Exploring the Structures and Properties of Cationic Aromatic Clusters

by

Weiqiang Fu

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Chemistry

Waterloo, Ontario, Canada, 2020

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner   Dr. Travis D. Fridgen
                                     Professor
                                     Memorial University, St. John's, Newfoundland
                                     and Labrador, Canada

Supervisor(s)   Dr. W. Scott Hopkins
                                     Associate Professor
                                     University of Waterloo, Waterloo, ON, Canada

Internal Member   Dr. Terry B. McMahon
                                     University Professor
                                     University of Waterloo, Waterloo, ON, Canada

Internal-external Member   Dr. Roger G. Melko
                                     Professor
                                     University of Waterloo, Waterloo, ON, Canada

Other Member(s)   Dr. Larry Campbell
                                     Adjunct Associate Professor
                                     University of Waterloo, Waterloo, ON, Canada

                                     Dr. Germán Sciaini
                                     Associate Professor & Canada Research Chair in
                                     Atomically Resolved Dynamics and Ultrafast
                                     High-Resolution Imaging
                                     University of Waterloo, Waterloo, ON, Canada

Author's Declaration

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

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

Statement of Contributions

Weiqiang Fu was the sole author for Chapters 1, 2, 5 and 6 which were written under the supervision of Dr. Scott Hopkins and were not written for publication.

Research presented in Chapter 3:

For [Phe/Ser+H]$^+$, this research was conducted at the University of Waterloo by Weiqiang Fu under the supervision of Dr. Scott Hopkins, which has been published in peer-reviewed journals.

Citation:

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

For [(x-Phe)$_2$+H]$^+$, the theoretical calculations were conducted at the University of Waterloo by Weiqiang Fu under the supervision of Dr. Scott Hopkins. Ce Zhou wrote the python scripts to help the data analysis. Dr. Patrick Carr, Michael Lecours, Dr. Michael Burt, Dr. Rick Marta, Estelle Loire, Dr. Eric Fillion, and Dr. Terry McMahon contributed to collect the IRMPD spectra.

Research presented in Chapter 4:

This research was conducted at the University of Waterloo by Weiqiang Fu under the supervision of Dr. Scott Hopkins. Para-substituted BP species were synthesized by Christian Ieritano.

Abstract

In this thesis, the structures and properties of protonated aromatic clusters have been investigated. The aromaticity of the phenyl ring can be influenced by the substituents. Electron withdrawing groups (EWGs) can reduce the electron density of the phenyl ring, while electron donating groups (EDGs) can donate their electron density to the π systems. Therefore, the charge density distribution of molecules containing the phenyl rings can be affected by the substituents, which can cause multiple potential protonation sites and the variation of structures and properties of the aromatic compound derivatives.

A variety of aromatic clusters have been studied in this thesis in combined experimental and computational methods. A computational study of the proton-bound heterodimer of Phe/Ser has been conducted. The cosine similarities between each isomer were calculated, which was applied to partition the potential energy surface (PES) via hierarchical clustering. The cosine distances between the calculated spectra and experimental UV/IR double-resonance spectra were also calculated to assign the structures to each experimental spectrum. The same computational method of hierarchical clustering and spectral comparison was applied to study proton-bound homodimers of Phe derivatives. By changing the positions and types of the substituents on the phenyl ring, the effects of substituents on the structures and binding motifs were investigated. Infrared multiple photon dissociation (IRMPD) was utilized to produce experimental spectra for these homodimers in the gas phase. Para-substituted benzylpyridinium (BP) ions were also studied. The UV photodissociation (UVPD) action spectra of these species were measured via our modified DMS instrument. The absorption energies were found to be tuned by the substituents, and correlate with Hammett constants and electrophilic substituent constants. Last, two protomers of para-aminobenzoic acid (PABA) were examined via the UVPD spectroscopy study. PABA can be protonated at carbonyl oxygen (O-protomer) or amine nitrogen (N-protomer) under different electrospray solvent conditions. These two protomers were separated by the DMS cell and then irradiated by UV light. The two protomers showed distinctive UVPD action spectra. The simulated electronic

absorption spectra were calculated using the vertical gradient Franck-Condon approach, and then compared with the experimental UVPD action spectra.

Acknowledgements

First and foremost, I would like to express my sincere appreciation to my supervisor, Professor W. Scott Hopkins, who has guided and encouraged me to be professional and achieve my goals. Without his persistent help, the projects would not have been realized. I would like to thank Dr. Terry McMahon for his advice on my research over the years. I would like to thank Dr. Larry Campbell for helping me with DMS studies. I would like to thank Dr. Germán Sciaini for being my committee and giving me suggestions for my study. I would also like to thank Dr. Marcel Nooijen for his guidance on many of my computational work. I would also like to thank my examining committees.

I wish to express my gratitude to the current and former members of Hopkins Lab and McMahon Lab whose assistance was a milestone in the completion of this project. Dr. Neville Coughlan has provided me with a lot of advice on my DMS project and taught me a lot on OPO laser. He has also proofread my thesis and manuscript. Dr. Jeff Crouse also helped me with my project and my writing. Mike Lecours, Ce Zhou, and Josh Featherstone developed many python scripts, which saved a lot of time on job submission and data analysis. Christian Ieritano synthesized the samples so that I can pursue my research. I wish to thank all the people in the C2-261D office, Fiorella Villanueva Heldmaier, Nour Mashmoushi, Su Ji Lim, Yuting Li for their help and friendship. It is my pleasure to work with all of you for so many years.

Last, but not least, I would like to thank my friends for their support and concern. I treasure our friendship and hope that it will last forever. I would acknowledge the support and great love of my parents. They have done so much for me and support me to study in Canada for eight years from undergraduate to Ph.D. Without their support, I would never achieve my goals and complete my degrees.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ACN	Acetonitrile
AIM	Atoms in molecules
AMBER	Assisted model building with energy refinement
B3LYP	Becke 3-parameter Lee-Yang-Parr
BCP	Bond critical point
BH	Basin-hopping
BO	Born-Oppenheimer
BP	Benzylpyridium
CAM-B3LYP	Coulomb attenuating method Becke 3-parameter Lee-Yang-Parr
CCP	Cage critical point
CCS	Collision cross section
CCSD	Coupled cluster with single and double excitations
CCSD(T)	Coupled cluster with single double and perturbative triple excitations
CHelpG	Charges from electrostatic potentials with grid sampling
CID	Collision induced dissociation
CLIO	Centre Laser Infrarouge d'Orsay
CV	Compensation voltage
DLPNO	Domain-based pair natural orbital
DFT	Density functional theory
DMS	Differential mobility spectrometry
EDG	Electron donating group
ESD	Excited state dynamics
ESI	Electrospray ionization
EWG	Electron withdrawing group
FAIMS	Field asymmetric ion mobility spectrometry
FC	Franck-Condon
FEL	Free electron laser
GM	Global minimum
GPB	Gas-phase basicity
HDX	Hydrogen-deuterium exchange
HF	Hartree Fock
His	Histidine
IR	Infrared
IRMPD	Infrared multiple photon dissociation
IVR	Intramolecular vibrational redistribution
LASER	Light amplification by stimulated emission of radiation
LDI	Laser desorption ionization
Linac	Linear accelerator
LSIMS	Liquid secondary ion mass spectrometry
MALDI	Matrix assisted laser desorption ionization
MeOH	Methanol

MDS	Multi-dimensional scaling
ML	Machine learning
MM	Molecular mechanics
MO	Molecular orbital
MS	Mass Spectrometry
NCI	Non-covalent interactions
NIST	National Institute of Standards and Technology
OPO	Optical parametric oscillators
PABA	Para-aminobenzoic acid
PES	Potential energy surface
Phe	Phenylalanine
QTAIM	Quantum theory of atoms in molecules
RCP	Ring critical point
RF	Radiofrequency
Ser	Serine
SI	Supplementary information
STEOM	Similarity transformed equation of motion
SV	Separation voltage
TD-DFT	Time-dependent density functional theory
Trp	Tryptophan
Tyr	Tyrosine
UFF	Universal force field
UV	Ultraviolet
UVPD	Ultraviolet photodissociation
VG	Vertical gradient
VMI	Velocity map imaging
WPGMA	Weighted pair group method with arithmetic means

Chapter 1

Introduction

The study of clusters is an interesting and challenging field that serves as a bridge between atomic/molecular systems and macroscopic matter.[1–3] The physical and chemical properties of clusters are often very dissimilar from bulk size regimes of the same composition. In fact, the current interest in clusters largely comes from the possibility of identifying and exploiting their unique characteristics.[4,5] These differences are caused by the fact that most bulk properties operate at scales much greater than the nanoscale range of a typical cluster system.[6] Unlike isolated atoms and molecules, clusters can be held together by non-covalent interactions such as Van der Waals forces, and hydrogen bonding. Therefore, a better understanding of intermolecular and intramolecular interactions that bind the molecules together can ameliorate the performance and characteristics of clusters for specific applications and provide some insights into the properties of matter.

The study of weakly bound clusters has grown tremendously.[2,7–9] Various experimental techniques for probing the properties of clusters have been developed to date. These include laser spectroscopic techniques including infrared multiple photon dissociation (IRMPD),[10–13] UV photodissociation (UVPD),[14,15] and velocity map imaging (VMI).[16,17] Ion chromatographic technologies to separate cluster species also exist, such as differential mobility spectrometry (DMS),[18–20] and linear ion mobility spectrometry.[21,22] In a single cluster study, multiple techniques are often applied. Techniques that incorporate mass spectrometry analyses have shown great advantages. The pairing of experimental and theoretical methods is a common practice in this field, which facilitates considerable theoretical methods for computing the properties of these clusters. On the one hand, small clusters are manageable at a high level of theory. On the other hand, in the absence of calculation support, the experimental results cannot be easily explained.

In this thesis, two techniques are applied to study the structure and properties of gas-phase aromatic ionic clusters: IRMPD and DMS-MS coupled with UVPD. The experimental IRMPD

spectra for these species were collected at the Centre Laser Infrarouge d'Orsay (CLIO) free-electron laser (FEL) facility of the University of Paris XI in Orsay, France. The low-energy isomers of the clusters were determined by density functional theory (DFT). Their calculated IR spectra were compared with the IRMPD spectra for validation. The properties such as binding motifs can be predicted via computational results. DMS is a tool that can be used for spatial separation of structural isomers, stereoisomers, and tautomers. Ions exhibit different mobilities under high- and low-electric fields of an asymmetric radiofrequency waveform that are dependent on their geometry, mass and charge differences. The DMS-MS setup is modified to couple with UVPD experiments, which allow the separation of isomeric species through spectroscopic interrogations. The vertical gradient Franck-Condon (VG-FC) model was implemented to simulate the isomers' vibronic absorption spectra, for comparison with UVPD action spectra obtained experimentally.

The aromatic compounds studied herein are species that contain electrons that populate a delocalized π system. Specifically, they conform to Hückel's rule, which states that the number of delocalized π electrons equals to $4n + 2$, where n is a positive integer. Aromaticity confers additional stability to the ring system. Benzene and its derivatives are the most well-known aromatic systems. Aromatic compounds play important roles in many aspects of biological systems. All five nucleobases, adenine, cytosine, guanine, thymine, and uracil, which are the fundamental units of genetic coding in DNA and RNA, contain aromatic purine or pyrimidine rings. The amino acids, histidine (His), phenylalanine (Phe), tyrosine (Tyr), and tryptophan (Trp), which are building blocks of proteins, also contain aromatic ring moieties. These aromatic amino acids not only serve as precursors for the synthesis of many biological compounds, but also stabilize the structures of polypeptides via interactions involving their π systems (*e.g.*, cation- π, anion-π, and π-π interactions).[23–27]

The electron density of aromatic groups is strongly correlated to substituents on the aromatic ring. For example, an electron donating group (EDG) is said to donate its electron density into the conjugated π systems through resonance or inductive effects. In contrast, an electron withdrawing group (EWG) has the opposite effect and can reduce the electron density from a π system. For molecules containing π systems, the electron density of the phenyl ring exhibits

a quadrupole charge distribution. A relatively high negative charge density exists above and below the phenyl ring, with low negative charge density surrounding the ring's edge. Therefore, positively charged cations are attracted to the face of the ring rather than the edge, leading to cation-π interactions. In terms of substituent effects, EDGs can strengthen cation-π interactions, while EWGs can weaken the interactions with respect to the positions of the substituents (*e.g.*, para, meta, ortho). Substituents with electron rich donor atoms also allow hydrogen bonding interactions. Therefore, non-covalent interactions such as hydrogen bonding and cation-π interactions can be adjusted by adding the substituents on the phenyl rings.[28,29] Sequentially, the binding motifs or the integral structures of molecules can be influenced by substituent characteristics. For example, the amine group is an EDG that can donate electrons from the nitrogen lone pair, whereas the carboxylic acid can accept the π-electron density owing to the π-conjugated sp^2 orbital in the carbonyl group.[30] When these two groups are on the para positions of benzene (*i.e.*, para-aminobenzoic acid), two possible protonation sites are available and are populated under different conditions due to the resonance effect. The positions of these two substituents (*e.g.*, para-, meta-, ortho- aminobenzoic acid) can also change the molecular basicity and its preferred protonation sites.

Chapter 3 presents a study on Phe-containing dimers. Phe is one of the aromatic amino acids that contain a phenyl group. To better understand how Phe interacts with its neighbouring amino acids, the research focus was directed away from the study of intramolecular interactions to the study of intermolecular interactions between the moieties of the dimers. These non-covalent interactions can facilitate dynamic biological processes, and give rise to secondary structure in large systems such as protein.[31] The study of dimer systems of amino acids provides a pared-down model to study these non-covalent interactions.

For the proton-bound heterodimer of Phe and serine (Ser), the experimental IR spectra of this dimer were obtained via UV/IR double-resonance spectroscopy by Lorenz and Rizzo in 2012.[32] The resulting five distinct IR spectra indicate the existence of at least five isomers of Phe-Ser dimers in the probed sample. A Monte Carlo conformational search was performed, which was followed by DFT treatment. However, a thorough search of the PES and assignment of the experimental spectra for this complex system is tedious. Therefore, to efficiently

partition the PES and quantitatively compare the spectra, a computational approach was developed to investigate the structures of Phe-Ser and assign those structures to the experimental spectra. A custom-written basin-hopping (BH) search algorithm can be employed to identify low-energy isomers, which are then treated with DFT calculations. Based on the cosine distance between the mass-weighted distance matrix of each isomer, hierarchical clustering was used to partition the PES of the cluster. The calculated IR spectra and experimental spectra were compared by using the cosine distance metric to assist in assigning the spectra obtained by Lorenz and Rizzo to specific structures.

For the proton-bound homodimers of Phe derivatives, a combined computational and experimental approach was exploited. It has been shown for the Phe monomer that an intramolecular cation–π interaction between the ammonium group and the phenyl ring can affect Phe geometries and properties such as gas-phase basicity and proton affinity.[33] This cation–π interaction can be tuned by varying the substituents on the phenyl ring, which subsequently influence the structures and properties.[33] These variations in geometries and properties caused by the substituents were found to correlate with the Hammett constants. To study the effects of substituents on the Phe homodimer structures, ten different derivatives of protonated Phe homodimer were investigated. The global minima along with low-energy isomers of these homodimer derivatives were obtained by BH and then treated with DFT level of theory. The calculated spectra were compared with IRMPD spectra. Hierarchical clustering and multi-dimensional scaling (MDS) were employed to categorize all Phe derivatives homodimers and explore the influence of the substituents on the structures.

Chapter 4 describes a study of benzylpyridinium (BP) derivatives, also known as thermometer ions. Thermometer ions have a simple fragmentation pattern, in which the C-N bond between the benzyl and pyridyl moieties breaks; thus, it can be used to probe the energy uptake from the activation process.[34] The dissociation energy of the C-N bond can also be tuned by the substituents on the benzyl group.[35] Therefore, para-substituted BP ions were investigated experimentally and computationally. UV photodissociation action spectra were measured in a modified DMS instrument by monitoring the intensities of parent ions and

daughter ions. The recorded absorption energies of the substituted BP ions were found to correlate with substituent Hammett constants.

Chapter 5 discusses the study of para-aminobenzoic acid (PABA) to exemplify the spectroscopic performance and resolving power of our instrument. PABA consists of a benzene ring with para positions substituted by carboxyl and amino groups. The choice of electrospray ionization (ESI) solvents can influence the protonation sites wherein O-protomer is protonated at carboxylic acid, and N-protomer is protonated at amine nitrogen.[36,37] We expanded on a previous study that explored the effect of ESI composition on the prevalence of the different PABA protomers.[37] Herein, we extended the use of our modified DMS device compared to Chapter 4 for spectroscopic studies and characterization of the DMS-separated protomers. The PABA protomers electrosprayed from different solvent conditions were separated in DMS and then identified by UVPD action spectra. The experimental spectra were compared with previous work by Dessent et al.[38]

Overall, this thesis provides a detailed investigation of the geometric and electronic structures of aromatic ionic clusters which range from biorelevant molecules to general organic compounds. These aromatic clusters exhibit unique properties due to the phenyl group, which turns out to be manipulated by the substituted functional groups. Computationally, the study provides a novel approach to compare the calculated structures based on the cosine distance metric. Experimentally, this study applied a well-established technique, IRMPD, which was combined with the computational results to explain the observed phenomena. DMS was used to separate species of identical m/z for study with UVPD, and UVPD was used to distinguish species separated by DMS to provide insight as to how different structures are separated by their differential ion mobility. Through these findings, we hope this research can provide new insights into the aromatic clusters in the gas phase and some cognitive techniques employed in future studies.

Chapter 2

Methods

2.1 Computational Methods

2.1.1 Basin-Hopping Routine

Determining the lowest-energy geometry of a molecule or ion is a useful first step when investigating molecular phenomena. Computational approaches are a promising alternative to investigate the structures of molecular systems since it can be difficult to determine their geometries experimentally. Typically, the initial guess of the molecular structures comes from the chemists' intuition based broadly on applicable chemical phenomena, including steric hindrance, resonance effects, and H-bonding. However, it is challenging to guess the structures of complex molecules or clusters, especially with increasing molecular size, as the number of possible permutations of the atoms for those molecules scales factorially. In some cases, the energetically low-lying geometries are also required for the studies. Basin-hopping (BH) has been shown to be a helpful tool to explore the potential energy surface (PES) and find the global minimum candidates.[39,40]

 BH is a custom written algorithm[41,42] that can be interfaced with programs such as the Gaussian software package;[43] it can be regarded as a modified Monte Carlo routine with minimization, in which a large number of geometries of the molecule or cluster in question are sampled. Figure 2.1 shows a flow chart that outlines how BH samples the PES. An initial guess structure of the molecule is optimized at the density functional theory (DFT) level of theory using the Becke 3-parameter Lee-Yang-Parr (B3LYP) functional[44–47] and 6-31G basis set. Atomic coordinates of the optimized structure are extracted as an input guess geometry for the BH algorithm. The algorithm then randomly distorts the molecular geometry for a set number of iterations. Distortion of the geometry is determined by several parameters defined by the user, including simulation temperature (*i.e.*, thermal energy), dihedral angle rotation, and rotation and translation of molecules in multimeric species.[41,42] Once the molecule is randomly

distorted, the new structure is optimized using a low-level computational method (*i.e.*, molecular mechanics, *vide infra*) after which its electronic energy is calculated. The energy of the new geometry is compared with that of the previously accepted global minimum to determine its candidacy for the new global minimum. If its energy is lower than the current global minimum, the structure will be accepted as the new global minimum. This newly accepted structure will then be used as the new input structure to undergo further distortion, and the process is reiterated for a set amount of times (usually, 10,000 iterations were selected in our studies). If, however, the energy of the optimized structure is higher than the current global minimum, this energy will be used to calculate a Boltzmann probability factor, which is compared with a randomly selected number from 0 to 1.[42] The equation for calculating the Boltzmann probability is shown in Equation 2.1:

$$\frac{P_i}{P_{GM}} = e^{-\left(\frac{E_i - E_{GM}}{k_B T}\right)} \tag{2.1}$$

where P_i and P_{GM} are the probabilities of two states: the current state, *i*, and the current accepted global minimum, *GM*. If the Boltzmann probability is lower than the randomly selected number, the structure will still be accepted and used as a new input structure. Otherwise, the structure will be discarded and the BH routine will distort the previously accepted structure again. A BH routine is repeated until a user-defined N_{BH} number of structures are accepted. Therefore, with an approximate acceptance ratio of 0.5, a standard BH procedure will produce and test approximately 20,000 structures.
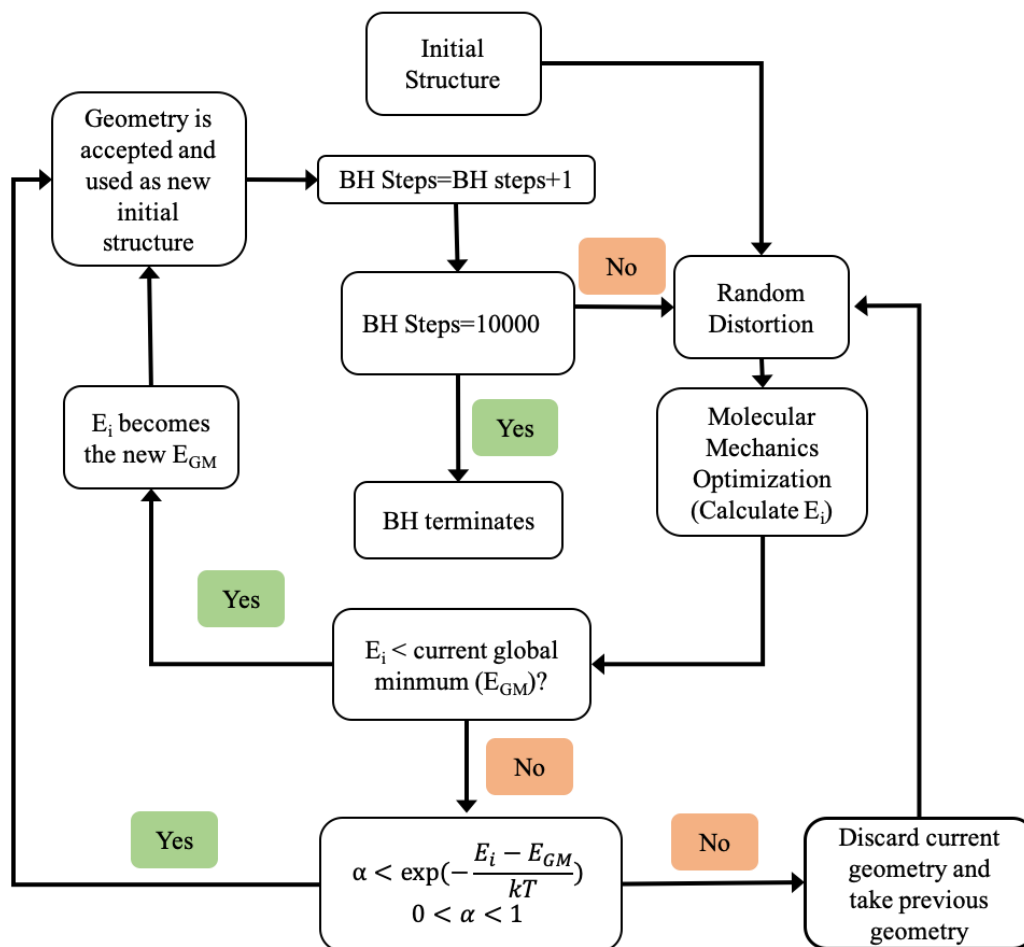
Figure 2.1 The BH routine. E_i is the energy of the distorted structure optimized by the molecular mechanics (MM) method. E_{GM} is the energy of current global minimum. The number of BH steps can be adjusted to any arbitrary value. The BH algorithm will iterate until the specified maximum number of steps is reached.

The search for geometries to map the PES includes thousands of geometry optimizations, which can be a very computationally demanding task. Molecular mechanics (MM), a relatively low level of theory, is employed for optimizing and calculating the energies of the structures produced by BH. This reduces the time required to perform such a rigorous geometry search while maintaining some level of accuracy, especially for large systems. MM employs classical physical models along with empirical potentials. It simulates the energy as a function of

conformation. The molecules are simulated as masses connected with springs. The equilibrium inter-particle distance of two given masses is set to an empirical equilibrium bond length. The total energy of the system in MM calculations is composed of covalent and noncovalent parts. Covalent interaction includes bond stretching, angular bending, and (dihedral) angular torsion.[48] MM uses a harmonic oscillator model to treat the bonds and bond angles for covalent interactions as shown in Equation 2.2.[49,50]

$$E_{covalent} = \sum_{bonds} K_b (r - r_0)^2 + \sum_{angles} K_\theta (\theta - \theta_0)^2 + \sum_{torsion} K_n [1 + \cos(n\tau - \varphi)]$$

(2.2)

In Equation 2.2, bond stretching and angle bending energies are estimated using Hooke's law. K_b and K_θ represent the force constants for each bond and angle, respectively. Furthermore, r/r_0 and θ/θ_0 are the input/equilibrium values for the bond length and bond angles, respectively. The torsion energy is periodic as a function of the dihedral angle, τ. K_n determines the amplitudes and n controls the periodicity. φ shifts the curve to synchronize the torsion potential. These parameters are unique and depend on bond types.

The energy calculated for noncovalent interactions is shown in Equation 2.3[49,50] and is given in two parts: the van der Waals forces and the Coulombic forces.[48]

$$E_{noncovalent} = \sum_{\substack{nonbonded \\ pairs\ i\ j}} \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} + \sum_{\substack{nonbonded \\ pairs\ i\ j}} \frac{q_i q_j}{4\pi \varepsilon_0 r_{ij}}$$

(2.3)

For the van der Waals forces, A and B determine the depth and position of the potential energy well for a pair of nonbonded atoms i and j, and r_{ij} is the interatomic distance. For the electrostatic potential, the Coulombic potential is employed where q_i and q_j are the partial atomic charges. ε_0 is the dielectric permittivity. These parameters are defined in the force field. The force field used here is the Universal Force Field (UFF). This model contains the force

field of all elements involved in the target molecules in the research studied here.[51] However, the choice of the force field is not especially important because BH is only used to generate candidate structures for geometry optimizations at higher levels of theory; as long as the force field yields a somewhat accurate/representative PES, the candidate structures will be close to their optimal geometries (for higher-level model chemistries).To simulate the non-covalent electrostatic interactions more precisely, partial charges for each atom were added in the MM calculations for each molecule. The partial charges were estimated by using the charges from electrostatic potentials with a grid-based method (CHelpG), which employs the least-squares method[52] and a point-selection algorithm to fit the atomic charges into the molecular electrostatic potential field.[53]

Once a BH search is completed, the unique structures are identified and low-energy isomers are carried forward for a more rigorous treatment at the semi-empirical, Hartree-Fock (HF), DFT, and/or CCSD(T) level of theory.[12,39]

2.1.2 Unsupervised Machine Learning

2.1.2.1 Geometric Cosine Similarity

Cluster systems can be compared and categorized based on their geometric similarities (*e.g.*, binding motifs). To compare the geometry of the low-lying isomers identified computationally, one can employ a similarity function to determine the similarity/difference between molecular coordinate systems. A vector of mass-weighted distances between the centre of mass, $\vec{R}_{COM}$ and each atomic position, R_i, is used to describe molecular geometries (see Equations 2.4 and 2.5):

$$\vec{R}_{COM} = \frac{\sum_{i=1}^{n} m_i \vec{r}_i}{\sum_{i=1}^{n} m_i} \ and \tag{2.4}$$

$$R_i = m_i \cdot d(\vec{r}_i, \vec{R}_{COM}) \tag{2.5}$$

where m_i and $\vec{r}_i$ are the mass and the Cartesian coordinates of the ith atom, respectively. The identity of the ith atom should be retained across all geometries being compared. Alternatively,

the mass-weighted vector can be sorted from smallest to largest R_i, which can be used as a unique identifier for each structure.

The similarity function applied for geometry comparison is cosine distance. The equation shows the cosine similarity between isomer A and isomer B as per Equation 2.6 and 2.7:

$$\text{distance} = \frac{\cos^{-1}(similarity)}{\pi} \quad and \tag{2.6}$$

$$\text{similarity} = \frac{\sum_{i=1}^{n} R_{Ai} R_{Bi}}{\sqrt{\sum_{i=1}^{n} R_{Ai}^2} \sqrt{\sum_{i=1}^{n} R_{Bi}^2}} \tag{2.7}$$

The cosine similarity ranges from 0 to 1, since the mass-weighted distances are always positive. Two identical structures will exhibit a cosine similarity equal to 1 and an angular distance equal to 0. The angular distance increases as the geometric difference between the two compared isomers increases.

Agglomerative hierarchical clustering can be used to group species with similar geometries once cosine distances have been calculated.[54] This analysis employs a tree-based representation, known as a dendrogram as shown in Figure 2.2. It builds a hierarchy of clusters to show which isomers have a closer distance on the PES with respect to the mass-weighted distance matrices. The Weighted Pair Group Method with Arithmetic Mean (WPGMA), developed by Michener and Sokal, is a common method for agglomerative hierarchical clustering,.[55,56] At each step of the WPGMA algorithm, the two closest isomers, A and B, are combined into a higher-level group, A∪B. The distance between this group and another isomer C is simply the arithmetic mean of the average distances between isomers A and C, and isomers B and C, as described in Equation 2.8.

$$d_{(A \cup B),C} = \frac{d_{A,C} + d_{B,C}}{2} \tag{2.8}$$
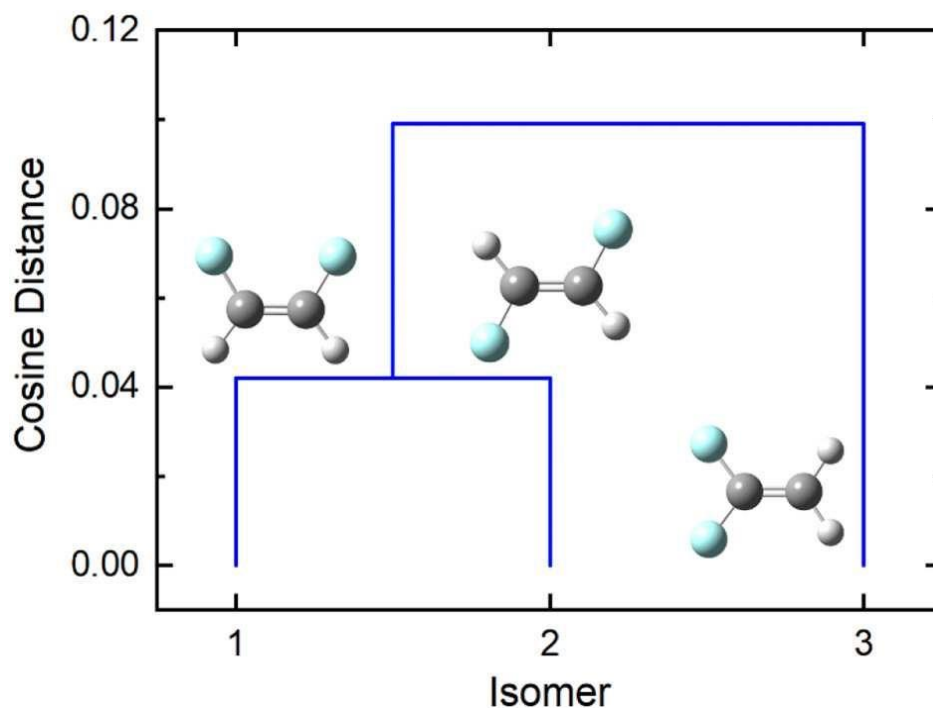
Figure 2.2 The illustration of cosine distance dendrogram of difluoroethene.[40] The structures of difluoroethene were optimized at the PM6 level of theory.

The multi-dimensional scaling (MDS) plot is a visual representation of similarities between compared objects. The similarity matrix employed here is cosine similarities. MDS assigns the objects based on the similarities into a low-dimensional space while retaining the largest possible intervals between them.[57,58]

2.1.2.2 Spectral Similarity

Similar to geometric comparisons outlined in the previous section 2.1.2.1, one can employ a similarity metric to compare the calculated IR spectra for a collection of isomers with the IR spectrum measured experimentally. In this work, experimental IR spectra were obtained via infrared multiple photon dissociation (IRMPD), as described below in Section 2.2.1. Prior to the comparison with the experiment, the calculated IR spectra must be scaled using an appropriate scaling factor to account for anharmonicity.[59] The experimental spectra were then interpolated at the same wavenumber values as the calculated spectra. The intensities of the experimental and theoretical spectra were then normalized such that the maximum intensity

for each in the given wavenumber range is equal to 1. The normalized intensity vectors could then be compared using a similarity metric such as the cosine similarity described above in Equation 2.7. Another option is to employ a scaled Euclidean similarity as described in Equation 2.9:

$$scaled\ similarity = \frac{d_i - d_{min}}{d_{max} - d_{min}} \tag{2.9}$$

where d_{max} and d_{min} are the maximum and minimum distance of the compared vectors. This treatment yields similarities that range from a minimum of 0 (worst match) and a maximum of 1 (best match).

2.1.3 Quantum Theory of Atoms in Molecules and the Non-Covalent Interactions Index

The quantum theory of atoms in molecules (QTAIM) was first introduced by Bader and has been used to extensively identify chemical bonds. The details of this theory have been explained elsewhere,[60–65] but a brief explanation pertinent to this thesis is useful. In QTAIM, the electron density of a molecule is used to understand the chemical bonding within that molecule. The topology of electron density confines the structure of a molecule and therefore correlates to many properties. The electron density can exhibit maxima, minima, and saddle points, which are critical points on the electron density surface.[61] At these critical points, the first derivative of the electron density is equal to 0, and the sign of the second derivative distinguishes the type of extremum. There are four types of critical points: (3, −3), (3, −1), (3, +1), and (3, +3). The first number in each pair, *i.e.*, 3, is the number of non-zero curvatures of electron density at the critical point. The second number is the algebraic sum of the signs of the curvatures. Nuclei exhibit local maxima of electron density. Therefore, nuclear positions exhibit (3, −3) critical points. Bond critical points (BCPs) are saddle points between the maxima of the nuclei. Thus, BCPs are (3, −1) type critical points because the electron density is at a minimum along the bonding axis and a maximum along the other perpendicular

directions. By the same logic, ring critical points (RCPs; those centred in a ring) are of type (3, +1), and cage critical points (CCPs) are of type (3, +3).

The Laplacian of the electron density ($\nabla^2 \rho$) plays a significant role in the characterization of chemical bonding. The Laplacian of electron density at critical points is given in Equation 2.10:[61]

$$\nabla^2 \rho(r_c) = \frac{\partial^2 \rho(r_c)}{\partial x^2} + \frac{\partial^2 \rho(r_c)}{\partial y^2} + \frac{\partial^2 \rho(r_c)}{\partial z^2} = \lambda_{1c} + \lambda_{2c} + \lambda_{3c} \tag{2.10}$$

where r_c is the critical point, and $\rho(r_c)$ is the electron density at the critical points. λ_{1c}, λ_{2c}, and λ_{3c} are the eigenvalues of the Hessian of $\rho(r_c)$, which are also referred to as the curvatures of $\rho(r_c)$. For covalent bonds, the electron density ρ at the BCP is usually greater than 10^{-1} a.u. with a negative value of $\nabla^2 \rho(r_c)$.[60,61] A negative value of the Laplacian at the BCPs, $\nabla^2 \rho(r_{BCP})$, indicates that the electronic charge is concentrated at the BCP. For non-covalent bonds, the Laplacian is positive and the magnitude of $\rho(r_{BCP})$ varies depending on the type of interaction. For hydrogen bonding interactions, $\rho(r_{BCP}) \approx 10^{-2}$, whereas $\rho(r_{BCP}) \approx 10^{-3}$ for van der Waals interactions.[61] For cation-π interactions (with the exception of protons), $\nabla^2 \rho(r_{BCP})$ is positive and ranges from 10^{-2} to 10^{-3} a.u.[66,67] In the case of proton-π interactions, the electron density is shared/donated because H^+ does not (formally) have electrons; therefore, $\nabla^2 \rho(r_{BCP})$ is negative and $\rho(r_{BCP}) \approx 10^{-1}$ a.u.[66] However, it is worth mentioning that Mohajeri and Karimi's study on cation-π interactions is solely concerned with free cations interacting with π systems. This situation is slightly different from the studies described herein where cation-π interactions exist in complex systems with partially positive cations and π systems constrained/tethered by covalent bonds.

Although in some cases BCPs are not identified with the QTAIM treatment, non-covalent interactions may still exist in those systems.[68,69] Occasionally, the QTAIM theory is too rigorous in its definition of the BCPs. An alternative method, the non-covalent interaction (NCI) index, has been shown to reveal non-covalent interactions in cases where QTAIM

fails.[70–72] The NCI index is a visualization index based on electron density (ρ) and reduced density gradient (s). The reduced density gradient is described in Equation 2.11:

$$s(r) = \frac{|\nabla \rho(r)|}{2(3\pi^2)^{\frac{1}{3}}\rho(r)^{\frac{4}{3}}}. \tag{2.11}$$

Non-covalent interactions occur at regions with low $\rho(r)$ and $s(r)$. The sign ($-/+$) of second eigenvalues of the Hessian of electron density (λ_2) multiplied by the electron density $\rho(r)$ can assess the strength and character of attractive or repulsive interactions. An NCI representation is given by a two-dimensional plot of $s(r)$ versus sign(λ_2)$\rho(r)$. A spike at low density and low gradient indicates the existence of non-covalent interactions as shown in Figure 2.3. If the spike reaches 0 (*i.e.* $s(r) = 0$), the spike is a critical point. The sign of λ_2 can determine whether it is a BCP (negative sign) or RCP (positive sign). This agrees with QTAIM results. Under some circumstances, the spike reaches slightly above 0, indicating no critical points are present. However, the non-covalent interactions still exist but with different strengths.[68,69] In a three-dimensional representation, the isosurfaces of the reduced density gradient can be used to visualize non-covalent interactions. The isosurfaces are determined by the value of sign(λ_2)$\rho(r)$ and are usually coloured with a blue-green-red heatmap;[70] large negative values indicate strong attractive interactions (coloured in blue) and large positive values indicate strong repulsive interactions (coloured in red).
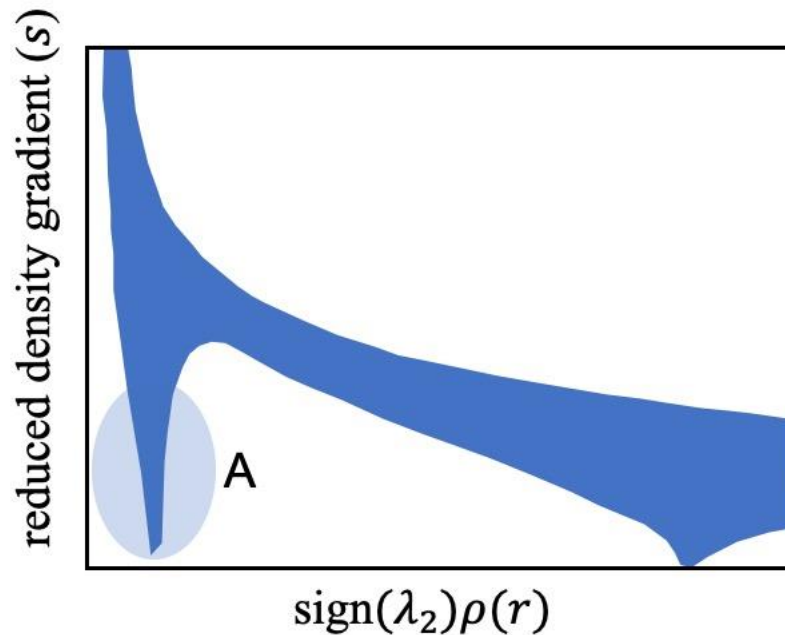
Figure 2.3 A schematic diagram of the plot of $s(r)$ versus $\text{sign}(\lambda_2)\rho(r)$. The spike is labelled as A at regions with low $\rho(r)$ and $s(r)$ in this figure.

2.1.4 Franck-Condon Principle

The Born-Oppenheimer (BO) approximation separates the electronic and nuclear parts of the molecular Hamiltonian because of the dramatic difference in timescales; nuclei are much heavier than electrons and therefore move much more slowly.[73] Using the BO approximation and applying the harmonic oscillator approximation, one can estimate the vibrational quantum structure associated with a given normal mode for a particular PES. The Franck-Condon principle assumes that the positions of the nuclei do not change during an electronic transition. This corresponds to "vertical transitions" between the populated ground state vibrational energy levels and excited vibronic states in absorption processes, as indicated by the blue arrow in Figure 2.4. To assess the probability of such a process occurring, one can calculate the overlap integral for the two corresponding vibronic wavefunctions,[73] as shown in Figure 2.4. The absolute square of the overlap integral is proportional to the intensity of a vibrational band.
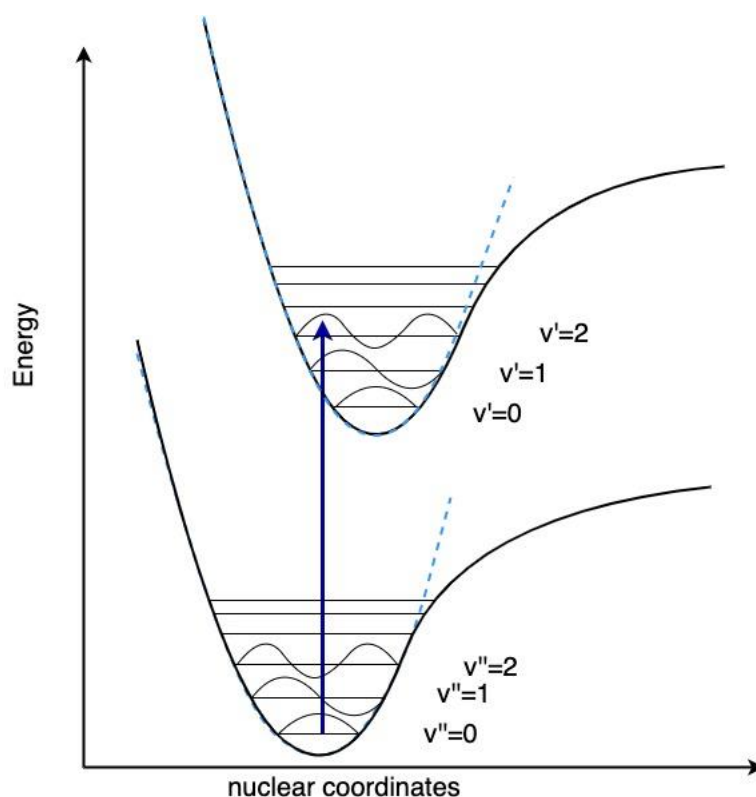
Figure 2.4 Vertical gradient Franck-Condon principle energy diagram. The excited state geometry is assumed to be that of the ground state during the vertical transition.[73]

There are two methods commonly used for determining the Franck-Condon integrals (*i.e.*, vibrational overlap integral) used in computing vibrational transition intensities: the adiabatic hessian and the VG-FC approximations. The adiabatic Franck-Condon approach typically involves the optimization of equilibrium geometries of both the ground state and the excited electronic state.[74–76] The vibrational overlap integral can subsequently be evaluated to determine the vibrational transition components that will be added to the spectrum. This provides a more accurate representation of the excited state geometry of the analyte; however, the spectrum generated using this approach is typically less reliable. Most intense transitions take place along the vertical excitation of a molecule due to the significant difference in timescale between nuclear and electronic transitions. The VG-FC approach provides a more accurate description of the transition intensities. This method evaluates the vibrational transition integrals using a Taylor expansion on the excited state PES around the initial state

17

equilibrium geometry (*viz.* that of the ground state) rather than optimizing the excited state geometry.[77,78] In this approach, the force constants of the excited state are assumed to be those of the ground state translated to match the equilibrium of the excited state through the calculation of energy gradients.[79] This approach is time-feasible and can be employed for simulating the spectra of larger molecules. While the VG-FC approach typically results in reliable spectra, this approach breaks down when a large shift in normal modes exists between the two states. A detailed explanation of these approaches can be found elsewhere.[75,77] In this study, the VG-FC approach, implemented in ORCA, was used to simulate the vibronic spectra.[80]

2.2 Experimental Methods

2.2.1 Infrared Multiple Photon Dissociation

In principle, each molecule has a characteristic vibrational spectrum. Unlike the absorption peaks in the ultraviolet (UV) and visible region, which arise from ro-vibronic transitions, absorption in the infrared (IR) region corresponds to changes in the ro-vibrational state of molecules. IR absorption is highly sensitive to molecule's functional groups, with different functional groups absorbing different frequencies of light in the IR region. Traditional IR absorption spectroscopy detects the transmittance of the light absorbed by a sample after irradiation. However, this method is not suitable for probing species in the gas phase where analyte concentration is far too low to absorb an appreciable number of incident photons to induce an observable change in the incident beam. IRMPD is an indirect measure of IR absorption that monitors the photodissociation yield of ions as a function of IR wavelength.[81] The sensitivity of mass spectrometers facilitates sensitive measurements of photodissociation yields. However, these signals are dependent on both the absorption cross section and the coupling efficiency to thermodynamic dissociative thresholds, introducing some ambiguity with respect to signal intensity. In this thesis, IRMPD was employed to measure the gas-phase vibrational spectra of proton-bound dimers of Phe derivatives. IRMPD has been described in detail in several publications.[12,33,82–87] Key aspects will be briefly discussed here.

18

IRMPD is a stepwise multiple photon absorption process that is used to determine the IR vibrational spectra of ions in the gas phase.[6] Figure 2.5 provided a schematic description of the IRMPD process. Excitation from the ground state v"=0 to v"=1 occurs when a molecule absorbs the energy of a single photon that is resonant with the vibrational energy gap. The absorbed photon energy is dissipated to bath states (*i.e.*, other low-energy states accessible from the excited v"=1 mode) through intramolecular vibrational redistribution (IVR),[88,89] allowing for subsequent single-photon absorption to occur via the same vibrational transition.[11] This single-photon absorption event repeats until the accumulated energy of the molecular ion reaches a thermodynamic dissociation threshold and the ion fragments. Figure 2.6 shows an example of the mass spectra for isolated $[(CH_2CH_3)_3NH \bullet B_{12}Cl_{12}]^-$.[90,91] Since the parent cluster does not absorb at 1613 cm^{-1}, fragmentation is negligible and only the parent ion can be observed. However, at 1461 cm^{-1} and 1033 cm^{-1}, irradiation results in different fragmentation channels. The IRMPD action spectrum can be acquired by monitoring the depletion of the parent ion and the production of fragment ions as a function of photon wavenumber. The IRMPD efficiency is calculated as per Equation 2.12:[92]

$$IRMPD_{efficiency} = -\log\left(\frac{\sum Intensity_{parent}}{\sum Intensity_{parent} + \sum Intensity_{fragment}}\right). \qquad (2.12)$$

Although IRMPD spectra are reliable in terms of vibrational frequencies, caution must be taken when interpreting spectral intensities. As stated above, intensities obtained using the IRMPD technique are heavily dependent on the photon absorption cross section, the efficiency of coupling to bath states and, ultimately, dissociative channels.[6]
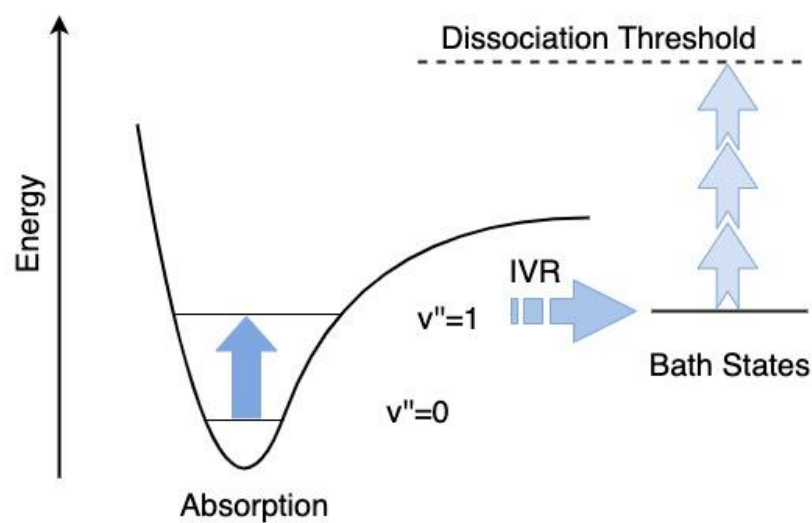
Figure 2.5 A schematic diagram of the IRMPD process. The absorbed energy is distributed to the bath states. The molecule will not dissociate until the energy exceeds a thermodynamic dissociation threshold.
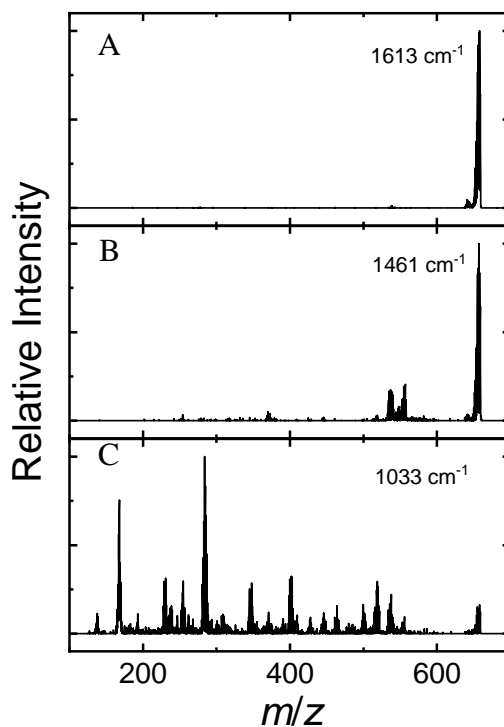
Figure 2.6 The mass spectra observed for isolated $[(CH_2CH_3)_3NH \cdot B_{12}Cl_{12}]^-$ at (A) 1613 cm^{-1} (non-resonant); (B) 1461 cm^{-1} (resonant); (C) 1033 cm^{-1} (resonant).[90,91]

To achieve IRMPD, the dissociation thresholds of the target cluster or ion must be relatively low, or the intensity of the IR source must be high enough to deliver sufficient photon energy to the analyte. Free electron lasers (FELs), which have wide tunability and high power, have become an important tool for IRMPD studies.[93,94] As Figure 2.7 shows, the main components of a FEL are a thermionic electron gun, a tunable 15–50 MeV radio-frequency (RF) accelerator, and an undulator. The electron gun can deliver electron pulses at intervals that can be regulated. The accelerator is a 3 GHz RF linear accelerator (linac), which can accelerate the electron beam to an energy 15–50 MeV. The energy of the electron beam is filtered by the energy-selecting slit before the electron beam passes through the undulator. Owing to the periodic arrangement of the magnets in the undulator shown in Figure 2.8, the electron beam oscillates during transit, resulting in the emission of synchrotron radiation. The emitted photons are captured in the optical cavity and interact with the electron beam in the undulator, leading to light amplification by stimulated emission of radiation (*i.e.*, LASER).[95–97] At the resonant wavelength λ_R, laser radiation is emitted by the FEL, as described by Equation 2.13: [95–98]

$$\lambda_R = \frac{\lambda_0}{2\gamma^2}(1 + \frac{K^2}{2}) \qquad (2.13)$$

where λ_0 is the magnetic period of the undulator, γ is the ratio between the energy of the electron beam and the energy of the light, and K is the deflection parameter, which is equal to $\frac{eB\lambda_0}{2\pi mc}$ (B is the magnetic field of the undulator, e is the electron charge, m is the electron mass, and c is the speed of light).[95] Therefore, modifying the undulator magnetic field can tune the wavelength of the emitted light. This is accomplished by changing the gap between permanent magnets.[97]
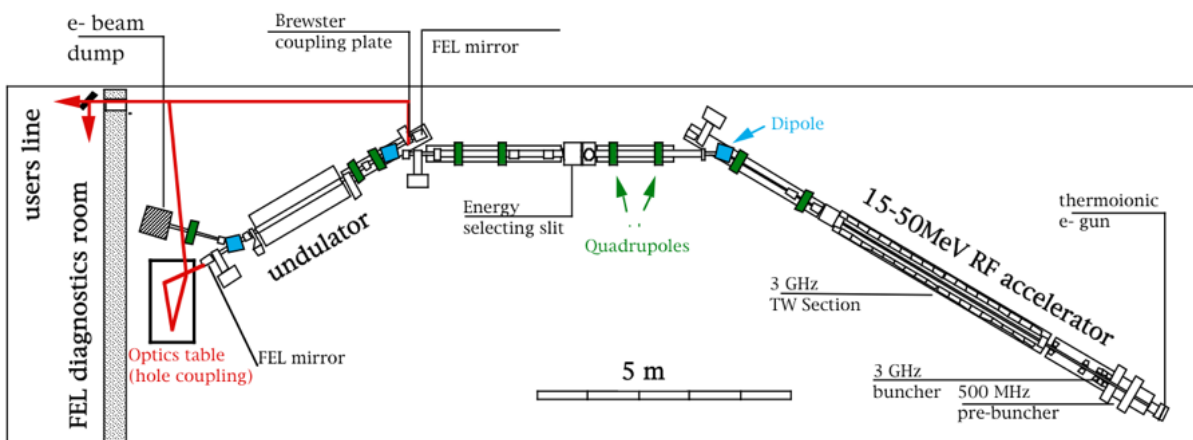
Figure 2.7 The schematic diagram of an IR FEL.[95]
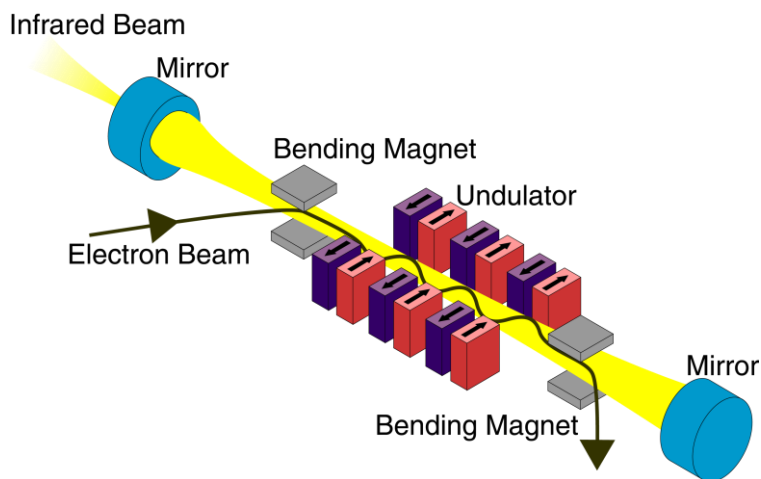
Figure 2.8 The schematic diagram of the FEL undulator[99]

To isolate clusters of interest, a quadrupole ion trap spectrometer (Bruker Esquire 3000+), as seen in Figure 2.9, is used. The analyte can be brought into the gas phase, and then ionized through electrospray ionization (ESI). The cap and annular electrodes are used to trap the ionized clusters for mass selection and interaction with the light.

22
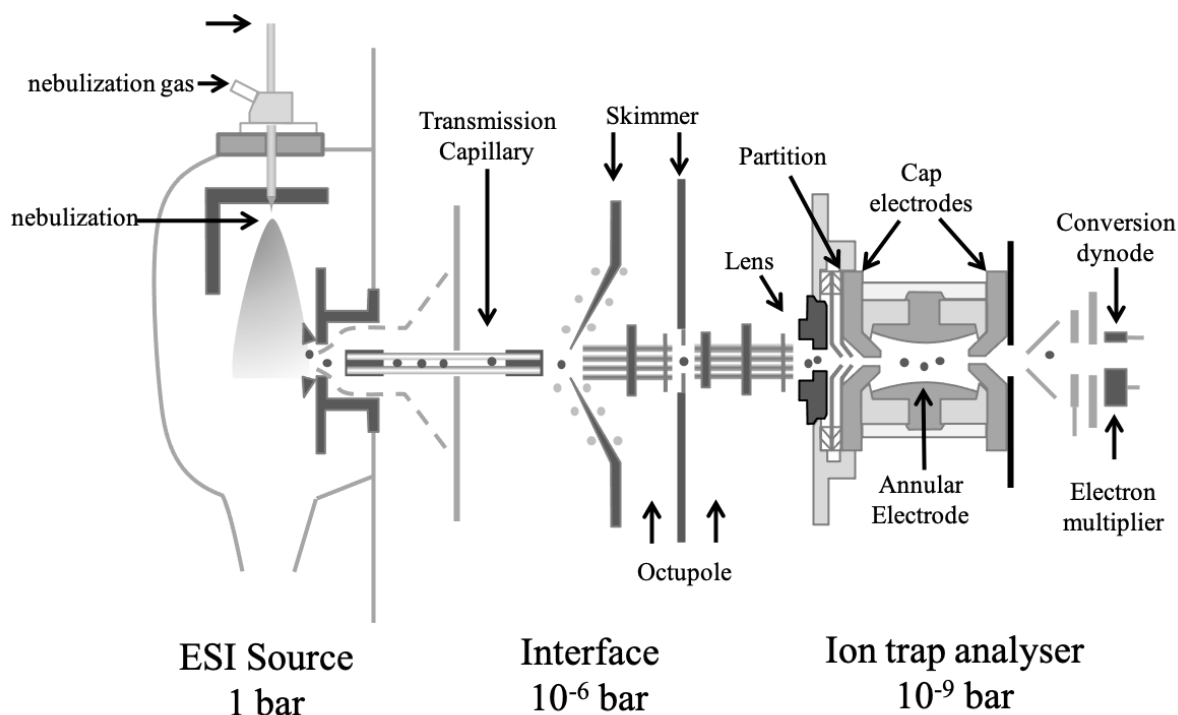
Figure 2.9 The schematic diagram of the Bruker Esquire 3000+ ion trap mass spectrometer[100]

2.2.2 Differential Mobility Spectrometry-Mass Spectrometry coupled with UV-Vis Photodissociation

Differential mobility spectrometry (DMS), is a tool used to spatially separate and characterize gas-phase ions based on their differential mobilities under high and low electric field conditions.[101] The coefficient of mobility K determines the mobility of an ion under a constant electric field and is given by Equation 2.14:[102]

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

(2.14)

where E is the strength of the applied electric field, and N is the density of the gas medium. Under low electric field conditions, K is approximately constant and independent of $(\frac{E}{N})$, and is described by $K(0)$ in Equation 2.14. As $(\frac{E}{N})$ increases, the dependence of K on the electric

23

field becomes increasingly important. The normalized function, $\alpha(\frac{E}{N})$, (also known as the alpha function), is characteristic of the analyte ions and describes their field-mobility dependence. The alpha function typically has a value less than 1.

