# Investigating the structures and properties of DMS-selected amino acid ions

by

Yuting Li

A thesis

presented to the University of Waterloo

in fulfillment of the

thesis requirement for the degree of

Master of Science

in

Chemistry

Waterloo, Ontario, Canada, 2020

## Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

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

# Abstract

This thesis outlines gas-phase studies of ionized amino acids in which their structures and interactions within ion cluster complexes are probed using mass spectrometry, differential ion mobility, and laser spectroscopy. Chapter 1 discusses the motivation and background of this thesis, while chapter 2 introduces the methods applied in the research project. In chapter 3, ultraviolet photodissociation (UVPD) spectroscopy of molecular ions following separation within a differential mobility spectrometry mass spectrometer (DMS-MS) is used to investigate the various gas-phase conformations of two protonated aromatic amino acids, tryptophan (Trp) and tyrosine (Tyr). Isomeric forms of target ions are isolated using DMS-MS, and UVPD is performed to map out their respective UVPD spectra. Vibronic spectra of the protonated amino acid conformers are simulated at the density functional theory (DFT) level to assist the identification of prototropic isomers (protomers) of ions observed in the experiments. Assigning the experimental spectra to the calculated vibronic spectra, we conclude that the most likely protonation site of Trp and Tyr is on the nitrogen of the amine in the gas phase.

In chapter 4, DMS-MS is employed to study the structural stability of protonated arginine $(Arg)_n$ clusters (n=1−4) resulting from salt-bridge interactions. The ion-solvent interactions between ions and various gaseous molecules monitored in the DMS cell are analyzed to define cluster stability under different gaseous environments. Further understanding is gained by simulating the electrostatic potential maps of protonated Arg monomer and clusters − this provides insight into the charge distribution around a molecule induced by relevant molecular interactions. Both experimental and computational findings suggest that the protonated Arg monomer is stabilized with the single protonation site on the nitrogen of the guanidino group, and that salt-bridges exist in the gas phase structures of protonated $(Arg)_n$ (n=2−4).

## Acknowledgements

Firstly, I would like to express my special thanks of gratitude to my research supervisor, Prof. Scott Hopkins, who gave me the opportunity to pursue graduate studies, and who provided invaluable help throughout my undergraduate and graduate studies. It is a great privilege and honour to work and study under his guidance. I am extremely grateful for what he has offered me.

I would like to thank Prof. Germán Sciaini and Dr. Rick Marta for being a part of my committee and spending time on my thesis. I really appreciate it. My special thanks also go to Prof. Marcel Nooijen for his guidance of simulations during the Chem 740 course.

My sincere thanks also go to the current and past members of the Hopkins group for the stimulating discussions, the time we have been working together, and the constant support. In particular, I am grateful to Dr. Neville Coughlan for the instructions he has given me experimentally and his expertise in thesis writing. I would like to acknowledge Dr. Jeff Crouse for his patience and help during my project. I would also like to thank people working in the international Hopkins office (C2-261D), Nour Mashmoushi, Fiorella Villanueva Heldmaier, Weiqiang Fu, and Suzy Lim. They are not only great lab mates during work but also perfect friends in life.

My special thanks also go to Feng Li and his wife, Fengjiao Zhu, who allowed me to work at their place and cooked me delicious food during my apartment construction.

Last but not least, I would like to thank my family and my friends for supporting me spiritually throughout my life.

# Table of Contents

# List of Figures

# List of Tables

# List of Abbreviations

| | |
|---|---|
| ACE | Acetone |
| ACN | Acetonitrile |
| Arg | Arginine |
| Asp | Aspartic acid |
| B3LYP | Becke 3-parameter LEE-Yang-Parr |
| BDE | Bond dissociation energy |
| BH | Basin hopping |
| CAM-B3LYP | Coulomb attenuating method Becke 3-parameter LEE-Yang-Parr |
| CCS | Collision cross section |
| CE | Collision energy |
| CHelpG | Charges from electrostatic potentials using a grid-based method |
| CID | Collision induced dissociation |
| CV | Compensation voltage |
| DFT | Density functional theory |
| DMS | Differential mobility spectrometry |
| DNA | Deoxyribonucleic acid |
| DC | Direct current |
| DP | De-clustering potential |
| ESI | Electrospray ionization |
| FAIMS | High-field asymmetric waveform ion mobility spectrometry |
| FC | Franck-Condon |
| Glu | Glutamic acid |
| HF | Hartree Fock |
| HOTO | Highest occupied transition orbital |
| IPA | Isopropyl alcohol |
| LIT | Linear ion trap |
| LUTO | Lowest unoccupied transition orbital |
| Lys | Lysine |
| MeOH | Methanol |
| MM | Molecular mechanics |
| MO | Molecular orbital |
| MS | Mass Spectrometry |
| *m/z* | Mass-to-charge ratio |

| | |
|---|---|
| NTO | Natural transition orbitals |
| OPO | Optical parametric oscillators |
| PCM | Polarizable continuum model |
| PES | Potential energy surface |
| Q1 | First quadrupole |
| Q2 | Second quadrupole |
| Q3 | Third quadrupole |
| RF | Radiofrequency |
| RNA | Ribonucleic acid |
| SV | Separation voltage |
| TD-DFT | Time-dependent density functional theory |
| UFF | Universal force field |
| UV | Ultraviolet |
| UVPD | Ultraviolet photodissociation |
| VG | Vertical gradient |

# Chapter 1: Introduction

Amino acids, the basic structural units of proteins, play important roles in various physiological functions.[1–4] Among the hundreds of known amino acids, research is focused on analyzing the twenty common amino acids involved in protein sequencing. They are the key molecules required for the synthesis and modification of deoxyribonucleic acid (DNA) and ribonucleic acid (RNA).[1–4] Additionally, they participate in most cell functions, such as metabolic regulation, osmoregulation, and cell signalling.[1–4] The mechanism of these essential biological reactions depends on the intermolecular interactions and biological activities of relevant amino acids, which are determined by their structural and chemical properties.[5–9] Structures of amino acids can be modified by protonation and proton transfer, which are associated with various biological functions such as energy transfer, acid-base reactions, and catalysis.[10–19] Multiple protonated states of an amino acid can exist owing to the presence of both acidic and basic functional groups such as the amine ($–NH_2$) and carboxyl (–COOH) functional groups, as shown in Figure 1.1.

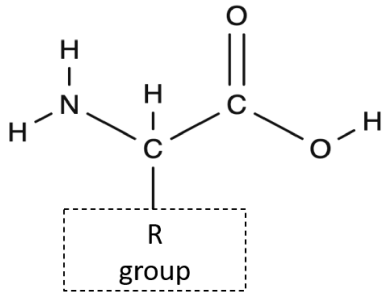


**Figure 1.1** An example of an $\alpha$-amino acid structure, where the R group refers to various side chains.

Amino acids tend to adopt zwitterionic structures in the aqueous phase across a wide pH range.[12,20,21] In contrast, most amino acid structures are stabilized by nonzwitterionic forms in the gas phase, resulting in the isolation of each site for protonation.[12,20,21] The relative proton affinity

of each site and gas-phase basicity determine the position of a proton attachment in the gas phase.[11,12,20,22–24] Nitrogen atoms, which reside on an amine or a basic side chain, are typically the most likely protonation site in amino acids and small peptides in the gas phase because they have relatively high proton affinities.[19,21] However, the proton is not always located on the site with the highest proton affinity but is capable of moving between protonation sites, as a proton tends to reside at the cleavage site in the course of dissociation of most peptides.[13,25,26] Additionally, oxygen-protonated amino acids can be induced when the peptide size increases.[27,28] Lorenz and Rizzo reported that different protonated sites and intermolecular proton transfers were observed for structures of phenylalanine/serine dimers in the gas phase − this was then confirmed by Fu and Hopkins with machine learning studies.[28,29] Their research indicated that there is more than one possible protonation site in these amino acid structures. Therefore, it is important to investigate structural changes of amino acids with different protonation sites, which give rise to distinct properties.

Photodissociation action spectroscopy has been demonstrated as a valuable tool to help explore the electronic and geometric structures and distinguish protomers of analytes in the gas phase, especially for molecules containing conjugated $\pi$-systems that have strong light absorptivity in the ultraviolet (UV) region.[30–35] For the purpose of this thesis, emphasis is placed on the two aromatic amino acids, tryptophan (Trp) and tyrosine (Tyr). Photodissociation action spectroscopy of aromatic amino acids in the gas phase has been performed by many groups.[30–33] Talbot and co-workers showed that protonated Trp had the appropriate complexity for being a benchmark to examine computational findings, providing a better understanding of electronic excited states and fragmentation mechanisms.[32] Pereverzev *et al.* compared the UV spectra between aromatic amino acids and protonated His in both the gas and aqueous phases, suggesting that the proton was not

likely to reside on the aromatic ring in structures of Trp and Tyr.[30] However, experiments conducted in previous studies did not include methods that allow for the isolation of individual protomers of an ion prior to UVPD. Accordingly, this project employs a technique called differential mobility spectrometry (DMS), which can spatially filter different protomers of an ion before irradiation.[36–43] In this way, it is possible to investigate individual conformations of aromatic amino acids with different protonation sites using UVPD in the gas phase. Electronic spectra of molecular ions, particularly those where the charge site is adjacent to the aromatic group, are beneficial for identifying the structures of amino acids with different protonation sites. Identification of various structures in the gas phase also provides clues about how proton transfer occurs in amino acid systems.

Ionic bonding can occur between charged amino acid side chains resulting in multiply charged centers in amino acid complexes. Accordingly, arginine (Arg), lysine (Lys), aspartic acid (Asp), and glutamic acid (Glu) can form what are known as salt-bridges between their side chains, which can involve both ionic and hydrogen bonding interactions, as shown in Figure 1.2a.[44–55]

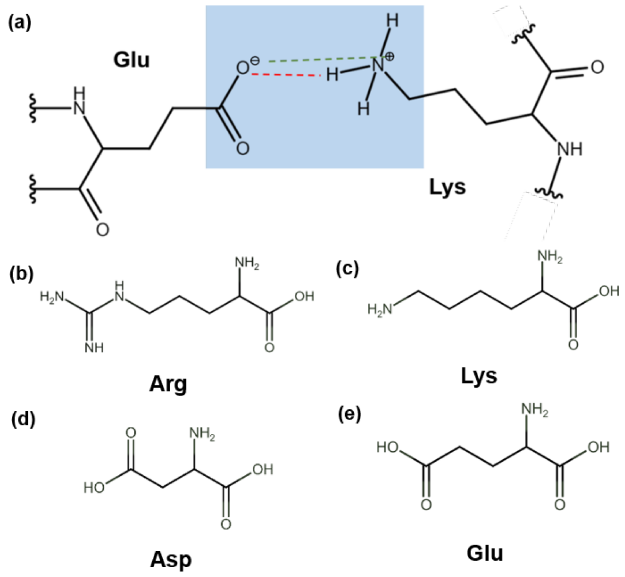


**Figure 1.2** (a) An example of a salt-bridge between Glu and Lys, of which region is highlighted as blue, containing hydrogen bonding (red dash) and ionic bonding (green dash); Structures of (b) Arg, (c) Lys, (d) Asp, and (e) Glu.

The salt-bridge interaction is a type of protein interaction, which has been broadly explored based on the specific functions that result from ion pairs, such as binding sites of enzymes, molecular recognition, and fundamental motifs of protein secondary and tertiary structures.[56–61] Studying the structural and functional roles played by salt-bridges provides knowledge into the relevant protein stability contributed from salt-bridge interactions.[26,56–62] Previous work of salt-bridge interactions involving relevant amino acids has mainly focused on their protonated monomers, dimers, and peptides. These studies indicated that the salt-bridge interactions stabilize amino acid complexes of dimers and peptide chains more than monomers.[48–55,63] It is expected that the number of salt-bridge interactions increases when additional salt-bridge-forming functional groups are present, which possibly influences their strengths. Protonated amino acid clusters that can form numerous salt-bridges have not been thoroughly investigated. Detailed investigation of these systems would aid in understanding the relationship between the structural stability and strength of salt-bridge interactions. Arg, which potentially owns the highest number of protonation sites compared to Lys, Asp, and Glu (see Figure 1.2), is used in this project to examine the presence and stability of different Arg protomers in the gas phase. Salt bridge interactions of protonated Arg clusters (*i.e.*, dimer, trimer, and tetramer) are explored and compared within the cluster sizes using DMS, which is helpful to explore the stability and charge shielding effect of ions through studying their ion-solvent interactions with various gas molecules.[36–43]

This thesis employs both experimental and computational methods to explore the geometries of molecules of interest, details of which are introduced in Chapter 2. Chapter 3 summarizes studies of the geometric and electronic structures of protonated Trp and Tyr, which are experimentally investigated with UVPD spectroscopy in the gas phase using DMS mass spectrometry (DMS-MS) coupled to a tunable probe laser system.[39] Computational work is

conducted in parallel with the experimental studies to provide a more comprehensive understanding of the electronic transitions and fragmentation pathways associated with each of the analyte species. In Chapter 4, DMS-MS is employed to determine the link between structural stability and salt-bridge interactions of protonated $(Arg)_n$ (n=1−4) by studying the ion-clustering behaviours in various gas-phase environments and the computational electrostatic potential maps. Overall, both projects are completed with a decent agreement between experimental and quantum chemical calculations conducted at the DFT level of theory.[64,65]

This thesis provides an investigation into the geometric and electronic structures of protonated amino acids and amino acid-containing clusters. The work presented demonstrates the utility of DMS-MS in the exploration and isolation of different target ion conformers formed in the gas phase. Altering the DMS cell conditions (*i.e.*, gas-phase environments and de-clustering potential voltages) provides a greater ability to comprehend the conformational stability and interactions of target ions with solvent molecules. Notably, coupling DMS-MS to a laser system allows for investigating the photodissociation and electronic spectroscopy of DMS-filtered ions. The combination of DMS-MS and action spectroscopy analysis along with computational work is beneficial to delivering a comprehensive physicochemical analysis of ions of interest. With the work mentioned here, we hope to provide better insight into the physicochemical properties of aromatic amino acids and the nature of salt-bridge interactions formed between relevant amino acids.

# Chapter 2: Introduction to methods

## 2.1 Differential mobility spectrometry

Differential mobility spectrometry (DMS), also known as high-field asymmetric waveform ion mobility spectrometry (FAIMS), is a useful technique for characterizing and separating gaseous molecular ions.[36,38,43,66] A DMS cell is comprised of two planar and parallel isolated electrodes between which ions navigate and are separated spatially according to their differential mobility between high and low electric field conditions in a carrier gas, as shown in Figure 2.1a.[36,43]

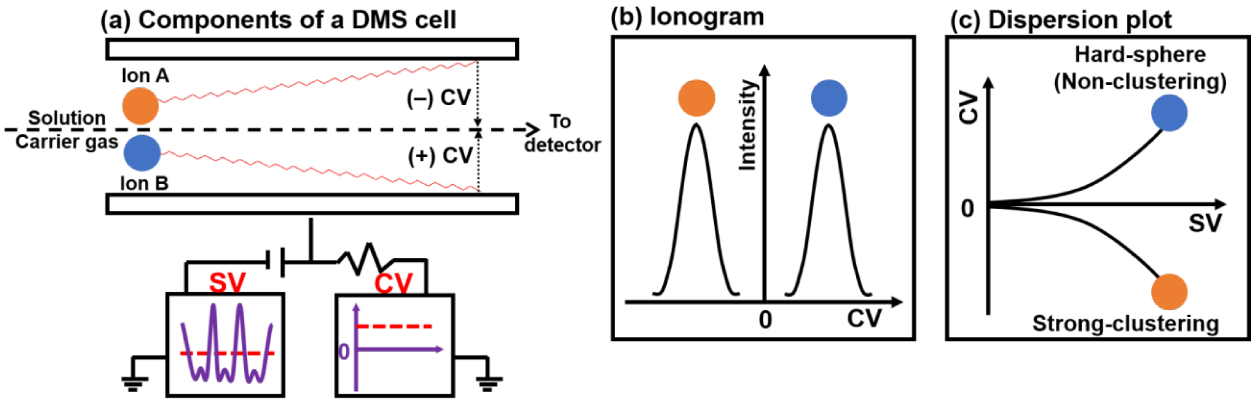


**Figure 2.1** (a) The components of a DMS cell, (b) an ionogram, and (c) a dispersion plot. Ions A and B are positively charged. Adapted from Mashmoushi.[67]

A neutral carrier gas (*i.e.,* $N_2$) drives ions to drift axially through the DMS cell. The ion drift velocity ($v_d$) is affected by the ion mobility ($K$) and the applied electric field ($E$) in the following manner:[38,43]

$$v_d = KE. \tag{2.1}$$

Ion mobility remains relatively unchanged over a wide range of low electric-field strengths; therefore, it can be considered constant under low field conditions. However, at higher field strengths, the ion mobility constant exhibits field dependence. The mobility of ions under the high-field condition is approximated by:[43]

$$K\left(\frac{E}{N}\right) = K_0\left[1 + \alpha\left(\frac{E}{N}\right)\right]. \quad (2.2)$$

$K_0$ refers to the mobility constant under low field, $E/N$ refers to the electric field strength in Townsend, $\alpha(E/N)$ − also known as the alpha parameter − refers to the normalized function describing the ion's field-mobility dependence, and $N$ refers to the density of the transport gas.[43] The ion mobility $K$ varies as a function of the electric field strength ($E/N$) when the electric field is high, as illustrated in Equation (2.2). The dependence of the ion mobility on the electric field strength is exploited in DMS. An asymmetric radiofrequency waveform, called the separation voltage (SV), is applied across the electrodes of the DMS cell. The waveform is comprised of both high and low electric field components that influence ion motion in the cell. The different ion mobilities under high and low field affect ion drift velocity under the relevant field conditions [Equation (2.1)]. This results in ions moving unequally under high and low electric fields in the DMS cell. The unequal displacements of the ion, as shown in Figure 2.1a, result in an off-axis trajectory in the DMS cell, leading to the neutralization of ions as they collide with one of the electrodes.[38,43] According to the various interactions between ions and carrier gas molecules in the DMS cell, contrasting paths can be observed for different ions, as shown in Figure 2.1a. To counteract the off-axis trajectory due to the SV, a direct current (DC) voltage known as the compensation voltage (CV), is applied across the two electrodes.[38,43] An appropriate CV can steer ions back towards the exit orifice of the DMS cell for subsequent detection.[38,43] An ionogram illustrates the detected ion intensity as the CV is scanned at a fixed SV value, where the optimal CV is determined as the point of maximum transmission of selected ions. An example of an ionogram is shown in Figure 2.1b, in which positively charged ions A and B have characteristic optimal CV values at a given SV value according to their unique clustering behaviours. Therefore, they can be selected and isolated in the DMS cell with their experimentally determined SV/CV

pairs chosen for separation. Hence, experimentally determined SV/CV pairs can be employed to distinguish structural isomers or protomers of an ion in the gas phase, wherein spatial separation of similarly structured ions is relatively attainable at higher SV values, as the differential mobility of these ions can be greatly exaggerated, resulting in distinct CV values.

The ion's optimal CV as a function of SV (Figure 2.1c) is known as its dispersion behaviour, usually visualized in a dispersion plot. The DMS behaviour can be rationalized in terms of one of three main behaviours with respect to solvent clustering, as shown in Figure 2.2.[36,38,43]

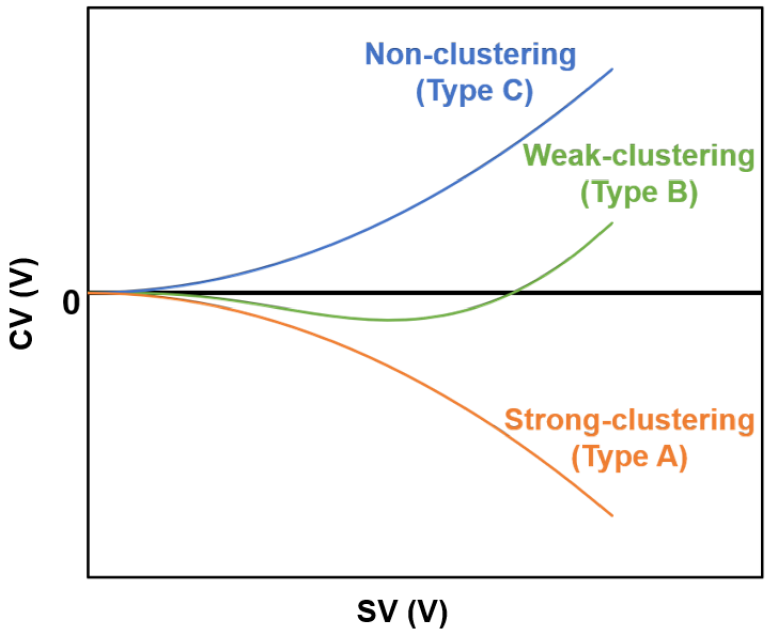


**Figure 2.2** A dispersion plot showing three typical clustering behaviours.

Type A behaviour is observed when ions cluster strongly with solvent molecules under the low-field portion of the waveform then de-cluster under high field conditions; the CV required for optimal ion transmission is observed to decrease as the SV increases.[36,43] When ions form clusters under the low-field component of the SV, their apparent size increases.[43] This results in a reduction of the ion mobility due to an artificially increased collision cross section (CCS) which in turn increases collision frequency in comparison to the bare ion. Under high-field conditions, relative ion mobility increases as a result of the smaller apparent CCS. Therefore, as shown in Figure 2.2,

negative CV values are required to correct the trajectory of strongly clustering ions. Non-clustering behaviour, or Type C behaviour, is associated with CV increasing as SV increases.[36,43] In Type C behaviour, the change in the ion's apparent size is relatively insignificant since no obvious dynamic clustering is observed between ions and solvents. Instead, the differential mobility of ions showing Type C behaviour is dictated by collision frequency of target ions and gas molecules. Collisions between ions and the carrier gas occur more frequently during the high-field component of the SV waveform, resulting in a reduced ion mobility compared to that under low field condition.[43] Consequently, contrasting trends are observed for strong-clustering (Type A) and non-clustering (Type C). The dispersion plot for Type B behaviour is observed for weakly clustering ions; it is an intermediate between Types A and C, as a small amount of ion-solvent clusters are formed under low SVs and eliminated at moderate field strengths.[36,43] Therefore, the ion experiences strong-clustering (indicated by a negative slope in the dispersion plot) until a critical SV point is reached, in which the ion is de-solvated leading to a non-clustering behaviour (positive slope). Distinct shapes of dispersion curves reveal different types of interactions between ions and gas molecules in the DMS cell.

Multiple interactions, such as ion-dipole, ion-induced dipole, hydrogen bonding, and van der Waals forces can exist between a target ion and the carrier gas which can be modified by seeding it with a low concentration of volatile solvent vapour. The addition of gas-phase modifiers induces dynamic ion-solvent clustering and de-clustering under alternating electric field, that is manifested as a change in ion mobility, as shown in Figure 2.3.[36,40] At low-fields, solvation of ions is observed as ion-solvent clusters are formed, whereas ion clusters are de-solvated under high-field conditions due to an increase in input energy and collisions with neutral gas molecules.[36,40] The dynamic clustering process indicates that a slight change in binding capability

between solvent molecules and ions can lead to notable diversity in dynamic ion-solvent clustering behaviour, resulting in a significant shift in CV values during an SV cycle.[36] Hence, the various gas-phase conditions presented in the DMS cell can induce notable changes in ions' clustering behaviours and thus enhance the resolving power for ions with similar structures.

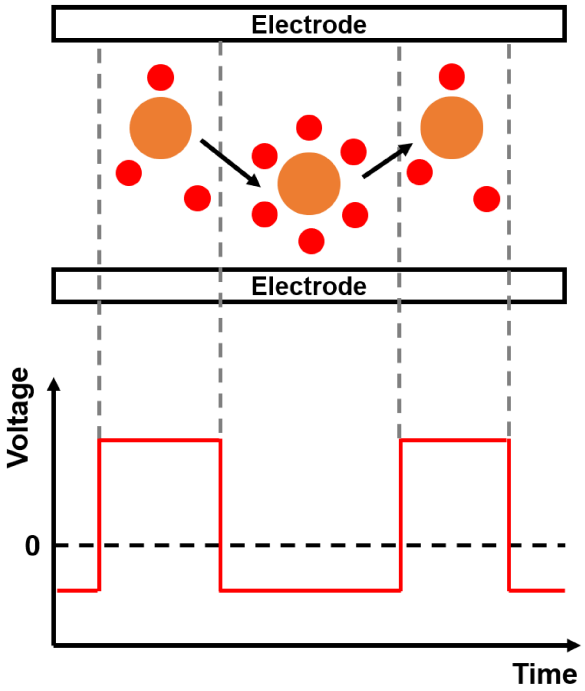


**Figure 2.3** Dynamic ion-solvent clustering at alternating electric field. Orange and red circles indicate the target ion and surrounding solvent molecules, respectively.

DMS generally plays an important role in the separation of ions with similar structures (*i.e.*, protomeric species) due to their distinct ion-solvent interactions. DMS also works as an ion filter for isomeric species, or ions with the same *m/z* ratio, and can distinguish them using experimentally determined SV/CV pairs. Additionally, geometric structures of target ions can be investigated through studying the ion-solvent interactions under various gas-phase conditions. A drawback of DMS-MS is its inability to conclusively identify some isomeric species. Accordingly, an additional characterization technique, photodissociation action spectroscopy, is coupled to DMS-MS for further structural identifications of target species.

## 2.2 UV Photodissociation

Photodissociation is the process in which a bound molecule fragments following absorption of one or more photons.[65] The photon energy is absorbed leading to bond breaking in the molecule, after which any excess photon energy is converted into the internal and translational energy of the products.[65] The photodissociation process for a diatomic molecule, AB, can be written as:[65]

$$AB + N_{photon}h\nu \overset{(1)}{\rightarrow} (AB)^* \overset{(2)}{\rightarrow} A + B + E_{trans} + E_{int}\,. \tag{2.3}$$

In Expression (2.3), $N_{photon}h\nu$ is the total energy of $N$ absorbed photons, $(AB)^*$ is an excited intermediate state of the molecule, A and B are fragmentation products of AB, and $E_{\text{trans}}$ and $E_{\text{int}}$ are translational and internal energies of the products, respectively. Hence, the photodissociation process involves two steps. In step 1, the excited complex is formed in a dissociative electronic state or a bound state which then couples to a dissociative state. Following dissociation, atoms A and B are produced with $E_{\text{trans}}$ and $E_{\text{int}}$ distributions consistent with the excess photon energy above the dissociation threshold, $D_0$:[65]

$$N_{photon}h\nu - D_0 = E_{trans} + E_{int}. \tag{2.4}$$

There are four photodissociation processes that can occur for diatomic molecules when a single photon is absorbed: direct photodissociation, predissociation, unimolecular decay, and spontaneous radiative dissociation. These are illustrated in Figure 2.4.

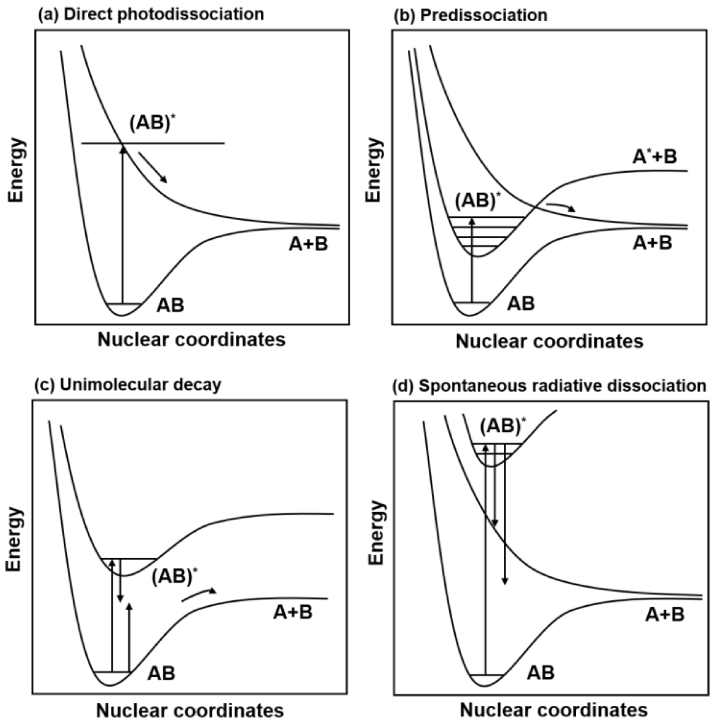


**Figure 2.4** Four photodissociation processes of a diatomic molecule AB, which are (a) direct photodissociation, (b) predissociation, (c) unimolecular decay, and (d) spontaneous radiative dissociation, respectively, when a single photon is absorbed. Adapted from Schinke and Dishoeck *et al.*[65,68]

Direct photodissociation (Figure 2.4a) is the simplest photodissociation process where the molecule is excited to a repulsive electronic state.[65,68] During direct dissociation, the dissociation of a molecule occurs faster than the spontaneous emission of a photon back to the ground state, meaning that all absorption events lead to fragmentation of the molecule.[68] In predissociation (Figure 2.4b), a molecule is first excited to a quasi-bound electronic excited state, which can couple to a repulsive excited state through non-radiative relaxation. The bound excited state may also go through a radiationless transition to the ground state potential, resulting in population above the ground state thermodynamic threshold and subsequent dissociation to form ground state products − this is known as unimolecular decay (Figure 2.4c).[65,68] Another case of unimolecular decay includes the single-photon excitation of overtone vibrations that leads to energy accumulation in the molecule populating a quantum state above the dissociation threshold. Photodissociation is thus induced for the AB molecule in its ground electronic state.[65] Finally, spontaneous radiative

dissociation (Figure 2.4d) involves spontaneous emission and relaxation into the continuum of a lower-lying electronic state of a molecule that is repulsive along the nuclear coordinate.[68]

UV photodissociation (UVPD) refers to the photodissociation process taking place following photon absorption in the UV region (190−400 nm). The high energy of the UV photons can encompass multiple electronic states, and thus all four photodissociation pathways (Figure 2.4) can be involved during UVPD of a molecule.[65,68] Electronic transitions *via* various UVPD processes can be visualized in a UVPD spectrum, which maps the absorption as a function of photon energy − an example is shown in Figure 2.5. In our experiment, the UVPD action spectrum is obtained by treating absorption as the parent ion depletion and product fragment enhancement as a function of wavelength (photon energy), details of which (*i.e.*, experimental setup and fragmentation efficiency) will be introduced later in section 3.2.1.

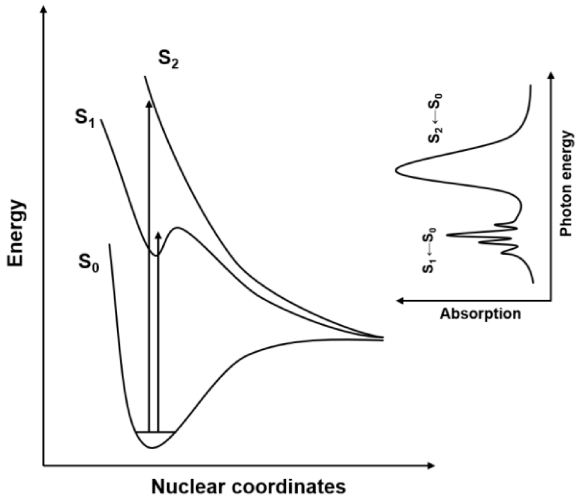


**Figure 2.5** An example of potential energy surface of the ground state ($S_0$) and excited states ($S_1$ and $S_2$) for a molecule (left); the photodissociation spectrum corresponded with electronic transitions of $S_0$ to $S_1$ and $S_0$ to $S_2$ (right). Adapted from Kelley.[69]

UVPD action spectroscopy is helpful for geometric identification since a structure owns unique electronic transitions during excitation. The intensity of an electronic transition can be interpreted by Franck-Condon factors, which also provide prediction in the shapes of photodissociation spectra.[64,69]

## 2.3 Franck-Condon factors

Franck-Condon (FC) factors describe the intensity of vibronic band absorptions, which includes simultaneous changes in electronic and vibrational energy.[64,69] Since nuclei are much heavier than electrons, the timescale of an electronic transition (*i.e.,* femtoseconds) is much shorter than that of vibrational motion (*i.e.,* picoseconds).[64,69] During the electronic transition, the nuclei can be considered fixed due to the relatively large timescale difference between electronic and vibrational transitions. This is known as the Born-Oppenheimer approximation.[64,69] Charge redistribution in a molecule induced by the electronic transitions leads to changing Coulombic forces acting on the nuclei. The geometry of the nuclei is distorted changing the energies of the vibrational states.[64] Consequently, it is necessary to consider both the electronic and vibrational states in the course of vibronic transitions.

$$\Psi_{elect,i}\,\psi_{vib,iI} \rightarrow \Psi_{elect,f}\,\psi_{vib,fF} \tag{2.5}$$

In Expression (2.5), $\Psi_{elect}$ represents the electronic wavefunction, while $\psi_{vib}$ denotes the vibrational wavefunction; *i* and *f* signify the initial and final electronic states, respectively; *I* and *F* represent the respective vibrational quantum numbers.[69] The electronic and vibrational wavefunctions are separable based on the Born-Oppenheimer approximation.[69] The electronic transition dipole moment between an initial and a final electronic state determines the transition intensity and is defined in Expression (2.6):[69]

$$\langle \Psi_{elect,f}\,\psi_{vib,fF}\,|\mu|\,\Psi_{elect,i}\,\psi_{vib,iI} \rangle. \tag{2.6}$$

In Expression (2.6), bra-ket notation describes an integral over all electronic and vibrational coordinates of the system, and $\mu$ refers to the dipole moment operator.[69] Employing the Born-Oppenheimer approximation and assuming the initial vibrational state lies close to the bottom of the well, the dipole moment operator can be approximated through a Taylor series as:

$$\mu = \mu^0 + \sum_j \left(\frac{\partial \mu}{\partial Q_j}\right)(Q_j - Q_{j0}) + \cdots, \tag{2.7}$$

where $\mu^0$ is the equilibrium transition moment and $Q$ refers to the coordinates of nuclei.[64,69] Inserting Equation (2.7) into (2.6) gives:

$$\left\langle \Psi_{elect,f}\, \psi_{vib,fF} \middle| \mu^0 + \sum_j \left(\frac{\partial \mu_z}{\partial Q_j}\right)(Q_j - Q_{j0}) \middle| \Psi_{elect,i}\, \psi_{vib,iI} \right\rangle, \tag{2.8}$$

which for fixed nuclei (*i.e.,* $Q_j$=$Q_{j0}$) gives:

$$\langle \psi_{vib,fF} | \psi_{vib,iI} \rangle \langle \Psi_{elect,f} | \mu^0 | \Psi_{elect,i} \rangle. \tag{2.9}$$

Transition intensities, are then determined by the square of the transition dipole, Expression (2.9):

$$Intensity \propto \left| \langle \Psi_{elect,f} | \mu^0 | \Psi_{elect,i} \rangle \right|^2 \left| \langle \psi_{vib,fF} | \psi_{vib,iI} \rangle \right|^2. \tag{2.10}$$

Expression (2.10) allows for the determination of transition intensities. The constant electronic transition $\Psi_{elect,i} \rightarrow \Psi_{elect,f}$, will only differ by the vibrational wavefunction overlap, $\left| \left\langle \psi_{vib,fF} \middle| \psi_{vib,iI} \right\rangle \right|^2$, defined as the Franck-Condon factor. Therefore, the overall intensity of transition between electronic states, termed the oscillator strength, is shown in Equation (2.11):

$$f_{i \rightarrow f} = \frac{2 m_e \hbar}{e^2} \bar{\nu} \left| \langle \Psi_{elect,f} | \mu^0 | \Psi_{elect,i} \rangle \right|^2, \tag{2.11}$$

where $m_e$, $e$ and $\bar{\nu}$ are the electron mass, electron charge and the average frequency of electronic transition, respectively.[64] Hence, the oscillator strength yields the intensity for the overall electronic transition between the two electronic states (*i.e.*, *i* to *f*). The FC factor describes the intensity exhibited by transitions among individual vibrational levels (*i.e.*, *iI* to *fF*). An illustrative energy diagram of the ground state, $S_0$, and the first singlet excited state, $S_1$, is shown in Figure **2.6**, where the occurrence of the most intense transition in the ($v'$, $v''$=0) progression is the (2,0)

band owing to the relatively large overlap between the wavefunctions of these two vibrational states.[70]

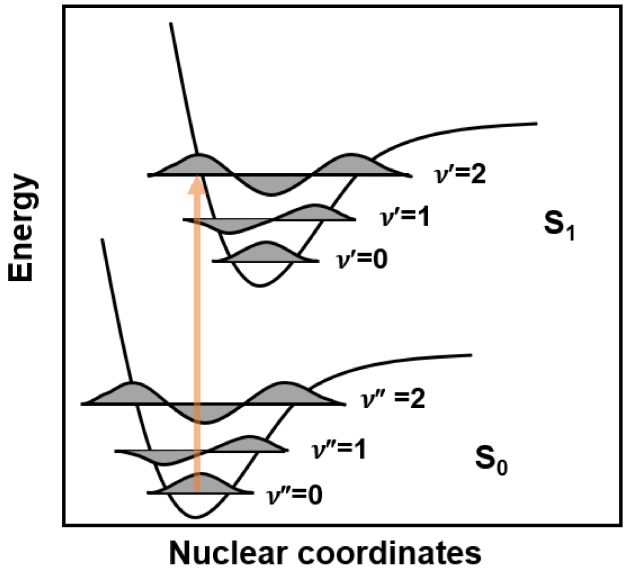


**Figure 2.6** The FC principle energy diagram from the ground state to the first excited state where the most favoured transition is represented by an orange vertical arrow between $\nu''= 0$ and $\nu'= 2$ for a diatomic molecule. Adapted from Nano.[70]

A challenge for simulating FC factors is determining the equilibrium geometry and force field of the excited state.[71,72] A linear coupling model or the vertical gradient (VG) approximation is employed for calculating the FC factors in this project, eliminating time-consuming steps (*i.e.,* excited state optimizations).[72] In the VG approximation, the excited state geometry is approximated by the excited state gradient evaluated at the ground state equilibrium geometry.[71,72] Therefore, the potential energy surface of an excited state is defined near the vertical excitation point, which induces the most intense vibronic band without changes in the normal coordinates and harmonic frequencies.[72] Vibronic transitions calculated using VG|FC factors can be interpreted by natural transition orbitals (NTOs) associated with each relevant excitation. Different from the canonical molecular orbitals (MOs) of which densities are complicated for characterization, NTOs simplify the study of an electronic state as a particle/hole depiction with unchanged transition density.[73,74] NTOs qualitatively describe electronic excitations *via* the orbital transformation

correspondence between the excited particle (occupied orbitals) and the empty hole (virtual orbitals) of an excited state.[73,74]

## 2.4 Basin-hopping

Basin-hopping (BH) is an unbiased algorithm employed in this research project as a global optimization approach to locate the local and global minimum configurations of interest on the potential energy surface (PES).[75,76] In the BH algorithm, the transformed energy $\tilde{V}(\boldsymbol{X})$ can be written as:[75]

$$\tilde{V}(\boldsymbol{X}) = min\{V(\boldsymbol{X})\}. \tag{2.12}$$

In Equation (2.12), $\boldsymbol{X}$ is the vector of nuclear coordinates and *min* refers to the minimization starting from $\boldsymbol{X}$. During the BH process, the energy, $V(\boldsymbol{X})$, on the PES is collected for each local minimum. The PES is transformed as a set of interpenetrating staircases, where each plateau represents the energy of each local minimum, as shown in Figure 2.7.[75,76] Relative energies of the global minimum and local minima are not altered due to the transformation of the PES since the energy at any point in a specific configuration range is always consistent with that of the local minimum.[75]

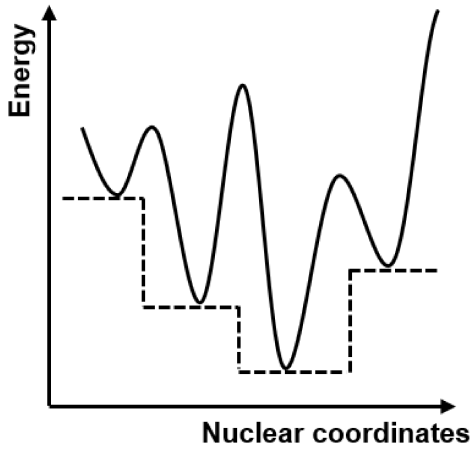


**Figure 2.7** A schematic diagram illustrating the transformation of PES during BH where solid lines are the energy of original PES $V(\boldsymbol{X})$ and dashed lines are the transformed energy $\tilde{V}(\boldsymbol{X})$, respectively. Adapted from Wales and Doye.[75]

During a BH simulation, an initial guess structure of the molecule of interest with energy $V_{old}$ is inputted and is distorted by changing dihedrals and/or translating molecules in a modified Monte Carlo approach to determine a new optimized geometry with energy $V_{new}$.[75,76] If $V_{new}$ is less than $V_{old}$, the new structure is accepted as a new input and $V_{old}$ is replaced with the energy of this optimized geometry. The new structure may still be accepted if $e^{\frac{V_{old}-V_{new}}{kT}}$, where $k$ is the Boltzmann constant, is larger than a random number chosen from [0,1]. In this case, however, $V_{old}$ remains unchanged since $V_{old} < V_{new}$ in the process of searching for new input geometries.[75,76] This step accepts structures to complete the PES curve, and distortion of these structures could result in lower-lying energy structures that better model the true global minimum. Otherwise, the conformation is rejected, and the previous input geometry is distorted and examined following the procedures as described and shown in Figure 2.8.

