AN ABSTRACT OF THE DISSERTATION OF

Benjamin James Bythell for the degree of Doctor of Philosophy in Chemistry presented

on December 5, 2007.

Title: Gas-phase Fragmentation Chemistry of Protonated Peptide Ions

Abstract approved:_____

Douglas F. Barofsky

The fragmentation characteristics of protonated peptides have been investigated using

tandem mass spectrometry (MS/MS) under various fragmentation regimes and detailed

density functional theory (DFT) reaction pathway calculations. The DFT calculations

predict novel salt-bridge stabilized transition structures as well as consecutive reactions

occurring in proton-bound dimers for some of the ions detected in the MS/MS

experiments. Four previously unknown protonated peptide fragmentation mechanisms are

described in this thesis based on the theoretical and experimental results. These results

have been evaluated using the mobile proton model and the more recently introduced

pathways in competition (PIC) protonated peptide fragmentation model. The mobile

proton model cannot adequately explain these experimental findings. The PIC model on

the other hand, considers post-cleavage as well as pre-cleavage and bond-cleavage

events, which enable these complicated protonated peptide fragmentation pathways to be

rationalized.

Gas-phase Fragmentation Chemistry of Protonated
Peptide Ions

by
Benjamin James Bythell

A DISSERTATION

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

Presented December 5, 2007
Commencement June 2008

Doctor of Philosophy dissertation of <u>Benjamin James Bythell</u> presented on <u>December 5, 2007</u>

<u>APPROVED</u>:

_____

Major Professor, representing Chemistry

_____

Chair of the Department of Chemistry

_____

Dean of the Graduate School

I understand that my dissertation will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my dissertation to any reader upon request.

_____

Benjamin James Bythell, Author

ACKNOWLEDGEMENTS

CONTRIBUTION OF AUTHORS

Dr. Béla Paizs contributed to the calculations, discussion and writing presented in papers Chapters 2 & 3. Franceso Pingitore took the hybrid sector-quadrapole tandem mass spectra in paper *J. Am. Soc. Mass Spectrom*., 200**7**, 7, 1291-1303 (Chapter 2). Prof. Chrys Wesdemiotis was a minor contributor to the writing of *J. Am. Soc. Mass Spectrom*., 2007, 7, 1291-1303 (Chapter 2).

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

LIST OF FIGURES

LIST OF FIGURES (Continued)

LIST OF TABLES

LIST OF SCHEMES

Gas-phase Fragmentation Chemistry of Protonated

Peptide Ions

1 Introduction

1.1 Overview of Thesis

The research described in this thesis focuses the use of tandem time-of-flight and ion trap mass spectrometry coupled with computational modeling to study the fragmentation characteristics of small protonated peptide ions in the gas phase. This thesis begins with a brief description of peptide, amino acid and peptide fragmentation nomenclature. This is followed by an introduction to why studying protonated peptides is important and how this is achieved by mass spectrometric and modeling techniques.

The first system investigated was protonated alanylglycylglycine, [AGG + H]$^+$ a relatively simple aliphatic tripeptide, with some novel fragmentation characteristics. Tandem mass spectrometry, isotope labeling, and theoretical calculations were used to investigate the effect of activation method and experimental timescale on the observed spectra. This is discussed in chapter 2 and has been published (*J. Am. Soc. Mass Spectrom.*, **2007**, 7, 1291-1303).

The second system studied using these techniques was protonated arginylglyclaspartic acid, [RGD + H]$^+$. This is a very different type of peptide as it contains both acidic and basic residues which can interact. This offers the possibility of internal solvation or salt bridge (protonated zwitterion) interactions stabilizing the ion and also potentially being involved in the fragmentation mechanisms. Tandem mass

spectrometry and theoretical calculations were used to investigate the effect of activation method and experimental timescale on the observed spectra. This work is discussed in Chapter 3 and has been submitted for publication in the Journal of the American Chemical Society.

Finally, Chapter 4 summarizes some of the principal findings from these studies and indicates where future investigation might be targeted.

1.2 Mass Spectrometry

1.2.1 In the beginning…

Mass Spectrometry (MS), which determines the mass-to-charge ratios (m/z) of ions, began with Sir J. J. Thomson's studies on electrical discharges in gases [1]. These studies led to the discovery of the electron in 1897 and subsequently to the first mass spectrometer for the determination of m/z of ions. Mass spectrometers comprise a source for generating the ions, a means of separating the ions and finally a means of detecting them. Thomson's spectrometer used discharge tubes to generate the ions, electric and magnetic fields to separate the different m/z ratios in parabolic trajectories ('parabola spectrograph' was the term Thomson used for his mass spectrometer), followed by detection on photographic plates or fluorescent screens. The Nobel Prize in Physics was awarded to Thompson in 1906 "in recognition of the great merits of his theoretical and experimental investigations on the conduction of electricity by gases" [2].

Many subsequent methods of achieving the measurement on the mass-to-charge ratios of ions have been developed resulting in several more Nobel Prizes in Physics and Chemistry, reflecting the growing importance of mass spectrometry. MS techniques and allied technologies are the subject matter of this thesis.

1.2.2 Ionization Methods

The massive improvements in mass spectrometry technology over the last twenty five years owe much to the development of new ionization methods. Fast-atom-

bombardment (FAB) [3], electrospray ionization (ESI) [4] and matrix-assisted-laser-desorption/ ionization (MALDI) [5-7], are so-called 'soft' ionization techniques that were all introduced in the 1980s. These complimentary techniques now enable a wide variety analytes to be ionized and introduced into the mass spectrometer simply and efficiently. ESI and MALDI in particular, require far less sample in order to perform analyses than earlier methods, while allowing much higher mass molecules to be investigated e.g. peptides, proteins and DNA. The importance of these techniques was recognized by the 2002 Nobel Prize in Chemistry being awarded (1/4 share each) to John Fenn and Koichi Tanaka for *"for their development of soft desorption ionisation methods for mass spectrometric analyses of biological macromolecules"* [2]. These techniques are used in some of my experimental work and will be described next.

1.2.2.1 Fast-Atom-Bombardment (FAB)

The first of these three 'soft' ionization techniques was FAB [3, 8] which enabled large polar organic molecules (up to ~20kDa) to be ionized in a mass spectrometer for m/z determination. This technique uses a solution of the analyte molecule in a low-volatility matrix (glycerol, sulfuric acid) which is bombarded with fast heavy atoms (Xe, Ar, Cs) in the source producing analyte ions as well as matrix oligimers and other adducts. The matrix reactions that occur in FAB have the disadvantage of producing chemical background signal. FAB is also generally less sensitive than either MALDI or ESI, so is not used as frequently any more.

1.2.2.2 Matrix-Assisted-Laser-Desorption/Ionization (MALDI)

The MALDI technique takes advantage of solvation to large molecules by an organic acid matrix. The analyte molecule is prepared in a solution of organic acid and volatile solvent where the concentration of the organic acid is on the order of 10^4 times that of the analyte molecule. The solution is then placed onto a stainless steel plate and the volatile solvent is allowed to evaporate, leaving crystals of the matrix with the analyte molecules contained. The sample plate is then introduced to the mass spectrometer and a pulsed laser is used to ablate the matrix and analyte sample. The resulting plume contains matrix ions and neutrals, but crucially, generates gas-phase analyte ions in significant quantities. The exact mechanism by which this occurs is a subject of some debate in the literature [9-12], but the result is usually singly charged ions resulting from the net gain of a proton by the analyte molecules in the MALDI process. These ions may then be accelerated, dispersed and detected based on their mass-to-charge ratio, m/z.

1.2.2.3 Electrospray Ionization (ESI)

ESI [4, 13] is the most gentle of the ionization techniques, so much so that it allows some biological complexes to remain intact during the ionization process so that they may be detected in the mass spectrometer. ESI utilizes the already present liquid-phase ions and transfers them to the gas-phase. A needle electrode is maintained at high voltage while the sample (dissolved in solvent) is pumped through it. A Taylor cone forms at the needle. Charged droplets formed from the analyte electrolytes dissolved in the solvent are formed at the end of the cone and an unstable 'mist' of these droplets is

created at the tip. These droplets shrink as solvent evaporates from them and repeated droplet fissions lead to small, highly charged droplets which eventually form gas-phase ions. Some of these ions make it into the mass spectrometer for subsequent analysis. Modern ESI sources used heated needles and co-axial gas flow to aid electrical nebulization and ion transport. This also has the advantage of enabling higher flow rates of sample to be run without destabilizing the electrospray.

1.2.3 Mass Spectrometry types and Instruments

1.2.3.1 Ion Trap Mass Spectrometry

The quadrupole ion trap (IT) employs three electrodes; two end-cap electrodes that are normally at ground potential and a ring electrode between them which has a radiofrequency (RF) voltage applied to generate a quadrupole electric field [14]. Once trapped, the ions present may be analyzed by adjusting the operating voltages to cause particular m/z ratios to adopt unstable trajectories. Adjusting the RF voltage to the ring electrode makes ions of increasing m/z adopt unstable trajectories and exit the IT before being detected to produce a mass spectrum. Alternatively, resonant ejection may be used where by manipulation of the ring RF voltage and a supplementary RF signal applied to the end caps. Ions successively come into resonance with the supplementary signal, exit the IT and get detected externally. This method requires less voltage to expel a given m/z, so extends the range of the instrument.

Tandem MS (MS/MS or MS^n) can be performed using mass-selective stability, which is the combination of DC and RF fields to make all ions except the desired m/z ions unstable (can be done via resonant ejection too). The isolated ions can now be

collided (CID/ CAD, see section 1.3.2) with helium bath gas (present continually to improve IT performance) using resonant excitation but not ejection. After an allotted time, the resulting ions are sequentially ejected to produce a MS/MS spectrum of the selected ion. This analysis can be successively repeated on the resulting ions giving a MS^n spectrum (where n=3,4,5…) as long as sufficient ions remain in the IT. These tandem experiments are tandem in time, whereas for most other instruments, these experiments are tandem in space (i.e. the MS/MS experiment occurs away from the original precursor ion isolation; see Section 1.2.2.3.1 for further explanation). This is a major advantage for this type of instrument but this analytical flexibility does come at the cost of increased time per spectrum.

1.2.3.2 Sector Mass Spectrometry

Sector instruments contain magnetic and electric sector that can be arranged and operated in a variety of configurations. In this thesis a hybrid-sector EBEhQ geometry instrument (E, electric sector; B, magnetic sector; h, RF-only (transmission) hexapole; Q, quadrupole mass filter) with a FAB source was used [15]. Only the sector section (EBE) was used, so I shall not describe the hexapole or quadrupole sections. Protonated peptide ions formed by the FAB process are accelerated to 8 keV and mass-selected by the EB sectors for measurement of their metastable decomposition or CID (section 1.3.2) tandem mass spectra in the field-free region (FFR) between EB and the subsequent electric sector. The product ions from these reactions were mass-analyzed by scanning the second electric sector. In MS^3 experiments, a specific fragment ion from metastable decomposition of the protonated peptide ion dissociating in the field-free region in front

of the first electric sector was transmitted through EB by using the E and B fields, and the corresponding 8keV CID spectrum was acquired.

1.2.3.3 Time-of-flight Mass Spectrometry

The basis of time-of-flight (TOF) mass spectrometry was described initially by W.E Stephens in 1946 [16]. His paper stated that:

"Advances in electronics seem to make practical a type of mass spectrometer in which microsecond pulses of ions are selected every millisecond from an ordinary low-voltage ion source. In traveling down the vacuum tube, ions of different m/z have different velocities and consequently separate into groups spread out in space. … This type of mass spectrometer should offer many advantages over present types. The response time should be limited only by the repetition rate (milliseconds)… Magnets and stabilization equipment would be eliminated. Resolution would not be limited by smallness of slits or alignment. Such a mass spectrometer should be well suited for composition control, rapid analysis, and portable use."

So the TOF is a very simple concept. Ions are formed in the source, an electric field is applied and the ions are accelerated to a constant energy (qV) as they enter the field-free drift region (D) and finally arrive at the detector having flight times (t) proportional to the square root of their masses (m) to first order [17]

$$t = (m/2qV)^{1/2}D$$

Ions of differing masses arrive at the detector at different times and a mass spectrum is created. Distributions in the initial velocity and position in the source mean that a kinetic energy distribution is also present. Consequently, ions of the same mass do

not necessarily arrive at the detector at exactly the same time which causes a decrease in resolution.

The situation was improved in 1955 by Wiley and McLaren who approached the problem of the initial spatial and velocity distributions by using a dual-stage source [18]. This was achieved by adjusting the two fields until the space focal plain was positioned at the detector. A delay was also introduced between the ion formation and the application of the accelerating fields (known as time-lag focusing). This allowed the ions to arrange their positions according to their initial velocities in the direction of the mass analyzer. Those ions with greatest initial velocity (and therefore energy) in this direction, consequently gain less energy from the extracting fields.

The introduction of the reflectron (or ion mirror) by B.A. Mamyrin and co-workers provided kinetic energy focusing that was independent of mass [19]. Those ions with greatest kinetic energy penetrate further into the reflectron before being turned around by the retarding electric field and reaccelerated to the same kinetic energy. This means that they spend longer in the reflectron than do ions of the same m/z which have less kinetic energy. The detector is positioned such that the drift length D is equal to 4d where d is the average depth of penetration of a particular m/z into the reflectron. Consequently the TOF is described by (to 1st order)

$$TOF = (m/2qV)^{1/2}[D + 4d]$$

in which D is the sum of the drift lengths before and after the reflectron [20].

MS/MS in TOF instruments is achieved in a variety of ways largely dependent on the ionization method used. A description of how this is achieved for the Applied Biosystems (AB) MALDI-TOF/TOF used in this thesis is described in the next section.

1.2.3.3.1 MALDI-TOF/TOF Mass Spectrometry: The AB 4700 Proteomics Analyzer

This instrument (Figure 2.1) operates using a MALDI source with a diode pumped Nd:YAG laser (355nm, 5-ns pulse width; Laser-Compact, Moscow, Russia) and can be operated in MS, MALDI-TOF, or MS^2, MALDI-TOF/TOF modes. In MS mode the ions created are accelerated to 8keV in the source, allowed to disperse based on m/z before being reflected in the ion mirror (reflectron) and focused at the detector. The MS^2 experiment can be viewed in stages; (1) a short linear time-of-flight analyzer (TOF1 section of the diagram), which generates the precursor ions, (2) the timed ion selector device (TIS), deceleration stack, collision cell and field free drift region prior to reacceleration (CID section of the diagram) and (3) a reflector time-of-flight analyzer (TOF2 section of the diagram) to provide the mass spectra of the fragment ions [21]. The velocity focus [18, 22] of the precursor ions
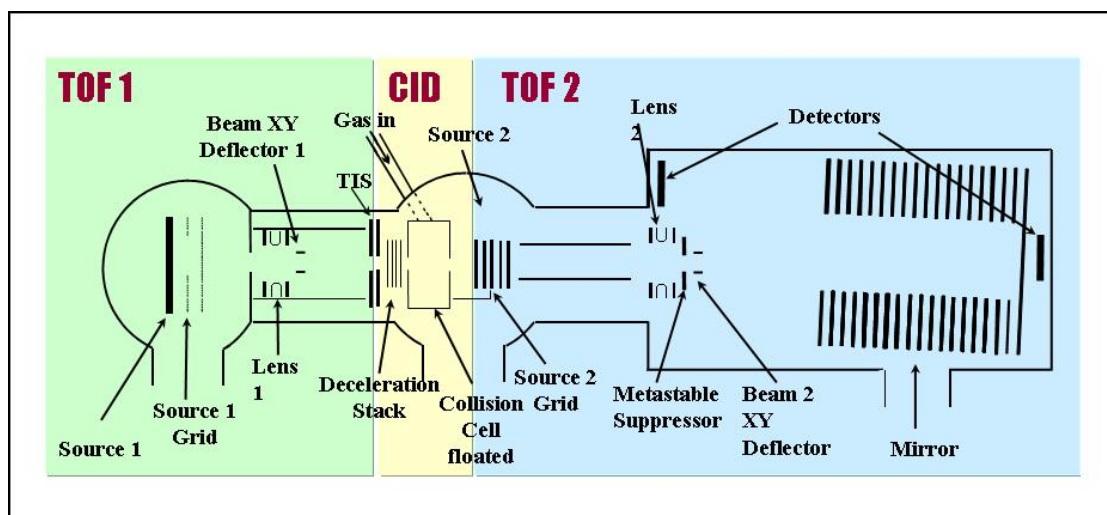
Fig.1. Schematic of the Applied Biosystems 4700 TOF/TOF^{tm} Proteomics Analyzer

is set at the center of the TIS, which consists of two sets of deflection plates, separated by 1cm. These plates are basically a door set to deflect all ions with m/z less than or greater

than the desired precursor m/ z. This provides a mass resolving power (m/Δm) of up to 200 at mass 1000, but is typically operated at m/Δm equal to 50 to prevent loss of sensitivity.

The collision cell (and field free drift region prior to reacceleration) is floated at 7keV making the laboratory frame collision energy 1keV. Consequently the selected precursor ions are decelerated substantially before entering the collision cell [21]. On exiting the collision cell, the remaining precursor and fragment ions have a velocity distribution that is independent of m/z to first order. This is of no use for fragment ion detection, so the m/z dependency on energy is reintroduced by reaccelerating these ions to 14keV in the second source, before reflecting and detecting them. This arrangement for MALDI-TOF-MS^2 is not the only means of achieving this goal and competing manufacturers and groups have alternate solutions to this problem [23].

1.2.4 Precursor ion selection in Time-of-flight MS^2

As was described in section 1.2.2.3.1 for the AB 4700, prior to MS^2 some sort of precursor selection is necessary. This is even more important when these types of analyses are undertaken with complex samples, with many possible analytes present in each sample i.e. proteomic samples from biological systems. Often as many as 10 analytes may be present in one sample well (an etched spot on the stainless steel MALDI plate on to which the analytes and matrix are placed). If these are of similar m/z, then a resolving power of 50 at 1000m/z (typical for the AB 4700) may not be sufficient to guarantee each recorded MS^2 spectrum is of just one analyte. Fragmenting more than one analyte at a time makes interpretation/ sequencing extremely difficult and violates one of

the assumptions built into each of the analysis software packages i.e. that the spectrum is

from one analyte.

1.3 Mass Spectrometry of protonated peptides

1.3.1 Peptides and Amino Acids

The mass spectrometric experiments described in this thesis were performed on peptides. Peptides are biological polymers made up of individual amino acid residues (monomers) linked via peptide bonds.
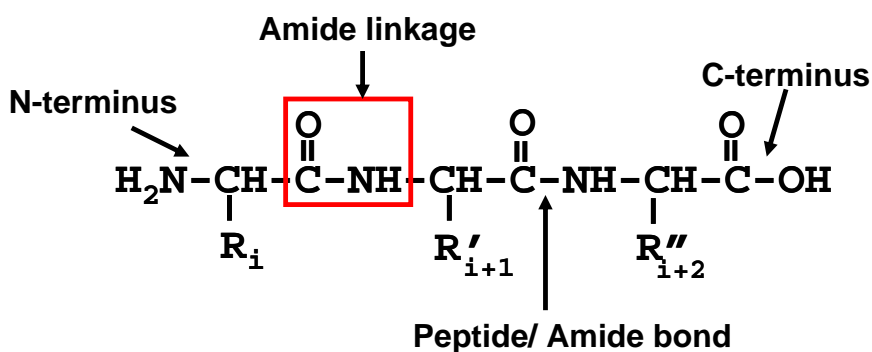
Figure 1.1: Tripeptide with peptide nomenclature indicated.

The different amino acids are defined by their R groups of which there are twenty naturally occurring in nature and are described by either three or one letter codes (see Appendix A). The types of R groups and therefore amino acids present in a protein or peptide determine its structure, reactivity and properties.

The majority of mass spectrometric sequencing of proteins is done by first chemically digesting the proteins of interest with an enzyme that cuts the protein at specific peptide bonds resulting in many smaller peptides. By far the most common enzyme in use is trypsin which cleaves at the C-terminus of lysine and arginine (except

when the next residue is proline), resulting in peptides with these basic peptides at the C-terminal end.

1.3.2 Fragmenting Protonated Peptides: Collision-Activated-Dissociation and Tandem Mass Spectrometry

Sequencing of protonated peptides is often achieved by means of fragmenting the peptide into diagnostic fragments. The most common method for fragmenting protonated peptides is collision-activated(induced)-dissociation (CAD or CID) [24, 25]. Gas-phase protonated peptides are isolated, activated by collision(s) with an inert gas, dissociated and detected. This is a tandem mass spectrometry experiment (MS/MS or MS^2) where the first mass spectrometry stage is used to isolate a single precursor ion which is then collided with an inert gas and the resulting fragment ions are detected and used to identify the precursor peptide. These experiments may be tandem in space or in time depending on the instrument configuration in use.

1.3.3 Energy Regimes in CAD

Tandem mass spectra are classified as either low or high energy. This describes the types of CAD processes involved in fragmenting the ions under study. The term 'low energy' refers to activation via multiple collisions (E_{LAB} < 300eV, more typically ~30eV) with an inert gas to induce fragmentation. This method allows a millisecond timescale and is performed predominantly in quadruple or ion trap instruments. 'High energy' refers to collisional activation (>5keV) [26] via one or two collisions on a microsecond timescale usually in magnetic sector or time-of-flight (TOF) instruments. It should also

be explicitly understood that the total energy transferred to the analyte ion in 'low energy' collisions may be equal to or greater than that from a 'high energy' collision despite the term 'low energy' activation (Appendix B). The distinction is often incorrectly used to describe the amount of energy transferred rather than the mechanism of activation. These mechanisms of activation result in characteristic spectra. 'High energy' activation conditions enable all of the product ions shown in Figure 1.2 to be potentially detected although not necessarily in every spectrum. Alternatively, 'low energy' activation generally only produces a subset of the total possible product ions.

1.3.4 Peptide Fragmentation Nomenclature

This thesis discusses the fragmentation characteristics and resulting species from peptide fragmentation; the Roepstorff nomenclature [27] modified by Biemann [28] will be used throughout.
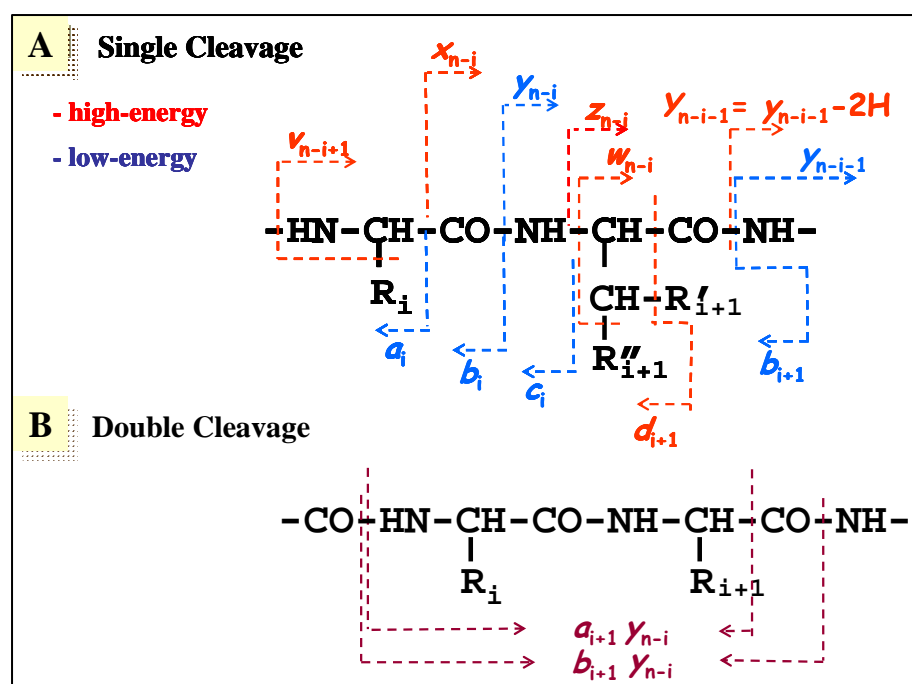
Figure 1.2: The nomenclature of peptide fragmentation.

Fragment ions containing the N-terminus are labeled a-c, or d while fragment ions containing the C-terminus are labeled x-z, v or w. Fragment ions with neither termini are the product of two backbone cleavages, one at each end of the protonated peptide ion and are termed internal ions. Internal ions contain two or more residues. Fragment ions with two backbone cleavages that result in internal ions containing only one amino acid, are termed immonium ions. These immonium ion peaks can be characteristic of a particular residue's presence within the peptide ion and is assigned the corresponding residue's letter (e.g. W for the tyrptophan immonium ion at 159 m/z).

1.3.5 Peptide sequencing and identification

In peptide identification, the most important ion peaks are the ones that define a run of sequence with minimum ambiguity. Sequencing algorithms are adjusted to reflect the expected spectra from the instrument used in the study as the ions produced by CID are strongly dependent on the instrument used as well as the sequence of the peptide, producing great variability in the mass and intensity of peaks in the spectra of a particular peptide [29]. The success of these computer-aided peptide sequencing approaches is directly related to the quality of the applied fragmentation models that summarize our present understanding of gas-phase ion peptide chemistry [30]. The difference in m/z of consecutive ions of a particular type (for example, b_n, b_{n-1}), corresponds to the m/z of an amino acid, so the longer or the more members of an ion series you have, the more likely you are to be able to identify the peptide under study. A good run of y ions, b ions, (a ions) or both, is preferable. These ion series are known as sequence ions and in CAD, b, y

and to a lesser extent a ions are by far the most common series. The remaining series (c, x, z) occasionally occur but this is analyte and fragmentation condition(s) specific.

Losses of small neutral molecules (NH_3, H_2O etc), are also observed from the protonated peptide under study, as well as the sequence and immonium ions. These losses are often peptide composition and/ or activation method specific [31]. Often these ions are a hindrance to peptide identification as they decrease the signal of the sequence ions and complicate the spectra observed.

The high energy d, v and w series usually do not result in differences of one amino acid as these ions involve cleavage of a side chain bond. Only when the same amino acid and therefore side chain is present in consecutive d, v, or w ions, will this be the case. These high energy ions have the major advantage of enabling the identification of isobaric amino acids in peptides. For example Leucine, L, and isoleucine, I, are isomers which are indistinguishable under low energy CAD as d, v, and w ions are not detected in low energy spectra.

Internal and immonium ions mainly provide confirmation that a particular sequence identified from sequence ions, is plausible. If the instrument configuration did not include them and they did occur in the instrument, they would be treated as noise peaks and would reduce the score or confidence in any assignments made.

1.3.6 Current sequencing approaches and the associated problems

There are several commercially available software packages that allow searching of protonated peptide fragmentation information (Mascot, SEQUEST, XProteo etc). In these packages candidate peptides and their theoretical MS/MS spectra are generated *in*

silico (by a computer) using protein and/or DNA databases and proprietary fragmentation models. The *in silico* spectra are then compared to the experimental MS/MS spectrum in order to find the most closely matching sequences. These can include a few post translational modifications (PTMs: changes to specific residues within a protein within its source environment e.g. phosphoylation or addition of a sugar molecule) too, if they are specifically expected. Each of these algorithms produce different results meaning that a particular sequence may be identified by some but not others of the algorithms as matching a particular spectrum. At present these algorithms completely ignore the ion intensity information and treat all peaks above the user defined threshold with the same weighting in their calculations.

Due to the limitations inherent in each of these software packages, it is sensible practice to use more than one to analyze protonated peptide fragment spectra; those spectra identified by more than one algorithm as having a particular sequence are thus more likely to be identified correctly.