As shown in Figure 2.10, the ions are conveyed through the DMS cell by a carrier gas (typically N_2). The DMS cell is composed of two planar electrodes positioned on opposite sides of the carrier gas flow axis. As ions pass through the cell, an RF asymmetric waveform known as the separation voltage (SV) is applied across the two electrodes to induce differential mobility of the ions. Under the influence of the SV, the ions adopt off-axis "zigzag" trajectories and migrate towards one of the electrodes, where they are neutralized. Only the ions with zero average velocity transverse to the carrier gas flow axis can pass through the gap between the electrodes. Typically, DMS is coupled with mass spectrometry (MS) to act as a detector and be used in further analysis (*e.g.,* hydrogen-deuterium exchange (HDX), collision-induced dissociation (CID), MS^N, etc.).[101,103] To correct off-axis trajectories and steer ions through the DMS cell, a direct current known as the compensation voltage (CV) can be applied along with the SV. Each molecular geometry adopts a different ion trajectory, as encoded in the SV/CV behaviour, owing to the fact that ion trajectories are determined by instrument parameters (*i.e.*, ion axial velocity, gas pressure and composition, and applied field). Critically, molecular parameters that are unique to each ion (*i.e.*, collision cross sections (CCS), mass, charge, and interaction potential with the bath gas environment) also play a role in the DMS behaviour observed.[104]
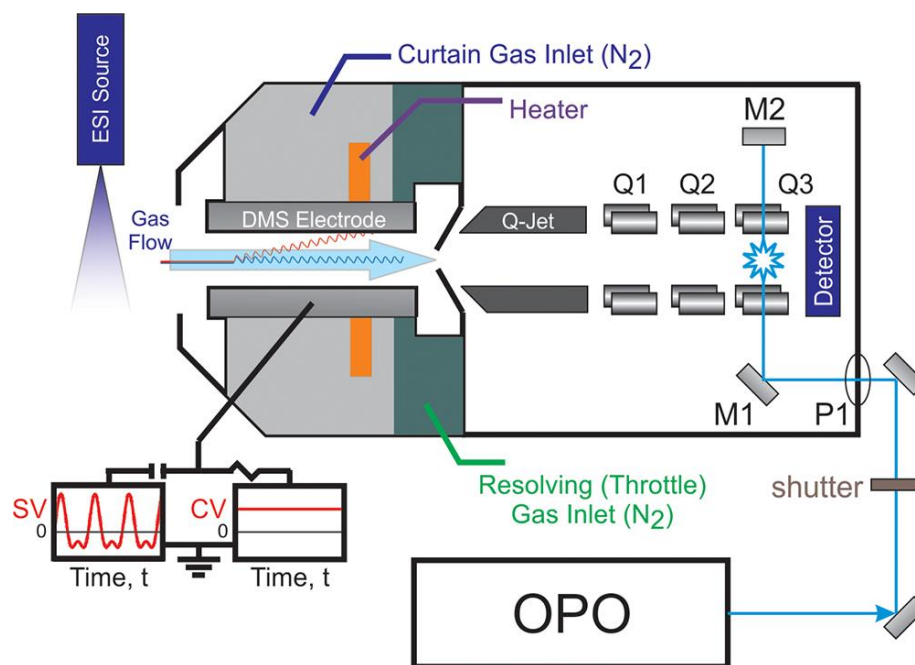
Figure 2.10 A schematic diagram of the DMS-MS coupled with UVPD. ESI: electrospray ionization. Triple quadrupole: Q1, Q2, Q3. M1: mirror 1. M2: mirror 2. P1: optical viewport. OPO: Nd:YAG-pumped optical parametric oscillator probe laser system.[105]

Doping the carrier gas with a small amount of volatile solvent vapour (*e.g.*, methanol, acetonitrile, etc.) enhances the differences in the mobility of similar ions and provides greater insight into the gas-phase properties of clusters by inducing dynamic ion-solvent clustering. Under the low-field period of the SV, ions are microsolvated – they are locally relatively cold and cluster with the solvent vapour – thereby artificially increasing their apparent CCS. This in turn reduces the ions' mobility. In contrast, under the high field portion of the SV waveform, ions are accelerated and are locally heated, inducing desolvation and decreasing their apparent CCS.[106,107] This subsequently increases the ions' mobility in comparison to that in the low field condition.[106,107] This dynamic clustering/de-clustering behaviour is unique to each geometric structure and, over the course of the *ca.* 20,000 duty cycles during DMS transit, can yield enhanced separation for geometrically similar species (*i.e.,* conformers, isomers, tautomer).[19,106,108,109]

25

At each SV, the CV can be scanned to produce a plot, referred to as an ionogram, as shown in Figure 2.11. The ionogram displays the peaks represented by ions and isomers separated by their differential mobility. A dispersion plot can be generated by graphing the CV of maximum ion transmission obtained from the ionograms as a function of SV (see Figure 2.12). Cations that cluster strongly with the solvents in the DMS cell exhibit Type A behaviour. In this scenario, CV decreases as SV increases. The ions undergo hard-sphere collisions if they do not dynamically cluster with the solvents. This is known as Type C behaviour where CV increases when SV increases. Weak-clustering ions initially exhibit Type A behaviour whereby CV decreases with increasing SV. However, at a critical extremum point, the behaviour breaks down to hard-sphere behaviour, whereby CV increases with increasing SV. This is referred to as Type B behaviour. Thus, the phenomenon of differential ion mobility provides an orthogonal means of separation prior to MS, while providing additional information into the ions' physicochemical properties.[18]
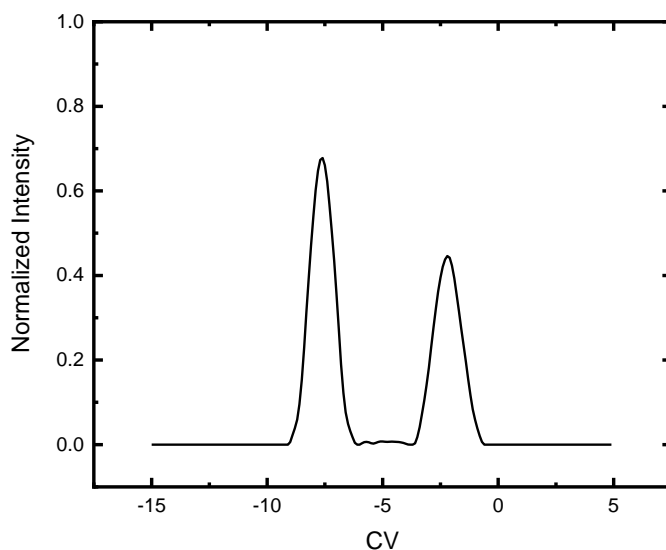
Figure 2.11 A DMS ionogram of PABA electrosprayed from ACN/H_2O 1:1 at SV = 3500 V
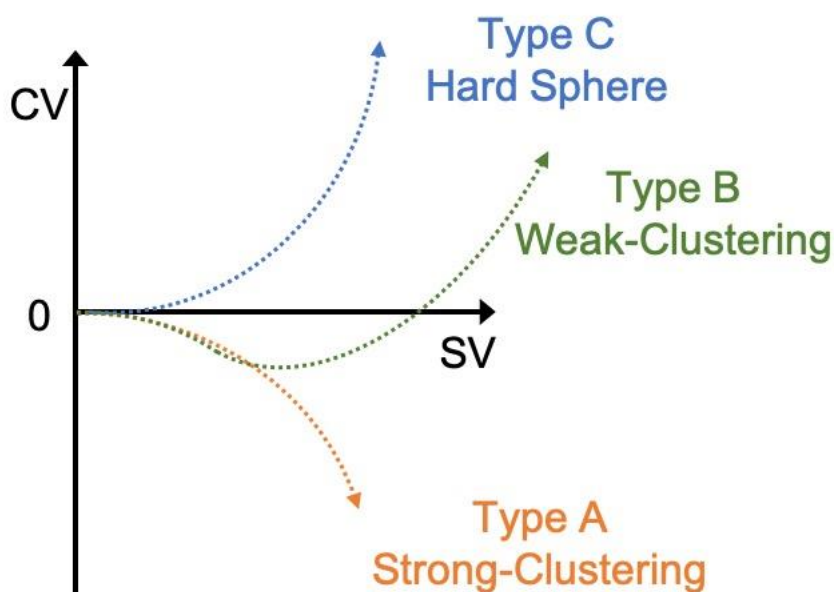
Figure 2.12 Dispersion plot of DMS behaviour: strong-clustering (Type A), weak-clustering (Type B), and non-clustering (Type C).[6]

Although a great deal of progress has been made with respect to understanding the chemistry and physics that underpins DMS, there are still challenges/ambiguities associated with the technique. In some cases, extra peaks are observed in the ionogram in addition to those that are expected. For example, protonated nucleobases presented additional ion signals that came from the same species clustering with solvents.[110] These separated isomers or tautomers cannot be distinguished by the fragmentation patterns determined via CID.[110] To overcome these limitations, we have coupled the DMS/MS instrument with an external Nd:YAG-pumped optical parametric oscillator (OPO) probe laser system such that we may perform in-trap spectroscopy on DMS-separated species. This provides an additional orthogonal characterization technique for unambiguously identifying isomeric and tautomeric species via measurement of their photodissociation action spectra and/or photofragmentation pathways.[105] This approach of photodissociation spectroscopy on DMS-separated species has been deployed in a few studies.[111–114] Adam J. Trevitt and his coworkers combined cylindrical geometry field-asymmetric waveform ion mobility spectrometry (FAIMS) with a tunable UV-Vis midband

OPO laser system,[111,112] while DMS with planar electrodes can provide higher resolving power.[115,116] Philippe Maitre, Gary Glish, *et al.* combined DMS with IRMPD.[113,114]

Unlike CID which is adjusted by collision energy, photodissociation is more easily modulated via parameters such as laser wavelength, photon flux, etc.[15,117] Electronic spectroscopy can be a reliable technique used to identify isomeric species. As stated previously, IRMPD results in vibrational excitation of the molecule in the electronic ground state via multiple photon absorption. On the other hand, UVPD is a single photon process that possesses sufficient energy to excite molecules to higher electronic states. This process allows one to access dissociative channels that are otherwise unattainable by CID and IRMPD, providing greater structural information on the target ion.[117,118] The excited ions may undergo direct fragmentation pathways if the electron is excited to a repulsive electronic state with respect to the nuclear coordinate.[119–121] The ions can also get excited into a bound excited electronic state, and then non-radiatively interact with an adjacent repulsive electronic state, a process called predissociation.[119,122] Alternatively, spontaneous radiative dissociation can occur wherein the photon energy interconverts into vibrational modes. The excited ions relax back to a lower-lying repulsive state or electronic ground state accompanied by vibrational excitation through IVR, which the latter leads to similar results as CID or IRMPD.[120,121] In principle, all of these processes can happen. For many small molecules such as OH, H_2O, the photodissociation occurs mostly through direct photodissociation.[119] For large molecules, as the size increase, this occurrence becomes less frequent.[119] Photodissociation action spectra can be recorded by monitoring the signal of the parent ion and fragment ions at each wavelength. The action spectra can be obtained by using Equation 2.15 as a function of wavelength and normalized to the laser power of each wavelength:

$$Fragment_{efficiency} = -\frac{\lambda}{p}\log(\frac{\sum Intensity_{parent}}{\sum Intensity_{parent} + \sum Intensity_{fragment}}) \qquad (2.15)$$

where λ is the wavelength and p is the laser power at each wavelength. For our experiments, a commercially available DMS-MS device has been modified to introduce UVPD through an optical viewport (P1) at the MS (see Figure 2.10). M1 and M2 were positioned on opposite

sides of the third quadrupole (Q3), such that incident light from P1 passes perpendicularly through Q3 and is retro-reflected back out of the instrument. In Q3, the parent ions are trapped and accumulated for irradiation by OPO light. Mass spectra of trapped ions are recorded at various wavelengths ranging from 208 nm to 400 nm. The laser power at each wavelength is recorded separately for normalization.

Chapter 3

The Structures of Phenylalanine-Containing Dimers

3.1 Introduction to Phe-Containing Dimers

Amino acids form the building blocks of proteins that play indispensable roles in biological processes. The non-covalent interactions within a protein can contribute to its structural integrity and have relatively low binding energies compared to covalent bonds, which can be broken more easily to facilitate dynamic biological processes. Non-covalent interactions include such as hydrogen bonds, Van der Waals, and interactions involving aromatic rings (aromatic interactions). Some diseases have been linked to protein misfolding.[123,124] Understanding how a polypeptide fold into a three-dimensional structure is necessary and significant. To study these non-covalent interactions in a protein, we need first investigate the non-covalent interactions present between amino acids.

Phe is one of the three essential amino acids containing phenyl groups in their side chains. It is a precursor for Tyr and has a relatively simple structure compared to Trp. The side chain of Phe can lead to aromatic interactions with neighboring molecules owing to the π system present in the phenyl group.[28,125] These interactions have been found to maintain the overall structures of proteins and play an indispensable role in various biological processes involving proteins.[126,127] Baker and Grant found that Phe binds selectively to the DNA chain in protein-DNA complexes, which is important in the transcription process.[128] Due to the presence of CH$\cdots\pi$ interactions, aromatic rings in the protein play an important role in carbohydrate recognition.[129] The Dougherty group also explored the cation-π interactions that contribute to the protein secondary structures and drug-receptor interactions.[125,130–134] However, there are less studies on the binding of Phe to other amino acids and identifying the binding motifs mathematically. Herein, the structures of Phe-containing dimers are investigated. The heterodimer of proton-bound Phe/Ser (Phe/Ser + H]$^+$) and the homodimer of proton-bound Phe derivatives ($[(x\text{-Phe})_2 + H]^+$) were studied. Serine has a relatively simple structure and its sidechain is often hydrogen bonded, which can provide a complex pattern of the intermolecular

interaction between the amino acids. Thus, [Phe/Ser + H]$^+$ were selected for investigation as a model system to study amino acid binding motifs by applying BH. The PES of the dimer was searched, and candidate structures were treated at the B3LYP/6-311++G(d,p) level of theory. Hierarchical clustering was used to partition the PES in terms of nuclear configurations. The calculated IR spectra are then compared with the UV/IR double-resonance spectra recorded by Lorenz and Rizzo[32] using the cosine distance metric. For [(x-Phe)$_2$ + H]$^+$, the influence of side chain derivatization on resultant binding motif was studied by varying substituent identity and positions on the phenyl ring side chain. Similar calculation methods were used to generate calculated spectra that were then compared with the experimental IRMPD spectra.

3.2 [Phe/Ser + H]$^+$

3.2.1 Introduction to [Phe/Ser + H]$^+$

Infrared spectroscopy of isolated molecular ions and ionic clusters has emerged as an important tool for determining molecular structure and properties.[88,135–139] However, as molecular ions and clusters become larger, it can be difficult to confidently assign a specific geometry as the carrier of an observed experimental spectrum. The challenge in this regard is two-fold: One must first be able to conduct an exhaustive search of the associated potential energy surface (PES) to identify likely candidate geometries, and then one must accurately calculate the vibrational spectra of the various isomers and/or conformers and determine which calculated spectrum best matches experiment. Many approaches have been suggested to solve the problem of global optimization and identification of local minima on PESs (*e.g.*, see refs 41,42,140,141). Monte Carlo-based methods, usually involving low-level model chemistry (*e.g.*, molecular mechanics), are most commonly employed in IR spectroscopic studies.[32,142–144] These searches are then followed by higher-level electronic structure calculations, often at the density functional level of theory, to predict IR absorption spectra for the various isomers/conformers identified by the search of the PES. The calculated IR spectra are then compared with the experimental spectrum (usually qualitatively by the researcher) to identify the structure(s) most likely to give rise to the experimental observations. To date, this approach has been highly successful. However, there are cases—typically associated with highly

complex potential energy landscapes—in which, despite the use of this common methodology, the spectra elude assignment. One such example is the protonated phenylalanine/serine dimer.[32]

In 2012, Lorenz and Rizzo published the results of a beautiful series of experiments wherein UV/IR double-resonance spectroscopy was used to obtain isomer-specific IR spectra of protonated 1:1 clusters of L-Phe and L-Ser.[32] In that work, the authors obtained five distinct IR depletion spectra when exciting various UV transitions in [Phe/Ser + H]$^+$, thus indicating the presence of at least five isomers in the probed ensemble (which they labeled A−E). Careful spectral analysis and ^{15}N isotopic substitution of the Phe moiety led the authors to conclude that species A−C were structures with protonation on the amino group of the Phe moiety, whereas species D and E were protonated on the Ser moiety. Lorenz and Rizzo also attempted to determine detailed structures by conducting a Monte Carlo conformational search using the AMBER force field as implemented in MacroModel, followed by DFT treatment of the candidate structures.[32] However, the PES search of this (very flexible) system proved to be a formidable task, and a definitive assignment for the observed spectra was not possible. Here, we pick up the gauntlet for the protonated phenylalanine/serine dimer. We employ a custom-written BH algorithm to search the cluster potential energy landscape and identify candidate structures for treatment at the B3LYP/6-311++G(d,p) level of theory.[39,142,145,146] The outcomes of these calculations are then treated with agglomerative hierarchical clustering to partition the cluster potential energy surface in terms of nuclear configurations. The calculated IR spectra and the UV/IR double-resonance spectra are then compared using the cosine distance metric to determine which isomers give rise to the spectra recorded by Lorenz and Rizzo.[32]

3.2.2 Methods

The BH algorithm has been described in detail elsewhere, and additional details are available in the Appendix A.[39,142] Briefly, two separate searches of the [Phe/Ser + H]$^+$ PES were conducted, one in which the site of protonation was on the Phe moiety and one with the protonation site located on the Ser moiety. The neutral and protonated amino acid moieties were first optimized individually at the B3LYP/6-311++G(d,p) level of theory to better

approximate their geometries within the cluster, and atomic partial charges were calculated using the ChelpG partition scheme.[147] The individual molecules were then combined to produce the 1:1 dimer for treatment with the BH algorithm. The cluster PES was modeled using the Universal Force Field.[51] To search the PES, the dihedral angles associated with single bonds were randomly distorted by $-5° \geq \theta \geq +5°$ at each iteration of the BH code. Simultaneously, the serine moiety was randomly rotated by $-5° \geq \phi \geq +5°$ about its centre of mass, and the serine centre of mass was randomly translated by -0.3 Å $\geq \eta \geq +0.3$ Å in the x, y, and z directions. In total, approximately 40000 geometries were sampled. Unique structures were then carried forward for geometry optimization at the B3LYP/6-311++G(d,p) level of theory, where normal-mode analyses were also conducted to predict IR spectra and to ensure that each structure was a local minimum on the PES. Unique cluster structures were identified based on zero-point-corrected energy and geometry. The optimized Cartesian coordinates for each atom in a cluster were converted to mass-weighted distances from the cluster centre of mass (*i.e.*, $m_i r_i^{com}$, where i indicates the atom index). The resulting column vector, which was then sorted in order of increasing values of $m_i r_i^{com}$, was used as a unique identifier for cluster structure. To compare the various cluster structures, cosine distances were calculated using the Orange Python package[148] according to the Equations 2.6 and 2.7.

In general, the cosine similarity ranges from $+1$ (meaning identical) to -1 (meaning exactly opposite), with a value of 0 indicating orthogonality. Thus, two identical structures will exhibit mass-weighted distance vectors with zero angular distance between them. The angular distances between mass-weighted distance vectors increase as the differences between the geometric structures of the associated isomers increase. This procedure resulted in the identification of 37 isomers for [Phe/Ser + H]$^+$ within 180 kJ·mol^{-1} of the global minimum. The Cartesian coordinates and calculated thermochemical data for these species are available in the Appendix A.

Calculation of the cosine distances between the various cluster structures facilitated analysis through agglomerative hierarchical clustering using the weighted pair group method with arithmetic mean (WPGMA), developed by Sokal and Michener.[55,56] The WPGMA algorithm can be used to prepare a dendrogram that reflects the structure present in the pairwise distance

matrix. For our purposes, this groups the [Phe/Ser + H]⁺ clusters based on structural similarity as defined by the cosine similarity of the mass-weighted distance vector, thereby providing a visual representation of the PES partitioned in terms of nuclear coordinates. At each step in the WPGMA algorithm, the nearest two clusters (*P* and *Q*) are combined into a higher-level group P ∪ Q, thereby reducing the m × m distance matrix by one column and one row. The distance between this group and another cluster R is the arithmetic mean of the distances between R and the members of P ∪ Q (as shown in Equations 2.8).

3.2.3 Results and Discussion

The dendrogram for [Phe/Ser + H]⁺ is shown in Figure 3.1. The top five groups based on structural similarity are highlighted. Note that isomers are numbered in order of increasing relative zero-point-corrected energy. The two major groupings—highlighted in green and yellow—are associated with compact cluster structures. In the green group, the Ser moiety is oriented in a bridging fashion with respect to the N and carbonyl O atoms of the Phe moiety. In general, this group is characterized by a N···H···N binding motif, however, there are examples wherein an internal rotation of the Ser moiety results in a N··· H···O or O···H···O binding motif (*e.g.*, isomers 16 and 25; *vide infra*). The structures of the isomers in the yellow group are similar to those in the green group, but in this case the Ser moiety is bound slightly out-of-plane with respect to the plane formed by the Phe N−C−C=O atoms. In contrast to the green group, the yellow group is predominantly characterized by a N···H···O binding motif, however, there are examples wherein an internal rotation of the Ser moiety results in an O··· H···O binding motif. The orange group also exhibits compact structures, but not with the Ser moiety oriented in a bridging fashion to the Phe N and carbonyl O atoms. Instead, the Ser moiety is oriented above the Phe ring (for isomers 1 and 36) or extending away from the Phe ammonium group, perpendicular to the plane of the ring (isomer 29). It is likely that isomers 1 and 36 are stabilized by partial charge/ring quadrupole interactions. The same is true for isomer 10 (the only member of the blue group), but the structure of isomer 10 is somewhat unique in comparison with the structures of the rest of the [Phe/Ser + H]⁺ isomers. In this case, the Ser moiety is oriented above the Phe ring with the plane of the Ser COOH group nearly

parallel to the ring plane. Isomer 10 is bound by a Phe—$NH_3^+\cdots O{=}C$ interaction. Finally, the isomers of the red group are all extended/elongated structures in which the COOH group of the Phe moiety binds with Ser in a bidentate fashion and the Ser moiety is oriented away from the Phe ring. All isomer structures are available in the Appendix A; those of important isomers as determined by spectroscopy are shown in Figure 3.1.
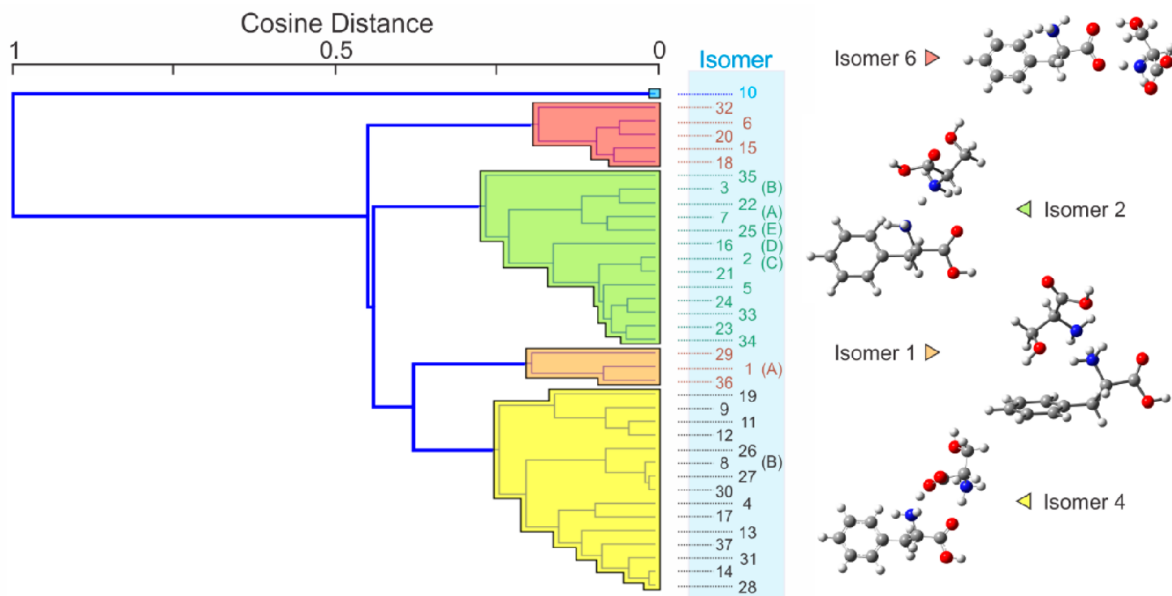
Figure 3.1 WPGMA dendrogram constructed from the cosine distances between the various cluster structures

Having mapped the PES of $[\text{Phe/Ser} + \text{H}]^+$, we turned our attention to the vibrational spectra recorded by Lorenz and Rizzo.[32] First, the experimental spectra that were plotted in their original *J. Am. Chem. Soc.* communication[32] were digitized and interpolated across the 2800−3750 cm^{-1} region to produce five *XY* data sets (A−E) that had intensity measurements at 0.5 cm^{-1} intervals. The calculated IR spectra for the 37 isomers identified by our BH search were convoluted with a Gaussian distribution of fwhm = 5 cm^{-1} and similarly interpolated. The intensities of the experimental and theoretical spectra were then individually normalized to a maximum of 1, and cosine distances were calculated for the various *Y* vectors (*i.e.*, columns of intensity values). Owing to the fact that the National Institute of Standards and Technology (NIST) recommends an anharmonic scaling factor of 0.967 ± 0.021 for harmonic frequencies

calculated at the B3LYP/6-311+ +G(d,p) level of theory, the process of calculating the cosine distance matrix was repeated several times as the scaling factor was stepped from 0.9200 to 1.0000. In this way, we were able to identify a scaling factor of ca. 0.952 for [Phe/Ser + H]$^+$ (see Appendix A for details). Using this scaling factor, the cosine distances between the experimental and calculated spectra were calculated, then renormalized across the [0,1] interval, and plotted as the heat map shown in Figure 3.2. The red-colored panels in Figure 3.2 indicate the closest distance (*i.e.*, best match) between experiment and theory, whereas the black panels indicate the greatest distance (*i.e.*, worst match).
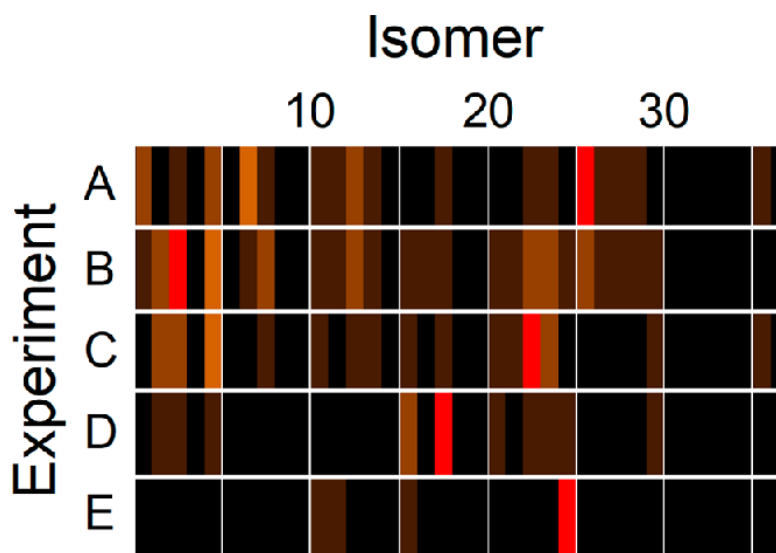
Figure 3.2 Heat map showing the relative distances between the experimental IR spectra recorded by Lorenz and Rizzo[32] and the calculated IR spectra for the 37 isomers of [Phe/Ser + H]$^+$ identified in our search. Red indicates the closest cosine distance between experiment and theory; black indicates the farthest cosine distance.

It is clear from Figure 3.2 that experimental spectrum E can be unambiguously assigned to isomer 25. A comparison of experimental spectrum E and the calculated spectrum for isomer 25 is shown in Figure 3.3. The calculated geometry of isomer 25 exhibits protonation on the amino group of the Ser moiety, which is consistent with expectations based on the ^{15}N isotopic substitution study conducted by Lorenz and Rizzo.[32] Experimental spectrum D matches well with the calculated spectra of isomers 16 and 18. The calculated spectra for both isomers are plotted along with experimental spectrum D in Figure 3.3. Although the calculated spectrum

for isomer 18 yields a slightly better cosine distance match (because of a slightly better alignment of major spectral features), we favour isomer 16 as the carrier for spectrum D because it exhibits protonation on the amino group of the Ser moiety, which is consistent with ^{15}N isotopic substitution results.[32] Note that isomer 16 and isomer 25 both belong to the green subgroup shown in Figure 3.1. Their geometries differ slightly in the orientation of the Ser moiety: Isomer 16 exhibits an intermolecular hydrogen bond between the Ser carboxylic acid OH and the Phe carbonyl O, whereas isomer 25 exhibits a hydrogen bond between the Ser carboxylic acid OH and the Phe amino group.
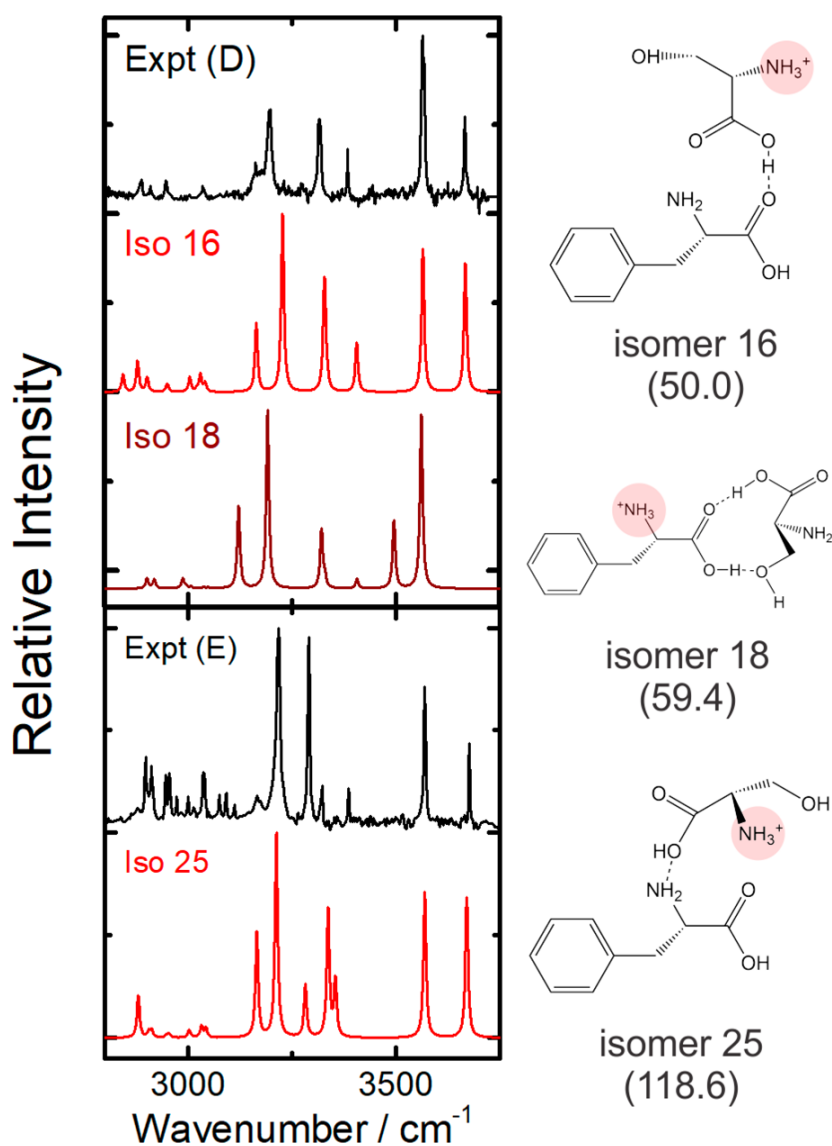
Figure 3.3 (Black traces) Experimental [Phe/Ser + H]$^+$ UV/IR double-resonance spectra D and E, adapted from ref 32. (Red traces) Calculated IR spectra for isomers 16, 18, and 25. Calculations were conducted at the B3LYP/6-311++G(d,p) level of theory. A scaling factor of 0.952 was applied to the calculated spectra. Zero-point-corrected energies (in parentheses; relative to the global-minimum structure) are reported in kJ mol^{-1}.

Figure 3.2 shows that the spectra of several isomers provide reasonable matches with experimental spectra A−C. Upon examining the five closest matches for experimental spectrum C (isomers 2, 3, 5, 23, 24), we find that all five of these isomers are associated with

the green subgroup of the dendrogram (Figure 3.1) and exhibit N···H···N binding. Consequently, these species all exhibit similar spectra. Even though isomers 23 and 5 have slightly smaller cosine distances to spectrum C, we favour the assignment of spectrum C to isomer 2 because of its lower calculated relative energy (assuming that structures within 3 kJ· mol^{-1} of the global minimum are populated). Note that the breadth of the spectral features and the peak at ca. 3605 cm^{-1}, which was identified by Lorenz and Rizzo as a contribution from spectrum A, suggests that multiple isomers might be contributing to the observed spectrum. Our analysis shows that multiple isomers are responsible for experimental spectra A and B. The five best matches for spectrum B are isomers 2, 3, 5, 8, and 24. With the exception of isomer 8, which belongs to the yellow subgroup, these isomers are all associated with the green subgroup of the dendrogram and exhibit N···H···N binding. The spectrum of isomer 3, which exhibited the best match to spectrum B, is plotted in Figure 3.4. Although the spectrum of isomer 3 is a good match in the NH and OH regions of spectrum B, there are additional features in the experimental spectrum that remain unassigned. These features are not well-predicted by isomers 2, 5, or 24, which all have structures and spectra similar to those of isomer 3. However, the spectrum for isomer 8 does capture these features and is overlaid with the spectrum of isomer 3 in Figure 3.4. In the case of experimental spectrum A, the top five matches were isomer 1 (orange group), isomers 5 and 7 (green group), and isomers 13 and 26 (yellow group). The spectrum of isomer 1, the global-minimum structure, is plotted in Figure 3.4 along with the spectrum of isomer 7. Although isomer 5 is calculated to be lower in energy than isomer 7, we discounted isomer 5 because its spectrum exhibits an additional feature to higher wavenumber of the highest observed band. A convolution of the calculated spectra of isomers 1 and 7 captures all of the features in spectrum A, except for the bands observed at 3115 and 3203 cm−1 (marked with an × in Figure 3.4). The spectrum of isomer 13 does not exhibit bands at either of these positions; however, the spectrum of isomer 26 (not plotted) does show a peak at 3200 cm^{-1}. Thus, similarly to spectrum B, spectrum A appears to be a convolution of spectra for multiple isomers. Note that all five of the calculated spectra shown in Figure 3.4 are associated with isomers that exhibit protonation at the Phe nitrogen atom or

shared between the amino groups of the two amino acids. This accords with expectations for spectra A−C based on the ^{15}N isotopic substitution study by Lorenz and Rizzo.[32]
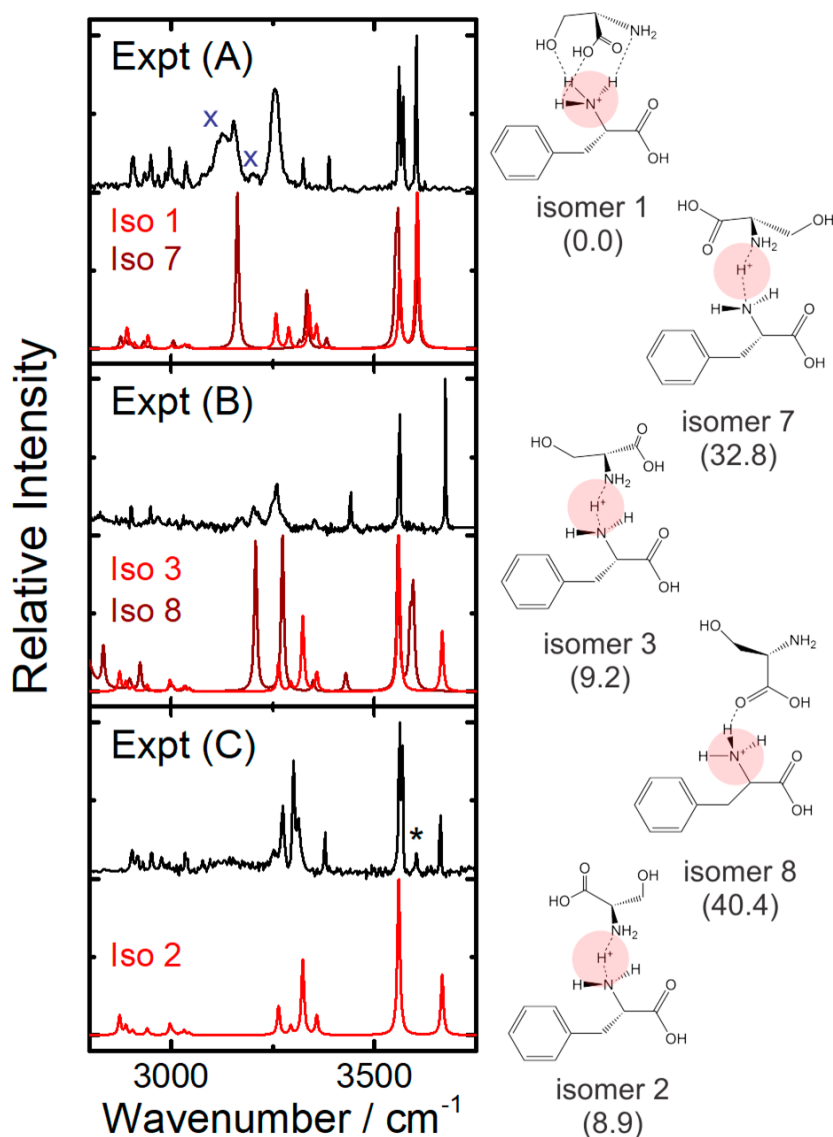
Figure 3.4 (Black traces) Experimental [Phe/Ser + H]⁺ UV/IR double-resonance spectra A−C, adapted from ref 32. (Red traces) Calculated IR spectra for isomers 1−3, 7, and 8. Calculations were conducted at the B3LYP/6-311++G(d,p) level of theory. A scaling factor of 0.952 was applied to the calculated spectra. Zero-point-corrected energies (in parentheses; relative to the global-minimum structure) are reported in kJ mol⁻¹. The band marked with an asterisk is attributed to spectrum A.[32] Bands marked with an × are attributed to higher-energy structures.

3.2.4 Conclusions for [Phe/Ser + H]⁺

By applying hierarchical clustering to the results of our BH search of the [Phe/Ser + H]⁺ PES, we are able to partition the PES in terms of nuclear configuration. Although the resulting dendrogram does not provide information on barriers to isomerization (as would a disconnectivity graph),[149,150] this exercise does provide some insight into likely regions of kinetic trapping. Indeed, when we compare our calculated spectra to the experimental IR spectra recorded by Lorenz and Rizzo,[32] we find that the spectral carriers are associated with three relatively different regions of the PES. It should be noted that the assignment of the experimental spectra was enabled by calculating cosine distances between the experimental spectra and calculated harmonic spectra. It is possible that the matching algorithm could be improved by using another distance metric or by introducing anharmonic corrections to the calculated frequencies. Such considerations might be necessary for larger, more complex molecular clusters. Nevertheless, this work does demonstrate the utility of introducing aspects of unsupervised machine learning to the analysis of vibrational spectra.

3.3 [(x-Phe)₂ + H]⁺

3.3.1 Introduction to [(x-Phe)₂ + H]⁺

Amino acids have been the subject of many studies to investigate their gas phase structures and properties such as sites of protonation and effects of non-covalent interactions.[82,93,151–156] Infrared multiple photon dissociation (IRMPD) has emerged as a powerful tool for these investigations owing to the relation between geometric structure and the vibrational spectrum.[11,12,158–162,33,82,86,87,93,142,155,157] Previous work has demonstrated that, in general, amino acid monomers are protonated on the amine nitrogen and that intramolecular hydrogen bonding occurs between the ammonium group and the carbonyl oxygen.[82,93,156] Several studies have also found that non-covalent hydrogen bonding and cation-π interactions influence the structures and properties of these systems.[24,28,33,163,164]

The literature pertaining to non-covalently bound amino acid clusters is relatively sparse. Some focus has been directed towards the serine octamer owing to its being implicated in

homochiral genesis.[165,166] Some examples of protonated amino acid heterodimers have also been reported.[12,32,159,160,167] Typically, these systems are bound by an intermolecular hydrogen bond between the carbonyl group of the neutral moiety and ammonium group of the protonated moiety. The protonated heterodimer of glycine and phenylalanine (Phe) is an example of N-H⋯N binding motif.[12] In contrast, the protonated heterodimer of glycine and pentafluorophenylalanine, F_5-Phe, exhibits the generally observed N-H⋯O intramolecular hydrogen bonding. This variation in binding motif has been shown to correlate with the difference in the gas phase basicities (GPB) of the two moieties. With closer GPBs of the two moiety, it tends to form N-H⋯O binding motif.[12]

A few studies have explored the structures of protonated amino acid homodimers.[86,154,157,158] For the homodimer of tryptophan (Trp), it was found that the global minimum structure is charge-solvated.[157] The protonated homodimers of serine (Ser), isoleucine (Ile), Phe, and tyrosine (Tyr) were also studied to investigate the effects of the side chain on their pairwise interactions.[158] It was found that the interaction between the side chain and the protonation site have an impact on the binding motif between the two amino acids.[158] The ammonium group binds with neutral amine group if the side chain interacts with the protonation site, whereas the ammonium group binds to the carbonyl group if the side chain does not (or weakly) interacts with the protonation site.[158]

To date, there have been no investigations to study how chemical derivatization of a side chain influences the binding motif of a protonated amino acid homodimer. One might hypothesize that derivatization could be used to control interactions such as π-π, cation-π, hydrogen bonding, which might then influence morphology and properties of larger systems. To explore this scenario, a series of protonated homodimers of Phe derivatives that differ by the identify and positions of substituents on the phenyl ring side chain were selected. Along with the protonated homodimer of Phe, the 3-fluoro, 4-fluoro, 2,5-difluoro, 3,5-difluoro, pentafluoro, 3-cyano, 3-trifluoromethyl, 3,4-dimethoxy, and 4-nitro derivatives were studied. These substitutions introduce potential H-bonding groups on the side chain and change the quadrupole moment of the phenyl ring, thus potentially tuning π-π and cation-π interactions.

To investigate the structure and bonding of the protonated Phe homodimer and its derivatives, a combined computational and experimental IRMPD approach was employed.

3.3.2 Experimental Methods

Experimental IRMPD spectra of homodimers of Phe derivatives were recorded at the Centre Laser Infrarouge d'Orsay (CLIO) free electron laser (FEL) facility at the University of Paris XI.[11,96] The experimental equipment has been described in detail previously.[142,151,155] 100 µmolL^{-1} electrospray ionization (ESI) solutions of methanol/water (50/50 vol%) with 0.1% formic acid (*ca*. 5 µmolL^{-1}) were prepared with stoichiometric quantities of the L-Phe derivatives (Alfa Aesar). The amino acids were used without further purification. The specific L-Phe derivatives chosen for study were Phe and the 3-fluoro-, 4-fluoro-, 2,5-difluoro-, 3,5-difluoro-, pentafluoro-, 3-cyano-, 3-trifluoromethyl-, 3,4-dimethoxy-, and 4-nitro-derivatives. Positive mode ESI was employed to produce the gas phase protonated homodimers of the L-Phe derivatives. Nascent ions were transferred to a Bruker Esquire 3000+ ion trap mass spectrometer, where they were mass-selected and subsequently irradiated by the tunable output of the FEL over the $1000 - 2000$ cm^{-1} range. The laser beam energy ranges from 32 to 60 MeV, and the pulse length is from ~2 ps.[98] Vibrational spectra were generated by recording the fragmentation efficiency of the protonated homodimers of Phe derivatives as a function of FEL wavenumber. The fragmentation efficiency is described by Equation 2.12.

3.3.3 Computational Methods

A custom-written BH algorithm coupled with Gaussian 09 was employed to map the potential energy surfaces (PESs) of the protonated homodimers of Phe and its derivatives.[39,41,42,151] Molecular geometries were optimized using molecular mechanics (MM) using Universal Force Field (UFF), which incorporated atomic centred charges as calculated at the B3LYP/6-31+G(d,p) level of theory using the CHelpG partition scheme.[51,145,147,168] For each step of the BH algorithm, each of the dihedral angles in the given Phe derivatives was given a random rotation of $-5° \leq \alpha \leq 5°$. In addition, the neutral moieties of the homodimers were given a random rotation of $-5° \leq \beta \leq 5°$ about its body-fixed *x*, *y* and *z* axes. For each cationic species, BH searches were carried out twice: once with the excess proton attached to the amine nitrogen

atom, and once with the proton attached to the carbonyl oxygen atom. Approximately 20,000 structures were sampled by the BH algorithm for each protonated structure (*i.e.*, *ca*. 40,000 for each derivative). Each unique structure identified by the BH routine was pre-optimized with the semi-empirical PM6 method prior to treatment at the B3LYP/6-311++G(d,p) level of theory.[169] Normal mode analyses were conducted to ensure that each isomer was a local minimum on the PES. This also served to calculate the harmonic vibrational spectrum for each structure, which was then used to assign the experimental IRMPD spectrum. The calculated IR spectra obtained from frequency calculations, were scaled by 0.9679 to correct for the anharmonicity.[59] The incomplete incorporation of electron correlation and the use of finite basis sets also contribute to error in the calculated IR spectrum.[59] DFT results and cluster *XYZ* coordinates are provided in the supplementary information. To compare all the isomers, geometric similarity has been calculated.[40,170] The Cartesian coordinates for each isomer can be converted to a mass-weighted distance vector. The cosine distance and similarity between each isomer is calculated using Equation 2.6 and 2.7. The larger the cosine distance, the more geometrically similar the structure is. The electron density and its Laplacian were calculated at BCPs to show non-covalent interactions, which are represented by gradient isosurfaces. These were generated using the Multiwfn program.[171]

3.3.4 Results and Discussion

3.3.4.1 Hierarchical clustering of Phe dimer

One can classify the geometric structures of the homodimers by their binding motifs as determined by unsupervised machine learning (ML). This can be accomplished by considering only the main chain of each Phe moiety (*i.e.* carbon atom 8, the amine group, and the carboxylic acid group of carbon 9 shown in Figure 3.5).
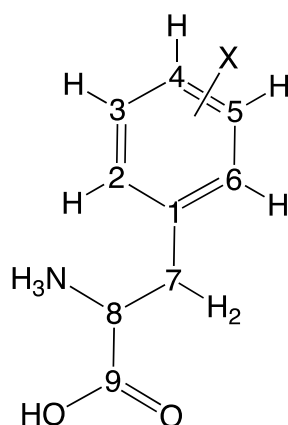
Figure 3.5 The skeleton structure of the protonated Phe derivatives investigated in this study. Carbon atom positions are numbered 1-9. X represents the different functional groups associated with the various Phe derivatives.

Previous work suggests that the phenyl rings can orient spatially via cation-π interaction with the protonation site.[33] However, the cation-π interaction is expected to be relatively weak compared to hydrogen bonding, which should be the strongest interaction associated with intermolecular bonding. Consequently, atoms associated with the side chains were excluded from hierarchical clustering analysis. A dendrogram illustrating the binding motifs of the protonated Phe homodimer, [Phe$_2$ + H]$^+$, is shown in Figure 3.6.[148] Hierarchical clustering employed the weighted pair group method with arithmetic mean.[56] The conformers/isomers are numbered in order of increasing relative Gibbs energy. Grouping illustrates the different intermolecular hydrogen bonding motifs, which may be classified into four groups: Motif 1 (blue) corresponds to NH⋯N binding, motif 2 (red) to bidentate binding, motif 3 (green) to OH⋯O binding, and motif 4 (yellow) to NH⋯O binding. Bidentate binding occurs when the ammonium group simultaneously forms intermolecular hydrogen bonds with the amino group and the carboxylic acid group of the neutral moiety or, alternatively, simultaneous N–H•••N and O–H•••O=C hydrogen bonding occurs. The different binding motifs may be viewed as different regions of the associated potential energy surface; species within a grouping share a binding motif and can interconvert via conformational changes associated with rotation around bonds. Thus, the different groupings in Figure 3.6 represent funnels on the potential energy

45

landscape of [Phe$_2$ + H]$^+$, with motifs 1 and 3 being relatively high in energy compared to motifs 2 and 4.
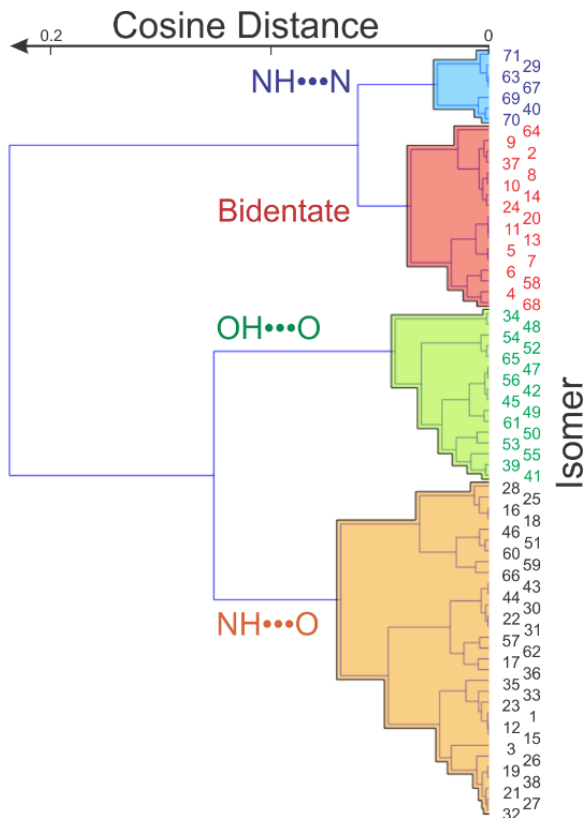
Figure 3.6 A dendrogram of the 71 low energy conformers and isomers of [Phe$_2$ + H]$^+$ constructed from cosine distances between the mass-weighted distance vectors of each isomer. The isomers are classified into four groups associated with N-H···N, bidentate, O-H···O, and N-H···O binding motifs.

3.3.4.2 Assigning IRMPD Spectra

The experimental IRMPD spectrum for [Phe$_2$ + H]$^+$ was recorded by monitoring fragmentation efficiency for the production of Phe•H$^+$ (*m/z* 166) as a function of FEL wavenumber (see Figure 3.7A). As expected, vibrational bands are observed at ~1750 cm-1 (C=O stretching), in the range 1400 – 1600 cm^{-1} (NH$_3$ umbrella, NH$_2$ scissor, COH bending), and at ~1100 cm^{-1} (phenyl ring vibrations). Experimental IRMPD spectra were compared with calculated IR spectra to determine the spectral carriers likely to be present in the probed ensemble. The

intensities of the experimental IRMPD spectrum and the computed IR spectra in the wavenumber range 900-1900 cm^{-1} were first normalized to a maximum of 1, then cosine distances between the experimental and computed intensity vectors for each spectrum were calculated as outlined in previous articles.[40,170] The computed similarities of the intensity vectors were then renormalized across the [0,1] interval. The resulting similarities are plotted in Figure 3.8. The blue-colored panel in Figure 3.8 indicates the isomers that have the closest distance (*i.e.* best match) with the experimental spectrum. The calculated spectra of isomers 1 (0 kJ·mol^{-1}), 19 (30.5 kJ·mol^{-1}), 35 (63.6 kJ·mol^{-1}), and 71 (241.8 kJ·mol^{-1}) all show reasonable agreement with the experimental IRMPD spectrum based on this analysis. Note that isomers 1, 19, and 35 all belong to the same group shown in Figure 3.6 (the N-H···O binding motif).

Figure 3.7B-E plot the calculated spectra for isomers 1, 19, 35, and 71, respectively. Based on the similarities plotted in Figure 3.8 , isomer 35 exhibits the best agreement with the experimental IRMPD spectrum. However, calculated features for isomer 35 in the range of 1200-1300 cm^{-1} do not appear in the experimental spectrum, eliminating it as the candidate structure. We attribute the high similarity value of isomer 35 to its greater resemblance in its transition intensities with respect to those of the experimental spectrum. Likewise, the calculated spectra for isomers 19 and 71 also exhibit features that are not observed in the experimental spectrum. Furthermore, calculated positions of the carbonyl stretching band (~1750 cm^{-1}) for isomer 71 and the phenyl ring vibration (~1100 cm^{-1}) for isomer 19 do not agree well with those observed experimentally. Moreover, since isomers 19 and 35 are both higher energy conformers associated with the potential energy funnel of isomer 1, we favour isomer 1 as being responsible for the observed experimental IRMPD spectrum of [Phe$_2$ + H]$^+$. The comparison between the experimental spectra and global minimum for other Phe derivative clusters are provided in the Appendix B. The Table B.2 has shown the similarities of all the Phe derivatives isomers. In general, the experimental IRMPD spectra can be assigned to the calculated global minimum of each derivative.
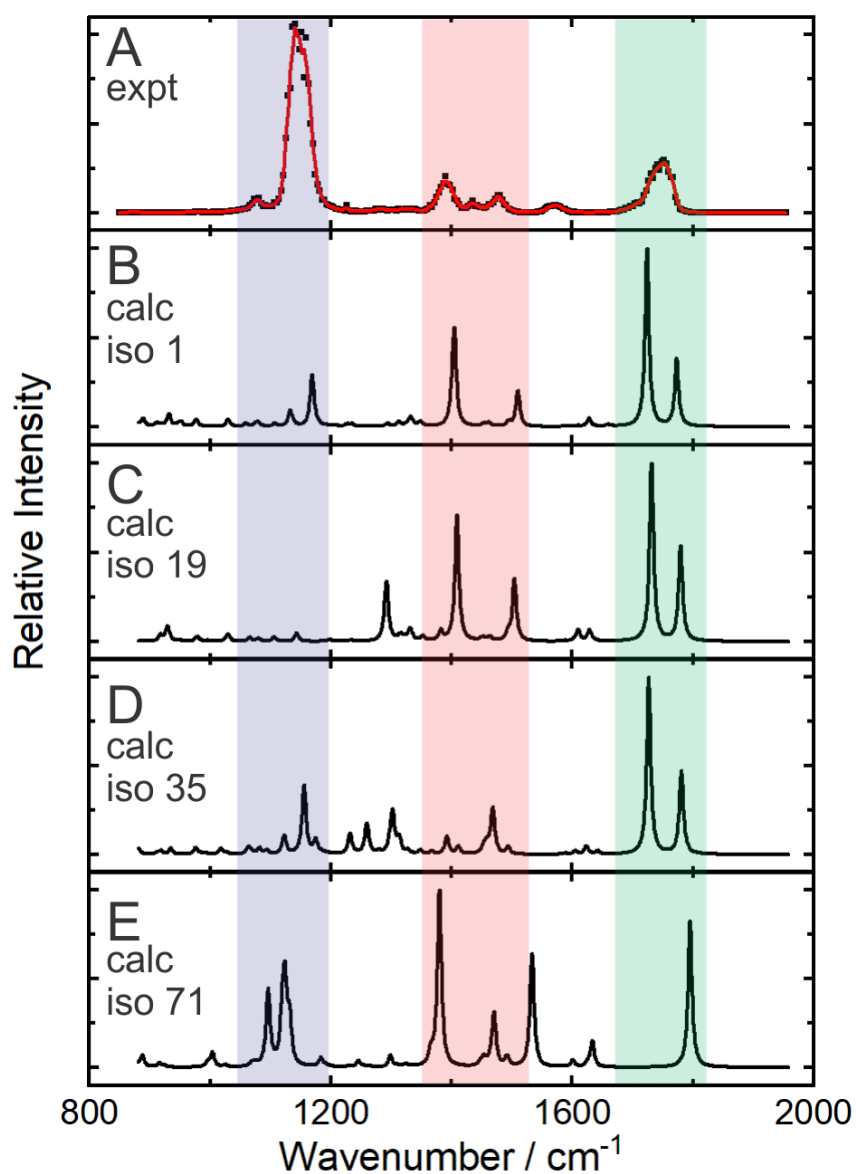
Figure 3.7 (A) The experimental IRMPD spectrum of $[Phe_2 + H]^+$ and the calculated IR spectra for (B) isomer 1, (C) isomer 19, (D) isomer 35, and (E) isomer 71. These calculated spectra yielded the closest match to experiment based on the unsupervised ML analysis shown in Figure 3.8.
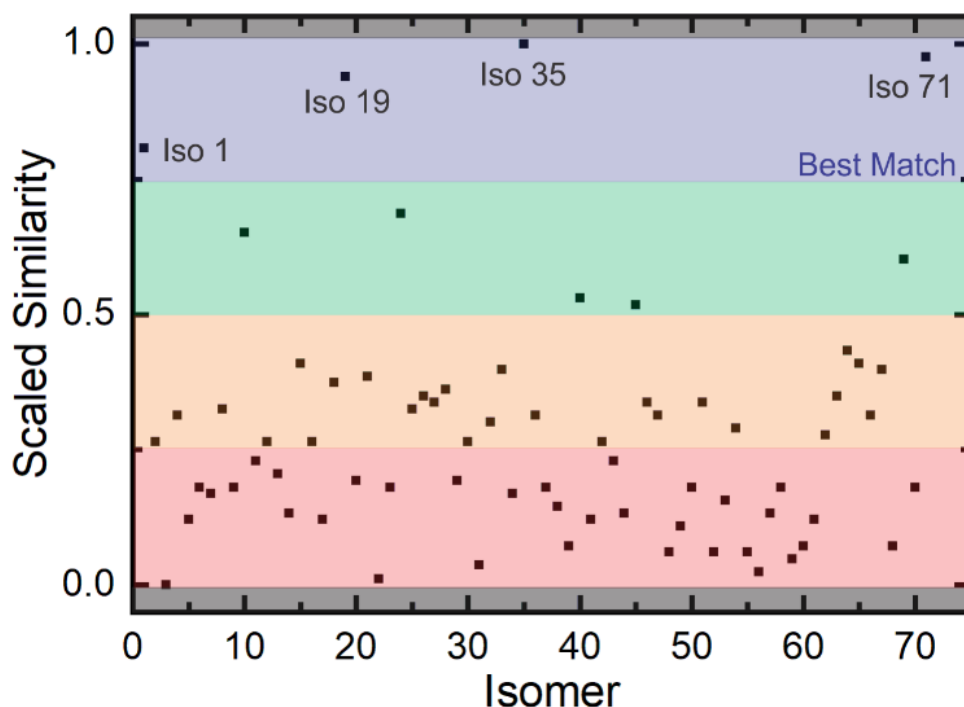
Figure 3.8 Spectral similarities between the calculated IR spectrum and the experimental IRMPD spectra of $[Phe_2 + H]^+$. Similarities were rescaled to the [0,1] interval. Experimental spectra that yield the closest match to the computed spectrum appear in the [0.75,1] range (highlighted in blue).

3.3.4.3 Quantum Theory of Atoms in Molecules and Non-Covalent Index Analyses

Following identification of the gas-phase structures of the homodimers, non-covalent interactions were calculated for the spectral carriers. Non-covalent interactions within and between moieties can be assessed by calculating the electron density and reduced density gradient.[70] To map the non-covalent interactions present in the various homodimers, the Multiwfn package was employed.[171] The gradient isosurfaces of $[Phe_2 + H]^+$ isomer 1 is shown in Figure 3.9. The color of the surface indicates the strength of the interaction; blue represents a strong attractive interaction, green a weak attractive interaction, red a strong repulsive interaction, and yellow a weak repulsive interaction. Point I on in Figure 3.9 shows a strong intermolecular hydrogen bond between the ammonium group of the protonated moiety and the carbonyl group of the neutral moiety (blue surface). Based on Bader's QTAIM theory,[60,62,65] the electron density (ρ) and Laplacian of the electron density ($\nabla^2\rho$) at the bond critical points

may be used to characterize chemical bonding and non-covalent interactions.[61] For weak interactions, such as H-bonding, ρ is typically less than ~10^{-2}, while for van der Waals interactions ρ ~10^{-3}.[61] The cation-π interactions have been also reported to be around ~10^{-2}.[66,67] The value of $\nabla^2\rho$ in these interactions should be positive.[61] For the intermolecular H-bond identified at region I (see Figure 3.9), $\rho = 0.0552$ and $\nabla^2\rho = 0.1452$, indicating a very strong hydrogen bond. This is consistent with calculations at the B3LYP/6-311++G(d,p) level of theory, which indicate a bond dissociation energy of 105.4 kJ mol^{-1}. Calculations also show a relatively strong intramolecular H-bond between the OH and NH$_2$ groups of the neutral moiety (region II; $\rho = 0.0418$ and $\nabla^2\rho = 0.1085$). The red surface associated with region II indicates strong steric repulsion within the 5-membered ring. QTAIM and NCI analysis also indicates the presence of cation-π (regions III and V), anion-π (region VI), and π-π (region IV) interactions. These regions all exhibit weakly attractive interactions with strengths on the order of those expected for van der Waals interactions. The results of the QTAIM analysis for [Phe$_2$ + H]$^+$ isomer 1 are reported in Table 3.1.

However, for the cation-π interaction in region III and VI do not have bond critical points, thus ρ and $\nabla^2\rho$ cannot be read, which does not necessarily mean that the these weak cation/anion-π interactions do not exist.[69] NCI analysis shows the green isosurfaces, which indicates the presence of the weak non-covalent interactions. The values of ρ for π-π stacking and T-shaped interactions were all reported to be around 7×10^{-3} a.u.[61] As can be seen, the weak π-π interaction in the [Phe$_2$ + H]$^+$ global minimum is 6×10^{-4} (label IV). It is much weaker than the reported values, which means the π-π interaction here is not dominant and is likely to have little influence on the structure.
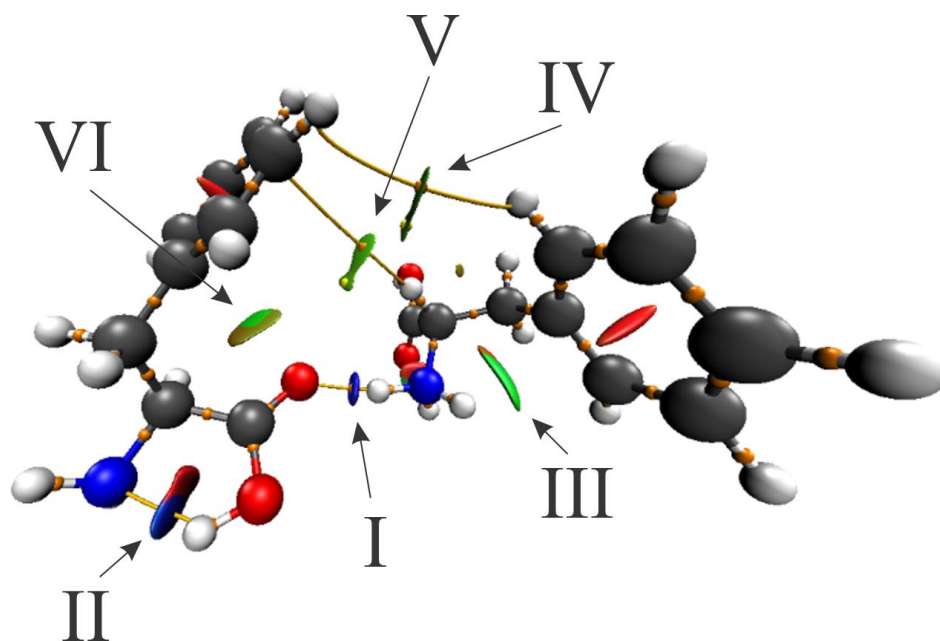
Figure 3.9 Bond critical points (orange points) and gradient isosurfaces for $[Phe_2 + H]^+$ isomer 1 as calculated at B3LYP/6-311++G(d,p) level of theory. Isosurfaces show non-covalent interactions with the strength of interaction being colour-mapped from strongly attractive (blue) to strongly repulsive (red). (I) intermolecular hydrogen bond, (II) intramolecular hydrogen bond and steric repulsion within the 5-membered ring, (III) cation-π interaction with ammonium group, (IV) π-π interaction, (V) cation-π interaction with CH group, (VI) anion-π interaction with O centre.

Table 3.1 Electron densities (ρ) and Laplacians ($\nabla^2\rho$) for the non-covalent interactions in $[Phe_2 + H]^+$ isomer 1 (shown in Figure 3.9). Calculations were conducted at B3LYP/6-311++G(d,p) level of theory. Bond critical points were not found for points III and VI, so values were not reported.

Critical Point	Electron Density (ρ)/a.u.	Laplacian ($\nabla^2\rho$)/a.u.
I	0.0552	0.1452
II	0.0418	0.1085
III	-	-
IV	0.0006	0.0021
V	0.0009	0.0029
VI	-	-

3.3.4.4 The structures of homodimers of Phe derivatives

The calculated global minima of all phenylalanine derivative homodimers are shown in Figure 3.10. The coordinates have been provided in Appendix B. In all cases, protonation was found to occur on the amine N centre. QTAIM analysis of the global minima indicates that both intramolecular and intermolecular hydrogen bonding contribute to the stabilization of the complexes. In seven of the ten homodimers studied, intramolecular H-bonding occurs in the neutral moiety. Intermolecular H-bonding occurs via NH···O interactions in all cases. For eight of the systems studied, the intermolecular NH···O interaction is between the ammonium group and the carbonyl oxygen of the neutral moiety. In two cases, $[3,4\text{-MeO-Phe}_2 + H]^+$ and $[4\text{-}NO_2\text{-Phe}_2 + H]^+$, the NH···O interaction occurs between the ammonium group and the side chain substituent. The $[3,4\text{-MeO-Phe}_2 + H]^+$ system is further stabilized by a second intermolecular H-bond between the OH group of the protonated moiety and the NH_2 group of the neutral moiety. Bidentate binding via a second intermolecular H-bond is also observed in the $[3\text{-CN-Phe}_2 + H]^+$ derivative, where the OH group of the protonated moiety interacts with the CN group of the neutral moiety, and in the $[3\text{-CF}_3\text{-Phe}_2 + H]^+$ and $[2,5\text{-F}_2\text{-Phe}_2 + H]^+$ species, where the ammonium group interacts with the carbonyl and amine groups of the neutral moiety.
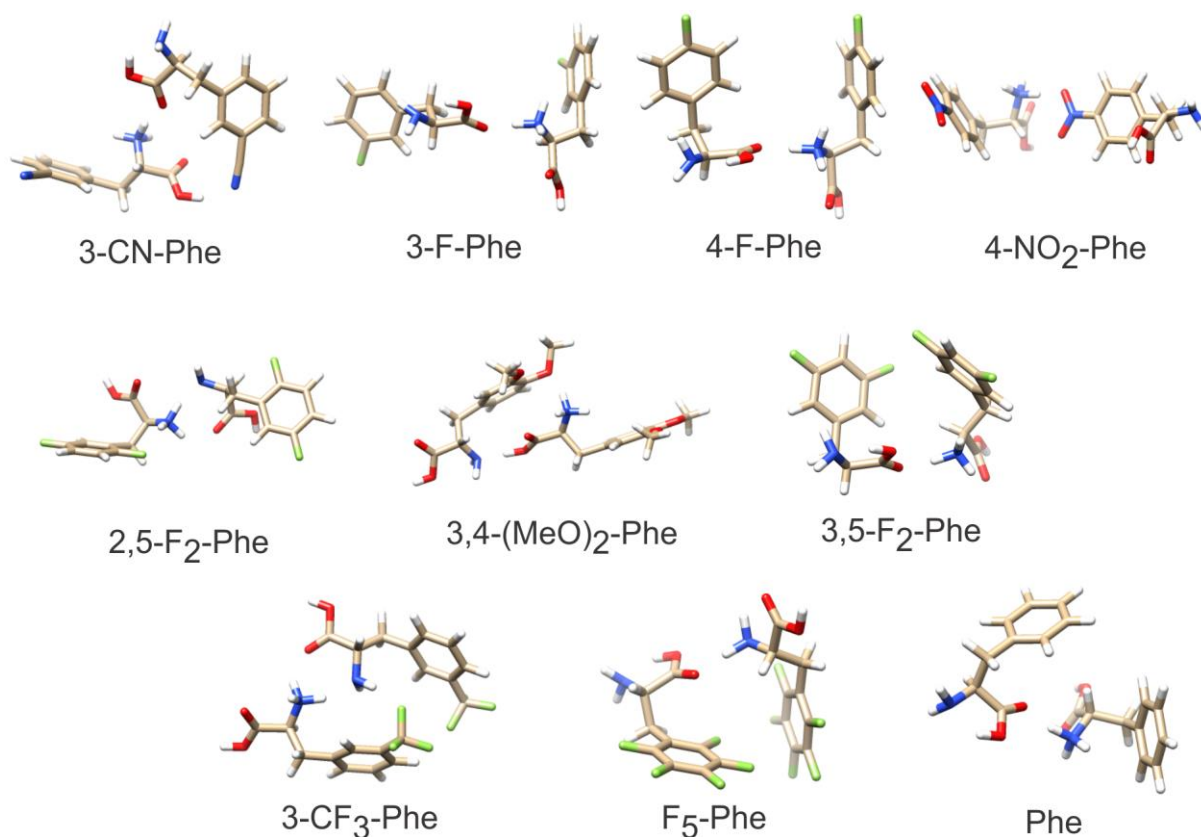
Figure 3.10 The global minimum structures of Phe derivative homodimers. The calculations were conducted at B3LYP/6-311++G(d,p) level of theory.