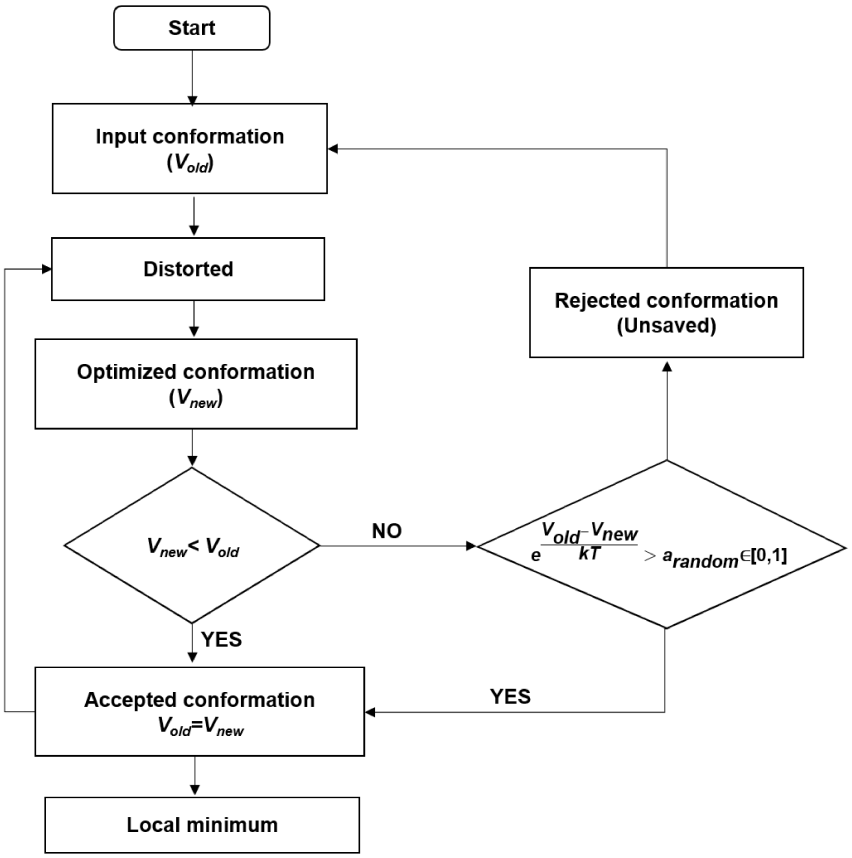


**Figure 2.8** A flow chart that shows the energy selection rules during BH optimizations.

An example of a BH simulation trajectory is shown in Figure 2.9.[76] From left to right. $C_1$, $C_2$ and $C_4$ geometries are accepted since their energies meet the criteria described in Figure 2.8. The $C_3$ structure exhibits a higher energy when compared to $C_2$, which does not meet the selection rule that $e^{\frac{V_{old}-V_{new}}{kT}} > a_{random} \in [0,1]$ and is thus discarded. $C_2$ is then perturbed again until a new conformation, $C_4$, is accepted.

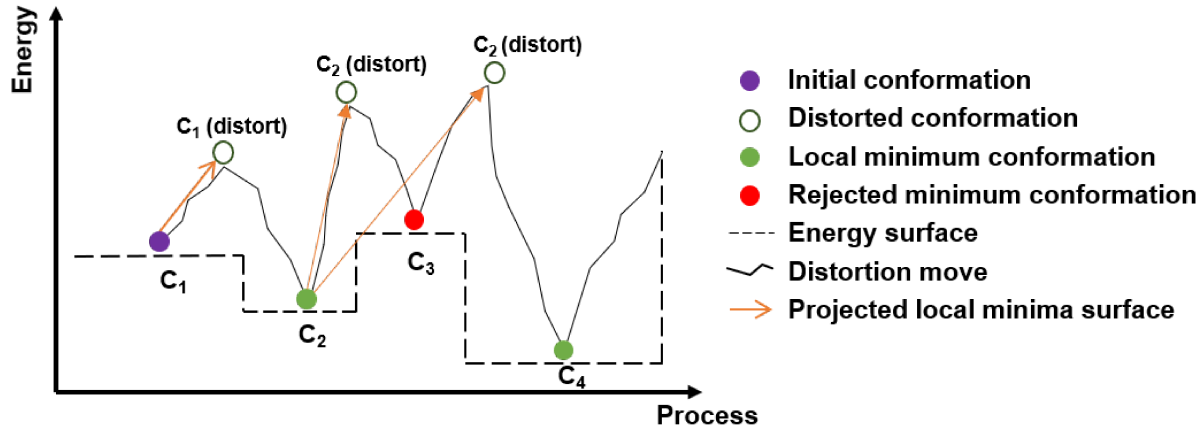


**Figure 2.9** An example of BH trajectory. Adapted from Olson *et al.*[76]

Many individual optimization steps (*ca.* 10,000) are involved in BH simulations for a single system and require time-efficient computational calculations. Therefore, molecular mechanics (MM) − a low level classical theory − is employed to estimate system energies. MM force fields include contributions from the harmonic oscillator model to model bonding interactions and Coulombic and van der Waals interactions to model non-bonding interactions.[74–77] MM simulations in this work employ the universal force field (UFF), a computationally inexpensive approach parametrized for all atoms of the periodic table that defines parameters based on the hybridization and connectivity of the element.[76–79] One of the disadvantages of UFF, however, is that it models electrostatic interactions poorly, and specific partial charges need to be calculated and added to BH input files manually.[78–80] Since the precision of MM simulations with UFF is

relatively poor, further optimizations of local minimum structures obtained from BH are computed at higher levels of theory.

## 2.5 Density functional theory

Density functional theory (DFT) has been used broadly for studying geometric, electronic, and spectroscopic properties of chemical systems owing to its relatively good trade-off between computational cost and calculation accuracy.[81–84] DFT applies electron density as the basic quantity since ground-state properties of a many-electron system can be well-determined by the total electron density.[86] Additionally, the use of electron density significantly reduces the scale factors of DFT simulations by only considering three spatially variables for each electron in many-electron wavefunctions compared to many-dimensional wavefunctions.[84] However, the exchange-correlation functionals in terms of density still remain unknown, and thus functionals for electron exchange and correlation need to be approximated for DFT.[82,84,85] The choice of the approximated exchange-correlation functionals can affect the accuracy of the DFT method. For example, Grimme reported that current functionals of DFT could not accurately perform weak interaction calculations, due to the incorrect exchange-correlation potential;[87] Zhang and Yang found that self-interaction error of approximate density functionals could lead to erroneous dissociation.[88] Thus, the deficiencies introduced by the approximated exchange-correlation functionals can limit the efficiency of the DFT method for some predicted properties.[82,85,89]

The DFT method applied with the hybrid Becke 3-parameter Lee Yang Parr (B3LYP) functional performs well for determining various chemical systems and properties, especially amino acid-containing systems like those studied in this work.[28,29] The hybrid functional improves the electron exchange correlation by combining a generalized gradient approximation in which both the electron density and gradient are dependent on exact Hartree Fock (HF) exchange.[27] The accuracy of the hybrid B3LYP falters when applied to time-dependent studies. This is seen in the case of time-dependent density functional theory (TD-DFT), an extension of the DFT method that

defines the properties of the excited states.[8,28,29] Instead, the Coulomb attenuating method B3LYP functional, CAM-B3LYP, is combined with the application of the TD-DFT method for simulations of excited states in this project.[29] CAM-B3LYP not only achieves the accuracy of the B3LYP functional but corrects the charge transfer excitations that B3LYP underestimates significantly.

# Chapter 3: UVPD and electronic spectroscopy of DMS-selected tryptophan and tyrosine ions

## 3.0 Overview

The geometric structures of two gas-phase aromatic amino acid ions (Trp and Tyr) were studied both experimentally and computationally in this chapter. DMS was used to examine their conformations (*i.e.*, protomers) and to study the clustering behaviour of protonated Trp and Tyr under different gas-phase environments. The electronic spectroscopy of DMS-selected ions was studied *via* UVPD to investigate their unique electronic structures, electronic transitions, and fragmentation pathways. Calculated vibronic spectra using the VG approximation at the DFT level showed good agreement with the experimental UVPD spectra for both protonated Trp and Tyr. The identical experimental UVPD spectra observed for two ion populations of protonated Trp indicated that the same structures were presented when looking at the $[Trp + H]^+$ protomers resulting from the DMS-resolved monomer and dimer cluster. This concluded that both the monomer and dimer Trp protomers exhibited the same protonation site. The distinct UVPD spectra of protonated Tyr generated with different fragmentation channels indicated that various fragmentation pathways were involved during the dissociation of protonated Tyr. Bond dissociation energies related to fragmentation processes of protonated Tyr were studied to provide knowledge into the dissociative mechanisms of relevant fragmentation pathways.

## 3.1 Introduction

Trp and Tyr are aromatic amino acids containing side chains composed of an indole ring and a hydroxyphenyl ring, respectively. Their structures are shown in Figure 3.1. These amino acids are related to multiple biological functions and genetic coding processes. [1,10] For example,

Trp plays a vital role in producing the neurotransmitter serotonin, while Tyr is important for signal transduction processes.[90,91] Both Trp and Tyr possess multiple acidic and basic functional groups, which can produce different protonated states in *vivo*, inducing different biological activities dependent on the biological pH.[10] As discussed in Chapter 1, investigating protomeric structures of a molecule helps reveal its chemical and structural properties and provides insight into the mechanism of relevant biological reactions.[5–9] Therefore, exploring the structures of these two aromatic amino acids helps rationalize the mechanisms of important reactions involving Trp and Tyr.

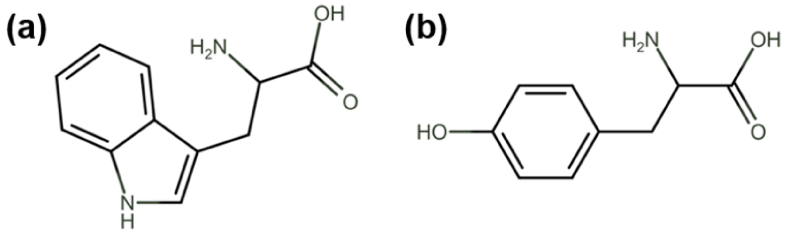


**Figure 3.1** The structures of (a) Trp and (b) Tyr.

UVPD combined with MS is a helpful tool to study the electronic structures of ions in the gas phase.[34,35] The work of Matthews and Dessent demonstrated that the two protomers of the protonated 4-aminobenzoic acid (PABA) system (see Figure 3.2a), which were promoted by two different ESI solvents, were distinguishable using a laser-interfaced mass spectrometer, as their UV photodepletion spectra illustrated different band characters.[34] Marlton *et al.* applied a FAIMS-MS device in tandem with UVPD and successfully assigned the protomers of protonated nicotine (see Figure 3.2b) in the gas phase based on the distinct electronic transitions of the [nicotine + H]$^+$ protomers.[35] The spectroscopic properties of the aromatic amino acids have also been investigated using tandem MS-UVPD. The aromatic amino acids are spectroscopically accessible to most tunable light sources; since they can undergo intense $\pi$ to $\pi^*$ transitions following absorption of a UV photon.[30–33] The electronic spectra of the protonated aromatic amino acids in the UV range

have been explored using photofragment mass spectrometry by several teams.[30,32,33,92] Boyarkin *et al.* reported the electronic spectra of protonated Trp and Tyr obtained at both ~298 K and at 10 K, and found that the cooler ion trap led to higher spectral resolution since low temperatures prevent the appearance of additional hot-band features and suppress spectral broadening by restricting vibrational and rotational motion.[33] Pereverzev *et al.* measured the photofragmentation spectra of protonated Trp and Tyr with an Orbitrap mass analyzer and DUV/UV optical parametric oscillator (OPO). Protonated aromatic amino acids were shown to absorb light in the region of 190 to 290 nm, with significant absorption observed at ~280 nm for protonated Trp, ~230 and ~275 nm for protonated Tyr.[30]

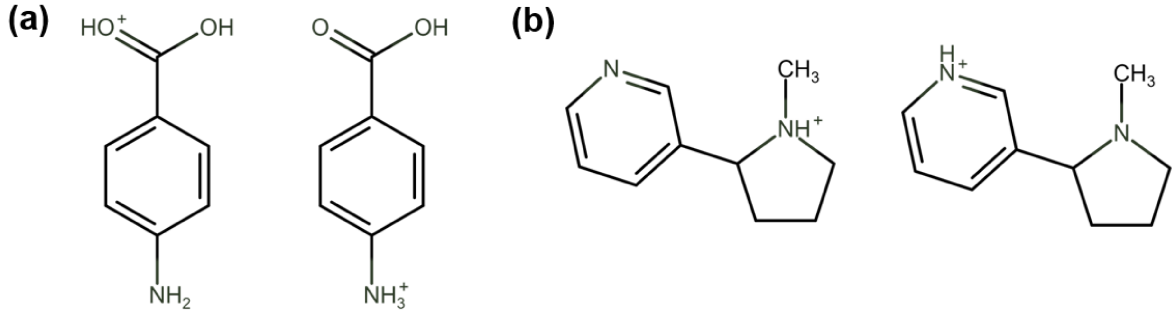


**Figure 3.2** The structures of (a) protomers of $[PABA + H]^+$ and (b) protomers of $[nicotine + H]^+$.

In this project, we explore the spectroscopic measurement of DMS-MS gated aromatic amino acid ions. The individual ion populations of targets are spatially separated in the DMS stage of the experiment before mass-selection in the first quadrupole (Q1), as shown in Figure 3.3. These ions are then trapped in the third quadrupole (Q3), where they are irradiated with tunable light from an Nd: YAG-pumped OPO. The conformation of an ion in the gas phase was identified by comparing the relative Gibbs energies. Electronic transitions were modelled using VG|FC factors for the analysis of experimental UVPD spectra. Electron density transfer during electronic transitions was visualized by computing the electronic orbitals. Bond dissociation energies were discovered to provide predictions of the fragmentation pathways during the UVPD process.

## 3.2 Methods

### 3.2.1 Experimental methods

Dried and solid tryptophan (99% purity) and tyrosine (99% purity) were purchased from Fisher Scientific and Sigma, respectively. The experimental setup employed in this project is illustrated in Figure 3.3. Samples of Trp and Tyr with concentrations of ~300 ng/mL were directed towards the inlet of a SelexION differential mobility spectrometer from an electrospray ion source. Dry carrier gas ($N_2$) carried the nascent ions through the DMS and towards a QTRAP 5500 mass spectrometer (DMS-MS), of which the gas-phase environment can be changed by adding volatile liquids. The output of an Nd: YAG-pumped OPO was directed toward the DMS-MS system such that the beam could pass through an optical port at the back of the QTRAP mass spectrometer and interact with trapped ions in Q3. The electrospray ionization (ESI) solvents and modifiers applied in this experiment are tabulated in Table 3.1.

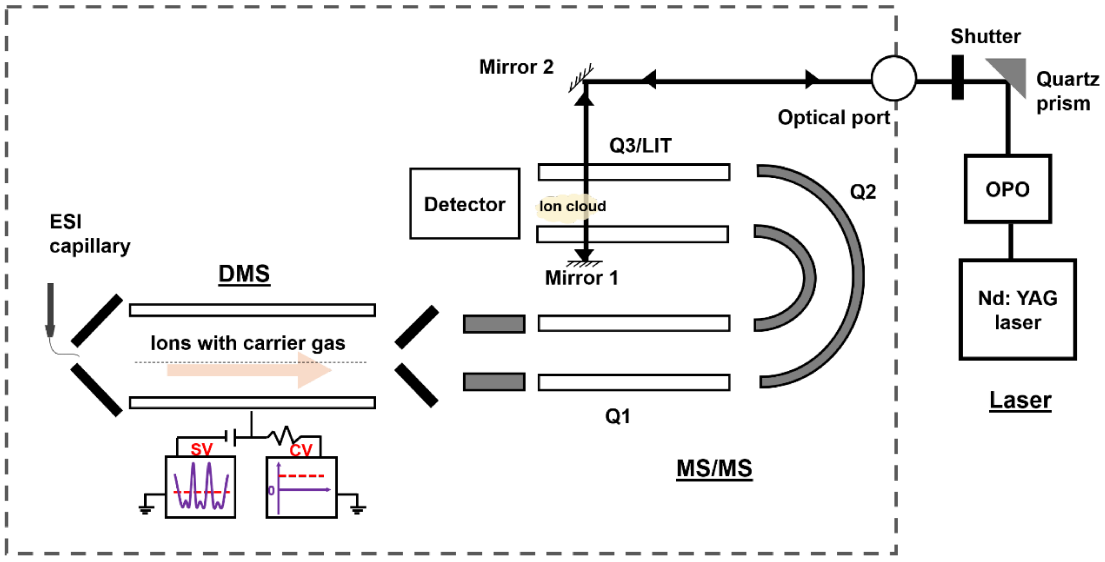


**Figure 3.3** A schematic diagram of the experimental setup, where electrospray ionization (ESI) source, first quadrupole (Q1), second quadrupole (Q2), third quadrupole (Q3), and linear ion trap (LIT) are labelled. DMS-MS is highlighted in the dashed frame.

**Table 3.1** ESI solvents and gas modifiers that were used in the DMS experiments of $[Trp + H]^+$.

| ***ESI solvents*** |
|---|
| *50:50 MeOH:$H_2O$ with 0.5% formic acid (MeOH/$H_2O$)* |
| *50:50 ACN:$H_2O$ with 0.5% formic acid (ACN/$H_2O$)* |
| ***Gas environments*** |
| *$N_2$* |
| *$N_2$ +1.5% (mole ratio) 2-propanol (IPA) vapour* |

In the DMS-MS experiments, the target ion was selected by its *m/z* in Q1, and the mass spectrum was recorded *via* mass axial ejection from Q3. The SV was scanned from 0 to 3000 V in 500 V increments, and from 3000 − 4000 V in 200 V increments. At each SV step, the CV was scanned from −60 to 25 V in 0.1 V increments. Mass-selected ion intensity was recorded as a function of SV and CV, yielding dispersion plots. The effect of the de-clustering potential (DP) was studied. The DP is a potential difference applied across the gap between the exit of the DMS cell and the Q-jet region and is used to de-cluster ions that have chances forming ion-solvent clusters as they enter the MS. In UVPD experiments, DMS-selected ions were mass-selected in Q1 before being trapped in Q3 for 3−5 ms. The wavelengths were scanned from 208 nm to 350 nm by 1 nm or 2 nm increments, which was chosen upon demands, and the intensities of the fragments and the parent ion in Q3 were recorded at each wavelength. The experimental UVPD spectra were generated by monitoring the fragmentation efficiency as a function of wavelength, as described in Equation (3.1):

$$I = ln\left(\frac{I_{parent,LasOn}}{I_{parent,LasOn}+\sum I_{frag,LasOn}}\right). \quad (3.1)$$

The intensity was then normalized with respect to the OPO output power, which was measured separately at each wavelength used in the experimental UVPD spectra. The process described by

Equation (3.1) does not account for parent ion fragmentation by collision induced dissociation (CID) in Q3. This background process simply results in the appearances of fragment ion signals in Q3 before irradiation. During the UVPD process, fragments generated by CID were dissociated by UV irradiation and simultaneously reproduced from parent ion dissociation in the trap. Hence, both the spectra of the parent ion and daughter ions formed by CID were taken by gating on the respective masses in Q1 and were compared to ensure that we were observing the pure vibronic spectra of the parent ion. All experiments were repeated at least twice for reproducibility.

A challenging aspect of this experiment was the inability to discern fragmentation as a result of parent ion dissociation or subsequent dissociation of daughter ions, which can introduce unexpected spectral features to the parent ion's UVPD spectrum when not accounted for in the data processing. As introduced in Chapter 2, the duration of vibronic transitions depends on the time of electron transfer and vibrational motion, which occur on timescales of femtoseconds and picoseconds, respectively. The timescale of vibronic transitions is much shorter than that of the pulse width (several nanoseconds), introducing the possibility of multiple photon excitation. This provides chances for resulting fragment ions to absorb photons leading to secondary fragmentation. Alternatively, the parent ion could fragment into a vibrationally-hot fragment ion, which can further dissociate with sufficient energy into other fragment ions. With our existing technology, it is difficult to determine whether fragmentations were rearranged from the parent ion directly or from the daughter ion *via* different dissociation pathways during the UVPD process.

### 3.2.2 Computational methods

Molecular models of the targeted molecular ions were generated manually using the Gaussview software.[93] The structures of the target ions were optimized at the B3LYP/6-31G(d,p) level of theory using the DFT method, and atomic partial charges were calculated with the charges

from electrostatic potentials using a grid-based method (CHelpG) at the same level of theory with Gaussian 16.[93] The optimized geometries and calculated partial charges were used to generate input files for a BH routine, as discussed in Chapter 2. A BH routine included 10,000 steps, which results in 10,000 structures being sampled for each molecule. Since the target ions could cluster with solvent molecules or form dimers in the DMS experiments, both monomer and cluster/dimer calculations were conducted. For the monomer, each internal dihedral angle was rotated by a random value between $-5^{\circ} \leq \alpha \leq 5^{\circ}$. For cluster calculations, additional parameters were modified, in which each moiety was rotated around its center of mass by a random quantity $\beta$ chosen from $-4^{\circ} \leq \beta \leq 4^{\circ}$ and translated by a random step size $\eta$ from $-0.5$ Å $\leq \eta \leq 0.5$ Å. The unique geometries obtained from BH calculations were optimized at the PM6 level of theory first, and then re-optimized with the DFT method at the CAM-B3LYP/6-311++G(d,p) level of theory.[44,94] Structures of protomers of an ion in the water solution were estimated by optimizing clusters consisting of the protomer and water molecules manually located at the hydrogen bonding sites using a $H_2O$ polarizable continuum model (PCM) at the CAM-B3LYP/6-311++G(d,p) level of theory. Thermodynamic corrections were performed to determine the Gibbs energy at the same level of theory based on the optimized structures in the final step. The global minimum structures of each protomer of desired ions were used to calculate the vibronic spectra using the VG|FC approximation in the ORCA suite.[95] First, the Hessian file of the ground state, which is essential for computing the VG|FC factors, was calculated at the CAM-B3LYP/6-311++G(d,p) level of theory. At the same level, electronic transitions between the ground state and each singlet excited state (*i.e.*, $S_1$, $S_2$, $S_3$…) were computed based on the ground state geometry of the global minimum with TD-DFT calculations, in which the VG|FC approximation was combined to determine the vibrational wavefunction overlap between the ground state and relevant excited state. The

computed spectra were generated with the sum of the intensities of all the calculated electronic transitions (*i.e.*, $S_1 \leftarrow S_0$, $S_2 \leftarrow S_0$, $S_3 \leftarrow S_0$…) as a function of wavelength. Triplet excited states were not included when simulating vibronic spectra as electronic transitions are more likely to occur between singlet states; additional electronic changes involved in singlet-to-triplet transitions can lead a longer lifetime compared to that of single-to-single transitions.[96]

## 3.3 Results and discussion

### 3.3.1 Protonated tryptophan

#### 3.3.1.1 Experimental results for $[Trp + H]^+$

The clustering behaviours of $[Trp + H]^+$ were studied with various ESI solvents and different gas environments, as tabulated in Table 3.1. Note that the DP applied for all DMS experiments of $[Trp + H]^+$ was 100 V. When $[Trp + H]^+$ was sprayed with ACN/$H_2O$ solvent, the dispersion plots (see Figure 3.4) indicated that only one conformation was resolved under the pure $N_2$ and IPA environments. Different clustering behaviours, as illustrated by the shape of the dispersion plots, were observed under different gas-phase environments. $[Trp + H]^+$ showed non-clustering (Type C) behaviour with pure $N_2$ gas and strong clustering (Type A) behaviour when IPA vapour was added. Conversely, two dispersion curves were observed when MeOH/$H_2O$ was used as the ESI solvent in pure $N_2$ (see Figure 3.5a and 3.5c). The dispersion plots of $[Trp + H]^+$ electrosprayed from MeOH/$H_2O$ solution in both pure $N_2$ and IPA were recorded and are compared in Figure 3.6.

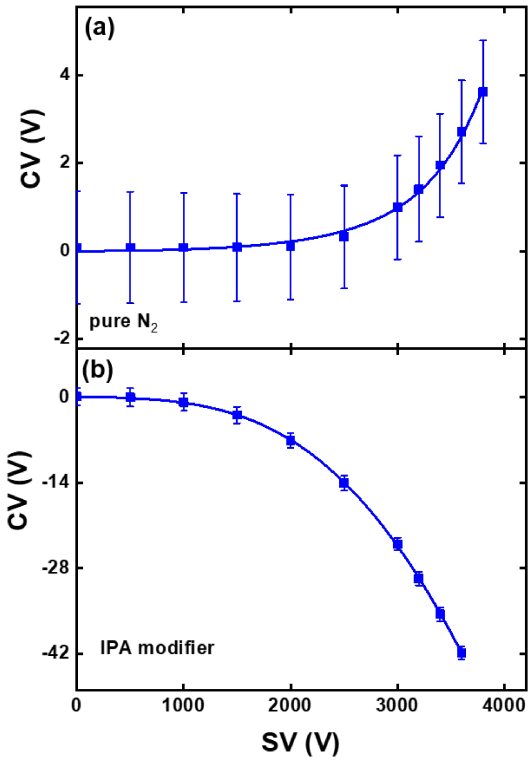


**Figure 3.4** The dispersion plots of [Trp + H]+ electrosprayed from ACN/$H_2O$ in (a) $N_2$ gas and (b) $N_2$ gas with the addition of 1.5% IPA vapour.

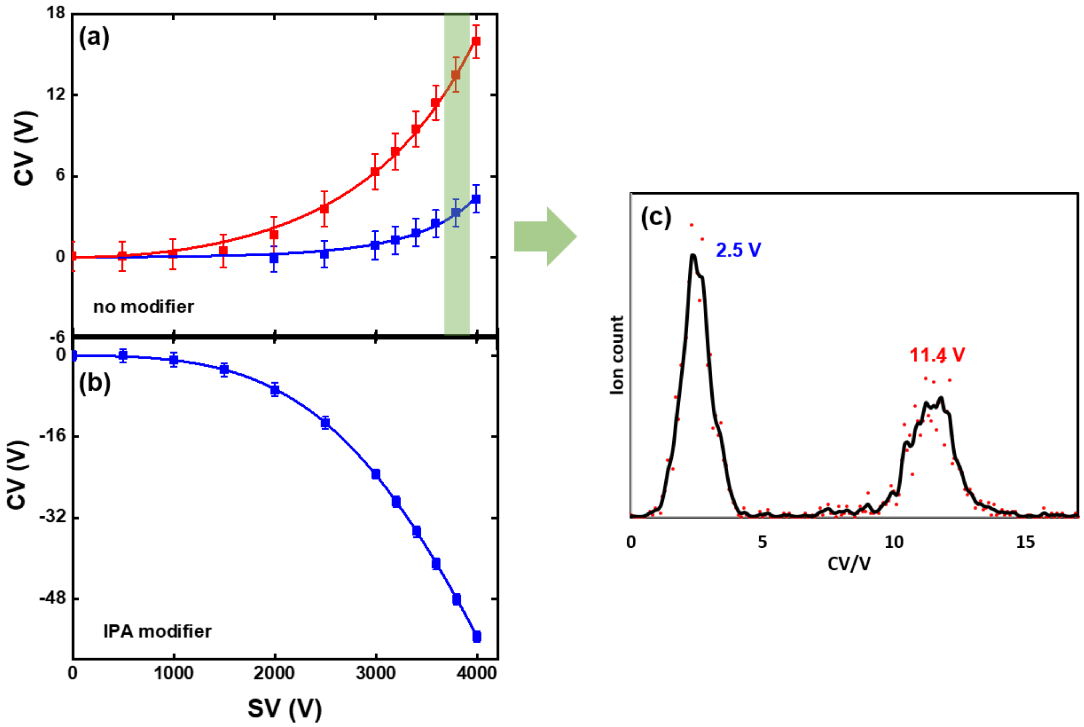


**Figure 3.5** The dispersion plots of [Trp + H]+ electrosprayed from MeOH/$H_2O$ solution in (a) $N_2$ gas and (b) $N_2$ gas-phase seeded with 1.5% IPA vapour. (c) The ionogram of $[Trp + H]^+$ electrosprayed from MeOH/$H_2O$ solution in $N_2$ gas at SV=3600 V.

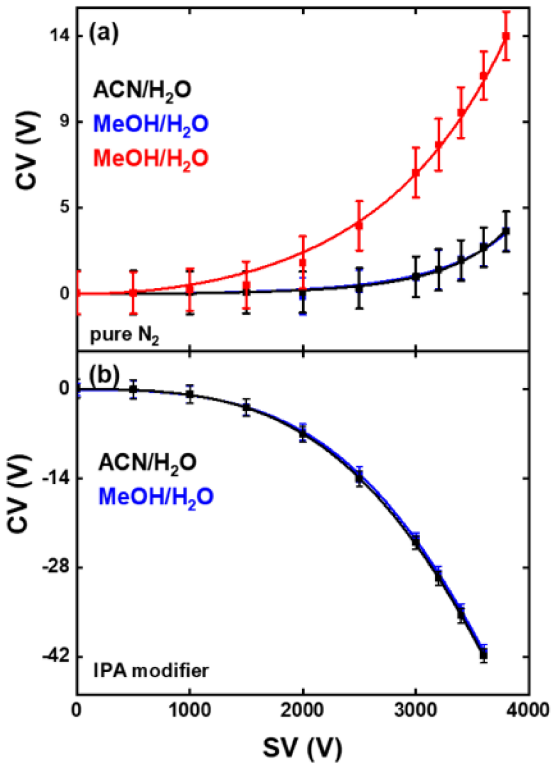


**Figure 3.6** The dispersion plots of $[Trp + H]^+$ electrosprayed from $MeOH/H_2O$ and $ACN/H_2O$ solutions in (a) $N_2$ gas and (b) $N_2$ gas seeded with 1.5% IPA.

Figure 3.6a shows the dispersion plots of $[Trp + H]^+$ using the different ESI solvents. One of the dispersion curves (blue) observed in the pure $N_2$ environment when sprayed from $MeOH/H_2O$ perfectly overlaps the Type C curve observed when sprayed from $ACN/H_2O$. The same dispersion plot of $[Trp + H]^+$ was observed in the IPA-modified carrier gas irrespective of the two ESI solvents, as shown in Figure 3.6b. As discussed in section 2.1, the same dispersion curves imply that identical differential ion mobility was detected in the DMS cell for the ion populations associated with the curves. The same ion mobility implies indistinguishable ion conformations, indicating structures associated with those two overlapped curves were identical.

The observation of multiple curves in the dispersion plot recorded following electrospray from $MeOH/H_2O$ could indicate one of two things: (1) the existence of different protomers of $[Trp + H]^+$ in the probed ensemble, or (2) the presence of (at least one) cluster that contains Trp and produces $[Trp + H]^+$ upon dissociation following transit through the DMS cell. As ion-solvent

clusters enter the vacuum cavity of the MS from the atmospheric conditions of the DMS cell, ion de-solvation typically occurs. While this is beneficial for quantitative analysis by MS, it introduces ambiguity with respect to distinguishing bare ions from ionic clusters that fragment to produce the bare ion or another species of the same nominal *m/z*. The mass spectra for the two DMS-separated species at CV=2.5 V and 11.4 V (SV=3600 V) were recorded as DP was scanned from 0 to 300 V in 5 V increments to determine the identity of the second dispersion curve (red) – this is shown in Figure 3.7. The precursor ion scan in Q1 for the DMS-gated ion population at CV=2.5 V demonstrates ion masses up to *m/z*=205, indicating that larger clusters were not fragmenting to yield $[Trp + H]^+$. However, for the ion population selected at CV=11.4 V, the signal of *m/z*=409 appeared, implying the presence of a protonated Trp dimer, $[(Trp)_2 + H]^+$, formed in the ESI process. At higher DP, the intensity of the *m/z*=409 signal decreased, while that of *m/z*=205 increased, indicating that $[(Trp)_2 + H]^+$ was fragmenting to produce $[Trp + H]^+$. These results indicate that both the protonated Trp monomer and dimer are formed when electrosprayed from $MeOH/H_2O$ and that these species adopt significantly different DMS behaviour in a pure $N_2$ environment.

**Figure 3.7** Precursor ion scans for $[Trp + H]^+$ electrosprayed from MeOH/$H_2O$ solutions and selected at CV=2.5 and 11.4 V at DP=0, 100 V, 200 V, 300 V.

For further investigation, the dispersion plot of $[(Trp)_2 + H]^+$ (*m/z*=409 selected in Q1) generated by ESI of MeOH/$H_2O$ solution and measured in a pure $N_2$ environment was generated and compared with that of $[Trp + H]^+$ in Figure 3.8. The DMS results of $[(Trp)_2 + H]^+$ indicate the presence of only one conformation of $[(Trp)_2 + H]^+$ in the gas phase (or potentially multiple unresolved conformers) that exhibits non-clustering behaviour. The similar (identical within error) dispersion curves of ions selected at *m/z*=409 (black) and *m/z*=205 (red) indicate that the second ion population detected for $[Trp + H]^+$ (red curve Figure 3.6) was associated with dissociation of $[(Trp)_2 + H]^+$.

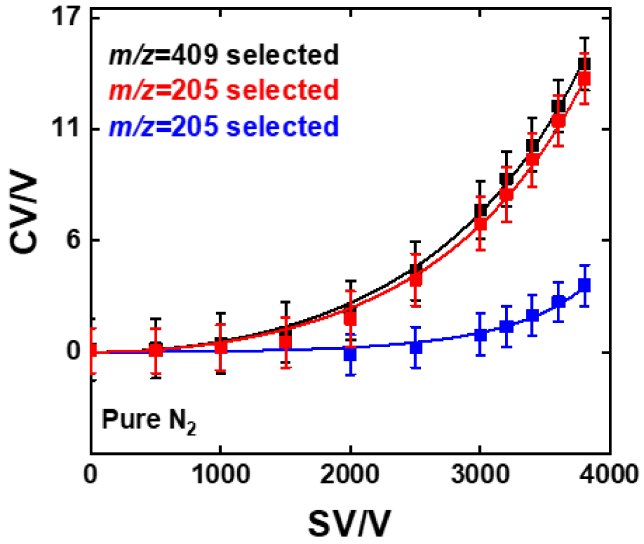


**Figure 3.8** The dispersion plots of $[(Trp)_2 + H]^+$ and $[Trp + H]^+$ electrosprayed from MeOH/$H_2O$ solution in pure $N_2$.

UVPD experiments were performed to characterize the electronic structures of the two ion populations of $[Trp + H]^+$ that resulted from the monomeric species of protonated Trp and the protonated Trp dimer in Q3 in pure $N_2$. During the UVPD process, the same photodissociation fragmentations, as tabulated in Table 3.2, were observed in Q3 for both ion populations of $[Trp + H]^+$ selected at CV=2.5V and 11.4 V as SV=3600 V (Figure 3.5c) in the DMS cell when gating on *m/z*=205 in Q1. The mass distribution of the parent ion (*m/z*=205) in Q3 with the laser off is shown in Figure 3.9.

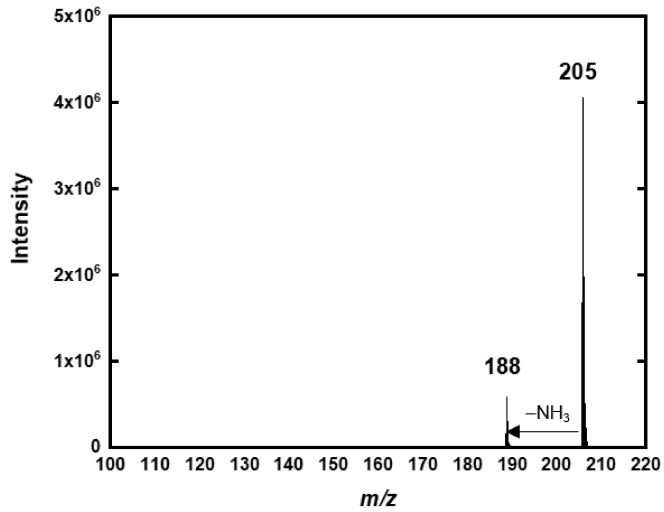


**Figure 3.9** The Q3 mass spectrum of $[Trp + H]^+$ as *m/z*=205 was selected in Q1 and subsequently underwent CID with background gases.

