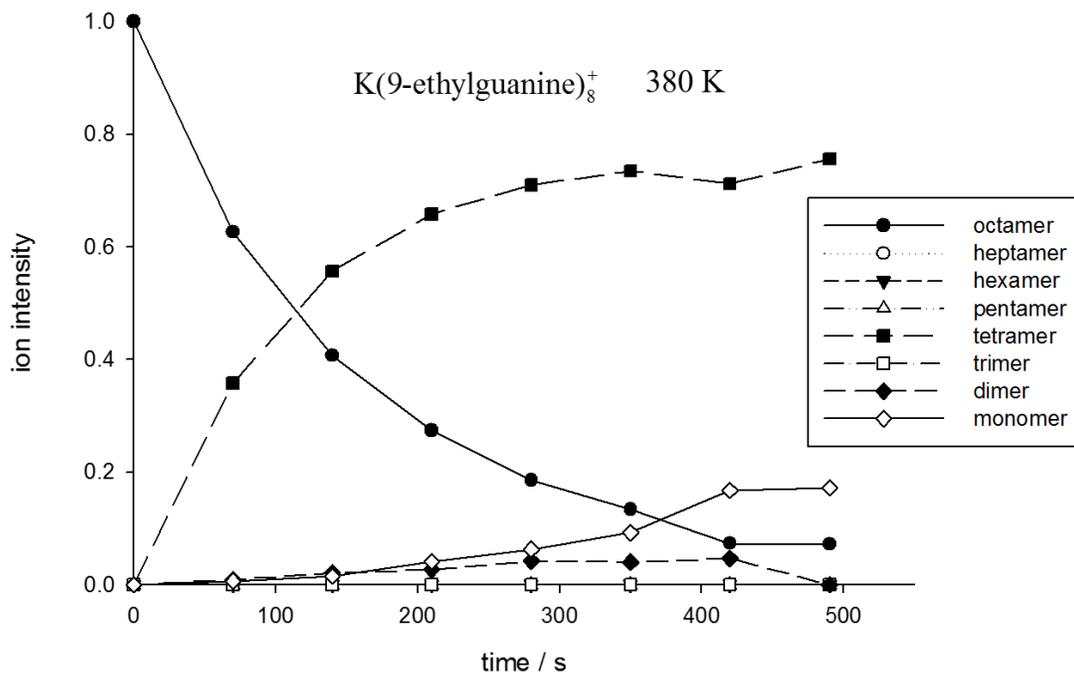
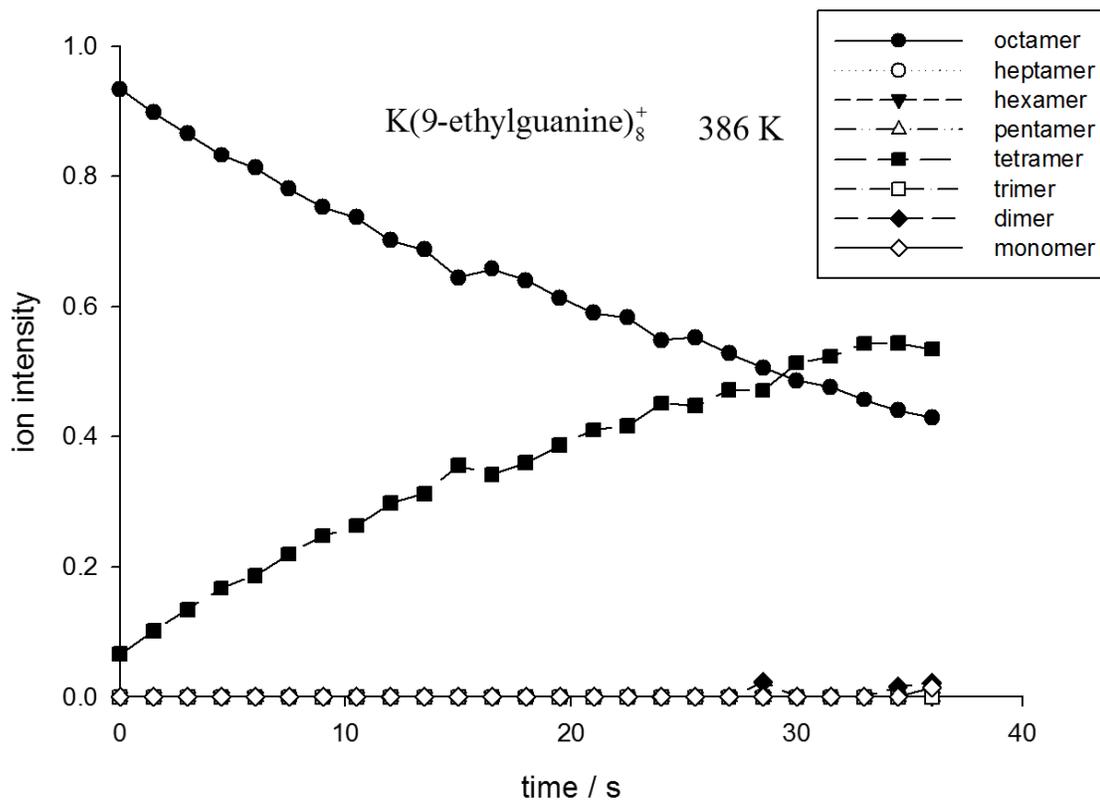
Na(9eG) ₈ ⁺	K(9eG) ₈ ⁺	Rb(9eG) ₈ ⁺	Cs(9eG) ₈ ⁺
$k_{322}=1.20(\pm 0.02)\times 10^{-2} \text{ s}^{-1}$	—	—	—
$k_{333}=4.45(\pm 0.04)\times 10^{-2} \text{ s}^{-1}$	—	—	—
$k_{341}=7.48(\pm 0.07)\times 10^{-2} \text{ s}^{-1}$	—	—	$k_{342}=1.98(\pm 0.09)\times 10^{-4} \text{ s}^{-1}$