To better visualize the binding observed across all Phe derivative homodimers studied, multi-dimensional scaling (MDS) was used to create a 2D plot of the clustered data (shown in Figure 3.11).[148] Based on this analysis, we find five groupings of related geometric structures (colour coded in Figure 3.11). In general, these grouping are based on the homodimer binding motif and the relative compactness of the geometry. Species that were observed experimentally are plotted as large data points and are labelled. In general, as one moves from left to right across the graph, the intermolecular H-bonding motif evolves from NH⋯R to [NH⋯R + OH⋯N] to [NH⋯O + OH⋯R] to NH⋯O to [NH⋯O + NH⋯N]. Thus, towards the left-hand side of Figure 3.11 one finds structures associated with the 4-NO$_2$-Phe, 3,4-MeO-Phe, and 3-CN-Phe derivatives, which can all participate in relatively strong H-bonding that involves the sidechain. Towards the right-hand side of the plot are structures associated with intermolecular NH⋯O hydrogen bonding between the amino acid backbones. The vertical dimension

provides information on the homodimer morphology. Structures located near the top of the plot are relatively extended (see, *e.g.*, [3F-Phe$_2$ + H]$^+$ in Figure 3.10), whereas those near the bottom of the plot are relatively compact (see, *e.g.*, [F$_5$-Phe$_2$ + H]$^+$ in Figure 3.10). Thus, the distribution of experimentally observed species in Figure 3.11 indicates that, for the most part, the protonated homodimers of Phe and its derivatives favour NH···O intermolecular hydrogen bonding and relatively open, extended geometries.
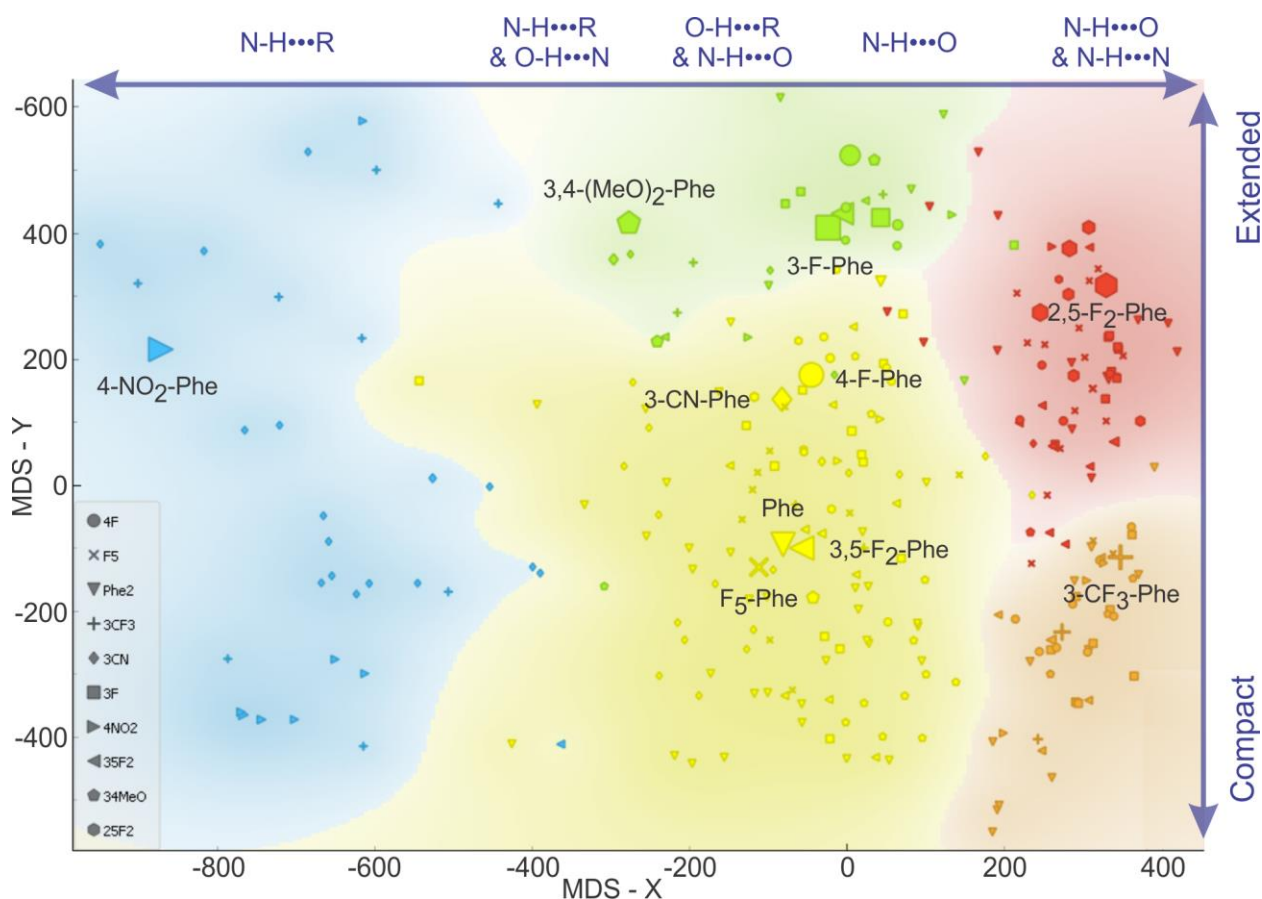
Figure 3.11 The multidimensional scaling (MDS) of all Phe derivatives isomers. The left legend shows the symbol of each derivative. The size of the symbols correlates with the stability of the isomer. The right legend shows the five colors of each clustering type. From left to right, only the side chain becomes only the backbone that participates in the intermolecular H-bonding motif. From top to bottom, structures are from relatively elongated to compact.

54

3.3.5 Conclusions for [(x-Phe)$_2$ + H]$^+$

A BH search algorithm has been used to map the PES of the protonated Phe derivative homodimers. Experimental IRMPD spectra have been compared with calculated IR spectra obtained at the DFT level of theory with cosine similarities. The protonation site is found to be at the amine nitrogen for the global minimum structures. Based on cosine distance, intermolecular hydrogen bonds between the protonated moiety and the neutral moiety can be visualized in the MDS plot, the intermolecular H-bonding motif evolves from NH···R to [NH···R + OH···N] to [NH···O + OH···R] to NH···O to [NH···O + NH···N]. In 4-NO$_2$-Phe and 3,4-(MeO)$_2$-Phe, amine prefers binding with a side chain oxygen (NH···R). The amino groups in 2,5-F$_2$-Phe and 3-CF$_3$-Phe favourably interact in a bidentate fashion with the carbonyl oxygen and the amine nitrogen of the neutral moiety, and the remaining six species bind through [NH···R + OH···N] to [NH···O + OH···R] to NH···O. The electron density and Laplacian values also indicate cation-π interactions occur, but π-π interactions might be too weak to influence the structure. This work provides an efficient way to categorize many clusters to reveal trends in binding motifs. Non-covalent interactions in the Phe derivatives homodimer clusters can also be proved quantitatively.

3.4 Conclusions for Phe-Containing Dimers

Taken as a whole, this chapter applied hierarchical clustering to categorize the structures of the proton-bound heterodimer of Phe/Ser and homodimer of Phe derivatives. BH algorithm has successfully searched the PES and found the global minimum candidates. The binding motifs were identified based on the structures. Especially for [(x-Phe)$_2$ + H]$^+$, we were able to find the correlations between the substituents' effect and the binding motifs. In addition, we compared the calculated IR spectra with the experimental UV/IR double resonance spectra recorded by Lorenz and Rizzo[32] for [Phe/Ser + H]$^+$ and compared the calculated IR spectra with the IRMPD spectra for [(x-Phe)$_2$ + H]$^+$. The assignment of the experimental spectra can be achieved by calculating cosine distances between the calculated spectra and experimental spectra.

Chapter 4

UV Photodissociation Spectroscopy of Para-Substituted Benzylpyridinium Ions

4.1 Introduction

Thermometer ions are the ions that can fragment in a well-defined pathway when their internal energies reach the threshold energies associated with the fragmentation pathway.[35] The requirements for being a thermometer ion are that the fragments can be obtained via a simple reaction with no competitive reactions, and the activation energies of the fragmentation threshold are known.[172] To meet these requirements, several compounds can be chosen as thermometer ions, such as tetraethylsilane ions,[172,173] transition-metal carbonyl complexes,[172] and substituted benzylpyridinium (BP) salts.[174] BP ions have a simple fragmentation pathway, which makes them suitable candidates for thermometer ions.[174,175] The fragmentation pathway of BP ions is the direct bond cleavage of the C-N bond between the pyridyl and benzyl moieties, which generates neutral pyridine and benzylium cation as shown in Figure 4.1. The energy required to drive the dissociation can be tuned by the type and position of the substituents on the benzyl moiety.[174,175] As a consequence, the bond dissociation energy for cleavage of this bond can be used to estimate internal ion temperatures.
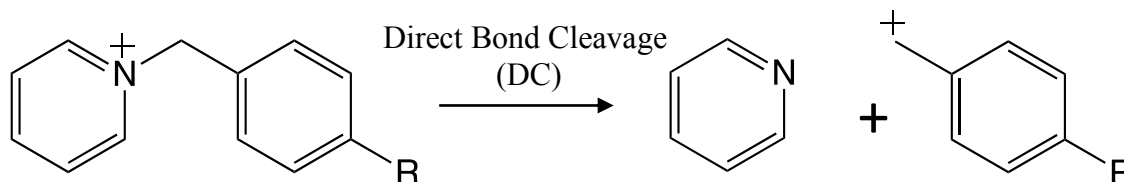
Figure 4.1 Fragmentation of para-substituted benzylpyridinium ions (M^+) to yield the substituted benzyl cation ($[M-79]^+$) and pyridine

BP ions have been used as thermometer ions in mass spectrometry for many studies.[34,35,176–178] All of these studies utilize dissociation on the BP ion ground state to assess internal energy and therefore temperature. De Pauw and coworkers first studied BP ions to estimate the internal

energy distribution of BP ions arising from a soft ionization technique — liquid secondary ion mass spectrometry (LSIMS).[174,179,180] Similarly, substituted BP ions have been used to estimate the internal energy of ions produced from electrospray ionization (ESI),[175,181–184] matrix-assisted laser desorption/ionization (MALDI),[185] and femtosecond laser desorption ionization (fs-LDI).[186,187] Most of the work investigated the fragmentation pathways of BP ions arising from in-source fragmentation, CID experiments, and IRMPD.[34,176,177,188] Recently, a series of para-substituted BP ions were used to characterize the effective temperature of ions in a differential mobility spectrometer, to estimate field-induced heating effects from the applied SV waveform.[189]

While extensive research has been conducted on these ions, their electronic spectra have not yet been investigated. Given our new capability for spectroscopy of trapped ions and the fact that these are aromatic species and should therefore absorb strongly in the UV region, a UVPD study can be conducted to better characterize BP and its derivatives. The UV spectroscopic study of the neutral methyl-substituted benzene demonstrates that the substituents can influence $\pi \rightarrow \pi^*$ electronic transitions of the benzene ring.[190] The substituent effect on the electronic transitions of the benzene cation has also been studied via UV-Vis photodissociation spectroscopy, which found that the absorption peaks shifted to lower energy when increasing the number of substituting methyl.[191] By comparing the UVPD spectra between all H and the substituted species, it will be interesting to see if each different chemical derivative will exhibit slightly different electronic spectra owing to the electronic effects of the EDGs and EWGs influencing the electronic transitions. In this work, the UVPD action spectra of the BP ions were explored in a custom-modified DMS-MS device. To eliminate other factors that can cause dissociation and rearrangement of BP ions in the DMS prior to the irradiation by the laser, the samples were studied with DMS voltages (SV and CV) set to 0 V (so as not to activate the ions). The thermometer ions studied here include BP and para-substituted BP ions (R = -H, -F, -Cl, -CN, -Me, -OMe, -OEt, -OiPr, -OtBu). Theoretical UV absorption spectra were calculated to compare with the experimental spectra and confirm the transition types.

4.2 Experimental and Theoretical Methods

4.2.1 Experimental Methods

The apparatus used in this study is based on a DMS-MS device[192,193] that has been modified to enable optical access into the Q3 quadrupole ion trap; the modifications are described in detail in reference 105. Samples were prepared with para-substituted benzyl chloride derivatives dropwise with pyridine. The details for the synthesis of the BP species can be found in reference 189. The samples were dissolved in acetonitrile and diluted in methanol. The samples were electrosprayed and ions were propelled through the DMS cell by a stream of dry N_2 carrier gas at atmospheric pressure DMS voltage, *i.e.*, the separation voltage (SV) and compensation voltage (CV), set to 0 V so as not to activate/dissociate the analyte ions. The DMS cell was heated to 423 K for all experiments involving DMS to maximize ion transmission through the cell. The ions were transferred into a triple-quadrupole mass spectrometer where they were mass-selected in the first quadrupole (Q1) before detection by a channeltron ion detector.

For photodissociation action spectroscopy experiments, each mass-selected substituted BP ion was accumulated in Q3 for approximately 5 ms and subsequently trapped for 500 ms. Trapped ions were irradiated with the output of a tunable Nd:YAG-pumped OPO system (Continuum Horizon II, 10 Hz), where they underwent photodissociation following absorption of light. The irradiated ion bunch was ejected from the trap using mass-selective axial ejection,[194] and the intensities of parent ions and any daughter ions were recorded as a function of the OPO output wavelength. The scanned wavelength ranged from 208 to 292 nm.

4.2.2 Theoretical Methods

Structures of protonated BP ions were optimized at B3LYP/6-311++G(d,p) level of theory using Gaussian 16 A03.[195] The BP ions are permanently charged on the pyridinium N atom. Normal mode analyses were conducted to ensure the optimized structures were local minima on the PES. The frequencies were re-calculated at CAM-B3LYP/6-311++G(d,p) level using the quantum chemistry program package ORCA.[80] The Excited State Dynamics (ESD) program was called to calculate the absorption spectra in ORCA with the ESD(ABS) keyword,

where ABS refers to a computed absorption spectrum. The first fifteen vertical excitation energies were calculated using TD-DFT at the CAM-B3LYP/6-311++G(d,p) level of theory on ORCA 4.2.1. The vibronic contribution of the first twelve excited transitions was individually calculated using the VG-FC module at the same level of theory and summed to create a vibronic spectrum. Vertical excitation energies were also calculated using the STEOM-DLPNO-CCSD framework implemented in ORCA 4.2.1. This method has been shown to provide a more accurate depiction of excited energy levels and dynamics, as well as predicted vibronic spectrum, in comparison to TD-DFT.[196] However, the STEOM-DLPNO-CCSD method is only available for calculating energies and transition moments, so cannot be used to calculate optimized geometries or frequencies. Thus, the vertical absorption energy calculated by the STEOM-DLPNO-CCSD associated with the TDDFT geometry and frequency were used.

4.3 Results and Discussion

4.3.1 Photodissociation Action Spectra

The photofragmentation mass spectrum of the benzylpyridinium ion at 215 nm is shown in Figure 4.2. The parent benzylpyridinium ion (*m/z* 170) yields a single dominant fragment ion at *m/z* 91. The loss channel, [M−79]$^+$, which corresponds to the loss of neutral pyridine, is characteristic of all the substituted benzylpyridinium ions described here.
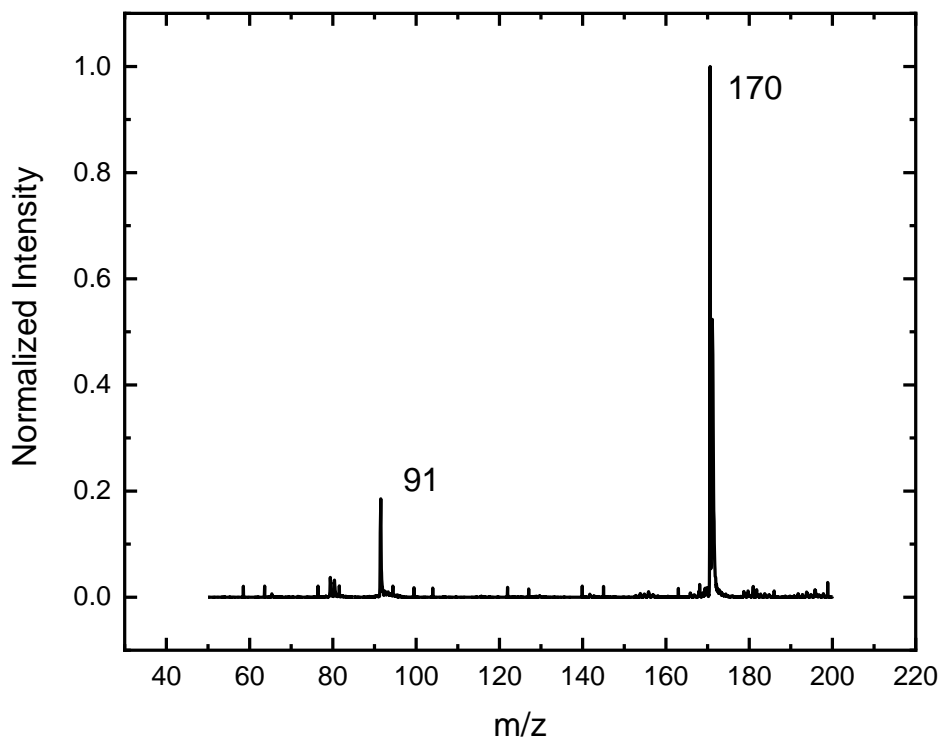
Figure 4.2 UVPD fragmentation mass spectrum of the benzylpyridinium cation.

There were some cases where additional fragments (*m/z* 107) can be observed for the alkoxy derivatives *i.e.*, -OEt, -OiPr, and -OtBu. The mass spectra for those derivatives at 237 nm are shown in Figure 4.3. As the number of methyl groups in the substituents increases, the relative intensities of the fragment of *m/z* 107 increase, while the relative intensities of fragments of [M−79]$^+$ decrease. Especially for BP-OtBu, the fragment of [M−79]$^+$ (*m/z* 164) can be barely seen. It means that the fragment of *m/z* 107 is more favoured than the fragment of [M−79]$^+$ when the number of methyl groups in the substituents increases. Also, for BP-OtBu, the intensity of the fragment of *m/z* 80 increases dramatically, which comes from the protonated fragment of pyridine. The fragmentation mechanism of the alkoxy groups substituted BP ions is shown in Figure 4.4.
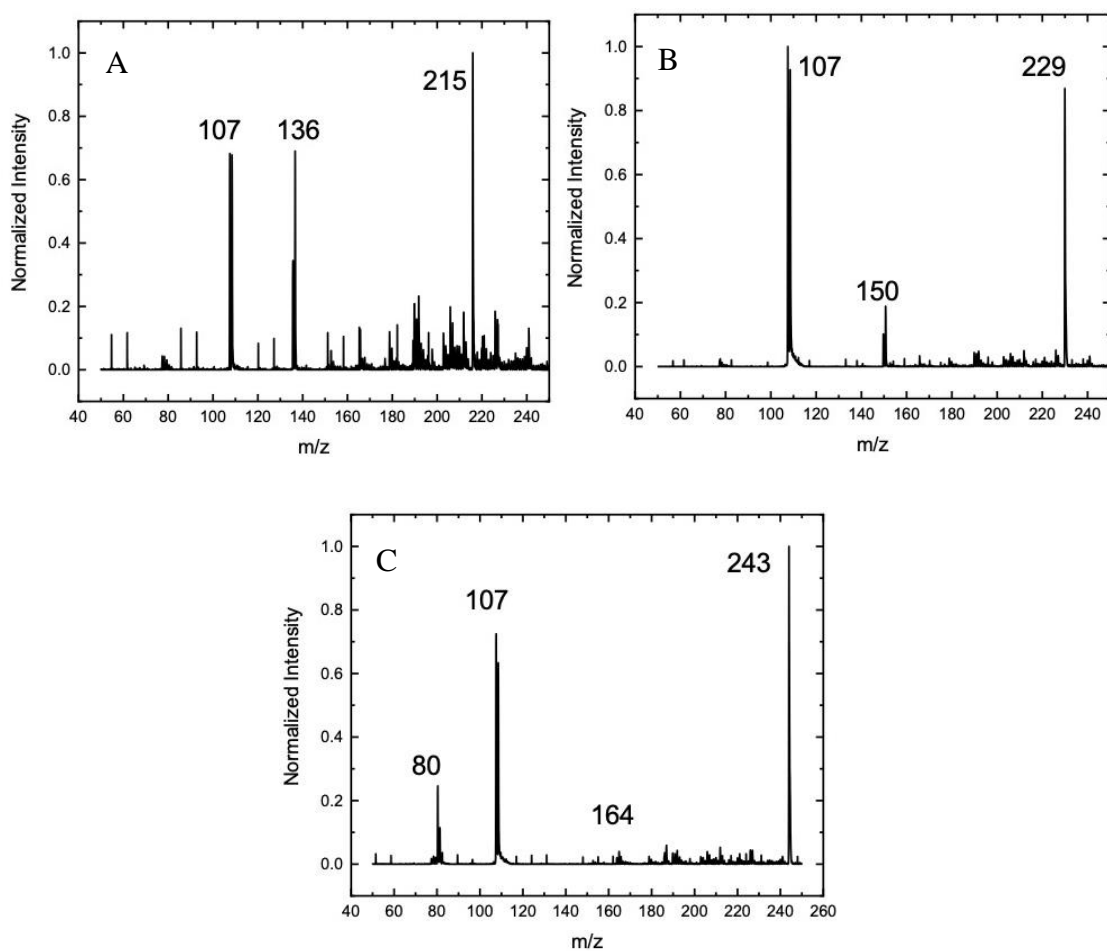
Figure 4.3 UVPD fragmentation mass spectrum of (A) BP-OEt cation, (B) BP-OiPr cation, and (C) BP-OtBu cation.
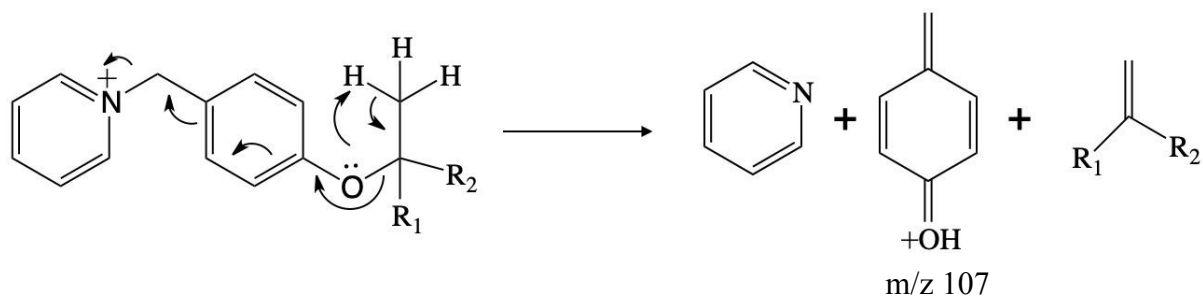
Figure 4.4 Fragmentation mechanism detailing the formation of m/z 107 from UVPD of the alkoxy-substituted BP ions

UVPD action spectra have been measured for nine para-substituted benzylpyridinium ions. The intensities were normalized with respect to the total ion count and the OPO power. As can be seen in Figure 4.5, two well-separated absorption features are observed in the spectra of the BP ion and its para-substituted derivatives. Species that are derivatized with EWGs (R = F, Cl, CN) exhibit a broad feature in the 4.6-5.0 eV region and a second feature in the 5.3-6.0 eV region. For the spectra of BP-Me, two features were observed for the BP ions substituted by Me. For BP-OMe ion, two observed features shift to shorter wavelengths. For BP ions are substituted with other alkoxy groups (R = OEt, OiPr, OtBu), the UVPD action spectra are shown in Figure 4.6. The second higher-energy feature has shifted out of the accessible range defined by the probe range. The action spectra of the fragments are also shown in Figure 4.6. For BP-OEt, the action spectra of the fragments of m/z 136 and m/z 107 are almost identical, which shows the absorption energies at 5.2 eV. For the action spectra of the fragment of m/z 136, the intensities slightly increase after 5.6 eV compared with the spectra of m/z 107, which means after 5.6 eV, the fragmentation of direct bond cleavage (as shown in Figure 4.1) is favoured. At 5.2 eV, BP-OEt has access to both of the two product channels. For BP-OiPr, the two fragments undergo different fragmentation processes. At 5.2 eV, the fragment of m/z 107 coming from the process of Figure 4.4 is favoured, while the fragment of $[M−79]^+$ (m/z 150) is favoured after 5.6 eV. For BP-OtBu, the product of m/z 107 can be accessed for both energies. The product of m/z 80 is favoured at 5.2 eV, whereas the product of m/z 164 prevails over the product of m/z 80 as the energy increases above 5.6 eV. The absorption energies of the observed features are summarized in Table 4.1.
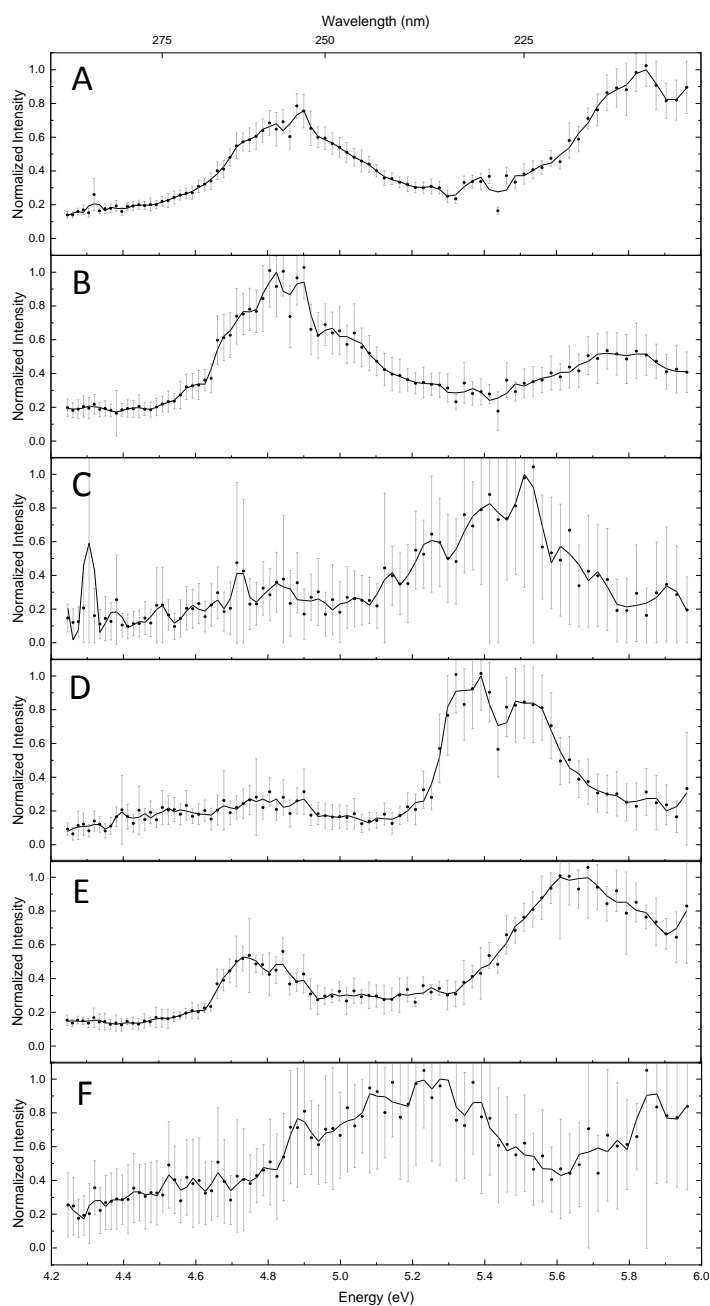
Figure 4.5 UVPD action spectra of para-substituted benzylpyridinium ions electrosprayed from methanol: (A) BP-H; (B) BP-F; (C) BP-Cl; (D) BP-CN; (E) BP-Me; (F) BP-OMe. The spectra represent the fragmentation channel corresponding to $[M-79]^+$. Error bars represent $+/- 0.5\sigma$ at each point, where variations in both ion intensity and laser power were considered.
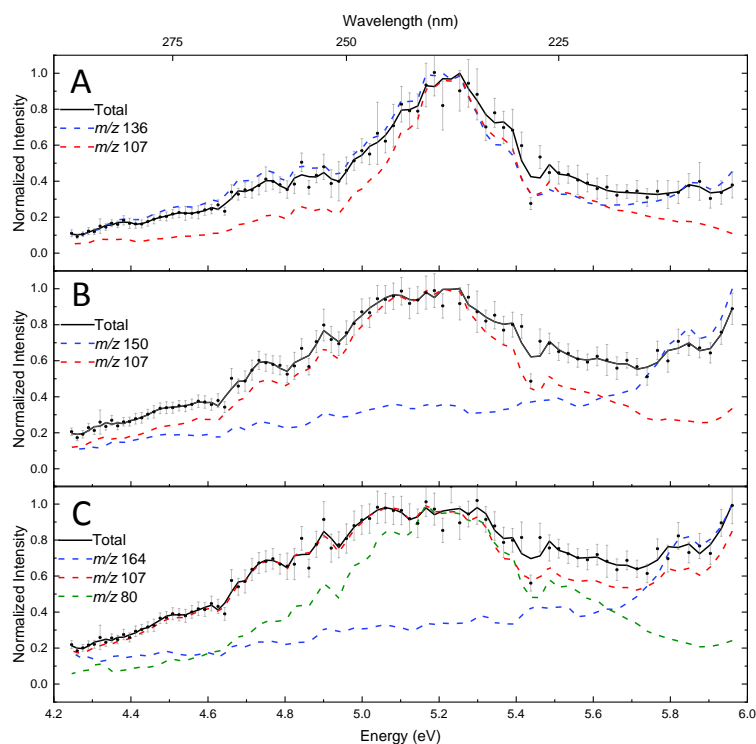
Figure 4.6 UVPD action spectra of para-substituted benzylpyridinium ions electrosprayed from methanol: (A) BP-OEt; (B) BP-OiPr; (C) BP-OtBu. The spectra in black represent the total fragmentation channels. The spectra in the blue dash line represent the fragmentation channel corresponding to [M−79]+. The spectra in the red dash line represent the fragmentation channel corresponding to *m/z* 107. The spectrum in the green dash line represents the fragmentation channel corresponding to *m/z* 80. Error bars represent +/− 0.5σ at each point, where variations in both ion intensity and laser power were considered.

Table 4.1 The experimental electronic transition band maxima of the para-substituted BP ions

	Energy (eV)		Wavelength (nm)	
	Band I	Band II	Band I	Band II
BP-H	4.87	5.86	254.5	211.5
BP-F	4.85	5.82	255.6	213.0
BP-Cl	4.79	5.44	258.9	228.0
BP-CN	4.53	5.44	273.5	228.0
BP-Me	4.76	5.63	260.5	220.4
BP-OMe	5.18	5.90	239.6	210.1
BP-OEt	5.21	N/A	238.2	N/A
BP-OiPr	5.15	N/A	240.8	N/A
BP-OtBu	5.14	N/A	241.3	N/A

4.3.2 Simulated Vertical Excitation Energies and Electronic spectra

The vertical transition energies and oscillator strengths for the first fifteen electronic transitions were calculated using the CAM-B3LYP 6-311++G(d,p) and STEOM-DLPNO-CCSD 6-311++G(d,p) methods (see Table 4.2). The first six states of the H-BP derivative as calculated by the CAM-B3LYP method, are expected to lie within the accessible experimental measurement range. At the CCSD level of theory, only the first four excited states lie within the accessible spectral range of the probe laser system, however, to err on the side of caution, the first 6 states were used to calculate the absorption spectrum of H-BP. The results of similar analyses for the remaining substituted BP derivatives are given in Appendix C.

Table 4.2 The calculated vertical excitation energies of the BP ions with the associated oscillator strengths (f_{osc}). The excitation energies are predicted using the CAM-B3LYP/6-311++G(d,p) and the STEOM-DLPNO-CCSD/6-311++G(d,p) level of theory.

BP-H	Energy (eV)		Wavelength (nm)		f_{osc}	
State	CAM-B3LYP	CCSD	CAM-B3LYP	CCSD	CAM-B3LYP	CCSD
1	4.65	4.83	266.7	256.9	0.0001179	0.0055094
2	4.77	4.94	260.1	251.1	0.0012016	0.0809100
3	5.33	5.29	232.7	234.2	0.0043323	0.0005281
4	5.45	5.58	227.5	222.2	0.0703338	0.0022347
5	5.72	6.24	216.6	198.7	0.0081083	0.1506005
6	5.74	6.41	215.9	193.6	0.0118857	0.0313142
7	6.06	6.41	204.7	193.5	0.1835774	0.0278487
8	6.27	6.45	197.9	192.1	0.0281198	0.1579475
9	6.87	6.62	180.5	187.2	0.8600453	0.6600045
10	6.92	7.19	179.0	172.4	0.3850504	0.0203005
11	7.00	7.26	177.2	170.8	0.0143647	0.2332850
12	7.09	7.88	174.9	157.3	0.0874119	0.0529414
13	7.28	7.92	170.3	156.6	0.0013464	0.0147046
14	7.30	7.92	169.9	156.6	0.1974482	0.0050653
15	7.43	8.03	166.8	154.4	0.3591463	0.0037928

UV absorption spectra for all BP derivatives were calculated using the VG-FC approach. The predicted electronic absorption spectrum of the BP-H ion is shown in Figure 4.7. The simulated spectra for the remaining substituted BP derivatives are given in Appendix C. As can be seen from the spectra and the calculated vertical excitation energies, the excited electronic states are nearly degenerate (with around 0.1 eV difference), which might give rise to the broad features in the experimental spectra. For BP-H, the calculated S_0–S_6 transition agrees with the experimental feature observed at 5.8 eV. The spectra for transitions of S_0 to S_1 and S_0 to S_2 match with the low-energy feature of the experimental spectrum of BP-H.
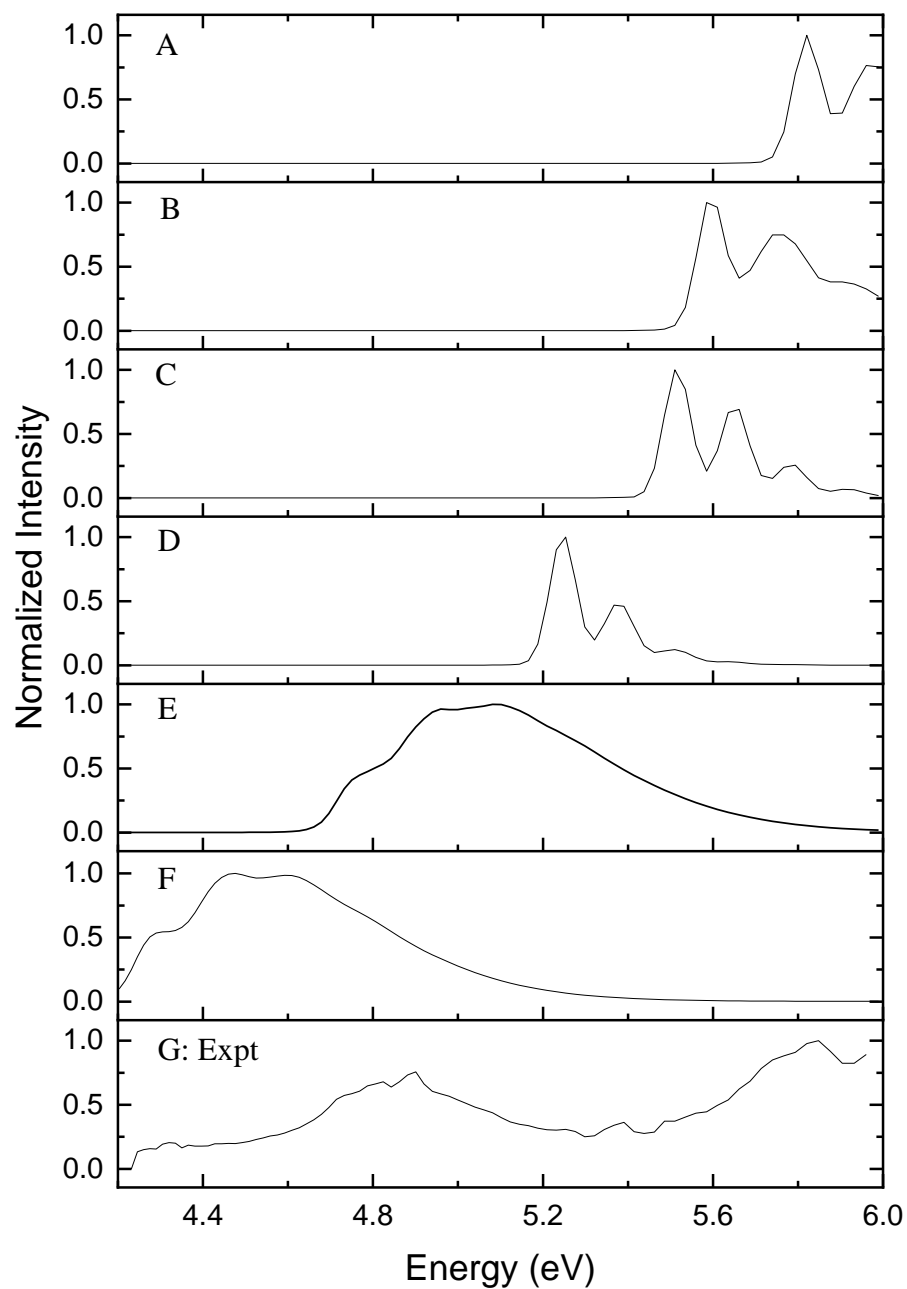
Figure 4.7 The experimental UVPD action spectra of BP-H ion (G) and the simulated UV vertical excitation spectra of (A) S_0-S_6, (B) S_0-S_5, (C) S_0-S_4, (D) S_0-S_3, (E) S_0-S_2, (F) S_0-S_1. The spectra were calculated using the CAM-B3LYP functional with the 6-311++G(d,p) basis set. All the spectra were normalized from 0 to 1.

As identified from the comparison between the simulated spectra and the experimental spectrum of BP-H, the experimental spectrum is comprised of three transitions: S_0 to S_1, S_0 to S_2, and S_0 to S_6. The molecular orbitals (MOs) with the highest weight for these transitions are shown in Figure 4.8. As can be seen from the figure, these transitions are $\pi \rightarrow \pi^*$ transitions wherein electron density. Through photoexcitation, the electron density is shifted to the pyridinium moiety and the charge centr is shifted to the C_7H_7 moiety. This destabilizes the molecule and leads to the production of neutral pyridine (C_5H_5N) and $[C_7H_7]^+$, which can adopt the tropylium form (*i.e.*, $C_7H_7^+$) or the benzyl cation form (*i.e.*, $C_6H_5CH_2^+$). Which of these two species is produced via photodissociation is an open question and requires further investigation.

Similar analyses have been conducted for all substituted BP derivatives studied (see Appendix C), which shows that the main transitions involve $\pi \rightarrow \pi^*$ transitions as well. The molecular geometries of Phe derivatives have been studied, which shows that the coordinates associated with cation-π interactions, correlate with Hammett constants since the substituents can influence the electron density of the phenyl ring.[33] Given that photoexcitation moves electron density from the chemically substituted benzyl moiety to the unsubstituted pyridinium moiety in each case, we hypothesized that the transition wavelengths are likely to be influenced by the electron withdrawing or donating nature of the substituents on the phenyl ring.
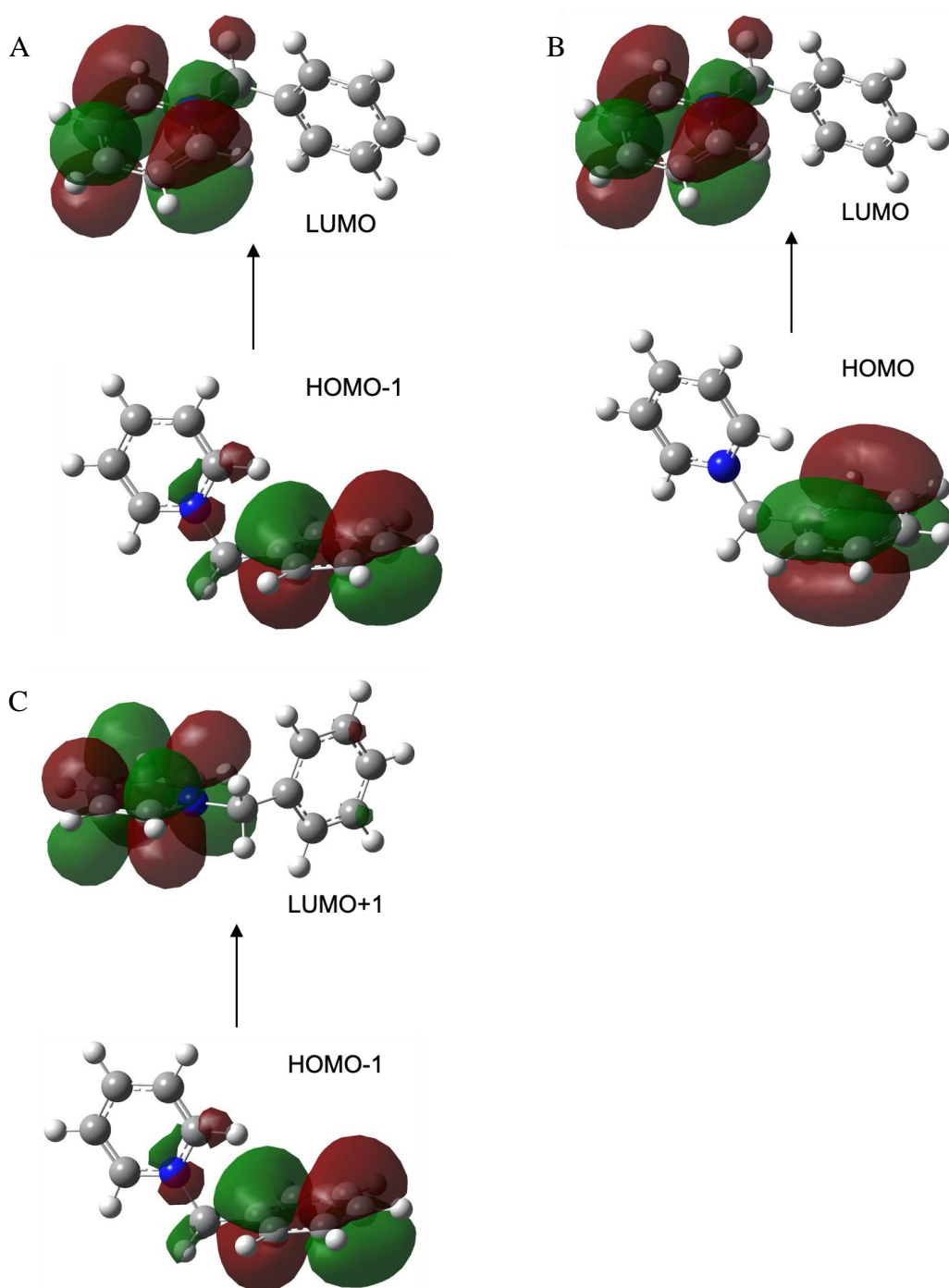
Figure 4.8 The calculated molecular orbitals of BP-H ion for transitions with percentage contributions: (A) S_0 to S_1 (96.8%), (B) S_0 to S_2 (97.9%), and (C) S_0 to S_6 (81.4%). S_0 is the ground state, and S_n is the nth excited state.

4.3.3 Experimentally Measured Transition Energies versus Hammett Constants

Hammett constants (σ) have been extensively applied to describe the movement of electrons and explain the trends in the reactivity of meta- or para-substituted benzoic acid derivatives and related compounds,[197] and were first developed and published by Louis P. Hammett.[198] Electronic excitation accompanied by electron density redistribution was proven to correlate with ground state Hammett constants in the optimized and vertical $\pi\pi^*$ states of para-substituted benzoic acids but with lower correlation coefficients compared to correlations found in the ground state.[199] Brown and Okamoto also reported electrophilic substituent constants (σ^+),[200,201] which are suitable for positive charge stabilization,[202] and have been applied to correlate with benzylpyridinium C-N bond dissociation.[188,203] The BP ions studied here were para-substituted with EWGs and EDGs. Since the calculated MOs indicated that the electronic transitions were $\pi \rightarrow \pi^*$ transitions and the substituents affect the electron density and stability of the ground state substituted rings, it is interesting to see if the electronic transitions of the substituted BP ions correlate with these constants. Table 4.3 shows the para-substituent constants σ, and σ^+ obtained from Ref. 204, 201 and 205, and the absorption features measured experimentally for x-BP.

Table 4.3 The experimental electronic transition energies of the para-substituted BP ions with Hammett constants (σ), and electrophilic substituent constants (σ^+). Only para-substituent σ, and σ^+ are reported here.

	σ	σ^+	Energy (eV)		Wavelength (nm)	
			Band I	Band II	Band I	Band II
BP-H	0	0	4.87	5.86	254.5	211.5
BP-F	0.06	-0.073	4.85	5.82	255.6	213.0
BP-Cl	0.23	0.114	4.79	5.44	258.9	228.0
BP-CN	0.66	0.659	4.53	5.44	273.5	228.0
BP-Me	-0.17	-0.311	4.76	5.63	260.5	220.4
BP-OMe	-0.27	-0.778	5.18	5.90	239.6	210.1
BP-OEt	-0.24	N/A	5.21	N/A	238.2	N/A
BP-OiPr	-0.45	N/A	5.15	N/A	240.8	N/A
BP-OtBu	N/A	N/A	5.14	N/A	241.3	N/A

Figure 4.9 shows the linear regression of the experimental absorption energies of the band I and band II against the para-substituent constants σ, and σ⁺. The absorption energies of the band I correlate with these two constants very well, whereas the linear regression fits the data of band II poorly. For the band I, the fitting line of electrophilic substituent constants (σ^+) has a higher correlation coefficient. For band II, the peaks cannot be determined with accuracy since some of them extend the experimental range. The transition energies show a higher correlation to the Hammett constants (σ) in comparison to electrophilic substituent constants. Overall, the two parameters correlate well with the absorption energies of band I. These correlations indicate that the excited transition energies can be predicted by these substituent constants.
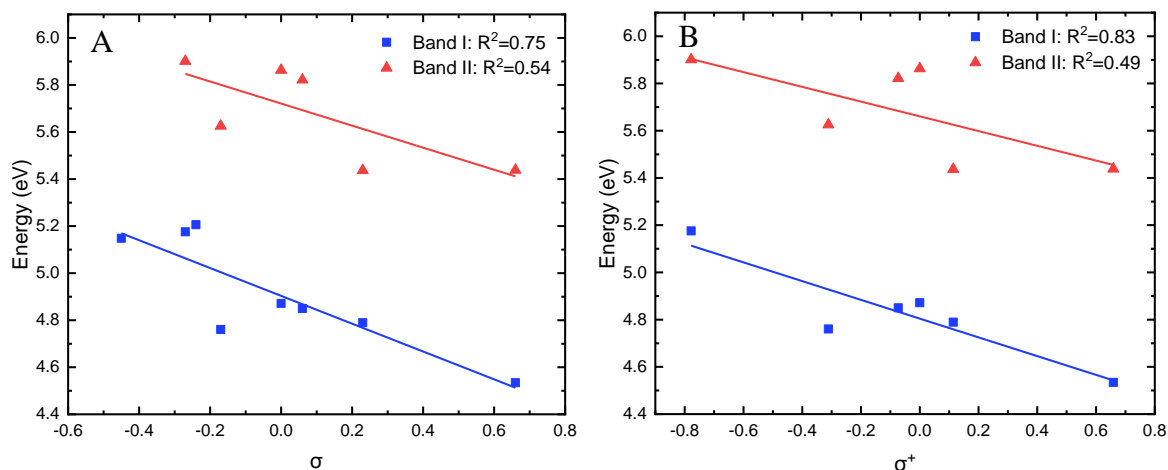
Figure 4.9 Experimentally determined absorption energies of the band I (blue squares) and band II (red triangles) versus (A) Hammett constants (σ), and (B) electrophilic substituent constants (σ^+). Straight lines represent linear regression models of the data.

4.4 Conclusions

This study illustrates that a modified DMS-MS instrument has been used to successfully measure the UVPD action spectra of para-substituted BP ions. In general, the spectra showed two features. Usually, the substituted BP ions fragment via the direct bond cleavage of the C-N bond between the pyridyl and benzyl moieties. However, for the BP ions substituted by alkoxy groups (R = OEt, OiPr, OtBu), the dissociation can produce the additional fragment

(*m/z* 107). The fragments of [M−79]$^+$ and *m/z* 107 exhibit different action spectra. The fragment of *m/z* 107 exhibit absorption at around 5.2 eV, while the fragment of [M−79]$^+$ exhibit more absorption after 5.6 eV than that of *m/z* 107. The VG-FC approach was able to predict the simulated absorption spectra of all the BP ions. The simulated spectrum of each transition allowed the experimental spectra to be assigned to specific electronic transitions. The MOs demonstrated the electronic transitions were predominantly π→π* character and can be affected by the substituents. The absorption energies of the band I better correlated with the Hammett constants and electrophilic substituent constants than that of band II. This correlation indicates that the Hammett constants and electrophilic substituent constants, which are used to describe the trends of ground state properties, can also be used to predict the electronic transition energies.

Chapter 5

Photodissociation Action Spectroscopy of DMS-Gated Para-Aminobenzoic Acid

5.1 Introduction

In mass spectrometry (MS) studies, a long-standing and frequently cited problem is the separation and identification of protomers (*i.e.*, prototropic isomers), which are isomeric ions that differ by the position of a proton.[206,207] In many cases, MS analysis probes mixtures of the protomers since these species are often not separable by MS techniques. In some cases, multiple protomers can be observed and their relative populations manipulated by varying the ionization conditions. The individual protomers may then be studied via diagnostic ion/exchange reactions or by spectroscopic measurement with the support of computational modelling.[36,38,208,209] Recently, researchers have demonstrated the use of ion mobility for separating protomers prior to MS analysis.[210–213] One form of the ion mobility which has shown the ability to separate isomers successfully is DMS.[18,106,110,193]

The relative stability of gas-phase protonation sites, and the resulting protomer populations, of a molecule can differ from those observed in the condensed phase.[214] Para-aminobenzoic acid (PABA) is a well-known example that demonstrates the formation of two protomers whose relative populations may be tuned based on environmental conditions.[36,215,216] The neutral PABA molecule can be protonated at two sites, the amine (N) centre or the carbonyl oxygen (O) centre of the carboxylic acid group (see Figure 5.1). In solution, the amino nitrogen is the most basic site and is expected to be the site of protonation.[30,217] However, in the gas phase, the carboxylic oxygen is the most basic site and is the preferred protonation site.[217]

The choice of ESI solvent dictates which protomer of PABA is formed (see Figure 5.1).[36] Formation of O-protonated PABA is favoured with a protic electrospray solvent (*i.e.*, MeOH/H_2O), whereas the N-protomer is formed when electrosprayed from an aprotic solvent (*i.e.*, acetonitrile).[36] These results have been verified using a series of different experimental

techniques.[30,36,37,108] The preferred protonation sites of PABA in the gas phase were first reported by Kass *et al.* who monitored the population of PABA protomers in a mass spectrometer under various electrospray solvent conditions using hydrogen/deuterium exchange (HDX).[36] The two protomers were subsequently investigated by Helden *et al.* using gas-phase IR spectroscopy combined with ion mobility mass spectrometry.[30] Campbell *et al.* also successfully separated and identified the two protomers of PABA using DMS in tandem with MS and isotope exchange experiments.[37,108] Different organic-solvent-to-water ratios were examined to study the influence of MeOH and ACN on the ions' behaviour in the DMS cell.[37] With MeOH/H_2O as the electrospray solvent, the production of the O-protomer was favoured over the N-protomer. The addition of a small amount of aprotic solvent (*i.e.*, ACN) to a primarily protic solvent has no effect on the resulting protonation site wherein the ratio of the two protomers is similar to that in MeOH/H_2O. However, a large percentage ($\geqq 50\%$) of ACN in the electrospray solution does promote the formation of the N-protomer. The N- and O-protomers can be separated at CV = −1.5 V (O-protomer) and CV = −7.5 V (N-protomer) when SV is fixed at 3500 V for the DMS instrument. The CID MS fragmentation patterns for each SV/CV pair was also consistent with the presence of these two protomers. The unique ions of *m/z* 121 (loss of ammonia) for N-protomer and a large amount of ion (*m/z* 92, loss of H_2O and CO) for the O-protomer can be observed in the fragmentation spectra.
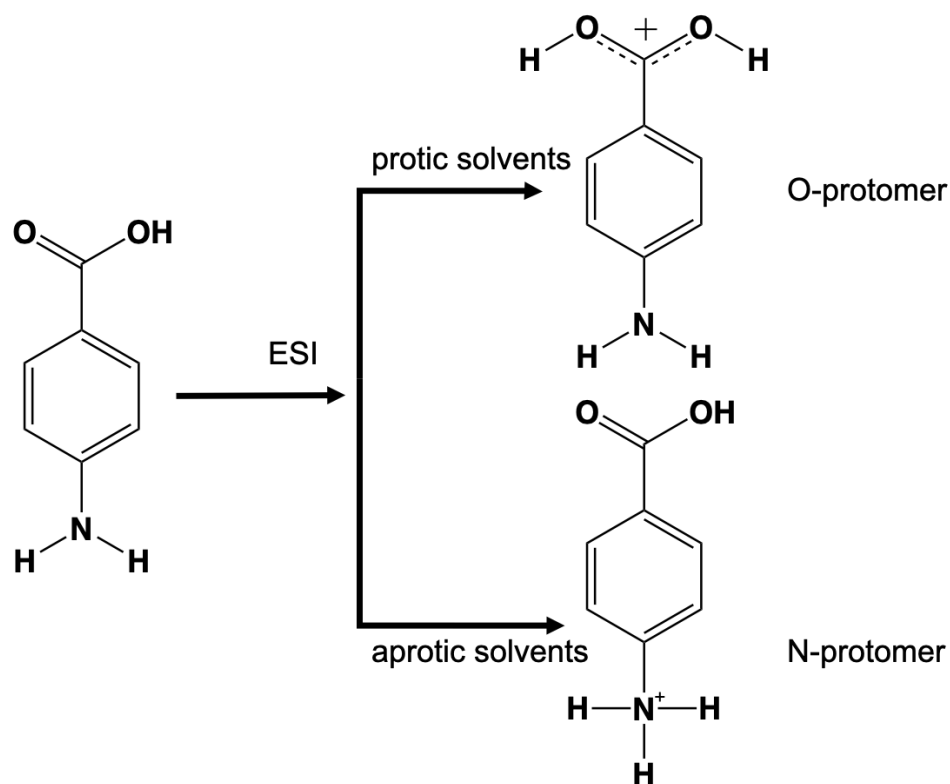
Figure 5.1 A schematic diagram showing two protonation sites for para-aminobenzoic acid

Compared to classic CID methods, a more sensitive approach for structural identification of isomers is action spectroscopy as discussed in Section 2.2.2. For action spectroscopy, the activation energy can be precisely controlled via the excitation wavelength and it probes the unique electronic structure of each molecule. Extensive studies have also been conducted on analogous systems using photoelectron spectroscopy,[153] IRMPD,[216,218,219] and UV spectroscopy,[38,220] and these works generally agree with the conclusions of Kass *et al*. Most recently, Dessent *et al*. have distinguished the two protomers of [PABA•H]⁺ via their distinct electronic spectra.[38] The UVPD action spectra were measured for protonated PABA electrosprayed from solutions of water and ACN.[38] The study revealed characteristic bands for each of the two protomers, allowing confident structural assignments from spectroscopic results. However, their spectra were possibly convoluted by contributions from each isomer,

as they had no facility to cleanly select isomer populations prior to spectroscopic interrogation.[38]

Here, we investigate the protomers of gas-phase [PABA•H]$^+$ cations using DMS-MS in tandem with UVPD action spectroscopy. With DMS alone, the protomers of protonated PABA can be resolved, although it remains difficult to identify each protomer based on their DMS trajectories alone. However, with UVPD the distinctive UV spectra of each DMS-gated protomer can be inspected. The effectiveness of this method has already been demonstrated by measuring the electronic spectra of two protomers of protonated quinoline.[105] Note also that other systems have been characterized by similar tandem ion-mobility/mass spectrometry based methods.[105] The experimental work is accompanied by Franck-Condon calculations to confirm the spectroscopic assignments, which provides some insights as to any differences between the electronic spectra of these two protomers.

5.2 Experimental and Theoretical Methods

5.2.1 Experimental Methods

The DMS-MS-UVPD experiments were conducted in a similar fashion to those described in Chapter 4 Section 4.2, which employed a modified DMS-MS device.[105,192,193] Samples of PABA were sourced from Sigma-Aldrich and used without further purification. The samples were dissolved in water, and varying ratios of protic and aprotic solvents were used including MeOH/H$_2$O, and ACN/H$_2$O, to facilitate selective protonation of either the amine or carboxyl sites during the electrospray process in accord with previous works.[37] In contrast to the method described in Chapter 4, the electrosprayed sample was DMS-selected prior to UVPD. A separation voltage (SV = 3500 V) was applied across the DMS electrodes and a static compensation voltage (CV) was scanned from −15 V to 5 V to identify the CV values promote optimal transmission of each protomer through the DMS cell. The DMS-gated PABA protomers (*m/z* 138) were transferred into a triple-quadrupole mass spectrometer where they were mass-selected in the first quadrupole (Q1) before detection by a channeltron ion detector.

Photodissociation action spectroscopy proceeded in the same fashion as for BP ions in Section 4.2.1. DMS-gated, mass-selected ions were accumulated in Q3 and then were irradiated with the output of a tunable Nd:YAG-pumped OPO system. The ions underwent photodissociation and the intensities of parent ions and any fragment ions were recorded as a function of the OPO output wavelength.

5.2.2 Theoretical Methods

Structures of protonated PABA were optimized at B3LYP/6-311++G(d,p) level of theory using Gaussian 16 revision C01.[195] Isomers with protonation at the oxygen and nitrogen atoms, as shown in Figure 5.1, were treated individually. For each protonation site, several different conformers were determined and optimized by changing the orientation of the NH_3^+ and COOH groups (N-protomer), or $C(OH)_2^+$ and NH_2 groups (O-protomer). The lowest energy structure for each protonation site was selected for further investigation. Normal mode analyses were conducted to ensure the optimized structures were local minima. The frequencies were re-calculated at the CAM-B3LYP/6-311++G(d,p) level using the quantum chemistry program package ORCA.[80] The Excited State Dynamics (ESD) program was called to calculate the absorption spectra in ORCA. Please refer to section 4.2.2 for detailed descriptions. The first six vertical excitation energies were calculated at the CAM-B3LYP/6-311++G(d,p) and STEOM-DLPNO-CCSD level of theory on ORCA 4.2.1. The vibronic contribution of the first three excited transitions was individually calculated at the CAM-B3LYP/6-311++G(d,p) level of theory and summed to create a vibronic spectrum. The energies and transition dipoles that obtained at STEOM-DLPNO-CCSD level of theory were added into the ESD module for each structure, so that ESD would not account for the values obtained by TD-DFT. The spectra of the STEOM-DLPNO-CCSD level of theory were added together to obtain the absorption spectra.

5.3 Results and Discussion

5.3.1 Differential Mobility Spectrometry

According to Campbell *et al*, electrospraying PABA from water promotes the formation of both protomers in a 10:1 (O:N) ratio.[37] DMS was used to separate mixtures of N-protomer and O-protomer electrosprayed from H_2O. Figure 5.2 shows the DMS ionograms of PABA in H_2O at two different SV values, SV = 0 V (A) and SV = 3500 V (B). As expected, there is a single peak at CV = 0 V for the SV = 0 V ionogram as shown in Figure 5.2A. In contrast, Figure 5.2B clearly shows that DMS separation is achieved at SV = 3500 V; there are two PABA ion populations at CV = −7.5 V and at CV = −2.5 V when electrosprayed from H_2O. Based on previous assignments, these populations correspond to the N-protomer (−7.5 V) and the O-protomer (−2.5 V).[37] However, in contrast to the results of Campbell *et al.* wherein nebulizing gas pressure was set to 30 psi and auxiliary gas pressure was 20 psi, the populations of the O-protomer and N-protomer are nearly equal when 0 psi nebulizing gas pressure was used in this thesis.[37]
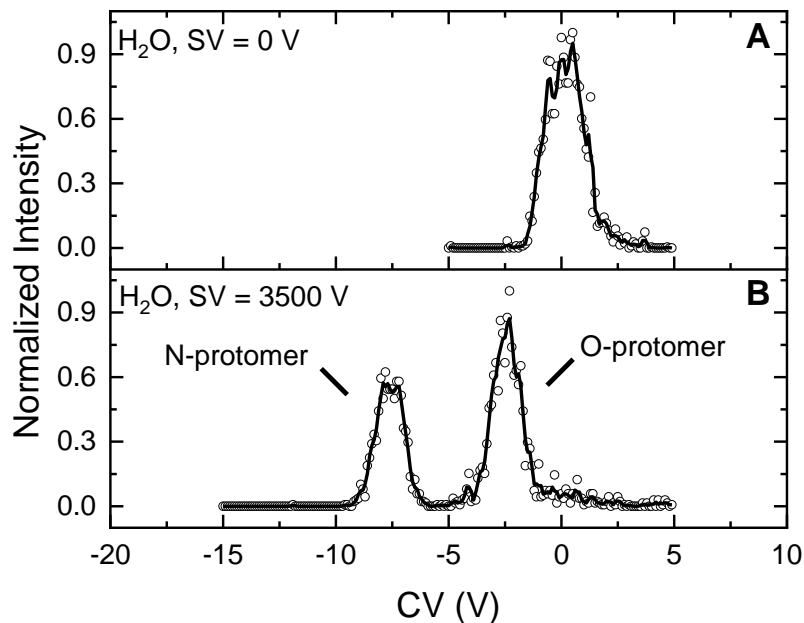
Figure 5.2 Ionograms for electrosprayed PABA dissolved in H_2O: (A) SV = 0 V and (B) SV = 3500 V. The raw data (open circles) are overlaid with a 3-point adjacent-average smoothed curve.

78

Figure 5.3 shows three ionograms for PABA, observed when electrosprayed from ACN/H_2O and MeOH/H_2O mixtures with varying solvent ratios. Figure 5.3A shows the ionogram for PABA when electrosprayed from a 50:50 mixture of acetonitrile and water. This solvent mixture leads to a 1:2 ratio of O- and N-protomer, which is similar to that observed when spraying from a pure H_2O solution. Figure 5.3B shows that when a 50:50 mixture of methanol and water is used as the ESI solvent, the O-protomer is dominant (3:1 O:N). Figure 5.3C shows that when a 95:5 mixture of ACN and H_2O is used as the ESI solvent, the N-protomer is formed almost exclusively. The populations of the separated protomeric forms of PABA in these solvents agree with the findings of Campbell *et al.*[37]
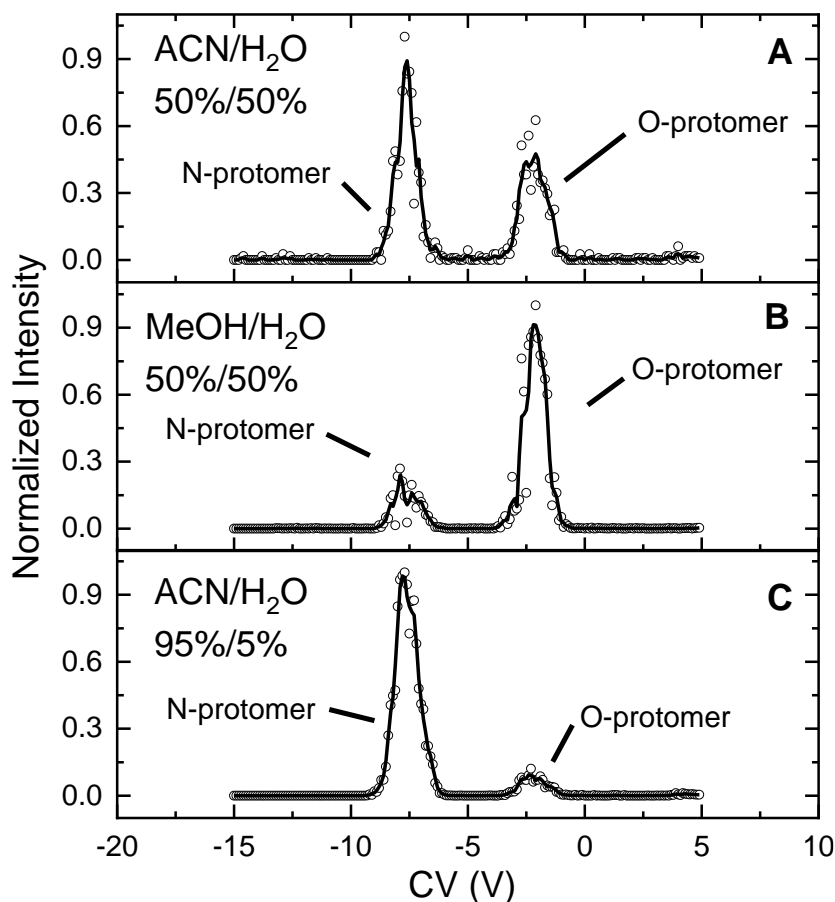
Figure 5.3 Ionograms with SV = 3500 V for electrosprayed PABA dissolved in: (A) solvent: ACN/H_2O 1:1; (B) solvent: MeOH/H_2O 1:1; (C) solvent: ACN/H_2O 95%:5%. The raw data (open circles) are overlaid with a 3-point adjacent-average smoothed curve.

