Structure, Energetics and Reactions of Metal Cation Complexes of Dipeptides in the Gas Phase

By

Maryam B. Moghaddam

A thesis submitted to the School of Graduate Studies

in partial fulfillment of the requirement for the degree of

Doctor of Philosophy in Chemistry

Department of Chemistry

Memorial University of Newfoundland

November 2015

St. John's Newfoundland and Labrador

Abstract

Structure, energetics and reactions of ions in the gas phase can be revealed by mass spectrometry techniques coupled to ions activation methods. Ions can gain enough energy for dissociation by absorbing IR light photons introduced by an IR laser to the mass spectrometer. Also collisions with a neutral molecule can increase the internal energy of ions and provide the dissociation threshold energy. Infrared multiple photon dissociation (IRMPD) or sustained off-resonance irradiation collision-induced dissociation (SORI-CID) methods are combined with Fourier Transform Ion Cyclotron Resonance (FT-ICR) mass spectrometers where ions can be held at low pressures for a long time. The outcome of ion activation techniques especially when it is compared to the computational methods results is of great importance since it provides useful information about the structure, thermochemistry and reactivity of ions of interest.

In this work structure, energetics and reactivity of metal cation complexes with dipeptides are investigated. Effect of metal cation size and charge as well as microsolvation on the structure of these complexes has been studied. Structures of bare and hydrated Na and Ca complexes with isomeric dipeptides AlaGly and GlyAla are characterized by means of IRMPD spectroscopy and computational methods. At the second step unimolecular dissociation reactions of singly charged and doubly charged multimetallic complexes of alkaline earth metal cations with GlyGly are examined by CID method. Also structural features of these complexes are revealed by comparing their IRMPD spectra with calculated IR spectra of possible structures. At last the unimolecular

dissociation reactions of Mn complexes are studied. IRMPD spectroscopy along with computational methods is also employed for structural elucidation of Mn complexes. In addition the ion-molecule reactions of Mn complexes with CO and water are explored in the low pressures obtained in the ICR cell.

Acknowledgments

First and foremost, I would like to express my deepest gratitude to my supervisor Dr. Travis D. Fridgen for his great support, guidance and tremendous patience through these years. I could not have imagined having a better supervisor.

I would also like to thank my supervisory committee members Dr. Raymond A. Poirier and Dr. Robert W. Davis for their valuable comments and suggestions.

I would especially like to thank all the staff at chemistry department who their help and support and heart-warming smiles made my PhD experience special.

I thank my fellow groupmates for their stimulating discussions and encouragements that help me go forward. In particular, I am grateful to Yasaman Jami-Alahmadi who helped me with the computational sections of my thesis.

Last but not the least my most sincere appreciation and love goes to my family, my beloved mother and father and my brothers who have been there for me all my life. I would also like to thank my dearest friends who showed unconditional love and support through these years.

Table of Contents

List of Figures

List of Tables

List of Schemes

List of Abbreviations and Terms

Ab initio	a Latin term for “from the beginning”
AlaGly	Alanylglycine
B3LYP	Becke-(3 parameter)-Lee-Yang-Parr
CW	continuous wave
CRM	charged residue model
CID	collision-induced dissociation
CS	charge-solvated
DFT	density functional theory
ECD	electron capture dissociation
ESI	electrospray ionization
ETD	electron transfer dissociation
FEL	free electron laser
FT	Fourier transform
FTICR	Fourier transform ion cyclotron resonance
FTICR-MS	Fourier transform ion cyclotron resonance mass spectrometry
GlyAla	glycylalanine
GlyGly	glycylglycine

ICR	ion cyclotron resonance
IEM	ion evaporation model
IR	infrared
IRMPD	infrared multiple-photon dissociation
IVR	intramolecular vibrational redistribution
KTP	potassium titanyl phosphate
Laser	light amplification by stimulated emission of radiation
LDA	local density approximation
MALDI	matrix-assisted laser desorption ionization
MS	mass spectrometry
MS/MS or MS^n	tandem mass spectrometry
Nd:YAG	neodymium-doped yttrium aluminum garnet, $Nd:Y_3Al_5O_{12}$
OPA	optical parametric amplifier
OPO	optical parametric oscillator
RF	radio frequency
SA	simulated annealing
SB	salt-bridged
SORI-CID	sustained off-resonance collision-induced
SVP	split valence polarization

Chapter 1 . Introduction

1.1. Mass spectrometry for biological systems

The study of gas phase biological ions has made it possible for scientists to access the intrinsic properties (without the complication of solvent) of biomolecules at the molecular level. In the gas phase it is even possible to probe, to some degree, environmental effects on these molecules, such as microsolvation. Nonetheless, the comparison between the solution and gas phase can lead to understanding the effect of the environment on these biomolecules.

Mass spectrometry is a direct analysis method that identifies molecular ions based on their mass to charge ratio (*m/z*) and is widely used in all areas of science In mass spectrometry, the investigation of detailed molecular structure becomes possible in the gas phase where the effect of solvent can be removed and temperature and pressure can be controlled.[1] Mass spectrometry has especially been shown to be a powerful and versatile technique with the introduction of soft ionization techniques such as electrospray ionization[2] (ESI) and matrix assisted laser desorption ionization[3,4] (MALDI). These techniques have enabled scientists to study large, intact, biomolecular ions in the gas phase devoid of solvent.[5] Since these soft ionization techniques produce molecular ions (e.g. protonated molecules) not much structural information can be derived from a single step mass spectrometry experiment. In this regard a need grew parallel to the discovery of ionization techniques for methods that can break intact molecular ions into their smaller parts to be analyzed in the mass spectrometer. Tandem mass spectrometry was developed as a versatile technique which employs ion activation methods to fragment large

molecular ions. There are several ion activation methods known so far such as infrared multiphoton dissociation (IRMPD), collision-induced dissociation (CID) and electron induced dissociation methods such as electron capture dissociation (ECD) and electron transfer dissociation (ETD). The activation techniques used in this study were IRMPD and CID and are discussed in more detail here.

IRMPD and CID are the most common activation techniques employed for structural studies of biological molecules in the gas phase. In these methods small amounts of energy are imparted into the ion and resulting in an increase in its internal energy until it surpasses the energy threshold for dissociation. IRMPD and CID differ in how the energy is imparted to the ions and the amount of energy that can be transferred upon activation. The IRMPD mechanism is based on irradiation of ions with an infrared laser which is coupled to the mass analyzer. In this case the trapped ion in the mass analyzer absorbs infrared photons and gains energy which becomes stored as internal energy until there is enough to observe decomposition. In CID, ions are accelerated and allowed to collide with an inert gas inside the mass analyzer. Upon each collision some of the ion's translational energy is converted to internal energy. The amount of internal energy gained in each collision is a function of the ion and target gas masses as well as the ion's initial kinetic energy.

The structure and energetics of metal complexes with amino acids have been extensively studied by IRMPD[6-8] and CID[9,10] methods. CID is a relatively simple ion activation technique, and has been especially beneficial in structural studies of biomolecules based on their fragmentation patterns. Protein identification and peptide sequencing, for example, have been widely studied with collision-induced dissociation

methods.[11-19] Sequences of amino acids or peptide residues can be extracted from extensive fragmentation obtained during CID experiments of enzymatically digested peptides and proteins.

Dipeptides are the smallest components of proteins that carry sequence information and have been studied as a template for peptide sequencing in proteins and peptide chains.[20] Based on nomenclature that was introduced by Roepstorff [21] and later by Biemann[22], there are three possible sites for bond dissociation in a peptide backbone; C-C, C-N (peptide bond) and N-C bond. Fragmentation of a protonated (charged) peptide at these bonds produces pairs of fragment ions, a/x, b/y and c/z, respectively (Figure 1.1). The difference in each pair comes from charge being localized at the peptide N-terminus (a, b and c) or C-terminus (x, y and z). The subscript "n" in a fragment ion (a_n for example) shows the number of residues left in that fragment ion.

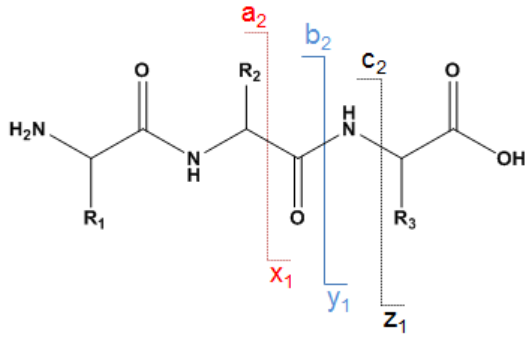


Figure 1.1. Scheme of a tripeptide fragmentation pattern based on Roepstorff nomenclature

Figure 1.2 shows the CID fragmentation mass spectrum of a doubly charged peptide (*m/z* 557) with the (Ser–Gly–Asn–Phe–Ser–Phe–Gln–Thr–Val–Lys) sequence.[23]

Main fragment ions in the 400-1100 Da mass range were observed at *m/z* 1027, 970, 865, 709, 622 and 475 which are all y type ions. For example the fragment at 1027 is y_9 which is a result of bond cleavage between Ser-Gly at the N-terminal of the peptide backbone. The mass differences of 57, 114, 147, 87 and 147 between these fragment ions from high to low corresponds to a Gly–Asn–Phe–Ser–Phe amino acid sequence in the peptide.

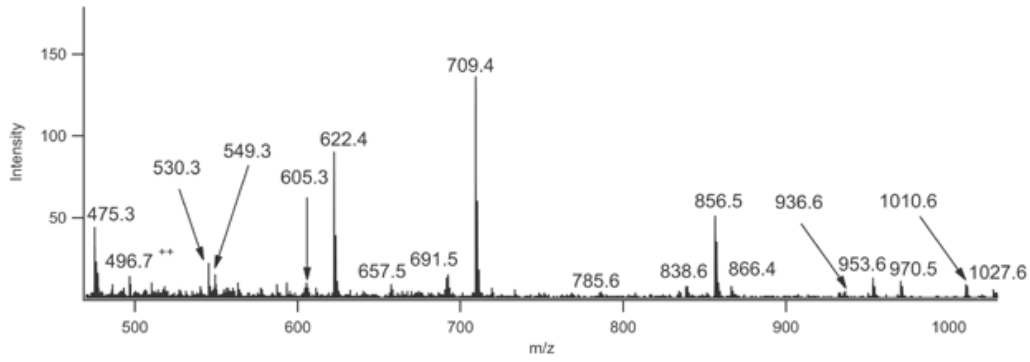


Figure 1.2. CID mass spectrum of doubly charged peptide Ser–Gly–Asn–Phe–Ser–Phe–Gln–Thr–Val–Lys with *m/z* 557. Reprinted with permission from *Mass Spectrom. Rev.* 2013, *34*, 43–63. Copyright (2013) John Wiley and Sons Ltd.[23]

1.1.1. IRMPD spectroscopy

IRMPD spectroscopy is based on the ability of a molecule to absorb multiple IR photons eventually leading to an internal energy in excess of some dissociation threshold, leading to unimolecular dissociation.[24] IRMPD spectroscopy is typically performed on ions trapped in a quadrupole ion trap or a Fourier transform ion cyclotron resonance mass spectrometer (FTICR), or some other ion trapping device. As in IR spectroscopy, in IRMPD spectroscopy vibrational modes in resonance with the IR photon frequency absorb IR photons. In the absence of collisions, vibrationally excited ions distribute the absorbed energy among other vibrational modes (intramolecular vibrational energy

redistribution or IVR) causing the absorbing mode to be depopulated and ready to absorb another photon if an intense enough laser is used. Continuous absorption of an IR photon followed by IVR increases the internal energy of the ion until it surpasses the dissociation threshold. In order to obtain an IRMPD spectrum, the light source must be tunable. CO_2 lasers were the first sources of intense IR radiation, but were of only limited tunability.[25] Currently, powerful and more broadly tunable IR sources such as free electron lasers at Free Electron Laser for Infrared eXperiments FELIX[26] and Centre Laser Infrarouge d'Orsay CLIO[27,28] and tuneable tabletop optical parametric oscillator (OPO)[24,29,30] lasers have replaced CO_2 lasers. Similar to conventional IR spectroscopy, the positions of the vibrational modes are very instructive since they provide information of possible functional groups in the structure of molecular ions. For example, an infrared absorption band above 3500 cm^{-1} is indicative of an O-H stretch. Much use is also made of computed infrared spectra and their comparison with the experimental spectrum. Structural features of many ions and ionic complexes have been determined by comparing their IRMPD efficiency spectrum with their theoretical IR spectra obtained from DFT calculations.

The study of bare and hydrated complexes of Pb^{2+} with conjugate bases of amino acids serves as a good example that shows the potential of IRMPD spectroscopy for structural elucidation of such systems.[31,32] Figure 1.3(a) compares the IRMPD spectrum and calculated IR spectra of some of the lowest energy structures of $[Pb(Leu-H)]^+$.[32] In the experimental spectrum two features were observed: one weak feature at 3410-3440 cm^{-1} and one much stronger feature at 3560 cm^{-1} corresponding to N-H stretch and carboxylic O-H stretch, respectively. The strong feature in the carboxylic acid O-H stretch region is only reproduced by the lowest energy structure Leu-A. This is in

accordance with Leu-A being deprotonated at the amine group. As is shown in Figure 1.3(b) Leu-B and Leu-C are easily ruled out because of the complete absence of carboxylic acid O-H stretch in their spectra. The computed spectrum for Leu-A is most consistent with the experimental IRMPD spectrum.

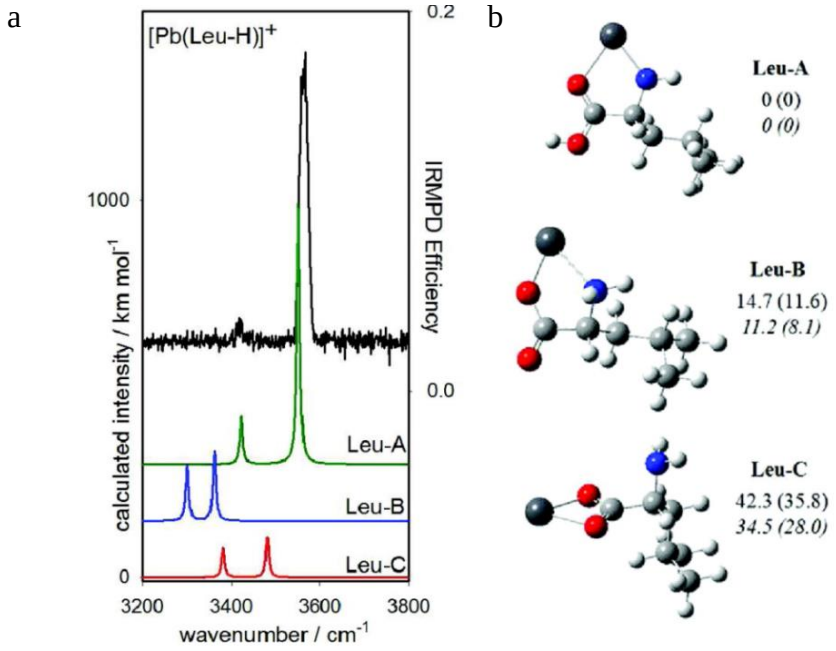


Figure 1.3. IRMPD and calculated IR spectra of $[Pb(Leu-H)]^+$ (a) and corresponding calculated structures A, B and C(b). The enthalpy and (Gibbs energies) are calculated at 298K using B3LYP/6-31+G(d,p) (nonitalicized) and MP2(full)/6-311++G(2d,2p)//B3LYP/6-31+G(d,p) (italicized) with LANL2DZ on Pb in both cases. Reprinted with permission from *J. Phys. Chem. B* **2011**, *115*, 11506-11518. Copyright (2011) American Chemical Society.[32]

1.1.2. CID

Collision induced dissociation (CID) relies on the inelastic collision of translationally excited ions with an inert molecule, normally argon or helium. Upon

collision of an ion with an inert gas molecule, a portion of the ion's kinetic energy is converted into the internal energy of the ion. The maximum translational energy of ions E_{com} that is convertible to internal energy after each collision is a function of kinetic energy of ions E_{lab} and masses of target gas (N) and the parent ion (m_p) as is shown in Equation 1.1:

$$E_{com} = E_{lab}\left(\frac{N}{m_p+N}\right) \qquad 1.1$$

If this newly attained internal energy is high enough, the parent ion will dissociate into fragment ions which are detected by the mass analyzer.[33,34] CID methods are categorized into three types based on the collision energy and activation time scale; high energy, low energy and very low energy.

In high energy CID, one (or very few) collisions occur with keV collision energy deposited on ions in one to ten microseconds time scale. This method is mainly carried out by mass analyzers such as time of flight (TOF) and in sector instruments. Higher energy fragmentation channels such as covalent bond dissociation, as well as secondary fragmentations are more possible in high energy CID. Examples of dissociation of peptides under high energy CID are side chain dissociation along with peptide bond cleavage.

In low energy CID around 1-100 eV collision energy becomes deposited through tens to hundreds of collisions and ion activation occurs in a few hundred microseconds to a few milliseconds. Low energy CID is typically carried out in triple quadrupole mass spectrometers and the fragmentation pathways are limited to the lowest energy

fragmentation pathways. The mass spectra obtained in low energy CID are less complex since high energy side chain cleavages are not observed as they are in high energy CID.

High energy and low energy CID mass spectra of a singly charged decapeptide (ACTH) are compared in Figure 1.4.[35] At high collision energy, dissociation occurs at almost all peptide backbone bonds and produces a, b, y and z type fragment ions (Fig 1.4(A)). More interestingly, higher energy fragmentations related to dissociation at the side chain of the amino acid (where the peptide backbone is fragmented) appears in the mass spectrum. For example formation of w (C-terminal) and d (N-terminal) fragment ions are stemmed from side chain C_β-C_γ (C_α being the carbon at the peptide backbone at each amino acid residue) bond cleavage and loss of the whole side chain produces v (C-terminal) type fragment ions. In contrast, in low energy CID the main dissociation pathway occurs at the peptide bond forming b and y fragment ions. Also loss of NH_3 and H_2O are other fragmentations that occur in lower abundances in low energy CID (Fig 1.4(B).

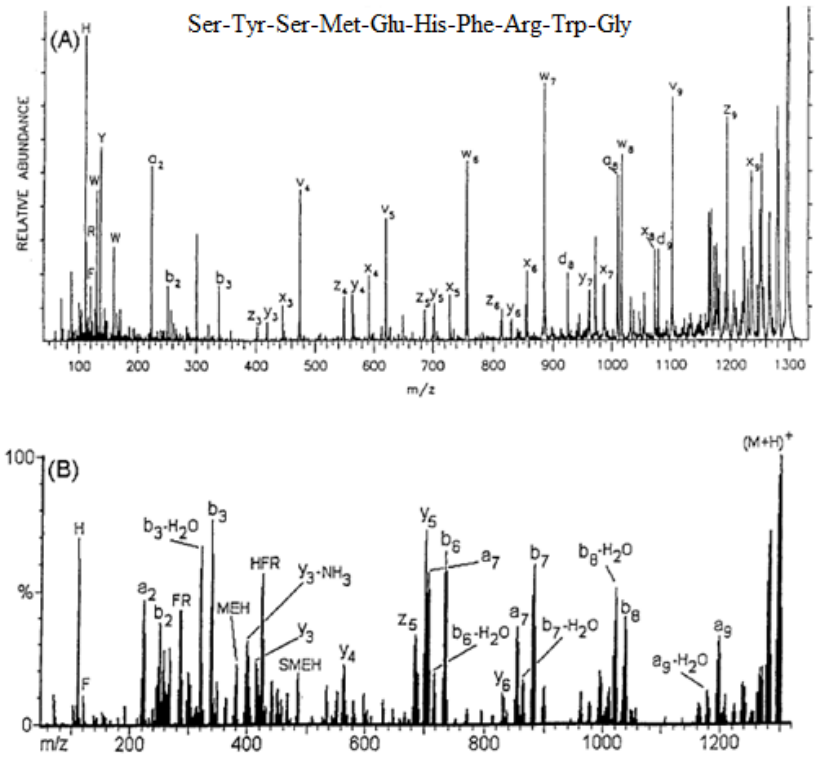


Figure 1.4. CID mass spectra of protonated decapeptide ACTH (singly charged) at 7 keV (high energy) acquired with a four-sector magnetic reflector mass analyzer with helium as collision gas (A) and 45 eV obtained by a triple quadrupole with argon as the collision gas (B). Reprinted with permission from *Mass Spectrom. Rev.* **2005**, *14*, 49-73. Copyright (2005) John Wiley and Sons Ltd. [35]

Very low collision energies may be attained in ion trapping devices such as the ICR and quadrupole ion trap. In these instruments only a few eV collision energies may be applied and hundreds of collisions are required and the ion activation time scale is much longer, hundreds of milliseconds to seconds. This method is sometimes referred to as a slow heating method. Very low energy CID is beneficial for study of weakly bound moieties such as non-covalent complexes, with very low energy fragmentation pathways.

Different fragmentation outcomes for each CID type is advantageous in structural studies since a wide range of information can be obtained with different collision energies.

1.1.3. Electron induced dissociation

There are mainly two types of electron induced dissociation techniques; electron capture dissociation (ECD) and electron transfer dissociation (ETD). In these techniques multiply positively charged ions become activated by site selective absorption of an electron that is produced by a hot filament in ECD or carried by an anion in ETD. After each electron absorption, charged ions form radical ion intermediates which are unstable and dissociate into fragment ions. Also, because of addition of an electron upon ion activation, the charge state of the parent ion differs from fragment ions. Compared to IRMPD and low energy CID methods, which mainly produce b and y fragment ions from dissociation of labile peptide bonds, electron based dissociation techniques do not necessarily follow the lowest energy dissociation pathway.[36] Site selectivity of these methods are especially beneficial in the study of peptides and proteins with labile post-translational modifications. For example in a phosphorylated peptide, loss of H_3PO_4 (98 Da) is the main dissociation pathway observed under low energy CID condition whereas in ETD, only peptide backbone fragmentation is observed showing that the phosphoryl group in the fragment ions is maintained.[37]

Besides fragmentation methods, ion-molecule reactions studies in the mass spectrometer can provide structural and thermochemical information complementary to

results of dissociation methods. Detailed discussion about electron-induced dissociation and ion-molecule reactions are beyond the scope of this study.

1.2. Metals in Biology

Metal ions are essential parts of all living systems since they play important structural and functional roles in biological molecules.[38-41] Large numbers of proteins are metal-binding proteins in which metal ions are the active centers or participate in stabilization of structural motifs that are responsible for biological processes inside living organisms. Metal ions can interact with charged and polar sites of proteins and nucleic acids and form Lewis acid/base pairs in the structure of these molecules. In biosystems, metal ions of alkali and alkaline earth metal groups such as Na^+, K^+, Mg^{2+} and Ca^{2+} can exist in substantial quantities. These metals are known to play an important part in many biological processes. For example, Na^+ and K^+ keep osmotic equilibrium by forming an ion gradient on both sides of cell membranes. These metal ions are necessary for nerve and muscle function. Also Na^+, K^+ and Ca^{2+} metal ions participate in intra- and extracellular signal transduction. Living cells use any increase in extracellular Ca^{2+} concentration as a signal which causes the opening of the membrane to import Ca^{2+} ions into the cell. Calmodulin is a Ca^{2+} binding EF-hand protein which participates in the activation of many cellular processes such as muscle contraction and neurotransmission. In the structure of calmodulin there are four metal binding sites available for Ca^{2+}. Binding of Ca^{2+} can induce conformational changes that become transferred to the target proteins.

Transition metals mostly exist in trace amounts in the human body but are extremely important nonetheless. Metal ions such as Fe^{2+}, Mn^{2+}, Zn^{2+} and Cu^{2+} are known for their role in many catalytic reactions and redox reactions. Fe^{2+} and Cu^{2+} are the electron transfer centers in the redox reactions in haemoglobin and hemocyanin, respectively. Transition metals such as Zn^{2+} as well as alkaline earth metals Mg^{2+} and Ca^{2+} are also known for their role in protein structure stabilization. Zinc fingers are good examples of metal mediated protein folding.[39]

1.2.1. Metal cation role in the structure of amino acids and peptides

Amino acids and peptides are the smallest building blocks of proteins. Whether individually or as parts of proteins, amino acids and peptides can serve as ligands which directly coordinate metal ions in metalloproteins/metalloenzymes. Functionalities such as the terminal amine, amide, and carboxylic groups, as well as heteroatomic side chains, are favorable charge solvating and binding sites for metal ions.[42-47]

A wide range of studies on metal/amino acid complexes has been focused on the effect of metal cations on the stability of zwitterionic versus non-zwitterionic conformations of amino acids.[48] The stability of zwitterionic versus non-zwitterionic forms of an amino acid depends on many factors. One of these factors is coordination of metal ions with amino acid functional groups. For example it has been predicted that the interaction of Cu^{2+} and Co^{2+} ions with glycine[49,50] in the gas phase stabilizes the zwitterionic conformer over the non-zwitterionic conformer.

The relative stabilities of zwitterion and non-zwitterion structures in metal cation-amino acid complexes is dependent on the identity of the metal ion,[51] due to the size and

charge density of the ion.[52-59] IRMPD spectrum of M^{2+}.His complexes, where M^{2+} is an alkaline earth divalent metal cation (Ca^{2+}/Ba^{2+}), is compared with calculated IR spectra of some low energy conformations of these complexes in Figure 1.5. The features at 1680 cm^{-1} (carbonyl stretching mode) and 1170 cm^{-1} (the OH bending mode) were related to a charge-solvated structure and the features around 1400-1550 cm^{-1} were found to correspond to carboxylate carbonyl stretches characteristic of a salt-bridged structure. The presence of a carbonyl stretch and the C-O-H bending mode in the IRMPD spectrum of Ca^{2+}.His is in agreement with a charge-solvated structure (Ca-CS1). In contrast these features are absent in Ba^{2+}.His in accordance with a salt bridged structure (Ba-SB1). Also, a comparison of the optimized structures of these complexes shows that smaller cations like Ca^{2+} favor a charge-solvated conformation whereas a larger cation such as Ba^{2+} stabilizes a salt bridged structure.[60] This is because smaller cations have higher charge densities which can be solvated more effectively in a tridentate CS structure by forming strong ion-dipole interactions.

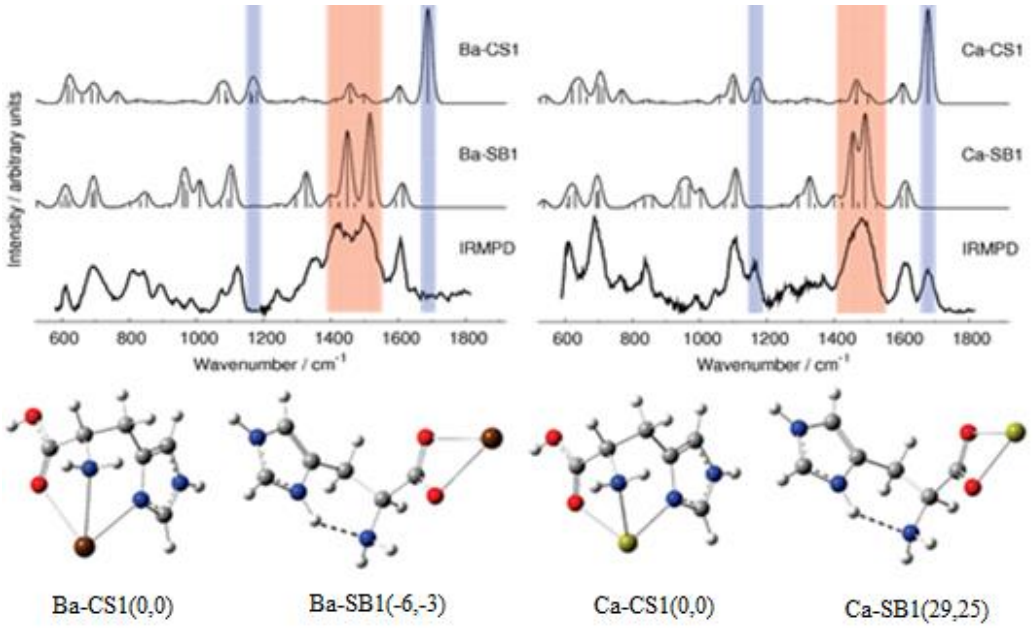


Figure 1.5. IRMPD and calculated IR spectra of lowest energy structures of Ba^{2+}.His and Ca^{2+}.His calculated at the MPW1PW91/6-311++G(d,p) level (A) Optimized structures of Charge-Solvated and Salt-Bridged complexes of Ba^{2+}.His and Ca^{2+}.His. The values (in parentheses) are enthalpies $\Delta H_0°$, and free energies $\Delta G_{298}°$ in kJ/mol. Reprinted with permission from *J. Phys. Chem. B* **2009**, *113*, 10403-10408. Copyright (2009) American Chemical Society.[60]

The metal ion coordination number in an amino acid complex is also dependent on metal cation size. IRMPD spectra of cationized cysteine with alkali metal ions revealed that the smaller alkali metal cations such as Li^+ and Na^+ form stable tridentate structures in which the metal cation is coordinated to the amine nitrogen, carboxylic carbonyl oxygen, and sulfur in the side chain. Strong interaction of the metal cation with the carbonyl oxygen results in a red shifted carbonyl stretch (shifted to lower energies) in the IRMPD spectrum in agreement with the calculated IR spectra of the Li^+ and Na^+ complexes. As the metal cation gets larger (K^+, Rb^+ and Cs^+), the carbonyl oxygen shows less participation in metal coordination. The main coordination site for these larger metal ions is the carboxylic acid group oxygens forming a bidentate structure around metal ions.

As a result the carbonyl stretch becomes blue shifted and appears at 1715, 1734, 1743 and 1756 cm^{-1} for Li^+, Na^+, K^+, Rb^+, respectively.[54]

1.2.2. Microsolvation

The interaction of biomolecules with their environment is also a determining factor in structure and reactivity of biological molecules. Solvation of macromolecules such as proteins as well as their subunits is one important example of such interactions. Protein structure is governed by inter- and intramolecular non-covalent interactions such as hydrogen bonding. Water is the most abundant molecule in nature and it exists as a solvent in biological systems and contributes to the protein's structure, enzymatic function, folding, and movement.[61-63] Water molecules participate in protein folding processes by linking ionic and polar functionalities of proteins via hydrogen bonding. In the secondary structure of proteins, hydrogen bonding links the amide hydrogen and carbonyl oxygen in the peptide backbone. For example, in an alpha helix protein domain, the n^{th} amide carbonyl oxygen, acting as an electron donor, solvates the $(n+4)^{th}$ amide hydrogen.[64] Water molecules can interrupt such interactions and form new hydrogen bonds. In the structure of collagen for example, water molecules stabilize the collagen triple helix by forming an intramolecular bridge between the carbonyl oxygen and amide hydrogen in the peptide chain.[65] Water molecules exist in two forms in the structure of proteins; internal and surface water.[61] Internal water molecules contribute to hydrogen bonding inside the protein and help in maintaining the folding and at the same time provide flexibility in the protein interior space. Surface water forms a network of hydrogen bonds with the functionalities at the protein surface as well as other water

molecules. This produces a hydration layer at the protein surface which governs the reactivity and dynamics of the protein inside biological media. Furthermore because of its nature, water can participate in proton/electron transfer reactions. Presence of water molecules as a ligand at the oxygen evolving center of photosystem II is a good example of water's role in this type of reaction.[66]

A diketone is a bifunctional molecule with two carbonyl groups that resemble the features of a peptide backbone which also contains repetitive units as carbonyl groups. The structure of protonated 2,4-pentanedione was taken as a model for the peptide chain.[67] The lowest energy structure of diketone in the presence of a proton is a closed chain in which the proton is located between two carbonyl oxygen atoms forming hydrogen bonds with these groups. In Figure 1.6 the optimized structures of the protonated 2,4-propanedione upon hydration with one(a), two(b) and three(c) water molecules are compared. It was shown that the lowest energy conformations for different degrees of hydration are PDW1I, PDW2I and PDW3I. These conformations each have a water mediated linkage between two carbonyl oxygen atoms. Upon addition of the first water molecule, three low energy conformers were formed. In the lowest energy structure (PDW1I) water has simply replaced the intrachain proton forming a one water molecule bridge. Second and third lowest energy structures were around 5.4 and 29.7 kJ mol^{-1} higher in energy. PDW1III is an open-chain structure and in PDW1II, H_3O^+ is linking two carbonyl oxygen atoms instead of a proton. There are four low energy structures reported for hydration with a second water molecule. In the first two lowest energy structures, the second water molecule is bound to the H_3O^+. Structure PDW2III has a two water bridge and PDW2IV is an open chain which is 17 kJ mol^{-1} higher in energy. In a similar way the

open chain structure is the highest in energy for the triply hydrated complex. Authors have reported the vibrational predissociation spectra for doubly and triply hydrated complexes and compared them with the calculated spectra of optimized structures in each case (Figure 1.7). In the experimental spectrum of the doubly hydrated complex four features were observed at 3694, 3647, 3737 and 3258 cm^{-1} related to free, symmetric, asymmetric and hydrogen bonded O-H stretches, respectively. Comparison of vibrational predissociation spectra of doubly and triply hydrated structures with the calculated spectra of these ions showed that open chain PDW2IV and PDW3IV are the best matches for the experimental results (Figure 1.6). However, authors could also observe triply hydrated isomers with a water bridge (PDW3III) at higher temperatures (200K). It was concluded that study of such systems provides a good example of hydration effects on unfolding of peptide/protein chains in a real biological system.

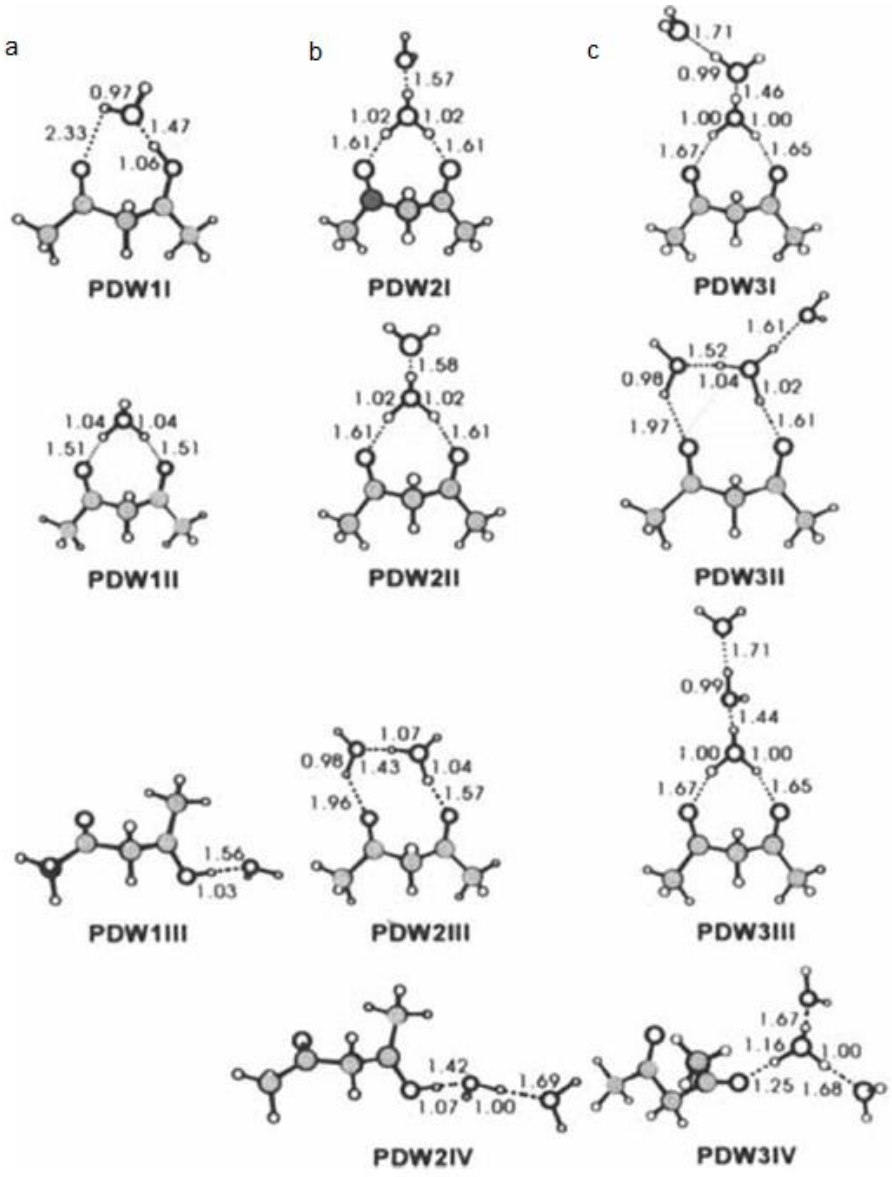


Figure 1.6. Structures of protonated 2,4-pentanedione hydrated with one (H^+[MeC(O)CH_2C(O)Me](H_2O))(a), two(H^+[MeC(O)CH_2C(O)Me]($H_2O)_2$)(b) and three H^+[MeC(O)CH_2C(O)Me]($H_2O)_3$(c) water molecules. Complexes are DFT optimized with (B3LYP/6-31+G*) level of theory. Reprinted with permission from *Mol. Phys.* **2003**, *101*, 1285-1295. Copyright (2003), with permission from Taylor & Francis group.[67]

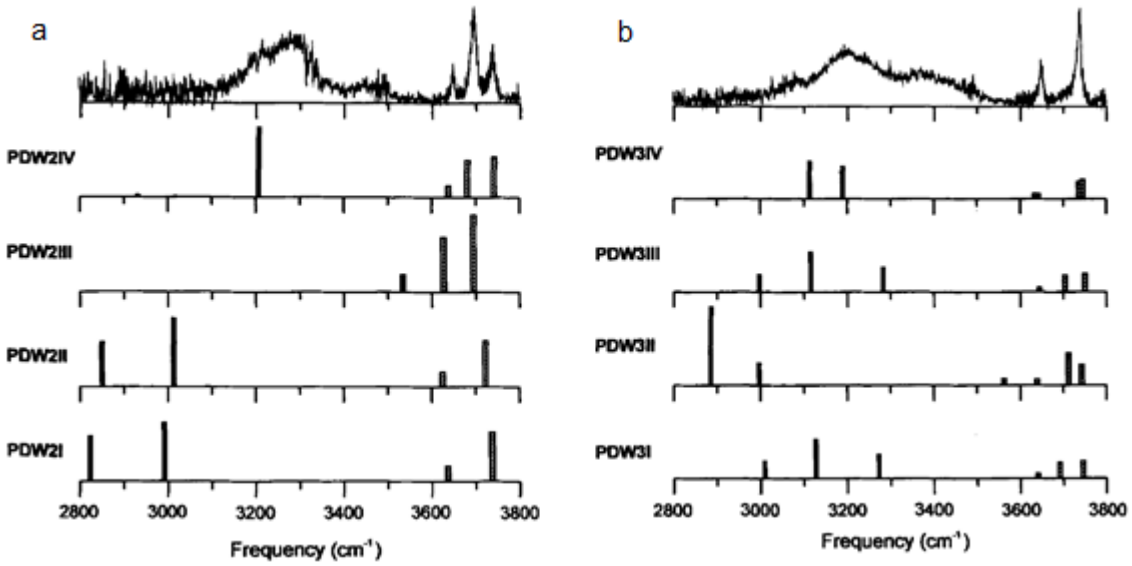


Figure 1.7. Vibrational predissociation spectra of $(H^+[MeC(O)CH_2C(O)Me](H_2O)_2)$(a) and $H^+[MeC(O)CH_2C(O)Me](H_2O)_3$(b) compared to calculated spectra of DFT optimized structures of these complexes in 2800-3800 cm^{-1}. Reprinted with permission from *Mol. Phys.* **2003**, *101*, 1285-1295. Copyright (2003), with permission from Taylor & Francis group.[67]

Microsolvation is a term used for the stepwise addition of water molecules to a molecular ion. The study of microsolvated ions is of great importance since it provides useful information about the effect of solvation on the structure and reactivity of biomolecules without the complication caused by bulk solvent. Microsolvated ions of biomolecules can mimic the structural features of many proteins in their active form which interact with only a limited number of water molecules. The stability of zwitterionic versus nonzwitterionic structures of amino acids[68-71] and their protonated and metallated[72,73] forms has been investigated upon microsolvation of these moieties. For example, the zwitterionic forms of alanine and glycine become stable after solvation with more than 4 and 6 water molecules, respectively.[68] It was also shown that amino acids

with a basic side chain such as lysine[74] or arginine[75] require fewer water molecules for the zwitterionic form to be stable compared to glycine or alanine. In solvation of ions with more than one water molecule, there is intermolecular hydrogen bonding between the water molecules as well. For example, in Figure 1.8, modelled structures of protonated arginine are presented with different degrees of hydration.[76] In the monohydrated structure there are two intramolecular hydrogen bonds between the carboxylic acid carbonyl oxygen and the terminal amine with guanidinium group. The water molecule also participates in intermolecular hydrogen bonding with the guanidinium group. The second water molecule inserts between the carbonyl oxygen and guanidinium NH. The intramolecular hydrogen bonds become replaced completely after solvation with four water molecules. Also it is shown that water-water hydrogen bonding occurs after the third water molecule insertion.[77]

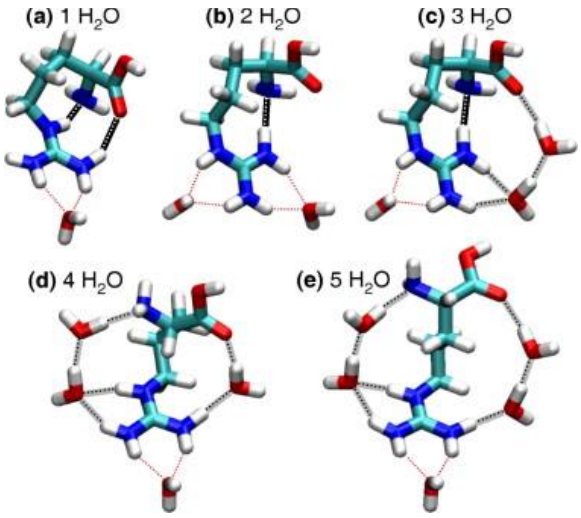


Figure 1.8. Structures of protonated arginine with one(a), two(b), three(c), four(d) and five(e) Hydration degree. Reprinted with permission from *Chem. Phys. Lett.* **2009,** *480*, 1–16. Copyright (2009), with permission from Elsevier.[76]

IR photodissociation of hydrated metal cations provides interesting information about ion-solvent and solvent-solvent interactions.[78-82] Investigating these structures using vibrational spectroscopy in the N-H/O-H region is useful to find out about the structure and hydrogen bonding of complexed water molecules. It was shown that hydration of Ca^{2+} with up to 6 water molecules occurs when forming a single hydration shell around the metal ion, Figure 1.9(6A). In the IR spectra of $Ca^{2+}(H_2O)_6$ there are two features related to the symmetric (3589–3595 cm^{-1}) and asymmetric (3665–3669 cm^{-1}) water OH stretches. By addition of one more water molecule to this cluster, new features start to appear between 3000-3500 cm^{-1} which correspond to hydrogen bonded O-H stretches of water molecules (Figure 1.10). It was shown that the lowest energy structure, 7B, contains one water molecule in the second hydration shell. Direct interaction of all seven water molecules with the Ca^{2+}, similar to 6A, results in a structure (7A) which is around 12 kJ mol^{-1} higher in energy. The water molecule in the first hydration shell participates as a double hydrogen bond donor and the water molecule in the second hydration shell as the hydrogen bond acceptor.[78]

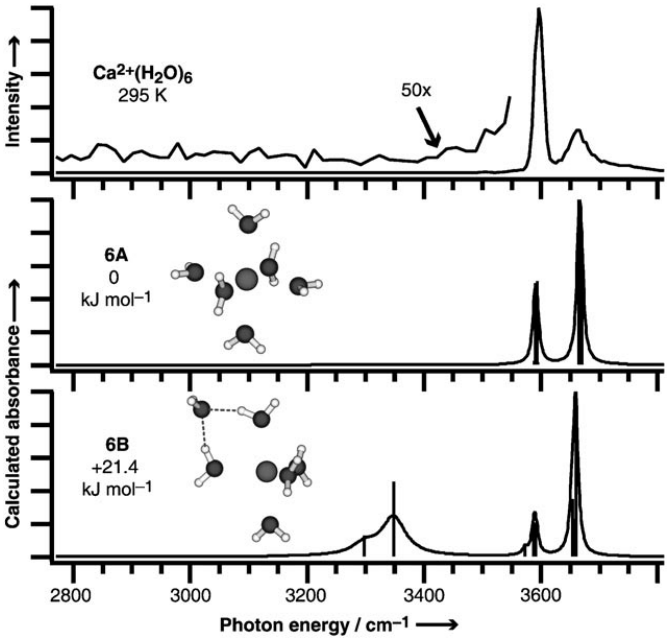


Figure 1.9. Infrared action spectrum compared to calculated IR spectra of optimized low energy structures of $Ca^{2+}(H_2O)_6$ (6A,6B). B3LYP/LACVP++** level of theory and basis set are used for vibrational frequency calculations. Reprinted with permission from *ChemPhysChem* **2007**, *8*, 2245 – 2253. Copyright (2007) John Wiley and Sons Ltd.[78]

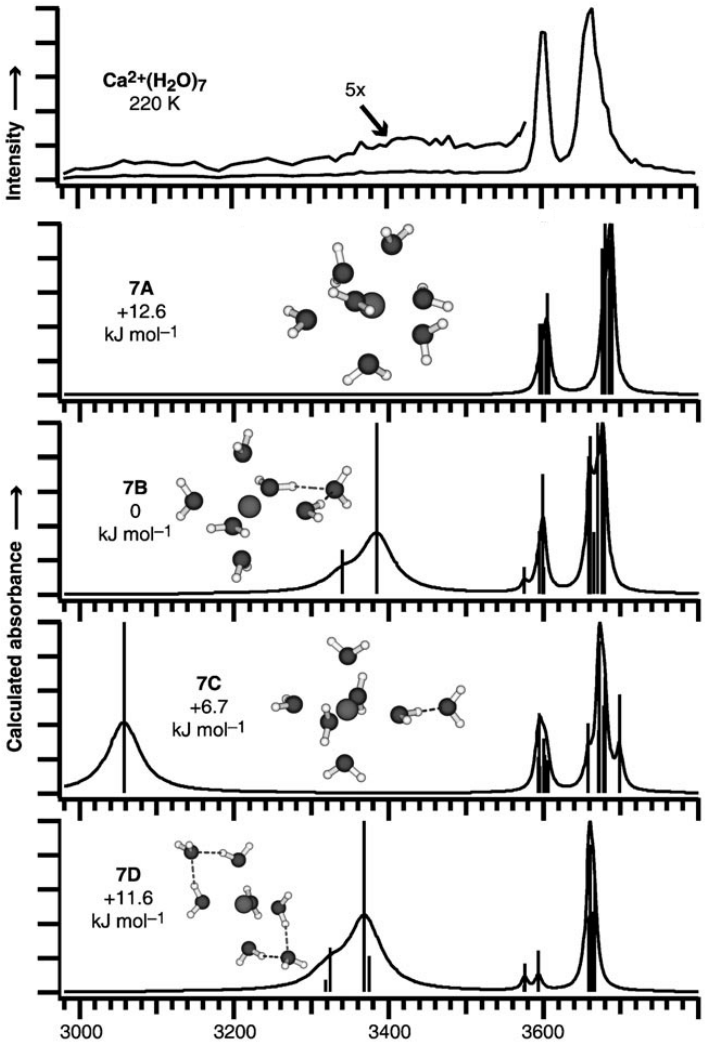


Figure 1.10. Infrared action spectrum compared to calculated IR spectra of optimized low energy structures of $Ca^{2+}(H_2O)_7$ (7A-7D). B3LYP/LACVP++** level of theory and basis set are used for vibrational frequency calculations. Reprinted with permission from *ChemPhysChem* **2007**, *8*, 2245 – 2253. Copyright (2007) John Wiley and Sons Ltd.[78]

Investigating the structural features of microsolvated metal cation complexes of amino acids and peptides is also of great interest. The effect of solvation on [Pb(Alanine-H)]$^+$ has been studied. It was shown that isomerization reactions occur upon solvation of these complexes with one water molecule.[32] The IRMPD spectrum of bare and hydrated

complexes are compared to calculated IR spectra of some low energy structures (Figure 1.11). The strong feature at 3560 cm^{-1} in the IRMPD spectrum of $[Pb(Alanine-H]^+$ was assigned to the carboxylic O-H stretch. Comparisons between calculated IR spectra of the lowest energy structures show that only Ala-A and Ala-D reproduced the carboxylic O-H stretch feature. Based on the optimized structures, Ala-A and Ala-D were both deprotonated at the amine group (Figure 1.12). However Ala-D was much higher in energy and metal was coordinated with the hydroxyl oxygen. Structures Ala-B and Ala-C were ruled out because of absence of the feature for the carboxylic acid group O-H stretch. Addition of one water molecule provided more possible structures. IRMPD spectra of $[Pb(Ala-H)H_2O]^+$ was compared to the calculated IR spectra of the lowest energy structure. Two features were observed at 3690 cm^{-1} and at 3580 cm^{-1} which were related to Pb^{2+} bound O-H and a carboxylic acid O-H stretch, respectively. Structures AlaW-Ai and AlaW-D were ruled out because of the absence of these features in their spectra. Presence of any free water O-H stretch was also discounted by isotopic experiments. It was shown that replacing H_2O with $H_2{}^{18}O$ does not shift both of the features observed in the water O-H stretching region which confirms absence of asymmetric and symmetric water O-H stretch and, as a result, any structure with a free water (AlaW-B). In another attempt H_2O was replaced with CH_3OH. Since there was a band around 3640 cm^{-1} which was in contrast with absence of Pb-OH stretch, the authors suggested that a mixture of AlaW-A and AlaW-C structures contributed to the experimental features.

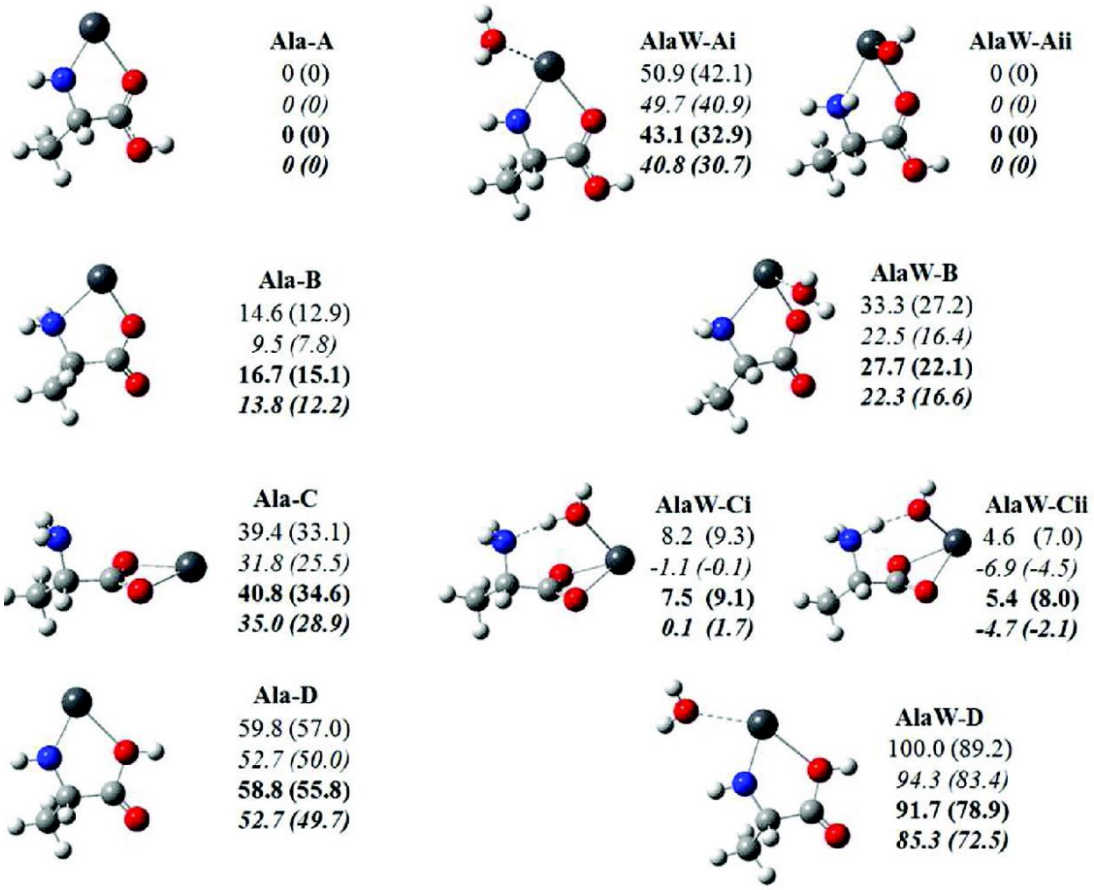


Figure 1.11. Structures and enthalpy (kJ mol^{-1}) (and Gibbs energies) of bare (left) and hydrated (right) $[Pb(Ala-H)]^+$ calculated with B3LYP/6-31+G(d,p) using LANL2DZ (unbolded) SDD (bolded) basis set and effective core potential on Pb. Italicized values are MP2(full)/6-311++G(2d,2p)//B3LYP/6-31+G(d,p) (with LANL2DZ or SDD on Pb). Reprinted with permission from *J. Phys. Chem. B* **2011**, *115*, 11506-11518. Copyright (2011) American Chemical Society.[32]

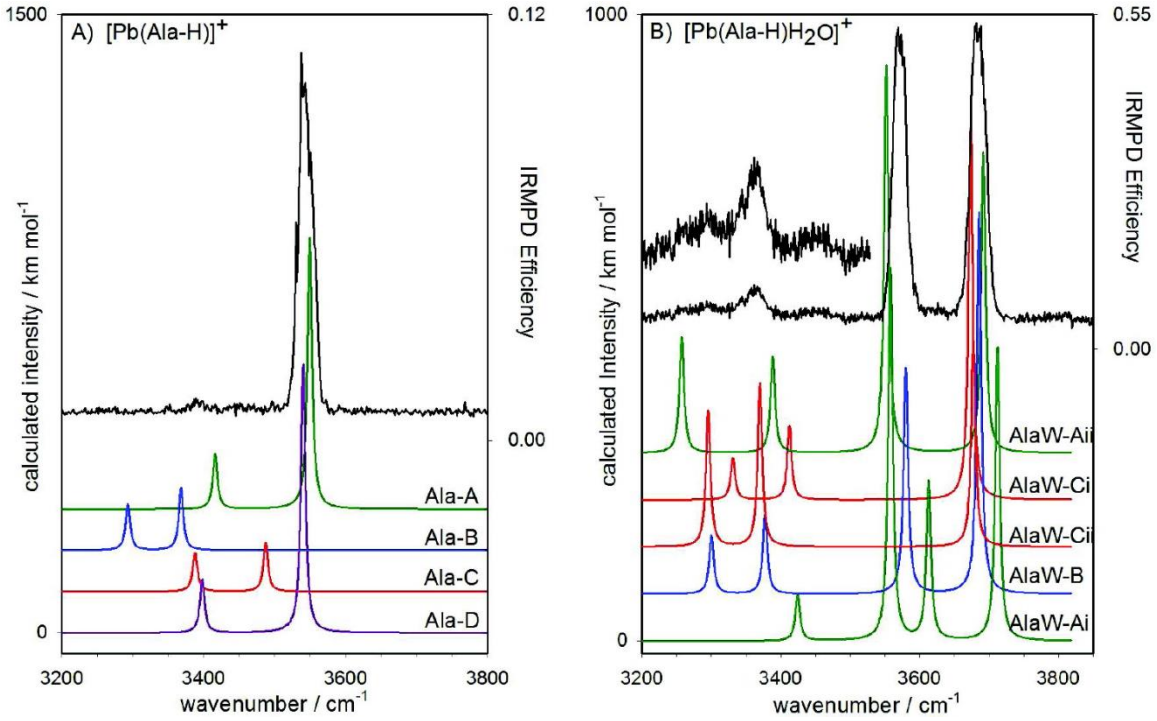


Figure 1.12. The IRMPD and Calculated IR spectra of different conformers of bare (A) and hydrated (B) [Pb(Ala-H)$^+$]. Reprinted with permission from *J. Phys. Chem. B* **2011**, *115*, 11506-11518. Copyright (2011) American Chemical Society.[32]

1.2.3. Effect of amino acid structures on metal ion complexes

Heteroatoms as well as aromatic rings in amino acid side chains are potential sites for metal ion coordination. Complexes of alkali metal cations with aromatic amino acids such as tryptophan[83] and phenylalanine[84] form stable charge-solvated structures. Smaller alkali metal cations Li^+, Na^+ and K^+ preferably form a tridentate complex in which metal cations interact with the aromatic ring as well as carbonyl oxygen and terminal amine nitrogen. Larger alkali metal cations, Rb^+ and Cs^+ mainly form a bidentate complex with metal being coordinated with the aromatic ring and the carbonyl oxygen. The higher charge density of small metal ions becomes stabilized through interactions with both

carbonyl oxygen and amine nitrogen as well as the pyrrole-ring side of the tryptophan indole ring whereas larger metal ions are more loosely bound to carbonyl oxygen and participate in electrostatic cation-π interactions with the indole benzene ring.

The presence of a basic side chain in the amino acid can also provide a coordination site for metal ions. Comparing the IRMPD efficiency and calculated IR spectra of lithiated and sodiated lysine with an amine side chain (Figure 1.13) shows a carboxylic acid carbonyl stretch band at 1720 cm^{-1} and NH_2 bending mode at 1595 cm^{-1}. These features were related to a non-zwitterionic structure in which the metal is triply coordinated with terminal and side chain amine nitrogen as well as carboxylic acid carbonyl oxygen.[85]

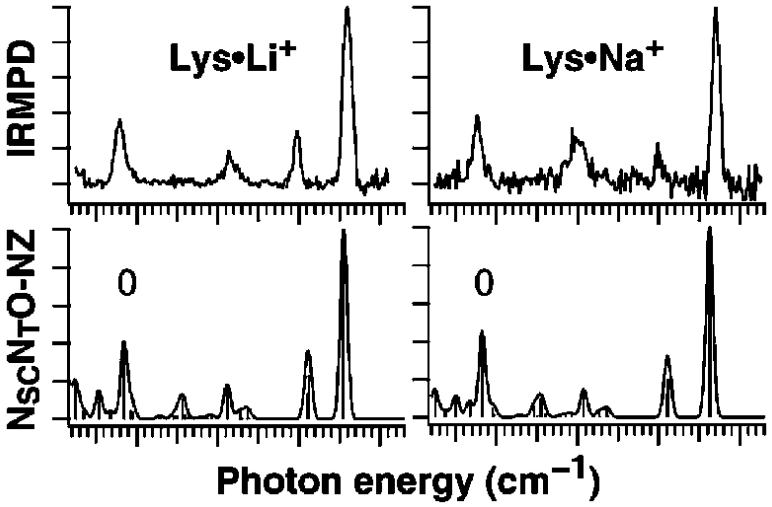


Figure 1.13. The IRMPD spectra compared to calculated IR spectra of lowest energy structures of Li^+.Lys and Na^+.Lys. DFT calculations are done by B3LYP/6-31++G** and Gibbs energies are reported at 298 K in kJ mol^{-1}. Reprinted (adapted) with permission from *J. Phys. Chem. A* **2007**, *111*, 7753-7760. Copyright (2007) American Chemical Society.[85]

In a similar study it was shown that interaction of Ca^{2+} with serine and cysteine results in a stable charge-solvated structure whereas Ca^{2+}.Gly is stable as a salt-bridged conformer in the gas phase. Heteroatoms such as oxygen or sulfur of OH or SH groups in the side chain of these amino acids can participate in metal cation coordination. The lowest energy structures of Ca^{2+} with Gly, Cys and Ser are compared in Figure 1.14. The most stable structure for Gly is a salt-bridged (b) in which metal is solvated with carboxylate oxygens. For Ser and Cys complexes the lowest energy structure is charge-solvated (c), but in this case, Ca^{2+} is solvated by carboxylic carbonyl and amino group as well as oxygen/sulfide from OH/SH in serine and cysteine, respectively. In the charge solvated structure (a), which is much higher in energy for all three amino acids, the metal is coordinated with amine nitrogen and carboxylic acid carbonyl oxygen. Authors have also discussed that the interaction of Ca^{2+} with oxygen is much stronger than sulfur so the energy difference between the SB and CS conformers of Ca^{2+}.Ser is higher compared to the Ca^{2+}.Cys.[86]

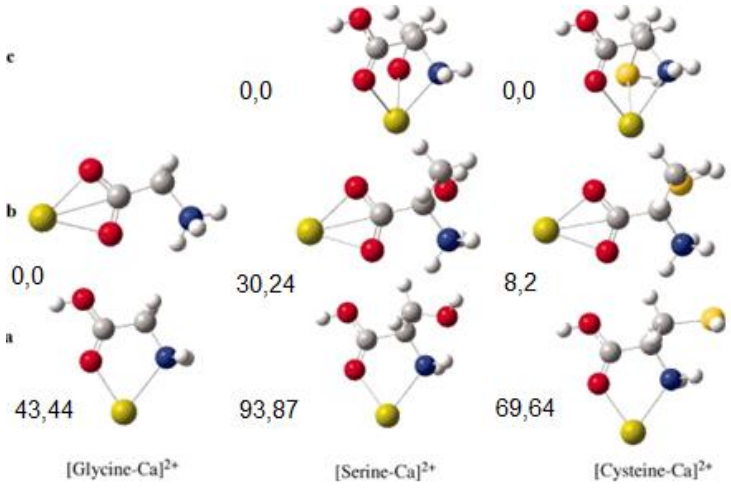


Figure 1.14. Structures of Gly, Ser and Cys with Ca^{2+} are DFT optimized using B3LYP/6-311G(d,p). Relative enthalpies and Gibbs energies (298K) are reported at kJ mol^{-1}. Reprinted with permission from *Int. J. Quantum Chem*. 2012, *112*, 2126–2134. Copyright (2012) John Wiley and Sons Ltd.[86]

1.2.4. Peptides as ligands

Besides amino acids, interaction of metal cations with peptides as ligands has been also widely studied. Peptides are the smallest part of proteins which carry sequence information. Along with metal cation size, peptide length[51,87] and sequence[88] are determining factors in the conformation of metal/peptide complexes. Both the effect of metal cation size and peptide sequence was studied experimentally and computationally on complexes of ArgGly and GlyArg with alkali metal cations. The results for H^+.ArgGly and H^+.GlyArg are taken as references for the spectral features. In Figure 1.15 IRMPD spectra of H^+.ArgGly and H^+.GlyArg are compared to calculated IR spectra of low energy structures of these complexes.[89] Spectral features for both isomeric dipeptides suggest that charge solvated structures (CS4 and CS1* respectively) mostly contribute to the experimental spectrum. In conclusion, the sequence of amino acids shows no effect on

stability of H^+.ArgGly and H^+.GlyArg complexes. The main experimental features observed in the spectrum were a free carboxyl C-O-H bend at 1100-1150 cm^{-1} and the carboxylic acid carbonyl stretch at 1750-1780 cm^{-1} that were related to a charge-solvated structure. The absence of these features for Li^+, Na^+ and Cs^+ complexes of GlyArg confirm a salt-bridged structure to be dominant. In contrast, the stability of isomeric ArgGly complex is sensitive to metal cation size. Figure 1.16 Compares the IRMPD and IR spectra of a mixture of salt-bridged and charge solvated conformers for Li^+ and Cs^+ complexed to ArgGly. Spectral features as well as the energy comparison demonstrate that smaller metal cations prefer a charge solvated conformation whereas a salt bridged structure is the lowest energy structure for larger metal cations.

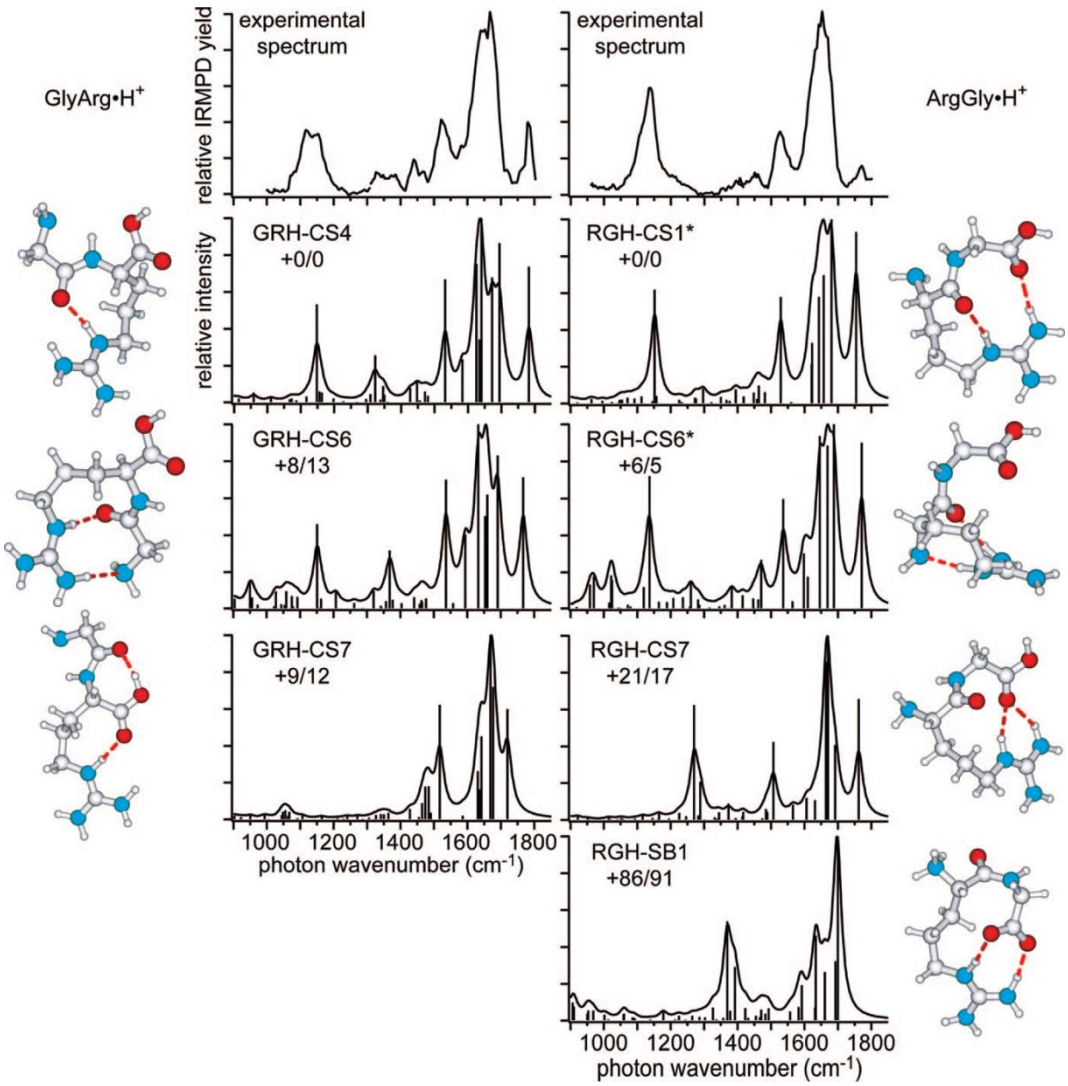


Figure 1.15. The IRMPD and calculated IR spectra of low energy conformers of H^+.ArgGly and H^+.GlyArg and the relative Gibbs energies (kJ mol^{-1}) at 0 and 298K. The features around 1100-1150 cm^{-1} and 1788 cm^{-1} are related to the bending of OH group and the carboxylic acid carbonyl stretch, respectively. Reprinted with permission from *J. Am. Chem. Soc.* 2009, *131*, 1232-1242. Copyright (2009) American Chemical Society.[89]

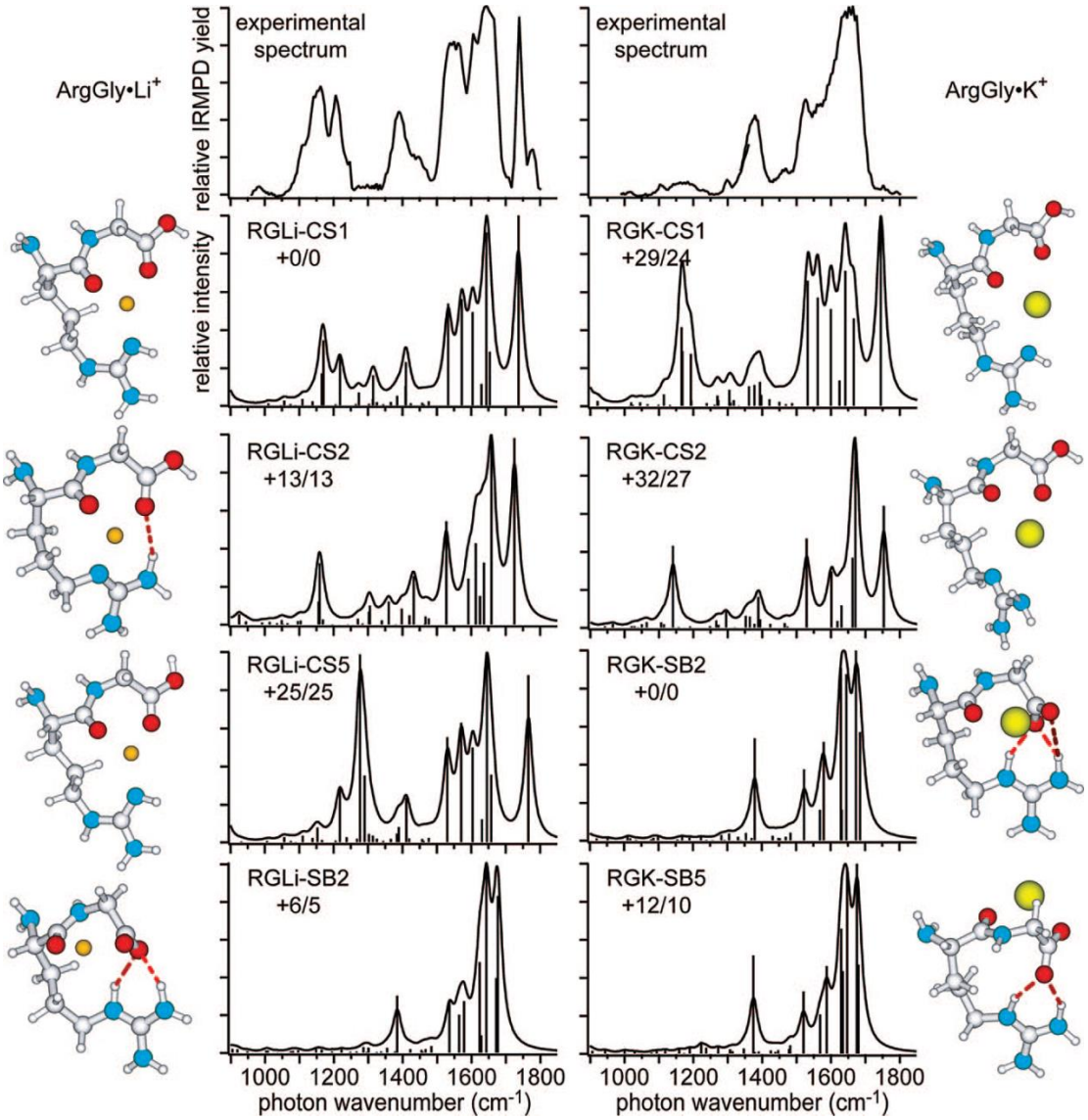


Figure 1.16. The IRMPD and calculated IR spectra of low energy conformers of Li^+.GlyArg and Cs^+.GlyArg. Reprinted with permission from *J. Am. Chem. Soc.* 2009, *131*, 1232-1242. Copyright (2009) American Chemical Society.[89]

Figure 1.17 compares the IRMPD spectra of isomeric dipeptides AlaPhe and PheAla complexed with Ba^{2+} with the calculated spectra of the lowest energy structure in each case. The features in the IRMPD spectrum of Ba^{2+}-PheAla was in agreement with a charge solvated conformation. The features between 1600-1700 cm^{-1} and at about 1150 cm^{-1} were related to a carboxylic acid carbonyl stretch and C-O-H bending modes, respectively which are characteristics of a charge solvated structure. In contrast experimental results for Ba^{2+}.AlaPhe were reproduced by a salt bridged structure. It was proposed that changing the sequence from AlaPhe to PheAla provides a more favorable metal solvation by the phenylalanine residue in a way that the aromatic side chain of the phenylalanine is able to solvate the metal charge through cation-π interactions. The geometry of the peptide in the Ba^{2+}.PheAla complex allows efficient interaction of the aromatic ring with the metal cation, and this induces stability of a charge-solvated versus the salt-bridged conformation. Effect of peptide chain length was also studied on complexes of Ba^{2+} with AlaAla and AlaAlaAla. In Figure 1.18 experimental spectra of Ba^{2+}.AlaAla and Ba^{2+}.AlaAlaAla are compared to calculated IR spectra of the charge solvated and salt bridged structures in each case. In a similar way, presence of carbonyl stretch and C-O-H bending modes around 1600-1700 cm^{-1} and 1150 cm^{-1}, respectively, were taken as references for a charge solvated conformation. These features were observed in the spectrum of Ba^{2+}.AlaAlaAla. In the optimized structure of Ba^{2+}.AlaAlaAla all carbonyl oxygens have participated in metal cation coordination whereas in Ba^{2+}.AlaAla the metal is interacting with carboxylate oxygens. Increasing the peptide chain length increases the chelating sites for the metal cation resulting in a preference for charge-solvated structure.[88]

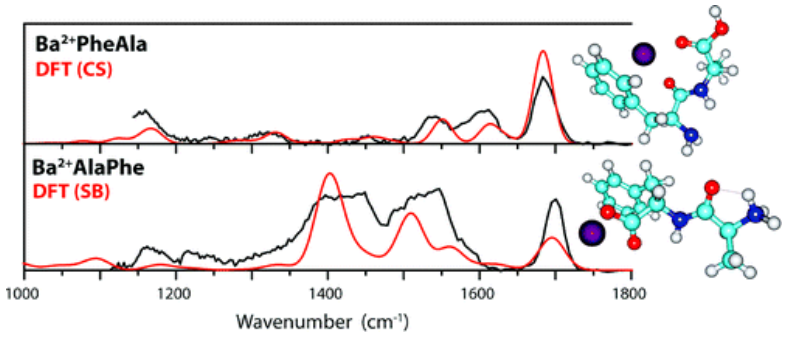


Figure 1.17. The IRMPD and calculated IR spectra of lowest energy conformers of Ba^{2+}.PheAla and Ba^{2+}.AlaPhe. Structures are optimized at B3LYP/6-31+G(d,p) (SDD ECP on Ba). IR frequencies are scaled by 0.975. Reprinted with permission from *J. Phys. Chem. B* 2009, *113*, 10552-10554. Copyright (2009) American Chemical Society.[88]

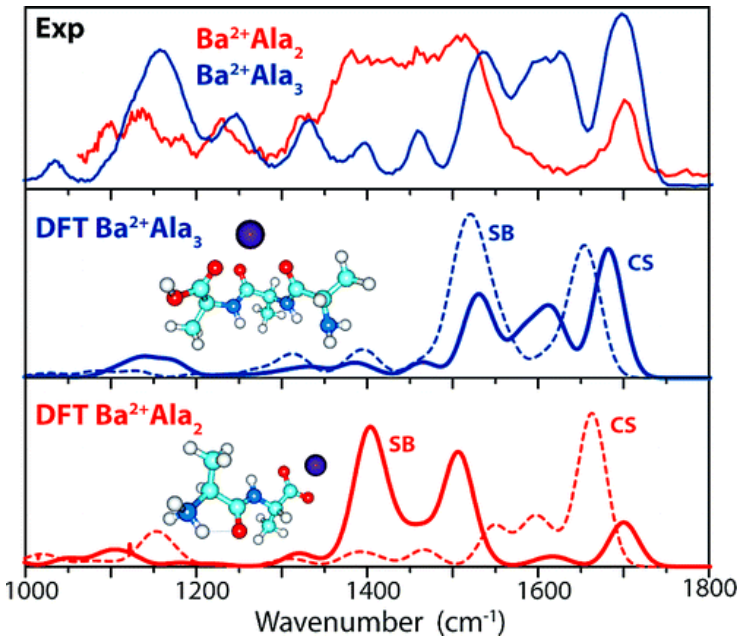


Figure 1.18. The IRMPD spectrum and calculated IR spectra of predicted Charge-Solvated and Salt-Bridged conformers of Ba.AlaAla^{2+} and Ba.AlaAlaAla^{2+}. Reprinted with permission from *J. Phys. Chem. B* 2009, *113*, 10552-10554. Copyright (2009) American Chemical Society.[88]

1.2.5. Metal cations amino acid/peptide complex dissociation

Metal ions can affect the dissociation pattern of metal cation biomolecular complexes. Since metal ions size affects the stability of a specific structure in an amino acid it can also influence the dissociation pathways of these complexes. For example CID of lithiated alanine leads to small neutral dissociation such as NH_3. Loss of NH_3 proceeds through zwitterionic intermediates (16 and 17) (Figure1.19) which are stabilized by metal cation coordination. [90]

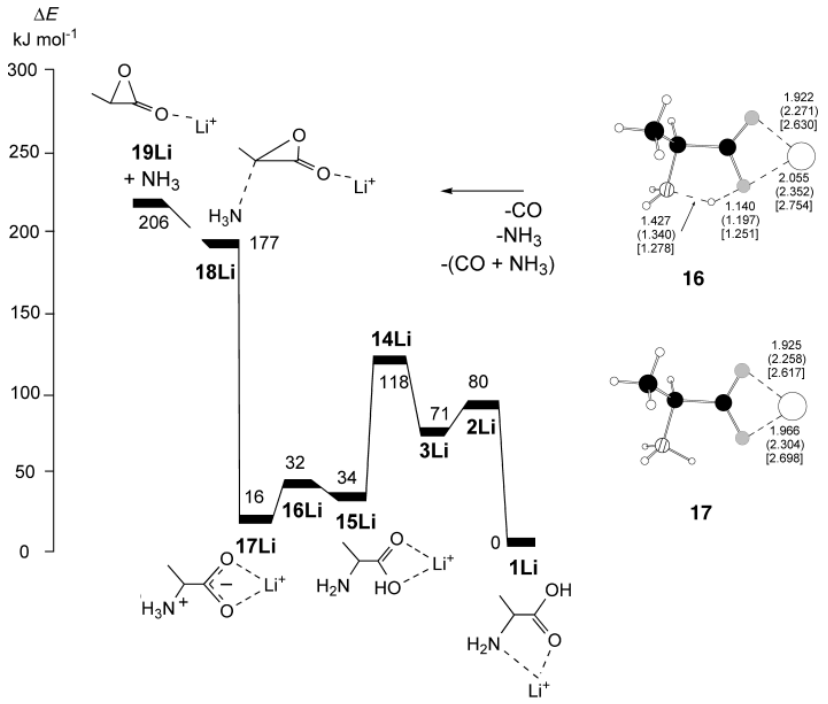


Figure 1.19. Potential energy surface diagram for dissociation of NH_3 from Li^+.Ala and DFT optimized structures 16 and 17 using B3LYP/6-31G(d). Relative enthalpies are calculated by single point calculations B3LYP/6-311+G(3df,2p) and reported at 0 K in kJ mol^{-1}. Reprinted with permission from *Chem. Eur. J.* 2005, *11*, 5289 – 5301. Copyright (2005) John Wiley and Sons Ltd.[90]

In another study it was found that metallated histidine complexes with alkali metal cations shows loss of an intact amino acid as the only dissociation for larger metal cations e.g. K^+, Cs^+ and Rb^+ whereas Na^+ induces smaller neutral molecule dissociations such as H_2O, CO_2 and NH_3 followed by a CO.[91]

Similar to amino acids, the dissociation mechanism of metal cation peptide complexes is also governed by the metal type and charge as well as the functional groups contributing to the metal solvation. In a study the effect of first row transition metals on dissociation of polyalanine was investigated. CID mass spectra of heptaalanine with Fe^{2+} shows high intensities of $[M+Fe-H]^+$ and $[M+Fe]^{2+}$ as precursor ions.[92] Dissociation of $[M+Fe-H]^+$ produces a complete series of metallated a, b and y fragments from peptide residues that can be distinguished by their mass to charge difference. In Figure 1.20 the CID mass spectra of $[M+Fe]^{2+}$, $[M+Co]^{2+}$ and $[M+Ni]^{2+}$ are compared. It is shown that the $[M+Fe]^{2+}$ precursor ion dissociates to a mixture of metallated and non-metallated fragments ions of a, b and c series. In contrast $[M+Co]^{2+}$ only produces metallated fragment ions that have a better peptide sequence coverage after the fourth residue. More interestingly, $[M+Ni]^{2+}$ forms a-type fragment ions as well as a couple of y-type ions ($[y_5+Ni]^+$ and $[y_6+Ni-H]^+$) and these are only observed for Ni^{2+}. This shows how different metal ions can influence the dissociation pattern of peptide complexes.