$k_{347}=1.42(\pm 0.02)\times 10^{-1} \text{ s}^{-1}$	—	—	$k_{351}=1.06(\pm 0.03)\times 10^{-3} \text{ s}^{-1}$
$k_{355}=2.60(\pm 0.02)\times 10^{-1} \text{ s}^{-1}$	$k_{355}=1.06(\pm 0.03)\times 10^{-4} \text{ s}^{-1}$	$k_{357}=2.13(\pm 0.16)\times 10^{-4} \text{ s}^{-1}$	$k_{355}=1.63(\pm 0.22)\times 10^{-3} \text{ s}^{-1}$
$k_{363}=4.47(\pm 0.04)\times 10^{-1} \text{ s}^{-1}$	$k_{364}=5.24(\pm 0.07)\times 10^{-4} \text{ s}^{-1}$	$k_{367}=6.88(\pm 0.32)\times 10^{-4} \text{ s}^{-1}$	$k_{361}=4.87(\pm 0.10)\times 10^{-3} \text{ s}^{-1}$
$k_{380}=1.21(\pm 0.02) \text{ s}^{-1}$	$k_{380}=5.70(\pm 0.02)\times 10^{-3} \text{ s}^{-1}$	$k_{380}=6.40(\pm 0.11)\times 10^{-3} \text{ s}^{-1}$	$k_{379}=3.97(\pm 0.11)\times 10^{-2} \text{ s}^{-1}$