In addition to the uncertainties in the chemistry prior to getting the digested peptides into the mass spectrometer, unidentified spectra may be due to; low signal to noise, non uniform fragmentation, too much fragmentation, the algorithm used to identify the protonated peptide spectra, unexpected (or included) PTMs or instrumental prejudices amongst other possibilities.

1.3.6.1 Current peptide fragmentation models: The Mobile Proton Model

The most common model currently used to describe the dissociation of protonated peptides upon low energy collisional activation is the 'mobile proton model' [32, 33] or

'heterogeneous population model' [34]. The majority of low energy fragmentation mechanisms involve charge-directed reactions. The 'mobile proton model' considers the relative stability and reactivity of the various protonation sites present within peptides which make up the possible precursor protonated peptide isomers. Naturally the protonation sites resulting in the most stable protonated peptide ions are initially the most likely to be populated. The mobile proton model states that energy is required to 'mobilize' the proton from a site of lower energy to a reactive protonation site i.e. for a charge-directed reaction to occur. The reactive protonation sites are usually the amide nitrogens; protonation of amide nitrogen leads to considerable weakening of the amide bond [35], but usually require more energy input than do the more stable N-terminal, amide oxygen and basic residue protonation sites. Amide nitrogen protonated species play a critical role in most of the peptide fragmentation pathways (PFPs) that lead to sequence-informative b, a, and y ions [30].

When using the mobile proton model to explain the dissociation of protonated peptides upon low energy collisional activation two scenarios need to be considered. Firstly, if no strongly basic residues (arginine, R, lysine, K, or histidine, H) are present in the protonated peptide, then several of the possible precursor isomers will be significantly populated. In this situation, there is a small energy range separating the precursor isomers from the reactive protonation sites. Thus, minimal activation of the protonated precursor isomers leads to formation of fragment ions remote from the initial site(s) of protonation as the mobilized proton initiates fragmentation. Alternatively, if a strongly basic residue is present, a single protonation site is usually significantly energetically and/or kinetically more favorable than any of the other protonation sites [30]. These protonated peptides

require a much larger amount of energy from collisional activation to mobilize the ionizing proton away from the basic protonation site to the less stable, reactive protonation sites than peptides that contain only non-basic residues. Generally, arginine containing peptides require more energy to dissociate than analogous lysine, then histidine and finally non-basic residue containing peptides [32, 33, 35]. This can lead to some charge-remote fragmentations being more favorable than proton mobilization in some cases. This model has been extended to include multiply charged peptides with multiple basic residues as well as residues like aspartic acid (aspartic acid effect) which promote charge-remote fragmentation [36]. If the number of arginines is greater than or equal to the number of ionizing protons, then selective charge-remote fragmentation can be expected. I shall only be discussing singly charged peptides in this thesis however.

The mobile proton model has found support from hydrogen/deuterium labeling experiments [37-40]. In singly deuterated peptides, the deuteron was found to have been redistributed [39] amongst the exchangeable protonation sites upon excitation to induce fragmentation of the precursor ion. Theoretical calculations [41-43] also support the mobile proton model where internal rotations and proton transfers have been modeled and the transition state determined. Approximate unimolecular reaction rates and timescales of reaction have been calculated for these processes leading to 'proton traffic maps' [41] for proton mobilization. These calculations showed that the ionizing proton could sample all of the protonation sites in dipeptide systems prior to fragmentation at internal energies well below the threshold for the lowest energy fragmentation pathway [30, 41, 43]. In summary, the mobile proton model enables prediction of whether a particular peptide in a particular state of protonation is likely to produce a sequence-

informative MS/MS spectrum. The mobile proton model is supported by theoretical and labeling studies. On the other hand, the mobile proton model says nothing about the cleavage or post-cleavage phases of peptide fragmentation [30] and does not explain everything seen routinely in peptide mass spectra. In order to have a more detailed understanding of why the spectra observed are that way and to enable more robust predictive capability, an improved model is necessary.

1.3.6.2 Other approaches

An alternative approach to peptide fragmentation prediction is based on systematic evaluation of large databases of MS/MS spectra of protonated peptides which have been identified by a specific commercial search algorithm (e.g. SEQUEST) [44-48]. In so doing, probabilities of each particular ion type occurring from each position within the peptide can be obtained. The predicted relative intensity of such peaks can also be obtained in this way. This approach has the advantage of giving a global view of protonated peptide fragmentation results, but is limited by its reliance on a particular search algorithm to initially screen the peptides. This means that whatever prejudices are present in the initial screening algorithm will be present in the probabilities generated. The probabilities calculated are also a function of instrument type, ionization method and collision energy so can't be taken as universal by any means. Protonated peptide m/z limitations in some of the studies also apply. Despite these current limitations, this technique offers a huge amount of useful information from which subsequent work can be based.

A kinetic model has recently been proposed to predict ion intensities for protonated peptide fragmentation in ion trap instruments [49]. This is based on a mixture of published fragmentation mechanisms, the mobile proton model, the author's own assumptions and training of his algorithm. Despite some dubious assumptions and the model being built based only for one company's specific mass spectrometer, the model did have some success. It is clear however that a lot more needs to be done to accurately predict the presence of and relative signal intensities in protonated peptide tandem MS.

1.3.6.3 The Pathways in Competition Model

The 'pathways in competition' (PIC) fragmentation model [30], provides a more general framework than the mobile proton model and takes into account events occurring before, during and after the bond cleavage phase. This enables the specific features of individual peptide fragmentation pathways (PFPs) and their interaction to be incorporated albeit at the cost of simplicity and time. The PIC model is a logical extension to the mobile proton model and classifies the PFPs according to a hierarchy (Figure 1.3).
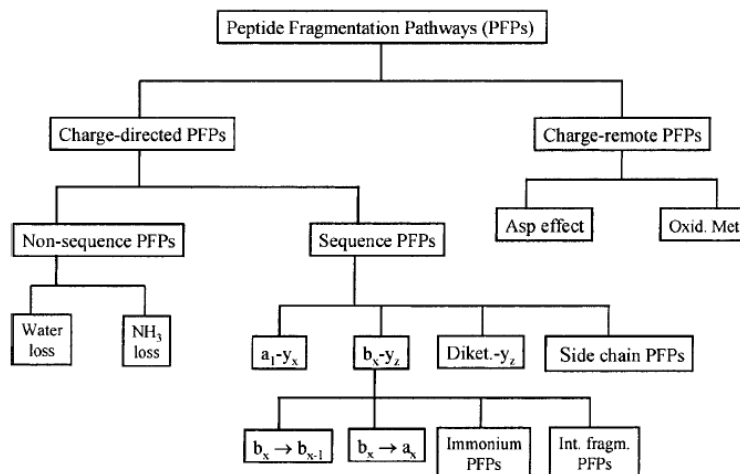
Figure 1.3. The peptide fragmentation hierarchy of the PIC model [30]

The majority of protonated peptide fragmentation reactions occur via charge-dirrected pathways. These reactions can then be further classified into sequence or non sequence ion forming reactions. As mentioned earlier, the useful b and y ions are formed following proton transfer(s) from energetically favorable protonation sites to amide nitrogens. This weakens the amide bond and makes the carbon atom of the protonated amide group a likely target for nucleophilic attack of nearby electron rich groups. With low energy fragmentation conditions, the majority of the amide bonds cleaved are due to complex rearrangement reactions rather than the more energetically demanding direct bond cleavage (a_1-y_m type reactions). This leads to several potential pathways for the formation of b and y ions. Once these ions have been formed they can subsequently react to form lower m/z b, a, immonium or internal ions if sufficient time and/or energy are available.

The energetics and kinetics of the competing PFPs determine which chemical bonds are cleaved upon excitation and, therefore, what kind of fragments appear as charged or neutral species in the mass spectrometer. Under low energy conditions, the thermodynamic product is usually formed; the fragment with the larger proton affinity, PA, or smaller gas-phase basicity, GB, will usually keep the added proton.

For the generic reaction; $B(g) + H^+(g) => BH^+(g)$, the GB is the negative of the free energy change of reaction, $-\Delta G°$, while the PA is the negative of the enthalpy change of reaction, $-\Delta H°$ [64]. $PA = GB - T\Delta S°$ where T is temperature and $\Delta S°$ is the entropy change of reaction.

Based on this rationale, simple free-energy relationships [51] can be used to predict the relative fragment ion intensities [30, 52] of some PFPs. The thermodynamic product is of course not always the product formed. A more complete understanding of

the mass spectra of protonated peptides requires an understanding of the kinetics involved in the complex reaction pattern. The PIC approach seeks to incorporate this post-cleavage phase. This phase of peptide fragmentation can feature rich chemistry including PFPs that lead to scrambling of primary sequence information [53] and rearrangements via fragment re-association in post-cleavage proton-bound dimers [31, 54]. The combination of the pre-dissociation (mobile proton), dissociation and also the post-dissociation events present in PFPs are described by the PIC model. The long term goal being semi-quantitative prediction of these three phases involved in protonated peptide fragmentation leading to prediction of the ion intensity relationships in the protonated peptide spectra.

1.4 Computational Chemistry and Modeling

1.4.1 Context

In 1929 Paul Dirac remarked that:

"The fundamental laws necessary for the mathematical treatment of a large part of physics and the whole of chemistry are thus completely known, and the difficulty lies only in the fact that application of these laws leads to equations that are too complex to be solved."

While the immense mathematical problem of exact quantum mechanical solution of all but the simplest systems is still present, the last 78 years have seen technological advances that render the later part of this statement, at least partially, inaccurate. The development of computers capable of performing calculations at hitherto unheralded speeds combined with increasingly sophisticated (i.e. faster and more efficient) numerical methods have transformed computational chemistry and physics into an indispensable predictive tool, a companion to experimentation and in some cases even replacement of experimentation e.g. initial evaluation of drug-substrate docking capability. Some of these advances were recognized by the 1998 Nobel Prize in Chemistry awarded to John A. Pople "for his development of computational methods in quantum chemistry" (developed the GAUSSIAN set of programs) and Walter Kohn for the "development of the density functional theory" [55].

The mathematical procedures developed seek to provide assistance in the qualitative interpretation of chemical phenomena and provide predictive capability [56]. In some smaller systems, higher level calculations can provide results as accurate and precise as experimental work. 'Chemical accuracy' is generally defined as 1-2kcal mol^{-1} ($\sim$4-8kJ mol^{-1}) [56-60] and current development is aimed at achieving this quality of result more frequently. The accuracy of computations when compared to experimental results is a function of the approximations used (i.e. the computational modeling method and basis-set), the species under study, and the time available for the calculations. The larger the system under investigation, the more time consuming (expensive) the calculations become. Generally, more complex and accurate methods require more time, so a balance has to be struck between result-quality and computational expense.

A second issue is whether the researcher is interested in absolute or relative values of calculated parameters. Some methods, while not producing the exact experimental value for a calculated parameter, do give the correct relative values between related species. For many applications this is nearly as important as having the calculated values match the experimentally observed ones.

1.4.2 The Potential Energy Surface

The potential energy surface (PES) or hypersurface is central to computational chemistry as it offers a means of intuitively visualizing mathematically complex concepts [61-64]. The PES utilizes the Born-Oppenheimer approximation [65], which allows separation of the nuclear and electronic components of the internal energy [66] with the additional consequence that molecules have shape resulting from the average positions of

their nuclei. The nuclei, of course, are in constant motion vibrating about equilibrium positions and 'see' the electrons as a smeared-out cloud of negative charge that binds the nuclei in "fixed" relative positions. Variation of the nuclear coordinates of a molecule should, as such, be accompanied by variation in the energy since the potential energy is a function of the position of the atoms. The PES is thus a construct of one's imagination; a topographical map describing the terrain on which the reactant molecules must move en route to a transition state and then onward toward a new geometrical arrangement or the product molecule(s) [67]. This topographical map can be drawn for simple di- and tri-atomics but more complex systems require more than three dimensions and cannot be accurately drawn. This leads to the terminology hypersurface that is sometimes used (potential energy hypersurface). The hypersurface can be defined by an equation $E = f(q_1,q_2, \ldots q_N)$, where f is the function that describes how E varies as a function of the q's (nuclei positions). This allows treatment of systems with any number of atoms.

1.4.3 Transition States and other PES features

The location of a particular structure on the PES determines whether this site is a minimum, maximum, a transition state, or a structure located on a slope. This, in turn, determines the properties of the structure with respect to change of geometric parameters. Stationary points are ones where:

$$\partial E/\partial q_i = 0$$

and

$$\partial^2 E/\partial q_i^2 > 0$$

these have finite lifetimes and are the energy minima each occupying the lowest energy point in their region of the PES [64]. The lowest energy minimum of the whole PES is the global minimum (GM). Energy values are often reported relative to this minimum energy. Transition states are saddle points on the PES, where:

$$\partial^2 E/\partial q_i^2 > 0$$

for all q_i except along the reaction coordinate (q_{RC}) where,

$$\partial^2 E/\partial q_{RC}^2 < 0$$

corresponding to the negative eigenvector (see below).

The minima and transition states on the PES can be calculated and corrected for zero-point energies to increase the accuracy of the PES. The level of calculation used defines the PES produced. Once a TS has been located, it should be verified that it connects the desired minima. The TS should have one imaginary frequency corresponding to the reaction coordinate and inspection of the corresponding motions using a visualization package such as GausView often offers a strong indication that it is a "correct" TS [60]. Proof of this requires calculation of the minimum energy path from the TS to the connecting minima. The calculation of the minimum energy path in mass-weighted coordinates is termed the Intrinsic Reaction Coordinate (IRC) [68]. This method calculates the minimum energy path connecting the reactants and the products by taking small steps along the negative gradient (one calculation towards products, one towards reactants) [68-70]. As these structures (close to the TS) are often a long way structurally from the reacting and product configurations, very small step sizes must be used to ensure the correct path is followed.

1.4.4 Geometry Optimization

Geometry optimization is the method by which stationary points are located on a PES. The larger the system of interest is, the larger the number of stationary points (minima, transition states, higher order saddle points) possible. These optimizations begin with an input structure; a set of coordinates describing the relative positions and identity of atoms within a system under investigation. From there an algorithm systematically changes the geometry of the input structure until it has located a stationary point. The more closely the input structure matches the stationary point, the quicker the optimization usually is. The curvature at the stationary point is then used to define what type of stationary point has been located. As such, the minimum located is directly affected by the input geometry, thus for more complex species many different input structures must be used if realistic description of the PES or the GM is desired.

1.4.5 Computational Strategy

Our strategy is to survey the PES of protonated peptides and calculate the transition and product structures and their energies. In doing this, the aim is to obtain as realistic a description (meaning in agreement with experiment) of the relative energies and rates of transitions using the least expensive method of calculation as possible. The reason for this minimalist strategy is to allow as large and/ or complicated a system as possible to be examined within a reasonable amount of time.

To begin a survey of the PES of a protonated peptide, a molecular mechanics (MM) conformational search engine developed in Heidelberg by my collaborator, Dr. Béla Paizs, is utilized [71]. This search engine generates a large number of candidate

structures for each of the potential protonation sites in the peptide under study, and runs a brief optimization search for local minima. Molecular mechanics treats all of the atoms as balls, and all bonds in the ion as springs. This method completely ignores the electrons and works on the basis of adjusting the balls and springs within some background "forcefield". An AMBER (Assisted model building with energy refinement) field modified to allow protonated peptides is used. This method, which is extremely fast and objective, allows wide sampling of the PES and eliminates the very labor intensive generation of candidate structures via direct coding or use of a molecular viewing program such as GausView or HyperChem. The generated input geometries are parsed into families of conformers grouped via bond angles and named accordingly. The input geometries are then optimized at the Hartree-Fock (HF) 3-21g level (for a discussion of basis sets and model chemistries, see section 1.4.6). The resulting geometries are run through an algorithm that removes degenerate geometries (and therefore energies) to prevent unnecessary duplication of computational effort. The output HF 3-21g geometries are then optimized using the B3LYP/6-31g(d) method. The resulting B3LYP/6-31g(d) geometries are also run through an algorithm to remove degenerate geometries and then optimized at the B3LYP/6-31+g(d,p) level of theory. Thus, fewer structures are optimized at each increasingly expensive level of theory, and the PES is defined at each level of theory.

Relative energies were calculated using the B3LYP/6-31+g(d,p) total energies and zero point energies vibrational frequencies were determined at the B3LYP/6-31g(d) level. This surveying technique has been shown to sample the PES as well or better than those

used by other groups of investigators [43, 52, 72, 73] giving equivalent or lower energy structures with a given level of theory.

After sampling the PES, transition state (TS) geometries are then calculated at the B3LYP/6-31g(d) and B3LYP/6-31+g(d,p) levels of theory. All TSs are checked carefully, so have one negative eigenvector corresponding to the reaction coordinate. The TSs were then also checked using IRC calculations to define which minima are connected by the TS investigated.

1.4.6 Density Functional Theory and the B3LYP model

1.4.6.1 The Hohenberg-Kohn Theorems

Density functional theory (DFT) is based on the electron density, $\rho(r)$, rather than the wavefunction, Ψ, used in quantum mechanics. A functional is a rule that transforms a function into a number. So if the function $f(x) = x^2$, the functional $F[f(x)]$

$$F[f(x)] = \int_0^2 f(x)dx = 8/3$$

In DFT, functionals of the ground state electron density are used to calculate ground state properties of the system under study. The first work on this theory produced the Hohenberg-Kohn (HK) and Kohn-Sham (KS) theorems [74, 75]. The first HK theorem states that *'the external potential* $V_{ext}(r)$ *is a unique functional of $\rho(r)$; since, in turn* $V_{ext}(r)$ *fixes* $\hat{H}$ *we see that the full many particle ground state is a unique functional of $\rho(r)$'.* i.e. $V_{ext}(r)$ is unique to the electron density of the ground state ρ_0 of the system under study and is fully defined by nuclei-electron (Ne) attraction. Thus, all ground state properties (e.g. energy, E_0) of a system defined by an external potential V_{ext} are determined by a functional F of the ground state electron density function (ρ_0).

$$E_0 = F[\rho_0] = E_0[\rho_0]$$

The components of the ground state energy are also functionals of the electron density.

$$E_0[\rho_0] = T[\rho_0] + E_{Ne}[\rho_0] + E_{ee}[\rho_0]$$

Separating the potential energy resulting from nuclei-electron (Ne) attraction from the other terms gives

$$E_0[\rho_0] = T[\rho_0] + E_{ee}[\rho_0] + \int\rho_0(\acute{r})V_{ext}d\,\acute{r} = F_{HK}[\rho_0] + \int\rho_0(\acute{r})V_{ext}d\,\acute{r}$$

which defines the Hohenberg-Kohn functional $F_{HK}[\rho_0]$, which contains the kinetic energy $T[\rho_0]$ and electron-electron repulsion $E_{ee}[\rho_0]$ terms and is completely independent of the system. Unfortunately, the explicit form of these functionals is not known. If they were known, the Schrödinger Equation could be solved exactly. There is some relief, however, in the fact that the classical Coulomb part $J[\rho]$ of the electron-electron interaction is known, but this still leaves the non-classical contribution $E_{ncl}[\rho]$ comprising the effects of self-interaction correction, exchange, and Coulomb correlation. Finding explicit expressions for $T[\rho]$ and $E_{ncl}[\rho]$ is the major challenge in density functional theory.

The second HK theorem states that any trial electron density (ρ_{trial}) will give a higher energy than the true ground state electron density.

$$E[\rho_{trial}] = T[\rho_{trial}] + E_{Ne}[\rho_{trial}] + E_{ee}[\rho_{trial}] \geq E_0\,[\rho_0]$$

Any trial density defines its own Hamiltonian $\hat{H}$, and consequently its own wavefunction ψ_{trial}. The wavefunction can now be used as the trial wavefunction for the Hamiltonian generated from the true external potential [58],

$$<\psi_{trial}|\hat{H}|\psi_{trial}> = T[\rho_{trial}] + V_{ee}[\rho_{trial}] + [\int\rho_{trial}(\acute{r})V_{ext}d\,\acute{r}] = E[\rho_{trial}] \geq <\psi_0|\hat{H}|\psi_0>$$

giving the variational result.

1.4.6.2 The Kohn-Sham Method

Initial attempts to calculate the kinetic energy using DFT produced results that were much worse than those obtained using wavefunction methods. Kohn and Sham realized that if the explicit functional for kinetic energy was not accurate, then an alternate means of computing some of the kinetic energy term exactly was required. This was achieved by calculating the non-interacting kinetic energy (no electron-electron interactions), T_S, of a system moving in an effective potential V_S with identical electron density as the real, interacting system and then dealing with the remainder in an approximate manner

$$T_S = -\tfrac{1}{2} \sum_{1}^{N} <\varphi_i| \nabla^2 | \varphi_i >$$

where φ_i are the spin orbitals. T_S is $\leq T$ of the interacting system and this is accounted for by

$$F[\rho(\acute{r})] = T_S[\rho(\acute{r})] + J[\rho(\acute{r})] + E_{XC}[\rho(\acute{r})]$$

where $J[\rho(\acute{r})]$ is the classical Coulomb part of the electron-electron interaction and E_{XC}, is the exchange-correlation energy defined as

$$E_{XC}[\rho(\acute{r})] = (T[\rho] - T_S[\rho]) + (E_{ee}[\rho] - J[\rho]) = T_C[\rho] + E_{ncl}[\rho]$$

The remaining part of the true kinetic energy, T_C, is added to the non-classical electrostatic contributions. The exchange-correlation energy E_{XC} is the functional that describes all the components that are unknown. E_{XC} contains the non-classical effects of self interaction correction, exchange, and correlation, which are contributions to the potential energy of the system as well as a portion of the kinetic energy.

The variation principle is now applied to minimize the energy expression under the usual constraint $<\varphi_i| \varphi_j > = \delta_{ij}$ resulting in [90]

$$\{-\tfrac{1}{2}\nabla^2 + V_{eff}(\acute{r})\}\ \varphi_i = \varepsilon_i\ \varphi_i$$

$V_{eff}(\acute{r})$ contains V_{XC}, the potential due to the exchange-correlation energy E_{XC} for which no exact form exists. If an exact form did exist the Kohn-Sham method would lead to the exact energy i.e. the correct eigenvalue of the Hamiltonian operator $\hat{H}$ of the Schrödinger equation. The approximations come in creating an explicit form of E_{XC} and V_{XC} and this is where DFT development is currently centered. The B3LYP model used in my work is discussed next.

1.4.6.3 The B3LYP model

The B3LYP [77] model is a hybrid functional used to describe the exchange-correlation energy, E_{XC} (recall that the remaining components (functionals) can be calculated exactly). Hybrid functionals incorporate some exact (Hartree-Fock) exchange along with pure density functionals for exchange. Hybrid functionals perform significantly better [78] than pure DFT functionals which perform poorly for proton transfer reactions (like those in this thesis), in that the calculated DFT barriers are much less [79] than the corresponding MP2 (second order Møller-Plesset perturbation theory) values.

The B3LYP functional is comprised of the Becke (B) exchange functional [80], E_x^B and the Lee, Yang and Parr (LYP) correlation functional [81], E_c^{LYP}, with three semiempirical coefficients used to determine the weighting of the various components

$$E_{xc} = (1\text{-}a)E_x^{LSDA} + aE_{XC}^{\lambda=0} + bE_x^B + (1\text{-}c)E_c^{LSDA} + cE_c^{LYP}$$

where $E_{XC}^{\lambda=0}$ describes the pure exchange functional, which ignores electron correlation entirely. $E_{XC}^{\lambda=1}$ would correspond to a fully interacting system, with exact exchange as

used in the Hartree-Fock method. E_x^{LSDA} and E_c^{LSDA} are the local spin-density approximation functionals for exchange and correlation respectively. The remaining parameters (a,b,c) were determined from regressions [77] using standard reference compounds for which experimental values were accurately known.

1.4.6.4 Rigorous theory versus pragmatism: The pros and cons of using DFT

DFT incorporates electron correlation into its theoretical framework rather than as an add-on, which occurs in many quantum and semi-empirical methods. This feature enables DFT to calculate geometries and relative energies with accuracy roughly equivalent to MP2 calculations in roughly the same amount of time required as for a simple HF calculation. DFT calculations benefit from being basis-set-saturated more readily than ab initio calculations. This enables smaller basis sets (see next section) to be used to get a given quality of result thereby saving time. DFT also scales favorably, meaning that the length of calculation increases less steeply with number of atoms for a given basis set ($\sim N^3$ for DFT vs. $\sim N^4$ for MP2). This enables results to be obtained for larger systems, like those dealt with in this thesis, in a much more reasonable amount of time [64]. DFT gives good energies and structures for hydrogen-bonded species [82-84], which is of particular relevance to protonated peptide ions.

On the other hand, the exact form of the exchange-correlation functional $E_{XC}[\rho_0]$ that should be used in the DFT expression is unknown and, thus far, the approximate form is not systematically improvable unlike in ab initio methods. The accuracy of DFT is generally improved by comparing the results with experiment and then modifying functionals accordingly, which makes DFT somewhat semiempirical. The functionals

used today are only approximate meaning that DFT is not variational unlike the original theorems meaning that the energies calculated can in theory be less than E_0. DFT is also mainly a ground-state theory making it unsuitable for calculations on electronically excited species. As with any modeling technique one should be cautious when applying it to vastly different systems or molecules than have been examined previously. For example, early DFT functionals are used a lot in solid state work, where DFT first became popular, but these functionals fail miserably for protonated peptide systems. They are just not realistic, so claims made using these functionals for the systems described in this thesis would very likely be incorrect. Fortunately, the B3LYP model in conjunction with the basis sets (see 1.4.7) used in this thesis offer a good compromise between accuracy and computational time.

In summary, for many purposes, "at present, DFT stands unrivalled as a practical, reasonably high-quality quantum chemical tool, and appropriately so" [85].

1.4.7 A brief guide to Basis Sets

A basis set is a set of mathematical functions (basis functions), linear combinations of which yield molecular orbitals [64]. These functions are usually centered on atomic nuclei in the system under study and describe the electron distribution around each atom (atomic wavefunctions). Most modern molecular ab initio programs use Gaussian functions ($\exp[-\alpha \mathbf{r}^2]$) to approximate the electron distribution (in the GAUSSIAN program and others) as these are much easier to calculate two-electron integrals for than Slater functions ($\exp[-\alpha \mathbf{r}]$). This results in shorter calculation times.

In GAUSSIAN, the model in use is listed first followed by the basis set used with that model. For example HF 3-21g corresponds to the restricted Hartree Fock model with a single contracted (i.e. large α, so large r terms equal zero) three-Gaussian basis function for the inner shell and a set of inner two-contracted along with a set of outer uncontracted (i.e. smaller α, so large r terms are nonzero) Gaussians for the valence shell of each atom. This is a so-called split valence basis.

3-21g is a fairly crude approximation of the electron distribution around each atom. Generally larger basis sets provide better approximations. I use larger basis sets in my other calculations; 6-31g(d) and 6-31+g(d,p) (or 6-31g* and 6-31+g** respectively). These consist of a single contracted six-Gaussian basis function for the inner shell and a set of inner three-contracted along with a set of outer uncontracted Gaussians for the valence shell of each atom [56]. In addition, 6-31g(d) incorporates polarization functions denoted (d) or * on all atoms beyond helium corresponding to a single set of uncontracted d-functions. The 6-31+g(d,p) basis set has d and p polarization functions and diffuse functions on non hydrogen atoms (denoted +). The p polarization functions are a single set of uncontracted p-functions which are used on each hydrogen atom. The diffuse functions consist of Gaussian functions with small values of α causing the function to fall off slowly with the distance r from the nucleus. These allow better descriptions of lone pair electrons, electronically excited electrons, and other relatively loosely held electrons that are on average at a larger distance from the nuclei than core or bonding electrons [56]. Diffuse functions on oxygen and nitrogen atoms are mandatory to obtain reliable energetics of hydrogen-bonded systems. However, I do not use a set of diffuse functions on the hydrogen atoms (this would be a 6-31++g(d,p) basis set) in my work as there is no

real physical interaction that should be described by the diffuse function on hydrogen atoms and with my method of calculation, these diffuse functions result in more pronounced basis set superposition error (BSSE). This leads to a spurious lowering of the total energy of the complex with respect to its fragments and thus to an artificial overestimation of the complexation energy [58, 86] in our hydrogen-bonded systems. In addition, calculations at this level are significantly more expensive.

