The Structures and Properties of Protonated Phenylalanine Derivatives

by
Weiqiang Fu

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

Waterloo, Ontario, Canada, 2016

Author's Declaration

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

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

Abstract

Non-covalent interactions influence the 3-dimentional structures of large biomolecules. Hydrogen bonds, π-π stacking and cation-π interactions play a prominent role in protein folding and molecular recognition. In clusters peptide, and proteins that contain phenylalanine (Phe), the cation-π interaction arises from the interaction of the quadrupole charge distribution of the phenyl group with a positively charged species. By varying substituents around the phenyl ring one can potentially tune the cation-π interaction, even going so far as to invert the ring's quadrupole moment, thus changing the cation binding motif. To explore the effects of electron donating/withdrawing groups on the non-covalent interactions of Phe clusters, ionic clusters including Phe, Phe derivatives, proton-bound dimers of Phe derivatives, and Phe dipeptide were investigated in a combined experimental and computational study. The low-lying isomers of various clusters of Phe were identified using the Basin-hopping (BH) search algorithm and optimized with density functional theory (DFT). The predicted harmonic vibrational spectra were then compared with experimental spectra obtained via infrared multiple photon dissociation (IRMPD) to determine cluster geometries.

Acknowledgements

First and foremost, I would like to acknowledge the help and guidance of my supervisor Dr. W. Scott Hopkins. Starting from my undergraduate project until now, it has been over three years with your constant and generous help not only in my studies, but also in my life.

I would also like to thank my committee members Terry McMahon and Larry Campbell for guiding me through my thesis and making time for my presentations.

I would also like to express my gratitude to Hopkins group former and current members for their contributions to my research. I would specially like to thank Stephen Walker, Patrick Carr, and Ce Zhou for reading my thesis and giving me helpful feedback. I really appreciate Ce Zhou's efforts in drawing the figure of Bruker Esquire 3000+ ion trap mass spectrometer. Also thanks go to Patrick Carr, Christian Ieritano, and Tina Lee for helping me with my research, Mike Lecours, Luke Melo and Ce Zhou for programing, and Zach Johnston, Sarah Ajami, Dalia Naser, and Shabna Mohideen for sharing their research results with me.

In addition, I would like to express my appreciation to Dr. W. Scott Hopkins, Dr. Eric Fillion, Dr. Terry McMahon, Dr. Michael Burt, Dr. Vincent Steinmetz, and Patrick Carr who did the experiments at the Centre de Laser Infrarouge d' Orsay at the Unversity of Paris XI.

Last, but not least, I would like to thank my family and friends. Special thanks to my parents for their support and encouragement.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

Arg	Arginine
B3LYP	Becke, three-parameter, Lee-Yang-Parr
BH	Basin hopping
CHELPG	Charges from electrostatic potentials using a grid based method
CLIO	Centre Laser Infrarouge D'Orsay
DFT	Density Functional Theory
EDGs	Electron donating groups
ESI	Electrospray ionization
ESP	Electrostatic potential
EWGs	Electron-withdrawing groups
FEL	Free electron laser
GM	Global minimum
GPB	gas-phase basicity
GTOs	Gaussian-type orbitals
HF	Hartree-Fock
IMHB	Intramolecular hydrogen bond
IR	Infrared
IRMPD	Infrared Multiple Photon Dissociation
IVR	Intramolecular vibrational redistribution
L-DOPA	3,4-dihydroxyphenylalanine
MM	Molecular mechanics
PA	proton affinity
PES	Potential Energy Surface
Phe	Phenylalanine
PhePhe	Phenylalanine dipeptide
PM6	Parameterization method 6
TMA	Trimethylamine
Trp	Tryptophan
Tyr	Tyrosine
UFF	Universal force field

Chapter 1

Introduction

1.1 Non-covalent interactions

Non-covalent interactions, which are usually very weak relative to covalent interactions, are crucial to maintaining the three-dimensional structures of large molecules.[1] Typically, non-covalent bonds have energies in the range of 1−5 kcal/mol.[1] Thus molecules can break non-covalent bonds relatively easily, facilitating, *e.g.*, dynamic biological processes. Even though non-covalent interactions are individually weak and exist ephemerally in biological systems, the cumulative energies of multiple non-covalent interactions can be significant and, combined, numerous non-covalent bonds can significantly stabilize molecular structures.

Non-covalent interactions refer to the collection of multipolar electrostatic interactions that may be present in molecular systems. Permanent dipole-dipole interactions, dipole-induced dipole interactions, induced dipole-induced dipole interactions (London dispersion forces), and charge-quadrupole interactions are all common examples of non-covalent interactions.[1] Interestingly, hydrogen bonds arise from both covalent and non-covalent interactions; they exhibit incipient chemical bonding via lone pair electron density donation to the D-H σ* molecular orbital, and directional interactions between the electronegative acceptor atom and the electropositive hydrogen atom that is covalently bound to the electronegative donor atom.[1] In

biological systems, the most common acceptor and donor atoms are nitrogen and oxygen, respectively. For proteins, peptides, and amino acids, the carbonyl oxygen atoms act as acceptors, while hydroxyl oxygen atoms and amine nitrogen atoms can be both acceptors and donors.

The cation-π interaction, which is a type of charge-quadrupole interaction, has also been identified as an important influence on protein secondary structure.[2] This interaction typically involves common metal cations (*e.g.*, Na^+, K^+) and amino acid residues that contain conjugated π-systems (*e.g.*, phenylalanine [Phe], tryptophan [Trp]).[3–5] To model these systems, researchers have investigated cation-π interactions between cationic species and small π-conjugated molecules such as benzene.[6–9] However, cation-π interactions in biological systems, are much less studied. Recently, the McMahon laboratory at the University of Waterloo measured the binding energies of cation-π systems containing aromatic residues of amino acids and free organic or metal cations. This work demonstrated enhancements in binding energies for Phe, Trp and tyrosine (Tyr) on account of cation-π interactions.[10] Here, we extend this work. Rather than exploring intermolecular cation-π systems, we investigate systems that have the potential to exhibit intramolecular charge-quadrupole interactions. Specifically, we focus on protonated Phe derivatives wherein substitution of the phenyl ring with electron donating groups (EDGs) or electron withdrawing groups (EWGs) affect the local quadrupole moment of the ring and, therefore, molecular geometry.

1.2 Quadrupole moments

Generally, a quadrupole can be thought of as two dipoles oriented in an antiparallel fashion

(Figure 1.1 a). For example, in benzene, the quadrupole moment is oriented such that the regions above and below the plane of the ring possess a relatively high electron density compared to the plane of the molecule (Figure 1.1 b). Consequently, cations are attracted to the high electron density of the ring faces, while anions bind to the ring edge.[2] The (perhaps surprisingly large) binding energies for alkali metal cations to benzene are summarized in Table 1.1.[2,4] The binding energy trend of $Li^+ > Na^+ > K^+ > Rb^+$ mirrors that of alkali cation binding to Cl^-.[4] This indicates that the electrostatic charge-quadrupole interaction is dominant over charge induced dipole interactions associated with molecular polarizability, dispersion, or charge-transfer, since these interactions would favour the larger Rb^+.[4] This conclusion is also supported by the fact that benzene binds alkali cations more strongly than does cyclohexane, even though cyclohexane is the more polarizable molecule.

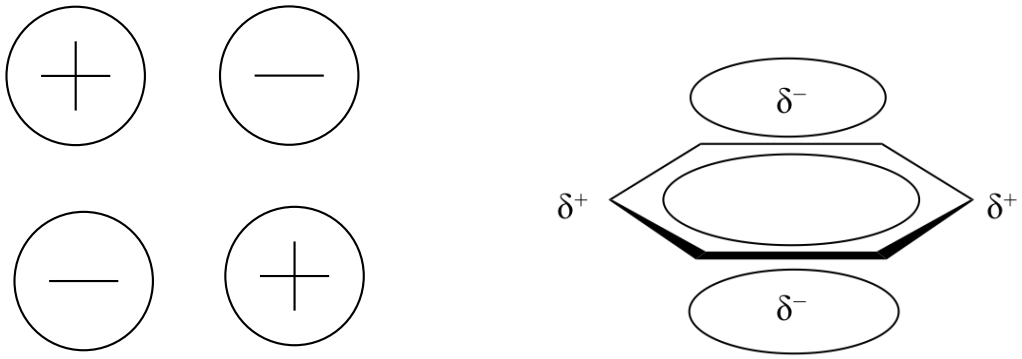


Figure 1.1 The representations of quadrupole moments: a) general quadrupole moments represented as two dipoles; b) the quadrupole moment in benzene

Table 1.1 The binding energies for alkali metal ions to benzene in gas phase.

Alkali metal ions	Binding Energies (kcal/mol)
Li^+	38
Na^+	28
K^+	19
Rb^+	16

As mentioned above for Phe derivatives, the quadrupole moment of benzene can be inverted if EWGs replace the hydrogen atoms. For example, the ring edge of hexafluorobenzene exhibits a relatively high electron density compared to the regions above and below the molecular plane, which is exactly opposite to the quadrupolar charge distribution of benzene. In this case, it is the face of hexafluorobenzene which binds anions, whereas cations interact with the ring edge.[11] The evolution of molecular structure as influenced by the varying charge-quadrupole interaction upon substitution of the π-system with EWGs (or EDGs) is an interesting open question.

1.3 Phenylalanine

The amino acids Phe, Trp, and Tyr, have been found to coordinate cations with their conjugated π-systems.[12,13] Since Phe is the smallest common amino acid containing a phenyl ring (see Figure 1.2),[14] it offers the best (*viz.* simplest) opportunity to study biologically-relevant cation-π interactions. In particular, owing to its relatively small size, Phe is the most

computationally tractable of the three naturally occurring aromatic amino acids. This is especially important since experimental outcomes will need to be supported with high-level quantum chemical calculations.

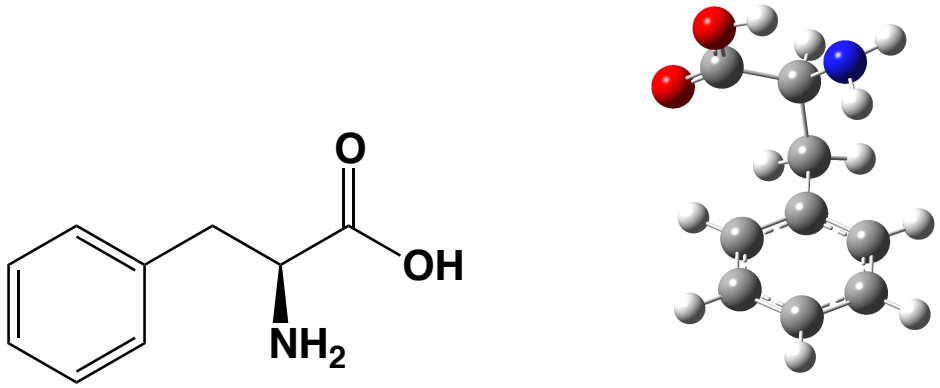


Figure 1.2 The structure of L-phenylalanine in gas phase

Many studies have described the cation-π interaction between amino acids and free organic or metal ions.[2–5,10] Here, protonated Phe derivatives were studied to investigate the effect of intramolecular cation-π interactions on molecular structure. Instead of a non-covalently bound metallic cation, either a protonated carbonyl group or an ammonium group of the Phe·H^+ derivative will be participating in the intramolecular cation-π interaction. Following characterization of the protonated monomers, the proton-bound dimers of the Phe derivatives were also studied. This increased complexity introduces additional non-covalent and hydrogen bonding interactions, which may be more representative of a true biological system. Moreover, it is well known that the amino acids exhibit a zwitterionic (charge-separated) structure in protic solution, and a canonical (charge-solvated) structure in the gas phase.[15–18] It has been shown that the zwitterionic form(s) of amino acids can be stabilized in the gas phase by interaction with

(some) small molecular ions.[15–18] Trimethylamine (TMA) has been reported to lead to zwitterionic-like structure of 3-cyanophenylalanine (3CN-Phe), which suggests that TMA is good for stabilizing zwitterionic structures.[15] Owing to the very high proton affinity (PA) of the guanidinium group in the side chain, arginine (Arg) has also been found to exhibit a zwitterionic structure in the Arg_2 neutral dimer.[19] The study of the proton-bound heterodimers of glycine with Phe and pentafluoro Phe was published recently.[20] It has also illustrated that the N-H^+•••O and N-H^+•••N intermolecular binding motifs correlate with the difference of gas-phase basicity (GPB) in the amino acid moieties. Accordingly, whether the zwitterionic (or other unexpected) structures of the proton-bound homodimers of Phe derivatives can be formed in the gas phase is well worth exploring.

Chapter 2

Introduction to Computational Methods

Global minimum candidates of the Phe derivatives were generated using a custom written basin hopping (BH) code[21,22] that was interfaced with the Gaussian 09 software package.[23] Each molecule and cluster underwent BH for structural sampling, which can be considered as Monte Carlo with minimization. A flow chart of the BH routine is shown in Figure 2.1. After an initial guess structure is submitted to BH, the geometry is randomly distorted and then optimized according to the parameters defined in BH input. These parameters include simulation temperature (*i.e.*, thermal energy), box size (*i.e.*, containment volume), rotation angles, and translation steps. After a new structure is obtained, it is geometrically optimized to identify the most stable isomer in that given region of the potential energy surface (PES). If the energy of this isomer is lower than the energy of current global minimum, the isomer is accepted as the new global minimum. The input geometry is then re-distorted and the algorithm repeats.[22] If the energy of the isomer is higher than the energy of the current global minimum, its energy is compared with an energy that is randomly selected from a user defined Boltzmann distribution.[22] A favourable comparison (lower isomer energy) signals acceptance of the randomly generated structure and its geometry is passed on for further distortion by the BH algorithm. If the isomer energy is higher than the Boltzmann energy, the newly distorted structure is rejected, and the BH routine re-distorts the previous structure. Typically, the BH algorithm undergoes several thousand iterations prior to completion, with an acceptance: rejection ratio of approximately $\varepsilon = 0.5$. In other words, a BH routine that conducts 10,000 steps will sample approximately 20,000 structures.

Owing to the large number of geometry optimizations that are required for a BH search of a PES, a relatively low level of theory is employed for BH calculations. Typically, molecular mechanics (MM) is used. MM employs classical mechanics to model molecular systems and treats molecules as classical balls connected by springs.[24] There are various contributions to the MM total energy, E_{total} (see Equation 2.1): bonding interactions (harmonic oscillator) including bond stretching ($E^{stretch}$), angle bending (E^{bend}), and angle torsion ($E^{torsion}$), and non-bonding (van der Waals and Coulombic) interactions ($E^{non\text{-}bonded}$).[25] In this work, the universal force field (UFF) was used for MM calculations because the force field contains parameters for all elements (necessary for some Phe derivatives).[26] The parameters of UFF contains a series of hybridization dependent atomic bond radii, hybridization angles, van der Waals parameters, torsional and inversion barriers, and nuclear charges.[26] Due to the relatively poor accuracy of MM calculations, the unique structures that are identified by MM BH searches must be re-optimized at a higher level of theory. Consequently, the accuracy of force field is not especially important since BH is being used simply to generate candidate structures for more accurate electronic structure calculations (e.g., DFT, MP2, CCSD). As MM calculations require partial charges for each atom in the molecule (to more accurately model non-covalent electrostatic interactions), CHelpG calculations were performed to calculate the partial charges for each atom in the monomeric species.[27]

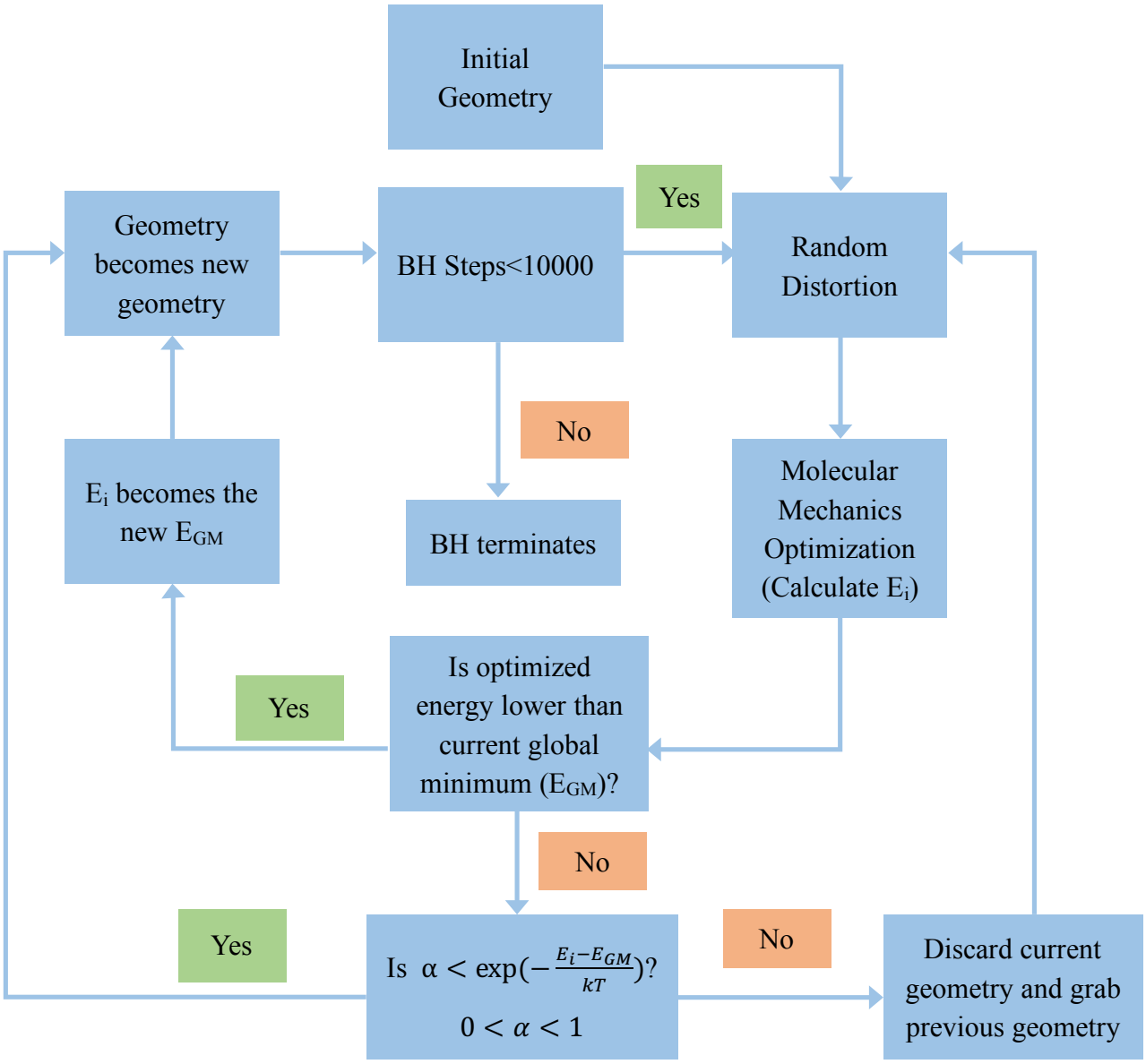


Figure 2.1 A flow chart of BH routine. This routine will iterate until the max number of steps are accomplished. E_i is the energy of a random structure optimized by MM. E_{GM} is the energy of the current global minimum. α is a random number chosen from 0 to 1.

$$E_{total} = \underbrace{\sum_{A}^{bonds} E_A^{stretch} + \sum_{A}^{bond\ angles} E_A^{bend} + \sum_{A}^{torsion\ angles} E_A^{torsion}}_{harmonic\ oscillator} + \sum_{AB}^{non-bonded\ atoms} E_{AB}^{non-bonded}$$

Equation 2.1

To map the PESs of the Phe derivatives monomers, 10,000 BH steps were conducted. The same low energy isomers, as determined by their energies and geometries, were identified many times. All of the low-energy unique isomers were carried forward for treatment using semi-empirical PM6 so as to better estimate relative energies and geometries. This semi-empirical method, which is based on Hartree-Fock (HF) theory (but which has more approximations and uses some empirical data), is less computationally expensive than HF without sacrificing much of the accuracy of HF.[24,28,29] Following treatment at the PM6 level of theory, the 10 lowest-energy structures were further optimized by density functional theory (DFT) using the B3LYP hybrid functional with the 6-31+G(d,p) basis set. The resulting structures were then re-optimized and their frequencies were obtained with the larger 6-311++G(d,p) basis set. The DFT method, which utilizes the total electron density to describe the system of interest, has not only better accuracy than the HF method, but also lower computational cost than the perturbation based methods due to the incorporation of density functionals to approximate the non-classical electron correlation.[24] Among the functionals, B3LYP, which stands for Becke, three-parameter, Lee, Yang, and Parr, has been chosen because it performs well for similar systems and is a common choice of functional.[24] Normal mode analyses were undertaken to ensure that the stationary points are local minima on the PES. The calculated frequencies were used to predict the vibrational spectra for these isomers.

Chapter 3

Introduction to Experimental Methods

In principle, each molecule has a unique vibrational spectrum that arises due to molecular symmetry and bonding. To measure the gas phase vibrational spectra of the protonated Phe derivatives and clusters, the infrared multiple photon dissociation (IRMPD) technique was employed. Many research papers have described IRMPD,[17,30–33] so a detailed description will not be provided here. Briefly, IRMPD is a stepwise absorption process that can be used to record the infrared (IR) vibrational spectra of ions (Figure 3.1). When a molecule absorbs the energy of a single photon via a resonant vibrational mode, excitation occurs from the ground v=0 state to the v=1 level of the associated normal mode. The absorbed photon energy is then distributed throughout the rest of the molecule via intramolecular vibrational redistribution (IVR).[34] Once the IVR process occurs, the excited vibrational level is depopulated and subsequent single-photon absorption can again occur via the resonant vibrational transition.[35] Through repeated absorptions, the molecule is heated one photon at a time until it reaches the first thermodynamic dissociation threshold and fragmentation occurs. By monitoring the depletion of the parent ion signal and production of fragment ions (via mass spectrometry) as a function of excitation wavenumber, an IRMPD action spectrum is generated.[34] The efficiency of the IRMPD process, which is plotted as a function of wavenumber, is calculated as per Equation 3.1. IRMPD spectra are relatively accurate in terms of the vibrational frequencies, but care must be taken when interpreting intensities since these are dependent on the efficiency of IVR and coupling to dissociative channels in addition to photon absorption cross section.

$$\mathrm{IRMPD}_{efficiency} = -\log\left(\frac{\sum Intensity_{parent}}{\sum Intensity_{parent}+\sum Intensity_{fragment}}\right) \qquad \text{Equation 3.1}$$

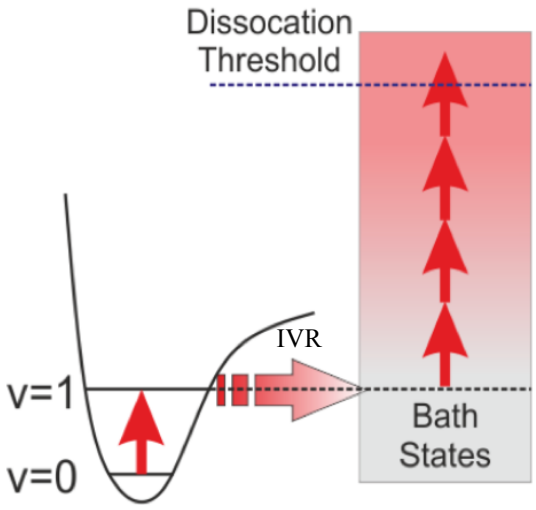


Figure 3.1 The Infrared multiple photon dissociation (IRMPD) process.[36] The absorbed energy is transferred from the absorbing local mode to the bath states. The cluster can dissociate until the energy reaches the threshold.