$k_{388}=1.73(\pm 0.02) \text{ s}^{-1}$	$k_{386}=2.10(\pm 0.03)\times 10^{-2} \text{ s}^{-1}$	$k_{389}=1.88(\pm 0.01)\times 10^{-2} \text{ s}^{-1}$	$k_{386}=9.19(\pm 0.09)\times 10^{-2} \text{ s}^{-1}$
$k_{397}=2.68(\pm 0.02) \text{ s}^{-1}$	$k_{398}=5.36(\pm 0.04)\times 10^{-2} \text{ s}^{-1}$	$k_{403}=7.77(\pm 0.06)\times 10^{-2} \text{ s}^{-1}$	$k_{399}=267(\pm 0.02)\times 10^{-1} \text{ s}^{-1}$

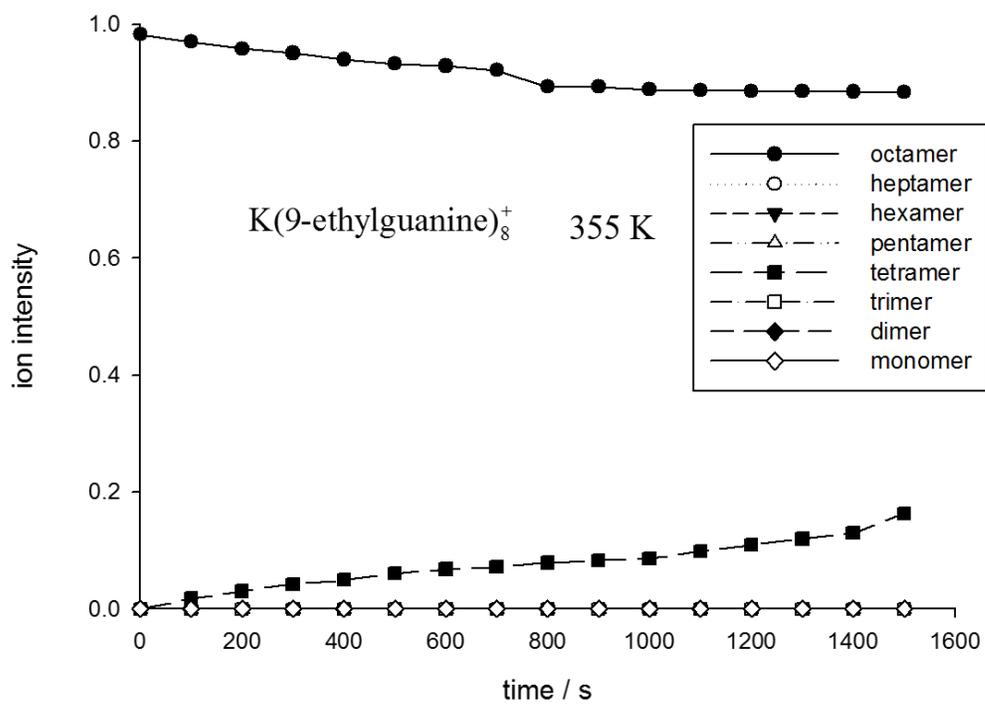
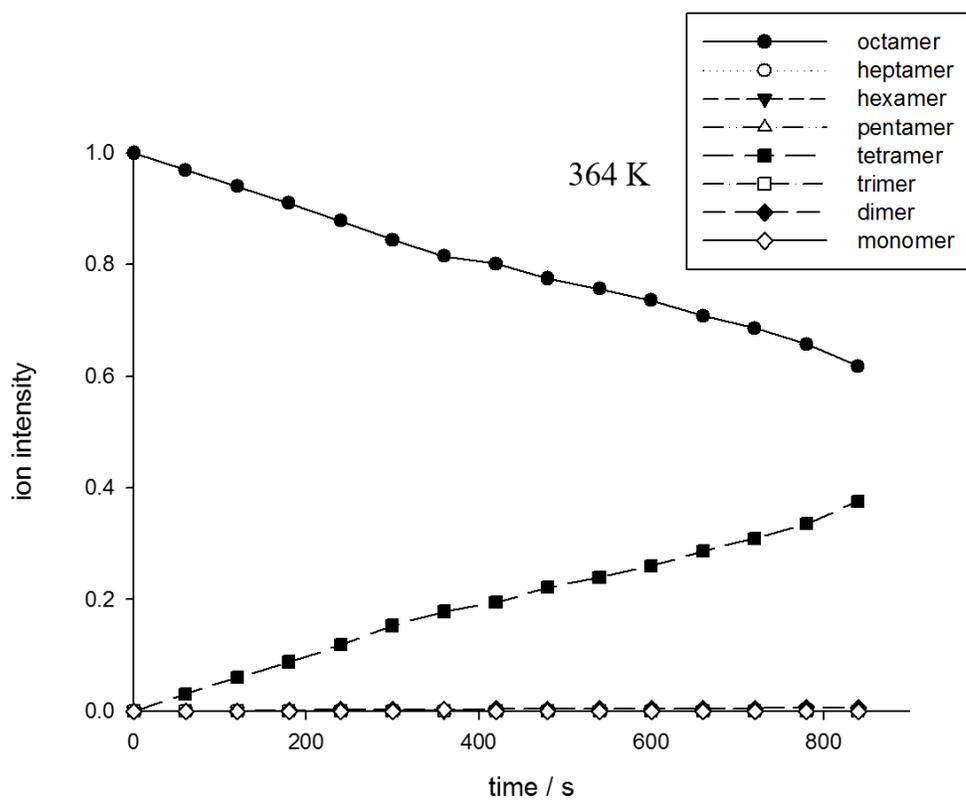


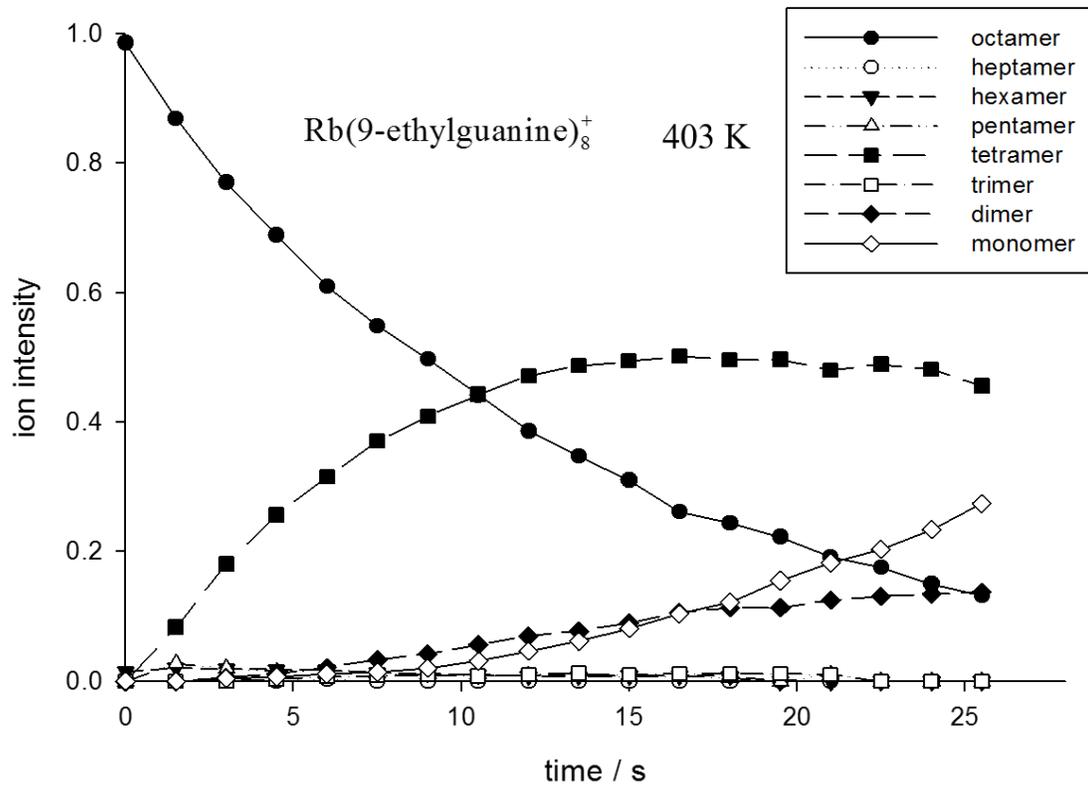
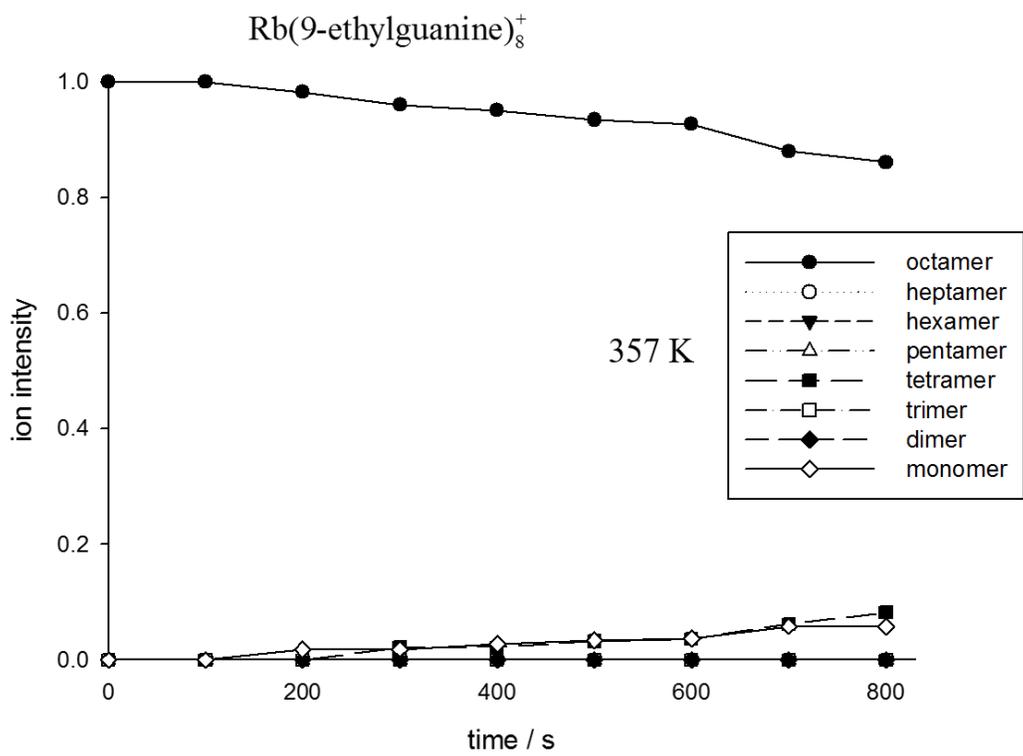


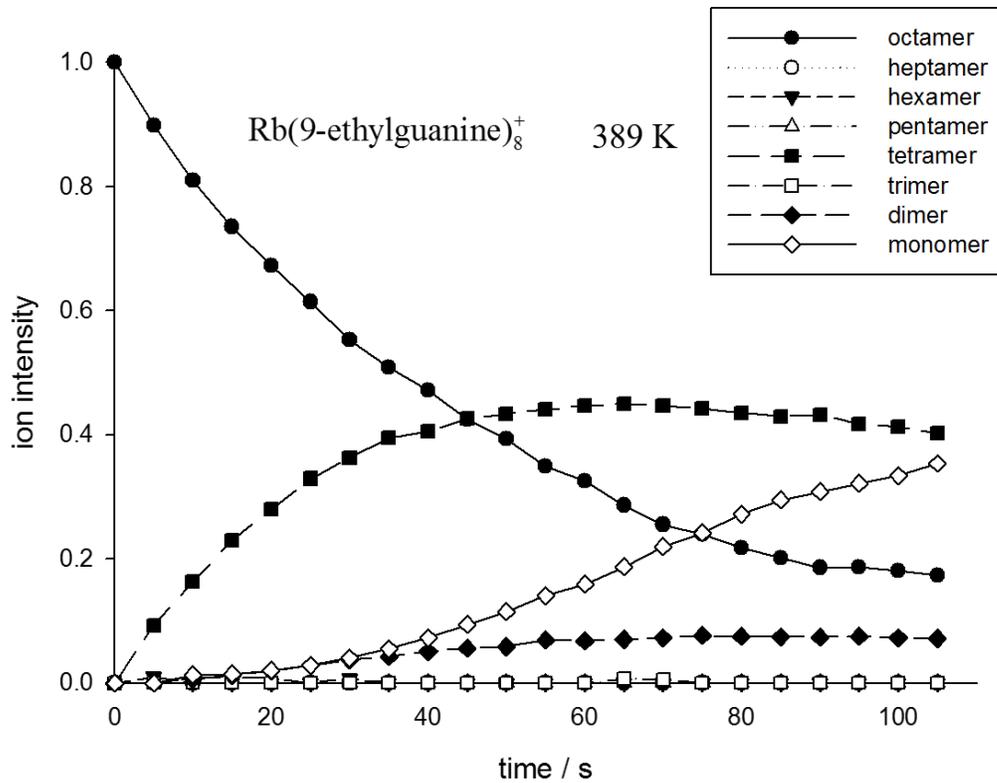
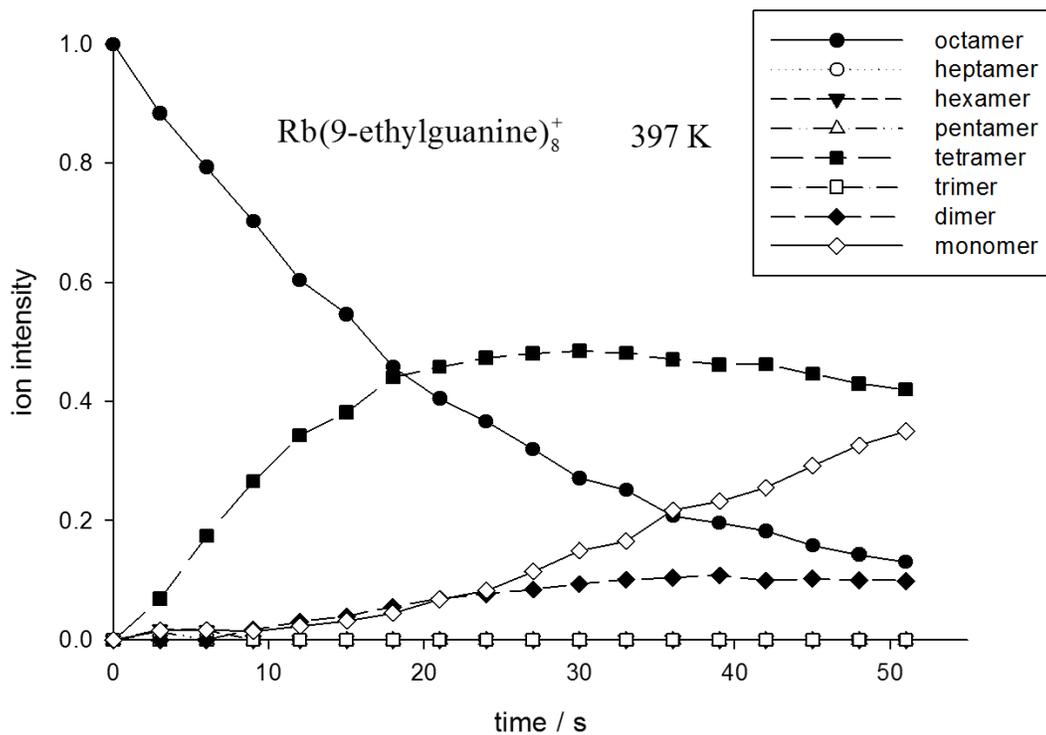