5.3.2 Photodissociation Action Spectra

The mass spectra that are derived from photodissociation at the highest intensity of the fragmentation efficiency of PABA ions of O-protomer following irradiation at 350 nm (A) and N-protomer following irradiation at 220 nm (B) are shown in Figure 5.4. In Figure 5.4A, the most prominent fragments have *m/z* 94 and 120, and are associated with the loss of CO_2 and H_2O, respectively. These fragments are consistent with the losses expected for the protonated carboxylic acid group of the O-protomer. Other major fragments include *m/z* 65, 77, 92, and 93, which correspond to the production of the $C_5H_5^+$, $C_6H_5^+$, $C_6NH_6^+$, and $C_6NH_7^+$ cations, respectively. Figure 5.4B shows the photofragment mass spectrum of the N-protomer photoexcited at 220 nm. The relative intensity of the daughter ions with respect to the parent ion is low (10% fragment efficiency) in comparison to the O-protomer experiment (90% fragment efficiency), which could indicate that the N-protomer has a low oscillator strength for this particular transition compared to the one probed for the O-protomer. The red one of Figure 5.4B shows the relative intensities of the photofragments at $100 \times$ magnification. The fragments produced correspond to *m/z* 65, 75, 77, 92, 93, 94, 103, 120, and 121, which is consistent with the results of CID obtained by Campbell *et al.*[37] In contrast to the O-protomer, the N-protomer exhibits additional fragments at *m/z* 75, 103, 121. The *m/z* 75 fragment is associated with $C_6H_3^{3+}$. The *m/z* 103 fragment is associated with loss of NH_3 and water, which is caused by the initial loss of NH_3 to form the 4-dehydrobenzoic acid ion and then lose water.[37] The *m/z* 121 fragment is associated with loss of NH_3, which indicates the presence of protonated NH_2.
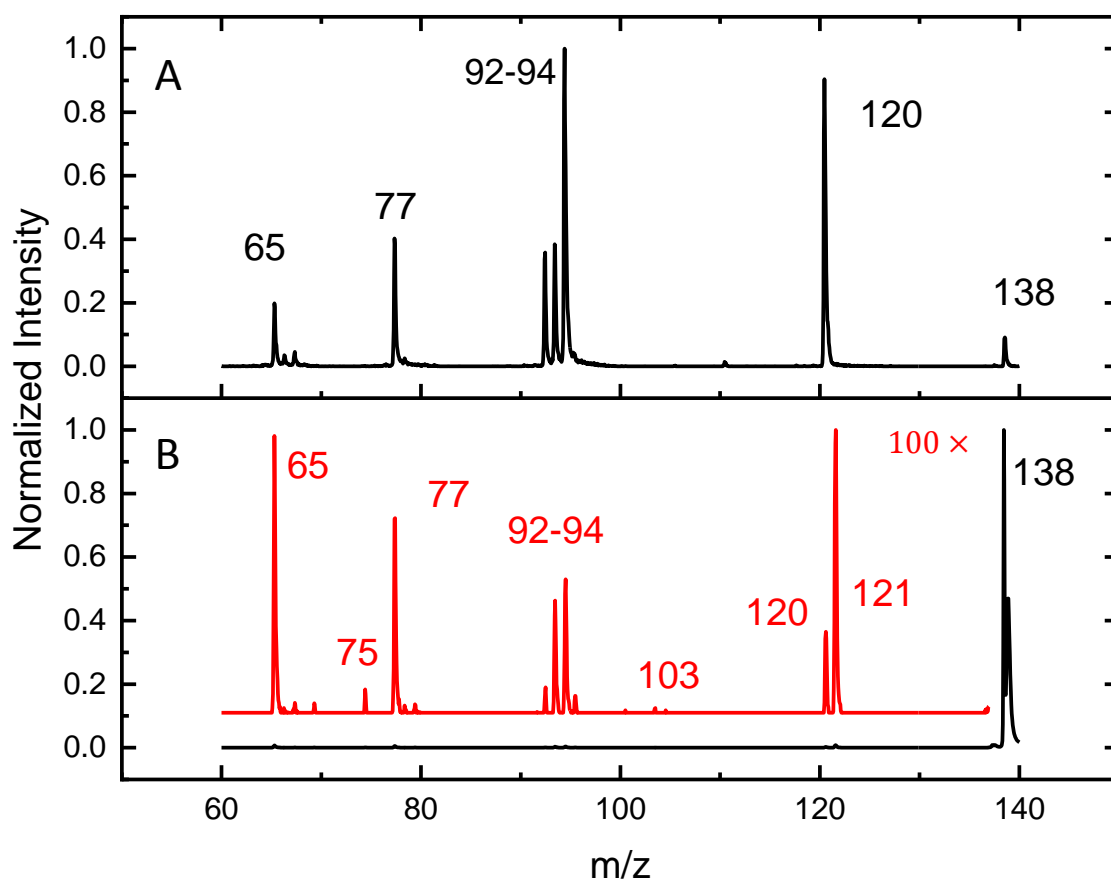
Figure 5.4 Photofragment mass spectra of (A) O-protomer excited at 350 nm; (B) N-protomer excited at 220 nm normalized to the parent ion (black) and fragment ions at 100 × magnification (red).

Edward and co-workers reported the first UV action spectra across the range 3.26-5.77 eV (380-215 nm) for the N-protomer and O-protomer in H_2O and ACN to selectively produce each protomer.[38] Based on their studies, only the O-protomer was formed when H_2O is used as the electrospray solvent. The spectrum of the O-protomer was obtained by irradiating the PABA ions in an ion trap, which were formed following electrospray ionization with H_2O as the ESI solvent. Two bands, one in the 3.4 - 4.1 eV region and a second in the 4.8 -5.4 eV region, were observed in the spectrum of O-protomer. The spectrum of the N-protomer was obtained by irradiating the PABA electrosprayed from acidified-ACN and only the fragments with m/z 121, 137, and 139 were plotted. One weak band was between 4.5 - 4.8 eV and another

band was observed with an onset 5.4 eV (the upper limit is beyond the limit of the OPO tuning range).

We also electrosprayed PABA from H_2O and measured three different photodissociation action spectra for the [PABA•H]$^+$. For each spectrum, mass-selected [PABA•H]$^+$ (m/z 138) were trapped in Q3 where they were irradiated with light from an OPO (208-400 nm). The wavelength was scanned in 1 nm intervals, and trapped ions were irradiated for 500 ms. Daughter ions arising from photodissociation of PABA ions were ejected from the trap using a mass-selective axial ejection protocol before detection at a channeltron ion detector. For PABA electrosprayed from H_2O, three spectra were recorded with the following DMS conditions: SV/CV = 0/0 V; SV/CV = 3500/−7.5 V, SV/CV = 3500/−2.5 V. When the SV and CV are both held at 0 V, the spectrum of [PABA•H]$^+$ recorded indicates that a mixture of the O-protomer and N-protomer is trapped, as expected. For SV = 3500 V, two peaks rise at CV = −7.5 V and CV = −2.5 V, corresponding to the N- and O-protomers, respectively.

Photodissociation action spectra for the major fragmentation channels of the protomers of [PABA•H]$^+$ are shown in Figure 5.5. The intensities have been normalized with respect to the total ion count and the OPO power. Figure 5.5A shows the spectrum of PABA electrosprayed from water without DMS separation and is consistent with the spectrum measured by Dessent *et al*;[38] both spectra have similar shapes and share the same two features at 3.6 eV and 5.0 eV. A small feature at around 4.5 eV is also observed in Figure 5.5A.

Figure 5.5B and C show the spectra of the DMS-isolated O- and N-protomers, respectively. The peak at approximately 5 eV in Figure 5.5B is lower in intensity than that seen in Figure 5.5A. It was initially believed that the N-protomer was isolated in Figure 5.5C; however, upon further analysis, this may not be the case. The spectrum contains features at approximately 3.6 eV, 5.6 eV, and a broad feature between 4.3-5.4 eV. The presence of the feature at 3.6 eV along with the other features, indicates that the N-protomer species is back converting to O-protomer post-separation due to H_2O vapour in the system.[108] Thus, the poor ion signal of N-protomer also leads to the noisy spectra in the range 4.4 eV-6.0 eV. The sample in C is possibly a mixture of the two protomers. A possible explanation for contamination of the trapped ion

population, even after DMS separation, is the breakdown of ion-solvent complexes due to the declustering potential (DP). The DP is applied after the DMS exit orifice before the ions enter Q1(here DP=100 V) and induces the declustering of ion-solvent complexes. If a [PABA-solvent]$^+$ complex elutes in nearby CV space to the selected PABA ion population, it is possible that the transmitted ion cloud will include the [PABA-solvent]$^+$ complex. At high DP, the solvent will be removed from the complex, allowing the previously solvated ion to pass through the mass filter in Q1. It is possible then, that the identity of the solvated ion is not the same as the selected ion population (*i.e.*, the PABA N-protomer), or that the declustering reaction could induce a proton migration and form the O-protomer since carboxylic oxygen is the preferred protonation site in the gas phase. Therefore, DMS-selection for the O-protomer and the N-protomer should be recorded with DP equal to 0 V. Comparing Figure 5.5A with the features present in spectra obtained for the isolated protomers, we conclude that the spectra shown in Figure 5.5A belong to a mixture of the O- and N-protomers with the O-protomer being more prevalent.
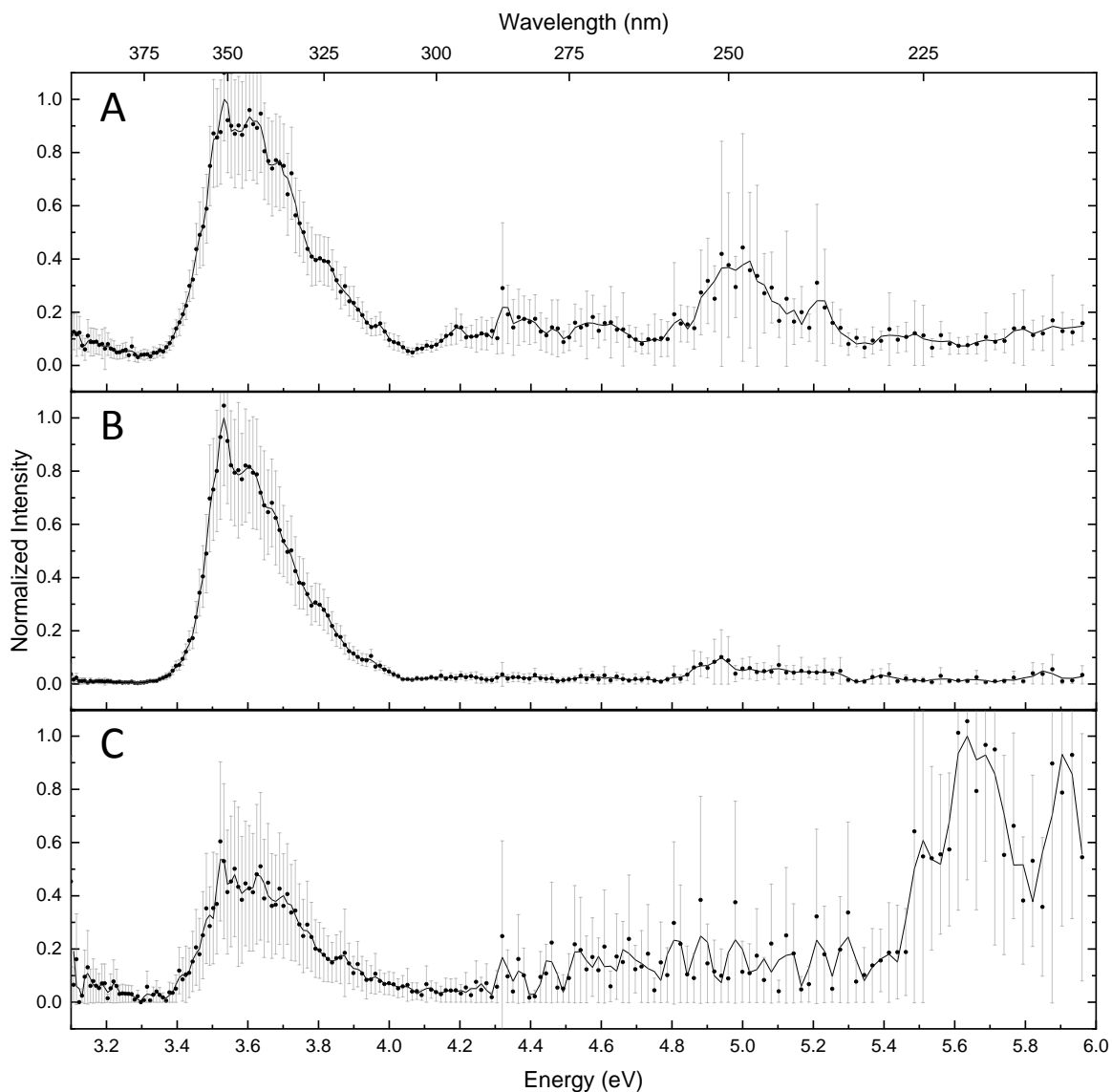
Figure 5.5 UVPD spectra of the [PABA•H]$^+$ when electrosprayed from H$_2$O. (A) SV=0 V, CV=0 V, (B) SV=3500 V, CV= −2.5 V, (C) SV=3500 V, CV= −7.5 V. The spectra represent the sum of the following fragmentation channels: A: *m/z* 65, 75, 77, 92, 93, 94, 103, 120, 121; B: *m/z* 65, 77, 92, 93, 94, 120; C: *m/z* 65, 75, 77, 92, 93, 94, 103, 120, 121. In all cases, DP = 100 V. Error bars represent +/− 0.5σ at each point, where variations in both ion intensity and laser power were considered.

Additional UVPD experiments were performed for PABA ions produced from different electrospray solvent ratios. Figure 5.6A shows the UVPD spectrum electrosprayed from an

ACN/H₂O 1:1 solution without prior DMS separation (a mixture of the two protomers, SV=0 V, CV=0 V). The spectrum recorded for the 1:1 ACN:H₂O mixture is similar to that obtained from the ion population produced by the pure water solution (Figure 5.5A). Two peaks at 3.6 eV and 5.0 eV, as well as two smaller features at approximately 4.5 eV and 5.75 eV, were observed in the spectrum. Figure 5.6B shows the spectrum for [PABA•H]⁺ when electrospraying from MeOH/H₂O 1:1 with SV = 3500 V, CV = −2.5 V and DP set to 0 V. The spectrum resembles that of the O-protomer and is consistent with that of the isolated O-protomer shown in Figure 5.5B. An intense spectral feature at 3.6 eV and a relatively weak feature at approximately 5.0 eV are again observed. Figure 5.6C shows the UVPD spectrum of the N-protomer in ACN/H₂O 95%:5% with DP =0 V. In this case, there are two spectral features associated with the N-protomer, at ~4.6 eV and ~5.6 eV, consistent with the spectrum of the N-protomer acquired by Dessent.[38] Although there is some evidence of back conversion when the region of the dominant band for the O-protomer is magnified (red line in Figure 5.6C), this shows that DMS selection and low H₂O vapour pressure (so as not to back-convert) can isolate the N-protomer for spectroscopic investigation.

The UVPD action spectra of each fragment are shown in Figure 5.7 (O-protomer) and Figure 5.8 (N-protomer). The six major fragments from O-protomer have almost the same action spectra (see Figure 5.7). However, for the N-protomer, the seven fragments have distinct action spectra. There is little amount of the fragments of *m/z* 75, and *m/z* 92 that is supposed to come from the loss of water and carbon monoxide of the O-protomer. The spectra of *m/z* 75 and 92 are very noisy, which are not shown here. The action spectra of fragments of *m/z* 65 and 103 are similar, which have feature after 5.4 eV. The action spectra of *m/z* 77, 93 and 121 have features at both 4.4-5.0 eV and 5.4-6.0 eV regions. The spectrum of *m/z* 77 has a strong absorption at both regions, but the spectrum of *m/z* 93 has much weaker absorption at 4.4-5.0 eV than that of *m/z* 77. The spectrum of *m/z* 121 has much weaker absorption at 5.4-6.0 eV than that of *m/z* 77. The action spectra of 94 and 120 have absorptions at three regions: 3.4-4.0 eV, 4.4-5.0 eV and 5.4-6.0 eV. The region of 3.4-4.0 eV belongs to O-protomer, which means these two fragments come from both protomers.
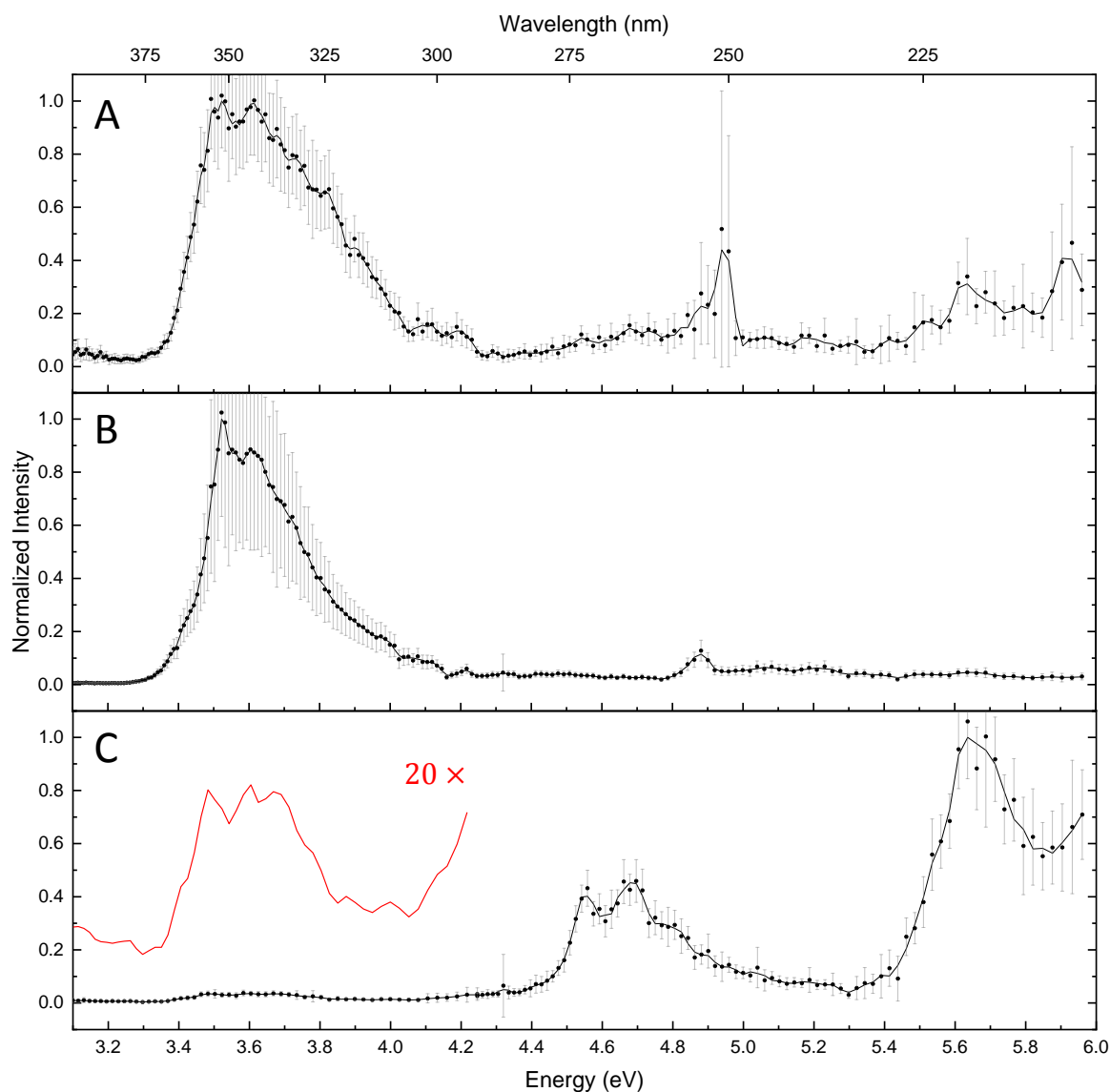
Figure 5.6 UVPD spectra of the [PABA•H]$^+$ ion electrosprayed from: (A) ACN/H$_2$O 1:1, SV=0 V, CV=0 V, DP =100 V; (B) MeOH/H$_2$O 1:1, SV=3500 V, CV= −2.5 V, DP = 0 V; (C) ACN/H$_2$O 95%:5%, SV=3500 V, CV= −7.5 V, DP = 0 V. The red line shows the 3.0-4.2 eV region at 20 × magnification. Fragmentation channels: A: *m/z* 65, 75, 77, 92, 93, 94, 103, 120, 121; B: *m/z* 65, 77, 92, 93, 94, 120; C: *m/z* 65, 75, 77, 92, 93, 94, 103, 120, 121. Error bars represent +/− 0.5σ at each point, where variations in both ion intensity and laser power were considered.
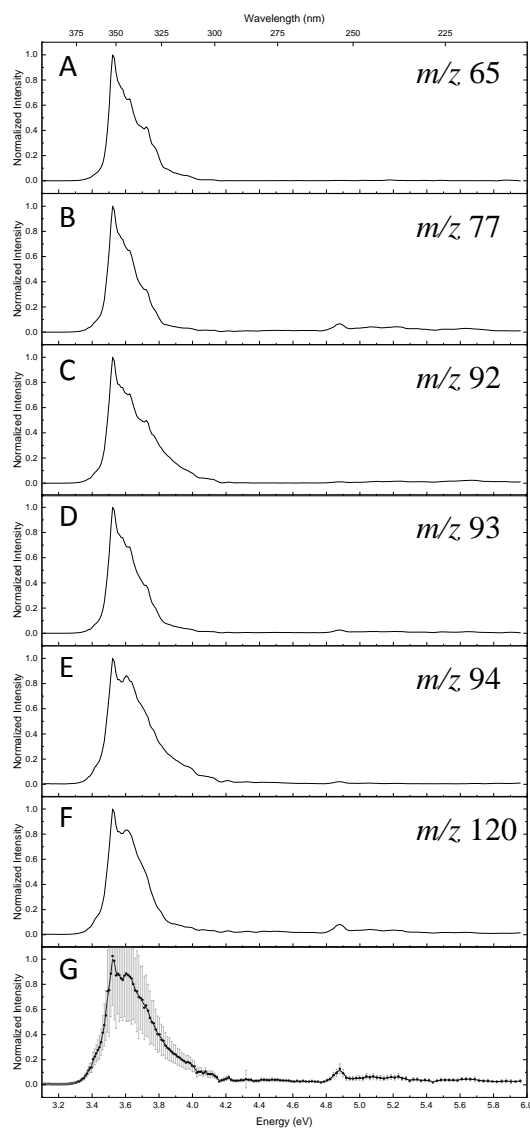
Figure 5.7 UVPD spectra of the O-protomer corresponding to the fragmentation channel: (A) *m/z* 65, (B) *m/z* 77, (C) *m/z* 92, (D) *m/z* 93, (E) *m/z* 94, (F) *m/z* 120, and (G) the sum of total fragments. Error bars represent +/− 0.5σ at each point, where variations in both ion intensity and laser power were considered.
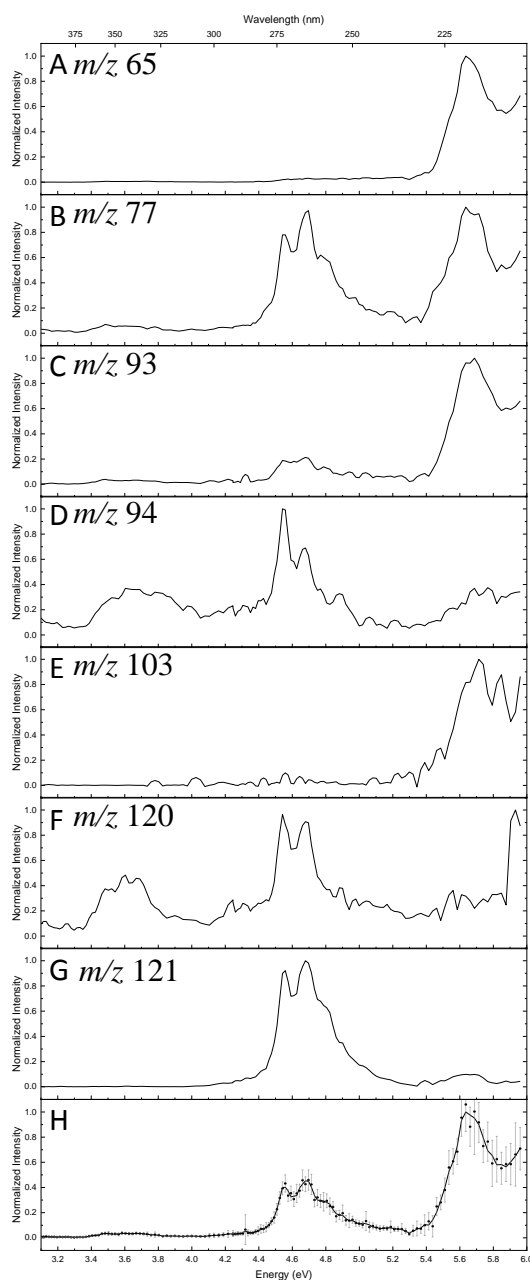
Figure 5.8 UVPD spectra of the N-protomer corresponding to the fragmentation channel: (A) *m/z* 65, (B) *m/z* 77, (C) *m/z* 93, (D) *m/z* 94, (E) *m/z* 103, (F) *m/z* 120, (G) *m/z* 121, and (H) the sum of total fragments. Error bars represent +/− 0.5σ at each point, where variations in both ion intensity and laser power were considered.

5.3.3 Simulated Electronic Spectra

Theoretical electronic absorption spectra of the O-protomer and N-protomer were calculated to compare with those measured experimentally. Table 4.2 and Table 5.2 list the calculated vertical transition energies with their oscillator strengths for the first six electronic transitions (*i.e.*, $S_0 - S_1$, through to $S_0 - S_6$) for the O- and N-protomers, respectively. For both protomers, the first three states are predicted to lie in the wavelength range accessible with our OPO system. For the O-protomer excitation energies as shown in Table 4.2, the first excited states at 4.21 eV (CAM-B3LYP) and 3.47 eV (CCSD) are the most intense transitions. Compared with the O-protomer experimental spectrum wherein the most intense peak is at 3.6 eV, the CCSD result agrees with the experimental spectrum. However, for both methods, the second excited state has slightly lower predicted energy than the experimental spectrum (4.9 eV). The third excited state is not observed in the experimental spectrum. In general, for the O-protomer, the vertical excitation energies given by CCSD were more accurate and were used for the simulation of the vibronic spectra. For the N-protomer excitation energies shown in Table 5.2, the CAM-B3LYP method predicts the first three excited states at 4.80, 5.28 and 5.84 eV with the third state stronger than the other two, which reproduces the trends observed in the experimental spectrum of the N-protomer (Figure 5.6C). In the experimental spectrum of N-protomer, 4.6 eV and 5.6 eV were observed and the peak at 5.6 eV has a higher intensity than the peak at 4.6 eV. However, the CCSD method predicts the first and second states are close in energy, and the third state at 6.19 eV has much higher energy than the one in the experimental spectrum. CAM-B3LYP is more accurate for the prediction of the transition energies of the N-protomer. In addition, the comparison of the calculated oscillator strength of the third state for the N-protomer with the first state in the O-protomer, the transition of the O-protomer is more intense than that of the N-protomer, which can be observed experimentally in the Figure 5.5A.

Table 5.1 The calculated vertical excitation energies of the O-protomer with the associated oscillator strengths (f_{osc}). The excitation energies are predicted using the CAM-B3LYP/6-311++G(d,p) and the STEOM-DLPNO-CCSD/6-311++G(d,p) level of theory.

S_1-S_6	Energy (eV)		Wavelength (nm)		f_{osc}	
	CAM-B3LYP	CCSD	CAM-B3LYP	CCSD	CAM-B3LYP	CCSD
1	4.21	3.47	294.2	356.8	0.6413646	0.7458487
2	4.69	4.37	264.2	283.5	0.0103013	0.0055783
3	5.46	5.50	226.9	225.4	0.0930116	0.1996886
4	6.19	6.24	200.4	198.7	0.0000441	0.0000395
5	6.64	6.35	186.6	195.2	0.0007608	0.0191427
6	6.67	6.61	185.9	187.4	0.0000001	0.0002165

Table 5.2 The calculated vertical excitation energies of the N-protomer with the associated oscillator strengths (f_{osc}). The excitation energies are predicted using the CAM-B3LYP/6-311++G(d,p) and the STEOM-DLPNO-CCSD/6-311++G(d,p) level of theory.

S_1-S_6	Energy (eV)		Wavelength (nm)		f_{osc}	
	CAM-B3LYP	CCSD	CAM-B3LYP	CCSD	CAM-B3LYP	CCSD
1	4.80	4.77	258.2	260.0	0.0000417	0.0002227
2	5.28	4.78	235.0	259.6	0.0217953	0.0173644
3	5.84	6.19	212.4	200.3	0.1385230	0.2910207
4	6.34	6.72	195.4	184.6	0.0000041	0.0030007
5	6.42	6.72	193.0	184.5	0.0000712	0.0000025
6	6.43	7.06	192.9	175.7	0.0008282	0.9435753

Since only the first three states are within the experimental measurement range, the first three energies and transition dipoles of the CCSD values were input directly into the ESD module. The predicted electronic absorption spectra of the O-protomers are shown in Figure 5.9. Figure 5.9A shows the simulated spectrum of the O-protomer predicted by CAM-B3LYP. Figure 5.9B shows the simulated spectrum using the CCSD method. These two methods both predicted two features, and the second features with higher energy have similar energy values. For both methods, the peaks at 5.3 eV and 5.5 eV (S_0-S_3) are blueshifted with respect to the experimental features (Figure 5.9C). The first feature at 4.2 eV (S_0-S_1) obtained by CAM-

B3LYP is also blueshifted, whereas the first feature at 3.5 eV (S_0-S_1) obtained by CCSD aligns well with the experimental spectrum.
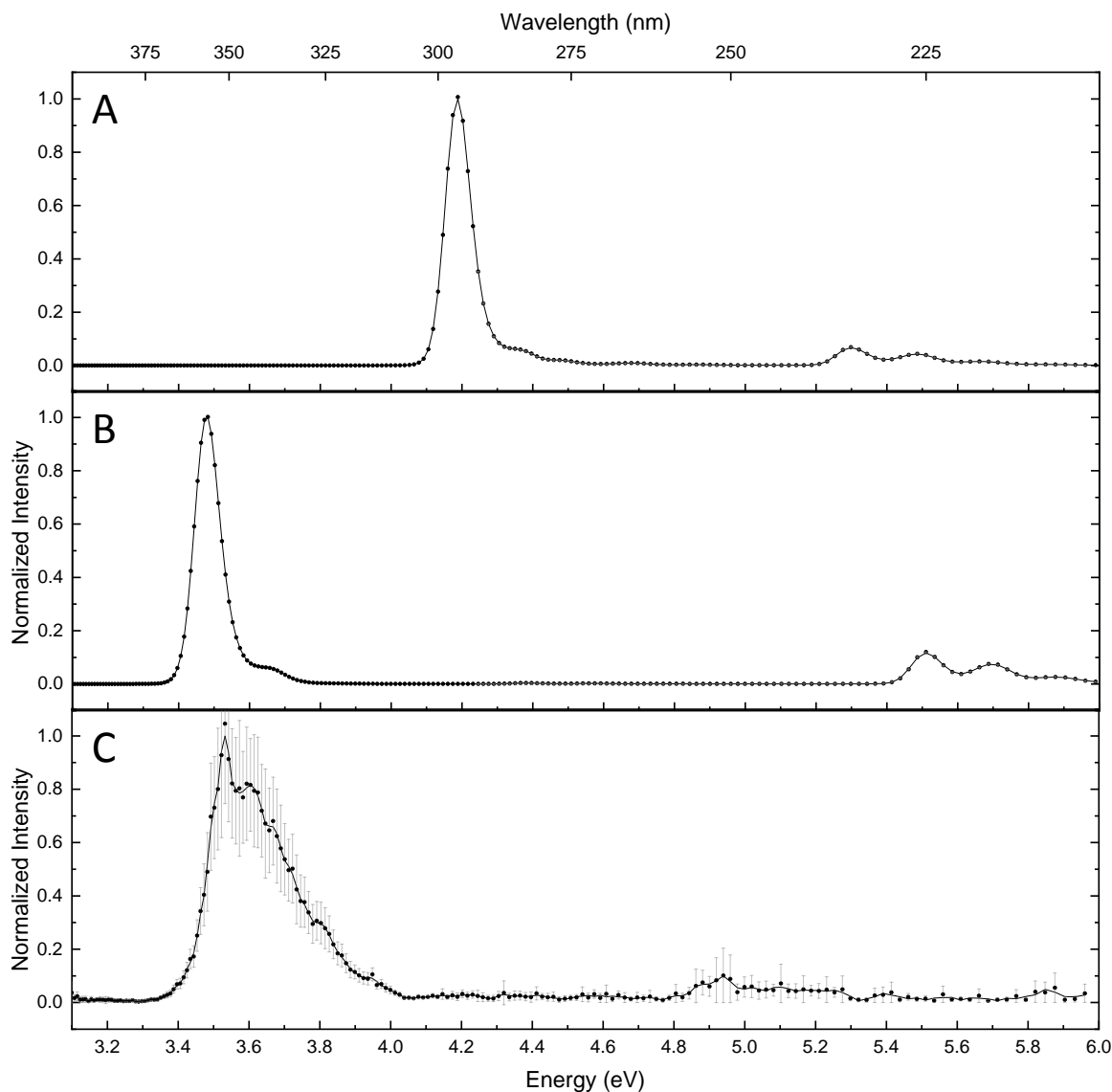
Figure 5.9 The simulated UV absorption spectra of O-protomer predicted by (A) CAM-B3LYP/6-311++G(d,p), (B) CCSD/6-311++G(d,p), and (C) experimental UVPD spectrum of the [PABA•H]$^+$ electrosprayed from H_2O, SV=3500 V, CV= −2.5 V.

Figure 5.10 shows the excitation spectra of the N-protomer calculated at the CAM-B3LYP (Figure 5.10A) and CCSD (Figure 5.10B) levels of theory, respectively. The excitation

spectrum of CAM-B3LYP (Figure 5.10A) has two features at 5.2 eV (S_0-S_2) and 5.6 eV (S_0-S_3). The first feature at 5.2 eV is blueshifted compared with the experimental spectrum (Figure 5.10C), while the second feature at 5.6 eV matches with the experimental spectrum. Figure 5.10B shows the CCSD-predicted spectrum with two peaks at 4.8 eV (S_0-S_2) and 6.2 eV (S_0-S_3). The first feature at 4.8 eV aligns well with the experimental spectrum. The second feature at 6.2 eV is blueshifted with respect to the experimental spectrum. Overall, the VG Franck Condon has correctly reproduced the shape and key trends of the experimental spectra. However, some of the predicted transition energies were blueshifted in comparison to experimental data.
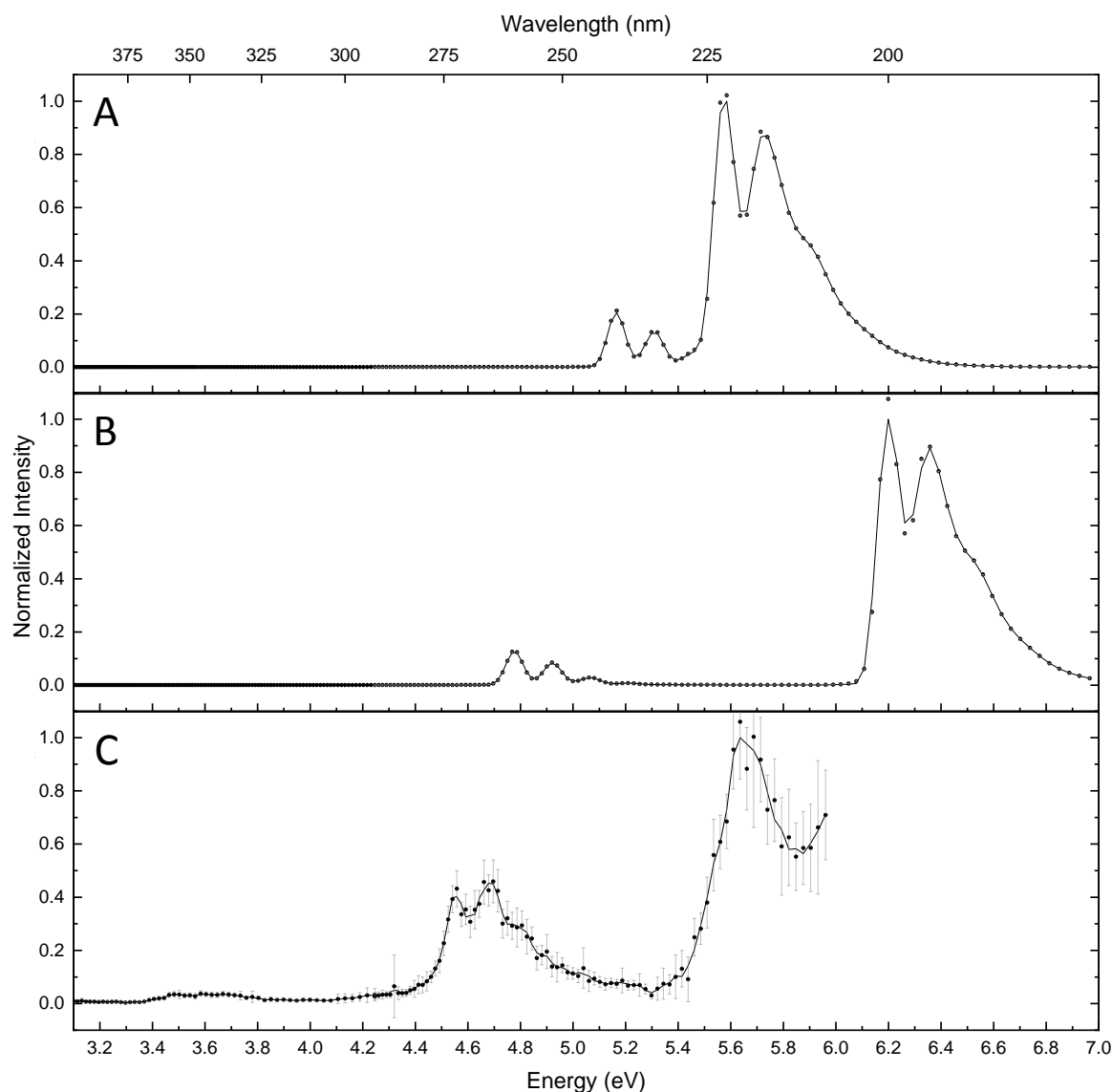
Figure 5.10 The simulated UV absorption spectra of N-protomer predicted by (A) CAM-B3LYP/6-311++G(d,p), (B) CCSD/6-311++G(d,p), and (C) experimental UVPD spectrum of the [PABA•H]⁺ electrosprayed from ACN/H₂O 95%:5%, SV=3500 V, CV= −7.5 V.

The calculated MOs with the highest contributions of the first three transitions for O-protomer and N-protomer are shown in Figure 5.11 and Figure 5.12. These transitions are $\pi \rightarrow \pi^*$ transitions wherein electron density.
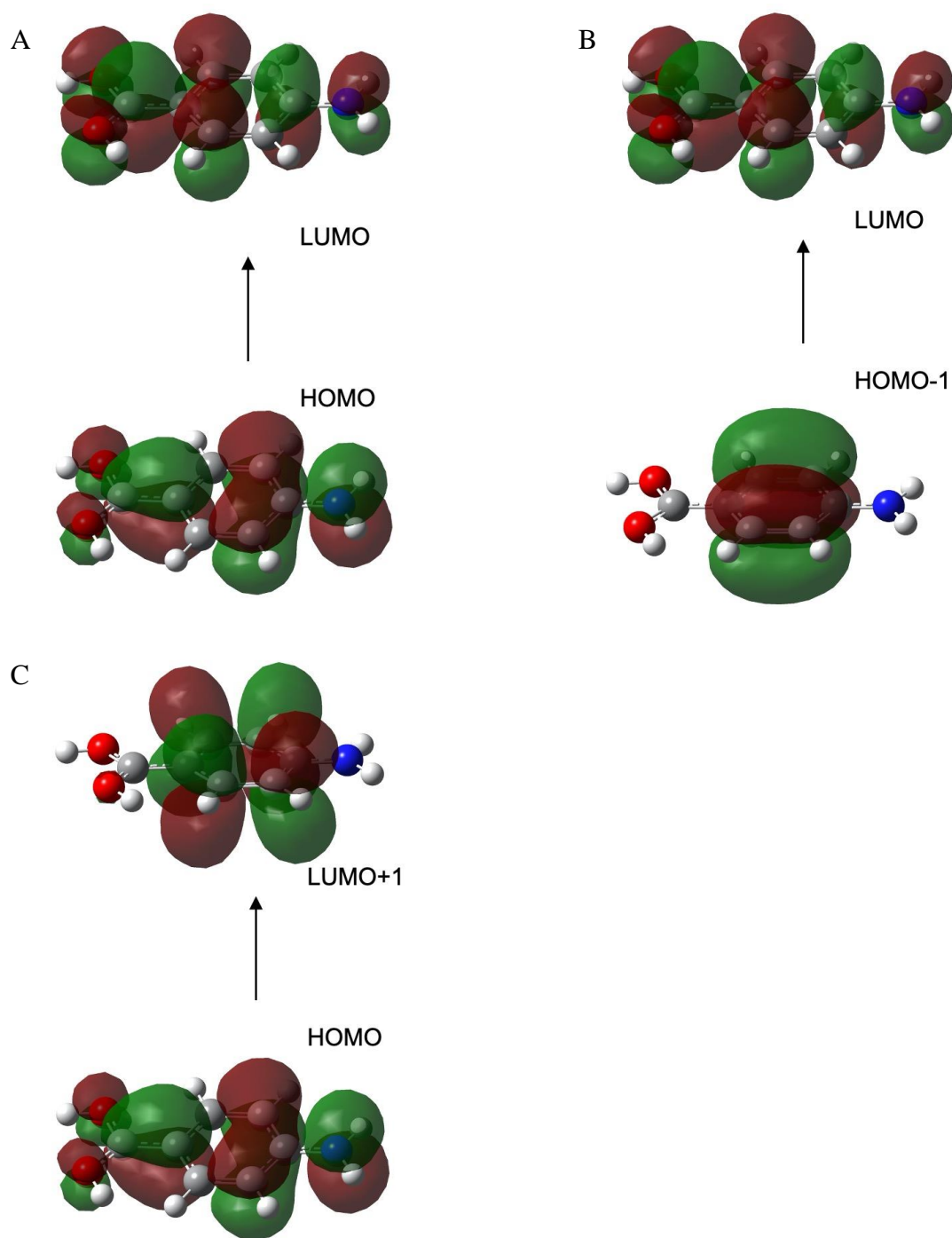
Figure 5.11 The calculated molecular orbitals of O-protomer for transitions with percentage contributions: (A) S_0 to S_1 (95.4%), (B) S_0 to S_2 (83.3%), and (C) S_0 to S_3 (82.1%). S_0 is the ground state, and S_n is the nth excited state.
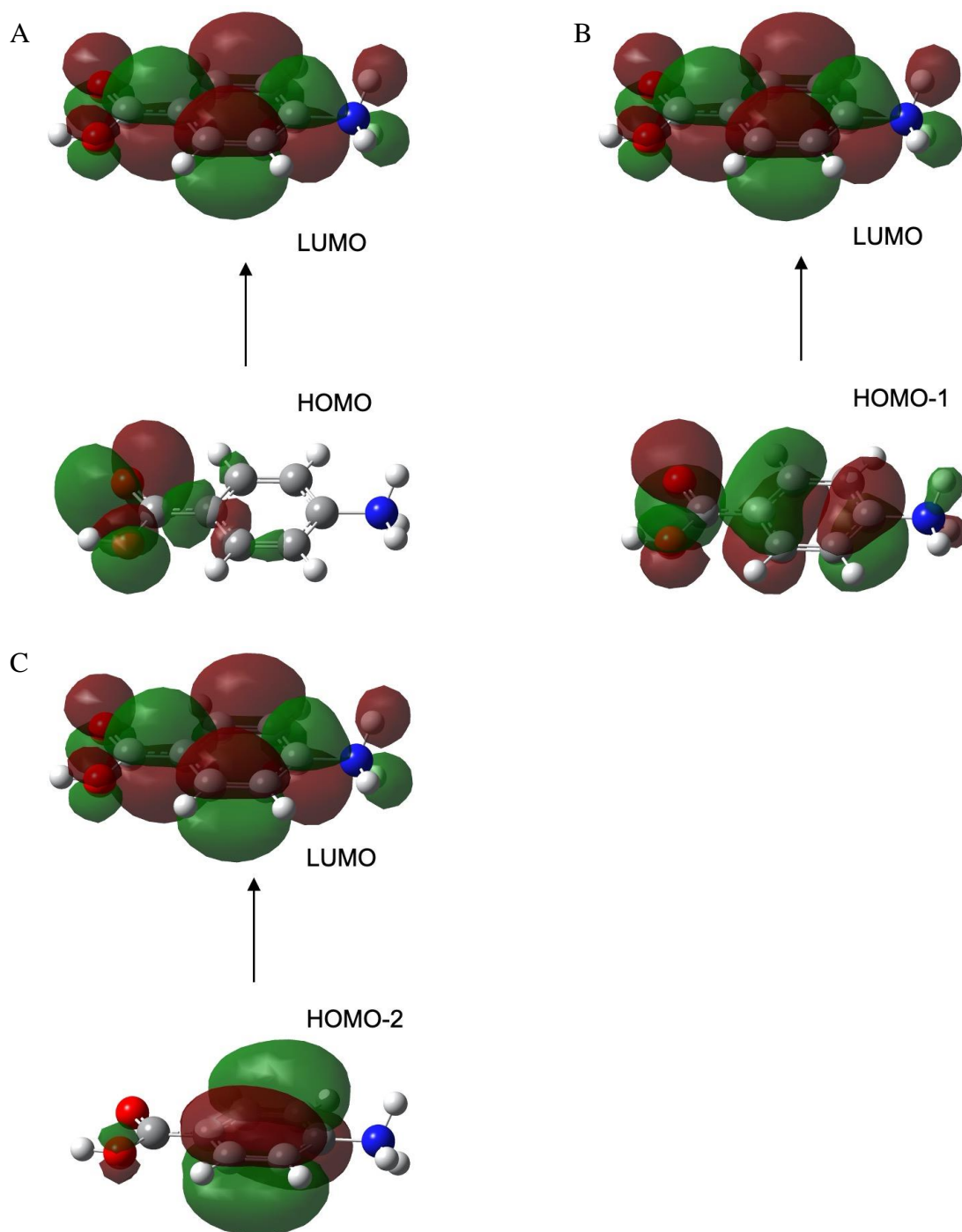
Figure 5.12 The calculated molecular orbitals of N-protomer for transitions with percentage contributions: (A) S_0 to S_1 (81.5%), (B) S_0 to S_2 (62.4%), and (C) S_0 to S_3 (71.1%). S_0 is the ground state, and S_n is the nth excited state.

5.4 Conclusions

Here, we demonstrate the use of DMS to separate the two protomers of protonated PABA for subsequent characterization via UVPD spectroscopy. The distinctly different electronic structures of the two protomers are manifested in the action spectra of the O- and N-protomers. The fragment at *m/z* 121 is associated with the loss of NH_3, which is only observed in the mass spectrum of N-protomer. The UVPD action spectrum of the N-protomer in ACN/H_2O is consistent with the previous work by Dessent et al.[38] However, the conversion of the N-protomer to O-protomer following selection happened in the sample electrosprayed from H_2O, which requires further investigations. The fragments observed from N-protomer exhibit different UVPD action spectra. For the O-protomer, the second feature has a much weaker intensity compared to Dessent's results. The differences that we observe suggests that the photodepletion spectrum of the O-protomer reported in Dessent's study is likely to be the mixture of the two protomers. Compared to conventional UVPD techniques, our method can provide an orthogonal degree of characterization and separation, which provides additional certainty in isomer assignments.

For the simulated electronic spectra, the general trends can be predicted by using the VG-FC method. These transitions are $\pi \rightarrow \pi^*$ transitions wherein electron density. However, the accuracy of the excitation energies can still be improved. The adiabatic Franck-Condon, in which the Hessians of the ground state and excited state are both required, can be used to provide a better precision towards experimental results.

Chapter 6
Conclusions

In this thesis, several aromatic clusters have been investigated using combined experimental and theoretical approaches. The IRMPD was used to measure the IR spectra of the probing ensemble. The DMS has been modified to couple with UVPD, which was able to separate tautomers prior to measuring the UVPD action spectra. The geometries of the clusters were found via BH and then treated with B3LYP/6-311++G(d,p) level of theory. The validity of the structures can be assessed by comparing the calculated IR spectra with the IRMPD spectra. The electronic spectra were simulated through the VG-FC method. With successful validation between the experimental and calculated results, additional properties can be explored, which provides further understanding of the behaviours of the clusters.

In Chapter 3, new methods that rely on techniques from machine learning were developed to interpret and assign spectra. A computational study of the proton-bound heterodimer of Phe/Ser was conducted. The low-energy isomers of the protonated Phe-Ser dimer came from the exhaustive search of the potential energy surface (PES) using the custom-written BH search algorithm, and 37 isomers are identified within $180 \, \text{kJ} \cdot \text{mol}^{-1}$ of the global-minimum structure. By calculating the mass-weighted distance matrix of each isomer, the cluster structures were categorized using hierarchical clustering to partition the PES in terms of nuclear configuration. Their calculated IR spectra are then compared with the experimental IR spectra recorded by Lorenz and Rizzo[32] by converting the spectra to unique matrices. The cosine distance between the experimental spectral matrices and calculated spectral matrices were calculated to facilitate the spectral assignment. The five experimental UV/IR double-resonance spectra were assigned to five different isomers. In addition, the binding motifs of each experimental spectrum were identified to show the structures populated in the electrospray ionization process.

By applying hierarchical clustering and spectral comparison that we developed for Phe-Ser heterodimer, proton-bound homodimer of Phe derivatives were studied experimentally and computationally. The computed structures obtained via BH were treated with hierarchical

clustering to partition the PES in terms of intermolecular binding motifs. MDS was used to visualize how chemical substitution of the Phe derivatives influenced homodimer bindings and geometries. The quantum theory of atoms in molecules (QTAIM) were used to analyze the electron topology, and quantitatively and qualitatively reveal the non-covalent interactions in the dimer system. Homodimer geometries were found to be influenced predominantly by hydrogen-bonding interactions. The calculated IR spectra were compared to the IRMPD spectra using cosine distance metrics to determine which geometric conformers/isomers were present in the probed ensemble.

In Chapter 4, the newly implemented UVPD spectroscopy study of BP and several of its derivatives has been conducted, which demonstrates the validity and sensitivity of our DMS-MS instrument for spectroscopy experiments. The absorption energies obtained experimentally were found to vary depending on the substituents. The correlation between band positions and the Hammett constants and electrophilic substituent constants for chemical substituents has been identified, which shows the effect of EWGs and EDGs on phenyl ring electronic stability. The VG-FC method was used to simulate electronic absorption spectra, which facilitates the assignment of the observed transitions to calculated transitions.

In Chapter 5, the newly implemented UVPD capabilities are tested in the context of DMS separation for two prototropic isomers of protonated PABA. The protonated PABA has two protomers: O-protomer (protonated at carbonyl oxygen) and N-protomer (protonated at amine nitrogen). The population of these two protomers varies depending on the solvent conditions. We show that distinct UVPD spectra are observed for each protomer and that conversion from one protomer to the other can occur post-DMS separation in environments that have high partial pressures of water. The VG-FC approach was also used to simulate the electronic spectra.

Taken as a whole, this thesis presents advances for experimental characterization and data analysis/assignment of ionic clusters that contain aromatic groups. These techniques can, of course, be extended to non-aromatic systems; the choice of aromatic species simplified experimentation owing to their spectral signatures, which appears in convenient spectral

regions. The study of aromatic clusters also provides opportunities to explore the effects of substituents on the structures and properties of clusters. These studies were carried out in a combined experimental and computational method. Experimentally, IRMPD can be used to verify the structures. UVPD can be used to study the electronic properties of the probed clusters. DMS coupled with UVPD provides the ability to distinguish and identify DMS-separated ions, and it can also reduce the spectral congestion when studying the mixtures of isomeric species. Computationally, we developed an efficient method to partition the PES and assign the spectra quantitatively when dealing with a large number of structures. The VG-FC can be used to simulate the electronic absorption spectra and assign the experimental transitions.

Many future directions can be explored for the investigations discussed herein. The matching algorithm discussed in Chapter 3 can be improved by using another distance metric. The current spectral comparison employed the intensities of the spectra as the unique matrix representing each spectrum. However, IRMPD is more reliable with respect to the wavelength of the peaks than the intensities. Future comparison can be adjusted to increase the proportion of the wavelength factor. In Chapter 4, BP ions can produce tropylium after fragment. Whether the rearrangement of the benzyl cation happened before or after the fragmentation still needs further investigations.

References

1.   Berry, R. S. Phases and Phase Changes of Small Systems. in *Theory of Atomic and Molecular Clusters* (ed. Julius Jellinek) vol. 430 1–26 (Springer-Verlag Berlin Heidelberg, 1999).

2.   Luo, Z. & Castleman, A. W. Special and general superatoms. *Acc. Chem. Res.* **47**, 2931–2940 (2014).

3.   Bačić, Z. & Miller, R. E. Molecular clusters: Structure and dynamics of weakly bound systems. *J. Phys. Chem.* **100**, 12945–12959 (1996).

4.   Castleman, A. W. & Keesee, R. G. Gas-phase clusters: Spanning the states of matter. *Science (80-.).* **241**, 36–42 (1988).

5.   Claridge, S. A. *et al.* Cluster-Assembled Materials. *ACS Nano* **3**, 244–255 (2009).

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

7.   Bondybey, V. E. & Beyer, M. K. How many molecules make a solution? *International Reviews in Physical Chemistry* vol. 21 277–306 (2002).

8.   Alexandrova, A. N., Boldyrev, A. I., Zhai, H. J. & Wang, L. S. All-boron aromatic clusters as potential new inorganic ligands and building blocks in chemistry. *Coordination Chemistry Reviews* vol. 250 2811–2866 (2006).

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

10.   Buck, U. & Huisken, F. Infrared spectroscopy of size-selected water and methanol clusters. *Chem. Rev.* **100**, 3863–3890 (2000).

11.   Aleese, L. Mac *et al.* Mid-IR spectroscopy of protonated leucine methyl ester performed with an FTICR or a Paul type ion-trap. *Int. J. Mass Spectrom.* **249–250**, 14–20 (2006).

12.   Fu, W. *et al.* The Structures of Proton-bound Dimers of Glycine with Phenylalanine and Pentafluorophenylalanine. *J. Mol. Spectrosc.* **330**, 194–199 (2016).

13.   Wu, R. & McMahon, T. B. Infrared multiple photon dissociation spectra of proline and glycine proton-bound homodimers. Evidence for zwitterionic structure. *J. Am. Chem. Soc.* **129**, 4864–4865 (2007).

14.   Pereverzev, A. Y. *et al.* Spectroscopic Evidence for Peptide-Bond-Selective Ultraviolet Photodissociation. *J. Phys. Chem. Lett* **11**, 17 (2020).

15.   Mistarz, U. H. *et al.* UV Photodissociation Mass Spectrometry Accurately Localize Sites of Backbone Deuteration in Peptides. *Anal. Chem.* **90**, 1077–1080 (2018).

16.   Hopkins, W. S., Hamilton, S. M., McNaughter, P. D. & Mackenzie, S. R. VUV photodissociation dynamics of diatomic gold, Au2: A velocity map imaging study at 157 nm. *Chem. Phys. Lett.* **483**, 10–15 (2009).

17.   Hopkins, W. S., Lipciuc, M. L., Gardiner, S. H. & Vallance, C. RG+ formation following photolysis of NO-RG via the Ã-X̃ transition: A velocity map imaging study. *J. Chem. Phys.* **135**, 034308 (2011).

18.   Walker, S. W. C. *et al.* Determining molecular properties with differential mobility spectrometry and machine learning. doi:10.1038/s41467-018-07616-w.

19. Barnett, D. A., Ells, B., Guevremont, R. & Purves, R. W. Separation of leucine and isoleucine by electrospray ionization-high field asymmetric waveform ion mobility spectrometry-mass spectrometry. *J. Am. Soc. Mass Spectrom.* **10**, 1279–1284 (1999).

20. Schneider, B. B., Covey, T. R., Coy, S. L., Krylov, E. V. & Nazarov, E. G. Control of Chemical Effects in the Separation Process of a Differential Mobility Mass Spectrometer System. *Eur. J. Mass Spectrom.* **16**, 57–71 (2010).

21. Ho, K. M. *et al.* Structures of medium-sized silicon clusters. *Nature* **392**, 582–585 (1998).

22. Furche, F. *et al.* The structures of small gold cluster anions as determined by a combination of ion mobility measurements and density functional calculations. *J. Chem. Phys.* **117**, 6982–6990 (2002).

23. Burley, S. K. & Petsko, G. A. Amino-aromatic interactions in proteins. *FEBS Lett.* **203**, 139–143 (1986).

24. Dougherty, D. A. Cation-π Interactions in Chemistry and Biology: A New View of Benzene, Phe, Tyr, and Trp. *Science (80-.).* **271**, 163–168 (1996).

25. Anderson, D. E., Hurley, J. H., Nicholson, H., Baase, W. A. & Matthews, B. W. Hydrophobic core repacking and aromatic-aromatic interaction in the thermostable mutant of T4 lysozyme ser 117 → phe. *Protein Sci.* **2**, 1285–1290 (1993).

26. Kannan, N. & Vishveshwara, S. Aromatic clusters: a determinant of thermal stability of thermophilic proteins. *Protein Eng.* **13**, 753–761 (2000).

27. Parthasarathy, A. *et al.* A Three-Ring Circus: Metabolism of the Three Proteogenic Aromatic Amino Acids and Their Role in the Health of Plants and Animals. *Front. Mol. Biosci.* **5**, 29 (2018).

28. Dougherty, D. A. Cation-π Interactions Involving Aromatic Amino Acids. *J. Nutr.* **137**, 1504S-1508S (2007).

29. Hunter, C. A., Low, C. M. R., Rotger, C., Vinter, J. G. & Zonta, C. Substituent effects on cation-π interactions: A quantitative study. *Proc. Natl. Acad. Sci.* **99**, 4873–4876 (2002).

30. Seo, J. *et al.* The impact of environment and resonance effects on the site of protonation of aminobenzoic acid derivatives. *Phys. Chem. Chem. Phys.* **18**, 25474–25482 (2016).

31. Lodish, H. *et al.* *Molecular Cell Biology. 4th edition. New York: W. H. Freeman.* (W. H. Freeman, 2000).

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

33. Fu, W. *et al.* Intramolecular cation-π interactions in protonated phenylalanine derivatives †. *Phys. Chem. Chem. Phys* **19**, 729 (2017).

34. Barylyuk, K. V., Chingin, K., Balabin, R. M. & Zenobi, R. Fragmentation of Benzylpyridinium 'Thermometer' Ions and Its Effect on the Accuracy of Internal Energy Calibration. *J. Am. Soc. Mass Spectrom.* **21**, 172–177 (2010).

35. Rahrt, R., Auth, T., Demireva, M., Armentrout, P. B. & Koszinowski, K. Benzhydrylpyridinium Ions: A New Class of Thermometer Ions for the Characterization of Electrospray-Ionization Mass Spectrometers. *Anal. Chem.* **91**, 11703–11711 (2019).

36. Tian, Z. & Kass, S. R. Gas-phase versus liquid-phase structures by electrospray ionization mass spectrometry. *Angew. Chemie - Int. Ed.* **48**, 1321–1323 (2009).

37. Campbell, J. L., Blanc, J. Le & Schneider, B. Probing electrospray ionization dynamics using differential mobility spectrometry: the curious case of 4-aminobenzoic acid. *Anal. Chem.* **84**, 7857–7864 (2012).

38. 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).

39. Lecours, M. J., Chow, W. C. T. & Hopkins, W. S. Density Functional Theory Study of Rh n S 0,± and Rh n+1 0,± (n = 1−9). *J. Phys. Chem. A* **118**, 26 (2014).

40. Zhou, C., Ieritano, C. & Hopkins, W. S. Augmenting Basin-Hopping With Techniques From Unsupervised Machine Learning: Applications in Spectroscopy and Ion Mobility. *Front. Chem.* **7**, 519 (2019).

41. 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).

42. Wales, D. J. & Scheraga, H. A. Global optimization of clusters, crystals, and biomolecules. *Science* **285**, 1368–1372 (1999).

43. Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., et al. Gaussian 09, Revision A.02. Gaussian, Inc.

44. Kim, K. & Jordan, K. D. Comparison of density functional and MP2 calculations on the water monomer and dimer. *Journal of Physical Chemistry* vol. 98 10089–10094 (1994).

45. Stephens, P. J., Devlin, F. J., Chabalowski, C. F. & Frisch, M. J. Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields. *J. Phys. Chem.* **98**, 11623–11627 (1994).

46. Becke, A. D. A new mixing of Hartree-Fock and local density-functional theories. *J. Chem. Phys.* **98**, 1372–1377 (1993).

47. Lee, C., Yang, W. & Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **37**, 785–789 (1988).

48. Cramer, C. J. *Essentials of Computational Chemistry Theories and Models*. *Essentials of Computational Chemistry* vol. 42 (2004).

49. Vanommeslaeghe, K., Guvench, O. & MacKerell, A. D. Molecular Mechanics. *Curr. Pharm. Des.* **20**, 3281–3292 (2014).

50. Hehre, W. J. *A Guide to Molecular Mechanics and Quantum Chemical Calculations*. (2003).

51. Rappe, A. K., Casewit, C. J., Colwell, K. S., Goddard III, 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).

52. Chirlian, L. E. & Francl, M. M. Atomic charges derived from electrostatic potentials: A detailed study. *J. Comput. Chem.* **8**, 894–905 (1987).

53. 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).

54.   Day, W. H. E. & Edelsbrunner, H. Efficient algorithms for agglomerative hierarchical clustering methods. *J. Classif.* **1**, 7–24 (1984).

55.   Michener, C. D. & Sokal, R. R. A Quantitative Approach to a Problem in Classification. *Evolution (N. Y).* **11**, 130–162 (1957).

56.   Sokal, R. R. & Michener, C. D. A statistical method for evaluating systematic relationships. *Univ. Kansas Sci. Bull.* **38**, 1409–1438 (1958).

57.   Wickelmaier, F. *An Introduction to MDS*. (Aalborg Universitetsforlag, 2003).

58.   Borg, I. & Groenen, P. J. F. *Modern Multidimensional Scaling*. *Modern Multidimensional Scaling* (Springer New York, 2005). doi:10.1007/0-387-28981-x.