We observed background CID in Q3, leading to a loss of $NH_3$ to give ions of *m/z*=188. The UVPD spectrum of the parent ion generated with the 188 fragmentation channels showed unexpected features (see Figure A.2 in Appendix A) due to the non-zero background, wavelength-dependent increases in fragment intensities according to UVPD of the parent ion, and wavelength-dependent depletion of fragment ions during UVPD (see Figure 3.11). As a result, the UVPD spectra of the two $[Trp + H]^+$ populations were generated with all the same fragmentations tabulated in Table 3.2 except for the *m/z*=188 fragment and are shown in Figure 3.10.[97] The UVPD spectra of $[Trp + H]^+$ monitored in each mass fragmentation channel (except for *m/z*=188) showed similar spectral features with the UVPD spectrum of $[Trp + H]^+$ − this can be found in Figure A.3 in Appendix A. The neutral fragment loss of $[Trp + H]^+$ at each fragmentation channel was cited from the fragmentation study of Zhang *et al.*[97]

**Table 3.2** The fragmentation channels of $[Trp + H]^+$ observed in UVPD process.[97]

| *m/z* | Neutral fragment(s) |
|---|---|
| 205 | - |
| 188 | $NH_3$ |
| 170 | $NH_3 + H_2O$ |
| 159 | $H_2O + CO$ |
| 146 | $NH_3 + CH_2CO$ |
| 144 | $NH_3 + CO_2$ |
| 132 | $H_2O + CO + HCN$ |
| 118 | $NH_3 + CH_2CO + CO$ |

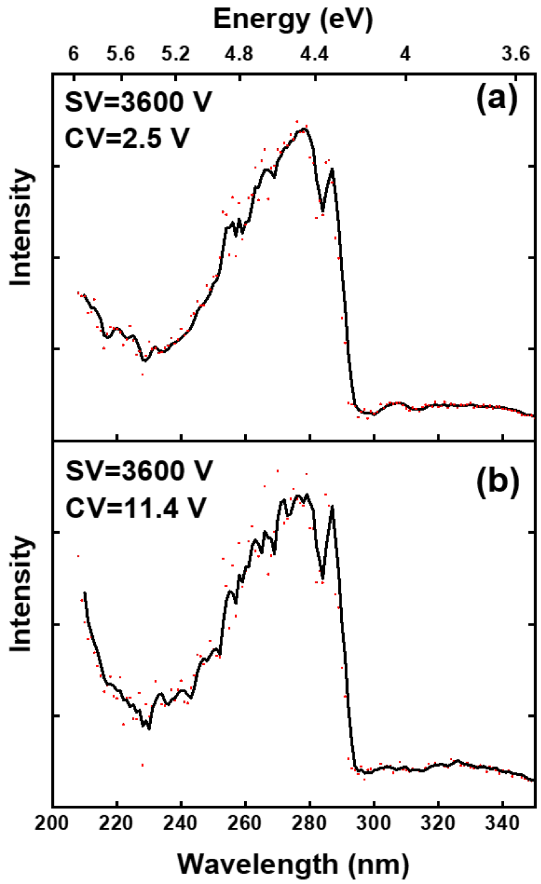


**Figure 3.10** The UVPD action spectra of $[Trp + H]^+$ electrosprayed from $MeOH/H_2O$ solution in $N_2$ gas with DP=100 V for **(a)** SV=3600 V, CV=2.5 V and **(b)** SV=3600 V, CV=11.4 V. The solid lines are three-point adjacent averages of the data points.

The same fragmentations were observed in both species during UVPD. The similarity of the two spectra indicates that the same electronic transitions in identical ion populations of $[Trp + H]^+$ were

probed *via* UVPD. This suggests that the $[Trp + H]^+$ fragment produced following in-source CID of $[(Trp)_2 + H]^+$ adopted an identical geometric structure to that of the bare ion population of $[Trp + H]^+$.

It is possible that the presence of the *m/z*=188 species, $[Trp - NH_3 + H]^+$, produced by CID in Q3 (see Figure 3.9) might contribute to unexpected artifacts to the UVPD spectrum of $[Trp + H]^+$. To quantify this, the UVPD spectrum was measured for the *m/z*=188 species following selection in Q1. The UVPD action spectra of $[Trp - NH_3 + H]^+$ and $[Trp + H]^+$ are overlaid in Figure 3.11. The spectrum of $[Trp - NH_3 + H]^+$ showed an intense electronic transition at ~245 nm and a band system from 300−350 nm that were not observed in the $[Trp + H]^+$ spectrum. The spectral difference between the *m/z*=205 and *m/z*=188 species indicates that the *m/z* 188 fragment generated by CID in Q3 does not contribute substantially to the measured UVPD spectrum of $[Trp + H]^+$.

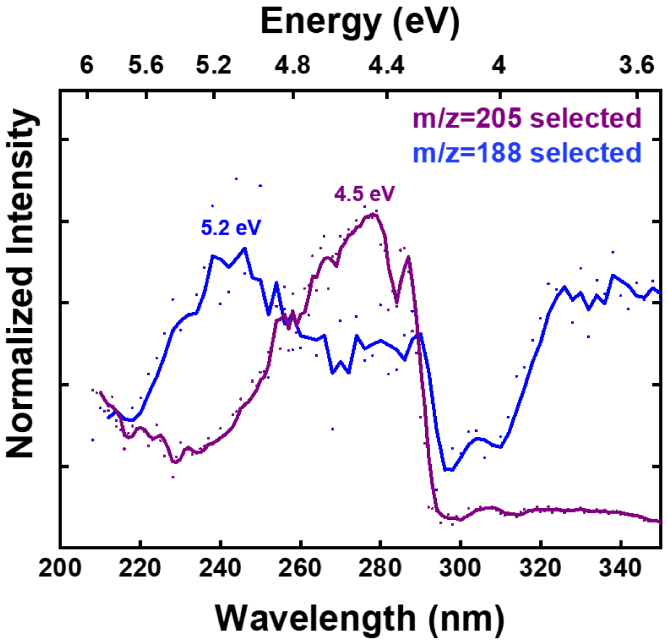


**Figure 3.11** The UVPD action spectra of $[Trp + H]^+$ (purple) and $[Trp - NH_3 + H]^+$ (blue). The solid lines were three-point adjacent averages of the data points.

The experimental UVPD action spectrum of $[Trp + H]^+$ was compared against the photofragmentation spectrum reported by Pereverzev *et al.*, which also considered all major charged fragments, as shown in Figure 3.12.[30] The experimental $[Trp + H]^+$ spectra of this experiment and Pereverzev *et al.* showed similar spectral shapes between 220 to 290 nm. A maximum intensity was observed at ~275 nm for the spectra obtained in the scope of this thesis. Pereverzev *et al.*, however, recorded the most intense spectral band within the range from 208 nm to 215 nm. The current experimental setup described in this thesis limits UVPD investigation to the 208 − 400 nm range; thus, UVPD findings below 208 nm could not be compared. While an intense transition from 208 nm to 215 nm was observed in the experimental findings of Pereverzev *et al.*, it cannot be conclusively stated that this difference is significant, as other factors (*i.e.*, instability of initial light signals) may contribute to this discrepancy. Despite those ambiguities, the lack of features from 208 nm to 215 nm of our experimental spectrum might result from the omitted 188 fragmentation channel when generating the parent spectrum, as indicated in Figure 3.12. An important note is that the observed transition from 208 to 215 nm in the reported spectrum by Pereverzev *et al.* and in our experimental spectrum monitored in the *m/z*=188 fragmentation channel matched with the calculated vibronic spectra reported in this thesis − this will be discussed further in the next section. It is concluded that the similar electronic transitions observed between 220 and 290 nm in both the experimental and literature spectra indicate that the same protomer of $[Trp + H]^+$ may be detected in both experiments.

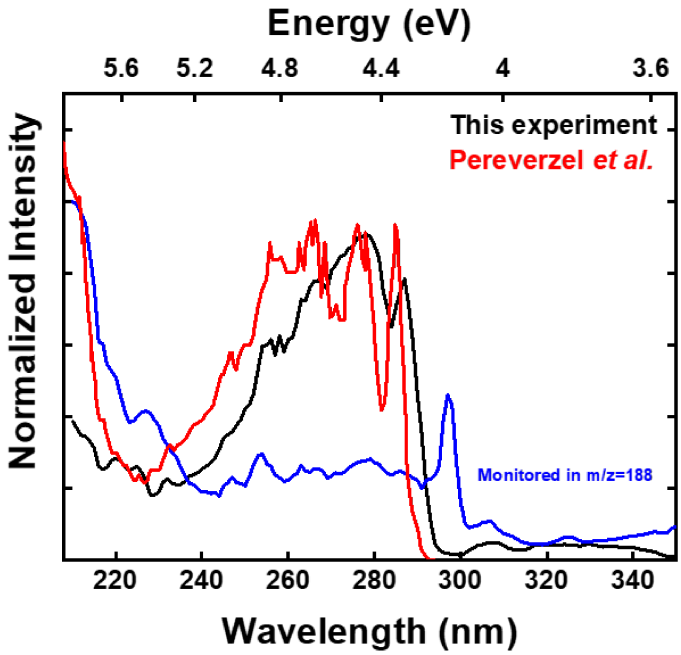


**Figure 3.12** The UVPD action spectra of our experiment (black) and Pereverzev *et al.* (red). Figure adapted from Pereverzev *et al.*[30]

### 3.3.1.2 Computational results for $[Trp + H]^+$

The electronic structures and transitions of protomers of protonated Trp were computationally investigated to assign the experimentally acquired action spectra. According to the resonance effects, the addition of proton on the oxygen of the −OH of the carboxylic group would lead to fragmentation (*i.e.*, loss of water) of the protonated Trp. Therefore, three likely protonation sites are considered for Trp, as shown in Figure 3.1: the amine nitrogen (N-amine), the indole nitrogen (N-indole), or the carboxyl oxygen (O-carboxyl). After BH and optimization at the CAM-B3LYP/6-311++G(d,p) level of theory, the lowest energy structures of $[Trp + H]^+$ with different protonation sites were obtained and are shown in Figure 3.13.

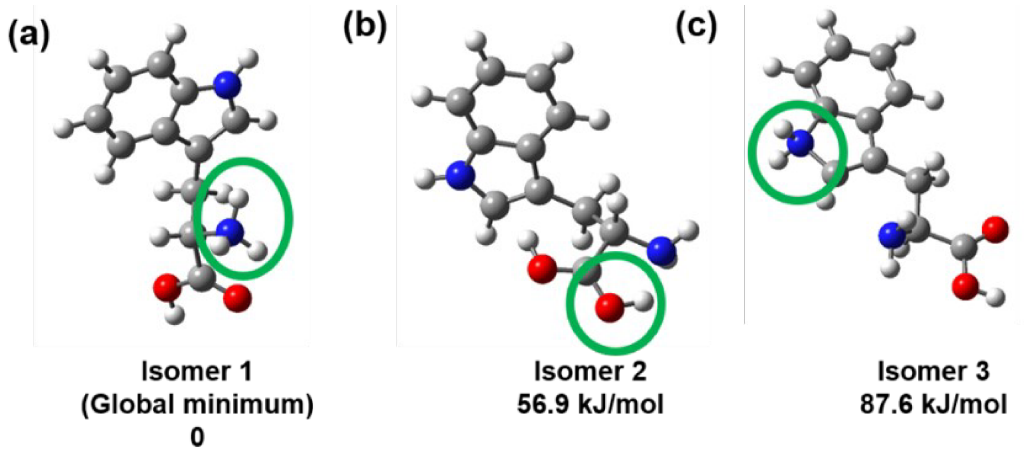


**Figure 3.13** The optimized structures of $[Trp + H]^+$ with protonation sites on (a) N-amine, (b) O-carboxyl, and (c) N-indole. The optimized structures and relative Gibbs free energy to that of the global minimum structure were determined at the CAM-B3LYP/6-311++G(d,p) level of theory.

The geometry of the N-amine protonated Trp had the lowest Gibbs energy among these three prototropic isomers. The relative Gibbs energies of isomers 2 and 3 to that of isomer 1 were higher than +50 kJ/ mol. Structures of the $[Trp + H]^+$ protomers in the water solution were estimated by optimizing clusters consisting of $[Trp + H]^+$ and water molecules manually located at the hydrogen bonding sites. The structures and Gibbs energies of three cluster systems containing three different $[Trp + H]^+$ protomers calculated at the CAM-B3LYP/6-311++G(d,p) can be found in Figure A.1 in Appendix A. Similarly, high relative Gibbs energies of isomer 2 and isomer 3 to that of isomer 1 (*i.e.*, >30 kJ/mol) were observed in the water solution. Assuming a thermodynamic equilibrium described by Boltzmann statistics, population of isomer 1 is likely to dominate the gas phase ensemble of $[Trp + H]^+$, whereas population of isomers 2 and 3 is expected to be negligible.

The vibronic spectra were calculated for each $[Trp + H]^+$ protomer (*i.e.*, isomers 1, 2, and 3) at the CAM-B3LYP/6-311++G(d,p) level of theory for comparison with the experimental UVPD spectrum of $[Trp + H]^+$. The computed spectra, which were red-shifted by ~30 nm for a better match with the experimental results, are shown in Figure 3.14. This energy shift between experimental and computational spectra can be rationalized by the limitations of the DFT method,

as discussed in Chapter 2. Besides, individual electronic transitions of the global minimum (isomer 1) (*i.e.*, $S_1 \leftarrow S_0$, $S_2 \leftarrow S_0$, $S_3 \leftarrow S_0$…) were compared with the experimental UVPD spectrum, as shown in Figure 3.15.

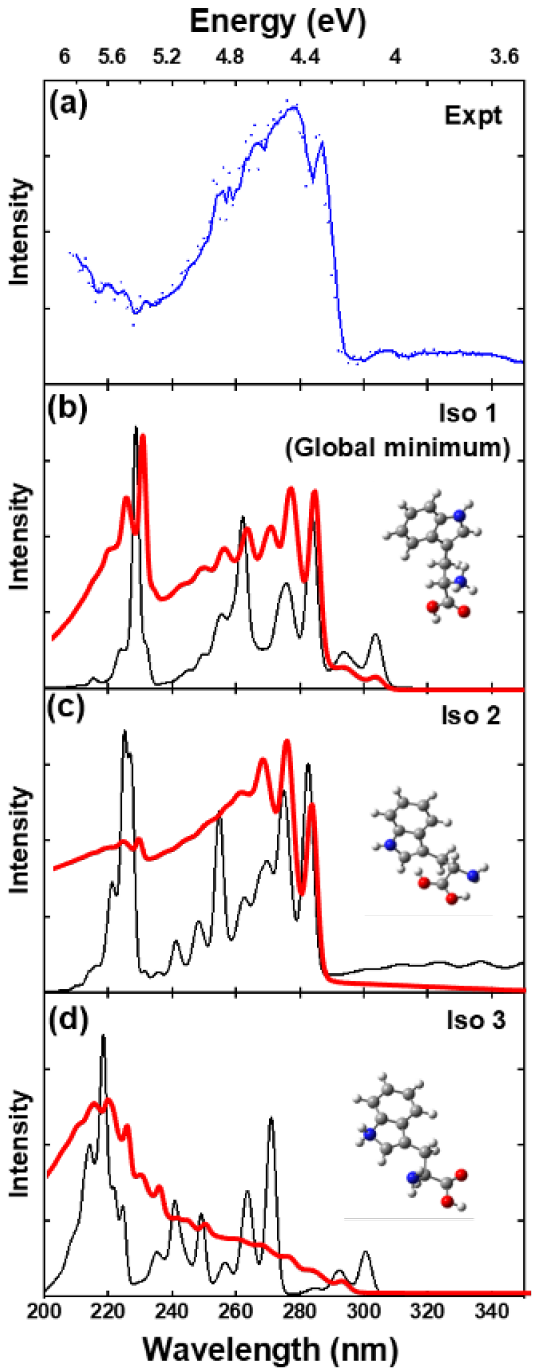


**Figure 3.14** (a) The experimental spectrum of $[Trp + H]^+$. The calculated vibronic spectra of (b) isomer 1, (c) isomer 2, and (d) isomer 3 of $[Trp + H]^+$ using the VG|FC approximation. Red lines are the vibronic spectra computed with VG|FC approximation and Duschinsky rotation.

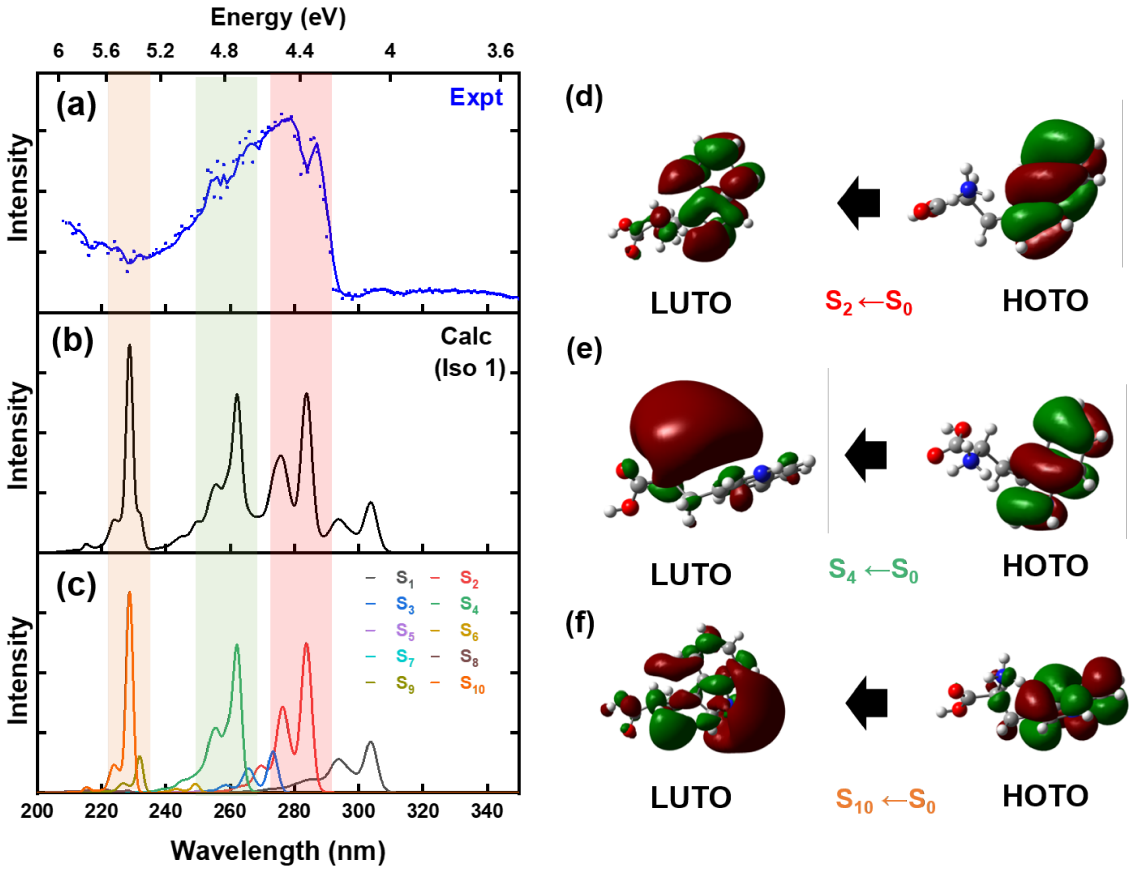


**Figure 3.15** (a) The experimental UVPD spectrum of $[Trp + H]^+$. The calculated vibronic spectra of $[Trp + H]^+$ of (b) a sum of electronic transitions from $S_0$ to $S_1$ to $S_0$ to $S_{10}$. (c) The calculated vibronic spectra of each electronic transition of $S_0$ to $S_1$, $S_2$ … $S_{10}$. The calculated LUTOs and HOTOs of (d) $S_0$ to $S_2$, (e) $S_0$ to $S_4$, and (f) $S_0$ to $S_{10}$.

The calculated vibronic spectra of the lowest energy conformers for the three prototropic motifs of $[Trp + H]^+$ (Figure 3.13) all exhibited light absorption in the wavelength range of 180 nm to 300 nm. The three calculated spectra showed similar band shapes for the predicted electronic transitions, and the computed spectra were all in decent agreement with the experimental spectrum. However, we have observed significant spectra shifts between experiment and DFT calculations; higher level calculations may be necessary. Based on the VG|FC calculations, the best match to the observed spectrum is the spectrum calculated for isomer 2. The major difference between the calculated spectra for isomer 1, 2, and 3 is the lack of intensity for the $S_1 \leftarrow S_0$ transition (280−350 nm) of isomer 2. Whether this transition is of low intensity or, instead, is shifted to

higher energy is currently under investigation. Despite these ambiguities, it is concluded that the intensity observed in the 300−340 nm region is not due to the photofragmentation of the parent ion; rather, it seems to be associated with the formation of $[Trp - NH_3 + H]^+$ (*via* CID in Q3) and its subsequent photodissociation, as indicated in Figure 3.11. In addition, the calculated spectra of isomer 1 and 2 are both reasonable after applying Duschinsky rotation. The intense transitions were observed at ~260 nm (4.6 eV) and ~280 nm (4.4 eV) in the experimental spectrum of $[Trp + H]^+$, which corresponded to the calculated $S_2 \leftarrow S_0$ and $S_4 \leftarrow S_0$ electronic transitions, respectively. The highest occupied transition orbital (HOTO) and lowest unoccupied transition orbital (LUTO) of the electronic transition from the ground state ($S_0$) and the second singlet excited state ($S_2$) shown in Figure 3.15d demonstrated that the $S_2 \leftarrow S_0$ excitation was mainly localized on the phenyl ring of the $[Trp + H]^+$ involving a $\pi^* \leftarrow \pi$ transition; the HOTO and LUTO involved in $S_4 \leftarrow S_0$ transition (Figure 3.15e) illustrated that $\sigma^* \leftarrow \pi$ characters were observed for the electronic transition observed at 4.8 eV. Moreover, the intense band shown in our experimental spectrum generated with the *m/z*=188 fragmentation channel and in the spectrum reported by Pereverzev *et al.* at 208 nm, as shown in Figure 3.12, is likely associated with the $S_{10} \leftarrow S_0$ electronic transition.[30]

Computational studies were also conducted for $[(Trp)_2 + H]^+$ to support the experimental findings. The structures of $[(Trp)_2 + H]^+$ associated with different protomers of $[Trp + H]^+$ were obtained from BH and then optimized at the CAM-B3LYP/6-311++G(d,p) level of theory. The lower energy structures of the protonated Trp dimer and their relative Gibbs energies are shown in Figure 3.16. In all cases, protonation was found to occur on the amine nitrogen of one of the Trp moieties. The binding sites involved in the lower energy isomers of $[(Trp)_2 + H]^+$ included hydrogen-bonding networks formed between the N-amine and O-carboxyl or the N-amine and N-amine groups, as indicated by the green dashes. The calculated lower energy isomers of

protonated Trp dimer demonstrated that the structure of $[(Trp)_2 + H]^+$ was stabilized by the proton on the N-amine, aligning with the computational findings for the $[Trp + H]^+$ monomer. The N-amine protonated $[Trp + H]^+$ found in the calculated structures of both the protonated Trp monomer and dimer also matched the experimental results that $[(Trp)_2 + H]^+$ produced the same $[Trp + H]^+$ protomer as that of the monomer.

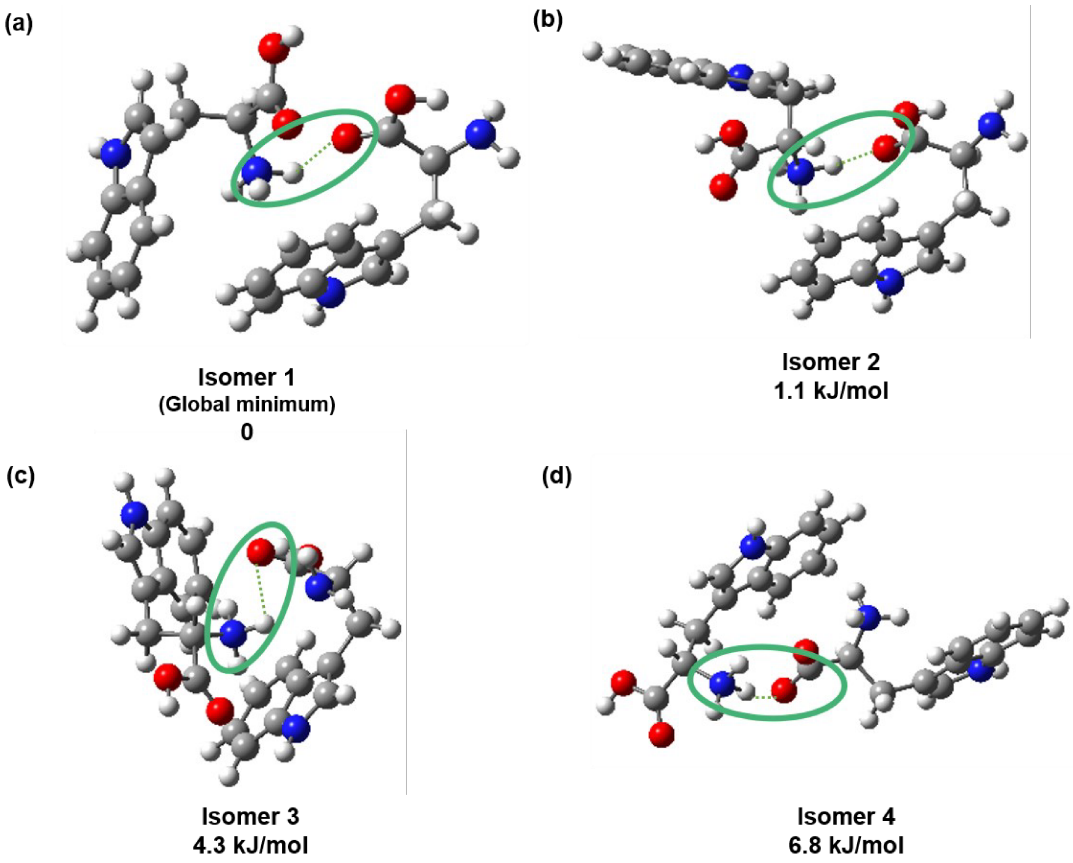


**Figure 3.16** The optimized structures of $[(Trp)_2 + H]^+$ isomers with Gibbs energies <10.0 kJ/mol. The optimized structures and Gibbs energy to that of the global minimum structure were determined at the CAM-B3LYP/6-311++G(d,p) level of theory. Hydrogen bonds are indicated with green dash lines.

However, the structure of the protonated moiety of $[(Trp)_2 + H]^+$ might change during the course of dissociation. To further investigate the dissociation process of the protonated dimer, the bond dissociation energy (BDE) thresholds of $[(Trp)_2 + H]^+$ producing protomers of $[Trp + H]^+$ and neutral Trp were calculated with Equation (3.2) and tabulated in Table 3.3

$$BDE = E_{[Trp+H]^+} + E_{Trp} - E_{[(Trp)_2+H]^+} \quad (3.2)$$

**Table 3.3** The calculated BDE thresholds of the protonated Trp dimer producing isomer 1, 2, and 3 at CAM-B3LYP/6-311++G(d,p) at T=298 K.

| | Production of isomer 1 | Production of isomer 2 | Production of isomer 3 |
|---|---|---|---|
| Calculated BDE (kJ/mol) | 101.54 | 204.08 | 235.82 |
| Calculated BDE (eV) | 1.05 | 2.12 | 2.44 |

The calculated BDEs of the protonated Trp dimer indicate that approximately 1 eV is required to produce isomer 1, and that this threshold is approximately 1 eV below the next lowest threshold. During experimental measurements, the collision energy (CE) was set at 10 V, which is just above the minimum of 5 V, to ensure optimal operation and stability. When CE is applied at 10 V, the estimated center of mass CE ($CE_{CoM}$) between the $N_2$ CID partner and the $[(Trp)_2 + H]^+$ ion is 0.61 eV, which is lower than all three BDE thresholds of $[(Trp)_2 + H]^+$ in production of $[Trp + H]^+$ at 298 K. However, $[(Trp)_2 + H]^+$ is not observed experimentally in Q3, despite the $CE_{CoM}$ being lower than the calculated threshold. This could be due either to inaccuracy of the calculations or, more likely, heating of the ions in *situ* leading to lower Gibbs corrected BDEs. For instance, heating of the $[(Trp)_2 + H]^+$ ion could result from applying a de-clustering potential (100 V) between the DMS and MS interface. This results in a significant reduction in its binding energy (*i.e.*, smaller than 0.61 eV). Accordingly, the temperature effects on BDE thresholds of $[(Trp)_2 + H]^+$ dissociation were studied and are shown in Table 3.4. The calculated BDE threshold drops to 0.59 eV as the temperature of ions is increased from 298 to 600 K. It suggests that the protonated Trp dimer ions were heated to above 600 K when transmitting from the DMS cell to Q1 at DP=100 V. The local heating in addition to the energy supplied in the collision cell align with computed BDE threshold for the dissociation of $[(Trp)_2 + H]^+$ into $[Trp + H]^+$.

**Table 3.4** The calculated BDE thresholds of the protonated Trp dimer producing isomer 1 at CAM-B3LYP/6-311++G(d,p) at various temperatures.

| Temperature (K) | 298 | 400 | 500 | 550 | 575 | 600 | 625 | 650 |
|---|---|---|---|---|---|---|---|---|
| BDE (eV) | 1.05 | 0.89 | 0.74 | 0.66 | 0.62 | 0.59 | 0.55 | 0.51 |

### 3.3.2 Protonated tyrosine

#### 3.3.2.1 Experimental results for $[Tyr + H]^+$

The dispersion plot recorded in $N_2$ for $[Tyr + H]^+$ following ESI from a 50:50 MeOH:$H_2O$ solution with 0.5% FA is shown in Figure 3.17. Weak clustering behaviour (Type B) was observed for $[Tyr + H]^+$ when sprayed from a methanolic solution. The DMS results showed only a single dispersion curve for $[Tyr + H]^+$, which suggest that either (a) only one protonation state of $[Tyr + H]^+$ was found in the gas phase, or (b) protomers of $[Tyr + H]^+$ were not resolved under the employed conditions.

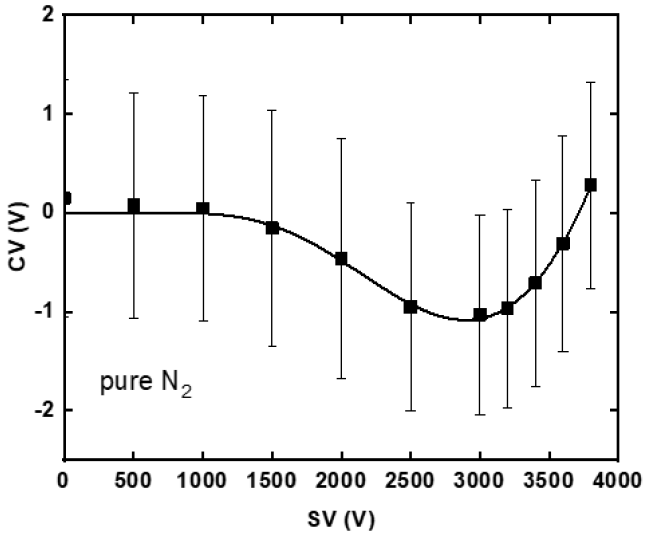


**Figure 3.17** The dispersion plot of $[Tyr + H]^+$ electrosprayed from 50:50 MeOH:$H_2O$ with 0.5% FA solution in $N_2$ gas.

UVPD experiments were performed to measure the electronic spectrum of the gas phase $[Tyr + H]^+$ ensemble. The $[Tyr + H]^+$ ions were selected at CV= −1.0 V and SV=3000 V in the DMS cell, and *m/z*=182 (*i.e.*, parent ion mass) was selected in Q1 prior to trapping and irradiation

in Q3. The mass spectrum observed following trapping of $[Tyr + H]^+$ in Q3 (but prior to irradiation) is plotted in Figure 3.18.

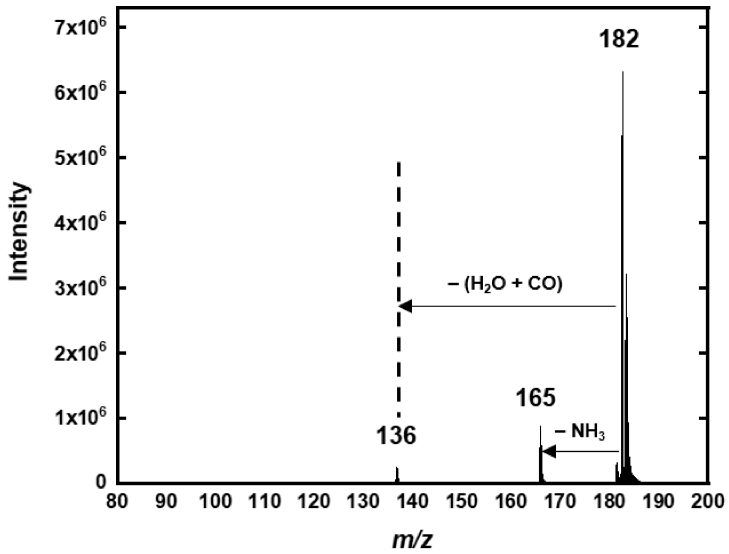


**Figure 3.18** The Q3 mass spectrum of $[Tyr + H]^+$ as *m/z*=182 was selected in Q1 and subsequently underwent CID with background gases.

Two major product channels, formation of *m/z*=136 and 165, are observed for $[Tyr + H]^+$ following fragmentation due to collision with background gases in Q3. The UVPD spectrum of the parent ion generated with those two fragmentation channels was unexpected (see Figure A.6 in Appendix A) due to the non-zero background, wavelength-dependent increases in fragment intensities according to UVPD of the parent ion, and wavelength-dependent depletion of fragment ions during UVPD (see Figure 3.21). As a result, the experimental UVPD action spectrum of $[Tyr + H]^+$ was generated using all the fragmented mass channels, excluding *m/z*=165 and 136, induced by UVPD (Table 3.5).[97] The neutral fragment loss of $[Tyr + H]^+$ at each fragmentation channel was cited from the fragmentation study of Zhang *et al.*[97] Notably, the measured UVPD spectrum is in excellent agreement with that reported by Pereverzev *et al.*, as shown in Figure 3.19, where two regions, 220−240 nm and 260−300 nm, exhibit appreciable intensity.

**Table 3.5** The fragmentation channels of $[Tyr + H]^+$ observed in UVPD process.[97]

| *m/z* | Neutral fragment |
|---|---|
| 182 | - |
| 165 | $NH_3$ |
| 147 | $NH_3 + H_2O$ |
| 136 | $H_2O + CO$ |
| 123 | $NH_3 + CH_2CO$ |
| 119 | $NH_3 + H_2O + CO$ |
| 107 | $H_2O + CO + CH_2NH$ |
| 95 | $NH_3 + CO_2 + C_2H_2$ |
| 91 | $H_2O + CO + NH_3 + CO$ |

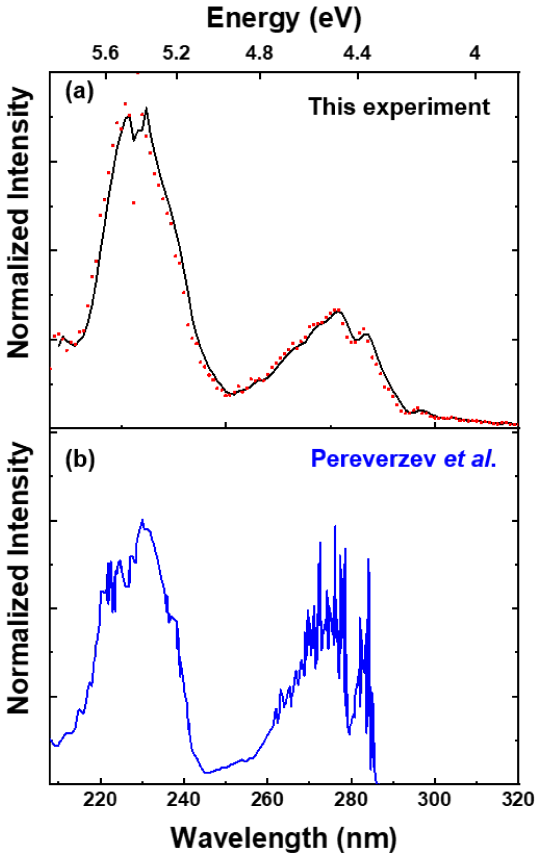


**Figure 3.19** (a) The experimental UVPD action spectrum of $[Tyr + H]^+$ in $N_2$ gas. The solid lines are three-point adjacent averages of the data points. (b) The reported photofragment spectrum of $[Tyr + H]^+$ by Pereverzev *et al.* Adapted from Pereverzev *et al.*[30]

To explore the fragmentation pathways that were involved in the UVPD process, the UVPD action spectra generated for each fragmentation channel of $[Tyr + H]^+$ (except for *m/z*=165 and 136) were studied and are shown in Figure 3.20.

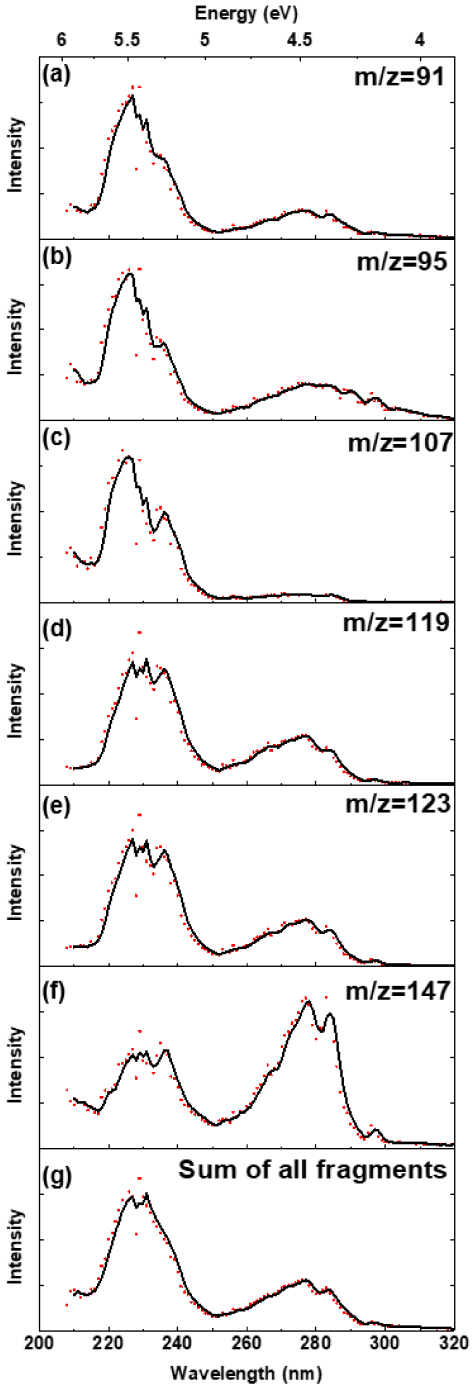


**Figure 3.20** The experimental UVPD action spectra of $[Tyr + H]^+$ generated with fragmentation channel of (a) *m/z*=91, (b) *m/z*=95, (c) *m/z*=107, (d) *m/z*=119, (e) *m/z*=123, (f) *m/z*=147, and (g) all fragmentation channels. The solid lines were three-point adjacent averages of the data points.