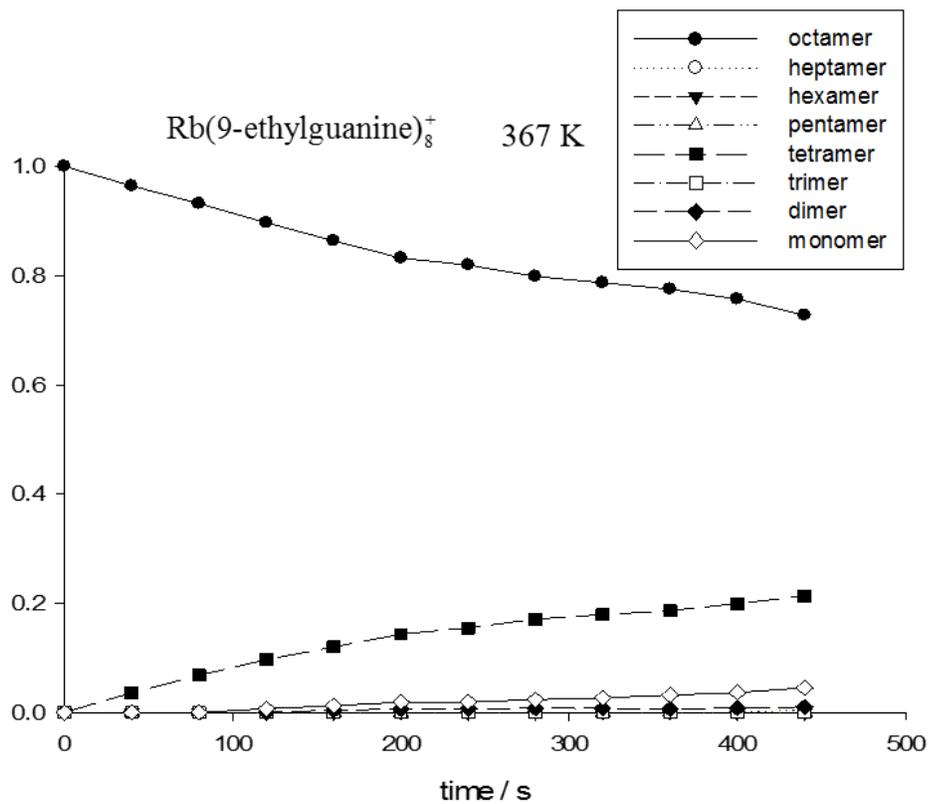
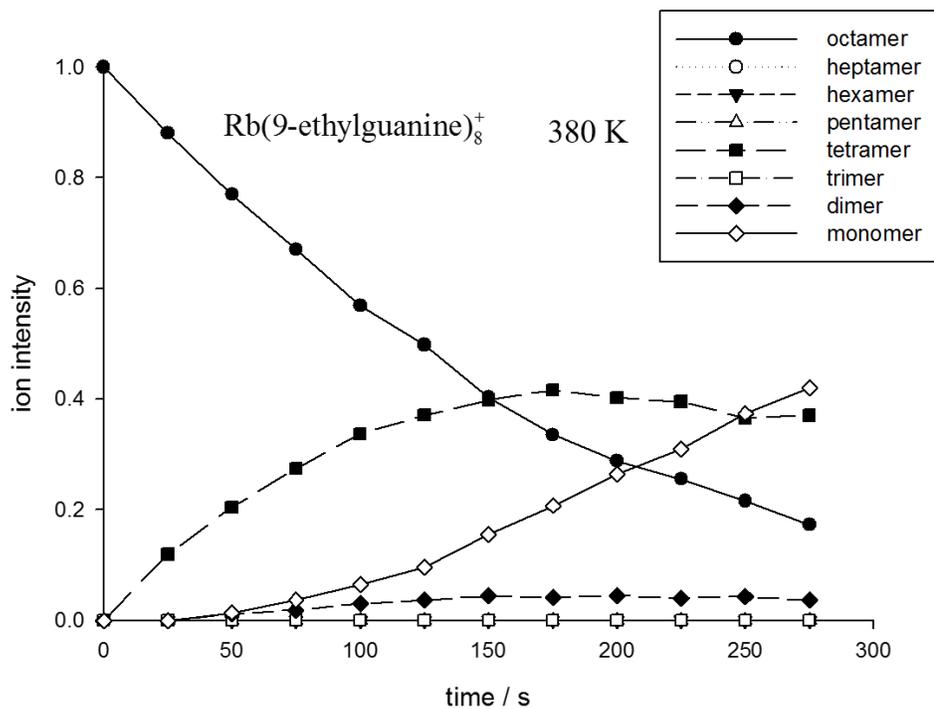


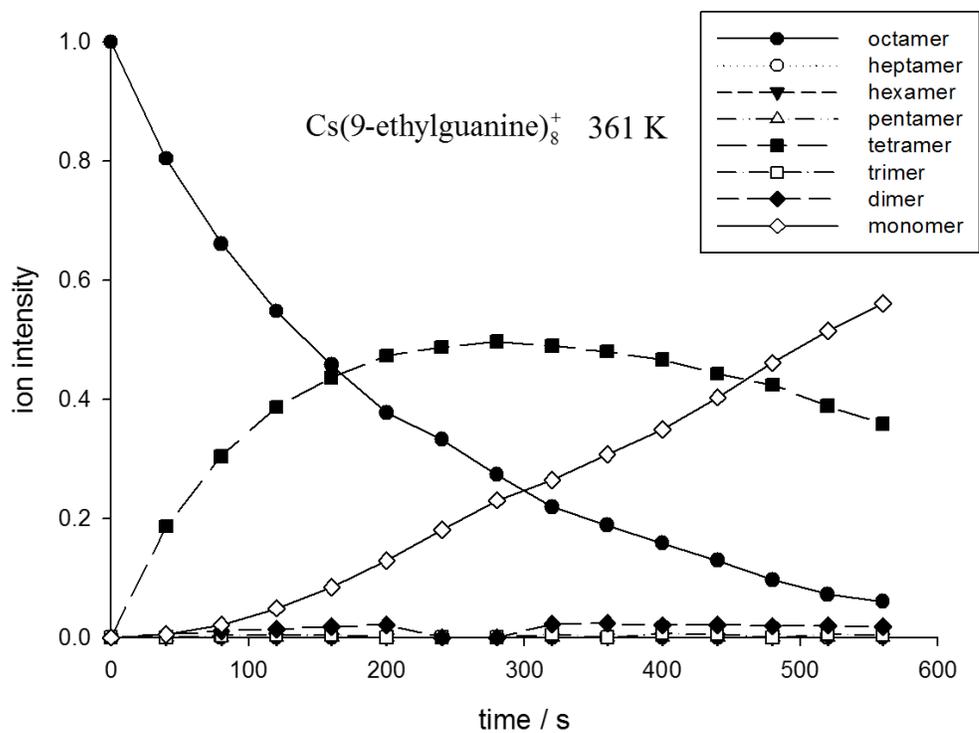
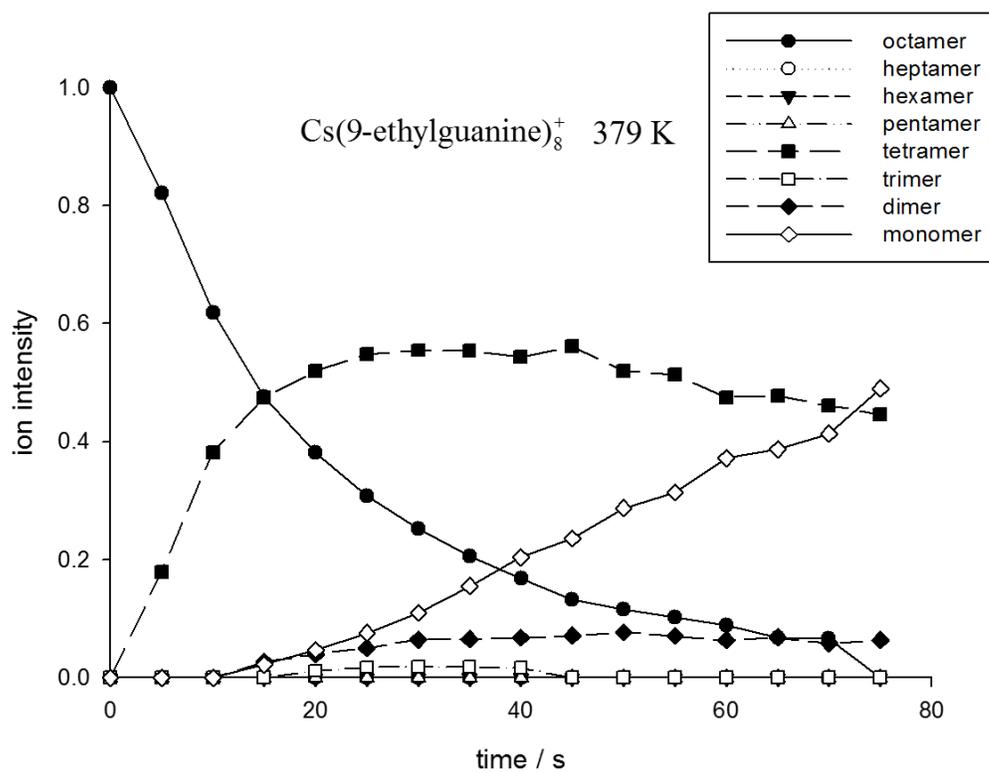


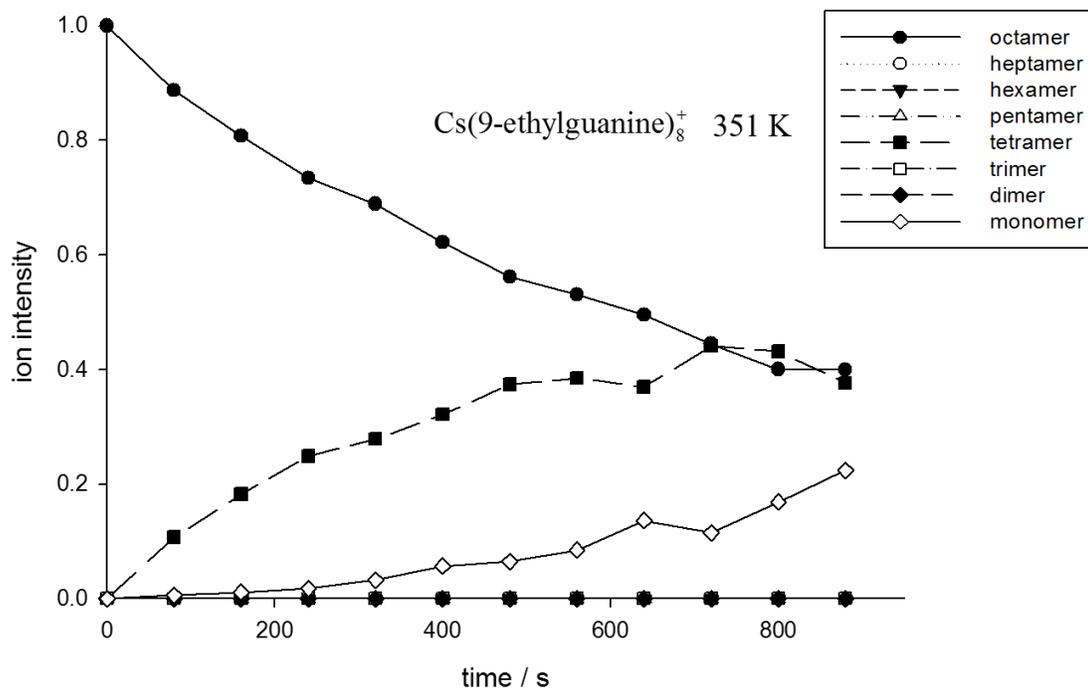
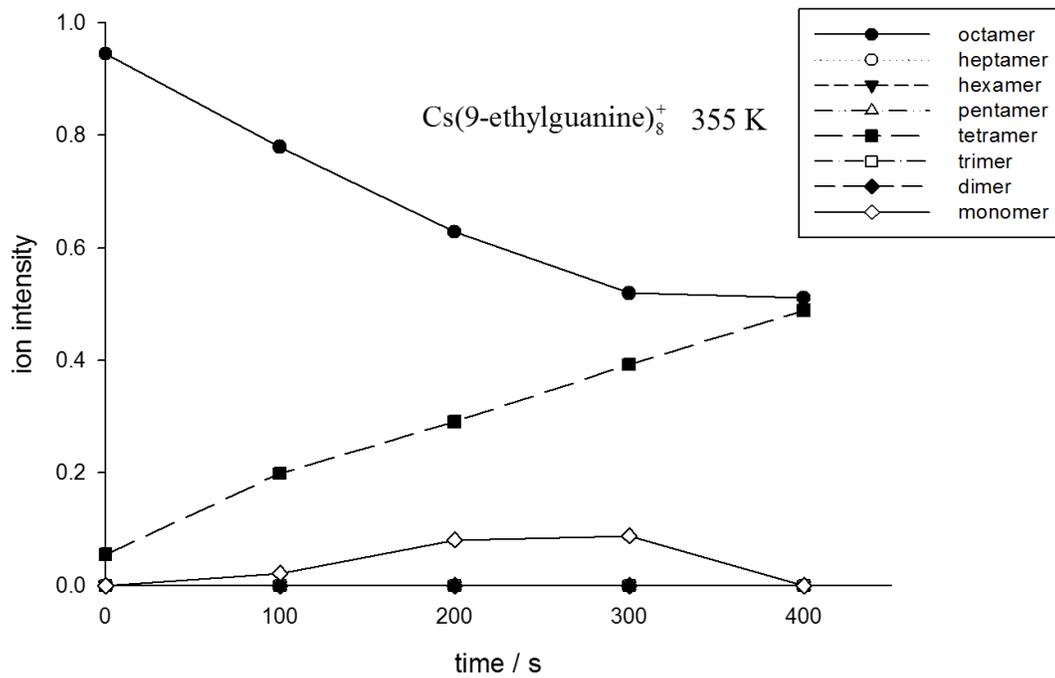


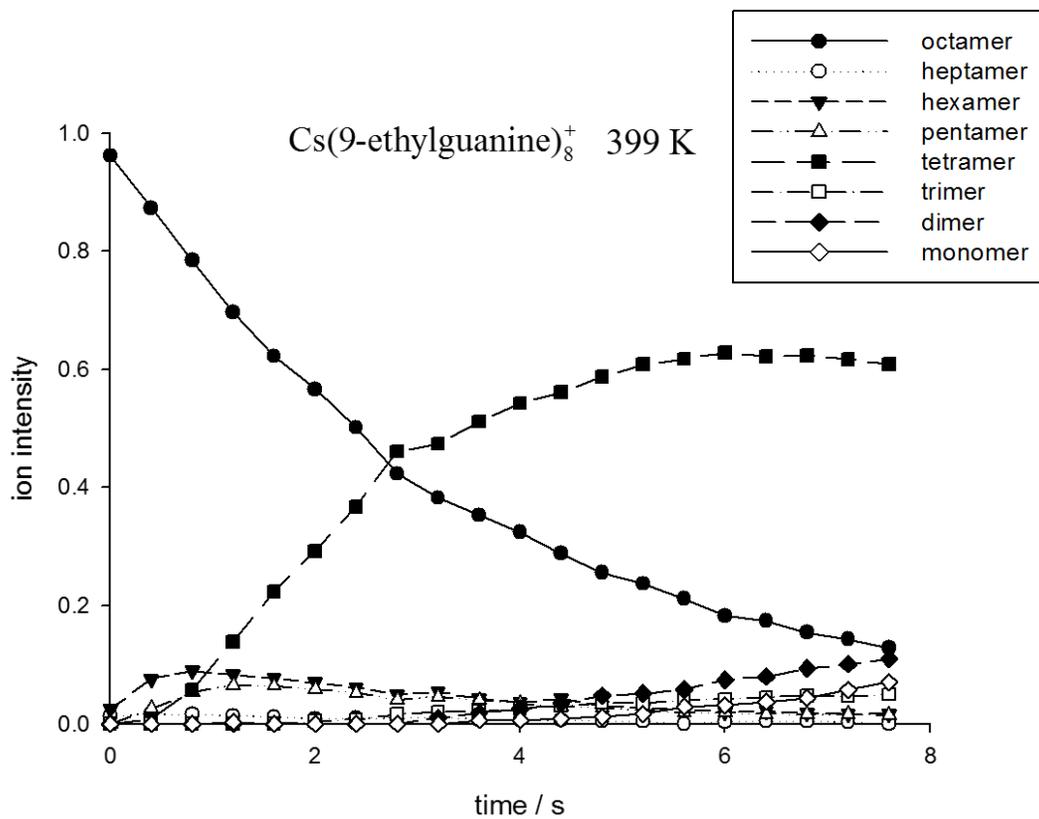
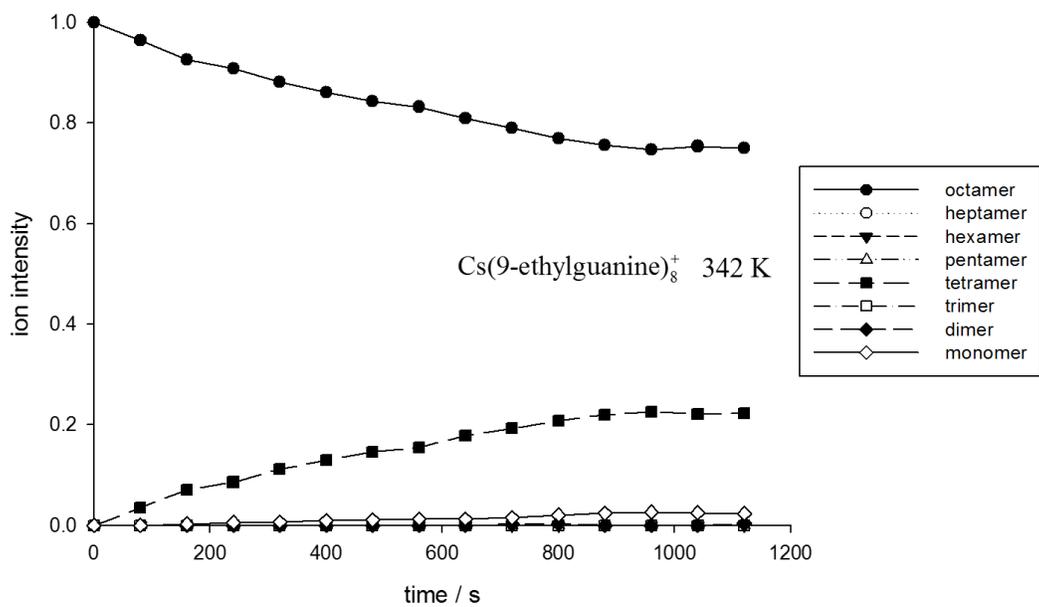