59.   Andersson, M. P. & Uvdal, P. New scale factors for harmonic vibrational frequencies using the B3LYP density functional method with the triple-zeta basis set 6-311+G(d,p). *J. Phys. Chem. A* **109**, 2937–41 (2005).

60.   Bader, R. F. W. & Essén, H. The characterization of atomic interactions. *J. Chem. Phys.* **80**, 1943–1960 (1984).

61.   Shyam, P., Kumar, V., Raghavendra, V. & Subramanian, V. Bader's Theory of Atoms in Molecules (AIM) and its Applications to Chemical Bonding. *Indian Acad. Sci. Spec. Issue Chem. Bond.* **128**, 1527–1536 (2016).

62.   R. F. W. Bader. *Atoms in Molecules: A Quantum Theory, Oxford University Press*. (Clarendon Press, 1990).

63.   Bader, R. F. W. W. Bond paths are not chemical bonds. *J. Phys. Chem. A* **113**, 10391–10396 (2009).

64.   Parr, R. G., Ayers, P. W. & Nalewajski, R. F. What is an atom in a molecule? *J. Phys. Chem. A* **109**, 3957–3959 (2005).

65.   Bader, R. F. W. A quantum theory of molecular structure and its applications. *Chem. Rev.* **91**, 893–928 (1991).

66.   Mohajeri, A. & Karimi, E. AIM and NBO analyses of cation-π interaction. *J. Mol. Struct. THEOCHEM* **774**, 71–76 (2006).

67.   Kheirjou, S., Fattahi, A. & Hashemi, M. M. The intramolecular cation-π interaction of some aryl amines and its drastic influence on the basicity of them: AIM and NBO analysis. *Comput. Theor. Chem.* **1036**, 51–60 (2014).

68.   Lane, J. R., Contreras-García, J., Piquemal, J. P., Miller, B. J. & Kjaergaard, H. G. Are bond critical points really critical for hydrogen bonding? *J. Chem. Theory Comput.* **9**, 3263–3266 (2013).

69.   Rosenberg, R. E. Can 2-X-Ethanols Form Intramolecular Hydrogen Bonds? *J. Phys. Chem. A* **123**, 29 (2019).

70.   Johnson, E. R. *et al.* Revealing Noncovalent Interactions. *J. Am. Chem. Soc.* **132**, 6498–6506 (2010).

71.   Saleh, G., Gatti, C. & Lo Presti, L. Non-covalent interaction via the reduced density gradient: Independent atom model vs experimental multipolar electron densities. *Comput. Theor. Chem.* **998**, 148–163 (2012).

72.   Boto, R. A., Contreras-García, J., Tierny, J. & Piquemal, J. P. Interpretation of the reduced density gradient. *Mol. Phys.* **114**, 1406–1414 (2016).

73.   Norman, P., Ruud, K. & Saue, T. *Principles and practices of molecular properties: Theory, modeling and simulations*. *Principles and Practices of Molecular Properties:*

Theory, Modeling and Simulations (wiley, 2018). doi:10.1002/9781118794821.

74.   Avila Ferrer, F. J. & Santoro, F. Comparison of vertical and adiabatic harmonic approaches for the calculation of the vibrational structure of electronic spectra. *Phys. Chem. Chem. Phys.* **14**, 13549–13563 (2012).

75.   Götze, J. P., Karasulu, B. & Thiel, W. Computing UV/vis spectra from the adiabatic and vertical Franck-Condon schemes with the use of Cartesian and internal coordinates. *J. Chem. Phys.* **139**, (2013).

76.   Li, S. L. & Truhlar, D. G. Franck-Condon Models for Simulating the Band Shape of Electronic Absorption Spectra. *J. Chem. Theory Comput.* **13**, 2823–2830 (2017).

77.   Hazra, A., Chang, H. H. & Nooijen, M. First principles simulation of the UV absorption spectrum of ethylene using the vertical Franck-Condon approach. *J. Chem. Phys.* **121**, 2125–2136 (2004).

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

79.   Domcke, W., Cederbaum, L. S., Köppel, H. & Von Niessen, W. A comparison of different approaches to the calculation of franck-condon factors for polyatomic molecules. *Mol. Phys.* **34**, 1759–1770 (1977).

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

81.   Cismesia, A. P., Bell, M. R., Tesler, L. F., Alves, M. & Polfer, N. C. Infrared ion spectroscopy: An analytical tool for the study of metabolites. *Analyst* **143**, 1615–1623 (2018).

82.   Wu, R. & McMahon, T. B. Protonation sites and conformations of peptides of glycine (Gly 1-5H+) by IRMPD spectroscopy. *J. Phys. Chem. B* **113**, 8767–8775 (2009).

83.   Dunbar, R. C. & McMahon, T. B. Activation of Unimolecular Reactions by Ambient Blackbody Radiation. *Science (80-.).* **279**, (1998).

84.   Gruene, P. *et al.* Structures of neutral Au7, Au19, and Au20 clusters in the gas phase. *Science (80-.).* **321**, 674–676 (2008).

85.   Martens, J. K. *et al.* Globule to helix transition in sodiated polyalanines. *J. Phys. Chem. Lett.* **3**, 3320–3324 (2012).

86.   Wu, R., Marta, R. A., Martens, J. K., Eldridge, K. R. & McMahon, T. B. Experimental and theoretical investigation of the proton-bound dimer of lysine. *J. Am. Soc. Mass Spectrom.* **22**, 1651–1659 (2011).

87.   Ziegler, B. E., Marta, R. A., Martens, S. M., Martens, J. K. & McMahon, T. B. Structure, energetics and vibrational spectra of protonated chlortetracycline in the gas phase: An experimental and computational investigation. *Int. J. Mass Spectrom.* **316–318**, 117–125 (2012).

88.   Oomens, J., Sartakov, B. G., Meijer, G. & von Helden, G. Gas-phase infrared multiple photon dissociation spectroscopy of mass-selected molecular ions. *Int. J. Mass Spectrom.* **254**, 1–19 (2006).

89.   Lehmann, K. K., Scoles, G. & Pate, B. H. Intramolecular Dynamics from Eigenstate-Resolved Infrared Spectra. *Annu. Rev. Phys. Chem.* **45**, 241–274 (1994).

90.   Carr, P. J. J. *et al.* The structure of proton-bound Triethylammonia (X = F, Cl) Clusters. *Mol. Phys.* **117**, 2972–2979 (2019).

91. Carr, James, Joseph, P. Computational and Spectroscopic Investigations of Intermolecular Interactions in Clusters. (University of Waterloo, 2019).

92. Nieckarz, R. J., Oomens, J., Berden, G., Sagulenko, P. & Zenobi, R. Infrared multiple photon dissociation (IRMPD) spectroscopy of oxazine dyes. *Phys. Chem. Chem. Phys.* **15**, 5049–56 (2013).

93. Wu, R. & McMahon, T. B. An investigation of protonation sites and conformations of protonated amino acids by IRMPD spectroscopy. *ChemPhysChem* **9**, 2826–2835 (2008).

94. MacAleese, L. & Maître, P. Infrared spectroscopy of organometallic ions in the gas phase: From model to real world complexes. *Mass Spectrom. Rev.* **26**, 583–605 (2007).

95. Ortega, J. M. The CLIO infrared FEL facility. *Synchrotron Radiat. News* **9**, 20–33 (1996).

96. Ortega, J. M., Glotin, F. & Prazeres, R. Extension in far-infrared of the CLIO free-electron laser. *Infrared Phys. Technol.* **49**, 133–138 (2006).

97. Couprie, M. E. Panorama of new generation of accelerator based short wavelength coherent light sources. *Nucl. Instruments Methods Phys. Res. Sect. B Beam Interact. with Mater. Atoms* **364**, 4–15 (2015).

98. Prazeres, R. *et al.* CLIO, an infrared free electron laser facility. *Nucl. Instruments Methods Phys. Res. Sect. B Beam Interact. with Mater. Atoms* **89**, 54–59 (1994).

99. Centre Laser Infrarouge d'Orsay (n.d.). Retrieved from: http://old.clio.lcp.u-psud.fr/clio_eng/FEL.html.

100. Agilent Technologies Agilent 6300 Ion Trap LC/MS Systems Concepts Guide. in (Agilent Technologies, Inc., 2006).

101. 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).

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

103. Campbell, J. L., Le Blanc, J. C. Y. & Kibbey, R. G. Differential mobility spectrometry: a valuable technology for analyzing challenging biological samples. *Bioanalysis* **7**, 853–6 (2015).

104. Haack, A., Crouse, J., Schlüter, F. J., Benter, T. & Hopkins, W. S. A First Principle Model of Differential Ion Mobility: the Effect of Ion-Solvent Clustering. *J. Am. Soc. Mass Spectrom.* **30**, 2711–2725 (2019).

105. 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).

106. 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).

107. Krylova, N., Krylov, E., Eiceman, G. A. & Stone, J. A. Effect of moisture on the field dependence of mobility for gas-phase ions of organophosphorus compounds at atmospheric pressure with field asymmetric ion mobility spectrometry. *J. Phys. Chem. A* **107**, 3648–3654 (2003).

108. Campbell, J. L., Yang, A. M.-C., Melo, L. R. & Hopkins, W. S. Studying Gas-Phase Interconversion of Tautomers Using Differential Mobility Spectrometry. *J. Am. Soc. Mass Spectrom.* **27**, 1277–1284 (2016).

109. Parson, W. B. *et al.* Rapid analysis of isomeric exogenous metabolites by differential mobility spectrometry - mass spectrometry. *Rapid Commun. Mass Spectrom.* **25**, 3382–3386 (2011).

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

111. 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).

112. Marlton, S. J. P. *et al.* Discrimination between Protonation Isomers of Quinazoline by Ion Mobility and UV-Photodissociation Action Spectroscopy. *J. Phys. Chem. Lett.* **11**, 4226–4231 (2020).

113. Hernandez, O., Isenberg, S., Steinmetz, V., Glish, G. L. & Maitre, P. Probing Mobility-Selected Saccharide Isomers: Selective Ion− Molecule Reactions and Wavelength-Specific IR Activation. *J. Phys. Chem. A* **119**, 19 (2015).

114. Berthias, F., Maatoug, B., Glish, G. L., Moussa, F. & Maitre, P. Resolution and Assignment of Differential Ion Mobility Spectra of Sarcosine and Isomers. *J. Am. Soc. Mass Spectrom.* **29**, 752–760 (2018).

115. Shvartsburg, A. A., Li, F., Tang, K. & Smith, R. D. High-resolution field asymmetric waveform ion mobility spectrometry using new planar geometry analyzers. *Anal. Chem.* **78**, 3706–3714 (2006).

116. Kolakowski, B. M. & Mester, Z. N. Review of applications of high-field asymmetric waveform ion mobility spectrometry (FAIMS) and differential mobility spectrometry (DMS). (2006) doi:10.1039/b706039d.

117. Brodbelt, J. S. Photodissociation mass spectrometry: New tools for characterization of biological molecules. *Chem. Soc. Rev.* **43**, 2757–2783 (2014).

118. Brodbelt, J. S., Morrison, L. J. & Santos, I. Ultraviolet Photodissociation Mass Spectrometry for Analysis of Biological Molecules. *Chem. Rev.* **120**, 3328–3380 (2020).

119. van Dishoeck, E. F. & Visser, R. Molecular Photodissociation. in *Laboratory Astrochemistry* 229–254 (Wiley-VCH Verlag GmbH & Co. KGaA, 2014). doi:10.1002/9783527653133.ch4.

120. Ly, T. & Julian, R. R. Ultraviolet Photodissociation: Developments towards Applications for Mass-Spectrometry-Based Proteomics. *Angew. Chemie Int. Ed.* **48**, 7130–7137 (2009).

121. Julian, R. R. The Mechanism Behind Top-Down UVPD Experiments: Making Sense of Apparent Contradictions. *J. Am. Soc. Mass Spectrom* **28**, 1823–1826 (2017).

122. Schinke, R. *Photodissociation Dynamics: Spectroscopy and Fragmentation of Small Polyatomic Molecules. Cambridge Monographs on Atomic, Molecular and Chemical Physics* (Cambridge University Press, 1993). doi:DOI: 10.1017/CBO9780511586453.

123. Gregersen, N., Bross, P., Vang, S. & Christensen, J. H. Protein Misfolding and Human Disease. *Annu. Rev. Genomics Hum. Genet.* **7**, 103–124 (2006).

124. Chiti, F. & Dobson, C. M. Protein Misfolding, Functional Amyloid, and Human Disease. *Annu. Rev. Biochem.* **75**, 333–366 (2006).

125. Gallivan, J. P. & Dougherty, D. A. *Cation-π interactions in structural biology. Chemistry* vol. 96 www.pnas.org. (1999).

126. Anjana, R. *et al.* Aromatic-aromatic interactions in structures of proteins and protein-DNA complexes: a study based on orientation and distance. *Bioinformation* **8**, 1220–1224 (2012).

127. Meyer, E. A., Castellano, R. K. & Diederich, F. Interactions with aromatic rings in chemical and biological recognition. *Angewandte Chemie - International Edition* vol. 42 1210–1250 (2003).

128. Baker, C. M. & Grant, G. H. Role of aromatic amino acids in protein-nucleic acid recognition. *Biopolymers* **85**, 456–470 (2007).

129. Asensio, J. L., Ardá, A., Cañada, F. J. & Jiménez-Barbero, J. Carbohydrate-aromatic interactions. *Acc. Chem. Res.* **46**, 946–954 (2013).

130. Ma, J. C. & Dougherty, D. a. The Cation-π Interaction. *Chem. Rev.* **97**, 1303–1324 (1997).

131. Kumpf, R. A. & Dougherty, D. A. A mechanism for ion selectivity in potassium channels: computational studies of cation-pi interactions. *Science* **261**, 1708–10 (1993).

132. Dougherty, D. A. The Cation-π Interaction. **17**, 49 (2012).

133. Beene, D. L. *et al.* Cation−π Interactions in Ligand Recognition by Serotonergic (5-HT 3A) and Nicotinic Acetylcholine Receptors: The Anomalous Binding Properties of Nicotine †. *Biochemistry* **41**, 10262–10269 (2002).

134. Xiu, X., Puskar, N. L., Shanata, J. A. P., Lester, H. A. & Dougherty, D. A. Nicotine binding to brain receptors requires a strong cation−π interaction. *Nature* **458**, 534–537 (2009).

135. Duncan, M. A. Infrared spectroscopy to probe structure and dynamics in metal ion-molecule complexes. *Int. Rev. Phys. Chem.* **22**, 407–435 (2003).

136. Hammer, N. I. *et al.* The vibrational predissociation spectra of the H 5O 2+·RG n(RG=Ar,Ne) clusters: Correlation of the solvent perturbations in the free OH and shared proton transitions of the Zundel ion. *J. Chem. Phys.* **122**, 244301 (2005).

137. Lemaire, J. *et al.* Gas phase infrared spectroscopy of selectively prepared ions. *Phys. Rev. Lett.* **89**, 273002 (2002).

138. Pribble, R. N. & Zwier, T. S. Size-specific infrared spectra of benzene-(H2O)n clusters (n = 1 through 7): Evidence for noncyclic (H2O)n structures. *Science (80-.).* **265**, 75–79 (1994).

139. Rizzo, T. R., Stearns, J. A. & Boyarkin, O. V. Spectroscopic studies of cold, gas-phase biomolecular ions. *Int. Rev. Phys. Chem.* **28**, 481–515 (2009).

140. Piela, L., Olszewski, K. A. & Pillardy, J. On the stability of conformers. *J. Mol. Struct. THEOCHEM* **308**, 229–239 (1994).

141. Scheraga, H. A. Some approaches to the multiple-minima problem in the calculation of polypeptide and protein structures. *Int. J. Quantum Chem.* **42**, 1529–1536 (1992).

142. Hopkins, W. S., Marta, R. A. & McMahon, T. B. Proton-bound 3-cyanophenylalanine trimethylamine clusters: Isomer-specific fragmentation pathways and evidence of gas-

phase zwitterions. *J. Phys. Chem. A* **117**, 10714–10718 (2013).

143. Martens, J., Grzetic, J., Berden, G. & Oomens, J. Structural identification of electron transfer dissociation products in mass spectrometry using infrared ion spectroscopy. *Nat. Commun.* **7**, 1–7 (2016).

144. Poutsma, J. C. *et al.* Infrared Multiple-Photon Dissociation Action Spectroscopy of the b2 + Ion from PPG: Evidence of Third Residue Affecting b2 + Fragment Structure. *J. Am. Soc. Mass Spectrom.* **28**, 1482–1488 (2017).

145. Becke, A. D. *Density-fnnctional exchange-energy approximation with correct asymptotic behavior*. *PHYSiCAL REVIEW A* vol. 38 https://journals.aps.org/pra/pdf/10.1103/PhysRevA.38.3098 (1988).

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

147. Wiberg, K. B. & Rablen, P. R. Comparison of atomic charges derived via different procedures. *J. Comput. Chem.* **14**, 1504–1518 (1993).

148. Demšar, J. *et al.* Orange: Data Mining Toolbox in Python. *J. Mach. Learn. Res.* **14**, 2349–2353 (2013).

149. Becker, O. M. & Karplus, M. The topology of multidimensional potential energy surfaces: Theory and application to peptide structure and kinetics. *J. Chem. Phys.* **106**, 1495–1517 (1997).

150. Wales, D. J., Miller, M. A. & Walsh, T. R. Archetypal energy landscapes. *Nature* **394**, 758–760 (1998).

151. Hopkins, W. S., Marta, R. A., Steinmetz, V. & McMahon, T. B. Mode-specific fragmentation of amino acid-containing clusters. *Phys. Chem. Chem. Phys.* **17**, 28548–28555 (2015).

152. 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).

153. Tian, Z., Wang, X.-B., Wang, L.-S. & Kass, S. R. Are Carboxyl Groups the Most Acidic Sites in Amino Acids? Gas-Phase Acidities, Photoelectron Spectra, and Computations on Tyrosine, p-Hydroxybenzoic Acid, and Their Conjugate Bases. *J. Am. Chem. Soc* **131**, 1174–1181 (2009).

154. Ieritano, C. *et al.* The structures and properties of anionic tryptophan complexes. *Phys. Chem. Chem. Phys.* **20**, 26532–26541 (2018).

155. Burt, M. *et al.* Assessing the impact of anion–π effects on phenylalanine ion structures using IRMPD spectroscopy. *Phys. Chem. Chem. Phys.* **16**, 24223–24234 (2014).

156. 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).

157. Feng, R., Yin, H. & Kong, X. Structure of protonated tryptophan dimer in the gas phase investigated by IRPD spectroscopy and theoretical calculations. *Rapid Commun. Mass Spectrom.* **30**, 24–28 (2016).

158. Seo, J. *et al.* Side-chain effects on the structures of protonated amino acid dimers: A gas-phase infrared spectroscopy study. *Int. J. Mass Spectrom.* (2018) doi:10.1016/j.ijms.2017.06.011.

159. Oh, H. Bin *et al.* Infrared Photodissociation Spectroscopy of Electrosprayed Ions in a Fourier Transform Mass Spectrometer. *J. Am. Chem. Soc.* **127**, 4076–4083 (2005).

160. Rajabi, K. & Fridgen, T. D. Structures of Aliphatic Amino Acid Proton-Bound Dimers by Infrared Multiple Photon Dissociation Spectroscopy in the 700-2000 cm-1 Region. (2008) doi:10.1021/jp0736903.

161. Lee, S.-S. *et al.* Infrared multiple photon dissociation spectroscopy and density functional theory (DFT) studies of protonated permethylated β-cyclodextrin–water non-covalent complexes. *Phys. Chem. Chem. Phys* **16**, 8376 (2014).

162. Martens, S. M., Marta, R. A., Martens, J. K. & McMahon, T. B. Tridentate ionic hydrogen-bonding interactions of the 5-fluorocytosine cationic dimer and other 5-fluorocytosine analogues characterized by irmpd spectroscopy and electronic structure calculations. *J. Phys. Chem. A* **115**, 9837–9844 (2011).

163. Kumar, K. *et al.* Cation–π interactions in protein–ligand binding: theory and data-mining reveal different roles for lysine and arginine. *Chem. Sci.* (2018) doi:10.1039/C7SC04905F.

164. Wu, R. & McMahon, T. B. Investigation of Cation-π Interactions in Biological Systems. *J. Am. Chem. Soc.* 12554–12555 (2008) doi:10.1021/ja802117s.

165. Scutelnic, V. *et al.* The Structure of the Protonated Serine Octamer. *J. Am. Chem. Soc.* **140**, 7554–7560 (2018).

166. Nanita, S. C. & Cooks, R. G. Serine Octamers: Cluster Formation, Reactions, and Implications for Biomolecule Homochirality. *Angew. Chemie Int. Ed.* **45**, 554–569 (2006).

167. Lorenz, U. J. & Rizzo, T. R. Structural Melting of an Amino Acid Dimer upon Intersystem Crossing. (2014) doi:10.1021/ja507981p.

168. Becke, A. D. Assessment of the Perdew-Burke-Ernzerhof exchange-correlation functional. *J. Chem. Phys.* **98**, 926 (1993).

169. 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).

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

171. Lu, T. & Chen, F. Multiwfn: A multifunctional wavefunction analyzer. *J. Comput. Chem.* **33**, 580–592 (2012).

172. Wysocki, V. H., Kenttämaa, H. I. & Cooks, R. G. Internal energy distributions of isolated ions after activation by various methods. *Int. J. Mass Spectrom. Ion Process.* **75**, 181–208 (1987).

173. Voyksner, R. D. & Pack, T. Investigation of collisional-activation decomposition process and spectra in the transport regions of an electrospray single-quadrupole mass spectrometer. *Rapid Commun. Mass Spectrom.* **5**, 263–268 (1991).

174. Derwa, F., De Pauw, E. & Natalis, P. New basis for a method for the estimation fo seconary ion internal energy distribution in 'soft' ionization techniques. *Org. Mass Spectrom.* **26**, 117–118 (1991).

175. Gabelica, V. & De Pauw, E. Internal energy and fragmentation of ions produced in electrospray sources. *Mass Spectrom. Rev.* **24**, 566–587 (2005).

176. Morsa, D., Gabelica, V., Rosu, F., Oomens, J. & De Pauw, E. Dissociation pathways of benzylpyridinium 'thermometer' ions depend on the activation regime: An IRMPD spectroscopy study. *J. Phys. Chem. Lett.* **5**, 3787–3791 (2014).

177. Zins, E. L. *et al.* Investigations of the fragmentation pathways of benzylpyridinium ions under ESI/MS conditions. *J. Mass Spectrom.* **44**, 1668–1675 (2009).

178. Zins, E. L., Pepe, C. & Schröder, D. Methylene-transfer reactions of benzylium/tropylium ions with neutral toluene studied by means of ion-trap mass spectrometry. *Faraday Discuss.* **145**, 157–169 (2010).

179. De Pauw, E., Pelzer, G., Marien, J. & Natalis, P. *Internal Energy Distribution of Ions Emitted in Secondary Ion Mass Spectrometry*. (Springer, Berlin, Heidelberg, 1986). doi:10.1007/978-3-642-82718-1_21.

180. Agnello, A., Derwa, F. & De Pauw, E. Evaluation of Secondary Ion Internal Energy Distribution in LSIMS. in 117–121 (Springer, Boston, MA, 1991). doi:10.1007/978-1-4684-7926-3_14.

181. Collette, C. & Pauw, E. De. Calibration of the internal energy distribution of ions produced by electrospray. *Rapid Commun. Mass Spectrom.* **12**, 165–170 (1998).

182. Collette, C., Drahos, L., Pauw, E. De & Vékey, K. Comparison of the internal energy distributions of ions produced by different electrospray sources. *Rapid Commun. Mass Spectrom.* **12**, 1673–1678 (1998).

183. Drahos, L., Heeren, R. M. A., Collette, C., De Pauw, E. & Vékey, K. Thermal energy distribution observed in electrospray ionization. *J. Mass Spectrom.* **34**, 1373–1379 (1999).

184. Gabelica, V., De Pauw, E. & Karas, M. Influence of the capillary temperature and the source pressure on the internal energy distribution of electrosprayed ions. *Int. J. Mass Spectrom.* **231**, 189–195 (2004).

185. Luo, G., Marginean, I. & Vertes, A. Internal energy of ions generated by matrix-assisted laser desorption/ionization. *Anal. Chem.* **74**, 6185–6190 (2002).

186. Milasinovic, S., Cui, Y., Gordon, R. J. & Hanley, L. Internal energy of thermometer ions formed by femtosecond laser desorption: Implications for mass spectrometric imaging. *J. Phys. Chem. C* **118**, 28938–28947 (2014).

187. Flanigan Iv, P. M., Shi, F., Archer, J. J. & Levis, R. J. Internal Energy Deposition for Low Energy, Femtosecond Laser Vaporization and Nanospray Post-ionization Mass Spectrometry using Thermometer Ions. *J. Am. Soc. Mass Spectrom* **26**, 716–724 (2015).

188. Carpenter, J. E., McNary, C. P., Furin, A., Sweeney, A. F. & Armentrout, P. B. How Hot are Your Ions Really? A Threshold Collision-Induced Dissociation Study of Substituted Benzylpyridinium "Thermometer" Ions. *J. Am. Soc. Mass Spectrom.* **28**, 1876–1888 (2017).

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

190. Stevenson, P. E. The ultraviolet spectra of aromatic hydrocarbons: Predicting substitution and isomerism changes. *J. Chem. Educ.* **41**, 234–239 (1964).

191. Teng, H. H. I. & Dunbar, R. C. Substituent effects in benzene cation spectroscopy. II. Photodissociation spectra of methylbenzene ions. *J. Chem. Phys.* **68**, 3133–3138

(1978).

192. Liu, C. *et al.* Using differential mobility spectrometry to measure ion solvation: an examination of the roles of solvents and ionic structures in separating quinoline-based drugs. *Analyst* **140**, 6897–6903 (2015).

193. Liu, C. *et al.* Assessing Physicochemical Properties of Drug Molecules via Microsolvation Measurements with Differential Mobility Spectrometry. doi:10.1021/acscentsci.6b00297.

194. Londry, F. A. & Hager, J. W. from a Linear Quadrupole Ion Trap. *Society* **0305**, (2003).

195. Frisch, M. J. *et al.* Gaussian 16, Revision C.01. Gaussian Inc. Wallingford CT (2016).

196. Berraud-Pache, R., Neese, F., Bistoni, G. & Izsák, R. Unveiling the Photophysical Properties of Boron-dipyrromethene Dyes Using a New Accurate Excited State Coupled Cluster Method. *J. Chem. Theory Comput.* **16**, 564–575 (2020).

197. Lewis, M., Bagwill, C., Hardebeck, L. K. E. & Wireduaah, S. THE USE OF HAMMETT CONSTANTS TO UNDERSTAND THE NON-COVALENT BINDING OF AROMATICS. *Comput. Struct. Biotechnol. J.* **1**, e201204004 (2012).

198. Hammett, L. P. The Effect of Structure upon the Reactions of Organic Compounds. Benzene Derivatives. *J. Am. Chem. Soc.* **59**, 96–103 (1937).

199. Sadlej-Sosnowska, N. & Kijak, M. Excited state substituent constants: To Hammett or not? *Struct. Chem.* **23**, 359–365 (2012).

200. Brown, H. C. & Okamoto, Y. Substituent Constants for Aromatic Substitution. *J. Am. Chem. Soc.* **79**, 1913–1917 (1957).

201. Brown, H. C. & Okamoto, Y. Electrophilic Substituent Constants. *J. Am. Chem. Soc.* **80**, 4979–4987 (1958).

202. Harrison, A. G. Linear free energy correlations in mass spectrometry†. *J. Mass Spectrom.* **34**, 577–589 (1999).

203. Gabelica, V., Lemaire, D., Laprévote, O. & De Pauwa, E. Kinetics of solvent addition on electrosprayed ions in an electrospray source and in a quadrupole ion trap. *Int. J. Mass Spectrom.* **210**–**211**, 113–119 (2001).

204. Hansch, C., Leo, A. & Taft, R. W. A Survey of Hammett Substituent Constants and Resonance and Field Parameters. *Chem. Rev.* **91**, 165–195 (1991).

205. Baldry, P. J. Substituent effects and excited state reactivity. *J. Chem. Soc. Perkin Trans. 2* 951–953 (1979) doi:10.1039/P29790000951.

206. Terent'ev, P. B. & Kalandarishvili, A. G. Application of mass spectrometry for the analysis of organic tautomeric compounds. *Mass Spectrom. Rev.* **15**, 339–363 (1996).

207. Furlong, J. J. P., Schiavoni, M. M., Castro, E. A. & Allegretti, P. E. Mass spectrometry as a tool for studying tautomerism. *Russian Journal of Organic Chemistry* vol. 44 1725–1736 (2008).

208. Almasian, M. *et al.* Non-equilibrium isomer distribution of the gas-phase photoactive yellow protein chromophore. *J. Phys. Chem. Lett.* **3**, 2259–2263 (2012).

209. Matthews, E., Cercola, R. & Dessent, C. Protomer-Dependent Electronic Spectroscopy and Photochemistry of the Model Flavin Chromophore Alloxazine. *Molecules* **23**, 2036 (2018).

210. Lapthorn, C., Dines, T. J., Chowdhry, B. Z., Perkins, G. L. & Pullen, F. S. Can ion

mobility mass spectrometry and density functional theory help elucidate protonation sites in 'small' molecules? *Rapid Commun. Mass Spectrom.* **27**, 2399–2410 (2013).

211. Lalli, P. M. *et al.* Protomers: formation, separation and characterization via travelling wave ion mobility mass spectrometry. *J. Mass Spectrom.* **47**, 712–719 (2012).

212. Kaufmann, A. *et al.* Are liquid chromatography/electrospray tandem quadrupole fragmentation ratios unequivocal confirmation criteria? *Rapid Commun. Mass Spectrom.* **23**, 985–998 (2009).

213. Xia, H. & Attygalle, A. B. Transformation of the gas-phase favored O-protomer of p-aminobenzoic acid to its unfavored N-protomer by ion activation in the presence of water vapor: An ion-mobility mass spectrometry study. *J. Mass Spectrom.* **53**, 353–360 (2018).

214. Tian, Z. & Kass, S. R. Does electrospray ionization produce gas-phase or liquid-phase structures? *J. Am. Chem. Soc.* **130**, 10842–10843 (2008).

215. Chang, T. M., Prell, J. S., Warrick, E. R. & Williams, E. R. Where's the charge? Protonation sites in gaseous ions change with hydration. *J. Am. Chem. Soc.* **134**, 15805–15813 (2012).

216. Schmidt, J., Meyer, M. M., Spector, I. & Kass, S. R. Infrared multiphoton dissociation spectroscopy study of protonated p-aminobenzoic acid: Does electrospray ionization afford the amino- or carboxy-protonated ion? *J. Phys. Chem. A* **115**, 7625–7632 (2011).

217. Kumar, R., Yerabolu, R. & Kenttämaa, H. I. Effects of Residual Water in a Linear Quadrupole Ion Trap on the Protonation Sites of 4-Aminobenzoic Acid. *J. Am. Soc. Mass Spectrom.* **31**, 124–131 (2020).

218. Steill, J. D. & Oomens, J. Gas-phase deprotonation of p-hydroxybenzoic acid investigated by IR spectroscopy: Solution-phase structure is retained upon ESI. *J. Am. Chem. Soc.* **131**, 13570–13571 (2009).

219. Patrick, A. L., Cismesia, A. P., Tesler, L. F. & Polfer, N. C. Effects of ESI conditions on kinetic trapping of the solution-phase protonation isomer of p-aminobenzoic acid in the gas phase. *Int. J. Mass Spectrom.* **418**, 148–155 (2017).

220. Kumler, W. D. & Strait, L. A. The Ultraviolet Absorption Spectra and Resonance in Benzene Derivatives-Sulfanilamide, Metanilamide, p-Aminobenzoic Acid, Benzenesulfonamide, Benzoic Acid and Aniline. *J. Am. Chem. Soc.* **65**, 2349–2354 (1943).

Appendix A
Proton-Bound heterodimers of Phenylalanine/Serine

A.1 Energies of of [Phe/Ser + H]⁺

Table A.1 The relative zero-point corrected energies (ZPE$_{rel.}$), enthalpy ($\Delta_{rel.}$H), entropy ($\Delta_{rel.}$S), and Gibbs free energies ($\Delta_{rel.}$G) of [Phe/Ser + H]$^+$. ZPE$_{rel.}$, $\Delta_{rel.}$H, and $\Delta_{rel.}$G at 298 K are given in kJ·mol^{-1}. $\Delta_{rel.}$S at 298 K is given in J·mol^{-1}. Calculations employed the B3LYP functional and 6-311++G(d,p) basis set.

Isomer List	ZPE$_{rel.}$ kJ·mol^{-1}	$\Delta_{rel.}$H kJ·mol^{-1}	$\Delta_{rel.}$S J·mol^{-1}	$\Delta_{rel.}$G kJ·mol^{-1}
Isomer 1	0.00	0.00	0.00	0.00
Isomer 2	8.89	9.85	24.11	2.66
Isomer 3	9.19	10.68	32.81	0.89
Isomer 4	18.68	19.81	26.04	12.04
Isomer 5	22.38	23.88	29.85	14.98
Isomer 6	29.30	29.74	18.64	24.19
Isomer 7	32.81	33.26	7.74	30.95
Isomer 8	40.36	42.92	46.57	29.04
Isomer 9	40.62	40.71	10.04	37.71
Isomer 10	41.26	43.37	28.03	35.01
Isomer 11	42.09	44.65	50.60	29.56
Isomer 12	42.30	45.06	77.69	21.90
Isomer 13	43.45	45.84	44.34	32.62
Isomer 14	43.66	46.25	54.54	29.99
Isomer 15	48.88	49.93	23.52	42.92
Isomer 16	50.03	51.60	28.50	43.11
Isomer 17	50.73	51.13	10.29	48.07
Isomer 18	59.45	60.25	17.18	55.12
Isomer 19	59.72	62.30	49.92	47.41
Isomer 20	65.75	68.00	45.41	54.46
Isomer 21	111.08	113.09	40.68	100.96
Isomer 22	111.88	113.98	45.27	100.48
Isomer 23	113.09	115.25	45.94	101.55
Isomer 24	117.72	119.62	39.64	107.80
Isomer 25	118.62	120.42	29.15	111.73
Isomer 26	135.53	138.60	63.49	119.67
Isomer 27	135.73	138.80	64.98	119.43
Isomer 28	142.06	145.29	61.40	126.98
Isomer 29	145.05	147.06	58.31	129.68

Isomer 30	148.07	151.32	68.55	130.88
Isomer 31	149.78	152.80	52.78	137.06
Isomer 32	151.25	153.28	41.12	141.02
Isomer 33	155.24	157.04	47.51	142.88
Isomer 34	155.54	157.05	40.25	145.05
Isomer 35	163.88	167.02	51.46	151.68
Isomer 36	183.91	187.26	35.33	176.73
Isomer 37	194.48	197.87	54.98	181.48

A.2 Determining the appropriate scaling factor

To determine the appropriate scaling factor for the calculated spectra, we began by employing a scaling factor of 0.95. This scaling factor is commonly used for calculated harmonic spectra in the $2800 - 3800$ cm^{-1} region. Using this initial guess to scale the calculated spectra, we used the cosine distances to judge which spectra best matched the various experimental spectra (as described in the article). Having identified several possible matches for each experimental spectrum, we then proceeded to calculate the cosine distances for the best matching calculated spectra as the scaling factor was varied from 0.92 to 1.0. Plots showing the cosine distances as a function of scaling factor are shown below for each experimental spectrum.
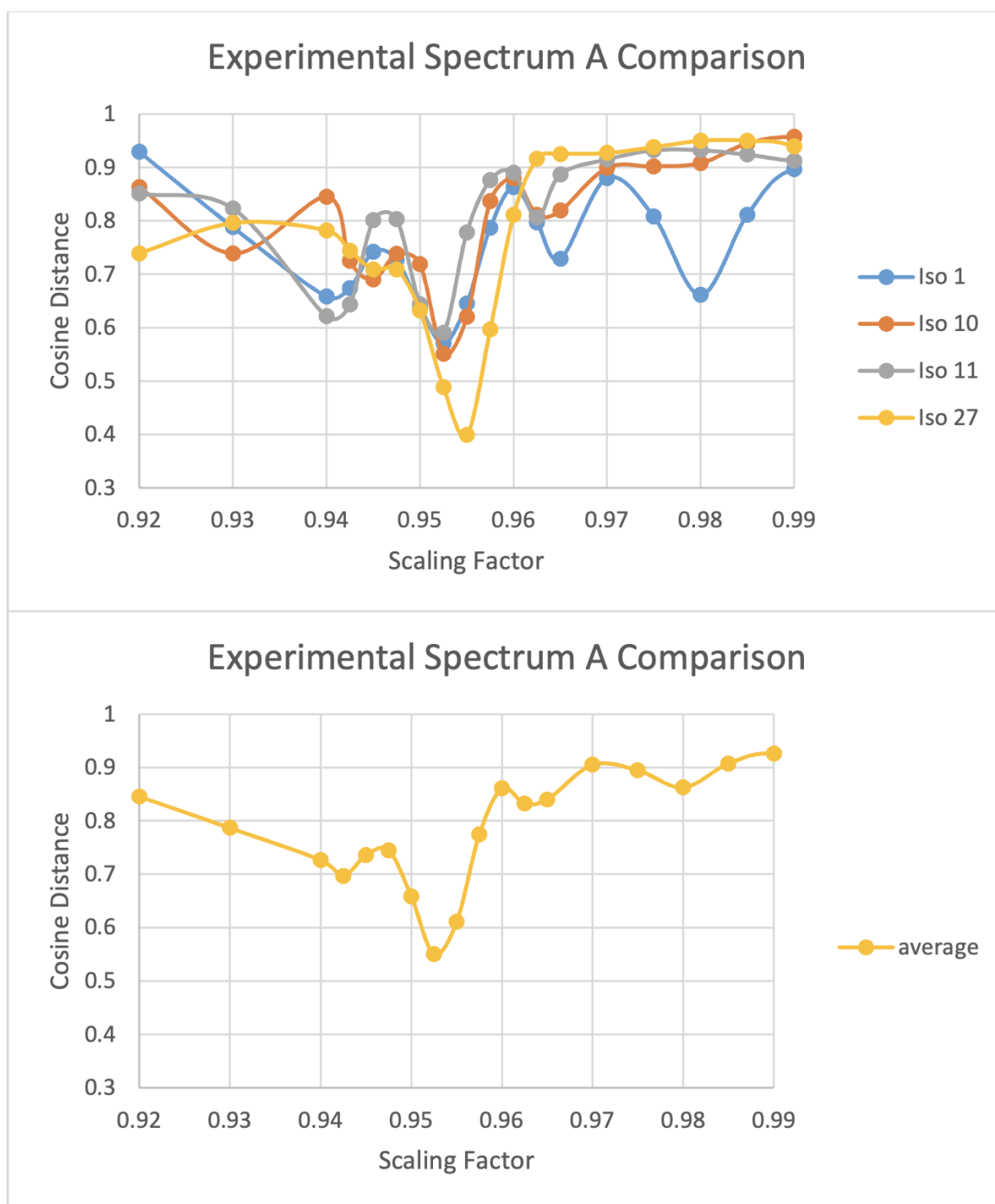
Figure A.1 The cosine distances for experimental spectrum A and the harmonic spectra of several isomers of [Phe/Ser + H]$^+$ plotted as a function of scaling factor.
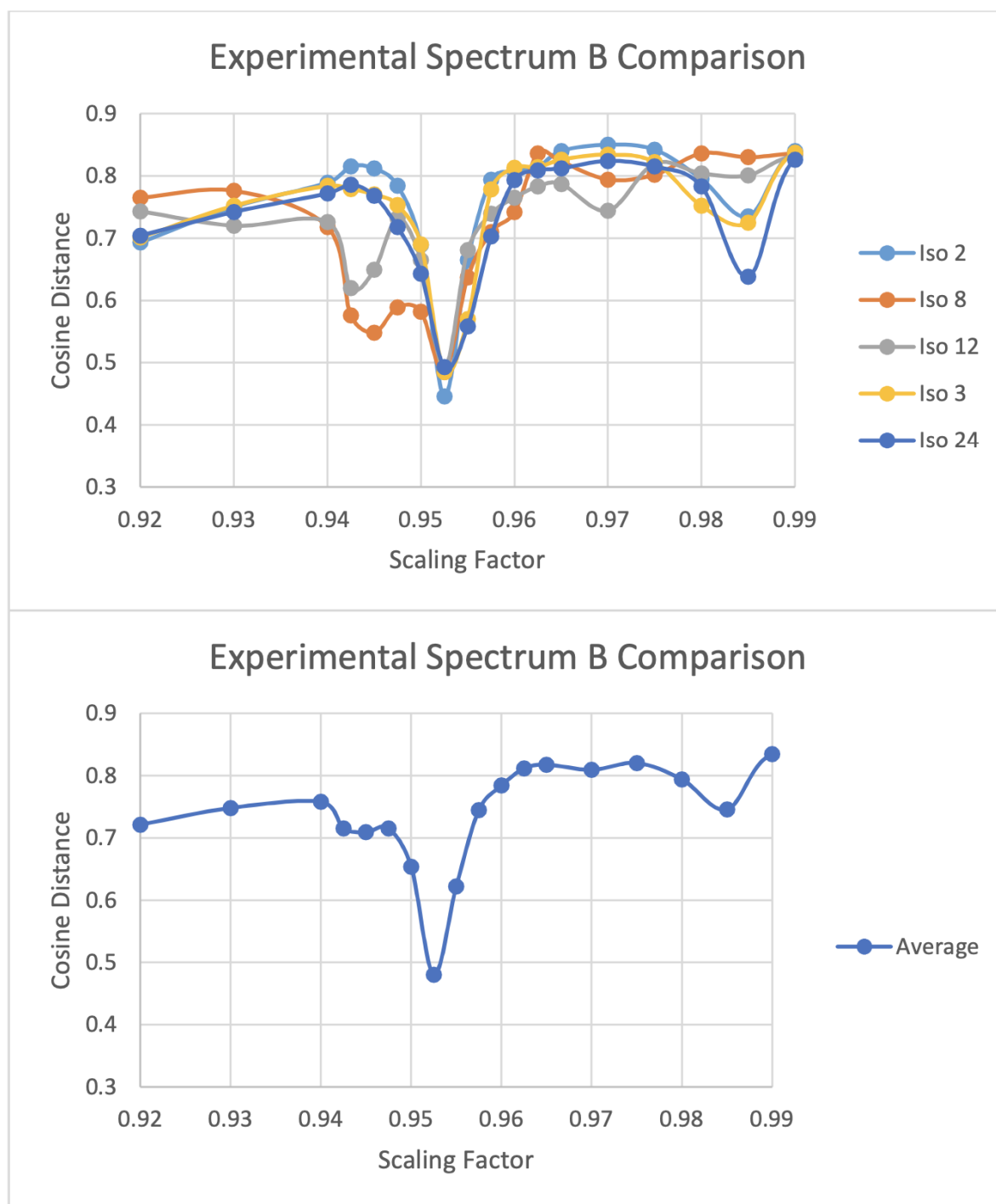
Figure A.2 The cosine distances for experimental spectrum B and the harmonic spectra of several isomers of [Phe/Ser + H]$^+$ plotted as a function of scaling factor.
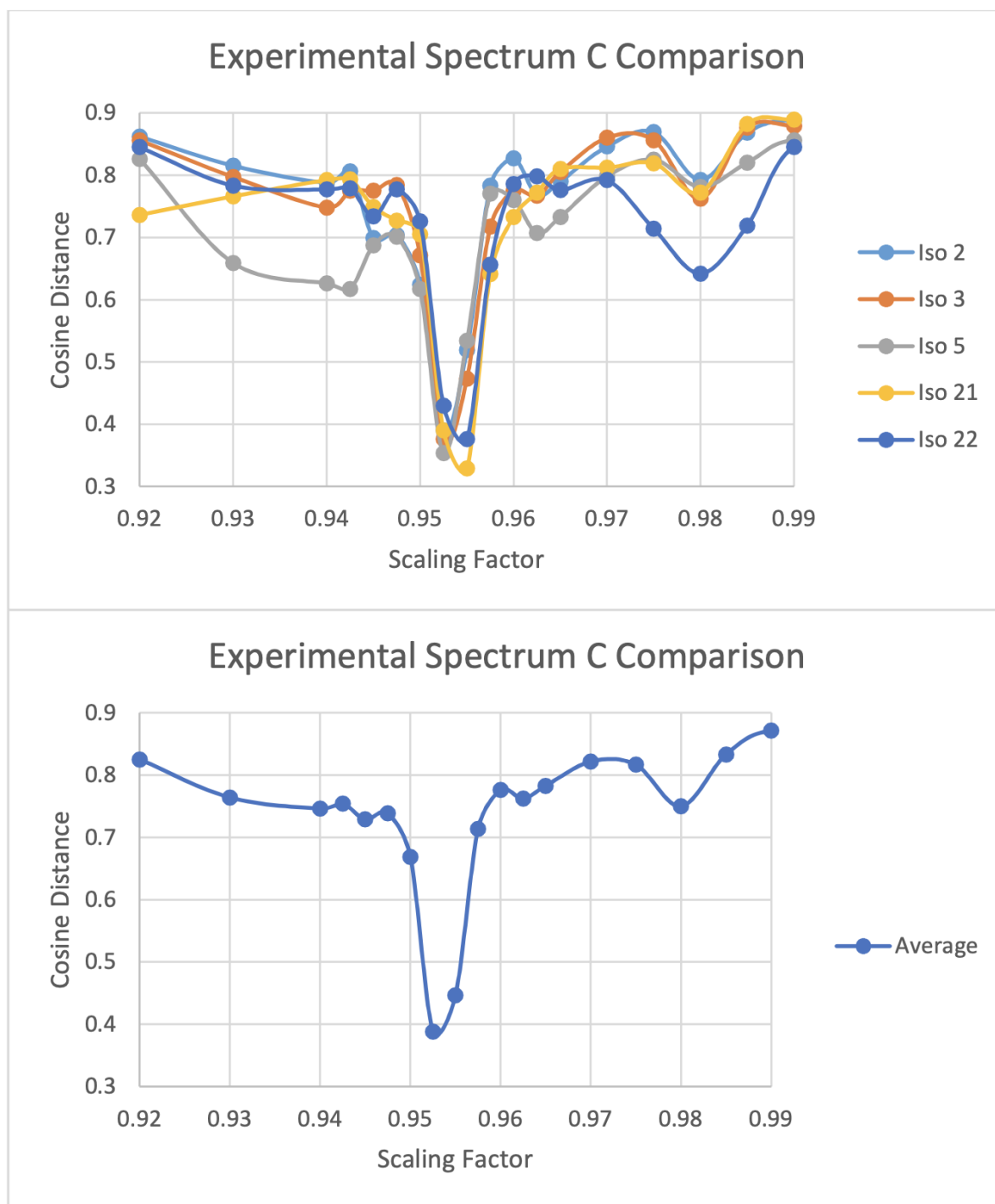
Figure A.3 The cosine distances for experimental spectrum C and the harmonic spectra of several isomers of [Phe/Ser + H]$^+$ plotted as a function of scaling factor.
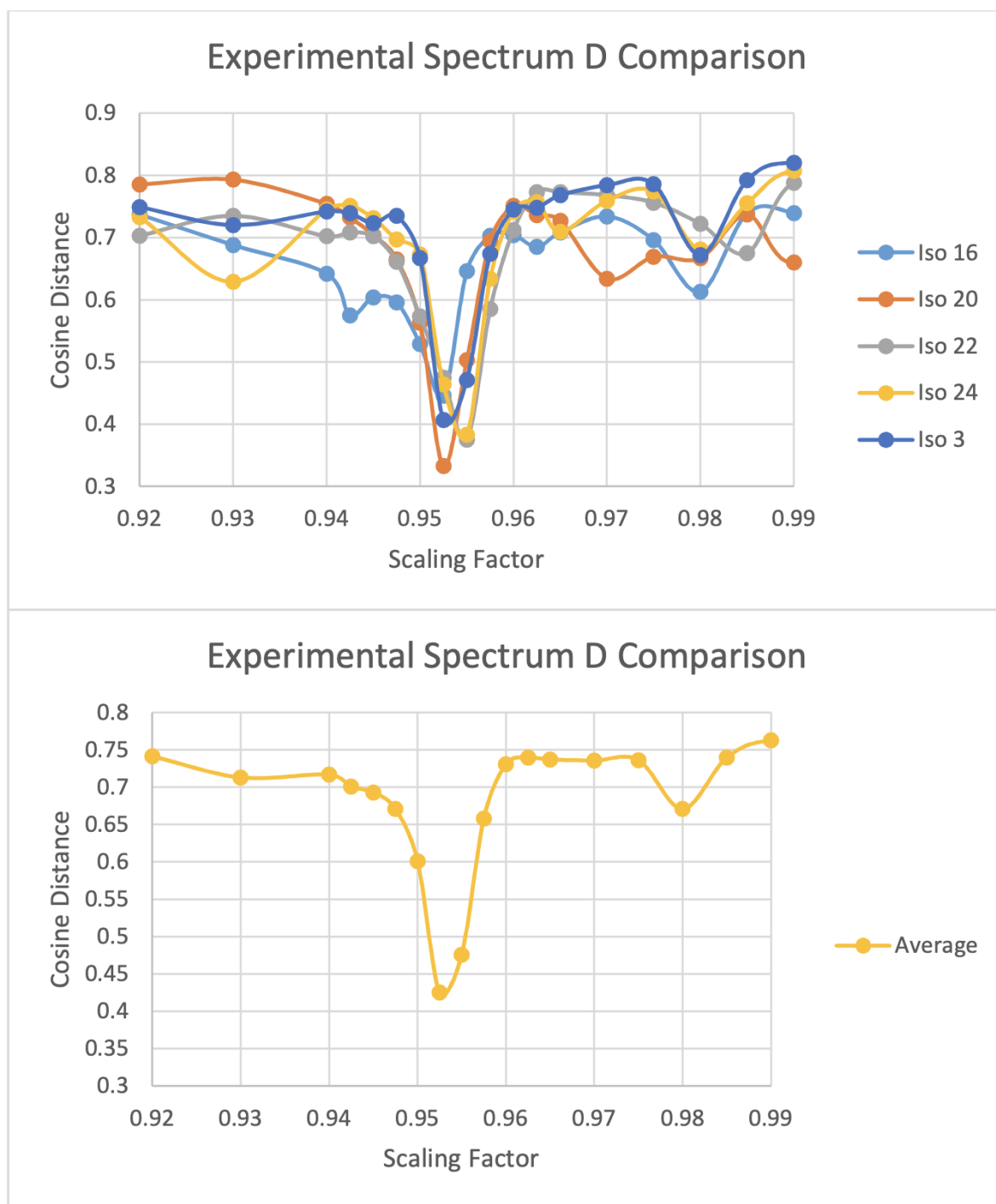
Figure A.4 The cosine distances for experimental spectrum D and the harmonic spectra of several isomers of [Phe/Ser + H]$^+$ plotted as a function of scaling factor.
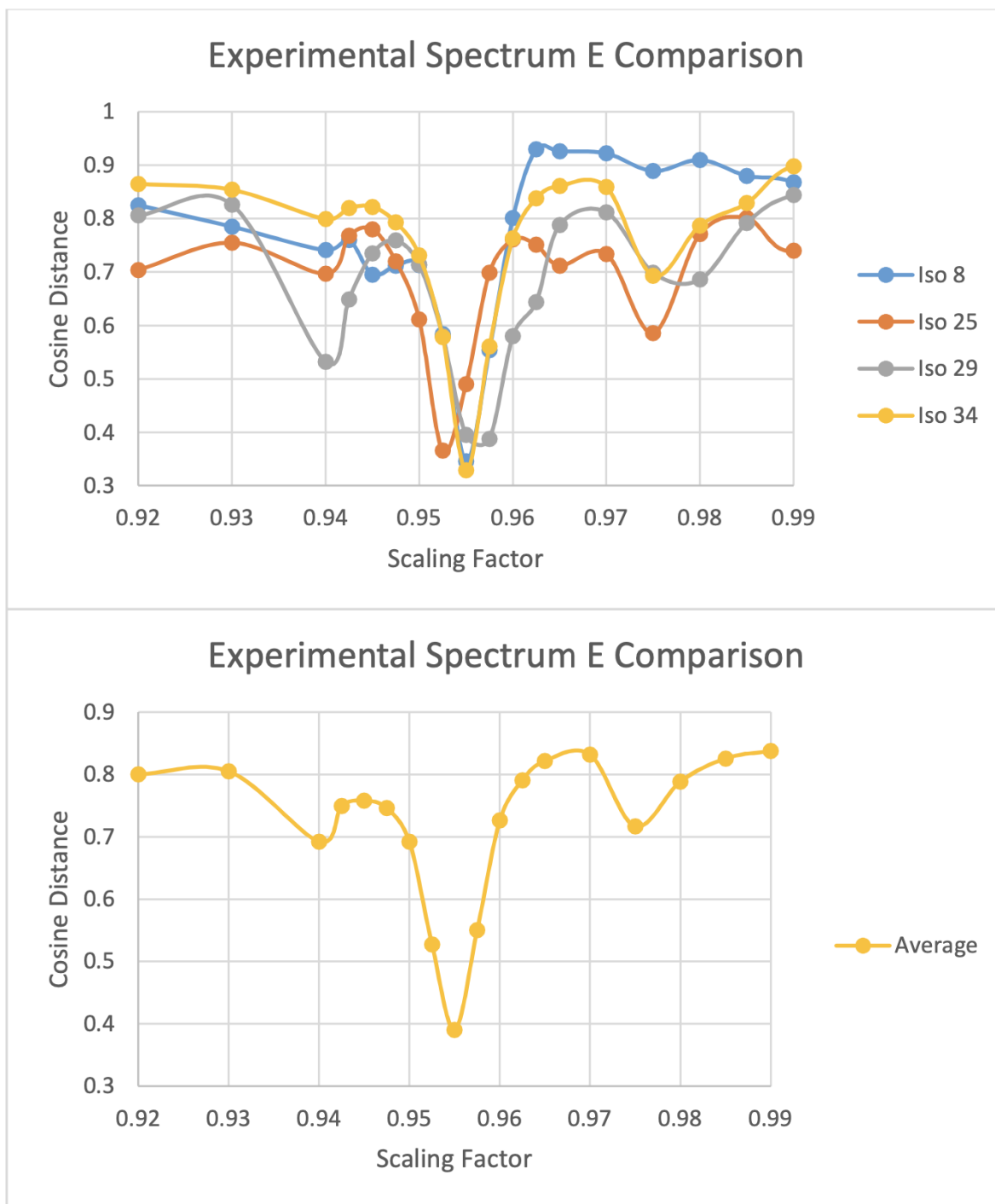
Figure A.5 The cosine distances for experimental spectrum E and the harmonic spectra of several isomers of [Phe/Ser + H]$^+$ plotted as a function of scaling factor.