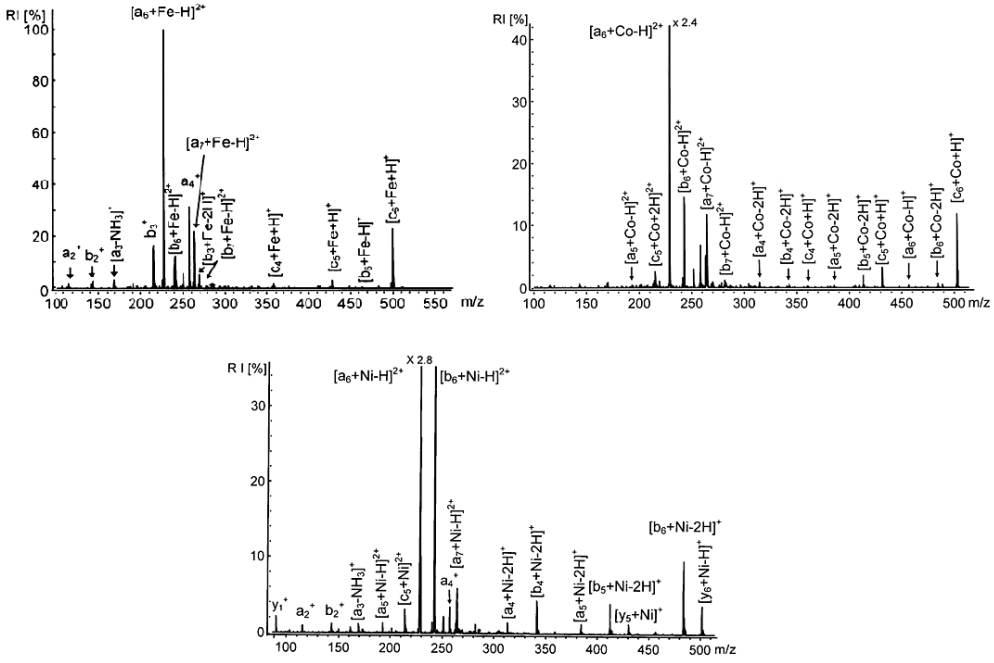


Figure 1.20. CID mass spectra of $[M+Fe]^{2+}$, $[M+Co]^{2+}$ and $[M+Ni]^{2+}$. Reprinted with permission from *J. Mass Spectrom.* 2011, *46*, 1099–1107. Copyright (2011) John Wiley and Sons Ltd.[92]

It was shown in another example that argentinated peptides produce a series of triplet fragment ions assigned to $[b_n+OH+Ag]^+$, $[b_n-H+Ag]^+$ and $[a_n-H+Ag]^+$ that their *m/z* values are different by masses of water and carbon monoxide.[93] The difference in the *m/z* value of each triplet with the next series corresponds to an amino acid or residue. Collision-induced dissociation mass spectra of argentiated enkephalin, a five residue peptide YGGFL, at *m/z* 662.2 shows sets of $[b_n+OH+Ag]^+$ ions as the most intense fragment ion at *m/z* 548.9, 402.1 and 345.0 which corresponds to sequence of Leucine/Isoleucine(L/I), phenylalanine(F), glycine(G) and glycine(G), respectively. Also authors explained that a higher intensity of $[y_n+H+Ag]^+$ is an evidence for presence of

amino acids with a basic side chain such as methionine, lysine, and arginine, and stronger interaction of metal ion with their basic site.

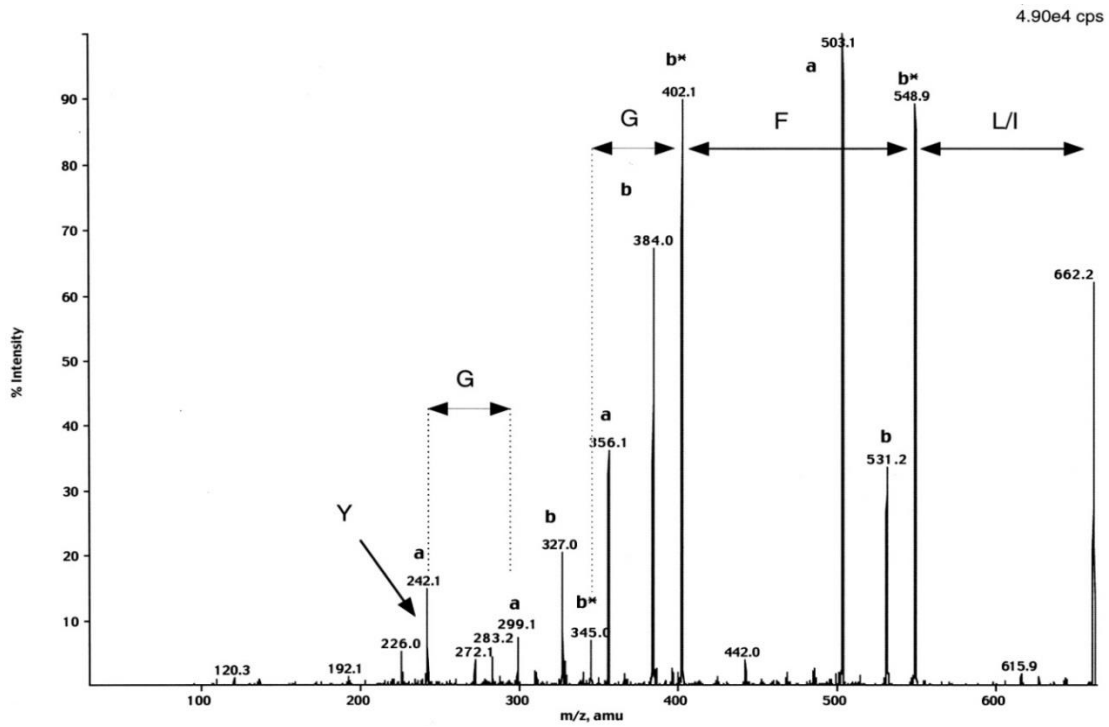


Figure 1.21. Fragmentation spectrum of argentinated leucine encephalin (*m/z* 662) YGGFL. Reprinted with permission from *Anal. Chem.* 1999, *71*, 2364-2372. Copyright (1999) American Chemical Society.[93]

1.3. Objectives of thesis

The focus of this work is on the structures and energetics of metal cation complexes of dipeptides in the gas phase by means of mass spectrometry techniques and computational methods. Structures of Na^+ and Ca^{2+} complexes with isomeric dipeptides AlaGly and GlyAla are revealed by their infrared spectroscopic characteristics. The possible structures of these complexes are optimized and their energies are compared in order to find the lowest energy structures. Structure and energetics of micro-solvated complexes of AlaGly and GlyAla with Na^+ and Ca^{2+} upon step-wise addition of one to three water molecules are also studied.

The unimolecular chemistry of Ca^{2+} clusters $[Ca_n(GlyGly\text{-}H)_{2n-1}]^+$ and $[Ca_{n+1}(GlyGly\text{-}H)_{2n}]^{2+}$ are also studied. Here it is shown that activation of singly charged and doubly charged clusters by collision-induced dissociation results in different fragmentation pathways. The potential energy surface is explored by simulated annealing as well as density functional theory (DFT) methods. Calculated IR spectra of the lowest energy structures are also compared to the IRMPD spectrum of $[Ca_3(GlyGly\text{-}H)_4]^{2+}$ and $[Ca_5(GlyGly\text{-}H)_8]^{2+}$ to determine the structure that is responsible for the experimentally observed features. The unimolecular fragmentation of $[M_{n+1}(GlyGly\text{-}H)_{2n+1}]^+$ and $[M_{n+1}(GlyGly\text{-}H)_{2n}]^{2+}$ clusters for $M=Mg^{2+}$, Sr^{2+} and Mn^{2+} were also explored. The effect of metal ion size and identity is studied by comparison of the dissociation reactions of these complexes.

References

(1) Wysocki, V. H.; Resing, K. A.; Zhang, Q.; Cheng, G. Mass spectrometry of peptides and proteins. *Methods* **2005**, *35*, 211-222.

(2) 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.

(3) Karas, M.; Hillenkamp, F. Laser desorption ionization of proteins with molecular masses exceeding 10000 daltons. *Anal. Chem.* **1988**, *60*, 2299-2301.

(4) Stults, J. T. Matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS). *Curr. Opin. Struct. Biol.* **1995**, *5*, 691-698.

(5) Ashcroft, A. E. Protein and peptide identification: The role of mass spectrometry in proteomics. *Nat. Prod. Rep.* **2003**, *20*, 202-215.

(6) 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.

(7) Polfer, N. C.; Oomens, J.; Moore, D. T.; Von Helden, G.; Meijer, G.; Dunbar, R. C. Infrared spectroscopy of phenylalanine Ag(I) and Zn(II) complexes in the gas phase. *J. Am. Chem. Soc.* **2006**, *128*, 517-525.

(8) Bush, M. F.; O'Brien, J. T.; Prell, J. S.; Saykally, R. J.; Williams, E. R. Infrared spectroscopy of cationized arginine in the gas phase: Direct evidence for the transition from nonzwitterionic to zwitterionic structure. *J. Am. Chem. Soc.* **2007**, *129*, 1612-1622.

(9) Talaty, E. R.; Perera, B. A.; Gallardo, A. L.; Barr, J. M.; Van Stipdonk, M. J. Elucidation of fragmentation pathways for the collision-induced dissociation of the binary Ag(I) complex with phenylalanine. *J. Phys. Chem. A* **2001**, *105*, 8059-8068.

(10) Ruan, C.; Rodgers, M. T. Cation-π interactions: Structures and energetics of complexation of Na^+ and K^+ with the aromatic amino acids, phenylalanine, tyrosine, and tryptophan. *J. Am. Chem. Soc.* **2004**, *126*, 14600-14610.

(11) Yates Iii, J. R. Mass spectrometry and the age of the proteome. *Journal of Mass Spectrometry* **1998**, *33*, 1-19.

(12) Aebersold, R.; Goodlett, D. R. Mass spectrometry in proteomics. *Chem. Rev.* **2001**, *101*, 269-295.

(13) Mitchell Wells, J.; McLuckey, S. A.: Collision-induced dissociation (CID) of peptides and proteins. *Methods in Enzymol.* **2005**, 402, 148-185.

(14) Medzihradszky, K. F.: Peptide sequence analysis. *Methods Enzymol*. **2005**, 402, 209-244.

(15) Hunt, D. F.; Yates Iii, J. R.; Shabanowitz, J.; Winston, S.; Hauer, C. R. Protein sequencing by tandem mass spectrometry. *Proc. Natl. Acad. Sci. U. S. A.* **1986**, *83*, 6233-6237.

(16) Johnson, R. S.; Martin, S. A.; Biemann, K. Collision-induced fragmentation of $(M+H)^+$ ions of peptides. Side chain specific sequence ions. *Int. J. Mass Spectrom. Ion Processes*, **1988**, *86*, 137-154.

(17) Sokolowska, I.; Wetie, A. G. N.; Woods, A. G.; Darie, C. C. Applications of mass spectrometry in proteomics. *Aust. J. Chem.* **2013**, *66*, 721-733.

(18) Hernandez, P.; Müller, M.; Appel, R. D. Automated protein identification by tandem mass spectrumetry: Issues and strategies. *Mass Spectrom. Rev.* **2006**, *25*, 235-254.

(19) Barr, J. M.; Van Stipdonk, M. J. Multi-stage tandem mass spectrometry of metal cationized leucine enkephalin and leucine enkephalin amide. *Rapid Commun. Mass Spectrom.* **2002**, *16*, 566-578.

(20) Crowe, M. C.; Brodbelt, J. S. Infrared multiphoton dissociation (IRMPD) and collisionally activated dissociation of peptides in a quadrupole ion trap with selective IRMPD of phosphopeptides. *J. Am. Soc. Mass Spectrom.* **2004**, *15*, 1581-1592.

(21) Roepstorff, P., Fohlman, J. Proposal for a common nomenclature for sequence ions in mass spectra of peptides. *Biomed. Mass Spectrom.* **1984**, *11*, 601.

(22) Biemann, K. Contributions of mass spectrometry to peptide and protein structure. *Biomed. Environ. Mass Spectrom.* **1988**, *1-12*, 99-111.

(23) Medzihradszky, K. F.; Chalkley, R. J. Lessons in de novo peptide sequencing by tandem mass spectrometry. *Mass Spectrom. Rev.* **2015**, *34*, 43-63.

(24) Eyler, J. R. Infrared multiple photon dissociation spectroscopy of ions in penning traps. *Mass Spectrom. Rev.* **2009**, *28*, 448-467.

(25) Bomse, D. S.; Beauchamp, J. L. Selective enhancement of bimolecular reaction rates by over three orders of magnitude using low intensity CW infrared laser radiation. *J. Am. Chem. Soc.* **1980**, *102*, 3967-3969.

(26) 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.

(27) Prazeres, R.; Glotin, F.; Insa, C.; Jaroszynski, D. A.; Ortega, J. M. Two-colour operation of a Free-Electron Laser and applications in the mid-infrared. *Eur. Phys. J. D* **1998**, *3*, 87-93.

(28) Ortega, J. M.; Berset, J. M.; Chaput, R.; Glotin, F.; Humbert, G.; Jaroszynski, D.; Joly, P.; Kergosien, B.; Lesrel, J.; Marcouillé, O.; Peremans, A.; Prazeres, R.; Tadjeddine, A. Activities of the CLIO infrared facility. *Nucl. Instrum. Methods Phys. Res., Sect. A: Accelerators, Spectrometers, Detectors and Associated Equipment* **1996**, *375*, 618-625.

(29) Fridgen, T. D. Infrared consequence spectroscopy of gaseous protonated and metal ion cationized complexes. *Mass Spectrom. Rev.* **2009**, *28*, 586-607.

(30) 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.

(31) 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.

(32) 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.

(33) Mitchell Wells, J.; McLuckey, S. A.: Collision-induced dissociation (CID) of peptides and proteins. *Methods Enzymol.* **2005**, *402*, 148-185.

(34) Sleno, L.; Volmer, D. A. Ion activation methods for tandem mass spectrometry. *J. Mass Spectrom.***2004**, *39*, 1091-1112.

(35) Papayannopoulos, I. A. The interpretation of collision-induced dissociation tandem mass spectra of peptides. *Mass Spectrom. Rev.* **1995**, *14*, 49-73.

(36) Zhurov, K. O.; Fornelli, L.; Wodrich, M. D.; Laskay, U. A.; Tsybin, Y. O. Principles of electron capture and transfer dissociation mass spectrometry applied to peptide and protein structure analysis. *Chem. Soc. Rev.* **2013**, *42*, 5014-5030.

(37) Syka, J. E. P.; Coon, J. J.; Schroeder, M. J.; Shabanowitz, J.; Hunt, D. F. Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 9528-9533.

(38) Lippard, S. J.: *Principles of Bioinorganic Chemistry*, University Science Books, Sausalito, **1994**.

(39) Crichton, R. R.: *Biological Inorganic Chemistry*, Elsevier, Boston, **2008**.

(40) Bertini, I. G., H. B., Lippard, S. J., Valentine, J. S.: *Bioinorganic Chemistry*, University Science Books , Mill Valley, Calif., **1994**.

(41) Permyakov, E. A.: *Metalloproteomics*, Wiley, Hoboken **2009**.

(42) Shoeib, T.; Siu, K. W. M.; Hopkinson, A. C. Silver ion binding energies of amino acids: Use of theory to assess the validity of experimental silver ion basicities obtained from the kinetic method. *J. Phys. Chem. A* **2002**, *106*, 6121-6128.

(43) Marino, T.; Russo, N.; Toscano, M. Interaction of Li^+, Na^+, and K^+ with the proline amino acid. Complexation modes, potential energy profiles, and metal ion affinities. *J. Phys. Chem. B* **2003**, *107*, 2588-2594.

(44) Jockusch, R. A.; Price, W. D.; Williams, E. R. Structure of cationized arginine ($Arg{\cdot}M^+$, H = H, Li, Na, K, Rb, and Cs) in the gas phase: Further evidence for zwitterionic arginine. *J. Phys. Chem. A* **1999**, *103*, 9266-9274.

(45) Dunbar, R. C. Complexation of Na^+ and K^+ to aromatic amino acids: A density functional computational study of cation-π interactions. *J. Phys. Chem. A* **2000**, *104*, 8067-8074.

(46) Siu, F. M.; Ma, N. L.; Tsang, C. W. Competition between π and Non-π Cation-Binding Sites in Aromatic Amino Acids: A Theoretical study of alkali metal cation (Li^+, Na^+, K^+)-phenylalanine complexes. *Chem.--Eur. J.* **2004**, *10*, 1966-1976.

(47) Talley, J. M.; Cerda, B. A.; Ohanessian, G.; Wesdemiotis, C. Alkali metal ion binding to amino acids versus their methyl esters: Affinity trends and structural changes in the gas phase. *Chem.--Eur. J.* **2002**, *8*, 1377-1388.

(48) Dunbar, R. C.; Polfer, N. C.; Oomens, J. Gas-phase zwitterion stabilization by a metal dication. *J. Am. Chem. Soc.* **2007**, *129*, 14562-14563.

(49) Bertrán, J.; Rodríguez-Santiago, L.; Sodupe, M. The different nature of bonding in Cu^+-glycine and Cu^{2+}-glycine. *J. Phys. Chem. B* **1999**, *103*, 2310-2317.

(50) Constantino, E.; Rodríguez-Santiago, L.; Sodupe, M.; Tortajada, J. Interaction of Co^+ and Co^{2+} with glycine. A theoretical study. *J. Phys. Chem. A* **2005**, *109*, 224-230.

(51) Dunbar, R. C.; Polfer, N. C.; Berden, G.; Oomens, J. Metal ion binding to peptides: Oxygen or nitrogen sites? *Int. J. Mass Spectrom.* **2012**, *330-332*, 71-77.

(52) Rodgers, M. T.; Armentrout, P. B.; Oomens, J.; Steill, J. D. Infrared multiphoton dissociation spectroscopy of cationized threonine: Effects of alkali-metal cation size on gas-phase conformation. *J. Phys. Chem. A* **2008**, *112*, 2258-2267.

(53) Armentrout, P. B.; Rodgers, M. T.; Oomens, J.; Steill, J. D. Infrared multiphoton dissociation spectroscopy of cationized serine: Effects of alkali-metal cation size on gas-phase conformation. *J. Phys. Chem. A* **2008**, *112*, 2248-2257.

(54) Citir, M.; Stennett, E. M. S.; Oomens, J.; Steill, J. D.; Rodgers, M. T.; Armentrout, P. B. Infrared multiple photon dissociation spectroscopy of cationized cysteine: Effects of metal cation size on gas-phase conformation. *Int. J. Mass Spectrom.* **2010**, *297*, 9-17.

(55) Carl, D. R.; Cooper, T. E.; Oomens, J.; Steill, J. D.; Armentrout, P. B. Infrared multiple photon dissociation spectroscopy of cationized methionine: Effects of alkali-metal cation size on gas-phase conformation. *Phys. Chem. Chem.Phys.* **2010**, *12*, 3384-3398.

(56) Heaton, A. L.; Bowman, V. N.; Oomens, J.; Steill, J. D.; Armentrout, P. B. Infrared multiple photon dissociation spectroscopy of cationized asparagine: Effects of metal cation size on gas-phase conformation. *J. Phys. Chem. A* **2009**, *113*, 5519-5530.

(57) Wyttenbach, T.; Witt, M.; Bowers, M. T. On the stability of amino acid zwitterions in the gas phase: The influence of derivatization, proton affinity, and alkali ion addition. *J. Am. Chem. Soc.* **2000**, *122*, 3458-3464.

(58) Strittmatter, E. F.; Lemoff, A. S.; Williams, E. R. Structure of cationized glycine, gly·M^{2+} (M = Be, Mg, Ca, Sr, Ba), in the gas phase: Intrinsic effect of cation size on zwitterion stability. *J. Phys. Chem. A* **2000**, *104*, 9793-9796.

(59) Drayß, M. K.; Armentrout, P. B.; Oomens, J.; Schäfer, M. IR spectroscopy of cationized aliphatic amino acids: Stability of charge-solvated structure increases with metal cation size. *Int. J. Mass Spectrom.* **2010**, *297*, 18-27.

(60) Dunbar, R. C.; Hopkinson, A. C.; Oomens, J.; Siu, C. K.; Siu, K. W. M.; Steill, J. D.; Verkerk, U. H.; Zhao, J. Conformation switching in gas-phase complexes of histidine with alkaline earth ions. *J. Phys. Chem. B* **2009**, *113*, 10403-10408.

(61) Mattos, C. Protein-water interactions in a dynamic world. *Trends Biochem. Sci.* **2002**, *27*, 203-208.

(62) Chaplin, M. Do we underestimate the importance of water in cell biology? *Nat. Rev. Mol. Cell Biol.* **2006**, *7*, 861-866.

(63) Ball, P. Water as an active constituent in cell biology. *Chem. Rev.* **2008**, *108*, 74-108.

(64) Voet, D., Voet, J.G.: *Biochemistry*; 4th ed.; John Wiley and Sons: New York, **1990**.

(65) Bella, J.; Brodsky, B.; Berman, H. M. Hydration structure of a collagen peptide. *Structure* **1995**, *3*, 893-906.

(66) Polander, B. C.; Barry, B. A. A hydrogen-bonding network plays a catalytic role in photosynthetic oxygen evolution. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 6112-6117.

(67) Wu, C. C.; Chaudhuri, C.; Jiang, J. C.; Lee, Y. T.; Chang, H. C. Hydration-induced conformational changes in protonated 2,4-pentanedione in the gas phase. *Mol. Phys.* **2003**, *101*, 1285-1295.

(68) Kim, J. Y.; Ahn, D. S.; Park, S. W.; Lee, S. Gas phase hydration of amino acids and dipeptides: Effects on the relative stability of zwitterion vs. canonical conformers. *RSC Advances* **2014**, *4*, 16352-16361.

(69) Tortonda, F. R.; Pascual-Ahuir, J. L.; Silla, E.; Tunon, I. Why is glycine a zwitterion in aqueous solution? A theoretical study of solvent stabilising factors. *Chem. Phys. Lett.* **1996**, *260*, 21-26.

(70) Ding, Y.; Krogh-Jespersen, K. The glycine zwitterion does not exist in the gas phase: results from a detailed ab initio electronic structure study. *Chem. Phys. Lett.* **1992**, *199*, 261-266.

(71) Kim, J. Y.; Lee, Y.; Lee, S. Effects of microsolvation on the relative stability of zwitterionic vs. Canonical proline. *Chem. Phys. Lett.* **2014**, *608*, 177-185.

(72) Bush, M. F.; Prell, J. S.; Saykally, R. J.; Williams, E. R. One water molecule stabilizes the cationized arginine zwitterion. *J. Am. Chem. Soc.* **2007**, *129*, 13544-13553.

(73) Ai, H.; Bu, Y.; Han, K. Glycine-Zn^{+}/Zn^{2+} and their hydrates: On the number of water molecules necessary to stabilize the switterionic glycine-Zn^{+}/Zn^{2+} over the nonzwitterionic ones. *J. Chem. Phys.* **2003**, *118*, 10973-10985.

(74) Hwang, T. K.; Eom, G. Y.; Choi, M. S.; Jang, S. W.; Kim, J. Y.; Lee, S.; Lee, Y.; Kim, B. Microsolvation of lysine by water: Computational study of stabilized zwitterion. *J. Phys. Chem. B* **2011**, *115*, 10147-10153.

(75) Im, S.; Jang, S. W.; Lee, S.; Lee, Y.; Kim, B. Arginine zwitterion is more stable than the canonical form when solvated by a water molecule. *J. Phys. Chem. A* **2008**, *112*, 9767-9770.

(76) Wyttenbach, T.; Bowers, M. T. Hydration of biomolecules. *Chem. Phys. Lett.***2009**, *480*, 1-16.

(77) Gao, B.; Wyttenbach, T.; Bowers, M. T. Protonated arginine and protonated lysine: Hydration and its effect on the stability of salt-bridge structures. *J. Phys. Chem. B* **2009**, *113*, 9995-10000.

(78) Bush, M. F.; Saykally, R. J.; Williams, E. R. Hydration of the calcium dication: Direct evidence for second shell formation from infrared spectroscopy. *ChemPhysChem* **2007**, *8*, 2245-2253.

(79) Iino, T.; Ohashi, K.; Inoue, K.; Judai, K.; Nishi, N.; Sekiya, H. Infrared spectroscopy of $Cu^{+}(H_2O)_n$ and $Ag^{+}(H_2O)_n$: Coordination and solvation of noble-metal ions. *J. Chem. Phys.* **2007**, *126*.

(80) Inokuchi, Y.; Ohshim, K.; Misaizu, F.; Nishi, N. Infrared photodissociation spectroscopy of $[Mg\cdot(H_2O)_{1-4}]^{+}$ and $[Mg\cdot(H_2O)_{1-4}\cdot Ar]^{+}$. *J. Phys. Chem. A* **2004**, *108*, 5034-5040.

(81) Kołaski, M.; Lee, H. M.; Choi, Y. C.; Kim, K. S.; Tarakeshwar, P.; Miller, D. J.; Lisy, J. M. Structures, energetics, and spectra of aqua-cesium (I) complexes: An ab initio and experimental study. *J. Chem. Phys.* **2007**, *126*.

(82) Walters, R. S.; Pillai, E. D.; Duncan, M. A. Solvation dynamics in $Ni^+(H_2O)_n$ clusters probed with infrared spectroscopy. *J. Am. Chem. Soc.* **2005**, *127*, 16599-16610.

(83) Polfer, N. C.; Oomens, J.; Dunbar, R. C. IRMPD spectroscopy of metal-ion/tryptophan complexes. *Physical Chemistry Chemical Physics* **2006**, *8*, 2744-2751.

(84) Dunbar, R. C.; Steill, J. D.; Oomens, J. Cationized phenylalanine conformations characterized by IRMPD and computation for singly and doubly charged ions. *Phys. Chem. Chem. Phys.* **2010**, *12*, 13383-13393.

(85) Bush, M. F.; Forbes, M. W.; Jockusch, R. A.; Oomens, J.; Polfer, N. C.; Saykally, R. J.; Williams, E. R. Infrared Spectroscopy of Cationized Lysine and ε-N-methyllysine in the Gas Phase: Effects of Alkali-Metal Ion Size and Proton Affinity on Zwitterion Stability. *J. Phys. Chem. A* **2007**, *111*, 7753-7760.

(86) Corral, I.; Lamsabhi, A. M.; Mó, O.; Yáñez, M. Infrared spectra of charge-solvated versus salt-bridge conformations of glycine-, serine-, and cysteine-Ca^{2+} complexes. *Int. J. Quantum Chem.* **2012**, *112*, 2126-2134.

(87) Balaj, O. P.; Kapota, C.; Lemaire, J.; Ohanessian, G. Vibrational signatures of sodiated oligopeptides (GG-Na^+, GGG-Na^+, AA-Na^+ and AAA-Na^+) in the gas phase. *Int. J. Mass Spectrom.* **2008**, *269*, 196-209.

(88) Dunbar, R. C.; Steill, J. D.; Polfer, N. C.; Oomens, J. Peptide length, steric effects, and ion solvation govern zwitterion stabilization in barium-chelated di- and tripeptides. *J. Phys. Chem. B* **2009**, *113*, 10552-10554.

(89) Prell, J. S.; Demireva, M.; Oomens, J.; Williams, E. R. Role of sequence in salt-bridge formation for alkali metal cationized glyArg and argGly investigated with IRMPD spectroscopy and theory. *J. Am. Chem. Soc.* **2009**, *131*, 1232-1242.

(90) Abirami, S.; Wong, C. C. L.; Tsang, C. W.; Ma, N. L. Dissociation of alkaliated alanine in the gas phase: The role of the metal cation. *Chem.--Eur. J.* **2005**, *11*, 5289-5301.

(91) Citir, M.; Hinton, C. S.; Oomens, J.; Steill, J. D.; Armentrout, P. B. Infrared multiple photon dissociation spectroscopy of cationized histidine: Effects of metal cation size on gas-phase conformation. *J. Phys. Chem. A* **2012**, *116*, 1532-1541.

(92) Watson, H. M., Vincent, J. B., Cassady, C. J. Effects of transition metal ion coordination on the collision-induced dissociation of polyalanines. *J. Mass Spectrom.* **2011**, *46*, 1099-1107.

(93) Chu, I. K.; Guo, X.; Lau, T. I.; Siu, K. W. M. Sequencing of argentinated peptides by means of electrospray tandem mass spectrometry. *Anal. Chem.* **1999**, *71*, 2364-2372.

Chapter 2 . Methods

2.1. Experimental and Computational Methods

In this work the chemistry of gas phase ions is studied by mass spectrometric techniques and computational methods. The purpose of this chapter is to introduce the instrumentation and techniques employed in this thesis and provide a comprehensive explanation of each method.

For analytical purposes, the Fourier Transform Ion Cyclotron Resonance (FT-ICR) mass spectrometer is known to be powerful because of high resolving power and mass accuracy. More importantly, the presence of a trapping cell as a mass analyzer in FT-ICR makes it a versatile instrument for study of structure, energetics, and reactions of ions in the gas phase. The ability to trap ions for long periods of time make it possible to study the fragmentations and reactions of the ions by different dissociation techniques such as infrared multi-photon dissociation (IRMPD) and low energy collision-induced dissociation (CID) which were briefly introduced in Chapter 1.

2.2. Instrumentation

The main experimental part of this thesis was done using a Bruker ApexQe7.0 FT-ICR-MS in our lab at Memorial University. A photo of this instrument is shown in Figure 2.1.

Figure 2.1. Photo of Bruker Apex-Qe7.0 T Fourier transform ion cyclotron mass spectrometer FT-ICR MS and a tabletop OPO laser at Memorial University.

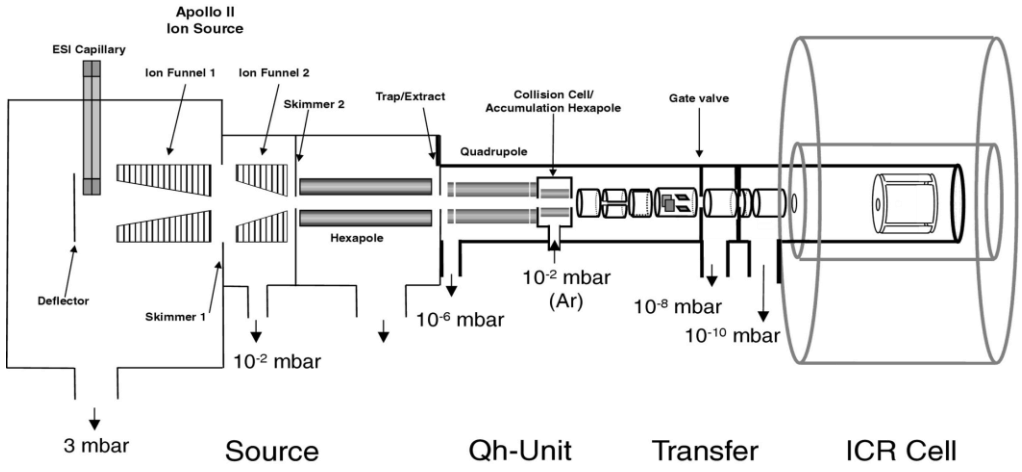


Figure 2.2. A detailed schematic of a FT-ICR mass spectrometer

Figure 2.2 shows a detailed schematic of the FT-ICR mass spectrometer in the lab at MUN.[1] There are four main subunits in the instrument: the ion source, quadrupole/hexapole region, ion transfer optics, and ICR cell. Gaseous ions are produced from aqueous solutions by means of electrospray ionization (ESI). Desolvated ions become accelerated through a series of mass filters consisting of quadrupoles and hexapoles with relatively low pressures (10^{-6} mbar) compared to ion source (3 mbar). At the hexapole collision cell or accumulation cell—held at ~10^{-2} mbar—it is possible to prepare hydrated ions by introducing water through a microvalve and decreasing the collision gas (Ar) pressure.[1] Bare or hydrated ions stored in the accumulation cell are directed to the ICR cell where they experience a very low pressure (10^{-10} mbar) and become isolated and detected. In the ICR cell ions can be trapped for long periods of time. The ICR cell can also act as a reaction chamber in which ions undergo fragmentation, initiated by energetic collisions or the absorption of IR photons, and ion-molecule reactions. It is also possible to introduce other small, neutral molecules such as CO or CO_2 to the ICR cell which can result in a pressure increase (10^{-10} - 10^{-5}) in the cell for the purpose of ion-molecule reaction studies.

2.2.1. Electrospray ionization (ESI)

The ion source used in the experimental part of this work is an Apollo II electrospray source. Electrospray ionization is a soft ionization method which produces ions from solution.[2-4] One of the features of ESI that has gained attention, especially from the biological point of view, is its ability to transfer large biomolecules in their intact

form to the gas phase without destroying their structure since even non-covalent interactions can be conserved during the electrospray ionization process.[5,6]

Figure 2.3[7] shows the ionization process that occurs inside an electrospray ion source. First a solution of the analyte is produced with conventional solvents such as water, methanol or acetonitrile. This solution then is introduced through a capillary to an electric field that is produced by applying a 2-4 kV potential between the capillary opening and a counter electrode.[3] As a consequence of the electric field, ions of the same charge (for example cations in positive ionization mode) spread through the tip because of Coulombic repulsion forces and stand as a liquid meniscus pointing outward forming a cone shape at the capillary tip, also known as a Taylor cone.[2,8] The number of positive charges increases at the tip until the electrostatic repulsion between the positive charges exceeds the surface tension. This eventually results in ejection of small droplets of ions toward the mass analyzer.[4,9]

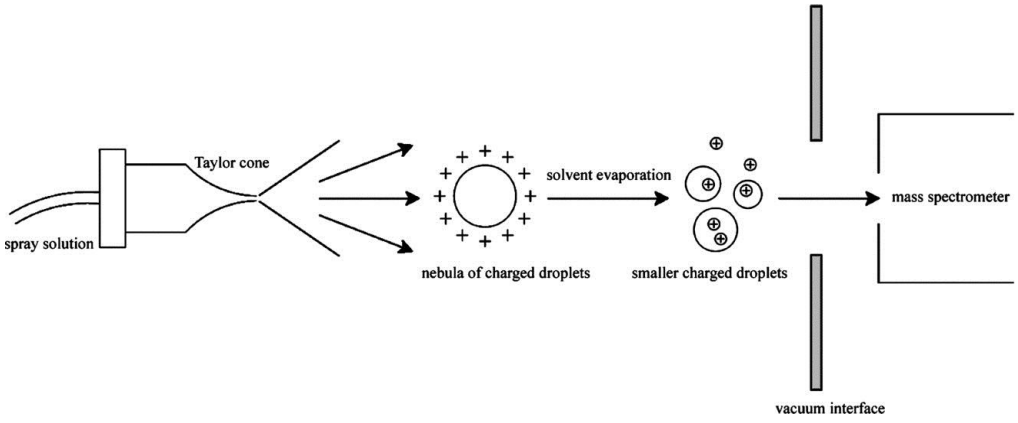


Figure 2.3. Schematic of electrospray ionization process. Reprinted with permission from *Appl. Spectrosc. Rev.* **2009**, *44*, 231-244. Copyright (2009) Taylor and Francis.[7]

Solvent evaporation from charged droplets can be aided by using volatile solvents or increasing the temperature inside the source. Continuous evaporation of solvent results in shrinkage of the droplet size and, at the same time, increasing the charge repulsion at the droplet surface. In general, there are two opposite forces acting on the droplet's stability: charge repulsion and surface tension. Coulombic fission occurs when the charge repulsion surmounts the surface tension and this produces primary charged droplets. This point is also known as the Rayleigh limit defined in the Equation 2.1.[2]

$$Q^2 = 64\pi^2\varepsilon_0\lambda r^3 \qquad 2.1$$

In equation 2.1, Q is the limiting charge, r is the droplet radius, λ is the surface tension and ε_0 is the electrical permittivity of the vacuum. The exact process for formation of gas phase ions from these droplets is unknown but it is suggested that there are two main mechanisms that govern this process: the ion evaporation model (IEM)[10] and the charge residue model (CRM).[11] The IEM predicts that gas phase ions form directly from the droplets. Ions become evaporated from the solvated droplets at the Rayleigh limit, when the droplet size decreases to around 20 nm (Figure 2.4(a)). In the charge residue model it is suggested that droplets undergo evaporation and fission cycles until they reach the point in which there is only one analyte ion left in the droplet. These droplets go through further desolvation leaving the charge behind which leads to formation of multiply charged ions (Figure 2.4(b)). The charge residue model is known to be the dominant mechanism for large biological molecules whereas the ion evaporation model is the mechanism for formation of gas phase ions of small molecules.[3,12] For a polyprotonated peptide or protein, for example, it is possible for one or more protonated sites to reach to

the droplet's surfaces which increase the Coulombic repulsion and contributes to instability of the droplet. Eventually a Taylor cone becomes formed that leads to small multiply charged droplet ions consisting of a single analyte.[2]

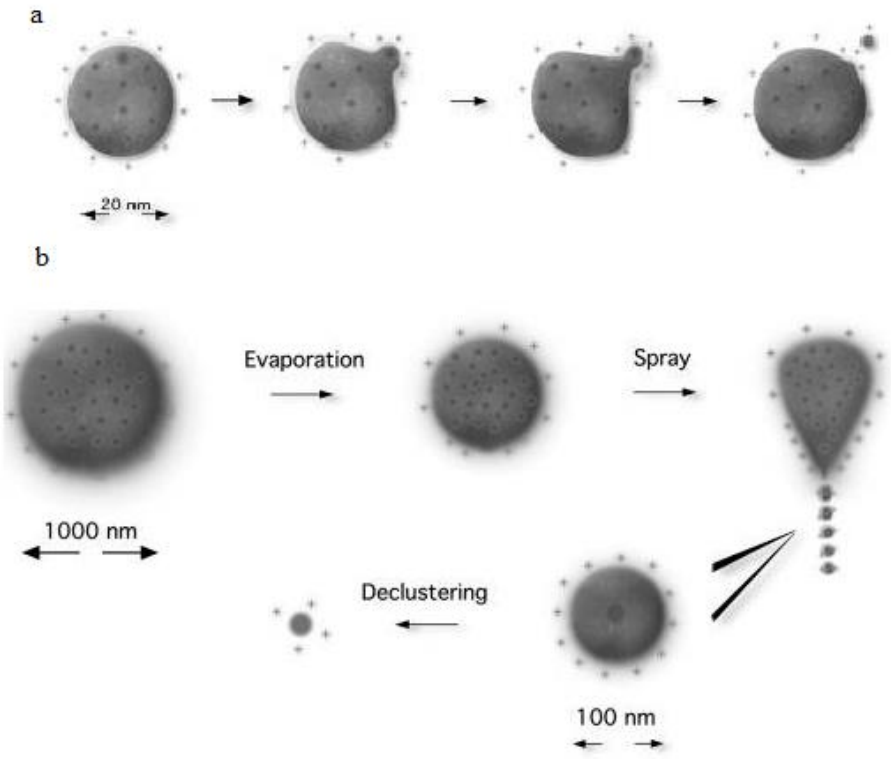


Figure 2.4. Mechanism of ionization in (a) ion evaporation model and (b) charged residue model. Reprinted with permission from *Analyst*, **2005**, *130*, 18–28. Copyright (2005) The Royal Society of Chemistry.[13]

2.2.2. Fourier Transform Ion Cyclotron Resonance Mass Spectrometry

The ion trap inside an FT-ICR mass spectrometer is a penning trap that is also called the ICR cell and is used to store ions for long periods of time. In Figure 2.5 a schematic diagram of a cylindrical ICR cell is shown.[13] The ICR cell is composed of three pairs of plates: end cap trapping plates, excitation plates and detection plates. In a cylindrical ICR cell, excitation and detection plates alternate with each other. Trapping plates are located at either ends of the ICR cell and an electric potential is applied to these plates prevent the loss of ions in the axial dimension along the magnetic field. The powerful superconducting magnetic field, in which the ICR cell sits, and the potential applied to the trapping plates confine ions inside the ICR cell for long periods of time.[14,15]

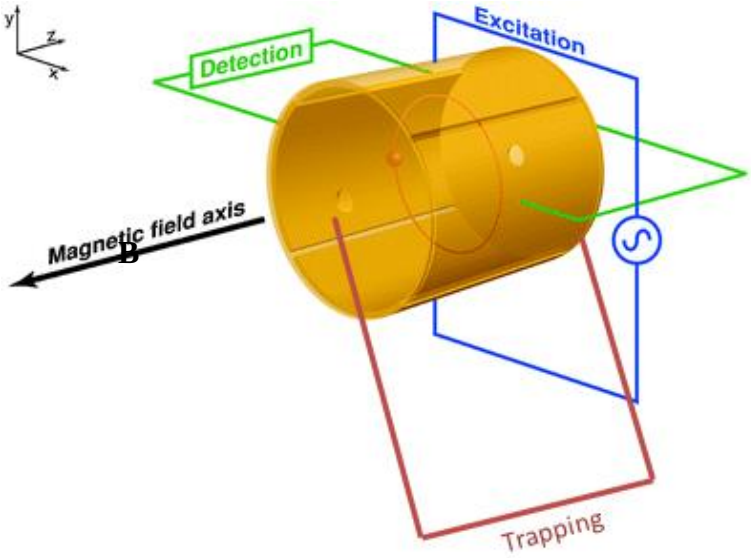


Figure 2.5. A schematic representation of excitation, detection and trapping electrodes in an ICR cell. Reprinted with permission from *Analyst*, **2005**, *130*, 18–28. Copyright (2005) The Royal Society of Chemistry.[13]

Ions become trapped inside the spatially uniform magnetic field inside the ICR cell. As a result, ions that are moving inside this magnetic field are subject to the Lorentz force (Equation 2.2).[16,17]

$$\vec{F} = qv \times \vec{B} \qquad 2.2$$

In which *F* is the Lorenz force and is perpendicular to the magnetic field, *B*, and cause the ions with velocity, v, to revolve in a circular path (Figure 2.6), also perpendicular to the magnetic field. This circular motion is known as cyclotron motion. Besides the Lorentz force, ions are also under the effect of centrifugal force (F'), Equation 2.3. The ion's cyclotron motion trajectory is stabilized only if the Lorentz force is equal to the centrifugal force, Equation 2.4 and 2.5.

$$F' = \frac{mv^2}{r} \qquad 2.3$$

$$F = F' \qquad 2.4$$

$$qvB = \frac{mv^2}{r} \qquad 2.5$$

For an ion with mass *m*, angular frequency, $\omega = \frac{v}{r}$, and charge, $q = ze$ it can be shown that the mass to charge ratio (*m/z*) is directly proportional to the magnetic field B and inversely related to the angular frequency.

$$\frac{m}{z} = \frac{eB}{\omega} \qquad 2.6$$

Equation 2.6 shows that cyclotron frequencies of all ions with the same *m/z* are identical regardless of their velocity.

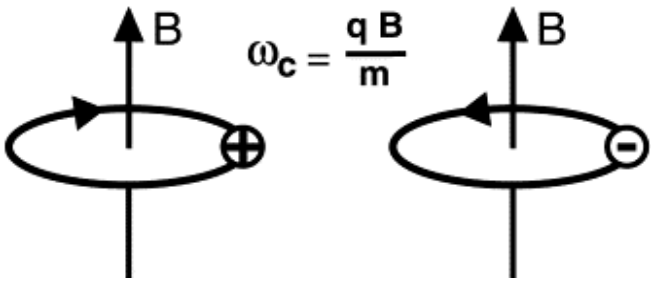


Figure 2.6. Cyclotron motion of a cation and anion in a uniform magnetic field of magnitude B. Reprinted with permission from *Int. J. Mass Spectrom.* **2002**, *215*, 59-75. Copyright (2002) Elsevier Ltd.[17]

Ambient temperature cyclotron motion is not useful on its own since the cyclotron radius at low kinetic energies is not large enough to be detected for most ions. Also ions of the same *m/z* move incoherently inside the ICR cell. Excitation of ions to a larger radius happens by applying a radio frequency (RF) voltage across the excitation plates. Ions of the same *m/z* have identical cyclotron frequencies and if their cyclotron frequency is in resonance with the applied RF voltage they become excited and gain kinetic energy (Figure 2.7).[17] By quickly sweeping the RF frequencies, ions of all *m/z* can be excited almost instantaneously. After absorption of energy, the ions of the same *m/z* move coherently in a larger radius packet, where they are closer to the detection plates. It is important to remember that the frequency before excitation is the same following excitation. The higher kinetic energy only translates into motion of larger radius.

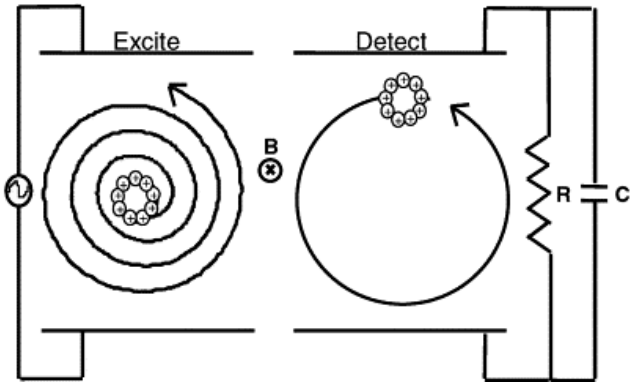


Figure 2.7. Excitation of ions moving incoherently (left) to a larger radius as a coherent packet to be detected. Reprinted with permission from *Int. J. Mass Spectrom.* **2002**, *215*, 59-75. Copyright (2002) Elsevier Ltd.[17]

When ions pass close to a detection plate, they induce an image charge in that plate. Since the two trapping plates are connected by a circuit, an image current is produced when the ion packet passes by one plate and travels closer to the second plate. The continuous circular motion of the ions by the detection plates produces a sinusoidal (waveform) signal with a frequency equal to the cyclotron frequency of the ions and amplitude that is reflective of the number of ions. This signal is called a time domain signal (Figure 2.8). Fourier transformation of the time domain signals results in a frequency domain signal which can be converted to a mass spectrum since *m/z* is inversely proportional to the cyclotron frequencies (Equation 2.6).

Besides excitation and detection of ions it is also possible to isolate an ion of interest by ejecting other ions from ICR cell. This happens by applying RF frequencies resonant with the cyclotron frequencies of all ions besides that of the ion being isolated, and exciting them to a radius larger than the cell dimension where they collide with the

cell walls, become neutralized and pumped away. Furthermore, carefully applying an RF can increase the translational energy of ions for collisional activation studies.[16,17]

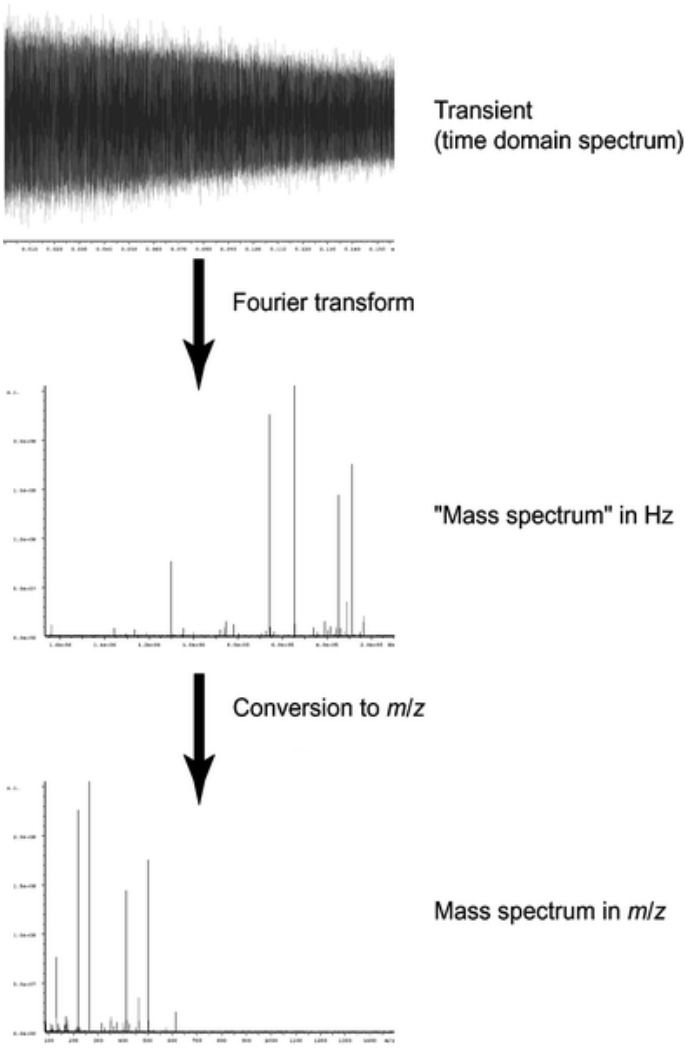


Figure 2.8. Time domain spectrum is converted to a frequency domain using Fourier transformation. Mass spectrum is derived from the Frequency domain spectrum. Reprinted with permission from *Analyst*, **2005**, *130*, 18–28. Copyright (2005) The Royal Society of Chemistry.[13]

2.3. Tandem mass spectrometry techniques

2.3.1. Infrared Multiple Photon Dissociation

2.3.1.1. History

Studies of gaseous ion fragmentation have become possible by applying low to medium powered IR lasers to the ions inside ion traps. The absorption of multiple IR photons by trapped ions, results in their dissociation and is known as infrared multiple photon dissociation (IRMPD). In early studies, CO_2 lasers were the first IR photon sources that were developed for gas phase ion unimolecular dissociation.[18-21] In 1971 Isenor and his coworkers investigated the dissociation of molecular gases such as CCl_2F_2, CH_3Cl, NH_3 and hydrocarbons with a high power (0.5MW) CO_2 laser.[19] They showed that absorption of multiple IR photons causes dissociation of these molecules to products that are in their electronic excited states. Beauchamp and his coworkers were pioneers in the study of the unimolecular dissociation of trapped ions induced by low powered IR CO_2 lasers.[22-24] In one study they applied a continuous wave CO_2 laser tuned to 10.6 μm to dissociate the proton bound dimer of diethyl ether. These dimer ions were activated by absorption of multiple IR photons which results in formation of $(C_2H_5)_2OH^+$ as the only product (Figure 2.9).[23]

$$[(C_2H_5)_2O)_2]H^+ + nh\nu \rightarrow (C_2H_5)_2O + (C_2H_5)_2OH^+ \qquad 2.7$$

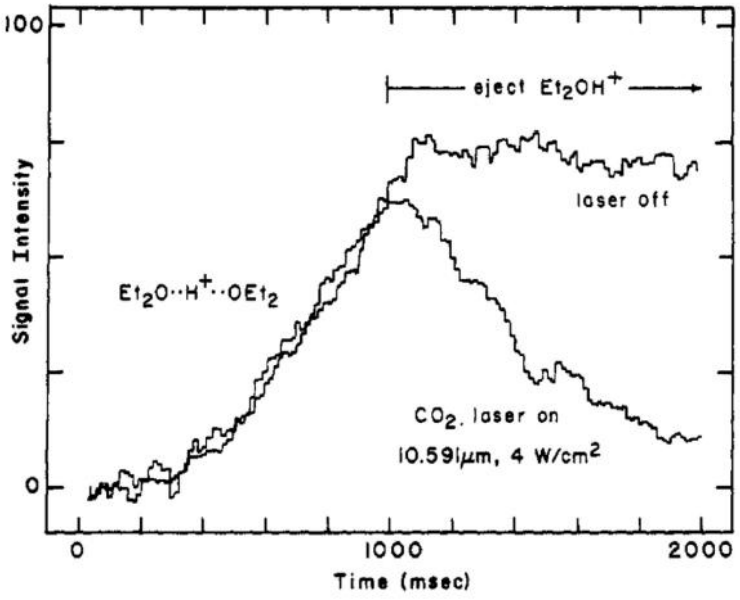


Figure 2.9. Ion intensity of parent and product ion for dissociation of proton dimer of diethylether. Reprinted with permission from *J. Am. Chem. Soc.* **1978**, *100*, 3248-3250. Copyright (1978) American Chemical Society.[23]

In another study Bensimon and coworkers could distinguish between the different isomers of heptene by investigating their dissociation rate constants as a function of IR laser wavenumber. They showed that the formation of same fragment ion happens with a different rate constant at a specific wavenumber for different isomers (Figure 2.10).[25]

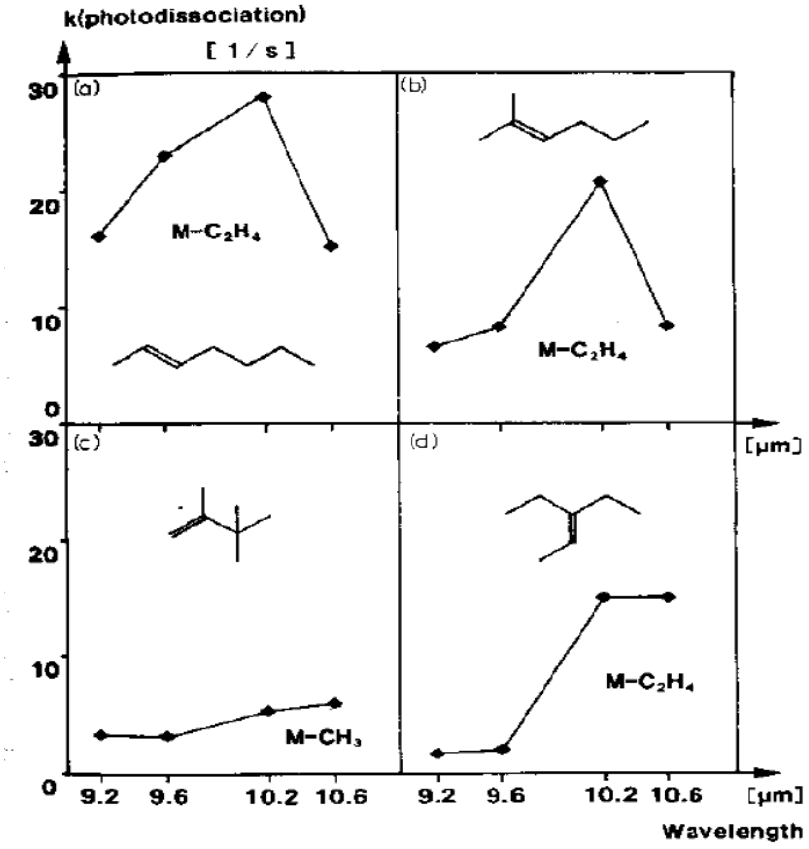


Figure 2.10. Fragmentation of isomeric heptene cations by CO_2 laser. Reprinted with permission from *Int. J. Mass Spectrom. Ion Processes* **1986**, *72*, 125-135. Copyright (1986) Elsevier Ltd.[25]

2.3.1.2. IRMPD spectroscopy

Due to low ion density in an ion trap and the insensitivity of conventional IR spectroscopy, it is hard to measure IR absorption of ions in the gas phase. IRMPD spectroscopy using intense tuneable IR lasers has become a powerful technique in which it is possible to monitor the fragmentation of ions as a function of IR wavelength and it can provide useful structural information.

2.3.1.3. The IRMPD mechanism

The IRMPD process starts with absorption of an IR photon. Typical IRMPD studies focus on the mid-infrared where in the 400-4000 cm^{-1} region corresponding to 5-48 kJ mol^{-1}. In this region of the infrared the radiation is in resonance with a vibrational mode of the trapped ion and corresponds to a transition from the ground vibrational state to the first vibrationally excited state (Figure 2.9) state $v=0 \rightarrow 1$. In a harmonic oscillator the energy difference between each successive vibrational state is identical (Figure 2.11(a)) so that sequential absorption of monochromatic infrared photons is possible. However in reality, chemical systems are anharmonic and the spacing between sequential energy levels decrease (Figure 2.11(b)) and become non-resonant with the frequency of the monochromatic radiation source. An important phenomenon in the IRMPD process is depopulation of the $v=1$ to the $v=0$ which happens through intramolecular vibrational-energy redistribution (IVR). The energy of the photon becomes distributed and stored in other internal modes of the molecular ion. As a result, the vibrational ground state is free again to absorb another photon.[26,27] This process continues until the internal energy of the ion is in excess of the lowest energy dissociation threshold, Figure 2.12, and is followed by dissociation.

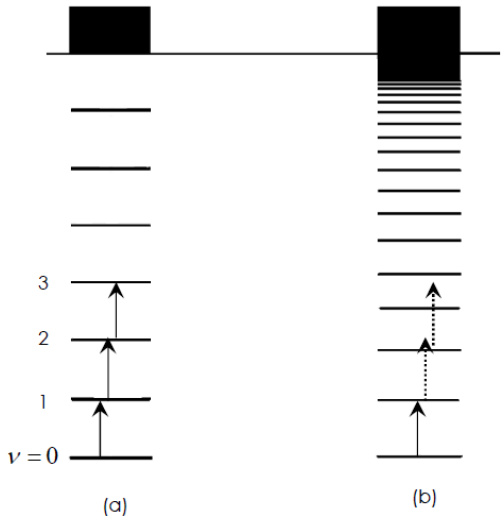


Figure 2.11. A Schematic of vibrational energy levels for an ideal/harmonic (a) and real/anharmonic (b) system. Modified from Fridgen and McMahon; "IRMPD". In *The Encyclopedia of Mass Spectrometry*; Gross, M. L., Caprioli, R. M., Eds.; Elsevier: New York, 2005; Vol. 4; pp 327. Copyright (2005) Elsevier.

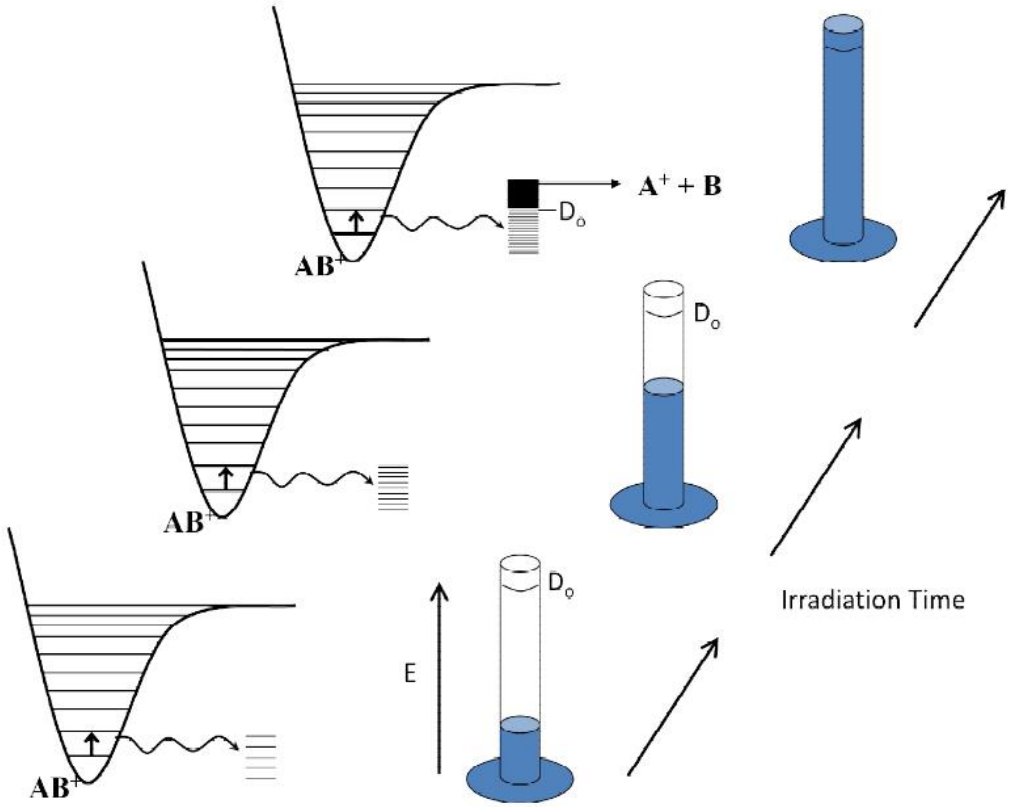


Figure 2.12. Unimolecular dissociation by IRMPD and the IVR process, Energy of the first absorbed photon is distributed through other vibrational modes. Internal energy of ion increases gradually until it passes the threshold energy for the lowest energy dissociation. Reprinted with permission from Dr. Travis D. Fridgen (Memorial University).

In IRMPD spectroscopy the dissociation of the parent ion and formation of product ion are monitored as a function of the IR laser frequency. The intensities of parent ion and product ions are recorded to give the IRMPD efficiency versus the IR wavenumber.

$$IRMPD\ efficiency = -\ln\left[\frac{I_p}{I_p + \sum_i I_{frag(i)}}\right] \quad 2.8$$

Where I_p and I_{frag} are the intensity of the parent and the fragment ions, respectively.

2.3.1.4. Lasers as IR photon source

2.3.1.4.1. CO_2 laser

The CO_2 laser is one of the most powerful molecular gas lasers and was introduced by Patel in 1964[28] and 1965.[29] It uses a mixture of CO_2, N_2 and He as the gas medium. The CO_2 laser is based on the transition between vibrational energy levels of a CO_2 molecule which falls in the vibrational-rotational mid-infrared region at 9.4 and 10.6 μm and produces outputs of up to 100 kW. In the first step N_2 molecule becomes vibrationally excited by electron impact and because of its homonuclear nature, N_2 stays vibrationally excited until it collides with a CO_2 molecule, transferring its energy to the ν_3 asymmetric stretching vibrational mode of the CO_2. In Figure 2.13 the energy level diagram of vibrational energy levels are shown for CO_2 and N_2. The vibrationally excited CO_2 relaxes by emitting photons of 9.6 and 10.6 μm retaining some of its energy in the ν_2 bending mode and ν_1 symmetric stretching modes, respectively. Collision of CO_2 with a He molecule causes CO_2 to return back to the ground state.

The CO_2 lasers operates in a narrow range of IR frequencies and this limits their application for spectroscopy purposes in which a wide range of frequencies need to be scanned. Because of their high intensities they have emerged as useful tools in industry for material laser cutting, especially metal cutting, welding, surface hardening, and surface metal alloying etc.[30]

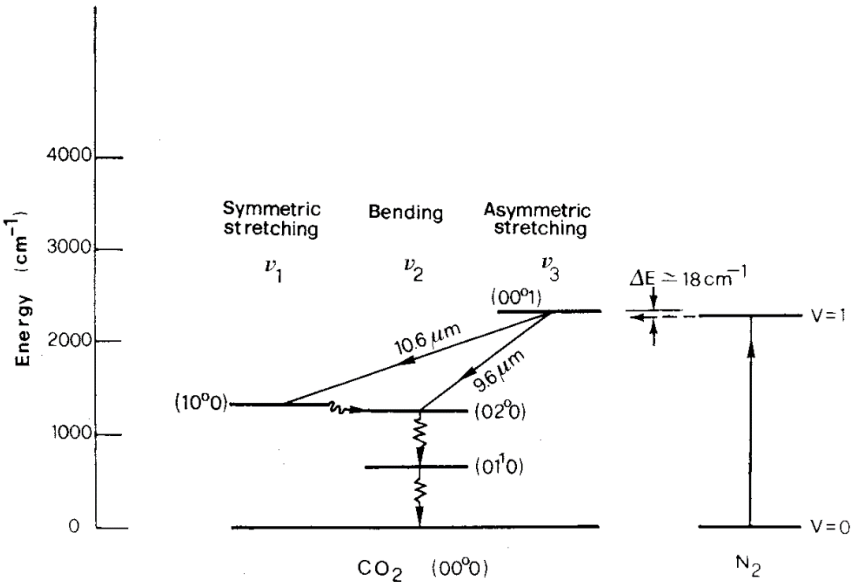


Figure 2.13. The lowest energy vibrational transitions of a N_2 and CO_2 molecule. Reprinted with permission from *SpringerLink* **2010**, Chapter 10, 444-454, Orazio Svelto. Copyright (2010) Springer Science and Business Media.[30]

2.3.1.4.2. Free Electron Laser (FEL)

The free electron laser facility at Centre Laser Infrarouge d'Orsay (CLIO) was used for the IRMPD spectroscopy studies in the 900-1900 cm^{-1} IR region. Free electron lasers work in a wide range of wavelengths from 5 to 110 μm and produce coherent infrared radiation. The source of radiation in free electron lasers is from oscillations of free electrons inside an alternating magnetic field (known as an undulator) which results in the wiggling motion of electrons inside the magnetic field. Before entering the magnetic field gigaamperes of electrons are accelerated to megaelectronvolts by an RF linear accelerator. The accelerated electrons, known as a relativistic electron beam (REB),

enter the changing magnetic field and, due to the Lorentz force, gain an oscillating motion. Figure 2.14 shows a schematic of free electron laser and the electrons pathway inside the oscillating magnetic field.[31] Oscillating electrons produce synchrotron radiation whose wavelength is a function of electron velocity, frequency, and the magnet period.[30] The mirrors at each end of the undulator amplify the laser beam radiation.

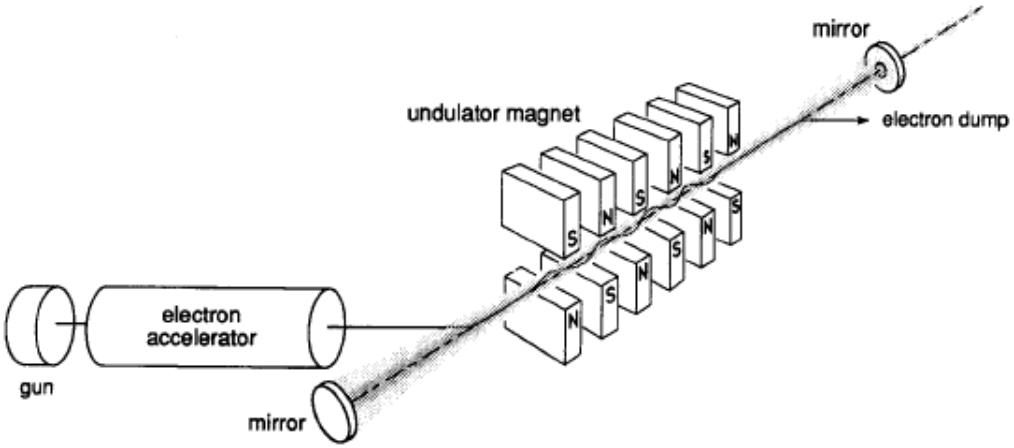


Figure 2.14. Free electron laser diagram. Electron beam is introduced to the undulator by electron accelerator. Electrons oscillate inside the undulator and emit photons. Mirrors at each end amplify the electron photon emission. Reprinted with permission from *Infrared Phys. Tech.* **1995**, *36*, 297-308. Copyright (1995) Elsevier.[31]

2.3.1.4.3. Optical Parametric Oscillators

Optical parametric oscillators are tuneable lasers that produce continuously tunable radiation over a wide range of IR wavelength. In the optical parametric oscillator used at Memorial University, a high frequency (1064 nm fundamental line) laser beam is supplied by a Nd:YAG laser. The input or pump beam with the frequency of ω_p is

introduced to an optical crystal such as lithium niobate ($LiNbO_3$) or potassium titanyl phosphate (KTP) where it is converted to two lower energy signal and idler beams, with frequencies at ω_s and ω_i, respectively.[32]

$$\omega_p = \omega_s + \omega_i \qquad 2.9$$

The signal and idler beams become amplified by two reflective mirrors, an optical cavity in which the crystal exists. The frequency of the output radiation, the signal and idler beams, can be tuned by changing the crystal angle.

2.3.1.4.4. Laser used in this work

Figure 2.1 shows the tabletop setup for the OPO laser in the Laboratory for the Study of the Energetics, Structures, and Reactions of Gaseous Ions at Memorial University. IRMPD spectroscopy experiments were performed by using two OPO laser setups. In the first OPO setup, used for experiments reported in chapter 3, a Brilliant B Nd:YAG laser is employed which emits in 6 ns (850 mJ) pulses at 10 Hz. The 1064 nm fundamental line of the Brilliant B Nd:YAG pumps light to a KTP crystal (potassium titanyl phosphate, $KTiOPO_4$) and produces two perpendicular beams: the signal and idler. The signal and idler beams are between 1.5 – 2.1 and 2.1 – 3.5 μm, respectively. The OPO beam is amplified by passing through a CaF_2 focusing window and BaF_2 window to the center of the ICR cell for IRMPD studies. This OPO has a 3 cm^{-1} bandwidth, and the output pulse varies as a function of the wavelength used. The most efficient IR pulse occurs at 2.6 μm and the pulse intensity decreases as the wavelength increases, Figure 2.15.

Currently at Memorial, a new OPO laser (LaserSpec) has been installed and was used for the works in Chapters 4 and 5. The laser is tuneable from 2200 to 7000 cm^{-1}, encompassing the C-H/N-H/O-H stretching region and beyond, with bandwidth of 2 cm^{-1} and is built around a periodically poled lithium niobate ($LiNbO_3$) crystal. The $LiNbO_3$ crystal is pumped using a diode-pumped 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 for the experiments reported in Chapter 4 and 5.

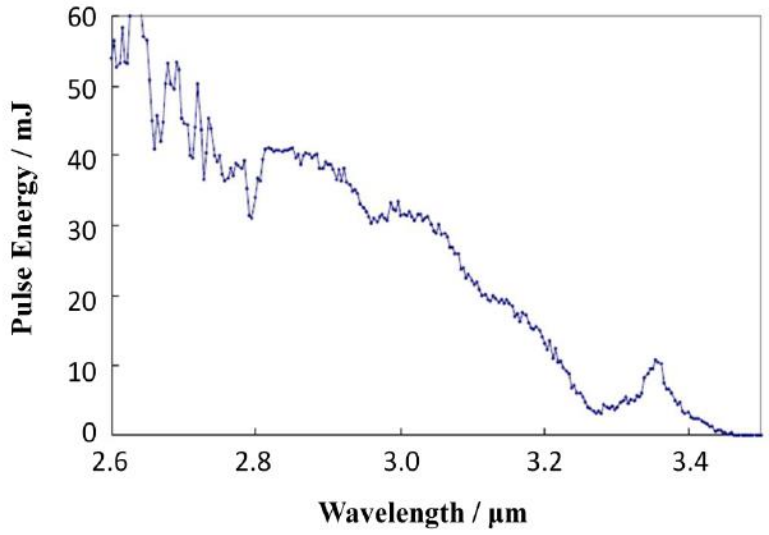


Figure 2.15. OPO laser signal intensity as a function of wavelength.

2.3.2. Collision-induced dissociation

Collision-induced dissociation (CID) is another ion activation/dissociation technique for structural studies of molecules. This method is based on the inelastic collision of translationally excited ions with slow moving, chemically inert gas molecules (targets) such as noble gases inside the ion trap. The kinetic energy of trapped ions

increases by applying an RF frequency across the excitation plates in the ICR cell. The translational energy of the accelerated ions becomes converted to internal energy of the ion during collisions with Ar, N_2 or He as target molecules, although in this work, only Ar targets were used. As a result, the internal energy of ions increases and passes the threshold energy for ion dissociation. The amount of convertible energy available after each collision is defined as the center of mass energy, Equation 2.10.[33-35]

$$E_{com} = E_{lab}\left(\frac{N}{N+m_i}\right) \qquad 2.10$$

Where E_{lab} equals the kinetic energy (lab frame) of the ion inside the ion trap and N is the mass of neutral gas molecule and m_i refers to the ion mass.

The CID method is typically divided into three categories of high, low, and very low energy, based on the translational energy of the ions.[33,35] In high energy CID the translational energy of ions is on the order of keV. The high amount of internal energy provided from one collision of these fast ions can easily overpass the threshold energy required for reactions such as covalent bond cleavage. In fact, the internal energy of the ions following a single collision is usually high enough that very high energy fragmentations or multiple fragmentations occur. The mass analyzers that are commonly used for high energy CID are time of flight (TOF) and sector instruments.

Low energy CID corresponds to translational energies in the 1-100 eV range. More collisions (10-100) are required in low energy CID to affect fragmentation of the precursor ion. The main fragmentation pathway is the lowest energy one since ions are slowly vibrationally excited until the internal energy of the ions is high enough that fragmentation effectively competes with further ion activation. The difference in the

fragmentation products in each CID energy regime is of great advantage since complementary structural information can be obtained about the ion of interest through comparison of dissociation reaction pathways. The low energy CID is typically performed in quadrupole mass analyzers such as triple quadrupoles or quadrupole hybrids and the inert gas used is typically Ar or N_2.

In the FT-ICR cell, sustained off-resonance irradiation collision-induced dissociation (SORI/CID) is commonly used and is classified as a very low energy CID process meaning that very small packets of energy, around hundreds of meV to 10 eV, are transferred to ions upon each collision.[36] Ions trapped in an ICR cell become excited by applying a RF pulse slightly higher (off-resonance) than the resonant ion cyclotron frequency. This cause ions to go through acceleration/deceleration cycles which provides longer activation times from 100 milliseconds up to seconds. In this regard, the internal energy of the ions increases very slowly and over a large number of collisions. The SORI/CID is known as a slow heating method of exciting ions, and mostly results in the lowest energy dissociation reaction pathways as well as rearrangement reactions which occur corresponding to the longer activation time and very low energy collisions in this method. The amount of energy for each collision in SORI/CID can be obtained through Equation 2.11.[37]

$$E_{lab} = \frac{\beta^2 q^2 V^2_{p-p}}{32\pi^2 m d^2 \Delta v^2} \qquad 2.11$$

In which β is a geometrical factor and d is the diameter of the ICR cell (0.92 and 6×10^{-2} m in this work, respectively), q and $V_{p\text{-}p}$ are the ion charge and peak to peak voltage,

respectively. Also m is the ion mass and Δv (500 Hz or s^{-1}) is the difference between off-resonance and on-resonance frequencies.

2.4. Computational methods

Computational quantum chemistry has grown to be a very powerful tool for the study of ions in the gas phase over the years. Computational modelling and simulation of complex biological systems has become possible due to ongoing development of powerful computers and computational methods. Very useful structural and thermochemical information can be obtained with different theoretical methods and it is of importance for systems that their empirical data is not available or hard to be investigated experimentally. It has also become important in gas-phase ion chemistry to support experimental findings with computed values, especially in areas such as structure elucidation and determining thermochemical parameters of chemical processes occurring in the mass spectrometer.