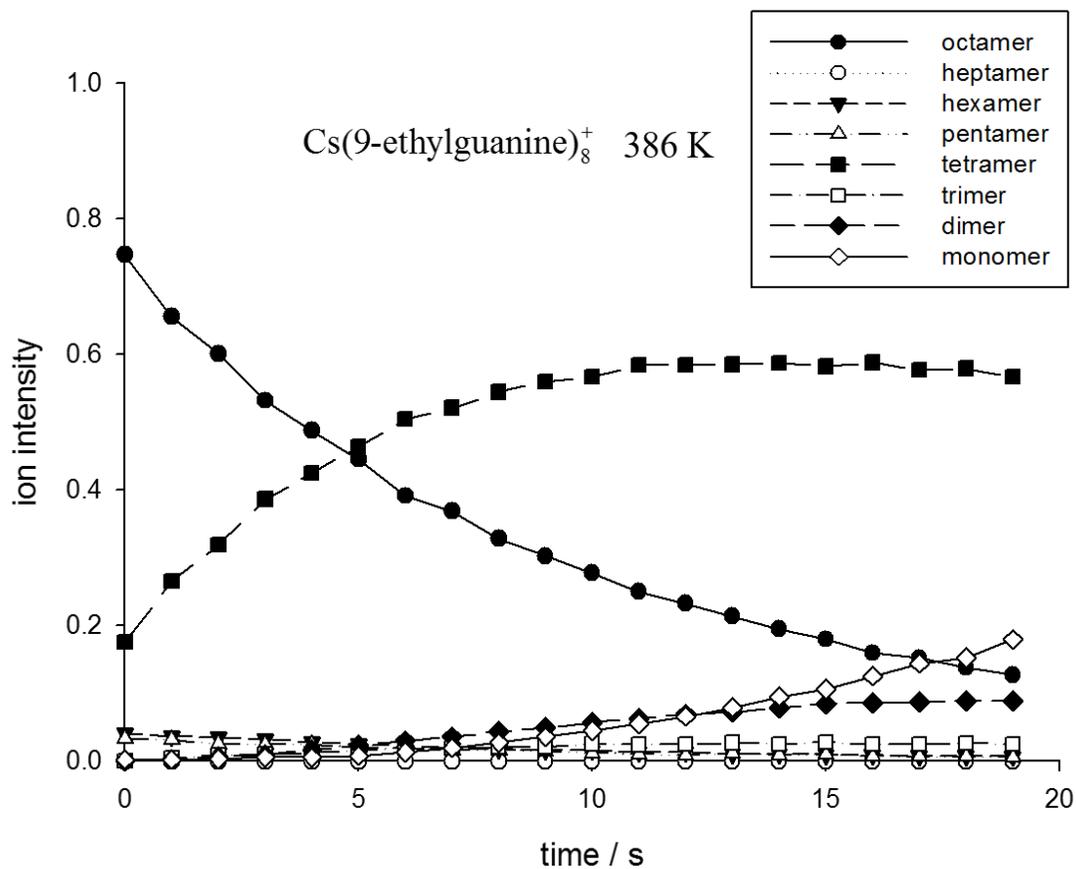
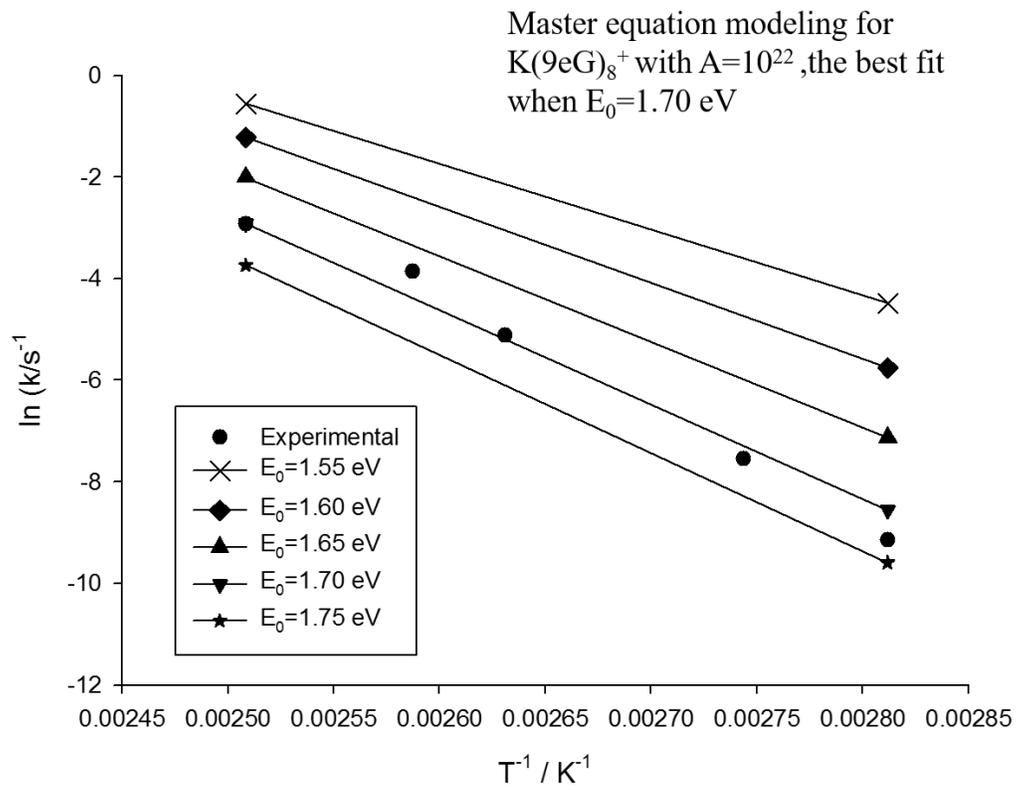
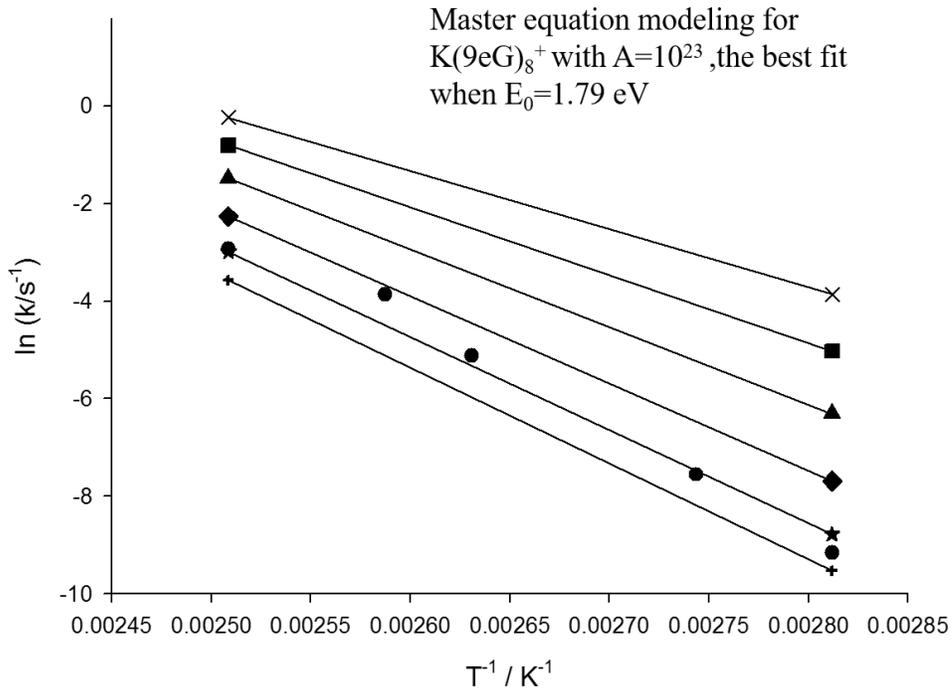
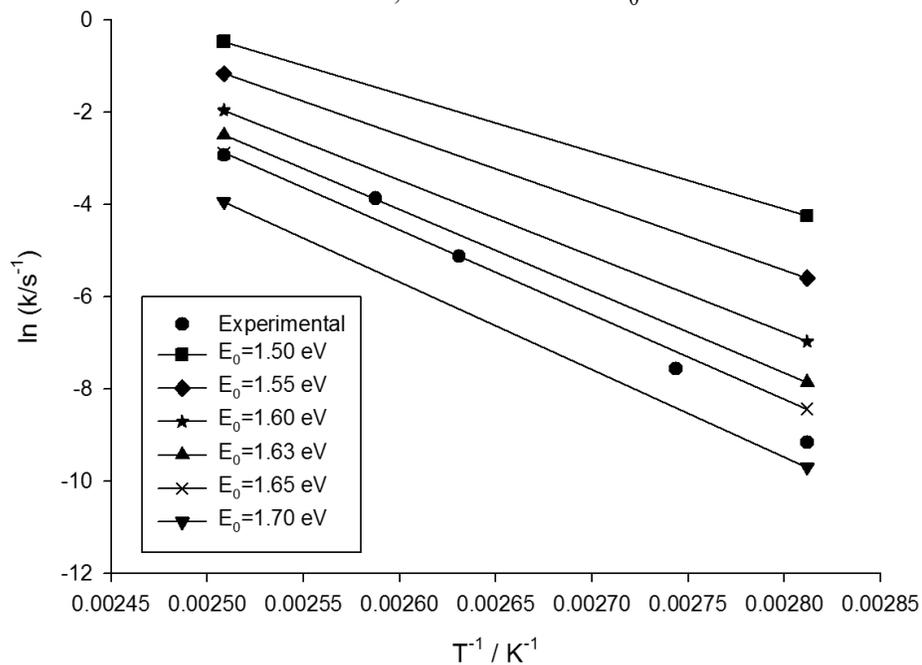


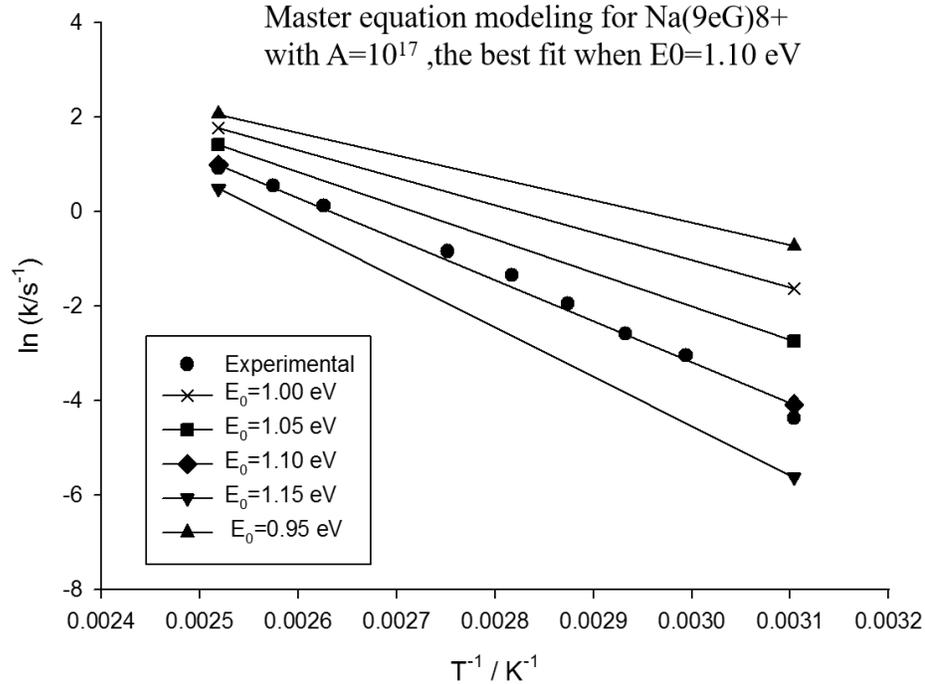
Figure S4.5. Kinetics plots for $\text{M}(9\text{eG})_8^+$ quadruplexes at different temperatures