A.3 The *XYZ* atomic coordinates and structures of [Phe/Ser + H]⁺

The *XYZ* atomic coordinates of all 37 isomers with relative zero-point corrected energies are provided. Calculations were conducted at the B3LYP/6-311++G(d,p) level of theory.
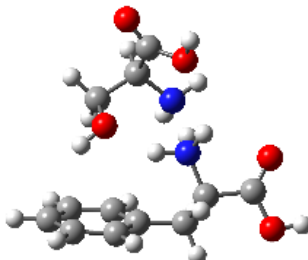
Isomer 1, 0.0 kJ·mol⁻¹

N	-0.131652	1.102907	-0.587218
H	0.365403	1.794129	-1.150232
H	0.048453	0.161148	-0.966506
C	-1.582130	1.417015	-0.458589
H	-2.035687	1.314421	-1.447137
C	-2.275769	0.465452	0.538987
C	-1.694335	2.876735	-0.031597
H	-1.951586	0.714352	1.553487
H	-3.346018	0.680125	0.490860
C	-1.999365	-0.997471	0.257695
O	-0.739196	3.571182	0.204051
C	-1.474009	-1.817755	1.259344
C	-2.243413	-1.549071	-1.005731
H	-1.299177	-1.409999	2.249838
C	-1.187665	-3.158934	1.006427
C	-1.945046	-2.885005	-1.268120
H	-2.682176	-0.942805	-1.792176
H	-0.798666	-3.787573	1.799321
C	-1.409646	-3.693627	-0.262673
H	-2.144347	-3.299196	-2.249500
H	-1.191779	-4.736392	-0.461506
O	-2.972095	3.255366	0.043481
H	-3.017476	4.186785	0.315387
C	2.661190	-0.230980	0.906861
H	3.461694	-0.467735	1.617499
C	2.077020	-1.571902	0.453895
H	1.545236	-2.035627	1.292518
H	2.903499	-2.220760	0.151542
O	1.183229	-1.358136	-0.644184
H	0.713594	-2.180581	-0.833779
C	3.376282	0.420463	-0.277129

O	2.771938	1.558025	-0.716372
H	3.329177	1.925272	-1.422364
O	4.370096	-0.033121	-0.764833
H	0.401442	1.064439	0.338601
H	1.264636	0.194643	2.348653
H	2.043878	1.521815	1.808658
N	1.629272	0.641436	1.510743
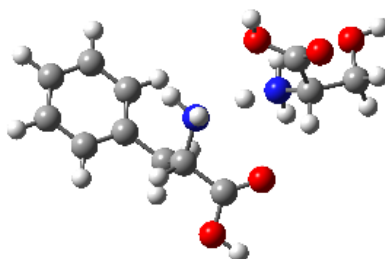
Isomer 2, 8.9 kJ·mol^{-1}

N	0.343825	0.128715	0.752637
H	-0.042177	0.175490	1.692843
H	0.993311	-0.657526	0.754049
C	1.095917	1.359603	0.443862
H	1.700000	1.692541	1.293508
C	2.048624	1.126001	-0.760693
C	0.127070	2.477369	0.089055
H	1.450467	0.886850	-1.645911
H	2.568438	2.064969	-0.962934
C	3.035947	0.015903	-0.481410
O	-1.008407	2.314188	-0.308276
C	2.876034	-1.248534	-1.057520
C	4.107923	0.229048	0.393824
H	2.064586	-1.422297	-1.758037
C	3.767093	-2.282241	-0.765336
C	4.997460	-0.800691	0.687112
H	4.256201	1.209206	0.836540
H	3.638312	-3.253573	-1.228557
C	4.827058	-2.060042	0.110028
H	5.829098	-0.619696	1.357996
H	5.523840	-2.858868	0.334515
O	0.699403	3.674636	0.222931
H	0.077954	4.364646	-0.060618
C	-3.124553	0.003683	-0.074333
H	-3.032048	1.014148	0.327085
C	-4.337402	-0.094511	-0.996735
H	-4.318927	0.726671	-1.724187
H	-5.241011	-0.012943	-0.387171
O	-4.219294	-1.361091	-1.636632

H	-5.030112	-1.585794	-2.104352
C	-3.279000	-0.958821	1.104787
O	-2.200265	-1.755905	1.269304
H	-2.376177	-2.362048	2.007712
O	-4.252956	-0.969989	1.797910
H	-1.941805	-1.196111	-1.289730
H	-1.773194	0.435636	-1.581875
H	-0.986520	-0.186012	-0.235066
N	-1.875235	-0.273246	-0.854066
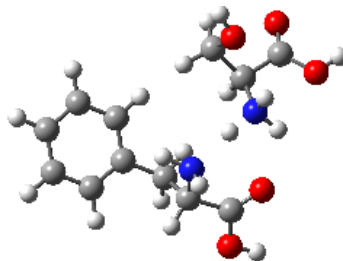
Isomer 3, 9.2 kJ·mol^{-1}

N	0.669382	-0.863910	-1.451309
H	0.683047	-1.308552	-2.367252
H	1.267440	-0.041522	-1.527192
C	1.263891	-1.771957	-0.449093
H	2.171590	-2.250633	-0.828565
C	1.630248	-0.996564	0.845980
C	0.270715	-2.866291	-0.084587
H	0.709761	-0.663366	1.335717
H	2.118498	-1.698462	1.526006
C	2.522275	0.190759	0.563603
O	-0.934170	-2.774522	-0.182904
C	2.038694	1.495795	0.699744
C	3.839885	0.002924	0.128485
H	1.030784	1.658264	1.069794
C	2.849353	2.592774	0.405894
C	4.650673	1.095405	-0.167766
H	4.239168	-1.002086	0.034130
H	2.465864	3.598798	0.531552
C	4.155870	2.393391	-0.033469
H	5.671416	0.935430	-0.494609
H	4.789944	3.242564	-0.258705
O	0.894351	-3.934405	0.416114
H	0.236039	-4.593412	0.690698
C	-2.192618	0.671099	0.156020
H	-1.506564	0.381571	0.955434
C	-1.839126	2.064610	-0.360453
H	-0.765858	2.121005	-0.577818

123

H	-2.091593	2.791651	0.415429
O	-2.615689	2.231420	-1.542128
H	-2.632954	3.153573	-1.817151
C	-3.604744	0.655266	0.752036
O	-4.334956	-0.377432	0.292934
H	-5.210962	-0.361657	0.712221
O	-3.968268	1.466523	1.552059
H	-2.539242	0.028671	-1.779588
H	-0.989311	-0.454460	-1.189034
H	-2.372539	-1.240418	-0.699741
N	-2.031156	-0.311529	-0.958825
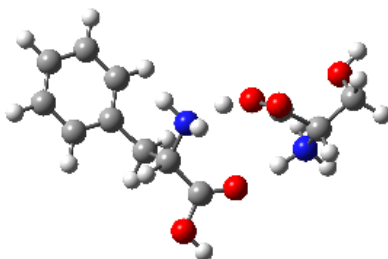
Isomer 4, 18.7 kJ·mol^{-1}

N	-0.527047	-0.106768	1.109275
H	-0.138459	-0.361612	2.015697
H	-1.293892	0.540994	1.286326
C	-1.059493	-1.280278	0.397204
H	-1.758456	-1.846439	1.019992
C	-1.812142	-0.826644	-0.886529
C	0.046467	-2.238016	-0.022147
H	-1.101081	-0.318636	-1.544366
H	-2.168913	-1.720012	-1.404552
C	-2.967751	0.088682	-0.552345
O	1.196223	-1.948035	-0.300962
C	-2.848772	1.474629	-0.700492
C	-4.162648	-0.439302	-0.049023
H	-1.933606	1.897112	-1.103715
C	-3.904203	2.318198	-0.352341
C	-5.216556	0.401348	0.298534
H	-4.274024	-1.513913	0.059040
H	-3.802551	3.389639	-0.479781
C	-5.088171	1.782886	0.149440
H	-6.139932	-0.020230	0.678174
H	-5.910237	2.436464	0.416169
O	-0.423324	-3.482232	-0.112186
H	0.273024	-4.072085	-0.442667
C	3.802241	0.073379	0.179147
H	4.178673	-0.778358	0.744879

C	4.875610	1.152687	0.055126
H	5.835742	0.700651	-0.226775
H	4.988994	1.628628	1.034321
O	4.412476	2.047985	-0.946467
H	4.955609	2.841972	-0.979409
C	2.515185	0.549594	0.872819
O	1.580621	0.909543	0.016248
H	0.651459	0.652196	0.432190
O	2.427410	0.512581	2.075088
H	3.163899	0.408949	-1.762827
H	4.182530	-0.907524	-1.666046
H	2.587919	-1.026431	-1.118983
N	3.417885	-0.413106	-1.202857
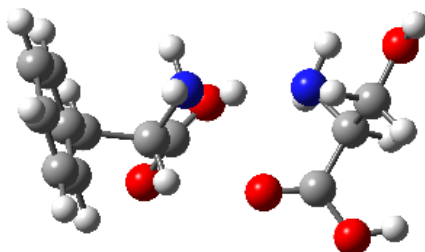
Isomer 5, 22.4 kJ·mol^{-1}

N	-0.244357	-0.318208	-0.947865
H	0.827368	-0.222274	-0.905365
H	-0.675816	0.590411	-0.760253
H	-0.485507	-0.587709	-1.904038
C	-0.791473	-1.315861	0.030084
H	-0.529875	-0.943473	1.020001
C	-2.323808	-1.434980	-0.111924
C	-0.056124	-2.626557	-0.193050
H	-2.567770	-1.922158	-1.061477
H	-2.670001	-2.096865	0.684158
C	-2.984844	-0.075571	-0.029824
O	0.788534	-2.785969	-1.041441
C	-3.536813	0.518761	-1.169669
C	-3.001960	0.630539	1.179660
H	-3.551878	-0.025695	-2.109271
C	-4.100431	1.793860	-1.103062
C	-3.563546	1.903179	1.246750
H	-2.588478	0.178999	2.076175
H	-4.537575	2.236767	-1.990286
C	-4.112408	2.488030	0.104563
H	-3.581776	2.434160	2.191218
H	-4.555439	3.475287	0.159619
O	-0.470028	-3.556755	0.666666

H	0.010010	-4.385067	0.503422
C	3.340295	0.489549	-0.063914
H	4.346271	0.065691	0.060528
C	3.507302	1.953575	-0.500897
H	4.095906	2.499846	0.245343
H	2.518471	2.425192	-0.577131
O	4.162641	1.909689	-1.758846
H	4.308545	2.801659	-2.091064
C	2.600865	0.449445	1.277767
O	3.321486	0.727791	2.371487
H	4.262506	0.833568	2.176312
O	1.420480	0.222408	1.366523
N	2.565828	-0.236305	-1.077140
H	2.926269	0.020766	-1.994744
H	2.691890	-1.242188	-0.985597
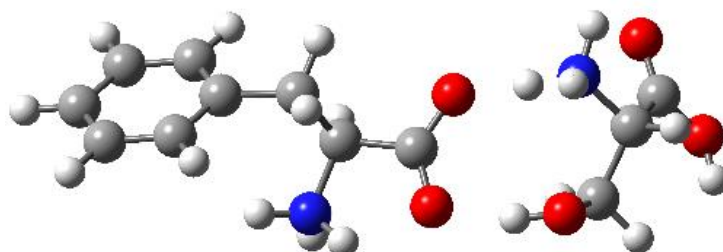
Isomer 6, 29.3 kJ·mol^{-1}

N	1.824398	1.171863	-1.293352
H	1.766134	0.992863	-2.295906
H	2.803015	1.379308	-1.055521
H	1.178187	1.958937	-1.067888
C	1.283990	0.016629	-0.469796
H	1.313474	-0.881647	-1.084514
C	2.137925	-0.169564	0.801604
C	-0.174264	0.388224	-0.114887
H	1.963444	0.674666	1.475924
H	1.756323	-1.062332	1.299827
C	3.609036	-0.289320	0.476895
O	-0.500026	1.580994	-0.337966
C	4.463891	0.806586	0.653800
C	4.125214	-1.467977	-0.076032
H	4.090630	1.713481	1.122223
C	5.808317	0.727725	0.281845
C	5.465744	-1.548065	-0.442619
H	3.479545	-2.331276	-0.204313
H	6.461462	1.578428	0.436892
C	6.308760	-0.448364	-0.269638
H	5.856634	-2.470197	-0.856556

H	7.352449	-0.515016	-0.552619
O	-0.853693	-0.531359	0.388889
H	-2.277994	-0.398596	0.989583
C	-4.295111	0.149279	0.508189
H	-5.120393	0.572270	1.088730
C	-3.681976	1.247177	-0.385483
H	-4.478651	1.729506	-0.964591
H	-2.977081	0.791767	-1.089384
O	-3.046302	2.187566	0.443293
H	-2.093557	2.178793	0.214601
C	-4.794770	-1.086617	-0.246971
O	-5.667224	-0.870026	-1.239723
H	-5.901571	0.061632	-1.344194
O	-4.424814	-2.191041	0.039708
N	-3.244963	-0.307746	1.472071
H	-3.136536	0.384677	2.213332
H	-3.492209	-1.223511	1.856598
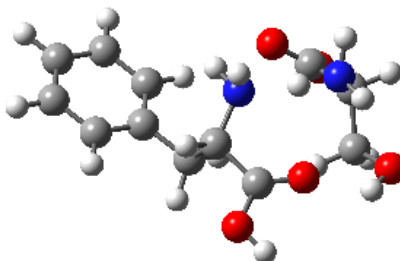
Isomer 7, 32.8 kJ·mol^{-1}

N	-0.122227	0.646989	-1.464489
H	-0.218862	1.037053	-2.399566
H	-0.742248	-0.162680	-1.430304
H	1.425060	0.033962	-1.338472
C	-0.568252	1.637079	-0.465287
H	-1.359896	2.281217	-0.858374
C	-1.122707	0.930797	0.803813
C	0.596540	2.525312	-0.061938
H	-0.349688	0.267600	1.204565
H	-1.321672	1.697337	1.556101
C	-2.381381	0.149952	0.498628
O	1.769418	2.215243	-0.151873
C	-2.345959	-1.237813	0.330217
C	-3.598927	0.823670	0.341485
H	-1.414246	-1.778204	0.456291
C	-3.508139	-1.940016	0.006280
C	-4.757772	0.123167	0.020533
H	-3.643498	1.899242	0.484187
H	-3.468615	-3.016151	-0.115745

127

C	-4.713675	-1.261541	-0.150667
H	-5.696146	0.654304	-0.088383
H	-5.617176	-1.806748	-0.396843
O	0.176175	3.677245	0.456899
H	0.937012	4.205919	0.747005
C	2.889650	-1.199872	-0.281022
H	3.837914	-1.671452	-0.555875
C	3.138854	-0.285699	0.939400
H	3.567564	-0.882235	1.750177
H	2.191419	0.133820	1.286779
O	4.053007	0.719616	0.559678
H	3.605106	1.574883	0.590819
C	1.817623	-2.259966	-0.022916
O	2.102042	-3.187103	0.897869
H	3.002934	-3.127481	1.243805
O	0.760249	-2.241997	-0.591994
N	2.450913	-0.354813	-1.431685
H	3.078245	0.453840	-1.485744
H	2.496582	-0.875713	-2.308054
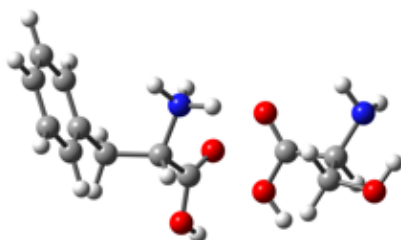
Isomer 8, 40.4 kJ·mol^{-1}

N	-0.801471	0.522174	1.167140
H	0.184687	0.149716	1.179060
H	-1.464185	-0.243577	1.336896
H	-0.876436	1.217513	1.917552
C	-1.104915	1.172533	-0.151941
H	-0.634895	0.565843	-0.925818
C	-2.631915	1.255357	-0.383668
C	-0.459487	2.550558	-0.120055
H	-3.067929	1.958654	0.332544
H	-2.781469	1.675431	-1.380118
C	-3.276905	-0.106294	-0.253958
O	-0.068859	3.069955	0.894956
C	-3.950273	-0.465154	0.920396
C	-3.156598	-1.048883	-1.282684
H	-4.085167	0.266204	1.712388
C	-4.491802	-1.744111	1.065286
C	-3.699292	-2.322920	-1.139372

H	-2.654351	-0.779222	-2.206726
H	-5.022684	-2.004816	1.973227
C	-4.363946	-2.674135	0.036871
H	-3.612990	-3.039345	-1.947817
H	-4.789808	-3.664548	0.144201
O	-0.437622	3.097119	-1.335254
H	-0.081485	3.998942	-1.276269
C	4.020886	-0.418599	0.215295
H	4.609159	0.438177	-0.131221
C	4.222123	-1.553442	-0.832877
H	3.820883	-1.268475	-1.809216
H	3.674323	-2.443066	-0.487306
O	5.587737	-1.806651	-1.011604
H	5.971139	-1.979911	-0.140533
C	2.552302	-0.021615	0.246397
O	2.073430	0.709656	-0.780712
H	2.768414	0.952164	-1.409087
O	1.775183	-0.369167	1.114176
N	4.556893	-0.866327	1.482095
H	4.773320	-0.102847	2.111535
H	3.922825	-1.502508	1.955126
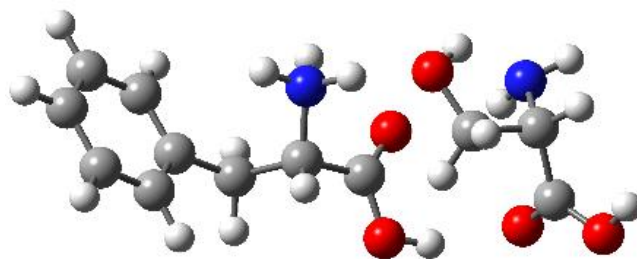
Isomer 9, 40.6 kJ·mol^{-1}

N	-0.896977	1.036149	-0.494268
H	-0.064433	1.548178	-0.084195
H	-1.787268	1.422189	-0.165067
H	-0.841615	1.151408	-1.510897
C	-0.807246	-0.428825	-0.155691
H	-0.758945	-0.498322	0.931275
C	-2.038720	-1.192091	-0.683983
C	0.501832	-0.916182	-0.768927
H	-2.015505	-1.194093	-1.778013
H	-1.938214	-2.229811	-0.357210
C	-3.321439	-0.573442	-0.172915
O	0.979567	-0.381167	-1.743622
C	-4.088964	0.256592	-0.998621
C	-3.725235	-0.774484	1.152808
H	-3.808204	0.394984	-2.038822

C	-5.239215	0.876626	-0.507584
C	-4.873742	-0.157444	1.642157
H	-3.151573	-1.432009	1.799061
H	-5.832847	1.505614	-1.160429
C	-5.630429	0.672394	0.813560
H	-5.184504	-0.331756	2.665509
H	-6.527100	1.146891	1.193851
O	1.014216	-1.928358	-0.094404
H	1.940062	-2.085982	-0.381359
C	3.707315	1.021050	0.266809
H	4.437099	1.747391	0.632498
C	2.369451	1.203204	1.075771
H	2.573462	1.645862	2.052281
H	1.891236	0.231860	1.231849
O	1.488205	2.058103	0.343705
H	1.977637	2.211287	-0.505572
C	4.204616	-0.417888	0.477149
O	5.193562	-0.626187	1.352214
H	5.553038	0.198010	1.705886
O	3.696755	-1.349665	-0.098860
N	3.443591	1.328548	-1.135325
H	4.269257	1.661342	-1.618383
H	3.064246	0.525355	-1.632787
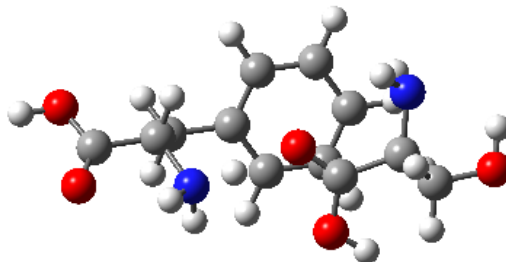
Isomer 10, 41.3 kJ·mol^{-1}

N	1.721401	1.351520	-0.476700
H	0.721383	1.272082	-0.127605
H	1.725565	1.107189	-1.469251
H	2.049274	2.322165	-0.410065
C	2.647124	0.460366	0.304899
H	2.278685	0.454226	1.333780
C	2.643498	-0.973095	-0.257536
C	4.022562	1.127772	0.296788
H	3.051111	-0.957765	-1.274141
H	3.339244	-1.554782	0.349688
C	1.264062	-1.595781	-0.244609
O	4.221987	2.204038	-0.207888
C	0.516850	-1.705559	-1.420175

C	0.712743	-2.059409	0.955388
H	0.940858	-1.381449	-2.365979
C	-0.760731	-2.265946	-1.399582
C	-0.558729	-2.624182	0.977982
H	1.287877	-1.998850	1.874005
H	-1.319291	-2.367326	-2.323176
C	-1.300970	-2.726107	-0.200564
H	-0.961095	-3.005931	1.909540
H	-2.281860	-3.188041	-0.187846
O	4.925639	0.378462	0.923027
H	5.783878	0.834368	0.924005
C	-3.090540	0.478611	0.162894
H	-3.444619	-0.289584	-0.533403
C	-4.070769	1.679149	0.042999
H	-4.032873	2.126549	-0.953623
H	-3.768880	2.449598	0.768382
O	-5.389512	1.247852	0.244281
H	-5.419493	0.773530	1.086431
C	-1.694271	0.894812	-0.274401
O	-1.482555	1.129313	-1.580502
H	-2.267200	0.944710	-2.115655
O	-0.766373	1.063454	0.495530
N	-3.170159	-0.048011	1.508065
H	-2.740262	-0.962738	1.579561
H	-2.703998	0.564710	2.170543
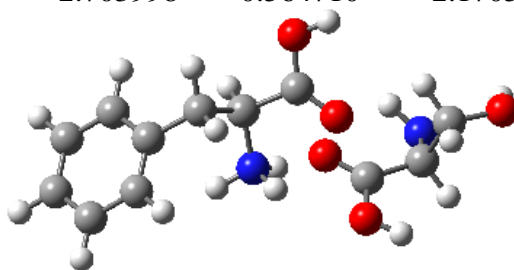
Isomer 11, 42.1 kJ·mol^{-1}

N	0.669319	-0.414936	-0.461224
H	-0.147113	-0.893089	0.007136
H	1.507512	-1.005943	-0.441560
H	0.383454	-0.267565	-1.434309
C	0.957119	0.898937	0.203821
H	0.991064	0.710885	1.278084
C	2.312384	1.471021	-0.273248
C	-0.207375	1.827274	-0.115199
H	2.232749	1.754622	-1.327400
H	2.491510	2.386190	0.294307
C	3.425539	0.466607	-0.071872

O	-1.033643	1.587754	-0.960486
C	3.911770	-0.282487	-1.150315
C	3.940128	0.227692	1.208453
H	3.545992	-0.086234	-2.154188
C	4.895040	-1.253777	-0.953062
C	4.922283	-0.739538	1.404738
H	3.586007	0.812672	2.051858
H	5.273538	-1.816916	-1.797929
C	5.398594	-1.484522	0.324756
H	5.322856	-0.906454	2.397641
H	6.166760	-2.232823	0.478916
O	-0.155863	2.924363	0.637454
H	-0.877820	3.525485	0.389877
C	-4.005969	-1.097847	0.107461
H	-4.593183	-1.709673	-0.585025
C	-4.386014	0.388348	-0.163210
H	-4.099428	0.695161	-1.171879
H	-3.831379	1.016533	0.550321
O	-5.772173	0.557563	-0.053418
H	-6.044623	0.187765	0.798250
C	-2.527219	-1.274604	-0.191407
O	-2.148665	-1.431325	-1.471525
H	-2.902793	-1.457946	-2.077101
O	-1.656960	-1.210864	0.658849
N	-4.396770	-1.427025	1.463187
H	-4.496428	-2.424050	1.612527
H	-3.732557	-1.066535	2.141156
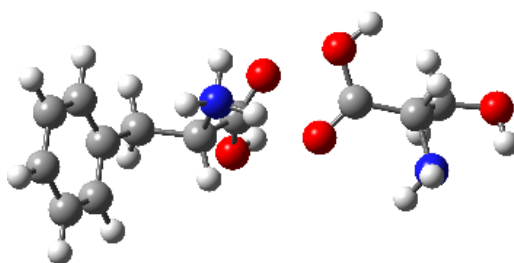
Isomer 12, 42.3 kJ·mol^{-1}

N	0.678237	-0.084985	-0.681475
H	-0.190894	-0.564382	-0.320309
H	1.454525	-0.751143	-0.760274
H	0.444252	0.280619	-1.610190
C	1.067464	1.041684	0.229842
H	0.983214	0.670195	1.252081
C	2.518785	1.497215	-0.048860
C	0.059200	2.160858	0.005096
H	2.567222	1.958703	-1.040025

H	2.756361	2.273623	0.680940
C	3.479926	0.332939	0.046665
O	-0.698522	2.186366	-0.932232
C	3.981658	-0.273691	-1.111352
C	3.832318	-0.192645	1.296072
H	3.745296	0.142450	-2.086529
C	4.820177	-1.386630	-1.022439
C	4.670116	-1.300945	1.384675
H	3.466123	0.277821	2.203507
H	5.213539	-1.838467	-1.925396
C	5.162441	-1.902136	0.225068
H	4.946978	-1.690510	2.357183
H	5.818661	-2.761277	0.296253
O	0.164570	3.091292	0.952158
H	-0.454521	3.816804	0.767219
C	-4.042541	-1.198233	0.033141
H	-4.587416	-1.826791	-0.679055
C	-4.704554	0.210968	-0.009358
H	-4.593706	0.672033	-0.994187
H	-4.193355	0.851966	0.725141
O	-6.077820	0.105037	0.244148
H	-6.185607	-0.393505	1.066326
C	-2.596583	-1.071957	-0.419741
O	-2.347712	-0.961903	-1.734265
H	-3.151598	-1.026119	-2.269107
O	-1.653738	-1.003847	0.349536
N	-4.218768	-1.744712	1.362574
H	-4.131533	-2.753574	1.388396
H	-3.555653	-1.348637	2.021248
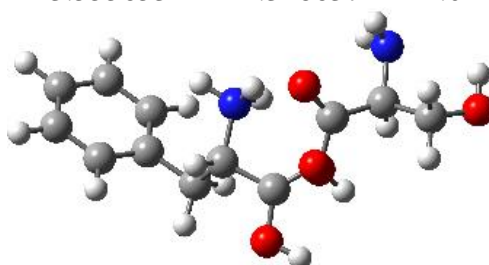
Isomer 13, 43.5 kJ·mol^{-1}

N	-0.825746	-0.548921	1.499471
H	-0.512768	-0.991122	2.365962
H	-1.592539	0.102689	1.693059
H	-0.009389	0.007910	1.124206
C	-1.273777	-1.571901	0.491735
H	-2.108155	-2.122556	0.933575
C	-1.742136	-0.869730	-0.800878

C	-0.108289	-2.534157	0.294337
H	-0.877484	-0.417538	-1.290946
H	-2.122448	-1.648665	-1.463750
C	-2.801908	0.168542	-0.503602
O	0.927193	-2.450510	0.904105
C	-2.457236	1.520967	-0.388599
C	-4.130278	-0.214544	-0.283001
H	-1.434860	1.835318	-0.577037
C	-3.423291	2.473154	-0.060167
C	-5.095043	0.735927	0.042777
H	-4.416380	-1.256932	-0.385922
H	-3.147177	3.518479	0.013653
C	-4.742435	2.081117	0.157447
H	-6.122681	0.429041	0.197641
H	-5.494988	2.820048	0.405341
O	-0.406633	-3.460870	-0.616139
H	0.335762	-4.081364	-0.700781
C	3.487829	1.200673	-0.366322
H	3.715046	1.459596	-1.406097
C	4.468834	0.059701	0.037834
H	4.291907	-0.843228	-0.552252
H	4.292679	-0.188053	1.095090
O	5.791053	0.454524	-0.203944
H	5.934345	1.296731	0.249613
C	2.063551	0.671841	-0.287168
O	1.634748	-0.128535	-1.279286
H	2.299730	-0.226292	-1.975647
O	1.316597	0.889483	0.648155
N	3.759555	2.352383	0.467141
H	3.415325	3.217858	0.069187
H	3.363404	2.245287	1.395560
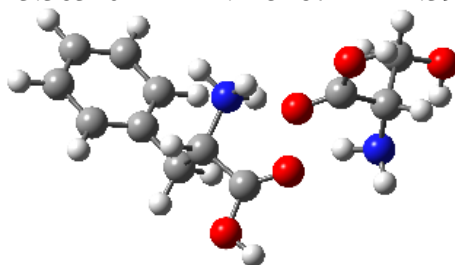
Isomer 14, 43.7 kJ·mol^{-1}

N	-0.899353	-0.781826	1.311236
H	-0.610329	-1.347716	2.111814
H	-1.475085	0.006087	1.620736
H	-0.020173	-0.391513	0.868783
C	-1.653325	-1.596918	0.297204

H	-2.590044	-1.918232	0.758987
C	-1.952322	-0.731405	-0.946104
C	-0.804467	-2.826585	-0.001634
H	-1.011074	-0.504397	-1.453139
H	-2.552351	-1.341333	-1.623315
C	-2.672914	0.541887	-0.559000
O	0.261479	-3.036673	0.519642
C	-1.968023	1.745136	-0.433487
C	-4.041084	0.519784	-0.262840
H	-0.910311	1.776407	-0.676551
C	-2.621727	2.906997	-0.020466
C	-4.693099	1.679858	0.148723
H	-4.603639	-0.402385	-0.373614
H	-2.070090	3.836280	0.060876
C	-3.983443	2.874802	0.273141
H	-5.755149	1.654371	0.362329
H	-4.492923	3.777830	0.587564
O	-1.405630	-3.610212	-0.896349
H	-0.857704	-4.395116	-1.061332
C	3.697013	0.239903	-0.220926
H	4.421893	-0.515927	0.100424
C	4.263256	1.621068	0.222854
H	4.331203	1.689764	1.311586
H	3.572901	2.404872	-0.124142
O	5.559161	1.794790	-0.279448
H	5.526275	1.652224	-1.235496
C	2.379555	-0.013842	0.495518
O	2.418541	-0.351875	1.793046
H	3.322259	-0.481673	2.113960
O	1.286683	0.126043	-0.026164
N	3.619778	0.219791	-1.666296
H	3.583076	-0.718282	-2.046256
H	2.816859	0.738441	-2.007491
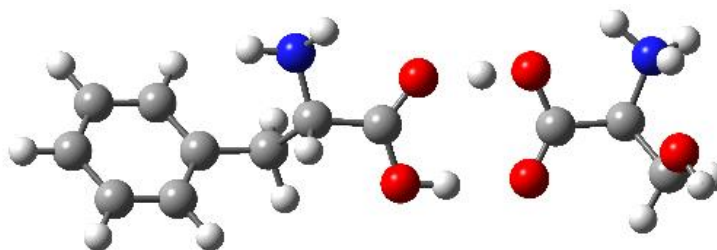
Isomer 15, 48.9 kJ·mol^{-1}

N	-1.960932	1.552328	0.919985
H	-1.416953	1.857939	1.718423
H	-2.945870	1.623693	1.150040

C	-1.638012	0.195620	0.520443
H	-1.833705	-0.567531	1.295405
C	-2.463392	-0.212057	-0.732277
C	-0.156993	0.074862	0.246726
H	-2.239605	0.499919	-1.531000
H	-2.139727	-1.202213	-1.057558
C	-3.945359	-0.219093	-0.435193
O	0.636708	0.995294	0.471270
C	-4.751426	0.874392	-0.767291
C	-4.528780	-1.313684	0.213777
H	-4.312766	1.726936	-1.274981
C	-6.111219	0.875544	-0.456018
C	-5.885996	-1.315042	0.525730
H	-3.918371	-2.175369	0.467395
H	-6.724320	1.728253	-0.724550
C	-6.680913	-0.218002	0.192534
H	-6.325088	-2.173150	1.021711
H	-7.738031	-0.219706	0.431434
O	0.208110	-1.103623	-0.214676
H	1.181035	-1.154678	-0.358163
C	5.051869	0.100242	-0.748134
H	5.137353	0.205533	-1.831384
C	5.849405	-1.101332	-0.245007
H	6.837267	-1.137534	-0.718607
H	5.290517	-2.004985	-0.499337
O	5.954885	-0.895290	1.161136
H	6.321440	-1.666452	1.606670
C	3.551870	-0.000556	-0.409383
O	3.055015	1.116939	0.038438
H	1.996813	1.043187	0.226859
O	2.962776	-1.042937	-0.611781
H	5.904403	1.092893	0.854326
H	6.413762	1.738565	-0.601542
H	4.861705	2.055803	-0.030238
N	5.610792	1.353995	-0.100364
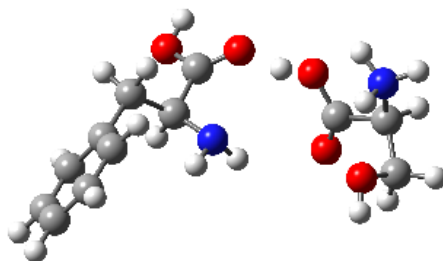
Isomer 16, 50.0 kJ·mol^{-1}

N	0.316837	0.174780	0.481374

H	-0.237435	0.250892	1.328687
H	0.835672	-0.696380	0.527966
C	1.246094	1.286887	0.350827
H	1.806238	1.514276	1.269181
C	2.288414	1.004812	-0.769048
C	0.486239	2.548682	-0.004103
H	1.751385	0.844573	-1.708112
H	2.914118	1.891950	-0.886962
C	3.141301	-0.199300	-0.441309
O	-0.677642	2.608305	-0.372371
C	2.919568	-1.430041	-1.066831
C	4.152326	-0.106129	0.522804
H	2.148706	-1.512575	-1.826436
C	3.689357	-2.545949	-0.735904
C	4.921368	-1.217834	0.855253
H	4.346036	0.845741	1.007733
H	3.512625	-3.491112	-1.236606
C	4.689992	-2.442482	0.227323
H	5.706744	-1.128212	1.596905
H	5.292864	-3.306335	0.481817
O	1.237317	3.641212	0.106503
H	0.715554	4.416791	-0.154815
C	-4.311704	-0.477556	0.104032
H	-5.154899	0.157824	0.381669
C	-4.368091	-1.818347	0.830198
H	-5.388592	-2.218620	0.833433
H	-4.037959	-1.653358	1.858423
O	-3.484422	-2.665097	0.097940
H	-3.319023	-3.490382	0.565120
C	-3.017445	0.310126	0.409683
O	-2.553585	0.897331	-0.675118
H	-1.718741	1.462389	-0.507282
O	-2.589432	0.364642	1.531012
H	-3.867252	-1.608980	-1.570107
H	-5.345758	-0.847176	-1.730240
H	-3.918169	0.034752	-1.881760
N	-4.388617	-0.736892	-1.390796

137
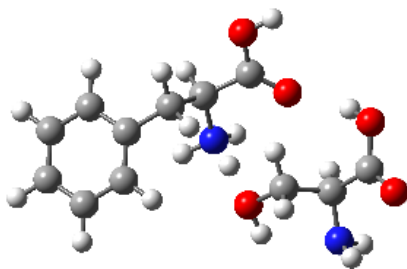
Isomer 17, 50.7 kJ·mol^{-1}

N	-0.769414	0.210162	-1.427924
H	-0.338107	0.481522	-2.313204
H	-1.661610	-0.262099	-1.610056
H	-0.131935	-0.505269	-0.950616
C	-1.027631	1.400605	-0.546136
H	-1.728558	2.054547	-1.070177
C	-1.659326	0.933180	0.787241
C	0.285715	2.144818	-0.353337
H	-0.918108	0.366475	1.357090
H	-1.895542	1.831351	1.360039
C	-2.891620	0.091685	0.535160
O	1.347673	1.745133	-0.777735
C	-2.810671	-1.306320	0.563507
C	-4.111221	0.700230	0.214973
H	-1.873101	-1.786912	0.826054
C	-3.934316	-2.082672	0.276534
C	-5.231871	-0.076436	-0.069947
H	-4.192247	1.783099	0.209473
H	-3.866137	-3.163518	0.314094
C	-5.143863	-1.468811	-0.042409
H	-6.175070	0.403603	-0.302348
H	-6.018178	-2.071254	-0.258005
O	0.099600	3.255536	0.344980
H	0.940789	3.727405	0.466589
C	3.047631	-0.957322	-0.368238
H	2.735121	-0.349396	-1.222103
C	1.856468	-1.052660	0.596035
H	1.555439	-0.080771	0.981740
H	2.104921	-1.702074	1.440521
O	0.730715	-1.601803	-0.117638
H	1.097593	-2.371578	-0.593603
C	4.215015	-0.259670	0.348319
O	4.112151	1.086507	0.437437
H	3.363554	1.417983	-0.080779
O	5.122019	-0.860218	0.846364
N	3.317236	-2.309981	-0.835791
H	3.837341	-2.307492	-1.706611
H	3.891208	-2.803557	-0.154585
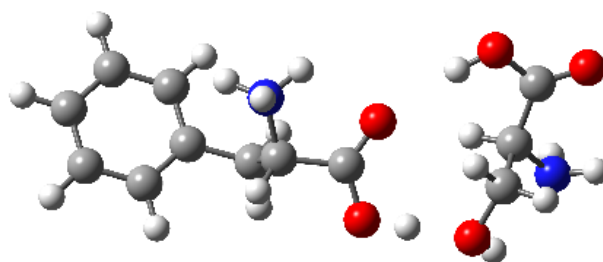
Isomer 18, 59.5 kJ·mol^{-1}

N	-1.640123	0.862791	1.347883
H	-1.625441	0.809260	2.368382
H	-2.560221	1.219629	1.043867
H	-0.884699	1.505072	1.063820
C	-1.370162	-0.479129	0.699150
H	-1.768612	-1.251072	1.356480
C	-2.068074	-0.537585	-0.681174
C	0.149980	-0.607460	0.572226
H	-1.555146	0.138241	-1.372426
H	-1.929492	-1.552534	-1.058018
C	-3.530782	-0.172961	-0.564683
O	0.855890	0.387151	0.685137
C	-3.966040	1.116652	-0.900343
C	-4.453938	-1.089886	-0.046837
H	-3.272510	1.824098	-1.347215
C	-5.302032	1.483201	-0.719377
C	-5.785811	-0.724528	0.128872
H	-4.135164	-2.097755	0.200518
H	-5.630634	2.477422	-0.997883
C	-6.210548	0.563891	-0.201868
H	-6.495324	-1.446359	0.515738
H	-7.248721	0.843187	-0.068355
O	0.510141	-1.821033	0.307871
H	1.534992	-1.960991	0.190454
C	3.852595	-0.024196	-0.569751
H	2.847816	0.139567	-0.971877
C	3.718315	-1.014364	0.595167
H	3.145443	-0.610196	1.427939
H	4.705012	-1.313754	0.958418
O	3.005167	-2.177209	0.130764
H	3.419768	-2.430179	-0.713160
C	4.380233	1.323132	-0.051396
O	3.481839	2.037455	0.669807
H	2.624954	1.586070	0.692110
O	5.496451	1.711615	-0.234801
N	4.655948	-0.669148	-1.595669

139

H	4.531994	-0.226947	-2.499927
H	5.644465	-0.589495	-1.366968
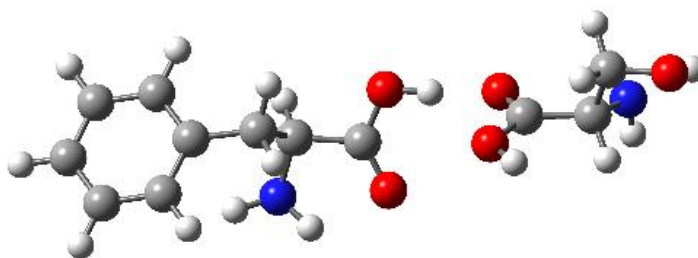
Isomer 19, 59.7 kJ·mol⁻¹

N	-2.465386	-1.825379	0.540526
H	-2.610292	-2.731514	0.092284
H	-3.373464	-1.472861	0.876663
H	-1.794818	-1.966941	1.318306
C	-1.817768	-0.822073	-0.391385
H	-2.069415	-1.099736	-1.414165
C	-2.332661	0.601189	-0.074850
C	-0.309715	-0.948066	-0.139704
H	-1.925236	0.923908	0.888068
H	-1.913493	1.259866	-0.837553
C	-3.843428	0.647080	-0.056374
O	0.095791	-1.541062	0.841457
C	-4.535527	0.630565	1.162196
C	-4.572542	0.628861	-1.251824
H	-3.985960	0.690311	2.097865
C	-5.932111	0.595477	1.184852
C	-5.964086	0.597499	-1.228376
H	-4.050621	0.658798	-2.203301
H	-6.456107	0.599387	2.133317
C	-6.646131	0.575515	-0.010096
H	-6.518338	0.598813	-2.159537
H	-7.729232	0.556102	0.004629
O	0.383659	-0.328802	-1.050356
H	1.381843	-0.370528	-0.869029
C	5.096489	-0.044244	0.309940
H	5.349380	-0.163245	1.369324
C	5.762515	1.277778	-0.166408
H	5.358733	2.141836	0.367682
H	5.544490	1.409919	-1.236801
O	7.140684	1.243863	0.092615
H	7.485224	0.423808	-0.288099
C	3.583098	0.076655	0.179853
O	2.937143	0.855578	1.058620
H	3.534595	1.211189	1.730944

140

O	2.949639	-0.457693	-0.710715
N	5.682917	-1.142644	-0.431014
H	5.560001	-2.034272	0.033862
H	5.281247	-1.215222	-1.360961
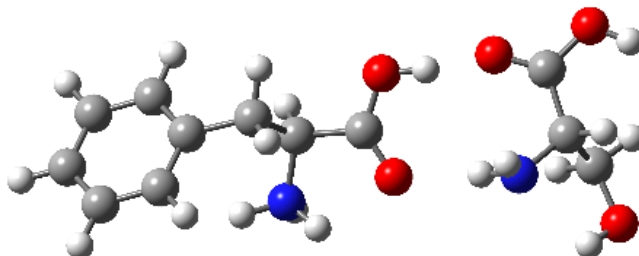
Isomer 20, 65.8 kJ·mol⁻¹

N	1.850367	-1.300823	1.019044
H	1.784484	-1.455324	2.026832
H	2.808618	-1.524783	0.709313
H	1.154149	-1.916974	0.567061
C	1.499440	0.121198	0.633329
H	1.751366	0.768245	1.472787
C	2.300438	0.535961	-0.624078
C	-0.008406	0.127868	0.370947
H	1.923608	-0.018184	-1.489241
H	2.083074	1.591417	-0.797561
C	3.778959	0.286547	-0.433827
O	-0.611418	-0.922511	0.231818
C	4.378947	-0.854052	-0.985189
C	4.551488	1.146948	0.356131
H	3.805784	-1.505482	-1.639662
C	5.728085	-1.130107	-0.749319
C	5.896817	0.872989	0.586906
H	4.105224	2.043703	0.774825
H	6.185071	-2.005818	-1.194734
C	6.485747	-0.268584	0.039264
H	6.489357	1.553586	1.186787
H	7.534110	-0.475980	0.217048
O	-0.476407	1.336524	0.283737
H	-1.466262	1.358609	0.091723
C	-4.533967	-0.258613	-0.561405
H	-5.410451	-0.264135	-1.218803
C	-4.955670	-0.943489	0.765872
H	-5.751066	-0.384472	1.264368
H	-4.082942	-0.961651	1.437157
O	-5.449363	-2.233820	0.512294
H	-4.821233	-2.665997	-0.083599
C	-4.146948	1.191754	-0.272570

141

O	-5.137137	2.073461	-0.119546
H	-6.002303	1.675882	-0.288807
O	-3.003712	1.576712	-0.128171
N	-3.499661	-1.066465	-1.175674
H	-3.354767	-0.839111	-2.151821
H	-2.613968	-0.994994	-0.682885
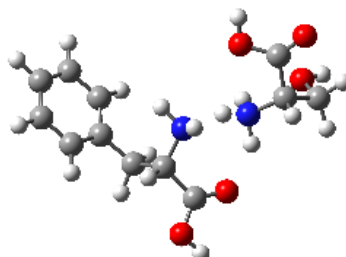
Isomer 21, 111.1 kJ·mol^{-1}

N	-0.378448	0.311559	-0.916625
H	-0.127240	0.395011	-1.899657
H	-0.930299	-0.543328	-0.846391
C	-1.241875	1.445723	-0.527241
H	-1.932667	1.718013	-1.331027
C	-2.088225	1.087686	0.726893
C	-0.402204	2.679333	-0.218175
H	-1.413722	0.878233	1.563035
H	-2.673678	1.969639	0.997049
C	-3.000045	-0.097225	0.487941
O	0.758107	2.679572	0.144728
C	-2.714978	-1.346194	1.050340
C	-4.144644	0.036390	-0.308743
H	-1.850019	-1.461732	1.696733
C	-3.549997	-2.440320	0.818738
C	-4.978061	-1.053883	-0.541585
H	-4.397469	1.002075	-0.735823
H	-3.322432	-3.398121	1.272181
C	-4.680785	-2.296391	0.019607
H	-5.865499	-0.932476	-1.151711
H	-5.334278	-3.142553	-0.156602
O	-1.117630	3.795921	-0.355651
H	-0.574321	4.558760	-0.098475
C	3.323103	-0.024476	-0.287377
H	3.354111	0.666969	-1.132524
C	4.537772	0.201567	0.613738
H	4.640513	1.268891	0.843307
H	5.430038	-0.140101	0.084065
O	4.275332	-0.558681	1.787659
H	5.057433	-0.607033	2.347201

142

C	3.336863	-1.446647	-0.858419
O	2.170174	-2.093307	-0.662332
H	2.247942	-2.987069	-1.035414
O	4.286582	-1.897082	-1.428228
H	2.098069	-0.286617	1.359488
H	1.995311	1.256849	0.714900
H	1.170002	0.080709	-0.047149
N	2.076361	0.255592	0.491543
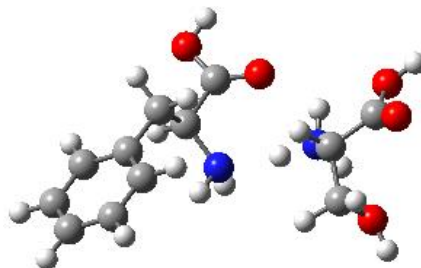
Isomer 22, 111.9 kJ·mol^{-1}

N	0.624527	0.523658	1.314522
H	0.590255	0.820147	2.288439
H	1.186350	-0.327791	1.307849
C	1.342912	1.542211	0.518029
H	2.184007	1.963596	1.077228
C	1.903987	0.923169	-0.791371
C	0.410020	2.694910	0.167317
H	1.069163	0.562986	-1.400945
H	2.391964	1.722253	-1.354190
C	2.877083	-0.206686	-0.527427
O	-0.801288	2.638233	0.134994
C	2.520891	-1.535077	-0.784080
C	4.153910	0.059043	-0.015570
H	1.548936	-1.757947	-1.214521
C	3.414650	-2.576730	-0.530716
C	5.046465	-0.978266	0.239282
H	4.460964	1.084008	0.168640
H	3.128213	-3.599259	-0.748039
C	4.677634	-2.299783	-0.014889
H	6.034153	-0.755578	0.625793
H	5.376061	-3.105575	0.177278
O	1.099851	3.794766	-0.143178
H	0.481533	4.496103	-0.405479
C	-2.594490	-0.821516	-0.197947
H	-1.862030	-0.887757	-1.006301
C	-2.653338	-2.144919	0.565346
H	-1.640458	-2.490629	0.802417
H	-3.148622	-2.884787	-0.066866

143

O	-3.388784	-1.860509	1.750574
H	-3.681685	-2.671226	2.179151
C	-3.947682	-0.504162	-0.849822
O	-4.307786	0.779203	-0.665485
H	-5.151591	0.941872	-1.118557
O	-4.565394	-1.316089	-1.473542
H	-2.678081	0.161858	1.613624
H	-1.103343	0.234757	0.944427
H	-2.306850	1.198181	0.356157
N	-2.153052	0.259093	0.738416
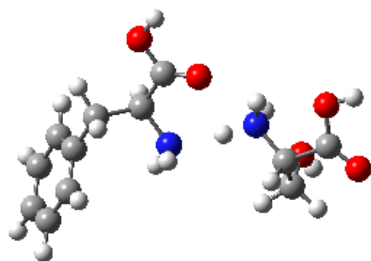
Isomer 23, 113.1 kJ·mol^{-1}

N	0.470545	-0.012909	-0.262553
H	1.015332	-0.849542	-0.058305
H	0.443081	0.043029	-1.280862
C	1.227181	1.149949	0.252006
H	1.388313	0.983729	1.322919
C	2.615179	1.342973	-0.416796
C	0.356026	2.397920	0.149755
H	2.467400	1.568638	-1.478406
H	3.090338	2.218295	0.030392
C	3.497972	0.122724	-0.263292
O	-0.797712	2.409375	-0.225223
C	3.726791	-0.740865	-1.339726
C	4.095339	-0.169108	0.969924
H	3.298660	-0.515460	-2.311870
C	4.528108	-1.873207	-1.188796
C	4.894321	-1.299388	1.123595
H	3.955139	0.502731	1.811706
H	4.705965	-2.524485	-2.036759
C	5.110270	-2.156355	0.044193
H	5.359661	-1.504117	2.080842
H	5.739343	-3.030826	0.161520
O	1.007097	3.493452	0.545817
H	0.412264	4.257679	0.475350
C	-3.143465	-0.865099	-0.399645
H	-2.738722	-0.931433	-1.412838
C	-3.117558	-2.240812	0.266479

H	-2.139338	-2.713837	0.121906
H	-3.889616	-2.860460	-0.194226
O	-3.368548	-1.992420	1.645670
H	-3.600967	-2.804784	2.106934
C	-4.582038	-0.344931	-0.531066
O	-4.680795	0.953660	-0.189954
H	-5.596723	1.249236	-0.321924
O	-5.478515	-1.027357	-0.930738
H	-2.475596	-0.070253	1.381764
H	-1.227761	-0.074738	0.201205
H	-2.448122	1.051741	0.147463
N	-2.276534	0.069920	0.385598
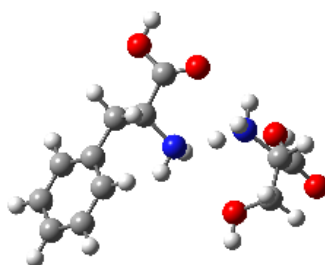
Isomer 24, 117.7 kJ·mol^{-1}

N	-0.022468	0.341103	-0.529753
H	0.399647	0.404345	-1.453662
H	-0.568686	-0.519440	-0.530267
C	-0.925938	1.487321	-0.305722
H	-1.439898	1.776799	-1.227380
C	-2.007086	1.141475	0.754958
C	-0.128800	2.696300	0.170221
H	-1.512366	0.906782	1.702774
H	-2.609892	2.037474	0.919469
C	-2.887477	-0.014970	0.330349
O	0.963920	2.657923	0.695461
C	-2.761354	-1.271458	0.932108
C	-3.844070	0.153802	-0.679176
H	-2.046813	-1.413579	1.737371
C	-3.567449	-2.338477	0.532407
C	-4.648021	-0.909311	-1.080184
H	-3.975395	1.126527	-1.143311
H	-3.467177	-3.302222	1.018298
C	-4.509560	-2.159848	-0.476742
H	-5.390583	-0.760247	-1.855401
H	-5.141365	-2.984657	-0.784404
O	-0.806124	3.830152	-0.028700
H	-0.294039	4.573812	0.328367
C	3.290926	-0.786832	0.747553

145

H	4.116022	-0.630975	1.449483
C	2.778503	-2.214961	0.902307
H	2.534270	-2.410694	1.954439
H	3.581292	-2.890565	0.597604
O	1.627516	-2.336424	0.076630
H	1.446672	-3.265884	-0.098320
C	3.872437	-0.559175	-0.652021
O	3.478654	0.616808	-1.182497
H	3.930775	0.739124	-2.033812
O	4.622555	-1.331460	-1.171116
H	1.310288	0.135898	0.446069
H	1.916393	0.102306	2.044522
H	2.501202	1.175102	0.943612
N	2.212515	0.200210	1.072190
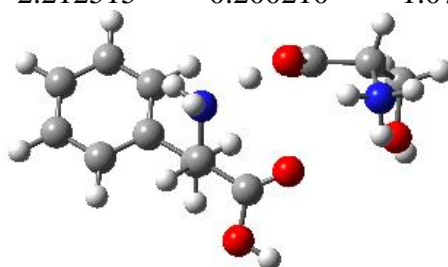
Isomer 25, 118.6 kJ·mol⁻¹

N	0.609520	0.687011	1.585455
H	0.585537	1.247122	2.435797
H	1.405556	0.053838	1.677050
C	0.847450	1.550684	0.406055
H	1.628559	2.285696	0.620447
C	1.308868	0.696322	-0.809057
C	-0.412563	2.325092	0.036754
H	0.511271	-0.001907	-1.074636
H	1.454069	1.375945	-1.651664
C	2.585258	-0.065035	-0.517938
O	-1.551218	2.024117	0.324340
C	2.545614	-1.434004	-0.227214
C	3.823047	0.590994	-0.521985
H	1.595336	-1.958436	-0.244172
C	3.720699	-2.131021	0.058919
C	4.994626	-0.104902	-0.236826
H	3.875173	1.647772	-0.766582
H	3.677567	-3.192667	0.273282
C	4.945208	-1.468029	0.057206
H	5.946685	0.412818	-0.253396
H	5.858220	-2.010061	0.273884
O	-0.100392	3.397321	-0.701061

146

H	-0.916267	3.852949	-0.962022
C	-3.472987	-1.265823	0.391691
H	-3.914986	-1.943659	1.124879
C	-3.922420	-1.613396	-1.023679
H	-4.976048	-1.914286	-1.036004
H	-3.302944	-2.441464	-1.374592
O	-3.725185	-0.430345	-1.794641
H	-3.744127	-0.625484	-2.736717
C	-1.933724	-1.279248	0.586465
O	-1.607397	-0.571961	1.642835
H	-0.615462	-0.117916	1.594705
O	-1.201660	-1.874314	-0.164464
H	-3.831709	0.713173	-0.123386
H	-4.892020	0.189838	1.059481
H	-3.276028	0.514166	1.436131
N	-3.927566	0.141658	0.728691
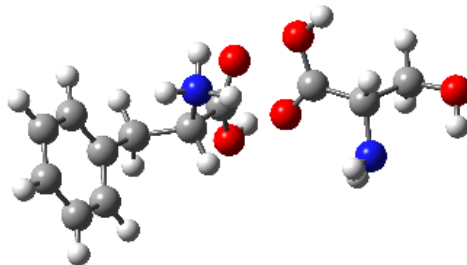
Isomer 26, 135.5 kJ·mol^{-1}

N	0.683671	0.486346	-0.828021
H	-0.249084	0.081404	-0.547980
H	1.326487	-0.278403	-1.066271
C	1.270888	1.323709	0.272847
H	0.953947	0.887726	1.220462
C	2.818226	1.347451	0.185566
C	0.676762	2.719351	0.110895
H	3.118728	1.906498	-0.705913
H	3.176698	1.908269	1.051187
C	3.406321	-0.047115	0.153256
O	0.132120	3.079463	-0.903134
C	3.853699	-0.599941	-1.053813
C	3.485319	-0.817487	1.320886
H	3.843337	-0.001777	-1.960780
C	4.365129	-1.898766	-1.094515
C	3.997407	-2.111232	1.280123
H	3.166966	-0.397963	2.270346
H	4.722219	-2.307806	-2.032305
C	4.434331	-2.655830	0.071557
H	4.067181	-2.691228	2.192852

147

H	4.839452	-3.660327	0.044071
O	0.876035	3.462793	1.195577
H	0.527173	4.357238	1.044451
C	-4.017191	-1.202592	0.036301
H	-4.460893	-1.984616	-0.589307
C	-4.996984	0.008619	0.003515
H	-5.093742	0.415472	-1.006054
H	-4.592425	0.799495	0.652209
O	-6.275531	-0.398080	0.404515
H	-6.183915	-0.846211	1.257075
C	-2.681558	-0.791456	-0.569899
O	-2.591368	-0.714077	-1.905646
H	-3.405231	-0.998256	-2.345859
O	-1.706631	-0.481900	0.093231
H	0.526655	1.084889	-1.648913
H	-3.641790	-2.657552	1.454446
H	-3.326035	-1.126652	1.970047
N	-3.944506	-1.692481	1.397717
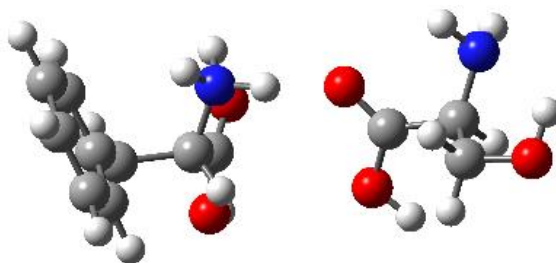
Isomer 27, 135.7 kJ·mol^{-1}

N	-0.810526	0.711493	1.103571
H	0.203931	0.442726	1.014246
H	-1.355585	-0.124265	1.347914
H	-0.881479	1.412502	1.852452
C	-1.337976	1.313367	-0.168472
H	-0.790708	0.862454	-0.996541
C	-2.857496	1.046642	-0.319154
C	-1.031933	2.805660	-0.093391
H	-3.399649	1.619982	0.439059
H	-3.154477	1.442066	-1.292627
C	-3.189184	-0.425499	-0.199756
O	-0.726362	3.355668	0.935531
C	-3.729744	-0.936048	0.987657
C	-2.932360	-1.305252	-1.259550
H	-3.979242	-0.262869	1.803209
C	-4.002764	-2.299582	1.115972
C	-3.207346	-2.663860	-1.132063
H	-2.537718	-0.924924	-2.196832

148

H	-4.434724	-2.678191	2.034823
C	-3.738718	-3.164284	0.057638
H	-3.018581	-3.331597	-1.964449
H	-3.958101	-4.221113	0.151653
O	-1.184881	3.395389	-1.276043
H	-1.030468	4.350373	-1.183208
C	4.104202	-0.261624	0.153260
H	4.742525	0.510047	-0.290882
C	4.294161	-1.549143	-0.703744
H	3.951218	-1.398566	-1.730502
H	3.692248	-2.354486	-0.257778
O	5.652178	-1.883413	-0.772623
H	5.986905	-1.928457	0.134052
C	2.653752	0.197147	0.067894
O	2.258192	0.831390	-1.050257
H	2.994471	0.982394	-1.660298
O	1.825060	-0.021662	0.931793
N	4.580118	-0.533209	1.493544
H	4.829159	0.306412	2.002732
H	3.898599	-1.050776	2.039547
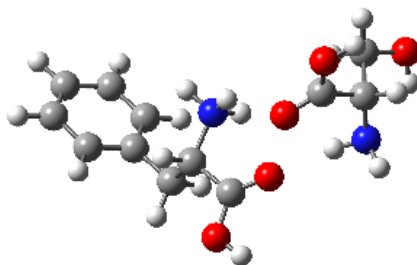
Isomer 28, 142.1 kJ·mol^{-1}

N	0.839461	0.792827	1.227261
H	0.583802	1.358954	2.039529
H	1.315871	-0.061281	1.532262
H	-0.066016	0.511004	0.756725
C	1.734717	1.557914	0.290163
H	2.647427	1.797802	0.842524
C	2.089142	0.689443	-0.934456
C	1.019700	2.867840	-0.029405
H	1.180062	0.488268	-1.508047
H	2.744571	1.288431	-1.569319
C	2.761605	-0.607326	-0.528134
O	-0.023847	3.188090	0.479984
C	2.045549	-1.810539	-0.518899
C	4.106072	-0.616762	-0.133832
H	1.012066	-1.826438	-0.851876
C	2.659546	-2.998552	-0.119016

149

C	4.718465	-1.802245	0.264774
H	4.686521	0.300713	-0.161178
H	2.097997	-3.925394	-0.129645
C	3.994977	-2.995092	0.275886
H	5.762704	-1.797805	0.554305
H	4.474752	-3.918030	0.579054
O	1.712714	3.592409	-0.907103
H	1.251178	4.432318	-1.066930
C	-3.888482	-0.200203	-0.174280
H	-4.649867	0.483321	0.217088
C	-4.353083	-1.643044	0.185588
H	-4.393532	-1.789888	1.267678
H	-3.624229	-2.355690	-0.227880
O	-5.645204	-1.867987	-0.305323
H	-5.641036	-1.648610	-1.247563
C	-2.565898	0.097610	0.520084
O	-2.593152	0.386639	1.830507
H	-3.495919	0.450259	2.173660
O	-1.481596	0.038132	-0.033067
N	-3.862777	-0.076525	-1.617220
H	-3.924499	0.884273	-1.932239
H	-3.029306	-0.494796	-2.017844
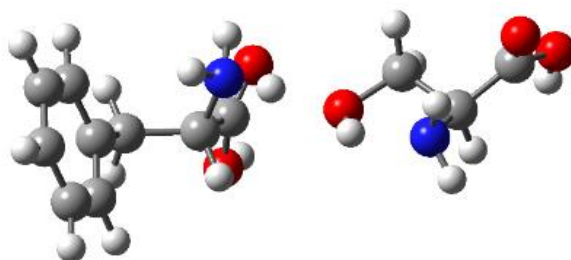
Isomer 29, 145.1 kJ·mol^{-1}

N	0.824202	0.520355	-0.789919
H	-0.127218	0.212610	-0.431999
H	1.382322	-0.307030	-1.032404
H	0.669033	1.094299	-1.627744
C	1.558275	1.347218	0.226788
H	1.288183	0.969562	1.213362
C	3.091133	1.247864	0.019622
C	1.053536	2.776669	0.060464
H	3.360439	1.747947	-0.915824
H	3.559110	1.810708	0.829822
C	3.566955	-0.189268	0.002999
O	0.432554	3.138671	-0.908406
C	3.868935	-0.822619	-1.209681
C	3.682579	-0.917919	1.194203

150

H	3.830162	-0.261146	-2.139063
C	4.272973	-2.159199	-1.232382
C	4.087797	-2.249539	1.171127
H	3.478587	-0.437430	2.146287
H	4.518259	-2.631696	-2.176214
C	4.379410	-2.873944	-0.042621
H	4.188209	-2.797599	2.100584
H	4.701510	-3.908300	-0.057041
O	1.416552	3.543914	1.084387
H	1.121155	4.456384	0.927326
C	-3.929761	-0.560184	0.600079
H	-4.045932	0.083907	1.476917
C	-2.867512	0.068852	-0.343376
H	-2.994974	1.147492	-0.441011
H	-2.934547	-0.391930	-1.334882
O	-1.575350	-0.186542	0.226510
H	-1.716816	-1.040238	0.705973
C	-5.259604	-0.698979	-0.161808
O	-6.034613	0.401525	-0.245468
H	-5.713565	1.117812	0.317021
O	-5.566861	-1.718520	-0.710031
N	-3.398562	-1.850394	1.038127
H	-3.726243	-2.095136	1.966063
H	-3.715408	-2.581612	0.402352
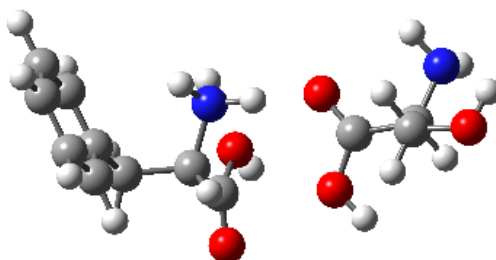
Isomer 30, 148.1 kJ·mol^{-1}

N	-0.781199	-0.682590	-1.085561
H	0.247718	-0.452217	-1.050902
H	-1.310429	0.178944	-1.273190
H	-0.937262	-1.338428	-1.855654
C	-1.257978	-1.274216	0.218069
H	-0.762871	-0.716022	1.012136
C	-2.794941	-1.132329	0.346223
C	-0.833482	-2.732476	0.345501
H	-3.282051	-1.770175	-0.397738
H	-3.064548	-1.523658	1.329562
C	-3.242476	0.304985	0.184579
O	-0.767094	-3.305075	1.392191

151

C	-3.812927	0.741955	-1.018178
C	-3.066283	1.227904	1.224249
H	-3.999499	0.031903	-1.819135
C	-4.195574	2.074983	-1.180894
C	-3.450009	2.556242	1.062314
H	-2.650445	0.902712	2.173025
H	-4.649936	2.395303	-2.111051
C	-4.011494	2.983151	-0.142022
H	-3.323038	3.256462	1.879570
H	-4.316103	4.015868	-0.262761
O	-0.603607	-3.292262	-0.865552
H	-0.392063	-4.233115	-0.749104
C	4.097743	0.314694	-0.091390
H	4.681277	-0.372017	0.531530
C	4.247677	1.731190	0.541793
H	3.820936	1.765090	1.547353
H	3.701320	2.450873	-0.085375
O	5.602819	2.061021	0.663072
H	6.009373	1.948718	-0.207634
C	2.635505	-0.109207	-0.045972
O	2.134911	-0.529215	1.128581
H	2.812526	-0.583914	1.817911
O	1.882155	-0.041517	-1.000643
N	4.683994	0.343539	-1.414933
H	4.958320	-0.573244	-1.747083
H	4.059288	0.760835	-2.097544
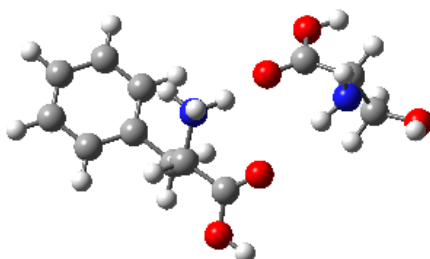
Isomer 31, 149.8 kJ·mol^{-1}

N	0.748012	0.612294	1.414056
H	0.519294	1.056369	2.305060
H	1.512062	-0.060066	1.553143
H	-0.094998	0.072126	1.100617
C	1.169571	1.620409	0.379105
H	1.974749	2.216312	0.815325
C	1.693216	0.891366	-0.880243
C	-0.027140	2.528887	0.119760
H	0.871386	0.336862	-1.341728
H	2.001724	1.663379	-1.587408

152

C	2.844236	-0.037724	-0.550172
O	-1.107281	2.366925	0.632464
C	2.625599	-1.411808	-0.384903
C	4.139013	0.469435	-0.379049
H	1.634775	-1.826763	-0.545325
C	3.682003	-2.260916	-0.049961
C	5.191775	-0.379127	-0.046847
H	4.332043	1.527484	-0.529017
H	3.502738	-3.323750	0.062085
C	4.964173	-1.745428	0.121504
H	6.191283	0.023121	0.069005
H	5.785962	-2.405151	0.373192
O	0.289846	3.502996	-0.728798
H	-0.482522	4.075019	-0.872907
C	-3.634271	-0.927809	0.349372
H	-4.357542	-1.651082	0.742034
C	-4.100902	-0.561487	-1.089459
H	-4.091688	-1.434063	-1.746790
H	-3.401777	0.181268	-1.503078
O	-5.417123	-0.081216	-1.063214
H	-5.461104	0.598583	-0.375756
C	-2.264740	-1.603754	0.285343
O	-2.221327	-2.871199	-0.132118
H	-3.105544	-3.233258	-0.285535
O	-1.212597	-1.047321	0.542558
N	-3.703343	0.267027	1.163966
H	-3.759275	0.071970	2.155798
H	-2.937093	0.909881	0.988883
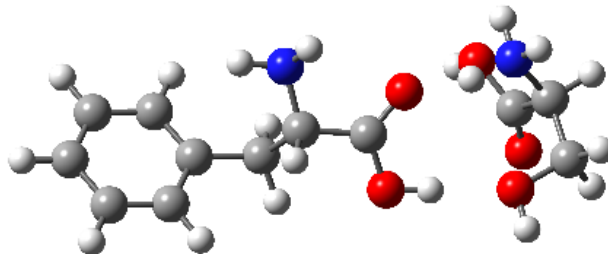
Isomer 32, 151.3 kJ·mol^{-1}

N	-1.544660	0.349648	1.867129
H	-1.152863	-0.028709	2.722045
H	2.594480	-0.020729	1.356044
H	-2.546603	0.448693	1.992085
C	-1.269956	-0.494406	0.718536
H	-1.626108	-1.535030	0.825751
C	-1.942247	0.085018	-0.557768
C	0.225772	-0.631693	0.533663

153

H	-1.564186	1.099791	-0.708999
H	-1.640281	-0.519251	-1.414950
C	-3.451750	0.108375	-0.447033
O	1.053365	-0.213779	1.342005
C	-4.131274	1.287493	-0.123470
C	-4.196090	-1.058994	-0.658325
H	-3.571283	2.204234	0.030100
C	-5.521432	1.299695	-0.007407
C	-5.583680	-1.049436	-0.542438
H	-3.689211	-1.980340	-0.929683
H	-6.033018	2.223253	0.238249
C	-6.250188	0.131069	-0.214166
H	-6.145664	-1.959815	-0.716879
H	-7.330505	0.140197	-0.128814
O	0.578942	-1.280288	-0.567000
H	1.551509	-1.390896	-0.584507
C	4.489037	0.197668	0.311728
H	5.478237	0.576801	0.588478
C	4.654741	-1.245767	-0.140406
H	5.197361	-1.813438	0.624124
H	5.245497	-1.233773	-1.059469
O	3.363819	-1.817917	-0.358668
H	3.467903	-2.701720	-0.732020
C	3.980971	1.117003	-0.811149
O	3.296375	2.156432	-0.299470
H	3.026357	2.756731	-1.014536
O	4.227911	0.938380	-1.965263
N	3.609578	0.308821	1.529540
H	3.989530	-0.230824	2.310046
H	3.533799	1.286677	1.822238
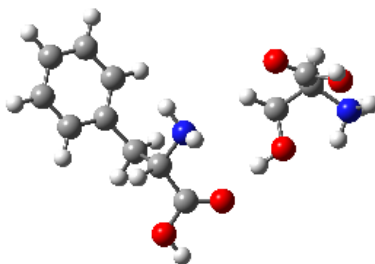
Isomer 33, 155.3 kJ·mol^{-1}

N	-0.416558	0.217700	-0.716124
H	0.041731	0.283164	-1.620643
H	-0.942818	-0.652288	-0.721459
C	-1.346910	1.328232	-0.526690
H	-1.974376	1.525620	-1.406591
C	-2.300559	1.045814	0.668690

154

C	-0.581530	2.614295	-0.262009
H	-1.694548	0.893879	1.566597
H	-2.915519	1.933642	0.830573
C	-3.182807	-0.160914	0.429680
O	0.598682	2.710723	0.019196
C	-2.931549	-1.375567	1.076106
C	-4.268511	-0.083127	-0.452116
H	-2.108440	-1.447851	1.780224
C	-3.740026	-2.489068	0.844260
C	-5.076233	-1.192535	-0.685834
H	-4.493835	0.855676	-0.948873
H	-3.536475	-3.420391	1.360370
C	-4.812216	-2.400449	-0.039472
H	-5.917422	-1.112896	-1.364783
H	-5.444686	-3.262253	-0.217408
O	-1.370271	3.687900	-0.362206
H	-0.847283	4.480772	-0.163199
C	3.732854	-0.943865	0.801657
H	3.832578	-1.752008	1.529082
C	2.459091	-0.119169	1.027605
H	2.301788	0.035535	2.103317
H	1.596910	-0.645769	0.613950
O	2.707146	1.099678	0.357623
H	1.876523	1.617101	0.218057
C	3.775479	-1.573117	-0.596354
O	4.990352	-1.425378	-1.170591
H	4.998513	-1.864551	-2.037186
O	2.850046	-2.154226	-1.074278
H	4.526712	0.927429	0.631838
H	5.166658	0.121214	1.958418
H	5.708087	-0.268525	0.423025
N	4.896604	0.016591	0.978621
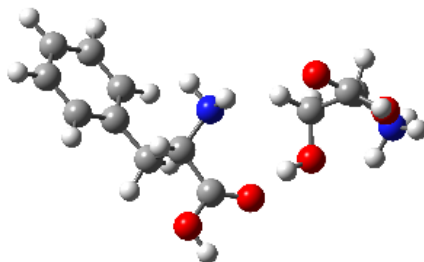
Isomer 34, 155.5 kJ·mol⁻¹

N	0.370655	-0.069491	0.625684
H	-0.142671	-0.005786	1.500731
H	1.061694	-0.805898	0.745781
C	1.068599	1.181159	0.336877

155

H	1.593328	1.596424	1.206984
C	2.122097	0.973550	-0.789349
C	0.083923	2.246953	-0.114296
H	1.605974	0.612862	-1.683512
H	2.560903	1.945098	-1.028468
C	3.207860	-0.002163	-0.388825
O	-1.029215	2.062199	-0.571793
C	3.232397	-1.301271	-0.906918
C	4.207933	0.378479	0.515213
H	2.478812	-1.605591	-1.626570
C	4.228286	-2.202351	-0.528016
C	5.201872	-0.518897	0.895571
H	4.219230	1.388419	0.913699
H	4.238483	-3.202259	-0.946637
C	5.213109	-1.813801	0.376270
H	5.973520	-0.205748	1.589408
H	5.990689	-2.510032	0.667879
O	0.590957	3.475013	0.025639
H	-0.045218	4.119330	-0.323683
C	-3.626164	-1.468722	-0.210295
H	-3.786374	-2.535070	-0.036465
C	-2.330821	-1.199693	-1.010292
H	-2.221580	-1.964418	-1.785737
H	-1.464711	-1.228320	-0.346337
O	-2.477681	0.065169	-1.624420
H	-1.871094	0.730109	-1.218146
C	-3.631179	-0.776321	1.156307
O	-4.804244	-0.150837	1.395825
H	-4.789709	0.247604	2.281901
O	-2.709972	-0.828323	1.914617
H	-4.315430	-0.122144	-1.562585
H	-5.065153	-1.615832	-1.768973
H	-5.551232	-0.623989	-0.511482
N	-4.753916	-0.936041	-1.072457
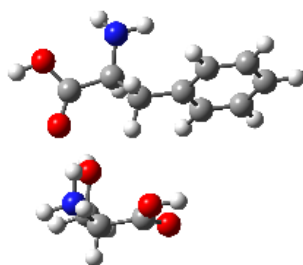
Isomer 35, 163.9 kJ·mol^{-1}