The parent spectra monitored in two mass fragments, $[Tyr - NH_3 - H_2O + H]^+$ (*m/z*=147) and $[Tyr - H_2O - CO - CH_2NH + H]^+$ (*m/z*=107), varied significantly from those based on the other fragments. The $[Tyr + H]^+$ spectrum as monitored *via* the *m/z*=147 product channel exhibited a relatively high band intensity for the feature at 4.5 eV. The observation that fragmentation efficiency for producing $[Tyr - NH_3 - H_2O + H]^+$ is greater at 4.5 eV than at 5.5 eV contrasts observations for all other observed product channels. This suggests that loss of $H_2O$ and $NH_3$ exhibits a relatively lower thermodynamic threshold during dissociation of $[Tyr + H]^+$ since more fragment ions (*i.e.*, $[Tyr - NH_3 - H_2O + H]^+$) were induced by absorbing smaller photon energies. The *m/z*=107 fragmentation channel of $[Tyr + H]^+$, corresponding to the loss of $H_2O$, CO, and $CH_2NH$, exhibited a single intense feature around 5.5 eV. It is predicted that the absorption at 4.5 eV, where fragment ions with *m/z*=91, 95, 119, 123, and 147 were generated, might not induce fragmentation of the $[Tyr - H_2O - CO - CH_2NH + H]^+$ ion. Accordingly, thermodynamic threshold calculations regarding to those two fragmentation processes are essential at this point − this will be discussed in the next section.

As mentioned, the presence of the fragment ions induced by CID in Q3 might contribute to unexpected artifacts in the UVPD spectrum of the parent ion. Hence, the UVPD spectrum for the *m/z*=165, $[Tyr - NH_3 + H]^+$, as produced from $[Tyr + H]^+$ by CID (see Figure 3.18) was recorded and compared with the spectrum of the parent $[Tyr + H]^+$ ion. This comparison is shown in Figure 3.21. The UVPD spectra of $[Tyr + H]^+$ and $[Tyr - NH_3 + H]^+$ exhibit significantly different spectral profiles. In particular, $[Tyr - NH_3 + H]^+$ displays distinct spectral features at 4.9 eV and 4.3 eV. However, both $[Tyr + H]^+$ and $[Tyr - NH_3 + H]^+$ absorbed light in the energy region of 4.3 eV to 4.7 eV, suggesting that the 165 fragmentation channel produced *via* CID could contribute to production of the lighter mass species (*i.e.*, *m/z*=147) during UVPD of $[Tyr + H]^+$.

This rationalizes the distinct $[Tyr + H]^+$ spectrum generated with *m/z*=147 fragment (Figure 3.20f) in which a more intense intensity was observed at 4.5 eV, and it also suggests that $[Tyr - NH_3 - H_2O + H]^+$ (*m/z*=147) can be produced sequentially from the $[Tyr - NH_3 + H]^+$ (*m/z*=165). Similarly, we hypothesize that $[Tyr - H_2O - CO + H]^+$ (*m/z*=136) exhibit an absorption at ~5.5 eV but not ~4.5 eV, and production of $[Tyr - H_2O - CO - CH_2NH + H]^+$ (*m/z*=107) is a stepwise process during UVPD. As a result, $[Tyr - H_2O - CO + H]^+$ can undergo additional photofragmentation at ~5.5 eV to yield $[Tyr - H_2O - CO - CH_2NH + H]^+$, rationalizing a single band observed at ~5.5 eV in the UVPD spectrum of $[Tyr + H]^+$ generated with *m/z*=107 channel. The fragmentation pathways of protonated Tyr postulated by Zhang *et al.* further supports that $[Tyr - NH_3 - H_2O + H]^+$ is produced from $[Tyr - NH_3 + H]^+$, and the formation of $[Tyr - H_2O - CO - CH_2NH + H]^+$ is based on $[Tyr - H_2O - CO + H]^+$.[97]

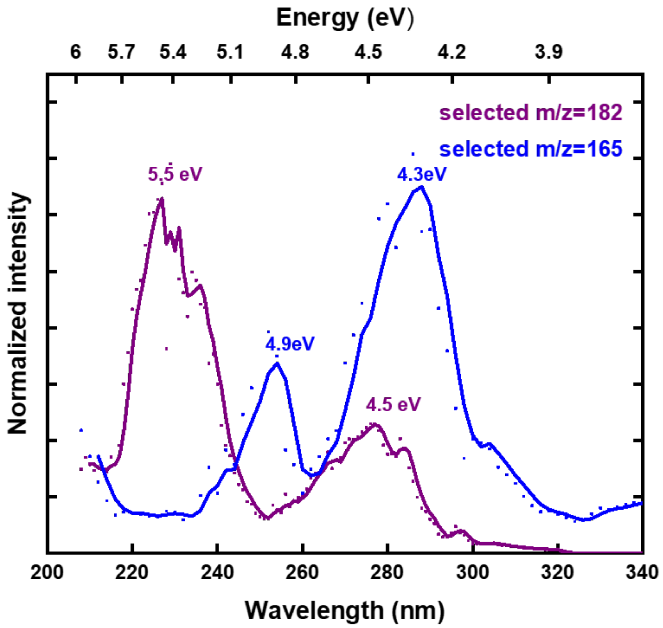


**Figure 3.21** The UVPD action spectra of $[Tyr + H]^+$ (purple) and $[Tyr - NH3 + H]^+$ (blue). The solid lines were three-point adjacent averages of the data points.

### 3.3.2.2 Computational results for $[Tyr + H]^+$

The lowest energy prototropic isomers of $[Tyr + H]^+$ are shown in Figure 3.22.

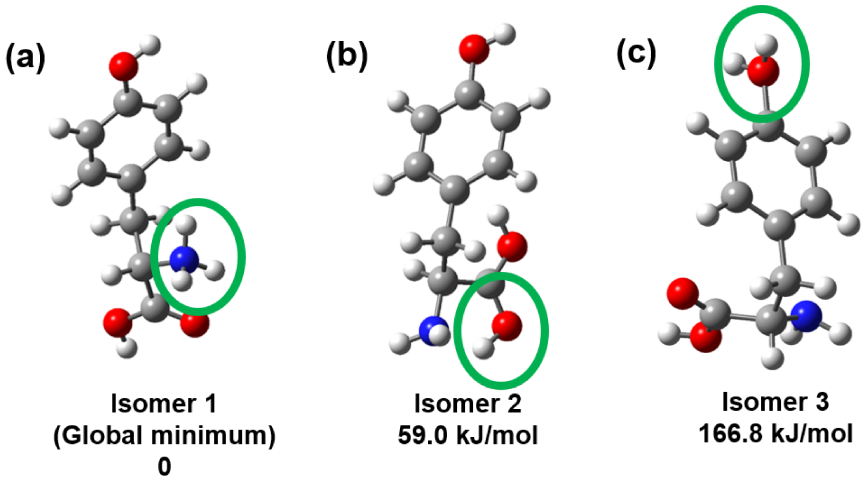


**Figure 3.22** The optimized structures of $[Tyr + H]^+$ with protonation sites on (a) N-amine, (b) O-carboxyl, and (c) O-hydroxyl. The optimizations and relative Gibbs free energy to that of the global minimum structure were determined at the CAM-B3LYP/6-311++G(d,p) level of theory.

The calculated Gibbs energies of isomers 2 and 3 were 59.0 and 166.8 kJ/mol, respectively, above the calculated global minimum. Structures of the $[Tyr + H]^+$ protomers in the water solution were estimated by optimizing clusters consisting of $[Tyr + H]^+$ and water molecules manually located at the hydrogen bonding sites. The structures and Gibbs energies of three cluster systems containing three different $[Tyr + H]^+$ protomers calculated at the same level can be found in Figure A**.5** in Appendix A. As was calculated for the gas phase bare ion, the lowest energy structure of $[Tyr + H]^+$ solvated by water was the N-amine protomer. The relative Gibbs energies of the O-carboxyl and O-hydroxyl prototropic clusters exceeded +90 kJ/mol. Assuming a thermodynamic equilibrium described by Boltzmann statistics, the population of isomers 2 and 3 in the gas phase ensemble is expected to be negligible. Consequently, one expects that the N-amine protomer dominates the gas phase ensemble.

Vibronic spectra were computed for isomers 1, 2 and 3 at the CAM-B3LYP/6-311++G(d,p) level of theory. The spectral simulations employed the VG approximation and results were red-shifted by 30 nm for better agreement with the experimentally acquired UVPD action spectrum of $[Tyr + H]^+$, as shown in Figure 3.23.

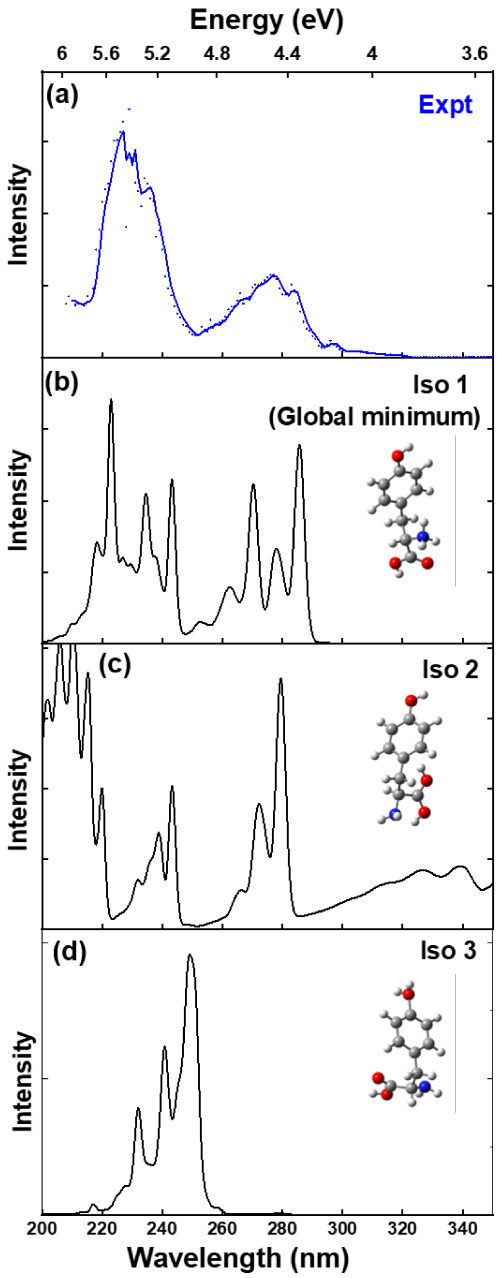


**Figure 3.23** (a) The experimental spectrum of $[Tyr + H]^+$. The calculated vibronic spectra of (b) isomer 1, (c) isomer 2, and (d) isomer 3 of $[Tyr + H]^+$ using the VG|FC approximation.

As mentioned previously in Chapter 2, the relatively shift of the computed spectrum compared to that measured experimentally is likely to be due to the limited accuracy of the DFT method. The calculated spectra of all three protomers of $[Tyr + H]^+$ exhibit absorption features in the 200−350 nm region. The calculated spectrum of isomer 3 exhibits only a single feature in the measures spectral region and does not well-reproduce the observed spectrum. The computed

spectra of isomers 1 and 2 both exhibit features that match those observed experimentally, but the calculated spectrum of isomer 2 shows an additional feature in the 300−340 nm region that was not seen experimentally. Given these results and the large relative Gibbs energies of isomers 2 and 3, we favour isomer 1 as the sole spectral carrier.

The calculated vibronic spectra for each electronic transition (*i.e.*, $S_1 \leftarrow S_0$, $S_2 \leftarrow S_0$, $S_3 \leftarrow S_0$…) of isomer 1 are plotted along with the experimental spectrum in Figure 3.24.

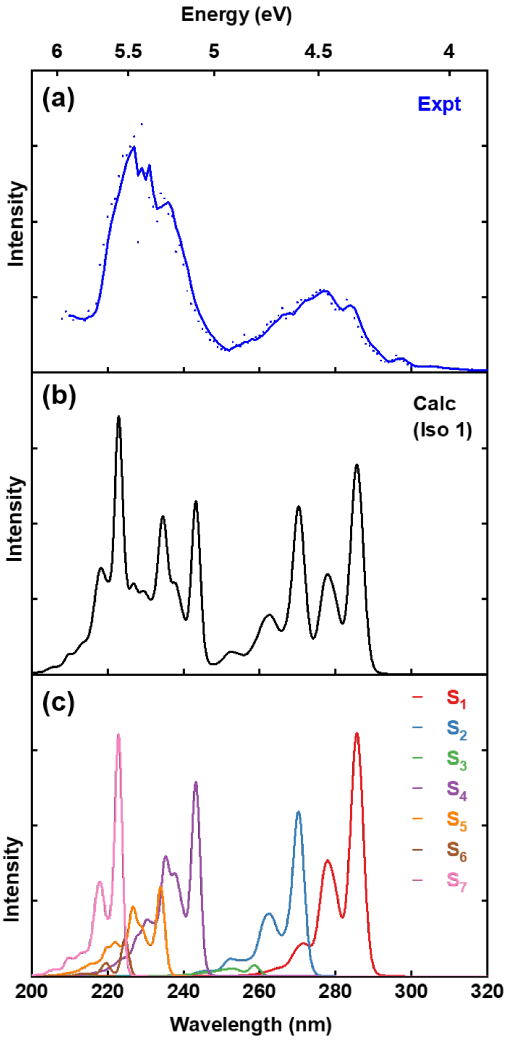


**Figure 3.24** (a) The experimental UVPD spectrum of $[Tyr + H]^+$. The calculated vibronic spectra of $[Tyr + H]^+$ of (b) a sum of $S_1$ to $S_7$ and (c) each singlet transition state from $S_1$ to $S_7$ using the VG|FC approximation.

Based on the simulations, we assign the feature at 4.5 eV to the $S_1 \leftarrow S_0$ and $S_2 \leftarrow S_0$ transitions, and the feature at 5.5 eV to the $S_4 \leftarrow S_0$ and $S_7 \leftarrow S_0$ transitions, respectively. The HOTOs and LUTOs associated with some featured electronic transition of $[Tyr + H]^+$ were computed and are shown in Figure 3.25.

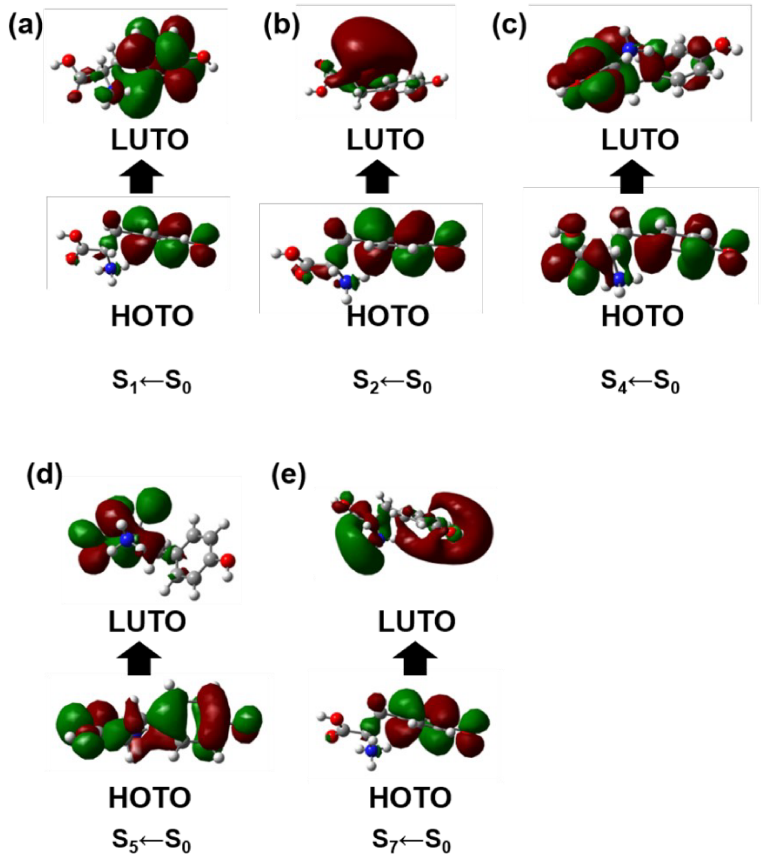


**Figure 3.25** The calculated LUTOs and HOTOs of (a) $S_1 \leftarrow S_0$, (b) $S_2 \leftarrow S_0$, (c) $S_4 \leftarrow S_0$, (d) $S_5 \leftarrow S_0$, and (e) $S_7 \leftarrow S_0$ transitions.

Calculations of the HOTOs and LUTOs indicate the $S_1 \leftarrow S_0$ transition to be of $\pi^* \leftarrow \pi$ characters with excitation localized on the aromatic ring system (Figure 3.25a). The $S_7 \leftarrow S_0$ transition, on the other hand, was found to be of $\sigma^* \leftarrow \pi$ characters (Figure 3.25e). The HOTOs and LUTOs of the $S_4 \leftarrow S_0$ and $S_5 \leftarrow S_0$ transitions illustrated that the electron density started moving from the phenyl ring to the $-NH_2$ and –COOH side as the photon energy increased. Meanwhile, the LUTO of $S_7 \leftarrow S_0$ transition showed anti-bonding characters on the hydroxy oxygen and the amino nitrogen

of $[Tyr + H]^+$. The electron density transfer might weaken these bonds and subsequently induce losses of $H_2O$, CO, and $CH_2NH$, leading to the production of $[C_7H_7O]^+$ (*m/z*=107).

To further investigate the dissociative mechanism(s) involved in the fragmentation of $[Tyr + H]^+$, energy thresholds were computed by determining the Gibbs energy difference between geometry optimized parent and fragment ions at the CAM-B3LYP/6-311++G(d,p) level of theory. Specifically, two fragmentation processes, production of $[Tyr - NH_3 - H_2O + H]^+$ (*m/z* 147) and $[Tyr - H_2O - CO - CH_2NH + H]^+$ (*m/z*=107), were investigated. The thresholds for generating $[Tyr - NH_3 - H_2O + H]^+$ and $[Tyr - H_2O - CO - CH_2NH + H]^+$ were calculated to be 1.01 eV and 1.93 eV, respectively, which are much lower than the photon energies utilized in the UVPD experiments (3.5−6.2 eV). This suggests that production of 147 and 107 fragments was achievable during UVPD of $[Tyr + H]^+$, and these two UVPD processes for $[Tyr + H]^+$ are single-photon processes, which provides chances for the photo-excited molecular cations undergoing internal conversion and dissociating on the ground state potential energy surface. This also proves that the intense band observed in the UVPD spectrum of $[Tyr + H]^+$ that was monitored in the *m/z*=147 fragmentation channel (Figure 3.20f) is most likely due to a stepwise fragmentation of $[Tyr - NH_3 + H]^+$ during UVPD.

**3.4 Conclusions**

$[Trp + H]^+$ and $[Tyr + H]^+$ were investigated in the gas phase with DMS, MS, and UVPD spectroscopy. Two dispersion curves were monitored for $[Trp + H]^+$ and exhibited hard-sphere (Type C) behaviour when electrosprayed from $MeOH/H_2O$ solvent under $N_2$. Ion populations associated with the two curves were shown to be the protonated Trp monomer and dimer species. Only one conformation of $[Tyr + H]^+$ was resolved in the DMS cell under a pure $N_2$ gas environment. Unlike the species observed for $[Trp + H]^+$, $[Tyr + H]^+$ exhibited weak-clustering,

or Type B, behaviour. This clustering behaviour indicated that the Tyr protomer experienced dynamic clustering interactions with the ESI solvent. Upon collision-induced dissociation in Q3, Trp monomer resulted in a prominent loss of $NH_3$, signified by a mass channel at *m/z*=188; fragment channels at *m/z*=165 (loss of $NH_3$) and *m/z*=136 (loss of $H_2O$ and CO) were produced from the $[Tyr + H]^+$ parent ion. These fragment ions produced *via* CID can contribute additional electronic transitions to the parent ion spectrum. The distinct shape of the UVPD spectrum of $[Tyr + H]^+$ monitored in the *m/z*=147 fragmentation channel suggests that the abundance of $[Tyr - NH_3 - H_2O + H]^+$ at 4.5 eV was due to UVPD of $[Tyr - NH_3 + H]^+$, which reveals that $[Tyr - NH_3 - H_2O + H]^+$ was produced from $[Tyr - NH_3 + H]^+$ sequentially during the UVPD process of $[Tyr + H]^+$. The UVPD spectra generated from both $[Trp + H]^+$ ion populations were identical, indicating that the protonated state fragmented from $[(Trp)_2 + H]^+$ was the same as that of the $[Trp + H]^+$ monomer observed in the DMS.

The calculated Gibbs energies of protonated Trp and Tyr reveal that their gas-phase ensembles contain a single protomer. In both cases, protonation takes place on the N-amine group. Computed electronic orbitals indicate that both $[Trp + H]^+$ and $[Tyr + H]^+$ showed strong electronic transitions involving $\pi^* \leftarrow \pi$ characters on the aromatic rings at 4.5 eV. The electronic transitions involving $\sigma^* \leftarrow \pi$ characters were observed for $[Trp + H]^+$ at 4.8 eV, whereas $[Tyr + H]^+$ showed a major electronic transition at 5.5 eV containing $\sigma^* \leftarrow \pi$ characters. The electron density change observed between HOTOs and LUTOs during electronic excitation rationalizes the observed $[Tyr - H_2O - CO - CH_2NH + H]^+$ product ion. Moreover, the calculated dissociation thresholds for the fragmentation pathways of $[Tyr + H]^+$ imply that UVPD of $[Tyr + H]^+$ was a single-photon process whereby the photo-excited molecular cations undergo internal conversion and dissociate on the ground state potential energy surface.

# Chapter 4: Investigating arginine clusters under various DMS conditions

## 4.0 Overview

Protonated arginine clusters $[(Arg)_n + H]^+$ (n=1−4) were investigated by exploring the dynamic ion-solvent interactions between the clusters and various solvent-modified nitrogen gas environments in the DMS cell. Computationally, the structures of target ions were optimized at the DFT level of theory to visualize the intramolecular interactions (*i.e.*, hydrogen bonding and ionic bonding) involved in the structures of $[(Arg)_n + H]^+$ (n=1−4). The electrostatic potential was mapped onto the total electron density distributions for the species of interest to study their intramolecular interactions *via* regions of high or low partial charge distributions. Thermodynamic properties were determined to investigate the relative stability of each species and to predict the most likely protonation site and binding motif for target ions in the gas phase.

## 4.1 Introduction

Arg is one of the twenty genetically encoded amino acids involved in protein biosynthesis and works as the precursor for the biosynthesis of nitric oxide, and it has been shown that salt-bridge interactions, which involve ionic and hydrogen bonding, are present in systems that contain protonated Arg.[1,10,26,48,50,63,98,99] Studying the salt-bridge interactions present in Arg-containing clusters may help in understanding protein-protein interactions that involve Arg moieties.[7,8,100,101] The presence of guanidino and carboxyl groups in Arg can induce intermolecular and intramolecular ionic bonding established in clusters containing Arg, producing salt-bridges.[48,50,51,54] Figure 4.1b gives an example of a salt-bridge structure between a protonated

guanidino group and a deprotonated carboxyl functional group, where both ionic bonding and hydrogen bonding occur.

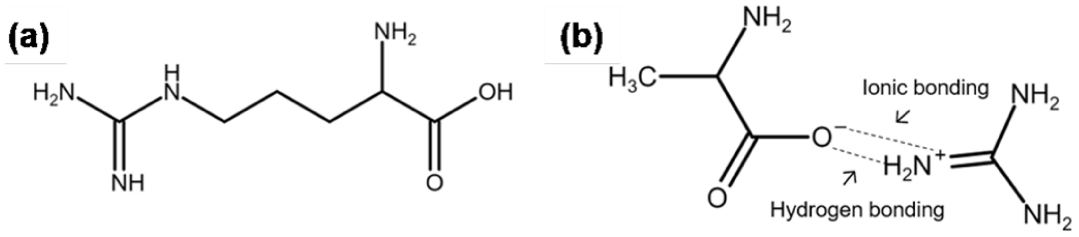


**Figure 4.1** (a) Arg structure; (b) an example of a salt-bridge between a protonated guanidino and a deprotonated carboxyl functional group.

Research concerned with the structures and properties of protonated Arg and its clusters has been conducted by several groups.[48,50,54] Computational work performed by Strittmatter and Williams suggested that the structures of protonated Arg dimer and Bradykinin (a peptide that contains Arg) were stabilized by salt-bridge interactions when compared to the conformation without ionic bonding interactions.[54] Price *et al.* investigated the dissociation kinetics of proton-bound dimers of Arg, glycine, and alanine in the gas phase and concluded that a protonated Arg dimer structure contained a salt-bridge, where ionic and hydrogen bonding existed between a protonated guanidino group and a deprotonated carboxyl group.[48] The authors found that additional salt-bridges could be formed when additional relevant functional groups were involved. Prell *et al.* studied the geometric structures (viz. vibrational spectroscopy) of protonated dipeptides of glycine, valine, proline, lysine, histidine, and Arg in the gas phase using infrared multiphoton dissociation (IRMPD) spectroscopy.[50] With the exception of the arginine dipeptide, a single formal charge site was detected for the protonated state of each dipeptide. The Arg dimer instead exhibited ionic bonding interactions and formed a salt-bridge structure.[50] The clusters of protonated Arg (*i.e.*, dimer, trimer, tetramer…), which can provide insight into the relationship between the number of existing salt-bridges and overall cluster structural stability, still remain unexplored.

Here, the structures of the protonated Arg clusters, $[(Arg)_n + H]^+$ (n=1−4), were investigated using DMS-MS and complementary quantum chemical calculations. Various gas-phase modifiers were introduced to explore ion-solvent interactions with the numerous clusters. Thermodynamic properties and electrostatic potential maps were calculated for the various protomeric forms of each cluster. The structural stability as related to the charge distribution of each Arg cluster is also explored.

## 4.2 Methods

### 4.2.1 Experimental methods

Arginine (98% purity) was purchased from Acros Organics. Arg standards were diluted to a concentration of 100 ng/mL using 50:50 MeOH:$H_2O$ (MeOH/$H_2O$) solution with 0.5% formic acid. The experimental setup (*i.e.*, DMS-MS) applied in this project is shown in Figure 3.3, as highlighted in the dashed frame. Electrosprayed ions were directed toward the inlet of a SelexION DMS from a SCIEX Turbo V electrospray ion source. Nascent ions were carried through the DMS and towards a QTRAP 5500 mass spectrometer by dry $N_2$. The carrier gas environment was modified by the addition of volatile solvents to the curtain gas line in experiments that employed a chemical modifier. The target ion was selected by its *m/z* in Q1, and the mass spectrum was recorded *via* mass axial ejection from Q3. The SV was scanned from 0−3000 V in 500 V increments and 3000−4000 V in 200 V increments. At each SV step, the CV was scanned from −80 to 25 V in 0.1 V increments. Mass-selected ion intensities were recorded as a function of SV and CV, yielding dispersion plots. A potential difference applied across the gap between the exit of the DMS cell and the Q-jet region, known as the de-clustering potential (DP), was ramped from 0 to 300 in 50 V increments to monitor the change in intensities of clusters and bare ions in Q1 as DP varied.

### 4.2.2 Computational methods

Molecular models of the targeted species were generated manually using Gaussview.[93] The geometries of the target ions were then optimized at the B3LYP/6-31G level of theory using DFT, and atomic partial charges were calculated using the CHelpG method as implemented in Gaussian 16.[93] The optimized geometries and calculated partial charges were used to generate input files for a basin hopping (BH) routine, as discussed in Chapter 2. The BH routines involved 10,000 steps, which result in 10,000 unique geometric structures being sampled for each molecule. With regard to the monomer, each internal dihedral angle ($\alpha$) was rotated by a random value between $-5^{\circ} \leq \alpha \leq 5^{\circ}$. For clusters with n > 1, additional parameters were introduced, in which each moiety was rotated around its center of mass by a random quantity $\beta$ chosen from $-4^{\circ} \leq \beta \leq 4^{\circ}$ and translated by a random step size $\eta$ from –0.5 Å $\leq \eta \leq$ 0.5 Å. The unique geometries obtained from BH calculations were first refined at the PM6 level of theory and re-optimized using DFT at the B3LYP/6-311++G(d,p) level of theory.[44,94] Normal mode analyses were performed to obtain thermodynamic corrections. The electrostatic potential was mapped onto the total electronic density for each target molecule at the B3LYP/6-311++G(d,p) level of theory.

## 4.3 Results and discussion

### 4.3.1 Experimental results in $N_2$

The DMS clustering behaviour of protonated Arg, $[Arg + H]^+$, was studied within a pure $N_2$ environment. The dispersion plots in $N_2$ recorded by monitoring the $[Arg + H]^+$ mass channel were recorded for different DP values are shown in Figure 4.2. A single Type C dispersion curve (corresponding on a single ionogram peak) was observed for $[Arg + H]^+$ when DP was set to 0 V. When DP was increased beyond 100 V, multiple Type C dispersion curves were observed. As discussed in Chapter 3, the appearance of multiple curves in the dispersion plots could result from

the resolution of different $[Arg + H]^+$ protomers or from larger cluster species fragmenting to produce $[Arg + H]^+$ (*i.e.*, *m/z*=175). To explore these possibilities, a DP scan was performed. The ionograms of $[Arg + H]^+$ recorded at SV=4000 V as DP was ramped from 0 to 300 V in 50 V increments and are shown in Figure 4.3.

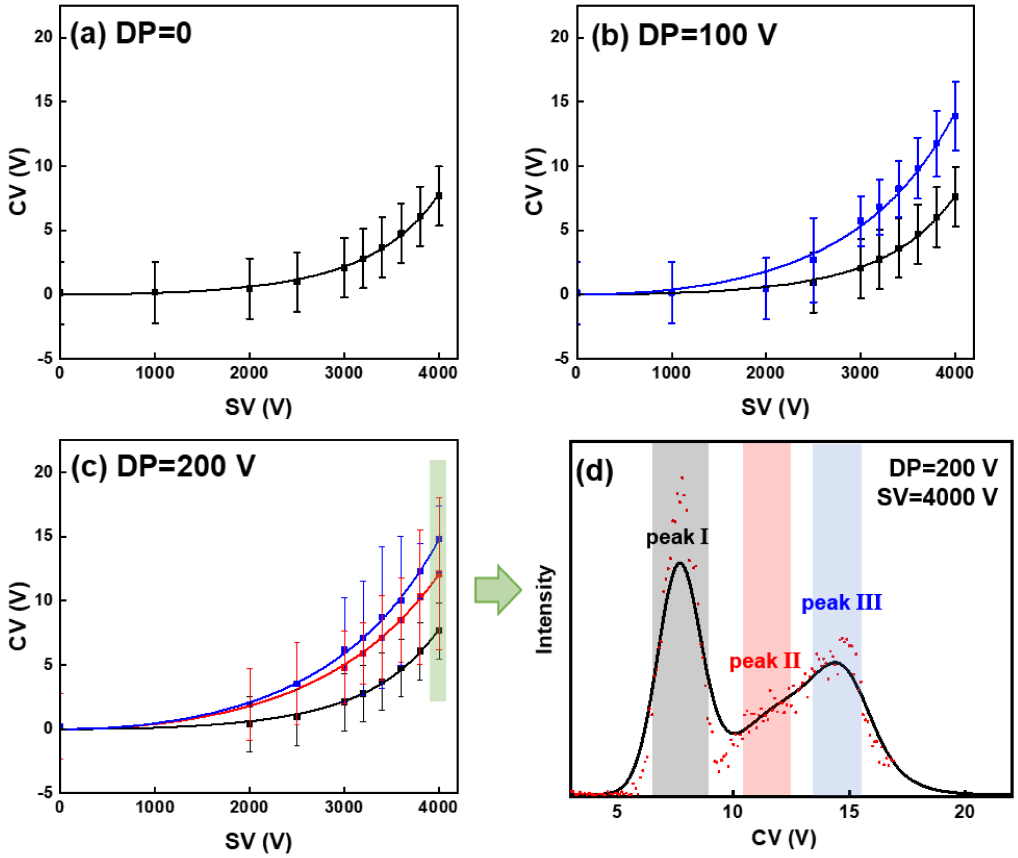


**Figure 4.2** The dispersion plots of $[Arg + H]^+$ electrosprayed from MeOH/$H_2O$ solution in pure $N_2$ at (a) DP=0 V, (b) DP=100 V, and (c) DP=200 V. (e) The ionogram of $[Arg + H]^+$ at DP=200 V and SV=4000 V.

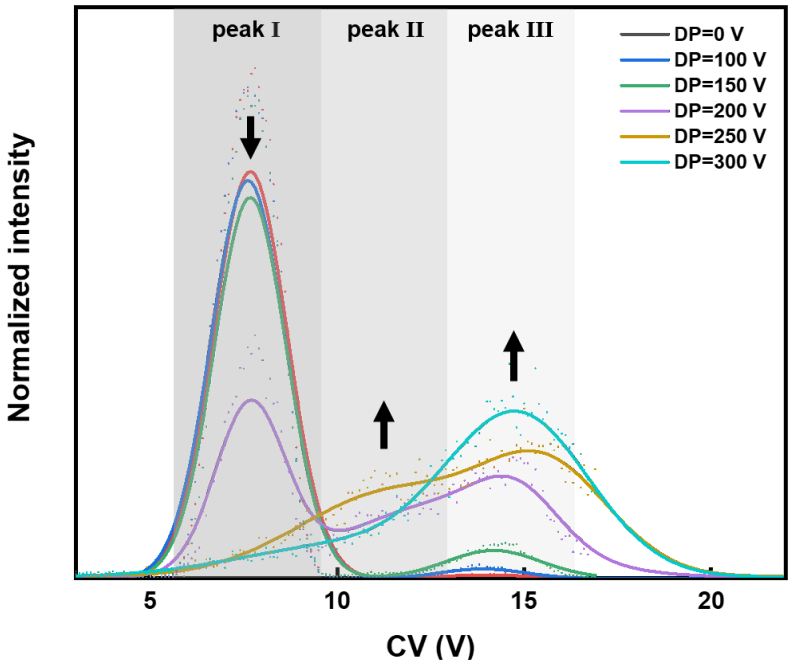


**Figure 4.3** The ionograms of $[Arg + H]^+$ at SV=4000 V in pure $N_2$ as DP was increased at DP=0, 100, 150, 200, 250, and 300 V.

Under low DP settings, peak I was the dominant feature in the observed ionogram. As the DP was increased, ion populations associated with peaks II and III increased in intensity. The simultaneous increase in the fraction of peaks II and III with rising DP suggests that the species associated with peaks II and III are fragmenting to yield $[Arg + H]^+$ ($m/z$=175). This suggests that peak I is associated with the bare $[Arg + H]^+$ ion, and peaks II and III were produced from larger Arg-containing clusters. To further investigate the species giving rise to ion populations I, II and III, precursor ion scans were conducted while gating the SV at 4000 V and selecting the ion populations at CV=7.6 V, 12.1 V, and 14.8 V, respectively. The precursor ion scans collected for DP=300 V are shown in Figure 4.5. As expected, $m/z$=175, $[Arg + H]^+$, is observed when gating on all three ionogram features. Mass peaks with $m/z < 175$ are attributed to product ions produced from fragmentation of $[Arg + H]^+$ or background ions, as tabulated in Table 4.1. Interestingly, mass peaks at $m/z$ 349, 523 and 697, were observed for all three ionogram peaks at high DP (300 V); these ions are assigned to $[(Arg)_n + H]^+$ (n=2−4), respectively. The relative abundance of

the $[(Arg)_n + H]^+$ (n=1−4) was monitored as a function of DP when setting CV=7.6 V, 12.1 V, and 14.8 V, respectively, with SV=4000 V. The evolution of ion intensity as a function of DP is shown in Figure 4.6.

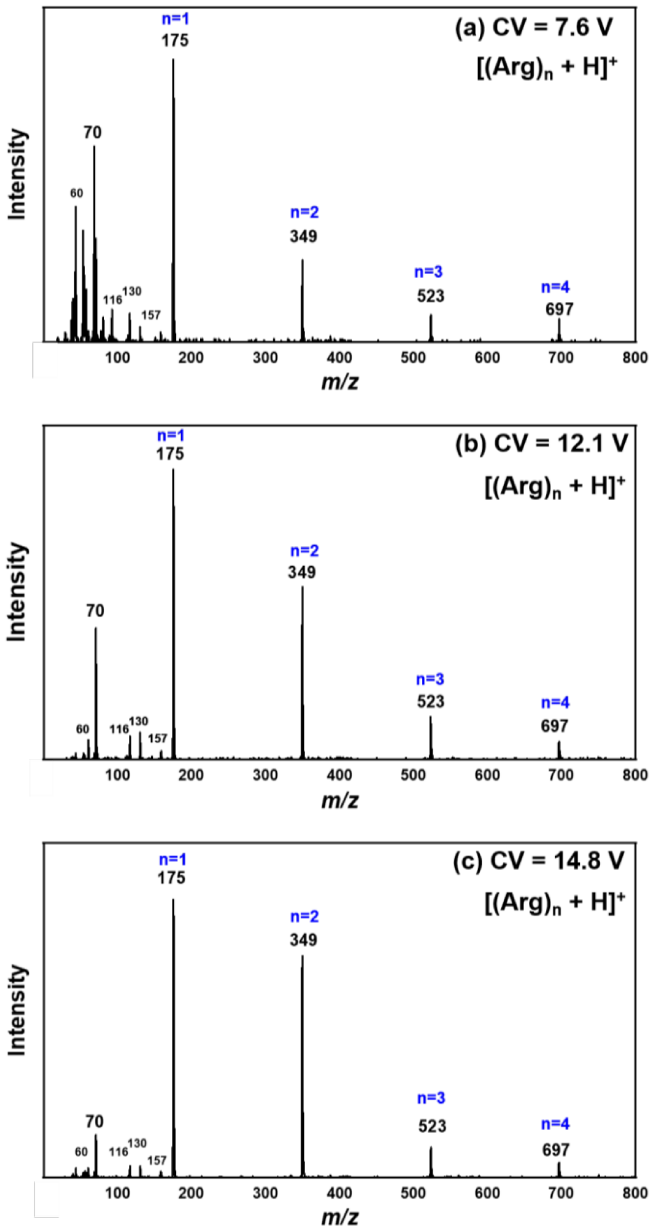


**Figure 4.4** Precursor ion scans for $[Arg + H]^+$ at (a) CV=7.6 V, (b) CV=12.1 V, (c) CV=14.8 V and with SV=4000 V and DP=300V. Peaks of $[(Arg)_n + H]^+$ are labelled.