Computational techniques such as ab initio and density functional theory (DFT) are examples of these methods that have been applied for computing structures, binding energies, infrared spectra, and exploring potential energy surfaces for unimolecular dissociations or ion-molecule reactions. Three dimensional structural investigations of large biomolecules such as proteins have also become plausible by computational methods such as molecular dynamics in which the atoms' motions and interactions are simulated by numerical methods. A better understanding of molecular structure can be achieved by a combination of theoretical and experimental results. For example an

IRMPD spectrum can provide structural information such as presence of certain functional groups and when it is compared to the calculated IR spectrum of suggested optimized structures provides some level of comfort with the structure of the ion under study. In this work we have used computational techniques to calculate the lowest energy geometries, relative thermochemistries, frequencies, and infrared spectra of metal bound complexes of dipeptides in the gas phase.

2.4.1. Electronic structure calculation

In a molecular ion the arrangement of nuclei and electrons and their interactions determine the overall electronic structure and as a result, energy of the ions. In order to find the lowest energy ion structure, the potential energy surface is explored by starting with many different isomers obtained either by chemical intuition or molecular dynamics simulations (simulated annealing). There are different computational methods that are employed to calculate the ion electronic structure by solving numerical equations e.g. Schrödinger equation.[38] The differences between these methods are the approximations that are made for a many electron system. In this work density functional theory is employed to compute molecular properties. A set of structures are introduced and become optimized by density functional theory. The frequency calculations are performed on the optimized structures with the same level of theory. Single point calculations are employed on the optimized structures using a larger basis set which is explained in the basis sets section. For all of the computational work, the Gaussian 09 suit of programs[39] was used.

The principle of DFT was first explained by the Hohenberg–Kohn theorems[38,40,41] based on the calculation of the energy as a functional of total electron density of a system.

This is in contrast with ab initio methods in which the wavefunction is calculated. The functional that explains the relation between the electron density and energy of a system was developed by Kohn-Sham in 1965.[42] In this notation E[ρ] and electron density ρ[r] are defined as a function of electron density and position, respectively. The Kohn-Sham theorem solves the energy functional for a system of interacting electrons with a set of basis functions for non-interacting electrons as well as a correction term known as exchange-correlation energy which is a functional of electron density. There are different approximations employed to model the exchange-correlation functional of a system. The local density approximation (LDA) is a well-known example in which the electron density is obtained from a homogeneous electron gas. By starting from an initial guess for electron density, initial orbitals Ψ_i are calculated.[43]

$$\rho(r) = \sum_{i=1}^{n} |\Psi_i(r)|^2 \qquad 2.12$$

From Equation 2.12 the electron density is then obtained from the calculated orbitals and this continues in a self-consistent way until the electron density stays the same. The energy functional is calculated from the final electron density that is obtained in this fashion. One well known functional in DFT is the B3LYP[44,45] (Becke, 3 parameter, Lee-Yang-Parr) hybrid functional. B3LYP is popular in geometry optimization calculations for small biological molecules which is the case for this study. This functional is empirical since it mixes the Hartree-Fock exchange energy with DFT exchange correlation energy terms to provide more accuracy.[43] Despite the fact that B3LYP is not very accurate[46], it has shown to be an efficient functional in this work.

2.4.2. Basis sets

Basis sets consist of functions that are used for B3LYP calculations. A linear combination of basis functions defines the molecular orbitals. These functions are known as basis sets which can vary in terms of type and number of functions. The basis sets that are used in this thesis for B3LYP calculations are 3-21G, 6-31+G(d,p)(sometimes shown as 6-31+G**) for optimization calculations and the larger 6-311++G(3df,3pd) basis set for single point calculations. These basis sets are split valence or Pople basis sets which introduce more than one basis function for every valence atomic orbital. This notation is because the valence atomic orbital electrons are the main participants in a bond formation. Each term in these basis sets corresponds to a function.[38,47,48]

The typical form for a split valence basis set is as follows:

X-YyG 2.13

In which X presents the number of Gaussian functions combined to form a core atomic orbital. This is an example of a double zeta basis function since Y and y both show the number of Gaussian functions for the valence atomic orbitals and their splitting. For example in chapter 4, computations is performed using a small basis set, the 3-21G(d,p) basis set, which is a split valence double zeta basis set. In this basis set three Gaussian functions are combined to form the core orbitals and the valence orbitals are split into two basis functions in such a way that first one is a combination of two and the second one consists of one Gaussian function. The + and ++ that sometimes appear in the basis set designation, define the addition of diffuse functions. These functions account for interaction of electrons with longer distances from the atom nuclei. Anions and ions with

lone pairs are examples of such systems. Diffuse interactions terms can be added to the basis set for non-hydrogen atoms or heavy atoms with a single + corresponding to a set of s and p (px, py, pz) functions. Furthermore addition of a diffuse s function for light atoms (hydrogen and helium atoms) is indicated as ++ in the basis set. Polarization functions are also common additions to the minimal basis sets which are shown as one or two asterisks, * and **, or as the basis functions type and number and an antisymmetric term for modification of molecular orbitals shape around the nuclei. A single * or simply a (d) represents addition of a set of d basis functions to the heavy atoms while in case of ** or (d,p) an extra set of p basis functions are accounted for hydrogen atoms as well.

The large basis set that is used for single point calculations in this work, 6-311++G(3df,3pd), is a triple zeta split valence in which six Gaussian functions are combined for the core atomic orbitals. Valence orbitals are described as three basis functions, first basis function is a combination of three and second and third basis functions are single Gaussian functions. The polarization functions as well as diffuse functions are added for both light and heavy atoms.

As part of this work the Ahlrichs type basis set[49,50], SVP (Split Valence Polarized) was used for optimization and frequency calculations of larger structures. SVP are the smallest and simplest form of split valence basis sets. The inner shell atomic orbitals are defined as single basis functions and each valence shell is explained by two basis functions as well as polarization functions.[48,51] Larger sets including polarization functions can be presented with SVPP.

2.4.3. Toward finding the global minimum for the complex systems

In larger/complex systems with considerable number of atoms there are more geometry possibilities; therefore, computational methods are needed to generate variety of structures which can be used as starting point for geometry optimization. In this work simulated annealing (SA), an energy optimization method, was used which is based on molecular mechanics. Simulated annealing uses molecular dynamics to generate different conformations to find the global minima.

2.4.3.1. Molecular Mechanics (MM) and Molecular Dynamics (MD)

In molecular mechanics the motions of the particles (e.g. atoms/molecules) are solved by applying Newton's law of motion. This method employs classical and statistical mechanics to calculate the thermodynamic properties. In molecular mechanics a partial charge is assigned to each atom with its atomic mass. Interactions between these point charges are calculated using Coulomb's law while the dispersion interaction is accounted for using a van der waals potential. Bonded interactions are also included using bond, angle, and dihedral potential.[38,47] From the classical mechanics the force is a function of potential gradient.[52]

$$F_i = -\nabla_{r_i} U(r_1, \ldots, r_N) \qquad 2.14$$

By knowing the forces (F_i) acting on the interacting atom (i) their velocities and as a result positions (r_i) can be derived over time (t).

$$F_i = m_i \frac{d^2 r_i(t)}{dt^2} \qquad 2.15$$

In molecular dynamics Equation 2.15 is solved numerically.[47,53] A discrete time step is assigned to solve the equation and at each time step, or after a number of time steps, velocities and/or forces are stored and named as trajectory. The trajectories obtained by this method are determined and dependent upon the simulation temperature which affects the kinetic energy of the system. If this energy is enough to overcome energy barriers in the potential energy surface, the system moves from one local minimum to the other. However based on the atomic motions time scale (e.g. femtoseconds for bond vibration) very short time intervals are required to solve the equations of motion which means that as a result simulation time is limited. In this regard, the simulation becomes confined to conformational search around a local minimum in a potential energy surface. As a result, to find the global minimum many initial conformations are needed.

2.4.3.2. Simulated annealing

Simulated annealing (SA) is designed for global optimization of many body systems. This method was first derived from the annealing technique in metallurgy and crystal growing. Kirkpatrick first applied simulated annealing for optimization purposes in 1983.[54] The simulated annealing algorithm is defined to explore the conformations at different temperatures. In this technique, optimization is based on heating structures to a high enough temperature and cooling them gradually in order to search the conformational space to find the lowest energy structure.[55] In each heating and cooling cycle a molecular dynamics simulation is performed to explore the conformational landscape.

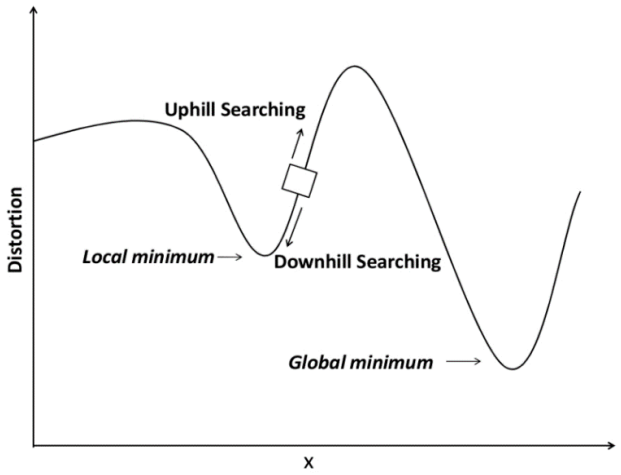


Figure 2.16. Surfing the potential energy surface with uphill and downhill motions. Reprinted with permission from *IEEE T. Consum. Electr.* **2011**, *57*, 1354-1361. Copyright (2011) Elsevier Ltd

Figure 2.16 demonstrates a potential energy surface with a global minimum which is the lowest energy between a numbers of local minima.[56] In simulated annealing, at each non zero temperature, the conformation that is lower in energy ($\Delta E < 0$ in Equation 2.16) is accepted and this downward conformational search process results in finding the local minimum in the potential energy surface. In order to find the optimal local minimum an upward search is needed to avoid trapping in a local minimum and its neighborhood. In this regard conformations with a higher energy ($\Delta E > 0$) are also accepted based on a probability which is an exponential function of temperature of the system (Equation 2.17). This sampling becomes more selective as the temperature decreases.[55]

$$\Delta E = E_2 - E_1 \qquad 2.16$$

$$P(\Delta E) = \exp(-\Delta E/kT) \qquad 2.17$$

A more detailed description of procedure is provided in Chapter 4 of this work.

2.4.4. Calculations

The electronic energy calculations of this work are done using density functional theory and B3LYP level of theory. The first steps were geometry optimization and frequency calculations of a series of predicted structures. These calculations provided the possible structures along with their IR spectra and thermochemistry data such as the relative enthalpies and Gibbs energies. The relative enthalpies *H* and Gibbs energies *G* are computationally calculated by applying the thermal corrections of enthalpy and Gibbs energy to the electronic energy, all obtained by B3LYP calculations.

$$H = E_{tot} + k_B T \qquad 2.18$$

$$G = H - TS_{tot} \qquad 2.19$$

In Equation 2.18 E_{tot} is the summation of vibrational, rotational, electronic and translational energies, Equation 2.20.

$$E_{tot} = E_v + E_{rot} + E_e + E_t \qquad 2.20$$

Similarly S_{tot} is the summation of vibrational, rotational, electronic and translational entropies, Equation 2.21.

$$S_{tot} = S_v + S_{rot} + S_e + S_t \qquad 2.21$$

The relative enthalpy ($\Delta_{rel}H$) and free energy ($\Delta_{rel}G$) of some optimized structures were calculated to estimate the low lying structures at this level of theory. The single point calculations were applied by using a larger basis set, as mentioned before, to obtain more accurate thermochemistry data. In this regard the electronic energies obtained from the

single point calculations were taken with the thermal corrections of enthalpy and free energy to obtain the single point enthalpies and Gibbs energies.

The frequencies calculated for the possible structures were scaled by common factors to match the anharmonic frequencies in the experiments.[57]

References

(1) 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.

(2) 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.

(3) Kebarle, P.; Verkcerk, U. H. Electrospray: From Ions in solution to Ions in the gas phase, what we know now. *Mass Spectrom. Rev.***2009**, *28*, 898-917.

(4) Gaskell, S. J. Electrospray: Principles and practice. *J. Mass Spectrom.* **1997**, *32*, 677-688.

(5) Loo, J. A. Studying noncovalent protein complexes by electrospray ionization mass spectrometry. *Mass Spectrom. Rev.* **1997**, *16*, 1-23.

(6) Ashcroft, A. E. Recent developments in electrospray ionisation mass spectrometry: Noncovalently bound protein complexes. *Nat. Prod. Rep.* **2005**, *22*, 452-464.

(7) Liu, Y.; Pan, Y. Application of FT-ICR MS for the study of protein complexes. *Appl. Spectrosc. Rev.* **2009**, *44*, 231-244.

(8) Wilm, M. Principles of electrospray ionization. *Mol. Cell. Proteomics* **2011**, *10*, 1-8.

(9) Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. Electrospray ionization-principles and practice. *Mass Spectrom. Rev.* **1990**, *9*, 37-70.

(10) Iribarne, J. V.; Thomson, B. A. On the evaporation of small ions from charged droplets. *J. Chem. Phys.* **1976**, *64*, 2287-2294.

(11) Dole, M.; Mack, L. L.; Hines, R. L.; Chemistry, D. O.; Mobley, R. C.; Ferguson, L. D.; Alice, M. B. Molecular beams of macroions. *J. Chem. Phys.* **1968**, *49*, 2240-2249.

(12) Cole, R. B. Some tenets pertaining to electrospray ionization mass spectrometry. *J. Mass Spectrom.* **2000**, *35*, 763-772.

(13) 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.

(14) Marshall, A. G. Fourier transform ion cyclotron resonance mass spectrometry. *Acc. Chem. Res.* **1985**, *18*, 316-322.

(15) Marshall, A. G.; Grosshans, P. B. Fourier transform ion cyclotron resonance mass spectrometry: The teenage years. *Anal. Chem.* **1991**, *63*, 215A-229A.

(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) Marshall, A. G.; Hendrickson, C. L. Fourier transform ion cyclotron resonance detection: Principles and experimental configurations. *Int. J. Mass Spectrom.* **2002**, *215*, 59-75.

(18) Isenor, N. R.; Merchant, V.; Hallsworth, R. S.; Richardson, M. C. CO_2 laser-induced dissociation of SiF_4 molecules into electronically excited fragments. *Can. J. Phys.* **1973**, *51*, 1281-1287.

(19) Isenor, N. R.; Richardson, M. C. Dissociation and breakdown of molecular gases by pulsed CO_2 laser radiation. *Appl. Phys. Lett.* **1971**, *18*, 224-226.

(20) Clark, P. O.; Smith, M. R. Pulsed operation of CO_2-N_2-He lasers. *Appl. Phys. Lett.* **1966**, *9*, 369-372.

(21) Frapard, C.; Roulot, M.; Ziegler, X. High peak power pulsed 10μ CO_2 laser. *Phys. Lett.* **1966**, *20*, 384-385.

(22) Bomse, D. S.; Woodin, R. L.; Beauchamp, J. L. 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.

(23) Woodin, R. L.; Bomse, D. S.; Beauchamp, J. L. Multiphoton dissociation of molecules with low power continuous wave infrared laser radiation. *J. Am. Chem. Soc.* **1978**, *100*, 3248-3250.

(24) Woodin, R. L.; Bomse, D. S.; Beauchamp, J. L. Multiphoton dissociation of molecules with low power cw infrared lasers: collisional enhancement of dissociation probabilities. *Chem. Phys. Lett.* **1979**, *63*, 630-636.

(25) Bensimon, M.; Rapin, J.; Gäumann, T. Comparison of infrared photodissociation in a Fourier transform mass spectrometer with metastable ion decay in a double-focusing mass spectrometer. *Int. J. Mass Spectrom. Ion Processes* **1986**, *72*, 125-135.

(26) Eyler, J. R. Infrared multiple photon dissociation spectroscopy of ions in penning traps. *Mass Spectrom. Rev.* **2009**, *28*, 448-467.

(27) Fridgen, T. D. Infrared consequence spectroscopy of gaseous protonated and metal ion cationized complexes. *Mass Spectrom. Rev.* **2009**, *28*, 586-607.

(28) Patel, C. K. N. Continuous-wave laser action on vibrational-rotational transitions of CO_2. *Phys. Rev.* **1964**, *136*, A1187-A1193.

(29) Patel, C. K. N.; Tien, P. K.; McFee, J. H. CW high-power CO_2-N_2-He laser. *Appl. Phys. Lett.* **1965**, *7*, 290-292.

(30) Svelto, O.: *Principles of lasers*; fifth ed., Springer, US: Boston, **2010**.

(31) 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.

(32) Zhou, J. X.; Hou, X.; Yang, K. X.; Tsai, S. J. J.; Michel, R. G. Lasers based on optical parametric devices: Wavelength tunability empowers laser-based techniques in the UV, visible, and near-IR. *Appl. Spectrosc.* **1998**, *52*, 176A-189A.

(33) Sleno, L.; Volmer, D. A. Ion activation methods for tandem mass spectrometry. *J. Mass Spectrom.* **2004**, *39*, 1091-1112.

(34) McLuckey, S. A.; Goeringer, D. E.; Glish, G. L. Collisional activation with random noise in ion trap mass spectrometry. *Anal. Chem.* **1992**, *64*, 1455-1460.

(35) Hayes, R. N.; Gross, M. L. Collision-induced dissociation. *Methods Enzymol.* **1990**, *193*, 237-263.

(36) McLuckey, S. A.; Goeringer, D. E. Slow heating methods in tandem mass spectrometry. *J. Mass Spectrom.* **1997**, *32*, 461-474.

(37) Peltz, C.; Drahos, L.; Vékey, K. SORI excitation: collisional and radiative processes. *J. Am. Soc. Mass Spectrom.* **2007**, *18*, 2119-2126.

(38) Lewars, E. G.: *Computational chemistry: Introduction to the theory and applications of molecular and quantum mechanics*, Springer, Netherlands: Dordrecht **2011**.

(39) Gaussian 09, R. D., M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox, Gaussian, Inc., Wallingford CT, 2009.

(40) Vignale, G.; Rasolt, M. Density-functional theory in strong magnetic fields. *Phys. Rev. Lett.* **1987**, *59*, 2360-2363.

(41) Hohenberg, P.; Kohn, W. Inhomogeneous electron gas. *Phys. Rev.* **1964**, *136*, B864-B871.

(42) Kohn, W.; Sham, L. J. Self-consistent equations including exchange and correlation effects. *Phys. Rev.* **1965**, *140*, A1133-A1138.

(43) Sousa, S. F.; Fernandes, P. A.; Ramos, M. J. General performance of density functionals. *J. Phys. Chem. A* **2007**, *111*, 10439-10452.

(44) Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **1993**, *98*, 5648-5652.

(45) 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.

(46) Zhao, Y.; Truhlar, D. G. Density functionals with broad applicability in chemistry. *Acc. Chem. Res.* **2008**, *41*, 157-167.

(47) Jensen, F.: Introduction to computational chemistry; 2nd eddition, Wiley, **2006**.

(48) Pople, J. A.; Schleyer, P. V.; Hehre, W. J.; Radom, L.: ab initio molecular orbital theory, **1986**.

(49) Weigend, F.; Ahlrichs, R. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297-3305.

(50) Weigend, F.; Furche, F.; Ahlrichs, R. Gaussian basis sets of quadruple zeta valence quality for atoms H-Kr. *J. Chem. Phys.* **2003**, *119*, 12753-12762.

(51) Schäfer, A.; Horn, H.; Ahlrichs, R. Fully optimized contracted Gaussian basis sets for atoms Li to Kr. *J. Chem. Phys.* **1992**, *97*, 2571-2577.

(52) Goldstein, H.; Classical Mechanics. Third edition, Addison Wesley. **2002**.

(53) Cai, W.; Li, J.; Yip, S.: *Molecular dynamics*. Comprehensive Nuclear Materials, **2012**, 1, 249-265.

(54) Kirkpatrick, S.; Gelatt Jr, C. D.; Vecchi, M. P. Optimization by simulated annealing. *Science* **1983**, *220*, 671-680.

(55) Wilson, S. R.; Cui, W.; Moskowitz, J. W.; Schmidt, K. E. Conformational Analysis of Flexible Molecules: Location of the Global Minimum Energy Conformation by the Simulated Annealing Method. *Tetrahedron Lett.* **1988**, *29*, 4373-4376.

(56) Shi, Z.; Fernando, W. A. C.; Kondoz, A. Adaptive direction search algorithms based on motion correlation for block motion estimation. *IEEE T. Consum. Electr.* **2011**, *57*, 1354-1361.

(57) Halls, M. D.; Velkovski, J.; Schlegel, H. B. Harmonic frequency scaling factors for Hartree-Fock, S-VWN, B-LYP, B3-LYP, B3-PW91 and MP_2 with the Sadlej pVTZ electric property basis set. *Theor. Chem. Acc.* **2001**, *105*, 413-421.

Chapter 3 . IRMPD Spectroscopic Study of Microsolvated $[Na(GlyAla)]^+$ and $[Ca(GlyAla-H)]^+$, and the Blue-Shifting of the Hydrogen Bonded Amide Stretch with Each Water Addition

This work has been published as M. B. Moghaddam and T. D. Fridgen., IRMPD Spectroscopic study of microsolvated $[Na(GlyAla)]^+$ and $[Ca(GlyAla-H)]^+$, and the blue-shifting of the hydrogen bonded amide stretch with each water addition, *J. Phys. Chem. B*, 2013, 117, 6157–6164. The experimental and computational work in this chapter was performed by me. The manuscript was co-written by me and my supervisor Dr. Travis D. Fridgen.

Abstract

In this study, the structures of $[Na(GlyAla)(H_2O)]^+$ and $[Ca(GlyAla\text{-}H)(H_2O)_n]^+$, (n = 1, 2, or 3) solvated ion-molecule complexes (as well as the AlaGly isomers) were investigated using infrared multiple photon dissociation (IRMPD) spectroscopy and with computational methods. Calculations showed that in the calcium clusters, the lowest energy complex is the one in which the peptide is deprotonated at the carboxylic acid end and that Ca^{2+} binds to both carboxylate oxygen atoms as well as the amide carbonyl oxygen. For the microsolvated structures, all three water molecules also bind directly to Ca^{2+}. For the singly, doubly, and triply solvated complexes these structures are supported by experimental infrared multiple-photon dissociation (IRMPD) spectra. For the $[Na(GlyAla)(H_2O)]^+$ complex both carbonyl oxygen atoms—one from the intact carboxylic acid and one from the amide group—as well as the water molecule were found to be bound to the Na^+. In all the spectra a strong band is observed between 3300 and 3400 cm^{-1} and is assigned to the amide N-H stretch which is red-shifted due to hydrogen bonding with the amine nitrogen. The position of the hydrogen bonded amide N-H stretch is experimentally and theoretically found to be sensitive to the number of water molecules; it is shown to blue shift upon successive hydrations.

Keywords: Calcium, Sodium, AlaGly, GlyAla, Infrared Multiple -Photon Dissociation (IRMPD) Spectroscopy, Mass Spectrometry, Peptides, Hydrogen Bonding, Solvation.

3.1. Introduction

Amino acids and peptides are the building blocks of the proteins which are among the most important and most complicated molecules in nature. In order to fully understand the roles and detailed activities of proteins, it is necessary to first study their smaller components. Amino acids and peptides as well as their complexes with metal ions in solution and in the gas phase have been widely investigated. From these studies we can begin to understand the intermolecular and intramolecular interactions that are of utmost importance in the larger biomolecules. Special interest has been given recently to metal-ion interactions with biomolecules such as amino acids and peptides in the gas phase.[1-5] Metal-ion complexes with these molecules are reported for a wide range of metal ions from alkali and earth alkaline metals, (Na^{+}[6-12], Li^{+}[8,10,13], K^{+}[7,10,12,14], Ca^{2+}[12,15,16]), transition metals (Zn^{2+}[17,18], Ag^{+}[19,20]), p-block metals such as Pb^{2+}[21-23], and even lanthanides such as La^{3+}, Ho^{3+}, and Eu^{3+}.[24] Sodium and calcium ions both play important roles in the structure and function of proteins and peptides in many biological systems. Sodium ions are also the center of neurotransmitter activity and are involved in the activation of many essential enzymes–in many biological molecules the presence of a metal interacting with the molecule stabilizes a specific conformation.[25] Calcium also takes part in protein activation and is involved in many other processes such as contributing to intercellular signaling and it acts as a messenger in many enzymatic processes in the body.[16]

Among structural studies of biological complexes with metals, recent interest in IRMPD spectroscopy has grown to be an integral part of many gas-phase studies and is found to be a powerful method in this area of research.[1,2,5,26] Other dissociation

techniques such as collision induced dissociation (CID), IRMPD activation, electron capture dissociation (ECD) and black-body infrared radiative dissociation (BIRD) techniques have also been used for structural studies of either protonated or metallated biomolecule complexes.

CID has shown great potential in peptide sequencing and also differentiation between isomeric peptides. Forbes et al. [27] were able to distinguish between protonated ArgGly and GlyArg isomeric dipeptides by CID in a QqTOF mass spectrometer. While $ArgGlyH^+$ and $GlyArgH^+$ produce similar fragmentation patterns in the quadrupole ion trap, distinctly different fragmentation patterns were observed when CID was performed in the QqTOF due to the higher excitation energies and shorter time to obtain a high internal energy, resulting in rapid fragmentation before isomerization to a common low energy structure.

In a systematic IRMPD spectroscopic investigation of gas phase protonated polyalanine structures, Ala_nH^+, Vaden et al. [28] showed that Ala_3H^+ was protonated at one of the backbone carbonyl oxygens and that the dominant structure was unfolded. The lowest energy, unfolded structure also predicted an amide N-H stretch where this amide was hydrogen bonded to the amine group and a characteristic band was observed about 3000 cm^{-1}, far red-shifted from a normal amide N-H stretch due to the hydrogen bond. The larger protonated peptides were found to adopt globular structures in which the peptide was protonated at the terminal amine and adopt cyclic or globular charge-solvated structures. In lanthanide-polyalanine complexes, $[La(Ala_n\text{-}H)]^{2+}$ it was determined that the metal cation was bound to all carbonyl oxygen atoms of the deprotonated peptide.[24]

Calculations showed that in all of the larger complexes terminal amine was also bound to the metal, but in $[La(Ala_2\text{-}H)]^{2+}$ the terminal amine was involved in hydrogen bonding with the amide NH on the side of the complex opposite the metal ion. It was difficult, however, to spectroscopically differentiate between the structure where the terminal amine was involved in binding to the metal or involved as a hydrogen bond acceptor in the 900 to 1800 cm^{-1} region. Similarly, in a IRMPD spectroscopy and computational study on sodiated di- and tri-glycine and -alanine[29] the authors were able to determine the number of carbonyl oxygens bound to Na^+ in the 1000 to 2000 cm^{-1} region, but were unable to determine whether the amine was also bound to Na^+ or was involved in hydrogen bonding with the amide NH. In the present work on the sodiated and calciated dipeptides it should be possible to tell if the amide is involved in hydrogen bonding with the amine group by examination of the spectroscopic region above 3000 cm^{-1}.

Gas-phase studies of biologically relevant complexes, such as metal amino acid and peptide ions are useful to help understand the structures of hydrophobic residues where water is excluded. The importance of water on the structures of amino acids and peptide complexes can also be studied in the gas phase and the study of such "microsolvated" complexes are able to bridge the gap between gas-phase and purely solution phase studies. For example, infrared-ultraviolet double resonance spectroscopy was used to determine that some addition of a single water molecule to certain conformers of gas-phase N-acetyl-phenylalanine (NAPA) causes a slight mechanical opening, increasing the hydrogen bond distance between the terminal NH_2 group and the acetyl oxygen which is observed in a blue-shift in the NH_2 symmetric stretch.[30] In another study, IRMPD

spectroscopy was used to determine that Li^+-bound dimers of either uracil (U) or thymine (T) are linear. However, addition of a single water molecule resulted in broad and intense absorption from about 2600-3200 cm^{-1} which was interpreted as N-H stretches which were involved in hydrogen bonding between the two DNA bases. The single addition of water slightly weakened and lengthened the Li^+-U(or T) bond allowing the bases to pair. The effect of water solvation on a number of gas phase metal cationized DNA base and amino acid complexes have been studied recently.[21-23,31-37] Kamariotis et al. [37] studied protonated and lithiated valine hydrated by one to four water molecules by IRMPD spectroscopy and computational chemistry. The study showed that valine was in the canonical form for all clusters and that the third water molecule completed the first solvation shell structures of the charge.

Many biological species exist in the isomeric form in the nature. Lemoff et al. [38] used BIRD kinetic data and master equation modeling to determine the threshold for loss of water from water-solvated α-methylated-proline (α-Me-Pro) complexes cationized with either Li^+ or Na^+. They compared these α-Me-Pro solvation energies with calculated binding energies to conclude their structures are zwitterionic in the gas phase. Blagojevic et al. [39] used differential mobility spectrometry (DMS) to distinguish between some protonated isomeric dipeptides such as AlaGly and GlyAla and SerGly and GlySer. The isomeric peptides were separated in the mobility spectrum when modifiers such as water, acetonitrile, or propanol were present in the DMS cell. One of the motivations of the present work was to determine if IRMPD spectroscopy could be used to distinguish between isomeric dipeptides.

Here we report on our investigation into the structures of the simple dipeptides, GlyAla and AlaGly, cationized with either Na^+ or Ca^{2+} and microsolvated by 1 – 3 water molecules. Electronic structure calculations are used to help interpret the IRMPD spectra recorded in the 3200 – 3900 cm^{-1} region to characterize the water-solvated complexes. Structures of many calcium bound proteins/peptides have been studied in solution and it is of great importance to characterize them in the gas phase.[16]

3.2. Methods

3.2.1. Experimental

$[Na(GlyAla)]^+$, $[Na(AlaGly)]^+$, $[Ca(GlyAla-H)]^+$, and $[Ca(AlaGly-H)]^+$ were electrosprayed from 2 mmol L^{-1} aqueous solutions of AlaGly or GlyAla to which a few drops of 1 mmol L^{-1} $CaCl_2$ or NaCl were added. The laboratory for the study of the energetics, reactions and structures of gaseous ions at Memorial University is home to an ApexQe Bruker FT-ICR mass spectrometer to which a tunable IR laser (OPO) has been mated. The coupling and details of the FT-ICR and OPO laser have been discussed previously.[21] Briefly, the laser power is at a maximum of 60 mJ at about 3800 cm^{-1} and decreases linearly to about 10 mJ at 3100 cm^{-1}. No corrections for laser power fluctuation are made to the IRMPD spectra presented. In order to obtain solvated complexes of calcium and sodium water vapour was leaked into the hexapole storage/collision cell as discussed previously.[40] Electrosprayed ions were stored in the hexapole storage cell for about 2 s before being transferred to the ICR cell where they were isolated by standard techniques. To obtain an IRMPD spectrum, ions were irradiated for between 0.5 and 2 s,

depending on the ease of dissociation of the particular ion, at approximately 1.5 cm^{-1} steps over the 3200 – 4000 cm^{-1} region.

3.2.2. Computational

For the calculations Gaussian 03 software was used.[41] Geometry optimizations, frequency calculations, and thermal corrections to the energies were done using B3LYP density functional theory and the 6-31+G(d,p) basis set. Calculations were done on many different tautomers and conformers of $[Na(GlyAla)]^+$, $[Na(AlaGly)]^+$, $[Ca(GlyAla-H)]^+$, and $[Ca(AlaGly-H)]^+$ and the solvated counterparts. Single-point calculations were done using the B3LYP/6-311+G(3df,3pd) method and basis set to obtain the electronic energies which were combined with the thermal corrections from the B3LYP/6-31+G(d,p) calculations to provide all reported thermochemistries. The wavenumber positions of the calculated spectra were scaled by 0.967.

The potential energy surface was explored with respect to different metal-peptide binding sites, conformers, and positions of the water molecules. The higher energy structures that are not shown in the main part of this report are supplied in the supplementary data and their computed IR spectra are compared to the experimental IRMPD spectra.

3.3. Result and discussion

Attempts at resonant IRMPD activation of the bare $[Na(GlyAla)]^+$, $[Na(AlaGly)]^+$, $[Ca(GlyAla-H)]^+$, and $[Ca(AlaGly-H)]^+$ ions resulted in no observed dissociation likely due to a dissociation threshold too high to be accessed using the OPO laser. All the solvated complexes, single water on Na^+ complexes and 1-3 water molecules on the Ca^{2+} complexes lost 1-3 water molecules upon resonant absorption of the OPO laser.

The IRMPD spectra in the 3200 – 4000 cm^{-1} region for $[Na(GlyAla)(H_2O)]^+$ and $[Na(AlaGly)(H_2O)]^+$ are presented in Figure 3.1(a) and those for $[Ca(GlyAla-H)(H_2O)_n]^+$ and $[Ca(AlaGly-H)(H_2O)_n]^+$ where n = 1, 2, and 3 are presented in Figure 3.1(b). Common to all eight spectra are strong, broad absorptions above 3600 cm^{-1} in the water O-H stretching region. Similarly, there is a strong absorption just below 3400 cm^{-1} in the sodiated complexes and below 3300 cm^{-1} for the calciated complexes. The bands in this lower energy region are lower in wavenumber position normally observed for N-H stretching vibrations for amides or amines and are more intense than normally observed.[21,22] They are also interesting, especially for the calciated complexes as there is a noticeable blue-shift in the bands on each addition of water, and this observation will be discussed further below. In the spectra of the Na^+ complexes, there is a strong band at about 3600 cm^{-1} (Figure 3.1) which suggests the presence of a carboxylic acid O-H stretch.[22,42] A carboxylic acid O-H stretch further suggests the structures of the sodiated complexes are canonical rather than zwitterionic. Furthermore, for the Ca^{2+} complexes, this strong band at 3600 cm^{-1} is not present, indicating that the peptide in these complexes is deprotonated at the carboxylic acid.

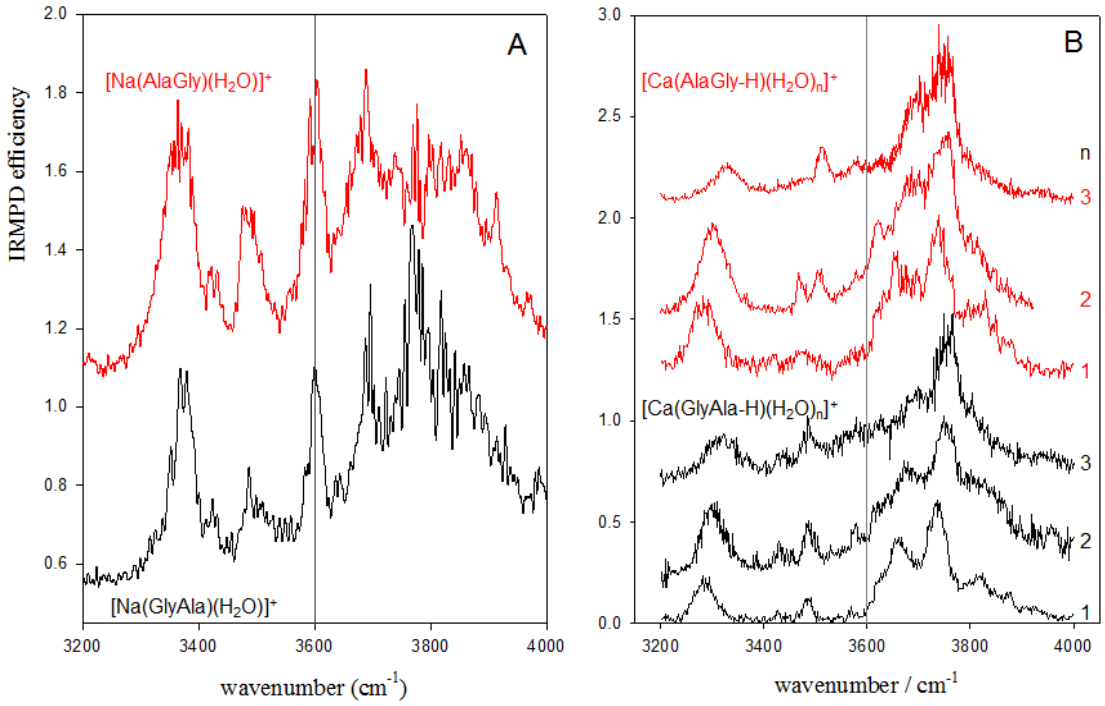


Figure 3.1. A comparison of the IRMPD spectra of a) $[Na(AlaGly)(H_2O)]^+$ and $[Na(GlyAla)(H_2O)]^+$ and b) $[Ca(AlaGly\text{-}H)(H_2O)_n]^+$ and $[Ca(GlyAla\text{-}H)(H_2O)_n]^+$; n=1, 2, or 3. The vertical grey line at 3600 cm^{-1} indicates the position of the COO-H stretch for $[Na(AlaGly)(H_2O)]^+$.

By comparing the GlyAla and AlaGly complexes in Figure 3.1, one quickly concludes that there is nothing in this spectral region that would distinguish the GlyAla complexes from the AlaGly. In the remainder of this paper we will focus on the microsolvated GlyAla complexes with both Na^+ and Ca^{2+}, however, similar structures were obtained with similar spectra for the AlaGly complexes.

3.3.1. $[Na(GlyAla)(H_2O)]^+$

The five lowest-energy structures for $[Na(GlyAla)(H_2O)]^+$ are shown in Figure 3.2 and in all water is bound to Na^+. The first four structures are canonical and the highest energy structure is zwitterionic, some 60 kJ mol^{-1} higher in Gibbs energy. Within about 1-2 kJ mol^{-1}, the relative Gibbs energies of the canonical structures do not change with the addition of a water molecule. The relative energy of the zwitterionic structure increases upon addition of a water molecule due to the solvation of Na^+ which decreases the ionic nature of the COO^-—Na^+ interaction. Like the bare structures shown in Figure A1 (Appendix A), respectively the lowest-energy structure (i-H_2O) has Na^+ bound to the two carbonyl oxygens and a hydrogen bond between the amide N-H bond and the amine group. A structure with similar Na^+ binding, iv-H_2O, where the amide-amine hydrogen bond is replaced by an amine N-H to carbonyl hydrogen bond, is about 22 kJ mol^{-1} higher in Gibbs energy. The second highest-energy structure, ii-H_2O, has Na^+ bound to the amide carbonyl and the amine group is about 16 kJ mol^{-1} higher in Gibbs energy. Structure iii-H_2O is similar to ii-H_2O except for a rotation of the carboxylic acid group, resulting in a structure 26 kJ mol^{-1} higher in Gibbs energy than the lowest energy structure.

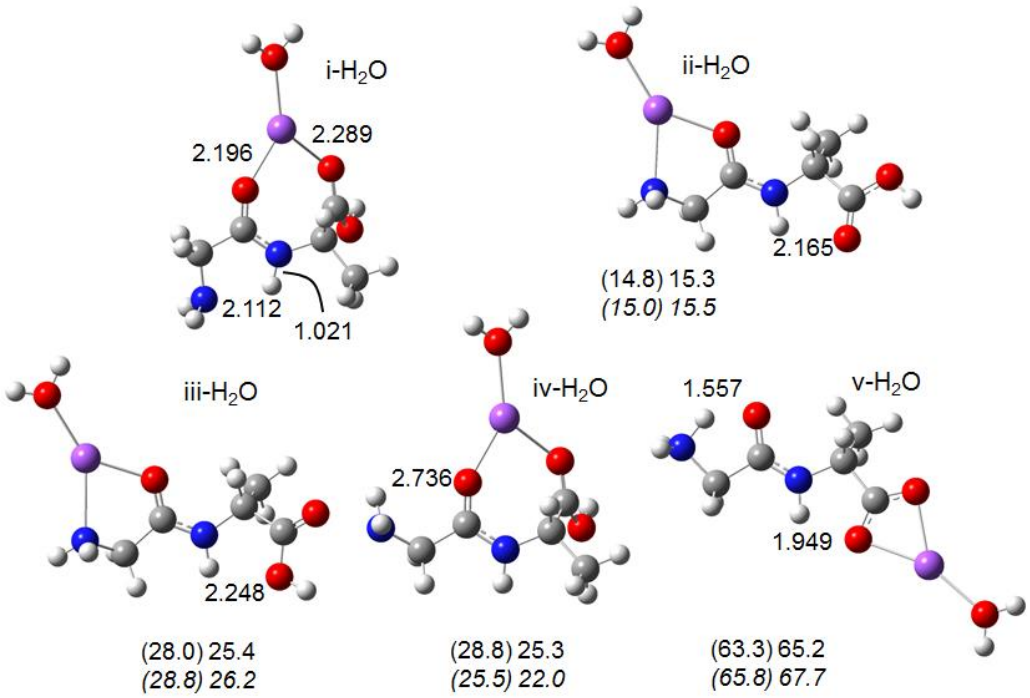


Figure 3.2. B3LYP/6-31+G(d,p) computed structures for $[Na(GlyAla)(H_2O)]^+$ along with relative enthalpies (in parentheses) and Gibbs energies computed at the B3LYP/6-31+G(d,p) (top numbers) or B3LYP/6-311++G(3df,3pd)//B3LYP/6-31+G(d,p) levels of theory (bottom, italicized numbers). Energies are in kJ mol^{-1} relative to the most thermodynamically stable structure, i-H_2O.

In Figure 3.3 the IRMPD spectrum for $[Na(GlyAla)(H_2O)]^+$ is compared to the computed IR spectra for those structures in Figure 2. The zwitterionic structure, v-H_2O, cannot solely account for the IRMPD spectrum since it lacks the carboxylic acid O-H stretch around 3600 cm^{-1}. Furthermore there is a strong feature predicted to occur below 3300 cm^{-1}, due to an N-H bond of the amine group hydrogen bonded to the neighboring carbonyl, which is not observed. Combined with its significantly higher relative energy it is reasonable to discount this structure as being responsible for the experimental spectrum. Structures ii-H_2O, iii-H_2O, and iv-H_2O do not reproduce the position of the strong band at 3370 cm^{-1}.

The lowest-energy structure, i-H_2O is most consistent with the IRMPD spectrum and does reproduce the main features of the IRMPD spectrum indicated in Figure 3.3. As mentioned previously, the water stretching region is very broad, likely due the fluxional nature of water bound to Na^+; there are no other interactions stopping the water molecule from rotating freely. The breadth of this region, and the feature just above 3800 cm^{-1}, is likely be due to combination bands with low-frequency modes associated with water wagging with respect to the rest of the molecule. The breadth of the features in the O-H stretching region, and the bands observed above 3800 cm^{-1} will be discussed further with respect to the Ca^{2+} complexes.

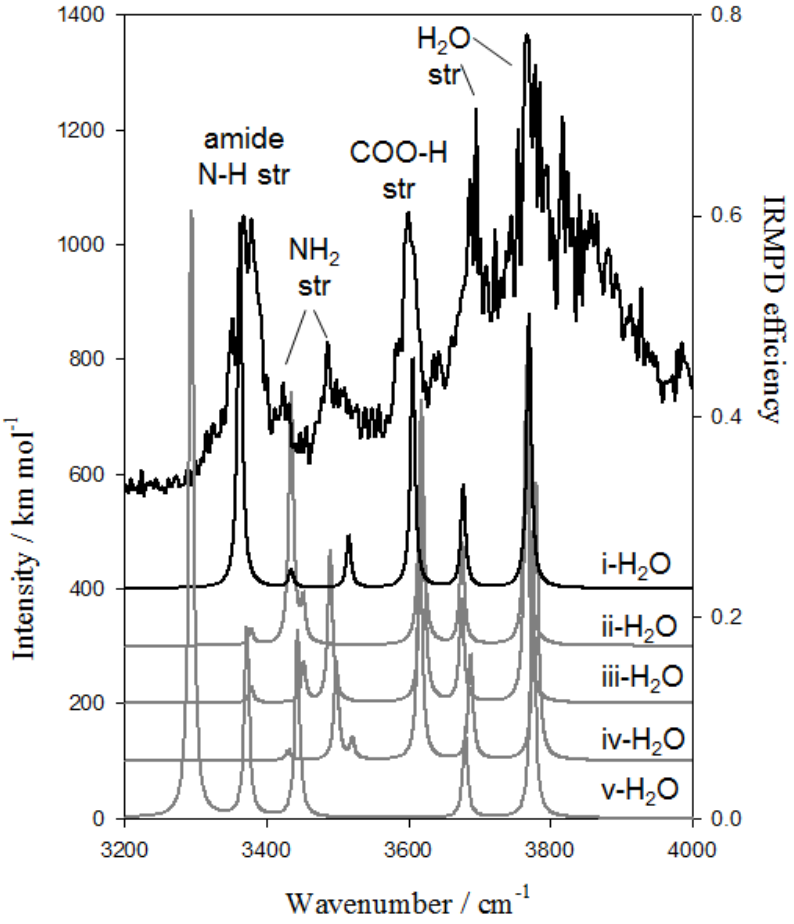


Figure 3.3. Comparison of the IRMPD spectrum of $[Na(GlyAla)(H_2O)]^+$ and the computed IR spectra for structures presented in Figure 3.2.

While we were not able to obtain the IRMPD spectrum for $[Na(GlyAla)]^+$, it is still interesting to look at the effect of adding a water molecule on the structure of this complex, even if only theoretically. Upon addition of a water molecule to the lowest energy structure the N-H---NH_2 hydrogen bond increases in length (Figures A1 and A2 in Appendix A) from 2.015 to 2.112 Å and a slight decrease in the N-H bond length from 1.023 to 1.021 Å. The effect of this is a predicted blue shift of the hydrogen bonded N-H stretching vibration by about 10 cm^{-1}. The blue shifting wit additions of water will be discussed more with respect to the Ca^{2+} complexes, where more experimental data is available.

3.3.2. $[Ca(GlyAla\text{-}H)(H_2O)_n]^+$

Displayed in Figure 3.4 are five computed structures for the bare $[Ca(GlyAla\text{-}H)]^+$ complex. The lowest-energy complex is a carboxylic acid deprotonated one where all three oxygen atoms are bound to calcium and, like the bare and singly solvated $[Na(GlyAla)]^+$ complexes has the amide N-H hydrogen bonded to the amine group. The difference in structure between A and B (27 kJ mol^{-1} higher in Gibbs energy) is that the amide-amine hydrogen bond is replaced by one between the amine and a carbonyl oxygen in B. In structure C the Ca^{2+} is bound to the amine N as well as all three oxygen atoms and is higher in Gibbs energy by 67 kJ mol^{-1}. Structures D and E are deprotonated at the amide N and amine N, respectively, and are considerably higher in energy. Since the IRMPD spectra of the microsolvated complexes do not have a COO-H stretch, the calculated spectra of the microsolvated and N-deprotonated complexes will not be discussed further.

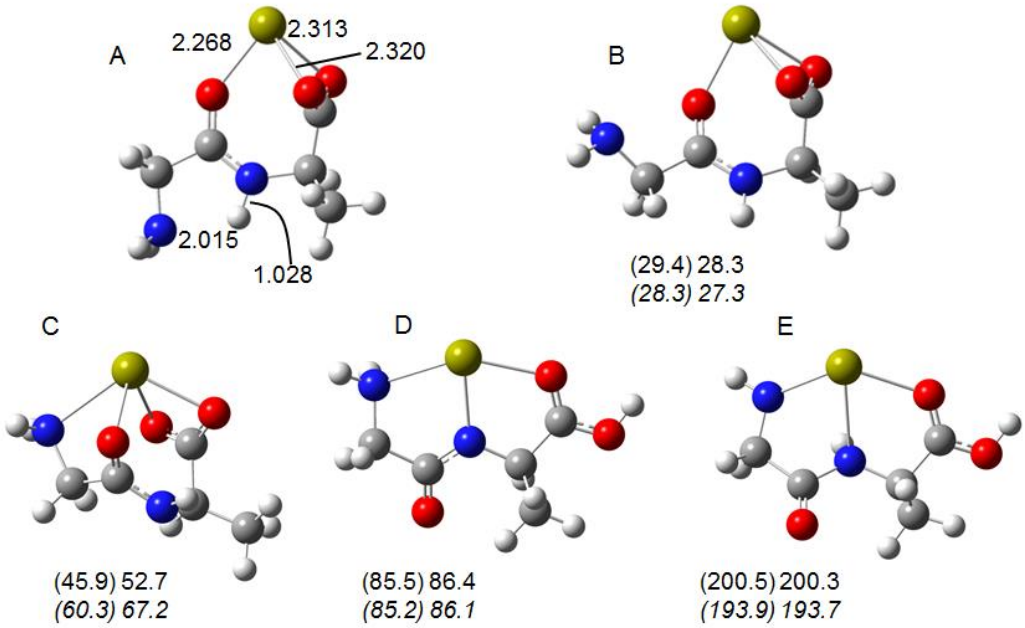


Figure 3.4. B3LYP/6-31+G(d,p) computed structures for $[Ca(GlyAla-H)]^+$ along with relative enthalpies (in parentheses) and Gibbs energies computed at the B3LYP/6-31+G(d,p) (top numbers) or B3LYP/6-311++G(3df,3pd)//B3LYP/6-31+G(d,p) levels of theory (bottom, italicized numbers).

The structures of the singly, doubly, and triply hydrated $[Ca(GlyAla-H)]^+$ complexes and their energies, relative to the lowest-energy structure, are displayed in Figure 3.5. Again, the lowest-energy structures all have the water molecules directly bound to calcium as would be expected. The lowest-energy structures are all A-type structures. The water molecule in A_1 is not hydrogen bound to the rest of the molecule, but simply sticks directly upward from the rest of the complex. All attempts to hydrogen bond it to the carbonyl oxygens were unsuccessful. However adding a second and third

water to form A_2 and A_3 results in all the water molecules being hydrogen bound to carbonyl oxygens, but for the A_2 complexes, a structure where one of the waters was not hydrogen bound was found. Attempts to have hydrogen bonding between the water molecules in the doubly and triply solvated A structures, were unsuccessful. Also, for the trihydrate, attempts at structures where a water was not hydrogen-bonded, but free like in A_1 and A_2-b, were unsuccessful. It is worth noting that for the C structures, single solvation does not thermodynamically stabilize the structure with respect to the A structures, but the doubly- and triply-solvated complexes are stabilized by about 10 kJ mol^{-1} due to a water inserting into the Ca-amide carbonyl bond. The B structures are significantly stabilized, with respect to the A structures, by addition of a water which is hydrogen bonded to the amine nitrogen.

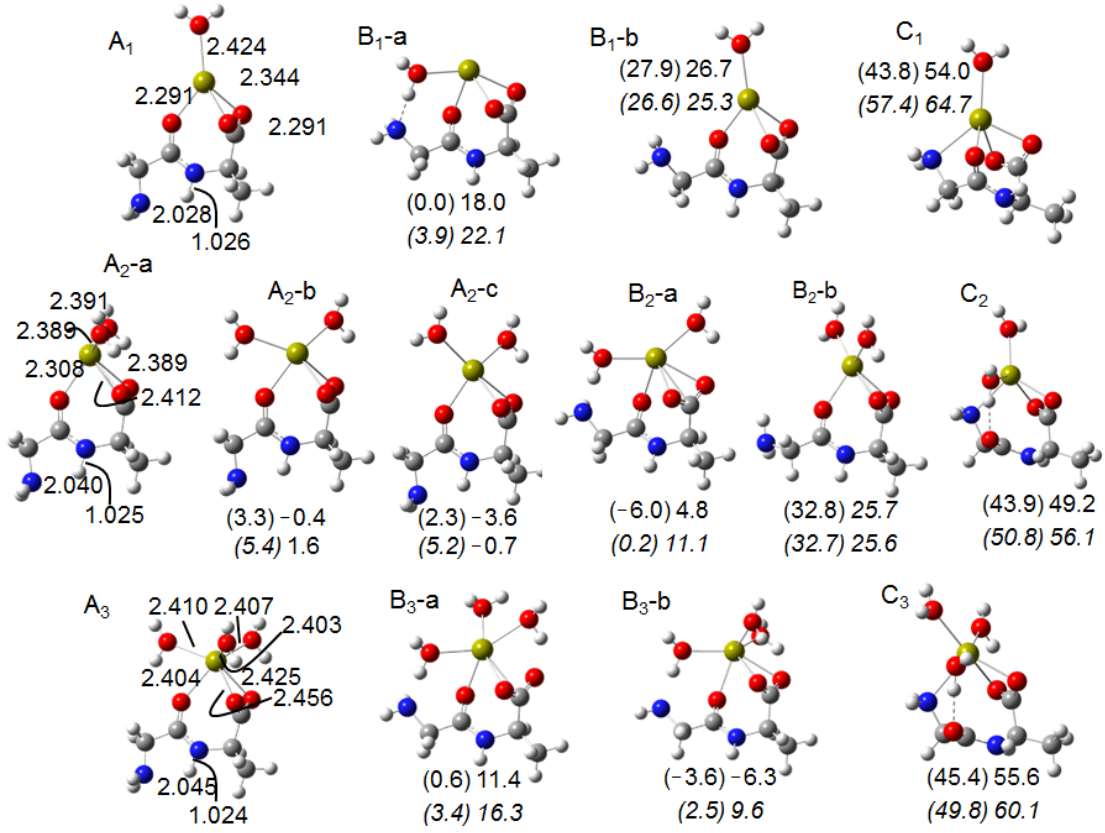


Figure 3.5. B3LYP/6-31+G(d,p) computed structures for $[Ca(GlyAla\text{-}H)(H_2O)_{1\text{-}3}]^+$ along with relative enthalpies (in parentheses) and Gibbs energies computed at the B3LYP/6-31+G(d,p) (top numbers) or B3LYP/6-311++G(3df,3pd)//B3LYP/6-31+G(d,p) levels of theory (bottom, italicized numbers). Energies are in kJ mol^{-1} relative to the most thermodynamically stable structure for each level of hydration.

In Figure 3.6 a), b), and c) the IRMPD spectra of the singly-, doubly-, and triply-solvated $[Ca(GlyAla-H)]^+$ complexes, respectively, are compared with IR spectra computed for the structures provided in Figure 3.5. For the singly-solvated complexes, the only structure with a strong band predicted in the 3300 cm^{-1} region is A_1. The position and expected large intensity of the observed band are well reproduced by theory and this band is assigned to the hydrogen bonded amide N-H stretch as it was for $[Na(GlyAla)(H_2O)]^+$. The predicted positions of the symmetric and antisymmetric N-H stretching absorptions (3437 and 3521 cm^{-1}) along with their lower intensities are in good agreement with the experimental bands at 3426 and 3489 cm^{-1}, respectively. Similarly the predicted positions and intensities of the two O-H stretching bands for A_1 at 3651 and 3723 cm^{-1} are also in excellent agreement with the experimentally observed bands at 3661 and 3736 cm^{-1}. It is fair to say that the best fit to the experimental spectrum is the IR spectrum predicted for structure A_1. The structured features in the IRMPD spectrum above 3800 cm^{-1} will be discussed below.

For $[Ca(GlyAla-H)(H_2O)_2]^+$ (Figure 3.6(b)), there are three A_2 structures which are virtually isoenergetic, differing only in the carbonyl oxygens to which the water molecules are hydrogen bound and one, A_2-c, has one water molecule that is not hydrogen bonded to the rest of the molecule. The predicted hydrogen-bonded amide stretch and amine N-H stretching bands (3298 cm^{-1}) are identical to one another for the three A_2 structures as are the predicted positions of the free O-H stretch of the water moieties 3763 cm^{-1}. The absorption at about 3680 cm^{-1} is only accounted for by the symmetric stretch of the non-hydrogen bonded water of A_2-c. However, the breadth and

other less pronounced features can only be accounted for by the hydrogen bonded O-H stretching vibrations which are predicted for all the A_2 structures. Based on the adequacy of all these A_2 structures to represent the experimental spectrum and the computed energetics it is very likely that all three of these structures are present.

For $[Ca(GlyAla\text{-}H)(H_2O)_3]^+$ (Figure 3.6(c)), there is only one A_3 structure and its predicted spectrum very well reproduces the hydrogen bonded amide N-H stretch, predicted to occur at 3309 cm^{-1}, the symmetric and asymmetric N-H stretching bands of the amine group (predicted at 3439 and 3523 cm^{-1} and observed at 3436 and 3482 cm^{-1}) as well as the intense band predicted for the three degenerate free O-H stretches of the water molecules (predicted and observed at 3760 and 3770 cm^{-1}, respectively). The very broad absorption to the red of the water free O-H stretching band is ascribed to the hydrogen bonded O-H stretches of water which manifest themselves as very broad features, as observed in the IRMPD spectra of the doubly- and triply-solvated complexes, and are not well reproduced by harmonic approximation calculations.[43,44] In Figure 6d) the extent of broadening on addition of water can be assessed somewhat by looking at the extent to which the baseline disappears. For example, in the n=1 cluster, the baseline, except for the O-H stretch region is flat right across the spectrum. For n=2, the baseline is fairly flat from about 3200 cm^{-1} to about 3450 cm^{-1}, but in the n=3 spectrum, the complex is seen to absorb almost in the entire spectral region.

Structures C_2 and C_3 for the doubly- and triply-solvated complexes, respectively, also have a band predicted in the 3300 cm^{-1} vicinity, although this is a hydrogen bonded O-H stretch for one of the water molecules which is greatly red-shifted due to the strength of

the hydrogen bonding interaction. Structure B_3-b also has a hydrogen bonded water O-H stretch in this 3300 cm^{-1} region. However, these structures are significantly higher in energy, cannot solely account for the IRMPD spectra and, therefore, are not likely contributors to the observed spectra.

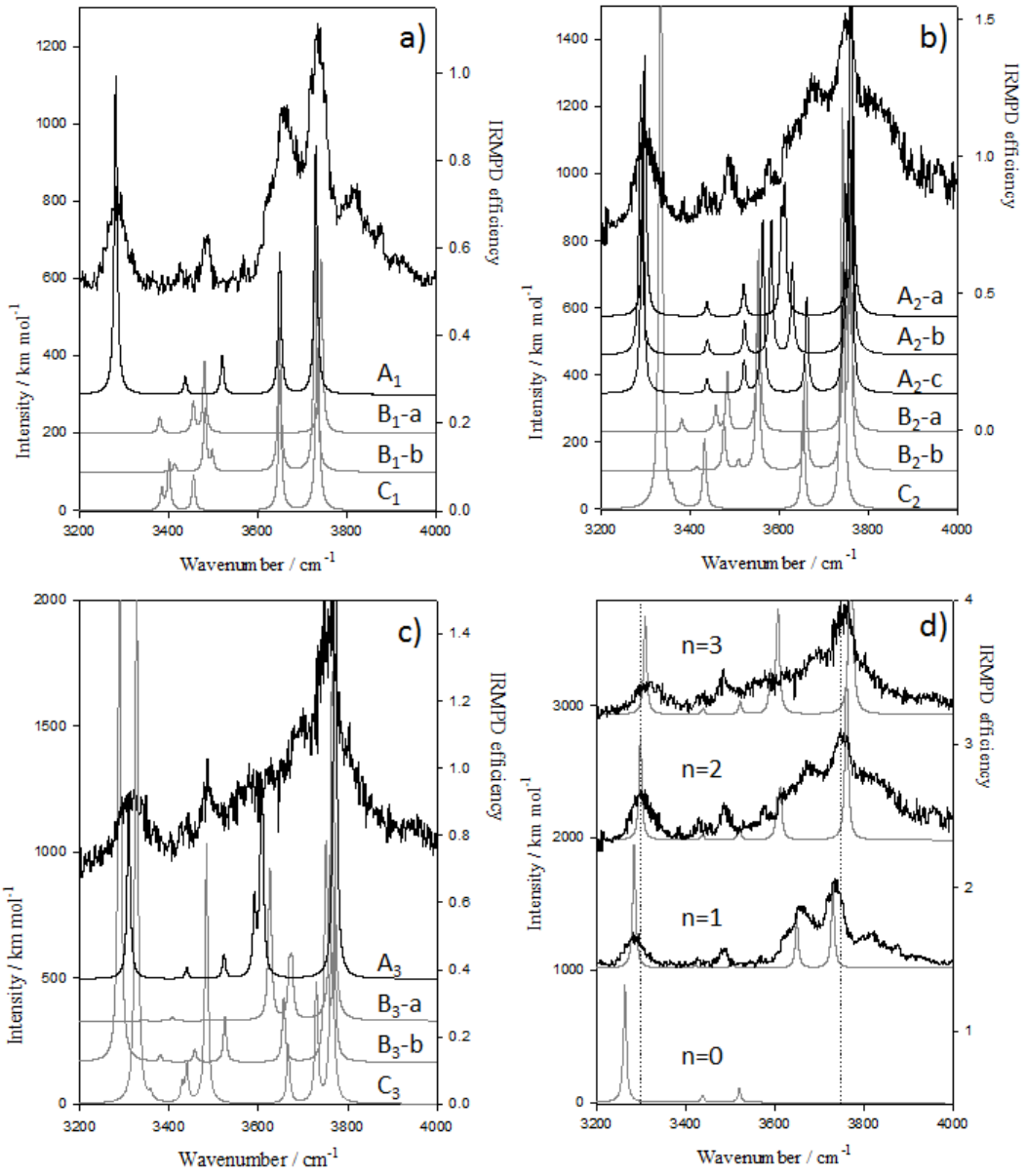


Figure 3.6. Comparison of the IRMPD spectra and computed IR spectra for a) $[Ca(GlyAla\text{-}H)(H_2O)]^+$, b) $[Ca(GlyAla\text{-}H)(H_2O)_2]^+$, and c) $[Ca(GlyAla\text{-}H)(H_2O)_3]^+$. The labels for the computed spectra refer to the structures in Figure 3.5. Also shown in panel d) is a comparison of the computed spectra for the lowest-energy structures for the non-, singly-, doubly, and triply-solvated structures as well as the IRMPD spectra for the latter 3.

In Figure 3.6 d) one can also see that both the predicted and observed hydrogen bonded amide stretch shifts to the blue slightly with greater solvation. The computed hydrogen bond length are 2.015, 2.028, 2.040, and 2.045 Å for the n=0, 1, 2, and 3 complexes, respectively, resulting in the N-H bond length decreasing slightly over the same series. The effect of this can be seen on the IRMPD spectra as the absorption maxima for the n=1, 2 and 3 amide N-H stretch are seen to shift from 3285 to 3300 to 3320 cm^{-1}, respectively; blue shifts which are almost perfectly predicted by the computed spectra. The cause of the blue shift is that upon each addition of water, the charge is less localized on the metal, so the hydrogen bond becomes less "ionic" in nature and weakens. It can be seen in Figures 3.4 and 3.5 that the carbonyl oxygen Ca-O bonds increase in length quite dramatically with each addition of water. It is also interesting to compare the structure and spectrum of $[Na(GlyAla)]^+$, where the N-H bond is shorter than in any of the Ca^{2+} complexes, and the hydrogen bond length is much larger at 2.112 Å (Figure 3.2). Furthermore, the natural bond order charge distributions (provided in Figure A2) show that the amide hydrogen becomes very slightly less positive, 0.474 for $[Ca(GlyAla-H)]^+$ to 0.471, 0.470 and 0.469 for the singly, doubly, and triply solvated complexes. They also show that on each addition of water, the positive charge on the metal decreases and the water molecule(s) bear some of the positive charge. The weaker hydrogen bond in the Na^+ complex results in a hydrogen bonded amide N-H stretch which is centered at 3380 cm^{-1}, much less red shifted from a free amide N-H stretch than observed for the Ca^{2+} complexes. The weaker hydrogen bond in the Na^+ complex can be explained by the decreased ionic nature of the hydrogen bond than in the Ca^{2+} complexes.

Finally, the absorptions in the IRMPD spectra to the blue of the O-H stretching features in the singly- and to a lesser extent, the doubly-solvated complex needs to be addressed. It was suggested above, when discussing the $[Na(GlyAla)(H_2O)]^+$ IRMPD spectra, that the absorptions to the blue of the O-H stretching region are combination bands, likely associated with the fluxional nature of water. It is unlikely that these can be anything but combination bands since nothing else can absorb in this high energy region, but their identity is speculation. However, for the Ca^{2+} complexes, it is seen with greater hydration, the absorptions above 3750 cm^{-1} are less pronounced and virtually disappear for the triply hydrated complex. This observation is consistent with combination bands being associated with water tortional or rotational modes because as the degree of hydration increases the water molecules are "locked" into place compared to the singly-hydrated species, and are less fluxional.

3.4. Conclusions

In this study, IRMPD spectroscopy in the gas phase was used in conjunction with electronic structure calculations to determine that $[Na(GlyAla)(H_2O)]^+$ and $[Ca(GlyAla-H)(H_2O)_n]^+$, n=1, 2, and 3, as well as the AlaGly isomers, are all complexes in which the metal ion is complexed to the available carbonyl oxygens. For the Na^+ complexes, there are two carbonyl groups since GlyAla is shown to be canonical based on the observation of a COO-H stretch around 3600 cm^{-1}. This COO-H stretch is absent in the Ca^{2+} complexes which shows that these complexes are deprotonated at the carboxylic acid group. In all complexes, the water molecules are bound directly to the metal cation. As

well, all complexes have an amide donor/amine acceptor hydrogen bond as is evidenced by the strong N-H stretching band observed around 3300 cm^{-1} for the Ca^{2+} complexes and just below 3400 for the Na^{+} complexes. Due to hydrogen bonding this amide N-H stretching absorptions are significantly red-shifted compared to a normal N-H stretch. These bands are also proof that the metal cation is not bound to the amine group. Furthermore, these amide N-H stretches are observed to blue shift upon each addition of water and this is due to the additional charge distribution among the extra water molecules. Structured absorptions observed above 3800 cm^{-1} for $[Na(GlyAla)(H_2O)]^{+}$ and $[Ca(GlyAla\text{-}H)(H_2O)]^{+}$ and to a lesser extent for $[Ca(GlyAla\text{-}H)(H_2O)_2]^{+}$ are explained as combination bands involving water tortional modes due to the fluxional nature of water. These bands are non-existent in $[Ca(GlyAla\text{-}H)(H_2O)_3]^{+}$ because the water molecules are locked into place due to sterics since Ca^{2+} is completely solvated.

References

(1) Fridgen, T. D. Infrared consequence spectroscopy of gaseous protonated and metal ion cationized complexes. *Mass Spectrom. Rev.* **2009**, *28*, 586-607.

(2) Polfer, N. C.; Oomens, J. Vibrational spectroscopy of bare and solvated ionic complexes of biological relevance. *Mass Spectrom. Rev.* **2009**, *28*, 468-494.

(3) Dunbar, R. C.; Polfer, N. C.; Oomens, J. Gas-phase zwitterion stabilization by a metal dication. *J. Am. Chem. Soc.* **2007**, *129*, 14562-14563.

(4) Carlton, D. D.; Schug, K. A. A review on the interrogation of peptide-metal interactions using electrospray ionization-mass spectrometry. *Anal. Chim. Acta.* **2011**, *686*, 19-39.

(5) Burt, M. B.; Fridgen, T. D. Structures and physical properties of gaseous Metal cationized biological ions. *Eur. J. Mass Spectrom.* **2012**, *18*, 235-250.

(6) Gapeev, A.; Dunbar, R. C. Cation−π Interactions and the Gas-Phase Thermochemistry of the Na^+/Phenylalanine Complex. *J. Am. Chem. Soc.* **2001**, *123*, 8360-8365.

(7) Ryzhov, V.; Dunbar, R. C.; Cerda, B.; Wesdemiotis, C. Cation-π effects in the complexation of Na^+ and K^+ with Phe, Tyr, and Trp in the gas phase. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 1037-1046.

(8) Marino, T.; Russo, N.; Toscano, M. Interaction of Li^+, Na^+, and K^+ with the proline amino acid. Complexation modes, potential energy profiles, and metal ion affinities. *J. Phys. Chem. B* **2003**, *107*, 2588-2594.

(9) Wyttenbach, T.; Witt, M.; Bowers, M. T. On the question of salt bridges of cationized amino acids in the gas phase: glycine and arginine. *Int. J. Mass Spectrom.* **1999**, *182*, 243-252.

(10) Forbes, M. W.; Bush, M. F.; Polfer, N. C.; Oomens, J.; Dunbar, R. C.; Williams, E. R.; Jockusch, R. A. Infrared spectroscopy of arginine cation complexes: Direct observation of gas-phase zwitterions. *J. Phys. Chem. A* **2007**, *111*, 11759-11770.

(11) Kapota, C.; Lemaire, J.; Maitre, 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.

(12) Dunbar, R. C.; Polfer, N. C.; Berden, G.; Oomens, J. Metal ion binding to peptides: Oxygen or nitrogen sites? *Int. J. Mass Spectrom.* **2012**, *330-332*, 71-77.

(13) Hoyau, S.; Ohanessian, G. Interaction of alkali metal cations (Li^+-Cs^+) with glycine in the gas phase: A theoretical study. *Chem.--Eur. J.* **1998**, *4*, 1561-1569.

(14) Polfer, N. C.; Paizs, B.; Snoek, L. C.; Compagnon, I.; Suhai, S.; Meijer, G.; von Helden, G.; Oomens, J. Infrared Fingerprint Spectroscopy and Theoretical Studies of Potassium Ion Tagged Amino Acids and Peptides in the Gas Phase. *J. Am. Chem. Soc.* **2005**, *127*, 8571-8579.

(15) Strittmatter, E. F.; Lemoff, A. S.; Williams, E. R. Structure of cationized glycine, Gly·M^{2+}(M = Be, Mg, Ca, Sr, Ba), in the gas phase: Intrinsic effect of cation size on zwitterion stability. *J J. Phys. Chem. A* **2000**, *104*, 9793-9796.

(16) Evenäs, J.; Malmendal, A.; Forsén, S. Calcium. *Curr. Opin. Chem. Biol.* **1998**, *2*, 293-302.

(17) Rogalewicz, F.; Ohanessian, G.; Gresh, N. Interaction of neutral and zwitterionic glycine with Zn^{2+} in gas phase: Ab initio and SIBFA molecular mechanics calculations. *J. Comput. Chem.* **2000**, *21*, 963-973.

(18) Rogalewicz, F.; Hoppilliard, Y.; Ohanessian, G. Structures and fragmentations of zinc(II) complexes of amino acids in the gas phase. III. Rearrangement versus desolvation in the electrospray formation of the glycine-zinc complex. *Int. J. Mass Spectrom.* **2001**, *206*, 45-52.

(19) Polfer, N. C.; Oomens, J.; Moore, D. T.; Von Helden, G.; Meijer, G.; Dunbar, R. C. Infrared spectroscopy of phenylalanine Ag(I) and Zn(II) complexes in the gas phase. *J. Am. Chem. Soc.* **2006**, *128*, 517-525.

(20) Schwabe, J. W. R.; Klug, A. Zinc mining for protein domains. *Nat. Struct. Biol.* **1994**, *1*, 345-349.

(21) 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.

(22) 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.

(23) Burt, M. B.; Fridgen, T. D. Gas-phase atructures of Pb^{2+}-cationized phenylalanine and glutamic acid determined by infrared multiple photon dissociation spectroscopy and computational chemistry. *J. Phys. Chem. A* **2013,** *117*, 1283-1290.

(24) 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* **2010**, *114*, 854-860.

(25) Lippard, S. J., Berg, J.M.: *Principles of Bioinorganic Chemistry*; University Science books: Mill Valley, 1994.

(26) Eyler, J. R. Infrared multiple photon dissociation spectroscopy of ions in penning traps. *Mass Spectrom. Rev.* **2009**, *28*, 448-467.

(27) Forbes, M. W.; Jockusch, R. A.; Young, A. B.; Harrison, A. G. Fragmentation of protonated dipeptides containing arginine. effect of activation method. *J. Am. Soc. Mass Spectrom.* **2007**, *18*, 1959-1966.

(28) Vaden, T. D.; De Boer, T. S. J. A.; Simons, J. P.; Snoek, L. C.; Suhai, S.; Paizs, B. Vibrational spectroscopy and conformational structure of protonated polyalanine peptides isolated in the gas phase. *J. Phys. Chem. A* **2008**, *112*, 4608-4616.

(29) Balaj, O. P.; Kapota, C.; Lemaire, J.; Ohanessian, G. Vibrational signatures of sodiated oligopeptides ($GG-Na^+$, $GGG-Na^+$, $AA-Na^+$ and $AAA-Na^+$) in the gas phase. *Int. J. Mass Spectrom.* **2008**, *269*, 196-209.

(30) Biswal, H. S.; Loquais, Y.; Tardivel, B.; Gloaguen, E.; Mons, M. Isolated monohydrates of a model peptide chain: Effect of a first water molecule on the secondary structure of a capped phenylalanine. *J. Am. Chem. Soc.* **2011**, *133*, 3931-3942.

(31) Gillis, E. A. L.; 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* **2009**, *113*, 824-832.

(32) Ali, O. Y.; Fridgen, T. D. Structures of electrosprayed Pb(Uracil-H)$^+$ complexes by infrared multiple photon dissociation spectroscopy. *Int. J. Mass Spectrom.* **2011**, *308*, 167-174.

(33) Burt, M. B.; Decker, S. G. A.; Fridgen, T. D. Water binding energies of [Pb(aminoacid-H)H_2O]$^+$ complexes determined by blackbody infrared radiative dissociation. *Phys. Chem. Chem. Phys.* **2012**, *14*, 15118-15126.

(34) Power, A. A.; Ali, O. Y.; Burt, M. B.; Fridgen, T. D. IRMPD spectroscopic and computational study of gas phase [M(Ura-H)(Ura)]$^+$ and [M(Ura-H)(H_2O)n]$^+$ (M=Sr, Ba; n=1, 2) complexes. *Int. J. Mass Spectrom.* **2012**, *330–332*, 233-240.

(35) Chang, T. M.; Prell, J. S.; Warrick, E. R.; Williams, E. R. Wheres the charge? Protonation sites in gaseous ions change with hydration. *J.Am. Chem. Soc.* **2012**, *134*, 15805-15813.

(36) Demireva, M.; O'Brien, J. T.; Williams, E. R. Water-induced folding of 1,7-diammoniumheptane. *J. Am. Chem. Soc.* **2012**, *134*, 11216-11224.

(37) Kamariotis, A.; Boyarkin, O. V.; Mercier, S. R.; Beck, R. D.; Bush, M. F.; Williams, E. R.; Rizzo, T. R. Infrared spectroscopy of hydrated amino acids in the gas phase: Protonated and lithiated valine. *J. Am. Chem. Soc.* **2006**, *128*, 905-916.

(38) Lemoff, A. S.; Bush, M. F.; Williams, E. R. Structures of cationized proline analogues: Evidence for the zwitterionic form. *J. Phys. Chem. A* **2005**, *109*, 1903-1910.

(39) Blagojevic, V.; Chramow, A.; Schneider, B. B.; Covey, T. R.; Bohme, D. K. Differential Mobility Spectrometry of Isomeric Protonated Dipeptides: Modifier and Field Effects on Ion Mobility and Stability. *Anal. Chem* **2011**, *83*, 3470-3476.

(40) 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.

(41) Frisch, M. J. e. a. Gaussian 03,revision C.02. **2004**.

(42) Atkins, C. G.; Rajabi, K.; Gillis, E. A. L.; Fridgen, T. D. Infrared multiple photon dissociation spectra of proton- and sodium ion-bound glycine dimers in the N-H and O-H stretching region. *J. Phys. Chem A* **2008**, *112*, 10220-10225.

(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) Ayotte, P.; Bailey, C. G.; Weddle, G. H.; Johnson, M. A. Vibrational Spectroscopy of Small $Br^-\cdot(H_2O)_n$ and $I^-\cdot(H_2O)_n$ Clusters: Infrared Characterization of the Ionic Hydrogen Bond. *J. Phys. Chem. A* **1998**, *102*, 3067-3071.

Chapter 4 . Structures and Energetics of $[M_n(GlyGly\text{-}H)_{2n-1}]^+$ and $[M_{n+1}(GlyGly\text{-}H)_{2n}]^{2+}$ Clusters (M=Mg, Ca and Sr) in the Gas Phase

This work has been published as M. B. Moghaddam, Yasaman Jami-Alahmadi and T. D. Fridgen, Self-assembled multimetallic/peptide complexes: structures and unimolecular reactions of $[Mn(GlyGly\text{-}H)_{2n-1}]^+$ and $[M_{n+1}(GlyGly\text{-}H)_{2n}]^{2+}$ clusters in the gas phase, *ChemPhysChem*, 2015, 16, 3290–3301. This work has been done and written by me and my supervisor Dr. Travis D. Fridgen. The simulated annealing calculation in this work was performed by Yasaman Jami-Alahmadi. The IRMPD spectral data in the fingerprint region were obtained by Travis D. Fridgen, Bryan D. Linford and Barry Power at the free electron laser facility at the Centre Laser Infrarouge d'Orsay (CLIO).