A tunable IR laser which has high peak power and relatively small bandwidth, is required for IRMPD. To meet these requirements of IRMPD experiments, a free electron laser (FEL), as shown in Figure 3.2, can be used. A FEL is composed of a thermoionic electron gun, a 15~50 MeV electron accelerator, an undulator, and other assembly units. The electron gun provides a beam of electrons, which is accelerated to near the speed of light by the accelerator. The beam of electrons then passes through the undulator (Figure 3.3), which consists of a periodic arrangement of magnets that forces the electron beam to oscillate periodically. This results in the emission of synchrotron radiation. Capturing the emitted photons in an appropriate optical cavity, stimulates emission of intense coherent light.[37,38] The wavelength of the emitted light can be tuned by modifying the undulator magnetic field, which can be achieved by changing either the gap between the magnets or the power supply current of the electromagnets.[38]

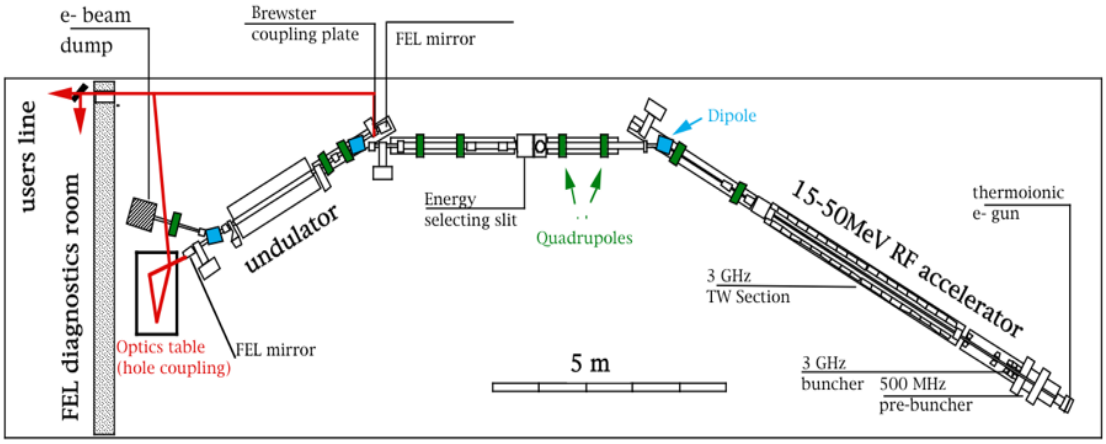


Figure 3.2 Schematic diagram of a Free Electron Laser (FEL)[39]

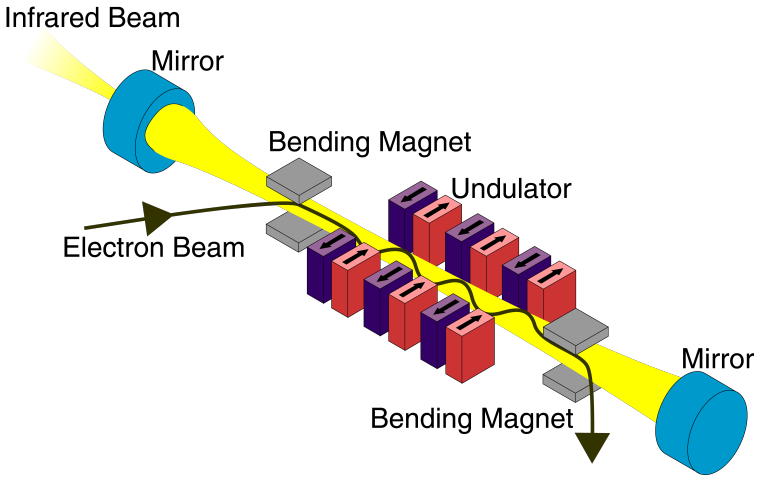


Figure 3.3 Schematic representation of the FEL undulator[39]

To produce and isolate clusters of interest, an electrospray ionization (ESI) source which is mated with ion trap spectrometer (Bruker Esquire 3000+) is utilized.[40] As shown in Figure 3.4, the analyte can be ionized, brought into the gas phase, mass-selected, and trapped prior to interrogation by the FEL.

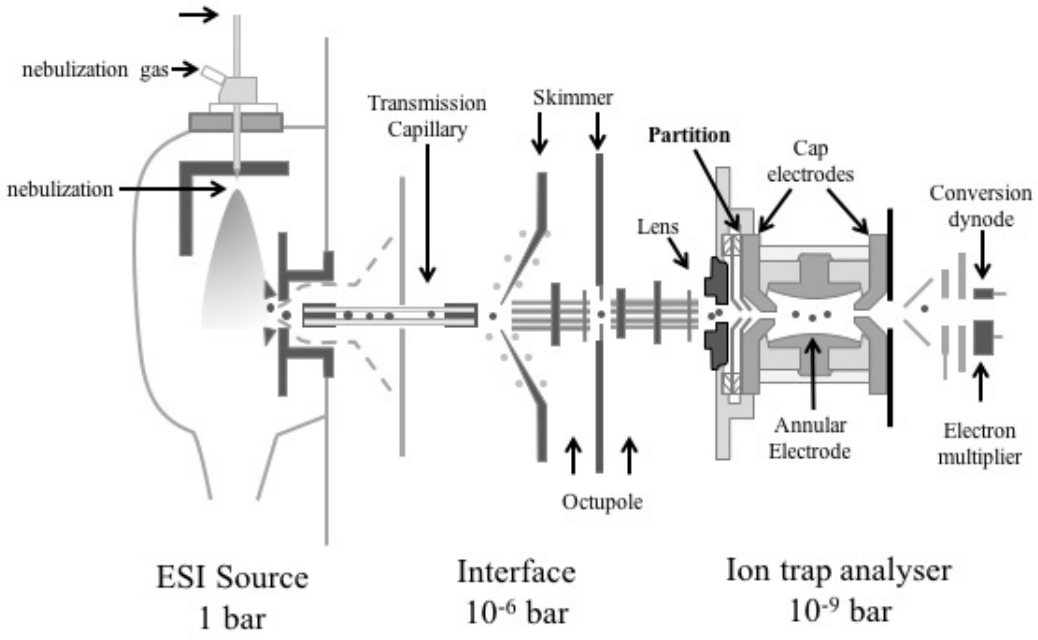


Figure 3.4 The Bruker Esquire 3000+ ion trap mass spectrometer[41]

Chapter 4

The Structures of Protonated Phe Derivatives

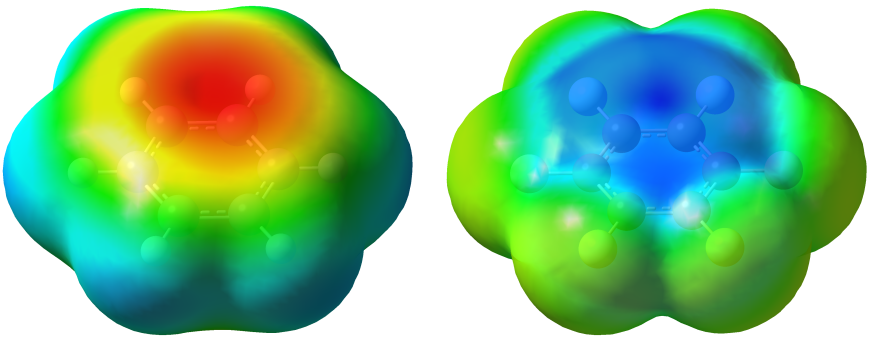
Overview

The structures and properties of ionic clusters of phenylalanine derivatives were investigated through a combined experimental and computational approach. The quadrupole charge distribution of the phenyl group creates a cation-π interaction that can affect the structure of phenylalanine (Phe) and its derivatives. The substituents on the aromatic ring, which can be electron-withdrawing or electron-donating, can influence the strength of the cation-π interaction, and even invert the quadrupole moment of the phenyl ring. In this project, the effects of aromatic substitution on the structures of protonated phenylalanine derivatives were explored. The low-energy structures of each phenylalanine derivative were identified using the basin-hopping (BH) algorithm and refined with density functional theory (DFT). Theoretical infrared spectra were obtained for the low energy structures and were then compared with the experimental spectra obtained via infrared multiple photon dissociation (IRMPD) at the Centre de Laser Infrarouge d' Orsay (CLIO). Proton affinities, and quadrupole charge distributions are also discussed.

4.1 Introduction

Species containing an aromatic phenyl ring contain a quadrupole moment owing to the delocalized π system of the phenyl group, which results in a relatively high negative charge

density above and below the ring, and a partial positive charge in the ring plane.[2] As described in Chapter 1, benzene, an ideal model for such a π system, has such a quadrupole charge distribution. This quadrupolar electron density distribution is visualized in Figure 4.1A.[4] Consequently, cationic species tend to interact with the face of the ring while anionic species tend to interact with the edge of the benzene ring. The effects of substituents on these non-covalent interactions were also studied. Electron donating groups (EDG) can strengthen the cation-π interaction, while electron withdrawing groups (EWG) weaken the interaction.[42] Under some extreme circumstances, where all the hydrogen atoms are replaced by strong EWGs, such as fluorine, the quadrupole moment can be inverted (Figure 4.1B). Phenylalanine (Phe), as an aromatic amino acid, plays an important role in biological processes. The quadrupole moment of the phenyl ring in Phe can participate in cation-π interactions. To investigate how the cation-π interaction influences the structures of the clusters of phenylalanine with different substituents, Phe derivatives were studied using an experimental and computational approach.



A B

Figure 4.1 The Electrostatic potential (ESP) surface of (A) benzene, and (B) hexafluorobenzene; ESP mapped onto total density; red color is for negative; blue color is for positive.

To examine the effects of EDGs and EWGs on the structure of protonated Phe derivatives and their clusters, fluoro, cyano, trifluoromethyl, and nitro were chosen as strong EWGs, and methoxy was chosen as a strong EDG for substitution on the phenyl ring. The effect of the different sites and the number of the substituents on the structures of Phe derivatives were also studied. Specifically, the molecules of interest are: phenylalanine (Phe), 3-fluorophenylalanine (3F-Phe), 3CN-Phe, 3-trifluoromethylphenylalanine (3CF_3-Phe), 4-fluorophenylalanine (4F-Phe), 4-nitrophenylalanine (4NO_2-Phe), 2,5-difluorophenylalanine (2,5F_2-Phe), 3,5-difluorophenylalanine (3,5F_2-Phe), 3,4-dimethoxyphenylalanine (3,4$(MeO)_2$-Phe), and pentafluorophenylalanine (F_5-Phe). Each Phe derivative has two possible protonation sites: the amine nitrogen (N-protonated) and the carbonyl oxygen (O-protonated). These are the most favourable protonation sites owing to the high electronegativity and valence lone pairs on N and O.[43] Neutral Phe derivatives were calculated as well to explore proton affinity.

4.2 Methods

Both computational and experimental methods were utilized in this project. BH and DFT electronic structure calculations, as described in Chapter 2, are applied. Phe derivatives are studied experimentally with IRMPD spectroscopy (Chapter 3).

4.2.1 Experimental methods

The experimental IRMPD spectra of the protonated Phe derivatives were acquired at the Centre de Laser Infrarouge d' Orsay (CLIO) free electron laser (FEL) facility at the University of

Paris XI.[37] A detailed description of the experimental apparatus has been previously reported.[17,32,33] 100 μM aqueous solutions of 50:50 methanol:water with additional 0.1% formic acid consisting of stoichiometric quantities of Phe derivatives (Alfa Aesar) were prepared, and chemicals were used without further purification.[15] Positive mode electrospray ionization (ESI) was employed to produce gas phase proton-bound Phe derivative clusters. The protonated Phe derivatives were then delivered to a Bruker Esquire 3000+ ion trap mass spectrometer for mass selection and interrogation by the FEL.[15] The power output from the FEL ranged from 800 mW to 1200 mW, and spectra were acquired over the 1000–2000 cm^{-1} range. IRMPD spectra were then collected as a function of FEL wavenumber by monitoring the depletion of parent ions and production of daughter ions.

4.2.2 Computational methods

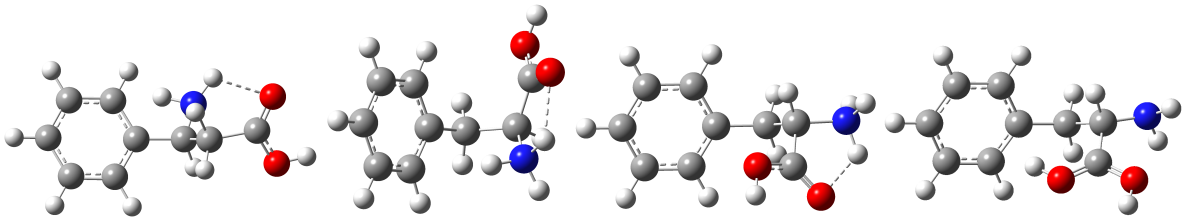
To carry out BH, each protonated Phe derivative was modeled using the Molecular Mechanics (MM) Universal Force Field (UFF) with partial charges that were calculated with the CHelpG partition scheme at the B3LYP/6-31G level of theory. For each BH step, an internal dihedral rotation was randomly chosen from the range $-5° \leq \varphi \leq 5°$. In total, 10,000 BH steps for each Phe derivative were obtained to sample potential energy surface (PES) of both the N-protonated and O- protonated tautomers, respectively. Unique isomers were defined by an energy variance of 10^{-5} hartree and a geometry variance of 0.5 Å. These structures were then further optimized by semi-empirical PM6, and then DFT. Frequencies were obtained to predict harmonic

vibrational spectra, which were later compared with IRMPD spectra, and thermodynamic corrections for each isomer.

4.3 Results and Discussion

4.3.1 Computed structures

The calculated global minimum (GM; lowest-energy) of phenylalanine and isomers that have relative Gibbs' energies within 100 kJ/mol of the global minimum are shown in Figure 4.2. Although both protonation sites (N- & O-) were tested, the three lowest-energy isomers (Isomer 1, 2, and 3) were all found to be N-protonated. Isomer 4 is the only O-protonated structure identified computationally, and its relatively high Gibbs' energy suggests that it is unlikely to be formed in the gas phase. The low-energy structures also clearly show that intramolecular hydrogen bonding contributes to the stabilization of the structures. Furthermore, it is apparent that the face of phenyl ring orients to interact with the site of protonation. Given the relatively low energies of the N-protonated isomers, it is possible that they are all present in the gas phase. To test this hypothesis, their calculated IR spectra can be compared with the experimental IRMPD spectra.

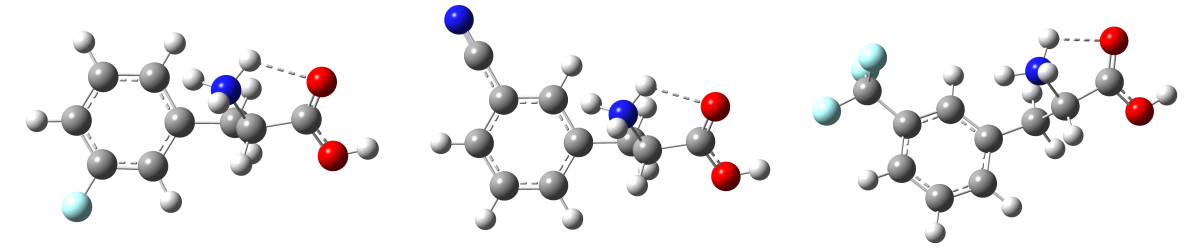


a) Isomer 1	b) Isomer 2	c) Isomer 3	d) Isomer 4
0	4.95	22.78	95.68

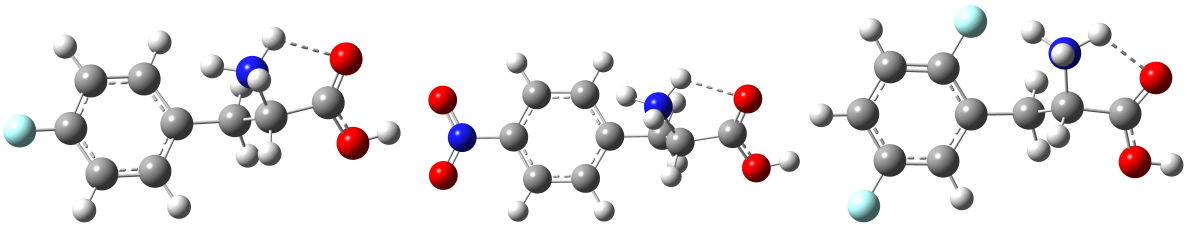
Figure 4.2 The four lowest-energy isomers of protonated phenylalanine. Relative Gibbs' energies including thermal corrections at 298 K are given in kJ/mol. Calculations used the B3LYP functional and 6-311++G(d,p) basis set.

The calculated Gibbs' energies of all of the Phe derivatives that were identified with our BH search and subsequent DFT treatment are available in Appendix I. Figure 4.3 only shows the protonated global-minimum structures. The GM structure of each molecule is N-protonated, and every GM structure exhibits an intramolecular hydrogen bond (IMHB) between the ammonium group and the carbonyl oxygen atom. The orientation of the phenyl ring in these structures is a good indication of the presence of a cation-π interaction since the rings are oriented to maximize the facewise or edgewise charge-quadrupole interactions. With one EWG, the phenyl ring orients to face the site of protonation. In the case of F_5-Phe, where the ring quadrupole has been inverted, it is the edge of the phenyl ring which interacts with the protonation site. While these observations are consistent with predictions based on the cation-π interaction, it is likely that other electrostatic and charge-transfer interactions also play a role. For example, the edgewise interaction in F_5-Phe can also arise from electrostatic and hydrogen bonding interactions between

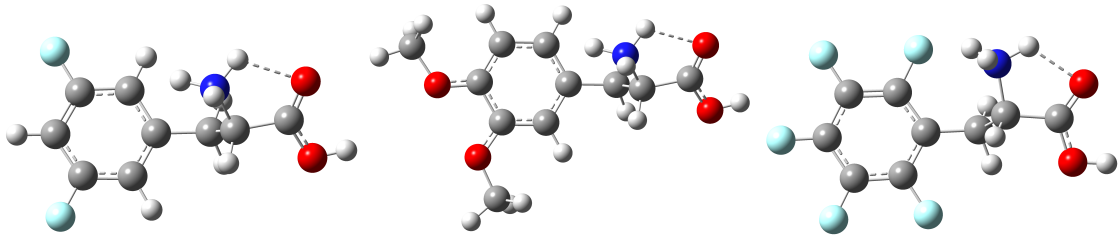
the ammonium group and the F atoms in the 2- and 6-positions on the phenyl ring. The structures of the 2,5- and 3,5-difluoro Phe derivatives demonstrate the influence of these additional interactions. Whereas both difluoro derivatives are expected to exhibit similar local quadrupoles about the phenyl ring, only the 2,5-difluoro derivative can exhibit direct $-NH_3^+$•••F interactions. These additional interactions result in a GM structure for $2,5F_2$-Phe wherein the phenyl ring is oriented in a more edgewise fashion than is the phenyl ring in the $3,5F_2$-Phe derivative. The structures of the Phe derivatives are discussed in more detail in section 4.3.3.



a) 3F-Phe b) 3CN-Phe c) $3CF_3$-Phe



d) 4F-Phe e) $4NO_2$-Phe f) $2,5F_2$-Phe



g) $3,5F_2$-Phe h) $3,4(MeO)_2$-Phe i) F_5-Phe

Figure 4.3 The global minima of a) 3-Fluorophenylalanine; b) 3-Cyanophenylalanine; c)

3-Trifluoromethylphenylalanine; d) 4-Fluorophenylalanine; e) 4-Nitrophenylalanine; f) 2,5-Difluorophenylalanine; g) 3,5-Difluorophenylalanine; h) 3,4-Dimethoxyphenylalanine; i) Pentafluorophenylalanine; calculated by DFT/B3LYP 6-311++G(d,p)

4.3.2 Proton Affinity and Gas Phase Basicity

PA, which is the change in enthalpy upon protonation, can be viewed as the energy released in the process shown in the Equation 4.1a. PA is closely related to GPB, which is the change in Gibbs' energy upon protonation. Besides enthalpy, Gibbs' energy also takes entropy into account. To calculate the PA and GPB of Phe and its derivatives, the enthalpies and Gibbs' energies of each neutral and protonated molecule were obtained from treatment at the DFT level of theory using a harmonic oscillator and rigid rotor model to calculate the internal energy level structures. We then employed Equations 4.1b (for PA) and 4.1c (for GPB). The calculated PA for Phe (926.0 kJ/mol) compares favourably with the value that was determined experimentally (922.9 kJ/mol).[44] Table 4.1 gives the calculated PAs and GPBs.

$$Phe + H^+ \rightarrow Phe \cdot H^+ \qquad \text{Equation 4.1a}$$

$$PA = \Delta H(Phe) + \Delta H(H^+) - \Delta H(Phe \cdot H^+) \qquad \text{Equation 4.1b}$$

$$GPB = \Delta G(Phe) + \Delta G(H^+) - \Delta G(Phe \cdot H^+) \qquad \text{Equation 4.1c}$$

Table 4.1 The calculated proton affinities and gas phase basicities of Phe and its derivatives at the B3LYP/6-311++G(d,p) level of theory.

Phe Derivatives	-ΔH (kJ/mol)	-ΔG(kJ/mol)
3,4$(MeO)_2$	951.13	915.31
Phe	926.02	896.32

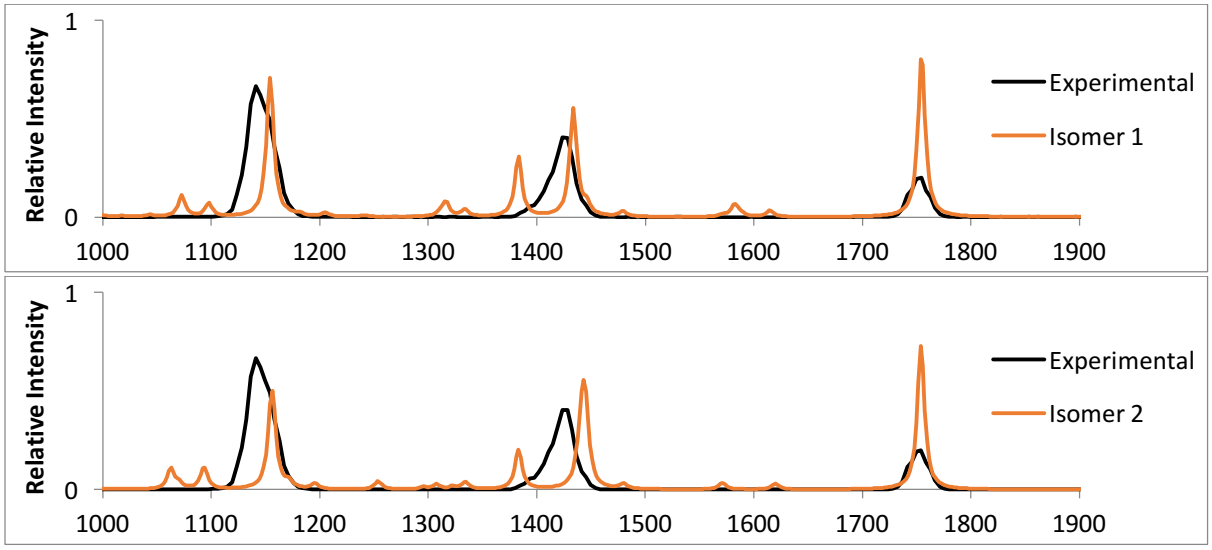
Phe Derivatives	-ΔH (kJ/mol)	-ΔG(kJ/mol)
$2,5F_2$	916.28	883.08
3F	914.74	884.65
4F	914.71	884.87
$3CF_3$	908.84	875.64
$3,5F_2$	903.98	873.96
3CN	897.21	867.76
$4NO_2$	891.75	862.61
F_5	890.37	854.93

In general, the GPBs of the Phe derivatives decrease with the addition of EWGs. Interestingly, the $2,5F_2$-Phe derivative has only a slightly lower GPB compared to 3F-Phe and 4F-Phe derivatives (*ca.* 2 kJ/mol less), whereas the $3,5F_2$-Phe derivative has a GPB that is *ca.* 11 kJ/mol lower. The ortho-positioning of the F in the $2,5F_2$-Phe derivative may be the cause of this behaviour. The –NH•••F interaction might donate electron density to the amine group, thereby enhancing GPB.