**Table 4.1** The fragmentation channels of $[Arg + H]^+$ and corresponding neutral fragment losses.

| *m/z* | Neutral fragment(s) |
|---|---|
| 175 | - |
| 158 | $NH_3$ |
| 157 | $H_2O$ |
| 130 | $NH_3 + CO$ |
| 116 | $CH_5N_3$ |
| 112 | $NH_3 + CO + H_2O$ |
| 70 | $CH_5N_3 + H2O + CO$ |
| 60 | $NH_3 + CO + C_2H_6O$ |

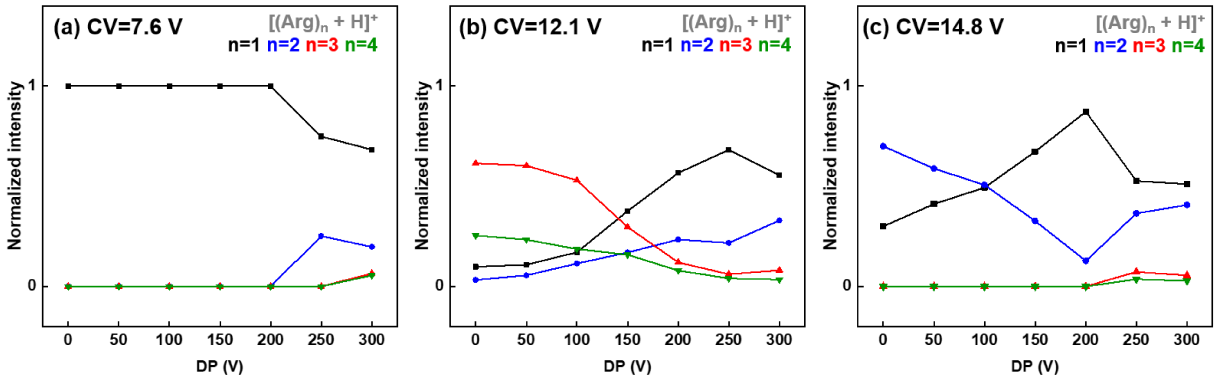


**Figure 4.5** The normalized intensity of *m/z*=175, 349, 523, and 697 in Q1 in $N_2$ at (a) CV=7.6 V, (b) 12.1 V, and (c) 14.8 V when SV=4000 V as DP was ramped from 0 to 300 V.

As expected, higher DP settings induced greater degrees of in-source CID owing to higher field-induced temperatures.[102] Figure 4.5a shows that the ion population for the CV=7.6 V feature is predominantly associated with the monomeric species $[Arg + H]^+$ at low DP. However, when DP > 200 V, in-source CID causes fragmentation of the larger clusters and their appearances in the *m/z* 175 mass channel. The intensity changes observed in Figure 4.5b show that the CV=12.1 V feature is predominantly associated with $[(Arg)_3 + H]^+$ and, to a lesser extent, $[(Arg)_4 + H]^+$. At higher DP, these species fragment to produce $[Arg + H]^+$ (major product) and $[(Arg)_2 + H]^+$ (minor product). Similarly, $[(Arg)_2 + H]^+$ is resolved at CV=14.8 V and, with increasing DP, it fragments to produce $[Arg + H]^+$ (see Figure 4.5c). The fact that the signal intensity increases at higher DP

for $[(Arg)_n + H]^+$ (n=2−4) at the CV=14.8 V setting indicates that a small amount of $[(Arg)_n + H]^+$ (n > 4) species are also present in the cluster distribution (and ultimately dissociate).

### 4.3.2 Experimental results in solvent-modified environments

To study ion-solvent interactions of the $[(Arg)_n + H]^+$ (n=1−4) species, their DMS behaviours were measured under different solvent-modified gas-phase environments. Table 4.2 gives the compositions of the various environments employed. Experimental data are shown for $H_2O$ modifier in Figures 4.6 and 4.7, for MeOH in Figures 4.8 and 4.9, for IPA in Figures 4.10 and 4.11, for ACN in Figures 4.12 and 4.13, and for ACE in Figures 4.14 and 4.15. The ion signal dropped below observable levels for $[Arg + H]^+$ under ACE modifier conditions at SV > 3000 V. It is noted that besides cluster peaks associated with $[(Arg)_n + H]^+$ (n > 1), the $[Arg + solvent + H]^+$ clusters also present under IPA, ACN, and ACE conditions, where strong ion-solvent interaction is exhibited. The mass channels that correspond to Arg-IPA ion-solvent clusters in Q1 all appeared at 1 mass unit higher than the *m/z* of $[(Arg)_n + IPA + H]^+$ (n= 1−4). This unexpected result is still under investigation. Accordingly, the precursor ion scan of ion-solvent clusters was only recorded under ACN- and ACE-modified environments, as shown in Figure 4.13 and Figure **4.15**.

Two peaks were observed under $H_2O$, MeOH, ACN and ACE conditions, while three peaks were observed under IPA-modified conditions. In the $H_2O$ modified environment (Figure 4.6), a major Type B sub-population and a minor Type C sub-population were observed. DP scans (*e.g.*, Figure 4.7) indicated that the major sub-population in $H_2O$, MeOH, and IPA is associated with the bare $[Arg + H]^+$ cation, whereas a mixture of $[Arg + ACN + H]^+$ and $[Arg + H]^+$ was detected in the major sub-population in ACN. Figure 4.7b shows that the minor peak observed in the $H_2O$-modified environment is predominantly associated with the dimer and, to a lesser extent, the trimer. This, too, is a trend across all modified environments except for ACE; the minor

features are associated with larger clusters. For all modified environments except IPA, a single minor feature associated with the dimer was observed. In the IPA-modified environment (Figure **4.10**b), two minor peaks were observed − one at CV= −2.0 V associated predominantly with the trimer and one at CV= −16.7 V associated predominantly with the dimer. The observations under ACE are varied from other gas-phase conditions. The major sub-population observed in ACE was the ion-solvent cluster, $[Arg + ACE + H]^+$, with a small amount of $[(Arg)_2 + H]^+$, whereas the bare protonated Arg ion was shown in the minor sub-population. This reveals that $[Arg + H]^+$ binds so strongly with acetone gas molecules that the $[Arg + ACE + H]^+$ cluster is the dominant species in ACE. The abundance of $[Arg + ACE + H]^+$ under ACE can explain the disappearance of the $[Arg + H]^+$ signal when SV reached 3000 V. As more energy was supplied at higher SVs, the acetone molecule tended to evaporate from the ion-solvent cluster, which could result in the fragmentation of $[Arg + ACE + H]^+$ in the DMS cell, thus, leading to failure in the detection of $[Arg + H]^+$ signal in Q1.

In examining the clustering behaviour of the ions, two trends are observed. Firstly, it is observed that the monomer species exhibits ion-solvent interactions that increase with the trend $H_2O$ < MeOH < IPA < ACN < ACE. This trend roughly correlates with the magnitudes of the solvent molecule dipole moments, except for ACN and ACE.[41] Secondly, it is observed that the larger clusters (*i.e.*, n > 2) exhibit weaker clustering interactions than does the monomer. This can be rationalized by charge shielding effects. Since the charge-carrying proton is bound between the Arg moieties in the larger clusters, it is shielded from the solvent molecules, thereby reducing the charge-dipole contribution to the ion-solvent interaction.[41]

**Table 4.2** Gas modifier conditions (mole ratio) used in the DMS study of Arg.

| ***Gas environments*** |
|---|
| *$N_2$* |
| *$N_2$ +1.5% $H_2O$ vapour* |
| *$N_2$ +1.5% methanol (MeOH) vapour* |
| *$N_2$ +1.5% 2-propanol (IPA) vapour* |
| *$N_2$ +1.5% acetonitrile (ACN) vapour* |
| *$N_2$ +1.5% acetone (ACE) vapour* |

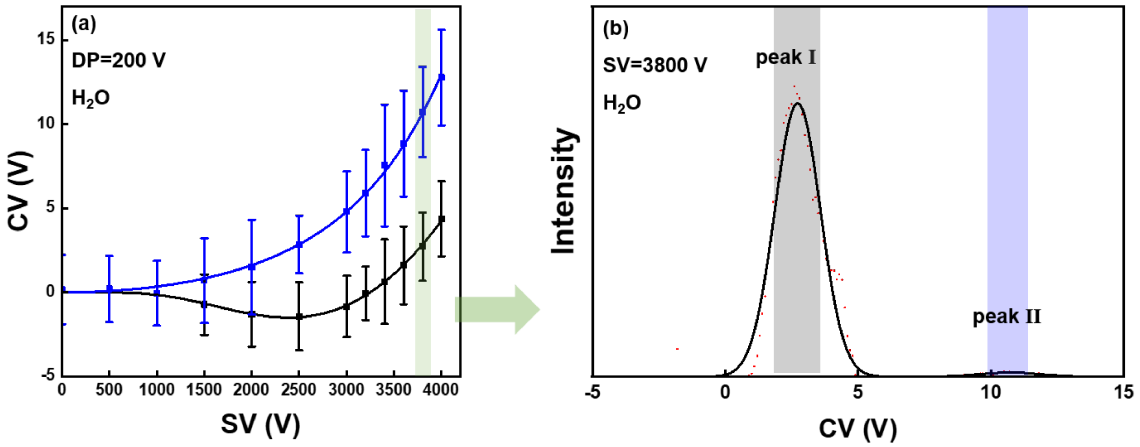


**Figure 4.6** (a) The dispersion plots at 200 V and (b) the ionogram at SV=3800 V and DP=200 V in $N_2$ seeded with 1.5% $H_2O$ vapour. The black and blue blocks labelled in the ionogram indicate peaks I and II, respectively.

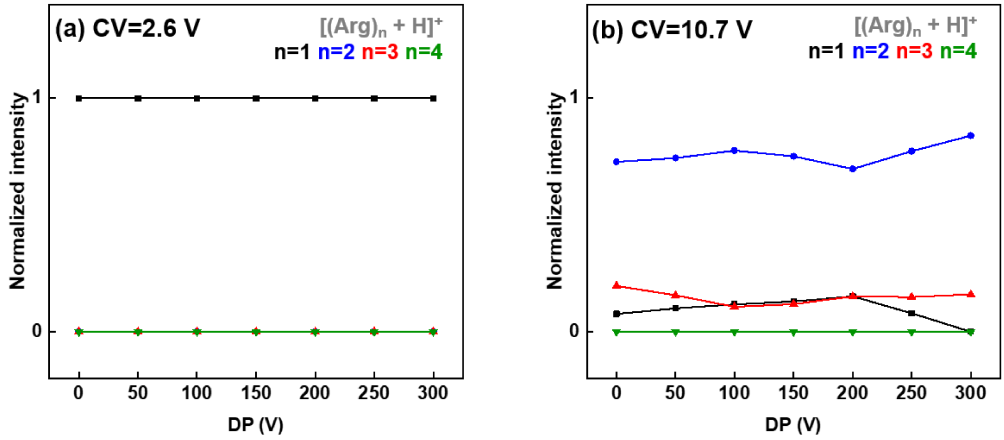


**Figure 4.7** The intensity of *m/z*=175, 349, 523, and 697 in Q1 at (a) CV=2.6 V and (b) CV=10.7V (SV=3800 V) as DP was ramped from 0 to 300 V in $N_2$ seeded with 1.5% $H_2O$ vapour.

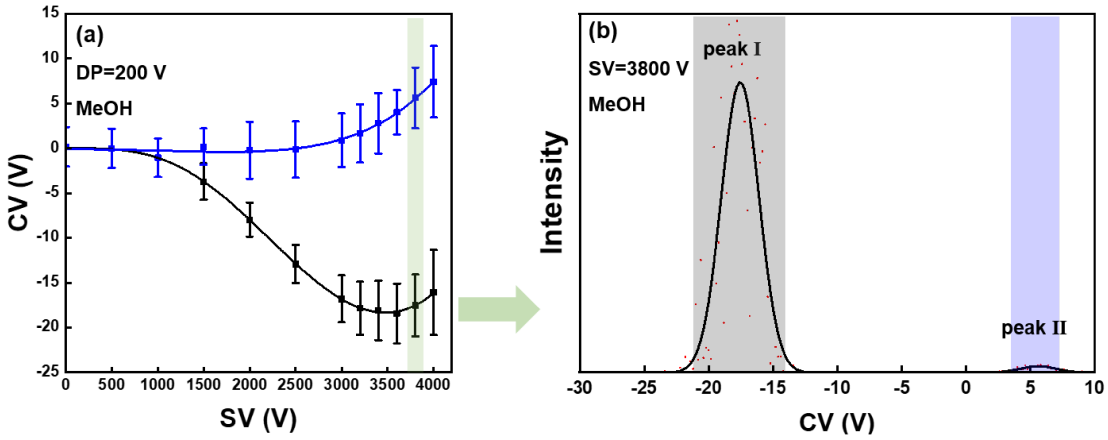


**Figure 4.8** (a) The dispersion plots at DP=200 V and (b) the ionogram at SV=3800V and DP=200V in $N_2$ seeded with 1.5% MeOH vapour. The black and blue blocks labelled in the ionogram indicate peaks I and II, respectively.

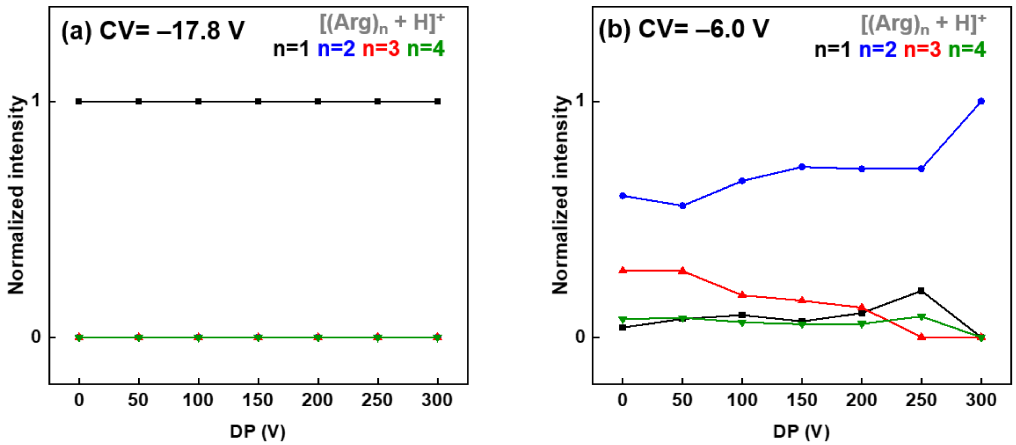


**Figure 4.9** The normalized intensity of *m/z*=175, 349, 523, and 697 in Q1 at (a) CV= 17.8 V and (b) CV=6.0 V (SV=3800 V) as DP was ramped from 0 to 300 V in $N_2$ seeded with 1.5% MeOH vapour.

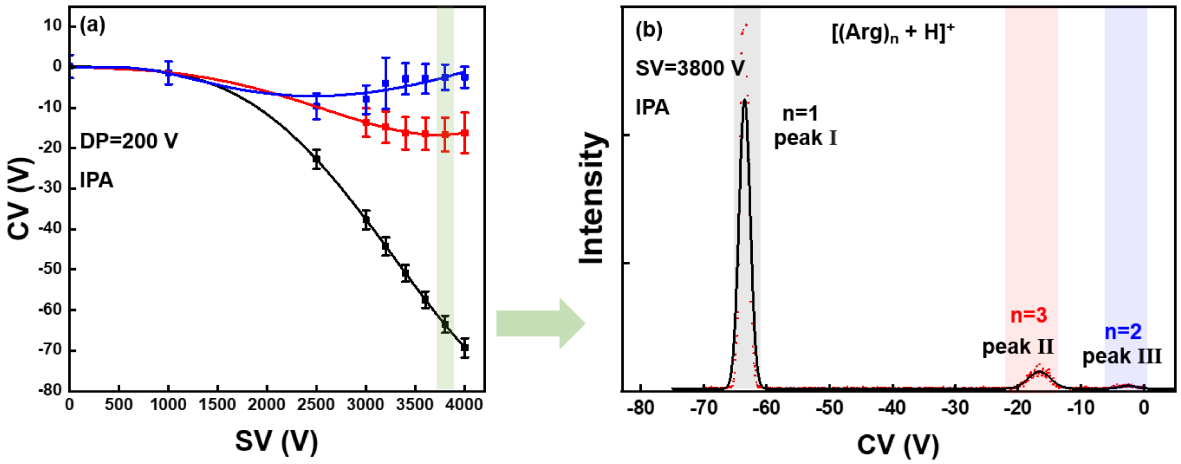


**Figure 4.10** (a) The dispersion plots at DP=200 V and (b) the ionogram at SV=3800V and DP=200V in $N_2$ seeded with 1.5% IPA vapour. The black, red, and blue blocks labelled in the ionogram indicate peaks I, II and III, respectively.

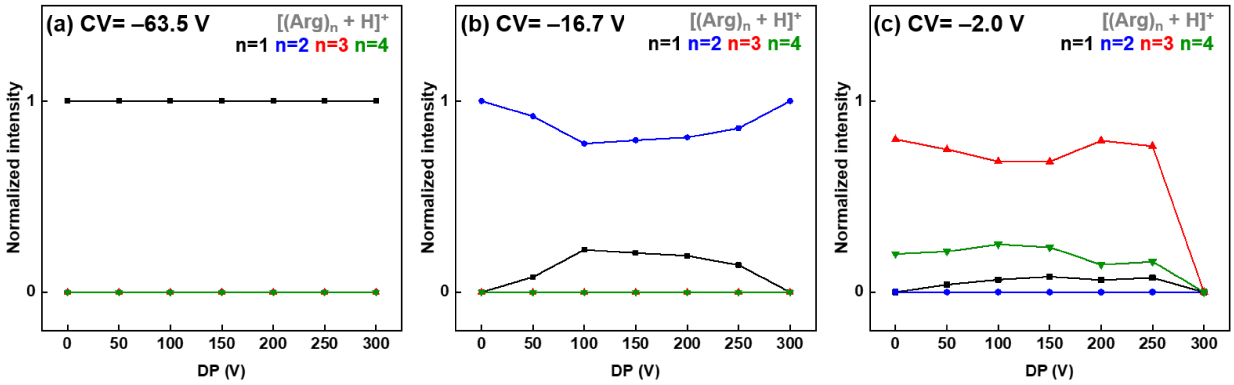


**Figure 4.11** The normalized intensity of *m/z*=175, 349, 523, and 697 in Q1 at (a) CV= −63.5 V, (b) −16.7 V and (c) CV= −2.0 V (SV=3800 V) as DP was ramped from 0 to 300 V in $N_2$ seeded with 1.5% IPA vapour.

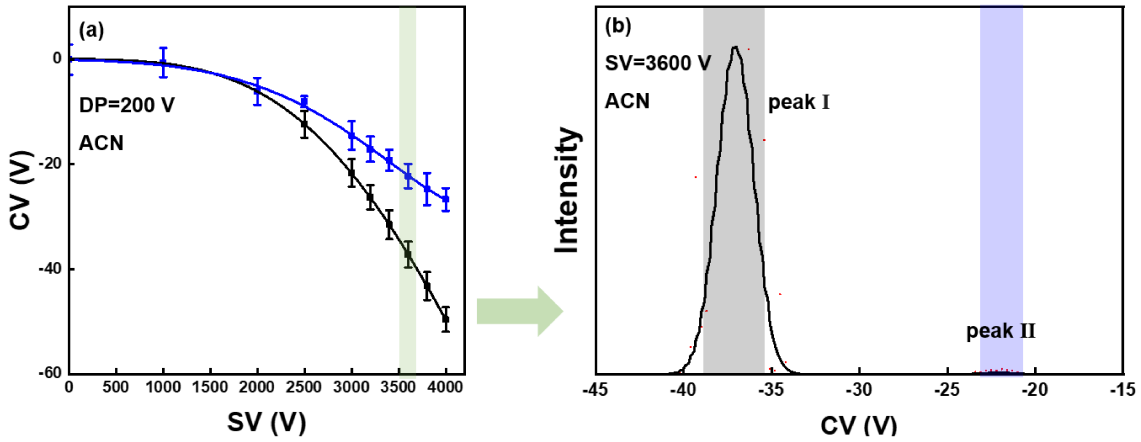


**Figure 4.12** (a) The dispersion plots at DP=200V and (b) the ionogram at SV=3800V and DP=200V in $N_2$ seeded with 1.5% ACN vapour. The black and blue blocks labelled in the ionogram indicate peaks I and II, respectively.

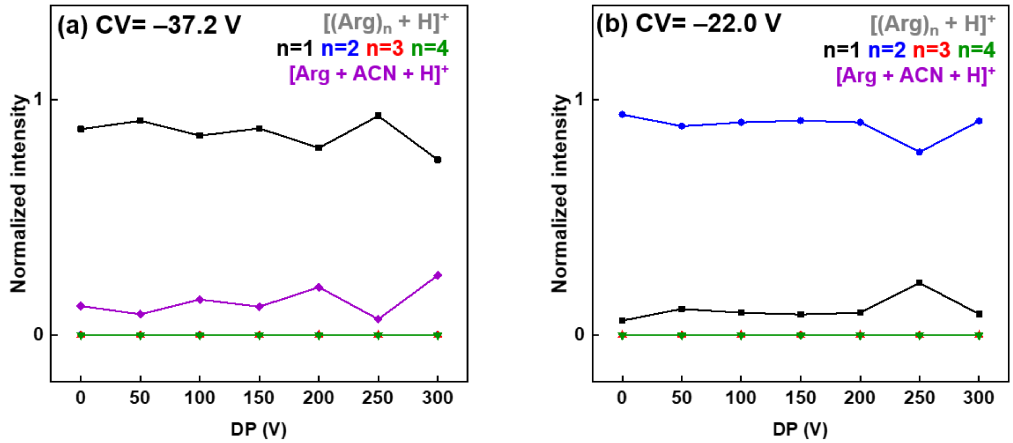


**Figure 4.13** The normalized intensity of *m/z*=175, 349, 523, and 697 in Q1 at (a) CV= −37.2 V and (b) CV= −22.0 V (SV=3800 V) as DP was ramped from 0 to 300 V in $N_2$ seeded with 1.5% ACN vapour.

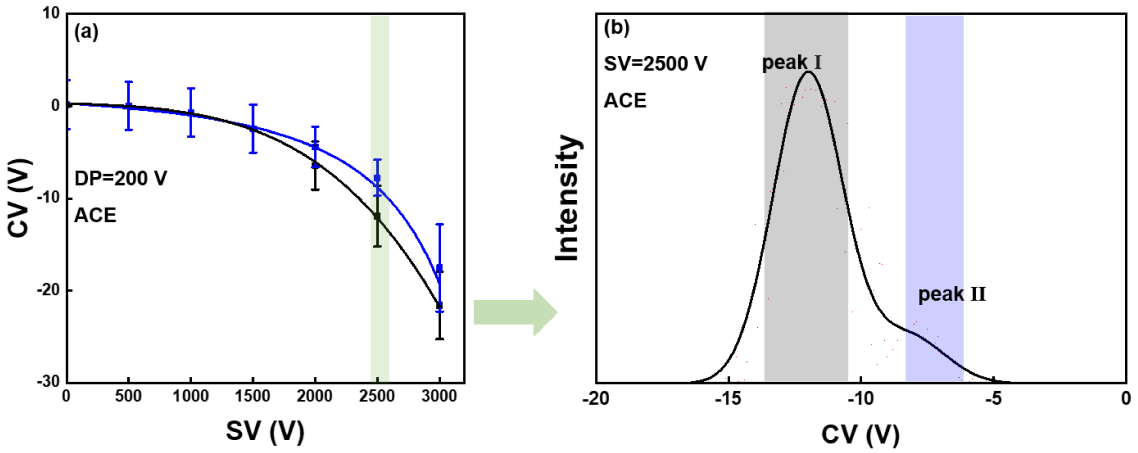


**Figure 4.14** The dispersion plots at DP=200 V and (b) the ionogram at SV=3000 V and DP=200V in $N_2$ seeded with 1.5% ACE vapour. The black and blue blocks labelled in the ionogram indicate peaks I and II, respectively.

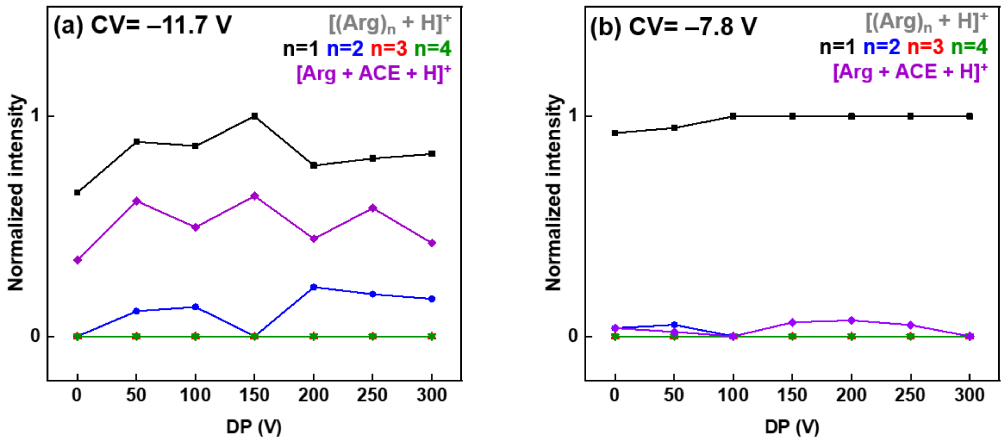


**Figure 4.15** The normalized intensity of *m/z*=175, 349, 523, and 697 in Q1 at (a) CV= −11.7 V and (b) CV= −7.8 V (SV=2500 V) as DP was ramped from 0 to 300 V in $N_2$ seeded with 1.5% ACE vapour.

Notably, when gating on the *m/z* of $[Arg + H]^+$, the intensity of $[(Arg)_3 + H]^+$ and $[(Arg)_4 + H]^+$ was not observed in Q1 under the ACE and ACN conditions, and signals of $[(Arg)_4 + H]^+$ were not detected in Q1 under $H_2O$. The diminishing intensities of protonated Arg clusters in Q1 indicated that these ions could be either fragmented when colliding with the solvents or be more stable under those gas-phase conditions; since a more strongly-bound structure was less expected to produce $[Arg + H]^+$ when the same DP was supplied. In order to investigate the presence and stability of the protonated Arg clusters, DMS behaviours of $[(Arg)_2 + H]^+$, $[(Arg)_3 + H]^+$, and $[(Arg)_4 + H]^+$ under the various modifier environments were studied by gating

on the parent mass of each cluster at *m/z* 349, 523, and 697, respectively in Q1. The dispersion plots and ionograms at SV=4000 V were recorded for $[(Arg)_n + H]^+$ (n=2−4) in various gas-phase environments and are shown in Figure 4.16 and Figure **4.17**, respectively. Besides, the comparison of $[Arg + H]^+$, $[(Arg)_2 + H]^+$, $[(Arg)_3 + H]^+$, and $[(Arg)_4 + H]^+$ ionograms at SV=4000 V in pure $N_2$ are shown in Figure 4.18.

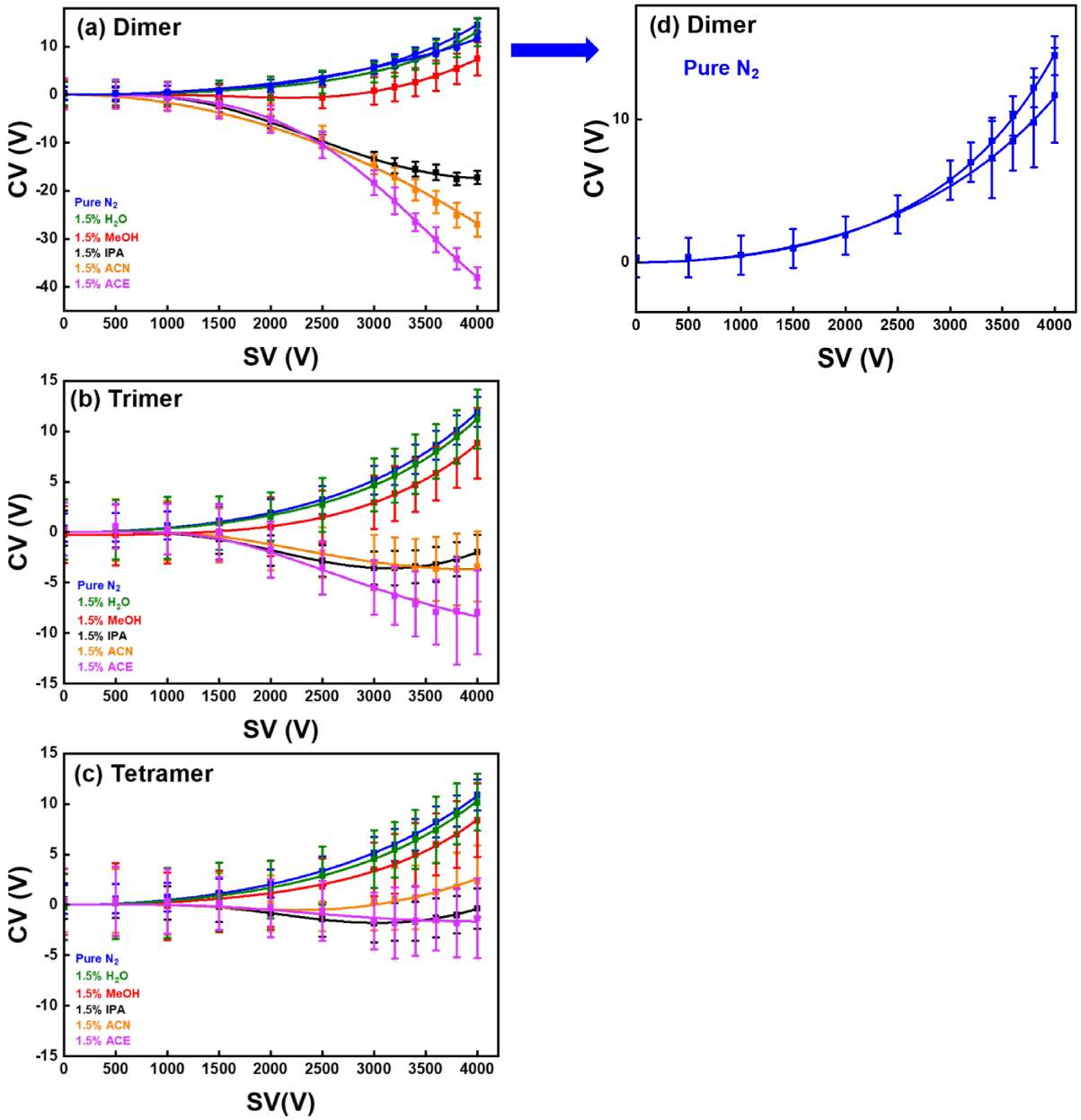


**Figure 4.16** The dispersion plots of the protonated Arg (a) dimer, (b) trimer, and (c) tetramer at DP=50 V electrosprayed from MeOH/$H_2O$ under various gas-phase modifier environments. (d) The dispersion plot of the protonated Arg dimer under $N_2$.

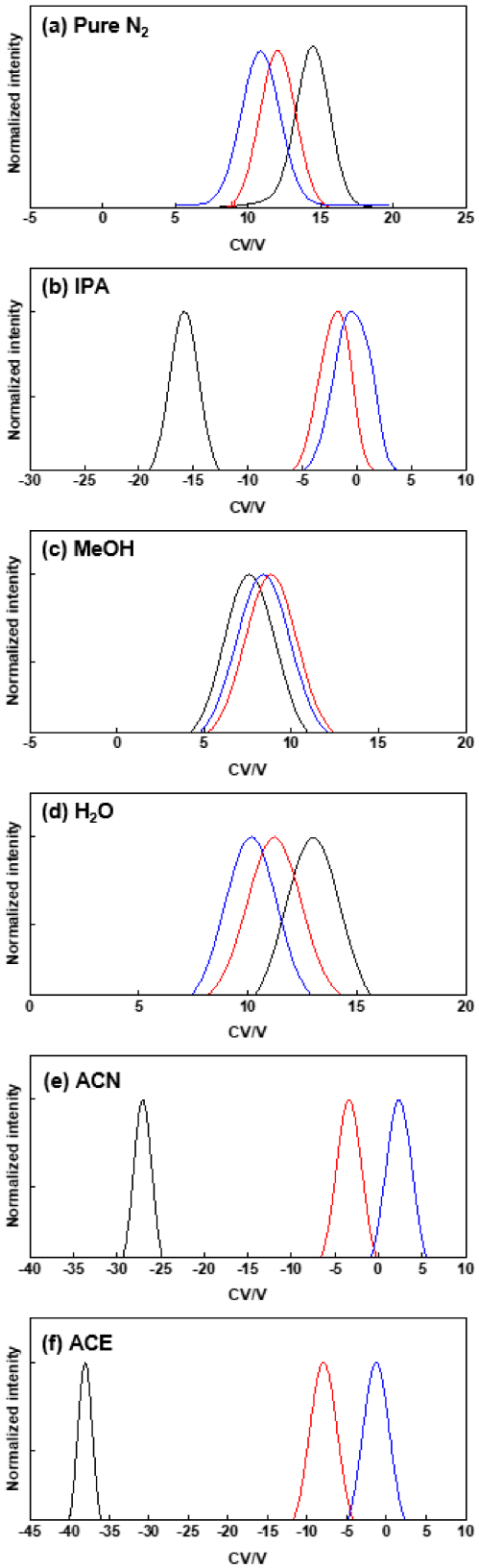


**Figure 4.17** The ionograms of the protonated Arg dimer (black), trimer (red), and tetramer (blue) at SV=4000 V under various gas-phase environments.

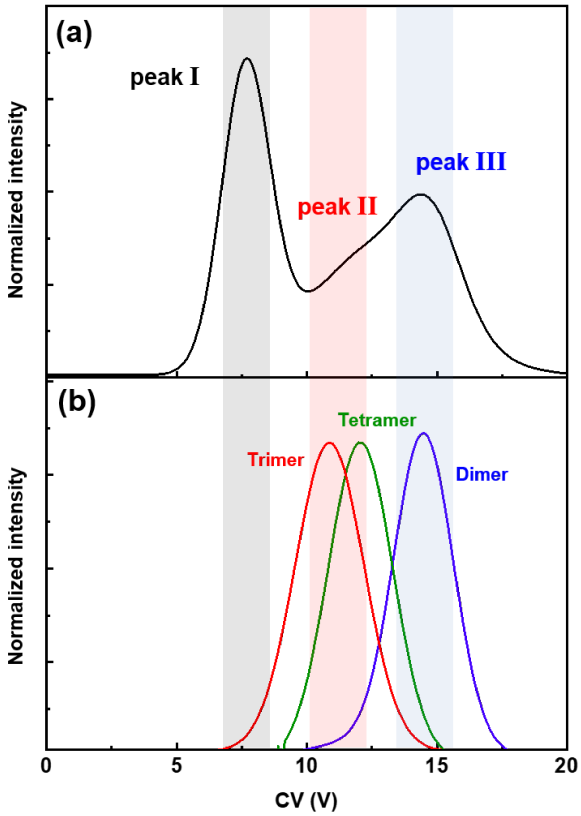


**Figure 4.18** The ionograms at SV=4000 V in pure $N_2$: (a) $[Arg + H]^+$ at DP=200 V and (b) $[(Arg)_2 + H]^+$ (blue), $[(Arg)_3 + H]^+$ (red) and $[(Arg)_4 + H]^+$ (green) at DP=50 V. The black, red, and blue blocks labelled in the ionogram indicate peaks I, II and III, respectively.

The DMS results of $[(Arg)_2 + H]^+$, $[(Arg)_3 + H]^+$ and $[(Arg)_4 + H]^+$ corresponded with the experimental findings observed in the dispersion plots of $[Arg + H]^+$. Two dispersion curves of $[(Arg)_2 + H]^+$ were resolved in pure $N_2$: one corresponding to $[(Arg)_2 + H]^+$ and the other corresponding to $[(Arg)_3 + H]^+$ according to the DMS clustering behaviour (SV/CV values). The observation of two curves in Figure 4.16d also corresponded to the existence of *m/z*=349 in both Figure 4.5b and 4.5c that the ion population of $[(Arg)_2 + H]^+$ appeared in Q1 at both CVs at 12.1 V and 14.8 V under $N_2$. Additionally, the peaks II and III observed in the ionogram of $[Arg + H]^+$ (Figure 4.18a) show excellent agreements with the peaks associated with each cluster (Figure 4.18b) – this conclusively proves that the monomer channel exhibits signal from the dimer, trimer, and tetramer due to fragmentation of the larger clusters. The overlapped ionogram peaks

shown in Figure 4.17a, 4.17c, and 4.17d suggest that the protonated Arg clusters were difficult to be separated under $N_2$, MeOH, and $H_2O$ − this matches with the observation of a combination of $[(Arg)_n + H]^+$ mass channels shown in Figure 4.7b, Figure **4.9**b, and Figure **4.11**b. The separation capability of the clusters was enhanced in ACE and ACN, as indicated by the isolated peaks shown in Figure **4.17**, revealing that the existences of the protonated Arg dimer, trimer, tetramer are independent rather than converted to each other in the gas phase. Non-clustering (Type C) behaviour was observed for protonated Arg trimer and tetramer in $N_2$, MeOH, and $H_2O$, suggesting that structures of $[(Arg)_3 + H]^+$ and $[(Arg)_4 + H]^+$ behaved like hard spheres within those gas-phase environments. Under the ACN and ACE modifier environments, strong ion-solvent interactions were observed between $[(Arg)_2 + H]^+$ and solvent molecules, resulting in strong clustering (Type A) behaviour. However, $[(Arg)_3 + H]^+$ and $[(Arg)_4 + H]^+$ showed weak-clustering behaviour in the ACN and ACE gas-phase environments (see Figure 4.16b and Figure **4.16**c). It is summarized that as the Arg cluster size increases, the clustering behaviour trends towards a non-clustering behaviour, suggesting that the charge was shielded greatly, resulting in weaker ion-solvent interactions. The protonated Arg trimer and tetramer detected in Q1 under all six gas-phase environments, as shown in Figure 4.16 and**Figure 4.17**, indicated that the vanishing of the *m/z*=523 and/or *m/z*=697 intensities when *m/z*=175 was selected in Q1 was due to the fact that $[(Arg)_3 + H]^+$ and/or $[(Arg)_4 + H]^+$ was more stable and tended to not fragment to $[Arg + H]^+$ when entering Q1 under $H_2O$, ACN and ACE modifier conditions. It is speculated that the molecular interactions (*i.e.,* ionic bonding and hydrogen bonding) exhibited by the protonated Arg cluster structure can potentially prevent hydrogen-bonding networks with solvent molecules, resulting in weak ion-dipole interactions. Computational studies are essential at this point to gain better insight into the

structures and interactions of protonated Arg clusters (*i.e.*, $[(Arg)_2 + H]^+$, $[(Arg)_3 + H]^+$ and $[(Arg)_4 + H]^+$) for understanding the clustering behaviours observed in the DMS cell.

### 4.3.3 Computational results

Protonated Arg structures with each possible protonation site (*i.e.*, the three nitrogens of the guanidino group, the nitrogen of amine group, and two oxygens of the carboxyl group) were manually modelled for initiating BH input files. Following BH, all accepted local minimum structures of protonated Arg were with protonation sites on nitrogens 1 or 2 of the guanidino group or the amine nitrogen, as shown in Figure 4.19. The lowest energy $[Arg + H]^+$ structures for protonation sites on $N_1$-guanidino, $N_2$-guanidino, and N-amine were optimized at the B3LYP/6-311++G(d,p) level of theory and are shown in Figure 4.20.