N	-1.914580	2.665711	0.790217
H	-1.692919	3.295657	1.552547

H	-2.692556	2.079551	1.072725
C	-0.786853	1.848810	0.385384
H	-0.365254	1.223058	1.192416
C	-1.210944	0.899167	-0.765448
C	0.396199	2.687505	-0.067294
H	-1.634048	1.510920	-1.566031
H	-0.316881	0.410080	-1.158299
C	-2.215946	-0.147456	-0.331504
O	1.438374	2.212968	-0.509569
C	-3.531767	-0.113239	-0.802991
C	-1.845981	-1.176633	0.545984
H	-3.836774	0.675811	-1.482220
C	-4.455147	-1.083662	-0.412169
C	-2.765731	-2.146724	0.937814
H	-0.834263	-1.217590	0.940354
H	-5.470554	-1.041432	-0.788846
C	-4.074741	-2.102557	0.457310
H	-2.462842	-2.934988	1.617635
H	-4.791420	-2.856555	0.760753
O	0.240853	3.994259	0.076886
H	1.048422	4.444947	-0.219452
C	3.284829	-1.132486	-0.364050
H	4.100994	-1.652121	-0.874592
C	3.696412	-0.822226	1.069109
H	4.670282	-0.315817	1.077813
H	3.790036	-1.774156	1.598508
O	2.675300	0.002325	1.605994
H	2.784881	0.089134	2.558892
C	2.064125	-2.062641	-0.419135
O	1.253807	-1.744317	-1.448077
H	0.497563	-2.355868	-1.468913
O	1.907473	-2.973292	0.336219
H	2.315905	0.787089	-0.660003
H	3.878568	0.689493	-1.232587
H	2.649627	-0.061697	-2.053649
N	3.020093	0.145036	-1.122494
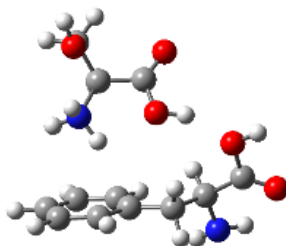
Isomer 36, 183.9 kJ·mol^{-1}

N	3.760729	0.516547	0.866056
H	4.393755	-0.103898	1.357880
H	3.501545	1.285365	1.471436
C	2.608968	-0.181488	0.322298
H	1.845660	-0.442553	1.076325
C	1.937143	0.660168	-0.789377
C	3.063879	-1.538253	-0.199252
H	2.744272	1.184095	-1.310334
H	1.467256	0.005709	-1.526498
C	0.899500	1.678272	-0.341404
O	4.169183	-1.985704	-0.156511
C	0.232768	2.425164	-1.329564
C	0.570635	1.929357	0.997179
H	0.483439	2.265371	-2.373659
C	-0.717707	3.388392	-0.993885
C	-0.383620	2.893665	1.339397
H	1.069965	1.392691	1.795752
H	-1.193844	3.968556	-1.776516
C	-1.033491	3.626511	0.347043
H	-0.594958	3.088426	2.385101
H	-1.747732	4.397369	0.613321
O	2.005646	-2.277089	-0.700676
H	2.361707	-3.135541	-0.982159
C	-2.560407	-0.717680	0.670770
H	-2.418262	-0.377746	1.698354
C	-3.976667	-1.257692	0.472123
H	-4.712868	-0.587685	0.931262
H	-4.035455	-2.238323	0.949866
O	-4.147578	-1.322967	-0.939795
H	-4.951557	-1.798687	-1.174397
C	-1.489564	-1.790634	0.400785
O	-0.551700	-1.339571	-0.435266
H	0.228165	-1.939284	-0.511507
O	-1.523101	-2.857914	0.939230
H	-2.835855	0.257947	-1.134606
H	-2.758844	1.333517	0.140933
H	-1.366320	0.645778	-0.427963
N	-2.366045	0.470361	-0.243600
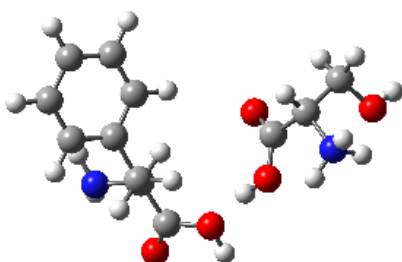
Isomer 37, 194.5 kJ·mol⁻¹

N	-3.073607	1.863891	0.859029
H	-3.115995	2.642048	1.506932
H	-3.512173	1.052725	1.278487
C	-1.716664	1.566079	0.440199
H	-1.089255	1.109704	1.226901
C	-1.741054	0.594172	-0.768945
C	-1.022064	2.871054	0.107409
H	-2.313648	1.067578	-1.569684
H	-0.722822	0.443929	-1.138064
C	-2.356149	-0.746468	-0.421872
O	-1.440531	3.976348	0.268173
C	-3.593861	-1.121601	-0.952154
C	-1.701845	-1.633511	0.443769
H	-4.118265	-0.442982	-1.616104
C	-4.163016	-2.354641	-0.632939
C	-2.269586	-2.863995	0.766681
H	-0.746668	-1.356433	0.880980
H	-5.123757	-2.627995	-1.053838
C	-3.502854	-3.229083	0.226585
H	-1.754178	-3.536495	1.443492
H	-3.946158	-4.185580	0.477839
O	0.270221	2.690674	-0.399579
H	0.620826	3.583933	-0.548893
C	3.314442	-1.185393	-0.335093
H	2.756788	-2.078048	-0.626083
C	4.331138	-1.503960	0.761730
H	4.844252	-2.449571	0.554222
H	3.792519	-1.583141	1.708315
O	5.240389	-0.407481	0.737773
H	5.835765	-0.422142	1.495045
C	2.284755	-0.126730	0.107536
O	2.058279	0.764762	-0.856927
H	1.353958	1.431277	-0.607435
O	1.753747	-0.186553	1.178218
H	4.886734	-0.155296	-1.198253
H	4.408038	-1.443321	-2.150270

159

H	3.472670	-0.050500	-2.092856
N	4.073083	-0.684827	-1.552270

Appendix B

Proton-Bound homodimers of Phenylalanine Derivatives

B.1   The *XYZ* atomic coordinates of global minima of [(x-Phe)₂ + H]⁺

Table B.1 The *XYZ* Cartesian coordinates of global minima

3-CN-Phe				3-F-Phe			
C	4.059845	0.910242	0.453768	C	-4.278099	0.734768	0.024394
C	3.456757	-0.270881	0.875202	C	-4.806589	0.051627	-1.077095
C	4.013625	-1.513536	0.540768	C	-6.145622	0.213601	-1.391724
C	5.200281	-1.588015	-0.199035	C	-6.993126	1.031315	-0.656889
C	5.81597	-0.407758	-0.599506	C	-6.461903	1.712725	0.433078
C	5.249692	0.825461	-0.280417	C	-5.11546	1.567145	0.773407
H	2.541902	-0.238296	1.452741	H	-4.193949	-0.589372	-1.7012
H	5.625589	-2.55102	-0.451482	H	-8.031976	1.129346	-0.945041
H	6.741203	-0.44803	-1.161235	H	-7.098779	2.366128	1.017301
H	5.751396	1.734139	-0.598403	H	-4.713417	2.125358	1.612511
C	3.296577	-2.688671	0.922785	C	-2.823219	0.558292	0.400159
N	2.627433	-3.584592	1.212739	H	-2.181345	0.674378	-0.475967
C	3.456056	2.259182	0.789328	H	-2.531397	1.331606	1.117936
H	2.848224	2.184797	1.696076	C	-2.528799	-0.833802	1.017795
H	4.267545	2.957502	1.012006	H	-2.763968	-1.601826	0.27506
C	2.615912	2.922129	-0.337963	C	-1.028574	-0.942953	1.307619
H	3.128079	2.762045	-1.292328	N	-3.241902	-1.052765	2.284261
C	1.221945	2.288937	-0.527786	H	-3.990335	-0.379928	2.418065
N	2.412754	4.369783	-0.163449	H	-3.657773	-1.976736	2.334945
H	2.494517	4.658522	0.806911	O	-0.714027	-1.008322	2.583341
H	3.06096	4.924222	-0.710257	O	-0.173733	-0.950029	0.429186
O	1.021301	1.083438	-0.556064	H	1.409951	-1.122058	0.748826
O	0.244781	3.155735	-0.693042	F	-6.639945	-0.447586	-2.459168
H	0.681833	4.048312	-0.606045	C	4.585129	0.519558	-0.269441
C	-3.883654	-1.082215	-0.350538	C	4.114156	1.516299	-1.132817
C	-4.228709	-0.283597	0.743085	C	4.181284	2.840606	-0.73
C	-5.260913	0.656425	0.63783	C	4.697957	3.222639	0.501602
C	-5.966551	0.796691	-0.566592	C	5.167115	2.229504	1.354006
C	-5.631196	-0.002982	-1.651537	C	5.110629	0.88531	0.975757

C	-4.595512	-0.933205	-1.547146
H	-3.713497	-0.392288	1.691291
H	-6.768853	1.52009	-0.63819
H	-6.181475	0.090415	-2.579947
H	-4.360943	-1.567057	-2.396647
C	-5.585249	1.476553	1.766042
N	-5.827674	2.143436	2.676967
C	-2.748065	-2.080602	-0.243155
H	-2.869046	-2.714991	0.638359
H	-2.746944	-2.744715	-1.112552
C	-1.367336	-1.402036	-0.118218
H	-1.314469	-0.80945	0.794952
C	-0.220491	-2.418639	-0.146368
N	-1.103691	-0.472908	-1.271376
H	-1.959378	0.007534	-1.562218
H	-0.361809	0.235281	-1.018422
O	-0.079343	-3.028281	1.012266
O	0.43106	-2.590486	-1.150634
H	-0.721482	-1.03448	-2.044449
H	0.758603	-3.556699	1.032178

H	3.719214	1.284437	-2.115648
H	4.737162	4.271571	0.766604
H	5.590674	2.502788	2.312938
H	5.517205	0.126844	1.63717
C	4.492499	-0.938635	-0.665916
H	4.818488	-1.087601	-1.697383
H	5.147818	-1.542654	-0.031434
C	3.044366	-1.479868	-0.565173
H	2.40075	-0.971544	-1.283648
C	2.994266	-2.987023	-0.799618
N	2.467229	-1.246212	0.80039
H	2.660731	-2.06892	1.385415
H	2.880923	-0.408795	1.224192
O	2.830267	-3.779122	0.094605
O	3.189649	-3.287531	-2.08081
H	3.191163	-4.253139	-2.190136
F	3.730778	3.790331	-1.569744
H	-1.599235	-1.008372	3.051714

4-F-Phe			
C	-2.906436	0.77872	-0.634673
C	-3.777341	-0.30373	-0.820319
C	-5.14181	-0.107629	-0.571232
C	-5.62977	1.122051	-0.131425
C	-4.731553	2.16076	0.052401
C	-3.374557	2.014482	-0.198107
H	-1.850906	0.662634	-0.854763
H	-5.844341	-0.915893	-0.748352
H	-6.684553	1.282233	0.053165
H	-2.711698	2.860119	-0.063517
C	-3.262428	-1.653371	-1.272839
C	-2.726193	-2.547127	-0.123567
C	-1.369746	-2.063148	0.411501
N	-3.619148	-2.706774	1.031888
O	-0.422776	-1.766902	-0.309392
O	-1.275428	-2.025458	1.722259

4-NO$_2$-Phe			
C	-4.593062	-0.002073	0.39359
C	-5.229557	0.585877	-0.707084
C	-5.171772	1.9633	-0.91444
C	-4.468691	2.740165	-0.005347
C	-3.837438	2.190598	1.105868
C	-3.902059	0.816168	1.298183
H	-5.806218	-0.026213	-1.393584
H	-3.422073	0.382115	2.168855
C	-4.628791	-1.503468	0.586513
H	-4.862173	-1.756923	1.622804
H	-5.412974	-1.945744	-0.035169
C	-3.271454	-2.169969	0.257356
H	-2.493236	-1.821525	0.936594
C	-3.36271	-3.693338	0.317653
N	-2.818986	-1.839174	-1.140295
H	-3.09238	-2.614901	-1.760408

H	-4.101543	-3.597274	1.038333
H	-4.309149	-1.961994	1.083477
H	-2.181113	-2.317879	2.037532
H	-2.45963	-1.534621	-2.003723
H	-4.064639	-2.207289	-1.76617
H	-2.513505	-3.535523	-0.545062
F	-5.189753	3.355792	0.474781
C	3.466069	1.658915	-1.219634
C	3.840036	0.964789	-0.060219
C	3.814983	1.642157	1.165735
C	3.417246	2.976996	1.243776
C	3.04639	3.623947	0.075878
C	3.069054	2.989951	-1.16045
H	3.503075	1.164805	-2.185236
H	4.142574	1.139607	2.070886
H	3.408058	3.513391	2.184192
H	2.791663	3.540338	-2.050705
C	4.242482	-0.492771	-0.126996
C	3.023953	-1.431598	-0.319426
C	3.427327	-2.899447	-0.217421
N	1.987922	-1.203123	0.741194
O	3.177223	-3.577814	0.747585
O	4.104562	-3.294457	-1.292729
H	2.174738	-1.835755	1.529503
H	1.007484	-1.418167	0.379874
H	4.381716	-4.218838	-1.178608
H	4.774851	-0.779616	0.784982
H	4.920925	-0.676729	-0.962697
H	2.552976	-1.242905	-1.284592
H	2.015596	-0.228722	1.059376
F	2.659878	4.909843	0.137536

H	-1.770724	-1.735852	-1.192627
O	-3.330818	-4.384028	-0.670785
O	-3.508191	-4.123981	1.566052
H	-3.591603	-5.092377	1.568564
H	-3.23939	-0.96181	-1.46209
H	-3.316044	2.837289	1.798537
H	-5.665029	2.433298	-1.754558
O	-4.952166	4.657628	-1.210504
O	-3.763384	4.861985	0.596595
N	-4.390162	4.2093	-0.223589
C	4.71112	-0.857308	-0.28999
C	3.996089	-0.991904	-1.492681
C	2.622737	-1.168304	-1.497123
C	1.954428	-1.212534	-0.26895
C	2.633354	-1.082956	0.947027
C	4.006682	-0.904783	0.924783
H	4.528547	-0.96855	-2.436744
H	4.547887	-0.788807	1.855092
C	6.206802	-0.665726	-0.317011
H	6.645369	-1.045198	0.608862
H	6.636831	-1.228886	-1.147773
C	6.65578	0.798038	-0.508613
H	6.165498	1.225162	-1.387163
C	6.291194	1.65996	0.703925
N	8.097969	0.825621	-0.75797
H	8.412148	1.748836	-1.036861
O	6.146508	1.256915	1.832919
O	6.216943	2.966185	0.371567
H	6.073614	3.477641	1.18356
H	8.626736	0.543156	0.062012
H	2.088394	-1.123687	1.879922
H	2.072368	-1.277568	-2.420865
O	-0.074459	-1.513026	-1.341235
O	-0.072113	-1.450203	0.822494
N	0.525461	-1.401705	-0.248964

2,5-F2-Phe			
C	-4.565312	2.439132	-0.893756

3,4-(MeO)2-Phe			
C	3.123814	-1.392652	-0.311958

C	-3.777522	1.666741	-0.061378	C	3.663448	-1.094111	-1.557401
C	-4.235577	0.548799	0.629742	C	4.9811	-0.639066	-1.672508
C	-5.580283	0.206425	0.445573	C	5.781936	-0.47197	-0.546002
C	-6.379052	0.969526	-0.393047	C	5.246154	-0.780399	0.73383
C	-5.89992	2.083331	-1.066921	C	3.929149	-1.227277	0.833473
H	-6.558906	2.656169	-1.706359	H	3.074489	-1.23043	-2.458319
C	-3.359813	-0.244658	1.58111	H	3.53518	-1.506783	1.803489
H	-4.001379	-0.884031	2.187943	C	1.685393	-1.83838	-0.163579
H	-2.844646	0.425968	2.275955	H	1.420437	-2.584569	-0.91605
C	-2.296334	-1.169585	0.937846	H	1.53024	-2.303224	0.814808
H	-1.936689	-1.860292	1.705038	C	0.687669	-0.665565	-0.326876
N	-1.095867	-0.422516	0.448083	H	0.700029	-0.294649	-1.349806
H	-0.418406	-1.033244	-0.10629	C	-0.735009	-1.07096	0.081163
H	-1.367067	0.388998	-0.11307	N	1.073916	0.480623	0.57301
F	-2.45473	2.024128	0.092497	H	0.572541	0.35282	1.463929
F	-7.6669	0.614443	-0.550564	H	2.0901	0.475655	0.726643
H	-4.141196	3.301808	-1.391304	O	-1.124347	-0.873186	1.217746
H	-6.012796	-0.642362	0.960804	O	-1.391659	-1.636421	-0.893877
O	-2.359588	-1.974226	-1.335052	H	-4.067413	-3.036409	-0.231702
C	-2.829313	-2.001069	-0.227553	H	0.785098	1.401291	0.168755
O	-3.857998	-2.758966	0.160849	H	5.381129	-0.427772	-2.654757
H	-4.180711	-3.271848	-0.598488	O	6.090527	-0.61879	1.774502
H	-0.514996	-0.099842	1.231711	O	7.060563	-0.04469	-0.555157
C	6.556821	0.101046	-0.824749	C	5.662132	-0.993588	3.080009
C	5.275365	-0.350573	-1.086372	H	4.816313	-0.382014	3.413587
C	4.128601	0.329632	-0.681703	H	5.392212	-2.053681	3.119773
C	4.305652	1.52515	0.020428	H	6.514243	-0.812054	3.731217
C	5.588496	1.980071	0.287787	C	7.685225	0.235619	-1.807331
C	6.720393	1.292102	-0.120911	H	8.699771	0.543007	-1.564193
H	7.703751	1.684624	0.103509	H	7.715675	-0.655724	-2.441407
C	2.75662	-0.207392	-1.010638	H	7.173651	1.048671	-2.332154
H	2.716948	-0.4713	-2.072431	C	-3.829785	0.962026	-0.763708
H	2.008001	0.570883	-0.83988	C	-3.121567	1.135713	-1.951432
C	2.349966	-1.481446	-0.222311	C	-1.883826	1.779355	-1.942105
H	3.120719	-2.241069	-0.369389	C	-1.351131	2.247173	-0.75167
N	1.022973	-1.941711	-0.691789	C	-2.079965	2.13475	0.447963
H	1.005786	-1.93435	-1.709675	C	-3.320372	1.502926	0.424283
H	0.88218	-2.914881	-0.427197	H	-3.513967	0.745478	-2.88325

F	5.116619	-1.5136	-1.774819
F	5.732249	3.136959	0.967817
H	7.408408	-0.470883	-1.170453
H	3.456609	2.112043	0.350574
O	1.327128	-0.651375	1.829045
C	2.276815	-1.165948	1.270533
O	3.400973	-1.495114	1.901264
H	3.341767	-1.214427	2.829644

H	-3.878421	1.381671	1.34375
C	-5.101741	0.145783	-0.717864
H	-5.894718	0.699068	-0.210308
H	-5.456459	-0.055139	-1.734073
C	-4.934499	-1.197544	0.034662
H	-4.507169	-0.988299	1.020378
C	-6.322419	-1.784342	0.325114
N	-4.002836	-2.113147	-0.656365
H	-4.303231	-2.2443	-1.621298
O	-7.190917	-1.176981	0.890965
O	-6.455871	-3.059355	-0.103252
H	-7.347097	-3.36273	0.133451
H	-2.397614	-1.793383	-0.690244
H	-1.30538	1.890981	-2.851992
O	-1.463188	2.606194	1.567009
O	-0.036766	2.707302	-0.717401
C	-2.178196	2.570091	2.808462
H	-1.522152	3.039237	3.538738
H	-3.109442	3.138451	2.738789
H	-2.387975	1.54035	3.110138
C	0.132188	4.143162	-0.626265
H	-0.330286	4.619868	-1.492705
H	-0.307855	4.519211	0.297993
H	1.205578	4.326043	-0.635541

3,5-F$_2$-Phe			
C	2.898444	0.985519	-0.422006
C	3.575025	1.514842	0.682612
C	3.10418	2.689339	1.262047
C	1.985498	3.357399	0.794593
C	1.337736	2.801107	-0.300363
C	1.762908	1.635208	-0.919776
H	4.481161	1.063702	1.070404
H	1.639414	4.27434	1.252634
H	1.217064	1.268319	-1.781093
C	3.376932	-0.307461	-1.048712
H	3.408745	-0.225508	-2.137073
H	4.393021	-0.534333	-0.713526

3-CF$_3$-Phe			
C	2.850354	2.501055	1.30466
C	2.223357	2.309144	0.065662
C	2.730391	1.336842	-0.795
C	4.443036	0.752189	0.812836
C	3.948991	1.732622	1.673118
H	2.488272	3.270608	1.979418
H	2.286519	1.193859	-1.774434
H	4.430181	1.899755	2.629201
C	1.019893	3.137118	-0.327174
C	-0.290088	2.662886	0.367606
C	-1.45923	3.576732	0.006746
N	-0.643887	1.25538	0.104989

C	2.447784	-1.49972	-0.712781	O	-2.537535	3.212912	-0.42126
H	1.465293	-1.342048	-1.157968	O	-1.171677	4.859205	0.228688
C	3.033928	-2.821337	-1.206781	H	0.875583	3.093163	-1.412112
N	2.255418	-1.654537	0.769234	H	1.182901	4.185683	-0.071861
H	1.26482	-1.968032	0.994093	H	-0.157331	2.769569	1.44974
H	2.417889	-0.765257	1.253554	C	4.315244	-0.502234	-1.371907
O	3.503613	-3.641454	-0.458277	H	5.300292	0.155087	1.09492
O	2.983005	-2.911544	-2.532093	C	3.830731	0.560932	-0.42159
H	3.394121	-3.744831	-2.817386	F	4.446317	-0.031483	-2.631381
F	3.765574	3.194167	2.315626	F	3.422577	-1.535277	-1.446781
H	2.912751	-2.370097	1.108761	F	5.492867	-1.031483	-1.007989
F	0.24804	3.424644	-0.780756	H	-0.492432	1.049357	-0.881865
C	-3.279449	-0.922316	-0.167182	H	-2.392952	0.614479	0.314321
C	-2.2072	-0.729217	-1.046581	H	0.012612	0.659279	0.605816
C	-2.200672	0.384322	-1.869546	H	-1.942502	5.402926	-0.002404
C	-3.216745	1.326017	-1.880675	C	-1.892105	-2.476305	-1.583745
C	-4.273275	1.099215	-1.012053	C	-2.380786	-2.128765	-0.316619
C	-4.329144	0.002507	-0.159393	C	-1.499442	-2.118399	0.763407
H	-1.399974	-1.446589	-1.10797	C	0.333624	-2.761268	-0.684259
H	-3.194912	2.185104	-2.537355	C	-0.550336	-2.796934	-1.763317
H	-5.203607	-0.130664	0.466629	H	-2.56898	-2.519079	-2.431362
C	-3.27955	-2.104729	0.777803	H	-1.860582	-1.893818	1.761674
H	-4.30469	-2.383394	1.031819	H	-0.1885	-3.082323	-2.743741
H	-2.818606	-2.970374	0.298324	C	-3.836518	-1.764596	-0.122104
C	-2.527373	-1.842568	2.108761	C	-4.172742	-0.312926	-0.536903
H	-2.599847	-2.753956	2.711702	C	-5.665417	-0.032175	-0.361651
C	-1.018304	-1.636967	1.893636	N	-3.449915	0.695168	0.301544
N	-3.011163	-0.711734	2.911607	O	-6.102091	0.633414	0.541683
H	-3.610999	-0.998656	3.675751	O	-6.386058	-0.639981	-1.303886
H	-3.506964	-0.024875	2.350273	H	-4.130746	-1.91239	0.922018
O	-0.331238	-2.344943	1.168479	H	-4.474149	-2.41637	-0.723281
O	-0.496614	-0.639497	2.576982	H	-3.884858	-0.141222	-1.575335
H	-1.276224	-0.243132	3.064332	C	0.784421	-2.292477	1.752243
F	-1.145997	0.557459	-2.702945	H	1.379832	-2.999868	-0.822883
F	-5.290142	1.97781	-0.995718	C	-0.148292	-2.420776	0.576461
				F	1.042933	-0.966148	2.002417
				F	1.966184	-2.886818	1.553803
				F	0.249665	-2.78801	2.884614

166

H	-3.823956	0.68717	1.254425
H	-3.568662	1.656353	-0.056773
H	-7.330454	-0.473459	-1.145987

F$_5$-Phe				Phe			
N	-3.508664	2.917525	-2.590051	N	-4.302015	-2.055057	-1.890941
H	-4.214884	3.610277	-2.367961	H	-4.742962	-1.717623	-2.739998
H	-3.843866	2.369263	-3.374723	H	-5.013641	-2.508054	-1.328397
C	-3.138779	2.09258	-1.429931	C	-3.597307	-0.99206	-1.157376
H	-3.769611	2.314348	-0.565422	H	-3.923204	0.010413	-1.447312
C	-3.257235	0.578222	-1.75883	C	-3.810647	-1.169668	0.372056
C	-1.700897	2.46759	-1.037295	C	-2.110781	-1.096922	-1.514837
H	-4.278233	0.418064	-2.112899	H	-4.89293	-1.118677	0.532557
H	-2.582414	0.335529	-2.582952	H	-3.497635	-2.181743	0.647953
C	-2.998169	-0.353356	-0.601945	C	-3.122241	-0.158761	1.261877
O	-1.143171	2.049007	-0.031331	O	-1.305109	-0.206262	-1.288628
C	-1.789748	-1.022572	-0.439674	C	-2.062669	-0.542804	2.08803
C	-3.979243	-0.594657	0.357661	C	-3.557225	1.170865	1.301918
C	-1.558144	-1.888832	0.618901	H	-1.728682	-1.575728	2.088928
C	-3.781873	-1.457344	1.428363	C	-1.444727	0.381083	2.930957
C	-2.559014	-2.108793	1.559144	C	-2.938868	2.097295	2.136909
O	-1.092969	3.280341	-1.875982	H	-4.395174	1.485562	0.68736
H	-1.785592	3.491647	-2.565216	H	-0.640578	0.060332	3.583938
N	1.447119	2.690536	0.437229	C	-1.878674	1.704729	2.954465
H	1.520048	3.672392	0.748602	H	-3.295081	3.120703	2.163475
H	1.975276	2.581402	-0.433286	H	-1.410281	2.421575	3.619086
C	1.996979	1.831694	1.547024	O	-1.751258	-2.235307	-2.075143
H	1.210305	1.139627	1.853014	H	-2.605027	-2.729399	-2.210145
C	3.256892	1.041606	1.127623	N	1.319209	0.109467	-1.667953
C	2.276956	2.787108	2.714386	H	1.504875	0.391042	-2.638308
H	4.007567	1.736255	0.73948	H	1.88447	-0.711051	-1.424467
H	3.669795	0.599323	2.033818	C	1.653323	1.257029	-0.760138
C	2.99733	-0.059147	0.122527	H	1.000611	1.184462	0.110399
O	2.142581	3.981734	2.617831	C	3.141415	1.215984	-0.328994
C	2.853726	-1.386683	0.525766	C	1.339976	2.526676	-1.544469
C	2.93107	0.183397	-1.246291	H	3.772095	1.427576	-1.197953
C	2.664015	-2.420317	-0.384348	H	3.289647	2.032941	0.380229
C	2.73714	-0.818525	-2.183758	C	3.526025	-0.114235	0.282685

167

C	2.604071	-2.133194	-1.74472	O	1.205261	2.535753	-2.742854
O	2.67049	2.131761	3.798986	C	3.149309	-0.435465	1.593202
H	2.837554	2.76135	4.520522	C	4.244285	-1.056897	-0.464037
H	0.427523	2.474851	0.233193	H	2.611283	0.29056	2.194759
F	3.049697	1.464836	-1.68986	C	3.482381	-1.67078	2.141601
F	2.667212	-0.532563	-3.481205	C	4.575294	-2.297301	0.084964
F	2.406753	-3.107079	-2.622298	H	4.582829	-0.809598	-1.466318
F	2.536048	-3.672694	0.035196	H	3.199085	-1.900942	3.162061
F	2.904399	-1.685692	1.830647	C	4.19239	-2.605944	1.387151
F	-0.791735	-0.823498	-1.325781	H	5.143551	-3.010856	-0.500044
F	-0.37464	-2.499447	0.746176	H	4.456367	-3.564088	1.81872
F	-2.344912	-2.931178	2.579833	O	1.285665	3.592638	-0.748764
F	-4.744366	-1.65996	2.323601	H	1.12797	4.385962	-1.287001
F	-5.162742	0.032149	0.25035	H	0.288014	-0.135755	-1.589941

B.2 The comparison between the experimental IRMPD and calculated spectra of global minima of $[(x\text{-}Phe)_2 + H]^+$

Figure B.1 The experimental IRMPD spectrum (black) and the calculated IR spectra (blue) for $[3\text{-}CN\text{-}Phe_2 + H]^+$ global minimum. The calculations were conducted at B3LYP/6-311++G(d,p) level of theory.

Figure B.2 The experimental IRMPD spectrum (black) and the calculated IR spectra (blue) for $[3\text{-}F\text{-}Phe_2 + H]^+$ global minimum. The calculations were conducted at B3LYP/6-311++G(d,p) level of theory.
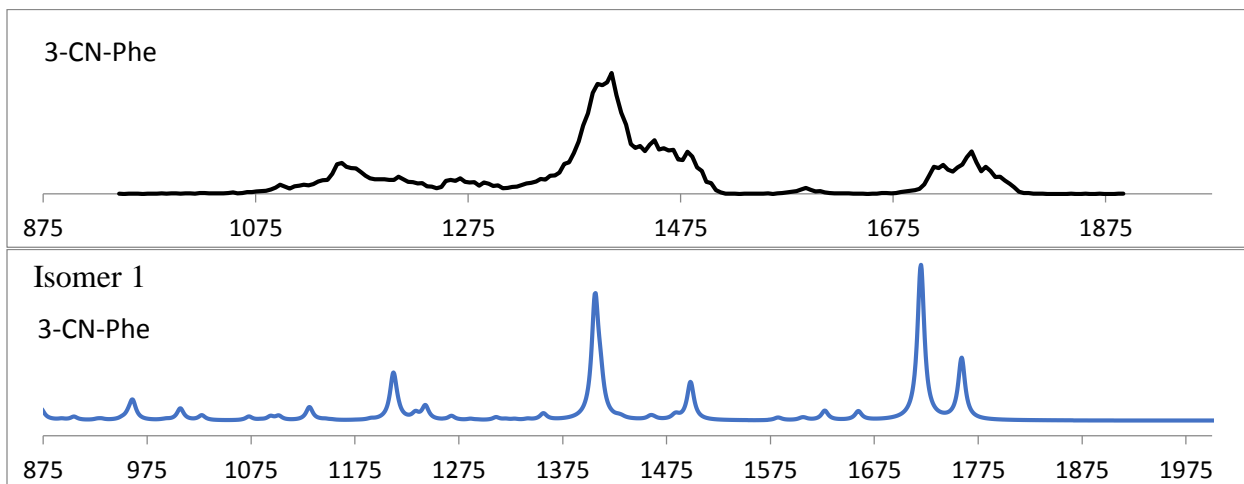
169

Figure B.3 The experimental IRMPD spectrum (black) and the calculated IR spectra (blue) for $[4\text{-F-Phe}_2+\text{H}]^+$ global minimum. The calculations were conducted at B3LYP/6-311++G(d,p) level of theory.

Figure B.4 The experimental IRMPD spectrum (black) and the calculated IR spectra (blue) for $[4\text{-NO}_2\text{-Phe}_2+\text{H}]^+$ global minimum. The calculations were conducted at B3LYP/6-311++G(d,p) level of theory.

Figure B.5 The experimental IRMPD spectrum (black) and the calculated IR spectra (blue) for [2,5-F$_2$-Phe$_2$+H]$^+$ global minimum. The calculations were conducted at B3LYP/6-311++G(d,p) level of theory.
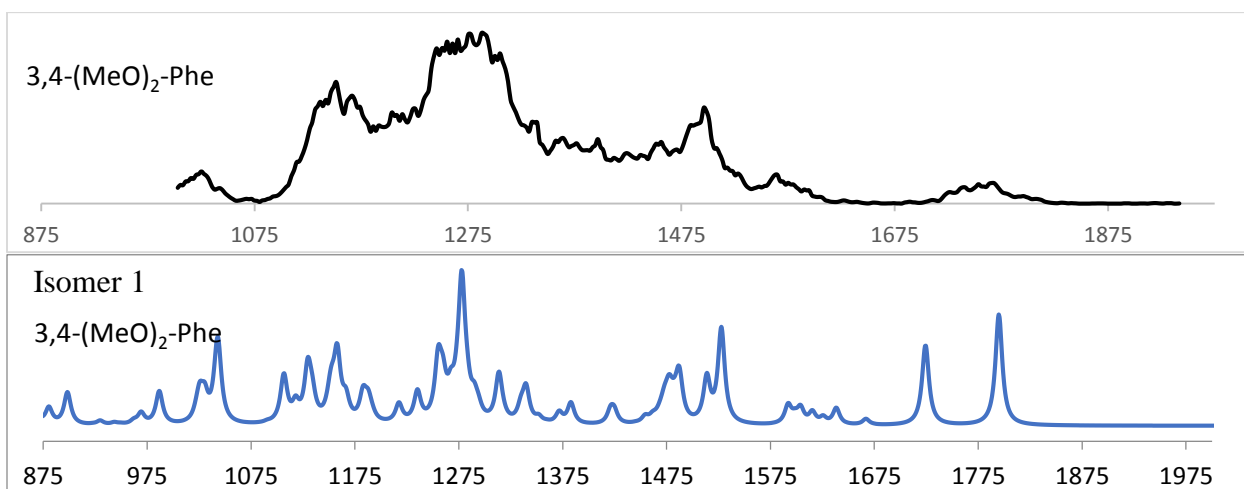
Figure B.6 The experimental IRMPD spectrum (black) and the calculated IR spectra (blue) for [3,4-(MeO)$_2$-Phe$_2$+H]$^+$ global minimum. The calculations were conducted at B3LYP/6-311++G(d,p) level of theory.
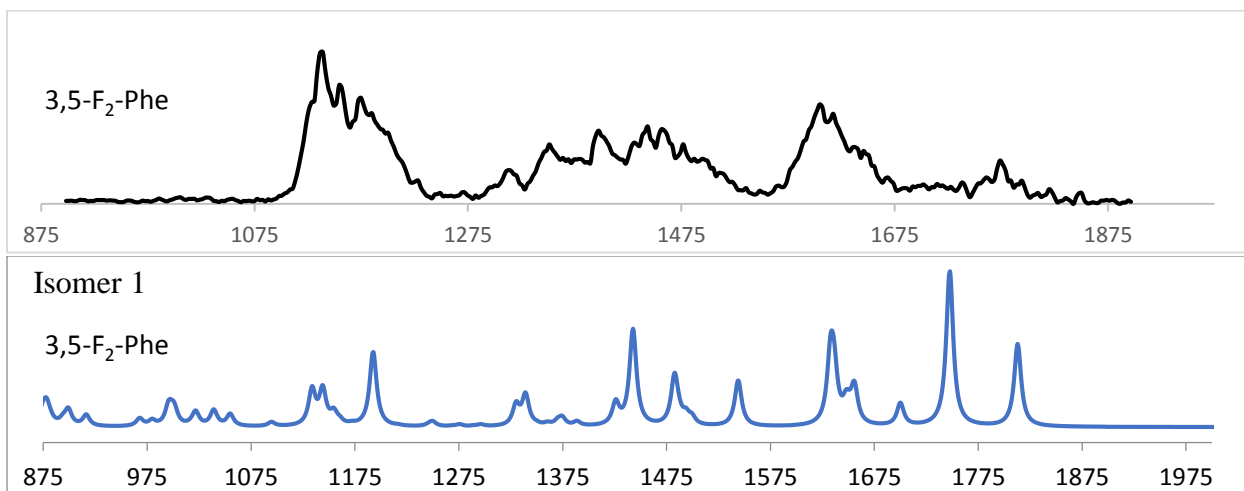
Figure B.7 The experimental IRMPD spectrum (black) and the calculated IR spectra (blue) for $[3,5\text{-}F_2\text{-}Phe_2+H]^+$ global minimum. The calculations were conducted at B3LYP/6-311++G(d,p) level of theory.
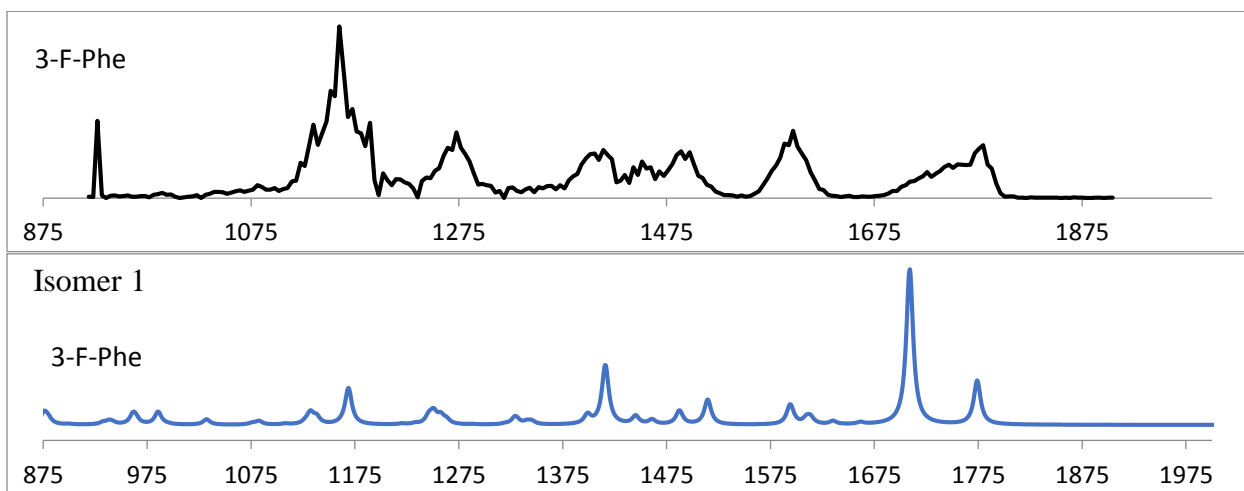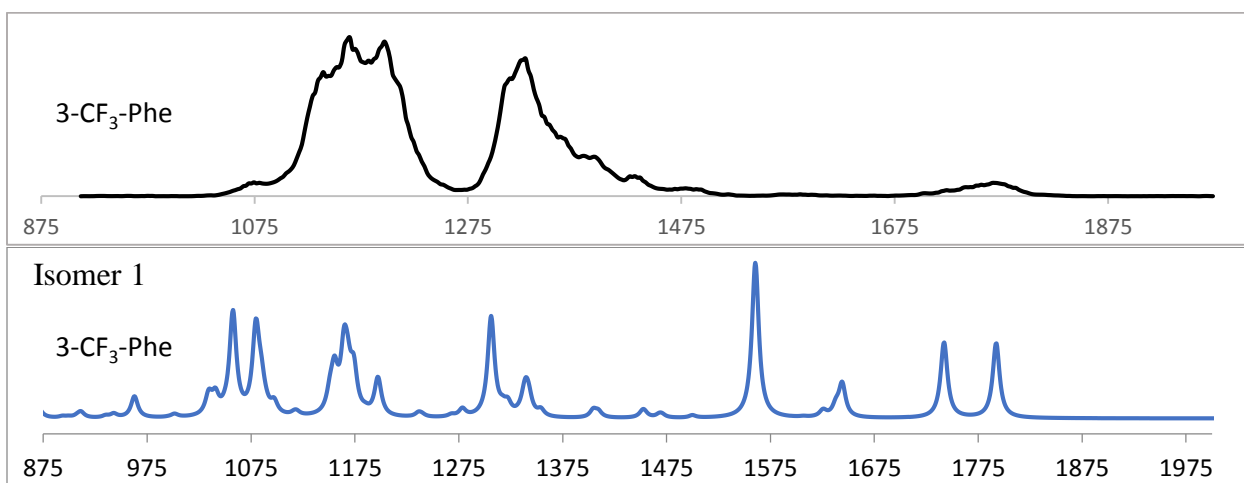

Figure B.8 The experimental IRMPD spectrum (black) and the calculated IR spectra (blue) for $[3\text{-}CF_3\text{-}Phe_2+H]^+$ global minimum. The calculations were conducted at B3LYP/6-311++G(d,p) level of theory.
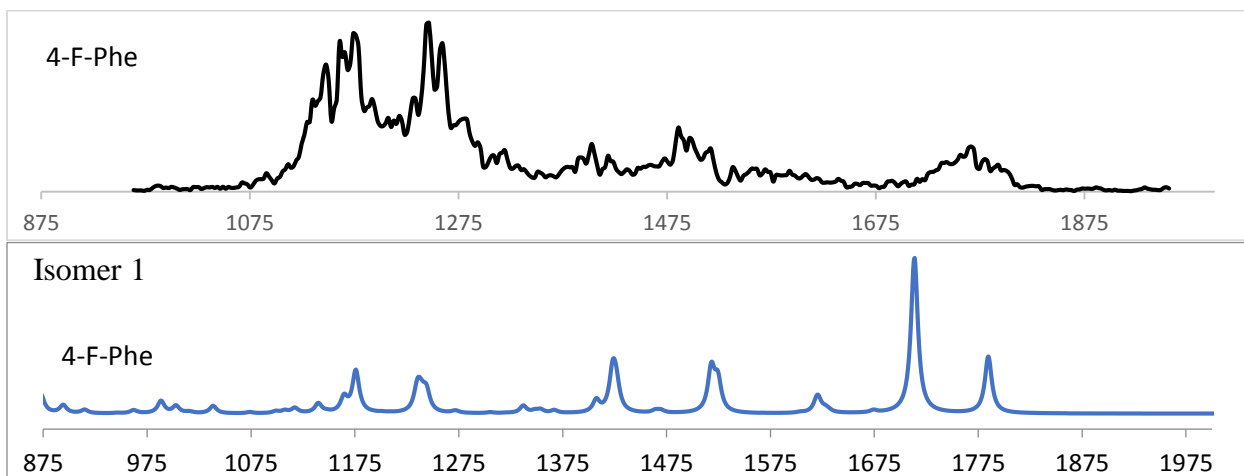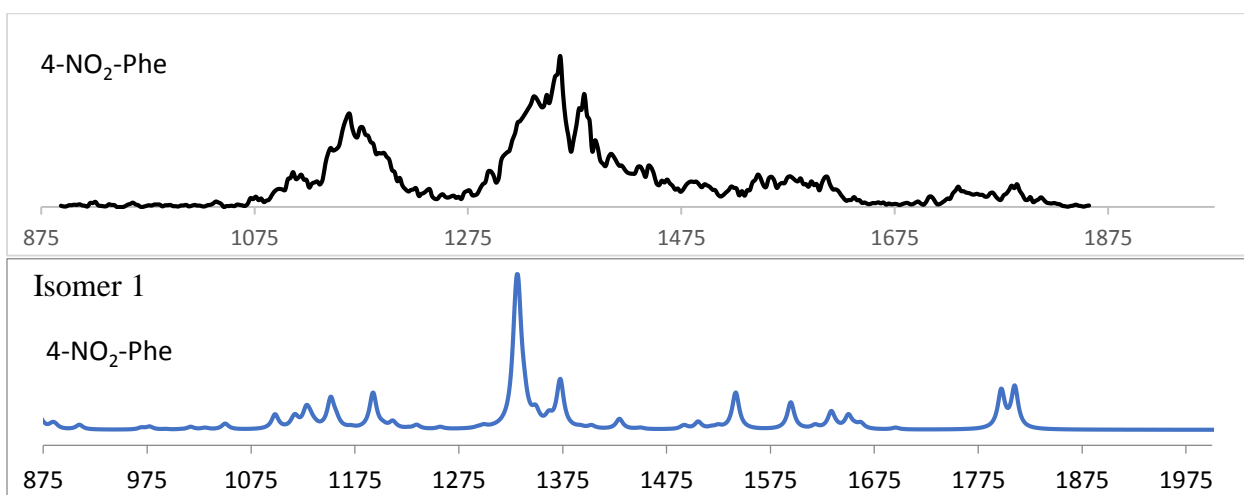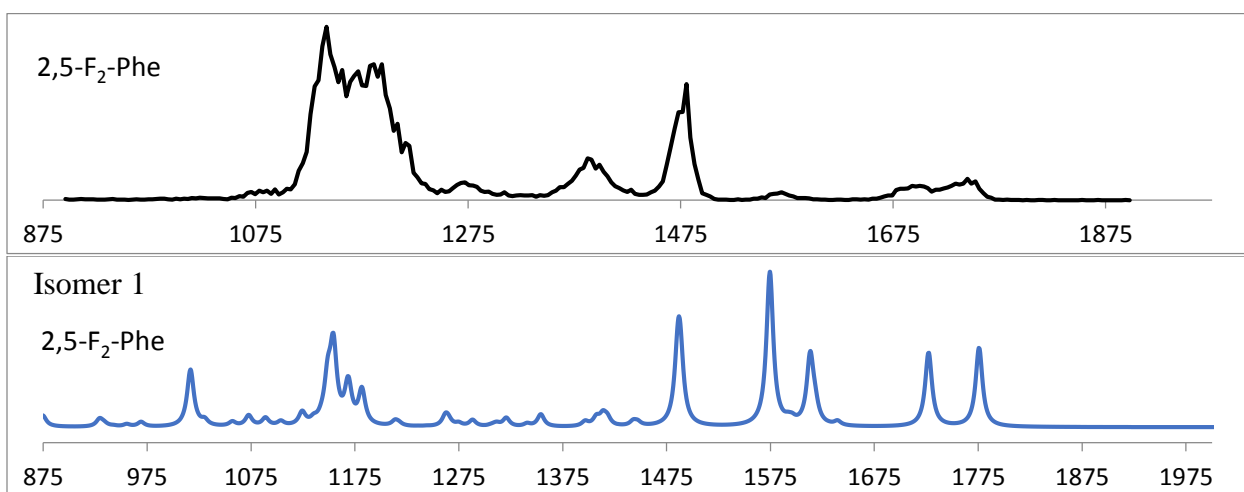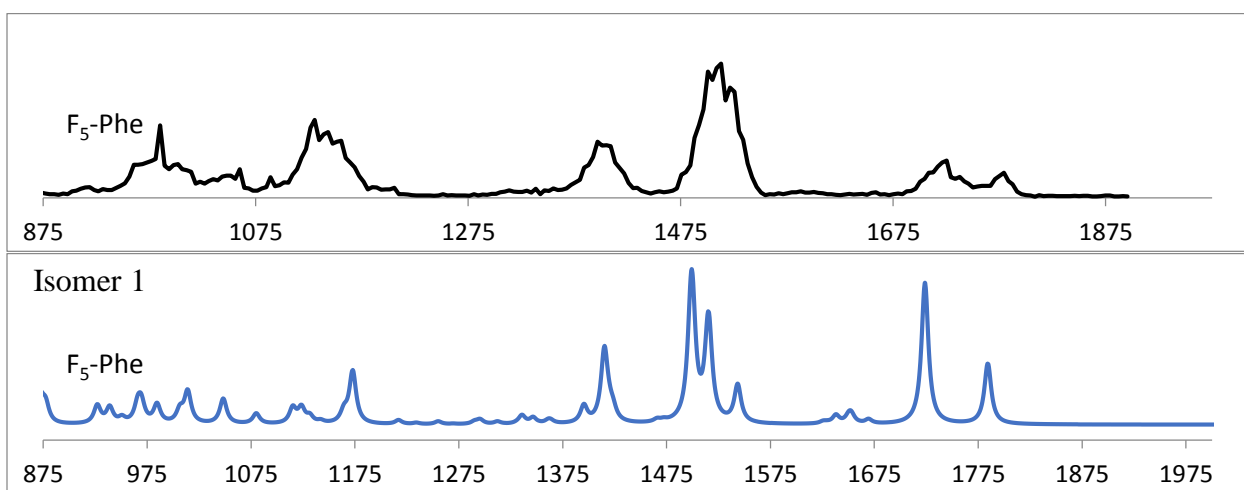
Figure B.9 The experimental IRMPD spectrum (black) and the calculated IR spectra (blue) for [F$_5$-Phe$_2$+H]$^+$ global minimum. The calculations were conducted at B3LYP/6-311++G(d,p) level of theory.

B.3   QTAIM and NCI analysis of global minima of [(x-Phe)$_2$ + H]$^+$

3-CN-Phe   3-F-Phe   4-F-Phe

4-NO$_2$-Phe   2,5-F$_2$-Phe

3,4-(MeO)$_2$-Phe   3,5-F$_2$-Phe

3-CF$_3$-Phe   F$_5$-Phe

Figure B.10 Bond critical points (orange points) and gradient isosurfaces for Phe derivatives global minima. The calculations were conducted at B3LYP/6-311++G(d,p) level of theory.

B.4 Spectra similarities of [(x-Phe)$_2$ + H]$^+$
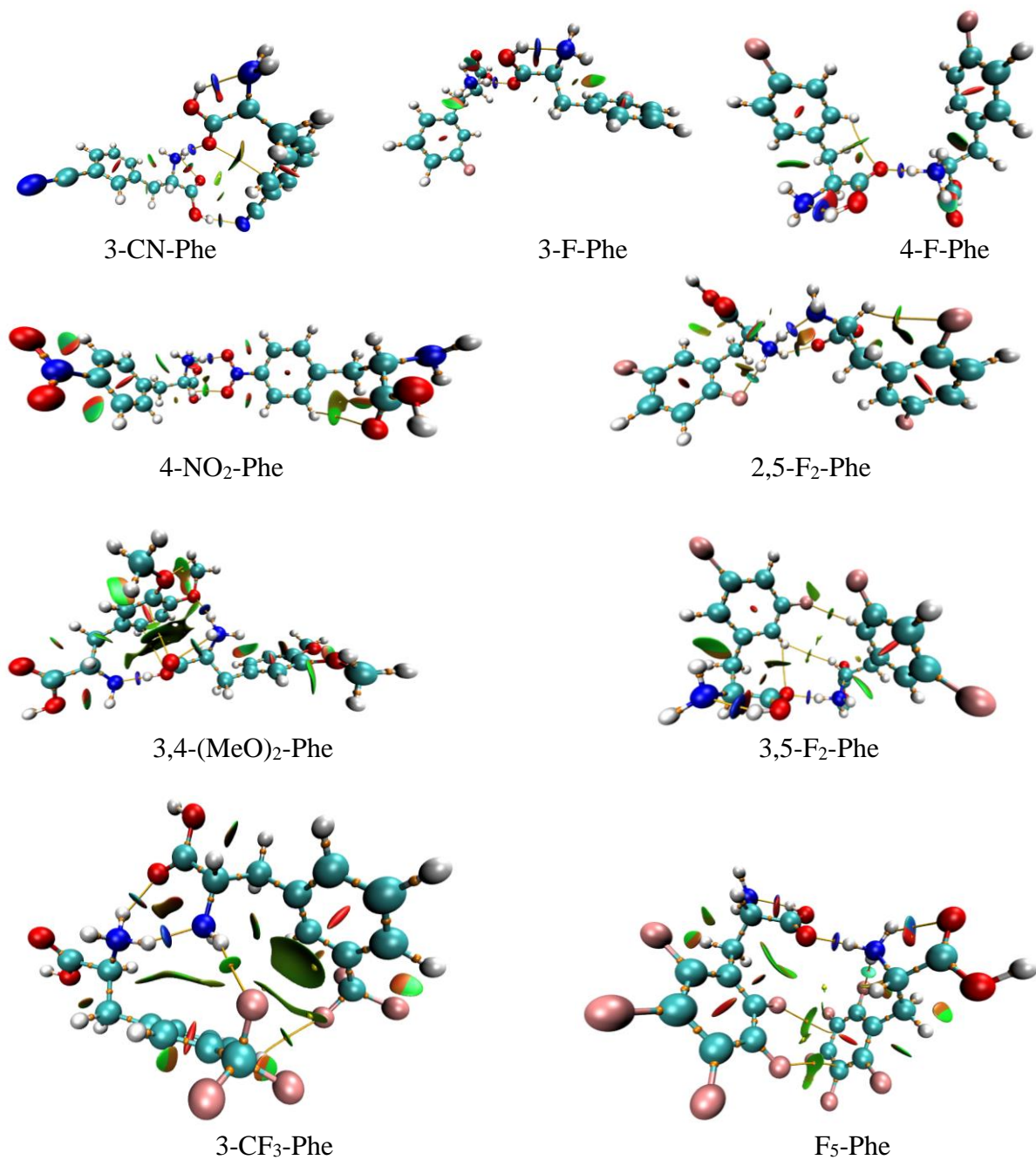
Table B.2 Spectral similarities between the calculated IR spectrum and the experimental IRMPD spectra of [Phe$_2$ + H]$^+$. Similarities were rescaled to the [0,1] interval.

Isomer	3F-Phe	4F-Phe	35F$_2$-Phe	25F$_2$-Phe	F$_5$-Phe	3CN-Phe	3CF$_3$-Phe	4NO$_2$-Phe	34(MeO)$_2$-Phe
1	0.29	0.18	0.61	0.40	0.79	1.00	0.22	0.09	1.00
2	0.43	0.16	0.61	1.00	0.59	0.93	0.71	0.15	0.21
3	0.57	0.38	1.00	0.82	0.57	0.86	0.84	0.16	0.22
4	0.48	0.23	0.92	0.68	1.00	0.54	0.63	0.21	0.58
5	0.00	0.39	0.28	0.90	0.31	0.95	0.42	0.42	0.43
6	0.94	0.87	0.91	0.88	0.62	0.86	0.17	0.39	0.54
7	0.74	0.94	0.93	0.40	0.56	0.51	0.35	0.19	0.17
8	0.21	0.89	0.78	0.00	0.55	0.70	0.00	1.00	0.41
9	0.85	0.54	0.99	0.42	0.29	0.43	0.70	0.44	0.33
10	0.88	0.41	0.81	0.31	0.56	0.55	0.64	0.45	0.22
11	0.13	1.00	0.87	0.43	0.51	0.74	0.15	0.21	0.94
12	0.15	0.93	0.84	0.47	0.43	0.48	0.09	0.12	0.41
13	0.16	0.44	0.80	0.22	0.69	0.75	0.32	0.29	0.19
14	0.99	0.00	0.95		0.51	0.45	0.46	0.00	0.73
15	1.00	0.91	0.49		0.49	0.43	0.46	0.15	0.65
16	0.17	0.93	0.74		0.52	0.32	1.00	0.46	0.29
17	0.86	0.97	0.71		0.70	0.66			0.43
18	0.13	0.82	0.30		0.43	0.52			0.00
19	0.75	0.83	0.80		0.51	0.29			0.39
20	0.68	0.24	0.59		0.31	0.44			0.44
21	0.51	0.91	0.00		0.32	0.37			
22	0.51	0.87	0.31		0.11	0.55			
23	0.86	0.96	0.29		0.79	0.68			
24	0.57	0.60	0.68		0.32	0.40			
25	0.55	0.92	0.63		0.23	0.29			
26	0.59	0.70	0.37		0.55	0.47			
27	0.66	0.43	0.69		0.17	0.40			
28	0.45	0.76	0.34		0.36	0.62			
29	0.49	0.71	0.67		0.59	0.00			
30	0.55				0.71	0.55			
31	0.80				0.00	0.31			

32	0.18	0.43	0.33
33	0.47	0.04	0.32
34	0.72	0.70	0.53
35		0.78	0.05
36			0.06
37			0.27
38			0.15
39			0.31

Appendix C
Para-Substituted Benzylpyridinium Derivatives

C.1 The calculated vertical excitation energies of the para-substituted BP ions with the associated oscillator strengths

Table C.1 The calculated vertical excitation energies of the para-substituted BP-F ions with the associated oscillator strengths (fosc). The excitation energies are predicted using the CAM-B3LYP/6-311++G(d,p) and the STEOM-DLPNO-CCSD/6-311++G(d,p) level of theory.

BP-F	Energy (eV)		Wavelength (nm)		f_{osc}	
State	CAM-B3LYP	CCSD	CAM-B3LYP	CCSD	CAM-B3LYP	CCSD
1	4.52	4.82	274.3	257.2	0.0001203	0.0020458
2	5.05	4.95	245.4	250.7	0.0026404	0.0829183
3	5.27	5.14	235.1	241.4	0.0023129	0.0002547
4	5.46	5.87	227.0	211.1	0.0712199	0.0019501
5	5.66	6.13	219.0	202.1	0.0018645	0.2837173
6	5.84	6.45	212.2	192.1	0.0382282	0.1118254
7	6.14	6.82	202.0	181.8	0.2218428	0.9467885
8	6.27	7.06	197.8	175.5	0.0343901	0.0157401
9	6.84	7.12	181.4	174.0	0.3993461	0.7658339
10	6.86	7.50	180.8	165.2	0.4615441	0.0009017
11	6.98	7.51	177.7	165.2	0.4105695	0.0019369
12	7.26	7.54	170.8	164.5	0.0037964	0.0414049
13	7.32	7.83	169.3	158.4	0.2326975	0.0050522
14	7.38	7.91	168.0	156.7	0.2320064	0.0032793
15	7.40	8.19	167.5	151.3	0.0032080	0.0375648

Table C.2 The calculated vertical excitation energies of the para-substituted BPF-Cl ions with the associated oscillator strengths (f$_{osc}$). The excitatF-Clion energies are predicted using the CAM-B3LYP/6-311++G(d,p) and the STEOM-DLPNO-CCSD/6-311++G(d,p) level of theory.

BP-Cl	Energy (eV)		Wavelength (nm)		f_{osc}	
State	CAM-B3LYP	CCSD	CAM-B3LYP	CCSD	CAM-B3LYP	CCSD
1	4.37	4.67	283.6	265.8	0.0000907	0.0004070
2	4.99	4.95	248.3	250.5	0.0019912	0.0897886

3	5.12	5.01	241.9	247.7	0.0002326	0.0002529
4	5.46	5.72	227.1	216.6	0.0740012	0.5273071
5	5.50	5.76	225.3	215.2	0.0037804	0.0010886
6	5.56	6.13	223.0	202.3	0.2469837	0.0238684
7	6.01	6.13	206.2	202.2	0.1978029	0.0244485
8	6.26	6.44	197.9	192.6	0.0401743	0.1214256
9	6.51	6.68	190.4	185.7	0.0000970	0.7372110
10	6.54	7.14	189.6	173.6	0.2849193	0.0441670
11	6.63	7.15	186.9	173.3	0.0318893	0.0188446
12	6.72	7.37	184.5	168.3	0.0000597	0.0528908
13	6.81	7.42	182.1	167.0	0.6702654	0.0674589
14	6.98	7.65	177.6	162.1	0.0003907	0.0023196
15	7.07	7.76	175.5	159.8	0.0017187	0.1562122

Table C.3 The calculated vertical excitation energies of the para-substituted BP-CN ions with the associated oscillator strengths (f_{osc}). The excitation energies are predicted using the CAM-B3LYP/6-311++G(d,p) and the STEOM-DLPNO-CCSD/6-311++G(d,p) level of theory.

BP-CN	Energy (eV)		Wavelength (nm)		f_{osc}	
State	CAM-B3LYP	CCSD	CAM-B3LYP	CCSD	CAM-B3LYP	CCSD
1	4.65	4.70	266.6	263.5	0.0038483	0.0130259
2	5.02	5.08	247.0	244.3	0.0028418	0.0967495
3	5.14	5.38	241.4	230.6	0.0085232	0.0156838
4	5.44	5.38	227.7	230.5	0.0297182	0.0200496
5	5.46	5.76	226.9	215.3	0.4028783	0.0288801
6	5.74	5.80	215.8	213.6	0.0321986	0.5892867
7	6.06	6.48	204.5	191.3	0.1175473	0.0596334
8	6.20	6.48	200.0	191.2	0.0000033	0.0670803
9	6.27	6.49	197.8	191.1	0.0444491	0.0656424
10	6.60	6.78	188.0	182.8	0.3105488	0.2418172
11	6.74	6.79	184.0	182.7	0.6255814	0.2249513
12	6.91	6.80	179.4	182.4	0.0004269	0.2811429
13	7.13	6.95	173.9	178.4	0.0464758	0.0015984
14	7.18	7.34	172.6	169.0	0.1498983	0.0949311
15	7.27	8.05	170.6	153.9	0.3806380	0.0295379

Table C.4 The calculated vertical excitation energies of the para-substituted BP-Me ions with the associated oscillator strengths (f_{osc}). The excitation energies are predicted using the CAM-B3LYP/6-311++G(d,p) and the STEOM-DLPNO-CCSD/6-311++G(d,p) level of theory.

BP-Me	Energy (eV)		Wavelength (nm)		f_{osc}	
State	CAM-B3LYP	CCSD	CAM-B3LYP	CCSD	CAM-B3LYP	CCSD
1	4.32	4.77	287.2	260.1	0.0001040	0.0006560
2	4.72	4.94	262.5	251.1	0.0011061	0.0002856
3	5.18	4.94	239.5	251.0	0.0000374	0.0791926
4	5.45	5.50	227.3	225.6	0.0695227	0.0022606
5	5.48	6.01	226.4	206.1	0.0052617	0.2934353
6	5.64	6.06	220.0	204.7	0.0061996	0.0231301
7	5.87	6.48	211.2	191.4	0.3165974	0.0540198
8	6.26	6.51	197.9	190.5	0.0329471	0.6593733
9	6.68	6.92	185.6	179.1	0.3364817	0.0165745
10	6.68	6.93	185.5	178.9	0.1835429	0.6407897
11	6.84	7.20	181.4	172.1	0.7504631	0.5706183
12	6.94	7.26	178.7	170.8	0.0019788	0.0192516
13	7.16	7.62	173.1	162.8	0.0266650	0.0168038
14	7.20	7.81	172.2	158.7	0.0592274	0.0086794
15	7.24	7.82	171.2	158.6	0.0003851	0.0083422

Table C.5 The calculated vertical excitation energies of the para-substituted BP-OMe ions with the associated oscillator strengths (f_{osc}). The excitation energies are predicted using the CAM-B3LYP/6-311++G(d,p) and the STEOM-DLPNO-CCSD/6-311++G(d,p) level of theory.

BP-OMe	Energy (eV)		Wavelength (nm)		f_{osc}	
State	CAM-B3LYP	CCSD	CAM-B3LYP	CCSD	CAM-B3LYP	CCSD
1	3.81	4.32	325.6	287.0	0.0002220	0.0006245
2	4.74	4.54	261.8	273.0	0.0059829	0.0108565
3	4.86	4.95	255.0	250.6	0.0009947	0.0855318
4	5.18	5.31	239.1	233.3	0.0130742	0.5090094
5	5.44	5.46	228.0	227.2	0.2667514	0.0132449
6	5.46	5.64	227.1	219.8	0.0880862	0.0013711
7	5.86	6.13	211.5	202.3	0.1823799	0.0229456
8	6.14	6.46	201.9	191.9	0.0042868	0.1542168

9	6.27	6.60	197.7	187.9	0.0378597	0.5246142
10	6.54	6.66	189.5	186.3	0.2882933	0.0200954
11	6.57	6.72	188.7	184.5	0.0111433	0.5859033
12	6.79	6.87	182.5	180.5	0.5600706	0.0223736
13	6.84	6.99	181.1	177.3	0.1442790	0.0352142
14	6.89	7.25	180.0	171.0	0.0013653	0.0316370
15	6.95	7.25	178.4	170.9	0.0626616	0.3013036

Table C.6 The calculated vertical excitation energies of the para-substituted BP-OEt ions with the associated oscillator strengths (f_{osc}). The excitation energies are predicted using the CAM-B3LYP/6-311++G(d,p) and the STEOM-DLPNO-CCSD/6-311++G(d,p) level of theory.