4.3.3 Comparison of calculated and experimental IR spectra

Figure 4.4 plots the calculated and experimental IRMPD spectra for the four lowest energy isomers of Phe (shown in 4.3.1). The calculated IR spectra were obtained from frequency calculations at the B3LYP/6-311++G(d,p) level of theory. Compared with the experimental spectra, the vibrational peaks of the calculated spectra are at slightly higher wavenumbers. This arises (predominantly) because the effects of anharmonicity are neglected in the harmonic

frequency calculations.[45] The incomplete incorporation of electron correlation and the use of finite basis sets also contribute to this error.[45] Hence, the calculated spectra shown in Figure 4.4 have been scaled by 0.9679 (a commonly used DFT scaling factor)[46] so as to account for these errors. Phe•H^+ undergoes IRMPD via one mass channel: production of the CO+H_2O (m/z = 46 amu) and cationic $C_8H_{10}N^+$ (m/z = 120 amu). Thus, the ammonium group transfers a proton to the carbonyl upon fragmentation. The calculated spectra of isomers 1, 2 and 3 all show good agreement with the experimental spectrum suggesting that these three isomers are likely present in the gas phase. This is not surprising since these three species are all low-energy conformers of the same general binding motif. The calculated spectrum of isomer 4, on the other hand, does not compare well with the experimental spectrum. We can therefore conclude that this species is unlikely to be present in the gas phase ensemble. Vibrational peak assignments are provided in Table 4.2.



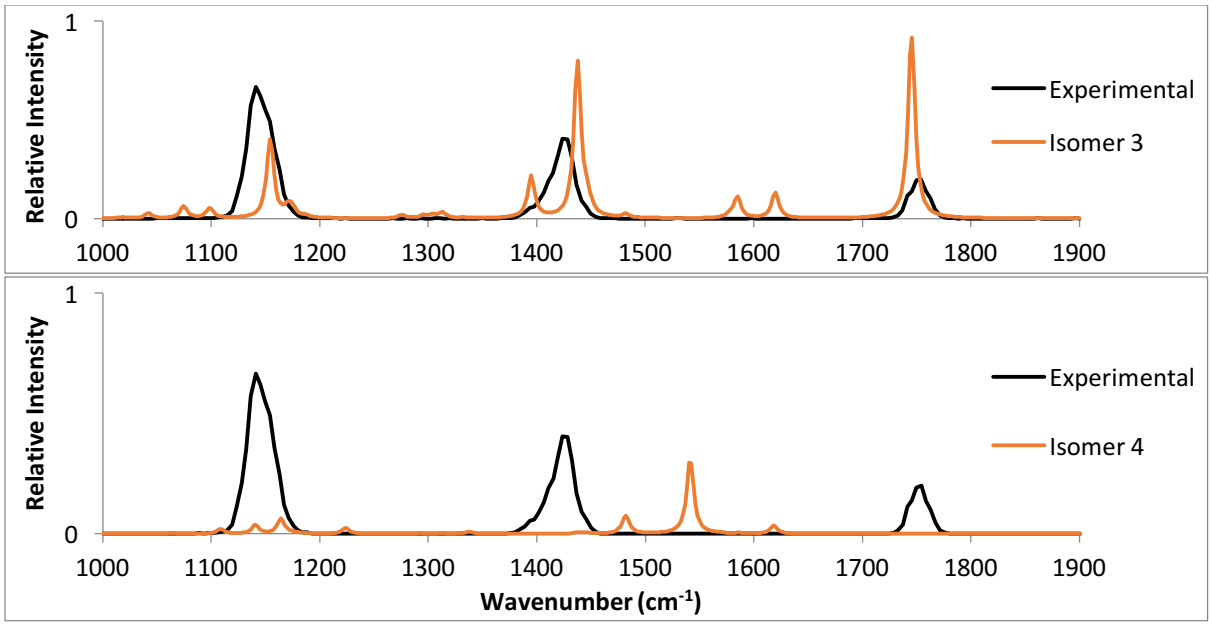


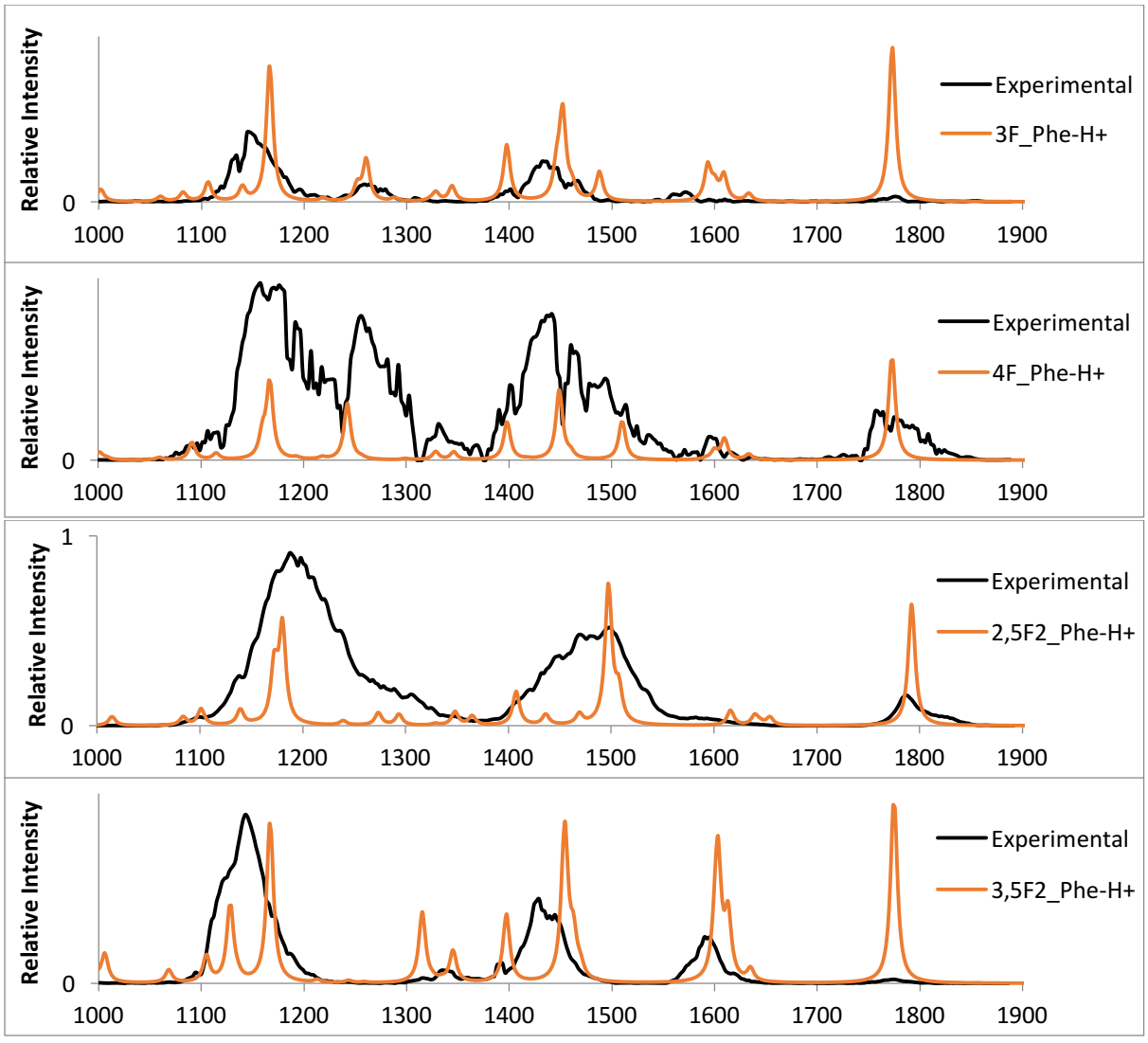
Figure 4.4 Comparison of the experimental IRMPD spectrum of Phe•H^+ with the calculated IR spectra of four lowest-energy isomers of Phe•H^+ as identified by BH and subsequent B3LYP/6-311++G(d,p) treatment.

Table 4.2 Assignments of vibrational modes of protonated phenylalanine, Phe•H^+.

	Vibrational modes (cm^{-1})			
Expt. IRMPD	**1780-1730**	**1460-1380**		**1190-1100**
isomers	C=O Stretch, NH_2 Scissor	NH_3 Umbrella	COH Bend, CH & NH_2 Wag	COH Bend, Ring H Scissor
Isomer 1	1754	1434	1383	1154
Isomer 2	1754	1443	1383	1156
Isomer 3	1746	1438	1395	1154
Isomer 4	1557	NA	NA	NA

Figures 4.5 and 4.6 compare the experimental and calculated IR spectra of the GM for the Phe derivatives. Again, calculated spectra are scaled by 0.9679.[46] Common to all derivatives, the vibrational bands at ~1800 cm^{-1} are assigned to C=O stretch, ~1500 cm^{-1} is the NH_3^+ umbrella, and the COH bend is at ~1200 cm^{-1}. Detailed vibrational mode assignments are given in Table

4.3 and 4.4. In all cases, the calculated spectra of the global minima are in good agreement with those acquired experimentally via IRMPD. As was the case with Phe•H^+, the protonated Phe derivatives dissociate predominantly to produce of CO+H_2O. The 3,4$(MeO)_2$-Phe•H^+ derivative, however, also accesses a NH_3 production channel, two unassigned production channels with m/z = 163 amu and 139 amu, respectively, and the CO+H_2O mass channel at a branching ratio of approximately 52:12:29:7.



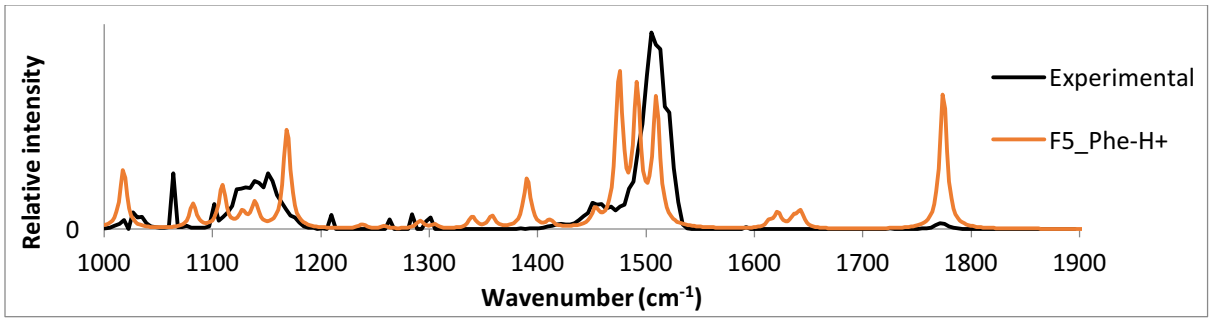
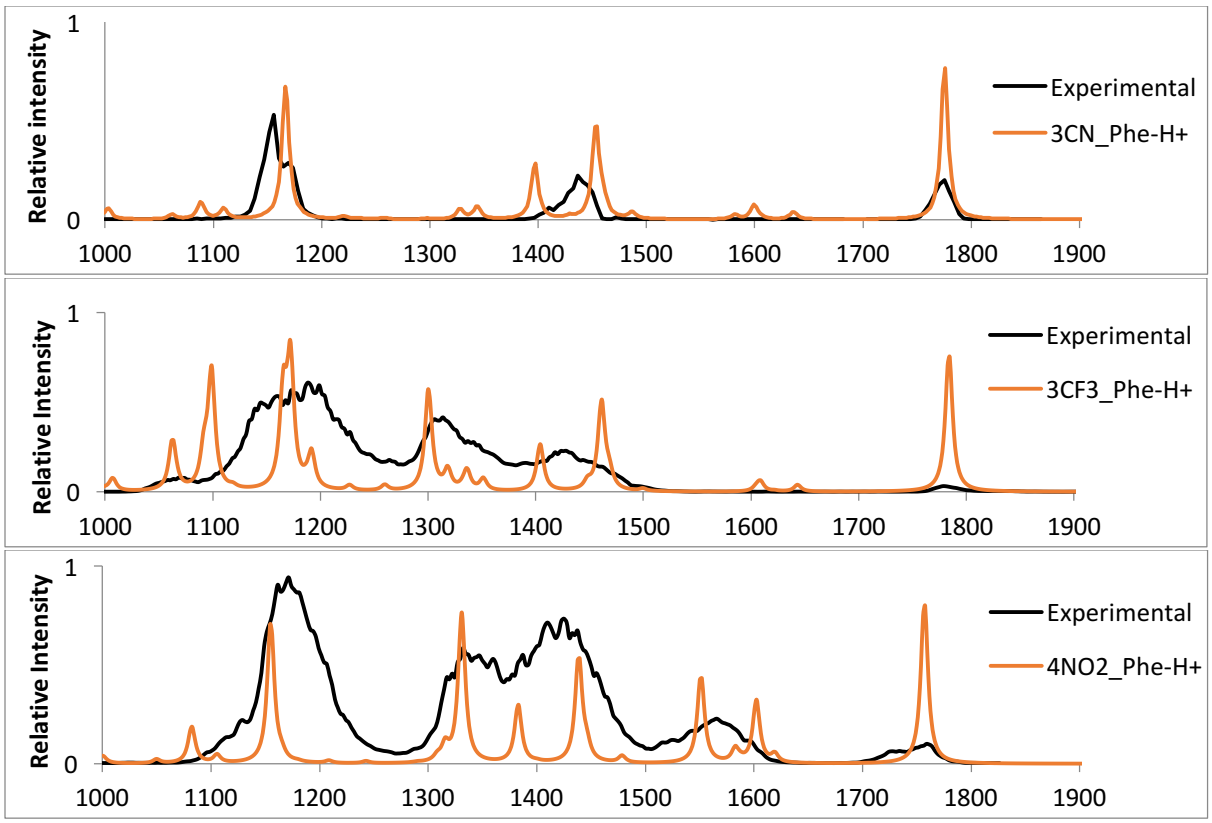


Figure 4.5 Comparison of the experimental IRMPD spectrum of fluorinated Phe•H^+ with the calculated IR spectra of global minima as identified by BH and subsequent B3LYP/6-311++G(d,p) treatment.



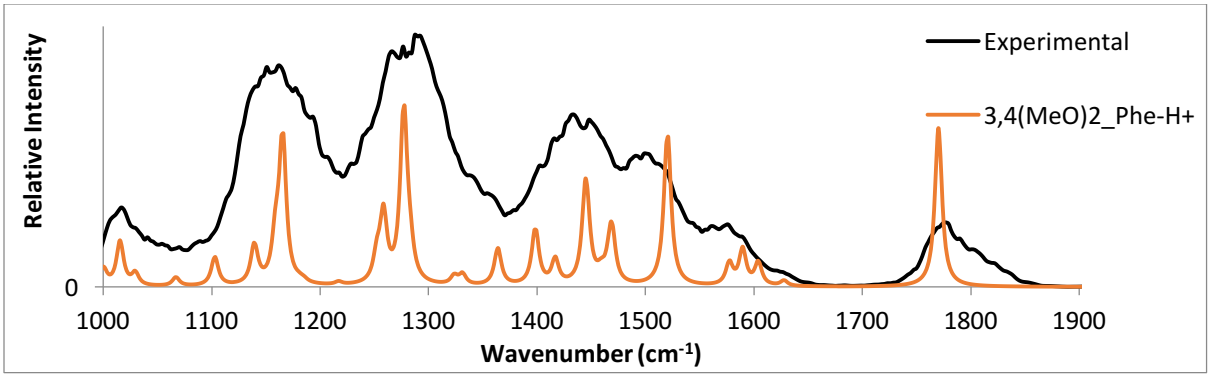


Figure 4.6 Comparison of the experimental IRMPD spectrum of protonated Phe derivatives (different EWGs/EDGs) with the calculated IR spectra of global minima as identified by BH and subsequent B3LYP/6-311++G(d,p) treatment.

Table 4.3 Assignments of vibrational modes of fluorinated Phe•H^+

Structure	Vibrational modes (cm^{-1})									
IRMPD	**1790-1745**	**1620-1540**		**1500-1360**			**1330-1080**			
3F-Phe-H^+	1774	1609	1593	1488	1452	1397	1260	1166		
	C=O Stretch, NH_2 Scissor	Ring Stretch		Ring Rock	NH_3 Umbrella	COH Bend CH Wag	C-F Stretch, Ring Stretch	COH Bend		
IRMPD	**1860-1740**	**1630-1563**		**1563-1375**			**1310 – 1063**			
4F-Phe-H^+	1774	1609		1509	1448	1397	1243	1166		
	C=O Stretch, NH_2 Scissor	Ring Stretch		Ring H Wag	NH_3 Umbrella	COH Bend, CH & NH_2 Wag	C-F Stretch, Ring Stretch	COH Bend		
IRMPD	**1850 – 1750**	**1650-1375**					**1375 – 1075**			
2,5F_2-Phe-H^+	1792	1497	1469	1436	1408		1382, 1168	1161		
	C=O Stretch, NH_2 Scissor	NH_3 Umbrella	CH_2 Scissor	Ring Stretch	COH Bend, CH & NH_2 Wag		COH Bend	Asym. C-F Stretch, Ring Stretch		
IRMPD	**1800 – 1740**	**1650-1550**		**1500-1300**					**1230 – 1080**	
3,5F_2-Phe-H^+	1774	1604		1454	1397	1345	1315	1167	1128	
	C=O Stretch, NH_2 Scissor	Ring Stretch		NH_3 Umbrella, CH_2 Scissor	COH Bend, CH, NH_2 & CH_2 Wag, & C-F Stretch			COH Bend	Asym. C-F Stretch, Ring H Wag	

IRMPD	**1790 – 1760**	**1540-1400**			**1200 – 1090**	
F_5-Phe-H^+	1774	1509	1492	1476	1168	1109
	C=O Stretch, NH_2 Scissor	Ring Stretch, CH_2 Twist		NH_3 Umbrella	COH Bend	Asym. C-F Stretch

Table 4.4 Assignments of vibrational modes of protonated Phe derivatives (different EWGs/EDGs)

Structure	Vibrational modes (cm^{-1})					
IRMPD	**1800-1750**	**1480-1390**				**1200-1090**
3CN-Phe-H^+	1776	1452		1397		1166
	C=O Stretch, NH_2 Scissor	NH_3 Umbrella		COH Bend, CH & NH_2 Wag		COH Bend
IRMPD	**1810-1700**	**1510-1270**				**1270-1040**
3CF_3-Phe-H^+	1776	1454	1397		1294	1166
	C=O Stretch, NH_2 Scissor	NH_3 Umbrella, CH_2 Scissor	COH Bend, CH & NH_2 Wag		Ring Stretch	COH Bend
IRMPD	**1790 – 1700**	**1630-1500**		**1500-1270**		**1270 – 1070**
4NO_2-Phe-H^+	1758	1602	1552	1439	1383	1331	1154
	C=O Stretch, NH_2 Scissor	Asymmetrical N=O Stretch, Ring Stretch		NH_3 Umbrella, CH_2 Scissor	COH Bend, CH & NH_2 Wag	Symmetrical N=O Stretch	COH Bend
IRMPD	**1860 – 1700**	**1670-1000**				
3,4$(MeO)_2$-Phe-H^+	1770	1521		1445	1278	1166
	C=O Stretch, NH_2 Scissor	Ring H Wag		NH_3 & CH_3 Umbrella	Ring Stretch	COH Bend

4.3.4 Assessing charge-quadrupole interactions

Having identified molecular geometries with IRMPD spectroscopy, we can now return to the calculated structures to look for trends as a function of substituents which might be indicative

of charge-quadrupole interaction. In general, we find that the position of the charge center relative to the face of the phenyl ring changes with the number and type of the substituents. To gauge the effect of the electron donating/withdrawing groups on the cation-π interaction, we defined the angles θ and β, which are tilt angles from the axis normal to the phenyl ring plane passing through the ring center (see Figure 4.7). The angle θ we define as the angle between the ammonium N atom, the ring center, and the ring normal as projected into the plane that is perpendicular to the R-phenyl bond axis (see Figure 4.7). The angle β we define as the angle between the ammonium N atom, the ring center and the ring normal as projected into the plane containing the R-phenyl bond. These are shown in Figure 4.7. We have also extracted the distances between the centre of the phenyl ring and the ammonium N atom, as well as the ammonium H atom that is closest to the ring center. These geometric parameters are tabulated in Table 4.5. Note that in alkali cation/benzene half-sandwich complexes, θ and β are both zero and the distance between the charge center and the ring is less than the sum of the respective van der Waals radii.[7,9]

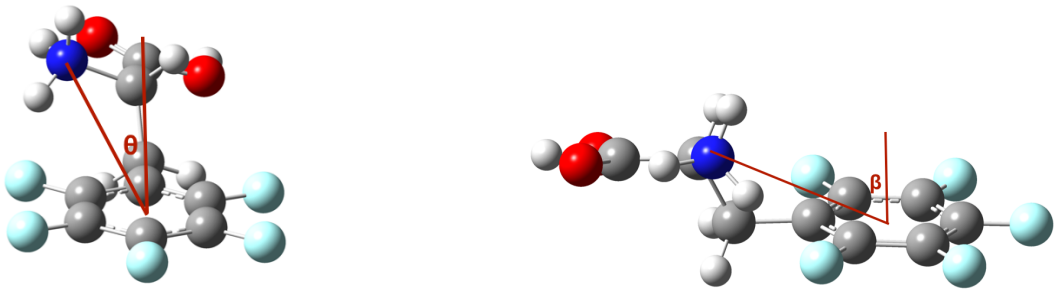


Figure 4.7 The representations of angles θ and β

In general, we find that the N–χ and H–χ distances (where χ is the ring center) are shorter

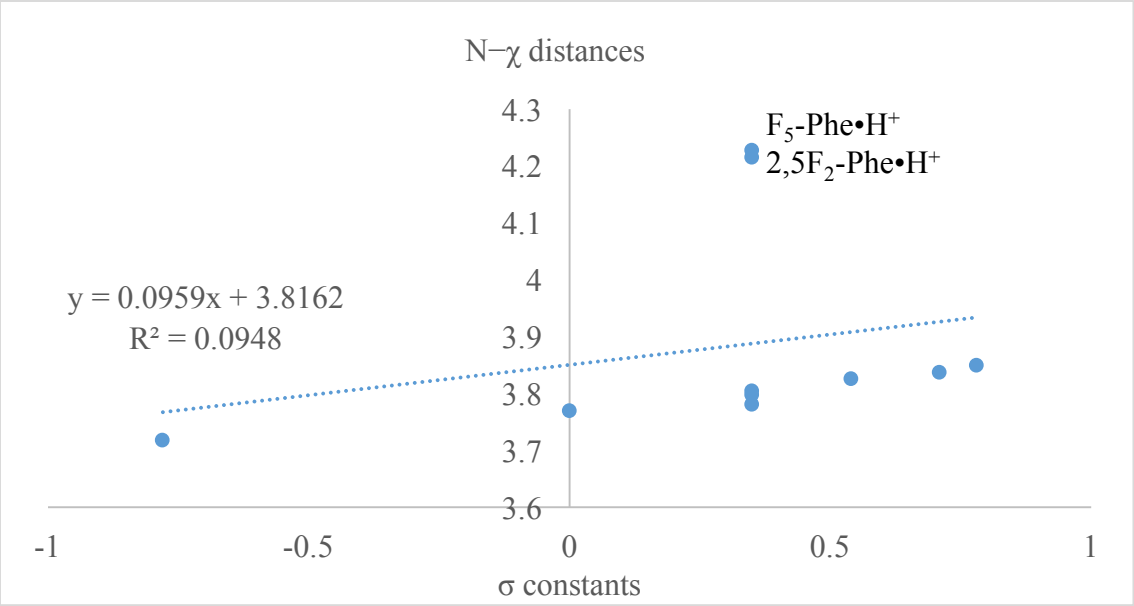
for Phe derivatives containing weak EDGs and EWGs than for those containing strong EWGs. This is consistent with expectations based on a charge-quadrupole model wherein strong EWGs are expected to reduce the ring quadrupole moment, thereby weakening the cation-π effect. A similar trend is observed for β. The N−χ and H−χ distances, and θ and β angles of 2,5F_2-Phe•H^+ are significantly larger than 3,5F_2-Phe•H^+ since the *ortho* F forms –NH_3^+•••F hydrogen bonding interactions.