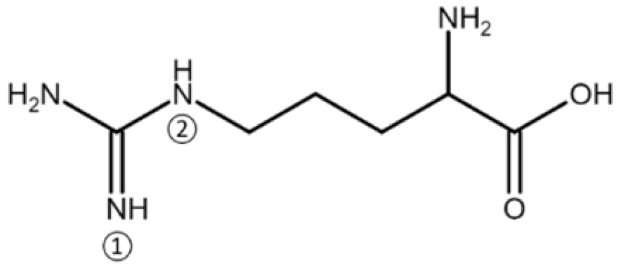


**Figure 4.19** The structure of Arg, where the double bond nitrogen and the secondary amine nitrogen on guanidino are labelled as 1 and 2, respectively.

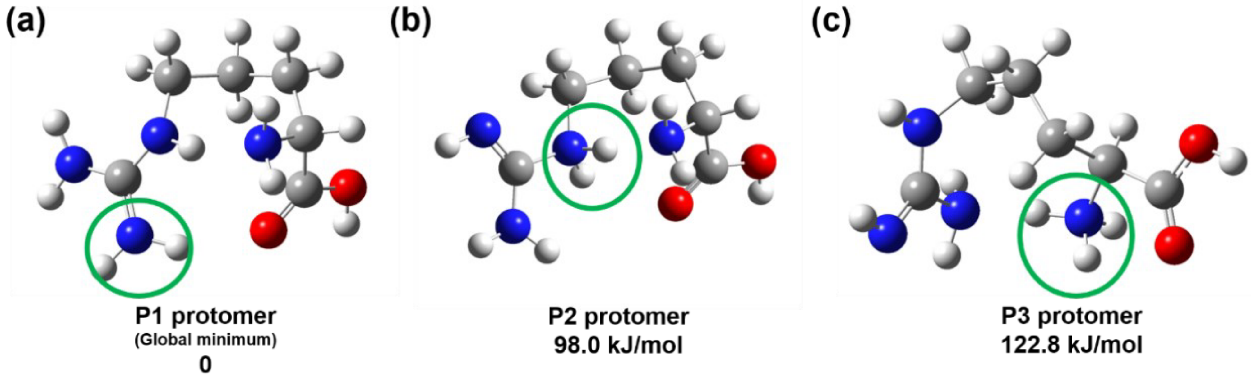


**Figure 4.20** The optimized lowest energy structures of $[Arg + H]^+$ with protonation sites on (a) $N_1$-guanidino, (b) $N_2$-guanidino and (c) N-amine, where green circles indicate single formal charge sites. The optimizations and relative Gibbs free energy to that of the global minimum structure were determined at the B3LYP/6-311++G(d,p) level of theory.

Calculations indicate that the global minimum structure of $[Arg + H]^+$ will exhibit protonation on the $N_1$-guanidino center in the gas phase. The relative Gibbs energies of P2 and P3 protomers were 98.0 kJ/mol and 122.8 kJ/mol, respectively. Thus, it is expected that these two isomers will not contribute to the gas phase ensemble owing to their negligible populations assuming a Boltzmann population distribution under equilibrium conditions. Taking the experimental DMS results into consideration, in which only one conformation of protonated Arg was resolved in the gas phase, it is expected that the geometry of $[Arg + H]^+$ is the $N_1$-guanidino protonated structure. The lowest-energy conformeric forms of the P1 protomer (see Figure 4.21) were then determined by basin hopping followed by geometry optimization at the B3LYP/6-311++G(d,p) level of theory. The $[Arg + H]^+$ conformers with relative Gibbs energies of less than 20.0 kJ/mol did not exhibit salt-bridge interactions; ionic bonding was not found and thus interactions between sites with formal positive and negative charges were not possible. The lowest energy structure of $[Arg + H]^+$ that contains a salt-bridge ($\Delta E$ = +59.7 kJ/mol) is shown in Figure **4.22**.

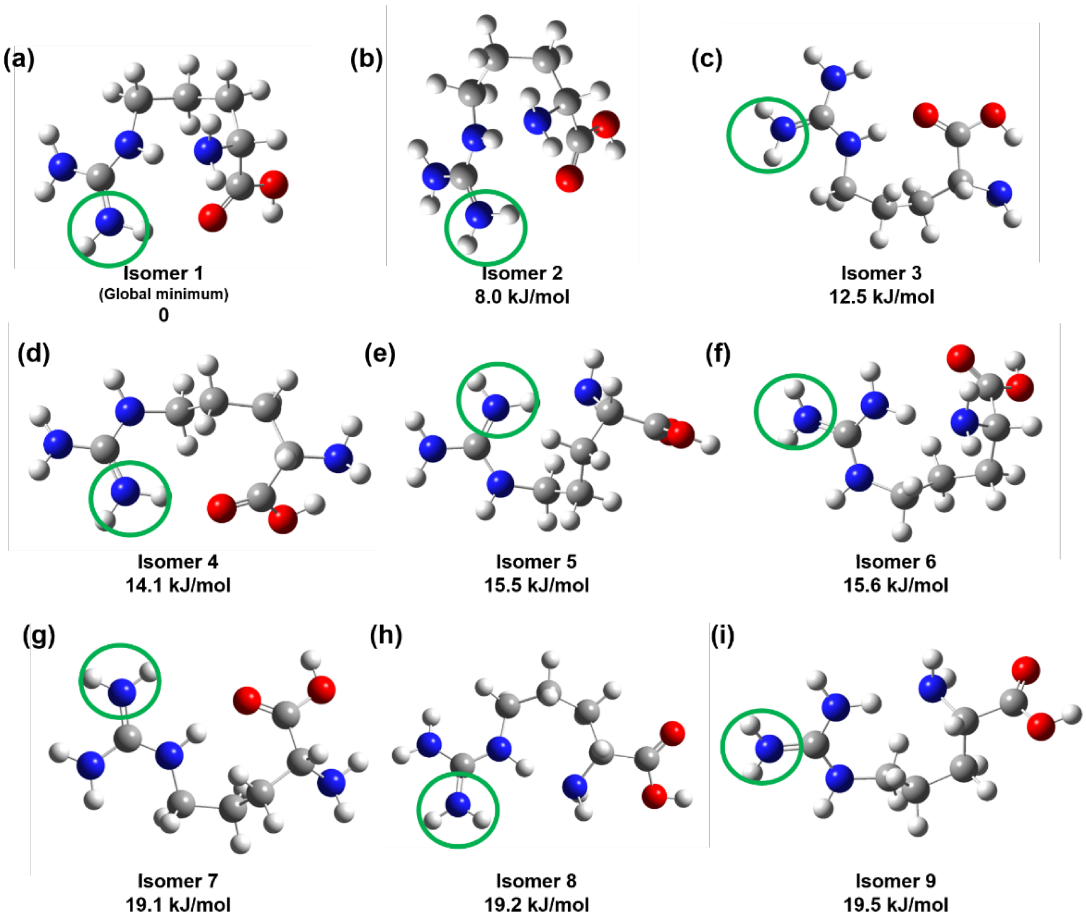


**Figure 4.21** The optimized structures of $[Arg + H]^+$ isomers with relative Gibbs free energies to that of the global minimum structure less than 20 kJ/mol at the B3LYP/6-311++G(d,p) level of theory. Green circles indicate single formal charge sites.

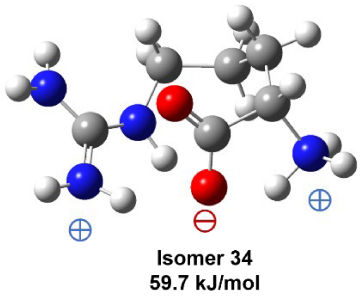


**Figure 4.22** The optimized lowest energy structure of $[Arg + H]^+$ which contains a salt-bridge with relative Gibbs free energies to that of the global minimum structure at the B3LYP/6-311++G(d,p) level of theory.

The geometries of the larger $[(Arg)_n + H]^+$ (n=2−4) clusters were also explored computationally at the B3LYP/6-311++G(d,p) level of theory. The optimized geometries of the low-energy isomers (within 40 kJ/mol of the computed global minimum) of $[(Arg)_2 + H]^+$ are

shown in Figure 4.23. Figure 4.24 and Figure **4.25** show the low-energy isomers of $[(Arg)_3 + H]^+$ and $[(Arg)_4 + H]^+$, respectively. As was the case for the monomer, the low-energy dimer and trimer species exhibited protonation at the $N_1$-guanidino sites. In the case of the tetramer, N-amine protonation was observed in addition to $N_1$-guanidino protonation for the global minimum structure. Intermolecular and intramolecular hydrogen bonding interactions were observed in all low-energy isomers of $[(Arg)_n + H]^+$ (n=2−4). Moreover, ionic bonding interactions were observed in the larger clusters owing to intracluster proton transfer between Arg moieties.

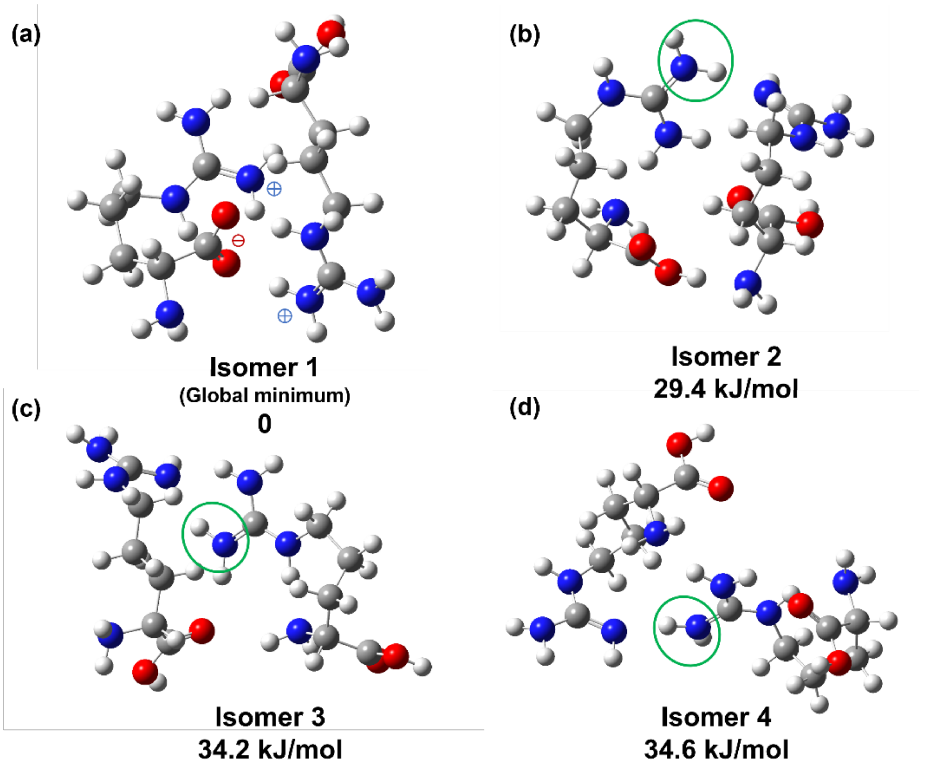


**Figure 4.23** The optimized structures of $[(Arg)_2 + H]^+$ isomers with relative Gibbs free energies to that of the global minimum structure less than 40.0 kJ/mol at the B3LYP/6-311++G(d,p) level of theory. Green circles indicate single protonation sites.

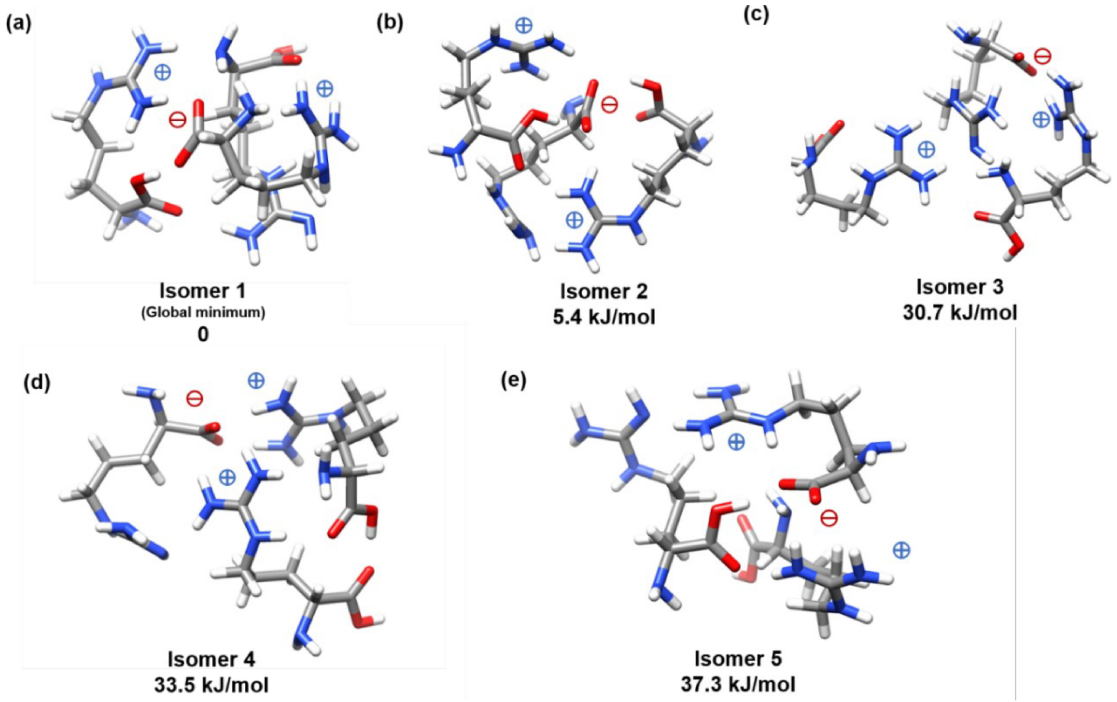


**Figure 4.24** The optimized structures of $[(Arg)_3 + H]^+$ isomers with relative Gibbs free energies to that of the global minimum structure less than 40.0 kJ/mol at the B3LYP/6-311++G(d,p) level of theory.

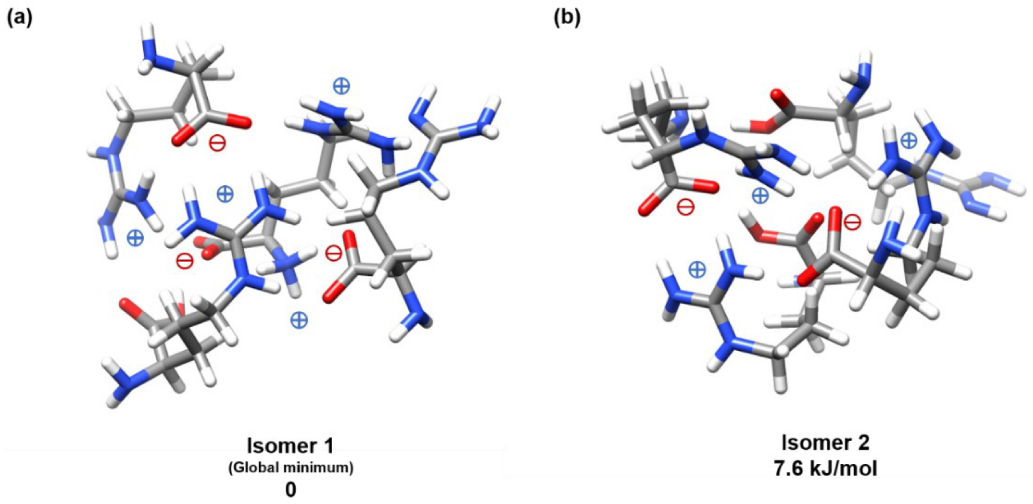


**Figure 4.25** The optimized structures of $[(Arg)_4 + H]^+$ isomers with relative Gibbs free energies to that of the global minimum structure less than 40.0 kJ/mol at the B3LYP/6-311++G(d,p) level of theory.

To investigate the stability of the $[(Arg)_n + H]^+$ (n=2−4) clusters, their bond-dissociation energies (BDEs) were calculated as per Equation (4.1), where *E* refers to the Gibbs energy of the respective moiety. BDE results are given in Table 4.3.

$$BDE = E_{[(Arg)_{n-1}+H]^+} + E_{Arg} - E_{[(Arg)_n+H]^+} \quad (4.1)$$

**Table 4.3** The calculated bond dissociation energy of protonated clusters at the B3LYP/6-311++G(d,p) level.

| ***Clusters*** | ***BDE (kJ/mol)*** |
|---|---|
| $[(Arg)_2 + H]^+$ | 102.2 |
| $[(Arg)_3 + H]^+$ | 79.5 |
| $[(Arg)_4 + H]^+$ | 132.0 |

The calculated BDEs for the $[(Arg)_n + H]^+$ (n=2−4) clusters lie in the range of 80−130 kJ/mol. The protonated Arg tetramer is the most stable among three Arg clusters, as indicated by its highest BDE. The stabilization of the $[(Arg)_4 + H]^+$ system could result from a greater number of salt-bridges presented in the protonated Arg tetramer structure compared to that of dimer and trimer. More BDE calculations need to be performed for $[Arg + ACN + H]^+$ and $[Arg + ACE + H]^+$ clusters in production of protonated Arg − this can provide a comparison of the Arg moiety binding affinity between $[(Arg)_{n-1} + H]^+$ and solvent molecules (*i.e.*, ACE and ACN).

One can also qualitative judge ion-solvent interactions by examining the electrostatic potential surfaces (ESPs) of the $[(Arg)_n + H]^+$ (n=1−4) species. Figure 4.26 shows the ESPs mapped onto the total electron density, thereby illustrating the charge distributions of the clusters. The monomer exhibits the region of highest partial positive (blue) charge amongst the four cluster systems. Charge shielding and delocalization due to the presence of the ionic bonding results in successively lower partial charges as n increases for $[(Arg)_n + H]^+$ (n=1−4). Consequently, ion-dipole interactions between the $[(Arg)_n + H]^+$ (n=1−4) clusters and solvent molecules decrease as the cluster size increases. This manifests as increasingly weak ion-solvent interactions, ultimately leading to hard-sphere behaviour for $[(Arg)_n + H]^+$ (n=1−4) as n increases (see Figure 4.16).

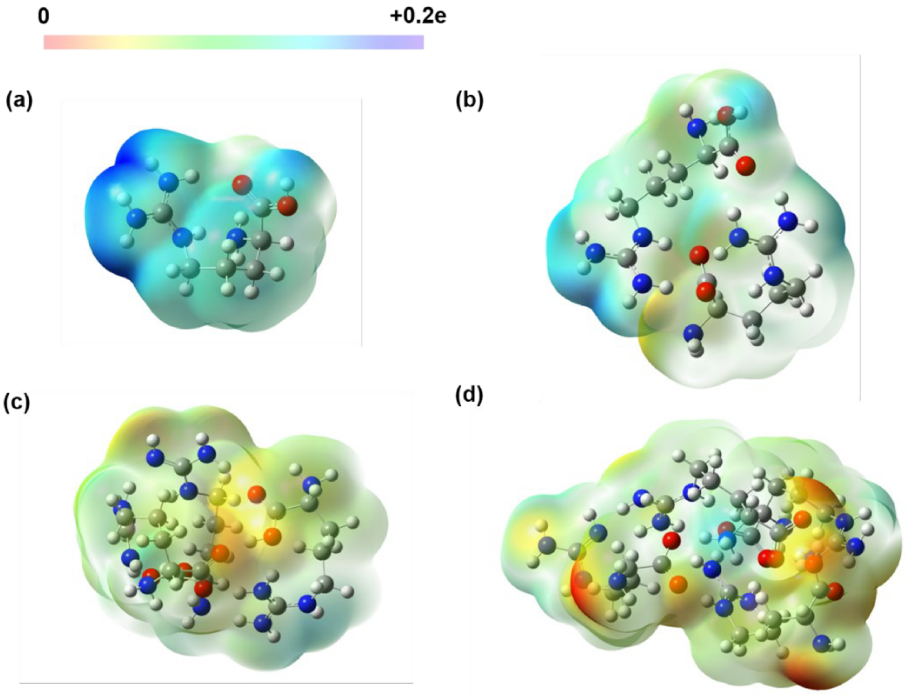


**Figure 4.26** The electrostatic potential maps of (a) $[Arg + H]^+$, (b) $[(Arg)_2 + H]_+$, (c) $[(Arg)_3 + H]_+$, and (d) $[(Arg)_4 + H]_+$ calculated at the B3LYP/6-311++G(d,p) level of theory, ranging from 0 (red) to +0.2e (blue).

## 4.4 Conclusions

In this chapter, the DMS behaviour of protonated Arg monomer, dimer, trimer, and tetramer was studied under $N_2$, $H_2O$, MeOH, IPA, ACN, and ACE. The ion-solvent interactions exhibited by $[Arg + H]^+$ increase with the trend $H_2O$ < MeOH < IPA < ACN < ACE. Multiple dispersion curves were observed for $[Arg + H]^+$ under all six environments. The major and minor sub-populations of $[Arg + H]^+$ were assigned to the bare $[Arg + H]^+$ ion and $[Arg + H]^+$ clusters (*i.e.*, dimer, trimer, or/and tetramer), respectively, under all those gas-phase conditions except for ACN and ACE, in which ion-solvent clusters were also detected. Both $[Arg + H]^+$ and $[Arg + ACN + H]^+$ contributed to the major sub-population in ACN. Unlike observations under other gas-phase conditions, in the ACE-modified environment, the predominant species was the $[Arg + ACE + H]^+$ cluster due to the great clustering affinity of ACE, whereas only a small

population of $[Arg + H]^+$ was monitored in the minor sub-ion population. A computational study of ion-solvent interactions would be valuable future work. Moreover, the protonated Arg tetramer tended to show a more hard-sphere (Type C) behaviour than the monomer, dimer, and trimer under the various gas-phase environments, suggesting charge shielding becomes more prominent as the size of the cluster increases.

The most likely gas-phase ensemble of $[Arg + H]^+$ contains a single protomeric form with protonation occurring at the $N_1$-guanidino site. Ionic bonding interactions were observed in the global minimum structures of $[(Arg)_2 + H]^+$, $[(Arg)_3 + H]^+$, and $[(Arg)_4 + H]^+$ but not $[Arg + H]^+$. The computed electrostatic potential maps illustrate that the charge is more shielded with increasing Arg cluster size due to an increase in ionic bonding interactions along with greater salt-bridge interactions, preventing ions from forming ion-solvent clusters. Hence, it is concluded that the salt-bridge interactions may contribute to the overall structural stability of Arg clusters.

# Chapter 5: Thesis conclusion

The structural and chemical properties of biological molecules can be obtained through the investigation of gas-phase chemistry. In this thesis, the structures and interactions of amino acids of interest were explored by studying their gaseous clustering behaviours with the application of DMS-MS. Spectroscopic studies of target ions were conducted with a recently modified DMS-MS instrument coupled to a probe laser system. The new UVPD modification provides an orthogonal method of structural characterization to aid in the determination of the geometric and electronic structures of DMS-selected ions. Quantum chemistry calculations at the DFT level were performed to justify experimental findings by modelling stable gas-phase structures of the ions of interest.

In Chapter 3, DMS and UVPD studies were performed to investigate the geometric and electronic structures of two aromatic amino acids (*i.e.*, Trp and Tyr) in the gas phase. Single protomeric species were identified for the DMS-resolved Trp and Tyr cations. It is predicted that the effective temperature of the $[(Trp)_2 + H]^+$ ion reached ~600 K before in-source CID of $[(Trp)_2 + H]^+$ producing $[Trp + H]^+$. The identical UVPD spectra of two different $[Trp + H]^+$ ion populations revealed that the protonated state produced from $[(Trp)_2 + H]^+$ was the same as that of the monomeric protonated Trp species. It is suggested that the protonation site of these species took place on the nitrogen atom of the amine group. The experimental UVPD spectra and computed vibronic spectra showed that electronic transitions involving $\pi^* \leftarrow \pi$ characters on the aromatic rings observed at 4.5 eV were monitored for both protonated Trp and Tyr. Intense transitions containing $\sigma^* \leftarrow \pi$ characters took place at 4.8 eV for $[Trp + H]^+$ and 5.5 eV for $[Tyr + H]^+$. Moreover, it is indicated that UVPD of $[Tyr - NH_3 + H]^+$ to produce $[Tyr - NH_3 - H_2O + H]^+$ is a stepwise dissociation process; as $[Tyr - NH_3 - H_2O + H]^+$ was most

abundant at 4.5 eV where the absorption of $[Tyr - NH_3 + H]^+$ was also intense. Furthermore, UVPD of $[Tyr + H]^+$ was a single-photon process whereby the photo-excited molecular cations underwent internal conversion and dissociated on the ground state potential energy surface.

In Chapter 4, the salt-bridge interactions of $[(Arg)_n + H]^+$ (n=1−4) were explored by studying the ion-solvent interactions between ions and various gas molecules (*i.e.*, $N_2$, $H_2O$, MeOH, IPA, ACN, and ACE) using DMS-MS. The interactions between $[Arg + H]^+$ and solvent molecules increase with the trend $H_2O$ < MeOH < IPA < ACN < ACE. In DMS experiments, multiple populations of $[Arg + H]^+$ were detected, of which the major one is assigned to the bare $[Arg + H]^+$ ion, and others are associated with $[(Arg)_2 + H]^+$, $[(Arg)_3 + H]^+$, and/or $[(Arg)_4 + H]^+$ under all six environments, with the exception of ACE. In ACE, the dominant species was attributed to $[Arg + ACE + H]^+$ due to strong ion-solvent interactions, and the bare $[Arg + H]^+$ cation was observed in the minor population. A small amount of $[Arg + ACN + H]^+$ also contributed to the major population in ACN. Meanwhile, less dynamic ion-solvent clustering interactions in the DMS cell were monitored for protonated Arg clusters as cluster size increased. Calculated Gibbs energies indicate that the most likely conformation of $[Arg + H]^+$ in the gas phase involved a single formal charge site on the nitrogen of the guanidino group instead of a salt-bridge. Conversely, salt-bridge interactions stabilized the structures of protonated Arg clusters as indicated by the prevalence of cluster formation in the DMS across different gas-phase environments. This is explained by the computed electrostatic potential maps in which the charge was more shielded in structures with greater salt-bridge interactions.

The gas-phase investigations into the geometric and electronic structures of amino acids and their clustering behaviours were accomplished by employing a combined experimental and computational approach. The methods proposed in this thesis are proven useful for structural

identification of amino acid ions and clusters by exploring their DMS behaviours and UVPD fragmentation processes. These techniques can be further applied to a wider field of biological systems for the exploration of their isomeric structures in the future.

# References:


1. Wu, G. Functional amino acids in growth, reproduction, and health. *Adv. Nutr.* **1**, 31–7 (2010).
2. Asif, M. *et al.* Amino acids: A review article. *Artic. J. Med. plant Res.* **5**, 3997–4000 (2011).
3. Kawashima, S. & Kanehisa, M. AAindex: Amino acid index database. *Nucleic Acids Res.* **27**, 368–369 (1999).
4. Ha, E. & Zemel, M. B. Functional properties of whey, whey components, and essential amino acids: Mechanisms underlying health benefits for active people (review). *J. Nutr. Biochem.* **14**, 251–258 (2003).
5. van Rossum, J. M. The relation between chemical structure and biological activity. *J. Pharm. Pharmacol.* **15**, 285–316 (1963).
6. Alberts B, Johnson A, L. J. The Shape and Structure of Proteins. in *Molecular Biology of the Cell.* (Garland Science, 2002).
7. Zeng, M. *et al.* Protein-protein interaction site prediction through combining local and global features with deep neural networks. *Bioinformatics* **36**, 1114–1120 (2020).
8. Zhang, L., Yu, G., Xia, D. & Wang, J. Protein–protein interactions prediction based on ensemble deep neural networks. *Neurocomputing* **324**, 10–19 (2019).
9. Ng, P. C. & Henikoff, S. SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Res.* **31**, 3812–3814 (2003).
10. Wagner, I. & Musso, H. New Naturally Occurring Amino Acids. *Angew. Chemie Int. Ed. English* **22**, 816–828 (1983).
11. Topol, I. A., Burt, S. K., Toscano, M. & Russo, N. Protonation of glycine and alanine: proton affinities, intrinsic basicities and proton transfer path. *J. Mol. Struct. THEOCHEM* **430**, 41–49 (1998).
12. 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.* **134**, 15805–15813 (2012).
13. Wysocki, V. H., Tsaprailis, G., Smith, L. L. & Breci, L. A. Mobile and localized protons: a framework for understanding peptide dissociation. *J. Mass Spectrom.* **35**, 1399–1406 (2000).
14. Kearley, G. J., Filiaux, F., Baron, M. H., Bennington, S. & Tomkinson, J. A new look at proton transfer dynamics along the hydrogen bonds in amides and peptides. *Science.* **264**, 1285–1289 (1994).
15. Green, M. K. & Lebrilla, C. B. Ion-molecule reactions as probes of gas-phase structures of peptides and proteins. *Mass Spectrom. Rev.* **16**, 53–71 (1997).
16. Huynh, M. H. V. & Meyer, T. J. Proton-coupled electron transfer. *Chem. Rev.* **107**, 5004–5064 (2007).
17. Holliday, G. L., Almonacid, D. E., Mitchell, J. B. O. & Thornton, J. M. The Chemistry of Protein Catalysis. *J. Mol. Biol.* **372**, 1261–1277 (2007).

18. Sibert, R. *et al.* Proton-coupled electron transfer in a biomimetic peptide as a model of enzyme regulatory mechanisms. *J. Am. Chem. Soc.* **129**, 4393–4400 (2007).

19. Wu, R. & McMahon, T. B. An Investigation of Protonation Sites and Conformations of Protonated Amino Acids by IRMPD Spectroscopy. *ChemPhysChem* **9**, 2826–2835 (2008).

20. Moser, A., Range, K. & York, D. M. Accurate proton affinity and gas-phase basicity values for molecules important in biocatalysis. *J. Phys. Chem. B* **114**, 13911–13921 (2010).

21. Campbell, S., Beauchamp, J. L., Rempe, M. & Lichtenberger, D. L. Correlations of lone pair ionization energies with proton affinities of amino acids and related compounds. Site specificity of protonation. *Int. J. Mass Spectrom. Ion Process.* **117**, 83–99 (1992).

22. Bouchoux, G., Desaphy, S., Bourcier, S., Malosse, C. & Bimbong, R. N. B. Gas-phase protonation thermochemistry of arginine. *J. Phys. Chem. B* **112**, 3410–3419 (2008).

23. Rak, J., Skurski, P., Simons, J. & Gutowski, M. Low-energy tautomers and conformers of neutral and protonated arginine. *J. Am. Chem. Soc.* **123**, 11695–11707 (2001).

24. Locke, M. J. & Mclver, R. T. Effect of Solvation on the Acid/Base Properties of Glycine. *J. Am. Chem. Soc.* **105**, 4226–4232 (1983).

25. Harrison, A. G. & Yalcin, T. Proton mobility in protonated amino acids and peptides. *Int. J. Mass Spectrom. Ion Process.* **165**–**166**, 339–347 (1997).

26. Barril, X., Alemán, C., Orozco, M. & Luque, F. J. Salt bridge interactions: Stability of the ionic and neutral complexes in the gas phase, in solution, and in proteins. *Proteins Struct. Funct. Genet.* **32**, 67–79 (1998).

27. Wu, R. & McMahon, T. B. Infrared multiple photon dissociation spectroscopy as structural confirmation for GlyGlyGlyH+ and AlaAlaAlaH+ in the gas phase. Evidence for amide oxygen as the protonation site. *J. Am. Chem. Soc.* **129**, 11312–11313 (2007).

28. Lorenz, U. J. & Rizzo, T. R. Multiple isomers and protonation sites of the phenylalanine/serine dimer. *J. Am. Chem. Soc.* **134**, 11053–11055 (2012).

29. Fu, W. & Hopkins, W. S. Applying Machine Learning to Vibrational Spectroscopy. *J. Phys. Chem. A* **122**, 167–171 (2018).

30. Pereverzev, A. Y., Kopysov, V. & Boyarkin, O. V. High Susceptibility of Histidine to Charge Solvation Revealed by Cold Ion Spectroscopy. *Angew. Chemie - Int. Ed.* **56**, 15639–15643 (2017).

31. Grégoire, G. *et al.* UV photoinduced dynamics in protonated aromatic amino acid. *Eur. Phys. J. D* **51**, 109–116 (2009).

32. Talbot, F. O., Tabarin, T., Antoine, R., Broyer, M. & Dugourd, P. Photodissociation spectroscopy of trapped protonated tryptophan. *J. Chem. Phys.* **122**, 3–7 (2005).

33. Boyarkin, O. V, Mercier, S. R., Kamariotis, A. & R., R. T. Electronic Spectroscopy of Cold, Protonated Tryptophan and Tyrosine. *J. Am. Chem. Soc.* **128**, 2816–2817 (2006).

34. Matthews, E. & Dessent, C. E. H. Experiment and theory confirm that UV laser photodissociation spectroscopy can distinguish protomers formed *via* electrospray. *Phys. Chem. Chem. Phys.* **19**, 17434–17440 (2017).

35. Marlton, S. J. P. *et al.* Selecting and identifying gas-phase protonation isomers of nicotineH+ using combined laser, ion mobility and mass spectrometry techniques. *Faraday Discuss.* **217**, 453–475 (2019).

36. Hopkins, W. S. Determining the properties of gas-phase clusters. *Mol. Phys.* **113**, 3151–3158 (2015).

37. Anwar, A. *et al.* Separating and probing tautomers of protonated nucleobases using differential mobility spectrometry. *Int. J. Mass Spectrom.* **429**, 174–181 (2018).

38. Krylov, E. V., Nazarov, E. G. & Miller, R. A. Differential mobility spectrometer: Model of operation. *Int. J. Mass Spectrom.* **266**, 76–85 (2007).

39. Coughlan, N. J. A. *et al.* Measuring Electronic Spectra of Differential Mobility-Selected Ions in the Gas Phase. *J. Am. Soc. Mass Spectrom.* **31**, 405–410 (2020).

40. Schneider, B. B., Covey, T. R., Coy, S. L., Krylov, E. V & Nazarov, E. G. Chemical effects in the separation process of a differential mobility/Mass spectrometer system. *Anal. Chem.* **82**, 1867–1880 (2010).

41. Campbell, J. L., Zhu, M. & Hopkins, W. S. Ion-molecule clustering in differential mobility spectrometry: Lessons learned from tetraalkylammonium cations and their isomers. *J. Am. Soc. Mass Spectrom.* **25**, 1583–1591 (2014).

42. 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.* **83**, 3470–3476 (2011).

43. Schneider, B. B., Nazarov, E. G., Londry, F., Vouros, P. & Covey, T. R. Differential mobility spectrometry/mass spectrometry history, theory, design optimization, simulations, and applications. *Mass Spectrom. Rev.* **35**, 687–737 (2016).

44. Stewart, J. J. P. Application of the PM6 method to modeling proteins. *J. Mol. Model.* **15**, 765–805 (2009).

45. Lataifeh, A. & Kraatz, H. B. Self-assembly of silver nanoparticles-low generation peptide dendrimer conjugates into poly-L-lysine. *Mater. Lett.* **254**, 353–356 (2019).

46. Bıçak, B., Gündüz, S., Kökçü, Y., Özel, A. E. & Akyüz, S. Molecular Docking and Molecular Dynamics Studies of L-Glycyl-L-Glutamic Acid Dipeptide. *Bilge Int. J. Sci. Technol. Res.* **3**, 1–9 (2019).

47. Wasileski, S. A. & Weaver, M. J. Electrode potential-dependent anion chemisorption and surface bond polarization as assessed by density functional theory. *J. Phys. Chem. B* **106**, 4782–4788 (2002).

48. Price, W. D., Jockusch, R. A. & Williams, E. R. Is arginine a zwitterion in the gas phase? *J. Am. Chem. Soc.* **119**, 11988–11989 (1997).

49. Wu, R. & McMahon, T. B. Stabilization of zwitterionic structures of amino acids (Gly, Ala, Val, Leu, Ile, Ser and Pro) by ammonium ions in the gas phase. *J. Am. Chem. Soc.* **130**, 3065–3078 (2008).

50. Prell, J. S., O'Brien, J. T., Steill, J. D., Oomens, J. & Williams, E. R. Structures of protonated dipeptides: The role of arginine in stabilizing salt bridges. *J. Am. Chem. Soc.*

**131**, 11442–11449 (2009).

51. 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.* **182**–**183**, 243–252 (1999).

52. Herrera-Zúñiga, L. D. *et al.* Molecular dynamics on laccase from Trametes versicolor to examine thermal stability induced by salt bridges. *Chem. Phys.* **517**, 253–264 (2019).

53. 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* **113**, 9995–10000 (2009).

54. Strittmatter, E. F. & Williams, E. R. Structures of protonated arginine dimer and bradykinin investigated by density functional theory: Further support for stable gas-phase salt bridges. *J. Phys. Chem. A* **104**, 6069–6076 (2000).

55. Hentzen, N. B., Islami, V., Köhler, M., Zenobi, R. & Wennemers, H. A Lateral Salt Bridge for the Specific Assembly of an ABC-Type Collagen Heterotrimer. *J. Am. Chem. Soc.* **142**, 2208–2212 (2020).

56. Barlow, D. J. & Thornton, J. M. Ion-pairs in proteins. *J. Mol. Biol.* **168**, 867–885 (1983).

57. Riordan, J. F., Mcelvany, K. D. & Borders, C. L. Arginyl residues: Anion recognition sites in enzymes. *Science.* **195**, 884–886 (1977).

58. Singh, J., Thornton, J. M., Snarey, M. & Campbell, S. F. The geometries of interacting arginine-carboxyls in proteins. *FEBS Lett.* **224**, 161–171 (1987).

59. Tormo, J. *et al.* Crystal structure of a human rhinovirus neutralizing antibody complexed with a peptide derived from viral capsid protein VP2. *EMBO J.* **13**, 2247–2256 (1994).

60. Raumann, B. E., Rould, M. A., Pabo, C. O. & Sauer, R. T. DNA recognition by β-sheets in the Arc represser-operator crystal structure. *Nature* **367**, 754–757 (1994).

61. Brown, B. M., Milla, M. E., Smith, T. L. & Sauer, R. T. Scanning mutagenesis of the Arc repressor as a functional probe of operator recognition. *Nat. Struct. Biol.* **1**, 164–168 (1994).

62. Lee, C. W., Wang, H. J., Hwang, J. K. & Tseng, C. P. Protein thermal stability enhancement by designing salt bridges: A combined computational and experimental study. *PLoS One* **9**, e112751 (2014).

63. Pylaeva, S., Brehm, M. & Sebastiani, D. Salt Bridge in Aqueous Solution: Strong Structural Motifs but Weak Enthalpic Effect. *Sci. Rep.* **8**, 13626 (2018).

64. Gupta, V. P. Interaction of Radiation and Matter and Electronic Spectra. in *Principles and Applications of Quantum Chemistry* 312–321 (2016). doi:10.1016/c2014-0-05143-x.

65. Schinke, R. Photodissociation Dynamics Spectroscopy and Pragmentation of Small Polyatomic Molecules. in *Cambridge monographs on atmonic, molecular, and chemical physics* 1–10 (1995).

66. Purves, R. W. & Guevremont, R. Electrospray ionization high-field asymmetric waveform ion mobility spectrometry-mass spectrometry. *Anal. Chem.* **71**, 2346–2357 (1999).