Master equation modeling for $K(9eG)_8^+$ with $A=10^{21}$, the best fit when $E_0=1.65$ eV



Master equation modeling for $Na(9eG)_8^+$ with $A=10^{17}$, the best fit when $E_0=1.10$ eV



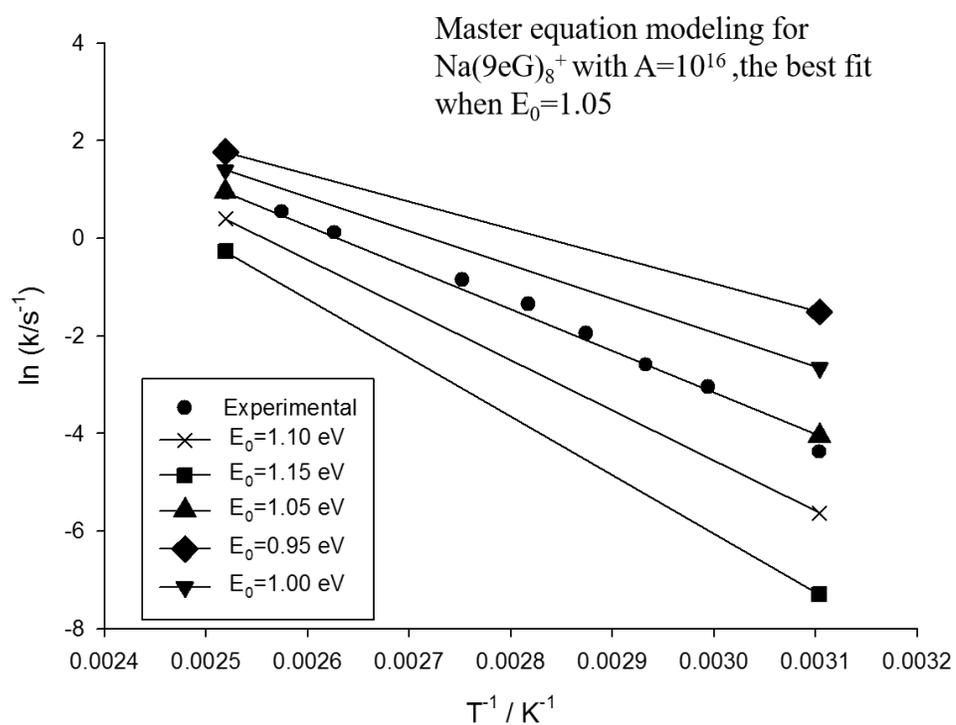
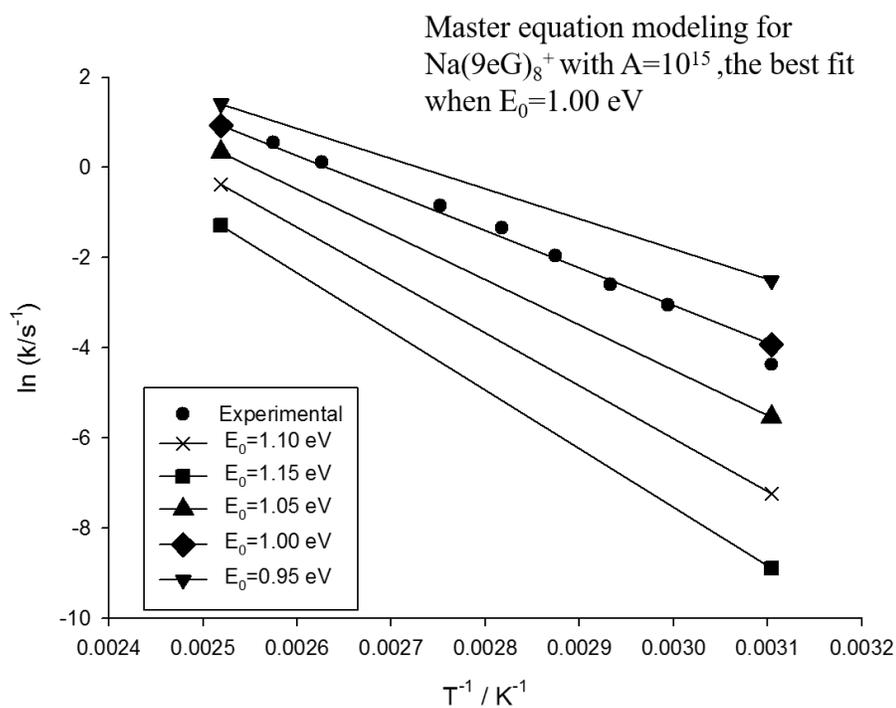


Figure S4.6. Master equation modeling results for $\text{K}(9\text{eG})_8^+$ and $\text{Na}(9\text{eG})_8^+$ quadruplex