1.4.8 Rate of reaction modeling

1.4.8.1 Rice-Ramsperger-Kassel-Markus (RRKM) Theory

The concepts for understanding unimolecular dissociation reactions from activated species were developed initially by Rice and Ramsperger [87, 88] and Kassel [89]. The system is treated as a group of s identical harmonic oscillators, one of which is truncated at the activation energy E_o for dissociation. If this critical oscillator has an energy greater than E_o the molecule dissociates. A fundamental assumption being that coupling between the various vibrational degrees of freedom is sufficiently strong for the excitation energy to be randomized rapidly amongst the active degrees of freedom on the timescale of reaction [90]. Thus the chance of finding the system with a particular arrangement of its internal energy is equivalent to any other [62]. The rate of dissociation is proportional to the frequency factor, ν, and the probability that for a particular energy E, the critical oscillator contains energy equal to or greater than E_o. In terms of energy this gives

$$k(E) = \nu[(E - E_o)/E]^{S-1}$$

This is known as the RRK expression and it predicts that rate is a strong function of the number of oscillators (s) and that the rate increases rapidly with excess energy, (E - E_o). Unfortunately this equation has major limitations. As the number of oscillators is often similar to rather than much less than the number of quanta, this function rises much too steeply. The function also ignores the zero point energy.

These problems were independently eliminated by Markus and Rice [60] and Rosenstock et al [91] giving the RRKM expression:

$$k(E) = \sigma N^{\ddagger}(E - E_o)/h\rho(E)$$

where $\rho(E)$ is the density of vibrational states of the reactive configuration at the energy E, $N^{\ddagger}(E - E_o)$ is the sum of the vibrational states from 0 to $E-E_o$ in the transition state, h is Planck's constant, and σ is the reaction symmetry factor.

1.4.8.2 Calculation method

The Beyer-Swinehart direct count algorithm [92] is used for rotational-vibrational treatment of both the reactant and the transition structure. This algorithm uses the 3N-6 harmonic vibrational frequencies of the reactive configuration and the 3N-7 frequencies of the TS and calculates how many ways a given amount of energy can be distributed among these modes. By summing over all such distributions for energy varying from zero to E, the algorithm determines the sum of the vibrational states. By taking the difference between the sum of the vibrational states (of the reactive configuration) for different energies, it determines $\rho(E)$ as a function of energy [93].

This is the most accurate method for determining the density of harmonic vibrational states [62] and is necessary as we are examining medium-sized systems where

the corresponding geometries can be quite different. The same method of calculation is used for all types of transition and no tunneling correction is included. This is because in most cases the rates are very high as soon as the energy is above the reaction threshold and the accuracy of our TSs is only a few kcal/mol, making a tunneling correction unreasonable [41]. The rate constant plots give a qualitative guide to the energy dependence of the various reactions (see below).

It should be explicitly understood that these RRKM calculations rely on the assumption of a harmonic force field applied in the calculation of the frequencies [41, 94], sums of states, and densities of states. The anharmonicity of the modes involved will in reality be significant, as will the coupling between them, making this approximation rather crude [41]. Detailed calculation of these frequencies, sums and densities of states is impractical due to the large amount of time and computing resources necessarily involved with even small peptide ions. Prior studies [41] using this methodology have indicated that the rate coefficients calculated are high relative to the inverse timescale of MS experiments. Thus, errors of 2 or 3 orders of magnitude would still result in the same overall conclusions. i.e. that protons are 'mobilized' under MS conditions allowing proton transfer reactions to make peptide ion fragmentation pathways available. This is in agreement with the experimental finding that different types of fragment ions are observed from the same nominal precursor peptide ion.

1.5 References

1) Borman, S. *Chem. Eng. News*, May 26, **1998**.

2) Nobel website http://nobelprize.org/

3) Barber, M.; Bordoli, R.A.; Sedgewick, R. D.; Tyler, A. N., *J. Chem. Soc., Chem. Commun*. **1981**, 7, 325-327.

4) Yamashita, M.; Fenn, J. B.; *J. Phys. Chem*. **1984**, 88, 4451-4459.

5) Karas, M.; Bachmann, D.; Bahr, U.; Hillenkamp, F. *Int. J. Mass Spectrom. Ion Proc.*, **1987**, 78, 53-68.

6) Karas, M.; Hillenkamp, F. *Anal. Chem.*, **1988**, 60, 2299-2301.

7) Tanaka, K. H.; Wake, H.; Ido, Y.; Akita, S.; Yoshida, Y.; Yoshida, I. *Rapid Commun. Mass Spectrom*. **1988**, 8, 2.

8) Barber, M.; Bordoli, R. S.; Elliot, G. J.; Sedgewick, R. D.;   Tyler, A. N. *Anal. Chem*., **1981**, 54, 645A.

9) Spengler, B.; Bahr, U.; Karas, M.; Hillenkamp, F. *Anal. Instrum*. **1988**; 17, 173.

10) Vestal, M. L.; Juhasz, P. ; Martin, S.A. *Rapid Commun. Mass Spectrom*. 1995; 9**, 1044- .

11) Brown, R.S.; Lennon, J.J. *Anal. Chem*. **1995**; 67: 1998-.

12) Karas, M. *J. Mass Spectrom*. 2000, **35**, 1–12.

13) Fenn, J.B.; Mann, M.; Meng, C.K.; Wong, S.F.; Whitehouse, C.M.; *Mass Spectrom. Rev*., **1990**, 9, 37-70.

14) Cooks, R.G.; Glish, G.; McLuckey, S.A.; Kaiser, R.E. *Chem. Eng. News*, March 25, **1991**.

15) Polce, M. J.; Cordero, M. M.; Wesdemiotis, C.; Bott, P. A. *Int. J. Mass Spectrom. Ion Proc.,* **1992**, 113, 35-58.

16) Stephens, W.E. *Phys. Rev.* 1946, **69**, 691.

17) Cotter, R.J., *Anal. Chem*., **1999**, 445A-451A.

18) Wiley, W.C.; McLaren, I.H. *Rev. Sci. Instrum*. **1955**, 26, 1150-1157.

19) Mamyrin, B.A.; Karataev, V.I.; Shmikk, D.V.; Zagulin, V.A. *Sov. Phys. JETP*, **1973**, 37, 45.

20) Cotter, R.J., Time-of-Flight Mass Spectrometry, **1997**, ACS, Washington.

21) Juhasz, P.; Campbell, J. M.; Vestal, M. L. MALDI-TOF/TOF technology for peptide sequencing and protein identification in *Mass Spectrometry and Hyphenated Techniques in Neuropeptide Research*, Edited by Silberring, J.; Ekman, R., **2002**, 375-413, Wiley-Interscience, 1st Edition

22) Vestal, M. L.; Juhasz, P. *J. Am. Soc. Mass Spectrom*., **1998**, 9, 892-.

23) Cotter, R. J.; Gardner, B. D.; Iltchenko, S.; English, R. D. *Anal. Chem*. **2004,** 76, 1976-1981.

24) Jennings, K. R. *Int. J. Mass Spectrom. Ion Phys*. **1968**, 2, 227.

25) McLafferty, F. W.; Schuddemage, H. D. R. *J. Am. Chem. Soc*., **1969**, 91, 1866.

26) Martin, S. A.; Johnson, R. S.; Costello, C. E.; Biemann, K., The Structure Determination of Peptides by Tandem Mass Spectrometry. In *Analysis of Peptides and Proteins by MassSpectrometry*; McNeal, C. J., Ed.; John Wiley and Sons: Chichester, England, **1988**, 135-150.

27) Roepstorff, P.; Fohlman, J. *Biomed. Mass Spectrom.* **1984**, 11(11), 601.

28) a)Biemann, K. *Methods Enzymol*., 1990, 193, 455-479. b) 351-360.

29) Khatun, J.; Ramkissoon, K.; Giddings, M.G.; *Anal. Chem*. **2007**, 79, 3032-3040.

30) Paizs, B.; Suhai, S. *Mass Spectrom. Rev*., **2005**, 24, 508-548.

31) Bythell, B. J.; Barofsky, D. F., Pingitore, F.; Polce, M. J., Wang, P., Wesdemiotis, C.; Paizs, B. *J. Am. Soc. Mass Spectrom.*, **2007**, 7, 1291-1303.

32) Dongre, A. R.; Jones, J. L.; Somogyi, A.; Wysocki, V. H., *J. Am. Chem. Soc.* **1996**, *118*, 8365-8374.

33) Wysocki, V. H.; Tsaprailis, G.; Smith, L. L.; Breci, L. A. *J. Mass Spectrom.,* **2000**, 35, 1399-1406.

34) Burlet, O.; Yang, C.Y.; Gaskell, S. J. *J. Am. Soc. Mass Spectrom.*, **1992**, 3, 337–344.

35) Somogyi, Á.; Wysocki, V.H.; Mayer, I. *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 704-717.

36) Tsaprailis, G.; Nair, H.; Somogyi, A.; Wysocki, V. H.; Zhong, W.; Futrell. J. H.; Summerfield, S. G.; Gaskell, S. J. *J. Am. Chem. Soc*., **1999**, 121, 5142–5154.

37) Tsang CW, Harrison AG. *J. Am. Chem. Soc*., **1978**, 98, 1301–1308.

38) Mueller, D.R.; Eckersley, M.; Richter, W. *Org. Mass Spectrom.* **1988**, 23, 217–222.

39) Johnson, R. S.; Krylov, D.; Walsh, K. A. *J. Mass Spectrom.,* **1995**, 30, 386–387.

40) Harrison, A.G.; Yalcin, T. *Int. J. Mass Spectrom. Ion Processes,* **1997**, 165, 339–347.

41) Csonka, I. P.; Paizs, B.; Lendvay, G.; Suhai, S. *Rapid Commun. Mass Spectrom*., **2000**, 14, 417–431.

42) Csonka, I. P.; Paizs, B.; Lendvay, G.; Suhai, S. *Rapid Commun. Mass Spectrom*., **2001**. 15, 1457–1472.

43) Paizs, B.; Suhai, S. *Rapid Commun. Mass Spectrom*., **2001**, 15, 651–663.

44) Huang, Y.; Wysocki, V.H.; Tabb, D. L.; Yates, J. R. *Int. J. Mass Spectrom. Ion Processes,* 2002, 219, 233-244.

45) Kapp, E. A.; Schutz, F.; Reid, G. E.; Eddes, J. S.; Moritz, R. L.; O'Hair, R. A. J.; Speed, T. P.; Simpson, R. J. *Anal. Chem,* 2003, 75, 6251–6264.

46) Huang, Y.; Triscari, J. M.; Pasa-Tolic, L.; Anderson, G. A.; Lipton, M. S.; Smith, R. D.; Wysocki, V. H. *J. Am. Chem. Soc.,* 2004, 126, 3034-3035.

47) Huang, Y.; Triscari, J. M.; Pasa-Tolic, L.; Anderson, G. A.; Lipton, M. S.; Smith, R. D.; Wysocki, V. H. *Anal. Chem.,* 2005, 77, 5800-5813.

48) Tabb, D. L.; Huang, Y.; Wysocki, V. H.; Yates, J. R., III, *Anal. Chem.* **2004**, 76, 1243-1248.

49) Zhang, Z. *Anal. Chem.*, **2004**, 76, 6374-6383.

50) Harrison, A.G. *Mass Spec. Rev.*, **1997**, 16, 201-217.

51) Harrison, A. G. *J. Mass Spectrom.*, **1999**, 34, 577-589.

52) Paizs, B.; Suhai, S. *Rapid Commun. Mass Spectrom.* **2002**, 16, 1699-1702.

53) Harrison, A. G.; Young, A. B.; Bleiholder, B.; Suhai, S.; Paizs, B. *J. Am. Chem. Soc.* **2006**, *128*, 10364-10365.

54) Cooper, T.; Talaty, E.; Grove, J.; Suhai, S.; Paizs, B.; Van Stipdonk, M. *J. Am. Soc. Mass Spectrom.* **2006**,

55) http://nobelprize.org/nobel_prizes/chemistry/laureates/1998/press.html   Nobel Prize website.

56) Pople, J.A. Nobel Lecture http://nobelprize.org

57) Wiliams, I.H. *Chem. Soc. Rev.*, **1993**, 22, 277-283.

58) Koch, W.; Holthausen, M.C.; A Chemist's Guide to Density Functional Theory, Second Edition, 2001 Wiley-VCH.

59) Jensen, F. Introduction to Computational Chemistry, Second Edition, 2007 John Wiley & Sons, Ltd.

60) Markus, R.A.; Rice, O.K.; *J. Phys. Colloid Chem.* **1951**, 55, 894-908.

61) Steinfeld, J.I.; Fransico, J.S.; Hase, W.L.; Chemical Kinetics and Dynamics, Second Edition, 1999.

62) Baer, T.; Hase, W.L.; Unimolecular Reaction Dynamics, Oxford, 1996.

63) Cramer, C.J. Essentials of Computational Chemistry, Second Edition, 2004, John Wiley & Sons, Ltd.

64) Lewars, E. Computational Chemistry, 2003, Kluwer Academic Publishers.

65) Born, M.; Oppenheimer, J.R.; *Ann. Physik.*, **1927**, 84, 457-484.

66) Eckart, C. *Physical Rev.*, **1935**, 46, 383-387.

67) Simons, J. Energetic Principles of Chemical Reactions, Jones and Bartlett (Boston), 1983.

68) K. Fukui. *Acc. Chem. Res.*, **1981**, 14, 363-368.

69) Gonzalez, C.; Schlegel, H. B.; *J. Chem. Phys.* **1989**, 90, 2154-2161.

70) Gonzalez, C.; Schlegel, H. B.; *J. Phys. Chem.* **1990**, 94, 5523-5527.

71) Paizs, B.; Lendvay, G.; Vékey, K.; Suhai, S. *Rapid Commun. Mass Spectrom.*, **1999**, 13, 525-533.

72) Rodriquez, C.F.; Cunje, A.; Shoeib, T.; Chu, I. K.; Hopkinson, A.C.; Siu, K. W. M. *J. Am. Chem. Soc.,* **2001**, 123, 3006-3012.

73) Balta, B.; Aviyente, V.; Lifshitz, C. *J. Am. Soc. Mass Spectrom.*, **2003**, 14, 1192-1203.

74) Hohenberg, P.; Kohn, W., *Phys. Rev.*, **1964**, 136, B864-B871.

75) Kohn, W.; Sham, L.J., *Phys. Rev.*, **1965**, 140, A1133-A1138.

76) Parr, R.G.; Yang, W.; Density-Functional Theory of the Electronic Structure of Molecules, Oxford University Press, New York, 1989.

77) Stephens, P.J.; Devlin, J.F.; Chabalowski, C.F.; Frish, M.J.; *J. Phys. Chem.* **1994**, 98, 11623-11627.

78) Barone, V.; Orlandini, L.; Adamo, C. *Chem. Phys. Lett.* **1994**, 231, 295-*300*.

79) Mijoule, C.; Latajka, Z.; Borgis, D. *Chem. Phys. Lett.* **1993**, 208, 364-368.

80) Becke, A.D., J. Chem. Phys., **1988**, 88, 1053-1062.

81) Lee, C.; Yang, W.; Parr, R.G.; *Phys. Rev. B*, **1988**, 37, 785-789.

82) Lozynski, M.; Rusinska-Roszak, D.; Mack, H.-G. *J. Phys. Chem. A,* **1998**, 102, 2899-2903.

83) Adamo, C.; Barone, V.; in "Recent Advances in DFT Methods. Part II," Chong, D.P., Ed., World Scientific, Singapore, 1997.

84) Sim, F.; St.-Amant, A.; Papai, I.; Salahub, D. R. *J. Am. Chem. Soc.*, **1992**, 114, 4391- 4400.

85) Ghosh, A. *Curr. Opin. Chem. Biol.* **2003**, 7, 110-112.

86) Paizs, B; Suhai, S., *J. Comput. Chem.*, **1998**, 19, 575-584.

87) Rice, O.K.; Ramsperger, H.C.; *J. Am. Chem. Soc*. 1927, 49, 1617- 1629.

88) Rice, O.K.; Ramsperger, H.C.; *J. A.m Chem. Soc*. 1928, 50, 617-620.

89) Kassel, L.S. *J. Phys. Chem.* 1928, 32, 225-242.

90) Gilbert, R.G.; Smith, S.C.; Theory of Unimolecular and Recombination Reactions, 1990, Blackwell Scientific Publications.

91) Rosenstock, H.M.; Wallenstein, M.B.; Wahrhaftig, A.L.; Eyring, H.; *Proc. Nat. Acad. Sci.* 38, 1952, 667-678.

92) Beyer, T.; Swinehart, D.R.; *ACM Commun.* **1973**, 16, 379.

93) Simons, J. An Introduction to Theoretical Chemistry, 2003, Cambridge University Press.

94) Ochterski, J.W. Vibrational Analysis in Gaussian, 1999, available at http://www.gaussian.com/

Chapter 2

Backbone Cleavages and Sequential Loss of Carbon Monoxide and Ammonia from Protonated AGG: A Combined Tandem Mass Spectrometry, Isotope Labeling, and Theoretical Study

Benjamin J. Bythell, Douglas F. Barofsky, Francesco Pingitore, Michael J. Polce, Ping Wang, Chrys Wesdemiotis and Béla Paizs

2.1 Abstract

The fragmentation characteristics of protonated alanylglycylglycine, $[AGG + H]^+$, were investigated by tandem mass spectrometry in MALDI-TOF/TOF, Ion Trap, and hybrid sector instruments. b_2 is the most abundant fragment ion in MALDI-TOF/TOF, Ion Trap, and hybrid sector metastable ion (MI) experiments while y_2 is slightly more abundant than b_2 in collision activated dissociation (CAD) performed in the sector instrument. The A-G amide bond is cleaved on the a_1-y_2 pathway resulting in a proton-bound dimer of GG and MeCH=NH. Depending on the fragmentation conditions employed, this dimer can then; (1) be detected as $[AGG + H - CO]^+$, (2) dissociate to produce y_2 ions, $[GG + H]^+$, (3) dissociate to produce a_1 ions, $[MeCH=NH + H]^+$, or, (4) rearrange to expel NH_3 forming a $[AGG + H - CO - NH_3]^+$ ion. The activation method and the experimental timescale employed largely dictate which of, and to what extent, these processes occur. These effects are qualitatively rationalized with the help of quantum chemical and RRKM calculations. Two mechanisms for formation of the $[AGG + H - CO - NH_3]^+$ ion were evaluated through nitrogen-15 labeling experiments and quantum chemical calculations. A mechanism involving intermolecular nucleophillic attack and association of the GG and imine fragments followed by ammonia loss was found to be more energetically favorable than expulsion of ammonia in an S_N2-type reaction.

2.2 Introduction

Soft ionization techniques such as fast-atom-bombardment (FAB) [1], electrospray ionization (ESI) [2] and matrix-assisted laser desorption/ ionization (MALDI) [3, 4], have enabled tandem mass spectrometry (MS/MS) to become the standard tool for elucidation of peptide sequence. Gas-phase protonated peptides can be isolated, activated (usually by collision(s) with an inert gas), dissociated and detected. The resulting spectra show that protonated peptides undergo backbone cleavages, dissociations in the side chains and losses of small neutrals (water, ammonia, carbon monoxide), or an amalgamation of these [5-8]. Experimental and computational studies have been undertaken on the backbone and side chain fragmentations [9-32] as well as neutral losses [11, 18, 19, 27-29, 31, 32].

Peptide fragment ion spectra are utilized to sequence peptides and proteins with the help of various bioinformatics tools. Candidate peptides and their theoretical MS/MS spectra are generated *in silico* using protein and/or DNA databases and fragmentation models. The *in silico* spectra are then compared to the experimental MS/MS spectrum in order to find the most closely matching sequences. The success of this computer-aided peptide sequencing approach is directly related to the quality of the applied fragmentation models that summarize our present understanding of gas-phase peptide chemistry [7].

The 'mobile proton' fragmentation model [26, 33] takes into account the energetics and reactivity of the various protonation sites of peptides. Upon excitation, the extra proton is transferred from a usually un-reactive site of higher gas-phase basicity (arginine, R, or lysine, K, side chain or the N-terminal amino group) to form an energetically less favored but reactive, backbone-amide- protonated species. Protonation

of the amide nitrogen leads to considerable weakening of the amide bond [34], and a species such as this plays a critical role in most of the peptide fragmentation pathways (PFPs) that lead to sequence-informative b, a, and y ions [7]. The mobile proton model enables prediction of whether a particular peptide in a particular state of protonation is likely to produce a sequence-informative MS/MS spectrum. For example, peptide ions whose number of added protons exceeds the number of R and K residues are expected to fragment at the various amide bonds forming b, a and y ions. On the other hand, substantially less sequence coverage is expected if the number of added protons is equal to the number of R residues, especially if the peptide contains aspartic acid, D, and/or glutamic acid, E, and the MS/MS spectrum is dominated by cleavages C-terminal to these residues (aspartic acid effect [7]).

The recently introduced "pathways in competition" (PIC) fragmentation model [7], which provides a more general framework, takes into account specific features of individual peptide fragmentation pathways (PFPs) and their interaction. Fragment ion abundances in the MS/MS spectra of peptides are determined by pre-cleavage, bond-cleavage, and post-cleavage events. The pre-cleavage phase involves the proton-transfers (mobile proton) and internal rotations necessary to populate fragmenting species; hence, PIC is a logical extension to the mobile proton model. The energetics and kinetics of the competing PFPs determine which chemical bonds are cleaved upon excitation and, therefore, what kind of fragments appear as charged or neutral species in the mass spectrometer. Since MS only detects charged species, the fate of the added proton is critical as this defines which fragments are detectable. Simple free-energy relationships [35] can often be used to explain and/or predict relative fragment ion abundances based

on fragment proton affinities (or gas-phase basicities) [7, 36]. Furthermore, the post-cleavage phase of peptide fragmentation can feature rather rich chemistry including PFPs that lead to scrambling of primary sequence information [37] and rearrangements via fragment re-association in post-cleavage proton-bound dimers [38]. While the mobile proton model does not consider the cleavage or post-cleavage phases of peptide fragmentation, the flexibility of PIC enables detailed understanding of the whole fragmentation process.

The present article has two major goals. First, recent experimental and modeling studies have mainly concentrated on low-energy fragmentation processes of protonated peptides. The emerging MALDI-TOF/TOF technique features ions energized both by the MALDI process and by 1keV (laboratory frame) collision(s); thus, potentially opening up fragmentation channels which are frozen in the usual ion trap (IT) or quadrupole-TOF instruments. In the present paper, MS/MS fragmentation characteristics of protonated AGG in MALDI-TOF/TOF, IT, and sector instruments are presented, and the effects of various internal energy distributions and time-scales on fragment ion abundances are compared. Second, the results of using nitrogen-15 (15-N) labeling and DFT calculations to investigate the kinetics and mechanism of sequential loss of CO and ammonia from protonated AGG are presented and discussed. Specifically, evidence has been found for the loss of CO occurring on the a_1-y_2 PFP and for the re-association of the proton-bound dimer (PBD) of GG and Me-CH=NH to form a species that can easily expel ammonia. While a similar reaction mechanism was proposed to account for combined loss of CO and ammonia from protonated glycinamide [39] and from b_3 of protonated GGGG [38,

see also 2.SI1c] this work is the first direct evidence that primary peptide fragments can undergo re-association type reactions in post-cleavage proton-bound dimers.

2.3 Methods and Materials

2.3.1 Materials

The solvents (HPLC-grade water and acetonitrile), trifluoroacetic acid, acetic acid, sulfuric acid, monoammoniumphosphate and the AGG peptide were purchased from Sigma-Aldrich (Milwaukee, WI). The labeled tripeptide A(^{15}N)GG was purchased from SynPep (Dublin, CA). All chemicals were used without further purification.

2.3.2 Tandem Mass Spectrometry (MS/MS) Experiments

The experiments were conducted on a Finnigan MAT LCQ (San Jose, CA) ion trap (IT) instrument with ESI [40, 41], an Applied Biosystems 4700 Proteomics Analyzer MALDI-TOF/TOF (Framingham, MA) [42-44] with MALDI, and a Micromass AutoSpec-Q hybrid tandem MS (Manchetser, UK) with FAB ionization [32, 45]. Samples for the IT experiments, were prepared by dissolving AGG in acetonitrile/water/acetic acid 30/70/0.1 (v:v) to form a 2×10^{-5} mol L^{-1} solution, which was infused into the ESI source at a rate of 10 µL min^{-1}. The entrance to the sampling capillary was set at -4kV and N$_2$ served as the nebulizing and drying gas (170°C). Collisionally activated dissociation (CAD) MS/MS of [AGG + H]$^+$, were performed by ejecting all ions except [AGG + H]$^+$, and then exciting the latter to fragment in the presence of He buffer gas (10^{-3} Torr) using a radiofrequency (RF) field (0.62V$_{p-p}$) [46]

under automated gain control (AGC) to optimize the quantity of ions accumulated in each scan. The reproducibility of the relative abundances from multiple, repeated scans was circa ±15%. Additional scans were performed where the excitation level (collision energy) was incrementally increased. These scans began from well below the threshold of product ion formation and continued until all product ions had been observed.

In the MALDI-TOF/TOF experiments, the matrix was prepared by dissolving α-cyano-4-hydroxycinnamic acid (α-CHCA) in acetonitrile/water/trifluoroacetic acid/monoammonium phosphate (6mg/ml) 47/47/0.1/6 (v:v) solution at a concentration of 2mg/mL. The AGG peptide was dissolved in acetonitrile/water/trifluoroacetic acid 50/50/0.1 (v:v) solution at a concentration of 100μg/mL. The sample solutions were then prepared by mixing the matrix/peptide solutions in a 1 to 1 ratio. 0.4μL of sample solution was applied to each spot on a 196-well target plate and allowed to air dry prior to introduction into the mass spectrometer. The MS/MS spectra consisted of 5000 laser shots per well, ten replicate wells, with laser fluence constant and at a level low enough to prevent signal saturation. Air was used as the collision gas at 2.7 x 10^{-8} ("no gas"), 7.3 x 10^{-7} and 1.6 x 10^{-6} mbar. The [AGG + H]$^+$ ions formed in the MALDI process are accelerated to 8keV, mass selected using timed ion selection (Resolving power = 200FWHM) then decelerated to 1keV (laboratory frame) for metastable ion (MI) decomposition or collisionally activated dissociation (CAD).