Abstract

Complexes of alkaline earth metal cations with deprotonated GlyGly are investigated by sustained off-resonance irradiation collision-induced dissociation (SORI-CID) and infrared multiple photon dissociation (IRMPD) spectroscopy, both conducted in a Fourier transform ion cyclotron resonance (FT-ICR) mass spectrometer. A comparison between dissociation patterns of the Mg^{2+}, Ca^{2+} and Sr^{2+} clusters under SORI-CID conditions is described here. The most intense ion observed in the electrospray mass spectrum of a mixture of either Mg^{2+}, Ca^{2+}, or Sr^{2+} with GlyGly was the $[M_2(GlyGly\text{-}H)_3]^+$ complex ion. Other singly charged ions observed were $[M(GlyGly\text{-}H)GlyGly]^+$, $[M_3(GlyGly\text{-}H)_5]^+$ and $[M_4(GlyGly\text{-}H)_7]^+$. Doubly charged complex ions were identified as $[M_3(GlyGly\text{-}H)_4]^{2+}$, $[M_5(GlyGly\text{-}H)_8]^{2+}$ and $[M_7(GlyGly\text{-}H)_{12}]^{2+}$. The main fragmentation pathway observed for almost all singly charged clusters was loss of 132 Da (GlyGly). Loss of 57 Da, NH_2CHCO, was a parallel dissociation pathway. Doubly charged Mg^{2+} clusters also showed similar dissociation patterns, losing GlyGly and NH_2CHCO. In contrast, the doubly charged clusters containing Ca^{2+} or Sr^{2+} form singly charged clusters through a charge separation pathway in competition with neutral loss of GlyGly and became more dominant for larger clusters. IRMPD spectroscopy and computational methods are used to obtain some idea of the structures of these complexes. Simulated annealing and DFT calculations were applied to search the potential energy surface for these complexes. Comparing the IRMPD and calculated IR spectra showed that the lowest energy structure of $[M_3(GlyGly\text{-}H)_4]^{2+}$ clusters for Mg^{2+}, Ca^{2+} and Sr^{2+} is consistent with deprotonation at the carboxyl group of all GlyGly molecules and the metal cations interacting with amide and carboxylate oxygen atoms. The N-H stretch

bands, observed at 3350 cm^{-1} is indicative of the terminal amine nitrogen atoms being hydrogen bonded to the amide hydrogen.

Keywords: Alkaline earth metal cation, GlyGly, Sustained off-resonance irradiation collision-induced dissociation (SORI/CID), Simulated annealing, Infrared multiple photon dissociation IRMPD

4.1. Introduction

Supramolecular assemblies or aggregation of molecules in an orderly manner under the control of non-covalent forces has been of great interest in many areas such as materials and biomaterials sciences.[1-3] Self-assembled complexes such as polynuclear metal complexes are supramolecular assemblies that occur spontaneously in nature. The presence of metal cations in biological systems and their strong binding to basic sites is a driving force for aggregation of biomolecules to highly ordered assemblies. In depth knowledge of their structure will advance our understanding of many biological processes.[4] The study of these structures and the interactions involved in their formation can help to provide a more complete picture of the chemistry occurring in nature. The knowledge obtained from these studies can be employed for synthesis of similar structures with properties of interest in many scientific areas.[5] Designing supramolecular assemblies means manipulating the chemical structure of these assemblies in a way that results in materials with specific specialized characteristics, such as novel magnetic[6-8] and electronic properties.[3]

Self-assembled complexes are also known for containing reactive sites since they can form cavities that can trap/bind target molecules of specific size or binding affinity.[9] Because of their selectivity to bind specific molecules, self-assembled complexes can act as catalysts or catalyst cavities for certain reactions.[10] Metal ions with different coordination capacities and ligands with functional groups of varying binding affinities for metal cations are important factors that can affect the architecture and stability, and as a result the reactivity, of supramolecular assemblies. For this reason studies on designing

assemblies have been focused on finding ligands[11] and metal cations[12] that provide structures with optimized characteristics.

The addition of metal cations to biomolecules has an effect on their structure and, therefore, function, and reactivity. For example, manganese ions of various oxidation states in photosystem II proteins act as redox centers. Alkaline earth metals are also known to be biochemically important since they participate in a wide range of biological processes.[13-17] EF-Hand motifs in calcium binding messenger proteins undergo conformational changes delivering calcium to sites where it regulates essential cellular functions.[18] Metal cations also induce protein folding—two well-known examples being the folding of azurin and cytochrome c upon binding to Cu and Fe, respectively.[19] Furthermore, the binding of metal ions to peptides may be responsible for forming amyloids which are known to be important contributors to many human diseases such as Parkinson's and Alzheimer's.[20] β-sheets are important components of amyloid like structures and their formation is dependent on hydrophobicity of the residues involved. In a study by Rufo *et al.* a combination of 7 hydrophilic amino acids such as LHLHLRL was found to be enzymatically active in the esterase (Enzymes) when it is bound to Zn^{2+}.[21]

Na^+ is also known to be an important metal in biological systems. In a study by Ohanessian and co-workers conformational changes in complexes of Gly_n-Na^+ with n=2-8 were investigated.[22] The presence of a non-hydrogen bonded carboxylic acid O-H bending mode around 1170 cm^{-1} for n=2, 3 and carboxylic acid O-H stretching mode around 3550-3560 cm^{-1} showed that the C-terminal group does not participate in hydrogen bonding. This was confirmed by computational methods. The sodium ion interacts with all carbonyl oxygen atoms in n=4, 5 and the terminal amine nitrogen binds

to the amide hydrogen of first residue through hydrogen bonding. In n=6, all 6 carbonyl oxygen atoms are bound to metal and the peptide length allows for the terminal amine and carboxylic groups to form hydrogen bonds. The presence of one single amide N-H stretch in the IR spectrum of these complexes shows that the terminal amine nitrogen is not hydrogen bonded to any amide hydrogen and all N-H bonds are identical. In bigger clusters such as n=7 and higher, because of the chance for diverse intramolecular hydrogen bonding, more complex structures become dominant. Also in large complexes such as n=8, folding of the structure allows proton transfer between terminal functional groups and as a result, zwitterionic complexes are observed.

Metal coordination by a biomolecule can affect the hydrogen bonding strength inside the structure and as a result, the stability of the structure. Oomens and co-workers investigated the structures and energetics of pentaalanine complexes with alkali (Na^+, K^+ and Cs^+) and alkaline earth (Ca^{2+} and Ba^{2+}) metal cations. They showed that in the IRMPD spectra of alkaline earth metal cation complexes there is a diagnostic peak around 1750 cm^{-1} which they related to the non-hydrogen bonded metal-bound carboxyl carbonyl stretching.[23] This was in accordance with their calculated lowest energy structure which showed that carboxyl carbonyl oxygen has participated in metal cations coordination. In contrast, for the larger monovalent metal cations, K^+ and Cs^+, the feature at 1750 cm^{-1} was weak and the ground state structure showed hydrogen bonding between the carboxylic acid and amine groups. However, the Na^+ complex showed a stronger peak at 1750 cm^{-1} compared to other alkali metal cation complexes but at the same time the lowest energy structure for Na^+ was in agreement with a non-hydrogen bonded carboxylic acid carbonyl stretch. Authors concluded that divalent metal cations form stronger

interactions compared to monovalent metal cations because of the higher positive charge density on the metal cation. In addition, as the smallest member of alkali metal cations studied, Na^+ has a high charge density and as a result shows intermediate properties. These results shows that the metal cation coordination is in competition with formation of intramolecular hydrogen bonding and the stronger the interaction of divalent metal cations with the carbonyl, the lower is the population of intramolecular hydrogen bonding in the structure of complex.

Collision induced dissociation is a widely used method for structure elucidation of complexes of metal cations with biomolecules such as amino acids[24-30], peptides[31-33] and nucleic acid bases.[34,35] The effect of metal cations on the dissociation pattern and activation energy of amino acid complexes were studied by CID method.[36,37] For example it was investigated that in the presence of metal cations, fragmentation channels are less energy demanding compared to the same dissociation in bare amino acid anions. In one study it was shown that dissociation of bare deprotonated glycine resulted in C-C bond cleavage at 6.9 eV in the centre of mass (com) frame whereas in the presence of Pb^{2+} the onset energy was much lower; around 0.49 eV. More interestingly, these studies provide useful information on the metal coordination site by investigating the fragmentation pathway in bare and metallated complexes.[37]

In the present study we explore the structures and unimolecular chemistry of singly and doubly charged multimetallic Ca^{2+} complexes with deprotonated GlyGly. These complexes have the following molecular formulae: $[Ca_n(GlyGly\text{-}H)_{2n-1}]^+$ and $[Ca_{n+1}(GlyGly\text{-}H)_{2n}]^{2+}$. Along with CID and IRMPD spectroscopy, simulated annealing

and electronic structure calculations were done on $[Ca_3(GlyGly\text{-}H)_4]^{2+}$ and $[Ca_5(GlyGly\text{-}H)_8]^{2+}$ clusters in an attempt to explore the structures of these complexes.

4.2. Methods

4.2.1. Experimental

All experiments were done on a Bruker ApexQe hybrid quadrupole-Fourier transform ion cyclotron resonance (FT-ICR) mass spectrometer. All chemicals were used without any purification. Solutions of $CaCl_2$ and GlyGly prepared with 18 MΩ (Millipore) water. Metal cationized complexes of GlyGly were electrosprayed from solutions of 1 mL of 5 mM GlyGly and a few drops of the 2 mM metal salt solution using an Apollo II electrospray ion source with flow rate 250 µl/h. For SORI-CID experiments Ar was pulsed to ~10^{-6} mbar inside the collision cell and SORI collision energies in the range of 0 - 1.4 eV in the centre of mass (com) frame for 250 ms. Fragmentation pathways were studied for the first and second generation ions by multi stage SORI-CID experiments.

IRMPD experiments were performed using two setups in two different regions of the IR; the fingerprint region, and the C-H/N-H/O-H stretching region. 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 Fourier-transform ion cyclotron resonance mass spectrometer (FT-ICR) coupled to an IR OPO, manufactured by LaserSpec. The OPO is 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 for the present experiments. In the fingerprint region, all experiments were performed using a FT-ICR-MS coupled to a mid-infrared free electron laser (FEL) (5 cm^{-1} bandwidth) at the Centre Laser Infrarouge d'Orsay (CLIO).

4.2.2. Computational

$[Ca_3(GlyGly\text{-}H)_4]^{2+}$ and $[Ca_5(GlyGly\text{-}H)_8]^{2+}$ cluster structures were initially optimized with B3LYP density functional theory and the 3-21G basis set with the Gaussian 09 suite of programs.[38] The initial geometries for the complexes were based on the known $Ca(AlaGly\text{-}H)^+$ structures determined previously.[39]

Simulated annealing[40] was performed on the optimized structures to further explore possible conformations of these clusters using generalized AMBER force field[41] (GAFF) and the AMBER9 software package.[42] For $[Ca_3(GlyGly\text{-}H)_4]^{2+}$ and $[Ca_5(GlyGly\text{-}H)_8]^{2+}$ clusters, calcium metal ions were positioned between the ligands in the cluster and an integer charge was assigned for each one. Minimized energy conformations were briefly equilibrated for 10 picoseconds (time step 0.5 fs) at 300 K before experiencing 2000 cycles of simulated annealing. Structures were drastically heated over 3 ps from 300 K to 750 K and equilibrated for 1 picosecond. This process was followed by cooling increments of 50K over 2 ps with 1 ps of equilibration at each temperature until 300 K. The lowest energy conformation of each simulated annealing cycle was then selected to pass to the subsequent round of simulated annealing. After each cycle, the low energy conformation was cooled to 0 K over 5 ps, minimized and the potential energy computed.

About 60 structures out of up to 2000 found by simulated annealing for $[Ca_3(GlyGly\text{-}H)_4]^{2+}$ were taken for further optimization with DFT. 50 unique structures out of 60 structures were identified. Optimization and frequency calculations were done using B3LYP/6-31+G(d,p) level of theory and basis set on these 50 unique structures to find the lowest energy structure. Initial structures of $[Mg_3(GlyGly\text{-}H)_4]^{2+}$ and $[Sr_3(GlyGly\text{-}H)_4]^{2+}$ were made using the optimized structure of the $[Ca_3(GlyGly\text{-}H)_4]^{2+}$ cluster. Five lowest energy structures of Ca^{2+} cluster were picked as starting geometries for Mg^{2+} and Sr^{2+}. Mg^{2+} clusters were optimized with 6-31+G(d,p). In the case of the Sr^{2+} complexes, the 6-31+G(d,p) was used on C, H, N and O and the Def2SVP basis set and effective core potential was used on Sr^{2+}. Also, for comparison optimization and frequency calculations on all complexes were carried out using the Def2SVP basis set. Single point calculations were done for all optimized structures using the B3LYP level of theory with 6-311++G(3df,3pd) basis set for Mg^{2+} and Ca^{2+}.

Furthermore $[M_5(GlyGly\text{-}H)_8]^{2+}$ cluster structure was investigated for Ca^{2+}. Optimizations with B3LYP/6-31+G(d,p) were done on 35 structures out of 2000 structures from the simulated annealing calculations. For frequency calculations the SVP basis set was used on all atoms. Around 30 unique structures were found from the optimized structures. All thermochemical data (enthalpy and Gibbs energy) are reported at 298 K and in kJ mol^{-1}. Wavenumber positions for calculated IR spectra were scaled by 0.966. All the higher energy structures and their enthalpy and Gibbs energies are supplied in the supporting information part of this paper.

4.3. Result and discussion

The mass spectra obtained upon electrospraying solutions of metal cation salts, $CaCl_2$, $MgCl_2$, $SrCl_2$ mixed with GlyGly is shown in Figure 4.1. Different clusters could be identified based on their *m/z* values and isotopic patterns. In these spectra the most intense peaks were observed at *m/z* 441.2, 473.1 and 569.0 for Mg^{2+}, Ca^{2+} and Sr^{2+}, respectively, and correspond to the singly charged $[M_2(GlyGly\text{-}H)_3]^+$ cluster but it is possible to tune the mass spectrometer to allow higher intensities at lower masses or at higher masses. In all cases the monomeric species, $[M(GlyGly\text{-}H)]^+$, could not be detected, presumably because of the high concentrations of GlyGly and metal salt solutions. The smallest complexes observed for the metal cations were $[M(GlyGly\text{-}H)GlyGly]^+$ (M = Mg^{2+}, Ca^{2+} and Sr^{2+}), at *m/z* 287.1, 303.1 and 351.0, respectively. $[M_3(GlyGly\text{-}H)_5]^+$ and $[M_4(GlyGly\text{-}H)_7]^+$ also were observed at *m/z* 727.3 and 1013.4 for Mg^{2+}, 775.2 and 1077.2 for Ca^{2+}, 919.1 and 1269.1 for Sr^{2+}. The smallest doubly charged clusters, $[M_3(GlyGly\text{-}H)_4]^{2+}$, were detected at *m/z* 298.1, 322.0, 394.0 for Mg^{2+}, Ca^{2+} and Sr^{2+}. Furthermore two other doubly charged clusters were observed at *m/z* 585.2, 624.1 and 744.0 corresponding to $[M_5(GlyGly\text{-}H)_8]^{2+}$ and *m/z* 871.2, 926.3 and 1094.1 corresponding to $[M_7(GlyGly\text{-}H)_{12}]^{2+}$, for Mg^{2+}, Ca^{2+} and Sr^{2+}, respectively. Other peaks in the mass spectra are either fragments (discussed later) or harmonics from higher mass peaks.

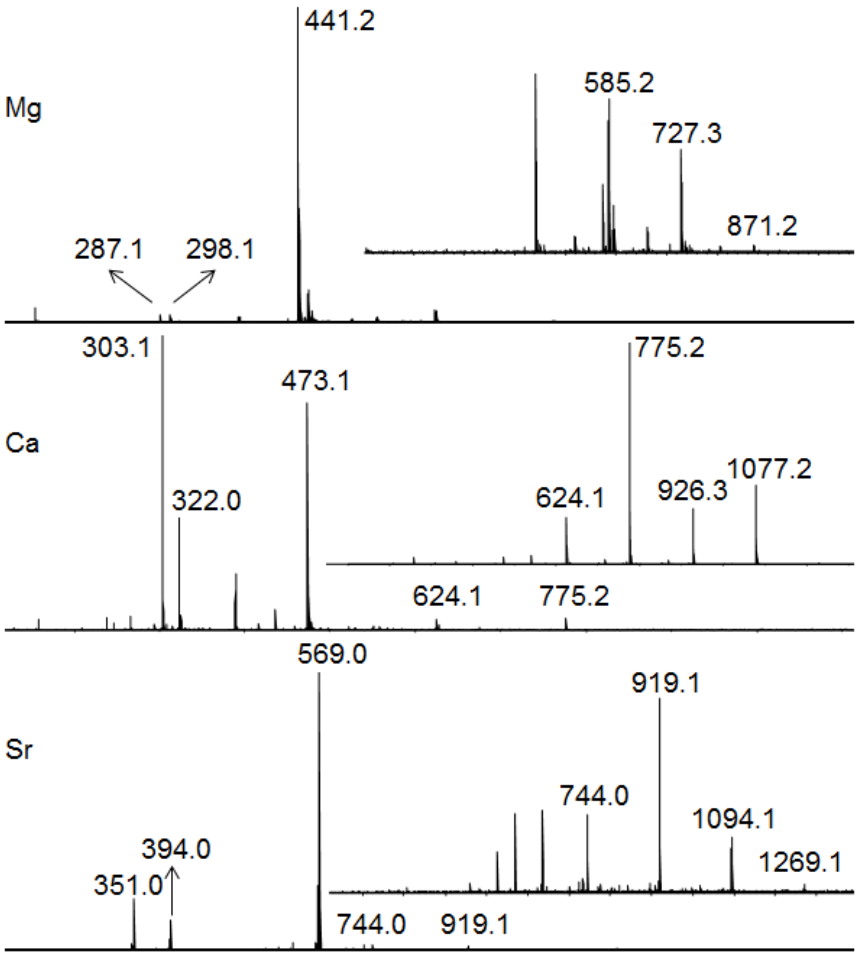


Figure 4.1. Full mass spectra of positive-ion electrosprayed aqueous solutions of $MgCl_2$, $CaCl_2$ and $SrCl_2$ with GlyGly are shown. $[M(GlyGly\text{-}H)GlyGly]^+$, $[M_3(GlyGly\text{-}H)_4]^{2+}$ $[M_2(GlyGly\text{-}H)_3]^+$, $[M_5(GlyGly\text{-}H)_8]^{2+}$, $[M_3(GlyGly\text{-}H)_5]^+$, $[M_7(GlyGly\text{-}H)_{12}]^{2+}$ and $[M_4(GlyGly\text{-}H)_7]^+$ clusters were observed at *m/z* 287.1, 298.1, 441.2, 585.2, 727.3 and 871.2 (largest singly charged cluster at *m/z* 1013.4 was not observed in this case) for Mg^{2+} and at *m/z* 303.1, 322.0, 473.1, 624.1, 775.2, 926.3 and 1077.2 for Ca^{2+} and at *m/z* 351.0, 394.0, 569.0, 744.0, 919.1, 1094.1 and 1269.1, for Sr^{2+}, respectively.

4.3.1 SORI/CID of $[M_n(GlyGly\text{-}H)_{(2n-1)}]^+$ and $[M_{(n+1)}(GlyGly\text{-}H)_{2n}]^{2+}$

4.3.1.1. Singly charged clusters

4.3.1.1.1. $[M(GlyGly)(GlyGly\text{-}H)]^+$

Of the three $[M(GlyGly)(GlyGly\text{-}H)]^+$ complexes the calcinated cluster provides the richest CID spectrum (Figure 4.2). SORI/CID of $[Ca(GlyGly)(GlyGly\text{-}H)]^+$ (*m/z* 303) resulted predominantly in fragment ions at *m/z* 215, 186, and 171, with lower intensity peaks observed at *m/z* 285, 259, and 246. The weak product ions at *m/z* 259 and 285 are clearly from the losses of CO_2 and H_2O, respectively. The *m/z* 246 peak is a loss of 57 Da which could be due to cleavage of the peptide bond and transfer of hydrogen to the remaining glycine. We will contemplate on the structure of the 57 Da loss below. MS/MS of *m/z* 246 resulted solely in the loss of 75 Da (glycine), presumably the rest of the peptide following the loss of 57 Da, at *m/z* 171 (Figure S4.1), This suggests that some of the *m/z* 171 observed in the CID of $[Ca(GlyGly)(GlyGly\text{-}H)]^+$ (*m/z* 303) may be secondary fragmentation product, but some direct loss of GlyGly, 132 Da cannot be ruled out. The *m/z* 215 fragment ion is interesting and corresponds to a loss of 88 Da. The simple explanation could be loss of two CO_2 molecules; however, the very low intensity of the *m/z* 259 ion would mean the primary product ion, following the initial loss of CO_2, is extremely unstable. Another explanation could be the direct loss of 88 Da which could mean the loss of a GlyGly ligand minus CO_2, which would leave CO_2 bound to Ca^{2+}. However, this would mean that CO_2 is very strongly bound to Ca^{2+} since MS/MS on *m/z* 215 does not result in loss of 44 Da, but loss of 29 Da, most likely CH_2NH, to form *m/z*

186 (Figure S4.2). B3LYP/Def2SVP calculations show that the enthalpy of binding of a $[Ca\text{-}O{=}C{=}O]^{2+}$ complex is 267 kJ mol^{-1} which may be a higher energy requirement than a concerted loss of CH_2NH. MS/MS of *m/z* 186 resulted in fragment ions at *m/z* 127 and 155 and 157, loss of 57 Da, 31 Da (CH_3NH_2) and 29 Da (CH_2NH), respectively (Figure S4.3).

The main fragmentation observed for $[Sr(GlyGly)(GlyGly\text{-}H)]^+$ is loss of GlyGly (132 Da) to produce *m/z* 219. Minor secondary fragmentations are observed at *m/z* 175 and 146. MS/MS shows that these losses of CO_2 and CO_2+CH_2NH stem from *m/z* 219 (Figure S4.4).

The main fragmentation product observed upon CID of $[Mg(GlyGly)(GlyGly\text{-}H)]^+$, *m/z* 287, is loss of 88 Da to form *m/z* 199. This product ion could be from the loss of two CO_2 molecules, however, there is no product ion corresponding to loss of one CO_2. As we speculate for the Ca^{2+} cluster, this dissociation could be the loss of GlyGly-CO_2. MS/MS of *m/z* 199 results in loss of 29 Da forming *m/z* 170 (Figure S4.5).

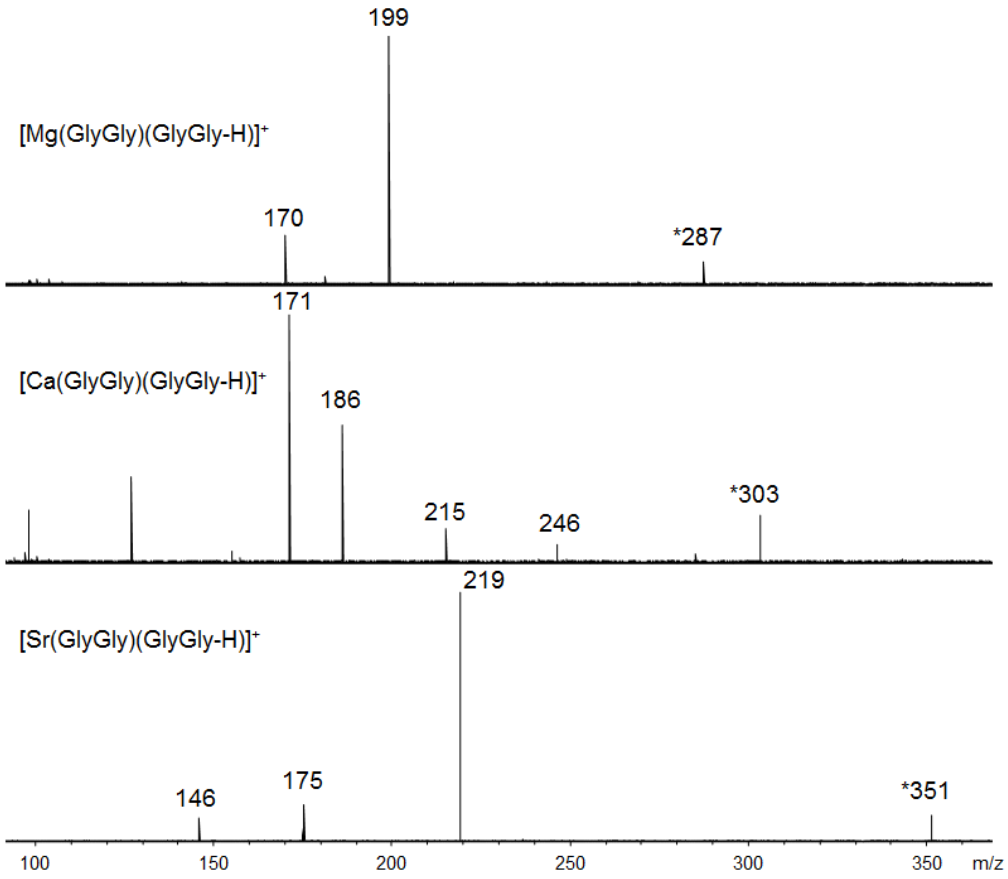


Figure 4.2. SORI/CID mass spectra of $[M(GlyGly)(GlyGly-H)]^+$ for Mg^{2+}, Ca^{2+} and Sr^{2+}.

4.3.1.1.2. $[M_2(GlyGly\text{-}H)_3]^+$

For the $[M_2(GlyGly\text{-}H)_3]^+$ complexes, the common fragment ions observed were due to loss of 57 and 132 Da (Figure 4.3). MS/MS on the fragment ions corresponding to the loss of 57 Da, showed the loss of 75 Da. Therefore, as was the case above, loss of GlyGly (132 Da) cannot be ruled out.

A less pronounced fragmentation for $[Ca_2(GlyGly\text{-}H)_3]^+$ was observed at *m/z* 359, which is also produced during MS/MS of *m/z* 416 and could be explained by loss of glycine-H_2O, following the loss of 57 Da (Figure S4.6). The fragment ion at *m/z* 323 is a secondary (and/or tertiary) loss of H_2O from *m/z* 341. MS/MS on *m/z* 341 did not result in any further losses of GlyGly, but extensive fragmentation of the peptide likely due to very strong Ca-GlyGly binding following the initial loss of a GlyGly molecule from $[Ca_2(GlyGly\text{-}H)_3]^+$. Instead MS/MS of *m/z* 341 produced fragment ions at *m/z* 323, loss of water; *m/z* 284, loss of NH_2CHCO; and *m/z* 281, loss of acetic acid. Other fragment ions observed were *m/z* 305, which is a secondary water loss from *m/z* 323; *m/z* 263 which is loss of acetic acid from *m/z* 323; and *m/z* 251, loss of CO_2 and CO from *m/z* 323 (a small peak is seen at *m/z* 279 in the MS^3 spectrum (Figure S4.7). Finally, *m/z* 297 and 269 in the MS/MS spectrum of *m/z* 341 are losses of CO_2 followed by loss of CO.

The fragment ion at *m/z* 380 in the mass spectrum of $[Sr_2(GlyGly\text{-}H)_3]^+$ is also produced during MS/MS on both *m/z* 512 and 437 and corresponds with a loss of 57 Da followed by loss of 132 Da and loss of 132 Da followed by loss of 57 Da, respectively (Figure S4.8, S4.9). Also the dominant fragment ion at *m/z* 437, related to loss of 132 Da, also resulted upon MS/MS of *m/z* 512. However as it was mentioned above this cannot

exclude the loss of 132 Da from primary fragment ions. Similar to $[Ca_2(GlyGly\text{-}H)_3]^+$, the fragment ion, *m/z* 437 did not show further dissociation of 132 Da. MS/MS on *m/z* 437 produced fragment ions from peptide chain dissociation to small neutral molecules; *m/z* 419, loss of water; 393, loss of CO_2; 380, loss of 57 Da; 377, loss of acetic acid; 365, loss of CO_2+CO and 349, loss of 88 Da related to loss of GlyGly-CO_2 which was also observed during MS/MS on *m/z* 380 upon loss of CH_3NH_2. We were unable to isolate *m/z* 393 to investigate if it shows further dissociation of CO_2 molecule to account for the fragment ion at *m/z* 349 through loss of a second CO_2 instead of acetic acid. The fragment ion *m/z* 512 produced *m/z* 455 which was related to loss of 57 Da. During MS/MS, the fragment ion at *m/z* 380 was dissociated to *m/z* 362, loss of water; *m/z* 349, loss of CH_3NH_2; *m/z* 323, loss of 57 Da; *m/z* 305 loss of glycine and *m/z* 277 related to loss of glycine followed by loss of CO (Figure S4.10).

In the mass spectrum for the Mg^{2+} complexes, a small feature at *m/z* 375 which is a loss of 66 Da corresponds to loss of GlyGly from Mg^{2+} doubly charged cluster ($[M_4(GlyGly\text{-}H)_6]^{2+}$). Less pronounced features observed at *m/z* 423 was related to loss of water and was followed by a second loss of water observed at *m/z* 405. MS/MS on the fragment ion *m/z* 384 also produced a fragment ion at *m/z* 309 upon loss of a glycine molecule (Figure S4.11). However as we observed for both Ca^{2+} and Sr^{2+} loss of 132 Da is one of the primary fragmentation pathways. Other fragment ions stemmed from MS/MS on the *m/z* 384 were *m/z* 366 and 348 related to consecutive losses of water. Other fragment ions were observed at *m/z* 340, loss of CO_2; 337, loss of water followed by loss of 29 Da; *m/z* 327 loss of 57 Da; *m/z* 322, loss of CO_2+water; *m/z* 309, loss of glycine and *m/z* 291 related to loss of glycine followed by loss of water. Furthermore the

fragment ion at *m/z* 280 seems to stem from *m/z* 337 corresponding to loss of 57 Da. Also fragment ion at *m/z* 267 could be loss of acetic acid from *m/z* 327. Another less pronounced feature at *m/z* 252 related to loss of 132 Da from *m/z* 384 was also observed during MS/MS of *m/z* 309 corresponding to loss of 57 Da. The fragment ion at *m/z* 291 was observed as both primary loss of GlyGly from *m/z* 309 followed by loss of water molecule from parent ion and secondary fragmentation during MS/MS on *m/z* 309 (Figure S4.12). Two features observed at lower masses, *m/z* 195 and 177, upon MS/MS on *m/z* 309 were related to loss of 57 Da followed by glycine-H_2O and glycine, respectively, from *m/z* 252, although we were not able to isolate *m/z* 252 to provide more evidence for this observation. MS/MS on both *m/z* 309 and 291 resulted in similar fragment ions. For example *m/z* 273 was observed upon MS/MS on *m/z* 309 was related to consecutive losses of two water molecules. Also other fragment ions observed upon MS/MS on *m/z* 291 were *m/z* 273, loss of water; *m/z* 247, loss of CO_2 and *m/z* 234, loss of NH_2CHCO; *m/z* 218 loss of CO_2+29 Da (Figure S4.13).

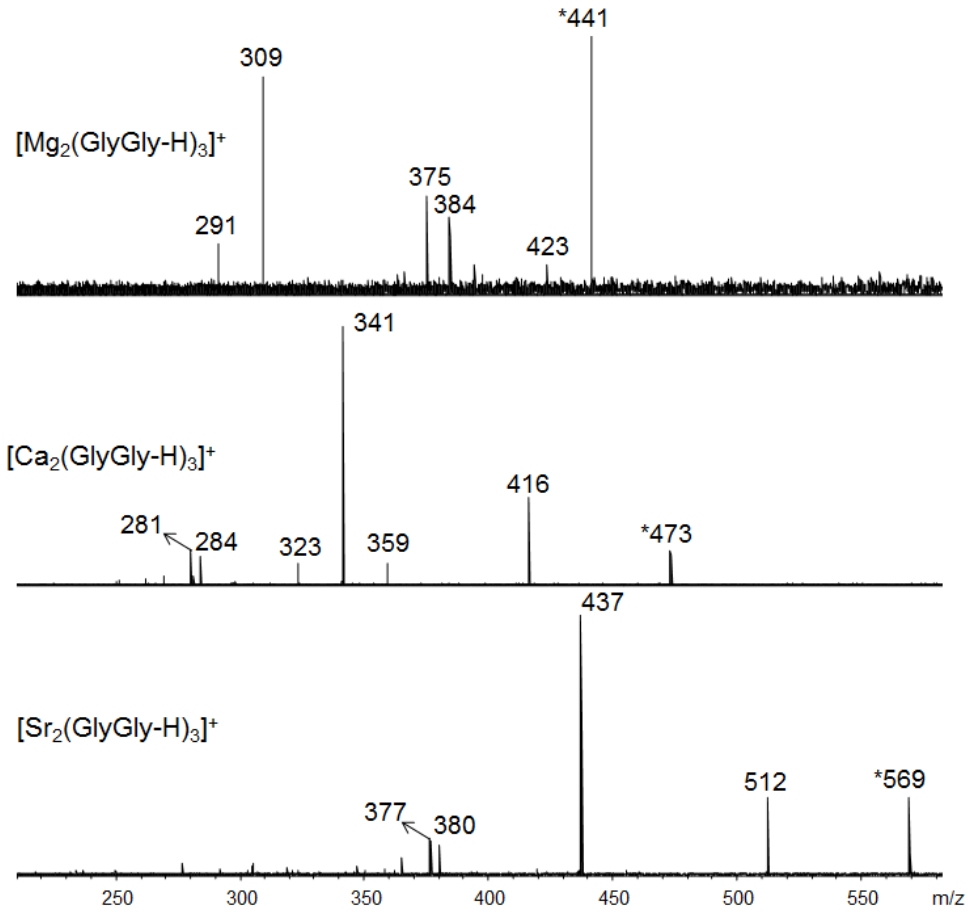


Figure 4.3. SORI/CID mass spectra of $[M_2(GlyGly\text{-}H)_3]^+$ for Mg^{2+}, Ca^{2+} and Sr^{2+}.

4.3.1.1.3. $[M_3(GlyGly\text{-}H)_5]^+$

CID mass spectra of $[M_3(GlyGly\text{-}H)_5]^+$ are compared for Mg^{2+}, Ca^{2+} and Sr^{2+} in Figure 4.4. $[M_3(GlyGly\text{-}H)_5]^+$ Clusters predominantly showed consecutive losses of GlyGly (132 Da) molecules. MS/MS on the fragment ion related to loss of one GlyGly resulted in loss of the second GlyGly.

In the Ca^{2+} clusters MS/MS on *m/z* 511 resulted in *m/z* 454 which was related to loss of 57 Da (Figure S4.14). The fragment ion at *m/z* 436 observed during MS/MS of *m/z*

511 corresponding to loss of glycine was also produced from *m/z* 454 upon loss of a water molecule (Figure S4.15). In a similar way, the fragment ion *m/z* 568 observed upon MS/MS on *m/z* 643 was related to loss of glycine (Figure S4.16). The small intensity of *m/z* 586 stemming from *m/z* 643, however, suggest that *m/z* 568 is also produced upon loss of 57 Da followed by loss of a water molecule from *m/z* 586. The fragment ion at *m/z* 379 is related to the loss of GlyGly molecule from *m/z* 511 however presence of the fragment ion at *m/z* 397 which is produced upon dissociation of 57 Da from *m/z* 454 suggest that *m/z* 379 could be stemmed from *m/z* 397 upon dissociation of a water molecule (Figure S4.16). Other less pronounced features were related to losses of water and CO_2 observed at *m/z* 625 and 599 during MS/MS of *m/z* 643 (Figure S4.16) and at *m/z* 493 and *m/z* 467 upon MS/MS of *m/z* 511 (Figure S4.14). Also the fragment ion at *m/z* 392 was suggested to be stemmed from *m/z* 454 upon loss of water+CO_2 (Figure S4.15).

Similar to the Ca^{2+} cluster, the Sr^{2+} cluster also showed consecutive losses of GlyGly. The fragment ion at *m/z* 598 was produced during MS/MS of both *m/z* 787 and 655 corresponding to consecutive losses of a 132 Da and 57 Da (Figure S4.17, S4.18). Dissociation of a third GlyGly molecule was observed at *m/z* 523 during MS/MS of *m/z* 655. Also a fragment ion observed at *m/z* 611 stemmed from *m/z* 655 related to loss of a CO_2. Other less pronounced fragment ions observed at *m/z* 505 and 595 were related to loss of water from *m/z* 523 and acetic acid from *m/z* 655. Unfortunately we were unable to follow further dissociation of *m/z* 523 to show if it accounts for *m/z* 505.

For the Mg^{2+} cluster, the main dissociation pathway was loss of one GlyGly at *m/z* 595. The other less pronounced fragment ion at *m/z* 463 corresponding to loss of two

GlyGly molecules observed at higher dissociation energies was also produced during MS/MS on the *m/z* 595 (Figure S4.19). Also other fragment ions were observed at *m/z* 577 and 559 upon MS/MS on *m/z* 595 and were related to consecutive losses of two water molecules. The fragment ions at *m/z* 520 and 502 stemmed from *m/z* 595 upon consecutive losses of glycine and water. The presence of a fragment ion at *m/z* 577 suggests that *m/z* 520 can be produced upon loss of 57 Da from 577. Similarly the fragment ion at *m/z* 502 can be produced from *m/z* 559 by loss of 57 Da. MS/MS on the *m/z* 463 resulted in *m/z* 445, 434 and 427 related to losses of one water molecule, 29 Da, and two water molecules (Figure S4.20).

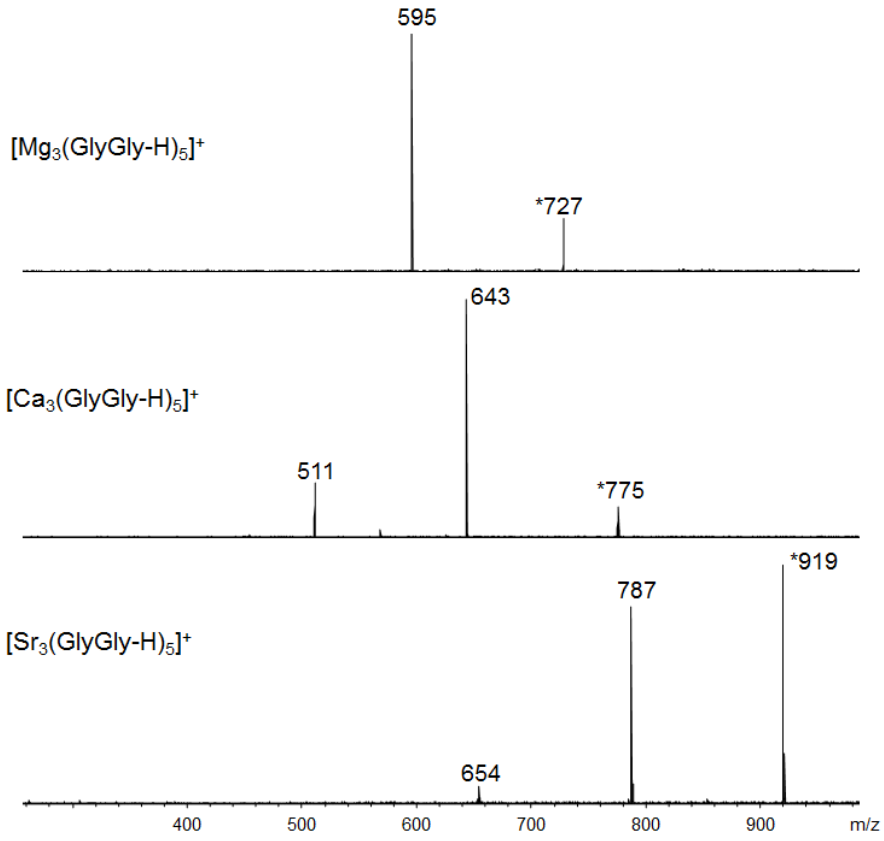


Figure 4.4. SORI/CID mass spectra of $[M_3(GlyGly\text{-}H)_5]^+$ for Mg^{2+}, Ca^{2+} and Sr^{2+}.

4.3.1.1.4. $[M_4(GlyGly\text{-}H)_7]^+$

Because of the low intensities of $[M_4(GlyGly\text{-}H)_7]^+$ for both Mg^{2+} and Sr^{2+} we were unable to study their dissociation reactions under SORI/CID conditions. Fragmentation of $[Ca_4(GlyGly\text{-}H)_7]^+$ predominantly resulted in one intense feature at *m/z* 945 corresponding to loss of one GlyGly and another less pronounced feature at *m/z* 813 corresponding to losses of two GlyGly. MS/MS on the *m/z* 945 resulted in loss of second GlyGly at *m/z* 813 (Figure S4.21). Another weak fragment ion was observed at *m/z* 870 corresponded to loss of glycine molecule. No loss of 57 Da was detected during MS/MS on the *m/z* 945 and this rule out the possibility of consecutive losses of 57 Da and water to account for fragment ion at *m/z* 870. MS/MS on *m/z* 813 also resulted in fragment ions at *m/z* 795, loss of water; *m/z* 738, loss of glycine and *m/z* 681, loss of GlyGly (Figure S4.22). It was obvious that compared to smaller singly charged clusters with diverse dissociation pathways as discussed before, $[Ca_4(GlyGly\text{-}H)_7]^+$ shows predominantly losses of GlyGly molecules.

4.3.1.2. Doubly charged clusters.

4.3.1.2.1. $[M_3(GlyGly\text{-}H)_4]^{2+}$

For the smallest doubly charged clusters studied, $[M_3(GlyGly\text{-}H)_4]^{2+}$ the main dissociation pathway was loss of GlyGly and 57 Da for all metal cations, Figure 4.5. Another fragmentation pathway observed was charge separation of Ca^{2+} and Sr^{2+} clusters to form two singly charged clusters, $[M(GlyGly\text{-}H)(GlyGly)]^+$ and $[M_2(GlyGly\text{-}H)_3]^+$ which will be discussed in more details in their related sections.

In the Ca^{2+} cluster loss of GlyGly and 57 Da were observed at *m/z* 293.5 and 256.0. A fragment ion at *m/z* 265.0 corresponding to loss of 57 Da followed by (glycine-H_2O) was produced as secondary fragmentation during MS/MS on *m/z* 293.5 (Figure S4.23). Additionally primary fragment ion at *m/z* 256.0 was also observed in MS/MS spectrum of *m/z* 293.5 corresponding to loss of 57 Da. The small intensity of a fragment ion at *m/z* 265.0 did not allow us to follow formation of *m/z* 256.0 through loss of water from *m/z* 265.0 however this can be another possible fragmentation pathway. Other less pronounced primary fragment ion at *m/z* 313.0 was related to loss of water molecule from parent ion at *m/z* 322.0. The fragment ion at *m/z* 284.5 corresponding to loss of glycine was also produced upon dissociation of *m/z* 293.5 and we suggested that *m/z* 284.5 is a secondary rather than a primary fragment ion. Another interesting observation was the secondary fragment ion at *m/z* 227.5 which was observed during MS/MS on both *m/z* 293.5 and 256.0 and was related to consecutive losses of GlyGly and 57 Da (Figure S4.23, S4.24). As mentioned briefly above another primary dissociation that was observed at higher dissociation energies was charge separation reaction which resulted in formation of singly charged clusters at *m/z* 303.0 and *m/z* 473.0 which can be explained by the high charge density of the doubly charged clusters that leads to dissociation of cluster to singly charged clusters. A more detailed discussion of this comparison is provided by Table 4.1.

Primary fragment ions related to losses of GlyGly and 57 Da were identified at *m/z* 328.0 and 365.5 for the Sr^{2+} cluster. The fragment ion *m/z* 328.0 was also produced during MS/MS on *m/z* 365.5 however this cannot exclude this fragment ion from primary fragmentation ions (Figure S4.25). Less pronounced features at *m/z* 299.5 and 262.0

observed upon MS/MS on *m/z* 328.0 were related to losses of 57 Da and GlyGly (Figure S4.26). Formation of singly charged clusters is considered to be a primary fragmentation in Sr^{2+} cluster and it is a more dominant dissociation pathway compared to Ca^{2+} cluster. The singly charged fragment ion at *m/z* 569.0 was further dissociated to *m/z* 512.0 and 437.0 related to loss of 57 Da and 132 Da as we showed in the singly charged clusters section.

The Mg^{2+} cluster losses of GlyGly and 57 Da as well as GlyGly + NH_2CHCO were observed at *m/z* 232.1, 269.6 and 203.6. In contrast to Ca^{2+} and Sr^{2+} clusters, no charge separation reaction was observed. The only explanation for this observation is that the higher charge density of Mg^{2+}, compared to Ca^{2+} and Sr^{2+} makes stronger interactions with the ligands. Since metal cations are coordinated with more than one ligand in these clusters, stronger metal-ligand interactions in Mg^{2+} cluster result in a strongly bound cluster that shows no charge separation reaction upon dissociation. Another fragment ion at *m/z* 289.1 was related to loss of a water molecule from the parent ion and the fragment ion at *m/z* 267.6 is likely loss of CO_2 from *m/z* 289.1. A less pronounced fragment ion at *m/z* 203.6 related to loss of 132 Da followed by 57 Da was also produced during MS/MS on *m/z* 232.1 (Figure S4.27). No further losses of 132 Da was observed as secondary fragmentation upon MS/MS on *m/z* 269.6 and the only fragmentations observed were at *m/z* 241.1 and 260.6 related to loss of glycine-H_2O and water, respectively (Figure S4.28).

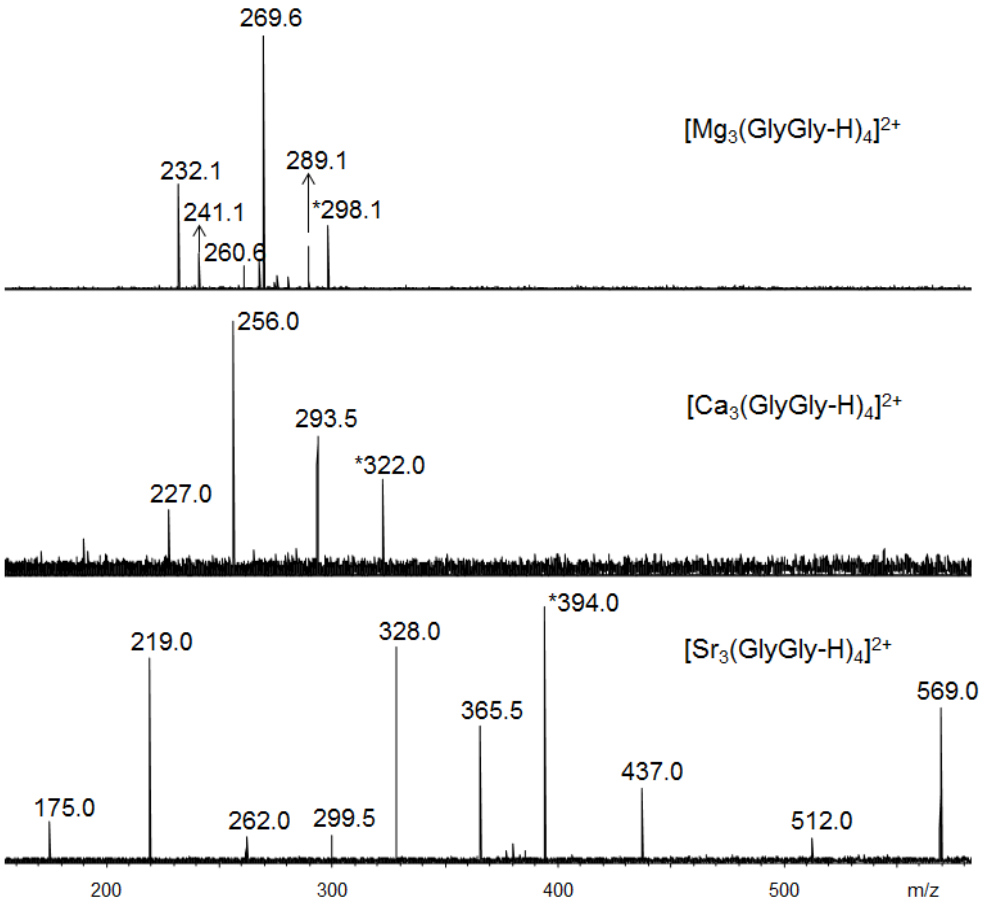


Figure 4.5. SORI/CID mass spectra of $[M_3(GlyGly\text{-}H)_4]^{2+}$ for Mg^{2+}, Ca^{2+} and Sr^{2+}.

4.3.1.2.2. $[M_5(GlyGly\text{-}H)_8]^{2+}$

Dissociation of the $[M_5(GlyGly\text{-}H)_8]^{2+}$ cluster for both Ca^{2+} and Sr^{2+}, predominantly resulted in two singly charged clusters, $[M_2(GlyGly\text{-}H)_3]^+$ and $[M_3(GlyGly\text{-}H)_55]+$, at m/z 473.1 and 775.0, respectively, for Ca^{2+} and m/z 569.0 and 919.1, respectively, for Sr^{2+}, Figure 4.6.

A less-pronounced feature was observed at m/z 558.1 was related to a loss of 132 Da for Ca^{2+}. MS/MS on the singly charged clusters at *m/z* 775.0 and 473.1 resulted in

primary fragmentation pathways that were discussed earlier for the singly charged clusters.

Other primary fragmentation observed for the Sr^{2+} cluster was related to further dissociation of the singly charged cluster *m/z* 919.0 to *m/z* 787.0 and 655.0 corresponding to consecutive losses of GlyGly as we observed in the SORI/CID of the singly charged clusters above. Also less pronounced feature at *m/z* 437.0 corresponding to loss of a GlyGly molecule was stemmed from singly charged cluster at *m/z* 569.0 (Figure S4.29).

In contrast to Ca^{2+} and Sr^{2+}, the Mg^{2+} cluster's main fragmentation pathway was observed to be loss of one GlyGly molecule at *m/z* 519.2, at higher SORI/CID a second loss of GlyGly is observed at *m/z* 453.2 which is also produced during MS/MS on *m/z* 519.2 and is followed by loss of water at *m/z* 444.2 (Figure S4.30).

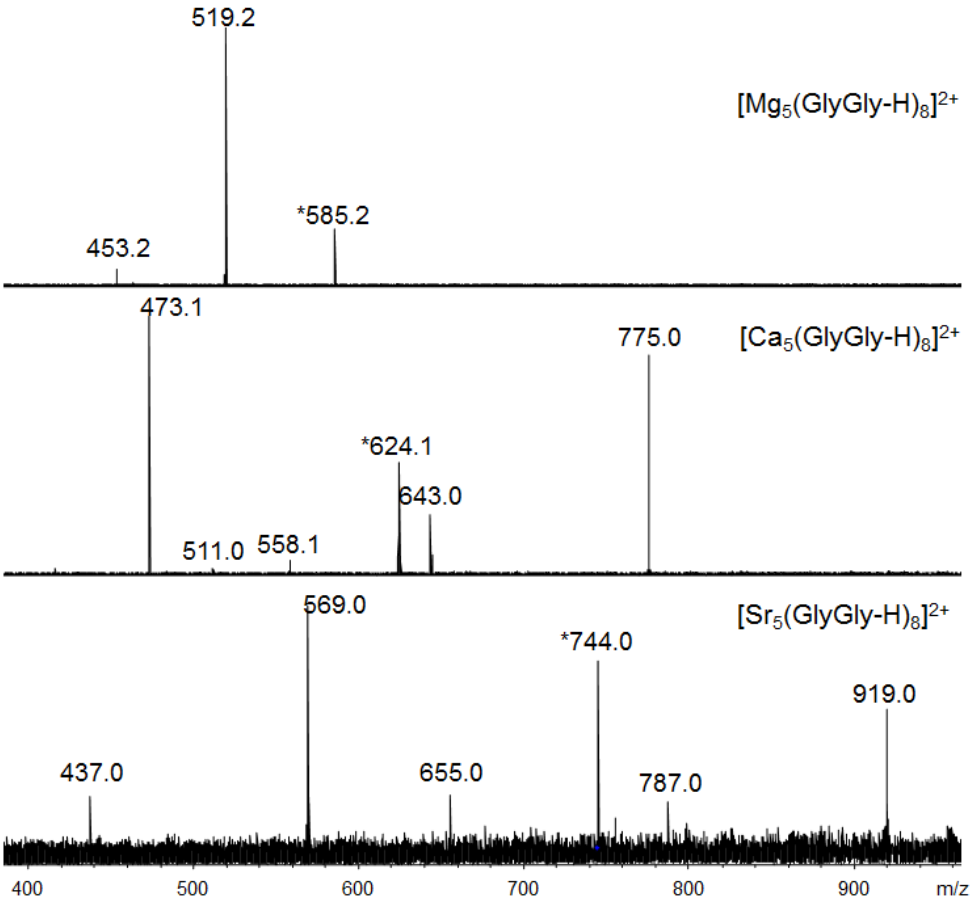


Figure 4.6. SORI/CID mass spectra of $[M_5(GlyGly\text{-}H)_8]^{2+}$ for Mg^{2+}, Ca^{2+} and Sr^{2+}.

4.3.1.2.3. $[M_7(GlyGly\text{-}H)_{12}]^{2+}$

The main dissociation pathway for the largest doubly charged cluster, $[M_7(GlyGly\text{-}H)_{12}]^{2+}$, for both Ca^{2+} and Sr^{2+}, was a charge separation reaction which resulted in formation of singly charged clusters $[M_3(GlyGly\text{-}H)_5]^+$ and $[M_4(GlyGly\text{-}H)_7]^+$ at *m/z* 775.1 and 1077.2 for Ca^{2+} and *m/z* 919.0 and 1269.1 for Sr^{2+}, respectively, Figure 4.7. In addition, losses of GlyGly from both parent ions as well as the fragment ions were observed at lower intensities.

In the Ca^{2+} cluster two peaks at *m/z* 860.3 and 794.3 were observed corresponding to sequential losses of GlyGly from *m/z* 926.3. Also, features at *m/z* 945.2 and 813.2 are consecutive losses of GlyGly from the singly charged fragment ion, *m/z* 1077.2. Similarly, *m/z* 643.1 is a fragment ion of *m/z* 775.1 formed by loss of one GlyGly molecule as we observed before.

In Sr^{2+} cluster, besides charge separation reactions, other primary fragmentation were sequential losses of GlyGly observed at *m/z* 1028.1 and 962.1. Other interesting features were the fragment ions at *m/z* 1137.1 and 1005.0 which originate from *m/z* 1269.1 corresponding to consecutive losses of GlyGly molecules (Figure S4.31). The fragment ion *m/z* 1005.0 was also produced during MS/MS on *m/z* 1137.1 (Figure S4.32). At higher collision energies further dissociation of the singly charged fragment ion, $[M_3(GlyGly\text{-}H)_5]^+$, at *m/z* 919.0 resulted in fragment ion at *m/z* 787.0 corresponding to loss of one GlyGly molecule.

Interestingly, the only dissociations observed for the Mg^{2+} cluster were consecutive losses of GlyGly observed at *m/z* 805.0, 739.0, 673.0, 607.0 and 541.0. Low intensity of these fragment ions limited further fragmentation studies to this point.

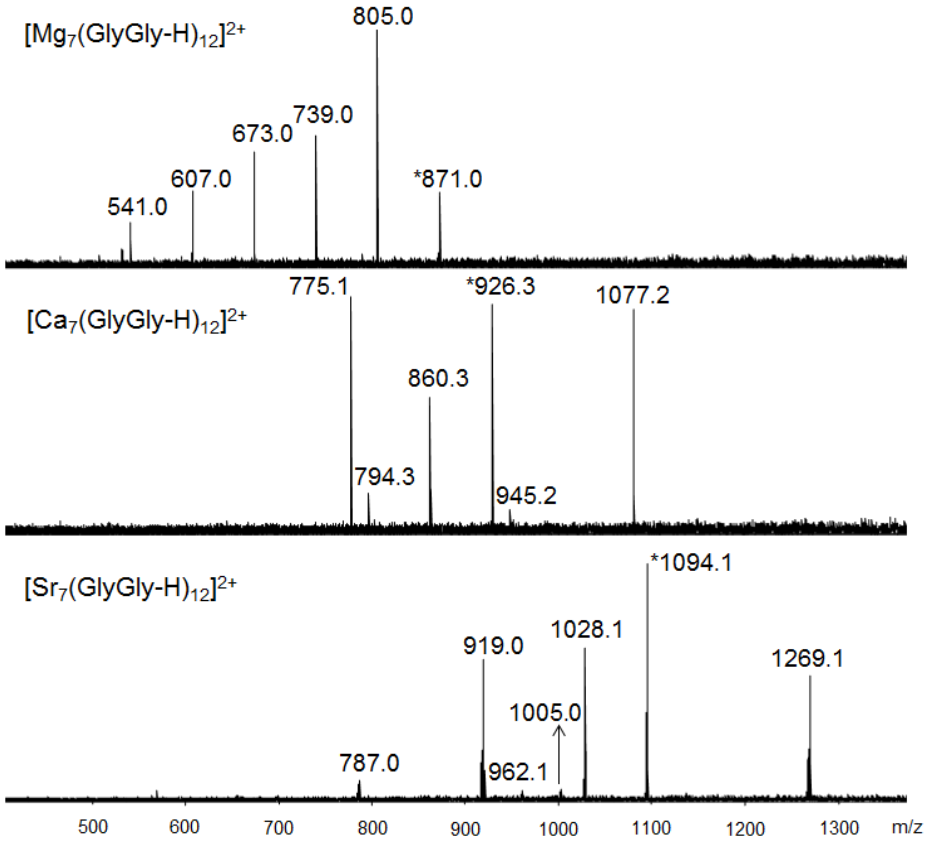


Figure 4.7. SORI/CID mass spectra of $[M_7(GlyGly\text{-}H)_{12}]^{2+}$ for Mg^{2+}, Ca^{2+} and Sr^{2+}.

4.3.1.3. Energy-resolved CID results

In Table 4.1 the collision energy (centre of mass) at half intensity of parent ion of two singly charged $[M_2(GlyGly\text{-}H)_3]^+$, $[M_3(GlyGly\text{-}H)_5]^+$ and two doubly charged $[M_3(GlyGly\text{-}H)_4]^{2+}$ and $[Ca_5(GlyGly\text{-}H)_5]^{2+}$ clusters are compared for all metal ions. Comparison between singly and doubly charged clusters for each of the metal cations showed that larger clusters in both cases have lower dissociation energies. One possible explanation for this observation is that larger clusters are weakly bound compared to

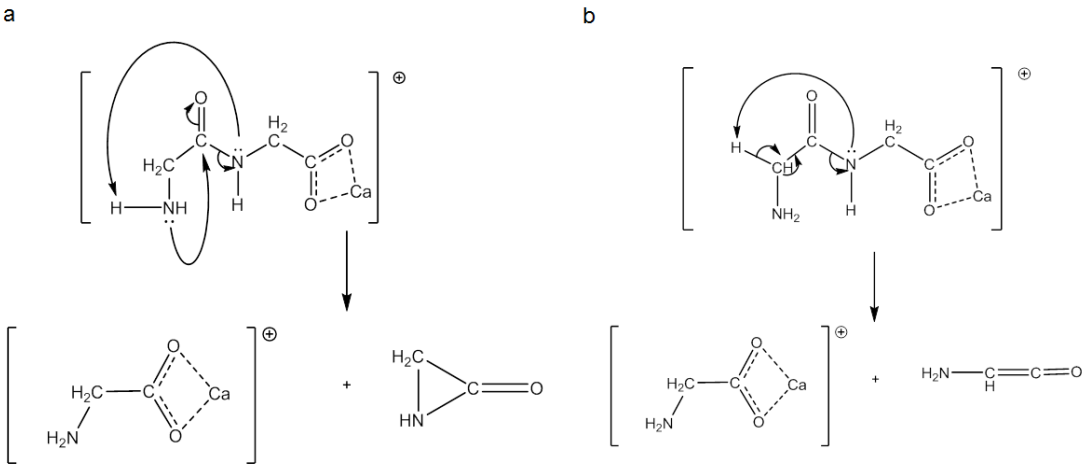
smaller clusters. Another interesting point is the effect of cluster charge on the dissociation energy of these clusters. It was also shown that doubly charged clusters more easily dissociate than singly charged clusters and this is because higher charge density localized on the cluster. Comparing the dissociation energies of these metal cation clusters also showed that the trend for collision energy for singly charged clusters was as Ca^{2+}>Mg^{2+}>Sr^{2+} whereas in the doubly charged clusters it showed the order of Sr^{2+}>Ca^{2+}>Mg^{2+}. In singly charged clusters there is an increase in the dissociation energy in Ca^{2+} cluster. We proposed that the geometry of singly charged clusters favors the Ca cation size however this experimental observation needs to be further investigated by computational methods in order to provide a more conclusive explanation. In contrast in doubly charged clusters however, the charge density of the metal cation seems to be the determining factor since, for example, extra positive charge density is localized on the metal cation in the case of Mg^{2+} rather than Ca^{2+} and Sr^{2+}.

Table 4.1. Dissociation energies (eV) at 50% of $[M_2(GlyGly\text{-}H)_3]^+$, $[M_3(GlyGly\text{-}H)_5]^+$ $[M_3(GlyGly\text{-}H)_4]^{2+}$ and $[Ca_5(GlyGly\text{-}H)_8]^{2+}$ M=Mg^{2+}, Ca^{2+} and Sr^{2+}.

Cluster	Mg^{2+}	Ca^{2+}	Sr^{2+}
$[M_2(GlyGly\text{-}H)_3]^+$	0.88	0.95	0.87
$[M_3(GlyGly\text{-}H)_4]^{2+}$	0.12	0.18	0.24
$[M_3(GlyGly\text{-}H)_5]^+$	0.40	0.6	0.15
$[M_5(GlyGly\text{-}H)_8]^{2+}$	0.06	0.12	0.14

4.3.1.3. NH_2CHCO fragment ion

So far we observed that the main and primary fragmentation of all singly charged clusters was loss of a GlyGly. Another widely seen fragment ion was loss of 57 Da, presumably NH_2CHCO as a result of peptide bond cleavage which is known to be a weak bond. In Scheme 4.1 two mechanisms are proposed based on studies by Kenny and coworkers on protonated peptide dissociation[43], for dissociation of C_2H_3NO from $[Ca(GlyGly\text{-}H)]^+$ for simplicity. In the mechanism a, one hydrogen is transferred from the amine group to the amide nitrogen parallel to nucleophilic attack of amine nitrogen at the carbonyl group which eventually leads to a cyclic product molecule. In mechanism b hydrogen is transferred from the neighboring CH_2 group to the amide nitrogen which results in a linear (aminoketene) product molecule.



Scheme 4.1. Two possible channels for loss of 57 Da from $[Ca(GlyGly\text{-}H)]^+$. Representation for dissociation mechanism for formation of a cyclic (a) and a linear (b) neutral product with C_2H_3NO formula.[43]

Computationally we observed that the loss of aminoketene is lower in energy than the cyclic form by 39.7 and 34.4 kJ mol^{-1} in enthalpy and 298 K Gibbs energy, respectively, in accordance with the findings of McCormack on formation of y ions from dissociation of protonated peptides.[44] However Kenny and his coworkers[43] showed by extensive isotopic labeling experiments on GGGGA peptide that it is the amine hydrogen that is transferred during dissociation. In order to find the dominant fragmentation pathway for loss of 57 Da as well as GlyGly, isotopic labeling experiments were performed on the singly charged $[Ca_2(GlyGly\text{-}H)_3]^+$ cluster at *m/z* 473.

A solution of GlyGly was prepared using D_2O instead of ultrapure water. $[Ca_2(GlyGly\text{-}H)_3]^+$ was observed at *m/z* 482 after substitution of all nine exchangeable hydrogen atoms with deuterium. Presumably the carboxyl hydrogens have been lost (*vide infra*) leaving the three labile amine and amide hydrogen atoms for H-D exchange in each of the three GlyGly's in the clusters. CID experiments on the labeled clusters produced fragment ions resulting from loss of 59 Da at *m/z* 423 which showed dissociation in which two of the labels were lost. However it is possible that another isotopic isomer at *m/z* 483 is still present upon isolation of *m/z* 482. Low intensity of *m/z* 424 from dissociation of *m/z* 482 is observed as well as *m/z* 423. It is therefore proposed that hydrogen is transferred from the CH_2 to the amide group and as a result loss of 59 Da is related to an aminoketene (mechanism b in Scheme 4.1).

Another observation from these isotopic substitution studies was a loss of 136 Da corresponding to a quadruply labelled GlyGly molecule. This result shows that loss of GlyGly is accompanied by a deuterium from a neighboring GlyGly moiety—a labile hydrogen from a neighboring GlyGly.

4.3.2. Computational study

For the computational aspect of this work, we only focused on the *m/z* 322 and 624 ions, $[M_3(GlyGly\text{-}H)_4]^{2+}$ and $[M_5(GlyGly\text{-}H)_8]^{2+}$.

4.3.2.1. $[M_3(GlyGly\text{-}H)_4]^{2+}$ structure

In a previous study we investigated the structures of complexes formed from Ca^{2+} bound to deprotonated AlaGly and GlyAla.[39] It was shown that the lowest energy structure of $[Ca(AlaGly\text{-}H)]^{+}$ is deprotonated at the carboxylic acid group and that the metal directly interacts with amide and carboxylate oxygen atoms. Comparing the experimental IR spectra with the results obtained from DFT calculations a red-shifted amide N-H stretch was concluded to be a result of the amine nitrogen hydrogen bonding with the amide hydrogen.

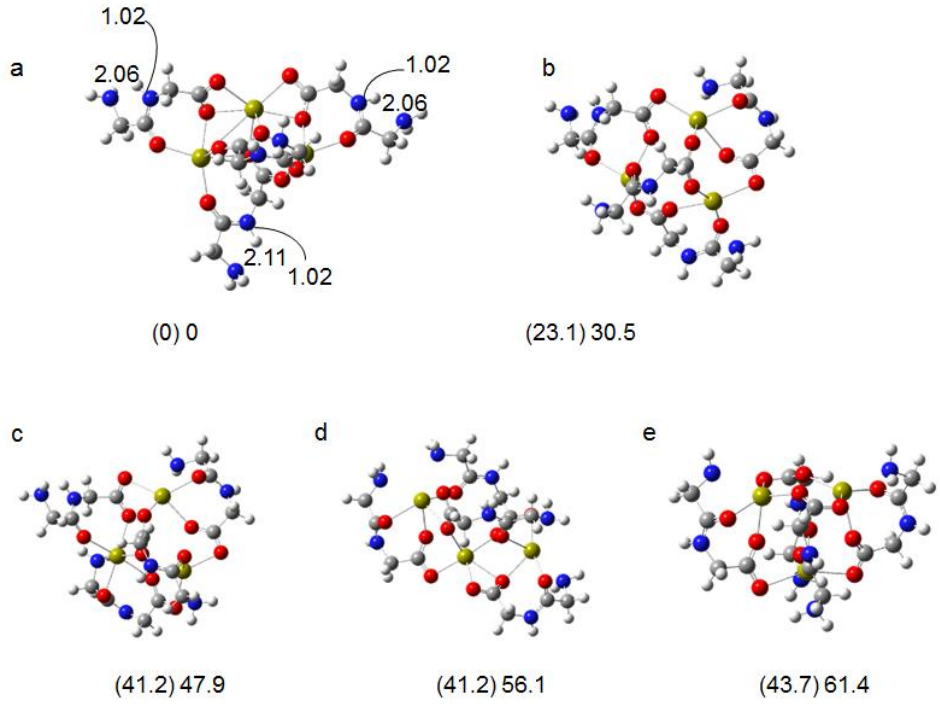


Figure 4.8. Optimized structures for $[Ca_3(GlyGly\text{-}H)_4]^{2+}$ cluster. 298 K relative enthalpies (left) and Gibbs energies (right) are reported in kJ mol^{-1} using B3LYP/6-311++G(3df,3pd)//B3LYP/6-31+G(d,p) level of theory and basis set.

In Figure 4.8 some optimized lowest energy structures obtained from simulated annealing and DFT calculations on $[Ca_3(GlyGly\text{-}H)_4]^{2+}$ are shown. The relative enthalpy and Gibbs energies of these complexes are reported from DFT calculations with B3LYP/ 6-311++G(3dp,3fd)//B3LYP/6-31+G(d,p) level of theory and basis set. In these complexes, the four $(GlyGly\text{-}H)^-$ ligands in this fall into two types, Figure 4.9. In type A both carboxylate oxygens are bound to the same Ca^{2+}, and type B where each carboxylate oxygen is bound to a different Ca^{2+}. The type A ligands are also oriented in the same plane made by all three calcium ions and are bound to only two of the three calcium ions whereas the type B ligands are oriented out of the plane and bound to all three Ca^{2+}. In the lowest energy structure obtained for $[Ca_3(GlyGly\text{-}H)_4]^{2+}$, the metal ions form an

obtuse and almost isosceles triangle with bond distances of 3.6, 3.7 and 5.5 Å. All carbonyl oxygen atoms are interacting with the metal and terminal amine nitrogen atoms are hydrogen bonded to the amide hydrogen atom of the same ligand, similar to the $[Ca(AlaGly-H)]^+$ complex. Also it is in agreement with the lowest energy structure of Na-AlaAla complexes reported by the Balaj group in which the metal cation is coordinated with all carbonyl oxygens.[45] One of the Ca^{2+}, at the apex of the obtuse triangle, is coordinated with six carbonyl oxygens; two pairs from the carboxylate groups of type A $(GlyGly-H)^-$ ligands and two oxygen atoms coming from the carboxylate carbonyl groups of both type B $(GlyGly-H)^-$ ligands. Five carbonyl oxygens participate in coordination of each of the other two Ca^{2+} centers; three carboxylate carbonyl group oxygen atoms, one from of a type A and two from type B ligands and two amide carbonyl oxygens one from a type A and one from a type B $(GlyGly-H)^-$ ligands. Another interesting feature is that one of the carboxylate oxygens in each of the four ligands bridges two Ca^{2+} ions. In the second lowest energy structure (b), which is 23.1 and 30.5 kJ mol^{-1} higher in enthalpy and Gibbs energy, the Ca^{2+} ions form a more or less equilateral triangle with bond lengths of 4.6, 4.5, and 4.2 Å. Three of the four amine nitrogens are interacting with the metal cation, completing an octahedral coordination shell with the other five members being oxygen atoms of the amide carbonyl and carboxylate oxygens. Similar to the second structure there is hydrogen bonding in structure (c), but the orientation of a GlyGly and interaction of carbonyl and amine nitrogen with the metal centers is different which results in a structure with relative enthalpy and Gibbs energy of 41.2 and 47.9 kJ mol^{-1}. In structure (d) two Ca^{2+} ions are solvated each with two amine nitrogen atoms as well as the

carbonyl oxygen atoms. The last structure (e) is similar to (d) in that all terminal amines are ligated to the metal cations and each cation is roughly octahedrally coordinated. Coordination of metal ions with terminal amine nitrogen atoms in structure (d) and (e) occurs by losing intramolecular hydrogen bondings and results in higher relative enthalpy and Gibbs energies of 41.2, 43.7, 56.1 and 61.4 kJ mol^{-1}, respectively.

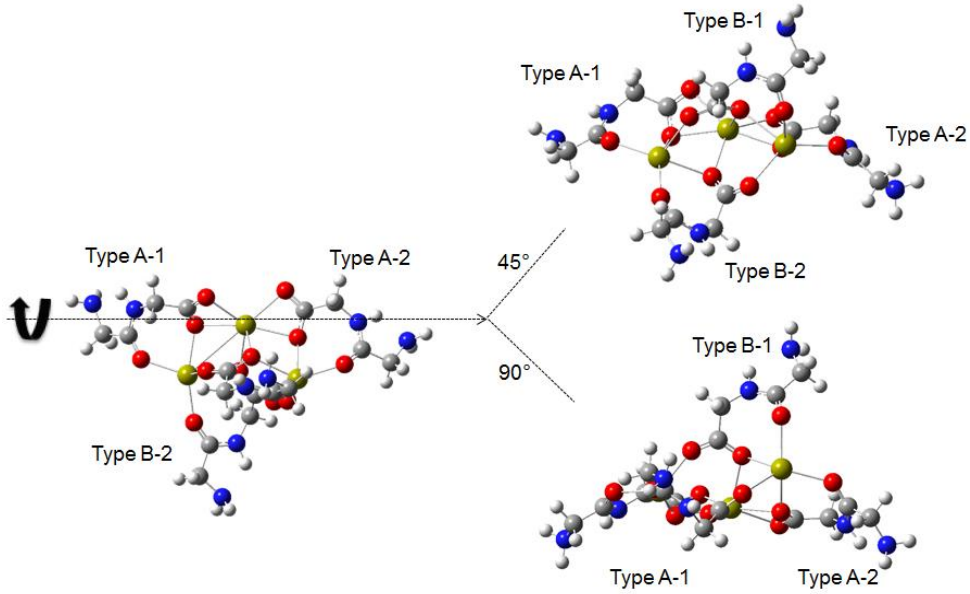


Figure 4.9. Presentation of GlyGly ligands types according to their positions in the $[Ca_3(GlyGly\text{-}H)_4]^{2+}$ structure. On the top right panel the structure is rotated by 45° around the horizontal axis that passes through both Type A GlyGly ligands and on the bottom right panel structure is rotated by 90°.

The five lowest energy structures of $[Ca_3(GlyGly\text{-}H)_4]^{2+}$ were taken as starting geometries for the Mg^{2+} and Sr^{2+} complexes. The lowest energy structures of Mg^{2+}, Ca^{2+} and Sr^{2+} are compared in Figure 4.10. Also all the five lowest energy structures for $[Mg_3(GlyGly\text{-}H)_4]^{2+}$ and $[Sr_3(GlyGly\text{-}H)_4]^{2+}$ are reported in supplementary data (Figure S4.33 and S4.34). In $[Mg_3(GlyGly\text{-}H)_4]^{2+}$ and $[Sr_3(GlyGly\text{-}H)_4]^{2+}$ clusters, similar to

$[Ca_3(GlyGly\text{-}H)_4]^{2+}$ cluster, the metal cations are interacting with all carbonyl oxygens and all amine nitrogens participate in hydrogen bonding to amide hydrogen in the lowest energy structures found. It was observed that the differences in these structures mainly come from the difference in metal ion radius which results in bond lengths involving the metal binding site to be in the order of $Mg^{2+} < Ca^{2+} < Sr^{2+}$ as expected. In addition the hydrogen bond length between amine nitrogen and amide hydrogen is also increased slightly with the same order as increasing the metal cation size. The hydrogen bond lengths for two hydrogen bondings are shown in Figure 4.9 for all three metal cations clusters. It can be seen the hydrogen bond lengths are 2.03 and 2.08 Å for the Mg^{2+} cluster, 2.06 and 2.11 Å for Ca^{2+}, and 2.07 and 2.12 Å in the Sr^{2+} cluster, respectively. This can be explained by the interaction of metal cation with carbonyl oxygens which is dependent on the metal cation size and charge density. A larger ion results in a less dense positive charge, less electron withdrawal from the peptide which results in a weaker and longer amide N-H to amine nitrogen hydrogen bond.

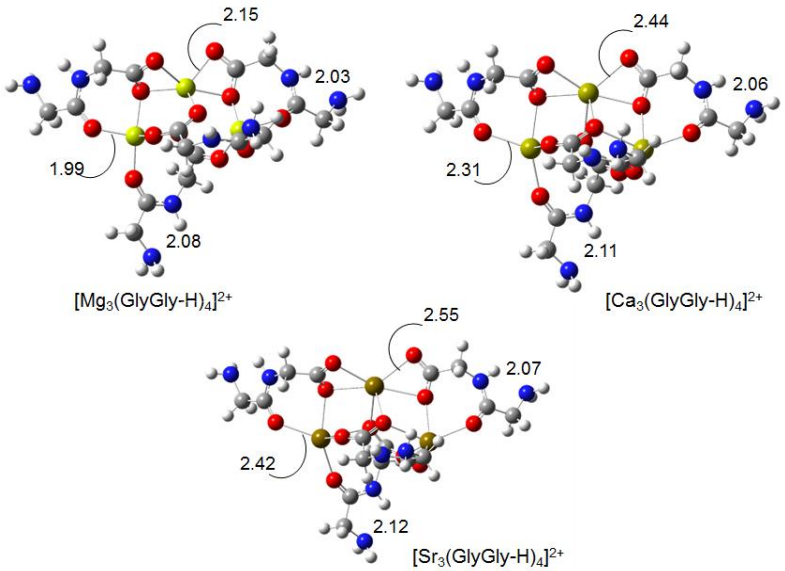


Figure 4.10. Lowest energy structures for $[M_3(GlyGly\text{-}H)_4]^{2+}$ cluster ($M=Mg^{2+}$, Ca^{2+} and Sr^{2+}) reported using B3LYP/6-31+G(d,p) on C, H, O and N and Def2SVP on all metal.

4.3.2.2. $[M_3(GlyGly\text{-}H)_4]^{2+}$ fragmentation mechanism

As we observed experimentally, the main fragmentation of the $[M_3(GlyGly\text{-}H)_4]^{2+}$ cluster is loss of neutral NH_2CHCO and GlyGly molecules. DFT calculations were performed using B3LYP/6-311++G(3df,3pd)//B3LYP/6-31+G(d,p) level of theory and basis set in order to find the lowest energy dissociation pathway for the $[Ca_3(GlyGly\text{-}H)_4]^{2+}$ cluster as a representative for all cluster cations. We assumed that the dissociating deprotonated GlyGly ligand leaves with a proton from the amide group of another deprotonated ligand GlyGly. Any attempt for optimizing structures in which the proton was stemmed from amine rather than amide group was unsuccessful and they eventually became deprotonated at amide group. The energetics of dissociation for the other three GlyGly ligands was also explored. It was observed that loss of the type B-1 GlyGly

(Figure 4.9) which stands perpendicular to the average molecular plane and closest to type A-1 and A-2 ligands is the least energy demanding compared with the loss of a type A GlyGly by 83.2 and 85.9 kJ mol^{-1} in enthalpy and Gibbs energies, respectively. Also loss of the type B-2 GlyGly which is in the molecular plane with other GlyGly moieties is less favored by 148.1 and 135.2 kJ mol^{-1} in enthalpy and Gibbs energy with respect to loss of type B-1 GlyGly. Calculations also show that the most energetically favored dissociation was when the proton was transferred from the type B-2 GlyGly to the type B-1 GlyGly.