BP-OEt	Energy (eV)		Wavelength (nm)		f_{osc}	
State	CAM-B3LYP	CCSD	CAM-B3LYP	CCSD	CAM-B3LYP	CCSD
1	3.76	4.25	329.7	291.4	0.0002124	0.0001897
2	4.69	4.54	264.2	273.2	0.0054074	0.0132615
3	4.83	4.94	256.6	250.7	0.0009304	0.0829344
4	5.16	5.31	240.0	233.6	0.0142892	0.5540568
5	5.41	5.39	229.0	230.1	0.3093073	0.0034532
6	5.46	5.62	227.1	220.8	0.0763154	0.0017794
7	5.83	6.08	212.6	204.0	0.1913552	0.0098905
8	6.09	6.45	203.5	192.1	0.0035870	0.1600326
9	6.27	6.57	197.7	188.8	0.0379200	0.5229576
10	6.52	6.62	190.0	187.3	0.1383059	0.0209242
11	6.53	6.71	189.9	184.9	0.1529380	0.5354215
12	6.75	6.81	183.7	182.1	0.0012100	0.0192710
13	6.77	6.88	183.1	180.1	0.3805464	0.0371337
14	6.81	7.22	182.1	171.7	0.2987593	0.0065399
15	6.91	7.22	179.5	171.7	0.1140755	0.0034134

Table C.7 The calculated vertical excitation energies of the para-substituted BP-OiPr ions with the associated oscillator strengths (f_{osc}). The excitation energies are predicted using the CAM-B3LYP/6-311++G(d,p) and the STEOM-DLPNO-CCSD/6-311++G(d,p) level of theory.

BP-OiPr	Energy (eV)		Wavelength (nm)		f_{osc}	
State	CAM-B3LYP	CCSD	CAM-B3LYP	CCSD	CAM-B3LYP	CCSD
1	3.69	4.14	336.3	299.3	0.0001728	0.0001526

	2	4.63	4.51	267.9	275.2	0.0046890	0.0159379
	3	4.81	4.94	258.0	250.8	0.0008354	0.0442567
	4	5.13	4.95	241.8	250.6	0.0140161	0.0430427
	5	5.36	5.28	231.1	234.7	0.3348252	0.0015961
	6	5.46	5.58	227.0	222.1	0.0740810	0.0011537
	7	5.80	5.99	213.6	206.9	0.1847981	0.0270754
	8	6.03	6.46	205.6	191.9	0.0039524	0.0784184
	9	6.27	6.53	197.7	189.8	0.0396154	0.3504988
	10	6.49	6.55	191.1	189.3	0.1422103	0.5363599
	11	6.50	6.65	190.9	186.4	0.1062778	0.1661713
	12	6.68	6.70	185.7	185.0	0.0008810	0.1561475
	13	6.70	6.71	185.0	184.9	0.1214653	0.2204465
	14	6.76	7.09	183.3	174.8	0.1602940	0.0544348
	15	6.77	7.19	183.1	172.5	0.2130909	0.7721010

Table C.8 The calculated vertical excitation energies of the para-substituted BP-OtBu ions with the associated oscillator strengths (f_{osc}). The excitation energies are predicted using the CAM-B3LYP/6-311++G(d,p) and the STEOM-DLPNO-CCSD/6-311++G(d,p) level of theory.

BP-OtBu	Energy (eV)		Wavelength (nm)		f_{osc}	
State	CAM-B3LYP	CCSD	CAM-B3LYP	CCSD	CAM-B3LYP	CCSD
1	3.63	4.05	341.6	306.2	0.0001598	0.0002136
2	4.57	4.47	271.3	277.5	0.0038324	0.0125523
3	4.77	4.94	260.1	250.9	0.0007860	0.0865760
4	5.11	5.13	242.8	241.8	0.0151544	0.5800348
5	5.32	5.19	233.1	238.9	0.3560390	0.0075982
6	5.46	5.53	227.0	224.4	0.0691636	0.0021020
7	5.76	5.88	215.2	210.9	0.1837410	0.0058168
8	5.97	6.44	207.8	192.5	0.0015850	0.2452509
9	6.27	6.47	197.7	191.5	0.0396039	0.0266278
10	6.43	6.58	192.7	188.3	0.0167821	0.0267464
11	6.46	6.59	192.0	188.1	0.2188078	0.4900581
12	6.56	6.70	189.0	185.2	0.0000893	0.0034363
13	6.66	6.70	186.1	185.1	0.0169682	0.0247738
14	6.67	7.01	185.9	176.9	0.0028148	0.0020303
15	6.72	7.14	184.5	173.7	0.2515474	0.1904482

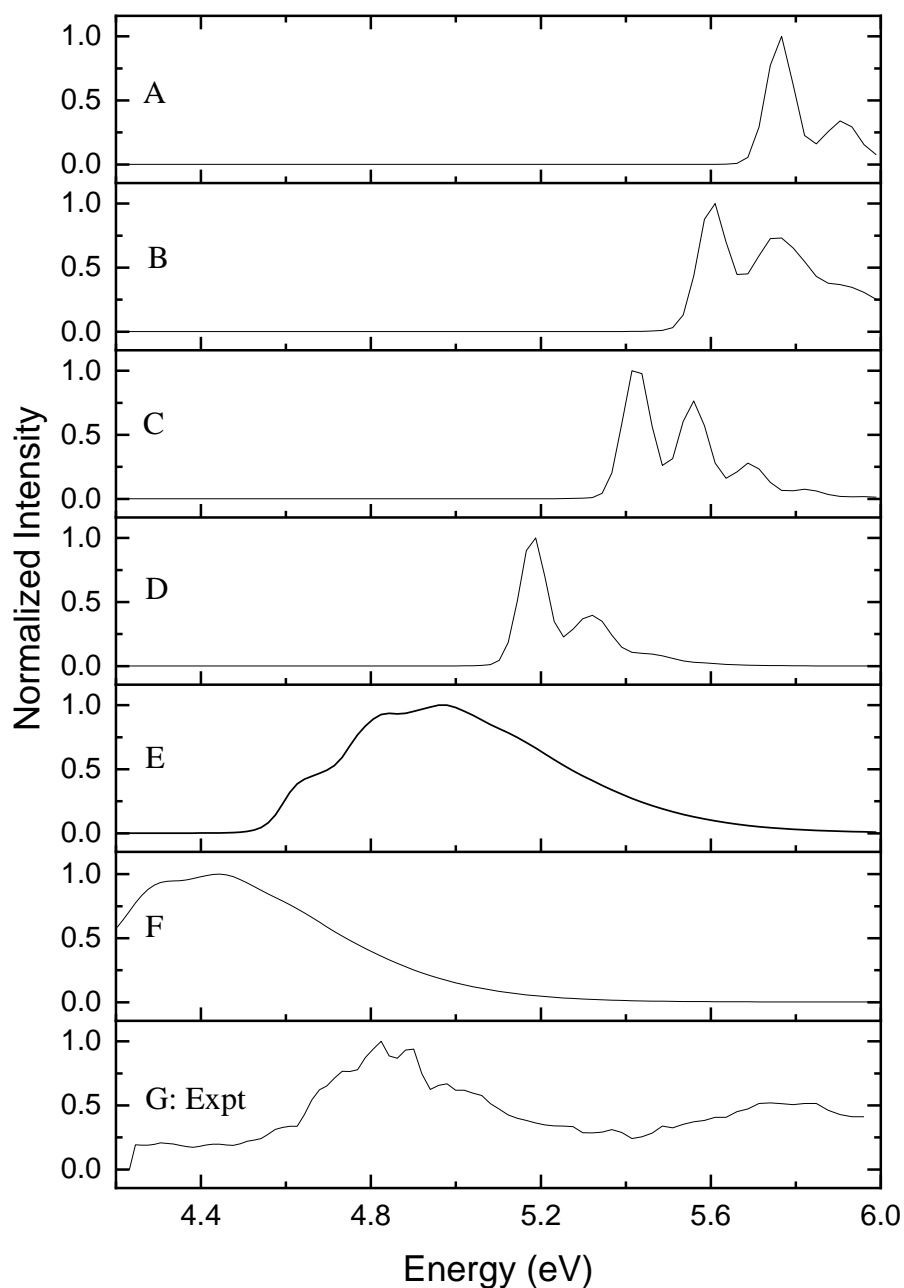## C.2   The spectra comparison of para-substituted BP ions

Figure C.1 The experimental UVPD action spectra of BP-F ion (G) and the simulated UV vertical excitation spectra of (A) S_0-S_6, (B) S_0-S_5, (C) S_0-S_4, (D) S_0-S_3, (E) S_0-S_2, (F) S_0-S_1. The spectra were calculated using the CAM-B3LYP functional with the 6-311++G(d,p) basis set. All the spectra were normalized from 0 to 1.
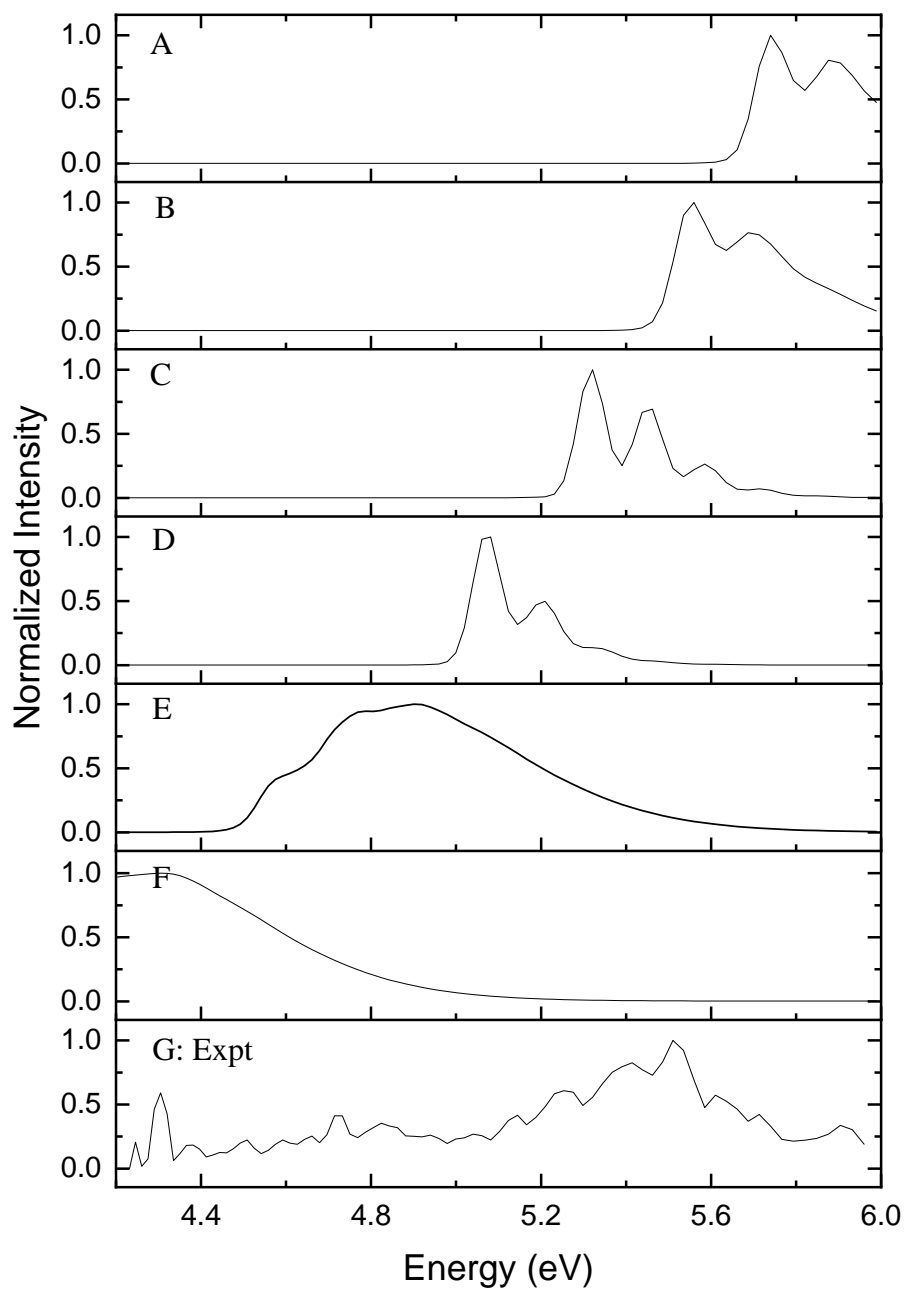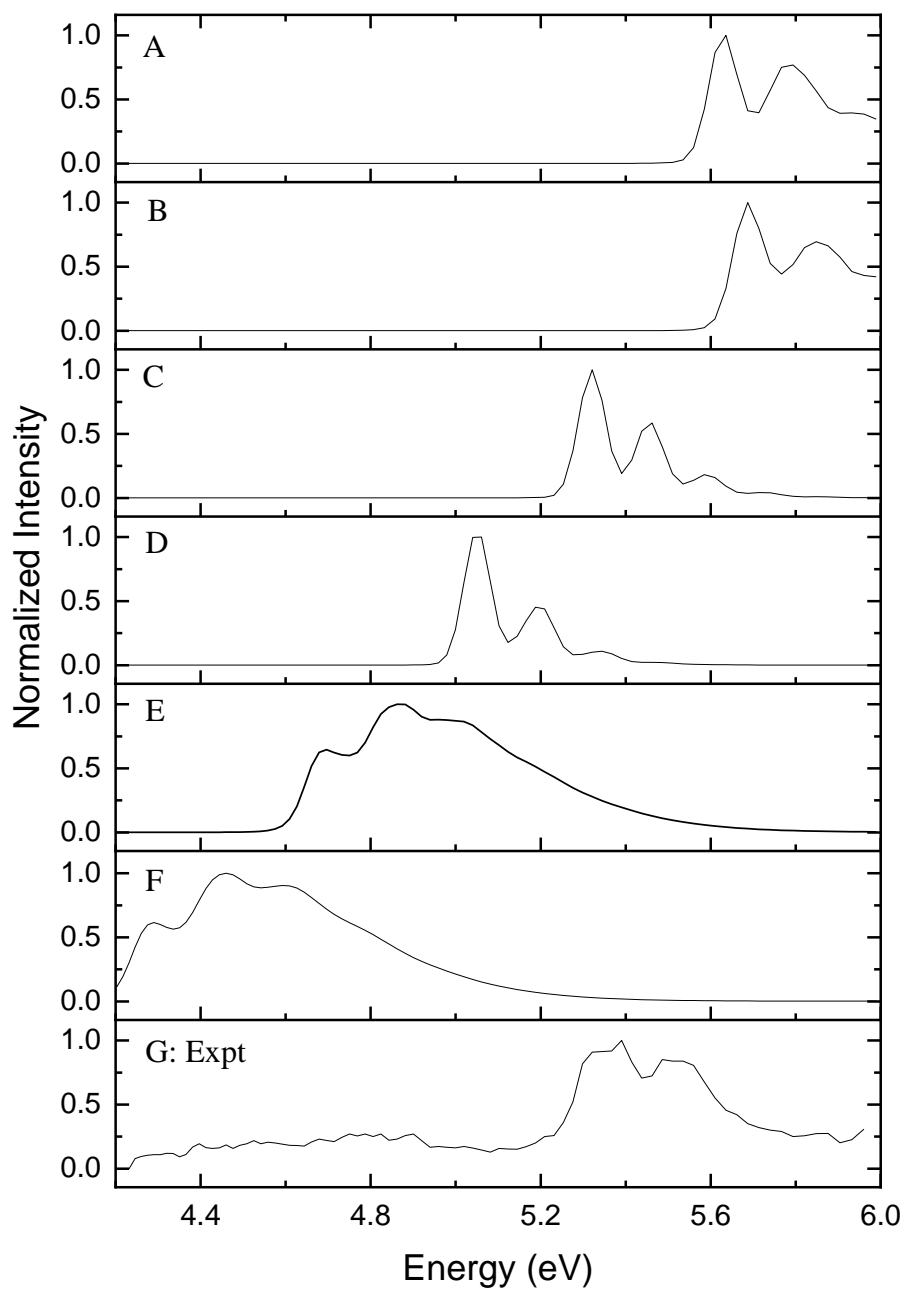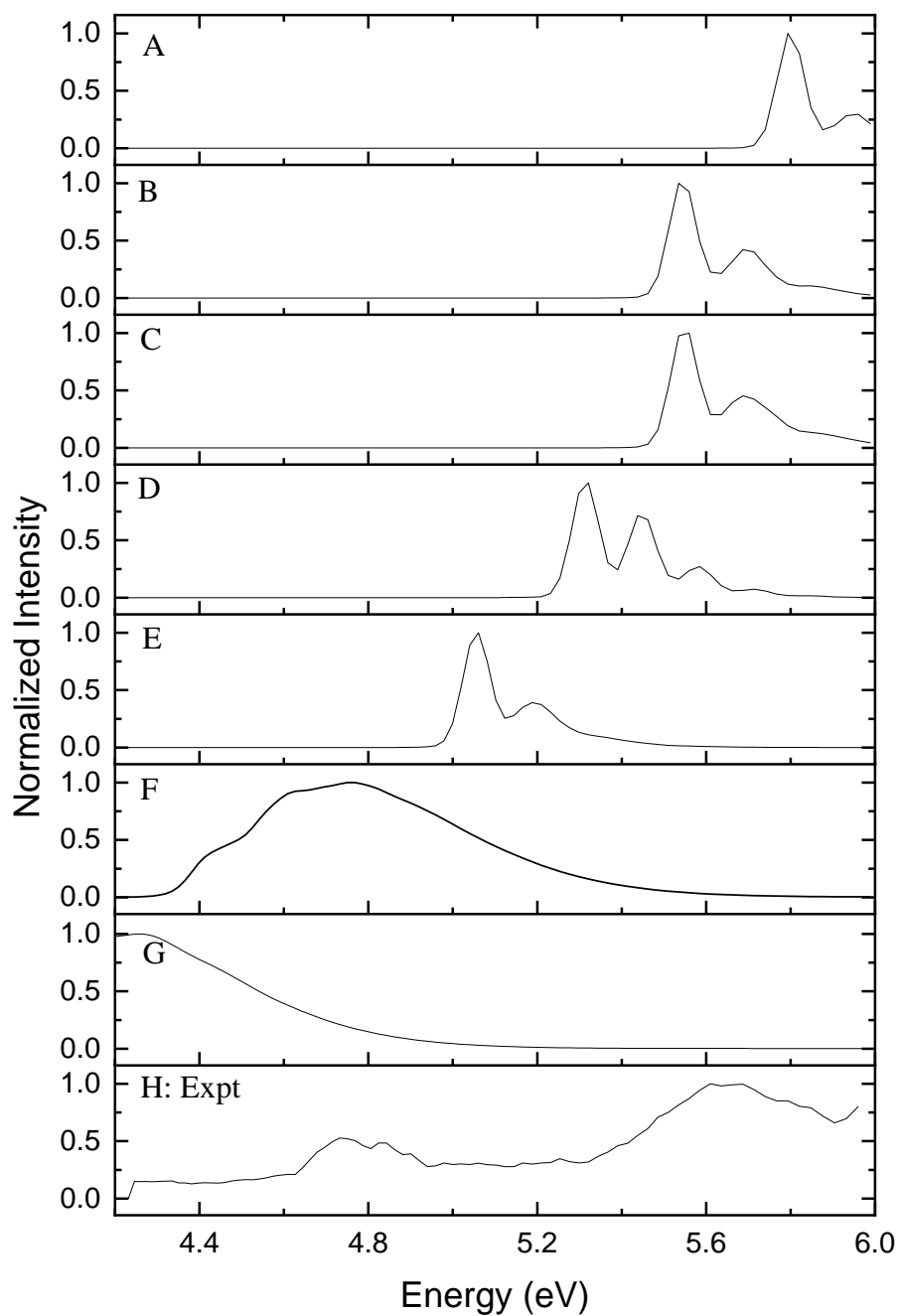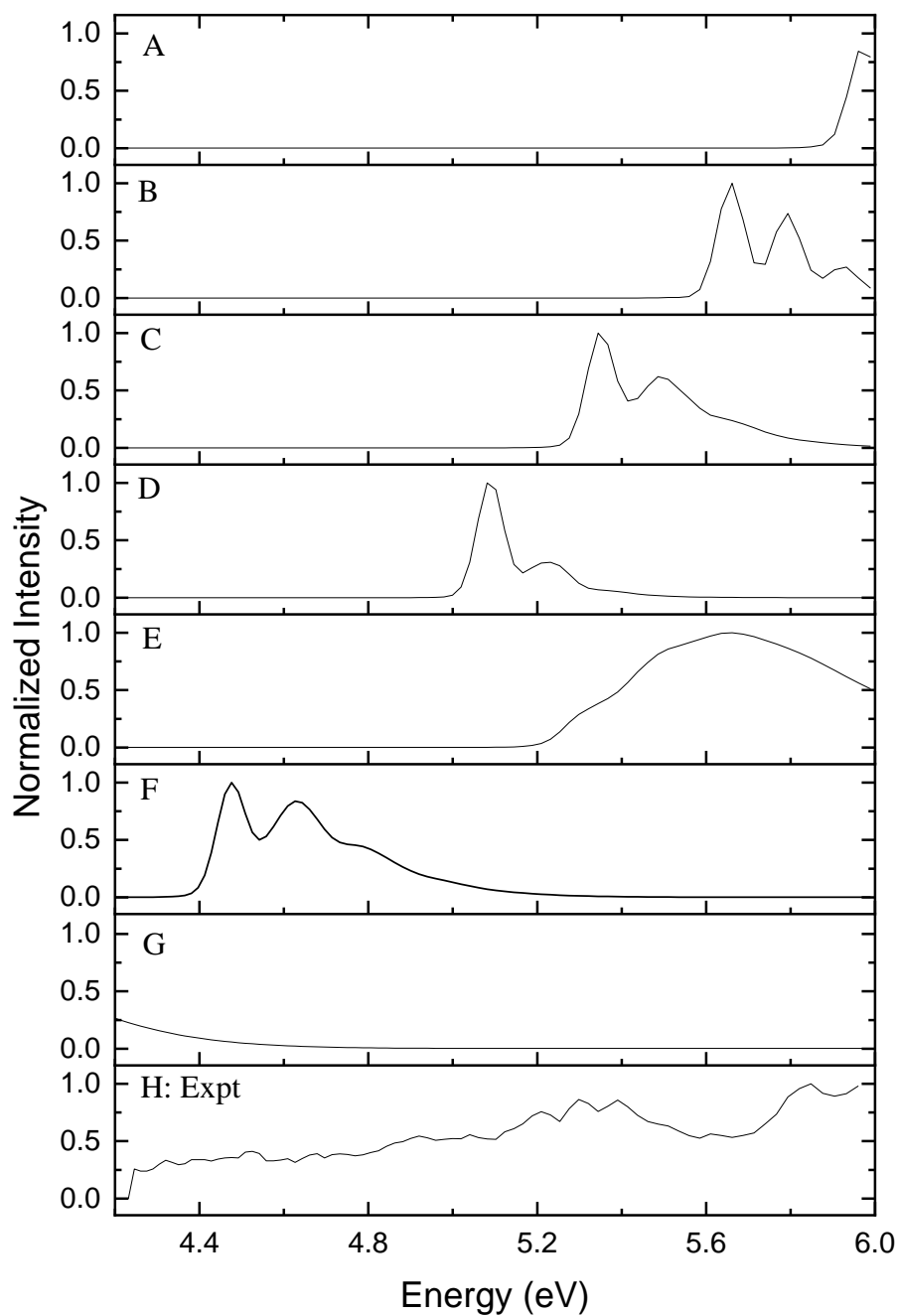
Figure C.2 The experimental UVPD action spectra of BP-Cl ion (G) and the simulated UV vertical excitation spectra of (A) S_0-S_6, (B) S_0-S_5, (C) S_0-S_4, (D) S_0-S_3, (E) S_0-S_2, (F) S_0-S_1. The spectra were calculated using the CAM-B3LYP functional with the 6-311++G(d,p) basis set. All the spectra were normalized from 0 to 1.
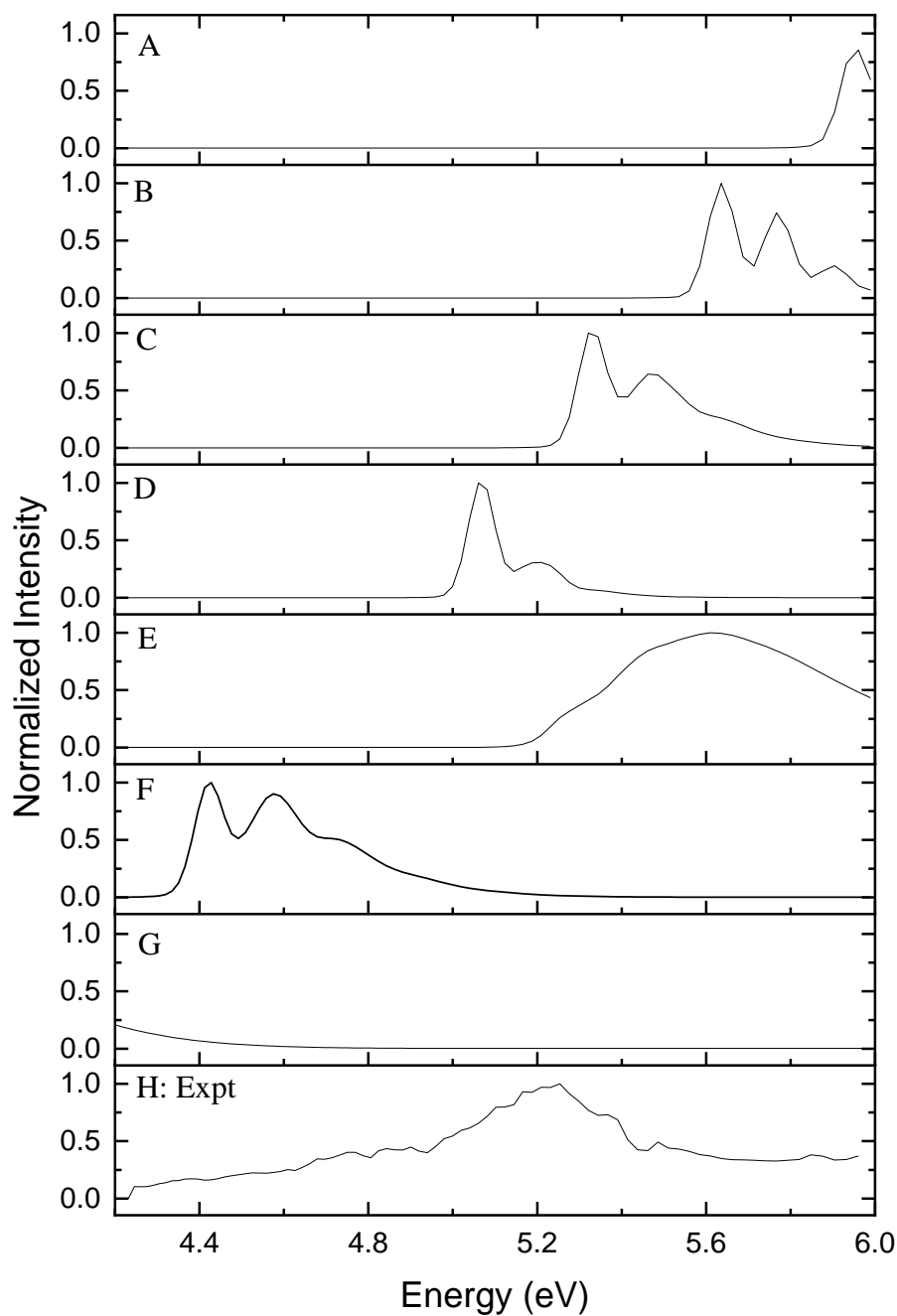
Figure C.3 The experimental UVPD action spectra of BP-CN ion (G) and the simulated UV vertical excitation spectra of (A) S_0-S_6, (B) S_0-S_5, (C) S_0-S_4, (D) S_0-S_3, (E) S_0-S_2, (F) S_0-S_1. The spectra were calculated using the CAM-B3LYP functional with the 6-311++G(d,p) basis set. All the spectra were normalized from 0 to 1.

Figure C.4 The experimental UVPD action spectra of BP-Me ion (H) and the simulated UV vertical excitation spectra of (A) S_0-S_7, (B) S_0-S_6, (C) S_0-S_5, (D) S_0-S_4, (E) S_0-S_3, (F) S_0-S_2, (G) S_0-S_1. The spectra were calculated using the CAM-B3LYP functional with the 6-311++G(d,p) basis set. All the spectra were normalized from 0 to 1.
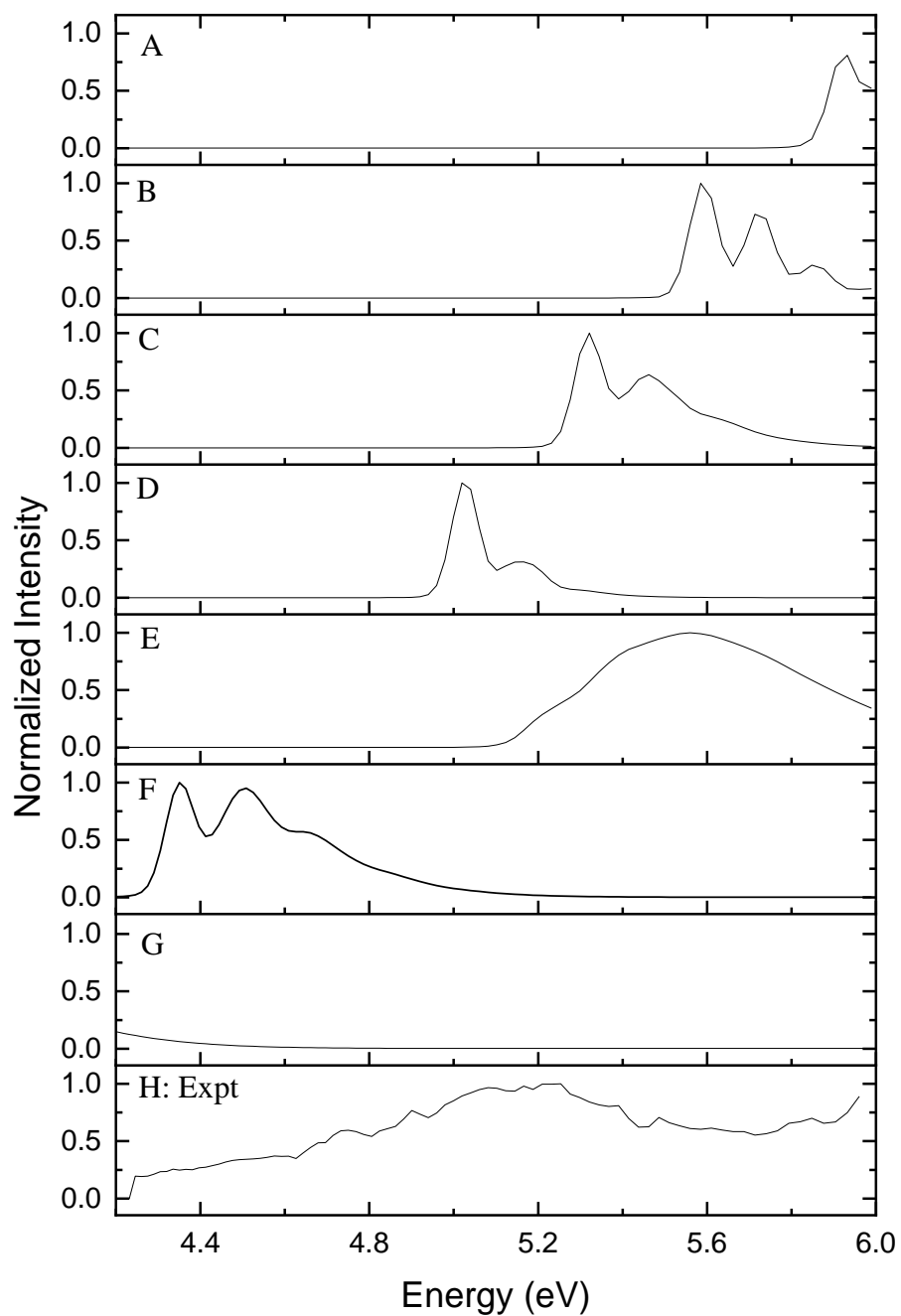
Figure C.5 The experimental UVPD action spectra of BP-OMe ion (H) and the simulated UV vertical excitation spectra of (A) S_0-S_7, (B) S_0-S_6, (C) S_0-S_5, (D) S_0-S_4, (E) S_0-S_3, (F) S_0-S_2, (G) S_0-S_1. The spectra were calculated using the CAM-B3LYP functional with the 6-311++G(d,p) basis set. All the spectra were normalized from 0 to 1.
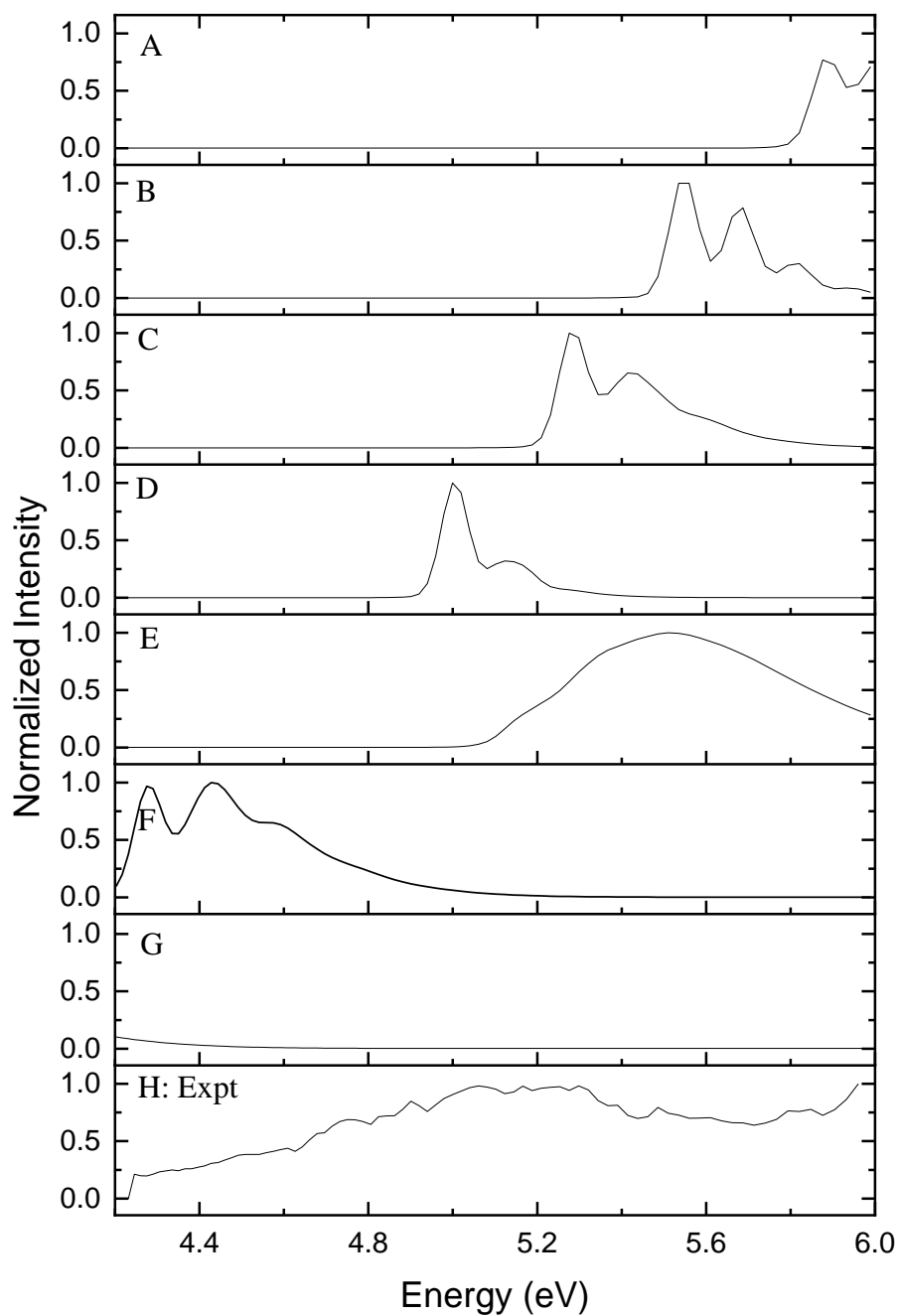
Figure C.6 The experimental UVPD action spectra of BP-OEt ion (H) and the simulated UV vertical excitation spectra of (A) S_0-S_7, (B) S_0-S_6, (C) S_0-S_5, (D) S_0-S_4, (E) S_0-S_3, (F) S_0-S_2, (G) S_0-S_1. The spectra were calculated using the CAM-B3LYP functional with the 6-311++G(d,p) basis set. All the spectra were normalized from 0 to 1.

Figure C.7 The experimental UVPD action spectra of BP-OiPr ion (H) and the simulated UV vertical excitation spectra of (A) S_0-S_7, (B) S_0-S_6, (C) S_0-S_5, (D) S_0-S_4, (E) S_0-S_3, (F) S_0-S_2, (G) S_0-S_1. The spectra were calculated using the CAM-B3LYP functional with the 6-311++G(d,p) basis set. All the spectra were normalized from 0 to 1.

Figure C.8 The experimental UVPD action spectra of BP-OtBu ion (H) and the simulated UV vertical excitation spectra of (A) S_0-S_7, (B) S_0-S_6, (C) S_0-S_5, (D) S_0-S_4, (E) S_0-S_3, (F) S_0-S_2, (G) S_0-S_1. The spectra were calculated using the CAM-B3LYP functional with the 6-311++G(d,p) basis set. All the spectra were normalized from 0 to 1.

C.3   The calculated molecular orbitals of para-substituted BP ions

A

LUMO

↑

HOMO

B

LUMO

↑

HOMO-1

C

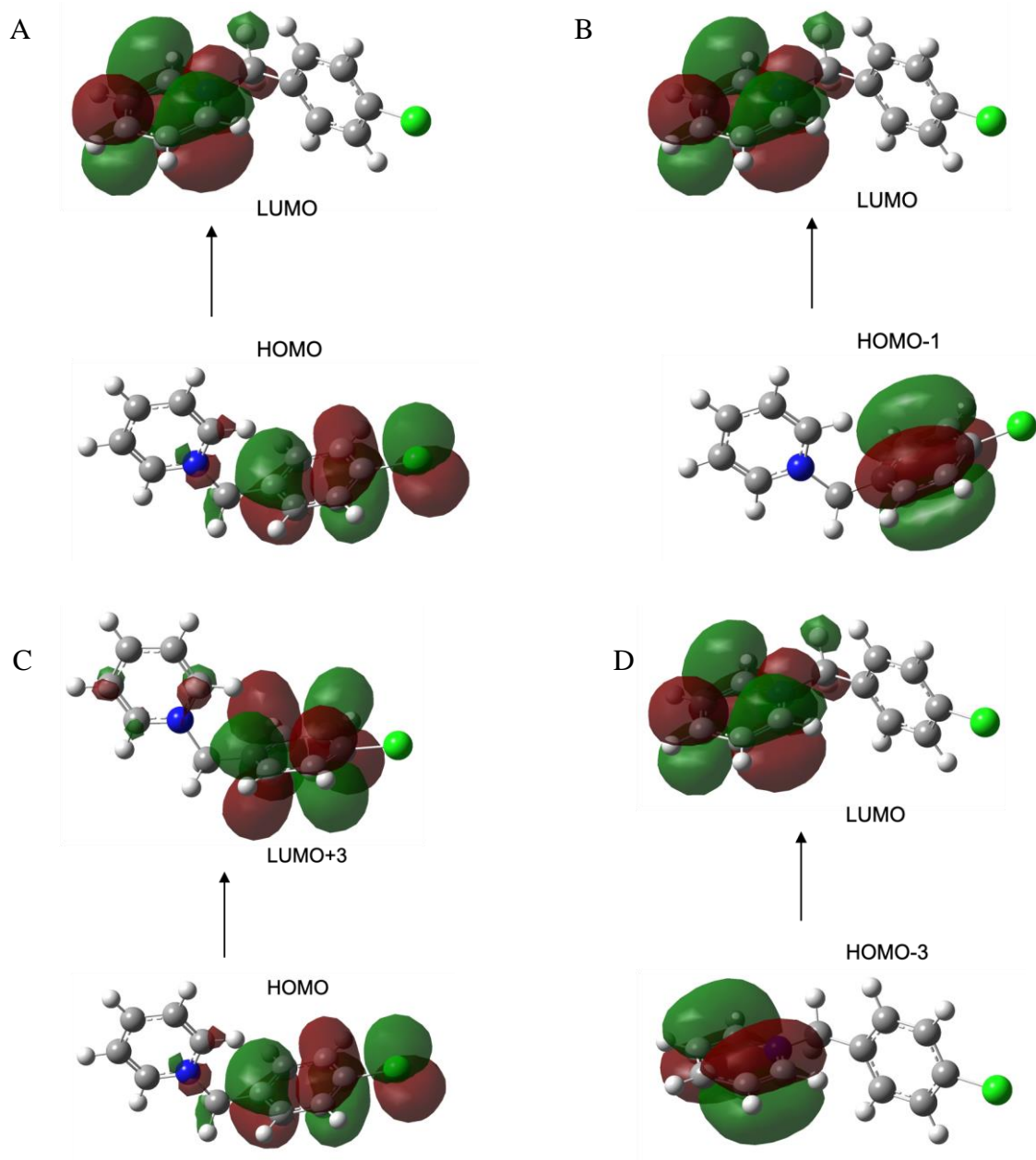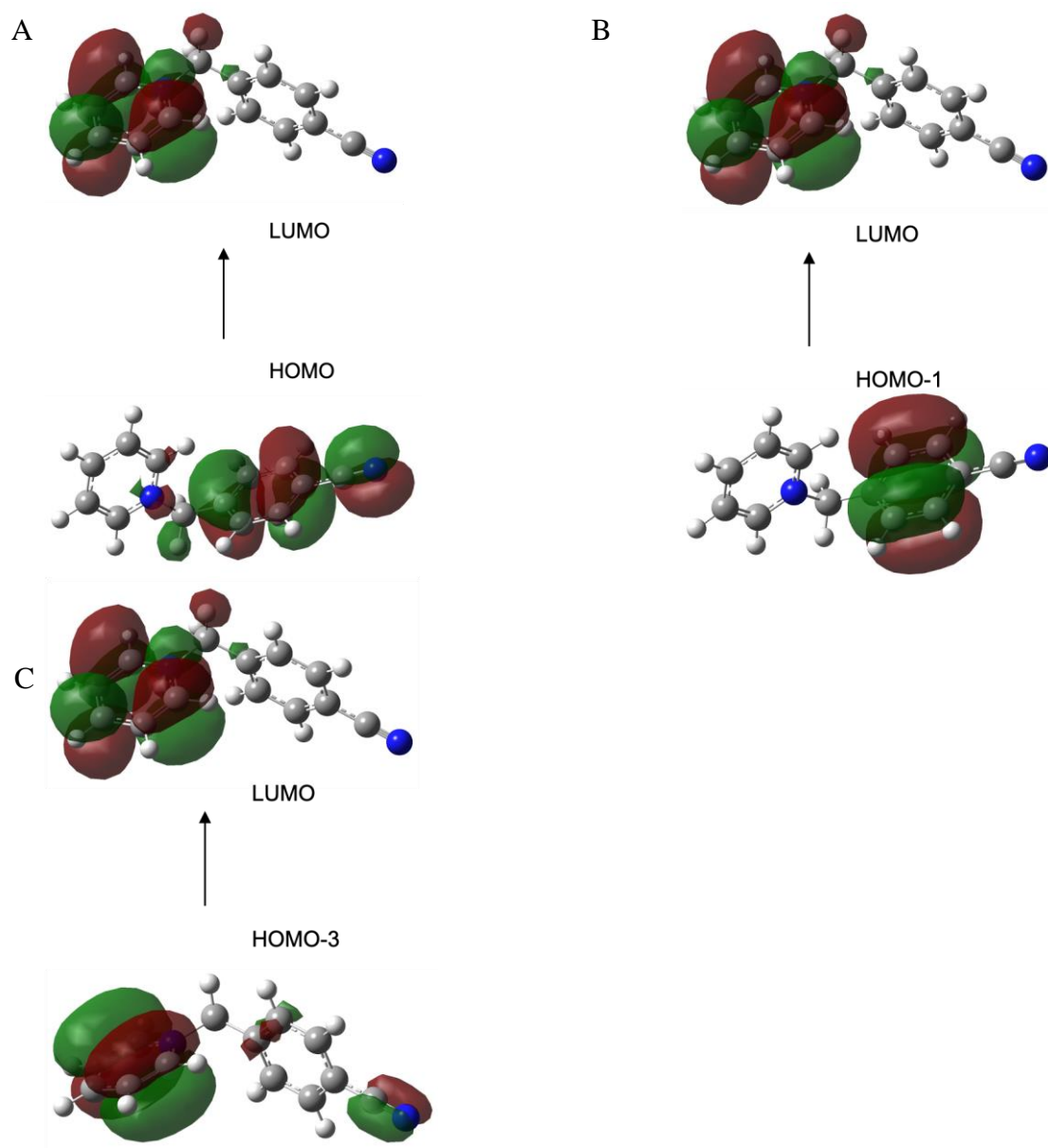LUMO+1

↑

HOMO

D

LUMO+2

↑

HOMO
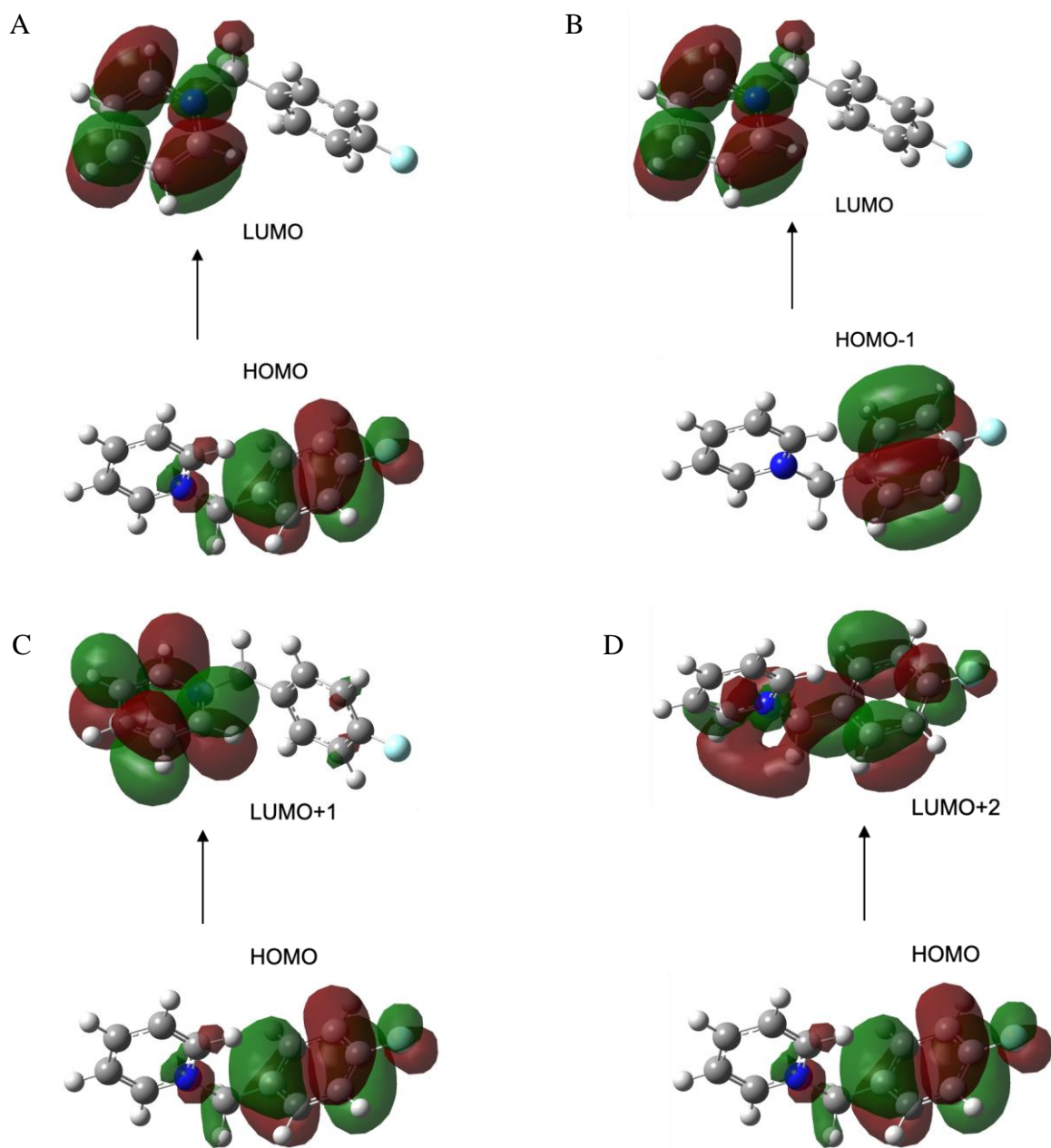
Figure C.9 The calculated molecular orbitals of BP-F ion for transitions with percentage contributions: (A) S_0 to S_1 (96.7%), (B) S_0 to S_2 (96.2%), (C) S_0 to S_5 (65.6%), and (D) S_0 to S_6 (45.9%). S_0 is the ground state, S_n is the nth excited state.
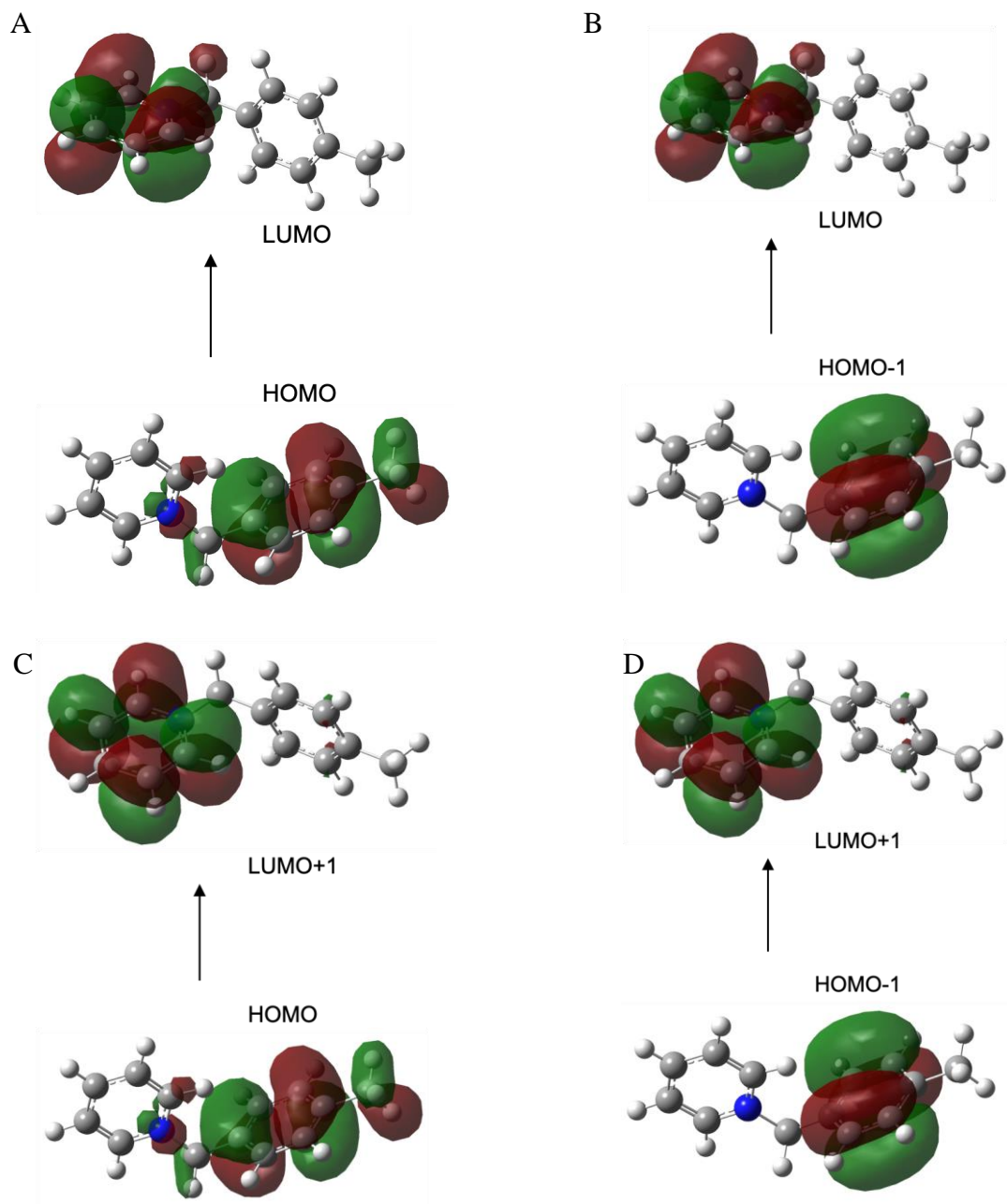
Figure C.10 The calculated molecular orbitals of BP-Cl ion for transitions with percentage contributions: (A) S_0 to S_1 (95.1%), (B) S_0 to S_2 (96.8%), (C) S_0 to S_3 (37.1%), and (D) S_0 to S_4 (81.9%). S_0 is the ground state, S_n is the nth excited state.

A

LUMO

↑

HOMO

B

LUMO

↑

HOMO-1

C

LUMO

↑

HOMO-3

Figure C.11 The calculated molecular orbitals of BP-CN ion for transitions with percentage contributions: (A) S_0 to S_1 (94.2%), (B) S_0 to S_2 (91.0%), and (C) S_0 to S_4 (70.9%). S_0 is the ground state, S_n is the nth excited state.
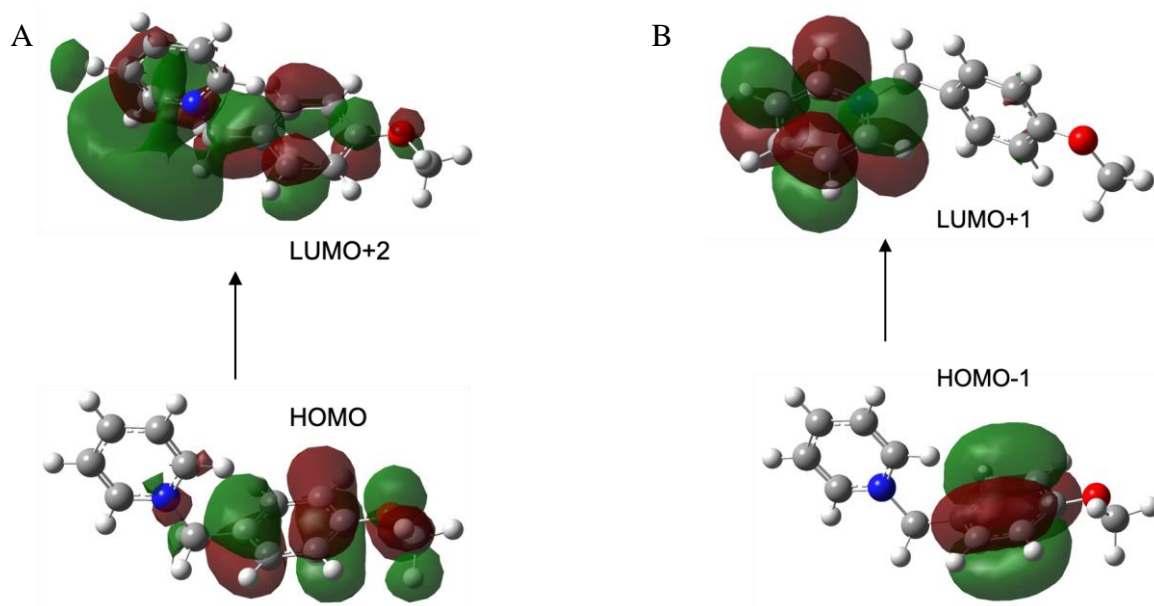
Figure C.12 The calculated molecular orbitals of BP-Me ion for transitions with percentage contributions: (A) S_0 to S_1 (96.6%), (B) S_0 to S_2 (97.9%), (C) S_0 to S_5 (49.0%), and (D) S_0 to S_6 (60.4%). S_0 is the ground state, S_n is the nth excited state.

Figure C.13 The calculated molecular orbitals of BP-OMe ion for transitions with percentage contributions: (A) S_0 to S_5 (40.6%), and (B) S_0 to S_7 (90.7%). S_0 is the ground state, S_n is the nth excited state.
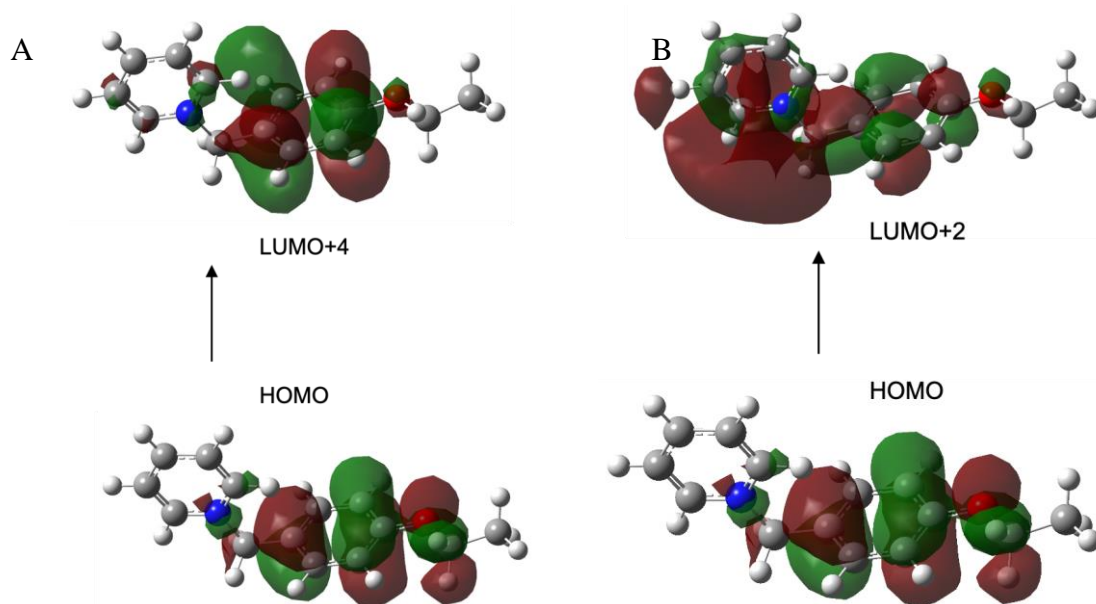

Figure C.14 The calculated molecular orbitals of BP-OEt ion for transitions with percentage contributions: (A) S_0 to S_4 (57.1%), and (B) S_0 to S_5 (37.0%). S_0 is the ground state, S_n is the nth excited state.
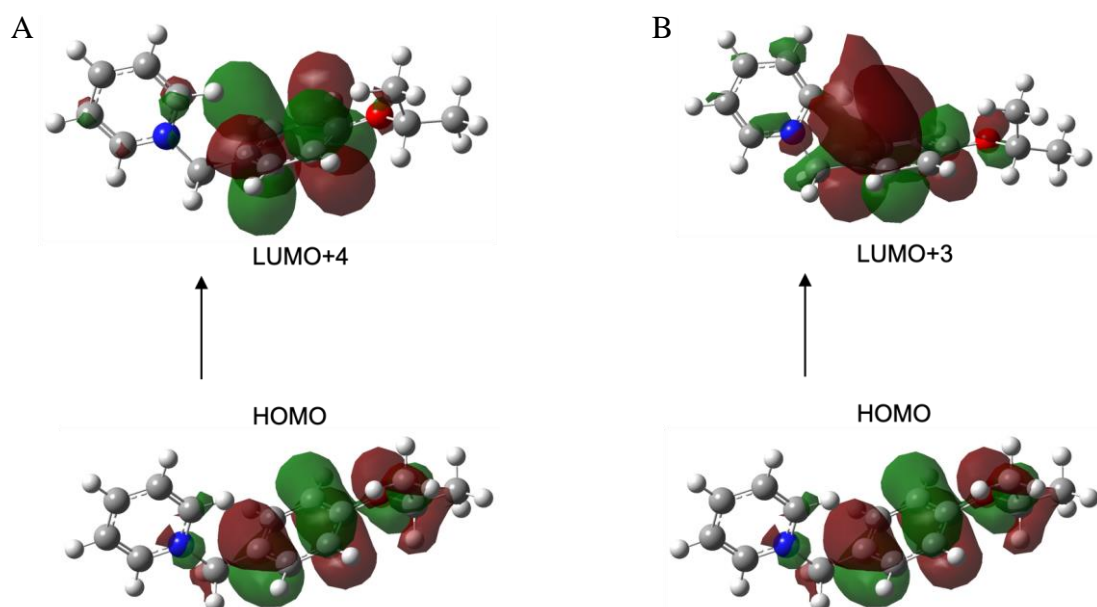
Figure C.15 The calculated molecular orbitals of BP-OiPr ion for transitions with percentage contributions: (A) S_0 to S_4 (64.9%), and (B) S_0 to S_5 (42.0%). S_0 is the ground state, S_n is the nth excited state.
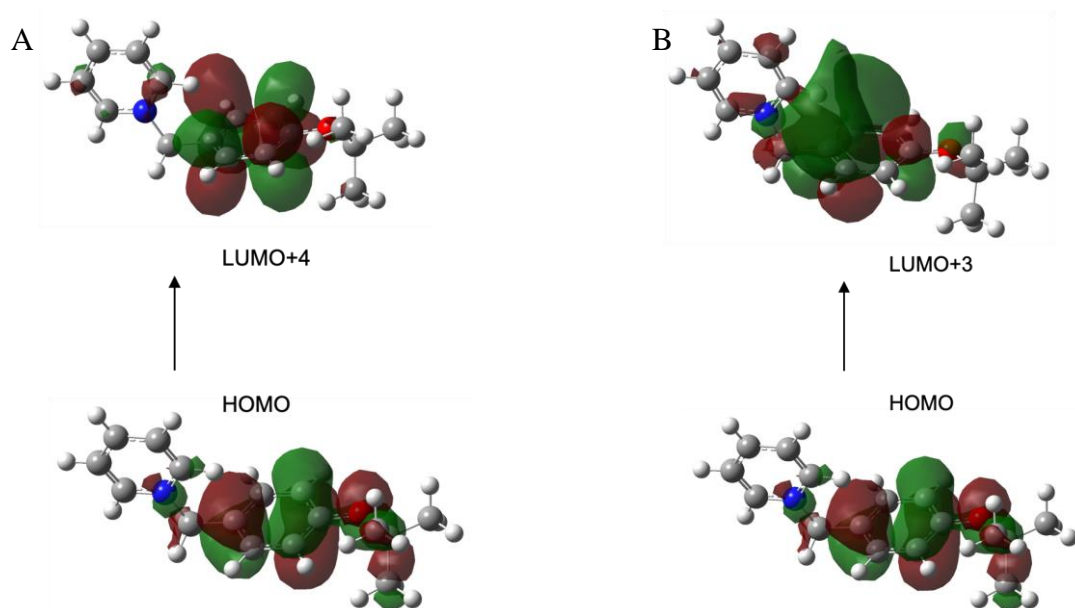

Figure C.16 The calculated molecular orbitals of BP-OtBu ion for transitions with percentage contributions: (A) S_0 to S_4 (66.9%), and (B) S_0 to S_5 (44.0%). S_0 is the ground state, S_n is the nth excited state.