67. Mashmoushi, N. *Employing Differential Mobility Spectrometry as a tool for Metabolomics*

*distinguishing and Characterizing the Properties of Lipid Species*. (2019).

68. Van Dishoeck, E. F. & Visser, R. Molecular Photodissociation. in *Laboratory Astrochemistry: From Molecules through Nanoparticles to Grains* 229–254 (Wiley-VCH, 2015).

69. Kelley, A. M. Molecular vibrations and infrared spectroscopy and electronic spectroscopy. in *Condensed-phase molecular spectroscopy and photophysics* 115–158 (2012).

70. Nano, A. Towards optical memories: switchable optical systems for electron and energy transfer processes. (Université de Strasbourg, 2015).

71. Rüger, R., Niehaus, T., Van Lenthe, E., Heine, T. & Visscher, L. Vibrationally resolved UV/Vis spectroscopy with time-dependent density functional based tight binding. *J. Chem. Phys.* **145**, 1–14 (2016).

72. Macak, P., Luo, Y. & Ågren, H. Simulations of vibronic profiles in two-photon absorption. *Chem. Phys. Lett.* **330**, 447–456 (2000).

73. Martin, R. L. Natural transition orbitals. *J. Chem. Phys.* **118**, 4775–4777 (2003).

74. Höfener, S. & Klopper, W. Natural transition orbitals for the calculation of correlation and excitation energies. *Chem. Phys. Lett.* **679**, 52–59 (2017).

75. Wales, D. J. & Doye, J. P. K. Global optimization by basin-hopping and the lowest energy structures of Lennard-Jones clusters containing up to 110 atoms. *J. Phys. Chem. A* **101**, 5111–5116 (1997).

76. Olson, B., Hashmi, I., Molloy, K. & Shehu, A. Basin Hopping as a General and Versatile Optimization Framework for the Characterization of Biological Macromolecules. *Adv. Artif. Intell.* **2012**, 1–19 (2012).

77. Hansson, T., Oostenbrink, C. & Van Gunsteren, W. F. Molecular dynamics simulations. *Curr. Opin. Struct. Biol.* **12**, 190–196 (2002).

78. Karplus, M. & McCammon, J. A. Molecular dynamics simulations of biomolecules. *Nat. Struct. Biol.* **9**, 646–652 (2002).

79. Rappé, A. K., Casewit, C. J., Colwell, K. S., Goddard, W. A. & Skiff, W. M. UFF, a Full Periodic Table Force Field for Molecular Mechanics and Molecular Dynamics Simulations. *J. Am. Chem. Soc.* **114**, 10024–10035 (1992).

80. Jász, Á., Rák, Á., Ladjánszki, I. & Cserey, G. Optimized GPU implementation of Merck Molecular Force Field and Universal Force Field. *J. Mol. Struct.* **1188**, 227–233 (2019).

81. Breneman, C. M. & Wiberg, K. B. Determining atom-centered monopoles from molecular electrostatic potentials. The need for high sampling density in formamide conformational analysis. *J. Comput. Chem.* **11**, 361–373 (1990).

82. Cohen, A. J., Mori-Sánchez, P. & Yang, W. Insights into current limitations of density functional theory. *Science.* **321**, 792–794 (2008).

83. Koch, W. & Holthausen, M. C. The defination of the Model. in *A Chemist's Guide to Density Functional Theory* 1–27 (Wiley, 2001).

84. Orio, M., Pantazis, D. A. & Neese, F. Density functional theory. *Photosynth. Res.* **102**, 443–

453 (2009).

85. Burke, K. Perspective on density functional theory. *J. Chem. Phys.* **136**, (2012).

86. Hohenberg, P. & Kohn, W. Inhomogeneous electron gas. *Phys. Rev.* **136**, B864 (1964).

87. Grimme, S. Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *J. Comput. Chem.* **27**, 1787–1799 (2006).

88. Zhang, Y. & Yang, W. A challenge for density functionals: Self-interaction error increases for systems with a noninteger number of electrons. *J. Chem. Phys.* **109**, 2604–2608 (1998).

89. Liu, J., Laachi, N., Delaney, K. T. & Fredrickson, G. H. Advantages and limitations of density functional theory in block copolymer directed self-assembly. *Altern. Lithogr. Technol. VII* **9423**, 94231I (2015).

90. Slominski, A. *et al.* Conversion of L-tryptophan to serotonin and melatonin in human melanoma cells. *FEBS Lett.* **511**, 102–106 (2002).

91. Yan, L. *et al.* Stress response to low temperature: Transcriptomic characterization in Crassostrea sikamea × Crassostrea angulata hybrids. *Aquac. Res.* **49**, 3374–3385 (2018).

92. Nolting, D., Marian, C. & Weinkauf, R. Protonation effect on the electronic spectrum of tryptophan in the gas phase. *Phys. Chem. Chem. Phys.* **6**, 2633–2640 (2004).

93. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratch, D. J. Gaussian 16, Revision B.01. (2016).

94. Stewart, J. J. P. Optimization of parameters for semiempirical methods V: Modification of NDDO approximations and application to 70 elements. *J. Mol. Model.* **13**, 1173–1213 (2007).

95. Neese, F. The ORCA program system. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2**, 73–78 (2012).

96. Cowan, D. O. & Drisko, R. L. The Triplet State. in *Elements of Organic Photochemistry* (ed. Cowan, D. O.) 205–266 (Springer US, 1976).

97. Zhang, P. *et al.* Revisiting Fragmentation Reactions of Protonated α-Amino Acids by High-Resolution Electrospray Ionization Tandem Mass Spectrometry with Collision-Induced Dissociation. *Sci. Rep.* **9**, 1–10 (2019).

98. Anslyn, E. V. & Dougherty, D. A. Organic Reaction Mechanisms. in *Modern Physical Organic Chemistry* 537–624 (University Science Books, 2006).

99. Jaeqx, S., Oomens, J. & Rijs, A. M. Gas-phase salt bridge interactions between glutamic acid and arginine. *Phys. Chem. Chem. Phys.* **15**, 16341–16352 (2013).

100. Scheiner, S., Kar, T. & Pattanayak, J. Comparison of various types of hydrogen bonds involving aromatic amino acids. *J. Am. Chem. Soc.* **124**, 13257–13264 (2002).

101. Guin, D. & Gruebele, M. Weak Chemical Interactions That Drive Protein Evolution: Crowding, Sticking, and Quinary Structure in Folding and Function. *Chem. Rev.* **119**, 10691–10717 (2019).

102. Ieritano, C. *et al.* How Hot Are Your Ions in Differential Mobility Spectrometry? *J. Am. Soc. Mass Spectrom.* **31**, 582–593 (2020).

# APPENDICES

## Appendix A − Supplementary information for Chapter 3

### Experimental parameters

**Table A.1** The experimental information of the gas source and DMS conditions used for protonated Trp and Tyr.

| Gas Source Conditions | |
|---|---|
| Polarity | Positive |
| Ion Source gas 1 | 20 psi |
| Ion Source gas 2 | 0 |
| Curtain Gas | 20 psi |
| Ion Spray Voltage | 5500 V |
| Collision Gas | High |
| Temperature | 34 ℃ |
| DMS Conditions | |
| Experimental type | Enhanced product ion |
| Temperature | 150 ℃ |
| Offset | −3.0 V |

**Table A.2** The experimental information of the compound conditions for dispersion plots.

| Compound Conditions for Dispersion Plots | | | |
|---|---|---|---|
| $[Trp + H]^+$ | | $[Tyr + H]^+$ | |
| De-clustering Potential | 100 V | De-clustering Potential | 100 V |
| Collision Energy | 10 V | Collision Energy | 10 V |
| Entrance Potential | 10 V | Entrance Potential | 10 V |

**Table A.3** The experimental information of the compound conditions for UVPD.

| Compound Conditions for UVPD | | |
|---|---|---|
| Sample | $[Trp + H]^+$ | $[Tyr + H]^+$ |
| De-clustering Potential | 0 | 100 V |
| Collision Energy | 5 V | 10 V |
| Entrance Potential | 10 V | 5 V |
| LIT Fixed Time | 5 ms | 3 ms |

## Geometries of $[Trp + H]^+$ with water molecules at the hydrogen bonding sites

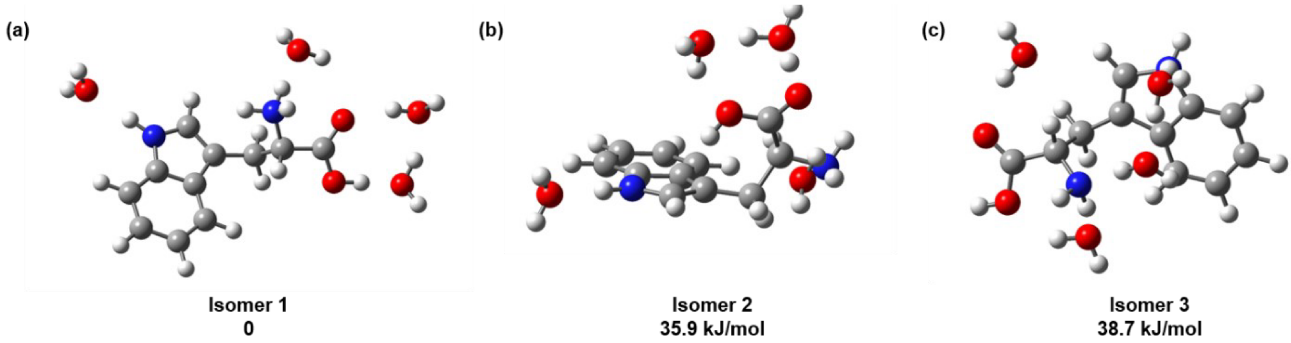


**Figure A.1** The structures of Trp protomers with water molecules at the hydrogen bonding sites using the PCM at the CAM-B3LYP/6-311++G(d,p) level of theory.

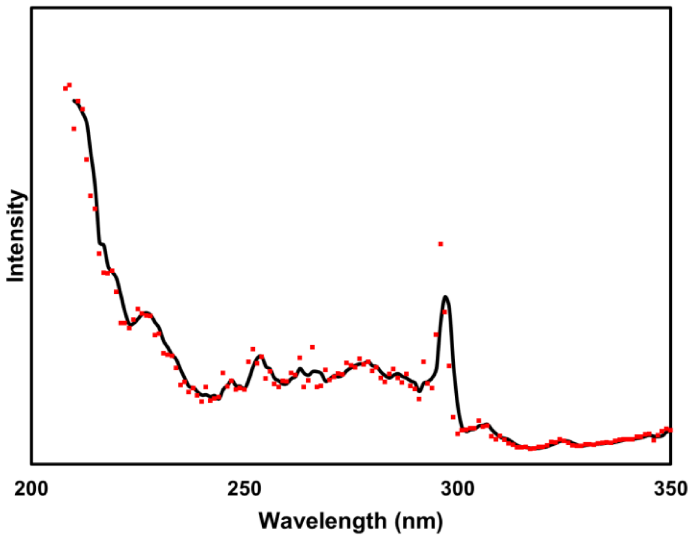


**Figure A.2** The UVPD spectrum of $[Trp + H]^+$ generated with the *m/z*=188 fragmentation channel.

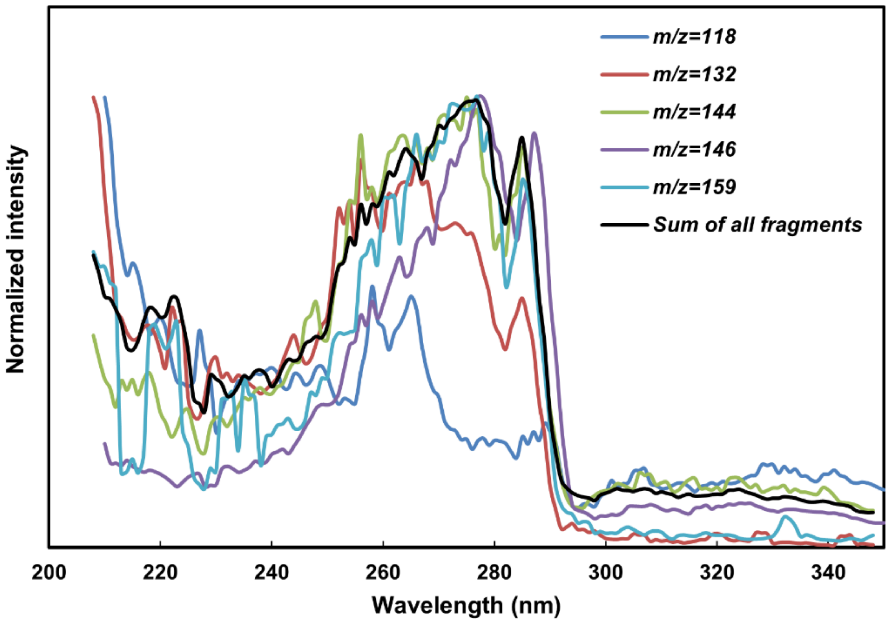


**Figure A.3** The UVPD spectra of $[Trp + H]^+$ generated with each fragmentation channel.

## UVPD spectra of $[Trp - NH_3 + H]^+$

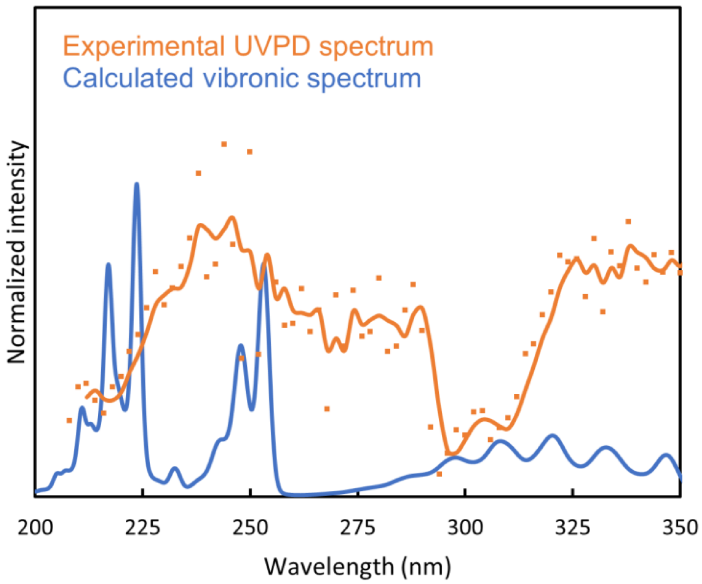


**Figure A.4** The experimental UVPD spectrum (orange) and the calculated vibronic spectrum (blue) of $[Trp - NH_3 + H]^+$ (*m/z*=188)

## Geometry coordinates of $[Trp + H]^+$

| | Isomer 1 | | | | Isomer 2 | | | | Isomer 3 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | X (Å) | Y(Å) | Z(Å) | | X (Å) | Y(Å) | Z(Å) | | X (Å) | Y(Å) | Z(Å) |
| **C** | -1.311365 | -0.138560 | -0.291853 | C | -1.179676 | -0.204991 | 0.345882 | C | 1.542932 | -0.274337 | 0.069647 |
| **C** | -2.430332 | 0.674851 | -0.028791 | C | -2.196599 | 0.716224 | 0.034708 | C | 2.502834 | 0.711140 | -0.094556 |
| **C** | -3.674283 | 0.144092 | 0.308365 | C | -3.434890 | 0.320446 | -0.467683 | C | 3.854398 | 0.480690 | -0.175674 |
| **C** | -3.780367 | -1.228917 | 0.371190 | C | -3.643025 | -1.030850 | -0.639851 | C | 4.251350 | -0.851204 | -0.077445 |
| **C** | -2.681388 | -2.061628 | 0.099449 | C | -2.650125 | -1.971570 | -0.317013 | C | 3.314848 | -1.866977 | 0.090240 |
| **C** | -1.452509 | -1.532979 | -0.232152 | C | -1.425091 | -1.574760 | 0.173081 | C | 1.952618 | -1.595198 | 0.164418 |
| **C** | -0.218323 | 0.748570 | -0.607978 | C | -0.053462 | 0.563875 | 0.827611 | C | 0.212298 | 0.357965 | 0.105845 |
| **C** | -0.712827 | 2.026407 | -0.526047 | C | -0.446227 | 1.884016 | 0.806740 | C | 0.359671 | 1.676119 | -0.026093 |
| **H** | -4.527406 | 0.781642 | 0.505104 | H | -4.206665 | 1.042013 | -0.705267 | H | 4.580279 | 1.273920 | -0.306919 |
| **H** | -4.732631 | -1.677144 | 0.624809 | H | -4.594321 | -1.377470 | -1.022963 | H | 5.304135 | -1.095452 | -0.132945 |
| **H** | -2.810882 | -3.135671 | 0.140302 | H | -2.859796 | -3.025314 | -0.449435 | H | 3.655958 | -2.891825 | 0.163264 |
| **H** | -0.626848 | -2.194548 | -0.470766 | H | -0.684464 | -2.320346 | 0.440569 | H | 1.237012 | -2.397110 | 0.293691 |
| **H** | -0.222469 | 2.966456 | -0.732258 | H | 0.090372 | 2.764041 | 1.129997 | H | -0.361191 | 2.476932 | -0.008098 |
| **C** | 1.182018 | 0.361810 | -0.954297 | C | 1.236860 | 0.018567 | 1.366198 | C | -1.054746 | -0.412649 | 0.278692 |
| **H** | 1.219497 | -0.342366 | -1.787760 | H | 1.829606 | 0.818236 | 1.820326 | H | -0.970618 | -1.364307 | -0.247777 |
| **H** | 1.753178 | 1.243907 | -1.257288 | H | 1.039190 | -0.714934 | 2.152031 | H | -1.154805 | -0.679476 | 1.337523 |
| **C** | 1.889067 | -0.319221 | 0.242754 | C | 2.084492 | -0.699850 | 0.292131 | C | -2.319423 | 0.298433 | -0.196714 |
| **H** | 1.484712 | -1.316734 | 0.404829 | H | 1.464195 | -1.472756 | -0.173664 | H | -2.190921 | 0.567156 | -1.251785 |
| **C** | 3.395170 | -0.368881 | 0.062258 | C | 2.462355 | 0.231091 | -0.845685 | C | -3.487542 | -0.689777 | -0.176135 |
| **O** | 3.749652 | -1.237503 | -0.862861 | O | 1.652553 | 1.038941 | -1.390937 | O | -3.377597 | -1.879506 | -0.078676 |
| **O** | 4.132000 | 0.346847 | 0.687497 | O | 3.651187 | 0.181072 | -1.268040 | O | -4.660345 | -0.058077 | -0.309269 |
| **H** | 2.473905 | 1.093421 | 1.630648 | H | 4.088832 | -0.494060 | -0.650397 | H | -5.368997 | -0.718802 | -0.322157 |
| **N** | 1.649009 | 0.490112 | 1.482156 | N | 3.338387 | -1.288017 | 0.743454 | N | -2.547854 | 1.535821 | 0.537372 |
| **H** | 0.790242 | 1.050187 | 1.335064 | H | 3.644764 | -0.929617 | 1.642093 | H | -2.728577 | 1.349388 | 1.519226 |
| **H** | 1.529609 | -0.088012 | 2.315426 | H | 3.316997 | -2.299301 | 0.788633 | H | -3.363570 | 2.018570 | 0.176888 |
| **N** | -2.033705 | 1.987121 | -0.179585 | N | -1.717953 | 1.977496 | 0.327695 | N | 1.804535 | 2.011017 | -0.167831 |
| **H** | -2.633353 | 2.792050 | -0.094997 | H | -2.228858 | 2.838416 | 0.211616 | H | 1.985905 | 2.485627 | -1.058247 |
| **H** | 4.712942 | -1.221833 | -0.985586 | H | 0.799514 | 1.071444 | -0.866488 | H | 2.106720 | 2.642947 | 0.580764 |

## Geometry coordinates of $[Trp + H]^+$ with water molecules at the hydrogen bonding sites

| | Isomer 1 | | | | Isomer 2 | | | | Isomer 3 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | X (Å) | Y(Å) | Z(Å) | | X (Å) | Y(Å) | Z(Å) | | X (Å) | Y(Å) | Z(Å) |
| **C** | -2.059972 | -0.909513 | -0.176839 | C | -1.801583 | -0.984923 | -0.006314 | C | 1.384101 | -0.398214 | -0.761547 |
| **C** | -3.329787 | -0.300081 | -0.067412 | C | -2.726284 | 0.081279 | -0.023423 | C | 2.507882 | 0.411335 | -0.682819 |
| **C** | -4.464444 | -1.016771 | 0.317056 | C | -3.847003 | 0.076921 | -0.855448 | C | 3.785806 | -0.066429 | -0.51808 |
| **C** | -4.309841 | -2.362042 | 0.587417 | C | -4.033995 | -1.025078 | -1.665981 | C | 3.921714 | -1.451091 | -0.443177 |
| **C** | -3.055898 | -2.98962 | 0.476894 | C | -3.131711 | -2.104994 | -1.652765 | C | 2.809176 | -2.288005 | -0.517594 |
| **C** | -1.934296 | -2.278547 | 0.0987 | C | -2.020873 | -2.094905 | -0.832992 | C | 1.525805 | -1.775216 | -0.67503 |
| **C** | -1.142052 | 0.118378 | -0.596884 | C | -0.783277 | -0.638913 | 0.954449 | C | 0.211074 | 0.467158 | -0.905789 |
| **C** | -1.881833 | 1.265092 | -0.718913 | C | -1.145121 | 0.584716 | 1.465294 | C | 0.617227 | 1.735936 | -0.917602 |
| **H** | -5.431239 | -0.534663 | 0.396639 | H | -4.544154 | 0.90583 | -0.858577 | H | 4.64113 | 0.592617 | -0.447191 |
| **H** | -5.171005 | -2.947256 | 0.88712 | H | -4.895102 | -1.062735 | -2.32239 | H | 4.907527 | -1.880497 | -0.319403 |
| **H** | -2.973442 | -4.048441 | 0.69091 | H | -3.316808 | -2.956632 | -2.296194 | H | 2.949001 | -3.359518 | -0.447727 |
| **H** | -0.976754 | -2.77942 | 0.011185 | H | -1.338179 | -2.93651 | -0.825285 | H | 0.6611 | -2.425965 | -0.717204 |
| **H** | -1.565323 | 2.250145 | -1.028237 | H | -0.651965 | 1.174206 | 2.223764 | H | 0.051913 | 2.651402 | -0.987796 |
| **C** | 0.325113 | -0.028302 | -0.847993 | C | 0.395 | -1.472334 | 1.35785 | C | -1.206795 | 0.005683 | -0.973382 |
| **H** | 0.534535 | -0.913577 | -1.450438 | H | 0.903752 | -1.01853 | 2.209425 | H | -1.750494 | 0.618751 | -1.690564 |
| **H** | 0.699391 | 0.83308 | -1.406974 | H | 0.057449 | -2.461523 | 1.672639 | H | -1.249908 | -1.032057 | -1.30843 |
| **C** | 1.13369 | -0.172744 | 0.45576 | C | 1.434321 | -1.727853 | 0.247106 | C | -1.898437 | 0.126219 | 0.39898 |
| **H** | 0.794153 | -1.049727 | 1.00657 | H | 0.906983 | -2.131711 | -0.625156 | H | -1.81099 | 1.162968 | 0.727049 |
| **C** | 2.621013 | -0.317197 | 0.193085 | C | 2.086613 | -0.458236 | -0.290077 | C | -3.384792 | -0.096055 | 0.168533 |
| **O** | 2.899275 | -1.428781 | -0.422367 | O | 1.356157 | 0.631389 | -0.428004 | O | -4.116783 | 0.762921 | -0.265631 |
| **O** | 3.419409 | 0.539646 | 0.526365 | O | 3.253032 | -0.447025 | -0.661185 | O | -3.803983 | -1.330091 | 0.448725 |
| **H** | 1.383925 | 1.87485 | 0.948099 | H | 3.864796 | 0.695932 | -1.213475 | H | -4.753867 | -1.39516 | 0.260158 |
| **N** | 0.907929 | 1.015774 | 1.322955 | N | 2.402427 | -2.706918 | 0.691379 | N | -1.267481 | -0.734785 | 1.388714 |
| **H** | -0.098849 | 1.201886 | 1.373825 | H | 3.036681 | -2.950888 | -0.060347 | H | -1.246113 | -1.700462 | 1.062586 |
| **H** | 1.249429 | 0.866991 | 2.273149 | H | 2.961539 | -2.299201 | 1.435861 | H | -1.791885 | -0.721296 | 2.258136 |
| **N** | -3.190786 | 1.023878 | -0.403635 | N | -2.301524 | 1.022016 | 0.882574 | N | 2.068583 | 1.804201 | -0.711444 |
| **H** | -3.936479 | 1.718203 | -0.42226 | H | -2.767286 | 1.906896 | 1.084751 | H | 2.550341 | 2.372559 | -1.41172 |
| **H** | 3.891317 | -1.540778 | -0.627603 | H | 0.458983 | 0.507637 | -0.03006 | H | 2.235903 | 2.222877 | 0.26731 |
| **O** | 6.194736 | 0.54861 | 0.123886 | O | -3.687194 | 3.475397 | 1.392792 | O | 2.126451 | 2.439613 | 1.831957 |
| **H** | 6.75064 | 0.649886 | 0.90293 | H | -4.239242 | 3.586232 | 2.174004 | H | 2.90128 | 2.687153 | 2.346373 |
| **H** | 5.276703 | 0.69929 | 0.411509 | H | -3.317834 | 4.345044 | 1.20725 | H | 1.798875 | 1.571352 | 2.180983 |
| **O** | 5.381534 | -1.733984 | -1.009903 | O | 4.29608 | 1.585188 | -1.597936 | O | -1.62531 | -3.255048 | -0.211683 |
| **H** | 5.899237 | -0.998705 | -0.608569 | H | 4.722758 | 1.42738 | -2.45232 | H | -1.476604 | -4.200864 | -0.112211 |
| **H** | 5.802584 | -2.560719 | -0.752321 | H | 3.536515 | 2.289477 | -1.68551 | H | -2.579044 | -3.127437 | -0.137815 |
| **O** | -5.298295 | 2.986481 | -0.459655 | O | 2.295965 | 3.065614 | -1.684787 | O | 1.182014 | 0.030067 | 2.476871 |
| **H** | -5.585195 | 3.558225 | 0.259684 | H | 1.648482 | 2.527576 | -1.204164 | H | 0.331655 | -0.236361 | 2.035685 |
| **H** | -5.822584 | 3.240454 | -1.225994 | H | 2.301779 | 3.942078 | -1.282277 | H | 1.805035 | -0.682416 | 2.300562 |
| **O** | 2.254163 | 3.247725 | 0.447991 | O | 3.24538 | -0.216373 | 2.290427 | O | -2.43928 | 2.990298 | -0.998239 |
| **H** | 3.101802 | 2.942029 | 0.10383 | H | 3.248359 | 0.03983 | 3.218428 | H | -2.824305 | 3.864134 | -1.112771 |
| **H** | 1.940404 | 3.931172 | -0.154048 | H | 4.084287 | 0.096346 | 1.935314 | H | -3.179704 | 2.402295 | -0.777865 |

## Geometries of $[Tyr + H]^+$ with water molecules at the hydrogen bonding sites

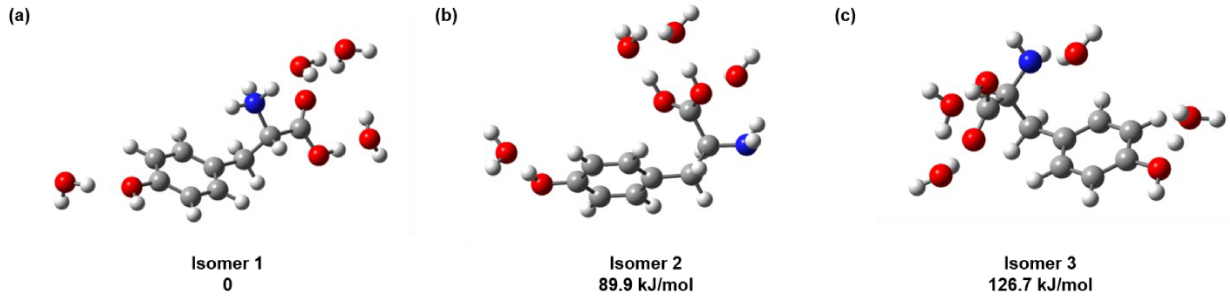


**Figure A.5** The structures of protonated Tyr protomers with water molecules at the hydrogen bonding sites using the PCM at the CAM-B3LYP/6-311++G(d,p) level of theory.

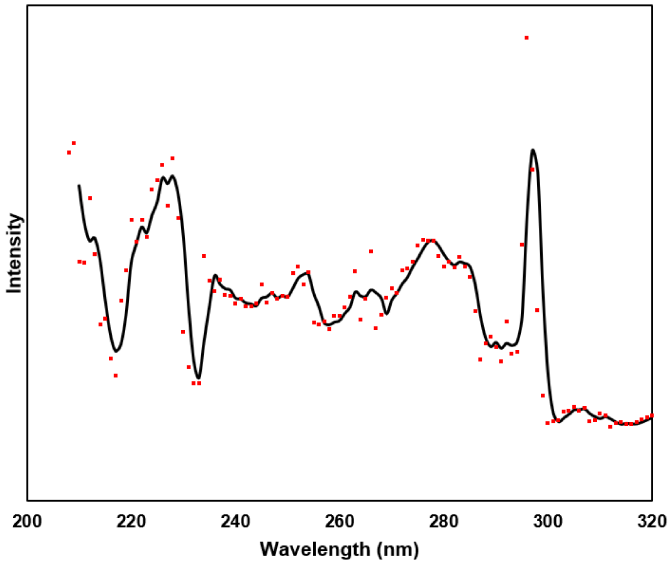


**Figure A.6** The UVPD spectrum of $[Tyr + H]^+$ generated with *m/z*=165 and *m/z*=136 fragmentation channels.

**UVPD spectra of $[Tyr - NH_3 + H]^+$**

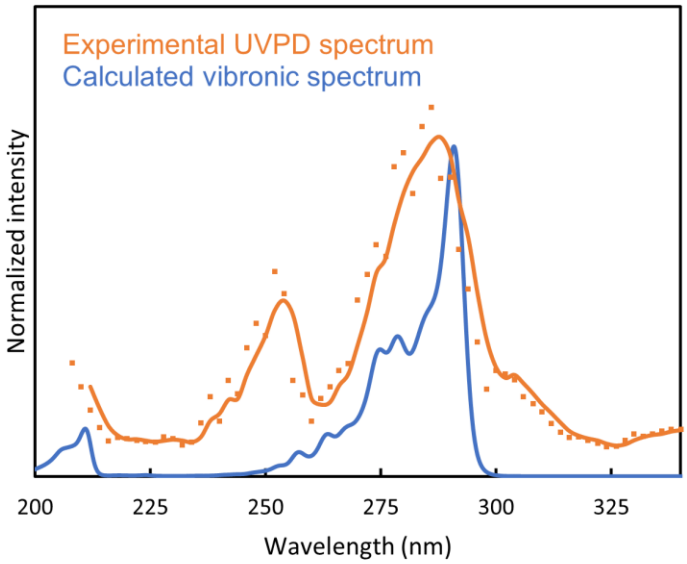


**Figure A.7** The experimental UVPD spectrum (orange) and the calculated vibronic spectrum (blue) of $[Tyr - NH_3 + H]^+$ (*m/z*=165).

**Table A.4** Thermodynamic parameters of the protonated Trp dimer at various temperatures.

| Temperature (K) | Sum of electronic and zero-point Energies (Hartree) | Zero-point correction (Hartree) | Thermal correction to Gibbs Free Energy (Hartree) | Sum of electronic and thermal Free Energies (Hartree) |
|---|---|---|---|---|
| **298** | -1372.462296 | 0.458955 | 0.396896 | -1372.524355 |
| **400** | -1372.462296 | 0.458955 | 0.363075 | -1372.558176 |
| **500** | -1372.462296 | 0.458955 | 0.324263 | -1372.596988 |
| **550** | -1372.462296 | 0.458955 | 0.302849 | -1372.618402 |
| **600** | -1372.462296 | 0.458955 | 0.280143 | -1372.641108 |
| **625** | -1372.462296 | 0.458955 | 0.268317 | -1372.652934 |
| **650** | -1372.462296 | 0.458955 | 0.256184 | -1372.665067 |

**Table A.5** Thermodynamic parameters of protonated Trp at various temperatures.

| Temperature (K) | Sum of electronic and zero-point Energies (Hartree) | Zero-point correction (Hartree) | Thermal correction to Gibbs Free Energy (Hartree) | Sum of electronic and thermal Free Energies (Hartree) |
|---|---|---|---|---|
| **298** | -686.381067 | 0.235673 | 0.195731 | -686.421009 |
| **400** | -686.381067 | 0.235673 | 0.175958 | -686.440782 |
| **500** | -686.381067 | 0.235673 | 0.153793 | -686.462947 |
| **550** | -686.381067 | 0.235673 | 0.141721 | -686.475019 |
| **600** | -686.381067 | 0.235673 | 0.12901 | -686.48773 |
| **625** | -686.381067 | 0.235673 | 0.122421 | -686.494319 |
| **650** | -686.381067 | 0.235673 | 0.115679 | -686.501061 |

**Table A.6** Thermodynamic parameters of neutral Trp at various temperatures.

| Temperature (K) | Sum of electronic and zero-point Energies (Hartree) | Zero-point correction (Hartree) | Thermal correction to Gibbs Free Energy (Hartree) | Sum of electronic and thermal Free Energies (Hartree) |
|---|---|---|---|---|
| **298** | -686.024449 | 0.221309 | 0.181085 | -686.064673 |
| **400** | -686.024449 | 0.221309 | 0.161161 | -686.084597 |
| **500** | -686.024449 | 0.221309 | 0.138855 | -686.106903 |
| **550** | -686.024449 | 0.221309 | 0.126719 | -686.119039 |
| **600** | -686.024449 | 0.221309 | 0.113953 | -686.131805 |
| **625** | -686.024449 | 0.221309 | 0.107339 | -686.138419 |
| **650** | -686.024449 | 0.221309 | 0.100574 | -686.145184 |

## Geometry coordinates of $[Tyr + H]^+$

| | Isomer 1 | | | | Isomer 2 | | | | Isomer 3 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | X (Å) | Y(Å) | Z(Å) | | X (Å) | Y(Å) | Z(Å) | | X (Å) | Y(Å) | Z(Å) |
| **N** | -1.367677 | -1.082891 | 1.096389 | N | 3.463092 | -1.030066 | 0.056019 | N | 2.121658 | 1.638073 | -0.598743 |
| **H** | -1.984809 | -1.835905 | 0.751947 | H | 3.814043 | -1.727686 | -0.588965 | H | 2.672901 | 1.447741 | -1.426763 |
| **C** | -1.777685 | 0.192127 | 0.413933 | C | 2.089094 | -0.624473 | -0.21533 | C | 2.252336 | 0.571788 | 0.384803 |
| **H** | -1.636529 | 1.012639 | 1.116128 | H | 1.794342 | -0.975113 | -1.20979 | H | 3.194387 | 0.635371 | 0.950984 |
| **C** | -0.913182 | 0.406078 | -0.845811 | C | 1.032811 | -1.098166 | 0.804727 | C | 1.112686 | 0.621482 | 1.416965 |
| **C** | -3.249048 | 0.018739 | 0.07429 | C | 2.143585 | 0.892275 | -0.333233 | C | 2.321593 | -0.795274 | -0.27254 |
| **H** | -1.205941 | -0.319036 | -1.610227 | H | 1.09367 | -2.18849 | 0.832423 | H | 1.315853 | -0.113171 | 2.195626 |
| **H** | -1.159358 | 1.394909 | -1.235605 | H | 1.308592 | -0.739734 | 1.800476 | H | 1.13948 | 1.607155 | 1.887193 |
| **C** | 0.556759 | 0.274496 | -0.537321 | C | -0.374434 | -0.656404 | 0.464909 | C | -0.254532 | 0.380313 | 0.833158 |
| **O** | -3.786398 | -1.055328 | 0.127938 | O | 3.284135 | 1.427016 | -0.323736 | O | 1.852032 | -1.798387 | 0.192239 |
| **O** | -3.800774 | 1.152515 | -0.305427 | O | 1.125236 | 1.641539 | -0.453606 | O | 3.017504 | -0.771421 | -1.416016 |
| **C** | 1.267928 | -0.858705 | -0.926142 | C | -1.027524 | -1.151985 | -0.673344 | C | -0.878165 | -0.857483 | 0.999107 |
| **C** | 1.236215 | 1.272577 | 0.170006 | C | -1.081252 | 0.210518 | 1.299889 | C | -0.905715 | 1.380989 | 0.110821 |
| **H** | 0.781609 | -1.630317 | -1.515436 | H | -0.523843 | -1.857779 | -1.326239 | H | -0.364672 | -1.645904 | 1.532106 |
| **C** | 2.615472 | -1.003498 | -0.617831 | C | -2.321483 | -0.783248 | -0.973929 | C | -2.134932 | -1.105918 | 0.465443 |
| **C** | 2.573343 | 1.141255 | 0.481122 | C | -2.384378 | 0.589594 | 1.007515 | C | -2.162703 | 1.167137 | -0.439988 |
| **H** | 0.718026 | 2.180891 | 0.460614 | H | -0.617352 | 0.589535 | 2.204484 | H | -0.419272 | 2.334199 | -0.03816 |
| **H** | 3.154442 | -1.885499 | -0.945333 | H | -2.82881 | -1.170925 | -1.847807 | H | -2.620904 | -2.065766 | 0.591994 |
| **C** | 3.271965 | -0.002687 | 0.088734 | C | -3.010376 | 0.092667 | -0.131176 | C | -2.71095 | -0.071329 | -0.229867 |
| **H** | 3.106556 | 1.918252 | 1.013595 | H | -2.91401 | 1.260441 | 1.674419 | H | -2.681409 | 1.931476 | -1.00302 |
| **O** | 4.574725 | -0.063676 | 0.423975 | O | -4.268426 | 0.407964 | -0.482389 | H | 3.081971 | -1.674545 | -1.759743 |
| **H** | 4.996215 | -0.85549 | 0.074566 | H | -4.685367 | 0.991026 | 0.160652 | H | 2.441768 | 2.518892 | -0.213549 |
| **H** | -0.378416 | -1.281232 | 0.880204 | H | 3.597986 | -1.366719 | 1.003863 | H | -4.102807 | -0.94004 | -1.546435 |
| **H** | -4.728735 | 1.008621 | -0.553902 | H | 0.288356 | 1.128349 | -0.330284 | H | -4.787787 | -0.406381 | -0.185835 |
| **H** | -1.480349 | -1.037975 | 2.110956 | H | 3.906947 | 0.630732 | -0.187689 | O | -4.06439 | -0.285913 | -0.826299 |