The enthalpy and Gibbs energy for loss of a GlyGly molecule from the $[Ca_3(GlyGly\text{-}H)_4]^{2+}$ clusters are also computed for Mg^{2+} and Sr^{2+}, Table 4.2. Interestingly, the order of binding energies of GlyGly to these clusters is $Ca^{2+}>Mg^{2+}>Sr^{2+}$, in accordance with the trend that we observed for collision energy at half intensity of the singly charged cluster $[M_2(GlyGly\text{-}H)_3]^{+}$. This can be explained with the same reason stating that Ca^{2+} can form stronger interactions with oxygen atoms compared to Mg^{2+} and Sr^{2+}. Furthermore, higher binding energy of Mg^{2+} compared to Sr^{2+} can be related to the stronger interaction of Mg^{2+}.

Table 4.2. DFT computed (using B3LYP/6-311++G(3df,3pd)//B3LYP/6-31+G(d,p) level of theory and basis set) binding enthalpies and Gibbs energies in kJ mol^{-1} at 298 K between GlyGly and $[M_3(GlyGly\text{-}H)_4]^{2+}$ ($M=Mg^{2+}$, Ca^{2+} and Sr^{2+}) Clusters.

M	ΔH (kJ mol^{-1})	ΔG (kJ mol^{-1})
Mg^{2+}	239.3	184.6
Ca^{2+}	250.3	199.9
Sr^{2+}	228.9	179.3

4.3.2.3. $[Ca_5(GlyGly\text{-}H)_8]^{2+}$ structure

The lowest energy structure found for $[Ca_5(GlyGly\text{-}H)_8]^{2+}$ was similar to what we showed for the smaller cluster $[M_3(GlyGly\text{-}H)_4]^{2+}$ in the sense that amine nitrogen atoms are hydrogen bonded to amide hydrogens and all carbonyl oxygen atoms are interacting with metal cations, (Figure S4.35). In the second lowest energy structure which is 63.3 and 78.4 kJ mol^{-1} higher in enthalpy and Gibbs energy, respectively, four terminal amine nitrogen atoms have participated in intramolecular hydrogen bonding with the amide hydrogen and two terminal amine nitrogen atoms are directly interacting with the metal ion and the other two are interacting with each other through intermolecular hydrogen bonding. The third lowest energy structure is 78.1 and 99.5 kJ mol^{-1} higher in enthalpy and Gibbs energies in which five amine nitrogens hydrogen bonded to amide hydrogens. Two other amine nitrogens stand close to each other in a way that they form intra- and intermolecular hydrogen bonding with amide hydrogens. Finally one of the amine nitrogens has participated in metal ion coordination. In the fourth lowest energy structure,

similar to the third lowest energy structure, five amine nitrogens are hydrogen bonded to amide hydrogens. One of the amine nitrogen is interacting with metal ion and two other amine nitrogens are at close distance and one forms intermolecular interaction with the other amine and amide hydrogens. This structure is 97.8 and 110.5 kJ mol^{-1} higher in enthalpy and Gibbs energies. At last, the fifth lowest energy structure which is 92.3 and 119.7 kJ mol^{-1} higher in enthalpy and Gibbs energy shows different intra- and intermolecular hydrogen bondings. Only three amine nitrogens are hydrogen bonded to the amide hydrogens. One of amine nitrogens is intermolecularly hydrogen bonded to neighboring amine hydrogen which has already participated in metal ion coordination. Two neighboring amine nitrogens have participated in intermolecular hydrogen bonding with the amide hydrogens. Furthermore amine nitrogen is interacting with metal ion.

4.3.3. IRMPD spectroscopic studies

4.3.3.1. $[M_3(GlyGly\text{-}H)_4]^{2+}$

The IRMPD spectra as well as calculated IR spectra of lowest energy structures of $[M_3(GlyGly\text{-}H)_4]^{2+}$ ($M=Mg^{2+}$, Ca^{2+} and Sr^{2+}) in the C-H/N-H/O-H stretching region (2700-3800 cm^{-1}) are presented in Figure 4.11(a-c). The IRMPD spectrum of Mg^{2+} complex contains features centered at 3300, 3390 and 3460 cm^{-1}. The feature at 3300 cm^{-1} corresponds with a red-shifted amide N-H stretch of the GlyGly-H. This was in agreement with what we observed for $[Ca(GlyAla\text{-}H)]^{+}$ cluster in a previous work.[39] The features at 3390 and 3460 cm^{-1} are indicative of symmetric and antisymmetric amine N-H stretches. In the IRMPD spectrum of Ca^{2+} complex there are two features around 3330

and 3460 cm^{-1}. The 3330 cm^{-1} band again is a red-shifted amide N-H stretch and is unresolved from the symmetric NH_2 mode. The antisymmetric NH_2 stretch corresponds to the 3460 cm^{-1} band. A similar spectrum was observed for the Sr^{2+} complex with bands occurring at 3350 and 3460 cm^{-1}. The computed IR spectra for the lowest energy structure for each complex are most consistent with the IRMPD spectra, especially for the Mg^{2+} complex. The computed spectra for the Ca^{2+} and to a greater extent the Sr^{2+} cluster show less resolution for the symmetric NH_2 stretching band from the amide N-H stretch and due to the blue-shifting of this experimental band (from 3300 to 3330, to 3350 cm^{-1}) merge of the two bands may be expected. Also, the blue shift itself is consistent with less densely charged metal cations, resulting in less positive charge on the ligand, a weaker amide to amine hydrogen bond, and ultimately a slightly stronger amide N-H bond. It is not possible to rule out the second lowest energy structures since they do share the same features as in the experimental spectrum.

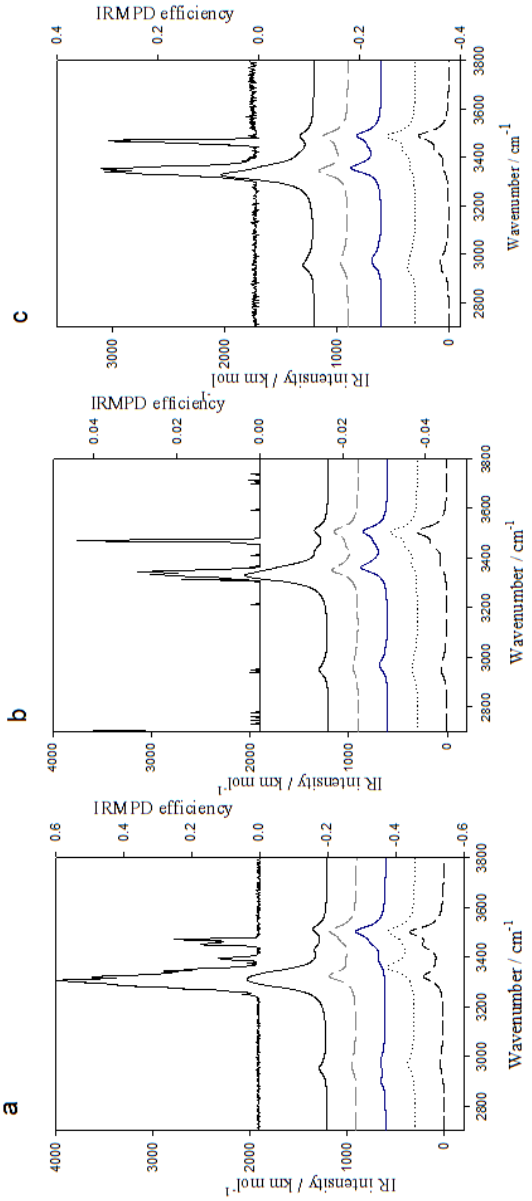


Figure 4.11. Comparison of the IRMPD spectra and the computed spectra for five lowest energy structures of $[M_3(GlyGly\text{-}H)_4]^{2+}$ obtained by B3LYP method 6-31+G(d,p) on Mg^{2+} (a) and Ca^{2+} (b) and B3LYP/6-31+G(d,p) on C, H, N and O and Def2SVP on Sr^{2+} (c).

4.3.3.2. $[M_5(GlyGly\text{-}H)_8]^{2+}$

The IRMPD spectra of $[Ca_5(GlyGly\text{-}H)_8]^{2+}$ are compared to the calculated IR spectra of five lowest energy structures in both fingerprint (900-2000 cm^{-1}) and N-H/O-H stretching region (2700-3800 cm^{-1}), Figure 4.12(a,b). In the finger print region there is a broad intense band between 1500 and 1700 cm^{-1} which corresponds to red-shifted carboxylic acid C=O stretch and amide C=O stretch (1690-1760 cm^{-1}). Also the N-H bending mode and $-NH_2$ scissoring mode were observed around 1500-1580 cm^{-1} respectively. The band around 1350-1450 cm^{-1} corresponds to $-CH_2$ scissoring mode. In addition a less pronounced band was observed around 1300 cm^{-1} which was related to the $-CH_2$ wagging mode. Comparing these spectra it was shown that the experimental features can be reproduced by all five low energy structures calculated IR spectra. In this regard it is impossible to distinguish between these structures based on their spectra in this region.

In the experimental IRMPD spectrum in N-H/O-H region there is one broad feature at 3370 cm^{-1}, Figure 4.12(b). This feature is well reproduced with the computed spectrum of the lowest-energy structure. These bands are related to the red-shifted amide N-H stretch as well as the symmetric/asymmetric NH_2 stretches which are all predicted to occur at roughly the same position.

The IRMPD spectra of $[M_5(GlyGly\text{-}H)_8]^{2+}$ ($M=Mg^{2+}$, Ca^{2+} and Sr^{2+}) are compared with the lowest energy IR spectra of Ca^{2+} cluster in N-H/O-H stretching region (2700-3800 cm^{-1}) in the supplementary data, Figure (S4.36). Similar to what was observed for Ca^{2+} cluster, there is a broad feature centered at 3350 and 3380 cm^{-1} for the Mg^{2+} and Sr^{2+}

complexes, respectively. In accordance with $[M_3(GlyGly\text{-}H)_4]^{2+}$ IRMPD spectra, a blue shift was observed for amide N-H band (3350, 3370 and 3380 cm^{-1}) related to the decrease in the positive charge density on the GlyGly ligands.

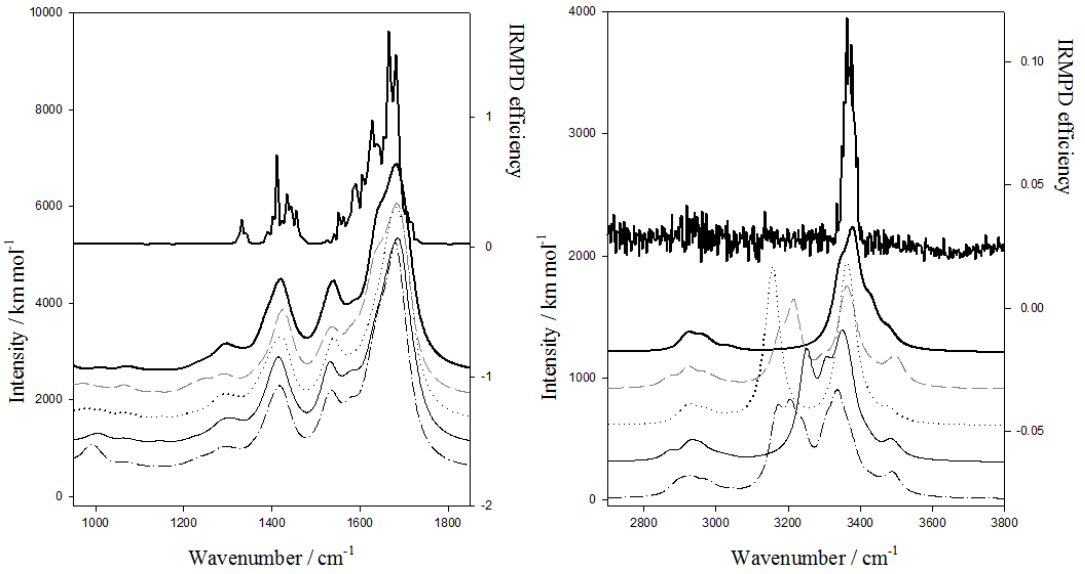


Figure 4.12. Comparison of the IRMPD spectra for $[Ca_5(GlyGly\text{-}H)_8]^{2+}$ and the computed IR spectra for five lowest energy structures in the fingerprint (a) and N-H/O-H stretch (b) regions. IR frequencies are computed using B3LYP/SVP method and basis set and scaled to 0.966.

4.4. Conclusion

Unimolecular dissociation reactions of singly and doubly charged clusters of Mg^{2+}, Ca^{2+} and Sr^{2+} with deprotonated GlyGly were studied. It was shown that singly charged clusters main dissociation pathway is loss of GlyGly as well as loss of 57 Da. In addition it was observed that other fragment ions related to loss of small neutral molecules such as water and CO_2 were more dominant for smaller clusters. Comparison

of Mg^{2+}, Ca^{2+} and Sr^{2+} cation clusters also showed that loss of GlyGly is more dominant for larger metal cations if it is not the only dissociation pathway whereas in Mg^{2+} loss of small neutral molecules occurs as a competing fragmentation pathway. In Ca^{2+} and Sr^{2+} doubly charged clusters, besides loss of GlyGly, another competing fragmentation pathway was charge separation reaction leading to formation of two singly charged clusters which becomes the main dissociation pathway for larger doubly charged clusters. In contrast for Mg^{2+} doubly charged clusters loss of GlyGly was the only dissociation pathway observed.

Energy-resolved CID results showed the order of dissociation energies to be $Ca^{2+}>Mg^{2+}>Sr^{2+}$ for singly charged clusters and $Sr^{2+}>Ca^{2+}>Mg^{2+}$ for doubly charged clusters. It was proposed that this is because of the higher order of metal-ligand interactions for Ca^{2+} in singly charged clusters. In doubly charged clusters, however, the charge density of the metal cation on the cluster could be the reason for the stability of these clusters.

To further investigate the nature of 57 Da as one of the main neutral fragmentation products, $[Ca_2(GlyGly\text{-}H)_3]^+$ cluster was isotopically labeled. Loss of 59 Da from $(GlyGly\text{-}H)^-$ ligand which was triply deuterated at amine and amide groups upon dissociation was in accordance with formation of aminoketene through hydrogen transfer from CH_2 rather than deuterated amine group to the amide nitrogen. Computationally it was also shown that aminoketene has less binding energy to the GlyGly compared to the alternative cyclic dissociation product molecule resulted from hydrogen being transferred from amine group. In addition the lowest energy dissociation pathway for loss of GlyGly was studied and it was shown that loss of different GlyGly ligands are not energetically

identical and more interestingly the type of GlyGly ligand of which the hydrogen is transferred to the leaving GlyGly was also important.

Computational methods were also employed to investigate the lowest energy structures of $[M_3(GlyGly\text{-}H)_4]^{2+}$ (M= Mg^{2+}, Ca^{2+} and Sr^{2+}) as well as $[Ca_5(GlyGly\text{-}H)_8]^{2+}$. In both cases it was shown that in the lowest energy structure metal cations are coordinated with all carboxylate and amide carbonyl oxygens from multiple GlyGly ligands and all amine nitrogens have formed intramolecular hydrogen bonding with amide hydrogens. Comparison of IRMPD spectrum with calculated IR spectra of some lowest energy structures was done on both clusters. In the case of $[M_3(GlyGly\text{-}H)_4]^{2+}$ clusters a feature was observed around 3350, 3330 and 3350 cm^{-1} for Mg^{2+}, Ca^{2+} and Sr^{2+}, respectively and was related to red-shifted amide N-H stretch. Another interesting point was the symmetric amine N-H stretch which was observed around 3390 cm^{-1} in Mg^{2+} cluster whereas it shows to be overlapped with (unresolved from) the red-shifted amide N-H stretching mode in Ca^{2+} and to more extent in Sr^{2+}, as a result of the blue-shift of amide N-H stretch to higher energies for these larger metal ions. This was explained by the less positive charged being localized by larger metal cations on the cluster and as a result on the amide group.

IRMPD spectra of $[Ca_5(GlyGly\text{-}H)_8]^{2+}$ in both fingerprint and N-H/O-H stretching region were reported. It was shown that all lowest energy structures calculated IR spectra reproduced the features at the fingerprint stretching region. However the results at the N-H/O-H region was mostly in accordance the lowest energy structure spectrum in which a single broad feature around 3370 cm^{-1} was related to the red-shifted amide as well as the symmetric and asymmetric amine N-H stretches.

References

(1) Lehn, J. M. From supramolecular chemistry towards constitutional dynamic chemistry and adaptive chemistry. *Chem. Soc. Rev.* **2007**, *36*, 151-160.

(2) Boncheva, M.; Whitesides, G. M. Making things by self-assembly. *Mater. Res. Bull.* **2005**, *30*, 736-742.

(3) Steed, J., Atwood, J.L.: Supramolecular Chemistry, Wiley, New York, 2000.

(4) You, C. C.; Dobrawa, R.; Saha-Möller, C. R.; Würthner, F.: Metallosupramolecular dye assemblies. *Top. Curr. Chem.* **2005**, *258*, 39-82.

(5) Ruben, M.; Rojo, J.; Romero-Salguero, F. J.; Uppadine, L. H.; Lehn, J. M. Grid-type metal ion architectures: Functional metallosupramolecular arrays. *Angew. Chem., Int. Ed.* **2004**, *43*, 3644-3662.

(6) Chakrabarty, R.; Mukherjee, P. S.; Stang, P. J. Supramolecular coordination: Self-assembly of finite two- and three-dimensional ensembles. *Chem. Rev.* **2011**, *111*, 6810-6918.

(7) Anwar, M. U.; Dawe, L. N.; Alam, M. S.; Thompson, L. K. μ-O bridged Mn_{10} assemblies with open O_6 sites for binding extra guests: Structural, magnetic, and surface studies. *Inorg. Chem.* **2012**, *51*, 11241-11250.

(8) Wei, Y.; Hou, H.; Fan, Y.; Zhu, Y. Transition metal ion directed self-assembly of polynuclear coordination complexes: Structural characterization and magnetic properties. *Eur. J. Inorg. Chem.* **2004**, *19*, 3946-3957.

(9) Johnson, D. W.; Raymond, K. N. The Role of Guest Molecules in the Self-assembly of Metal-ligand Clusters. *Supramol. Chem.* **2001**, *13*, 639-659.

(10) Davis, A. V.; Yeh, R. M.; Raymond, K. N. Supramolecular assembly dynamics. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 4793-4796.

(11) Dawe, L. N.; Shuvaev, K. V.; Thompson, L. K. Polytopic ligand directed self-assembly - Polymetallic [n×n] grids versus non-grid oligomers. *Chem. Soc. Rev.* **2009**, *38*, 2334-2359.

(12) Carnes, M. E.; Collins, M. S.; Johnson, D. W. Transmetalation of self-assembled, supramolecular complexes. *Chem. Soc. Rev.* **2014**, *43*, 1825-1834.

(13) Crichton, R. Biological Inorganic Chemistry: An Introduction., Elsevier, Boston, **2012**.

(14) Sigel, H. Calcium Role in Biology. Metal Ions in Biological Systems Vol. 17, CRC Press, Boca Raton, FL, **1984**.

(15) Hughes, M. N. The Inorganic Chemistry of Biological Processes. Wiley, New York, **1981**.

(16) Lippard, S. J. Principles of Bioinorganic Chemistry, University Science Books, Sausalito, **1994**.

(17) Permyakov, E. A. Metalloproteomics, Wiley, Hoboken, **2009**.

(18) Crivici, A.; Ikura, M. Molecular and structural basis of target recognition by calmodulin. *Annu. Rev. Biophys. Biomol. Struct.* **1995**, *24*, 85-116.

(19) Gomes, C. M. Protein Folding and Metal Ions:Mechanisms, Biology and disease, CRC Press, Boca Raton, FL, **2011**.

(20) Travaglia, A.; Pietropaolo, A.; La Mendola, D.; Nicoletti, V. G.; Rizzarelli, E. The inorganic perspectives of neurotrophins and Alzheimer's disease. *J. Inorg. Biochem.* **2012**, *111*, 130-137.

(21) Rufo, C. M.; Moroz, Y.S.; Moroz, O.V.; Stohr, J.; Smith, T.A.; Hu, X.; DeGrado, W.F.; Korendovych, I.V. Short peptides self-assemble to produce catalytic amyloids. *Nat. Chem.* **2014**, *6*, 303-309.

(22) Balaj, O. P.; Semrouni, D.; Steinmetz, V.; Nicol, E.; Clavaguera, C.; Ohanessian, G. Structure of sodiated polyglycines. *Chem.--Eur. J.* **2012**, *18*, 4583-4592.

(23) Dunbar, R. C.; Steill, J. D.; Polfer, N. C.; Oomens, J. Metal cation binding to gas-phase pentaalanine: Divalent ions restructure the complex. *J. Phys. Chem. A* **2013**, *117*, 1094-1101.

(24) Gholami, A.; Fridgen, T. D. The unimolecular chemistry of [Zn(amino acid)$_2$-H]$^+$ in the gas phase: H_2 elimination when the amino acid is a secondary amine. *Phys. Chem. Chem. Phys.* **2014**, *16*, 3134-3143.

(25) Rogalewicz, F.; Hoppilliard, Y.; Ohanessian, G. Structures and fragmentations of zinc(II) complexes of amino acids in the gas phase. I. Electrosprayed ions which are structurally different from their liquid phase precursors. *Int. J. Mass Spectrom.* **2000**, *201*, 307-320.

(26) Shoeib, T.; Zhao, J.; Aribi, H. E.; Hopkinson, A. C.; Siu, K. W. M. Dissociations of complexes between monovalent metal ions and aromatic amino acid or histidine. *J. Am. Soc. Mass Spectrom.* **2013**, *24*, 38-48.

(27) Shoeib, T.; Hopkinson, A. C.; Siu, K. W. M. Collision-induced dissociation of the Ag^+-proline complex: Fragmentation pathways and reaction mechanisms - A synergy between experiment and theory. *J. Phys. Chem. B* **2001**, *105*, 12399-12409.

(28) Hoppilliard, Y.; Rogalewicz, F.; Ohanessian, G. Structures and fragmentations of zinc(II) complexes of amino acids in the gas phase. II. Decompositions of glycine-Zn(II) complexes. *Int. J. Mass Spectrom.* **2001**, *204*, 267-280.

(29) Armentrout, P. B.; Citir, M.; Chen, Y.; Rodgers, M. T. Thermochemistry of alkali metal cation interactions with histidine: Influence of the side chain. *J. Phys. Chem. A* **2012**, *116*, 11823-11832.

(30) Armentrout, P. B.; Yang, B.; Rodgers, M.T. Metal cation dependence of interactions with amino acids: Bond energies of Rb^+ and Cs^+ to Met, Phe, Tyr, and Trp. *J. Phys. Chem. B* **2013**, *117*, 3771-3781.

(31) Watson, H. M.; Vincent, J. B.; Cassady, C. J. Effects of transition metal ion coordination on the collision-induced dissociation of polyalanines. *J. Mass Spectrom.* **2011**, *46*, 1099-1107.

(32) Chu, I. K.; Shoeib, T.; Guo, X.; Rodriquez, C. F.; Hopkinson, A. C.; Siu, K. W. M.; Lau, T. C. Characterization of the product ions from the collision-induced dissociation of argentinated peptides. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 163-175.

(33) Rozman, M.; Gaskell, S. J. Non-covalent interactions of alkali metal cations with singly charged tryptic peptides. *J. Mass Spectrom.* **2010**, *45*, 1409-1415.

(34) Ali, O. Y., Randell, N. M., Fridgen, T. D. Primary fragmentation pathways of gas phase [M(Uracil-H)(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.

(35) Ali, O. Y.; Fridgen, T. D. Structures of electrosprayed Pb(Uracil-H)$^+$ complexes by infrared multiple photon dissociation spectroscopy. *Int. J. Mass Spectrom.* **2011**, *308*, 167-174.

(36) Banu, L.; Blagojevic, V.; Bohme, D. K. Dissociation of deprotonated glycine complexes with Pb^{2+} and five transition-metal dications (Fe^{2+}, Co^{2+}, Ni^{2+}, Cu^{2+}, Zn^{2+}): The importance of metal bond activation. *Int. J. Mass Spectrom.* **2012**, *330-332*, 168-173.

(37) Banu, L.; Blagojevic, V.; Bohme, D. K. Bond activation in complexes of Pb(II) with 15 deprotonated amino acids: Correlation with gas-phase acidity. *Int. J. Mass Spectrom.* **2012**, *316-318*, 23-30.

(38) Gaussian 09, R. D., M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox, Gaussian, Inc., Wallingford CT, 2009.

(39) Moghaddam, M. B.; Fridgen, T. D. IRMPD spectroscopic study of microsolvated [Na(GlyAla)]$^+$ and [Ca(GlyAla-H)]$^+$ and the blue shifting of the hydrogen-bonded amide stretch with each water addition. *J. Phys. Chem. B* **2013**, *117*, 6157-6164.

(40) Wilson, S. R.; Cui, W. Applications of simulated annealing to peptides. *Biopolymers.* **1990**, *29*, 225-235.

(41) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and testing of a general Amber force field. *J. Comput. Chem.* **2004**, *25*, 1157-1174.

(42) D.A. Case, T. A. D., T.E.C. III, C.L. Simmerling, J. Wang, R.E. Duke, R. Luo, K.M. Merz, D.A. Pearlman, M. Crowley, R.C. Walker, W. Zhang, B. Wang, S. Hayik, A. Roitberg, G. Seabra, K.F. Wong, F. Paesani, X. Wu, S. Brozell, V. Tsui, H. Gohlke, L. Yang, C. Tan, J. Mongan, V. Hornak, G. Cui, P. Beroza, D.H. Mathews, C. Schafmeister, W.S. Ross, P.A. Kollman, AMBER 9, University of California, San Francisco, CA. **2006**.

(43) Kenny, P. T.; Nomoto, K.; Orlando, R. Fragmentation studies of peptides: the formation of y ions. *Rapid Commun. Mass Spectrom.* **1992**, *6*, 95-97.

(44) McCormack, A. L.; Somogyi, Á.; Dongre, A. R.; Wysocki, V. H. Fragmentation of protonated peptides: Surface-induced dissociation in conjunction with a quantum mechanical approach. *Anal. Chem.* **1993**, *65*, 2859-2872.

(45) Balaj, O. P.; Kapota, C.; Lemaire, J.; Ohanessian, G. Vibrational signatures of sodiated oligopeptides (GG-Na^+, GGG-Na^+, AA-Na^+ and AAA-Na^+) in the gas phase. *Int. J. Mass Spectrom.* **2008**, *269*, 196-209.

Chapter 5 . Structures, Unimolecular Fragmentations, and Reactivities of Self-Assembled Multimetallic/Peptide Complexes: $[Mn_n(GlyGly\text{-}H)_{2n-1}]^+$ and $[Mn_{n+1}(GlyGly\text{-}H)_{2n}]^{2+}$

Abstract

Complexes of Mn^{2+} with deprotonated GlyGly are investigated by sustained off-resonance irradiation collision-induced dissociation (SORI-CID) and infrared multiple photon dissociation spectroscopy. Singly charged, $[Mn_n(GlyGly\text{-}H)_{2n-1}]^+$, and doubly charged, $[Mn_{n+1}(GlyGly\text{-}H_{2n}]^{2+}$, clusters are formed from aqueous solutions of $MnCl_2$ and GlyGly by electrospray ionization. The most intense ion produced was the singly charged $[M_2(GlyGly\text{-}H)_3]^+$ cluster. Singly charged clusters show extensive fragmentations of small neutral molecules such as water and carbon dioxide as well as dissociation pathways related to the loss of NH_2CHCO and GlyGly. For the doubly charged clusters, however, loss of GlyGly is observed as the main dissociation pathway. Structure elucidation of $[Mn_3(GlyGly\text{-}H)_4]^{2+}$ clusters has also been done by IRMPD spectroscopy as well as DFT calculations. It is shown that the lowest energy structure of $[Mn_3(GlyGly\text{-}H)_4]^{2+}$ cluster is deprotonated at all carboxylic acid groups and metal ions are coordinated with carbonyl oxygen atoms and all amine nitrogens are hydrogen bonded to the amide hydrogen. A comparison of calculated high spin (sextet) and low spin (quartet) state structures of $[Mn_3(GlyGly\text{-}H)_4]^{2+}$ are provided. Also, the gas phase reactivity of these complexes towards neutral CO and water was investigated.

Keywords: Transition metal, GlyGly, sustained off-resonance irradiation collision-induced dissociation (SORI-CID), infrared multi-photon dissociation (IRMPD), ion-molecule reaction.

5.1. Introduction

Metal ions are important for the function of many biological molecules. Alkaline earth metals Mg^{2+} and Ca^{2+} are the most abundant divalent metal ions in biological systems–they play an essential role in both catalytic reactions and structural motifs. For example Mg^{2+} exists in large amounts inside the human body and it participates in processes that need phosphoryl transportation such as muscle contraction. Ca^{2+} is responsible for a range of intra- and extracellular signaling which is carried out by Ca^{2+}-binding proteins that are known as messenger proteins. Mg^{2+} and Ca^{2+} are also the main component of biominerals such as bones in the human body.[1]

Information gained in the study of small peptide complexes with metal ions can be used in models of metal ion-biomolecule interactions in much larger systems. The investigation of metal cation coordination by such systems can reveal possible binding sites of metal cations to amino acids in peptide chains and as a result in proteins. Intrinsic properties of metal cation complexes of peptides have been widely studied by mass spectrometric techniques.[2-17] It has been shown that characteristics such as metal ion size and charge can affect the conformation of metal ion-peptide complexes. In a study by Carron and coworkers the structure and fragmentation of alkaliated tripeptides GlyGlyGly as well as PheGlyGly complexes were investigated by collision-induced dissociation (CID) mass spectrometry.[18] It was shown that Li^+ and Na^+ complexes of these peptides mainly dissociate to b and y type fragment ions whereas the main dissociation pathway for Rb^+ complexes is formation of a_2 fragment ions and loss of the metal ion. For the cesiated complex, the metal ion is the only fragment observed. Authors suggested that K^+

has diverse fragmentation pathways, which is a result of having intermediate properties between the larger and smaller alkali metal cations. They showed that all metal cations interact with carbonyl oxygen atoms and the bond lengths increase as the metal cation becomes larger. In a similar study, Shin and coworkers investigated complexes of these deprotonated peptides with divalent metal cations.[19] Comparing the fragmentation pathways of the alkaline earth metal series, they observed more dissociation channels for larger metal cation complexes. Complexes of $[Ca(GlyGlyGly\text{-}H)]^+$ mainly produce metallated a-type fragments as well as their dehydrated (a-H_2O) and deamoniated (a-NH_3) forms whereas Sr^{2+} and Ba^{2+} complexes form x-type fragment ions along with a-type fragment ions. The more weakly bound is the metal ion, the more freely the cation can migrate through the peptide chain which results in diverse dissociation reactions. In the same study it was shown that transition metals with smaller cationic size compared to alkaline earth metal ion are more strongly bound to the ligand and this was confirmed by less varied fragmentations observed for the transition metal ion complexes. The higher binding energy of transition metals was explained by the partly covalent nature of their interactions with the peptide. The diversity in the dissociation pattern of different metal cation complexes can especially be useful in sequencing peptides. For example Fe^{2+} favorably binds to cysteine residues of peptides.[20] The strong interaction of Fe^{2+} with the sulfur of cysteine residues in a peptide chain, results in specific fragmentation pathways that can be applied to find the cysteine position in the peptide sequence. Compared to protonated peptides containing Cys residue which fragment almost at all peptide residues, Fe^{2+} induce fragmentation around cysteine residue forming a_n* ions (* means metal bound fragment).

Transition metals are of great importance in many biological reactions. Zn^{2+} for example, participates in the activation of the oxytocin receptor binding site by forming a strongly bound octahedral complex in which the metal ion is coordinated with all carbonyl oxygens leaving the terminal amine nitrogen and side chains free for binding to the target protein.[21] Manganese, in various oxidation states, in photosystem II (PS II) is an interesting example of transition metal ions that exists as the reactive center in biological systems. Clusters of Mn_4CaO_5 (Mn(III), Mn(IV) and Mn(V)) in the oxygen evolving centre (OEC) of PS II participate in water oxidation.[22] Transition metals are also known for their contributions in bond activation reactions. Fragmentation studies of $[Zn(Proline)_2\text{-}H]^+$ complex showed that the main dissociation pathway leads to dehydrogenation of deprotonated proline in which hydrogen atoms are detached from amine and C_5 groups.[23]

In previous work (chapter 4) we showed that alkaline earth metal cations, Mg^{2+}, Ca^{2+} and Sr^{2+} form singly and doubly charged clusters with deprotonated GlyGly. It was shown that the main dissociation pathway for the singly charged clusters is loss of GlyGly competing with the loss of NH_2CHCO. In doubly charged clusters, the loss of GlyGly as well as charge separation are the main fragmentation channels. Larger doubly charged clusters, however, mainly dissociate to the singly charged clusters. The lowest energy structure of the doubly charged $[M_3(GlyGly\text{-}H)_4]^{2+}$ and $[M_5(GlyGly\text{-}H)_8]^{2+}$ clusters were assigned using computational chemistry and IRMPD spectroscopy methods. It was shown that in these clusters the metal ion is solvated with carbonyl oxygens. Also, there is hydrogen bonding between the terminal amine nitrogen and amide hydrogen in each GlyGly. These results were in agreement with the structure of $[Ca(GlyAla\text{-}H)]^+$ which

were studied in more detail by IRMPD spectroscopy and computational methods.[24] Here the unimolecular fragmentation reactions of $[Mn_n(GlyGly\text{-}H)_{2n-1}]^+$ and $[Mn_{n+1}(GlyGly\text{-}H_{2n}]^{2+}$ complexes are investigated. Attempts to characterize the structure of $[Mn_3(GlyGly\text{-}H)_4]^{2+}$ clusters by means of IRMPD spectroscopy were carried out. Ion-molecule reactions involving $[Mn_2(GlyGly)_3]^+$ were also investigated.

5.2. Methods

5.2.1. Experimental

All experiments were done on a Bruker ApexQe hybrid quadrupole-Fourier transform ion cyclotron resonance (FT-ICR) mass spectrometer. All chemicals were used without any purification. Solutions of 2 mM of the metal salt $MnCl_2{\cdot}4H_2O$ and GlyGly were prepared with 18 MΩ Millipore water. Metal cationized complexes of GlyGly were electrosprayed from solutions of 1 mL of 5 mM GlyGly and a few drops of the metal salt solution using an Apollo II electrospray ion source with flow rate 250 μl/h. For SORI-CID experiments Ar was pulsed to ~10^{-6} mbar inside the ICR cell. Collision energies up to 1.9×10^{-5} - 1.8 eV in the center of mass (com) frame were applied in SORI pulses 0.25 s long. Fragmentation pathways were studied with multistage SORI-CID mass spectrometry.

IRMPD experiments were performed in the C-H/N-H/O-H stretching region. 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 a Fourier-transform ion cyclotron resonance mass spectrometer (FT-ICR) coupled to an IR

OPO, manufactured by LaserSpec. The OPO is tuneable from 1.4 to 4.5 μm, with bandwidth of 2 cm^{-1}. The OPO consists of a periodically poled lithium niobate crystal that is pumped by a diode pumped 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 for the present experiments.

Also, a CO_2 laser was applied to activate Mn^{2+} clusters in order to obtain extensive fragmentation. This was performed to expose metal centers for ion-molecule association reaction by removing ligands that have covered Mn^{2+}. IRMPD activation was performed using a continuous 25 W CO_2 laser to activate ions trapped inside the ICR cell. Ions were irradiated with the CO_2 laser operated at 85% of maximum power for duration of 1.5 s to expose the Mn ions in the complex.

5.2.2. Computational

All calculations were performed using Gaussian 09 suite of programs.[25]. Initial geometries were taken from the lowest energy structure of $[Ca_3(GlyGly\text{-}H)_4]^{2+}$ which was fully discussed in a previous work. Mn is a transition metal and can have different electron configurations and as a result different spin multiplicities. For Mn^{2+}, with there are three possible spin multiplicities, a doublet (S=1/2), a quartet (S=3/2), and a sextet (S=5/2). For a complex containing three Mn^{2+}, and assuming all have the same spin, the total multiplicities will be 4, 10, and 16 for each Mn^{2+} in the doublet, quartet, and sextet, respectively. From now on we use doublet, quartet, and sextet terms to describe the spin multiplicities of the $[Mn_3(GlyGly\text{-}H)_4]^{2+}$ complex. Geometry optimization and frequency calculations were done using the B3LYP method. The all electron 6-31+G(d,p) basis set

was used for C, H, N and O and the Def2SVP basis set and effective core potential was used on the metals. All reported thermochemical data are 298 K values in kJ mol^{-1} and calculated IR frequencies were scaled to 0.966.

5.3. Result and Discussion

The full electrospray mass spectrum of a solution containing Mn^{2+} and GlyGly is shown in Figure 5.1. The mass to charge ratio as well as the isotopic pattern of these clusters were used to confirm cluster identification. The most intense peak for Mn^{2+} was observed at *m/z* 503 corresponding to the singly charged $[M_2(GlyGly\text{-}H)_3]^+$ cluster. The monomeric species, $[Mn(GlyGly\text{-}H)]^+$, could not be detected, presumably because of the high concentrations of GlyGly and metal salt solutions. The smallest complex observed for the Mn^{2+} was $[Mn(GlyGly)_2\text{-}H)]^+$ at *m/z* 318. Other singly charged clusters were also observed at *m/z* 820 and 1137 related to $[Mn_3(GlyGly\text{-}H)_5]^+$ and $[Mn_4(GlyGly\text{-}H)_7]^+$. The smallest doubly charged cluster, $[Mn_3(GlyGly\text{-}H)_4]^{2+}$, was detected at *m/z* 344 and two other doubly charged clusters were observed at *m/z* 661 and 978 corresponding to $[Mn_5(GlyGly\text{-}H)_8]^{2+}$ and $[Mn_7(GlyGly\text{-}H)_{12}]^{2+}$, respectively.

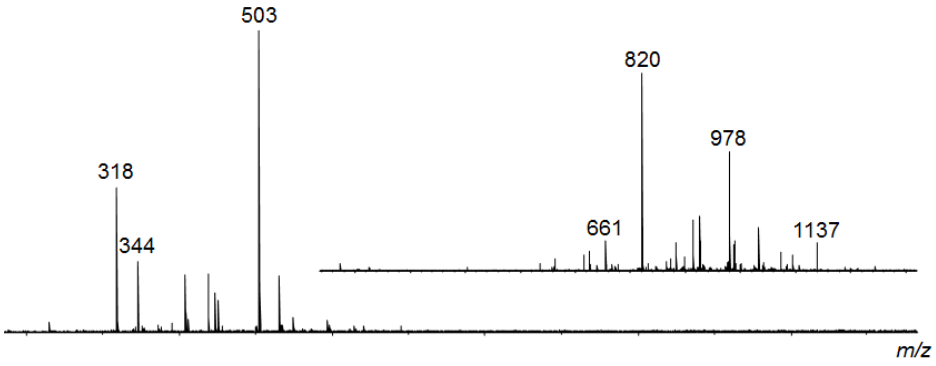


Figure 5.1. Full mass spectrum of positive-ion electrosprayed aqueous solutions of $MnCl_2$ with GlyGly is shown. $[M(GlyGly\text{-}H)GlyGly]^+$, $[M_3(GlyGly\text{-}H)_4]^{2+}$, $[M_2(GlyGly\text{-}H)_3]^+$, $[M_5(GlyGly\text{-}H)_8]^{2+}$, $[M_3(GlyGly\text{-}H)_5]^+$, $[M_7(GlyGly\text{-}H)_{12}]^{2+}$ and $[M_4(GlyGly\text{-}H)_7]^+$ clusters were observed at *m/z* 318, 344, 503, 661, 820, 978 and 1137.

5.3.1. SORI-CID fragmentation

5.3.1.1. $[Mn_n(GlyGly\text{-}H)_{2n-1}]^+$ (n=2, 3 and 4)

5.3.1.1.1. $[Mn(GlyGly\text{-}H)(GlyGly)]^+$

In Figure 5.2 collision-induced dissociation mass spectra of singly charged clusters of Mn^{2+} are compared. $[Mn(GlyGly\text{-}H)(GlyGly)]^+$ dissociates to *m/z* 274 and 230 corresponding to consecutive losses of carbon dioxide. MS/MS on *m/z* 230 resulted in *m/z* 212, loss of water; *m/z* 201, loss of 29 Da (previously seen $NHCH_2$); *m/z* 199, loss of 31 Da (NH_2CH_3) and *m/z* 144, loss NH_2CHCO following $NHCH_2$ loss (Figure S5.1 in supplementary data). Also, the fragment ion at *m/z* 201 dissociated to *m/z* 144 losing NH_2CHCO (Figure S5.2) confirming this loss following $NHCH_2$ loss.

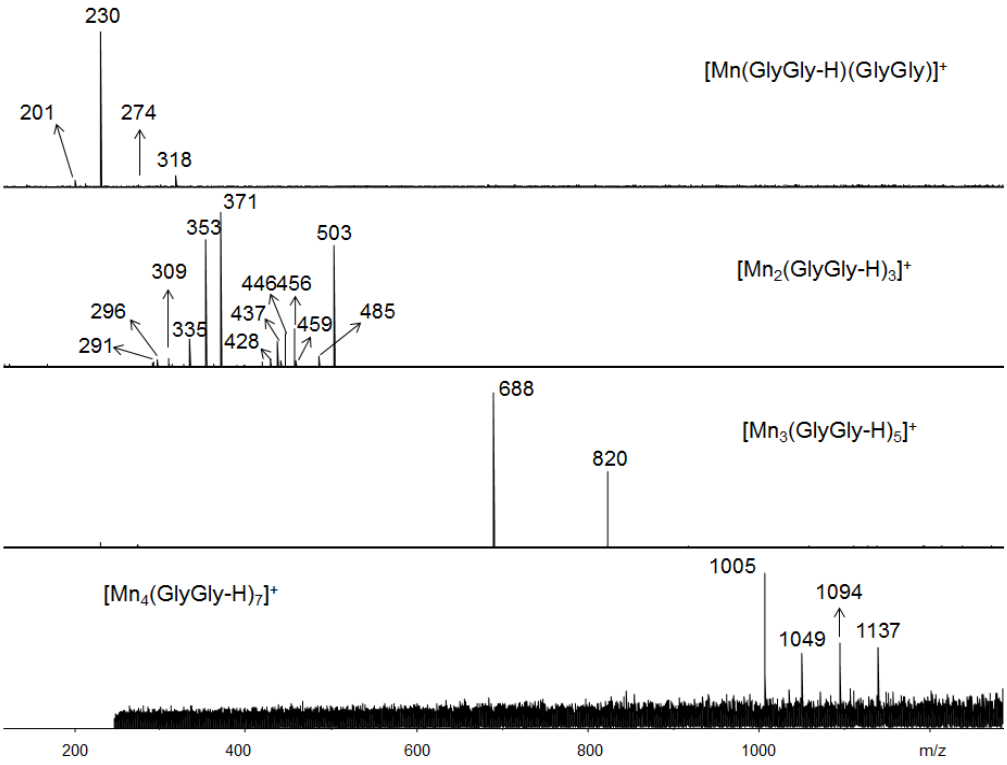


Figure 5.2. SORI/CID mass spectra of $[Mn_n(GlyGly\text{-}H)_{2n-1}]^+$ (n=2, 3 and 4).

5.3.1.1.2. $[Mn_2(GlyGly\text{-}H)_3]^+$

The dissociation products observed for the $[Mn_2(GlyGly\text{-}H)_3]^+$ cluster (*m/z* 503) were *m/z* 485, loss of water; 459, loss of CO_2; 456, loss of water + $NHCH_2$; and 446, loss of NH_2CHCO. The ion at *m/z* 437 is most likely the loss of GlyGly from the doubly charged cluster $[Mn_4(GlyGly\text{-}H)_6]^{2+}$ which is isobaric with the ion of interest, but based on the isotopic distribution is a minor component of *m/z* 503. The fragment ions at *m/z* 428, loss of glycine; 371, loss of GlyGly; 353, loss of GlyGly + water and 335, loss of GlyGly + two water molecules were also observed during MS/MS on *m/z* 446 (Figure S5.3). The fragment ions at *m/z* 353 and 335 are also produced during MS/MS on *m/z* 371 upon consecutive losses of two water molecules (Figure S5.4).

Other fragment ions upon dissociation of *m/z* 446 were observed at *m/z* 399, loss of water + $NHCH_2$; 389, loss of NH_2CHCO; 384, water + CO_2; 314, loss of GlyGly. The ion observed at *m/z* 296 could be related to loss of GlyGly + H_2O was also produced during MS/MS on *m/z* 353 (formed by loss of GlyGly+H_2O from *m/z* 503) upon loss of NH_2CHCO so we suggested *m/z* 296 has more than one origin. Another fragment ion observed upon dissociation of *m/z* 353 was *m/z* 309 corresponding to loss of CO_2 (Figure S5.5). In addition, the fragment ion, *m/z* 278 produced during MS/MS on *m/z* 353 was also observed upon dissociation of one water molecule from *m/z* 296 (Figure S5.6). MS/MS on fragment ion *m/z* 335 resulted in fragment ions at *m/z* 317, loss of water; 306, loss of $NHCH_2$; 263, loss of water + $NHCH_2$; 239, loss of ($2NH_2CHCO$-H_2O) and 221 loss of two NH_2CHCO molecules (Figure S5.7). Also the fragment ion at *m/z* 195 was assumed to come from *m/z* 239 upon loss of CO_2. The fragment ion at *m/z* 248 was related to loss of GlyGly-CO_2. However it could also stem from *m/z* 291 upon loss of a second CO_2. We were not able to follow further dissociation of *m/z* 291 through MS^3 to confirm consecutive losses of CO_2 from *m/z* 335 that account for the presence of the fragment ion at *m/z* 248.

5.3.1.1.3. $[Mn_3(GlyGly\text{-}H)_5]^+$

The sole fragmentation observed for $[Mn_3(GlyGly\text{-}H)_5]^+$ was the loss of a GlyGly at *m/z* 688. MS/MS on the *m/z* 688 resulted in loss of second GlyGly molecule at *m/z* 556 (Figure S5.8) as well as loss of water at *m/z* 670. At higher energies where fragment ion *m/z* 688 is completely dissociated, new fragments appeared at *m/z* 538, loss of water; 520,

loss of two water molecules; 496, loss of acetic acid and 478 related to loss of acetic acid + water from the *m/z* 556 fragment ion (Figure S5.9).

5.3.1.1.4. $[Mn_4(GlyGly\text{-}H)_7]^+$

$[Mn_4(GlyGly\text{-}H)_7]^+$ clusters mainly lose GlyGly to form the ion at *m/z* 1005. Other fragment ions observed at *m/z* 1093 and 1049 were related to loss of CO_2 and GlyGly-CO_2, respectively. Any attempt for MS/MS on the fragment ions was unsuccessful in this case due to very small ion intensities.

5.3.1.2. $[Mn_{n+1}(GlyGly\text{-}H_{2n}]^{2+}$ (n=2, 4 and 6)

Doubly charged clusters of Mn^{2+} show similar fragmentation patterns to the Mg^{2+} clusters observed in previous works of our group.[26] In contrast to the Ca^{2+} and Sr^{2+} complexes no charge separation processes were observed in both Mg^{2+} and Mn^{2+} cases. This observation can be explained by the similar cationic size of these metals. Compared to Ca^{2+} and Sr^{2+} with ionic radii of 1.14 and 1.32 Å (for coordination number of six), both high (0.97 Å) and low spin (0.81 Å) Mn^{2+} cations show similar metal ion size to the Mg^{2+} (0.86 Å).[27] This suggests that these complexes are electrostatic in nature and that, the smaller cations form stronger interactions and the positive charge of the metal cations is more delocalized or solvated compared to complexes involving larger cations, Ca^{2+} and Sr^{2+}, where the charge is more delocalized on the metal cation.

The main dissociation of $[Mn_3(GlyGly\text{-}H)_4]^{2+}$ was loss of GlyGly at *m/z* 278.5, and smaller fragmentations corresponding to losses of NH_2CHCO at *m/z* 316.0 and water at *m/z* 335.5, Figure 5.3. Other less pronounced features were observed at *m/z* 269.5, 256.5

and 250.0 corresponding to loss of GlyGly + water, GlyGly + CO_2 and GlyGly+NH_2CHCO molecules, respectively. MS/MS on the *m/z* 278.5 fragment resulted in *m/z* 269.5, loss of water; 256.5, loss of CO_2; 250.0, loss of NH_2CHCO; 247.5, loss of water + CO_2; 241.0, loss of glycine; 238.5, loss of two water molecules + CO_2; 232.0, loss of glycine + water; 228.0, loss of NH_2CHCO+CO_2; 223.0, loss of glycine + two water molecules and 212.5 related to loss of GlyGly respectively (Figure S5.10). Additionally, the fragment ion at *m/z* 316.0 dissociated to *m/z* 278.5 related to loss of glycine upon MS/MS (Figure S5.11). This means that the *m/z* 278.5 peak may be due to loss of GlyGly and/or loss of glycine following the loss of NH_2CHCO. Further dissociation of fragment ions with lower intensities was unsuccessful.

For both the larger clusters, $[Mn_5(GlyGly\text{-}H)_8]^{2+}$ and $[Mn_7(GlyGly\text{-}H)_{12}]^{2+}$ at *m/z* 661.5 and 978.5 respectively, the only fragmentations observed were single and multiple losses of GlyGly. MS/MS on *m/z* 595.5 resulted in loss of one GlyGly at *m/z* 529.5 (Figure S5.12) as well as a loss of water at *m/z* 520.5. Any attempt on dissociation of *m/z* 529.5 was unsuccessful because of the low intensities of this ion. In addition MS/MS on both *m/z* 846.5 and 780.5 resulted in *m/z* 780.5 and 714.5, again consecutive losses of GlyGly (Figure S5.13, S5.14).

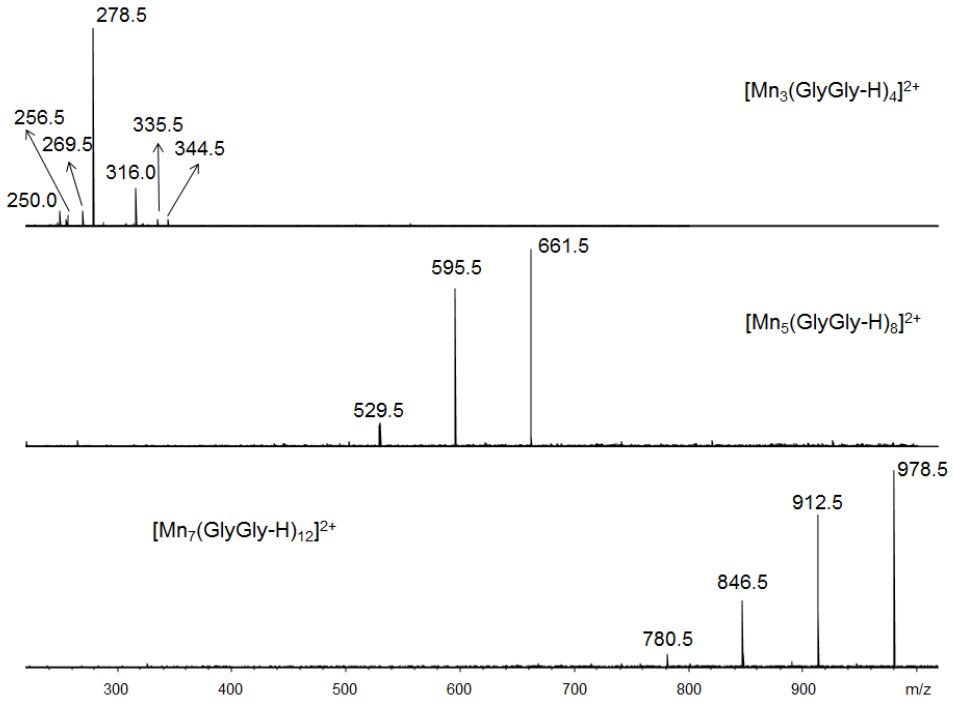


Figure 5.3. SORI/CID mass spectra of $[Mn_{n+1}(GlyGly\text{-}H)_{2n}]^{2+}$ (n=2, 4 and 6).

5.3.2. IRMPD spectroscopic study

The IRMPD spectra of $[M_3(GlyGly\text{-}H)_4]^{2+}$ ($M=Mg^{2+}$, Ca^{2+}, Sr^{2+} and Mn^{2+}) complexes are compared in Figure 5.4(a). Previously we determined the lowest energy structure of this cluster for Mg^{2+}, Ca^{2+} and Sr^{2+} by comparing the experimental spectra with the calculated IR spectra of some low energy structures.[26] The experimental spectra were most consistent with the lowest energy structures in which metal ions were coordinated with all carbonyl oxygens and amine nitrogens were hydrogen bonded to amide hydrogens. The IRMPD spectrum of Mg^{2+} complex shows features at 3300, 3390 and 3460 cm^{-1}. The features at 3300 and 3390 cm^{-1} are indicative of red-shifted amide and symmetric amine N-H stretches. Also the feature at 3460 cm^{-1} is related to asymmetric amine N-H stretch. Red-shifted amide N-H and symmetric amine N-H stretches were

observed around 3330 cm^{-1} for Ca^{2+} and around 3350 cm^{-1} for Sr^{2+} complex and asymmetric amine N-H stretch appeared at 3460 cm^{-1} for both metal cations. The IRMPD spectrum obtained for the Mn^{2+} complex showed similar features, in fact a similar spectrum as the Mg^{2+} complex; red shifted amide N-H stretch and symmetric amine N-H stretch around 3300 and 3390 cm^{-1} and asymmetric amine N-H stretch around 3470 cm^{-1}. This interesting resemblance is in accordance with the similarity of the radii of Mg^{2+} and Mn^{2+} which leads to a similar electrostatic interaction between metal and ligand in both cases. As a result the positive charge localized on the carbonyl oxygens and on the whole ligand is similar. In this regard, the position and shape of the features, especially the hydrogen bonded N-H stretching; in the IRMPD spectra of Mg^{2+} and Mn^{2+} reflect these similarities. Besides these features, in the Mn^{2+} complex, there is also a resolved band around 3345 cm^{-1} which is observed as a shoulder to the blue side of the broad feature around 3300 cm^{-1} in the Mg^{2+} spectrum. We propose that it is also related to the amide N-H stretches. Based on the results from comparing the IRMPD spectra of alkaline earth metal cations and Mn^{2+} complex, it is suggested that the lowest energy structure for $[M_3(GlyGly\text{-}H)_4]^{2+}$ have the same features as the optimized lowest energy structures for alkaline earth metal cation complexes.

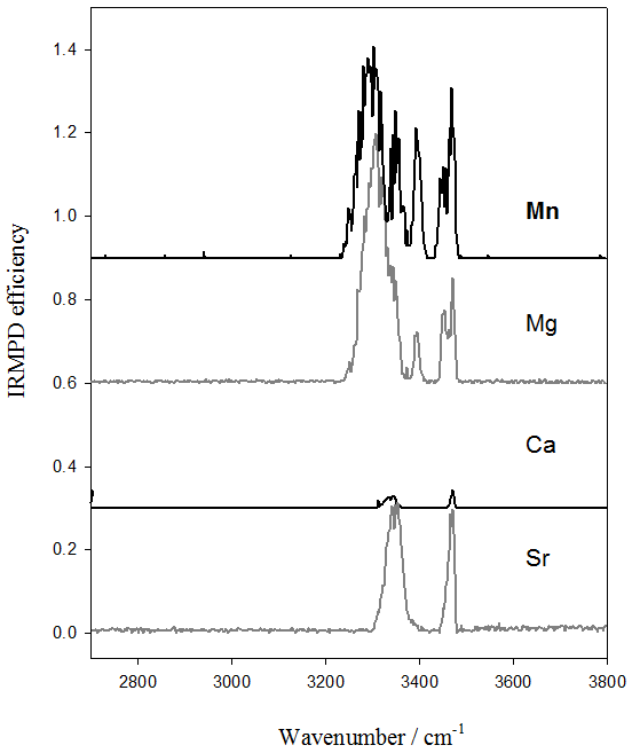


Figure 5.4. Comparison of IRMPD spectra of $[M_3(GlyGly\text{-}H)_4]^{2+}$ for Mg^{2+}, Ca^{2+}, Sr^{2+} and Mn^{2+}.

5.3.3. Computational study of $[Mn_3(GlyGly\text{-}H)_4]^{2+}$ structure

Figure 5.5 shows the DFT optimized structures of $[Mn_3(GlyGly\text{-}H)_4]^{2+}$ cluster for two possible spin multiplicity states of Mn; quartet (S=3/2) and sextet (S=5/2) which are termed high and low spin states, respectively. The attempt to optimize the doublet structure was not successful because of the much higher energy (714.56 kJ mol^{-1} higher electronic energy) that it showed compared to quartet and sextet spin multiplicity state structures. Both structures shared similar features with all alkaline earth metal structures; metal ions solvated with carbonyl oxygens at the carboxylate and amide carbonyl groups,

hydrogen bonding between amine nitrogens and amide hydrogens. However, there were slight differences between these two spin multiplicity states structures considering the hydrogen bonding as well as the metal-oxygen bond lengths. For example, as is shown in Figure 5.5 the hydrogen bond length for two hydrogen bonding were 2.04 and 2.06 Å for the quartet and 2.05 and 2.09 Å for the sextet spin state structures.

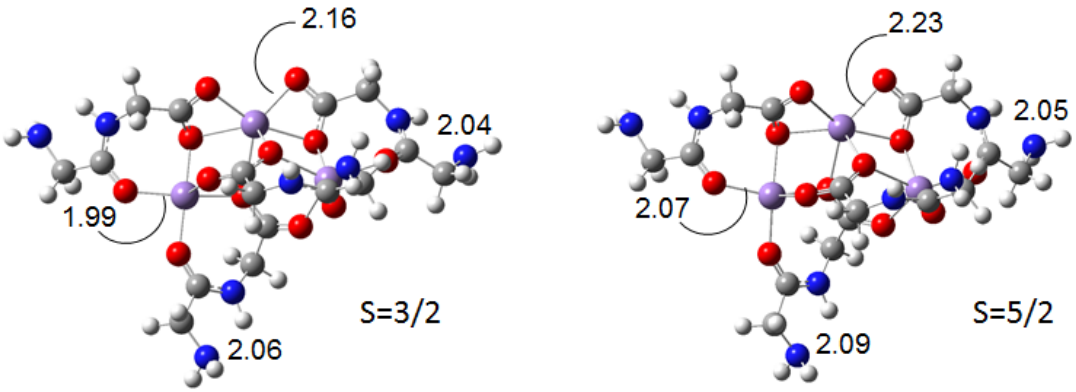


Figure 5.5. DFT optimized lowest energy structures of $[Mn_3(GlyGly\text{-}H)_4]^{2+}$ for high spin (S=5/2) and low spin (S=3/2) spin states using B3LYP/6-31+G(d,p) on C, H, N and O and Def2SVP on Mn^{2+}.

In Figure 5.6 Mn and oxygen atoms are labeled as Mn(1-3) and O(1-12). It is shown that Mn1 is coordinated with three pairs of carboxylate oxygen atoms which two pairs are of the same ligand and one is from two different GlyGly molecules. In contrast Mn2 and Mn3 are interacting with five oxygen atoms, two pairs are from carboxylate and amide carbonyl groups of different GlyGly ligands (O2, O3, O11 and O12 for Mn2 and O5, O6, O9, O10 for Mn3) and one is a carboxylate oxygen of the third GlyGly molecule (O7 for Mn2 and O8 for Mn3). The Mn-O bond lengths are provided in supplementary data, Table S5.1.

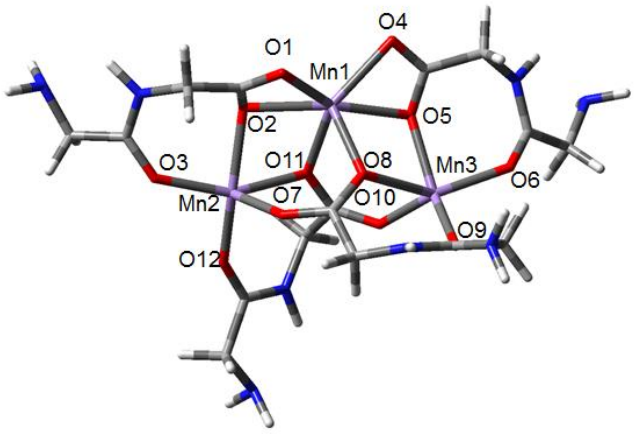


Figure 5.6. Representation of Mn-Oxygen bonds for a DFT optimized structure of $[Mn_3(GlyGly\text{-}H)_4]^{2+}$ using B3LYP/6-31+G(d,p) on C, H, N and O and Def2SVP on Mn^{2+}.

In a study by Khodabandeh and coworkers[28] on Mn^{2+} complexes of glycine it was shown that in each of the high spin (S=5/2), intermediate spin (S=3/2), and low spin (S=1/2) multiplicity states, the lowest energy structures are zwitterionic forms in which Mn^{2+} is interacting with carboxylate oxygen and amine hydrogens are hydrogen bonded to carbonyl oxygen. As well, the lowest energy structure among all the different spin states is a high spin state structure. This is in accordance with the results obtained here. In comparison of the calculated thermochemistry data, a large energy difference was observed, between two spin multiplicity states structures, showing that the high or sextet spin state structure is more favorable by 494.7 and 500.9 kJ mol^{-1} in enthalpy and Gibbs energy, respectively.

In Figure 5.7 calculated IR spectra of the structures (shown in Figure 5.5) are compared to the IRMPD result. It is shown that both quartet and sextet spin state structures reproduced the IRMPD spectrum. There are two distinct features around 3410

and 3490 cm^{-1} corresponding to the symmetric and asymmetric amine NH_2 stretches. Also two broad features around 3230-3370 cm^{-1} are related to the red-shifted amide N-H stretches and shows the amide N-H stretches appeared at slightly different wavenumbers for each of four GlyGly ligand. More interestingly, the slightly larger computed red-shift of amide N-H stretch in the quartet spin state cluster can be explained by the stronger Mn-O interaction through short range non-covalent bondings of metal cations to oxygens.

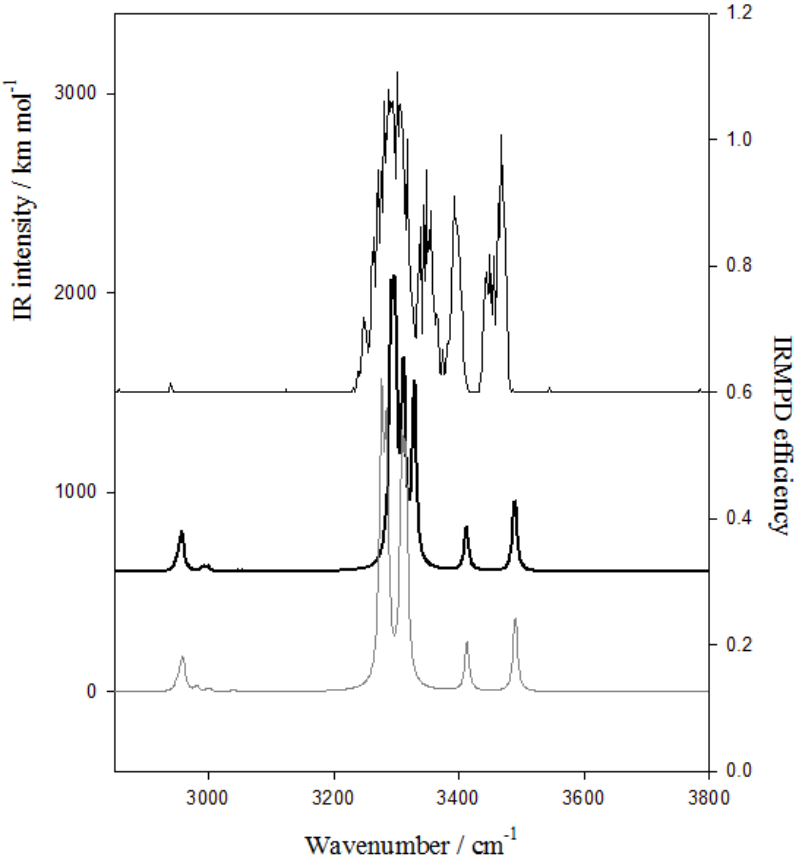


Figure 5.7. Comparison of IRMPD spectrum (top) with the calculated IR spectra for $[Mn_3(GlyGly\text{-}H)_4]^{2+}$ for sextet (middle) and quartet (bottom) spin multiplicity states obtained using B3LYP/6-31+G(d,p) on C, H, N and O and Def2SVP on Mn^{2+}.

5.3.4. Ion-molecule reactions of $[Mn_2(GlyGly\text{-}H)_3]^+$ and carbon monoxide—and water.

Ion-molecule reactions of Mn clusters with CO and water molecule were investigated. Many transition metal ions, in their bare and/or ligated form, are known to work as catalysts. For example a binuclear Mn cluster in the methionine aminopeptidase (metalloenzyme), in which Mn ions are linked by a hydroxyl group, binds to the target peptide at the N- and C-terminal and result in cleavage of N-terminal methionine residues.[29] In addition transition metals clusters are investigated as catalysts for bond activation of hydrocarbons and many important neutral molecules such as CO, O_2 and N_2O.[30] The importance of working with multimetallic clusters lies in the physical and chemical properties that can be achieved through formation of such systems. It has been shown that metal ions can form stable multinuclear clusters by linking to specific ligands. For example binuclear Ag clusters were shown to be stabilized through binding to thioglycolic acid molecules.[31] In this work we could successfully form singly and doubly charged mutimetallic clusters of Mn^{2+} with deprotonated GlyGly molecules as ligands in the gas phase. The main challenge was to investigate the reactivity of these systems, and more specifically the metal centers that are attached by carbonyl oxygens and buried inside by the ligands. In the previous studies in our research group it was shown that binding of transition metals to ligands such as carbon monoxide had an effect on the reactivity of their complexes. Gholami and coworkers showed that $[Ru(bipy)_2]^{2+}$ reaction with propane and 2-methylpropane only produce adduct ion-molecule $[Ru(bipy)_2(C_4H_{10})]^{2+}$complex. In contrast $[Ru(bipy)_2(CO)]^{2+}$ forms an activated complex

which leads to dehydrogenation and demethanation of adduct ion-molecule $[Ru(bipy)_2CO(C_4H_{10})]^{2+}$.[32] In these studies, it was concluded that the binding of CO to Ru, chemically activated the $[Ru(bipy)_2]^{2+}$ complex since the binding energy between the ion and the neutral could not be efficiently quenched in the low pressure confines of the FT-ICR cell.

Ion-molecule reactions were studied between $[Mn_2(GlyGly\text{-}H)_3]^+$ and carbon monoxide. The cluster ion was isolated and allowed to react with monoxide which was introduced to the ICR cell via pulsed valves. All attempts on addition of carbon monoxide to the parent ion were unsuccessful, Figure S5.15. It is likely that the metal cations are buried close to the center of the cluster where the three GlyGly-H ligands limit the interaction of metal cation with carbon monoxide.

To fragment cluster ions in an attempt to expose the metals, the cluster was isolated and irradiated with a CO_2 laser to extensively fragment ions which resulted in the mass spectra shown in Figure 5.8 (a-c). IRMPD was beneficial compared to CID since it was possible to fragment both parent and fragment ions at the same time by applying a powerful laser. However in other attempts we also tried to activate these clusters by multi stage CID method. Low ion intensities after multiple step CID experiments limited any further investigations of fragmentation products to this point. As was discussed earlier, the fragment ion at *m/z* 335 results from the loss of one GlyGly molecule followed by losses of two water molecules. The fragment ion *m/z* 335 was isolated and allowed to react with one 200 ms pulse of carbon monoxide. The mass spectrum showed a feature at *m/z* 353 which corresponds to the addition of one water molecule to the fragment ion and no carbon monoxide attachment was observed, Figure 5.8 (d,e). Water present in the ICR

cell was most likely coming along with pulses of CO from CO source which possibly had water as impurity.

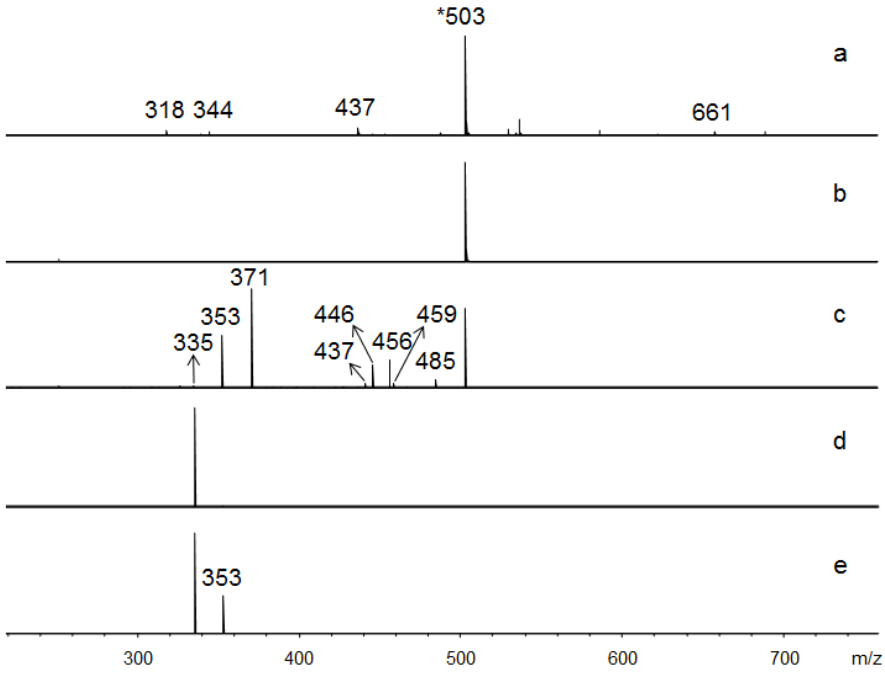


Figure 5.5. Full mass spectrum obtained from electrospraying $MnCl_2$ and GlyGly solution (a). Mass spectra obtained from isolation (b), CO_2 laser IRMPD at 85% laser power (c) of $[Mn_2(GlyGly\text{-}H)_3]^+$ cluster and isolation of the fragment ion at *m/z* 335 from *m/z* 503 (d) and Reaction of *m/z* 335 with one pulse of CO (e).

In a subsequent experiment *m/z* 335 was dissociated to a set of fragment ions. The fragment ion at *m/z* 239 (related to loss of $2NH_2CHCO\text{-}H_2O$) was selected for further investigation. Similarly in this case the only addition was one water molecule observed at *m/z* 257. Also another feature started to grow at *m/z* 275 which was related to addition of the second water molecule, Figure 5.9. These results show that in contrast to CO, water molecules can strongly bind to manganese ions. Further exposing the metal ion center to the ligands for interaction, result in binding of more water molecules. These results are in

accordance with the observations for hydrated $[Ca(AlaGly-H)]^+$ and $[Ca(GlyAla-H)]^+$ complexes in which water was directly interacting with metal cation.[24]

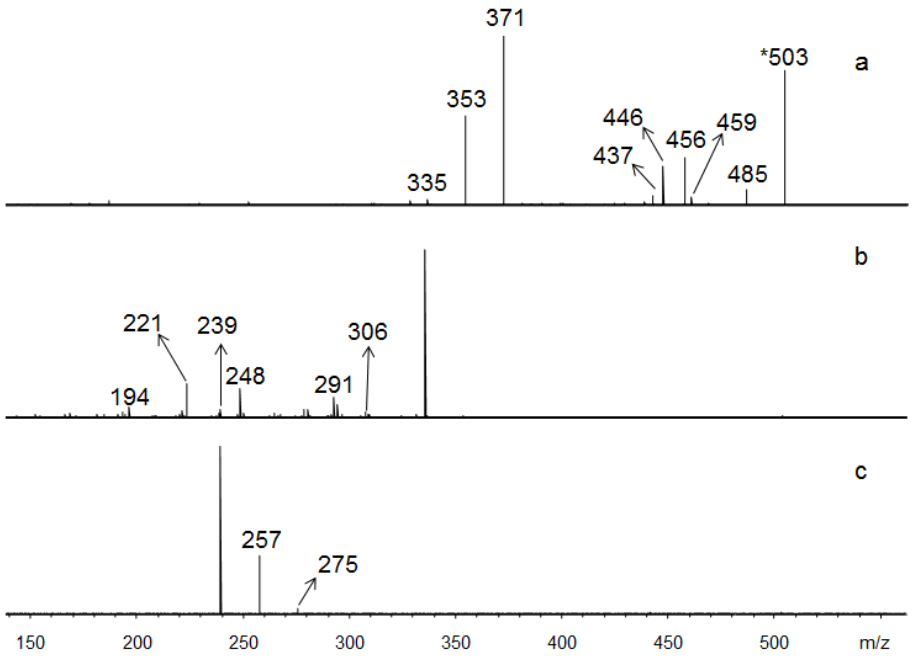


Figure 5.6. Mass spectra obtained from CO_2 laser IRMPD at 85% laser power of $[Mn_2(GlyGly-H)_3]^+$ (*m/z* 503) (a) and the fragment ion at *m/z* 335 (b) and Reaction of fragment ion at *m/z* 239 from *m/z* 335 with a pulse of CO (c).

5.4. Conclusion

Unimolecular fragmentation reactions of $[Mn_n(GlyGly-H)_{2n-1}]^+$ and $[Mn_{n+1}(GlyGly-H_{2n}]^{2+}$ complexes are investigated. It is shown that dissociation of singly charged clusters mainly result in losses of small neutral molecules such as water and carbon dioxide. Also other competing fragmentation pathways were loss of NH_2CHCO, glycine and GlyGly which were more dominant in larger singly charged clusters. Single or consecutive losses of GlyGly were the main dissociation observed for doubly charged clusters. Comparing

IRMPD spectra of $[M_3(GlyGly\text{-}H)_4]^{2+}$ for (Mg^{2+}, Ca^{2+}, Sr^{2+} and Mn^{2+}) showed that there is interesting similarity between the spectrum that was reported previously for Mg^{2+} cluster and Mn^{2+} cluster IRMPD spectrum both sharing the same features; two feature around 3300-3350 and 3390 cm^{-1} related to red-shifted amide N-H stretch and symmetric amine N-H stretch respectively. Also in both cases the asymmetric amine N-H stretch was observed around 3460 cm^{-1}. Computationally, we also showed that Mn^{2+} cluster lowest energy structure has all amine nitrogens hydrogen bonded to amide hydrogen and metal cations are coordinated with carbonyl oxygens at deprotonated carboxylic group as well as amide groups.

Detailed structural features of two possible spin multiplicity states, quartet and sextet, were studied. It was shown that the sextet spin state structure is much lower in energy. However it is not possible to distinguish between these structures based on the IRMPD spectroscopic studies in the N-H/O-H stretch region.

In this work we also showed that ion-molecule reaction between Mn^{2+} cluster and carbon monoxide only lead to addition of water molecules to the cluster and no carbon monoxide was attached to the cluster and cluster fragment ions.

References

(1) Crichton, R. R.: Biological Inorganic Chemistry, 2008.

(2) Lin, T.; Payne, A. H.; Glish, G. L. Dissociation pathways of alkali-cationized peptides: Opportunities for C-terminal peptide sequencing. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 497-504.

(3) Cerda, B. A.; Cornett, L.; Wesdemiotis, C. Probing the interaction of alkali and transition metal ions with bradykinin and its des-arginine derivatives via matrix-assisted laser desorption/ionization and postsource decay mass spectrometry. *Int. J. Mass Spectrom.* **1999**, *193*, 205-226.

(4) Dunbar, R. C.; Berden, G.; Oomens, J. How does a small peptide choose how to bind a metal ion? IRMPD and computational survey of CS versus Iminol binding preferences. *Int. J. Mass Spectrom.* **2013**, *354*, 356-364.

(5) Dunbar, R. C.; Polfer, N. C.; Berden, G.; Oomens, J. Metal ion binding to peptides: Oxygen or nitrogen sites? *Int. J. Mass Spectrom.* **2012**, *330-332*, 71-77.

(6) Dunbar, R. C.; Steill, J. D.; Polfer, N. C.; Berden, G.; Oomens, J. Peptide Bond Tautomerization Induced by Divalent Metal Ions: Characterization of the Iminol Configuration. *Angew. Chem., Int. Ed.* **2012**, *51*, 4591-4593.

(7) Dunbar, R. C.; Steill, J. D.; Polfer, N. C.; Oomens, J. Peptide Length, Steric Effects, and Ion Solvation Govern Zwitterion Stabilization in Barium-Chelated Di- and Tripeptides. *J. Phys. Chem. B* **2009**, *113*, 10552-10554.

(8) Kish, M. M.; Wesdemiotis, C. Selective cleavage at internal lysine residues in protonated vs. metalated peptides. *Int. J. Mass Spectrom.* **2003**, *227*, 191-203.

(9) Feng, W. Y.; Gronert, C.; Fletcher, K. A.; Warres, A.; Lebrilla, C. B. The mechanism of C-terminal fragmentations in alkali metal ion complexes of peptides. *Int. J. Mass Spectrom.* **2003**, *222*, 117-134.

(10) Lee, S. W.; Kim, H. S.; Beauchamp, J. L. Salt bridge chemistry applied to gas-phase peptide sequencing: Selective fragmentation of sodiated gas-phase peptide ions adjacent to aspartic acid residues. *J. Am. Chem. Soc.* **1998**, *120*, 3188-3195.