The sector MS/MS experiments were conducted on a Micromass AutoSpec- Q tandem mass spectrometer of EBEhQ geometry (E, electric sector; B, magnetic sector; h, RF-only hexapole; Q, quadrupole mass filter). Only the sector section (EBE) was used in this study. The protonated AGG was formed by fast atom bombardment (FAB)

ionization, using 12 keV Cs^+ ions as bombarding particles and sulfuric acid as the matrix. A few µL of a saturated solution of AGG in the matrix were introduced into the ion source and bombarded by Cs^+. The peptide $[AGG + H]^+$ ions formed in this process were accelerated to 8 keV and mass-selected by the EB sectors for measurement of their MI and CAD tandem mass spectra at high kinetic energy in the field-free region (FFR) between EB and the subsequent electric sector. The product ions from these reactions were mass-analyzed by scanning the second electric sector. In CAD mode, one of the collision cells situated in the FFR was pressurized with argon to effect 80% transmittance of the $[AGG + H]^+$ beam. In MS^3 experiments, a specific fragment ion from metastable $[M + H]^+$ ions dissociating in the field-free region in front of the first electric sector was transmitted through EB by proper adjustment of the E and B fields, and the corresponding high-energy CAD spectrum was acquired using the above mentioned collision cell. For the fragments in the MI spectra of $[AGG + H]^+$ measured at high kinetic energy, the accompanying kinetic energy releases were calculated using fragment peak widths at half height ($T_{0.5}$); the quoted $T_{0.5}$ values were corrected for the main beam width using established procedures. Approximately 100–200 scans were summed per MI, CAD, or MS^3 experiment, depending on the intensity of the main beam. The reproducibility of relative abundances was better than ±15%.

2.3.3 Computational Methods

A conformational search engine devised specifically to deal with protonated peptides, was used to scan the potential energy surface (PES) of protonated alanylglycylglycine, $[AGG + H]^+$ [22, 23, 28-30]. These calculations began with

molecular dynamics simulations on various forms of $[AGG + H]^+$ using the InsightII program (Biosym Technologies, San Diego, USA) in conjunction with the AMBER force field modified by Paizs et al. [22, 23, 28-30] in order to manage amide nitrogen and oxygen protonated species. During the molecular dynamics simulations, structures were regularly saved for further refinement by full geometry-optimization using the same force fields. In the next stage of the process, these structures were analyzed by a conformer-family search program [22, 23, 28-30]. This program is able to group optimized structures into families based on similarity of the most important characteristic torsion angles. The most stable species in these families were then fully optimized (in Heidelberg and Corvallis) at the HF/3-21G, B3LYP/6-31G(d), and the B3LYP/6-31+G(d,p) levels.

Having scanned the PES, transition structures (TSs) corresponding to various fragmentation pathways of $[AGG + H]^+$ and reactions of the proton-bound dimers of MeCH=NH and GG formed on the a_1-y_2 pathway were then sought. These were calculated at the B3LYP/6-31G(d) and B3LYP/6-31+G(d,p) levels of theory. In most of the cases, the resulting transition structures were checked using intrinsic reaction coordinate (IRC) calculations to unambiguously define which minima are connected by the TS investigated. Post-reaction complexes and proton-bound dimers were fully optimized at the B3LYP/6-31G(d) and B3LYP/6-31+G(d,p) levels of theory in a manner similar to that used for the various AGG protonation sites and transition structures. Relative energies were calculated by using the B3LYP/6-31+G(d,p) total energies and zero-point energy corrections (ZPE) determined at the B3LYP/6-31G(d) level. The Gaussian [47] program was used for all *ab initio* calculations.

The unimolecular rate coefficients for transitions involving the a_1-y_2, water-loss, and b_2-y_1 TSs were calculated using the results of the DFT calculations (relative energies, vibrational frequencies, rotational constants) via the RRKM method [48] over a grid of energies up to a limit well exceeding the calculated threshold energy of the lowest-energy fragmentation. The sum and density of states were calculated using the Beyer-Swinehart direct count algorithm [48].

2.4 Results and Discussion

2.4.1 Tandem mass spectra (MS/MS) of protonated AGG

MALDI-TOF/TOF experiments yielded fragment ions similar to those observed in the ESI IT and sector experiments, but with large differences in fragment ion abundance (Table 2.1). Fragmentation under MALDI-TOF/TOF conditions was insensitive to pressure changes in the CAD cell over two orders of magnitude. This indicates that MI dissociation played a large part in the fragmentation of the $[AGG + H]^+$ ion. The base peak belongs to the b_2 ion by a significant margin in all experiments except the sector CAD where the y_2 ion was slightly more abundant than the b_2. All spectra show a y_2 ion peak, the relative abundance of which varies from 4 to 100 % depending on the experimental setup (Table 2.1). It is one of the least abundant fragment ions in the IT, whereas it is the most abundant fragment in the sector CAD. Furthermore, the relative abundance of the y_2 ion exhibited in the MI sector spectra was lower than in the MALDI-TOF/TOF spectra. The intensity of the H_2O loss channel is low in all experiments except those performed by MALDI-TOF/TOF where the related peak is nearly totally missing. This channel has the largest relative abundance in the sector CAD spectra. The a_2 ion was

observed in the sector CAD spectra at medium and in the IT at low abundance. It should

be noted that the sum of the b_2 and a_2 ion intensities (a_2 is formed from b_2 [7, 11, 12, 49]

by losing CO) exceeds the intensity of the y_2 (formal base peak) in the sector CAD

spectrum.

Loss of CO, $[AGG + H - CO]^+$ is observed in all instruments except the IT.

Additionally, a fragment ion peak at m/z 159 is observed in the IT, MALDI-TOF/TOF,

and sector CAD spectra, but not in the MI sector spectrum. For protonated AGG, this

peak can be assigned as x_2 or $[AGG + H - CO - NH_3]^+$. Loss of 45u occurs for other

aliphatic peptides in the MALDI-TOF/TOF where glycine was the second residue (see

Supporting information (Figure 2.SI1) for GGA, GGL, GGGG, YGG spectra). However,

it does not occur for aliphatic peptides where glycine was not the second amino acid (e.g.

AAA, GPGG (Figure 2.SI2)). Ordinarily, x_n ions are not present in IT spectra, and no

other m/z values nominally corresponding to "x_n ions" were observed in the MALDI-

TOF/TOF (or sector CAD) spectra. The m/z 159 peak was also observed in MS^3

experiments performed on $[AGG + H - CO]^+$ (Figure 2.SI3) along with y_2 and a_1 (m/z 44)

ions [50]. This is strong evidence for the existence of the $[AGG + H - CO - NH_3]^+$ ion.

The a_1 ions were not present in the MALDI-TOF/TOF or IT data as these

instruments do not detect ion signals below m/z 69 or 55, respectively. However, the

CAD spectrum obtained on the sector instrument does contain a small a_1 fragment.

Relative appearance-energies for the primary fragmentation channels were

determined with the IT instrument by incrementally increasing the parent excitation level

(collision energy). These scans indicated that the b_2 ion peak appears at the lowest

excitation level while $[AGG + H - H_2O]^+$ and $[AGG + H - CO - NH_3]^+$ peaks appear

next with the latter having much higher abundance. The y_2 peak appears next followed lastly by the a_2 peak. It is worth noting here that, due to the relatively long time-scale of the IT instrument, kinetic shifts are small [51], so the order of the appearance energies likely reflects the order of the true threshold energies.

2.4.2 Fragmentation pathways of [AGG + H]$^+$

In the following sections theoretical data are presented for the most important fragmentation channels of protonated AGG. Then the computed energetics and kinetics are used to explain the effect of the various experimental energy distributions and time-scales in the MALDI-TOF/TOF, IT, and sector instruments on the fragmentation characteristics observed.

2.4.3 Protonation Energetics, Transition States and Peptide Fragmentation Pathways of [AGG + H]$^+$

The protonation energetics of AGG along with the energetics of the a_1-y_2, b_2-y_1, and water-loss TSs are shown in Table 2.2. As has been previously shown for GGG [52], the energetically most favored protonation site (Figure 2.1a) of AGG is the AG amide oxygen. The relative energy of the N-terminal amino protonated species is 2.2 kcal/mol while protonation at the amide nitrogens requires at least 18-19 kcal/mol internal energy, which is consistent with the literature [7].

2.4.4 The b₂-y₁ pathway

Cleavage of the C-terminal amide bond of $[AGG + H]^+$ on the b_2-y_1 pathway (Scheme 2.1) results in a b_2 ion with an oxazolone structure [11, 12, 53, 54]. The first step on the b_2-y_1 PFP is mobilization of the added proton to the nitrogen of the C-terminal (GG) amide bond (18.9 kcal/mol relative energy (E_{rel}), Table 2.2). Cleavage of the protonated C-terminal amide bond and simultaneous formation of the oxazolone ring take place through **TS_b₂_y₁** shown in Figure 2.1b (E_{rel} at 29.4 kcal/mol, activation entropy at -1.0 cal/mol K). After leaving **TS_b₂_y₁** various proton-bound dimers of AG_{oxa} and G are formed. As the PA of the oxazolone structure (217.9 kcal/mol, calculated at the B3LYP/6-31+G(d,p) level) is higher than that of G (211.4 kcal/mol [55, 56]), formation of b_2 is clearly favored. In agreement with this prediction, no y_1 ions were experimentally observed from $[AGG + H]^+$ under any of the fragmentation conditions employed in this study.

2.4.5 Loss of water from protonated AGG

Loss of water [31] from protonated AGG is initiated from an amide O protonated species (**RC_H₂O_loss**, Figure 2.1c, Scheme 2.2). This reactive configuration (RC) can be formed from the global minimum by rotations around the N-C_α(2)-CO-N and C_α(2)-CO-N-C_α(3)C torsions (Figure 2.1c) without mobilization of the extra proton. Once **RC_H₂O_loss** is populated, transfer of the extra proton to the C-terminal OH and nucleophilic attack of the GG amide oxygen on the COOH carbonyl are required to reach **TS_H₂O_loss** (Figure 2.1d, Scheme 2.2). While the relative energy of this 'tight' TS is

relatively low at 33.2 kcal/mol, the corresponding transition is entropically disfavored (activation entropy at -6.0 cal/mol K) due to the drastic structural changes needed to reach **RC_H₂O_loss** and **TS_H₂O_loss** from the global minimum. It should be mentioned that **RC_H₂O_loss** is not a local minimum on the B3LYP/6-31g(d) surface and could be optimized only at B3LYP/6-31+g(d,p). Therefore, the ZPE of this species was calculated at the latter level in contrast to other structures for which ZPEs derived at B3LYP/6-31g(d) were used. To validate the adjustment introduced by this approach we computed the ZPEs of the global minimum at the two model chemistries. These calculations indicate a ZPE decrease of 0.71 kcal/mol (<0.5%) which fully supports the use of the B3LYP/6-31+g(d,p) value for **RC_H₂O_loss**.

2.4.6 The a₁-y₂ PFP

The a_1-y_2 pathway (Scheme 2.3a) is initiated by mobilization of the extra proton to the AG amide nitrogen (E_{rel} at 17.8 kcal/mol, Table 2.2). This weakens the CO-NH and (H_2NMe)CH-CO bonds, which allows CO to be expelled and a proton-bound dimer of MeCH=NH and GG to be formed. **TS_a₁-y₂**, (Figure 2.1e, E_{rel} at 38.5 kcal/mol) is energetically the least favored of the investigated TSs, but, it has 'loose' character [31], so entropic factors (activation entropy at 11.0 cal/mol K) favor it relative to the 'tight' b_2-y_1 or water loss pathways.

2.4.7 RRKM calculations on the primary PFPs of protonated AGG

To approximate the time-scale [57] of the primary fragmentation channels of protonated AGG, RRKM calculations were performed using the energetics, vibrational frequencies, and rotational constants derived from the modeling. The calculated unimolecular rate constants for the b_2-y_1, water loss, and the a_1-y_2 PFPs are plotted in Figure 2.2. The threshold energy on the b_2-y_1 pathway (29.4 kcal/mol) is lower than that of the water loss (33.2 kcal/mol) or the a_1-y_2 (38.5 kcal/mol) PFPs (Table 2.2). This order of the threshold energies is in line with the results of IT experiments where the excitation level was scanned to evaluate relative appearance energies of the main fragments (see preceding). The RRKM calculations indicate that the b_2-y_1 PFP is favored at low internal energies while the a_1-y_2 pathway becomes increasingly preferred at high internal energies.

The RRKM calculations (Figure 2.2) further suggest that the water loss PFP is kinetically disfavored compared to b_2-y_1 or a_1-y_2. In agreement with this prediction the relative abundance of $[AGG + H - H_2O]^+$ only exceeds 5 % of the base peak in the sector CAD experiment (Table 2.1). It must be noted that the RRKM calculations probably overestimate the branching ratio of the b_2-y_1 and water loss PFPs. Use of the global minimum and the corresponding TSs, to estimate the unimolecular rate constants in the RRKM calculations assumes that the extra proton is similarly mobile for each case and that the real time determining step is the dissociation. It is likely that this assumption is not fully valid since the proton mobilization pathways are much more complex for b_2-y_1 or a_1-y_2 than for the water loss PFP. Actually, the latter does not involve proton mobilization to populate the reactive configuration which is produced in a complicated

one-step process from the global minimum. On the other hand, the b_2-y_1 or a_1-y_2 PFPs require numerous proton transitions [58, 59]. Consequently, over interpretation/ extrapolation of the RRKM plot should be avoided.

2.4.8 Sequential loss of CO and ammonia from protonated AGG: formation of the [AGG + H – CO – NH$_3$]$^+$ ion

The proton-bound dimers of MeCH=NH and GG formed on the a_1-y_2 PFP can be detected as the [AGG + H - CO]$^+$ peak, can dissociate to produce y_2 ions, [GG + H]$^+$, or a_1 ions, [MeCH=NH + H]$^+$, and can rearrange and expel NH$_3$ to form [AGG + H - CO - NH$_3$]$^+$ ions. Determining which of these processes occur and to what extent depends on the activation methods and the experimental timescale involved. MS^3 experiments on [AGG + H - CO]$^+$ ion in the sector instrument (Figure 2.SI3) show peaks with *m/z* 133 (y_2), *m/z* 44 (a_1), and *m/z* 159 ([AGG + H - CO - NH$_3$]$^+$).

There exist many different proton-bound dimers (PBD) of MeCH=NH and GG with substantially different intermolecular bonding patterns. The PBD which is most energetically favored is an imine protonated species (**I**, Figure 2.SI4, E_{rel} at 11.3 kcal/mol) despite the PA of GG, being higher than that of MeCH=NH (221.9 & 217.9 kcal/mol respectively, calculated at the B3LYP/6-31+G(d,p) level).

Two major pathways can be envisaged for the formation of *m/z* 159 from the proton-bound dimer of MeCH=NH and GG. On Path 1, **I** rearranges to form **II** (E_{rel} at 14.1 kcal/mol, Table 2.3, Scheme 2.3b, Figure 2.SI4) that features a strong N$^+$-H···O H-bond and a C-H···NH$_2$ interaction. In such a geometrical arrangement, the amino nitrogen

of GG can attack the partially positively charged carbon of the imine. Formation of the new N-C bond takes place through **TS II_III** (E_{rel} at 23.4 kcal/mol, Scheme 2.3b, Figure 2.SI4) to form **III**. Proton transfers to the GG amide oxygen through TSs **III_IV** (E_{rel} at 24.7 kcal/mol, Scheme 2.3b, Figure 2.SI4), and further to the N-terminal amino group via **IV_V** (E_{rel} at 28.8 kcal/mol, Scheme 2.3b, Figure 2.SI4) lead to structure **V** (E_{rel} at 14.5 kcal/mol, Scheme 2.3b, Figure 2.SI4), which can expel ammonia via TS **V_VI** (E_{rel} at 21.5 kcal/mol, Scheme 2.3b, Figure 2.SI4). Note that the highest energy structure on Path 1 is TS **IV_V**, at 28.8 kcal/mol relative energy, which is lower than the a_1-y_2 threshold energy at 38.5 kcal/mol.

An alternative pathway (Path 2, Scheme 2.3b) involves rearrangement of **I** and proton transfer to GG to form **VII** (E_{rel} at 15.9 kcal/mol, Scheme 2.3b, Figure 2.SI4). **VII** then rearranges to form **VIII** (E_{rel} at 26.2 kcal/mol, Scheme 2.3b, Figure 2.SI4), which can undergo an S_N2-type reaction to expel ammonia in a one-step process via TS (**VIII_VI** (E_{rel} at 51.5 kcal/mol, Scheme 2.3b, Figure 2.SI4)) to form the same final product as Path 1, namely structure **VI**. Path 2 requires a TS (**VIII_VI**) with energy considerably greater than the a_1-y_2 TS energy. Path 2 would therefore require a significantly greater input of energy to be active than Path 1. It should be noted that while performing well for TS and reaction-path geometries, density functional methods perform less well for activation barrier height calculations for S_N2 reactions [60] and underestimates their barrier heights [61]. As such it is reasonable to expect the E_{rel} of **TS_VIII_VI** (Path 2) to be a lower bound thus making Path 1 even more likely in terms of energy. These findings are also in agreement with the 15-N labeling results presented in this paper (see below).

The energetics of Paths 1 and 2 are summarized in Figure 2.3, which also displays the relative energies of the "a_1 + GG + CO" and "MeCH=NH + y_2 + CO" a_1-y_2 exit channels. The latter is nearly equi-energetic with the a_1-y_2 TS in line with the low kinetic energy release ($T_{0.5}$ at 0.1 eV) observed for the formation of the y_2 ion in MI experiments on protonated AGG. Path 1 is clearly energetically favored, over Path 2, considering formation of the *m/z* 159 ion. However, Path 1 is a multi-step process that can be kinetically controlled in experiments with short timescales.

CAD experiments on the sector instrument were also carried out with the labeled $A(^{15}N)GG$ peptide. The CAD spectrum of $[A(^{15}N)GG + H - CO]^+$ (m/z 177) shows fragments at m/z 159 (–18 u, corresponding to loss of $^{15}NH_3$), 133 (–44u, corresponding to loss of $^{15}NH=CH-CH_3$), and 45 (corresponding to ion $[^{15}NH_2=CH-CH_3]^+$), which provide corroborating evidence for the dominance of path 1 over path 2 of Scheme 2.3b in agreement with the theoretical results.

2.4.9 Effect of internal energy distribution and fragmentation timescale on relative fragment ion abundances

The DFT and RRKM calculations indicate that the branching ratios of the fragment ions from the b_2-y_1 and a_1-y_2 PFPs are a reasonable measure of the parent excitation level. That is, dominance of the former indicates mild while dominance of the latter suggests harsher excitation. These abundance ratios are 102/14, 100/65, 100/28, and 121/137 (Table 2.1) for the IT, MALDI-TOF/TOF, MI sector, and CAD sector instruments, respectively. As expected, the sector CAD precursor ions are the most energized. The MALDI-TOF/TOF precursor ions appear to be more energized than those

produced in the IT or MI sector experiments but less excited than those produced in the sector CAD.

This relative degree of precursor ion energization is also consistent with the final CAD products arising from the a_1-y_2 PFP. Here the y_2/ $[AGG + H - CO - NH_3]^+$ ion ratios are 4/10, 35/23, 100/28 for the IT, MALDI-TOF/TOF and sector CAD instruments, respectively (see Table 2.1). The increasing ratio reflects the increasing energization of the precursor ions making dissociation of the PBD to form y_2 ions more likely [62]. This trend is also supported by the decreasing time-scale of the experiment, which leaves less time for the PBD to rearrange and form $[AGG + H - CO - NH_3]^+$.

The high-energy CAD experiments performed in the instrument using FAB ionization deposit a broad range of internal energies in a single collision. The precursor ion dissociates in less than a few microseconds ($\sim 2\mu s$); under these conditions, several competitive dissociations can take place simultaneously, depending on the amount of energy gained by a particular precursor ion as well as the dissociation kinetics [63]. Consequently, each of the b_2-y_1, a_1-y_2, and water loss PFPs are active with medium to high fragment ion abundances (Table 2.1).

This is in direct contrast to the low-energy CAD in the IT, which utilizes multiple collisions each depositing a small amount of internal energy to activate the precursor ion. This makes millisecond time windows (~ 10 ms) available for dissociation [41, 42, 46]. Once the critical energy for a fast reaction is reached, the precursor ion reacts by this channel leading to an accumulation of the corresponding fragment ion. Any competing reaction is obstructed unless its critical energy and kinetics are very similar to the channel accumulating fragment ions [32]. This effect is clearly observed in the high b_2-y_1/a_1-y_2

abundance ratio in the IT experiment. Multiple collision conditions at low energy may, however, promote consecutive fragmentations if the corresponding energetics are favorable [27]. As a consequence, a weak a_2 ion signal is observed due to activity of the $b_2 \rightarrow a_2$ pathway. It is worth noting here that no $[AGG + H - CO]^+$ ions were observed in the IT experiment. This is very likely due to the millisecond time-scale of the fragmentation that allows the corresponding PBDs to rearrange and fragment by losing NH_3 to form ion m/z 159 (10% of the base peak, Table 2.1). Note, that the energetics of Path 1 (Figure 2.3) clearly favor this multi-step transition.

The fragment ion abundances observed in the spectra produced with the MI sector are similar to the corresponding IT values including the branching ratio of the fragment ions from the b_2-y_1 and a_1-y_2 PFPs (100/28 and 102/14, respectively). There are two differences between the IT and MI sector fragment abundances: no CO loss peak is observed in the IT (see preceding for explanation) while the ion with m/z 159 is not formed in the MI sector experiment. The latter can be explained by the short (~2 μs) time scale of the MI experiment and the rather narrow internal energy distribution of the corresponding parent population.

The MALDI-TOF/TOF instrument lies between two extremes as internal energy is acquired by the precursor ion during the ionization processes and subsequently in the collision cell through which it passes at 1 keV. Consequently, precursor ions in the collision cell (~6μs for m/z 204) can decay metastabily, as a result of CAD or a mixture of the two. The resulting ensemble of ions is then reaccelerated, focused, and detected. The higher internal energy imparted by the MALDI-TOF/TOF leads to increased fragmentation on the a_1-y_2 pathway relative to the IT and MI sector experiments. The

MALDI-TOF/TOF timescale is still sufficiently long to enable observation of species formed from chemical reactions other than direct dissociation and their intermediates (see Table 2.1). Interestingly, no evidence for a water loss peak is observed in the MALDI-TOF/TOF spectrum, but evidence for this reaction is seen at the 3-15 % level in the spectra from the other experiments. Ionization by MALDI is more energized than is that by ESI or FAB. As mentioned above, the water loss PFP is initiated from a structure (**RC_H₂O_loss**, E_{rel} at 13.6kcal/mol, Figure 2.1c) that is formed from the global minimum by backbone rotations without transfer of the ionizing proton. It can easily happen that the highly energized MALDI-TOF/TOF precursors populate protonation states like the N-terminal amino group at 2.2 kcal/mol relative energy (Table 2.2), therefore suppressing the water-loss PFP that requires protonation at the A-G amide oxygen.

2.5 Conclusions

A mixed ab initio, 15-N labeling and RRKM technique was successfully used to model fragmentation of $[AGG + H]^+$ under a variety of fragmentation conditions. The different fragmentation patterns observed in the three mass spectrometers reflected differences in internal energy distributions and timescales available for dissociation. Combined tandem MS experiments offer a powerful tool for probing the complex gas-phase chemistry of protonated peptides. Accompanying computational and labeling studies can provide important mechanistic details not readily observed from tandem mass spectra alone.

Similar reactions to the formation of $[AGG + H - CO - NH_3]^+$ peak could easily occur for other protonated peptides and consequently lead to unexpected fragment ions in MS/MS spectra. This, in turn, could lead to erroneous peptide and protein identification using current bioinformatics tools. As such, the post-cleavage phase of peptide fragmentation warrants further study so that the PIC model may be improved and implemented into software to enhance the effectiveness of peptide sequencing algorithms.

2.6 Acknowledgements

Table 2.1: Relative abundance in % of base peak intensity, using peak areas of $[AGG + H]^+$ fragment ions.

Instrument	H_2O loss	CO loss	$CO + NH_3$ loss	y_2	b_2	a_2
ESI ion trap	3	0	10	4	100	2
MALDI-TOF/TOF	<1	7	23	35	100	0
MI FAB AutoSpec-Q	5	17	0	11	100	0
CAD FAB AutoSpec-Q	14	9	28	100	98	23

Table 2.2: Relative (kcal mol^{-1}) and total (Hartree) energies of various protonated forms of [AGG + H]$^+$ and the b$_2$-y$_1$, a$_1$-y$_2$ and water loss TSs. Total energies, E$_{tot}$ are reported at the B3LYP/6-31G+(d,p) theoretical level. The relative energies, E$_{rel}$ (corrected for zero-point energy calculated from B3LYP/6-31G(d) vibrational frequencies) were determined with respect to the global minimum on the PES of [AGG + H]$^+$. Relative enthalpies (ΔH_{298}) and Gibbs free energies (ΔG_{298}) at 298K are in kcal/mol, relative entropies (ΔS_{298}) are in cal/mol K. $^{\#}$ Zero-point energy calculated using B3LYP/6-31G+(d,p).

Species	E$_{tot}$	E$_{rel}$	Species	E$_{tot}$	E$_{rel}$	ΔH_{298}	ΔG_{298}	ΔS_{298}
AG amide O protonated	-740.209780	0	AG amide O protonated	-740.209780	0	0	0	0
GG amide O protonated	-740.198967	7.2	**TS_b$_2$-y$_1$**	-740.162341	29.4	29.4	29.7	-1.0
N-term. Amino protonated	-740.209444	2.2	**TS_H$_2$O_loss**	-740.154810	33.2	33.0	34.8	-6.0
AG amide N protonated	-740.181374	17.8	**TS_a$_1$-y$_2$**	-740.144462	38.5	39.7	36.4	11.0
GG amide N protonated	-740.179625	18.9	**RC_H$_2$O_loss**	-740.187797	13.6$^{\#}$			

Table 2.3. Relative (kcal mol^{-1}) and total (Hartree) energies of proton-bound dimers and related species shown in Scheme 2.3. Total energies are reported at the B3LYP/6-31G+(d,p) theoretical level. The relative energies (corrected for zero-point energy calculated from B3LYP/6-31G(d) vibrational frequencies) were determined with respect to the global minimum on the PES of [AGG + H]$^+$, taking into account the ZPE-corrected total energy of CO (E_{tot} at -113.312292H, ZPE at 0.00503 H) eliminated on the a_1-y_2 PFP.