## Geometry coordinates of $[Tyr + H]^+$ with water molecules at the hydrogen bonding sites

| | Isomer 1 | | | | Isomer 2 | | | | Isomer 3 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | X (Å) | Y(Å) | Z(Å) | | X (Å) | Y(Å) | Z(Å) | | X (Å) | Y(Å) | Z(Å) |
| **N** | -3.112753 | -2.114626 | -0.658096 | N | 0.956472 | 1.175691 | -0.429409 | N | -0.810017 | 2.135511 | 1.223594 |
| **H** | -3.25101 | -1.822472 | -1.619524 | H | 1.208637 | 1.176443 | -1.420409 | H | -1.460508 | 2.262446 | 1.989053 |
| **C** | -1.870441 | -1.596492 | -0.131746 | C | 1.165853 | -0.169343 | 0.168125 | C | -1.508163 | 1.76345 | 0.004831 |
| **H** | -1.714092 | -1.980273 | 0.877296 | H | 1.060942 | -0.060348 | 1.247775 | H | -2.291205 | 2.486234 | -0.275604 |
| **C** | -0.68854 | -2.039199 | -1.022985 | C | 0.140165 | -1.184875 | -0.355443 | C | -0.540457 | 1.695748 | -1.189303 |
| **C** | -1.832538 | -0.0927 | -0.008956 | C | 2.588292 | -0.597272 | -0.135701 | C | -2.27554 | 0.461551 | 0.156353 |
| **H** | -0.728376 | -3.129472 | -1.044819 | H | 0.273494 | -1.310397 | -1.432538 | H | -1.116658 | 1.434994 | -2.076691 |
| **H** | -0.877657 | -1.688277 | -2.041041 | H | 0.374307 | -2.141881 | 0.111696 | H | -0.141678 | 2.700271 | -1.339418 |
| **C** | 0.658034 | -1.559365 | -0.548833 | C | -1.276462 | -0.770741 | -0.045185 | C | 0.598689 | 0.726081 | -1.004968 |
| **O** | -1.248581 | 0.407221 | 0.989329 | O | 3.290441 | 0.007919 | -0.92048 | O | -2.701927 | -0.143684 | -0.814244 |
| **O** | -2.339777 | 0.572879 | -0.973267 | O | 2.935303 | -1.66922 | 0.520849 | O | -2.47859 | 0.090167 | 1.397809 |
| **C** | 1.290105 | -2.166563 | 0.536084 | C | -1.772447 | -0.864123 | 1.254191 | C | 0.472276 | -0.60834 | -1.383647 |
| **C** | 1.288996 | -0.474396 | -1.148531 | C | -2.111518 | -0.262395 | -1.037412 | C | 1.801024 | 1.148446 | -0.439472 |
| **H** | 0.822767 | -3.017836 | 1.019995 | H | -1.146367 | -1.264974 | 2.043819 | H | -0.453966 | -0.952666 | -1.826168 |
| **C** | 2.506455 | -1.705614 | 1.00993 | C | -3.062569 | -0.462128 | 1.558919 | C | 1.510753 | -1.511463 | -1.192534 |
| **C** | 2.509308 | 0.001761 | -0.686277 | C | -3.405396 | 0.146314 | -0.747586 | C | 2.85441 | 0.265759 | -0.253568 |
| **H** | 0.822701 | 0.014363 | -1.997704 | H | -1.751687 | -0.18885 | -2.057731 | H | 1.911117 | 2.185883 | -0.147719 |
| **H** | 2.994759 | -2.18313 | 1.850714 | H | -3.437643 | -0.545739 | 2.572494 | H | 1.401905 | -2.550667 | -1.479869 |
| **C** | 3.123129 | -0.614201 | 0.401026 | C | -3.8762 | 0.045537 | 0.552855 | C | 2.683498 | -1.051474 | -0.630694 |
| **H** | 2.990986 | 0.846434 | -1.163869 | H | -4.054051 | 0.537876 | -1.520909 | H | 3.794367 | 0.593243 | 0.172156 |
| **O** | 4.313339 | -0.197837 | 0.899782 | O | -5.158059 | 0.46023 | 0.793276 | H | -3.032479 | -0.745484 | 1.461999 |
| **H** | 4.655252 | 0.567986 | 0.394489 | H | -5.394737 | 0.34372 | 1.720488 | H | -0.30635 | 3.006261 | 1.076594 |
| **H** | -3.890972 | -1.764174 | -0.108641 | H | -0.023849 | 1.454517 | -0.348367 | H | 4.202046 | -2.045833 | 0.769666 |
| **H** | -2.327715 | 1.577597 | -0.925793 | H | 3.876019 | -1.936782 | 0.287641 | H | 3.67941 | -2.756954 | -0.901117 |
| **H** | -1.155221 | 1.449601 | 1.071334 | H | 1.562319 | 1.901208 | 0.047953 | O | 3.767493 | -1.925847 | -0.41376 |
| **O** | 5.237424 | 1.942617 | -0.52652 | O | -7.086375 | 0.259232 | -1.284025 | O | -3.977945 | -2.557697 | -0.991585 |
| **H** | 6.020977 | 1.829262 | -1.075678 | H | -7.308615 | -0.670808 | -1.384222 | H | -4.761168 | -2.546253 | -1.550057 |
| **H** | 5.356806 | 2.781194 | -0.067375 | H | -6.462253 | 0.295179 | -0.54219 | H | -3.557493 | -1.677828 | -1.076685 |
| **O** | -0.93607 | 2.856372 | 1.240438 | O | 2.630771 | 2.949799 | 0.722334 | O | -3.87396 | -2.118526 | 1.651383 |
| **H** | -1.446775 | 3.352947 | 0.576176 | H | 2.623832 | 3.132953 | 1.667022 | H | -4.087389 | -2.443044 | 0.74773 |
| **H** | -1.12569 | 3.230849 | 2.108584 | H | 3.560672 | 2.756065 | 0.466828 | H | -4.702756 | -2.068566 | 2.138282 |
| **O** | -2.412678 | 3.185142 | -1.017471 | O | 5.398438 | -2.284876 | -0.07053 | O | 1.049911 | 4.444039 | 0.365442 |
| **H** | -1.989472 | 3.566137 | -1.798578 | H | 5.708968 | -1.923071 | -0.908646 | H | 0.747238 | 5.147481 | -0.218459 |
| **H** | -3.311792 | 3.539005 | -0.993671 | H | 5.704575 | -3.198339 | -0.037675 | H | 1.619647 | 4.878938 | 1.008637 |
| **O** | -4.205866 | -0.137191 | 1.349086 | O | 4.998782 | 2.135854 | -0.305544 | O | 4.582051 | -2.074101 | 1.838445 |
| **H** | -4.201906 | -0.250902 | 2.305227 | H | 4.698275 | 1.264374 | -0.607468 | H | 4.164746 | -2.757771 | 2.385025 |
| **H** | -5.023664 | 0.330534 | 1.150262 | H | 5.81914 | 1.991671 | 0.176688 | H | 5.546092 | -2.152619 | 1.907596 |

## Appendix B − Supplementary information for Chapter 4

### Experimental conditions

**Table B.1** The experimental information of the gas source and DMS conditions used for protonated Arg monomer, dimer, trimer, and tetramer.

| Gas Source Conditions | |
|---|---|
| Polarity | Positive |
| Ion Source gas 1 | 20 psi |
| Ion Source gas 2 | 0 |
| Curtain Gas | 20 psi |
| Ion Spray Voltage | 5500 V |
| Collision Gas | High |
| Temperature | 34 ℃ |
| DMS Conditions | |
| Experimental type | Enhanced product ion |
| Temperature | 150 ℃ |
| Offset | −3.0 V |
| Collision energy | 10 V |
| Entrance potential | 10 V |

## Precursor ion scan in Q1 showing ion-solvent clusters

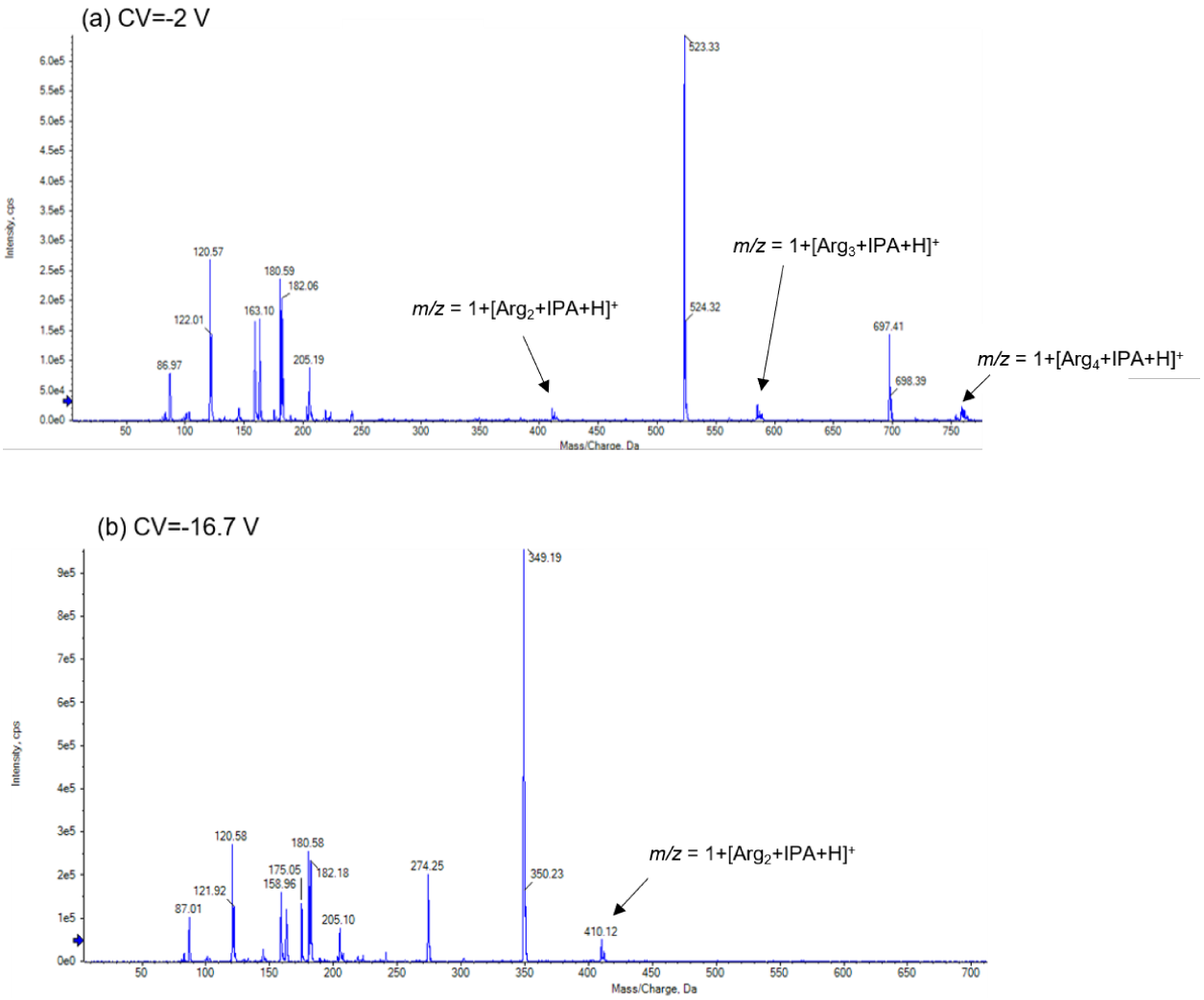


**Figure B.1** The Q1 mass spectrum at (a) CV= −2 V and (b) CV= −16.7 V as SV=3800 V when protonated Arg was electrosprayed from MeOH/$H_2O$ in IPA-modified environment.

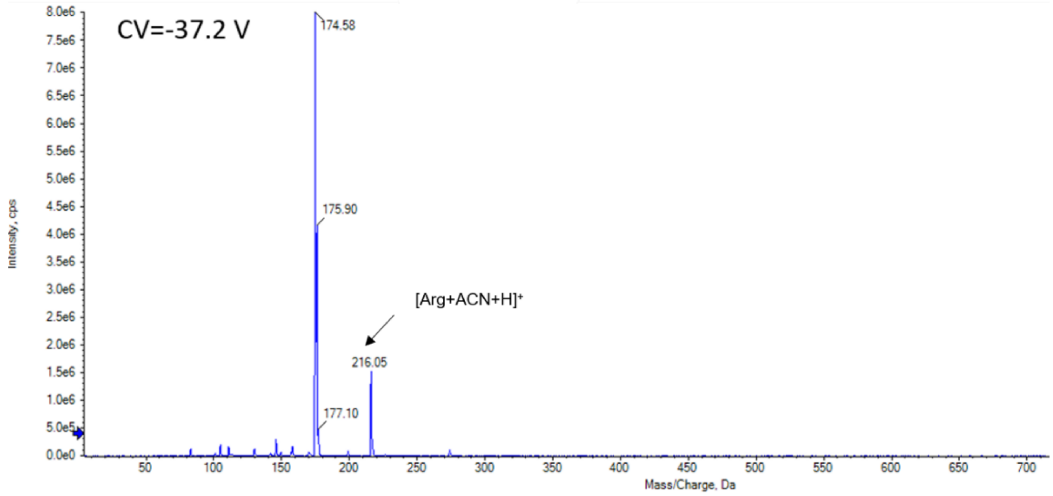


**Figure B.2** The Q1 mass spectrum at SV=3800 and CV= −37.2 V when protonated Arg was electrosprayed from MeOH/$H_2O$ in ACN-modified environment.

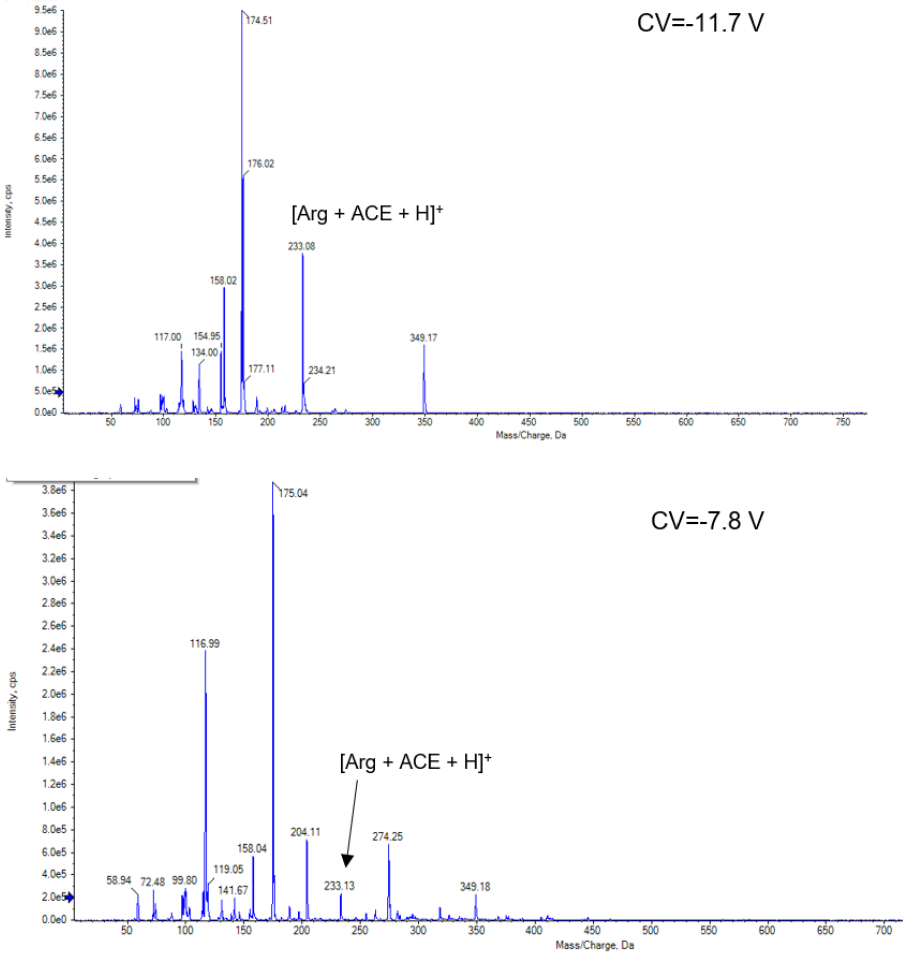


**Figure B.3** The Q1 mass spectrum at CV= −11.7 V and −7.8 V as SV=3800 V when protonated Arg was electrosprayed from MeOH/$H_2O$ in ACE-modified environment.

## Geometry coordinates of $[Arg + H]^+$

| | Isomer P1 | | | | Isomer P2 | | | | Isomer P3 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | X (Å) | Y(Å) | Z(Å) | | X (Å) | Y(Å) | Z(Å) | | X (Å) | Y(Å) | Z(Å) |
| **N** | 0.715817 | -1.270084 | -0.710523 | N | 2.832306 | -0.665585 | -0.677844 | N | -0.60646 | -0.214089 | 1.714119 |
| **H** | 1.085967 | -1.67203 | -1.575011 | H | 3.128516 | 0.119186 | -1.257395 | H | -0.753229 | -1.130334 | 2.132386 |
| **C** | 1.376053 | 0.051725 | -0.409705 | C | 2.073966 | -0.267139 | 0.580399 | C | -1.787261 | 0.203462 | 0.924589 |
| **H** | 1.474439 | 0.596576 | -1.3505 | H | 2.656792 | -0.585945 | 1.443651 | H | -2.683446 | 0.277685 | 1.546936 |
| **C** | 0.583098 | 0.888002 | 0.627535 | C | 1.816384 | 1.242132 | 0.666498 | C | -1.548381 | 1.590744 | 0.274145 |
| **C** | 2.77258 | -0.3019 | 0.11156 | C | 0.751654 | -1.085786 | 0.508956 | C | -2.051852 | -0.846899 | -0.149859 |
| **H** | 0.007381 | 0.212385 | 1.269329 | H | 1.368847 | 1.409836 | 1.648692 | H | -2.516393 | 1.950577 | -0.08192 |
| **H** | 1.312554 | 1.36723 | 1.282916 | H | 2.78465 | 1.753896 | 0.684328 | H | -1.246106 | 2.278197 | 1.072157 |
| **C** | -0.31 | 1.998164 | 0.046289 | C | 0.934046 | 1.908284 | -0.416769 | C | -0.552171 | 1.682894 | -0.90061 |
| **O** | 3.062673 | -1.411119 | 0.481818 | O | -0.001909 | -1.014108 | 1.469067 | O | -1.200465 | -1.568693 | -0.623334 |
| **O** | 3.571887 | 0.760211 | 0.1157 | O | 0.581904 | -1.669464 | -0.621298 | O | -3.326568 | -0.842345 | -0.536677 |
| **H** | -0.611309 | 2.64059 | 0.879685 | H | 1.189302 | 2.970693 | -0.413548 | H | -0.873617 | 1.038347 | -1.726791 |
| **H** | 0.287183 | 2.634206 | -0.617071 | H | 1.193994 | 1.563104 | -1.426675 | H | -0.620183 | 2.701704 | -1.292675 |
| **C** | -1.578761 | 1.564576 | -0.707737 | C | -0.597825 | 1.841546 | -0.23959 | C | 0.934611 | 1.455743 | -0.614712 |
| **H** | -2.078692 | 2.460072 | -1.084822 | H | -0.844245 | 1.88861 | 0.824054 | H | 1.537715 | 1.709651 | -1.485478 |
| **H** | -1.331061 | 0.974035 | -1.592389 | H | -1.040795 | 2.710877 | -0.733525 | H | 1.297713 | 2.050412 | 0.221693 |
| **N** | -2.538963 | 0.821172 | 0.1081 | N | -1.18888 | 0.619442 | -0.802237 | C | 2.651054 | -0.315104 | -0.02151 |
| **H** | -2.962237 | 1.340924 | 0.864312 | H | -0.631379 | 0.110039 | -1.474496 | H | -0.463895 | 0.437127 | 2.483118 |
| **C** | -2.599949 | -0.547069 | 0.256902 | C | -2.170772 | -0.102738 | -0.250292 | H | -3.45123 | -1.494619 | -1.246099 |
| **H** | 0.979545 | -1.930102 | 0.033232 | N | -2.260005 | -1.398616 | -0.562755 | N | 3.397634 | 0.643072 | 0.31059 |
| **H** | 4.438984 | 0.516483 | 0.481725 | N | -3.068021 | 0.478408 | 0.556606 | H | 4.359307 | 0.373101 | 0.503854 |
| **N** | -3.05341 | -1.235245 | 1.225129 | H | -3.037918 | -1.940266 | -0.219105 | N | 2.890377 | -1.643798 | -0.214404 |
| **H** | -3.343242 | -0.674878 | 2.02386 | H | -1.371011 | -1.88622 | -0.709356 | H | 2.143291 | -2.31585 | -0.304526 |
| **N** | -2.068203 | -1.279856 | -0.862612 | H | -3.146521 | 1.479985 | 0.621201 | H | 3.799116 | -2.015144 | 0.016572 |
| **H** | -2.359111 | -2.246754 | -0.726544 | H | -3.643036 | -0.07617 | 1.170468 | N | 1.225191 | 0.014934 | -0.273521 |
| **H** | -0.362033 | -1.230492 | -0.793887 | H | 3.643456 | -1.251542 | -0.483357 | H | 0.626469 | -0.204003 | 0.610271 |
| **H** | -2.45966 | -0.953882 | -1.744704 | H | 2.056867 | -1.250843 | -1.148684 | H | 0.837558 | -0.594431 | -0.999156 |

## Geometry coordinates of $[Arg + H]^+$

| | Isomer 10 | | |
|---|---|---|---|
| | X (Å) | Y(Å) | Z(Å) |
| **N** | 2.880473 | -0.547658 | -0.730155 |
| **H** | 3.177243 | 0.271148 | -1.260975 |
| **C** | 2.120341 | -0.232064 | 0.55202 |
| **H** | 2.736558 | -0.541332 | 1.395224 |
| **C** | 1.785917 | 1.261664 | 0.693292 |
| **C** | 0.844287 | -1.122663 | 0.472617 |
| **H** | 1.339951 | 1.370014 | 1.685018 |
| **H** | 2.729412 | 1.817363 | 0.730004 |
| **C** | 0.871585 | 1.941932 | -0.357452 |
| **O** | 0.144891 | -1.193288 | 1.470985 |
| **O** | 0.641725 | -1.604517 | -0.701863 |
| **H** | 1.090702 | 3.011421 | -0.305957 |
| **H** | 1.137552 | 1.654042 | -1.383341 |
| **C** | -0.656842 | 1.824236 | -0.187823 |
| **H** | -0.918231 | 1.895091 | 0.871991 |
| **H** | -1.119987 | 2.666682 | -0.710385 |
| **N** | -1.191708 | 0.566574 | -0.729521 |
| **H** | -0.550099 | -0.007768 | -1.266419 |
| **C** | -2.225238 | -0.121955 | -0.234712 |
| **N** | -2.294361 | -1.431064 | -0.4883 |
| **N** | -3.184239 | 0.498465 | 0.467045 |
| **H** | -3.090884 | -1.968302 | -0.18387 |
| **H** | -1.408665 | -1.921384 | -0.626733 |
| **H** | -3.268008 | 1.501721 | 0.466646 |
| **H** | -3.873076 | -0.026997 | 0.98103 |
| **H** | 3.694287 | -1.140423 | -0.566975 |
| **H** | 2.118868 | -1.109967 | -1.239377 |

## Geometry coordinates of $[(Arg)_2 + H]^+$

| | Isomer 1 | | |
|---|---|---|---|
| | X (Å) | Y(Å) | Z(Å) |
| **N** | 3.922293 | 1.745589 | 1.079142 |
| **H** | 4.710375 | 1.815288 | 1.714864 |
| **C** | 3.198523 | 0.48359 | 1.271662 |
| **H** | 2.862427 | 0.44344 | 2.310127 |
| **C** | 4.038619 | -0.774251 | 0.959742 |
| **C** | 1.936852 | 0.588417 | 0.403595 |
| **H** | 4.406512 | -0.686921 | -0.069703 |
| **H** | 4.927331 | -0.730825 | 1.595362 |
| **C** | 3.313316 | -2.127833 | 1.171295 |
| **O** | 0.811817 | 0.581859 | 0.949447 |
| **O** | 2.11201 | 0.707285 | -0.860468 |
| **H** | 3.893643 | -2.765473 | 1.843768 |
| **H** | 2.344299 | -1.968081 | 1.656748 |
| **C** | 3.091461 | -2.914519 | -0.127486 |
| **H** | 4.042824 | -3.097277 | -0.632232 |
| **H** | 2.661115 | -3.89413 | 0.084789 |
| **N** | 2.231744 | -2.190485 | -1.074066 |
| **H** | 2.584214 | -1.279732 | -1.360734 |
| **C** | 0.8979 | -2.282047 | -1.139016 |
| **N** | 0.208547 | -1.261466 | -1.659032 |
| **N** | 0.245561 | -3.379017 | -0.725185 |
| **H** | 0.647587 | -0.340066 | -1.581788 |
| **H** | -0.800092 | -1.332454 | -1.655827 |
| **H** | 0.731493 | -4.246406 | -0.576189 |
| **H** | -0.765933 | -3.383476 | -0.763719 |
| **H** | 4.277593 | 1.799588 | 0.12904 |
| **H** | 1.406882 | 2.227109 | -1.384767 |
| **N** | -4.207879 | -0.781299 | 1.890651 |
| **H** | -5.100649 | -0.558778 | 1.462787 |
| **C** | -3.151864 | -0.893195 | 0.902684 |
| **H** | -2.240363 | -1.238101 | 1.401366 |
| **C** | -2.841583 | 0.482732 | 0.271607 |
| **C** | -3.387864 | -1.882188 | -0.24163 |
| **H** | -2.041898 | 0.356382 | -0.462071 |
| **H** | -3.726795 | 0.835135 | -0.272706 |
| **C** | -2.426448 | 1.520119 | 1.319743 |
| **O** | -2.500066 | -2.347536 | -0.93763 |
| **O** | -4.674603 | -2.170539 | -0.445462 |
| **H** | -3.262668 | 1.725652 | 1.988894 |
| **H** | -1.617127 | 1.11814 | 1.93685 |
| **C** | -1.963002 | 2.833799 | 0.687471 |
| **H** | -2.728462 | 3.207062 | -0.002809 |
| **H** | -1.813844 | 3.585126 | 1.471856 |
| **N** | -0.710731 | 2.617383 | -0.043628 |
| **H** | -0.137811 | 1.801022 | 0.262601 |
| **C** | -0.173159 | 3.464354 | -0.908931 |
| **N** | 0.937593 | 3.129311 | -1.577614 |
| **N** | -0.767044 | 4.650981 | -1.165834 |
| **H** | 1.478667 | 3.847139 | -2.032018 |
| **H** | -0.397874 | 5.269432 | -1.868742 |

| | | | |
|---|---|---|---|
| **H** | -1.479202 | 5.016808 | -0.557149 |
| **H** | -4.325351 | -1.63475 | 2.424361 |
| **H** | -4.747239 | -2.770257 | -1.205506 |

## Geometry coordinates of $[(Arg)_3 + H]^+$

| | | **Isomer 1** | |
|---|---|---|---|
| | X (Å) | Y(Å) | Z(Å) |
| **N** | 8.241839 | -0.203986 | -1.164953 |
| **H** | 8.88154 | 0.539867 | -1.42373 |
| **C** | 7.7825 | -0.049093 | 0.221861 |
| **H** | 8.612417 | -0.013832 | 0.947226 |
| **C** | 6.850226 | -1.199638 | 0.61964 |
| **C** | 7.096399 | 1.309114 | 0.335711 |
| **H** | 7.441893 | -2.119273 | 0.5935 |
| **H** | 6.570322 | -1.056853 | 1.668223 |
| **C** | 5.588667 | -1.341821 | -0.262012 |
| **O** | 5.950424 | 1.508646 | 0.675633 |
| **O** | 7.939017 | 2.305034 | 0.018269 |
| **H** | 5.579057 | -0.561004 | -1.0244 |
| **H** | 5.600747 | -2.287609 | -0.809154 |
| **C** | 4.299776 | -1.260728 | 0.568098 |
| **H** | 4.181703 | -2.153759 | 1.18532 |
| **H** | 4.348666 | -0.406835 | 1.248107 |
| **N** | 3.095373 | -1.152169 | -0.252209 |
| **H** | 2.557237 | -1.993891 | -0.421991 |
| **C** | 2.631635 | -0.007293 | -0.76901 |
| **N** | 3.251308 | 1.156727 | -0.587509 |
| **N** | 1.496246 | -0.038381 | -1.49547 |
| **H** | 4.144428 | 1.216825 | -0.113509 |
| **H** | 2.705389 | 2.024853 | -0.780589 |
| **H** | 0.79663 | -0.773027 | -1.32878 |
| **H** | 1.151534 | 0.860145 | -1.807064 |
| **H** | 8.735754 | -1.084107 | -1.272266 |
| **H** | 7.466753 | 3.149286 | 0.090318 |
| **N** | -0.666072 | -2.091185 | -1.009052 |
| **H** | -0.307191 | -2.844768 | -1.59076 |
| **C** | -0.953025 | -2.60742 | 0.33599 |
| **H** | -1.728603 | -3.386157 | 0.316677 |
| **C** | -1.463878 | -1.499623 | 1.283873 |
| **C** | 0.287819 | -3.293118 | 0.881784 |
| **H** | -2.440503 | -1.187834 | 0.904424 |
| **H** | -1.640882 | -1.958926 | 2.257516 |
| **C** | -0.531469 | -0.277285 | 1.419443 |
| **O** | 1.348156 | -3.394897 | 0.302939 |
| **O** | 0.072298 | -3.820781 | 2.097157 |
| **H** | -0.502808 | 0.058227 | 2.459898 |
| **H** | 0.494939 | -0.560269 | 1.158569 |
| **C** | -0.981626 | 0.886557 | 0.535295 |
| **H** | -1.164929 | 0.524262 | -0.482773 |
| **H** | -1.930093 | 1.277849 | 0.908035 |
| **N** | 0.021041 | 1.957221 | 0.519572 |

| | | | |
|---|---|---|---|
| **H** | 0.942106 | 1.678383 | 0.825013 |
| **C** | 0.097408 | 2.856809 | -0.52888 |
| **N** | 1.281631 | 3.141715 | -1.004895 |
| **N** | -1.087159 | 3.323066 | -0.997928 |
| **H** | 1.257519 | 3.917323 | -1.660829 |
| **H** | -1.985506 | 3.217938 | -0.504764 |
| **H** | -1.069738 | 3.962364 | -1.775146 |
| **H** | -1.554036 | -1.810845 | -1.418555 |
| **H** | 0.87926 | -4.276452 | 2.382626 |
| **N** | -4.517006 | 0.274914 | -1.795117 |
| **H** | -3.774827 | -0.084842 | -2.387046 |
| **C** | -5.573888 | -0.723776 | -1.620964 |
| **H** | -6.107075 | -0.937572 | -2.560528 |
| **C** | -6.644624 | -0.259832 | -0.599809 |
| **C** | -4.948625 | -2.05104 | -1.230711 |
| **H** | -6.880703 | 0.784083 | -0.828905 |
| **H** | -7.552106 | -0.836698 | -0.790521 |
| **C** | -6.295458 | -0.396711 | 0.889107 |
| **O** | -3.762573 | -2.260817 | -1.09777 |
| **O** | -5.881191 | -3.003965 | -1.06607 |
| **H** | -7.203666 | -0.189421 | 1.466206 |
| **H** | -6.039049 | -1.4375 | 1.116745 |
| **C** | -5.167403 | 0.498241 | 1.413599 |
| **H** | -4.934053 | 0.20515 | 2.445261 |
| **H** | -4.266397 | 0.342676 | 0.824393 |
| **N** | -5.498613 | 1.918661 | 1.370533 |
| **H** | -6.239594 | 2.225794 | 1.98321 |
| **C** | -4.65852 | 2.898542 | 0.879144 |
| **N** | -3.74639 | 2.730298 | -0.017076 |
| **N** | -4.913233 | 4.160486 | 1.386607 |
| **H** | -3.817805 | 1.806209 | -0.441295 |
| **H** | -5.115007 | 4.240218 | 2.372696 |
| **H** | -4.294003 | 4.87433 | 1.032471 |
| **H** | -4.899528 | 1.095047 | -2.256172 |
| **H** | -5.43707 | -3.837988 | -0.84939 |

## Geometry coordinates of $[(Arg)_4 + H]^+$

| | | **Isomer 1** | |
|---|---|---|---|
| | X (Å) | Y(Å) | Z(Å) |
| **N** | 0.386391 | -1.629509 | 1.91179 |
| **H** | 0.770684 | -1.719926 | 0.952843 |
| **C** | 1.246251 | -0.55959 | 2.531447 |
| **H** | 1.482991 | -0.8389 | 3.557725 |
| **C** | 0.545714 | 0.798642 | 2.463543 |
| **C** | 2.539617 | -0.509204 | 1.673756 |
| **H** | 1.274388 | 1.55295 | 2.775096 |
| **H** | 0.316108 | 0.997027 | 1.412682 |
| **C** | -0.723201 | 0.909988 | 3.317314 |
| **O** | 3.588243 | -0.130504 | 2.225649 |
| **O** | 2.37753 | -0.863682 | 0.472261 |
| **H** | -1.369027 | 0.043374 | 3.162204 |

| | | | |
|---|---|---|---|
| **H** | -0.459193 | 0.92859 | 4.379903 |
| **C** | -1.523439 | 2.178511 | 2.995349 |
| **H** | -2.418583 | 2.237437 | 3.617233 |
| **H** | -0.921915 | 3.065231 | 3.214758 |
| **N** | -1.920316 | 2.281248 | 1.590836 |
| **H** | -1.251241 | 2.656191 | 0.91062 |
| **C** | -3.104629 | 1.943273 | 1.082088 |
| **N** | -4.01916 | 1.274416 | 1.830082 |
| **N** | -3.419395 | 2.300287 | -0.162409 |
| **H** | -3.644774 | 0.489658 | 2.350753 |
| **H** | -4.88031 | 1.062592 | 1.327848 |
| **H** | -2.746462 | 2.818877 | -0.711411 |
| **H** | -4.324916 | 1.993928 | -0.542882 |
| **H** | 0.490563 | -2.533931 | 2.368126 |
| **H** | 2.969956 | 1.797884 | -1.308894 |
| **N** | -4.222108 | -4.284143 | 1.180722 |
| **H** | -3.592504 | -4.852691 | 0.62022 |
| **C** | -3.997366 | -2.873194 | 0.877465 |
| **H** | -4.535182 | -2.271907 | 1.616338 |
| **C** | -4.572766 | -2.5585 | -0.518483 |
| **C** | -2.522846 | -2.442393 | 0.9213 |
| **H** | -4.186556 | -3.298707 | -1.227835 |
| **H** | -5.65275 | -2.718029 | -0.450603 |
| **C** | -4.25352 | -1.143082 | -1.011976 |
| **O** | -1.662778 | -3.186546 | 0.398096 |
| **O** | -2.246115 | -1.307571 | 1.455442 |
| **H** | -4.277481 | -0.448279 | -0.174105 |
| **H** | -3.235633 | -1.101631 | -1.410923 |
| **C** | -5.200804 | -0.623695 | -2.096174 |
| **H** | -4.85061 | 0.342975 | -2.465357 |
| **H** | -5.22081 | -1.306762 | -2.949486 |
| **N** | -6.583208 | -0.45656 | -1.654841 |
| **H** | -7.194549 | -1.253988 | -1.729173 |
| **C** | -7.003447 | 0.569412 | -0.864061 |
| **N** | -6.16638 | 1.44947 | -0.389318 |
| **N** | -8.360841 | 0.556 | -0.601233 |
| **H** | -6.667123 | 2.229788 | 0.02924 |
| **H** | -8.979193 | 0.200237 | -1.315728 |
| **H** | -8.735232 | 1.336266 | -0.084644 |
| **H** | -4.011598 | -4.484121 | 2.153403 |
| **H** | -0.656569 | -1.425481 | 1.822055 |
| **N** | 5.050905 | -4.561803 | -1.411445 |
| **H** | 5.282734 | -3.788412 | -2.027683 |
| **C** | 4.263626 | -4.089994 | -0.279068 |
| **H** | 4.341345 | -4.820411 | 0.529168 |
| **C** | 2.760293 | -3.977626 | -0.653917 |
| **C** | 4.731937 | -2.747823 | 0.279402 |
| **H** | 2.398105 | -4.997344 | -0.81558 |
| **H** | 2.214448 | -3.558444 | 0.195584 |
| **C** | 2.49665 | -3.144852 | -1.911216 |
| **O** | 5.230027 | -1.875623 | -0.410656 |
| **O** | 4.503924 | -2.627115 | 1.584089 |
| **H** | 2.894132 | -2.136751 | -1.769005 |
| **H** | 3.034271 | -3.595476 | -2.750629 |

| | | | |
|---|---|---|---|
| **C** | 1.011965 | -3.078731 | -2.308538 |
| **H** | 0.915683 | -2.768317 | -3.354531 |
| **H** | 0.564837 | -4.072727 | -2.245162 |
| **N** | 0.181806 | -2.200725 | -1.477751 |
| **H** | -0.491616 | -2.628479 | -0.829368 |
| **C** | 0.147599 | -0.869089 | -1.603066 |
| **N** | 0.920031 | -0.227493 | -2.490561 |
| **N** | -0.69818 | -0.143026 | -0.861328 |
| **H** | 1.669438 | -0.72184 | -2.94284 |
| **H** | 1.005847 | 0.797822 | -2.415172 |
| **H** | -1.290863 | -0.574682 | -0.159844 |
| **H** | 5.929433 | -4.959244 | -1.097463 |
| **H** | 4.53897 | -1.678061 | 1.855575 |
| **N** | 1.625938 | 5.206877 | -2.170403 |
| **H** | 1.412485 | 4.769111 | -3.059638 |
| **C** | 0.782972 | 4.669748 | -1.096352 |
| **H** | -0.201972 | 5.153013 | -1.050357 |
| **C** | 1.466599 | 4.885244 | 0.278536 |
| **C** | 0.507967 | 3.172918 | -1.302485 |
| **H** | 1.795549 | 5.929335 | 0.322307 |
| **H** | 0.712122 | 4.76632 | 1.063203 |
| **C** | 2.643966 | 3.948784 | 0.599606 |
| **O** | -0.506638 | 2.693249 | -0.708709 |
| **O** | 1.301825 | 2.501179 | -2.003978 |
| **H** | 3.027207 | 4.187502 | 1.597306 |
| **H** | 2.280736 | 2.917858 | 0.655446 |
| **C** | 3.805015 | 4.015535 | -0.405027 |
| **H** | 3.445758 | 3.880561 | -1.423082 |
| **H** | 4.272712 | 5.000499 | -0.382607 |
| **N** | 4.861794 | 3.035812 | -0.106199 |
| **H** | 5.758953 | 3.377437 | 0.200417 |
| **C** | 4.745434 | 1.711648 | -0.291433 |
| **N** | 3.692682 | 1.193456 | -0.923512 |
| **N** | 5.730101 | 0.899076 | 0.123003 |
| **H** | 3.383096 | 0.259884 | -0.633791 |
| **H** | 6.285808 | 1.183051 | 0.914249 |
| **H** | 5.608373 | -0.108635 | -0.00813 |
| **H** | 1.493574 | 6.207523 | -2.265718 |
| **H** | -0.643739 | 0.883893 | -0.880313 |