Table 4.5 The angles θ and β (see Figure 4.7) and distances from the phenyl ring center (χ) to the ammonium N and closest H atom.

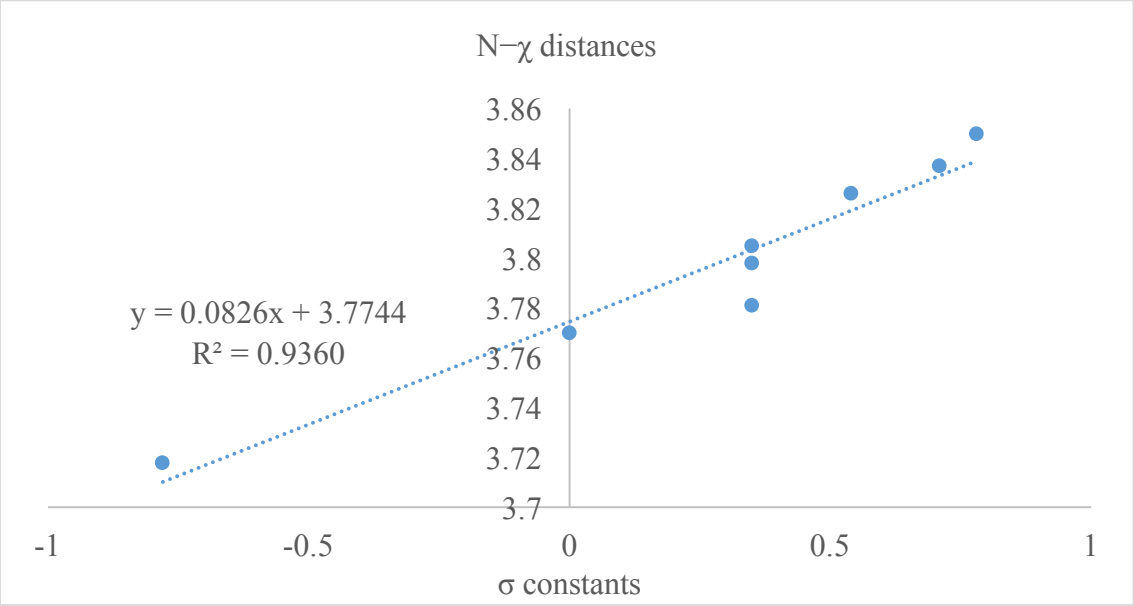
	Distance(Å)		**Angle(°)**	
	N---χ	H---χ	θ	β
3,4$(MeO)_2$-Phe	3.718	2.874	16.258	47.532
Phe	3.770	2.955	14.554	48.075
3F-Phe	3.781	2.980	14.538	48.272
4F-Phe	3.798	2.990	13.098	48.483
3,5F_2-Phe	3.805	3.020	16.093	48.744
3CF_3-Phe	3.826	3.040	17.619	48.948
4NO_2-Phe	3.837	3.063	13.973	48.789
3CN-Phe	3.850	3.071	16.695	49.488
2,5F_2-Phe	4.215	3.698	43.012	62.560
F_5-Phe	4.228	3.748	41.314	62.435

To provide an index for relative quadrupole moment, we can employ Hammett parameters. The Hammett σ constants (also called substituent constants) are indicative of linear free energy

relationships, which depend on the specific substituents.[47] Linear regression of the N–χ distance against the σ constants for the various substituents are plotted in Figure 4.8. When 2,5F_2-Phe•H^+ and F_5-Phe•H^+ are excluded, the R^2 value dramatically increases to close to unity, which demonstrates that the N–χ distances of Phe derivatives (except 2,5F_2-Phe•H^+ and F_5-Phe•H^+) exhibit strong linear correlation with the σ constants. Note that the 2,5F_2-Phe•H^+ and F_5-Phe•H^+ outliers both exhibit $–NH_3^+$•••F hydrogen bonding interactions unlike the other derivatives. Further to this, the H–χ distances, β angles, PAs, and GPBs of the Phe derivatives also exhibit strong correlations with the Hammett parameters (see Appendix II).



(1)



(2)

Figure 4.8 The linear regression between N–χ distances and σ constants for (1) all Phe derivatives; (2) Phe derivatives except 2,5F_2-Phe•H^+ and F_5-Phe•H^+.

4.4 Conclusions

The effects of substituents on the cation-π interaction were studied with a combination of quantum chemical calculations and IRMPD experiments. In order to find the global minimum, BH was used to sample the PESs of the various Phe derivatives and unique structures were further treated at the B3LYP/6-311++G(d,p) level of theory to predict the structures and IR spectra. The calculated IR spectra for the GM structures of the protonated Phe derivatives agree well with the experimental IRMPD spectra, which provides some support for the calculated GM structures being a significant population in the gas phase ensemble. However, it should be noted that oftentimes other low-energy conformers of the GM binding motif (N-protonated with IMHB) also exhibit calculated IR spectra that are a good match with experimental data.

Through comparing the structures of the global minima of the Phe derivatives, we find evidence that cation-π interactions vary with chemical substitution of EDGs and EWGs. This is exemplified by the strong correlations between the N−χ and H−χ distances and β angles with Hammett parameters for the various substituents. The PAs and GPBs were also found to correlate with Hammett parameters. Interestingly, the 2,5-F_2-Phe•H^+ and F_5-Phe•H^+ were outliers to this trend – a result which suggests that the structures of these *ortho*-substituted species are more strongly influenced by $–NH_3^+$•••F hydrogen bonding interactions than charge-quadrupole interactions.

Chapter 5

Proton-bound Dimers of Phenylalanine Derivatives

Overview

A combination of infrared multiple photon dissociation (IRMPD) and density functional theory calculations has been employed to explore the structures of proton-bound dimers of Phe, 3F-Phe, and F_5-Phe. Three structural motifs are identified: (1) canonical (charge-solvated), (2) zwitterionic (charge-separated), and (3) Phe-bridged. The bridged structures can be divided into two subclasses wherein the bridging Phe binds in a monodentate fashion or a bidentate fashion. In all three cases studied, the global minima are canonical structures whose calculated harmonic vibrational spectra accord well with the IRMPD spectra. The zwitterionic structures are unlikely in the gas phase ensemble due to the poor agreement between their calculated spectra and the IRMPD spectra. Charge-quadrupole interactions are found to influence the structures of these protonated homodimers. Interestingly, a quadrupole-quadrupole interaction is also found in the proton-bound dimer of F_5-Phe, which exhibits a Tee-shaped ring geometry.

5.1 Introduction

As established Chapter 4, charge-quadrupole interactions play an important role in determining the structures of protonated phenylalanine derivatives. To extend the exploration of the effects of the cation-π interaction on the Phe and its derivatives, the proton-bound homodimers of these species are investigated.

Naturally occurring amino acids exist as zwitterionic structures in solid state and aqueous solution within a wide pH range. However, in the gas phase, individual amino acid molecules are observed only in the canonical form.[15–18] However, the zwitterionic form of the amino acids can be stabilized by other molecules or ions in gas phase, as illustrated in many previous studies.[15–18] Therefore, it is of interest to explore the possibility that the proton-bound homodimers of Phe derivatives can exist as a zwitterionic structure in gas phase. Owing to the geometric complexity of the homodimers of Phe and its derivatives, a thorough search of the associated PESs is necessary to find the low-energy isomers. To do this, the BH search is employed (as described in Chapter 2).[21] To experimentally verify computational results, IRMPD spectra were acquired (as described in Chapter 3) for all of the proton-bound homodimers of all of the Phe derivatives described in Chapter 4. Due to the complexity of these dimers, the BH search is extremely time consuming. To date, the BH analysis and electronic structure calculations of proton-bound homodimers of Phe, 3F-Phe, and F_5-Phe have finished, and these calculated results compared with the experimental results later in this chapter. Calculations of the homodimers of the other species discussed in Chapter 4 are in process.

5.2 Methods

Please refer to Chapters 2 & 3 for a detailed description of the combined computational and experimental methods that were employed to study the proton-bound homodimers of Phe and its derivatives. Information that is pertinent to these specific studies is provided below.

5.2.1 Experimental methods

The IRMPD experiments were conducted in a similar fashion to those described previously in Chapter 4 Section 4.2.1. Approximately 100 μM solutions of 50:50 methanol:water with additional 0.1% formic acid were prepared from stoichiometric quantities of Phe, 3F-Phe, and F_5-Phe, etc. Chemicals were used without further purification.[15,16] Proton-bound homodimers of the Phe derivatives were then generated by positive mode ESI, and after these clusters were mass-selected and trapped in the Bruker Esquire 3000+ quadrupole ion trap mass spectrometer, they were irradiated with the tunable output of the CLIO FEL to generate IRMPD spectra in the 1000–2000 cm^{-1} region.

5.2.2 Computational methods

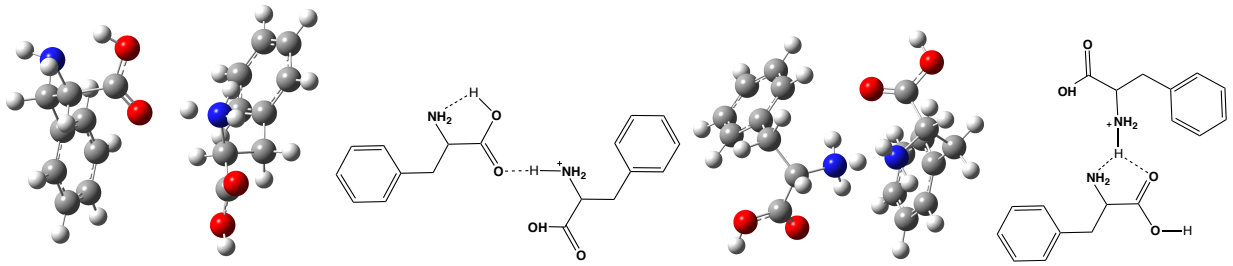
Please refer to Chapter 2 for a detailed description of computational methods. Owing to the geometric complexity of the proton-bound Phe homodimers, several degrees of freedom were included in the BH search in addition to those required for the monomers. Specially, the neutral Phe moiety was given a random translational step size of $-0.2 \leq \eta \leq 0.2$ Å, a random internal dihedral rotation of $-2° \leq \theta \leq 2°$, and a random rotation about the molecular center of mass of $-2° \leq \varphi \leq 2°$. The protonated moiety had a fixed center of mass and was given a random internal dihedral rotation of $-2° \leq \theta \leq 2°$. The low-energy isomers were identified for each proton-bound dimer, and normal mode analyses were undertaken. The harmonic vibrational spectra were predicted from the harmonic frequencies as calculated at the B3LYP/6-311++G(d,p) level of

theory.

5.3 Results and Discussion

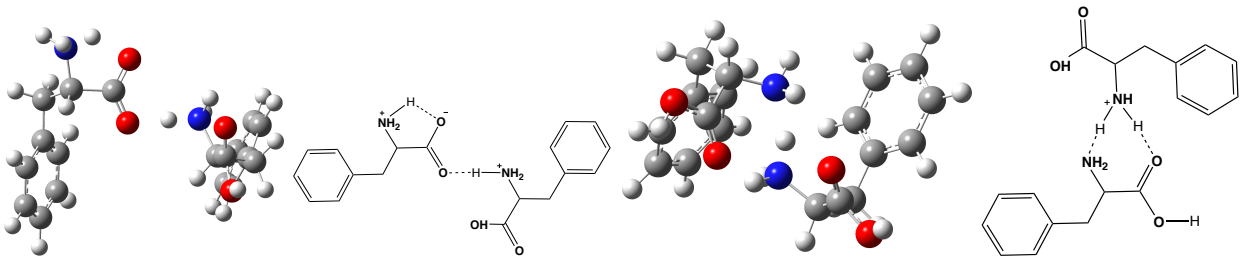
5.3.1 The Proton-bound Dimer of Phenylalanine, $(Phe)_2 \bullet H^+$

The low-energy structural motifs of the proton-bound homodimer of Phe, $(Phe)_2 \bullet H^+$, are shown in Figure 5.1. Four structural motifs were identified within 20 kJ/mol of the global minimum: (1) canonical (charge-solvated), (2) monodentate-bridged, (3) zwitterionic (charge-separated), and (4) bidentate-bridged. The structures and relative energies of all low energy isomers of $(Phe)_2 \bullet H^+$ are provided in Appendix I & III. In the monodentate-bridged structures, the Phe•H+ moiety orients so as to share one of the ammonium protons between the amine and carbonyl oxygen atoms of the neutral Phe moiety. The bidentate-bridged structure, on the other hand, shares two ammonium protons; one with the amine group and another with the carbonyl oxygen atom. As was the case with $Phe \bullet H^+$, the phenyl rings in all of the lowest energy structures of $(Phe)_2 \bullet H^+$ orient to face the protonation site, thus implying the influence of cation- π interactions.



a) Isomer 1 (Global minimum)
Canonical Structure
0 kJ/mol

b) Isomer 2
Monodentate-bridged Structure
11.7 kJ/mol



c) Isomer 4
Zwitterionic Structure
17.7 kJ/mol

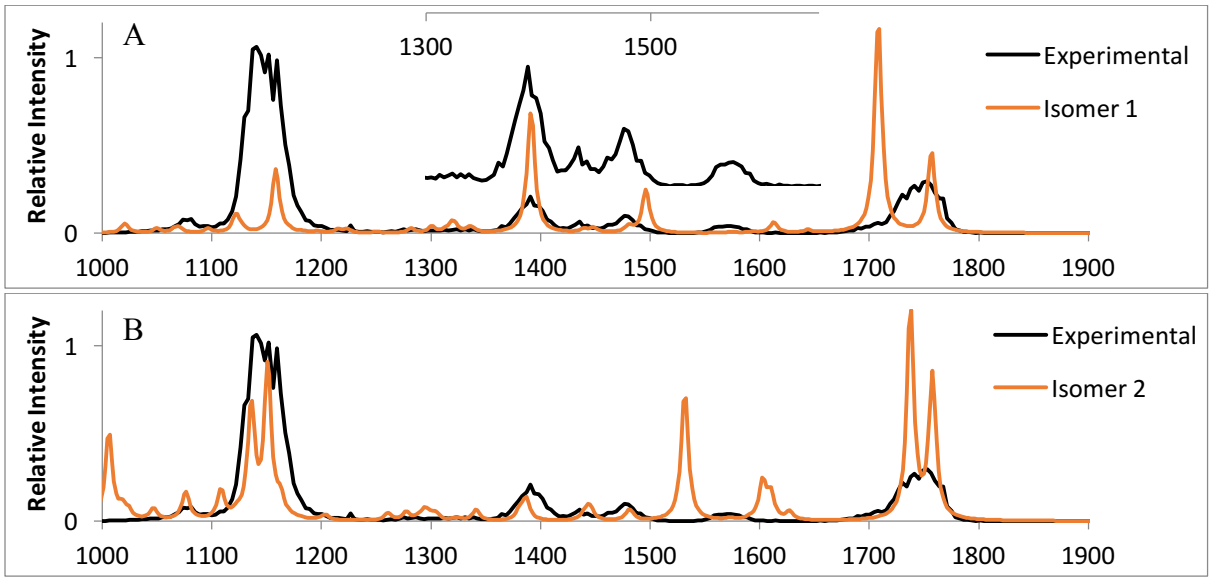
d) Isomer 5
Bidentate-bridged Structure
18.2 kJ/mol

Figure 5.1 The canonical (global minimum; isomer 1), monodentate-bridged (isomer 2), zwitterionic (isomer 4), and bidentate-bridged (isomer 5) structures of proton-bound dimer of Phe. Relative Gibbs' energies at 298 K are given in kJ/mol. Calculations used the B3LYP functional and 6-311++G(d,p) basis set.

Previous studies of Phe-containing proton-bound dimers have identified similar structural motifs.[15,16,20] However, in these studies the zwitterionic motif was found to be more stable than the bridged structure.[15,16] Here, the zwitterionic motif of $(Phe)_2 \bullet H^+$ is found to have a higher energy than the lowest-energy monodentate-bridged structure. This observation is consistent with a trend that has been observed in relative PAs (or GPBs) for the two moieties in proton-bound dimers; alkylated amines with significantly lower PAs are favourable for stabilizing gas phase zwitterionic amino acids.[17] Since the species being investigated are homodimers, there is no difference in PA between the two species. Consequently, zwitterionic structures are not favoured.

IRMPD of $(Phe)_2 \bullet H^+$ results in production of $Phe \bullet H^+$ (m/z = 166 amu). Monitoring this production channel as a function of FEL wavenumber yields the experimental IRMPD spectrum that is shown in Figure 5.2. The calculated harmonic vibrational spectra for the four lowest energy structural motifs are overlaying the experimental IRMPD spectrum in Figures 5.2A-D.

Note that the calculated harmonic vibrational spectra have been scaled by 0.9679, a commonly employed scaling factor.[46] By inspection, it is obvious that the calculated spectrum of the GM structure is a very good match with the experimental IRMPD spectrum. The calculated spectra of the bridged structures also yield a relatively good match. However, the relatively intense $-NH_3^+$ umbrella mode predicted at *ca.* 1530 cm^{-1} for the bridged species is not observed experimentally. It is clear that the zwitterionic structure (isomer 4; Figure 5.2C) does not exhibit a vibrational spectrum that matches well with experimental data. It should be noted that the widths of the observed vibrational peaks (*ca.* 70 cm^{-1}) are much broader than the 25 cm^{-1} band widths normally observed in CLIO IRMPD spectra.[35] This suggest that the observed IRMPD spectrum is a convolution of at least two closely related isomers/conformers. Given these observations, we favour assignment of the $(Phe)_2 \bullet H^+$ IRMPD spectrum to low energy conformers of canonical, charge-solvated structures. A complete assignment of all of the observed IRMPD vibrational bands is provided in Table 5.1.



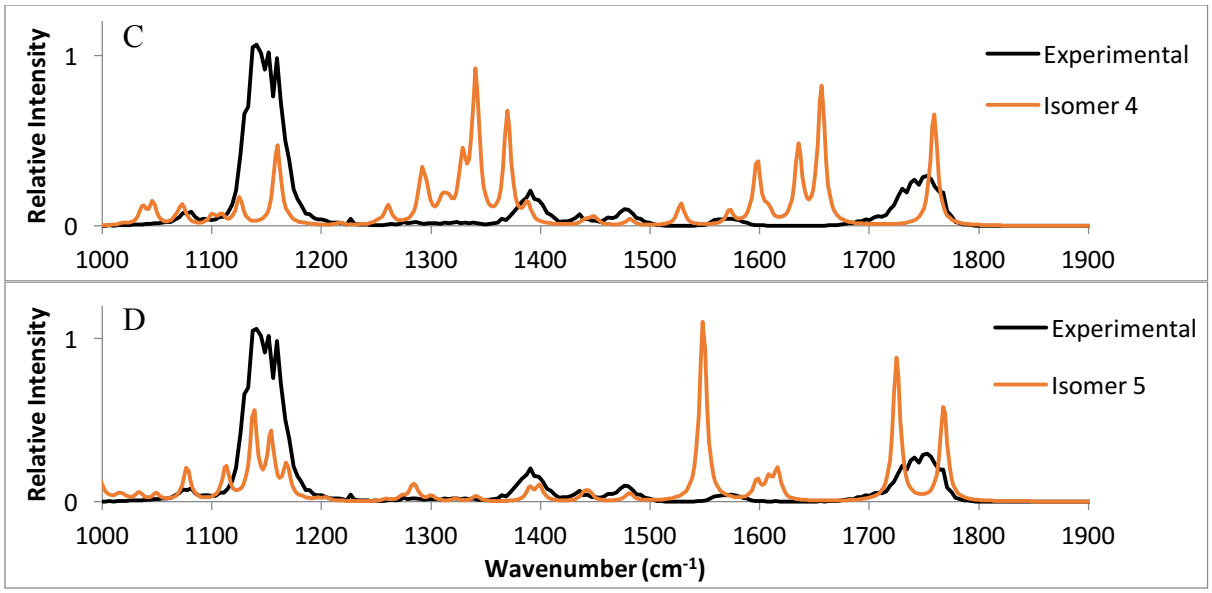


Figure 5.2 Comparison of the experimental IRMPD spectrum of $(Phe)_2 \bullet H^+$ with the calculated harmonic vibrational spectra of: (A) isomer 1, (B) isomer 2, (C) isomer 4, and (D) isomer 5. See text for details. Calculations were conducted at the B3LYP/6-311++G(d,p) level of theory and frequencies were scaled by a factor of 0.9679.[46]

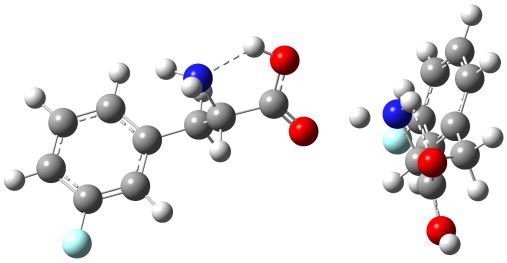
Table 5.1 Assignments of vibrational modes of proton-bound dimer of Phe, $(Phe)_2 \bullet H^+$.

	Vibrational modes (cm^{-1})			
Expt. IRMPD	**1785-1685**	**1500-1360**		**1210-1060**
isomers	C=O Stretch, NH_3 Scissoring	NH_3 Umbrella	COH Bend, CH & NH_3 Wag	COH Bend, Ring H Scissor
Isomer 1	1758, 1707	1498	1402	1158
Isomer 2	1758, 1738	1531	1387	1151
Isomer 4	1760	1340	1265	1160
Isomer 5	1767, 1725	1548	1282	1154

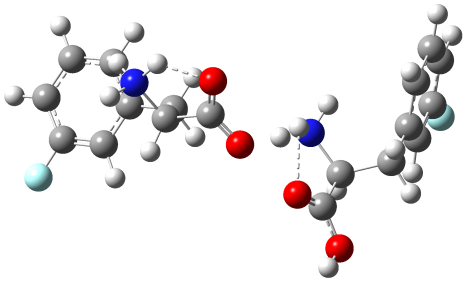
5.3.2 The Proton-bound Dimer of 3-fluorophenylalanine, $(3F\text{-}Phe)_2 \bullet H^+$

The same four structural motifs as were identified for $(Phe)_2 \bullet H^+$ were also found for

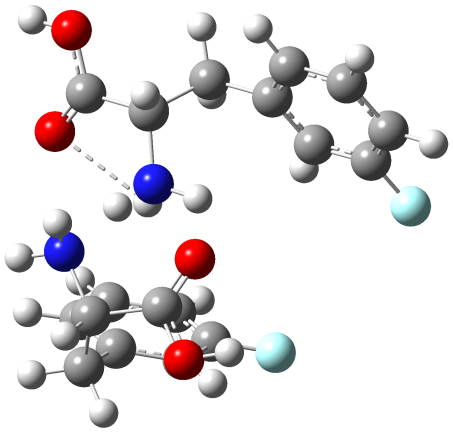
$(3F\text{-}Phe)_2 \bullet H^+$. Again, the GM was found to be a canonical structure with an IMHB contributing to structural stabilization. In this case, isomer 2 is the zwitterionic structure, and isomers 3 and isomer 12 are the first monodentate-bridged and bidentate-bridged structures, respectively (see Figure 5.3). The relative stabilization of the zwitterionic structure (by *ca.* 7 kJ/mol) in $(3F\text{-}Phe)_2 \bullet H^+$ compared to $(Phe)_2 \bullet H^+$ is unexpected and warrants further investigation. In our study of the protonated monomer species, the 2,5-difluoro and pentafluoro derivatives showed $-NH_3^+ \bullet\bullet\bullet F$ hydrogen bonding interactions (geometric restriction prevented the formation of similar hydrogen bonds in the 3F- and 4F-Phe derivatives). $(3F\text{-}Phe)_2 \bullet H^+$ also exhibits structures with $-NH_3^+ \bullet\bullet\bullet F$ and/or $-OH \bullet\bullet\bullet F$ hydrogen bonds, but at relatively high energies (shown in Appendix III). In all four low energy structural motifs, the phenyl rings appear to be oriented so as to maximize the cation-π interaction.