(11) Chu, I. K.; Guo, X.; Lau, T. C.; Siu, K. W. M. Sequencing of argentinated peptides by means of electrospray tandem mass spectrometry. *Anal. Chem.* **1999**, *71*, 2364-2372.

(12) Teesch, L. M.; Adams, J. Intrinsic interactions between alkaline-earth metal-ions and peptides - a gas-phase study. *J. Am. Chem. Soc.* **1990**, *112*, 4110-4120.

(13) Teesch, L. M.; Adams, J. Fragmentations of gas-phase complexes between alkali-metal ions and peptides - metal-ion binding to carbonyl oxygens and other neutral functional-groups. *J. Am. Chem. Soc.* **1991**, *113*, 812-820.

(14) Teesch, L. M.; Orlando, R. C.; Adams, J. location of the alkali-metal ion in gas-phase peptide complexes. *J. Am. Chem. Soc.* **1991**, *113*, 3668-3675.

(15) Dunbar, R. C.; Oomens, J.; Berden, G.; Lau, J. K. C.; Verkerk, U. H.; Hopkinson, A. C.; Siu, K. W. M. Metal ion complexes with HisGly: Comparison with PhePhe and PheGly. *J. Phys. Chem. A* **2013**, *117*, 5335-5343.

(16) Dunbar, R. C.; Steill, J. D.; Oomens, J. Chirality-induced conformational preferences in peptide-metal ion binding revealed by IR spectroscopy. *J. Am. Chem. Soc.* **2011**, *133*, 1212-1215.

(17) Shi, T. J.; Siu, K. W. M.; Hopkinson, A. C. Generation of La(peptide) (3+) complexes in the gas phase: Determination of the number of binding sites provided by dipeptide, tripeptide, and tetrapeptide ligands. *J. Phys. Chem. A* **2007**, *111*, 11562-11571.

(18) Carron, J. D.; Shin, J. W. Tandem mass spectrometric analysis of alkali metal cation-tripeptide clusters. *Chem. Phys. Lett.* **2014**, *598*, 102-107.

(19) Shin, J. W. Structure characterization of electrosprayed $[A\cdot PheGlyGly\text{-}H]^+$ (A = Ca, Sr, Ba, Cu, Fe, Zn, Mn, and Eu) clusters through collision-induced dissociation and extended DFT studies. *Int. J. Mass Spectrom.* **2014**, *372*, 39-45.

(20) Nemirovskiy, O. V.; Gross, M. L. Gas phase studies of the interactions of Fe^{2+} with cysteine-containing peptides. *J. Am. Soc. Mass Spectrom.* **1998**, *9*, 1285-1292.

(21) Liu, D., Seuthe, A. B., Ehrler, O. T., Zhang, X., Wyttenbach, T., Hsu, J. F., Bowers, M. T. Oxytocin-receptor binding: Why divalent metals are essential. *J. Am. Chem. Soc.* **2005**, *127*, 2024-2025.

(22) Brudvig, G. W. Water oxidation chemistry of photosystem II. *Philos. T. R. Soc. B: Biological Sciences* **2008**, *363*, 1211-1218.

(23) Gholami, A.; Fridgen, T. D. The unimolecular chemistry of $[Zn(amino\ acid)_2\text{-}H]^+$ in the gas phase: H_2 elimination when the amino acid is a secondary amine. *Phys. Chem. Chem. Phys.* **2014**, *16*, 3134-3143.

(24) Moghaddam, M. B., Fridgen, T. D. IRMPD spectroscopic study of microsolvated $[Na(GlyAla)]^+$ and $[Ca(GlyAla\text{-}H)]^+$ and the blue shifting of the hydrogen-bonded amide stretch with each water addition. *J. Phys. Chem. B* **2013**, *117*, 6157-6164.

(25) Gaussian 09, R. D., M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox, Gaussian, Inc., Wallingford CT, 2009.

(26) Moghaddam, M. B., Jami-Alahmadi, Y., Fridgen, T. D. Structures and Energetics of $[M_n(GlyGly\text{-}H)_{2n-1}]^+$ and $[M_{n+1}(GlyGly\text{-}H)_{2n}]^{2+}$ Clusters (M=Mg, Ca and Sr) in the Gas Phase,. **2015**, Unpublished manuscript.

(27) Shannon, R. D.; Prewitt, C. T. Effective ionic radii and crystal chemistry. *J. Inorg. Nucl. Chem.* **1970**, *32*, 1427-1441.

(28) Khodabandeh, M. H.; Davari, M. D.; Zahedi, M.; Ohanessian, G. Complexation of glycine by manganese (II) in the gas phase: A theoretical study. *Int. J. Mass Spectrom.* **2010**, *291*, 73-83.

(29) Leopoldini, M.; Russo, N.; Toscano, M. Which one among Zn(II), Co(II), Mn(II), and Fe(II) is the most efficient ion for the methionine aminopeptidase catalyzed reaction? *J. Am. Chem. Soc.* **2007**, *129*, 7776-7784.

(30) Operti, L.; Rabezzana, R. Gas-phase ion chemistry in organometallic systems. *Mass Spectrom. Rev.* **2006**, *25*, 483-513.

(31) Bellina, B.; Antoine, R.; Broyer, M.; Gell, L.; Sanader, Z.; Mitric, R.; Bonacic-Koutecky, V.; Dugourd, P. Formation and characterization of thioglycolic acid-silver cluster complexes. *Dalton Trans.* **2013**, *42*, 8328-8333.

(32) Gholami, A.; Fridgen, T. D. Dehydrogenation and demethanation of 2-methylpropane and propane in the gas-phase by the 16-electron complex $[Ru(bipy)_2(CO)]^{2+*}$ chemically activated by the association of $[Ru(bipy)_2]^{2+}$ and CO. *Dalton Trans.* **2013**, *42*, 3979-3985.

Chapter 6 . Summary and Future work

6.1. Summary and outlook

Mass spectrometry is a robust tool for the study of structures and energetics of ions in the gas phase especially when it is combined with powerful ion activation methods such as photodissociation and collisional dissociation. Well-known examples of such methods are IRMPD and CID that are widely used to elucidate structural features of ions. A Fourier transform ion cyclotron mass spectrometer is beneficial for ion activation purposes. In an ICR cell it is possible to store ions for long periods of time in a very low pressure environment and this makes it suitable for applying slow ion activation techniques, such as IRMPD, and furthermore investigating ion-molecule reactions. Additionally, applying soft ionization techniques, such as ESI is of great interest since even species held together by non-covalent interactions, such as the complexes which were the focus of this work, can be maintained intact during the ionization process. With the instrumentation used in this work, it was also possible to allow the interaction between the ions and water vapor in the high pressure region of the hexapole storage cell which results in hydrated ions.

In Chapter 3 of this thesis, the structure and thermochemistry of bare and hydrated $[Na(AlaGly)]^+$, $[Na(GlyAla)]^+$, $[Ca(AlaGly-H)]^+$ and $[Ca(GlyAla-H)]^+$ complexes were investigated using IRMPD mass spectrometry along with computational methods. Structures of these metal cation complexes were revealed by comparing their IRMPD spectra with the calculated IR spectra of some possible geometries. The isomeric dipeptide complexes with either Na^+ or Ca^{2+} were not distinguished from each other using the IRMPD spectroscopic technique since they both shared similar features in their

IRMPD spectra. It was shown that Na^+ complexes are in their canonical form rather than the zwitterionic, whereas Ca^{2+} complexes are deprotonated at the carboxylic acid group and these were confirmed by the presence of a carboxylic acid O-H stretch in the spectra of Na^+ complexes and the absence of this feature in Ca^{2+} complexes. Also calculations showed that in the lowest energy structure of Na^+ complexes, the metal cation is coordinated with the carboxylic as well as the amide carbonyl oxygens. In contrast, in the Ca^{2+} complexes, all carbonyl oxygens participated in metal coordination. In both cases the amine nitrogen was found to be intramolecularly hydrogen bonded to the amide hydrogen. Along with the bare complexes, the structure and energetics of hydrated complexes were also studied. It was shown that in all cases water directly interacts with the metal cation. More interestingly it was observed that the addition of more water molecules to Ca^{2+} complexes weakens the hydrogen bonding between amine nitrogen and amide hydrogen as a result of less positive charge being located on the complex and this was observed as a blue-shift of amide N-H stretch upon addition of each water molecule.

In Chapter 4, the structure and energetics of multimetallic complexes of deprotonated GlyGly were investigated. The unimolecular dissociation of $[M_n(GlyGly\text{-}H)_{2n-1}]^+$ n being 2, 3 and 4 and $[M_{n+1}(GlyGly\text{-}H)_{2n}]^{2+}$, n being 2, 4 and 6, complexes (M=Mg, Ca and Sr) were studied by CID. CID results showed that the size of the metal cation as well as the size and charge of complex can affect the stability and reactivity of such complexes. In singly charged complexes the main fragmentation observed was loss of a GlyGly and 57 Da. However it was observed that other fragment ions related to loss of small neutral molecules such as water and CO_2 were more dominant for smaller

complexes, especially for Mg^{2+}. Interestingly Ca^{2+}, and to a greater extent Sr^{2+}, doubly charged complexes showed charge separation reactions as another competing dissociation pathway and became the dominant (or sole) fragmentation reaction for larger (bigger n) complexes. The result of such charge separation reactions were formation of two singly charged complexes as fragment ions. In contrast to Ca^{2+} and Sr^{2+}, the main dissociation pathway for Mg^{2+} doubly charged complexes was loss of GlyGly and 57 Da, similar to singly charged complexes, and no charge separation reaction was observed in this case.

The nature of 57 Da loss was also studied both experimentally and computationally in Chapter 4. $[Ca_2(GlyGly\text{-}H)_3]^+$ was isotopically labeled as an example. It was shown that loss of 59 Da from triply deuterated complex is in accordance with formation of an aminoketene through hydrogen transfer from CH_2 to the amine nitrogen. More interestingly, binding energy of an aminoketene molecule to the rest of ligand was computationally found to be less than the binding energy of the alternative cyclic product resulted from hydrogen being transferred from the amine group. In addition, the study of GlyGly loss showed that loss of the GlyGly ligand perpendicular to the molecular plane was more favorable and a result of hydrogen transfer from a neighboring GlyGly with identical metal coordination features.

The effect of metal cation size as well as the complex charge was studied on both singly and doubly charged complexes using energy-resolved CID experiments. The results showed that the order of dissociation energies is $Ca^{2+}>Mg^{2+}>Sr^{2+}$ for singly charged complexes whereas it is $Sr^{2+}>Ca^{2+}>Mg^{2+}$ for doubly charged complexes. It was proposed that this was because of the higher order of metal-ligand interactions for Ca^{2+} in

singly charged complexes. In doubly charged complexes, however, the charge density of the metal cation on the complex could be the reason for the stability of these complexes.

Optimized lowest energy structures of $[M_3(GlyGly\text{-}H)_4]^{2+}$ (Mg, Ca and Sr) as well as $[Ca_5(GlyGly\text{-}H)_8]^{2+}$ were found to be similar to what we observed for $[Ca(AlaGly\text{-}H)]^+$ and/or $[Ca(GlyAla\text{-}H)]^+$ in Chapter 3; all GlyGly ligands deprotonated at carboxylic group and metal cations coordinated with carbonyl oxygens of more than one ligand and amine nitrogens were intramolecularly hydrogen bonded to amide hydrogens. IRMPD and calculated IR spectra of some lowest energy structures of $[M_3(GlyGly\text{-}H)_4]^{2+}$ complexes were combined to characterize the structural features of these complexes. The spectra of $[M_3(GlyGly\text{-}H)_4]^{2+}$ (Mg, Ca and Sr) complexes showed features at 3300, 3350 and 3370 cm^{-1}, respectively, which was related to the red-shifted amide N-H stretches. The interesting point here was the blue-shifting of this feature in Ca^{2+} and to an even larger extent in the Sr^{2+} complex which was related to less charge density being localized by larger metal cations on the complex and as a result on the amide group. This means that the ionic hydrogen bond was weakened in these cases strengthening the amide N-H bond, resulting in its stretching mode shifting to higher energies. Another interesting result was the feature around 3390 cm^{-1} related to the symmetric amine NH_2 stretch in the spectrum of these complexes which was more resolved in the Mg^{2+} complex and less intense in Ca^{2+} and almost unresolved in Sr^{2+} and this was most likely because of the blue shift of the amide N-H stretch in the spectra of two later metal cations complexes resulting in overlap of amide and amine stretching modes. The blue shift of the amide N-H stretch was also observed in the IRMPD spectra of larger doubly charged complex

$[M_5(GlyGly\text{-}H)_8]^{2+}$ consistent with the results for the $[M_3(GlyGly\text{-}H)_4]^{2+}$ complex. However, in all metal cations only one feature was observed relating to red-shifted amide N-H stretch as well as the symmetric and asymmetric amine NH_2 stretches.

In Chapter 5 of this work the unimolecular fragmentation reactions of $[Mn_n(GlyGly\text{-}H)_{2n-1}]^+$ and $[Mn_{n+1}(GlyGly\text{-}H_{2n}]^{2+}$ complexes were investigated. An interesting similarity was observed between the CID results of Mn^{2+} complexes with what we showed for Mg^{2+} complexes in the previous Chapter. This was related to the similar ionic size of these metal cations. Smaller singly charged complexes showed extensive fragmentation of small neutral molecules whereas in other cases the main fragmentation was loss of GlyGly and 57 Da. In contrast, the main dissociation observed for doubly charged complexes were single or consecutive losses of GlyGly molecules.

More evidence was provided for similarity of Mg^{2+} and Mn^{2+} from comparison of IRMPD spectra of the $[M_3(GlyGly\text{-}H)_4]^{2+}$ for M=Mg, Ca and Sr with Mn complexes. It was shown that the position of the red-shifted amide stretch as well as the symmetric and asymmetric amine NH_2 stretches matched well between Mg^{2+} and Mn^{2+} complexes.

Ion-molecule reactions of Mn complexes were studied in Chapter 5. Reactivity of the parent ion as well as the fragment ions stemming from this complex were investigated toward CO molecule. No association reaction was observed for the parent ion complex. Furthermore, the only addition that was observed for the fragment ions was water molecules.

The main goal of this work was to characterize the structure and energetics of metal cation bound complexes of dipeptides, a first step in the research efforts of the

Fridgen group. Deep insight into the structural and thermochemical features of smaller systems that can represent/reproduce the behavior of biologically relevant molecules as well as study of the effect of metal cations in formation of such highly ordered systems is of great importance in many areas such as chemistry, biochemistry and especially in material sciences and nanotechnology in order to be able to design new systems with modified/enhanced chemical and physical properties. In this work we could get useful structural information about small metal cation complexes with the simplest dipeptides and, more interestingly, in Chapter 4 and 5, we could extend this knowledge for similar systems in larger self-assembled forms. In Chapter 4 we initiated the detailed structure and reactivity studies of most biologically relevant metal cations and compared these complexes when they are bound to different metal cations. In Chapter 5 it was tried to investigate to see if a transition metal cation complex structure and reactivity is different for example if it can participate in ion-molecule reactions e.g. association or oxidation-reduction.

6.2. Future work

The work presented here provided a detailed discussion on structures, energetics and uni- and bimolecular reactivity of metal ion complexes of dipeptides in the gas phase. More importantly we were able to form and identify multinuclear metal clusters of dipeptides and study their ion-molecule reactions in the gas phase. This is an interesting area of research since these clusters show structural properties similar to numerous biological systems in which metal ions act as the reaction center. Also a deep

understanding of these highly ordered systems can provide useful information to design and modify materials that can be employed in different areas of research and technology, such as nanotechnology as well as biomaterials sciences.

As the next step of this work our goal is to develop a more detailed knowledge of the chemistry of multimetallic clusters of transition metals. Transition metal clusters can offer catalytic characteristics and they are especially well-known for their oxidation-reduction reactivity. So far, we have investigated bimolecular reactions of Mn clusters with small inorganic molecules such as water and CO and we were able to attach water to the metal centers of these clusters. In this regard it is also worthwhile investigating the interaction of small organic molecules with these clusters in order to follow their chemical reactions in the gas phase. Hydrocarbons can be transformed to other materials through dehydrogenation/demethanation via catalytic reactions. These reactions are of great importance in the petroleum field especially in the case of alkanes which are known to have very low reactivity. Transition metals in their bare and ligated form have shown great catalytic characteristics towards hydrocarbon by participating in C-C/C-H bond activation of these molecules. For future work it would be interesting to investigate the reactivity of small hydrocarbons (such as propane) towards Mn clusters to see if it could lead to a chemical reaction. Possible loss of hydrogen or methane could be traced as dissociation products under collision-induced dissociation and related to propane reformation as was the case in a study in our group on Ru complexes.[1] Also it is of relevance to compare the results for clusters of different transition metals. Studying metal clusters of first row transition metals, for example, can help understanding the effect of metal electronic structure in the formation, stability and reactivity of metal clusters. The

first step, however, would be structure characterization and unimolecular reaction investigation of all of the transition metal clusters by means of mass spectrometry, IRMPD spectroscopy combined with calculations.

In addition, the effect of metal cluster size (number of metal cations) on the reactivity of these systems toward neutral molecules will be worth studying since it has been shown that metal clusters with different number of metal ions can have different reactivity.[2,3] Studies have also shown that merging different metal ions together can modify/enhance the catalytic reactivity of metal clusters. Schwarz and coworkers investigated formation of HCN from CH_4 and NH_3 reactions by bimetallic clusters in the gas phase. They showed that homonuclear transition metal clusters, Pt_2^+, Ag_2^+, Au_2^+ and Cu_2^+ are not reactive in both of C-H bond activation and C-N bond formation whereas their mixed metal clusters, $Pt\text{-}Cu^+$, $Pt\text{-}Au^+$ and $PtAg^+$ show efficient catalytic properties.[4,5] In the next step of this work it is worth trying to form clusters with a mixture of transition metals by introducing the solutions of ligand (GlyGly) as well as a mixture of transition metal salts to the ionization source to see if stable metal clusters can be formed by this method and, if they can, how would Fe(II) cations for example affect the binding of neutral molecules such as CO, O_2 and hydrocarbons to the Fe-Mn clusters.

In photosystem II manganese cations with different oxidation states form a cluster which is known as the oxygen evolving center (or complex) and participates in oxidation-reduction reactions in which water become oxidized and form oxygen.[6] Accordingly another step in our research could be applying different oxidation states of transition metals to form multimetallic clusters which can be taken for further investigation of the

oxidation-reduction reactivity of such systems. For example, if they can oxidize the water molecules which we could successfully attach to metal centers in Mn clusters.

At last this could be of great interest to look at the metal complexes formed using known biomolecules as ligands. For example, carnosine (beta-alanyl-L-histidine) and anserine (beta-alanyl-N-methylhistidine) are biologically occurring dipeptides that exist as antioxidants in the living systems. The result of studies on such systems can image the differences in structure, energetics and both uni- and bimolecular reactions of metal ion complexes in the presence of diverse ligands. For example, how could replacing GlyGly with carnosine/anserine affect the metal ion coordination? Does the structure form intramolecular hydrogen bonding? What dissociation mechanism or loss of what neutral molecule is the lowest energy fragmentation pathway and what kind of chemical reactions do they participate in? It is also essential to try to computationally predict the structure and thermochemistry of the systems under study however the more complex the systems, the bigger is the challenge.

References

(1) Gholami, A.; Fridgen, T. D. Dehydrogenation and demethanation of 2-methylpropane and propane in the gas-phase by the 16-electron complex $[Ru(bipy)_2(CO)]^{2+*}$ chemically activated by the association of $[Ru(bipy)_2]^{2+}$ and CO. *Dalton Trans.* **2013**, *42*, 3979-3985.

(2) Harding, D.; Ford, M. S.; Walsh, T. R.; Mackenzie, S. R. Dramatic size effects and evidence of structural isomers in the reactions of rhodium clusters, $Rh_n^{+/-}$, with nitrous oxide. *Phys. Chem. Chem. Phys.* **2007**, *9*, 2130-2136.

(3) Adlhart, C.; Uggerud, E. C–H activation of alkanes on Rh_n^+(n=1–30) clusters: Size effects on dehydrogenation. *J. Chem. Phys.* **2005**, *123*, 214709.

(4) Koszinowski, K.; Schroder, D.; Schwarz, H. Probing cooperative effects in bimetallic clusters: Indications of C-N coupling of CH_4 and NH_3 mediated by the cluster ion $PtAu^+$ in the gas phase. *J. Am. Chem. Soc.* **2003**, *125*, 3676-3677.

(5) Koszinowski, K.; Schroder, D.; Schwarz, H. Coupling of Methane and Ammonia by Dinuclear Bimetallic Platinum - Coinage-Metal Cations PtM^+. *Angew. Chem., Int. Ed.* **2004**, *43*, 121-124.

(6) Brudvig, G. W. Water oxidation chemistry of photosystem II. *Philos. T. R. Soc. B: Biological Sciences* **2008**, *363*, 1211-1218.

Appendix

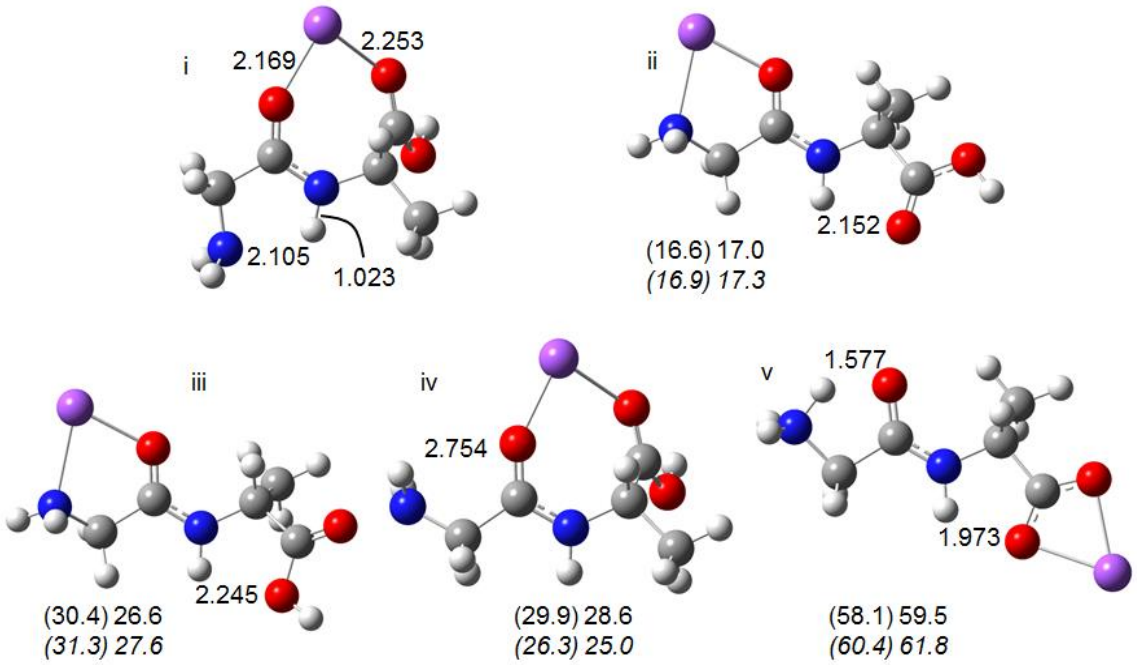


Figure S3.1. B3LYP/6-31+G(d,p) computed structures for $[Na(GlyAla)]^+$ along with relative enthalpies (in parentheses) and Gibbs energies computed at the B3LYP/6-31+G(d,p) (top numbers) or B3LYP/6-311++G(3df,3pd)//B3LYP/6-31+G(d,p) levels of theory (bottom, italicized numbers). Energies are in kJ mol^{-1} relative to the most thermodynamically stable structure, i.

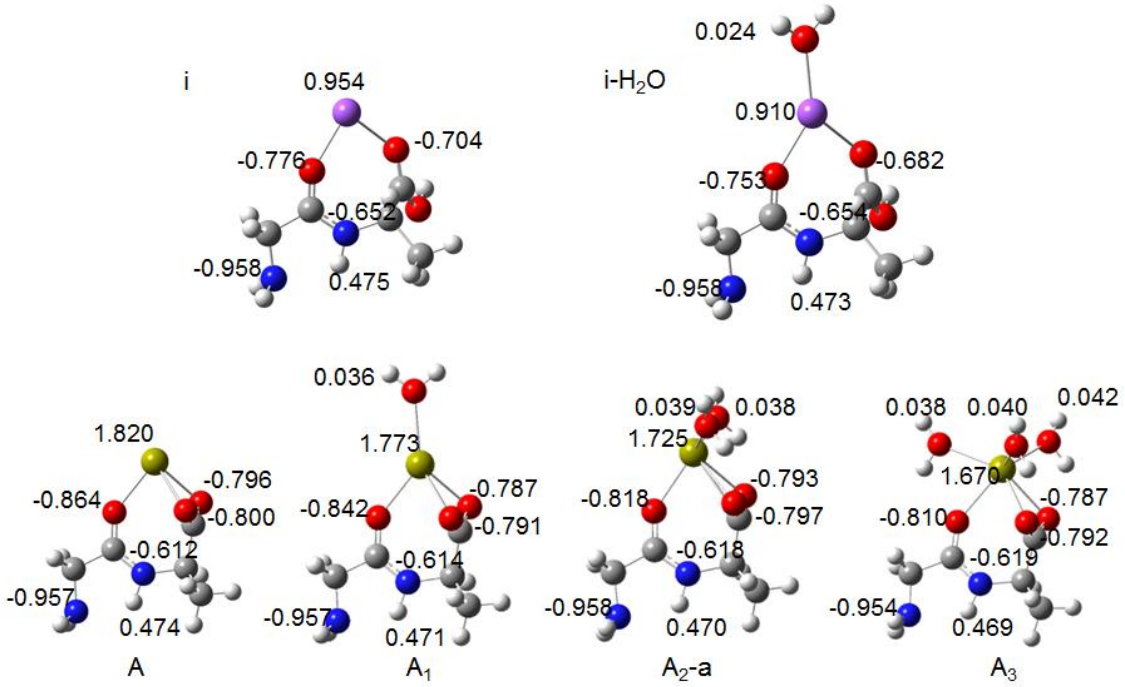


Figure S3.2. Computed NBO Charge distributions for the lowest energy $Na(GlyAla)^+$ and $[Ca(GlyAla-H)]^+$ complexes as well as the lowest energy hydrated complexes.

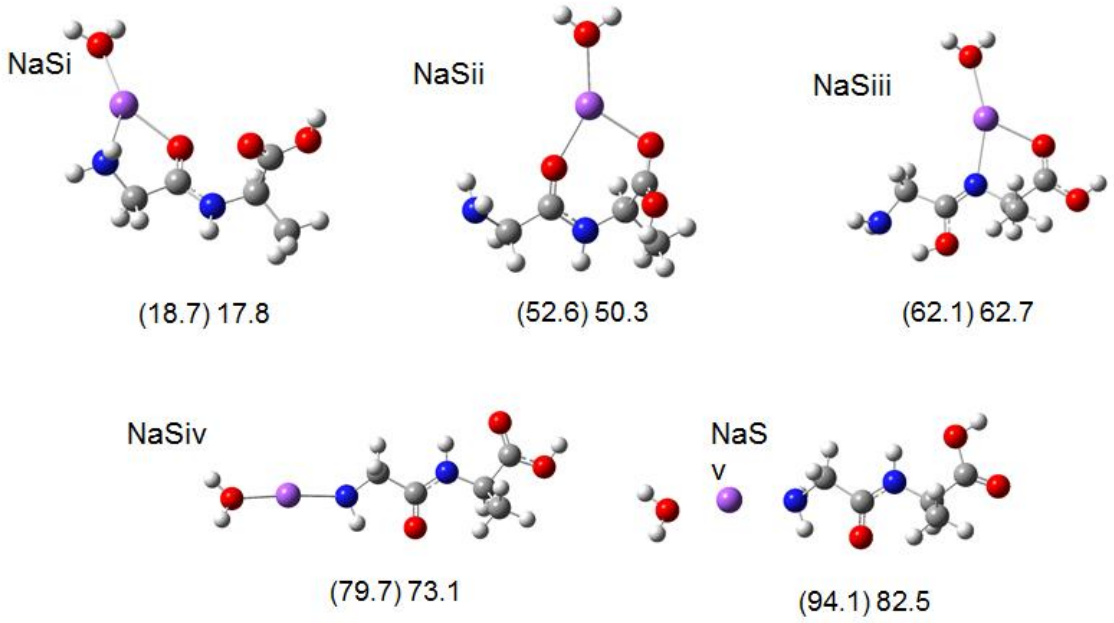
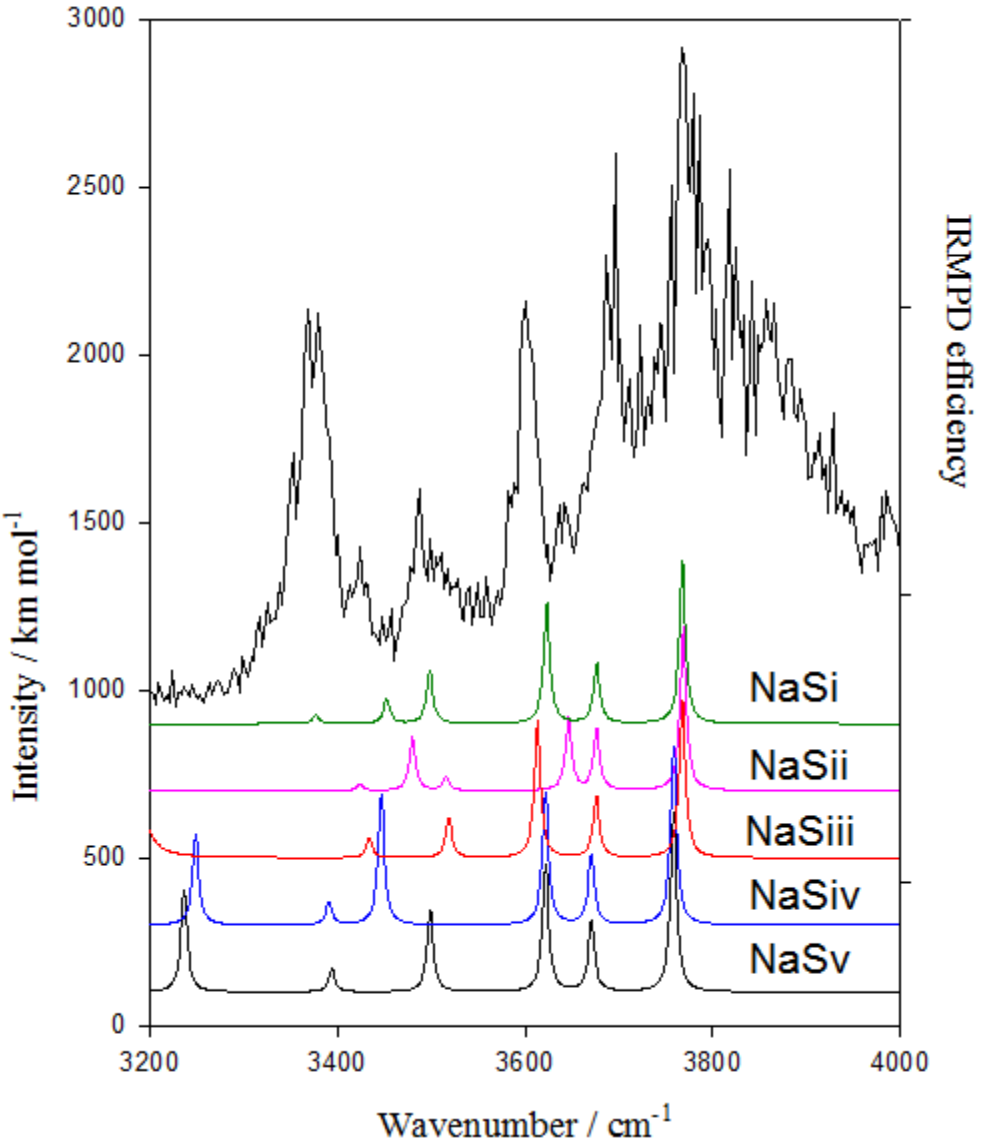


Figure S3.3. Supplementary B3LYP/6-31+G(d,p) computed structures for $[Na(GlyAla)H_2O]^+$ along with relative enthalpies (in parentheses) and Gibbs energies computed at the B3LYP/6-31+G(d,p). Energies are in kJ mol^{-1} relative to the most thermodynamically stable structure in Figure 3.2. Computed IR spectra for these structures are compared to the experimental IRMPD spectrum in Figure S3.4

Figure S3.4.


3000
2500
2000
1500
1000
500
0
Intensity / km mol-1
IRMPD efficiency
NaSi
NaSii
NaSiii
NaSiv
NaSv
3200
3400
3600
3800
4000
Wavenumber / cm-1

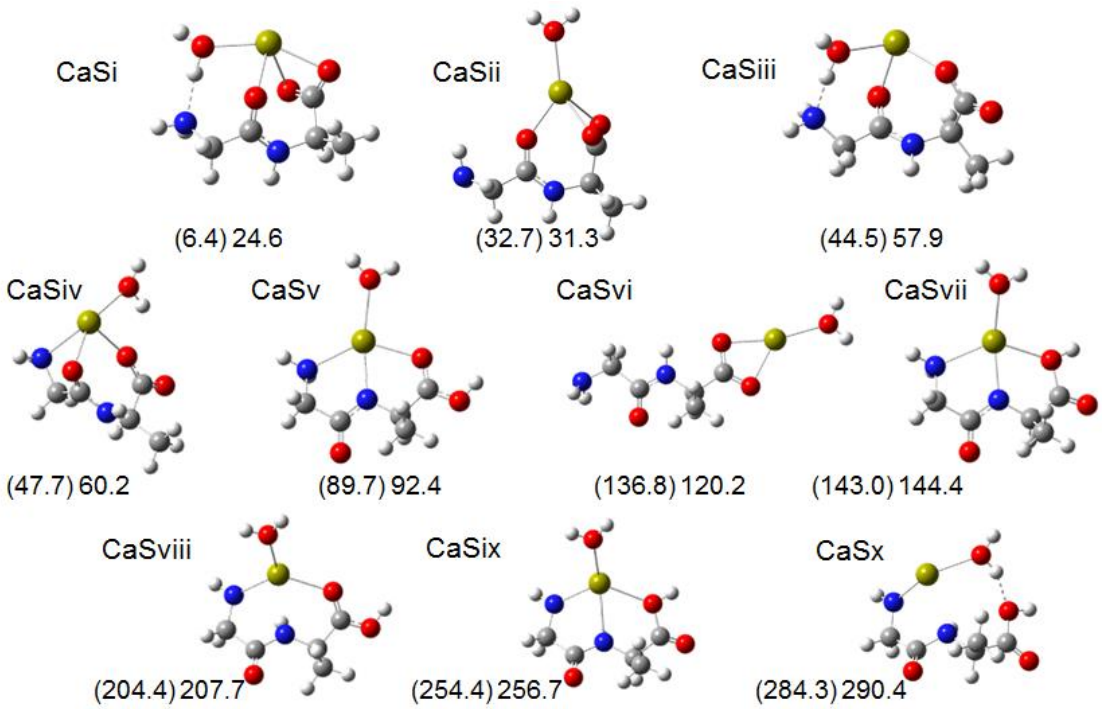
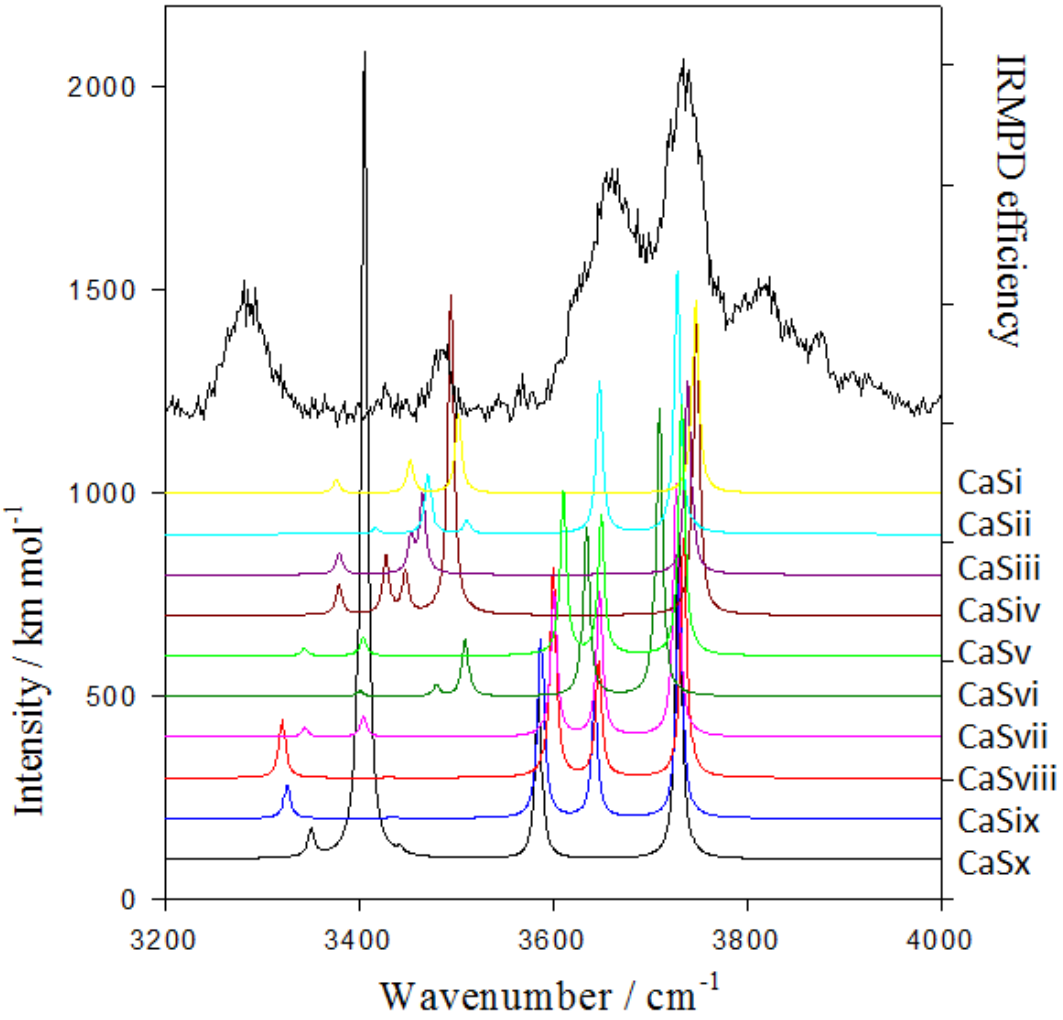


Figure S3.5. Supplementary B3LYP/6-31+G(d,p) computed structures for $[Ca(GlyAla)H_2O]^+$ along with relative enthalpies (in parentheses) and Gibbs energies computed at the B3LYP/6-31+G(d,p). Energies are in kJ mol^{-1} relative to the most thermodynamically stable structure in Figure 3.5. Computed IR spectra for these structures are compared to the experimental IRMPD spectrum in Figure S3.6.

Figure S3.6.


Intensity / km mol-1
2000
1500
1000
500
0
3200
3400
3600
3800
4000
Wavenumber / cm-1
IRMPD efficiency
CaSi
CaSii
CaSiii
CaSiv
CaSv
CaSvi
CaSvii
CaSviii
CaSix
CaSx

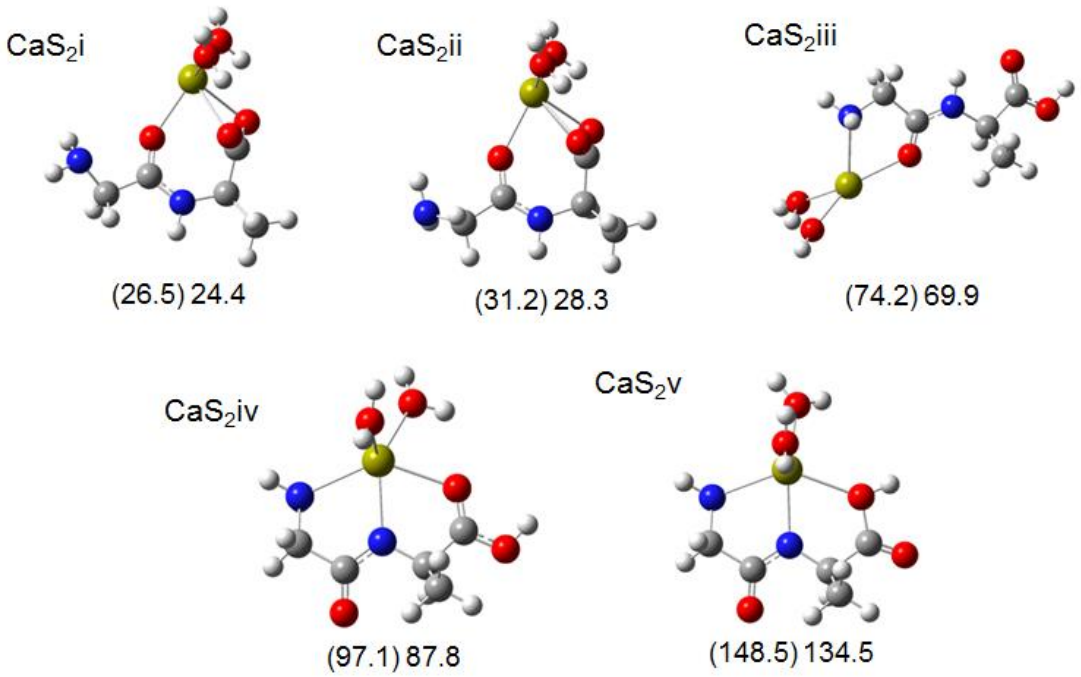
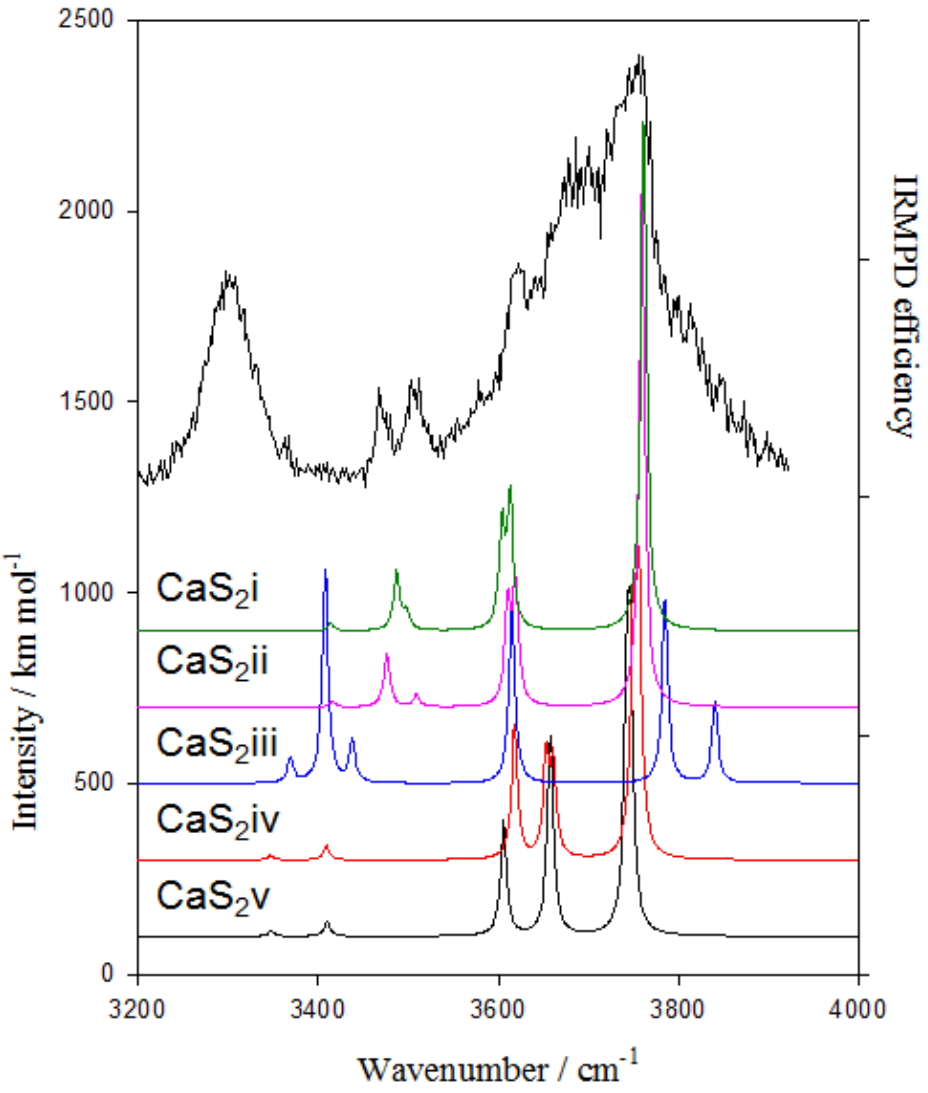


Figure S3.7. Supplementary B3LYP/6-31+G(d,p) computed structures for $[Ca(GlyAla)(H_2O)_2]^+$ along with relative enthalpies (in parentheses) and Gibbs energies computed at the B3LYP/6-31+G(d,p). Energies are in kJ mol^{-1} relative to the most thermodynamically stable structure in Figure 3.5. Computed IR spectra for these structures are compared to the experimental IRMPD spectrum in Figure S3.8.

Figure S3.8.


2500
2000
1500
1000
500
0
Intensity / km mol-1
IRMPD efficiency
CaS2i
CaS2ii
CaS2iii
CaS2iv
CaS2v
3200
3400
3600
3800
4000
Wavenumber / cm-1

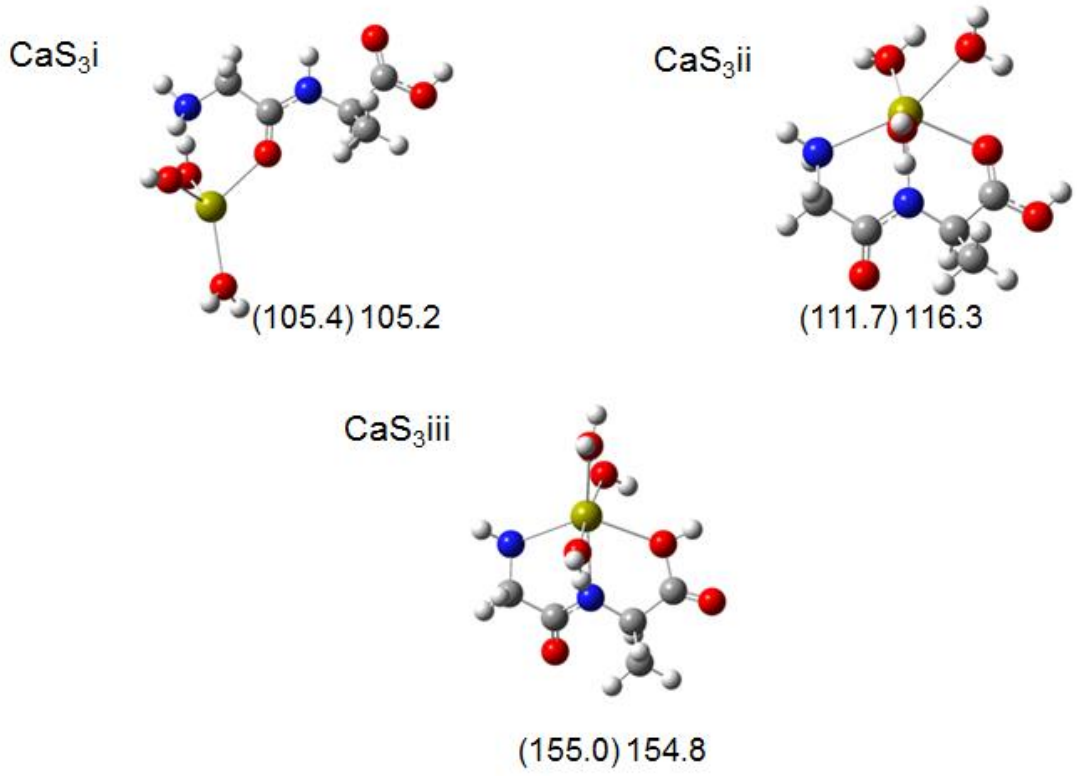
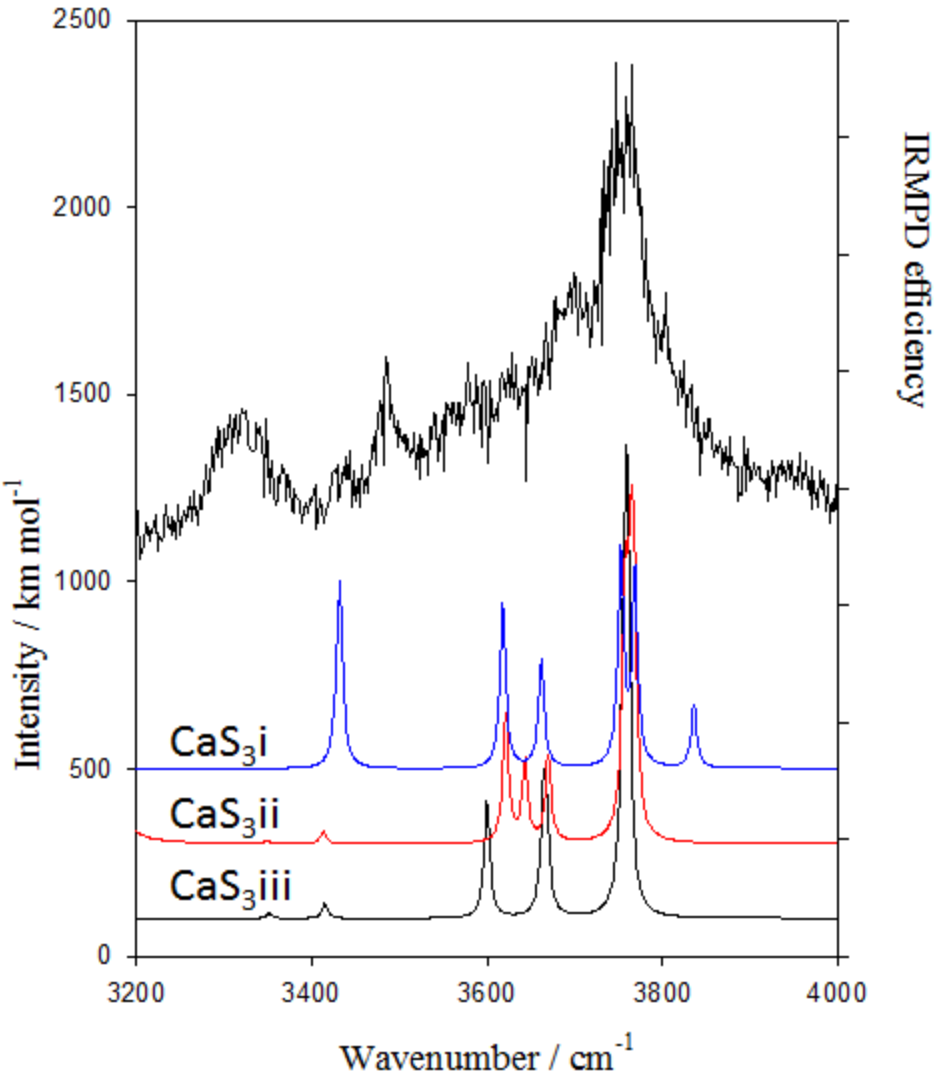


Figure S3.9. Supplementary B3LYP/6-31+G(d,p) computed structures for $[Ca(GlyAla)(H_2O)_3]^+$ along with relative enthalpies (in parentheses) and Gibbs energies computed at the B3LYP/6-31+G(d,p). Energies are in kJ mol^{-1} relative to the most thermodynamically stable structure in Figure 3.5. Computed IR spectra for these structures are compared to the experimental IRMPD spectrum in Figure S3.10.

Figure S3.10.


2500
2000
1500
1000
500
0
Intensity / km mol-1
IRMPD efficiency
CaS3i
CaS3ii
CaS3iii
3200
3400
3600
3800
4000
Wavenumber / cm-1

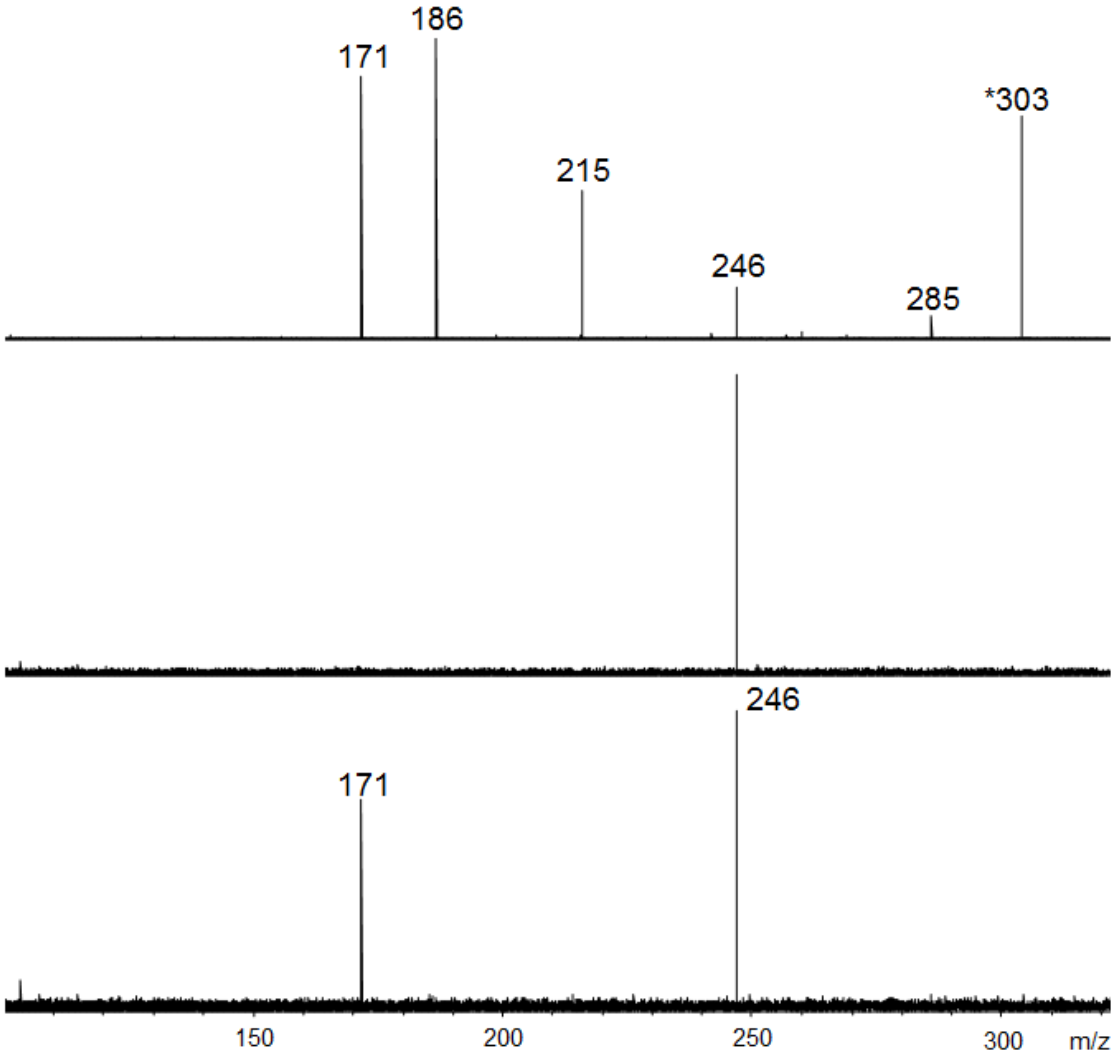


Figure S4.1. SORI-CID spectrum of $[Ca(GlyGly)(GlyGly-H)]^+$ (*m/z* 303) (a) and the mass spectra following isolation (b) and dissociation (c) of fragment ion *m/z* 246 formed by SORI-CID of *m/z* 303.

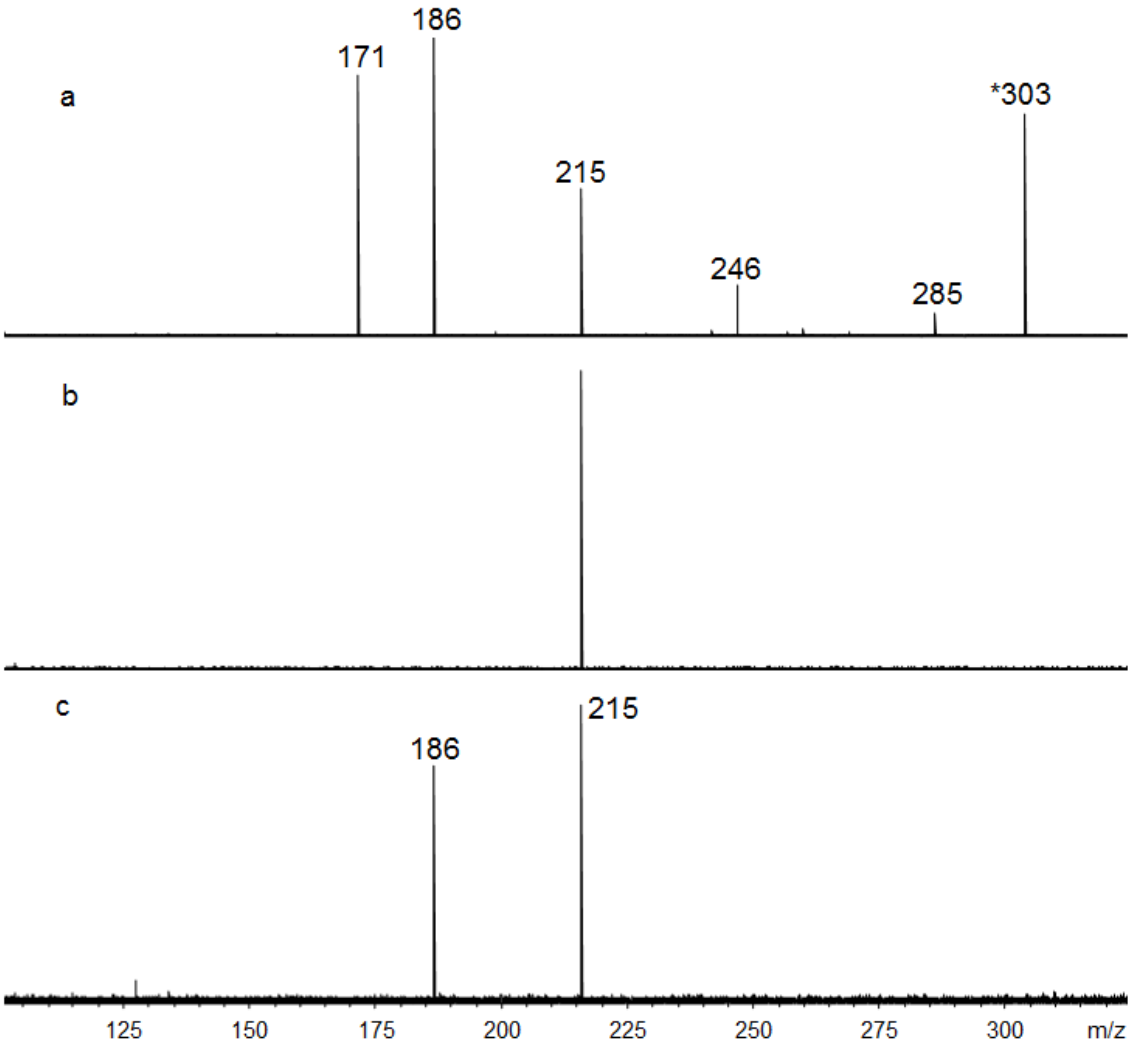


Figure S4.2. SORI-CID spectrum of $[Ca(GlyGly\text{-}H)(GlyGly)]^+$ (*m/z* 303) (a) and the mass spectra following isolation (b) and dissociation (c) of fragment ion *m/z* 215 formed by SORI-CID of *m/z* 303.

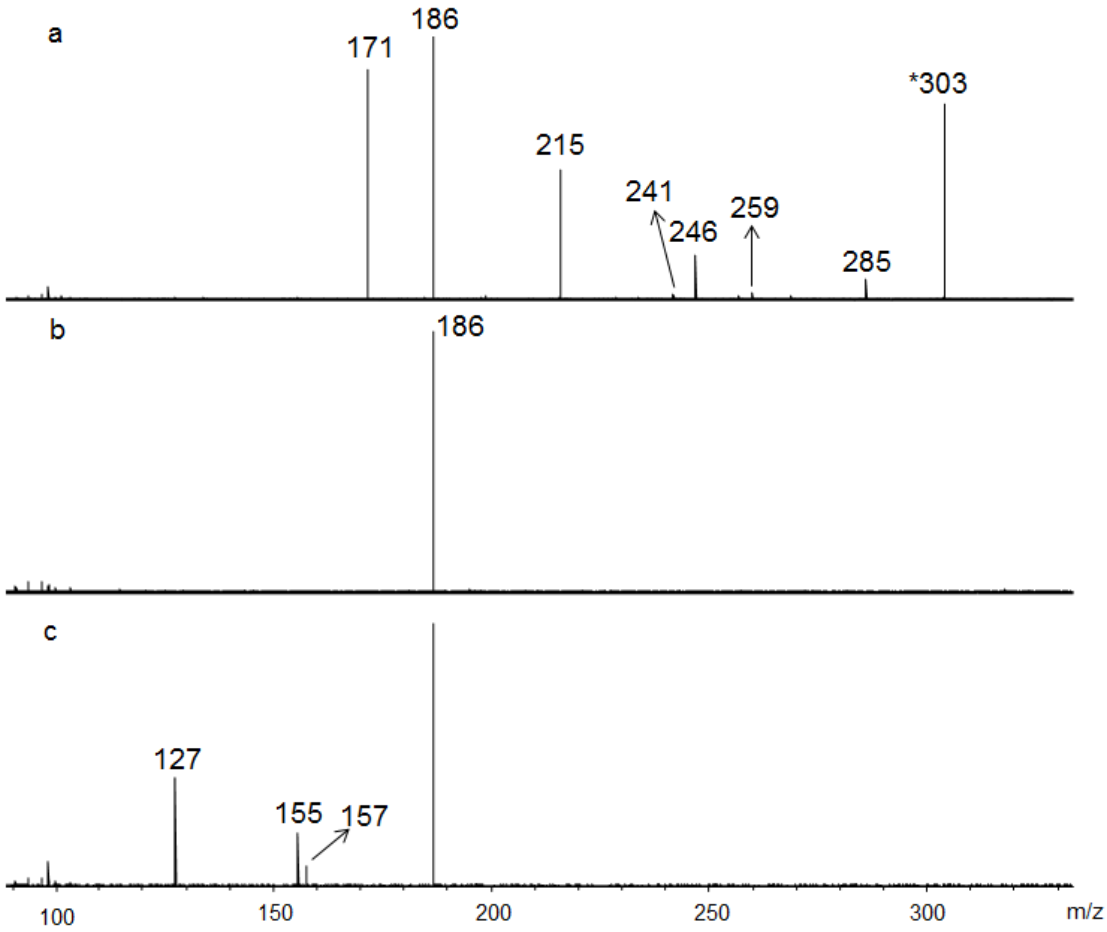


Figure S4.3. SORI-CID spectrum of $[Ca(GlyGly\text{-}H)(GlyGly)]^+$ (*m/z* 303) (a) and the mass spectra following isolation (b) and dissociation (c) of fragment ion *m/z* 186 formed by SORI-CID of *m/z* 303.

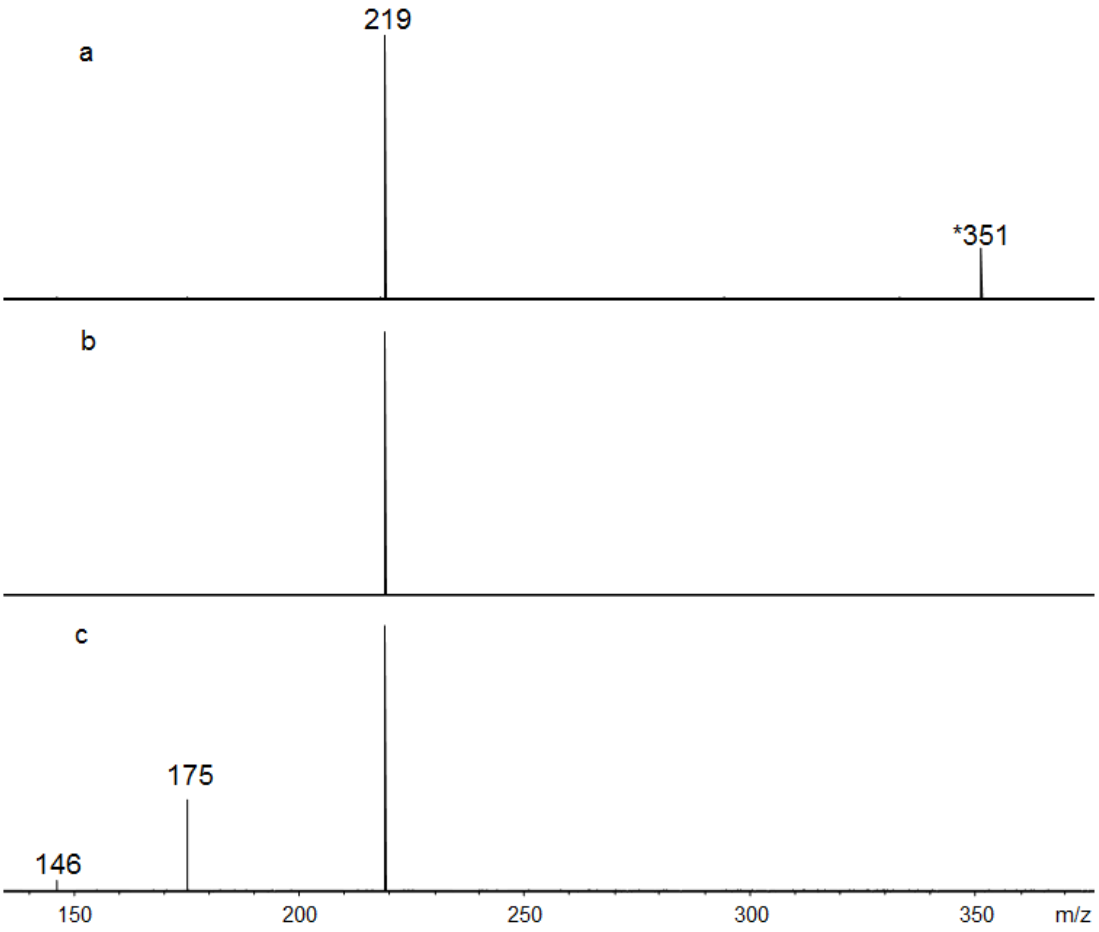


Figure S4.4. SORI-CID spectrum of $[Sr(GlyGly\text{-}H)(GlyGly)]^+$ (*m/z* 351) (a) and the mass spectra following isolation (b) and dissociation (c) of fragment ion *m/z* 219 formed by SORI-CID of *m/z* 351.

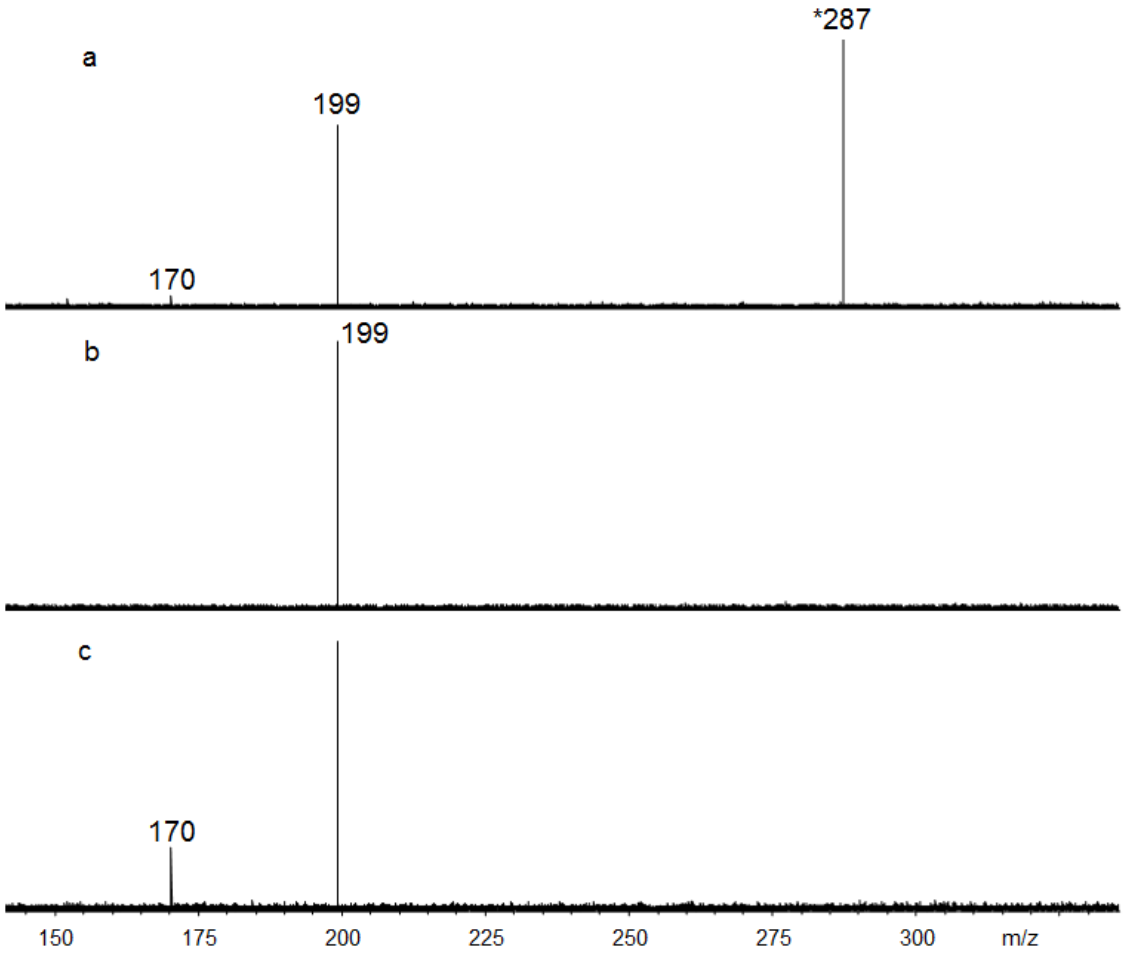


Figure S4.5. SORI-CID spectrum of $[Mg(GlyGly-H)(GlyGly)]^+$ (*m/z* 287) (a) and the mass spectra following isolation (b) and dissociation (c) of fragment ion *m/z* 199 formed by SORI-CID of *m/z* 287.

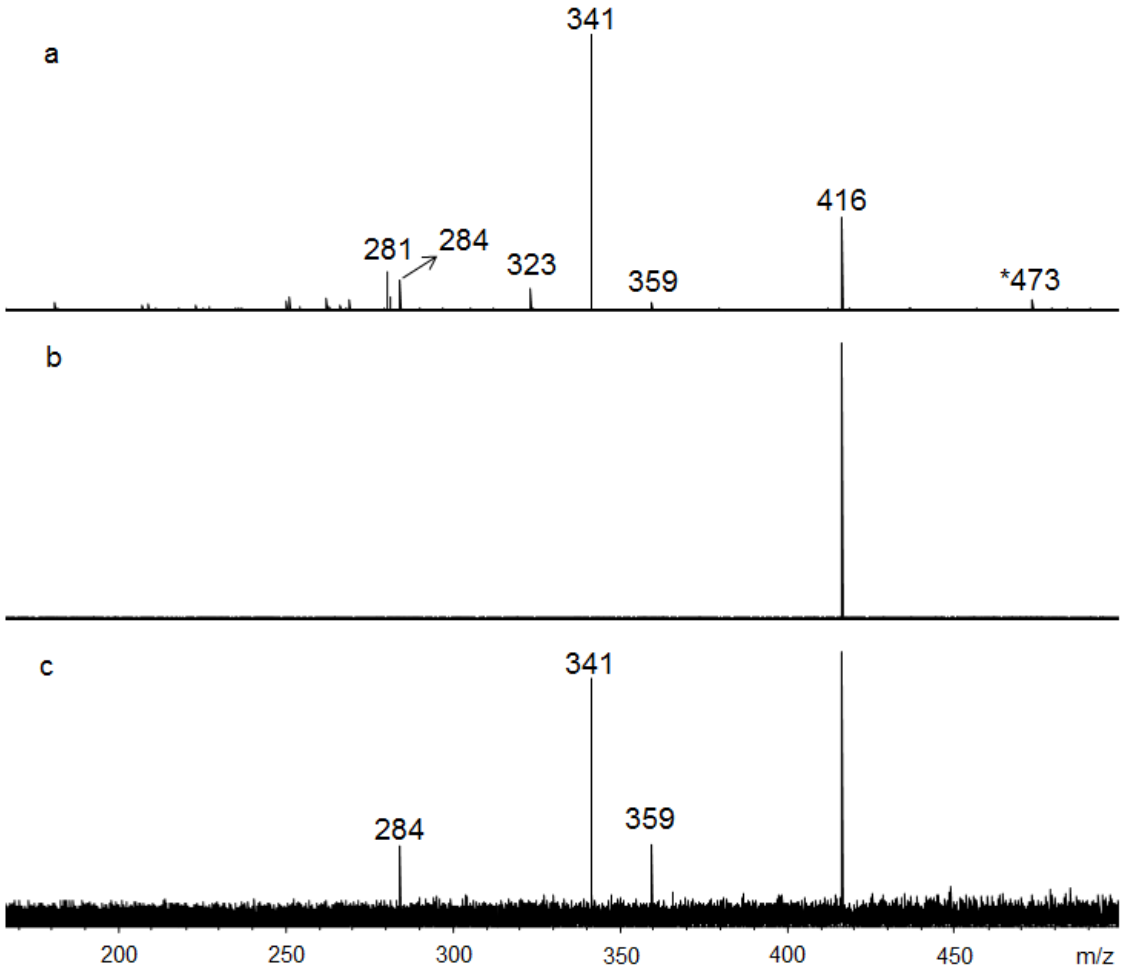


Figure S4.6. SORI-CID spectrum of $[Ca_2(GlyGly\text{-}H)_3]^+$ (*m/z* 473) (a) and the mass spectra following isolation (b) and dissociation (c) of fragment ion *m/z* 416 formed by SORI-CID of *m/z* 473.

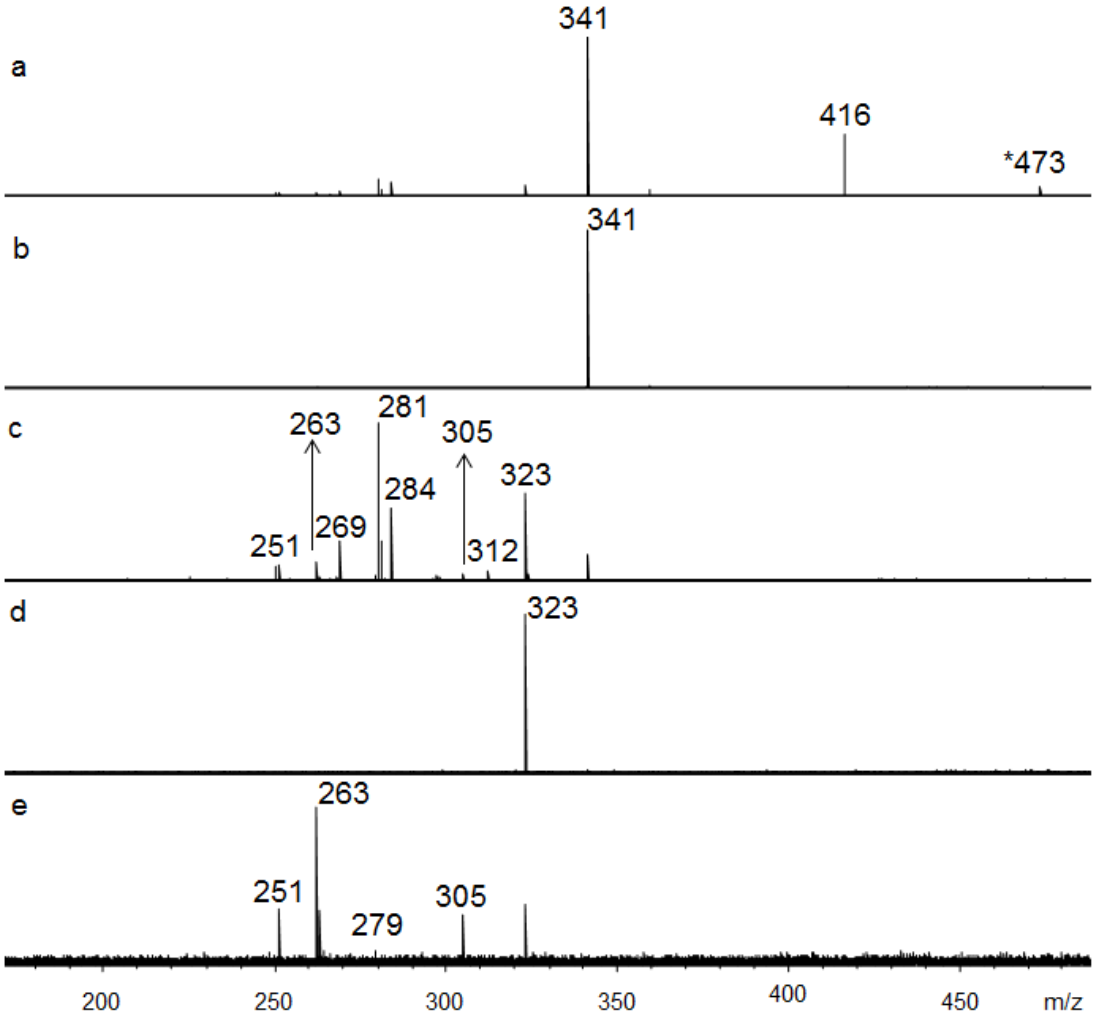


Figure S4.7. SORI-CID spectrum of $[Ca_2(GlyGly\text{-}H)_3]^+$ (*m/z* 473) (a), isolation and dissociation mass spectra of fragment ion *m/z* 341 formed by SORI-CID of *m/z* 473 (b,c) and isolation and dissociation mass spectra of *m/z* 323 formed by SORI/CID on the fragment ion at *m/z* 341(d,e).