Species	E_{tot}	$E_{rel.}$	Species	E_{tot}	E_{rel}
I	-626.867503	11.3	II	-626.862438	14.1
II_III	-626.849058	23.4	III	-626.855829	20.7
III_IV	-626.844640	24.7	IV	-626.845254	26.9
IV_V	-626.841334	28.8	V	-626.865704	14.5
V_VI	-626.850038	21.5	VI NH$_3$ comp	-626.864984	12.3
VI and NH$_3$	-626.836471	27.7	VII	-626.859314	15.9
VIII	-626.843598	26.2	VIII_VI	-626.801695	51.5

Figure 2.1: Selected structures on the potential energy surface of protonated AGG: a) global minimum (**GM**), b) **TS_b$_2$-y$_1$** on the b$_2$-y$_1$ PFP, c) RC_**H$_2$O_loss** and d) **TS_H$_2$O_loss** on the water loss PFP**,** and e) **TS_a$_1$-y$_2$** on the a$_1$-y$_2$ PFP.
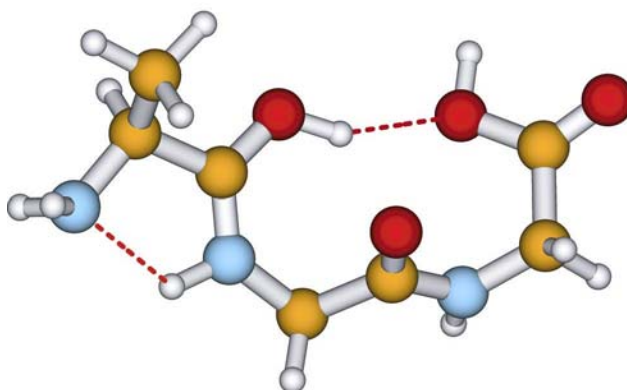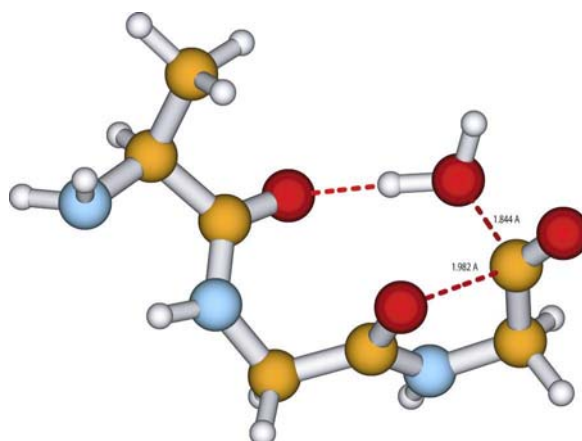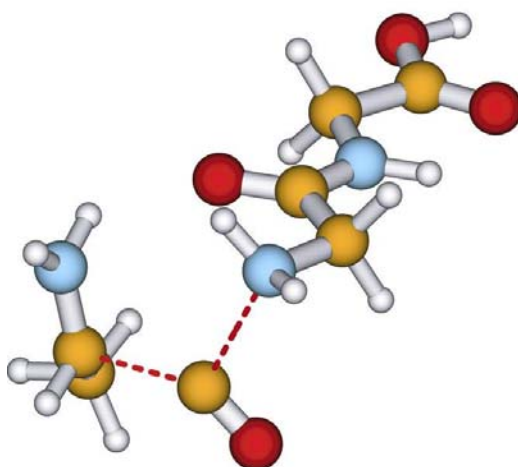
a)
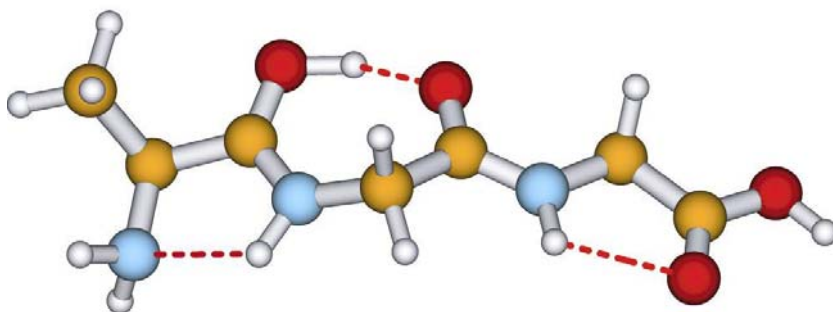
b)
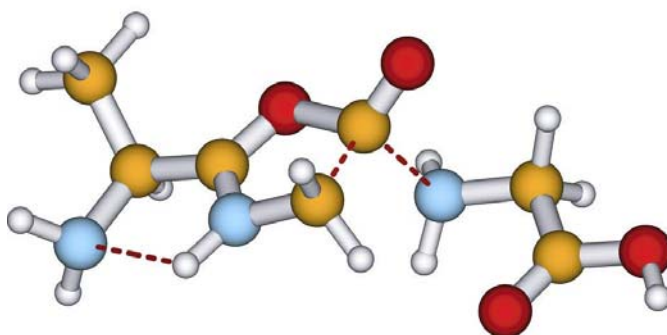
c)

d)

e)

Figure 2.2: Unimolecular rate constants (s^{-1}) calculated by the RRKM formalism for the b_2-y_1, water-loss, and a_1-y_2 PFPs.
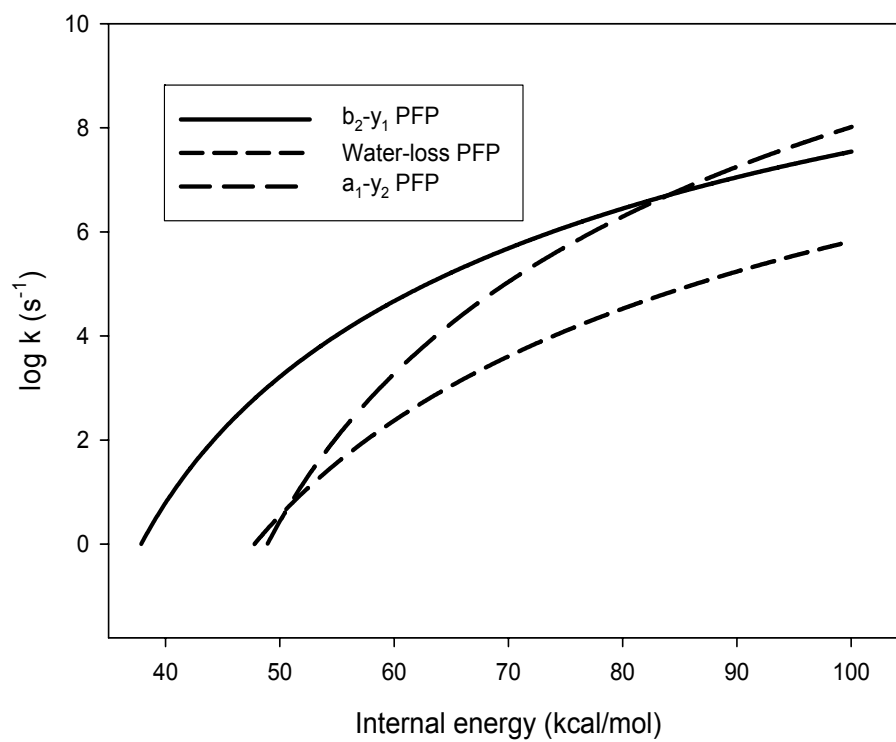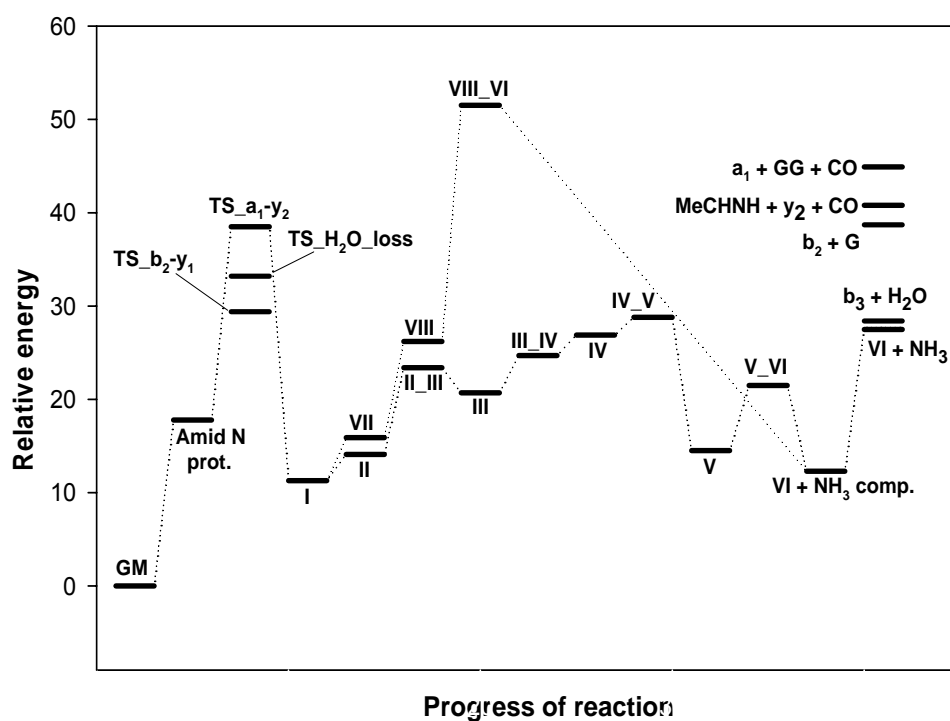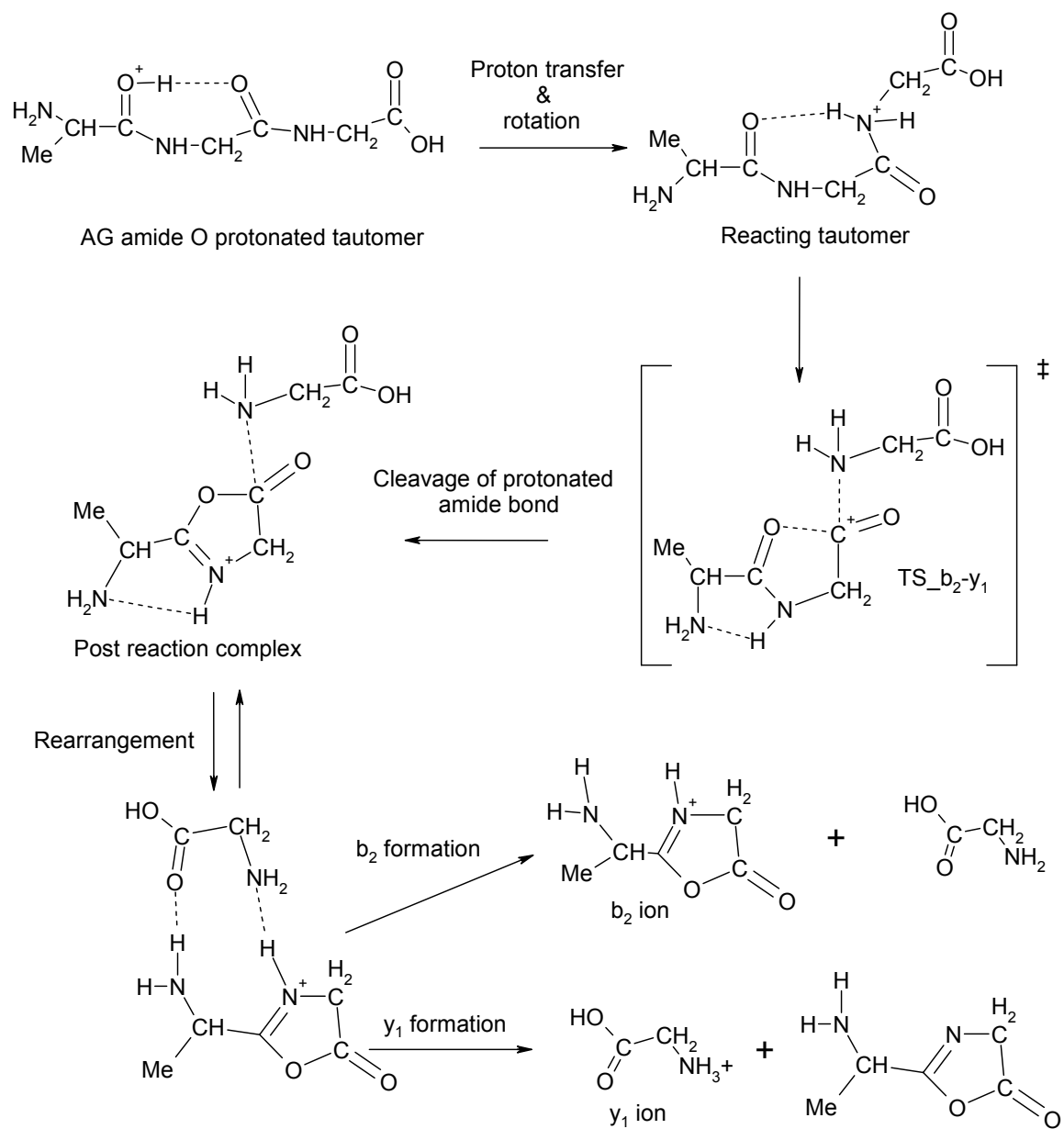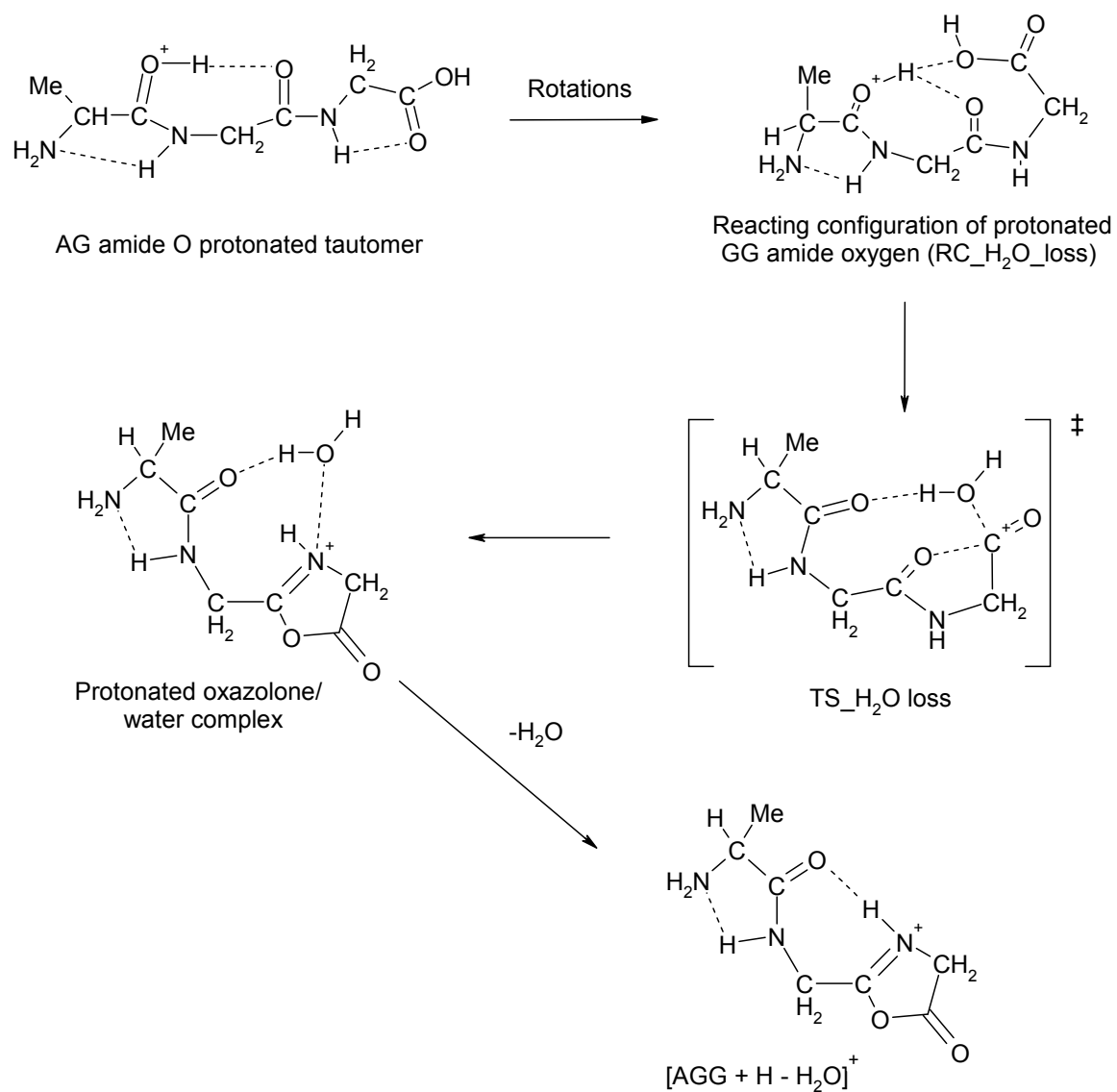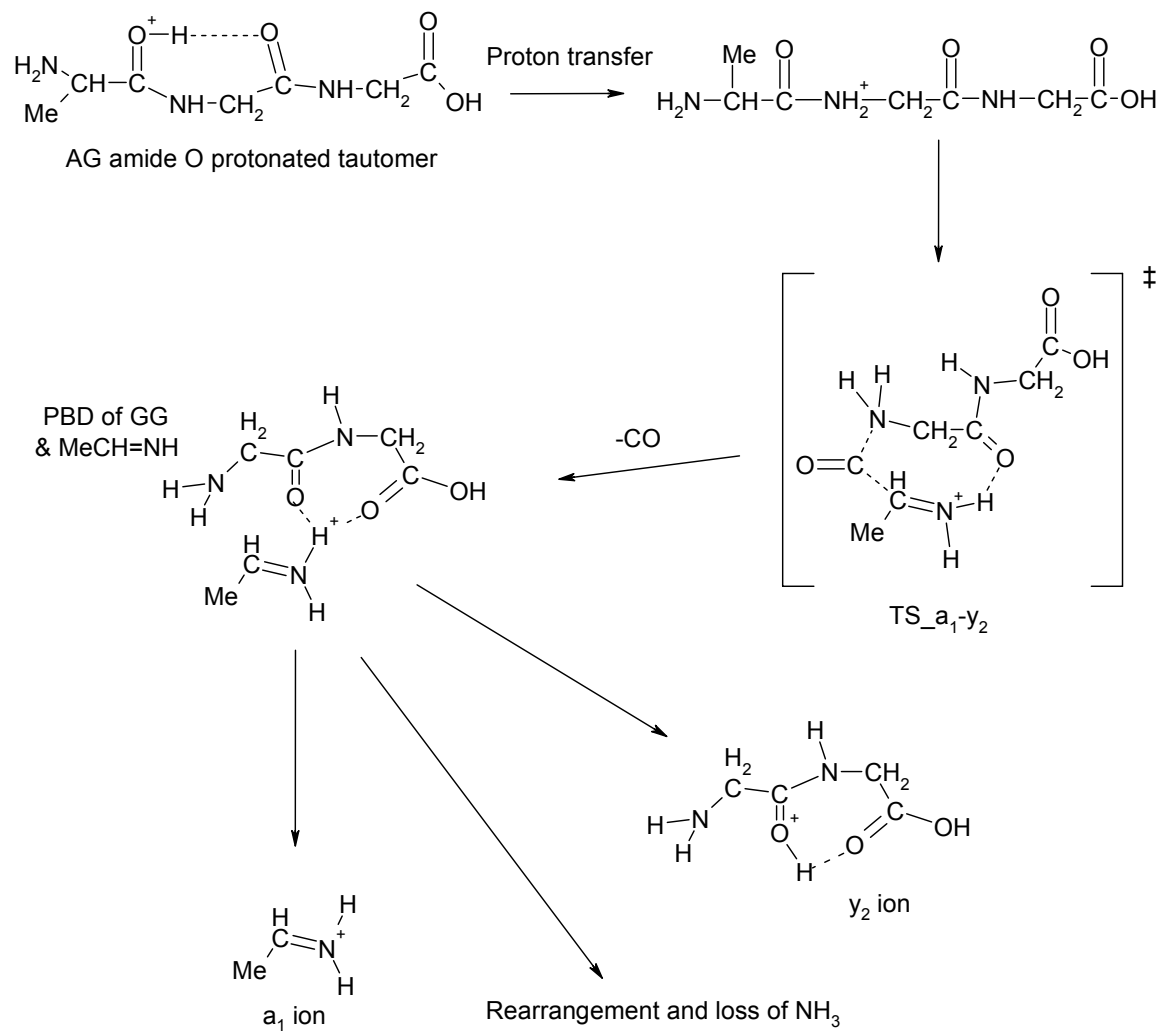
Figure 2.3: Relative energetics of paths 1 and 2 for $[AGG + H - CO - NH_3]^+$ ion formation. The relative energies (derived from B3LYP/6-31+G(d,p) total energies corrected for zero-point energy at B3LYP/6-31G(d)) were determined with respect to the global minimum on the PES of $[AGG + H]^+$. The energetics of the b_2-y_1 and water loss TSs and exit channels are shown for comparison.

Scheme 2.1 The b_2-y_1 PFP of protonated AGG
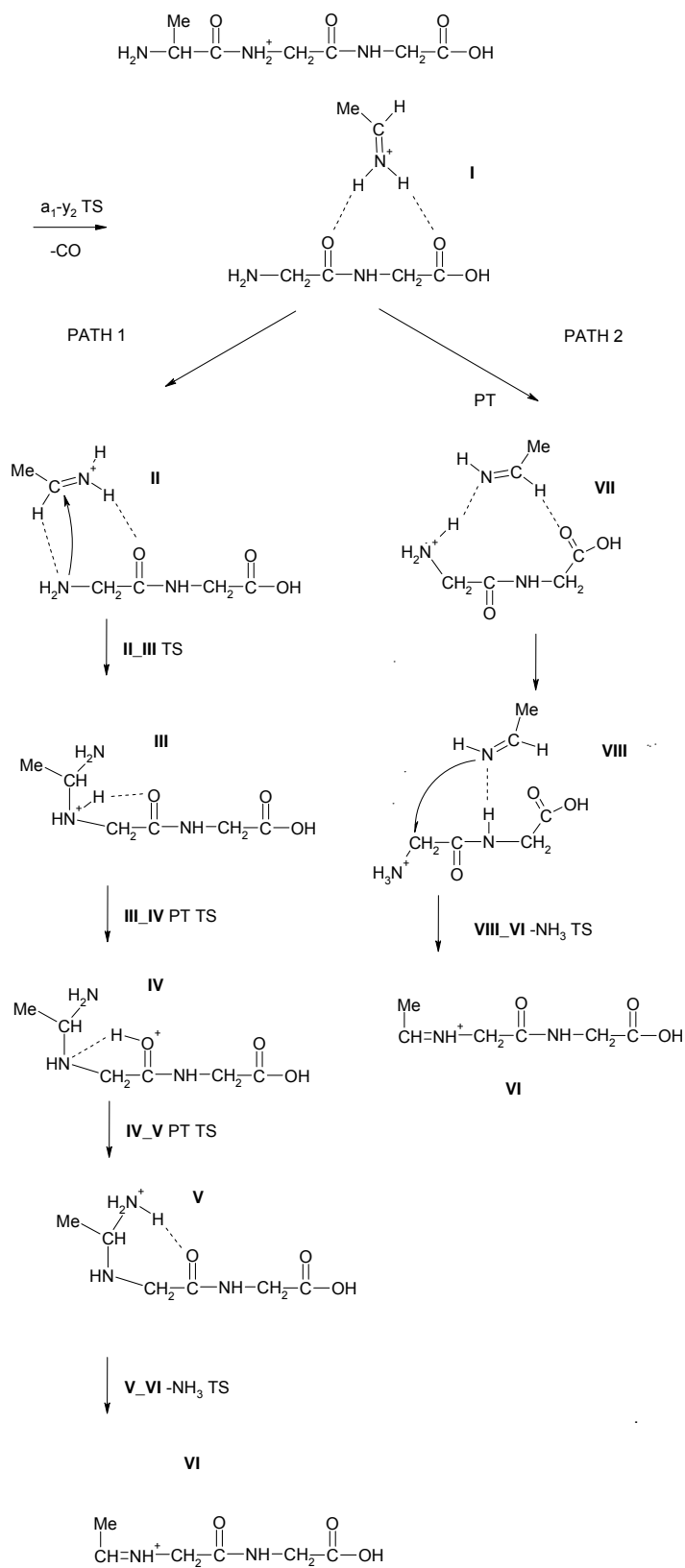

AG amide O protonated tautomer

Reacting tautomer

Post reaction complex

TS_b_2-y_1

Proton transfer & rotation

Cleavage of protonated amide bond

Rearrangement

b_2 formation

y_1 formation

b_2 ion

y_1 ion

Scheme 2.2 The water-loss PFP of protonated AGG

AG amide O protonated tautomer

Rotations

Reacting configuration of protonated
GG amide oxygen (RC_H$_2$O_loss)

TS_H$_2$O loss ‡

Protonated oxazolone/
water complex

-H$_2$O

[AGG + H - H$_2$O]$^+$

Scheme 2.3a The a_1-y_2 PFP of protonated AGG

AG amide O protonated tautomer

Proton transfer

TS_a_1-y_2

PBD of GG & MeCH=NH

-CO

y_2 ion

a_1 ion

Rearrangement and loss of NH_3

Scheme 2.3b Sequential loss of CO and NH₃ from protonated AGG

2.7 References

1) Barber, M.; Bordoli, R.A.; Sedgewick, R. D.; Tyler, A. N. Fast atom bombardment of solids (F.A.B.): a new ion source for mass spectrometry, *J. Chem. Soc., Chem. Commun.* 1981, 7, 325-327.

2) Yamashita, M.; Fenn, J. B.; Electrospray ion source: Another variation on the free-jet theme. *J. Phys. Chem.* **1984**, 88, 4451-4459.

3) Karas, M.; Bachmann, D.; Bahr, U.; Hillenkamp, F. Matrix-assisted ultraviolet laser desorption of nonvolatile compounds. *Int. J. Mass Spectrom. Ion Proc.*, **1987**, 78, 53-68.

4) Karas, M.; Hillenkamp, F. Laser desorption ionization of proteins with molecular masses exceeding 10,000 Daltons. *Anal. Chem.*, **1988**, 60, 2299-2301.

5) a)Biemann, K. Sequencing of peptides by tandem mass spectrometry and high-energy collision-induced dissociation, *Methods Enzymol.*, 1990, 193, 455-479. b) Biemann, K. Peptides and proteins: overview and strategy, *Methods Enzymol.*, **1990**, 193, 351-360.

6) Papayannopoulos, I. A. The interpretation of collision-induced dissociation tandem mass spectra of peptides, *Mass Spectrom. Rev.*, **1995**, 14(1), 49-73.

7) Paizs, B.; Suhai, S. Fragmentation pathways of protonated peptides, *Mass Spectrom. Rev.*, **2005**, 24, 508-548.

8) Wang,P.; Kish, M.M.; Wesdemiotis, C. Fragmentation mechanisms of peptide ions. Encycl. Mass Spectrom. 2005, 2, 139-151

9) Ballard, K.D.; Gaskell, S. J. Dehydration of Peptide [M H]$^{+}$ Ions in the Gass Phase. *J. Am. Soc. Mass. Spectrom.*, **1993**, 4, 477-481.

10) Cordero, M. M.; Houser, J. J.; Wesdemiotis, C. The neutral products formed during backbone fragmentations of protonated peptides in tandem mass spectrometry, *Anal. Chem.* **1993**, 65, 1594-1601.

11) Yalcin, T.; Khouw, C.; Csizmadia, I. G.; Peterson, M. R.; Harrison, A. G. Why are b ions stable species in peptide spectra? *J. Am. Soc. Mass. Spectrom.,* **1995**, 6(12), 1165-1174.

12) Yalcin, T.; Csizmadia, I. G.; Peterson, M. R.; Harrison, A. G. The structure and fragmentation of B_n (n >= 3) ions in peptide spectra, *J. Am. Soc. Mass. Spectrom.*, **1996**, 7, 233-242.

13) Dongré, A. R.; Somogyi, A.; Wysocki, V. H. Surface-induced dissociation: an effective tool to probe structure, energetics and fragmentation mechanisms of protonated peptides. *J. Mass Spectrom.* **1996**, 31(4), 339-350.

14) Ambihapathy, K.; Yalcin, T; Leung, Hei-Wun; Harrison, A. G. Pathways to immonium ions in the fragmentation of protonated peptides. *J. Mass Spectrom.* **1997**, 32, 209-215.

15) Nold, M.J.; Wesdemiotis, C.; Yalcin, T.; Harrison, A. G.; Amide bond dissociation in protonated peptides. Structures of the N-terminal ionic and neutral fragments. *Int. J. Mass Spectrom. Ion Proc.* **1997**, 164, 137-153.

16) Klassen, J.S.; Kebarle, P. Collision-Induced Dissociation Threshold Energies of Protonated Glycine, Glycinamide, and Some Related Small Peptides and Peptide Amino Amides. *J. Am. Chem. Soc.* **1997**, 164, 6552-6563.

17) Reid, G. E.; Simpson, R. J.; O'Hair, R. A. J.   A mass spectrometric and ab initio study of the pathways for dehydration of simple glycine and cysteine-containing peptide $[M+H]^+$ ions. *J. Am. Soc. Mass Spectrom.*, **1998**, 9(9), 945-956.