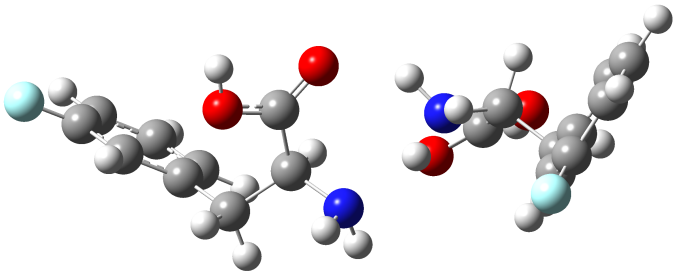
a) Isomer 1 (Global minimum)
Canonical Structure 0 kJ/mol



b) Isomer 2
Zwitterionic Structure 10.4 kJ/mol



c) Isomer 3
Monodentate-bridged Structure 14.9 kJ/mol



d) Isomer 12
Bidentate-bridged Structure 33.0 kJ/mol

Figure 5.3 The canonical (global minimum; isomer 1), zwitterionic (isomer 2), monodentate-bridged (isomer 3), bidentate-bridged (isomer 4) structures of proton-bound dimer of 3F-Phe. Relative Gibbs' energies at 298 K are given in kJ/mol. Calculations used the B3LYP functional and 6-311++G(d,p) basis set.

IRMPD of $(3F\text{-}Phe)_2 \bullet H^+$ yields $3F\text{-}Phe \bullet H^+$ (m/z = 184 amu). Monitoring this fragmentation channel as a function of FEL wavenumber generates the IRMPD spectrum shown in Figure 5.4. Also plotted in Figure 5.4 are the calculated harmonic vibrational spectra for the four lowest energy isomers that were identified via BH and subsequent DFT optimization. Based on band positions, the calculated spectrum for the GM structure yields the best agreement with experiment. As was the case with $(Phe)_2 \bullet H^+$, the observed vibrational peak widths are significantly broader than the *ca*. 25 cm^{-1} band width normally observed in CLIO FEL experiments.[35] Again, we attribute this to the presence of other, closely related structures in the probed ensemble. These are most likely to be higher energy conformers based on the same structural motif as the GM isomer. However, isomer 3 (monodentate-bridged) also exhibits a

calculated spectrum that agrees well with the experimental IRMPD spectrum. Assignments of the observed vibrational bands are provided in Table 5.2.

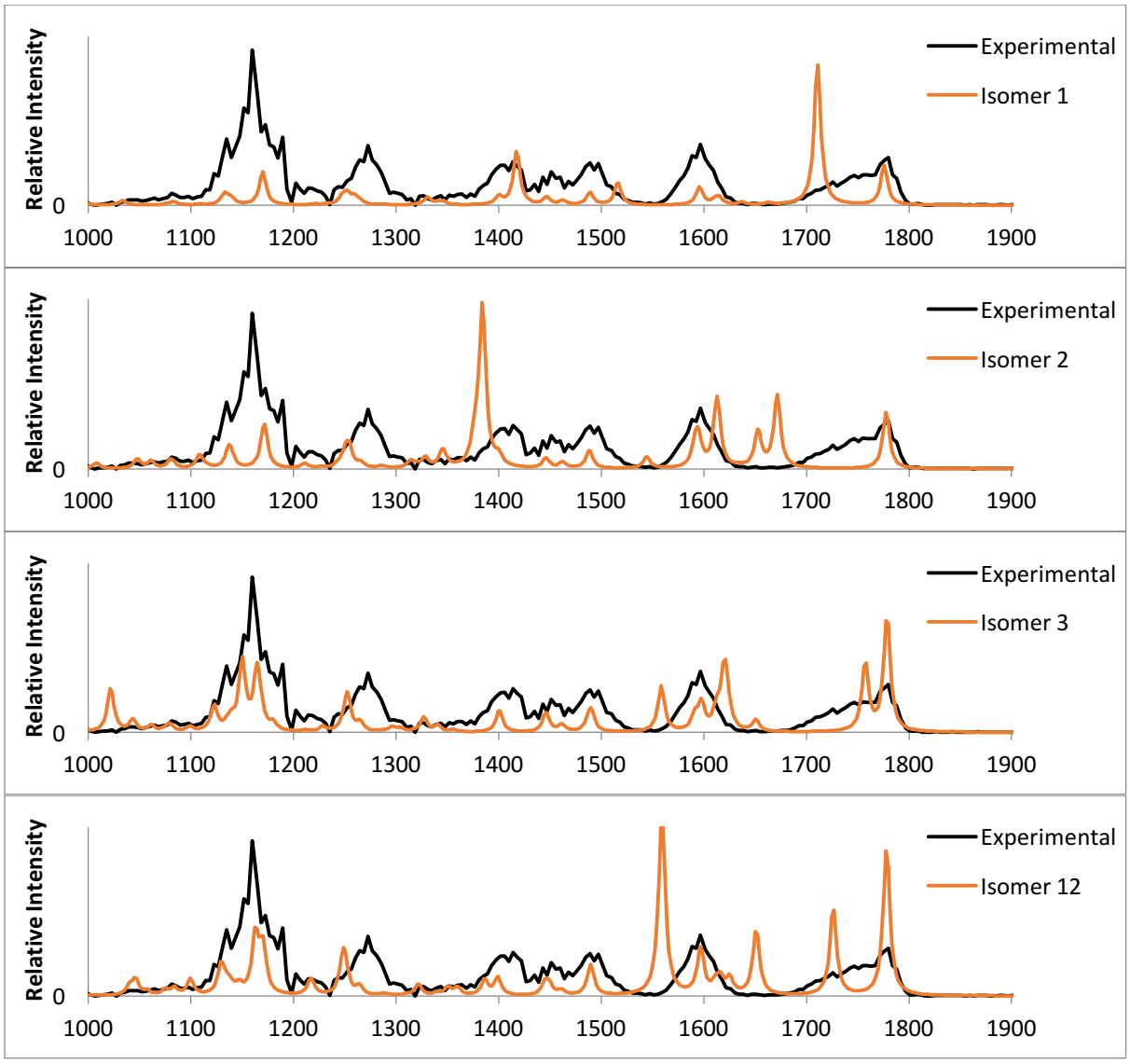


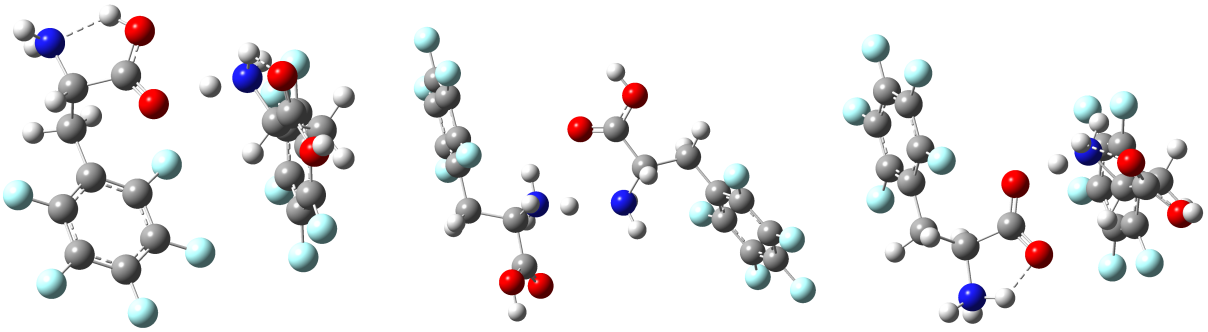
Figure 5.4 Comparison of predicted IR spectra for the four isomers (1, 2, 3, and 12) of $(3F\text{-}Phe)_2 \bullet H^+$ and the experimental IRMPD spectrum. Calculations were conducted at the B3LYP/6-311++G(d,p) level of theory and frequencies were scaled by a factor of 0.9679.[46]

Table 5.2 Assignments of vibrational modes of proton-bound dimer of 3F-Phe, $(3F\text{-}Phe)_2 \bullet H^+$.

	Vibrational modes (cm^{-1})				
Expt. IRMPD	**1800-1670**	**1540-1360**		**1310-1100**	
isomers	C=O Stretch, NH_3 Scissor	NH_3 Umbrella	COH Bend, CH & NH_3 Wag	C-F Stretch, Ring H wag	COH Bend, Ring H Scissor
Isomer 1	1776, 1711	1516	1418	1252, 1248	1170
Isomer 2	1778	1544, 1384	NA	1254, 1250	1171
Isomer 3	1779, 1757	1558	1327	1253	1164
Isomer 12	1778, 1727	1560	1399	1248	1170

5.3.3 The Proton-bound Dimer of Pentafluorophenylalanine, $(F_5\text{-}Phe)_2 \bullet H^+$

The calculated GM of $(F_5\text{-}Phe)_2 \bullet H^+$ is a canonical structure (see Figure 5.5) in which the pentafluorophenyl rings are oriented to maximize the edgewise charge-quadrupole interaction. Interestingly, the pentafluorophenyl rings are also oriented in a Tee-shaped fashion to one another, as one would expect to see for a quadrupole-quadrupole interaction. The first higher energy isomer is a bidentate-bridged structure. Unlike the $(Phe)_2 \bullet H^+$ and $(3F\text{-}Phe)_2 \bullet H^+$ analogues, our computational search did not identify any low energy monodentate-bridged structures for $(F_5\text{-}Phe)_2 \bullet H^+$. The first zwitterionic structure (isomer 5) was found to lie 15.0 kJ/mol above the global minimum.



a) Isomer 1 (Global minimum)
Canonical Structure
0 kJ/mol

b) Isomer 2
Bidentate-bridged Structure
6.4 kJ/mol

c) Isomer 5
Zwitterionic Structure
15.0 kJ/mol

Figure 5.5 The canonical (global minimum; isomer 1), lowest energy bidentate-bridged (isomer 2), and lowest energy zwitterionic (isomer 5) structures of $(F_5\text{-Phe})_2 \bullet H^+$. Relative Gibbs' energies at 298 K are given in kJ/mol. Calculations used the B3LYP functional and 6-311++G(d,p) basis set.

The proton-bound dimer of F_5-Phe produces $F_5\text{-Phe} \bullet H^+$ (m/z = 256 amu) upon IRMPD. Monitoring this product channel as a function of FEL wavenumber yields the experimental spectrum that is plotted in Figure 5.6. Overlaying the experimental spectrum in Figure 5.6 are the calculated IR spectra of canonical and bridged structures, both of which agree well with the experimental spectrum. However, the spectrum of the bridged species (isomer 2) does exhibit a relatively intense peak at *ca.* 1620 cm^{-1} ($-NH_3^+$ scissoring) which is not observed in the experimental spectrum. Consequently, we favour assignment of the experimental spectrum to the global minimum canonical structure. A complete assignment of the observed vibrational bands is provided in Table 5.3.

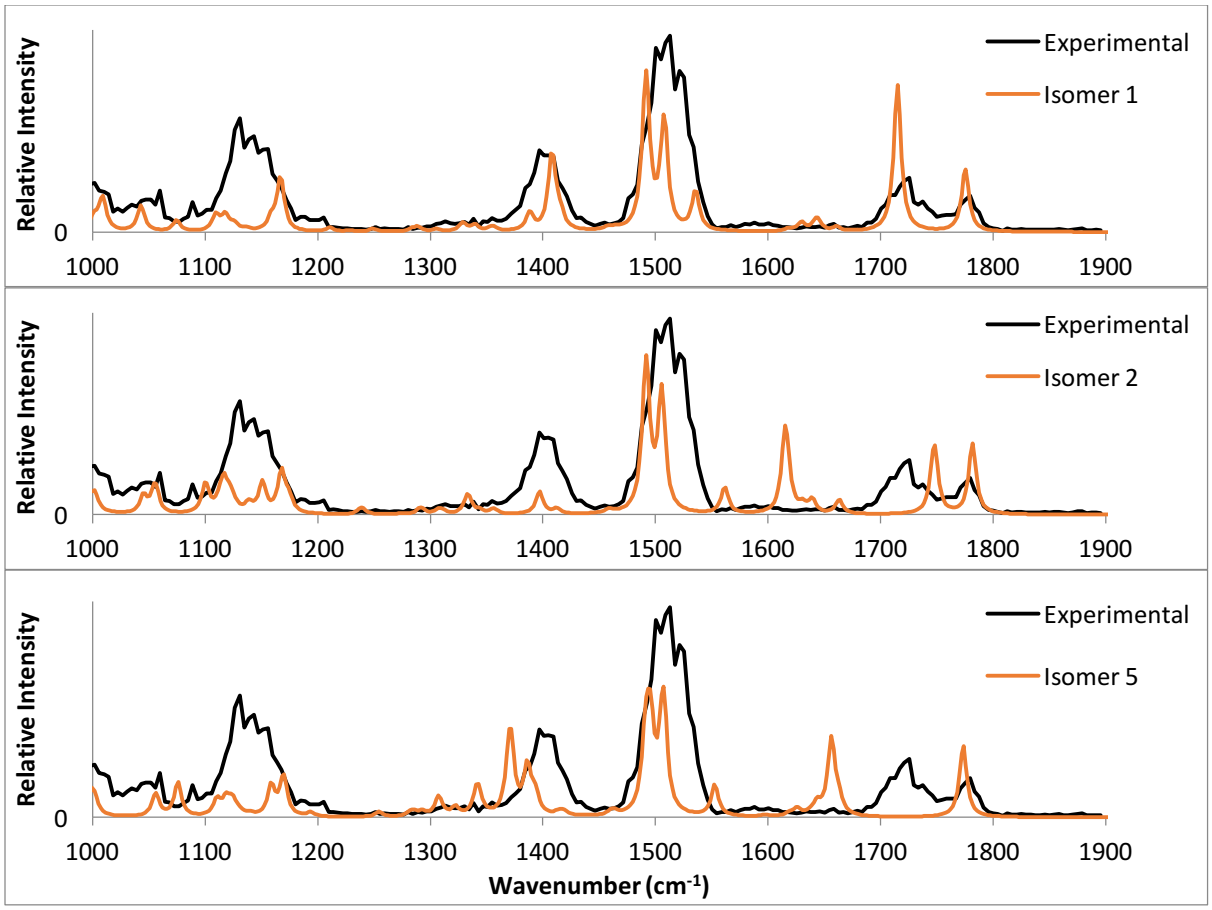


Figure 5.6 Comparison of the experimental IRMPD spectrum of $(F_5\text{-Phe})_2 \bullet H^+$ with the calculated harmonic vibrational spectra of isomer 1, isomer 2, and isomer 5. See text for details. Calculations were conducted at the B3LYP/6-311++G(d,p) level of theory and frequencies were scaled by a factor of 0.9679.[46]

Table 5.3 Assignments of vibrational modes of proton-bound dimer of F_5-Phe, $(F_5\text{-Phe})_2 \bullet H^+$.

	Vibrational modes (cm^{-1})				
Expt. IRMPD	**1800-1680**	**1560-1300**		**1210-1000**	
isomers	C=O Stretch, NH_3 Scissor	NH_3 Umbrella	COH Bend, CH & NH_3 Wag	COH Bend, Ring H Scissor	C-F Stretch
Isomer 1	1776, 1715	1536	1408	1167	1118
Isomer 2	1782, 1748	1562	1333	1168	1117
Isomer 5	1773	1553, 1371	1344	1170	1118

5.4 Conclusions

A thorough search of the potential energy landscapes of $(Phe)_2•H^+$ and $(3F\text{-}Phe)_2•H^+$ identified four structural motifs: (1) canonical (charge-solvated), (2) zwitterionic (charge-separated), (3) monodentate-bridged, and (4) bidentate-bridged. The same structural motifs, with the exception of the monodentate-bridged structure, were identified for $(F_5\text{-}Phe)_2•H^+$. In all cases, the global minima were canonical structures whose calculated IR spectra agreed well with those recorded experimentally via IRMPD. The presence of zwitterionic structures in the gas phase ensemble could be clearly discounted based on the poor match between their predicted spectra and the experimental IRMPD spectra. Bridged structures were discounted due to the strong calculated $-NH_3^+$ umbrella and scissoring motions which were not observed experimentally.

In all three cases studied, the cation-π interaction is found to influence the structure of the protonated homodimers. In $(Phe)_2•H^+$ and $(3F\text{-}Phe)_2•H^+$, the phenyl rings were found to orient so as to face the site of protonation, while in $(F_5\text{-}Phe)_2•H^+$, the inverted ring quadrupole moment results in an edgewise interaction. Calculations also suggest the presence of a quadrupole-quadrupole interaction between the pentafluorophenyl rings, which manifests as a Tee-shaped ring geometry whereby the face of one ring interacts with the edge of the other. The evolution of the geometries of Phe derivative homodimers as a function of EDG and EWG substituents is an open question that is currently under investigation. IRMPD spectra have been acquired for the homodimers of all of the species that were discussed in Chapter 4 and electronic

structure calculations are currently underway.

Chapter 6

Protonated Phenylalanine Dipeptide, (PhePhe)•H^+

Overview

A combined experimental IRMPD and computational study of Phe dipeptide (PhePhe) is conducted to determine the gas phase structure. Based on agreement between the calculated and experimental spectra, the nitrogen atom of the primary amine group is identified as the most favourable protonation site. Other PhePhe tautomers do not yield calculated spectra that match well with experiment. Furthermore, it is found that the presence (or absence) of intramolecular hydrogen bonds (IMHBs) significantly affects the IR spectra. The existence of structures exhibiting IMHBs can be identified through comparing the predicted and experimental spectra.

6.1 Introduction

To continue in our efforts to study the effects of non-covalent interactions on the structures of Phe-containing systems, the Phe dipeptide (PhePhe) was investigated. The covalent amide bond that connects the two Phe residues in the dipeptide is expected to impose additional geometric constraints on the molecular structure (compared to the proton-bound homodimer of Phe), and it is unclear as to if this will disrupt potential charge-quadrupole interactions. Furthermore, the site of protonation for (PhePhe)•H^+ and whether or not the proton is shared (de-localized) or localized to a single basic site is not obvious upon inspection.[48] Hence, a

thorough search of the PES and spectroscopic characterization of the molecule is warranted.

6.2 Methods

Please refer to Chapters 2 & 3 for a detailed description of the combined computational and experimental methods that were employed to study the Phe dipeptide. Information that is pertinent to these specific studies is provided below.

6.2.1 Experimental methods

The procedure of the IRMPD experiments has been described previously in Section 4.2.1. Approximately 100 μM solutions of 50:50 methanol:water with additional 0.1% formic acid were prepared from stoichiometric quantities of PhePhe. Chemicals were used without further purification.[15,16] (PhePhe)•H^+ is generated by positive mode ESI, then mass-selected and trapped in the Bruker Esquire 3000+ ion trap mass spectrometer. The FEL utilized electron beam energies of 46 MeV, and scanned over a wavenumber range of 1000−2000 cm^{-1}.

6.2.2 Computational methods

The computational routine, as described in Chapter 2, was also used to study (PhePhe)•H^+. Since PhePhe has four possible protonation sites (two amine nitrogen atoms and two carbonyl oxygen atoms), the four variants have to first be studied to determine the site of protonation. The (PhePhe)•H^+ tautomers were pre-optimized and modeled using MM/UFF, then partial charges were calculated by using CHelpG at DFT B3LYP/6-31G level of theory. Each (PhePhe)•H^+

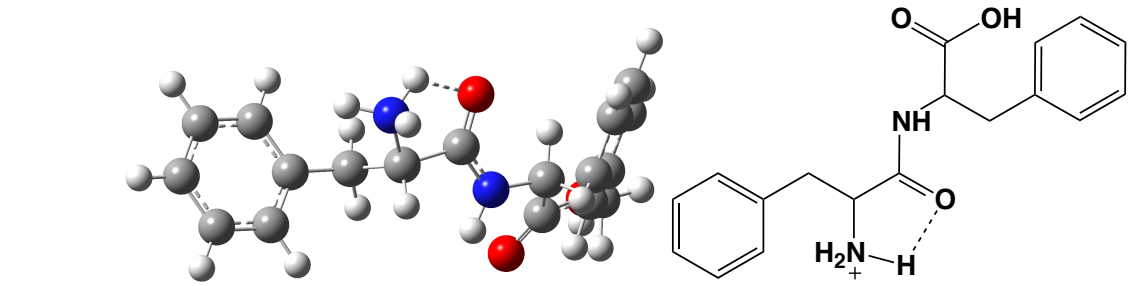
tautomer was then subjected to the BH algorithm wherein an internal dihedral rotation was randomly chosen from the range $-5° \leq \varphi \leq 5°$ for 10,000 BH steps. The unique isomers were identified and the low-energy species were re-optimized, first by PM6, then at the B3LYP/6-311++G(d,p) level of theory. Normal mode analyses were conducted to ensure that the stationary points were true minima on the PES, and harmonic vibrational spectra were predicted from the calculated frequencies.

6.3 Results and Discussion

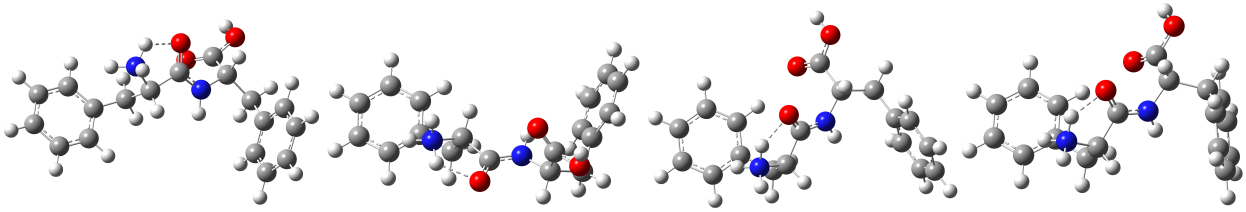
The calculated global minimum and five low-energy isomers (all exhibiting protonation on the amine group) are shown in Figure 6.1. As was found for Phe•H^+, the most favourable site for protonation was found to be the primary amine nitrogen. Note that in each of these structures, the protonation site is stabilized via IMHB formation with the amide carbonyl oxygen atom. The geometries of the low energy species also show a facial alignment of (at least one of) the phenyl rings with the charge site, thus indicating that cation-π interactions influence the overall structure of the protonated Phe dipeptide. In the case of the GM structure, both phenyl rings are oriented to maximize charge-quadrupole interactions with the charge center.

The relatively low energies of the first five isomers suggests that all of these species might be present in the gas phase ensemble. Isomer 33, on the other hand, can likely be discounted owing to its relatively high Gibbs' energy. The step-change in relative energy between isomers 5 and 55 arises due to a subtle change in geometry; although isomer 33 has the same protonation site as the lower energy species, the ammonium group instead forms an IMHB with the carbonyl

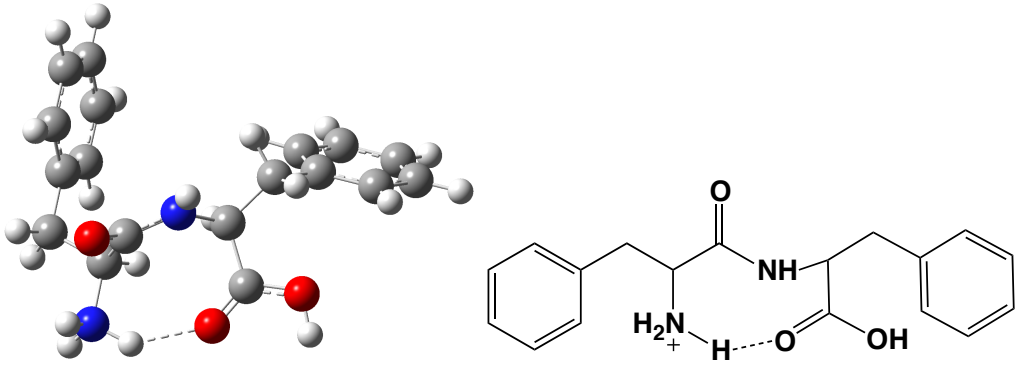
oxygen atom of the carboxylic acid. To determine which (if any) of these low energy structures are present in the gas phase ensemble, IRMPD experiments were conducted.