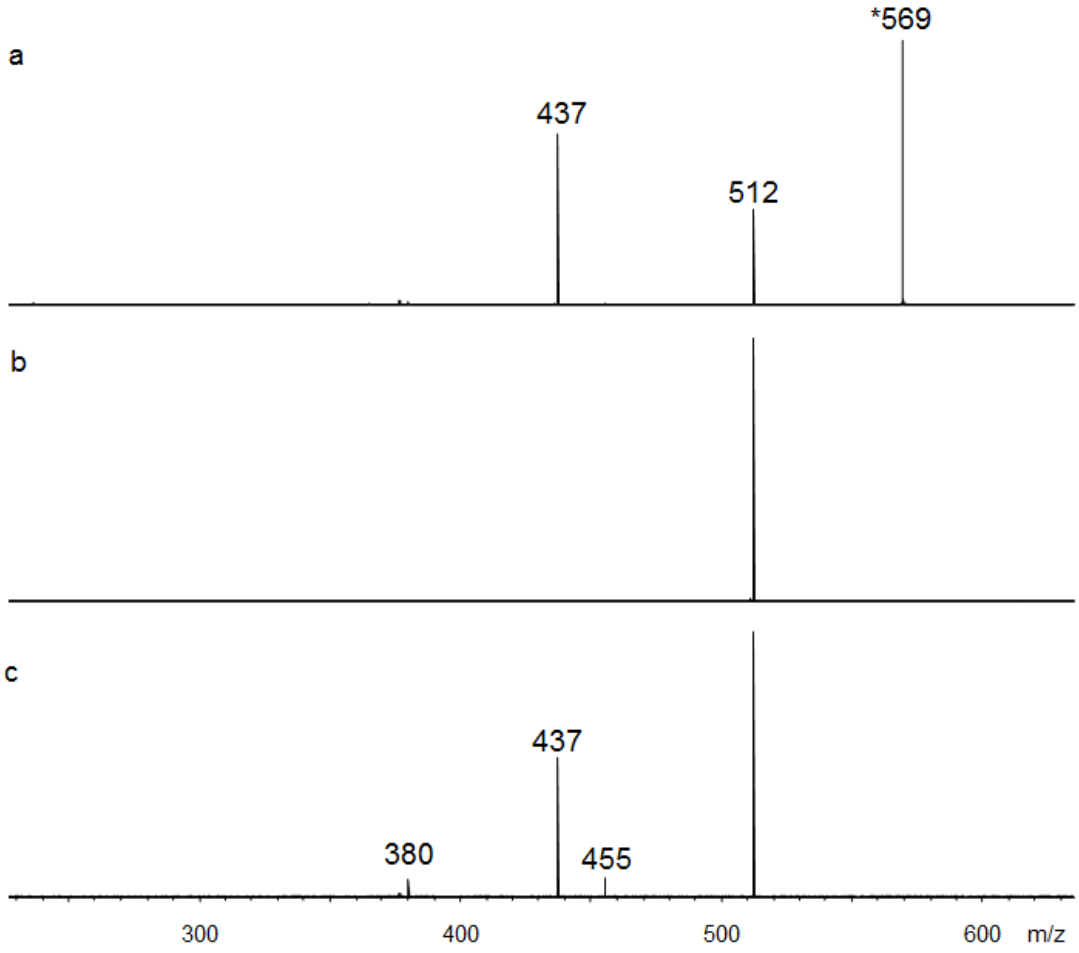


Figure S4.8. SORI-CID spectrum of $[Sr_2(GlyGly\text{-}H)_3]^+$ (*m/z* 569) (a) and the mass spectra following isolation (b) and dissociation (c) of fragment ion *m/z* 512 formed by SORI-CID of *m/z* 569 (b,c).

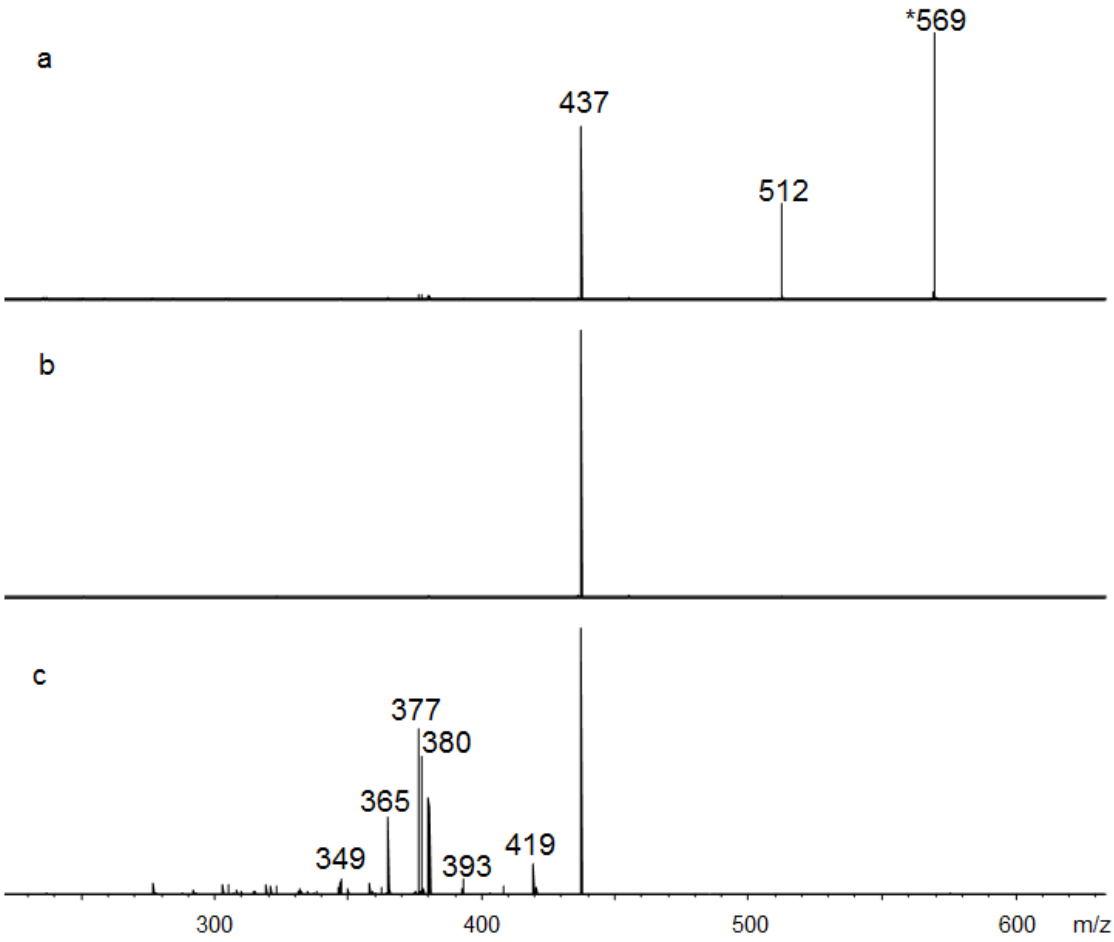


Figure S4.9. SORI-CID spectrum of $[Sr_2(GlyGly\text{-}H)_3]^+$ (*m/z* 569) (a) and the mass spectra following isolation (b) and dissociation (c) of fragment ion *m/z* 437 formed by SORI-CID of *m/z* 569.

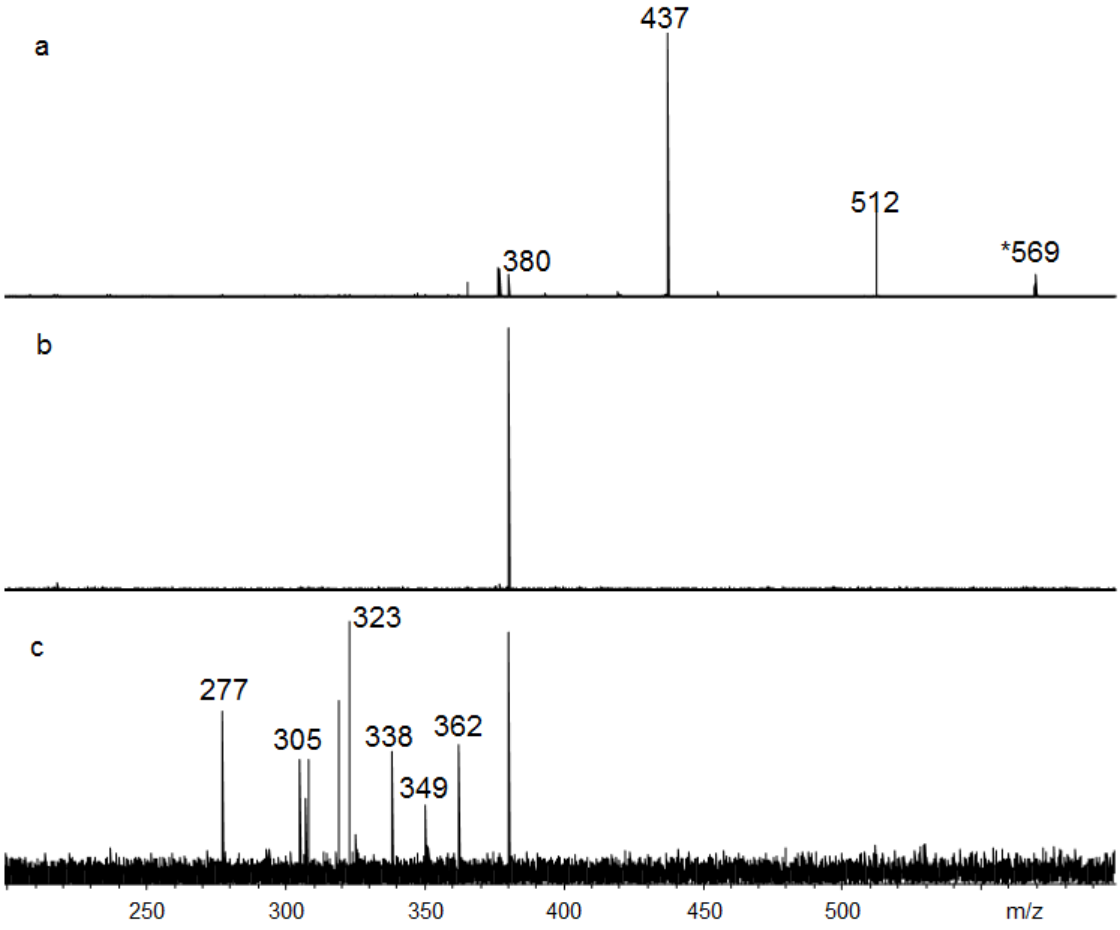


Figure S4.10. SORI-CID spectrum of $[Sr_2(GlyGly\text{-}H)_3]^+$ (*m/z* 569) (a) and the mass spectra following isolation (b) and dissociation (c) of fragment ion *m/z* 380 formed by SORI-CID of *m/z* 569.

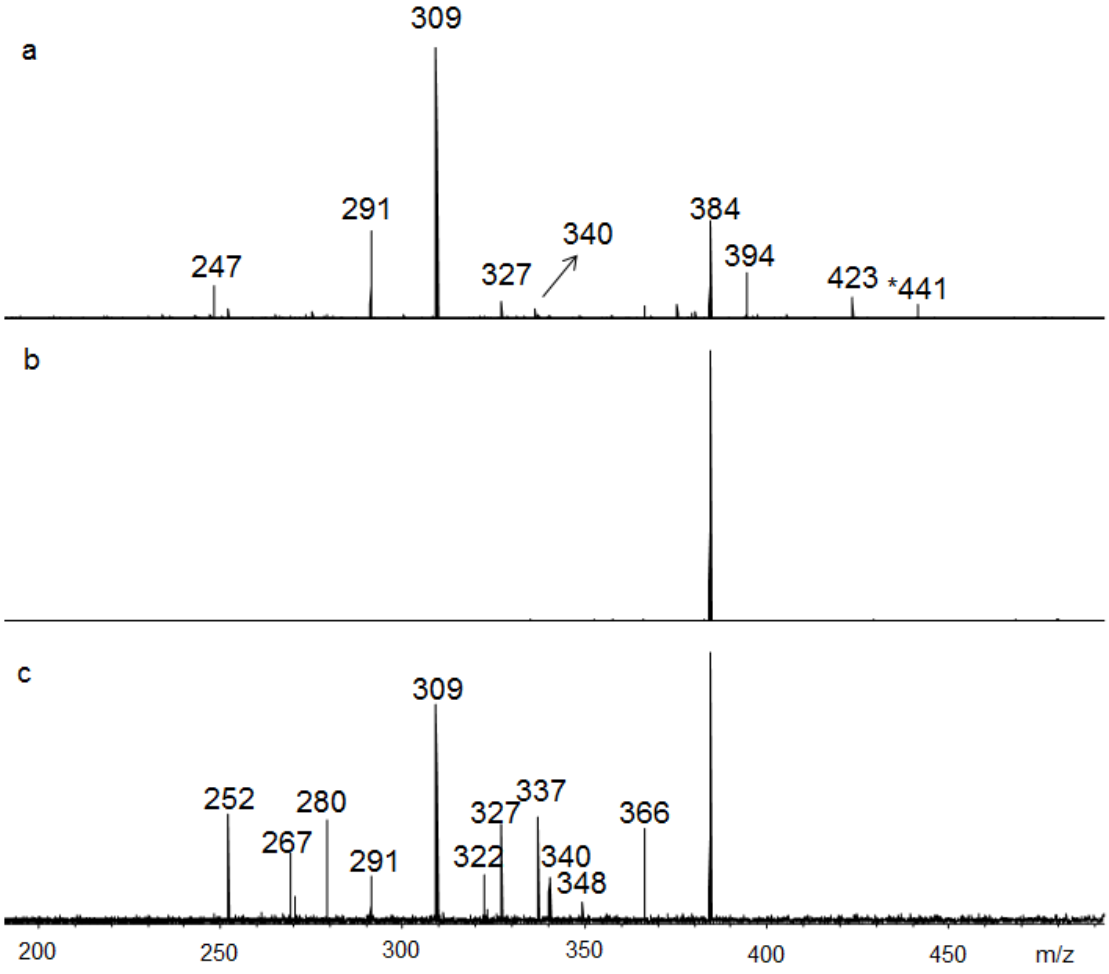


Figure S4.11. SORI-CID spectrum of $[Mg_2(GlyGly\text{-}H)_3]^+$ (*m/z* 441) (a) and the mass spectra following isolation (b) and dissociation (c) of fragment ion *m/z* 384 formed by SORI-CID of *m/z* 441.

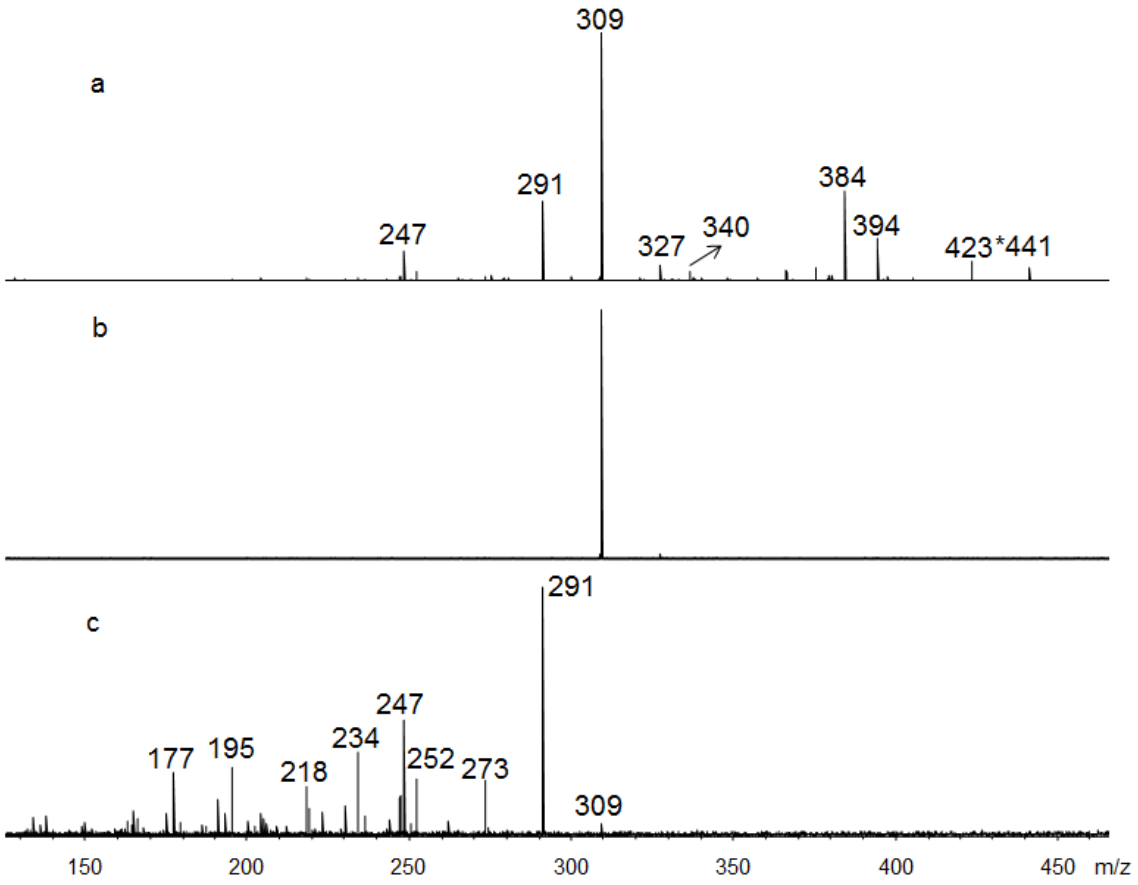


Figure S4.12. SORI-CID spectrum of $[Mg_2(GlyGly\text{-}H)_3]^+$ (*m/z* 441) (a) and the mass spectra following isolation (b) and dissociation (c) of fragment ion *m/z* 309 formed by SORI-CID of *m/z* 441.

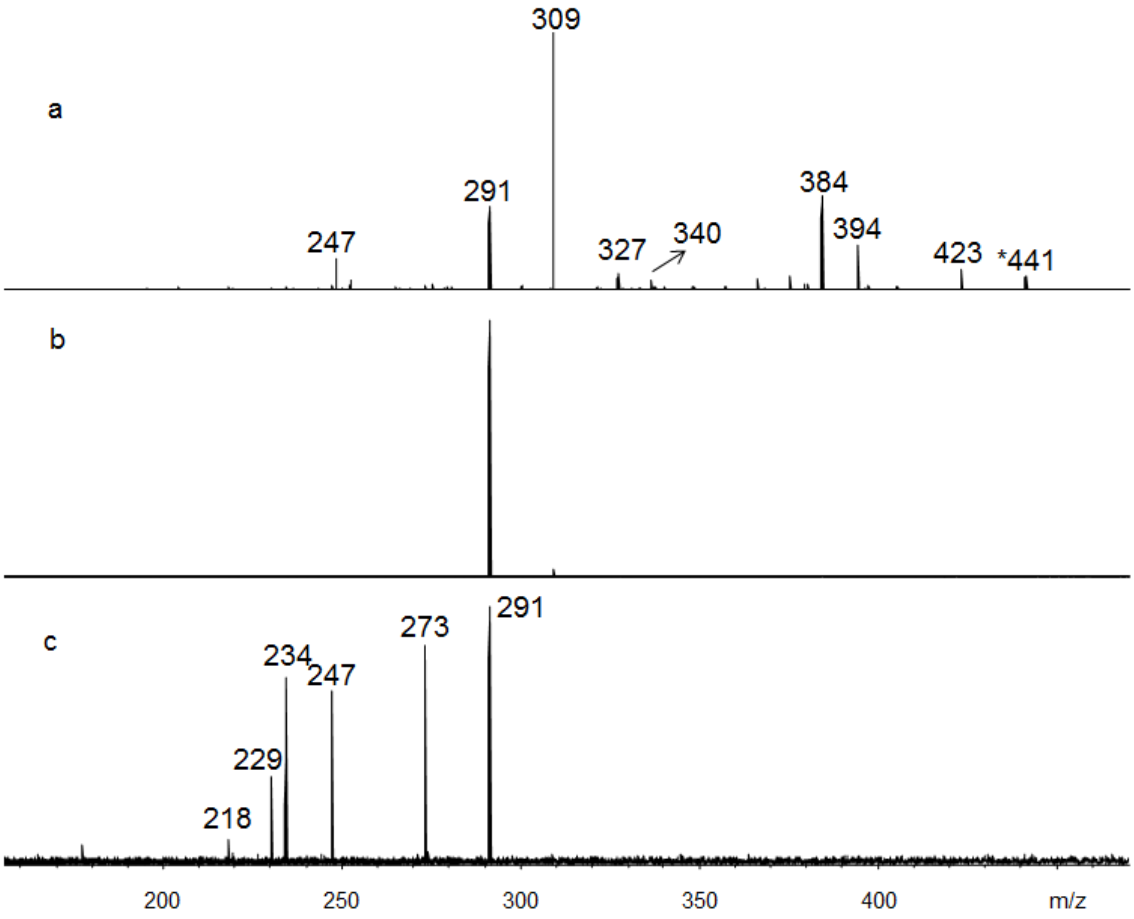


Figure S4.13. SORI-CID spectrum of $[Mg_2(GlyGly\text{-}H)_3]^+$ (*m/z* 441) (a) and the mass spectra following isolation (b) and dissociation (c) of fragment ion *m/z* 291 formed by SORI-CID of *m/z* 441.

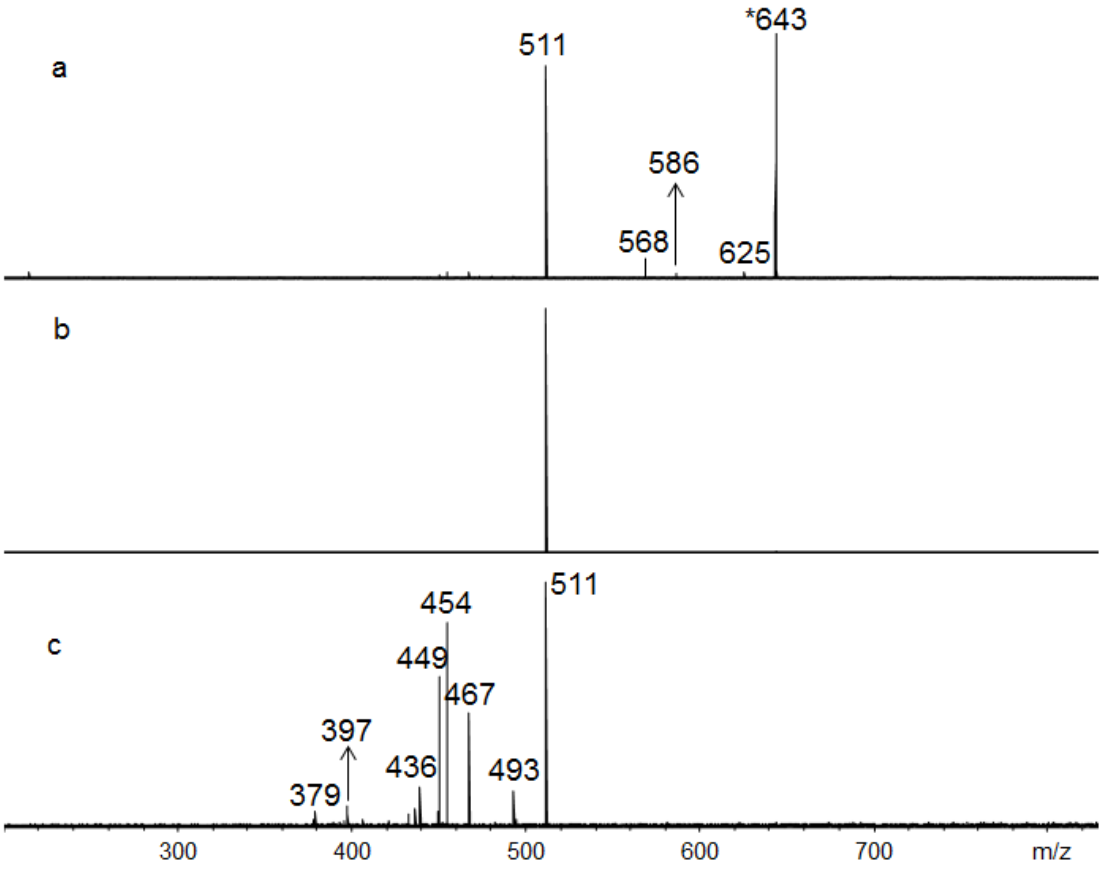


Figure S4.14. SORI-CID spectrum of $[Ca_3(GlyGly\text{-}H)_5]^+$ (*m/z* 775) at high collision energies (a) and the mass spectra following isolation (b) and dissociation (c) of fragment ion *m/z* 511 formed by SORI-CID of *m/z* 775.

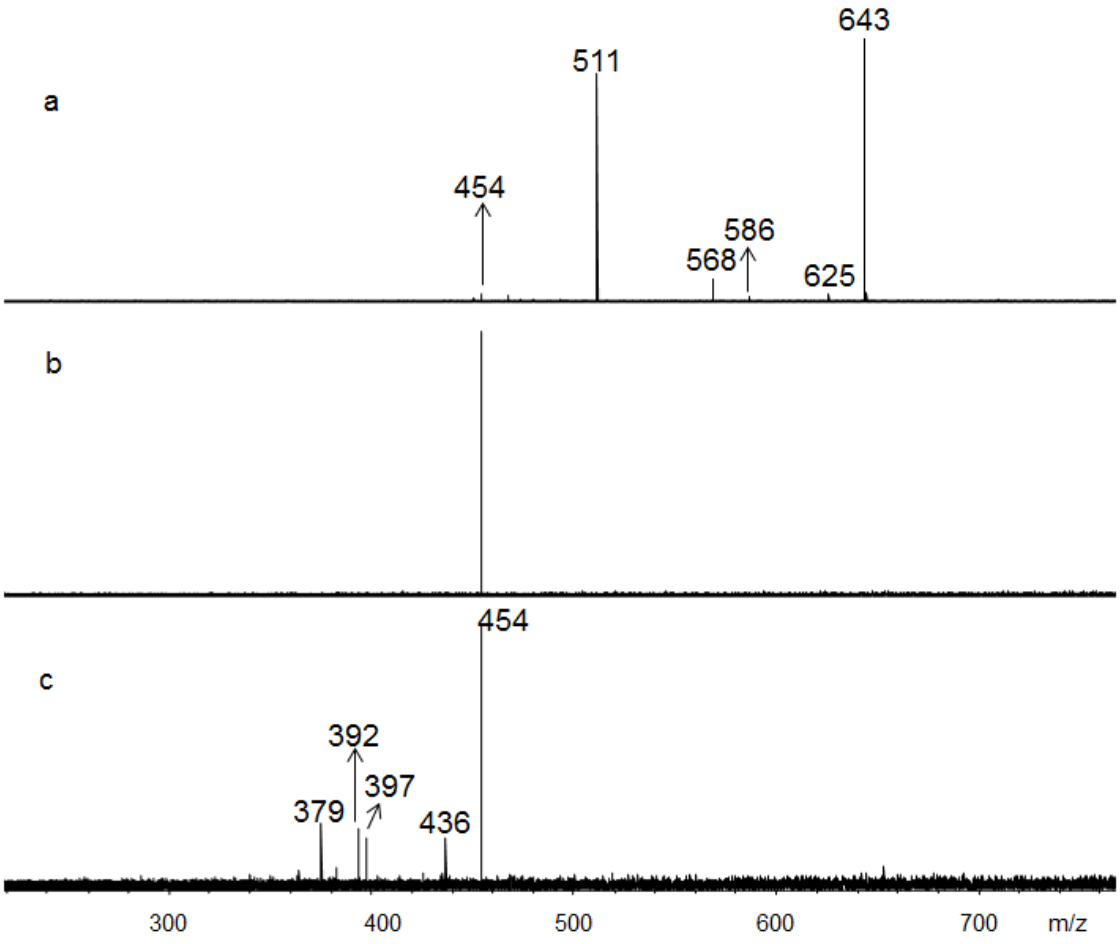


Figure S4.15. SORI-CID spectrum of $[Ca_3(GlyGly\text{-}H)_5]^+$ (*m/z* 775) at high collision energies (a) and the mass spectra following isolation (b) and dissociation (c) of fragment ion *m/z* 454 formed by SORI-CID of *m/z* 775.

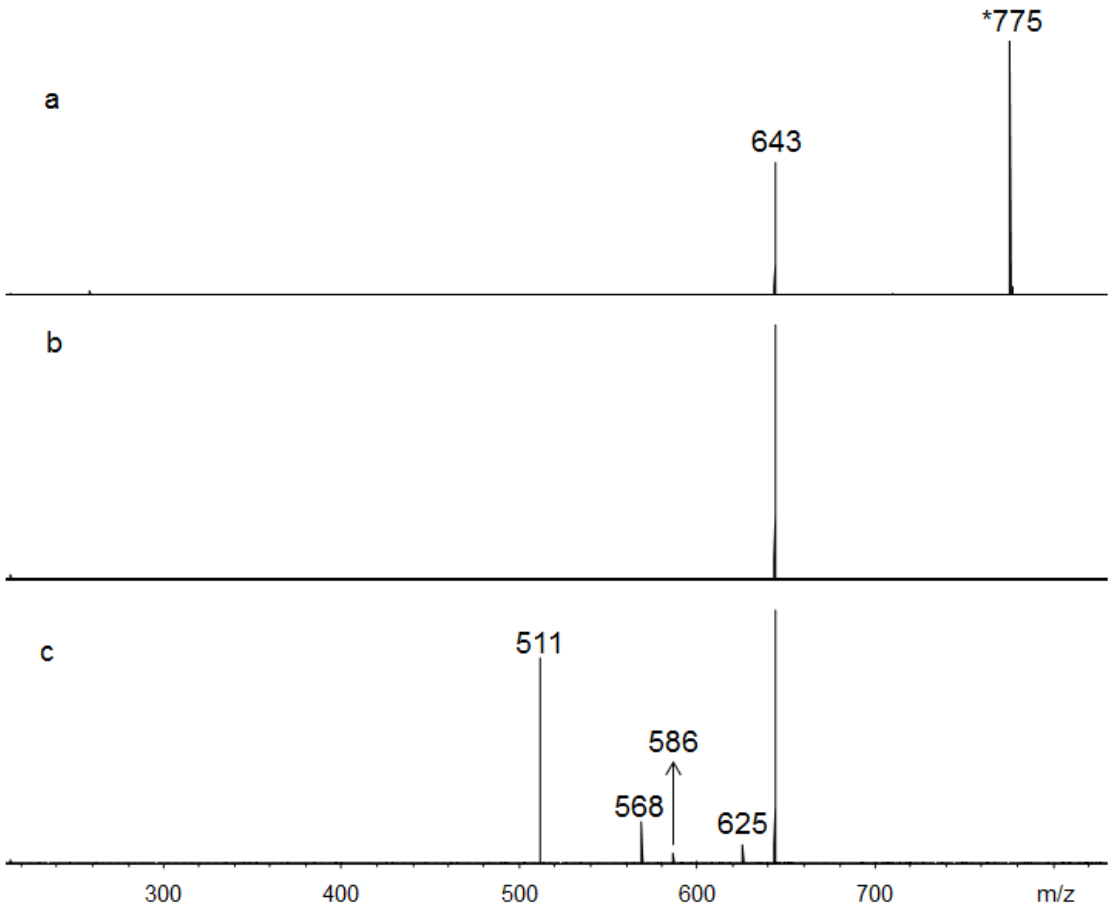


Figure S4.16. SORI-CID spectrum of $[Ca_3(GlyGly\text{-}H)_5]^+$ (*m/z* 775) (a) and the mass spectra following isolation (b) and dissociation (c) of fragment ion *m/z* 643 formed by SORI-CID of *m/z* 775.

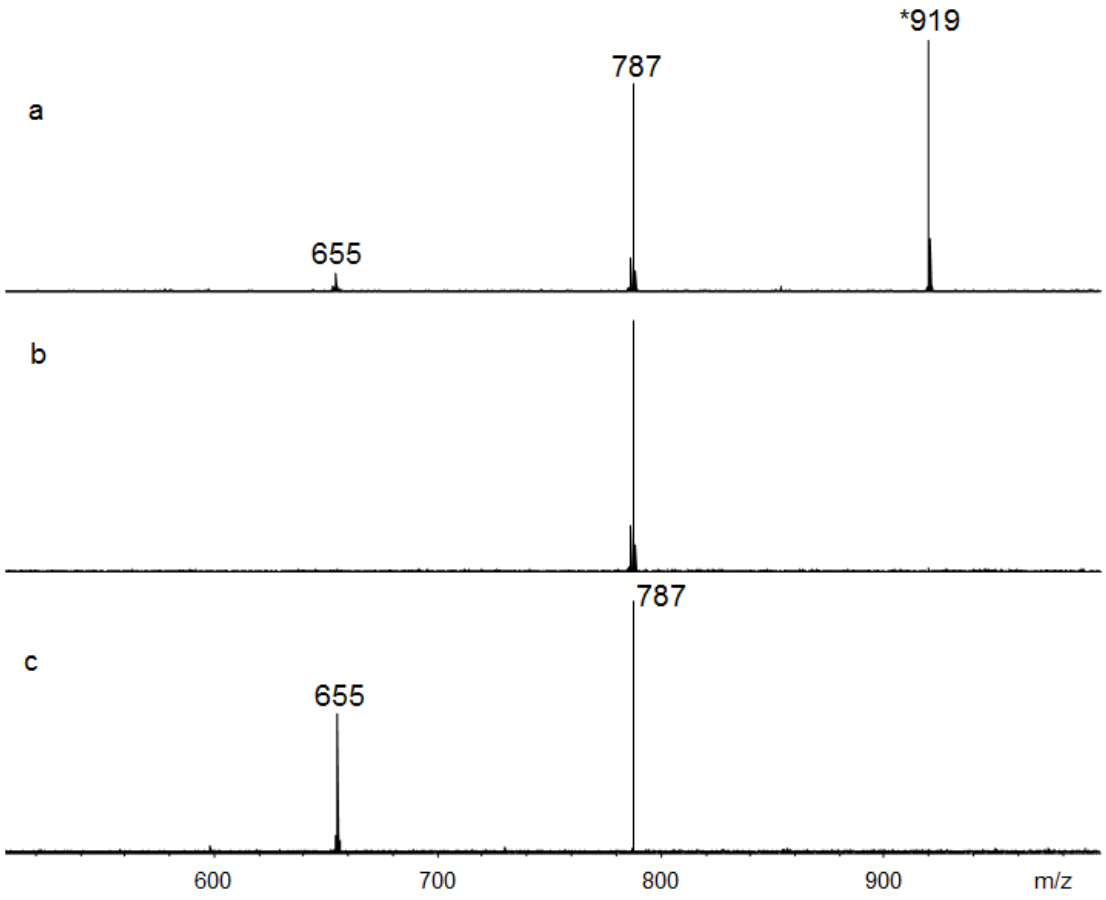


Figure S4.17. SORI-CID spectrum of $[Sr_3(GlyGly\text{-}H)_5]^+$ (*m/z* 919) (a) and the mass spectra following isolation (b) and dissociation (c) of fragment ion *m/z* 787 formed by SORI-CID of *m/z* 919.

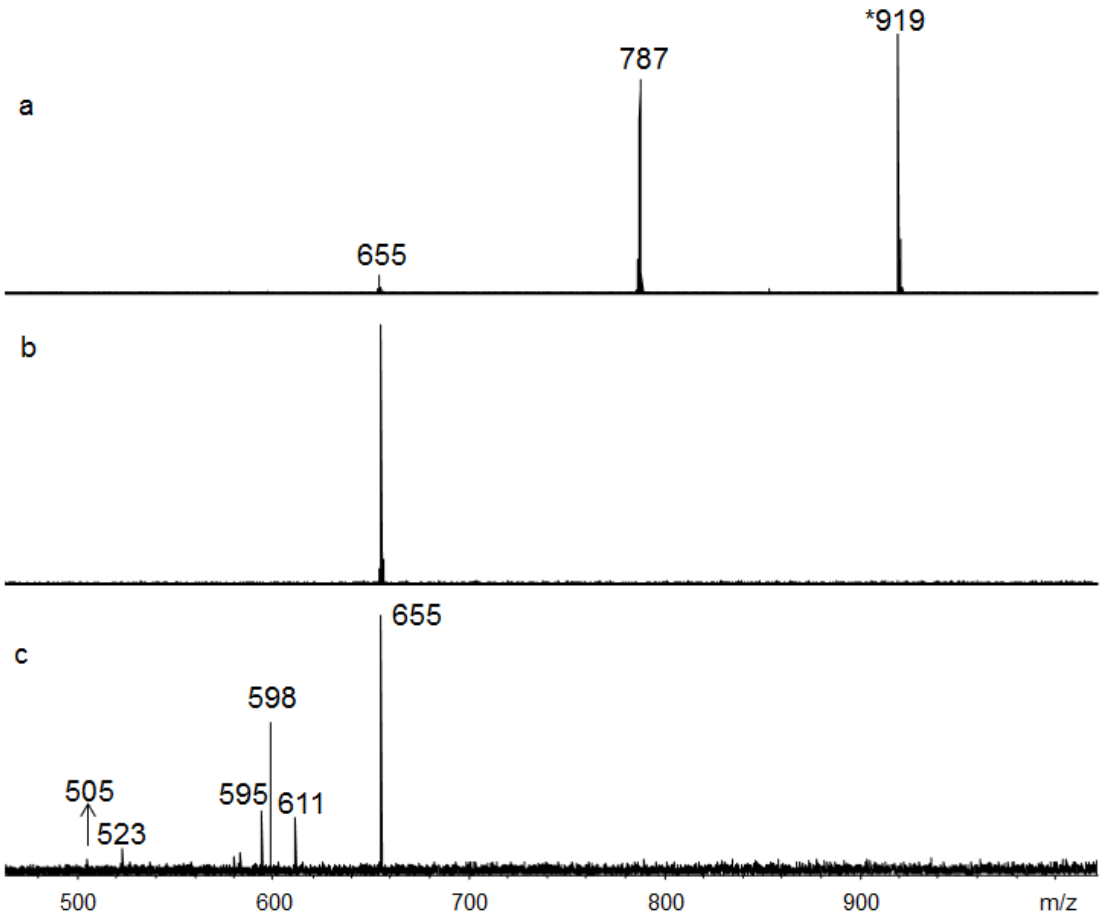


Figure S4.18. SORI-CID spectrum of $[Sr_3(GlyGly\text{-}H)_5]^+$ (*m/z* 919) (a) and the mass spectra following isolation (b) and dissociation (c) of fragment ion *m/z* 655 formed by SORI-CID of *m/z* 919.

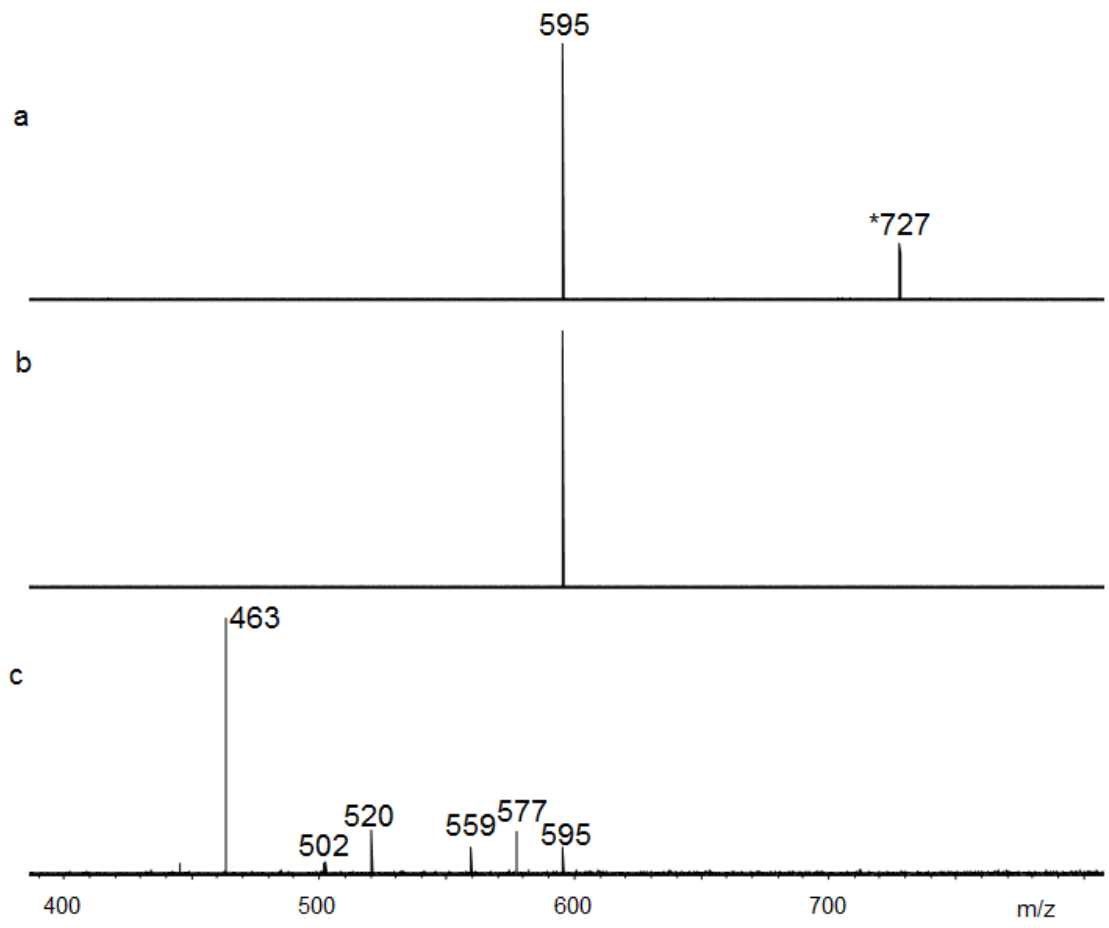


Figure S4.19. SORI-CID spectrum of $[Mg_3(GlyGly\text{-}H)_5]^+$ (*m/z* 727) (a) and the mass spectra following isolation (b) and dissociation (c) of fragment ion *m/z* 595 formed by SORI-CID of *m/z* 727.

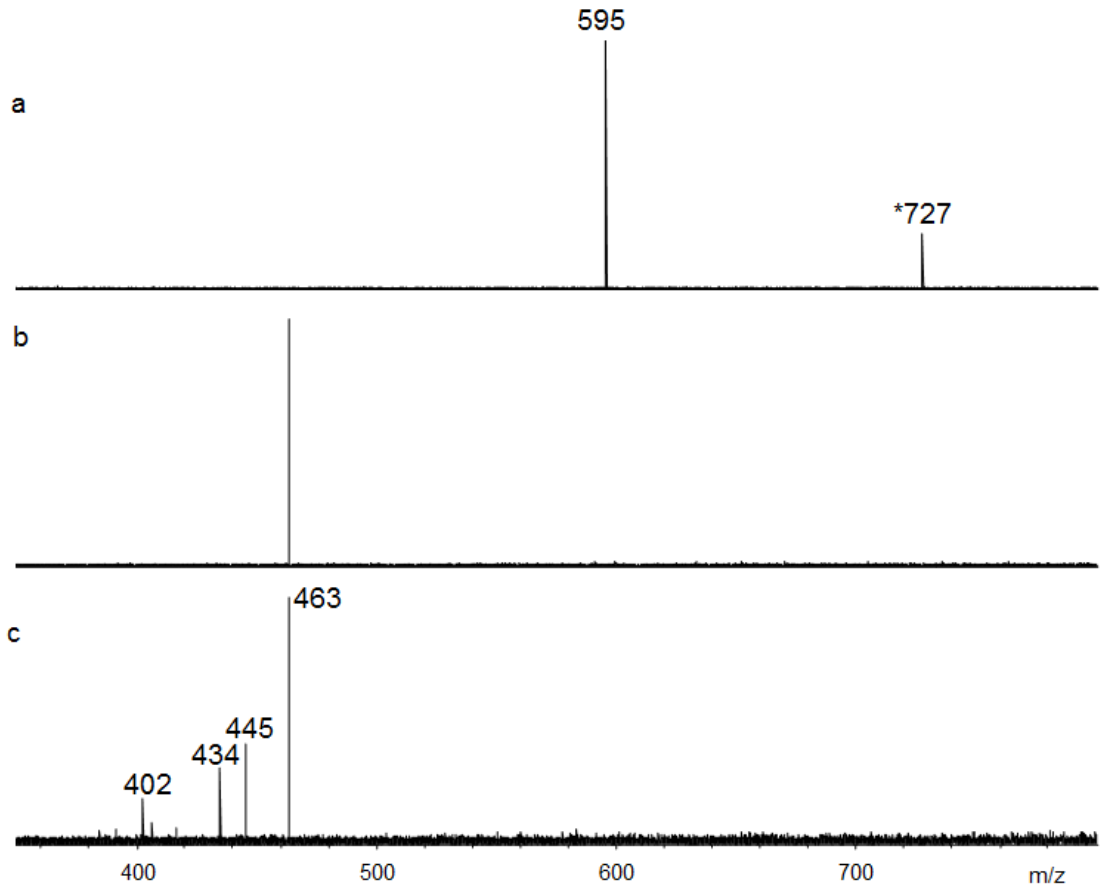


Figure S4.20. SORI-CID spectrum of $[Mg_3(GlyGly\text{-}H)_5]^+$ (*m/z* 727) (a) and the mass spectra following isolation (b) and dissociation (c) of fragment ion *m/z* 463 formed by SORI-CID of *m/z* 727.

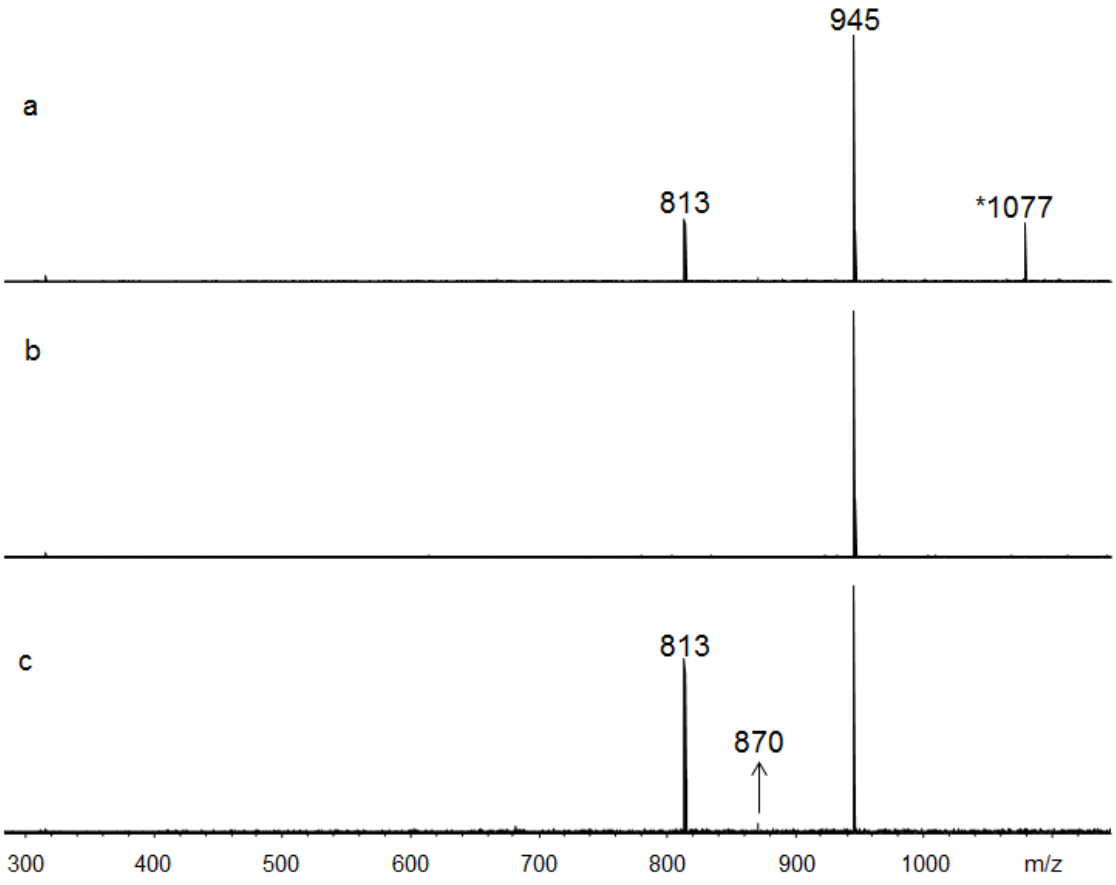


Figure S4.21. SORI-CID spectrum of $[Ca_4(GlyGly\text{-}H)_7]^+$ (*m/z* 1077) (a) and the mass spectra following isolation (b) and dissociation (c) of fragment ion *m/z* 945 formed by SORI-CID of *m/z* 1077.

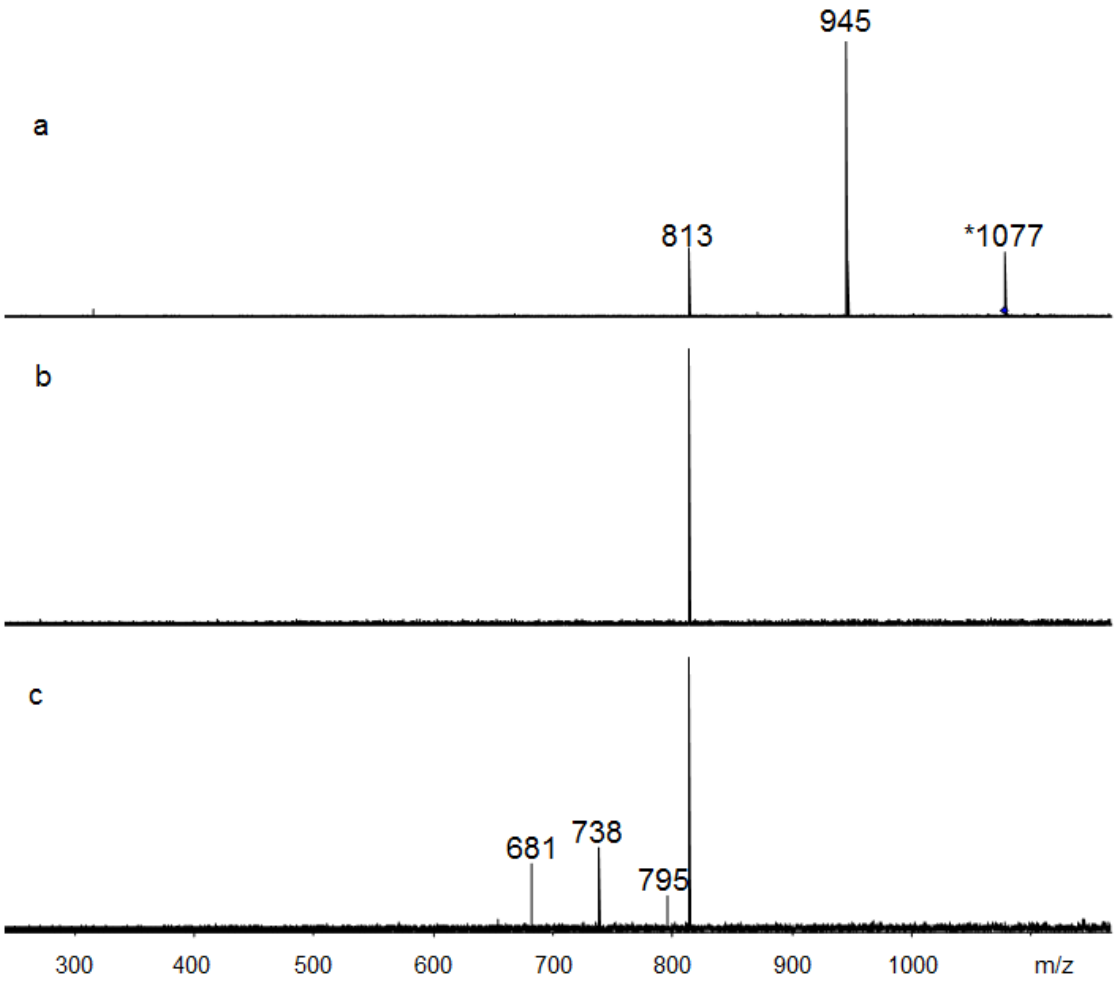


Figure S4.22. SORI-CID spectrum of $[Ca_4(GlyGly\text{-}H)_7]^+$ (*m/z* 1077) (a) and the mass spectra following isolation (b) and dissociation (c) of fragment ion *m/z* 813 formed by SORI-CID of *m/z* 1077.

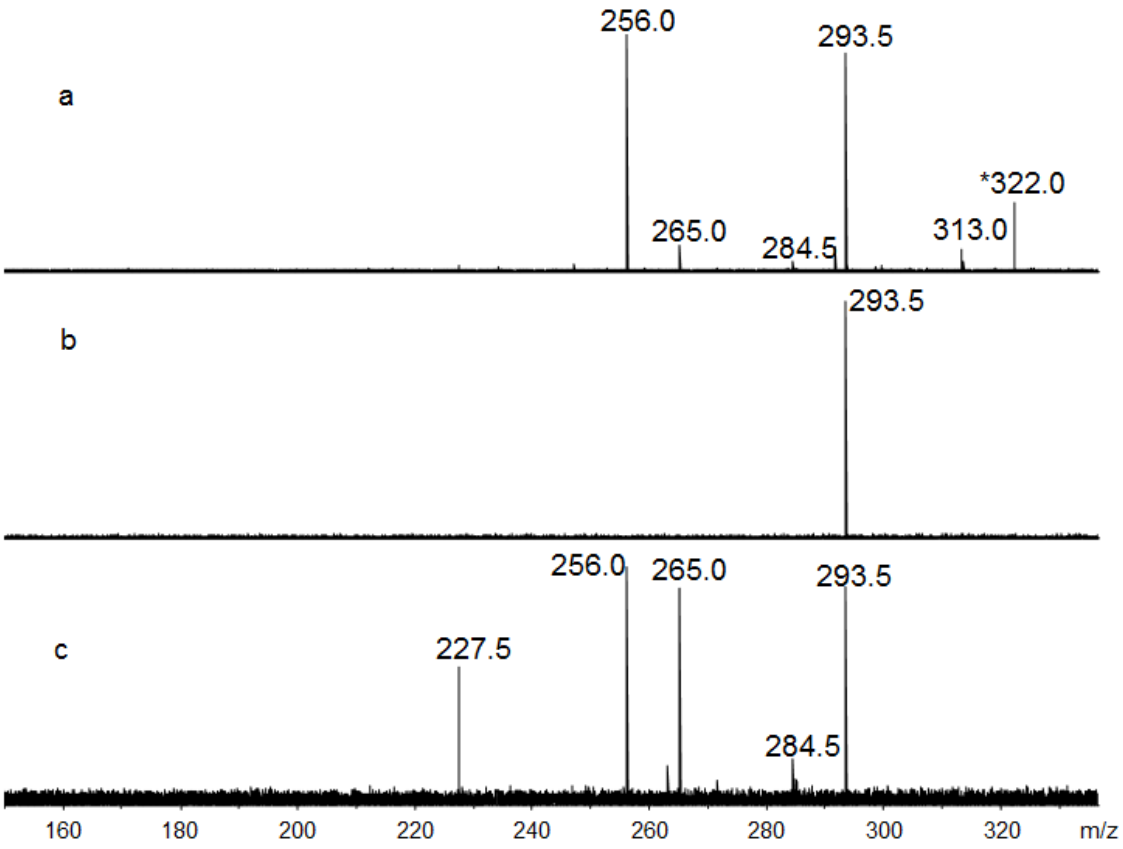


Figure S4.23. SORI-CID spectrum of $[Ca_3(GlyGly\text{-}H)_4]^{2+}$ (*m/z* 322.0) (a) and the mass spectra following isolation (b) and dissociation (c) of fragment ion *m/z* 293.5 formed by SORI-CID of *m/z* 322.0.

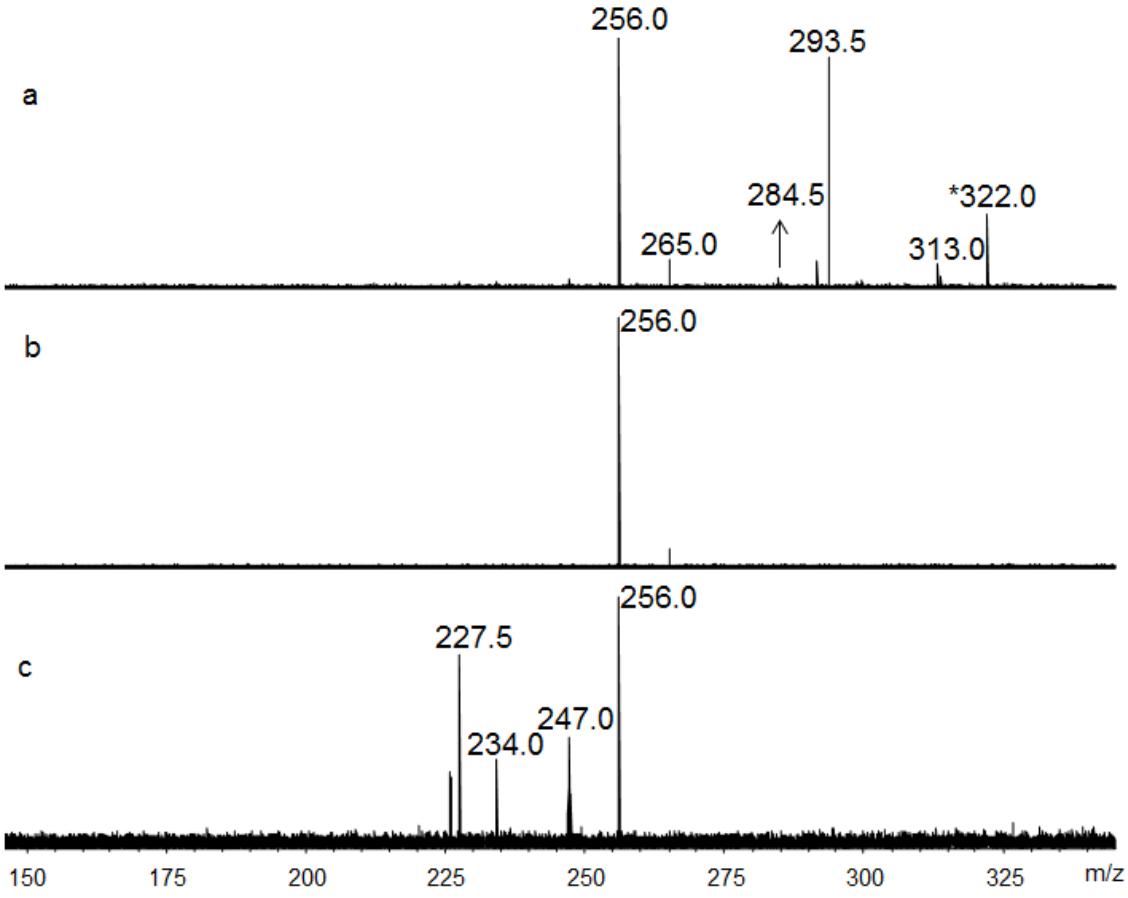


Figure S4.24. SORI-CID spectrum of $[Ca_3(GlyGly\text{-}H)_4]^{2+}$ (*m/z* 322.0) (a) and the mass spectra following isolation (b) and dissociation (c) of fragment ion *m/z* 256.0 formed by SORI-CID of *m/z* 322.0.

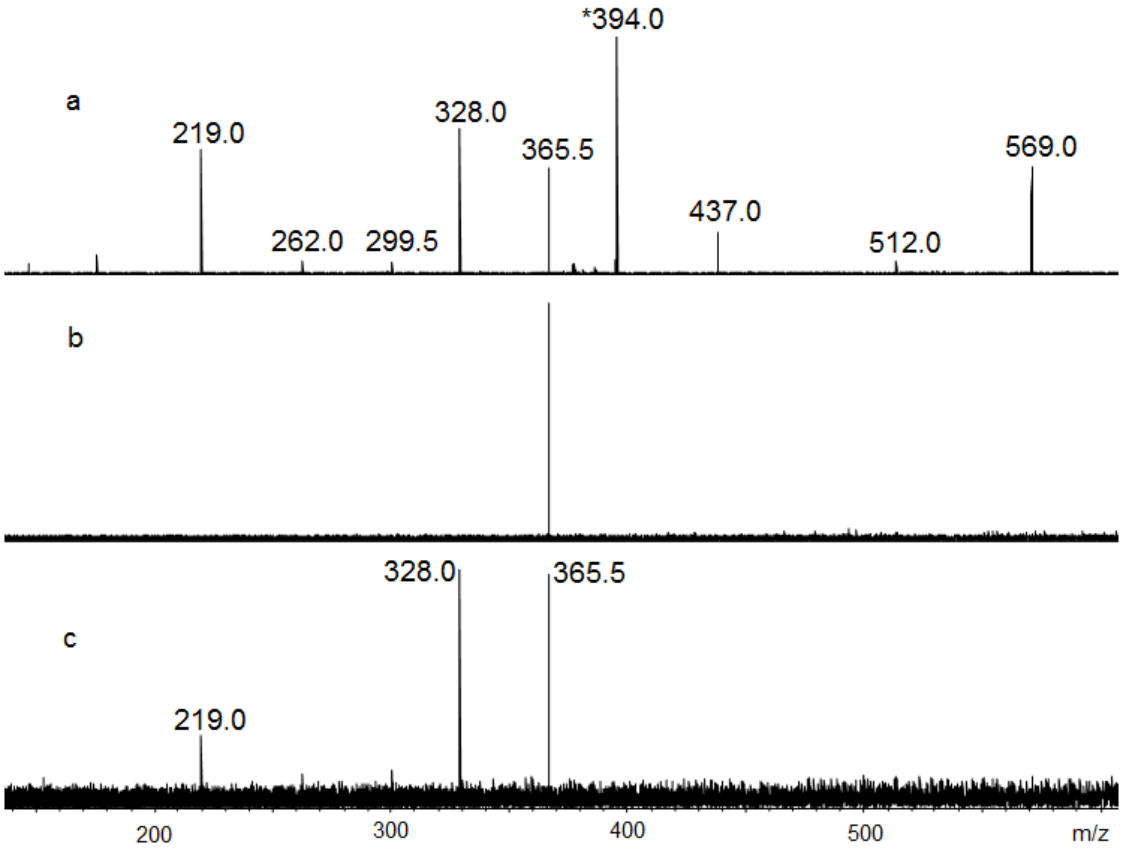


Figure S4.25. SORI-CID spectrum of $[Sr_3(GlyGly\text{-}H)_4]^{2+}$ (*m/z* 394.0) (a) and the mass spectra following isolation (b) and dissociation (c) of fragment ion *m/z* 365.5 formed by SORI-CID of *m/z* 394.0.

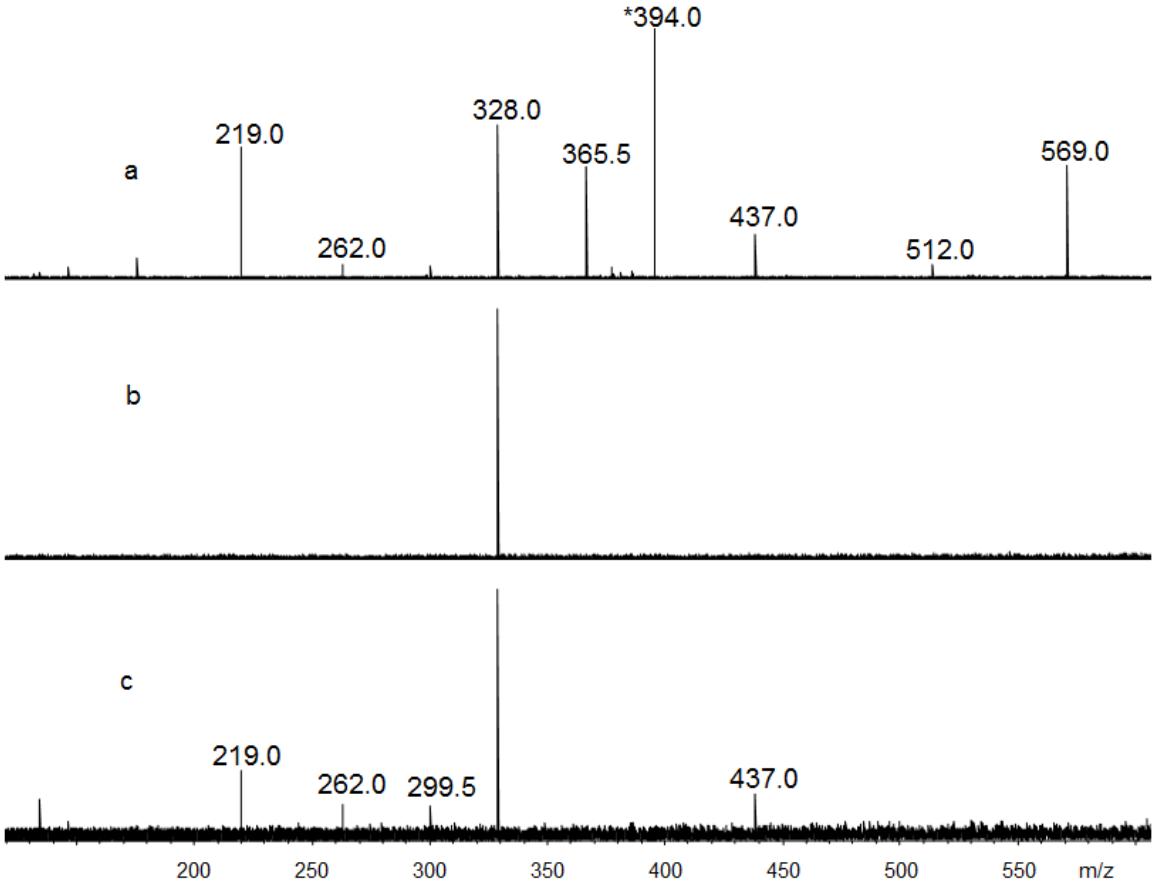


Figure S4.26. SORI-CID spectrum of $[Sr_3(GlyGly\text{-}H)_4]^{2+}$ (*m/z* 394.0) (a) and the mass spectra following isolation (b) and dissociation (c) of fragment ion *m/z* 328.0 formed by SORI-CID of *m/z* 394.0.

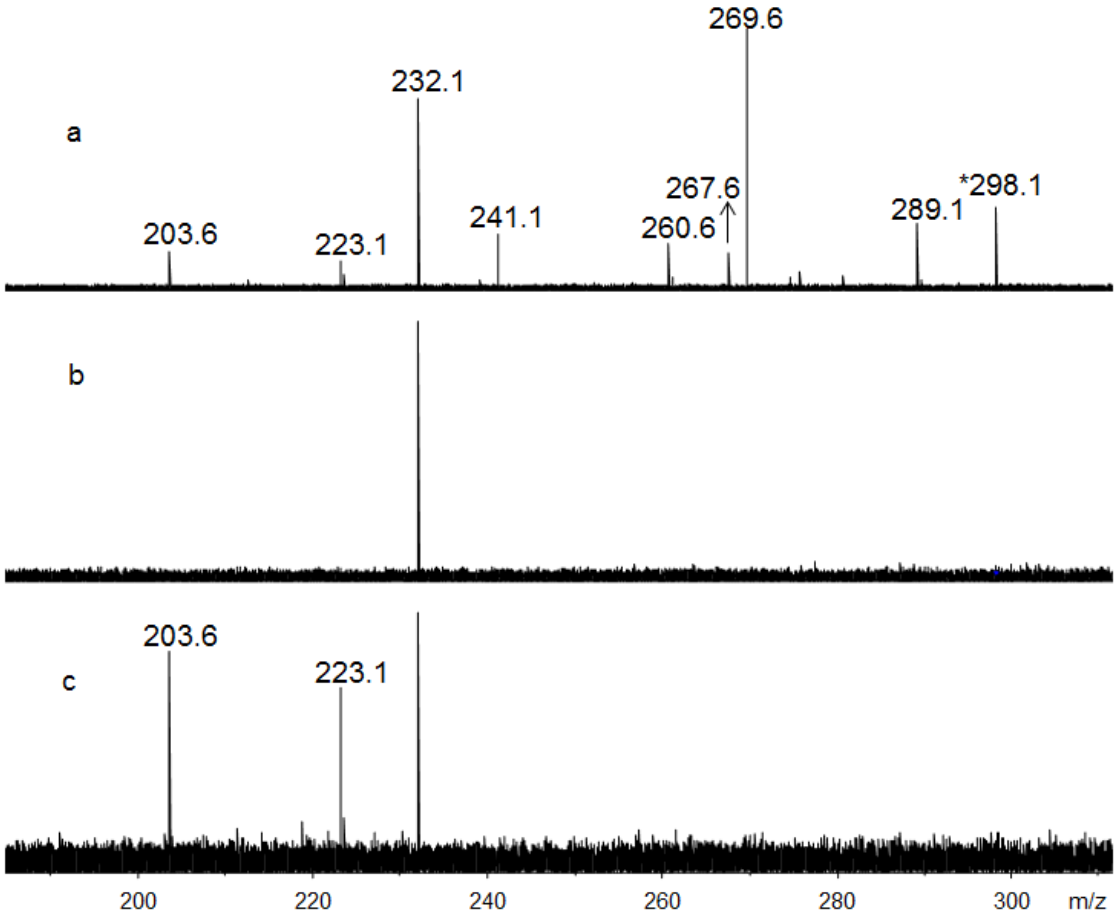


Figure S4.27. SORI-CID spectrum of $[Mg_3(GlyGly\text{-}H)_4]^{2+}$ (*m/z* 298.1) (a) and the mass spectra following isolation (b) and dissociation (c) of fragment ion *m/z* 232.1 formed by SORI-CID of *m/z* 298.1.

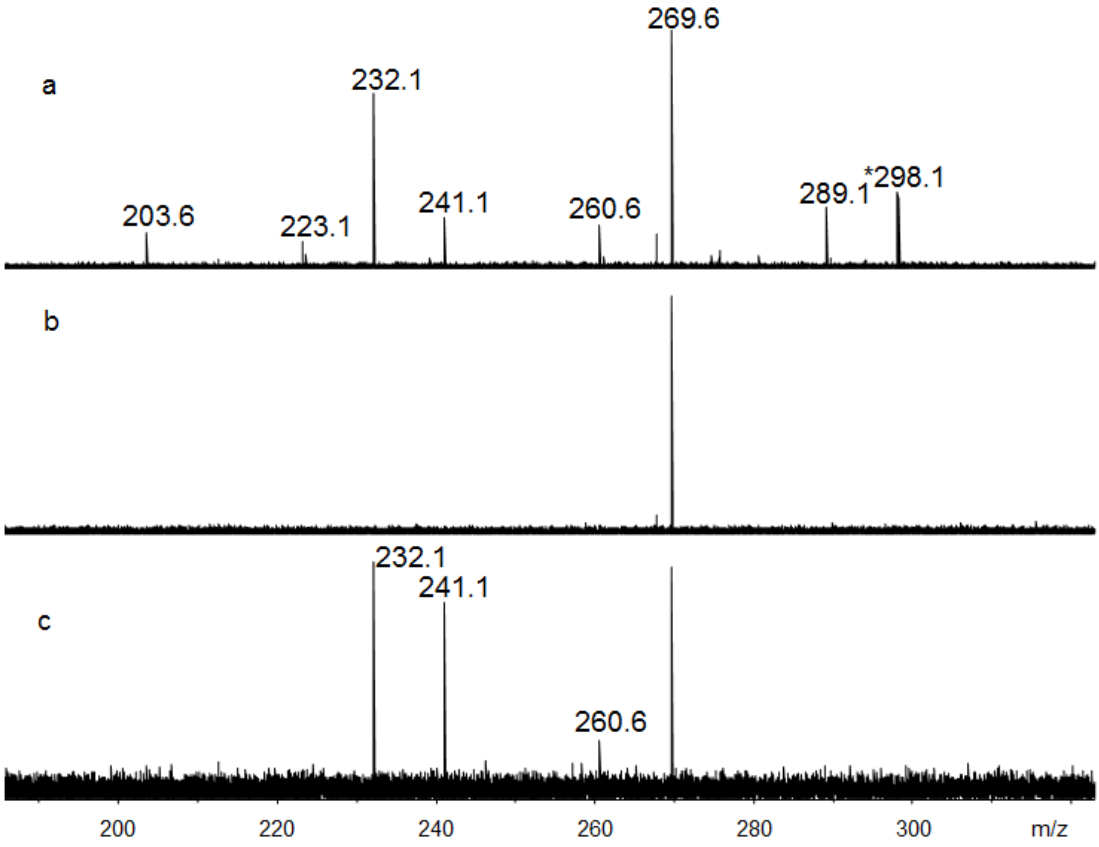


Figure S4.28. SORI-CID spectrum of $[Ca_3(GlyGly\text{-}H)_4]^{2+}$ (*m/z* 298.1) (a) and the mass spectra following isolation (b) and dissociation (c) of fragment ion *m/z* 269.6.1 formed by SORI-CID of *m/z* 298.1.

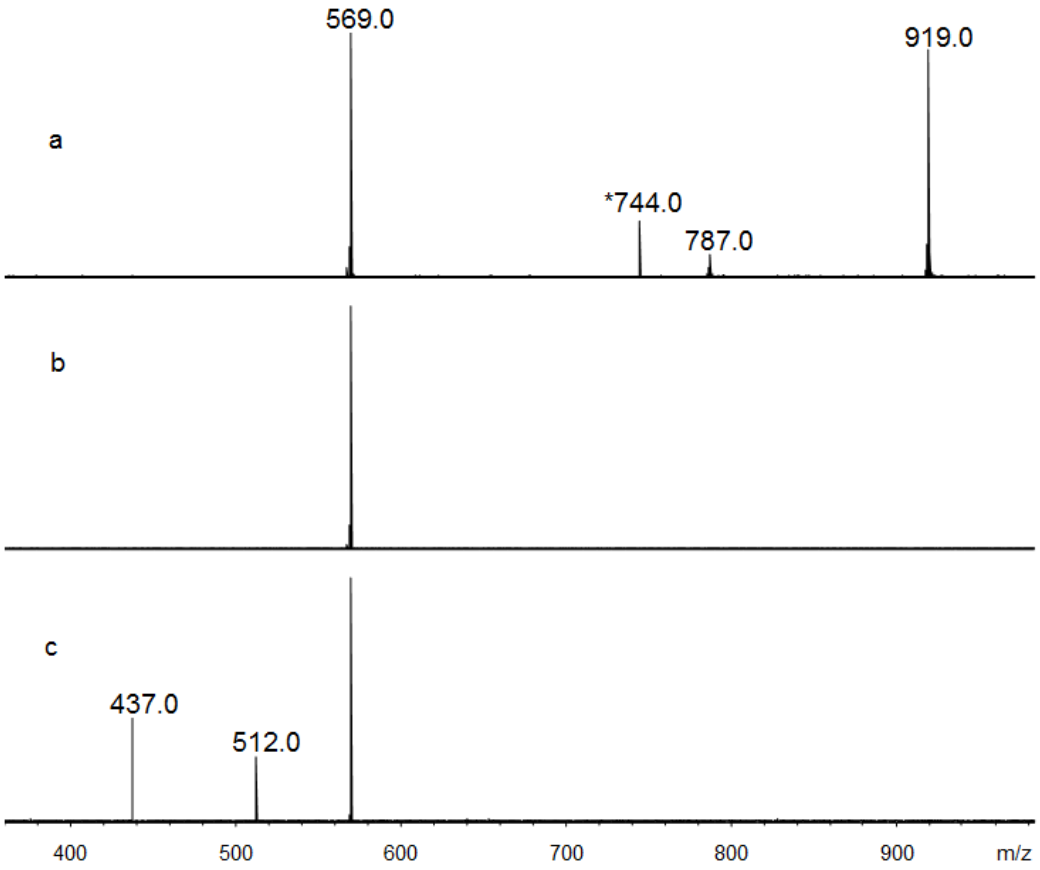


Figure S4.29. SORI-CID spectrum of $[Sr_5(GlyGly\text{-}H)_8]^{2+}$ (*m/z* 744.0) (a) and the mass spectra following isolation (b) and dissociation (c) of fragment ion *m/z* 569.0 formed by SORI-CID of *m/z* 744.0.