18) O'Hair, R. A. J.; Styles, M. L.; Reid, G. E.   Role of the sulfhydryl group on the gas phase fragmentation reactions of protonated cysteine and cysteine containing peptides.   *J. Am. Soc. Mass Spectrom,* **1998**, 9(12), 1275-1284.

19) Reid, G.E.; Simpson, R. J.; O'Hair, R. A. J. Probing the fragmentation reactions of protonated glycine oligomers via multistage mass spectrometry and gas phase ion molecule hydrogen/deuterium exchange. *Int. J. Mass Spectrom. Ion Proc.*, **1999**, 190/191, 209-230.

20) Nold, M. J.; Cerda, B. A.; Wesdemiotis, C. Proton affinities of the N- and C-terminal segments arising upon the dissociation of the amide bond in protonated peptides. *J. Am. Soc. Mass Spectrom.,* **1999**, 10, 1-8.

21) Tsaprailis, G.; Nair, H.; Somogyi, A.; Wysocki, V. H.; Zhong, W.; Futrell, J. H.; Summerfield, S. G.; Gaskell, S. J. Influence of Secondary Structure on the Fragmentation of Protonated Peptides. *J. Am. Chem. Soc.,* **1999**, 121, 5142-5154.

22) Paizs, B.; Lendvay, G.; Vékey, K.; Suhai, S. Formation of b_2^+ ions from protonated peptides: an ab initio study. *Rapid Commun. Mass Spectrom.*, **1999**, 13, 525-533.

23) Paizs, B.; Szlavik, Z.; Lendvay, G.; Vékey, K.; Suhai, S. Formation of a_2^+ ions of protonated peptides. An ab initio study. *Rapid Commun. Mass Spectrom.*, **2000**, 14, 746-755.

24) Harrison, A. G.; Csizmadia, I. G.; Tang, T.-H..; Tu, Y.-P. Reaction competition in the fragmentation of protonated dipeptides. *J. Mass Spectrom.,* **2000**, 35, 683-688.

25) Polce, M. J.; Ren, D; Wesdemiotis, C. Dissociation of the peptide bond in protonated peptides. *J. Mass Spectrom.,* **2000**, 35(12), 1391-1398.

26) Wysocki, V. H.; Tsaprailis, G.; Smith, L. L.; Breci, L. A. Mobile and localized protons: a framework for understanding peptide dissociation. *J. Mass Spectrom.,* **2000**, 35, 1399-1406.

27) Laskin, J.; Denisov, E.; Futrell, J. H. A Comparative Study of Collision-Induced and Surface-Induced Dissociation. 1. Fragmentation of Protonated Dialanine. *J. Am. Chem. Soc.,* **2000**, 122, 9703-9714.

28) Paizs, B.; Suhai, S. Theoretical study of the main fragmentation pathways for protonated glycylglycine. *Rapid Commun. Mass Spectrom.*, **2001**, 15, 651-663.

29) Paizs, B.; Suhai, S. Combined quantum chemical and RRKM modeling of the main fragmentation pathways of GGG II. Formation of b_2, y_1 and y_2 ions. *Rapid Commun. Mass Spectrom*., **2002**, 16, 375-389.

30) Paizs, B.; Suhai, S.; Harrison, A. G. Experimental and theoretical investigation of the main fragmentation pathways of protonated H-Gly-Gly-Sar-OH and H-Gly-Sar-Sar-OH. *J. Am. Soc. Mass Spectrom.,* **2003**, 14, 1454-1469.

31) Balta, B.; Aviyente, V.; Lifshitz, C. Elimination of Water from the Carboxyl Group of GlyGlyH[+]. *J. Am. Soc. Mass Spectrom*., **2003**, 14, 1192-1203.

32) Pingitore, F.; Polce, M. J.; Wang, P.; Wesdemiotis, C.; Paizs, B. Intramolecular condensation reactions in protonated dipeptides: carbon monoxide, water, and ammonia losses in competition. *J. Am. Soc. Mass. Spectrom*., **2004**, 15(7), 1025-1038.

33) Dongre, A. R.; Jones, J. L.; Somogyi, A.; Wysocki, V. H. Influence of Peptide Composition, Gas-Phase Basicity and Chemical Modification on Fragmentation Efficiency: Evidence for the Mobile Proton Model, *J. Am. Chem. Soc.* **1996**, *118*, 8365-8374.

34) Somogyi, Á.; Wysocki, V.H.; Mayer, I. The Effect of Protonation Site on Bond Strengths in Simple Peptides: Application of Ab Initio and MNDO Bond Orders and MNDO Energy Partitioning *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 704-717.

35) Harrison, A. G. Linear free energy correlations in mass spectrometry. *J. Mass Spectrom.*, **1999**, 34, 577-589.

36) Paizs, B.; Suhai, S. Towards understanding some ion intensity relationships for the tandem mass spectra of protonated peptides. *Rapid Commun. Mass Spectrom.* **2002**, 16, 1699-1702.

37) Harrison, A. G.; Young, A. B.; Bleiholder, B.; Suhai, S.; Paizs, B. Scrambling of Sequence Information in Collision-Induced Dissociation of Peptides. *J. Am. Chem. Soc.* **2006**, *128*, 10364-10365.

38) Cooper, T.; Talaty, E.; Grove, J.; Suhai, S.; Paizs, B.; Van Stipdonk, M. Isotope Labeling and Theoretical Study of the Formation of a_3* Ions from Protonated Tetraglycine. *J. Am. Soc. Mass Spectrom.* **2006**, in press available online August 24th.

39) Kinser, R. D.; Ridge, D. P.; Hvistendahl, G.; Rasmussen, B.; Uggerud, E. The unimolecular chemistry of protonated glycinamide and the proton affinity of glycinamide mass spectrometric experiments and theoretical model, *Chem.-Eur.J.*, **1996**, 2, 1143-1149.

40) March, R. E. An Introduction to Quadrupole Ion Trap Mass Spectrometry. *J. Mass Spectrom*. **1997**, 32, 351-369.

41) March, R. E. Quadrupole Ion Trap Mass Spectrometry: Theory, Simulations, Recent Developments and Applications. *Rapid Commun. Mass Spectrom.* **1998**, 12, 1543-1554.

42) Medzihradszky, K. F.; Campbell, J. M.; Baldwin, M. A.; Falick, Arnold M.; Juhasz, P.; Vestal, M. L.; Burlingame, A. L. The Characteristics of Peptide Collision-Induced Dissociation Using a High-Performance MALDI-TOF/TOF Tandem Mass Spectrometer. *Anal. Chem.*, **2000**, 72(3), 552-558.

43) Juhasz, P.; Campbell, J. M.; Vestal, M. L. MALDI-TOF/TOF technology for peptide sequencing and protein identification in *Mass Spectrometry and Hyphenated Techniques in Neuropeptide Research*, Edited by Silberring, J.; Ekman, R., **2002**, 375-413, Wiley-Interscience, 1st Edition.

44) Bienvenut, W. V.; Deon, C.; Pasquarello, C.; Campbell, J.M.; Sanchez, J.; Vestal, M.L.; Hochstrasser, D.F.   Matrix-assisted laser desorption/ ionization-tandem mass spectrometry with high resolution and sensitivity for identification and characterization of proteins,   *Proteomics*, **2002**, 2(7), 868-876.

45) Polce, M. J.; Cordero, M. M.; Wesdemiotis, C.; Bott, P. A. A New Trisector Tandem Mass Spectrometer for Neutralization-Reionization Studies. *Int. J. Mass Spectrom. Ion Proc.,* **1992**, 113, 35-58.

46) McLuckey, S. A.; Van Berkel, G. J.; Goeringer, D. E.; Glish, G. L. Ion Trap Mass Spectrometry of Externally Generated Ions. *Anal. Chem*. **1994**, 66,689A-696A.

47) Gaussian 03, Revision C.02, Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, Jr., J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone,

V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; and Pople, J. A.; Gaussian, Inc., Wallingford CT, **2004**.

48) Baer, T.; Hase, W. L. *Unimolecular Reaction Dynamics*, **1996**, Oxford University Press, Oxford.

49) El Aribi, H.; Orlova, G.; Rodriquez, C.F.; Almeida, D. R. P.; Hopkinson, A.C.; Siu, K. W. M. Fragmentation Mechanisms of Product Ions from Protonated Tripeptides, *J. Phys. Chem. B,* 2004, 108(48),18743 -18749.

50) Pingitore, F.; Wesdemiotis, C.; Paizs, B. Competitive Loss of Small Molecules in Peptides, *Proceedings of the 52*nd *ASMS Conference on Mass Spectrometry and Allied Topics*, Nashville, Tennessee, May 23-27, **2004**.

51) Lifshitz, C. Kinetic shifts. *Eur. J. Mass Spectrom.*, **2002**, 8, 85-98.

52) Rodriquez, C.F.; Cunje, A.; Shoeib, T.; Chu, I. K.; Hopkinson, A.C.; Siu, K. W. M. Proton migration and tautomerism in protonated triglycine, *J. Am. Chem. Soc.,* **2001**, 123, 3006-3012.

53) Paizs, B.; Suhai, S. Combined quantum chemical and RRKM modeling of the main fragmentation pathways of GGG. I. *Cis-trans* isomerization around protonated amide bonds, *Rapid Commun. Mass Spectrom.*, **2001**, 15, 2307-2323.

54) Polfer, N. C.; Oomens, J.; Suhai, S.; Paizs, B. Spectroscopic and Theoretical Evidence for Oxazolone Ring Formation in Collision-Induced Dissociation of Peptides *J. Am. Chem. Soc.* **2005**, *127*, 17154-17155.

55) Harrison, A. G.; The gas-phase basicities and proton affinities of amino acids and peptides. *Mass Spectrom. Rev.* **1997**, 116, 201-217.

56) Bleiholder, C.; Suhai, S.; Paizs, B. Revising the Proton Affinity Scale of the Naturally Occurring α-Amino Acids *J. Am. Soc. Mass Spectrom.* **2006**, 17, 1275-1281.

57) Lifshitz, C. Some recent aspects of unimolecular gas phase ion chemistry. *Chem. Soc. Rev.*, **2001**, 30, 186-192.

58) Csonka, I. P.; Paizs, B.; Lendvay, G.; Suhai, S. Proton mobility in protonated peptides: a joint molecular orbital and RRKM study. *Rapid. Commun. Mass Spectrom.* **2000**, 14, 417-431.

59) Paizs, B.; Csonka, I. P.; Lendvay, G.; Suhai, S. Proton mobility in protonated glycylglycine and N-formylglycylglycinamide: a combined quantum chemical and RKKM study. *Rapid Commun. Mass Spectrom.* 2001, 15, 637-650.

60) Laerdahl, J. K.; Uggerud, E. Gas phase nucleophilic substitution *Int. J. Mass Spectrom*, **2002**, 214,277-314.

61) Gritsenko, O. V.; Ensing, B.; Schipper, P. R. T.; Baerends, E. J.; Comparison of the Accurate Kohn-Sham Solution with the Generalized Gradient Approximations (GGAs) for the S_2N Reaction F- + CH_3F -> FCH_3 + F-: A Qualitative Rule To Predict Success or Failure of GGAs, *J. Phys. Chem. A.* **2000**, *104,* 8558-8565.

62) Tu, Y. -P.; Harrison, A. G. Fragmentation of Protonated Amides Through Intermediate Ion–Neutral Complexes: Neighboring Group Participation. *J. Am. Soc. Mass Spectrom.,* **1998**, 9(5), 454-462.

63) McLafferty, F. W.; Turecek, F. *Interpretation of Mass Spectra*, 4th ed. University Science Books: Mill Valley, CA, **1993,** 51-84.

2.8 Supplementary Information For Chapter 2: *J. Am. Soc. Mass Spectrom.*, **2007**, 7, 1291-1303.

Backbone Cleavages and Sequential Loss of Carbon Monoxide and Ammonia from Protonated AGG: A Combined Tandem Mass Spectrometry, Isotope Labeling, and Theoretical Study

Benjamin J. Bythell and Douglas F. Barofsky

Department of Chemistry, Oregon State University, Corvallis, Oregon, USA

Francesco Pingitore, Michael J. Polce, Ping Wang and Chrys Wesdemiotis

Department of Chemistry, The University of Akron, Akron, Ohio, USA

Béla Paizs*

German Cancer Research Center, Department of Molecular Biophysics, Heidelberg, Germany

Figure 2.SI1: MALDI-TOF/TOF spectra showing loss of CO and NH_3 for protonated aliphatic peptides; (a) GGA, (b) GGL, (c) GGGG and (d) YGG
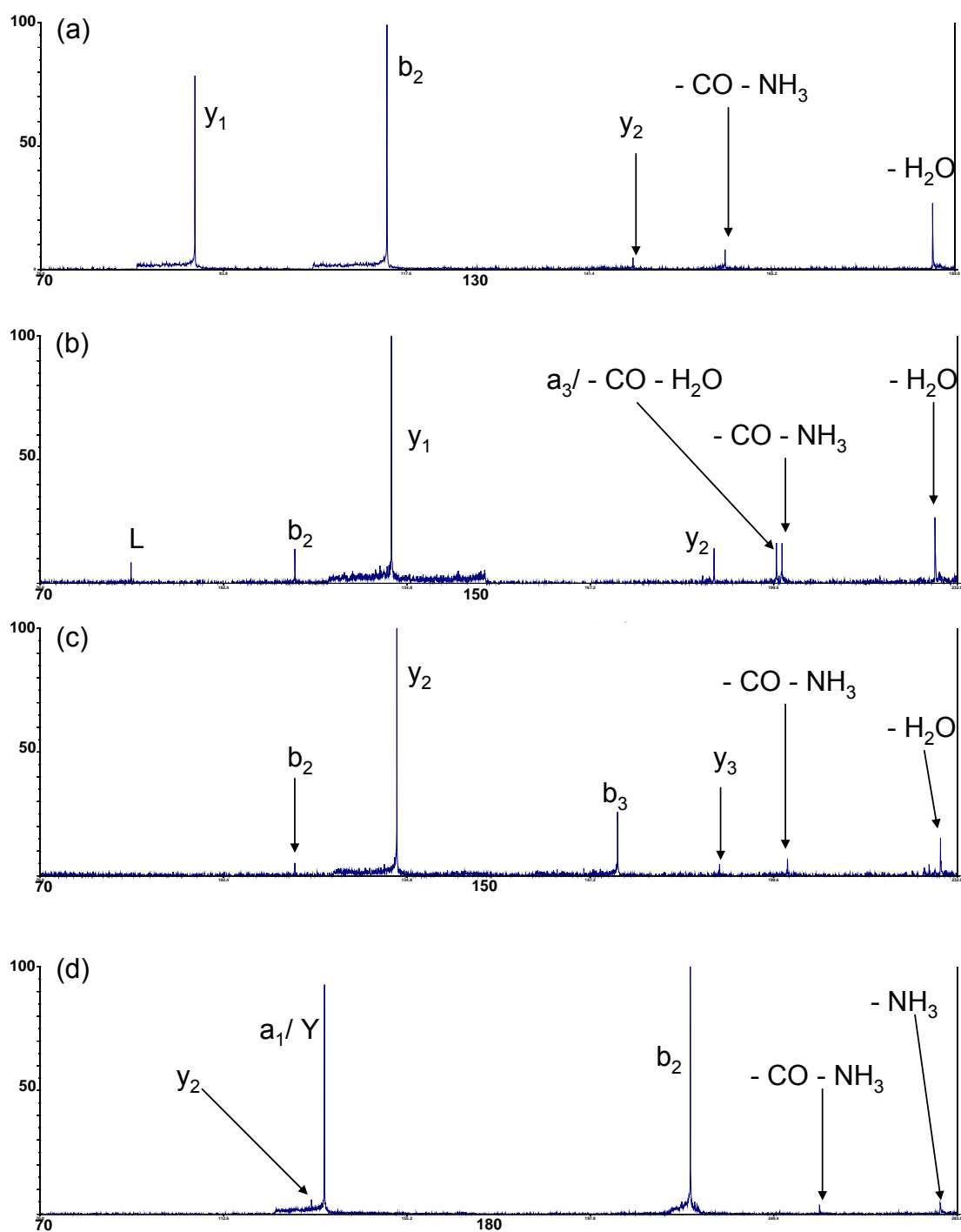
Figure 2.SI2: MALDI-TOF/TOF spectra showing no loss of CO and NH_3 for protonated aliphatic peptides that lack G as residue two; (a) AAA, (b) GPGG
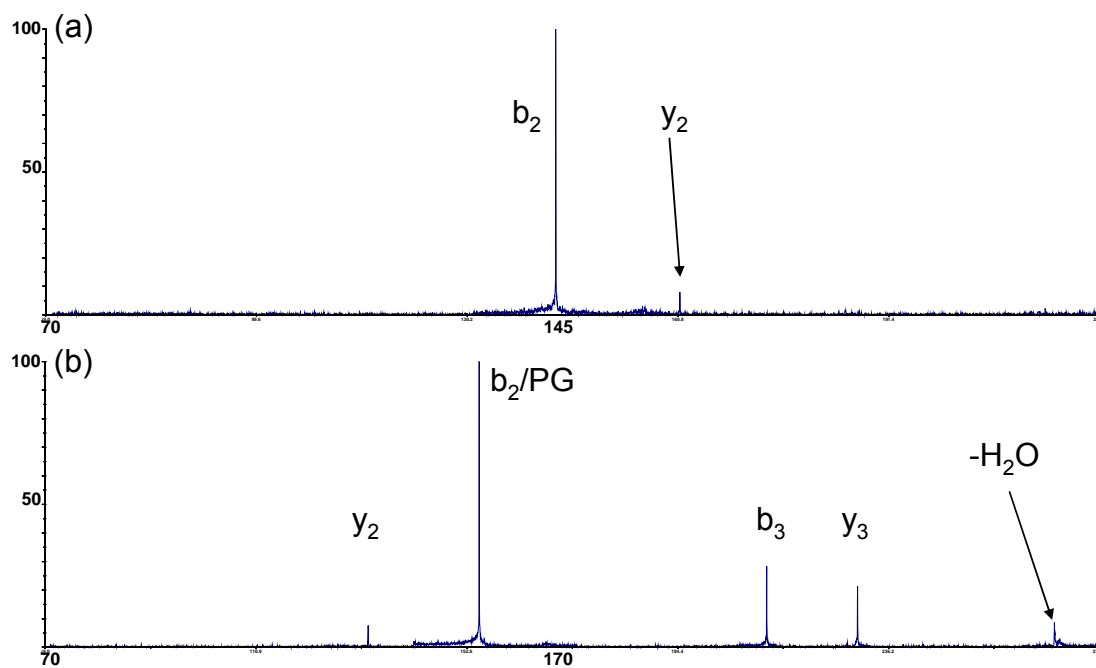
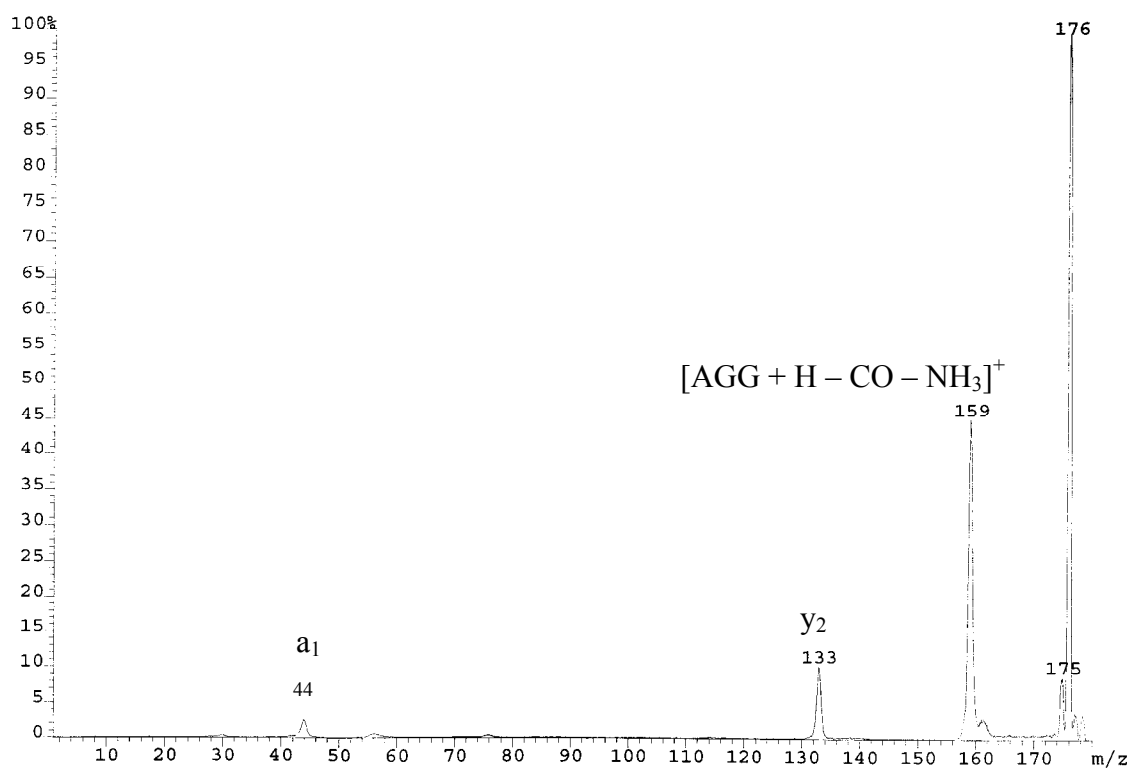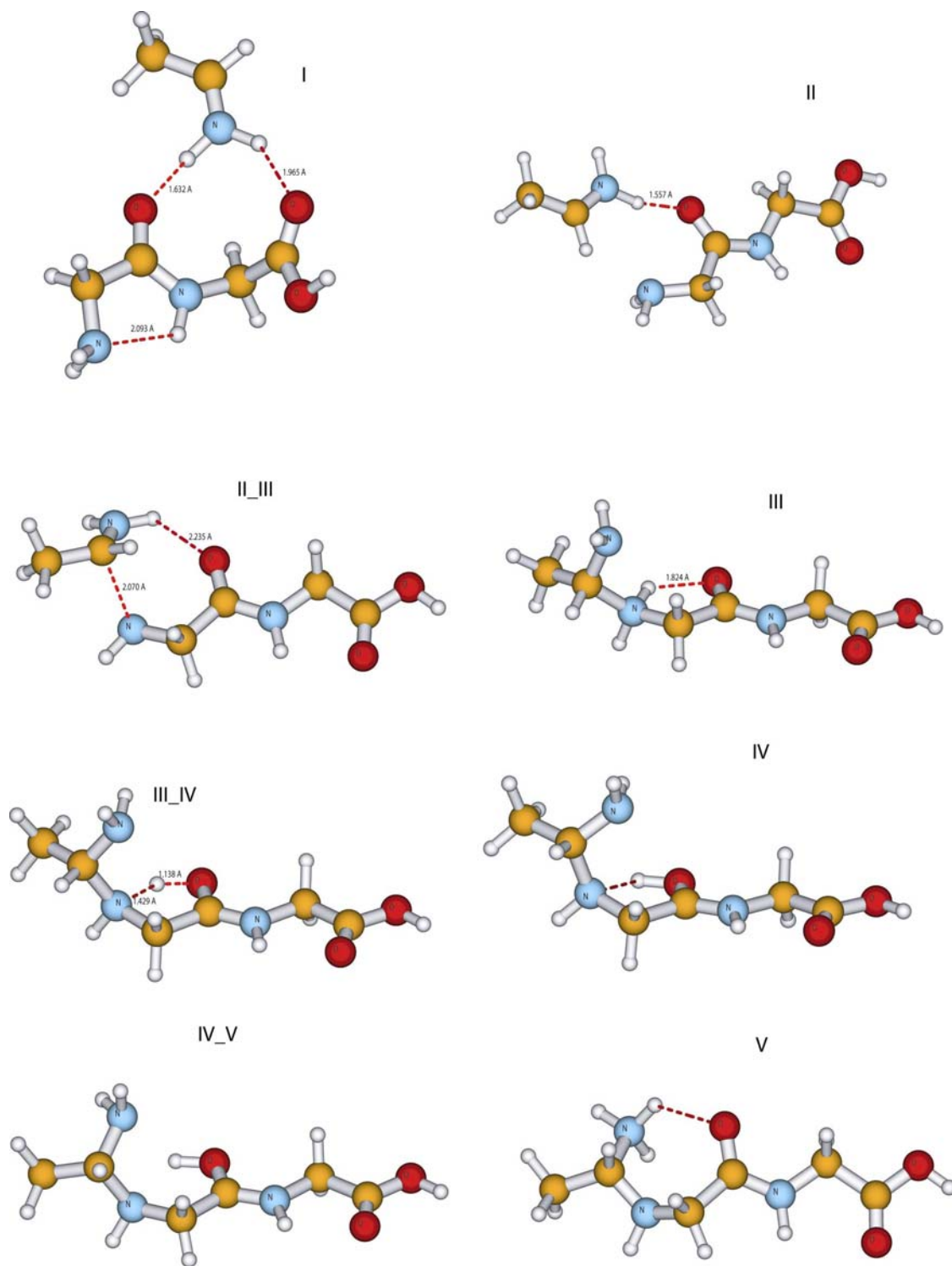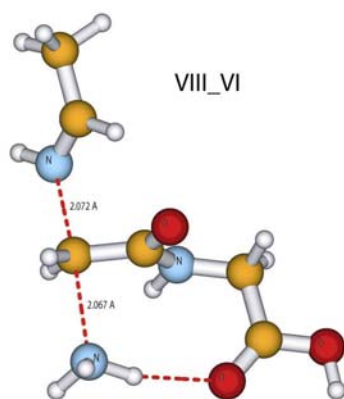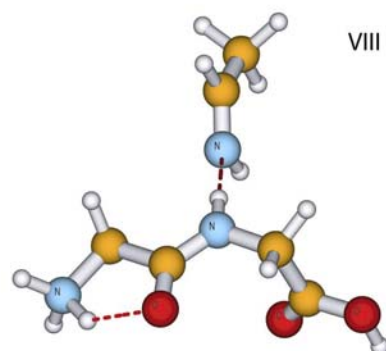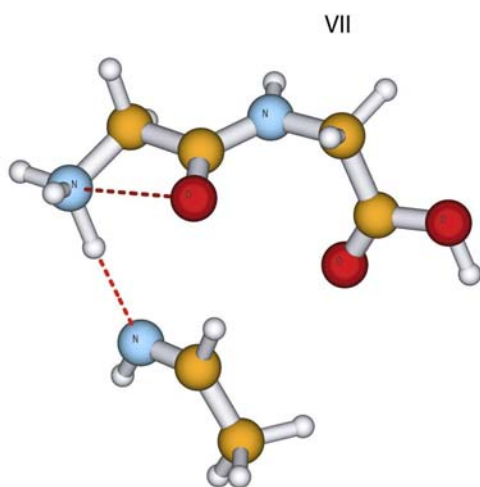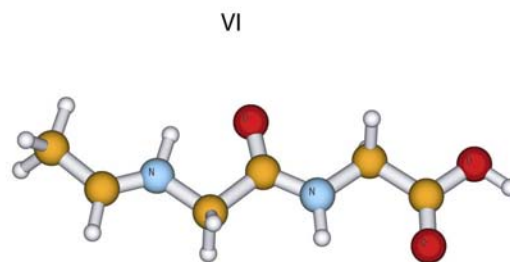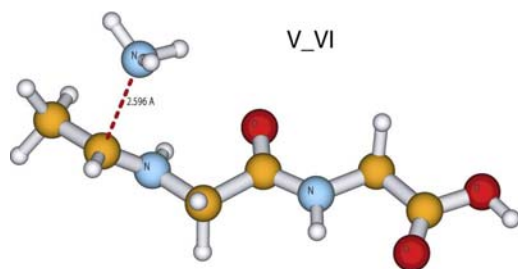Figure 2.SI3: Sector CAD MS^3 (176m/z) spectrum of [AGG + H – CO]$^+$

Figure 2.SI4: Species depicted in Scheme 2.3

Chapter 3

Novel Salt-Bridge Stabilized Peptide Fragmentation Mechanisms in Protonated RGD

Benjamin J. Bythell, Béla Paizs and Douglas F. Barofsky

3.2 Introduction

Tandem mass spectrometry (MS/MS) using collision-activated-dissociation (CAD) to induce bond cleavages is the standard technique used to determine the primary amino acid sequence of peptides [1-3] derived from proteins by enzymatic digestion. Under low energy collision conditions, protonated peptides fragment to predominantly form N-terminal b and C-terminal y fragment ions [4, 5]. Unfortunately, these are not the only fragment ion types observed. Incomplete sequence coverage and/or unusual dissociation processes are reasonably common. This makes peptide sequencing using MS/MS a much more complicated problem [6, 7]. Variation in peptide fragmentation chemistry is one of the reasons why current algorithms that search databases for matches to MS/MS spectra, such as SEQUEST [8] and Mascot, [9] mainly make use of the fragment ion m/z values and neglect their intensities completely. An improved understanding of the underlying dissociation chemistry would enable the incorporation of mechanistic information on the peptide fragmentation pathways (PFPs) [10, 11] into these algorithms and thereby facilitate more accurate identification of peptide sequence.