Isomer 1 0 kJ/mol



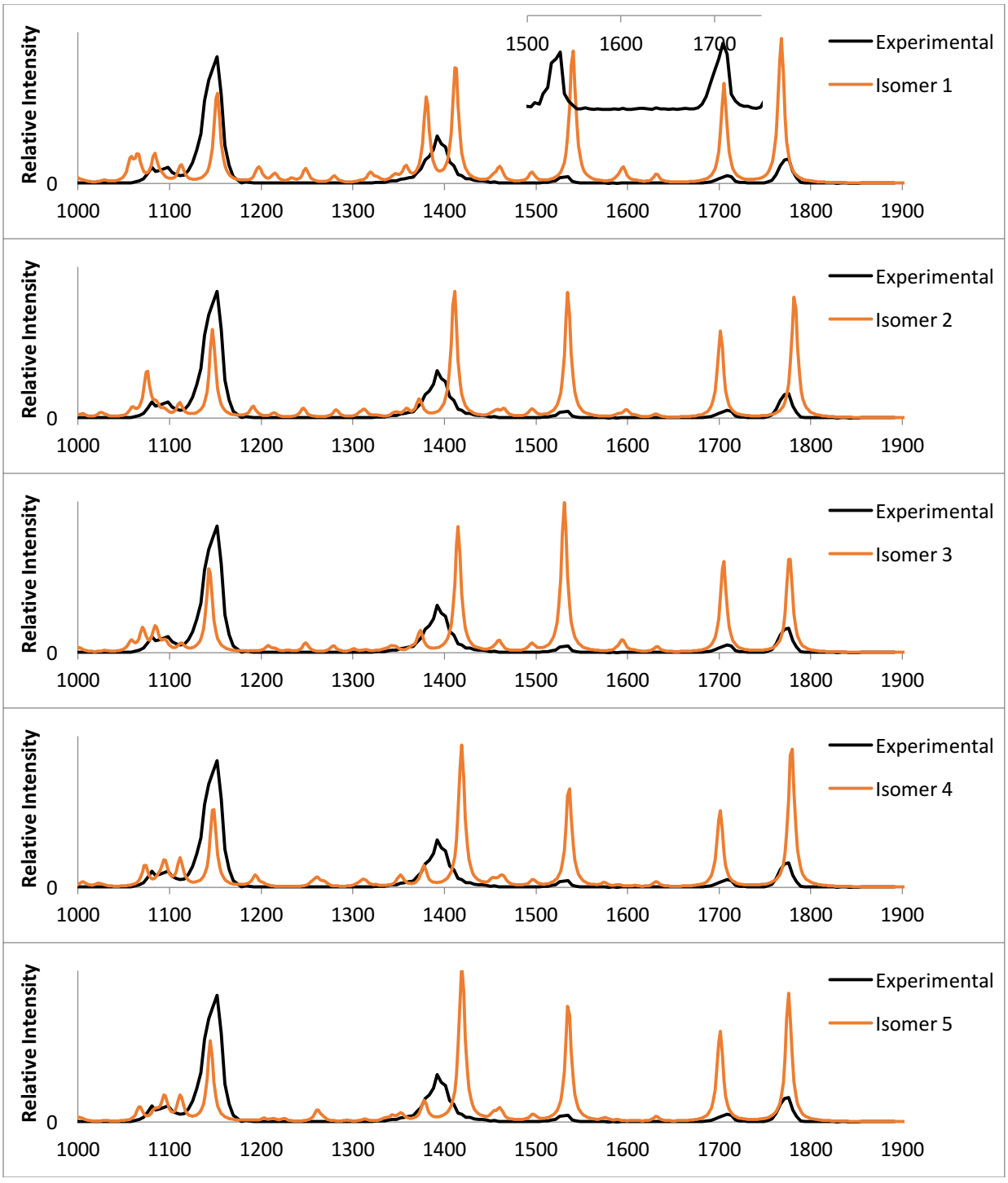
Isomer 2	Isomer 3	Isomer 4	Isomer 5
0.003	1.0	3.7	4.6



Isomer 33 30.9 kJ/mol

Figure 6.1 The isomers of Phenylalanine dipeptide. Relative Gibbs' energies at 298 K are given in kJ/mol. Calculations used the B3LYP functional and 6-311++G(d,p) basis set.

IRMPD of (PhePhe)•H^+ produces neutral $C_8H_{10}N$ (m/z = 120 amu) and $C_{10}H_{11}O_3N^+$ (m/z = 193 amu). Figure 6.2 plots the IRMPD spectrum that was recorded by monitoring this product channel as a function of FEL wavenumber. Also plotted in Figure 6.2 are the calculated harmonic vibrational spectra for the six isomers shown in Figure 6.1. The calculated spectra of isomers 1 – 5 (which are all conformers of the same binding motif) all agree well with the experimental spectrum. Given the relative Gibbs' energies of these species, it is likely that all five are present in the gas phase ensemble. The calculated spectrum of isomer 33, on the other hand, is a poor match to the experimental spectrum. Given this poor match and the relatively high energy of isomer 33, we can conclude that it is not present in the gas phase ensemble. A complete assignment of the observed vibrational bands is provided in Table 6.1.


Relative Intensity
0
1000
1100
1200
1300
1400
1500
1600
1700
1800
1900
1500
1600
1700
Experimental
Isomer 1
Experimental
Isomer 2
Experimental
Isomer 3
Experimental
Isomer 4
Experimental
Isomer 5

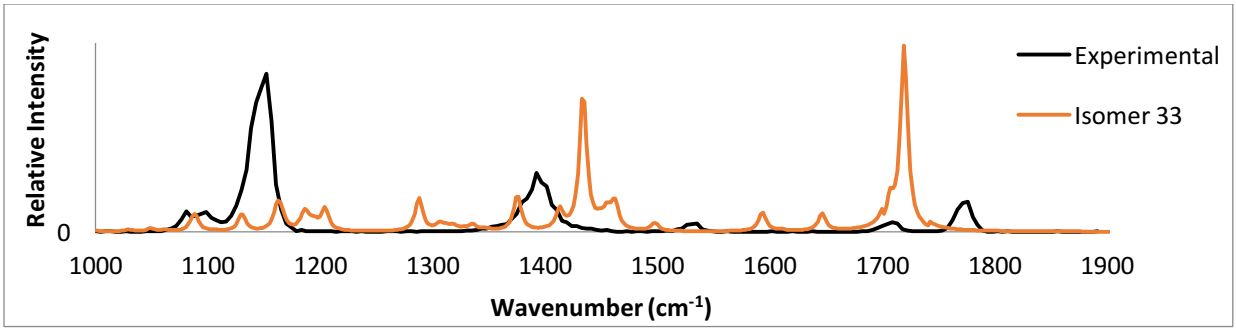


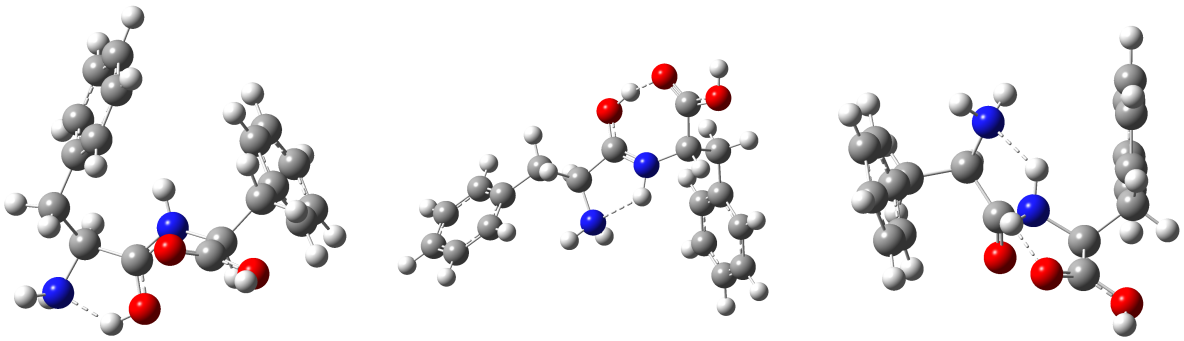
Figure 6.2 Comparison of calculated IR spectra for the six isomers of (PhePhe)•H^+ and the experimental IRMPD spectrum. Calculations were conducted at the B3LYP/6-311++G(d,p) level of theory and frequencies were scaled by a factor of 0.9679.[46]

Table 6.1 Assignments of vibrational modes of protonated Phe dipeptide, (PhePhe)•H^+

	Vibrational modes (cm^{-1})						
IRMPD	**1790-1760**	**1720-1690**	**1550-1500**		**1460-1330**		**1180-1050**
Isomers	C=O Stretch	C=O Stretch	NH Wag	CH_2 Scissor	NH_3 Umbrella	CH Wag	COH Bend
Isomer 1	1768	1705	1540	1461	1412	1380	1152
Isomer 2	1782	1701	1535	1465	1411	1373	1147
Isomer 3	1776	1705	1531	1460	1415	1374	1143
Isomer 4	1779	1701	1537	1461	1419	1377	1148
Isomer 5	1776	1701	1535	1462	1419	1378	1145
Isomer 33	1719	1707	1449	1462	1434	1413	1204

In addition to the amine-protonated species, low energy oxygen-protonated and amide-nitrogen-protonated tautomers were also identified by our computational treatment. Examples of the oxygen protonated and amide-nitrogen-protonated species are shown in Figure 6.3. Isomer 10 is lowest energy protonated on the amide carbonyl oxygen and exhibits an N•••H^+–O IMHB to the amine nitrogen atom. Isomer 16 is also protonated at amide carbonyl oxygen, but instead exhibits two IMHBs; one N•••H–N and one O•••H^+–O interaction. The lowest energy amide-nitrogen-protonated structure, isomer 40, exhibits one N•••H^+–N and one

O•••H^+–N IMHB. The calculated IR spectra for these three species are compared with the experimental IRMPD spectrum in Figure 6.4. The calculated spectra of isomers 16 and 40 are clearly a poor match to the experimental data, but the calculated spectrum for isomer 10 does compare relatively well. This is not necessarily surprising since isomer 10 has a structure that is very similar to the lower energy amine-protonated species. The vibrational wavenumber of protonated carbonyl stretch in isomer 10, which one might expect to shift dramatically when compared to the unprotonated carbonyl, remains relatively unshifted due to the electron density provided by the amine group via the IMHB. Consequently, we are unable to discount isomer 10 from the gas phase (PhePhe)•H^+ ensemble.



Isomer 10	Isomer 16	Isomer 40
11.7	18.4	67.8

Figure 6.3 The lowest energy structure that is protonated on the amide carbonyl oxygen and has an N•••H^+–O IMHB (isomer 10), the lowest energy structure that is protonated on the amide carbonyl oxygen and has an N•••H–N and an O•••H^+–O IMHB (isomer 16), and the lowest energy amide-nitrogen-protonated structure (isomer 40). Relative Gibbs' energies at 298 K are given in kJ/mol. Calculations used the B3LYP functional and 6-311++G(d,p) basis set.

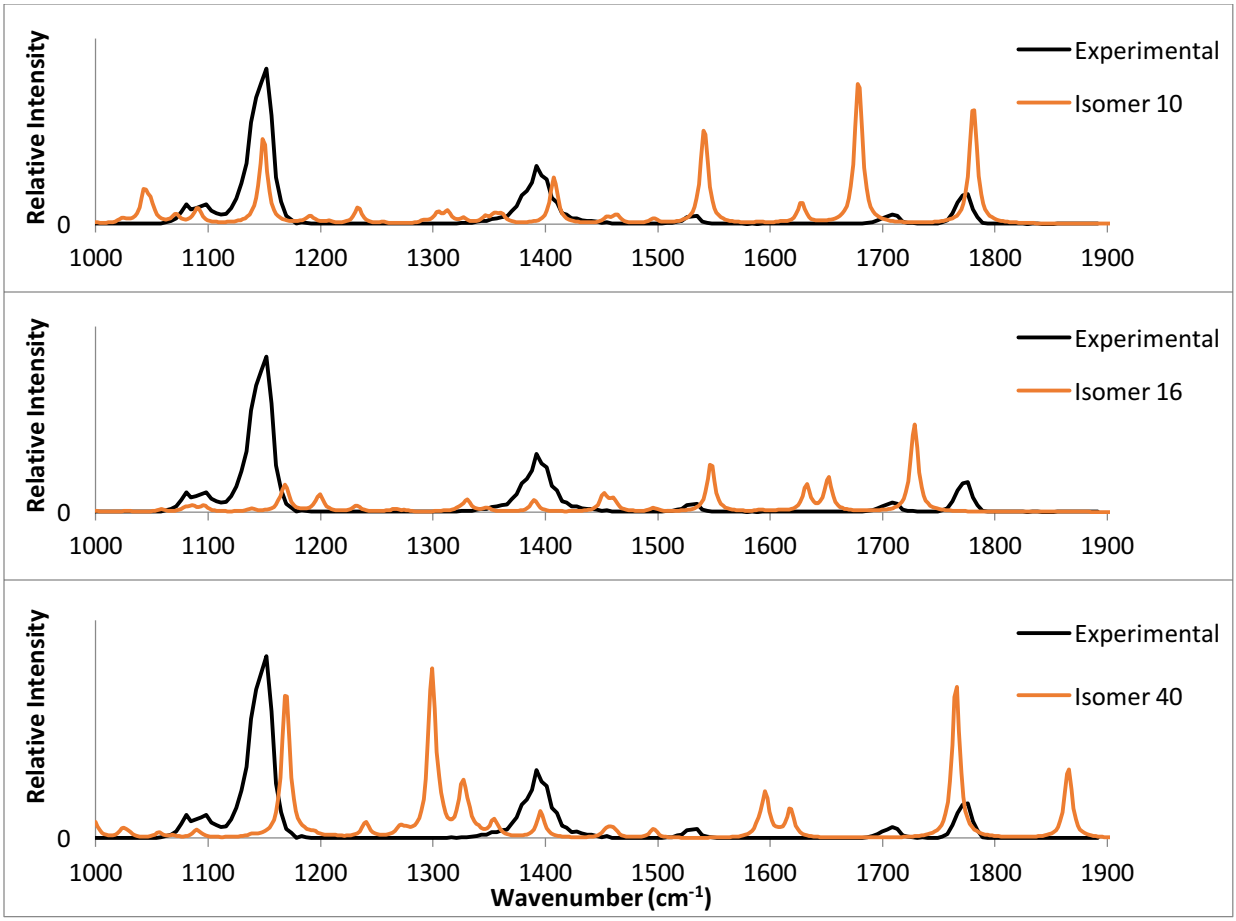


Figure 6.4 Comparison of calculated IR spectra for the isomer 10, isomer 16, and isomer 40 of (PhePhe)•H^+ and the experimental IRMPD spectrum. Calculations were conducted at the B3LYP/6-311++G(d,p) level of theory and frequencies were scaled by a factor of 0.9679.[46]

6.4 Conclusions

An exhaustive search of the (PhePhe)•H^+ PES indicates that the conformers which lie within *ca*. 10 kJ/mol of the global minimum are all protonated on the amine N atom, and that these low energy structures all exhibit IMHBs between the ammonium group and the amide carbonyl O atom. The experimental IRMPD spectrum is consistent with the calculated harmonic vibrational spectra for the five lowest energy conformers of this structural motif. For the most part, the calculated IR spectra of other (PhePhe)•H^+ tautomers and species which exhibit

different IMHB binding schemes do not match with experimental results. However, there is one low energy isomer (isomer 10), similar to amine-protonated isomers, but with protonation on the amide carbonyl O atom, which does produce a calculated spectrum that closely matches experiment, reminiscent of isomer 1 following proton-transfer to C=O. Consequently, isomer 10 cannot be discounted as a member of the probed ensemble, but given its relative Gibbs' energy of 11.7 kJ/mol it is likely to account for only 0.9 % of the total population at a temperature of 298 K. Given the ambiguity as to which low energy isomers contribute to the observed spectrum, it would be beneficial to study the (PhePhe)•H^+ system in the 2750 – 3750 cm^{-1} spectral region where there are more diagnostic CH, NH, and OH stretching modes.

All of the (PhePhe)•H^+ isomers that were identified as potential spectral carriers exhibit geometries wherein at least one of the phenyl rings orient so as to face the site of protonation and maximize non-covalent cation-π interactions. Extrapolating to larger peptide/protein systems which contain aromatic amino acid residues, these results suggest that pH-induced conformational changes might, at least in part, arise from intramolecular non-covalent charge-quadrupole interactions. It would therefore be interesting to continue this direction of research with larger chemically substituted aromatic peptides.

Chapter 7

Concluding Remarks

The secondary structures of large biomolecules are influenced by non-covalent interactions.[1] In the case of protein folding and molecular recognition, π-π stacking and cation-π interactions play a prominent role.[1] Here, we have begun a program of research that is designed to investigate cation-π interactions in systems that contain Phe. The cation-π interaction in the Phe derivatives described herein arises from the interaction of the quadrupole charge distribution of the phenyl group with a positively charged species. By varying substituents around the phenyl ring one can potentially tune the cation-π interaction, even going so far as to invert the ring's quadrupole moment, thus changing cation binding motifs and molecular/cluster structure.

To explore the effects of electron donating/withdrawing groups on the non-covalent interactions of Phe derivatives, protonated Phe derivatives, homodimers, and Phe dipeptide were investigated in a combined experimental and computational approach. The low-lying isomers of various species were identified using the Basin-hopping (BH) search algorithm,[21,22] and density functional theory (DFT) was used to calculate molecular geometries and properties. Calculated harmonic vibrational spectra were then compared with experimental spectra obtained via IRMPD to determine cluster geometries. In all cases, protonation was found to occur on the primary amine N atom, which then participated in intramolecular hydrogen bonding with the (closest) carbonyl O atom.

The calculated structures of Phe•H^+ and its derivatives exhibited geometric variations as a function of EWG/EDG substitution that were consistent with expectations based on non-covalent cation-π interactions. The phenyl ring of unsubstituted Phe oriented its face towards the site of protonation, whereas the pentafluorophenyl derivative interacted in an edgewise fashion. By plotting geometric changes as a function of Hammett parameters for the EWGs/EDGs, we were able to visualize the influence of the cation-π interaction on molecular geometry. However, 2,5F_2-Phe•H^+ and F_5-Phe•H^+ were outliers to the observed trends owing to apparent -NH_3^+•••F IMHB interactions with substituents in the *ortho* position (relative to the amino acid moiety). This electron density donation via the IMHB might also explain the fact that the proton affinity (PA) and gas-phase basicity (GPB) of 2,5F_2-Phe are significantly higher than those of 3,5F_2-Phe.

Studies of the proton-bound dimers of Phe, 3F-Phe and F_5-Phe have identified four low energy structural motifs: (1) canonical (charge-solvated), (2) zwitterionic (charge-separated), (3) monodentate-bridged and (4) bidentate-bridged. In all three cases, global minima were canonical structures. In $(Phe)_2$•H^+ and $(3F\text{-}Phe)_2$•H^+, the cation-π interaction causes the face of the phenyl rings to orient towards the site of protonation. In $(F_5\text{-}Phe)_2$•H^+, it is the edge of the phenyl rings that interacts with the protonation site. Interestingly, an additional non-covalent quadrupole-quadrupole interaction between the pentafluorophenyl rings in $(F_5\text{-}Phe)_2$•H^+ results in a local Tee-shaped geometry. The evolution of this structure as a function of EDG/EWG substitution for the remaining homodimers of the Phe derivatives characterized in Chapter 4 is currently under investigation. IRMPD spectra have been acquired for all of these species and

DFT calculations are currently running.

In our efforts to extend our studies of cation-π interactions to a more biologically-relevant system, protonated Phe dipeptide was also investigated. Our combined experimental and computational approach show that protonation of PhePhe occurs on the primary amine N atom and that an IMHB is formed between the ammonium N atom and the O atom of the amide carbonyl. As was the case with the protonated monomers and the proton-bound dimers, the global minimum of (PhePhe)•H^+ exhibits a geometry in which the phenyl rings orient so as to maximize cation-π interactions within the geometric restricts imposed by covalent bonding. This result has important bearing on our understanding of biological systems since it implies that pH-induced conformational changes in proteins and peptides might, at least in part, be driven by non-covalent charge-quadrupole interactions. The fact that a similar pH-driven intermolecular cation-π interaction is known to occur between nicotine and Trp residues in the nicotinic acetylcholine receptor provides some support for this hypothesis.[49,50] The research that is presented in this thesis only scratches the surface of the topic of cation-π interactions in biological systems and, ultimately, an extended program of research which includes studies of larger Phe polypeptides and Trp and Tyr derivatives is warranted for a more thorough exploration of this fascinating subject.

References

1. Lodish, H. *et al.* Molecular Cell Biology. 4th edition. New York: W. H. Freeman. Section 2.2, Noncovalent Bonds. (2000).
2. Dougherty, D. A. Cation-π Interactions Involving Aromatic Amino Acids. *J. Nutr.* **137,** 1504S–1508S (2007).
3. Dougherty, D. A. *et al.* Cation-π interactions in chemistry and biology: a new view of benzene, Phe, Tyr, and Trp. *Science* **271,** 163–8 (1996).
4. Ma, J. C. & Dougherty, D. a. The Cation-π Interaction. *Chem. Rev.* **97,** 1303–1324 (1997).
5. Zacharias, N. & Dougherty, D. A. Cation–π interactions in ligand recognition and catalysis. *Trends Pharmacol. Sci.* **23,** 281–287 (2002).
6. 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).
7. Sunner, J., Nishizawa, K. & Kebarle, P. Ion-solvent molecule interactions in the gas phase. The potassium ion and benzene. *J. Phys. Chem.* **85,** 1814–1820 (1981).
8. Guo, B. C., Purnell, J. W. & Castleman, A. W. The clustering reactions of benzene with sodium and lead ions. *Chem. Phys. Lett.* **168,** 155–160 (1990).
9. Tsuzuki, S., Yoshida, M., Uchimaru, T. & Mikami, M. The Origin of the Cation/π Interaction: The Significant Importance of the Induction in Li + and Na + Complexes. doi:10.1021/jp003287v
10. Wu, R. & Mcmahon, T. B. Investigation of Cation-π Interactions in Biological Systems. *J. Am. Chem. Soc.* 12554–12555 (2008). doi:10.1021/ja802117s
11. Quiñonero, D. *et al.* Anion-π interactions: Do they exist? *Angew. Chemie - Int. Ed.* **41,** 3389–3392 (2002).
12. Burley, S. K. & Petsko, G. A. Amino-aromatic interactions in proteins. *FEBS Lett.* **203,** 139–143 (1986).
13. Singh, J. & Thornton, J. M. SIRIUS. An automated method for the analysis of the preferred packing arrangements between protein groups. *J. Mol. Biol.* **211,** 595–615 (1990).
14. Voet, D. & Voet, J. G. *Biochemistry*. *4th Edition. Chapter 4, 67-81* (John Wiley & Sons, 2011).
15. 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).

16. 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).

17. 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).

18. 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).

19. Julian, R. R., Beauchamp, J. L. & Goddard, W. A. Cooperative salt bridge stabilization of gas-phase zwitterions in neutral arginine clusters. *J. Phys. Chem. A* **106,** 32–34 (2002).

20. Fu, W. *et al.* The Structures of Proton-bound Dimers of Glycine with Phenylalanine and Pentafluorophenylalanine. *J. Mol. Spectrosc.* 2–7 (2016). doi:10.1016/j.jms.2016.07.004

21. Wales, D. & 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).

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

23. 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.

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

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

26. 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).

27. 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).

28. 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).

29. Kayi, H. Parameterization of the AM1* semiempirical molecular orbital method for the first-row transition metals and other elements. Friedrich-Alexander University. (2009).

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

31. Gruene, P. *et al.* Structures of neutral Au7, Au19, and Au20 clusters in the gas phase. *Science* **321,** 674–6 (2008).