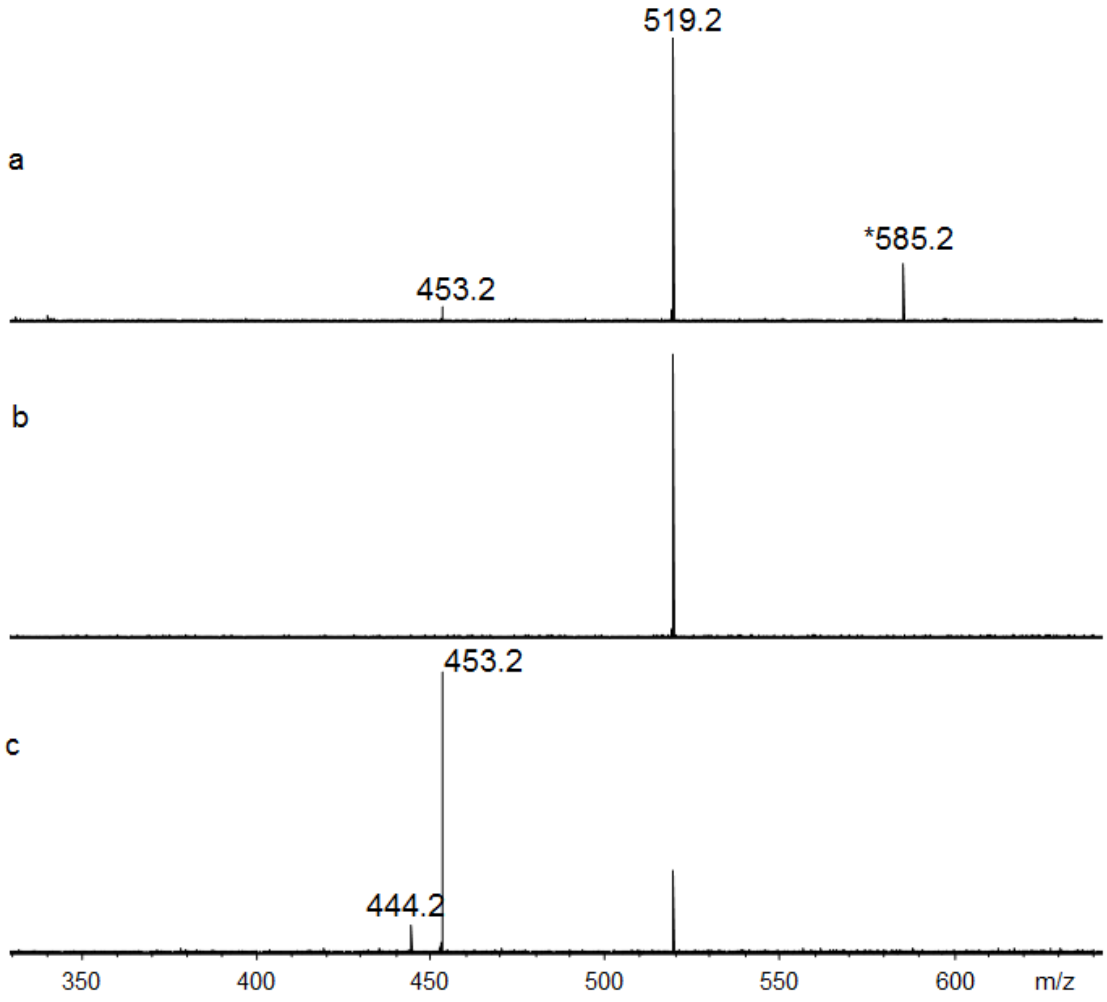


Figure S4.30. SORI-CID spectrum of $[Mg_5(GlyGly\text{-}H)_8]^{2+}$ (*m/z* 585.2) (a) and the mass spectra following isolation (b) and dissociation (c) of fragment ion *m/z* 519.2 formed by SORI-CID of *m/z* 585.2.

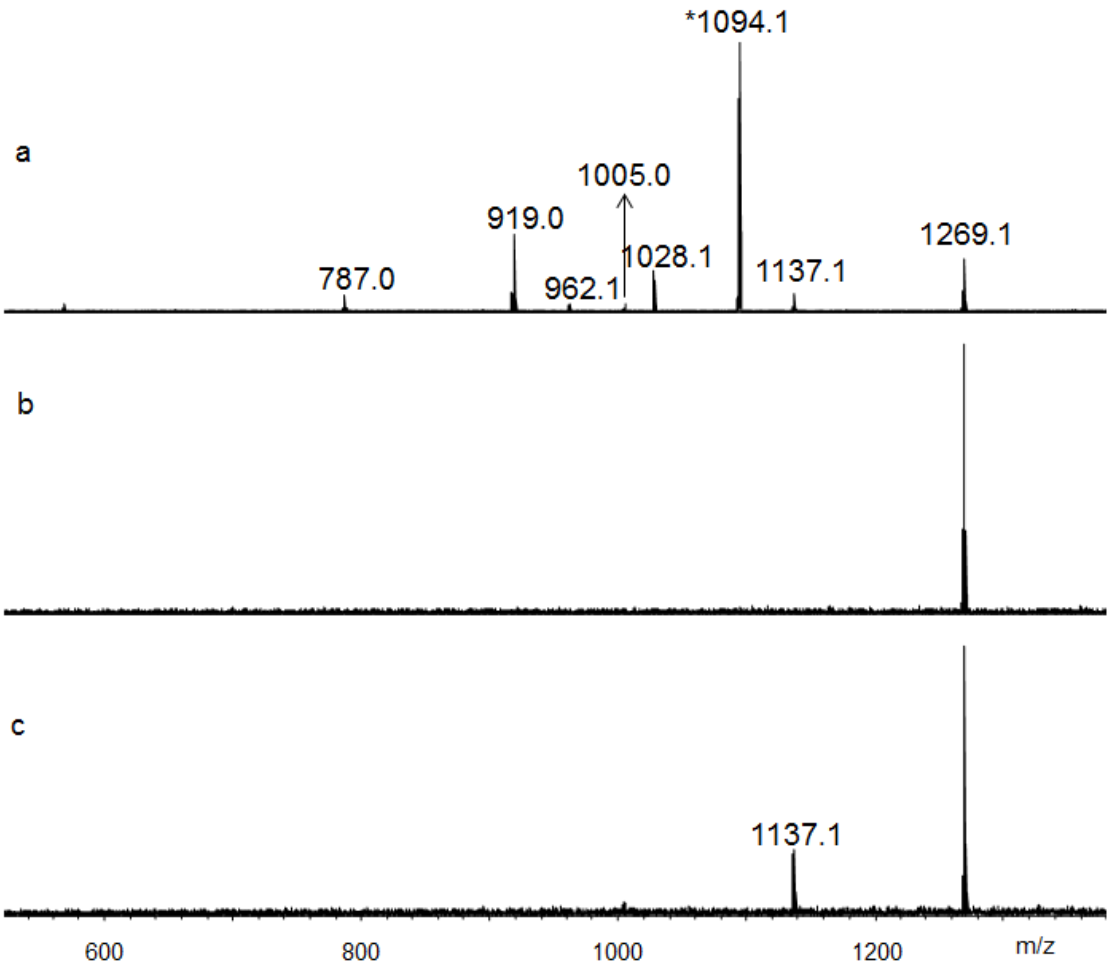


Figure S4.31. SORI-CID spectrum of $[Sr_7(GlyGly\text{-}H)_{12}]^{2+}$ (*m/z* 1094.1) (a) and the mass spectra following isolation (b) and dissociation (c) of fragment ion *m/z* 1269.1 formed by SORI-CID of *m/z* 1094.1.

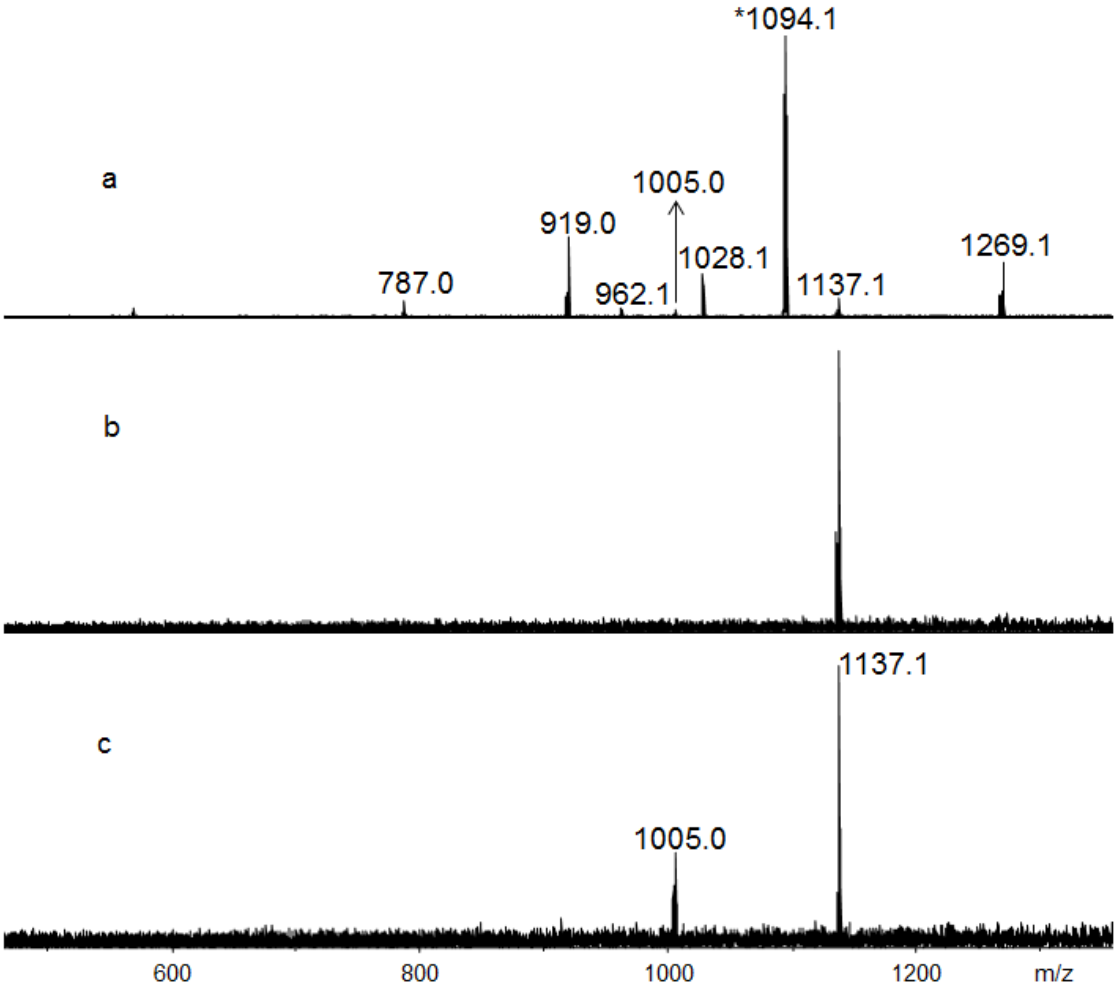


Figure S4.32. SORI-CID spectrum of $[Sr_7(GlyGly\text{-}H)_{12}]^{2+}$ (*m/z* 1094.1) (a) and the mass spectra following isolation (b) and dissociation (c) of fragment ion *m/z* 1137.1 formed by SORI-CID of *m/z* 1094.1.

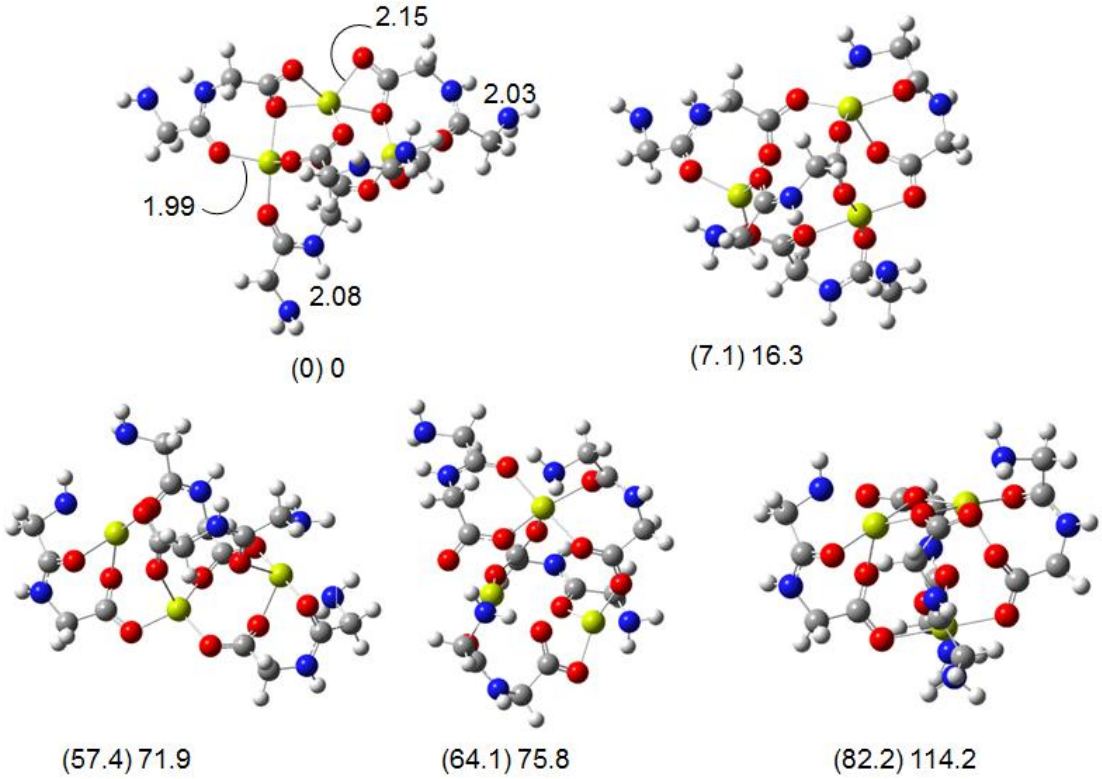


Figure S4.33. Optimized structures for $[Mg_3(GlyGly\text{-}H)_4]^{2+}$ cluster. 298 K relative enthalpies (in parentheses) and Gibbs energies are reported in kJ mol^{-1} using B3LYP/6-311++G(3df,3pd)//B3LYP/6-31+G(d,p) level of theory and basis set.

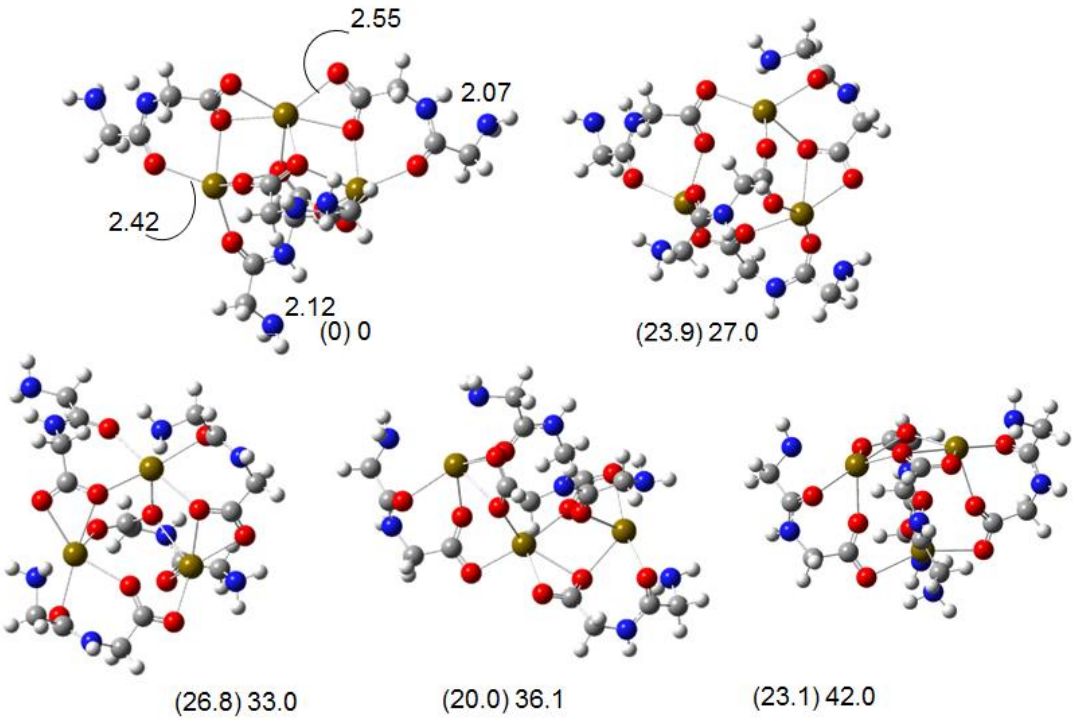


Figure S4.34. Optimized structures for $[Sr_3(GlyGly\text{-}H)_4]^{2+}$ cluster. 298 K relative enthalpies (in parenthesis) and Gibbs energies are reported in kJ mol^{-1} using B3LYP level of theory with 6-31+G(d,p) basis set on C, H, N, and O and Def2SVP basis set on Sr^{2+}

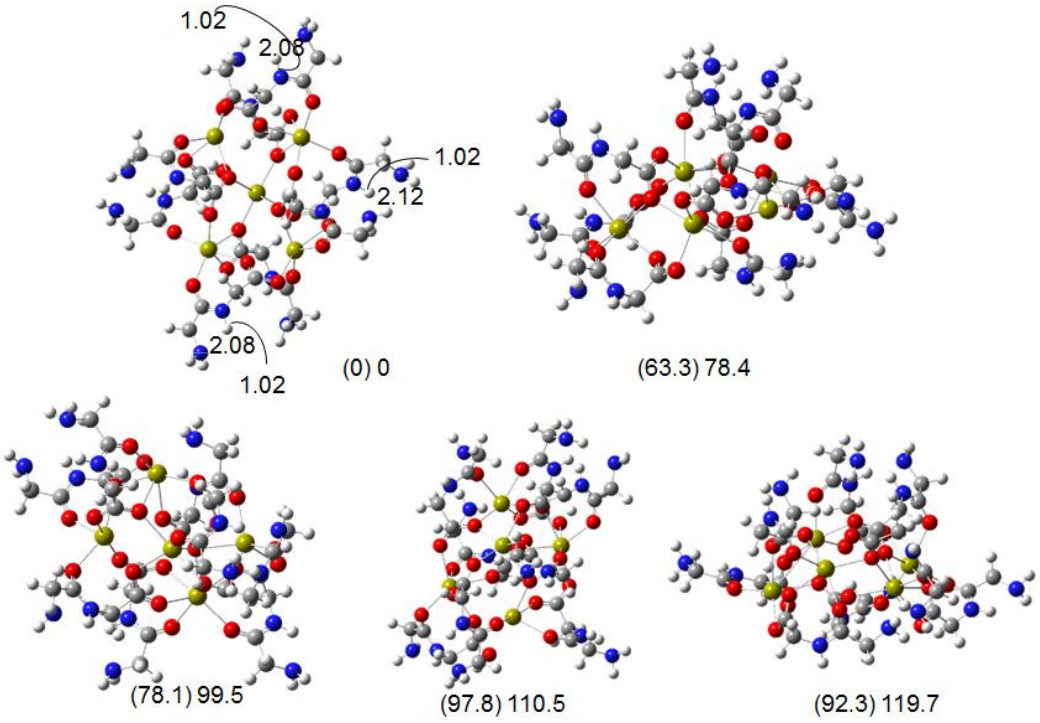


Figure S4.35. Optimized structures for $[Ca_5(GlyGly\text{-}H)_8]^{2+}$ cluster. 298 K relative enthalpies (in parenthesis) and Gibbs energies are reported in kJ mol^{-1} using B3LYP level of theory and SVP basis set.

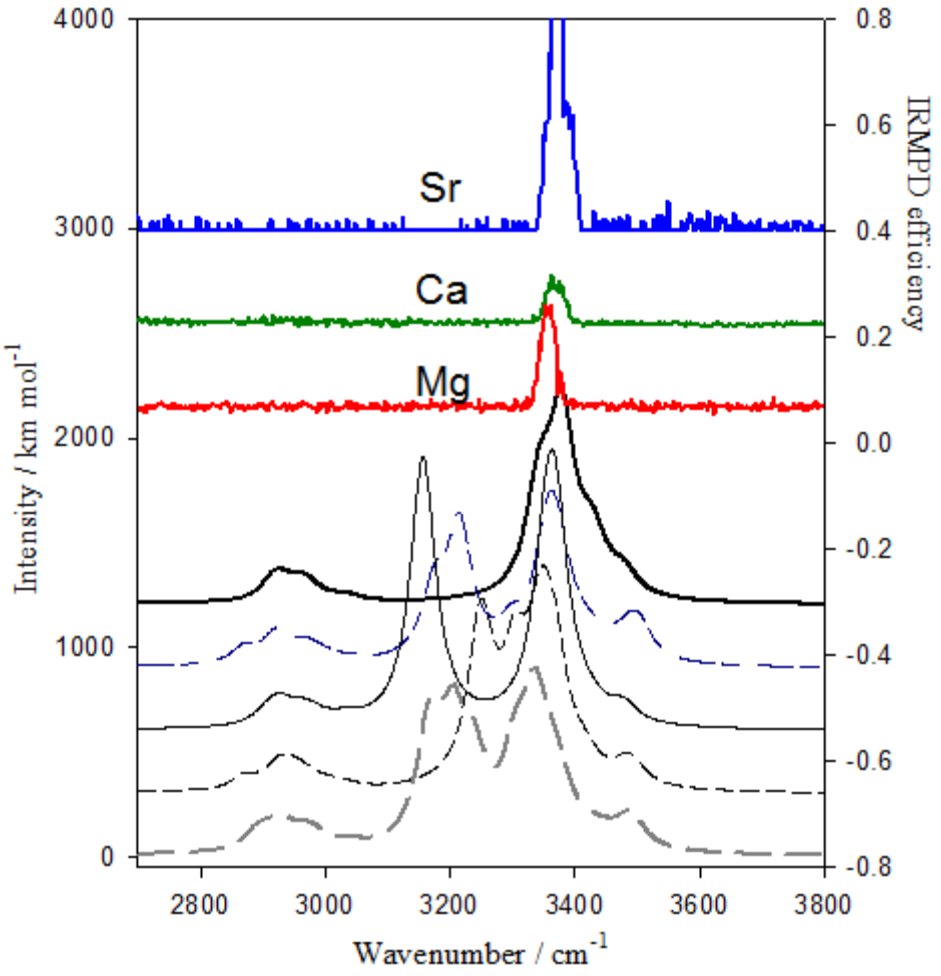


Figure S4.36. Comparison of the IRMPD spectra for $[M_5(GlyGly\text{-}H)_8]^{2+}$ ($M=Mg^{2+}$, Ca^{2+} and Sr^{2+}) and the computed spectra for five lowest energy structures of $[Ca_5(GlyGly\text{-}H)_8]^{2+}$ calculated at 298 K using B3LYP/SVP. IR frequencies are computed using B3LYP/SVP method and basis set.

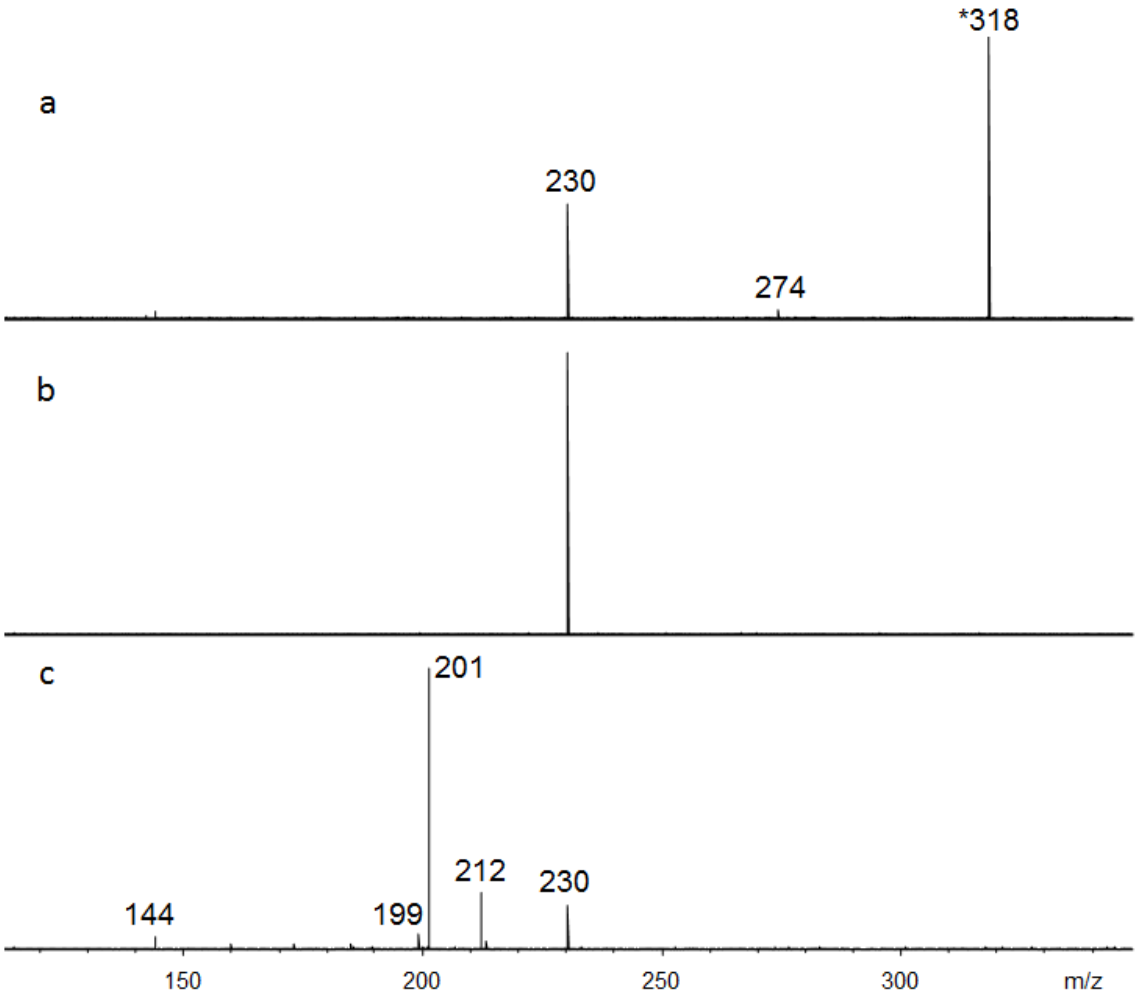


Figure S5.1. SORI-CID mass spectrum of $[Mn(GlyGly-H)(GlyGly)]^+$ (*m/z* 318) (a) and the mass spectra following isolation (b) and dissociation (c) of fragment ion *m/z* 230 formed by SORI-CID of *m/z* 318.

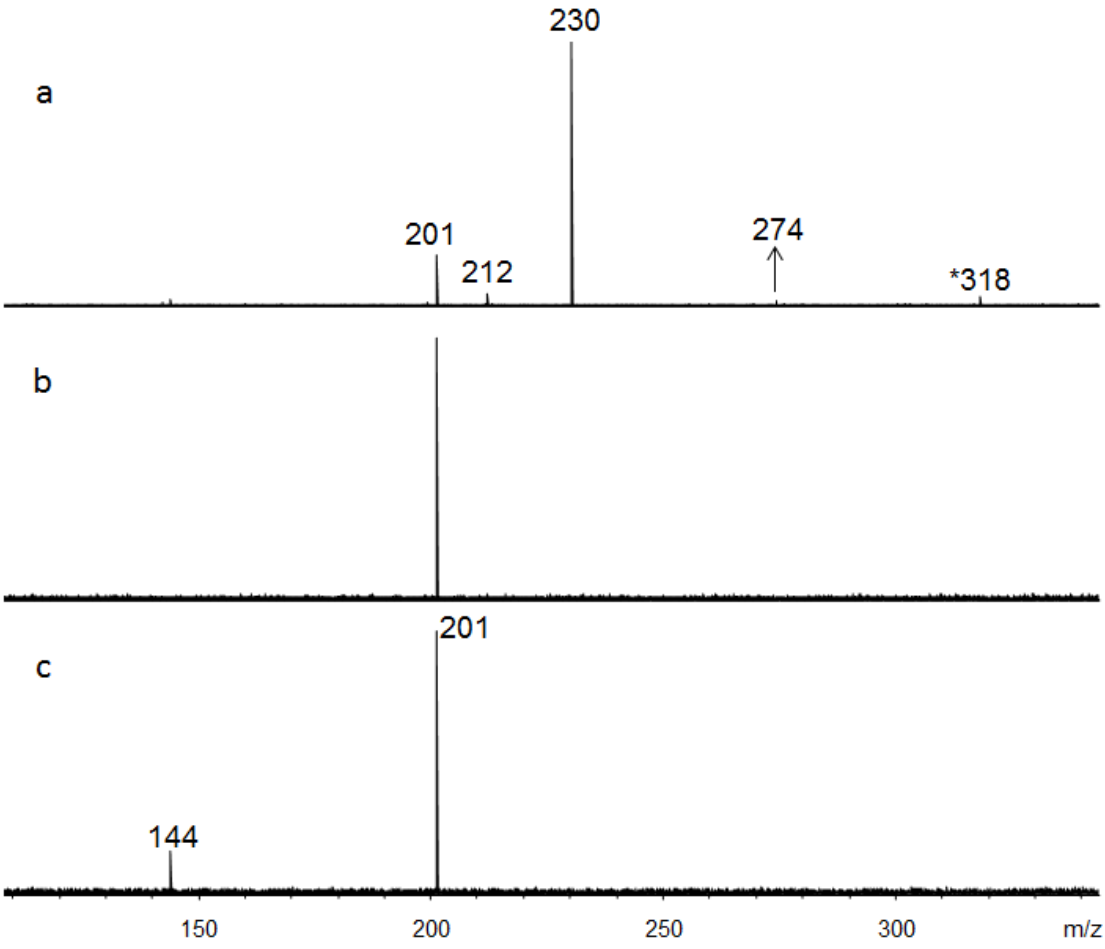


Figure S5.2. SORI-CID mass spectrum of $[Mn(GlyGly-H)(GlyGly)]^+$ (*m/z* 318) (a) and the mass spectra following isolation (b) and dissociation (c) of fragment ion *m/z* 201 formed by SORI-CID of *m/z* 318.

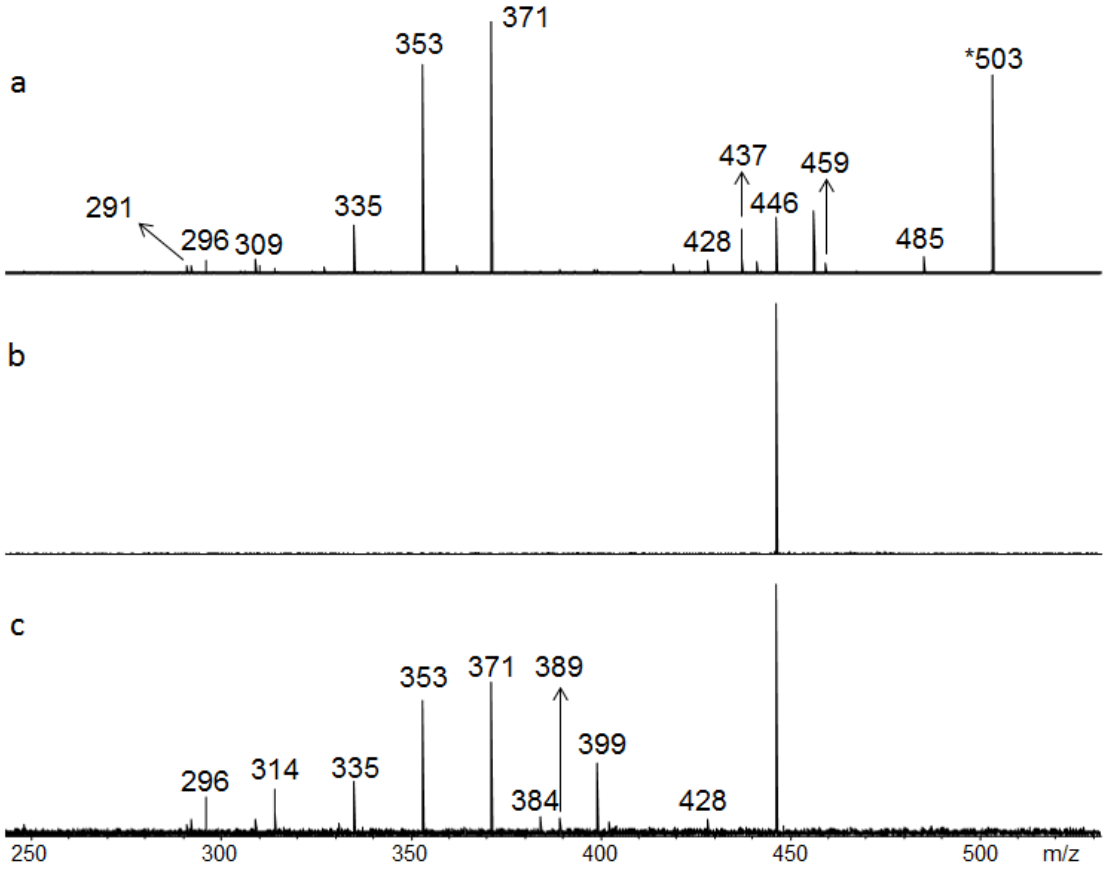


Figure S5.3. SORI-CID mass spectrum of $[Mn_2(GlyGly\text{-}H)_3]^+$ (*m/z* 503) (a) and the mass spectra following isolation (b) and dissociation (c) of fragment ion *m/z* 446 formed by SORI-CID of *m/z* 503.

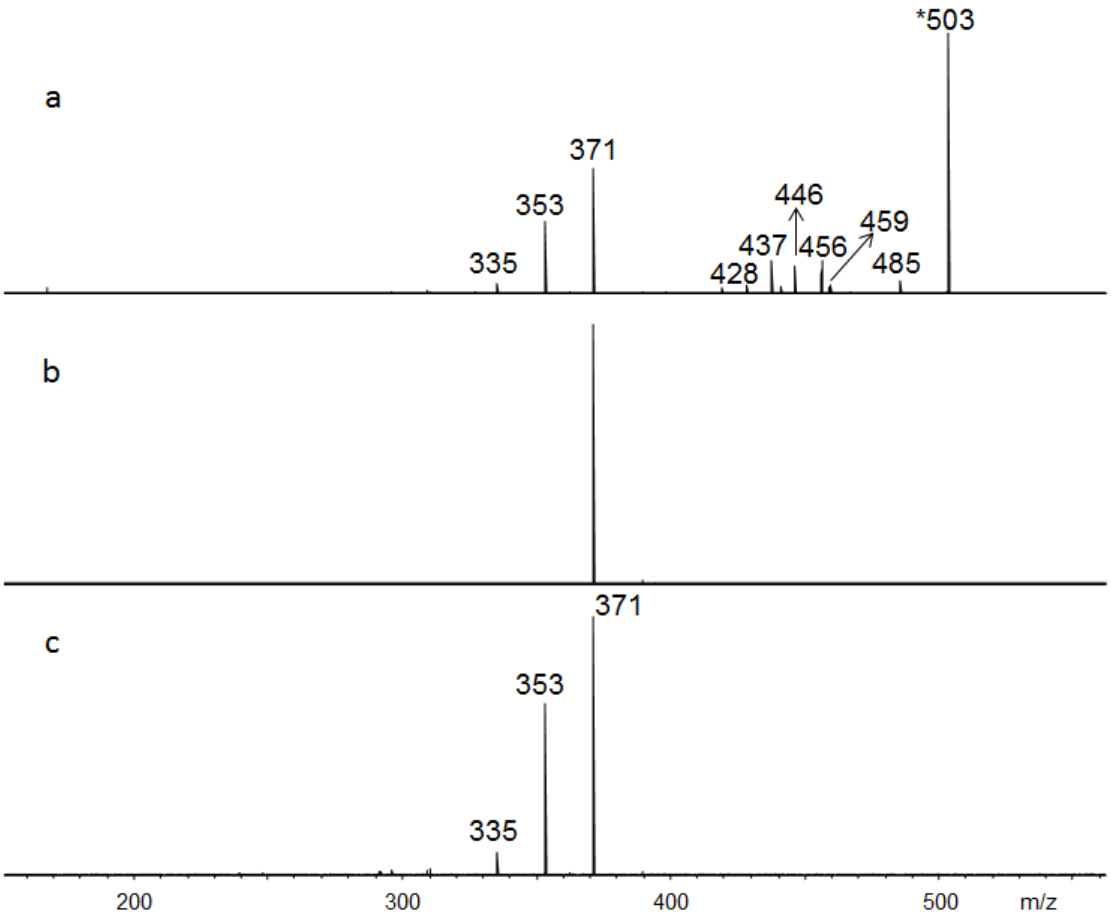


Figure S5.4. SORI-CID mass spectrum of $[Mn_2(GlyGly\text{-}H)_3]^+$ (*m/z* 503) (a) and the mass spectra following isolation (b) and dissociation (c) of fragment ion *m/z* 371 formed by SORI-CID of *m/z* 503.

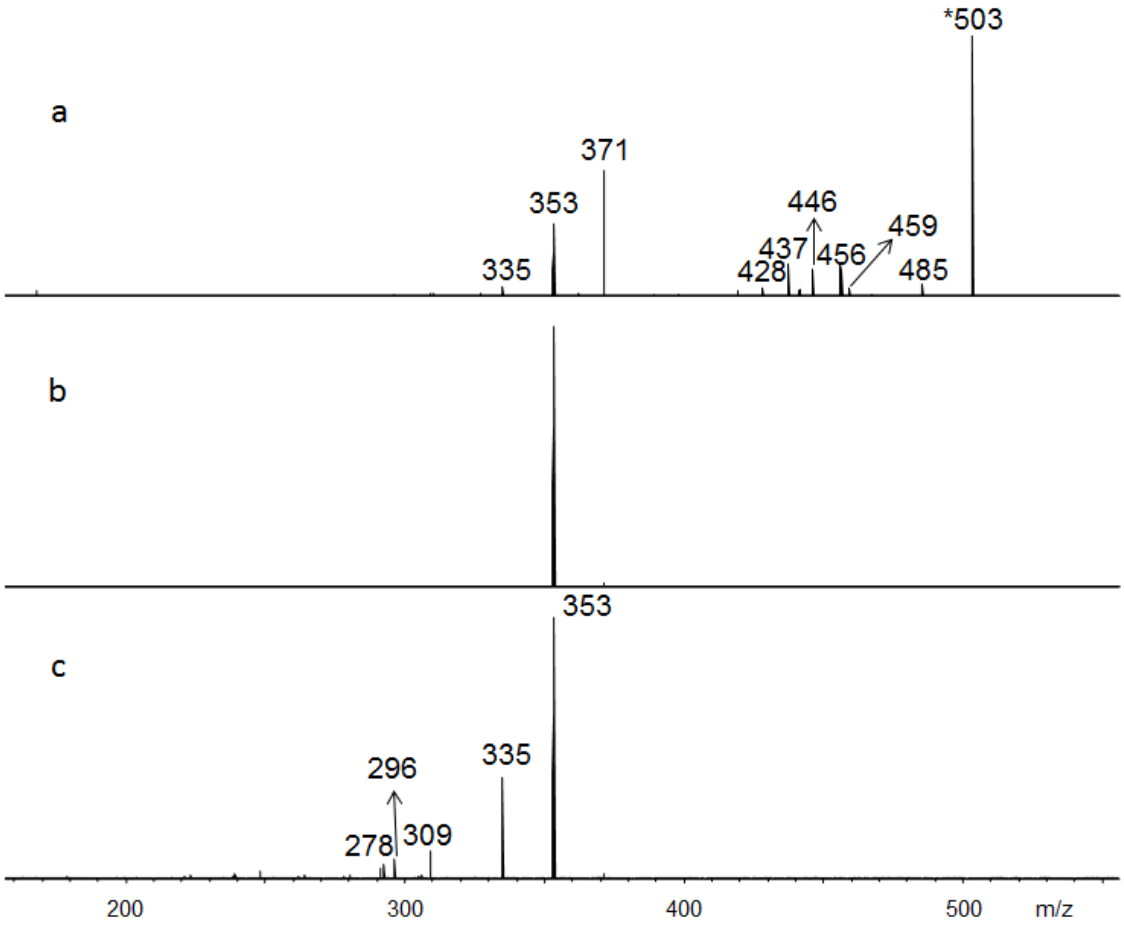


Figure S5.5. SORI-CID mass spectrum of $[Mn_2(GlyGly\text{-}H)_3]^+$ (*m/z* 503) (a) and the mass spectra following isolation (b) and dissociation (c) of fragment ion *m/z* 353 formed by SORI-CID of *m/z* 503.

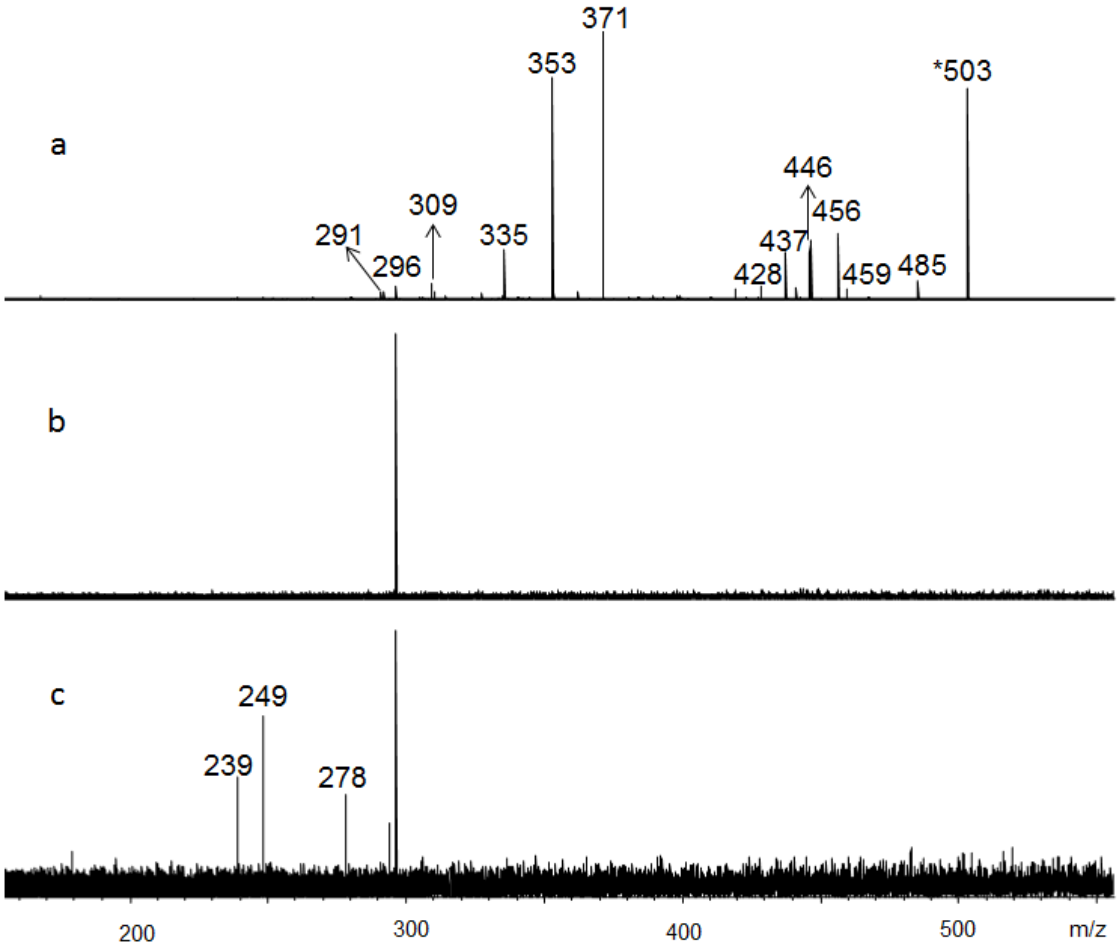


Figure S5.6. SORI-CID mass spectrum of $[Mn_2(GlyGly\text{-}H)_3]^+$ (*m/z* 503) (a) and the mass spectra following isolation (b) and dissociation (c) of fragment ion *m/z* 296 formed by SORI-CID of *m/z* 503.

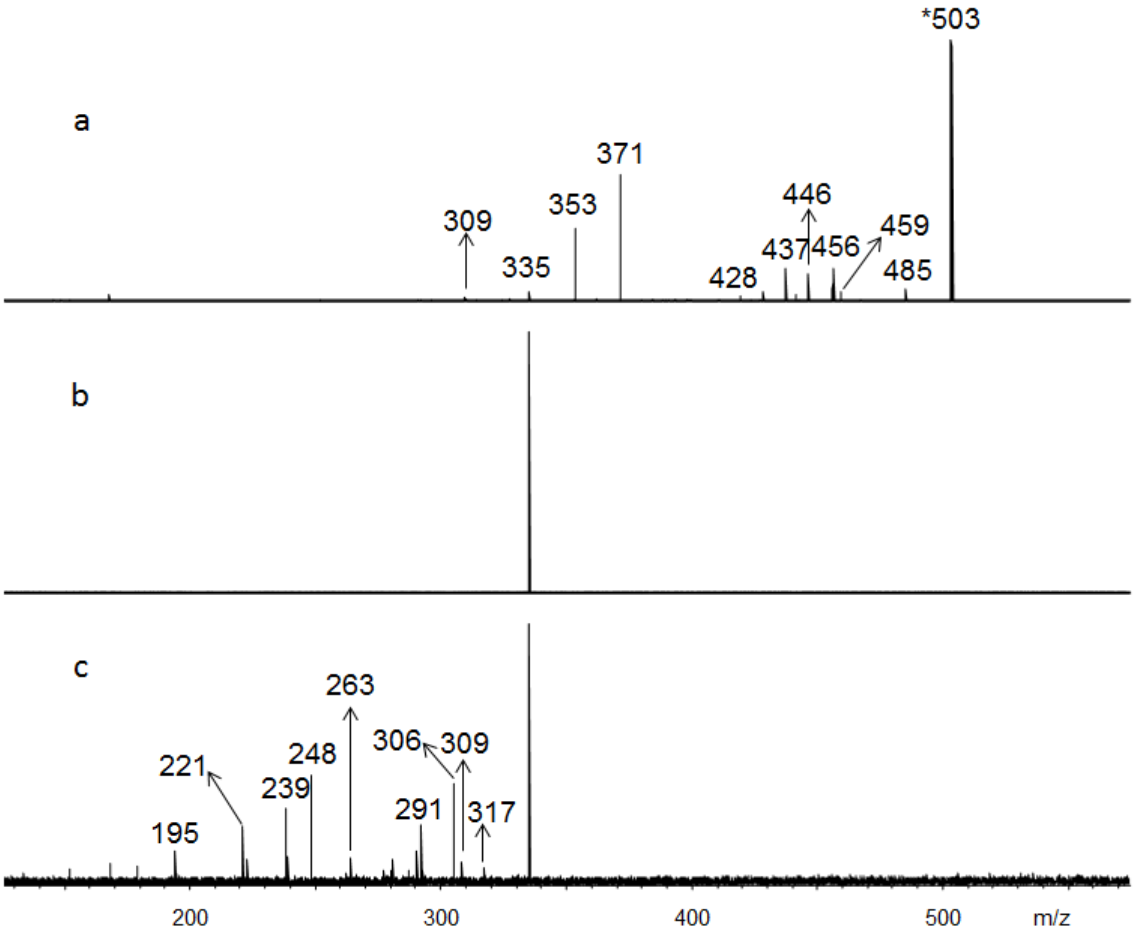


Figure S5.7. SORI-CID mass spectrum of $[Mn_2(GlyGly\text{-}H)_3]^+$ (*m/z* 503) (a) and the mass spectra following isolation (b) and dissociation (c) of fragment ion *m/z* 335 formed by SORI-CID of *m/z* 503.

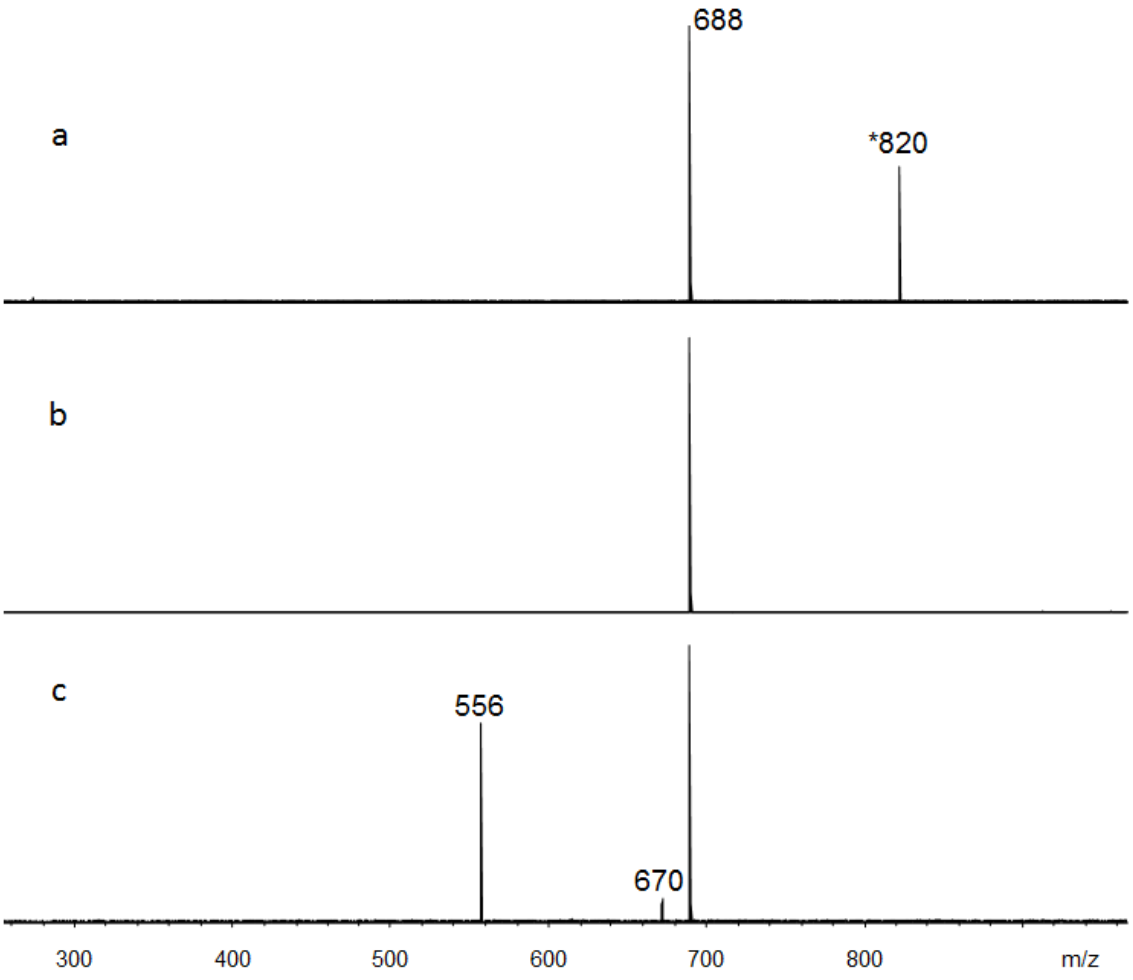


Figure S5.8. SORI-CID mass spectrum of $[Mn_3(GlyGly\text{-}H)_5]^+$ (*m/z* 820) (a) and the mass spectra following isolation (b) and dissociation (c) of fragment ion *m/z* 688 formed by SORI-CID of *m/z* 820.

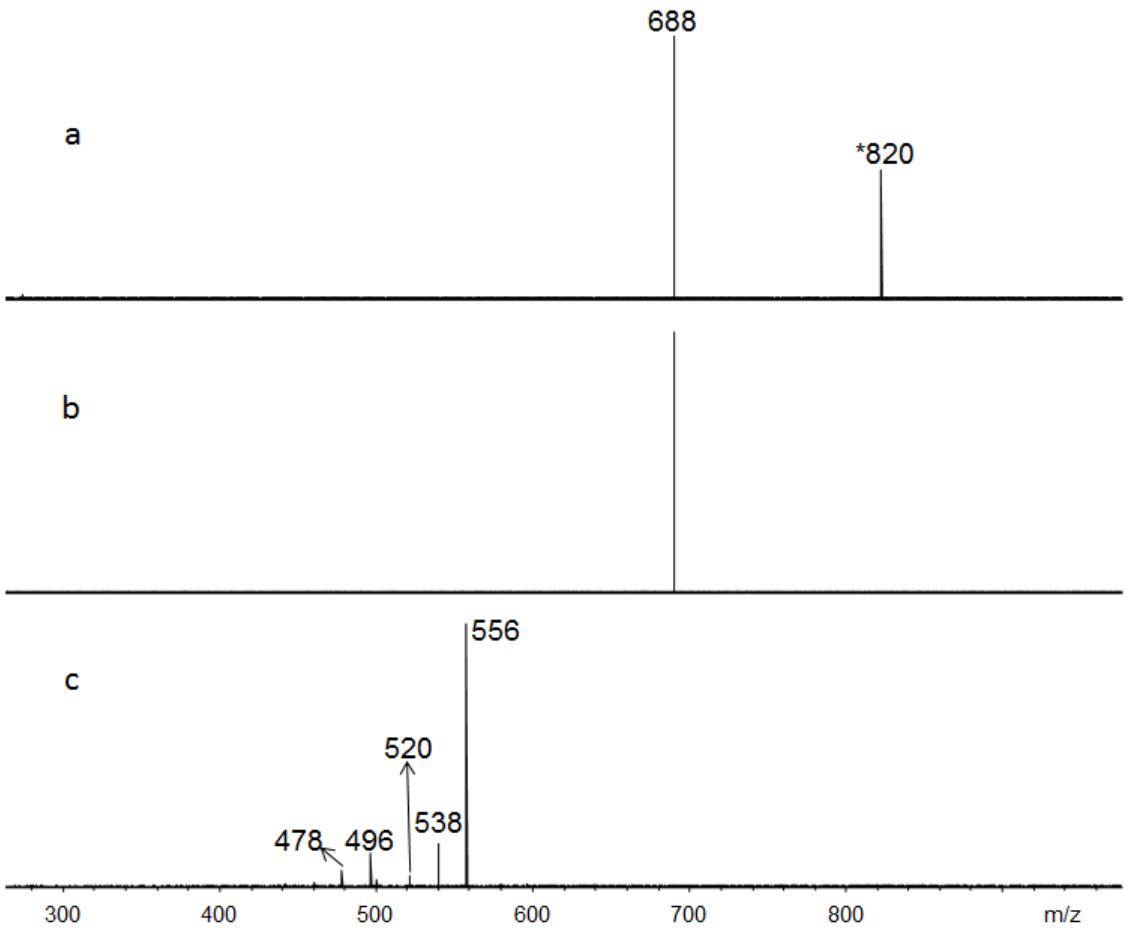


Figure S5.9. SORI-CID mass spectrum of $[Mn_3(GlyGly\text{-}H)_5]^+$ (*m/z* 820) (a) and the mass spectra following isolation (b) and dissociation (c) of fragment ion *m/z* 556 formed by SORI-CID of *m/z* 820.

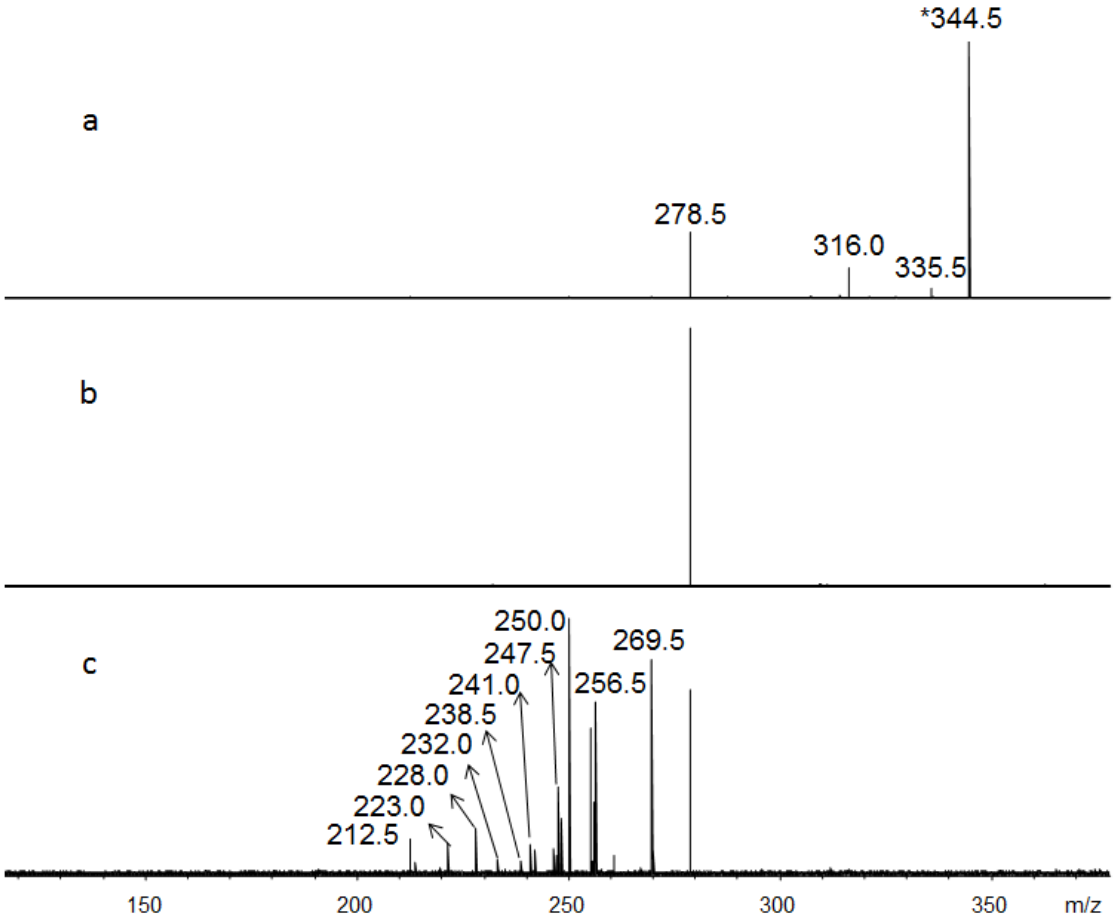


Figure S5.10. SORI-CID mass spectrum of $[Mn_3(GlyGly\text{-}H)_5]^+$ (*m/z* 344.5) (a) and the mass spectra following isolation (b) and dissociation (c) of fragment ion *m/z* 278.5 formed by SORI-CID of *m/z* 344.5.

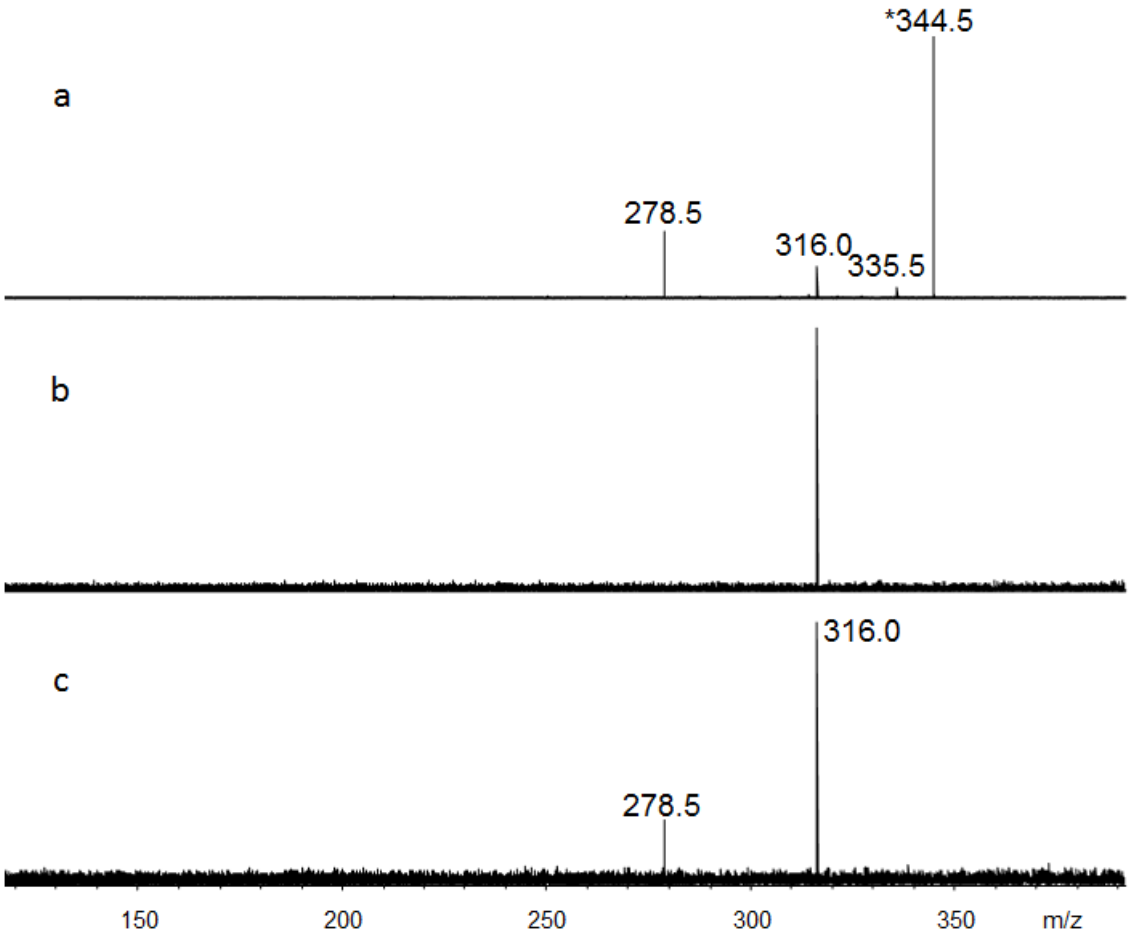


Figure S5.11. SORI-CID mass spectrum of $[Mn_3(GlyGly\text{-}H)_4]^{2+}$ (*m/z* 344.5) (a) and the mass spectra following isolation (b) and dissociation (c) of fragment ion *m/z* 316.0 formed by SORI-CID of *m/z* 344.5.

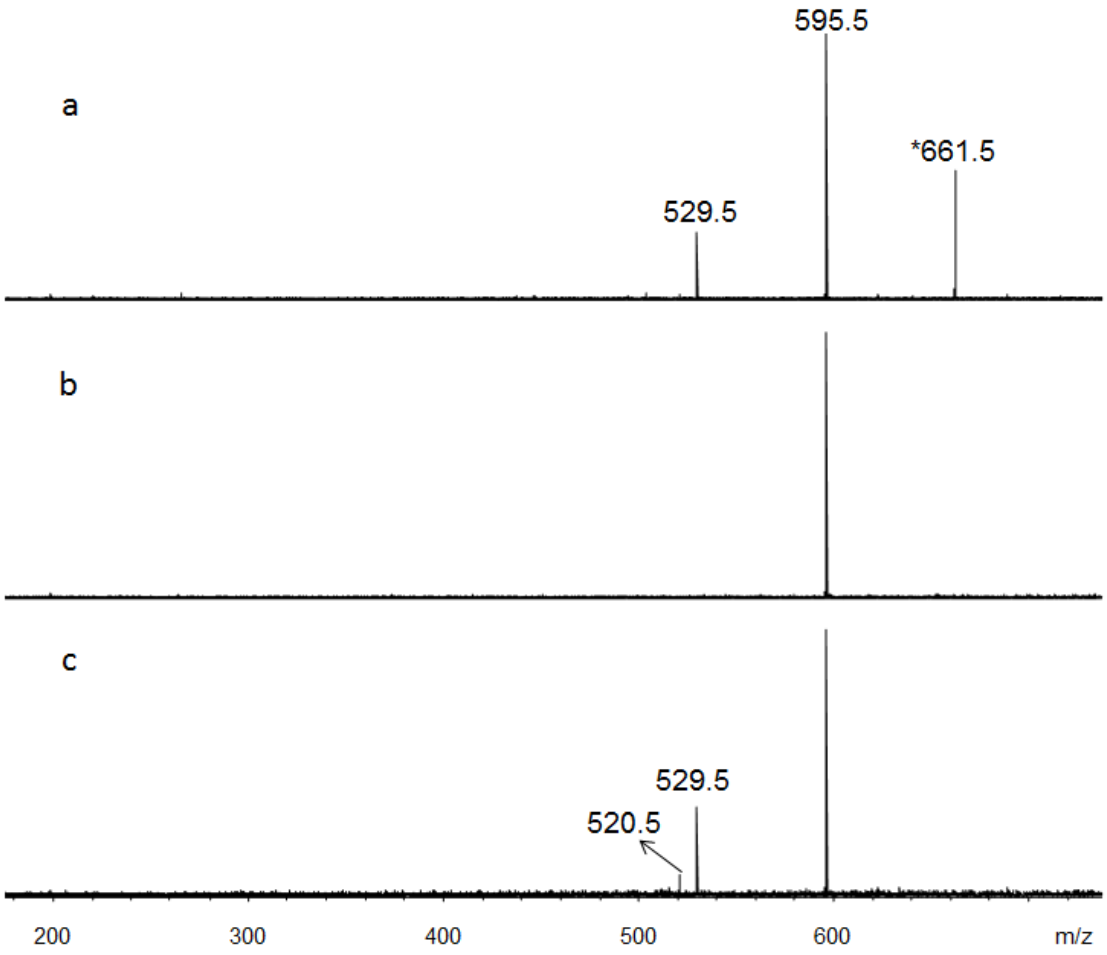


Figure S5.12. SORI-CID mass spectrum of $[Mn_5(GlyGly\text{-}H)_8]^{2+}$ (*m/z* 661.5) (a) and the mass spectra following isolation (b) and dissociation (c) of fragment ion *m/z* 595.5 formed by SORI-CID of *m/z* 661.5.

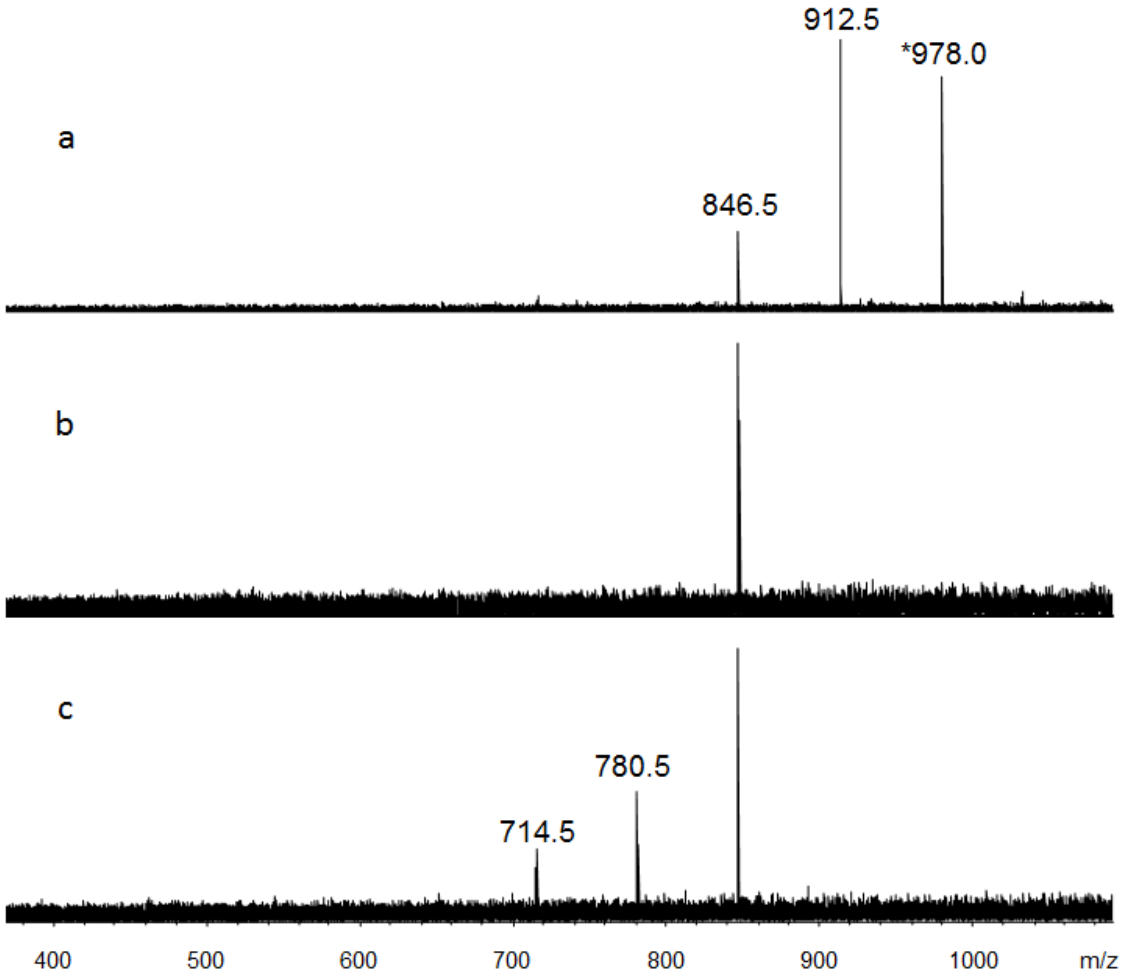


Figure S5.13. SORI-CID mass spectrum of $[Mn_7(GlyGly\text{-}H)_{12}]^{2+}$ (*m/z* 978.0) (a) and the mass spectra following isolation (b) and dissociation (c) of fragment ion *m/z* 846.5 formed by SORI-CID of *m/z* 978.0.

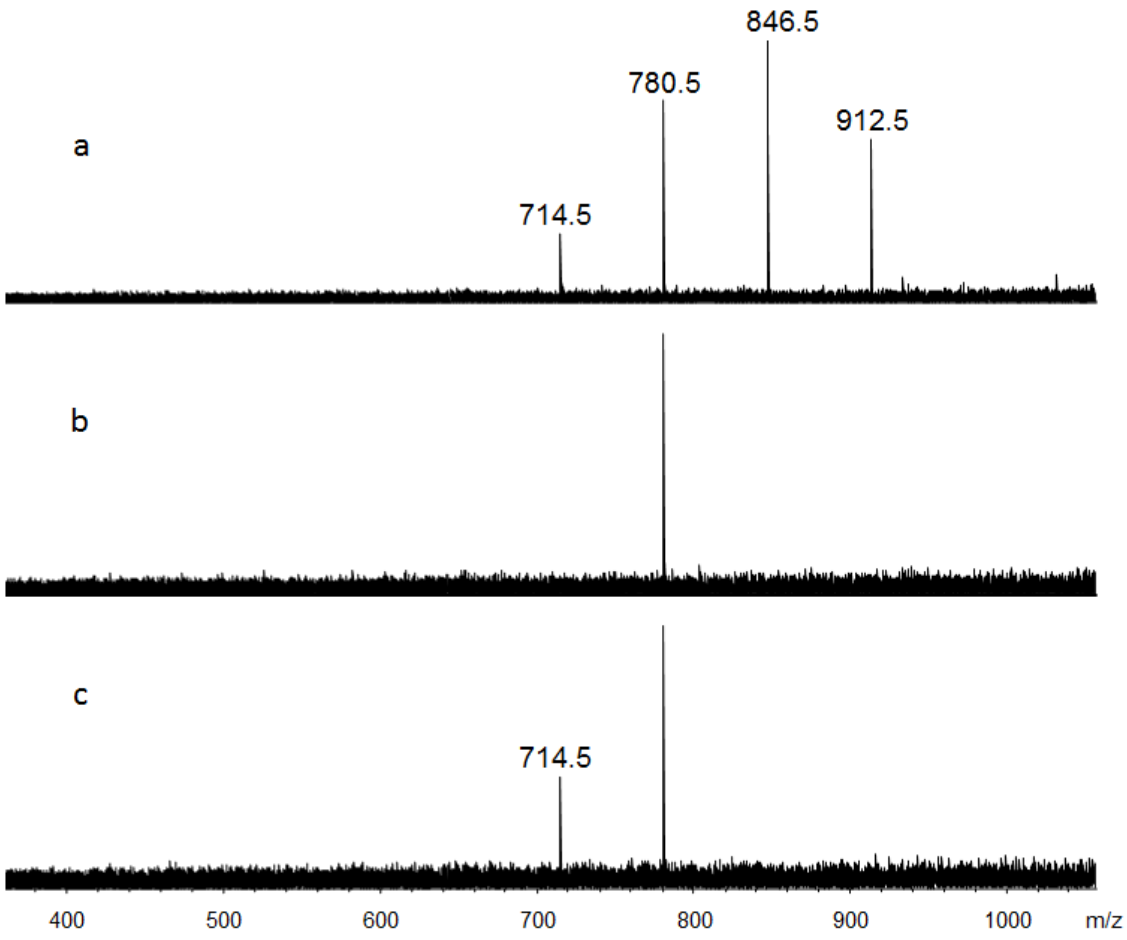


Figure S5.14. SORI-CID mass spectrum of $[Mn_5(GlyGly\text{-}H)_8]^{2+}$ (*m/z* 978.0) (a) and the mass spectra following isolation (b) and dissociation (c) of fragment ion *m/z* 780.5 formed by SORI-CID of *m/z* 978.0.

Table S5.1. Computed metal-oxygen bond lengths (Å) for high (S=5/2) and low (S=3/2) spin multiplicity state structures of $[Mn_3(GlyGly\text{-}H)_4]^{2+}$ using B3LYP/6-31+G(d,p) on C, H, N and O and Def2SVP on Mn^{2+}.

Metal-oxygen bond	S=3/2	S=5/2
Mn1-O1	2.11	2.26
Mn1-O2	2.27	2.21
Mn1-O4	2.16	2.23
Mn1-O5	2.18	2.25
Mn1-O8	2.12	2.19
Mn1-O11	2.50	2.35
Mn2-O2	2.02	2.17
Mn2-O3	1.99	2.07
Mn2-O7	2.28	2.13
Mn2-O11	2.00	2.25
Mn2-O12	2.00	2.06
Mn3-O5	2.07	2.13
Mn3-O6	2.04	2.08
Mn3-O8	2.28	2.40
Mn3-O9	2.00	2.06
Mn3-O10	2.08	2.10

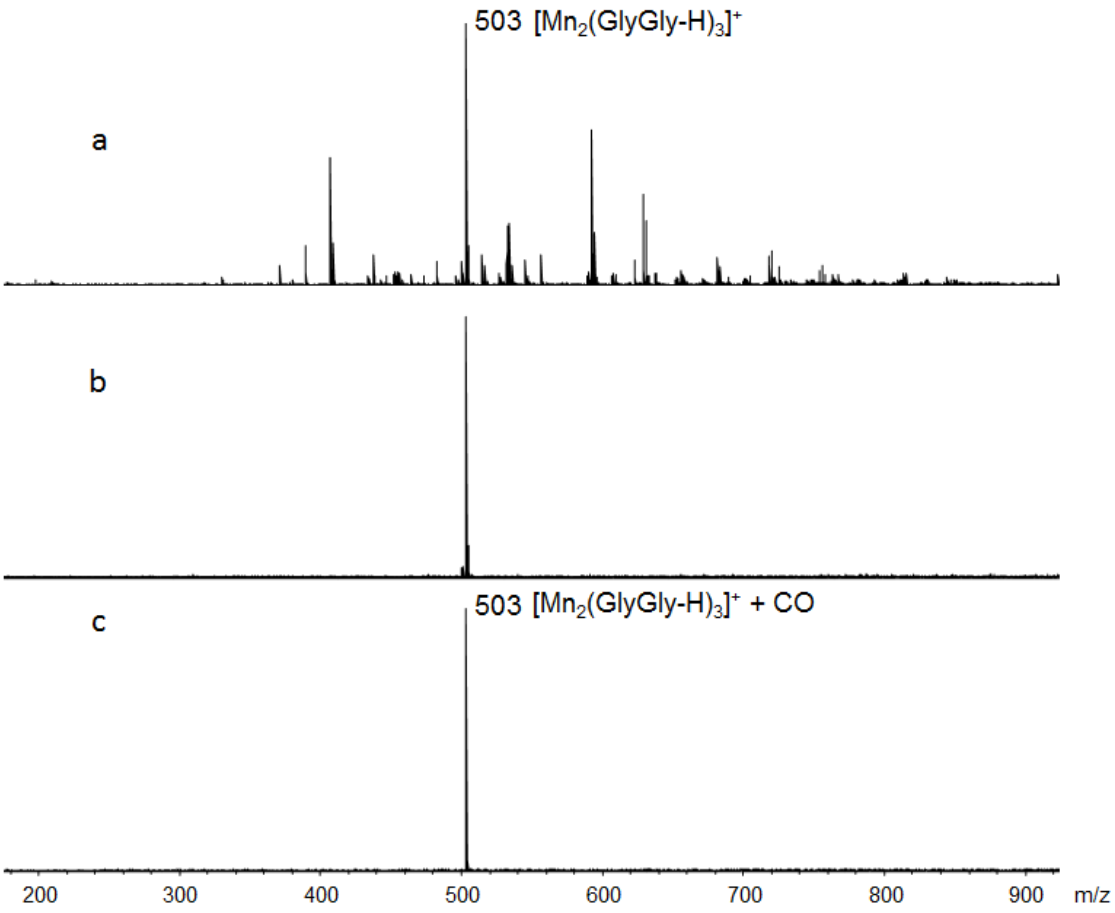


Figure S5.15. Full mass spectrum obtained from electrospraying $MnCl_2$ and GlyGly solution (a). Mass spectra obtained from (b) isolation of $[Mn_2(GlyGly\text{-}H)_3]^+$ cluster at *m/z* 503 and (c) reaction of *m/z* 503 with one pulse of CO.