The model of peptide fragmentation most commonly in use is the mobile proton model [12, 13]. This model states that upon excitation the ionizing proton is transferred from an unreactive site of higher gas-phase basicity (e.g. R or K side chain or the N-terminal amino group) to a less energetically favored, but reactive backbone amide-protonated species. Protonation of the amide nitrogen leads to considerable weakening of the amide bond [14], and a species such as this plays a critical role in most of the PFPs that lead to sequence-informative b, a, and y fragment ions [10].

The mobile proton model is based on a broad range of empirical peptide dissociation [16-27] and modeling [28-30] studies. The model allows one to predict whether a given peptide in a particular protonation state is expected to produce a sequence-informative MS/MS spectrum, or whether selective cleavage at some amide bonds are likely to hinder MS-based peptide sequencing [31]. For example, peptide ions whose number of added protons exceeds the number of R and K residues are expected to fragment at the various amide bonds forming b, a and y ions. On the other hand, substantially less sequence coverage is expected if the number of added protons is equal to the number of R or K residues. This is especially so when the peptide contains aspartic acid, D, or glutamic acid, E, because cleavage C-terminal to these residues (the aspartic acid effect) is mainly observed. The mobile proton model does have limitations. Most significantly in this regard, it cannot give an accurate prediction of the relative intensities of fragment ions. In order to make such predictions, a deeper understanding of peptide dissociation is necessary.

These issues have led to the development of the pathways in competition (PIC) model [10]. This provides a more general framework to understand gas-phase peptide chemistry by taking into account the specific features of the individual peptide fragmentation pathways and their interaction. The PIC model states that fragment ion abundances in the MS/MS spectra of peptides are determined by pre-cleavage, amide bond-cleavage, and post-cleavage events. The pre-cleavage phase involves the proton-transfer (PT) reactions and/or internal rotations necessary to produce fragment-prone species. The main tenet of the mobile proton model is the activity (or inactivity) of proton-transfer pathways which result in peptide fragmentation. The activity of these

pathways, and the related predictive rules (see preceding), are directly incorporated into the PIC model along with the additional considerations of the bond-cleavage and post-cleavage phases of peptide fragmentation. In other words, PIC is a logical extension of the mobile proton model which considers other reactions other than the initial proton transfer to predict which fragment ions are likely to be formed and their relative abundances in MS/MS spectra.

The PIC model has had recent success in explaining that fragments held together in proton-bound dimers can undergo various transitions, including association reactions and subsequent dissociation of the reorganized parent ion [32, 33]. Harrison et al. also showed that linear b ions with a C-terminal oxazolone ring can rearrange to form a cyclic peptide [34]; fragmentation of this isomer could lead to scrambling of the primary sequence information. The related fragments have been termed non-direct sequence ions [34]. Inclusion of the post-cleavage phase of peptide fragmentation, which previously attracted surprisingly limited research activity, facilitated these findings. The added flexibility of PIC enables greater understanding of the rich chemistry of the post-cleavage phase of peptide fragmentation. Analogously, the fate of the added proton (that is to say, which fragment survives the dissociation as the charged species) can often be predicted considering fragment proton affinities (or gas-phase basicities) [35].

In the present article, we report on the fragmentation characteristics of protonated RGD, a peptide with basic and acidic groups in close proximity. Peptides with N-terminal arginine have been relatively understudied due to the almost exclusive use of trypsin (which cleaves at the C-terminal of R and K except when followed by P) to digest

proteins prior to mass spectrometric analysis. However, peptides with an N-terminal arginine can still be readily formed from a single missed cleavage at the C-terminal of R occurring in the sequence XRRZ or XKRZ (where X & Z are residues not cleaved C-terminally by trypsin), which is a fairly common occurrence in proteins (if Z = P, this process will be enhanced). Arginine location has a significant effect on the fragment ions observed [15, 17-21], and N-terminal arginine can promote less common PFPs [36-38]. Additionally, the presence of R and D offers the possibility of chemically relevant salt bridge (SB) structures being amongst the minima and fragmentation pathways under investigation.

We have investigated how the canonical charge-solvated (CS), and zwitterionic salt-bridge structures compare energetically and, in addition, what role, if any, they play in the charge transfer and dissociation chemistry of this peptide. Specifically, we examined the formation of six different ion types (b_1, b_2, b_2+H_2O, $[RGD+H-NH_3]^+$, $[RGD+H-H_2O]^+$, and $[RGD+H-59]^+$) using ion trap (IT) and MALDI-TOF/TOF tandem MS, DFT modeling, and RRKM rate calculations. We describe new mechanisms for the formation of the b_1, b_2, and b_2+H_2O ions, and clarify the energetics of the other pathways for comparison. The b_2+H_2O PFP involves salt bridge interactions and anhydride formation prior to b_2+H_2O ion formation in a multi-step reaction. Most significant, however, is the discovery of new salt-bridge stabilized b_2-ion forming pathways without which our experimental findings could not be explained. Both the C-terminus salt-bridge and the D side chain salt bridge structures result in transition structures (TSs) that are much lower in energy (≥ 8kcal mol^{-1}) than the normal charge solvated TSs. Thus, this discovery is likely to apply generally to peptides with R at the N-terminus.

In addition, neither the b_2 ion nor the b_2+H_2O ion forming mechanisms can be explained by the mobile proton model since post cleavage events are fundamental to both mechanisms producing the detected product ions. These findings support the more detailed PIC model, which considers post cleavage as well as pre-cleavage and bond-cleavage events. The additional features of the PIC model enable complicated fragmentation pathways, like those of $[RGD+H]^+$ to be rationalized.

3.3 Methods and Materials

3.3.1 Experimental Methods

Experiments were conducted on a Finnigan MAT ESI- LCQ (San Jose, CA) ion trap (IT) [39, 40] and an Applied Biosystems 4700 Proteomics Analyzer MALDI-TOF/TOF (Framingham, MA) [41-43]. All chemicals were purchased from Sigma-Aldrich (Milwaukee, WI) and used without further purification. Samples for the IT experiments were prepared by dissolving RGD in acetonitrile/water/acetic acid 30/70/0.1 (v:v) to form a $2x10^{-5}$ mol L^{-1} solution, which was infused into the ESI source at a rate of 10 µL min^{-1}. The entrance to the sampling capillary was set at -4kV, and nitrogen served as the nebulizing and drying gas (170°C). CAD of $[RGD + H]^+$, was performed by ejecting all ions except $[RGD + H]^+$, and then exciting the latter to fragment in the presence of He buffer gas (10^{-3} Torr) using a radiofrequency field ($0.78V_{p-p}$) [44] under automated gain control. The reproducibility of the relative abundances from multiple, repeated scans was better than ±10%. Additional scans were performed where the excitation level was incrementally increased. These scans began from well below the threshold of product ion formation and continued until all product ions had been observed.

The MALDI matrix was prepared by dissolving α-cyano-4-hydroxycinnamic acid in acetonitrile/water/trifluoroacetic acid/monoammonium phosphate (6mg/ml) 47/47/0.1/6 (v:v) solution at a concentration of 2mg/mL. The RGD peptide was dissolved in an acetonitrile/water/trifluoroacetic acid 50/50/0.1 (v:v) solution at a concentration of 100μg/mL. The sample solutions were then prepared by mixing the matrix/peptide solutions in a 1 to 1 ratio; 0.4μl of sample solution was applied to each spot on a 196-well target plate and allowed to air dry prior to introduction into the mass spectrometer. The MS/MS spectra consisted of 5000 laser shots per well, twelve replicate wells, with laser fluence constant and at a level low enough to prevent signal saturation. Air was used as the collision gas at 2.0×10^{-8} ("no gas") and 7.1×10^{-7} mTorr. The $[RGD + H]^+$ ions formed in the MALDI process were selected (200 FWHM) then decelerated to 1keV (laboratory frame) for metastable ion (MI) decomposition or CAD.

3.3.2 Computational Methods

A conformational search engine devised specifically to deal with protonated peptides was used to scan the potential energy surface (PES) of protonated arginylglycylaspartic acid, $[RGD + H]^+$. These calculations began with molecular dynamics (MD) simulations on various forms of $[RGD + H]^+$ using the InsightII program (Biosym Technologies, San Diego, USA), in conjunction with the AMBER force field modified by Paizs [45] to allow amide nitrogen and oxygen protonated species. During the MD simulations, structures were regularly saved for further refinement by full geometry-optimization using the same force fields. In the next stage of the process, these structures were analyzed by a conformer-family search program. This program groups

optimized structures into families based on similarity of the most important characteristic torsion angles. The most stable species in these families were then fully optimized at the HF/3-21G, B3LYP/6-31G(d), and the B3LYP/6-31+G(d,p) levels.

Having scanned the PES, TSs corresponding to the various fragmentation pathways of [RGD + H]$^+$ were then sought. These were calculated at the B3LYP/6-31G(d) and B3LYP/6-31+G(d,p) levels of theory (at OSU or the DKFZ). The resulting TSs were checked using intrinsic reaction coordinate (IRC) calculations to unambiguously define which minima are connected by the TS investigated. Post-reaction complexes and proton-bound dimers were fully optimized at the B3LYP/6-31G(d) and B3LYP/6-31+G(d,p) levels of theory in a manner similar to that used for the various RGD protonation sites and transition structures (at OSU or the DKFZ). Relative energies were calculated by using the B3LYP/6-31+G(d,p) total energies and zero-point energy corrections (ZPE) determined at the B3LYP/6-31G(d) level. The Gaussian [46] program was used for all *ab initio* calculations.

3.4 Results and Discussion

3.4.1 Protonation Energetics of [RGD + H]$^+$

The lowest-energy conformer of protonated RGD is a charge solvated (CS) structure (Table 3.1, SI 3.1). Despite the presence of the acidic D residue, the lowest-energy zwitterionic conformer (SB_D) is less stable by 12.6 kcal mol^{-1}.

3.4.2 Tandem mass spectra (MS/MS) of protonated RGD

The IT and MALDI-TOF/TOF mass spectra are summarized in Table 3.2. These experiments yielded similar N-terminal fragment ions but with large differences in fragment ion abundance. The CAD MALDI-TOF/TOF experiment imparts more energy than the MI experiment. This is characterized by the shift in the product spectrum to include: large immonium ion peaks, increased overall ion current (values not shown), and altered relative intensities of the fragment ions. The nominally 'high energy' d_3 ion is also observed in the MALDI CAD experiment. Water loss and b_2 - NH_3 peaks are observed in the IT but not in the MALDI spectra. Surprisingly, a small d_3 peak was observed in the IT spectra. This assignment was supported by MS^3 of this peak which led to b_2, a_2, N-terminal peaks. The d_3 peak was also the major product from MS^3 of the water loss peak indicating a mechanism involving consecutive losses of small neutrals (data not shown) must exist.

3.4.3 Fragment Peak Appearance Energies

Relative appearance-energies for the primary fragmentation channels were determined with the IT instrument by incrementally increasing the parent excitation level (collision energy). These scans showed that the b_2 ion peak appears immediately as a large peak at the lowest excitation level. The water and ammonia losses are next, but not initially as large peaks; the ammonia-loss peak becomes much larger as the excitation level is increased. Appearing next are the b_1 and b_2+H_2O ions with the b_1 peak being larger initially. Finally, the 59-loss (guanidine), b_2-NH_3, and the d_3 ion peaks are observed (In summary, b_2 < -H_2O ≈ -NH_3 < b_1 ≈ b_2+H_2O < 59-loss ≈ b_2-NH_3 ≈ d_3). In

these experiments, the relatively long time-scale (~10ms) of the IT mean kinetic shifts are small [49], so the order of the appearance energies is likely to reflect the order of the true threshold energies.

3.4.4 Fragmentation pathways of [RGD + H]$^+$

3.4.4.1 Salt-bridge stabilized b$_2$-y$_1$ peptide fragmentation pathways

Cleavage of the C-terminal amide bond of [RGD + H]$^+$ on the b$_2$-y$_1$ pathway results in a b$_2$ ion with an oxazolone structure [50-53]. Typically, this process involves canonical charge-solvated structures (Scheme 3.1a) with the R side chain being neutral. However for protonated RGD, there are lower energy b$_2$-y$_1$ pathways with zwitterionic transition states. These salt-bridge stabilized pathways contain a protonated R side chain and either a deprotonated C-terminus or a deprotonated aspartic acid side chain (Schemes 3.1b and 3.1c respectively). As these transition states are more stable by 8 or more kcal mol^{-1} than the canonical structures, they offer a plausible explanation for the spectral peak intensities (Table 3.2) and order of peak appearance energy experiment. This finding should offer insight into b-ion formation in other peptides with N-terminal arginine (e.g. bradykinin and its analogues). For comparison, the lowest b$_2$-y$_1$ transition states of each family are illustrated in Figure 3.1.

The first step on the b$_2$-y$_1$ pathway is formation of SB_Cterm by transferring the C-terminal proton to the N-terminal amine group. This proton is then transferred to the nitrogen of the C-terminal amide bond (SB_N2). Cleavage of the protonated C-terminal amide bond and simultaneous formation of the oxazolone ring take place through a TS

stabilized by the salt bridge between the deprotonated C-terminus and the protonated R side chain (Scheme 1b, Figure 3.1). After leaving the SB_Cterm_b_2-y_1 TS, various salt-bridge bound dimers of $[RG_{oxa} + 2H]^{2+}$ and $[D – H]^-$ are formed. Eventually, because the proton affinity of the $[D – H]^-$ is much greater than that of the oxazolone ring, proton transfer to the aspartic acid anion occurs resulting in the formation of the normal b_2 oxazolone ion and neutral aspartic acid.

3.4.4.2 Loss of water from protonated RGD

Loss of water occurred only in the IT despite requiring only 40.7 kcal mol^{-1} to initiate. This reaction begins with a charge solvated species formed from the global minimum CS conformer by rotation of the aspartic acid side chain and the C-terminus without mobilization of the ionizing proton (RC_H$_2$O_loss, Figure 3.2). Once RC_H$_2$O_loss is populated, nucleophilic attack by an aspartic acid carboxyl oxygen on the C-terminal carbonyl group and transfer of the proton from the aspartic acid side chain to the C-terminal OH are required to reach the H$_2$O_loss TS. From the H$_2$O_loss TS, a cyclic anhydride is formed as the water molecule is expelled in a concerted manner (Figure 3.2).

The alternative mechanism of water loss begins with proton transfer from the arginine side chain to the C-terminal OH. This pathway results in the normal b_3 oxazolone-forming [33, 54] H$_2$O loss TS with the water loss being stabilized by hydrogen bonding between the C-terminal proton and the R carbonyl oxygen. However, this TS (H$_2$O_loss_Ox, Table 3.1, Figure SI3.1) is much less stable than H$_2$O_loss TS, and is therefore, less likely to be active.

3.4.4.3 Loss of ammonia from [RGD + H]$^+$

Loss of ammonia from the R side chain is the largest peak under the majority of spectral conditions. To evaluate this fragmentation pathway, both CS and SB structures (with the preformed NH$_3$ moiety on the R side chain) were examined as potential reactive configurations (SI Figure 3.1). Transition structures were then sought from each of these configurations. CS structures (AMM_CS) were found to be more favored for both the reactive configurations and transition states. The lowest energy TS (CS_NH$_3$_loss, Figure 3.3, Table 3.1) was formed from an AMM_CS_B reactive configuration.

3.4.4.4 The b$_1$-y$_2$ PFP

Normally, b$_1$ peaks are only observed when the N-terminal residue is K or R; this experimental finding indicates indicating that these ions have an unusual structure. A cyclization reaction involving nucleophilic attack on the arginine carbonyl carbon by the arginine side chain was previously proposed to explain the b$_1$ ion formation in RG-NH$_2$ [55]. Calculations performed to evaluate this proposed mechanism (Scheme 3.2) for protonated RGD confirmed that it was energetically feasible at 42.6 kcal mol^{-1} (CS_b$_1$-y$_2$, Table 3.1, Figure 3.4) and thus, that the b$_1$ ion is likely to be a lactam structure.

3.4.4.5 The b$_2$+H$_2$O PFP

Formation of so-called b$_{n-1}$ + H$_2$O ion (where n = total number of residues in the protonated peptide) usually involves the loss of the C-terminal amino acid residue to give a product ion having a carboxyl group at the new C-terminus [36]. Loss of more than one C-terminal residue can occur, but it is much less common and usually leads to minor

peaks. Formation of b_{n-1} + H_2O ion is promoted by the presence of basic amino acids usually in a non-C-terminal position. The mechanism by which this fragment ion type is formed has been debated in the literature for some time [36-38, 56-58]. Experimental, statistical, and theoretical work has led to wide ranging mechanistic speculation (in some cases involving cyclizations and/ or salt-bridge interactions); however, no transition structure calculations have been published in support of these proposals. In an attempt to resolve this controversy, we have investigated these proposed mechanisms. In light of our findings, we propose a new mechanism involving salt bridges and formation of an anhydride reactive configuration (RC) supported by a detailed description of the energetics involved from our DFT calculations.

The b_2+H_2O forming mechanism (Scheme 3.3, Figure 3.5) begins with transfer of a proton from the C-terminus to the GD amide nitrogen to form a salt bridge structure of the SB_N2 family. From this reactive configuration nucleophilic attack by a C-terminal oxygen on the glycine carbonyl carbon with simultaneous breaking of the GD amide bond (SB_b_2+H_2O_TS1) requires 28.8 kcal mol^{-1} and results in an anhydride structure (b_2+H_2O_Anhydride type). The anhydride is the reactive configuration for b_2+H_2O formation and a salt-bridge stabilized transition state (SB_b_2+H_2O_TS2) can be formed facilely from this anhydride. The reaction expels CO and transfers a proton from the D immonium ion to complete the formation of the new C-terminus of the b_2+H_2O ion.

The b_2+H_2O forming anhydride structure (b_2+H_2O_Anhydride type) is the same type as that proposed by Farrugia and O'Hair [57] and Mussinan et al. [58]; however, the mechanisms of formation are quite different. These authors proposed forming this

anhydride from an oxazolone alcohol via a four-center proton transfer which is potentially rate-limiting, with concerted amide bond cleavage (Appendix D1). This mechanism has two problems which make it unlikely to be active: the TS necessary to form the oxazolone alcohol requires more energy (31.2kcal mol^{-1}) than SB_b$_2$+H$_2$O_TS1, and four-center proton transfer reactions are very energetically unfavorable [28, 29]. The mechanism proposed by Thorne et al. [37] was also found to be extremely energetically unfavorable (Appendix D2).

3.4.4.6 Loss of guanidine from [RGD + H]$^+$

This relatively uncommon ion results from a S$_N$2 reaction with a TS at 44.9 kcal mol^{-1} (Figure SI Figure 3.3). This TS value should be viewed as a lower bound though because density functional methods perform less well for activation barrier height calculations for S$_N$2 reactions [59] than for other TS and reaction-path calculations and generally underestimate their barrier heights [60].

3.4.5 Effect of internal energy distribution and fragmentation timescale on relative fragment ion abundances

Although not the lowest energy reaction, the NH$_3$-loss pathway is entropically more favorable (positive entropy of activation, Table 3.1) than either the water-loss or b$_2$-y$_1$ pathways. Consequently, increasing excitation should make this pathway more favorable as was observed in the appearance energy experiment. The b$_1$ ion, loss of guanidine ion and b$_2$+H$_2$O ion formation pathways involve higher energy transition

states, but are also entropically favorable (Table 3.1). Increasing excitation should therefore increase the relative favorability of these pathways too. This is also supported by the results of the appearance energy and tandem MS experiments. This also indicates that despite some complicated multi-step fragmentation pathways, neither instruments' timescale (IT ~10ms, MALDI experiments ~8µs) is significantly rate limiting.

3.5 Conclusions

(1) A new energetically feasible b_1-y_2 mechanism has been discovered for b_1 ion formation for protonated peptides with an N-terminal arginine residue. Based on this mechanism, b_1 ions are protonated lactam structures.

(2) The role of salt bridge structures in protonated peptide fragmentation has been previously underestimated. The new salt-bridge stabilized b_2-y_1 TSs for ions containing an N-terminal arginine residue are much lower in energy than the normal CS TSs. These salt-bridge stabilized species offer an explanation for other b_n ions observed from protonated peptides with N-terminal arginine residues (e.g. bradykinin and its analogues). The controversy surrounding the mechanism of formation of b_{n-1}+H_2O ions has been resolved. The multi-step mechanism involves salt bridge structures and anhydride formation, but no cyclization.

(3) In addition, the b_2 ion and the b_2+H_2O ion forming mechanisms cannot be explained by the mobile proton model as post cleavage events are fundamental to both mechanisms giving the detected product ions. These findings offer strong support to the PIC model.

3.6 Acknowledgements

Table 3.1: Relative (kcal mol^{-1}) and total (Hartree) energies of various protonated forms of [RGD + H]$^+$ and the TSs of the various PFPs. Zero-point energy corrected total energies, $E_{tot + ZPE}$ are reported at the B3LYP/6-31G+(d,p) theoretical level. The relative energies, E_{rel} (corrected for zero-point energy calculated from B3LYP/6-31G(d) vibrational frequencies) were determined with respect to the global minimum on the PES of [RGD + H]$^+$. Relative enthalpies (ΔH_{298}) and Gibbs free energies (ΔG_{298}) at 298K are in kcal/mol, relative entropies (ΔS_{298}) are in cal/mol K.

Species Type	$E_{total + ZPE}$/ Hartrees	$\Delta E_{el + ZPE}$/ kcal/mol	ΔH_{298}/ kcal/mol	ΔG_{298}/ kcal/mol	ΔS_{298}/ cal/mol
CS	-1250.592653	0.0	0.0	0.0	0.0
SB_NG	-1250.572800	12.5	13.0	9.6	11.4
SB_D	-1250.572540	12.6	12.9	11.7	3.7
SB_Cterm	-1250.570174	14.1	14.1	13.9	0.6
SB_N2	-1250.564529	17.7	18.0	16.7	4.2
CS_NH$_3$_loss	-1250.525000	42.5	43.4	41.0	8.0
H$_2$O_ loss / b$_3$	-1250.527732	40.7	40.7	40.1	2.1
H$_2$O_loss_Ox	-1250.512748	50.1	50.4	49.3	3.5
SB_Cterm_b$_2$-y$_1$	-1250.528495	40.3	40.4	39.3	3.9
SB_D_b$_2$-y$_1$	-1250.526632	41.4	41.4	41.8	-1.5
SB_NG_ b$_2$-y$_1$_1	-1250.510875	51.3	51.9	48.5	11.5
SB_NG_ b$_2$-y$_1$_2	-1250.514040	49.3	49.5	48.2	4.5
SB_NG_ b$_2$-y$_1$_3	-1250.512781	50.1	50.5	47.3	10.8
CS_b$_1$-y$_2$	-1250.524725	42.6	42.8	40.7	7.1
SB_b$_2$+H$_2$O_TS1	-1250.546724	28.8	28.8	28.6	0.5
b$_2$+H$_2$O_Anhydride	-1250.563577	18.2	18.9	16.7	7.3
SB_b$_2$+H$_2$O_TS2	-1250.523030	43.7	44.9	42.0	9.7
59_loss_S$_N$2	-1250.521069	44.9	45.3	42.8	8.3

Table 3.2: Relative abundance in % of base (^{12}C) peak area, using peak areas of [RGD + H]$^+$ fragment ions. Values in 'curly' brackets, {}, correspond to relative abundance of CAD MALDI-TOF/TOF normalized to the [M+H-NH$_3$]$^+$ peak.