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

33. 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).

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

35. 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).

36. Hopkins, W. S. Atomic & Molecular Clusters. *Chem 450 course notes,* University of Waterloo, Waterloo, Canada (2014).

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

38. 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).

39. Centre Laser Infrarouge d'Orsay (n.d.). Retrieved from: http://clio.lcp.u-psud.fr/clio_eng/accel.html 2014 Nov.

40. March, R. E. SPECIAL FEATURE : An Introduction to Quadrupole Ion Trap Mass Spectrometry. *J. Mass Spectrom.* **32,** 351–369 (1997).

41. Arnaud Droit, Universite Laval, 2007. Retrieved from: http://theses.ulaval.ca/archimede/fichiers/24220/ch02.html.

42. Bania, K. K., Guha, A. K., Bhattacharyya, P. K. & Sinha, S. Effect of substituent and solvent on cation-π interactions in benzene and borazine: a computational study. *Dalton Trans.* **43,** 1769–84 (2014).

43. Fuster, F. & Silvi, B. Determination of protonation sites in bases from topological rules. *Chem. Phys.* **252,** 279–287 (2000).

44. Hunter, E. P. & Lias, S. G. Evaluated gas-phase basicities and proton affinities of molecules: An update. *J. Phys. Chem. Ref. Data* **27,** 413–656 (1998).

45. Merrick, J. P., Moran, D. & Radom, L. An evaluation of harmonic vibrational frequency

scale factors. *J. Phys. Chem. A* **111,** 11683–11700 (2007).

46. 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).

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

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

49. 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).

50. 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).

Appendix I: Energy Summary

Protonated Phe and Phe derivatives: summary of Gibbs' energies and relative energies at 298 K are given in kJ/mol. Calculations used the B3LYP functional and 6-311++G(d,p) basis set.

Table A1 The Gibbs' Energies and Relative energies of Phe

Phe Isomer	Gibbs (hartree)	Relative energy (kJ/mol)
Isomer 1	-555.1621	0.00
Isomer 2	-555.1602	4.95
Isomer 3	-555.1535	22.78
Isomer 4	-555.1257	95.68
Isomer 5	-555.1253	96.63
Isomer 6	-555.1252	96.86
Isomer 7	-555.1248	98.12
Isomer 8	-555.1162	120.52
Isomer 9	-555.1157	121.93
Isomer 10	-555.1122	130.97
Isomer 11	-555.1111	133.91
Isomer 12	-555.1095	138.06
Isomer 13	-555.1094	138.53
Isomer 14	-555.1094	138.56
Isomer 15	-555.1072	144.17

Table A2 The Gibbs' Energies and Relative energies of 3F-Phe

3F-Phe Isomer	Gibbs (hartree)	Relative energy (kJ/mol)
Isomer 1	-654.4359	0.00
Isomer 2	-654.4356	0.64
Isomer 3	-654.4345	3.53
Isomer 4	-654.4343	4.05
Isomer 5	-654.4306	13.85
Isomer 6	-654.4305	13.96
Isomer 7	-654.4295	16.71
Isomer 8	-654.4293	17.09
Isomer 9	-654.4281	20.36

Isomer 10	-654.4223	35.48
Isomer 11	-654.4220	36.34
Isomer 12	-654.4217	37.14
Isomer 13	-654.4216	37.42
Isomer 14	-654.4212	38.52
Isomer 15	-654.4209	39.30

Table A3 The Gibbs' Energies and Relative energies of 3CN-Phe

3CN-Phe Isomer	Gibbs (hartree)	Relative energy (kJ/mol)
Isomer 1	-647.4213	0.00
Isomer 2	-647.4197	4.15
Isomer 3	-647.4164	12.86
Isomer 4	-647.4160	13.93
Isomer 5	-647.4152	15.96
Isomer 6	-647.4075	36.28
Isomer 7	-647.4071	37.42
Isomer 8	-647.4066	38.65
Isomer 9	-647.4057	41.07

Table A4 The Gibbs' Energies and Relative energies of 3CF_3-Phe

3CF_3-Phe Isomer	Gibbs (hartree)	Relative energy (kJ/mol)
Isomer 1	-892.3060	0.00
Isomer 2	-892.3055	1.24
Isomer 3	-892.3051	2.32
Isomer 4	-892.3041	4.98
Isomer 5	-892.3010	13.15
Isomer 6	-892.3004	14.78
Isomer 7	-892.3003	14.96
Isomer 8	-892.2991	18.20
Isomer 9	-892.2990	18.37
Isomer 10	-892.2929	34.30
Isomer 11	-892.2921	36.54
Isomer 12	-892.2915	37.96
Isomer 13	-892.2915	38.12
Isomer 14	-892.2907	40.16
Isomer 15	-892.2901	41.63
Isomer 16	-892.2822	62.51

Table A5 The Gibbs' Energies and Relative energies of 4F-Phe

4F-Phe Isomer	Gibbs (hartree)	Relative energy (kJ/mol)
Isomer 1	-654.4356	0.00
Isomer 2	-654.4342	3.64
Isomer 3	-654.4304	13.52
Isomer 4	-654.4293	16.64
Isomer 5	-654.4281	19.54
Isomer 6	-654.4221	35.47
Isomer 7	-654.4213	37.45
Isomer 8	-654.4210	38.43
Isomer 9	-654.4123	61.13
Isomer 10	-654.3993	95.32
Isomer 11	-654.3905	118.31
Isomer 12	-654.3900	119.82
Isomer 13	-654.3869	127.74
Isomer 14	-654.3858	130.73
Isomer 15	-654.3838	135.91
Isomer 16	-654.3829	138.21

Table A6 The Gibbs' Energies and Relative energies of $4NO_2$-Phe

$4NO_2$-Phe Isomer	Gibbs (hartree)	Relative energy (kJ/mol)
Isomer 1	-759.7142	0.00
Isomer 2	-759.7126	3.98
Isomer 3	-759.7091	13.25
Isomer 4	-759.7080	16.16
Isomer 5	-759.7079	16.56
Isomer 6	-759.7018	32.39
Isomer 7	-759.7001	37.00
Isomer 8	-759.6995	38.56
Isomer 9	-759.6988	40.25

Table A7 The Gibbs' Energies and Relative energies of $2,5F_2$-Phe

$2,5F_2$-Phe Isomer	Gibbs (hartree)	Relative energy (kJ/mol)
Isomer 1	-753.7125	0.00
Isomer 2	-753.7121	0.96
Isomer 3	-753.7080	11.66
Isomer 4	-753.7077	12.47
Isomer 5	-753.7048	20.09

Isomer 6	-753.7030	24.88
Isomer 7	-753.7022	26.98
Isomer 8	-753.6982	37.36
Isomer 9	-753.6977	38.77
Isomer 10	-753.6973	39.74
Isomer 11	-753.6849	72.36
Isomer 12	-753.6734	102.62
Isomer 13	-753.6724	105.16
Isomer 14	-753.6704	110.42
Isomer 15	-753.6701	111.16
Isomer 16	-753.6638	127.69
Isomer 17	-753.6615	133.81
Isomer 18	-753.6612	134.69
Isomer 19	-753.6597	138.49
Isomer 20	-753.6593	139.56
Isomer 21	-753.6587	141.26

Table A8 The Gibbs' Energies and Relative energies of 3,5F_2-Phe

3,5F_2-Phe Isomer	Gibbs (hartree)	Relative energy (kJ/mol)
Isomer 1	-753.7084	0.00
Isomer 2	-753.7072	3.30
Isomer 3	-753.7032	13.61
Isomer 4	-753.7023	16.18
Isomer 5	-753.7015	18.16
Isomer 6	-753.6952	34.73
Isomer 7	-753.6946	36.23
Isomer 8	-753.6934	39.49
Isomer 9	-753.6933	39.69

Table A9 The Gibbs' Energies and Relative energies of 3,4$(MeO)_2$-Phe

3,4$(MeO)_2$-Phe Isomer	Gibbs (hartree)	Relative energy (kJ/mol)
Isomer 1	-784.2176	0.00
Isomer 2	-784.2170	1.46
Isomer 3	-784.2169	1.71
Isomer 4	-784.2165	2.89
Isomer 5	-784.2162	3.67
Isomer 6	-784.2157	4.98
Isomer 7	-784.2153	6.04

Isomer 8	-784.2151	6.47
Isomer 9	-784.2143	8.71
Isomer 10	-784.2140	9.44
Isomer 11	-784.2119	15.06
Isomer 12	-784.2118	15.21
Isomer 13	-784.2113	16.54
Isomer 14	-784.2109	17.68
Isomer 15	-784.2106	18.37
Isomer 16	-784.2102	19.31
Isomer 17	-784.2092	22.08
Isomer 18	-784.2072	27.36
Isomer 19	-784.2068	28.43
Isomer 20	-784.2067	28.59
Isomer 21	-784.2067	28.65
Isomer 22	-784.2044	34.56
Isomer 23	-784.2042	35.28
Isomer 24	-784.2013	42.79

Table A10 The Gibbs' Energies and Relative energies of F_5-Phe

F_5-Phe Isomer	Gibbs (hartree)	Relative energy (kJ/mol)
Isomer 1	-1051.5014	0.00
Isomer 2	-1051.5008	1.52
Isomer 3	-1051.4970	11.40
Isomer 4	-1051.4955	15.45
Isomer 5	-1051.4942	18.91
Isomer 6	-1051.4916	25.72
Isomer 7	-1051.4876	36.22
Isomer 8	-1051.4875	36.39
Isomer 9	-1051.4741	71.61
Isomer 10	-1051.4606	107.00
Isomer 11	-1051.4584	112.68
Isomer 12	-1051.4549	122.02
Isomer 13	-1051.4546	122.88
Isomer 14	-1051.4538	124.99
Isomer 15	-1051.4535	125.72
Isomer 16	-1051.4533	126.28
Isomer 17	-1051.4527	127.74

Proton-bound homodimers: summary of Gibbs' energies and relative energies at 298 K are given in kJ/mol. Calculations used the B3LYP functional and 6-311++G(d,p) basis set.

Table A11 The Gibbs' Energies and Relative energies of $(Phe)_2$

$(Phe)_2$ Isomer	Gibbs (hartree)	Relative energy (kJ/mol)
Isomer 1	-1109.9981	0.00
Isomer 2	-1109.9937	11.65
Isomer 3	-1109.9925	14.83
Isomer 4	-1109.9914	17.73
Isomer 5	-1109.9912	18.21
Isomer 6	-1109.9911	18.49
Isomer 7	-1109.9911	18.51
Isomer 8	-1109.9865	30.46
Isomer 9	-1109.9862	31.39
Isomer 10	-1109.9853	33.67
Isomer 11	-1109.9849	34.72
Isomer 12	-1109.9838	37.77
Isomer 13	-1109.9821	42.06
Isomer 14	-1109.9774	54.40
Isomer 15	-1109.9769	55.77
Isomer 16	-1109.9767	56.26
Isomer 17	-1109.9764	57.04
Isomer 18	-1109.9762	57.58
Isomer 19	-1109.9756	59.11
Isomer 20	-1109.9739	63.57

Table A12 The Gibbs' Energies and Relative energies of $(3F\text{-}Phe)_2$

$(3F\text{-}Phe)_2$ Isomer	Gibbs (hartree)	Relative energy (kJ/mol)
Isomer 1	-1308.5532	0.00
Isomer 2	-1308.5492	10.43
Isomer 3	-1308.5475	14.86
Isomer 4	-1308.5475	14.93

Isomer 5	-1308.5470	16.12
Isomer 6	-1308.5469	16.43
Isomer 7	-1308.5455	19.99
Isomer 8	-1308.5440	23.94
Isomer 9	-1308.5437	24.95
Isomer 10	-1308.5434	25.62
Isomer 11	-1308.5418	29.83
Isomer 12	-1308.5406	32.97
Isomer 13	-1308.5405	33.12
Isomer 14	-1308.5384	38.83
Isomer 15	-1308.5376	40.77

Table A13 The Gibbs' Energies and Relative energies of $(F_5\text{-Phe})_2$

$(F_5\text{-Phe})_2$ Isomer	Gibbs (hartree)	Relative energy (kJ/mol)
Isomer 1	-2102.6903	0.00
Isomer 2	-2102.6878	6.43
Isomer 3	-2102.6877	6.67
Isomer 4	-2102.6854	12.64
Isomer 5	-2102.6845	15.00
Isomer 6	-2102.6841	16.14
Isomer 7	-2102.6839	16.62
Isomer 8	-2102.6837	17.32
Isomer 9	-2102.6835	17.82
Isomer 10	-2102.6829	19.28
Isomer 11	-2102.6829	19.36
Isomer 12	-2102.6827	19.98
Isomer 13	-2102.6823	20.88
Isomer 14	-2102.6823	21.04
Isomer 15	-2102.6803	26.19
Isomer 16	-2102.6800	26.92
Isomer 17	-2102.6792	29.07
Isomer 18	-2102.6784	31.25
Isomer 19	-2102.6763	36.76

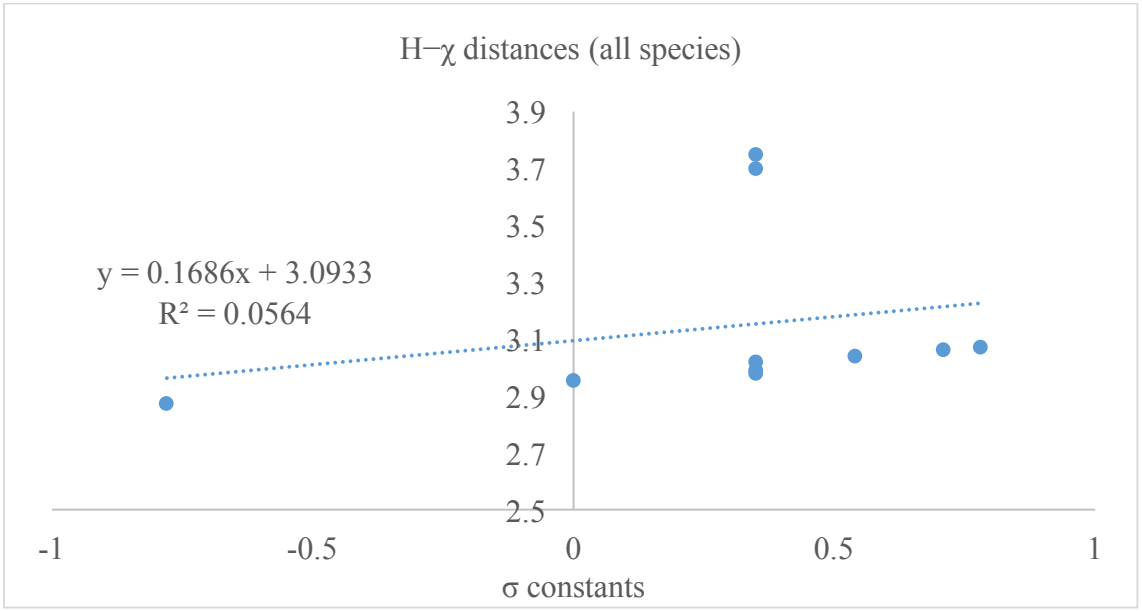
Protonated Phe dipeptide: summary of Gibbs' energies and relative energies at 298 K are given in kJ/mol. Calculations used the B3LYP functional and 6-311++G(d,p) basis set.

Table A14 The Gibbs' Energies and Relative energies of PhePhe

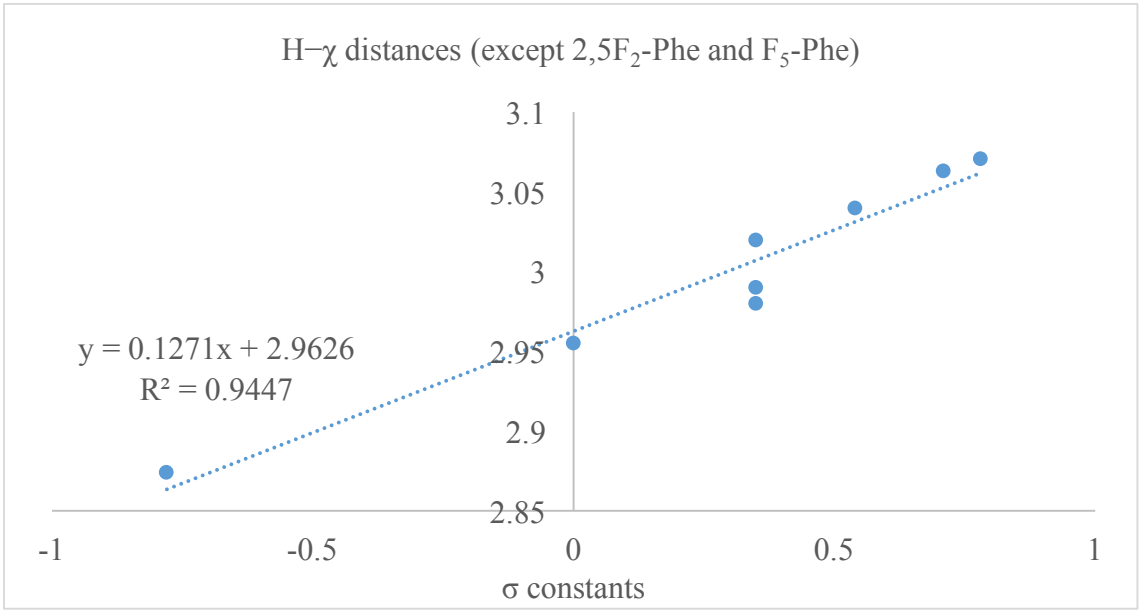
PhePhe Isomer	Gibbs (hartree)	Relative energy (kJ/mol)
Isomer 1	-1033.5358	0.00
Isomer 2	-1033.5358	0.003
Isomer 3	-1033.5354	1.00
Isomer 4	-1033.5344	3.67
Isomer 5	-1033.5340	4.59
Isomer 6	-1033.5336	5.68
Isomer 7	-1033.5323	9.03
Isomer 8	-1033.5322	9.27
Isomer 9	-1033.5319	10.16
Isomer 10	-1033.5313	11.67
Isomer 11	-1033.5301	14.91
Isomer 12	-1033.5300	15.12
Isomer 13	-1033.5299	15.30
Isomer 14	-1033.5290	17.67
Isomer 15	-1033.5289	18.11
Isomer 16	-1033.5288	18.39
Isomer 17	-1033.5283	19.56
Isomer 18	-1033.5281	20.18
Isomer 19	-1033.5279	20.56
Isomer 20	-1033.5277	21.31
Isomer 21	-1033.5274	21.94
Isomer 22	-1033.5273	22.22
Isomer 23	-1033.5271	22.72
Isomer 24	-1033.5270	22.94
Isomer 25	-1033.5268	23.51

Isomer 26	-1033.5266	24.01
Isomer 27	-1033.5263	24.94
Isomer 28	-1033.5261	25.33
Isomer 29	-1033.5259	25.85
Isomer 30	-1033.5257	26.37
Isomer 31	-1033.5250	28.29
Isomer 32	-1033.5245	29.62
Isomer 33	-1033.5240	30.93
Isomer 34	-1033.5238	31.41
Isomer 35	-1033.5236	31.93
Isomer 36	-1033.5233	32.82
Isomer 37	-1033.5229	33.69
Isomer 38	-1033.5227	34.23
Isomer 39	-1033.5226	34.63
Isomer 40	-1033.5100	67.79
Isomer 41	-1033.5084	71.79
Isomer 42	-1033.5080	72.80
Isomer 43	-1033.5075	74.11
Isomer 44	-1033.5074	74.58
Isomer 45	-1033.5072	74.96
Isomer 46	-1033.5059	78.46
Isomer 47	-1033.5051	80.42
Isomer 48	-1033.5046	81.91
Isomer 49	-1033.5044	82.30
Isomer 50	-1033.5040	83.31
Isomer 51	-1033.5034	84.87
Isomer 52	-1033.5023	88.00
Isomer 53	-1033.5017	89.53
Isomer 54	-1033.5005	92.56
Isomer 55	-1033.4984	98.24
Isomer 56	-1033.4949	107.31

Appendix II: Linear Regression with Hammett Parameters

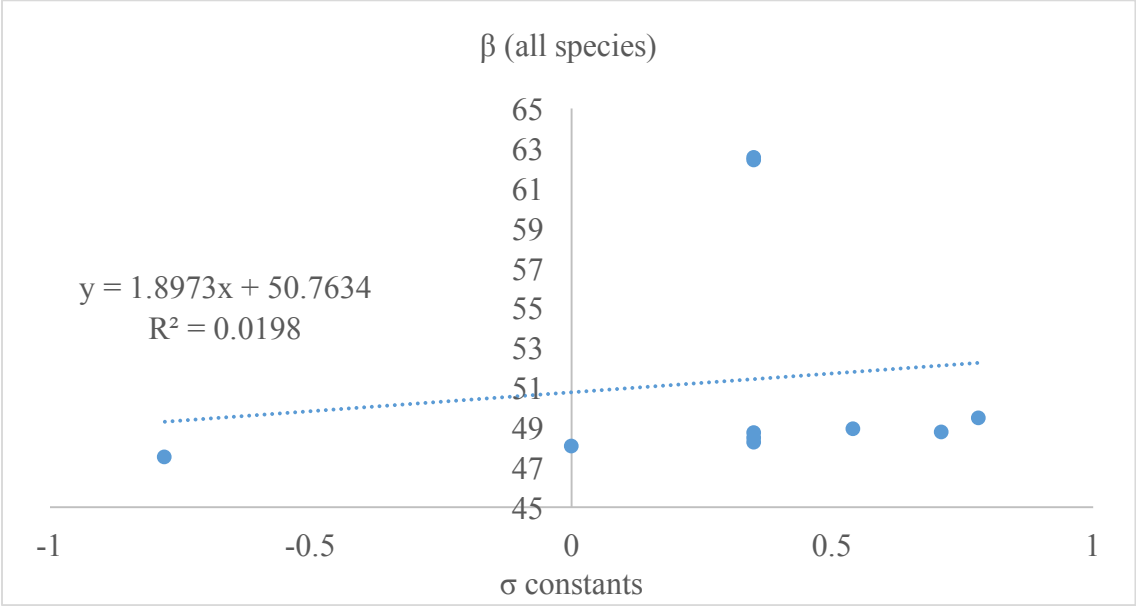


(1)

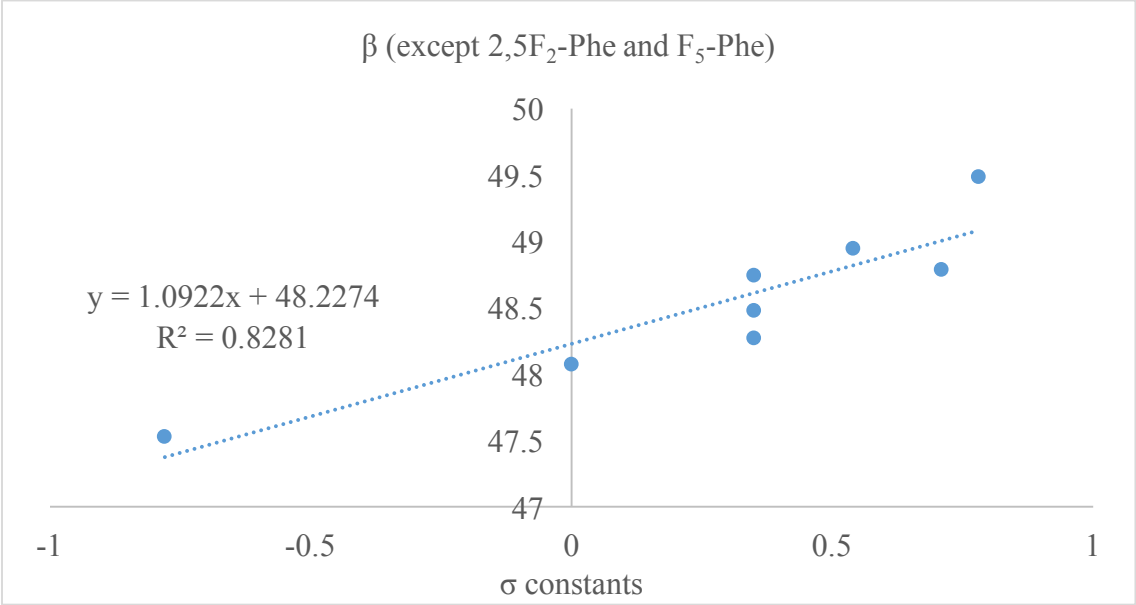


(2)

Figure A1 The linear regression between H–χ distances and σ constants for (1) all Phe derivatives; (2) Phe derivatives except 2,5F_2-Phe•H^+ and F_5-Phe•H^+.

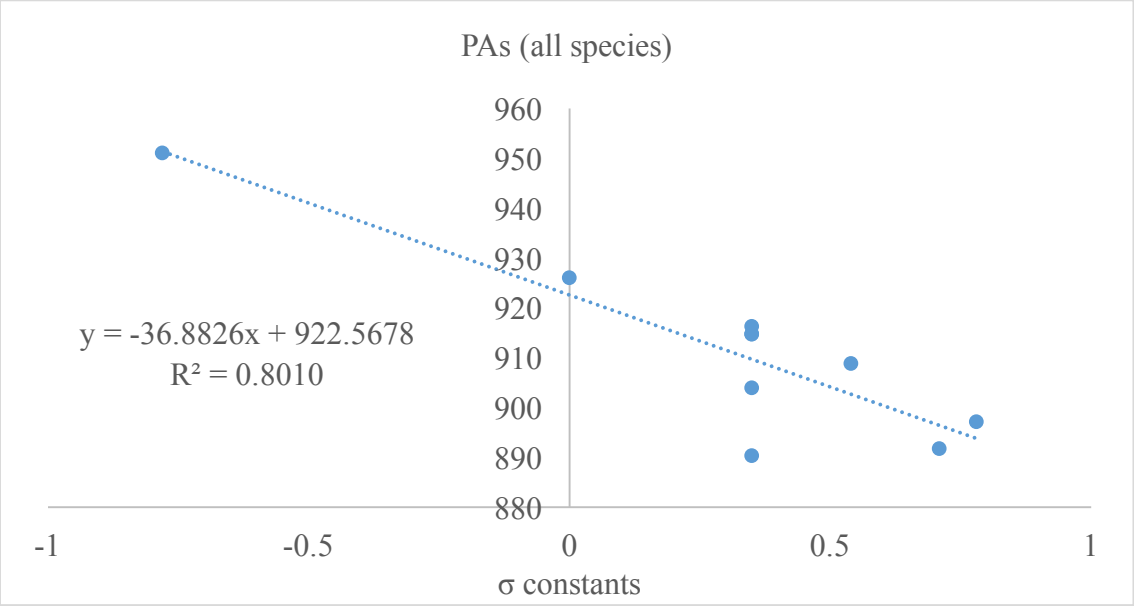


(1)

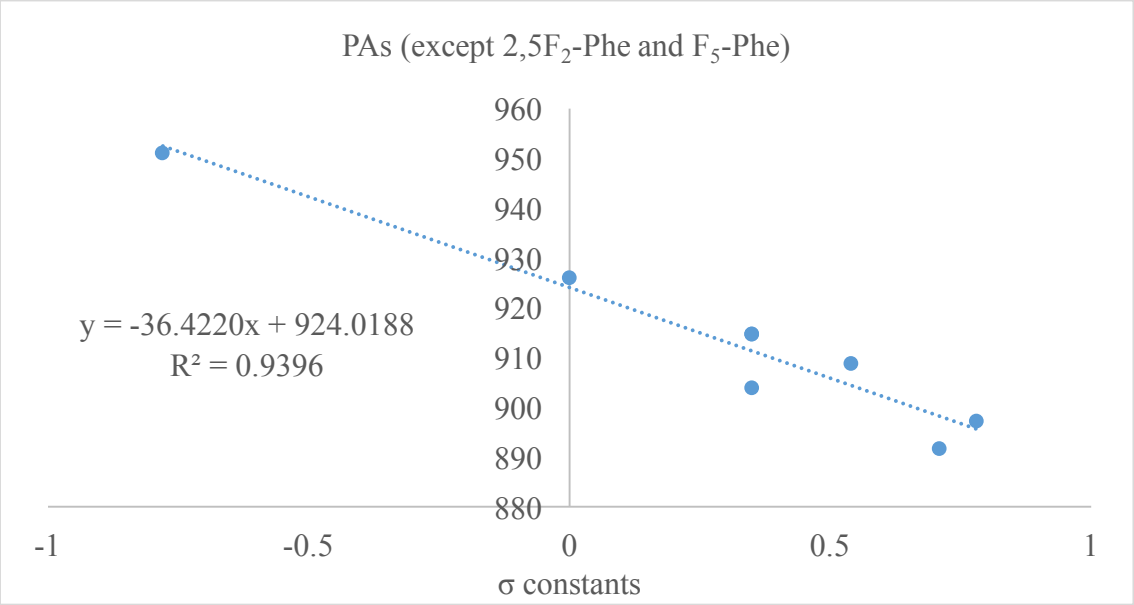


(2)

Figure A2 The linear regression between β angles and σ constants for (1) all Phe derivatives; (2) Phe derivatives except 2,5F_2-Phe•H^+ and F_5-Phe•H^+.

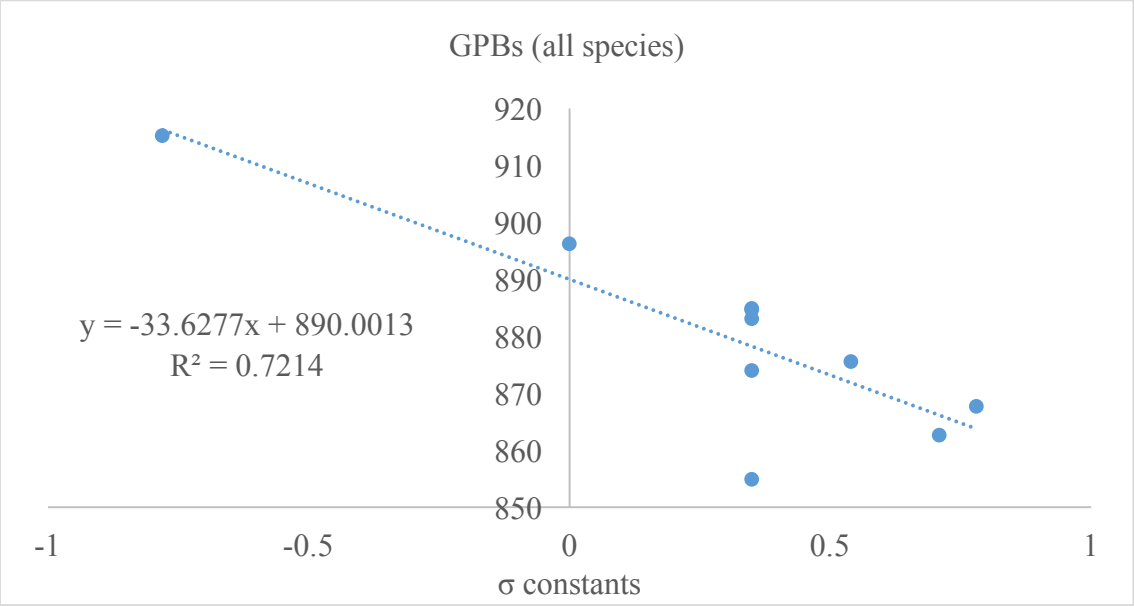


(1)

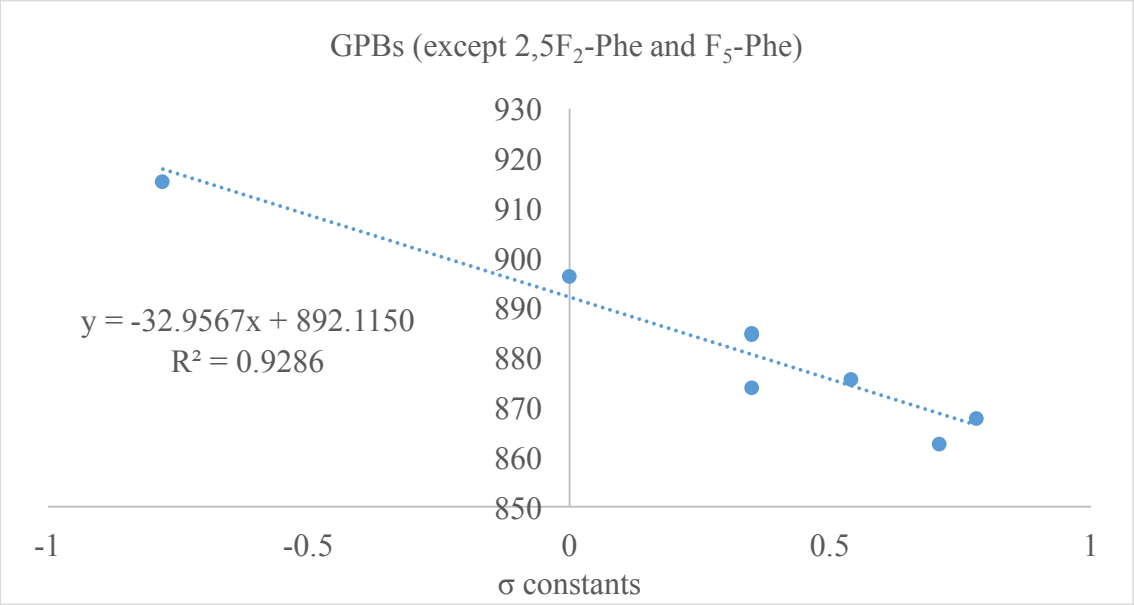


(2)

Figure A3 The linear regression between PAs and σ constants for (1) all Phe derivatives; (2) Phe derivatives except 2,5F_2-Phe•H^+ and F_5-Phe•H^+.



(1)



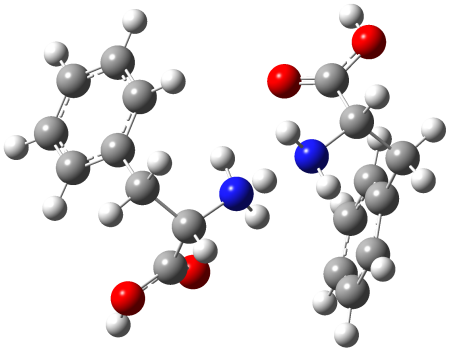
(2)

Figure A4 The linear regression between GPBs and σ constants for (1) all Phe derivatives; (2) Phe derivatives except 2,5F_2-Phe•H^+ and F_5-Phe•H^+.

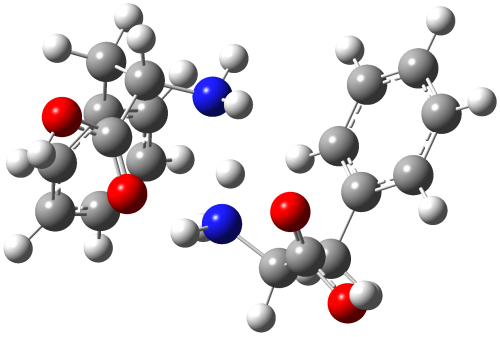
Appendix III: Structures

The Proton-bound Dimer of Phenylalanine, $(Phe)_2 \bullet H^+$

Isomer 3, 7, and 8 are the second higher energy monodentate-bridged, and bidentate-bridged, and canonical structures.



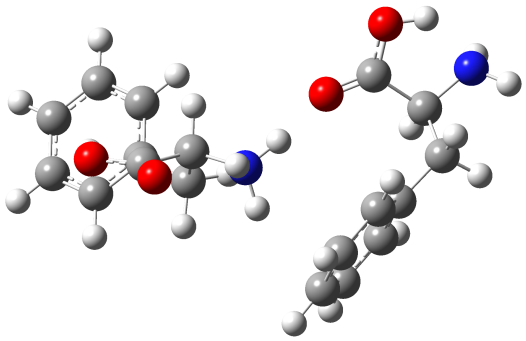
Isomer 3 monodentate-bridged
14.8 kJ/mol



Isomer 7 bidentate-bridged
18.5 kJ/mol

N	0.151401	0.404636	0.378552	-0.032159	0.458624	-1.11027
H	-0.283871	1.069058	-0.27508	0.327602	-0.41551	-1.502997
H	-0.601696	-0.058039	0.965589	-0.071523	0.345559	-0.05117
H	0.584475	-0.334074	-0.188203	0.708091	1.176733	-1.183401
C	1.175544	1.0998	1.220616	-1.309764	0.87804	-1.764751
H	0.651826	1.515202	2.086273	-1.076189	1.139703	-2.799723
C	2.26563	0.117147	1.713091	-2.36177	-0.259442	-1.77709
C	1.706124	2.283972	0.411846	-1.800228	2.141952	-1.062509
H	2.938217	0.687156	2.356008	-3.213218	0.116044	-2.348747
H	1.779517	-0.629101	2.348103	-1.942853	-1.088148	-2.356182
C	3.033092	-0.56345	0.597709	-2.810831	-0.750073	-0.415901
O	1.297889	2.575394	-0.682232	-1.363821	2.555512	-0.019843
C	4.18127	0.028049	0.054945	-3.817639	-0.081129	0.291451
C	2.602553	-1.793864	0.084228	-2.246495	-1.899084	0.148827
H	4.549345	0.963709	0.4617	-4.300791	0.786934	-0.144886
C	4.872117	-0.588728	-0.984703	-4.234631	-0.538263	1.539234
C	3.293064	-2.40941	-0.960268	-2.664078	-2.360907	1.396713
H	1.739529	-2.293867	0.513292	-1.49697	-2.457864	-0.402639

H	5.764903	-0.12432	-1.386885	-5.023513	-0.015677	2.067495
C	4.426351	-1.806093	-1.498666	-3.656006	-1.677668	2.09711
H	2.95191	-3.365197	-1.340412	-2.232655	-3.265586	1.810101
H	4.968609	-2.286962	-2.30417	-3.991865	-2.042115	3.060825
O	2.633663	2.960732	1.094495	-2.780033	2.718749	-1.766976
H	2.922453	3.721006	0.563548	-3.075177	3.518589	-1.301968
N	-1.867619	-0.914766	1.994049	0.504689	0.630658	1.639991
H	-1.354763	-1.568873	2.581662	-0.029976	1.482701	1.800423
H	-2.376785	-0.304046	2.628512	0.161279	-0.059235	2.303262
C	-2.834261	-1.656323	1.159024	1.939506	0.866813	1.875957
H	-3.390269	-2.401258	1.740903	2.130824	1.266942	2.877618
C	-3.889205	-0.701955	0.532296	2.75677	-0.45548	1.760375
C	-2.062494	-2.437053	0.100935	2.469446	1.904638	0.892022
H	-4.53915	-1.30365	-0.106276	3.798396	-0.219835	1.990432
H	-4.513864	-0.330737	1.349906	2.40482	-1.114775	2.559318
C	-3.309354	0.46495	-0.23709	2.657782	-1.166904	0.42775
O	-0.880579	-2.313846	-0.128155	2.010546	2.154114	-0.203266
C	-2.8595	0.30948	-1.554728	1.739639	-2.20684	0.236994
C	-3.202279	1.7282	0.357735	3.486801	-0.808868	-0.643745
H	-2.97208	-0.646722	-2.055077	1.117129	-2.53067	1.064522
C	-2.288866	1.37778	-2.246451	1.631529	-2.851939	-0.995978
C	-2.63713	2.801083	-0.332157	3.379703	-1.445823	-1.879016
H	-3.58584	1.884289	1.361501	4.237375	-0.036801	-0.51048
H	-1.953943	1.240708	-3.268148	0.935726	-3.675038	-1.116246
C	-2.168573	2.625935	-1.634106	2.445339	-2.465947	-2.062119
H	-2.578703	3.775324	0.139781	4.039023	-1.160938	-2.690741
H	-1.730528	3.457033	-2.173241	2.376449	-2.978081	-3.014889
O	-2.857277	-3.296176	-0.54795	3.557508	2.511695	1.374587
H	-2.329832	-3.781682	-1.202284	3.886264	3.14306	0.71407



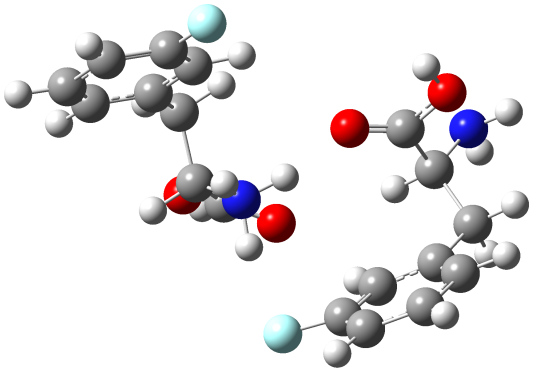
Isomer 8 Canonical
30.5 kJ/mol

N	0.025386	-0.915685	0.772249
H	0.682953	-0.103333	0.763107
H	0.244651	-1.507513	-0.037452
C	-1.409698	-0.474336	0.794857
H	-1.430955	0.465405	1.3533
C	-1.935101	-0.240157	-0.636386
C	-2.157004	-1.530952	1.641327
H	-1.20037	0.382935	-1.156994
H	-1.979807	-1.196335	-1.167209
C	-3.292573	0.43864	-0.680394
O	-3.477659	-1.443832	1.708832
C	-4.410803	-0.235264	-1.188138
C	-3.443312	1.760738	-0.236098
H	-4.309941	-1.252763	-1.552086
C	-5.655292	0.394305	-1.240617
C	-4.68503	2.387208	-0.288916
H	-2.585773	2.311561	0.138553
H	-6.510309	-0.139277	-1.638211
C	-5.794028	1.704003	-0.789558
H	-4.786716	3.409887	0.054418
H	-6.759518	2.193764	-0.831263
O	-1.535882	-2.371593	2.241417
H	0.178921	-1.502013	1.605273
N	4.57248	2.94328	-0.246662
H	5.521209	2.661897	-0.466395

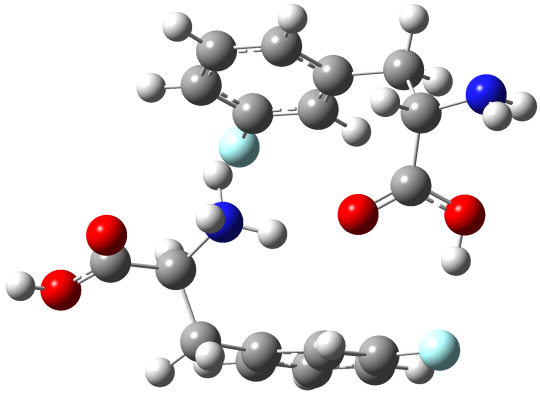
H	4.349002	3.760039	-0.804762
C	3.598642	1.855725	-0.413065
H	2.991488	1.976512	-1.314939
C	4.332424	0.489371	-0.476061
C	2.628094	1.904608	0.781186
H	5.037287	0.559009	-1.311764
H	4.932773	0.375925	0.431701
C	3.467953	-0.739129	-0.670321
O	1.574216	1.289324	0.811608
C	3.431794	-1.744182	0.303072
C	2.741792	-0.928849	-1.853384
H	4.005121	-1.626994	1.21684
C	2.691825	-2.911112	0.103218
C	1.993975	-2.088266	-2.055281
H	2.777357	-0.179115	-2.637409
H	2.698936	-3.689635	0.857813
C	1.966885	-3.086283	-1.07646
H	1.460949	-2.228944	-2.989023
H	1.415493	-4.004171	-1.24685
O	3.018997	2.660135	1.781778
H	3.859746	3.086576	1.469761
H	-3.838201	-0.755256	1.121511

The Proton-bound Dimer of 3-fluorophenylalanine, $(3F\text{-}Phe)_2 \bullet H^+$

Isomer 14 and 15 are the structures with $-NH_3^+ \bullet\bullet\bullet F$ and/or $-OH \bullet\bullet\bullet F$ hydrogen bonds.



Isomer 14
38.8 kJ/mol



Isomer 15
40.8 kJ/mol

C	3.529324	-0.623378	-0.395442	-2.273422	1.734989	-0.148876
C	2.705147	0.077845	-1.285072	-1.182238	2.425653	-0.679688
C	2.192003	-0.584968	-2.382462	-0.526973	3.347658	0.118643
C	2.498989	-1.894484	-2.703577	-0.92856	3.66149	1.402624
C	3.357188	-2.572687	-1.839108	-2.042161	2.992124	1.912312
C	3.849468	-1.952409	-0.689918	-2.701807	2.030332	1.151295
H	2.443619	1.11668	-1.12267	-0.827862	2.25218	-1.68785
H	2.082575	-2.36044	-3.587412	-0.391569	4.404051	1.978966
H	3.629494	-3.59846	-2.056648	-2.391577	3.223435	2.911094
H	4.490831	-2.510031	-0.016884	-3.561948	1.518003	1.568922
C	4.00789	0.043547	0.876726	-2.945614	0.647534	-0.958301
H	4.52809	-0.685004	1.503434	-4.021916	0.647346	-0.774294
H	4.726549	0.837198	0.64567	-2.805665	0.815428	-2.030716
C	2.871114	0.705422	1.715708	-2.439081	-0.768781	-0.593316
H	2.417334	1.50721	1.132331	-2.51284	-0.919481	0.484314
C	1.748462	-0.289477	1.977749	-3.237829	-1.849978	-1.323699
N	3.281265	1.274252	2.993542	-1.00303	-0.990828	-0.989091
H	3.800121	0.610758	3.559622	-0.38778	-0.157818	-0.957515
H	3.85216	2.101346	2.860753	-0.563203	-1.701473	-0.385575
O	0.655053	-0.286923	1.429471	-2.780775	-2.487976	-2.239265
O	2.077277	-1.207763	2.880605	-4.470876	-1.956907	-0.837949
H	1.32934	-1.810446	3.020454	-4.964014	-2.62399	-1.344554
F	1.275588	0.09761	-3.156113	0.590081	3.948095	-0.39491

C	-3.410135	0.374355	0.304737	2.374629	-1.628303	0.774681
C	-3.09903	-0.834047	0.940519	1.516743	-1.109552	1.75078
C	-3.924973	-1.933077	0.740905	0.377055	-1.81967	2.103524
C	-5.050609	-1.881261	-0.065131	0.044025	-3.037396	1.525095
C	-5.355339	-0.673965	-0.691731	0.909135	-3.557194	0.555053
C	-4.544005	0.445083	-0.513369	2.056573	-2.860122	0.183032
H	-2.24841	-0.92494	1.607205	1.731406	-0.173344	2.252932
H	-5.670908	-2.760514	-0.184088	-0.830486	-3.579622	1.864253
H	-6.238388	-0.60557	-1.315565	0.692683	-4.521197	0.109321
H	-4.809797	1.38065	-0.994013	2.724342	-3.289084	-0.555822
C	-2.507024	1.578041	0.482783	3.625497	-0.883834	0.36716
H	-3.093163	2.496841	0.537342	3.874934	-0.125656	1.114918
H	-1.940589	1.498129	1.414694	4.471543	-1.573043	0.318204
C	-1.519893	1.743602	-0.696415	3.567795	-0.207022	-1.034283
H	-2.077637	1.882965	-1.625322	3.276779	-0.964281	-1.766533
C	-0.565196	2.919446	-0.516644	2.440964	0.828476	-1.051461
N	-0.686014	0.504687	-0.857378	4.88984	0.276461	-1.379542
H	-1.306525	-0.309019	-0.924638	4.942308	0.600844	-2.338631
H	-0.097914	0.539225	-1.699145	5.192406	1.03279	-0.77484
O	0.629481	2.799216	-0.411403	1.281728	0.561795	-1.333137
O	-1.233482	4.071634	-0.490541	2.83701	2.048533	-0.702987
H	-0.608418	4.806121	-0.373093	2.087735	2.672883	-0.711085
F	-3.613365	-3.08714	1.363378	-0.445777	-1.300022	3.042555
H	-0.062044	0.345648	-0.027148	-0.986759	-1.372033	-1.944646