	MI MALDI-TOF/TOF	CAD MALDI-TOF/TOF	ESI IT MS^2
[M+H-NH$_3$]$^+$	100	21 {100}	100
b$_2$	45	15 {74}	58
b$_2$ + 18	41	16 {76}	28
b$_1$	34	14 {65}	20
-59/guanidine	24	11 {51}	6
d$_3$	-	9 {44}	3
[M+H-H$_2$O]$^+$/ b$_3$	-	-	25
b$_2$ - NH$_3$	-	-	17
Immonium ions	-	70(100%), 87(37%), 112(17%)	-

Figure 3.1: b$_2$-y$_1$ TSs of protonated RGD

SB_Cterm_b$_2$-y$_1$
40.3 kcal mol^{-1}

SB_D_b$_2$-y$_1$
41.2 kcal mol^{-1}

SB_NG_ b$_2$-y$_1$_1
51.3 kcal mol^{-1}

SB_NG_ b$_2$-y$_1$_2
49.3 kcal mol^{-1}

SB_NG_ b$_2$-y$_1$_3
50.1 kcal mol^{-1}
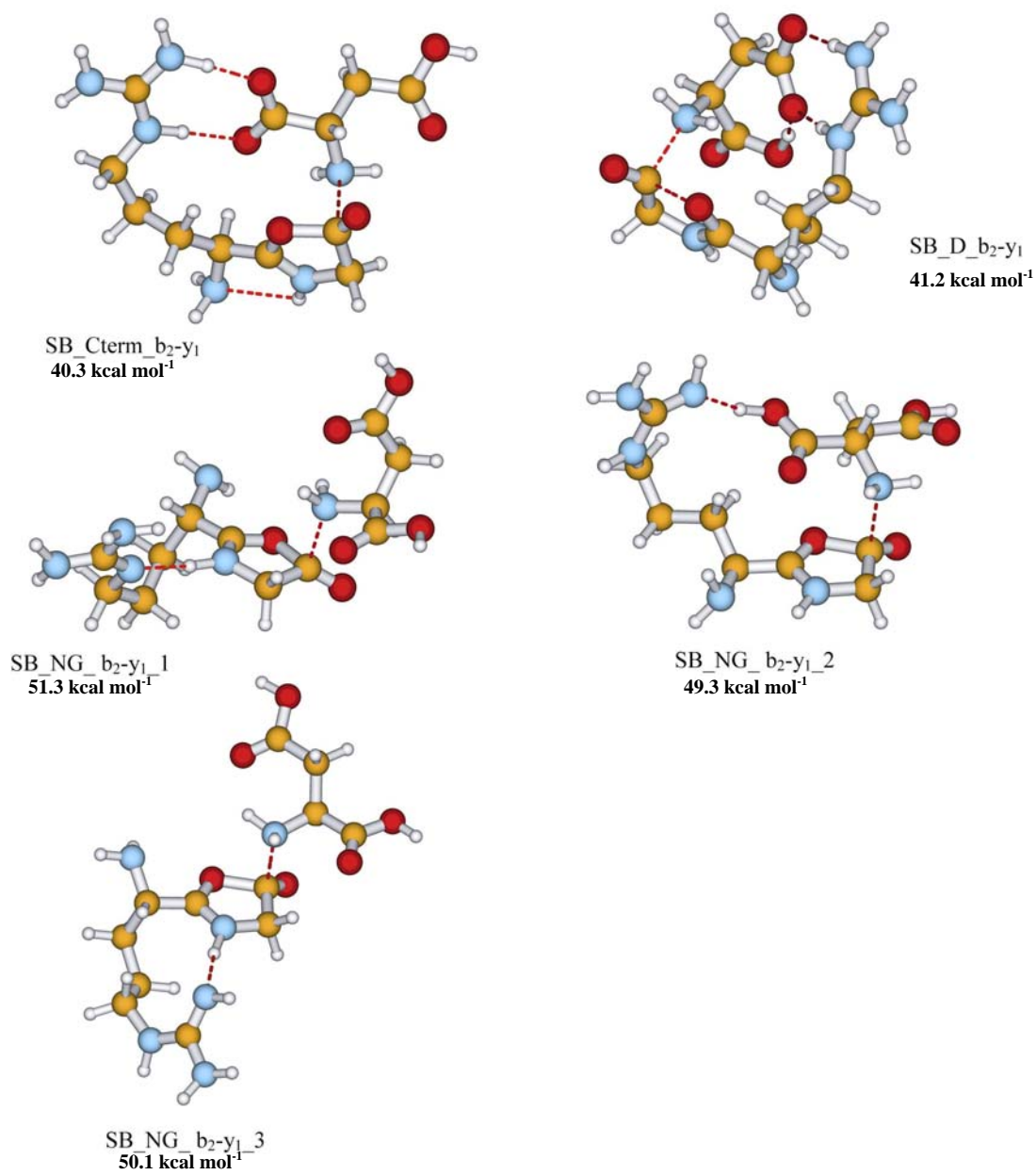
Figure 3.2 Structures on the water loss PFP

RC_H$_2$O_loss

16.7 kcal mol^{-1}

H$_2$O_loss TS

40.7 kcal mol^{-1}

H$_2$O_loss PBD

16.7 kcal mol^{-1}
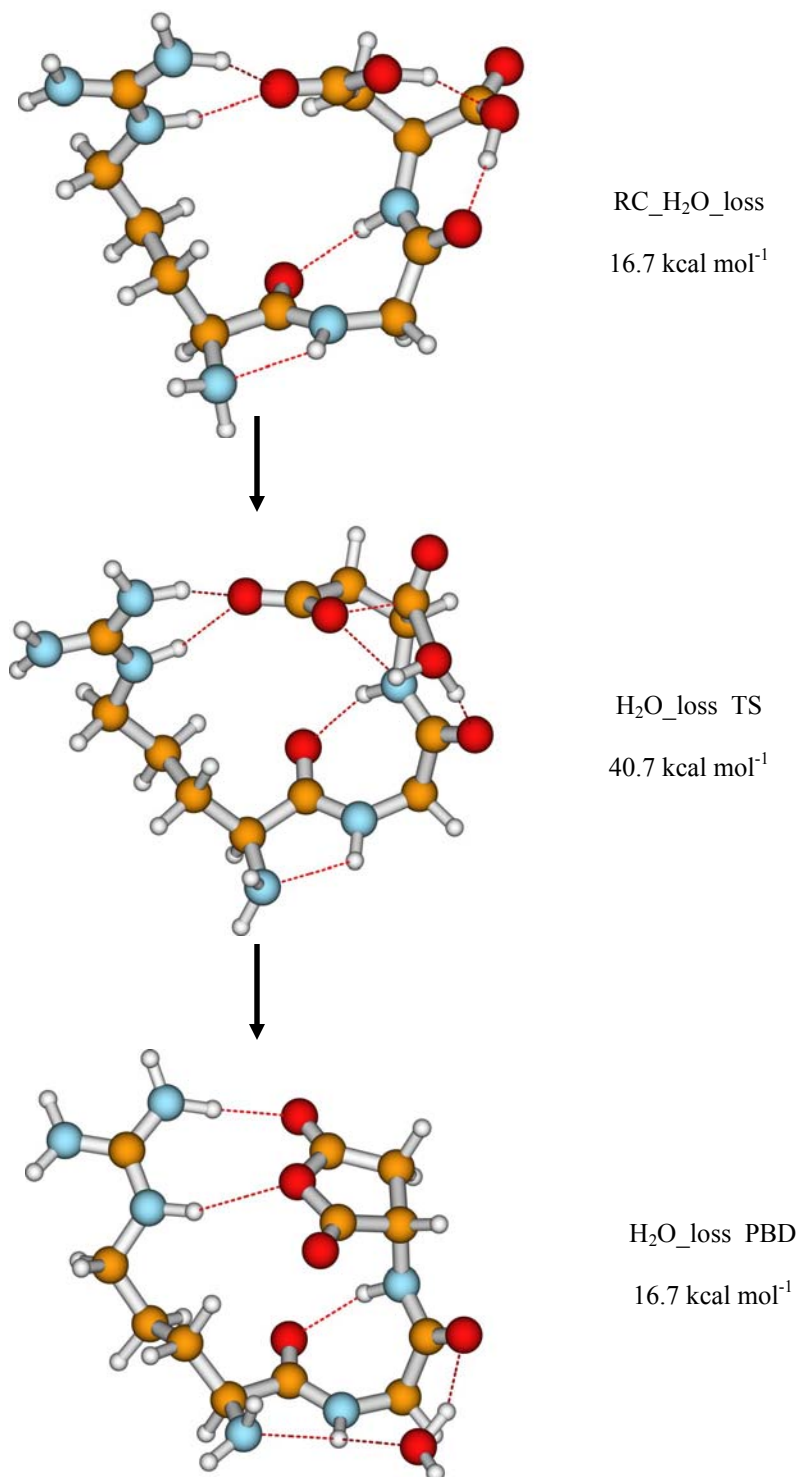
Figure 3.3 The NH$_3$_loss Transition Structure
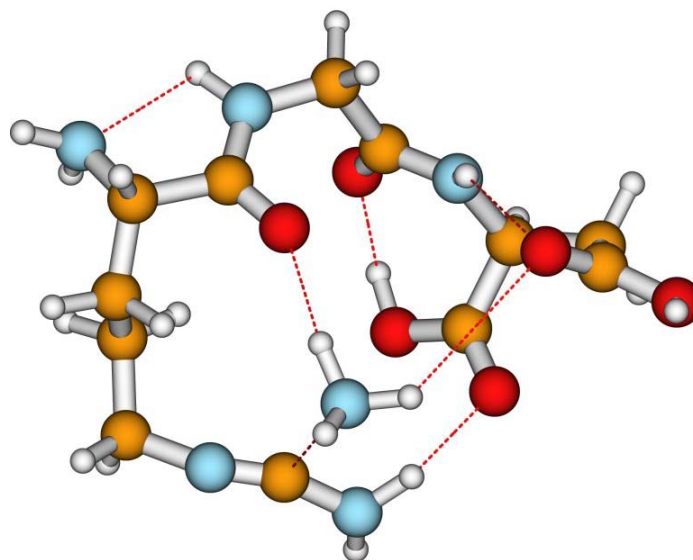
Figure 3.4 The CS_b_1-y_2 Transition Structure
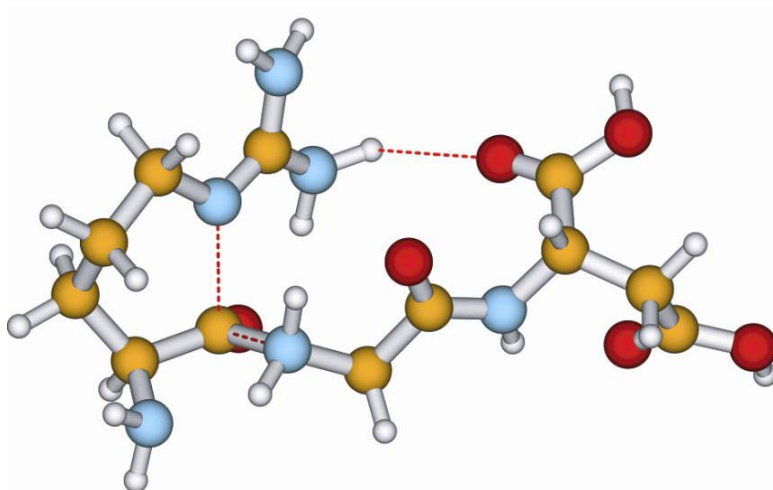
Figure 3.5a: Anhydride formation on the b_2+H_2O pathway

RC_Anhydride

28.3 kcal mol^{-1}

SB_b_2+H_2O_TS1

28.8 kcal mol^{-1}

Anhydride

23.1 kcal mol^{-1}
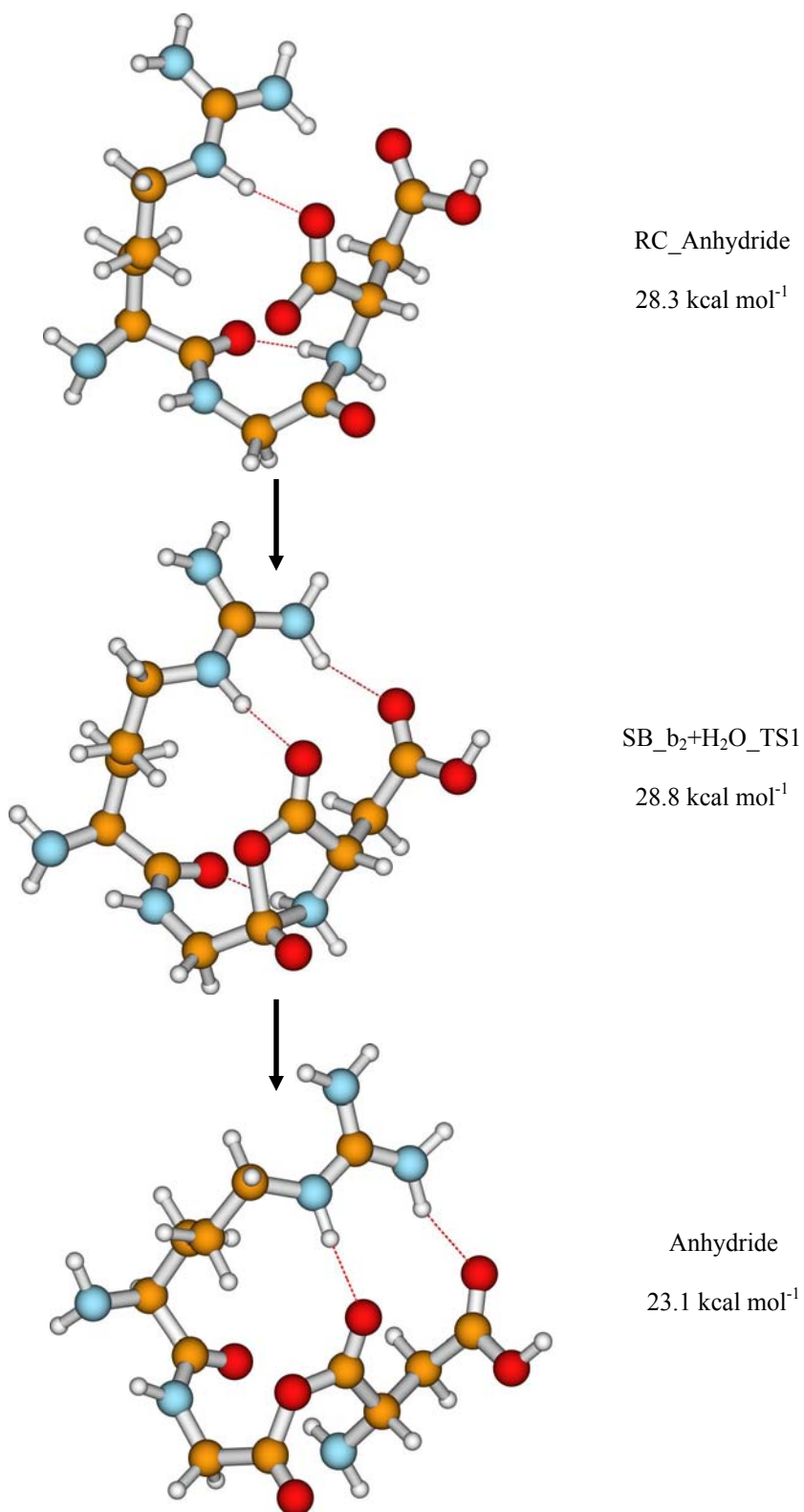
Figure 3.5b: Formation on the b_2+H_2O ion from an anhydride structure

RC_SB_b_2+H_2O_TS2

19.3 kcal mol^{-1}

SB_b_2+H_2O_TS2

43.7 kcal mol^{-1}

b_2+H_2O_PBD + CO

21.3 kcal mol^{-1}
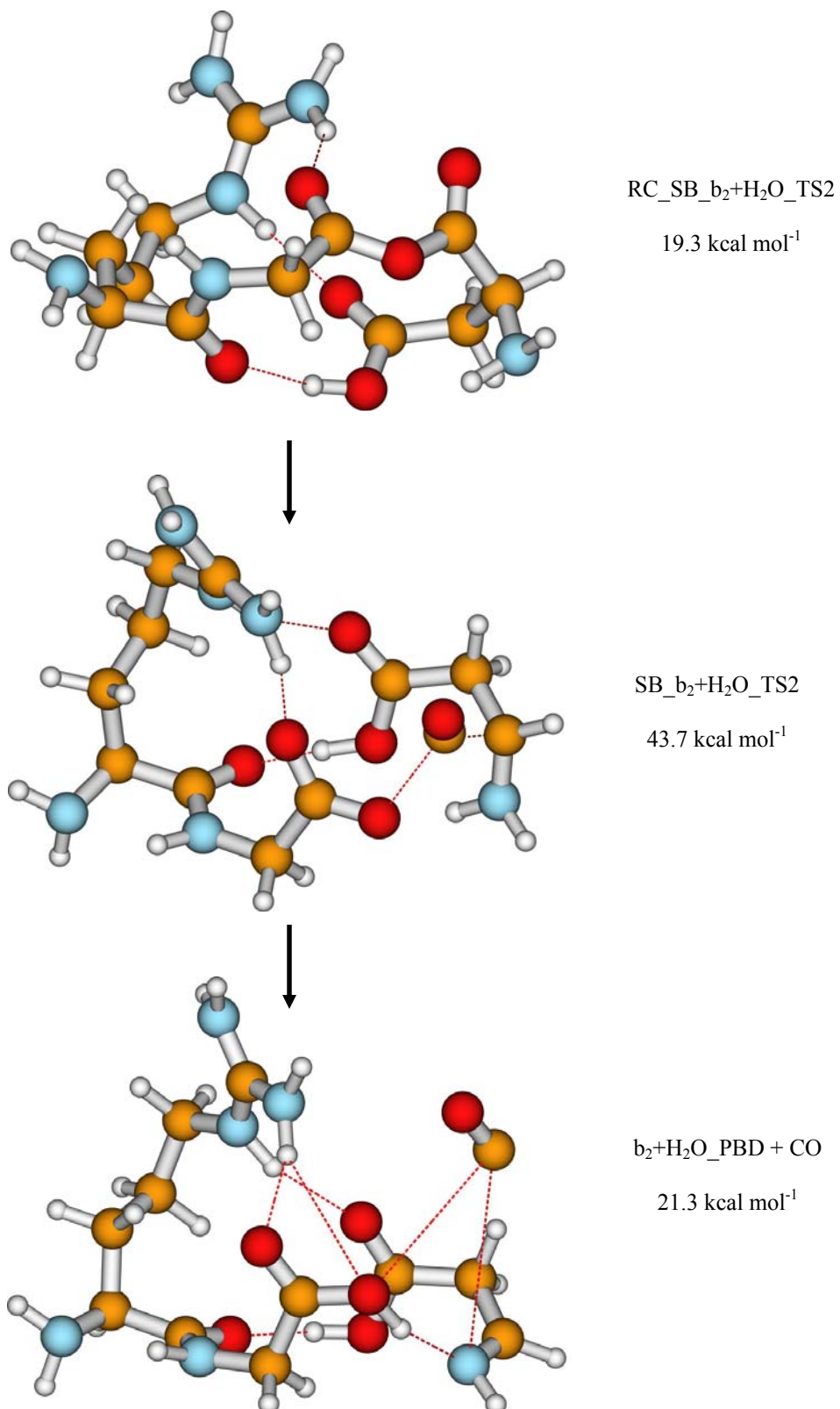
Scheme 3.1a. b_2-y_1 amide bond cleavage for species with neutral R side chain.

12.5 kcal mol^{-1} NG

PT

49.3 kcal mol^{-1} SB_NG_ b_2-y_1_2   TS

+

PT
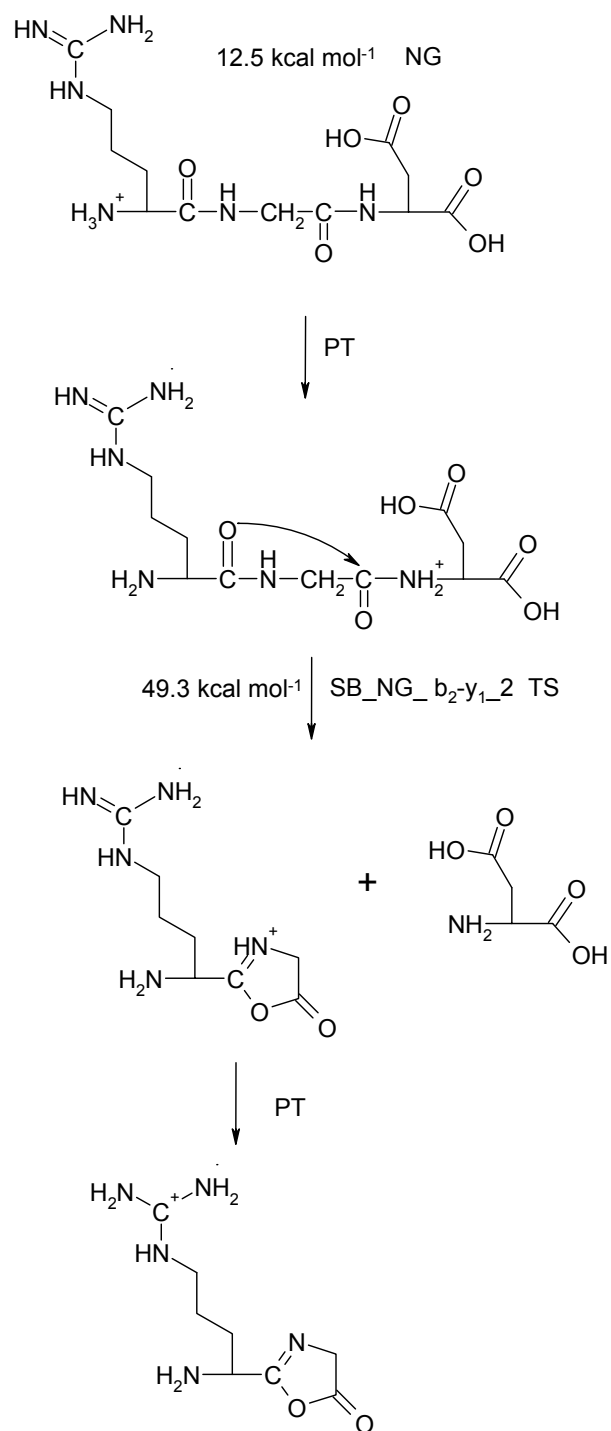
Scheme 3.1b. b_2-y_1 amide bond cleavage for SB species with deprotonated C terminus.
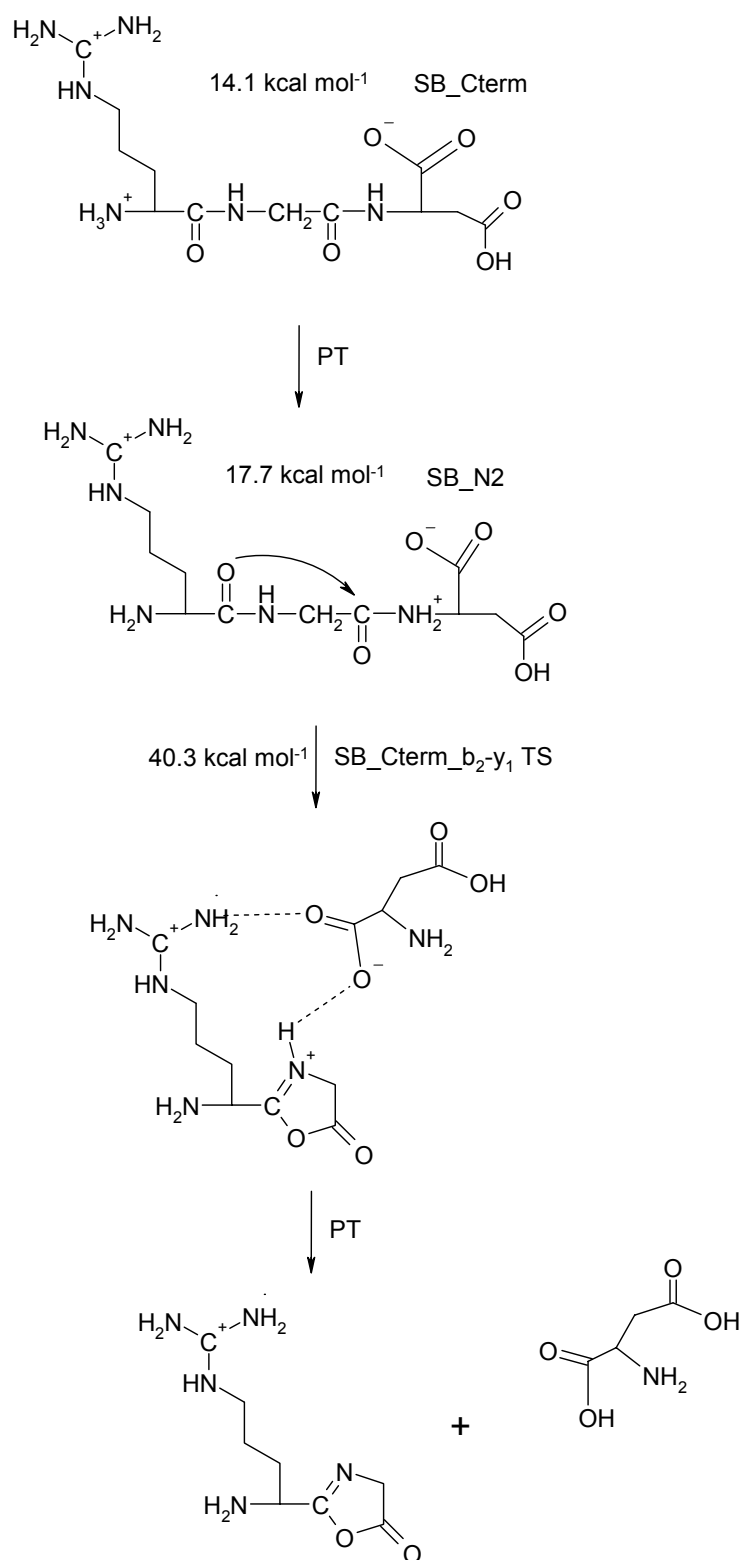
Scheme 3.1c. b_2-y_1 amide bond cleavage for SB species with deprotonated D side chain.
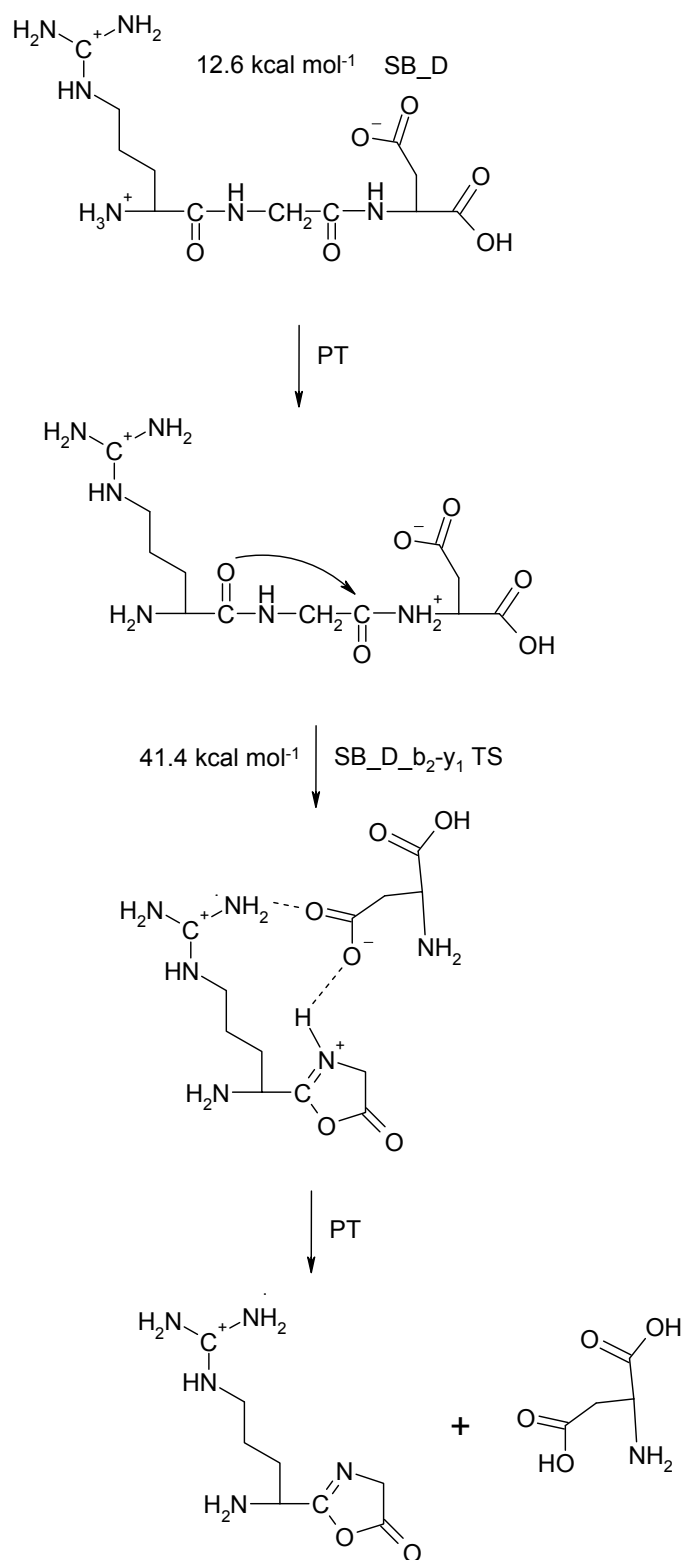
Scheme 3.2 The b_1-y_2 lactam-forming PFP

$42.6 \text{kcal mol}^{-1}$ b_1-y_2 TS

Scheme 3.3 The b_2+H_2O PFP

CS

SB_N2

28.8 kcal mol^{-1}

SB_b_2+H_2O_TS1

b_2+H_2O_Anhydride

43.7 kcal mol^{-1}

SB_b_2+H_2O_TS2

PT

CO_2H

CO

b_2+H_2O ion
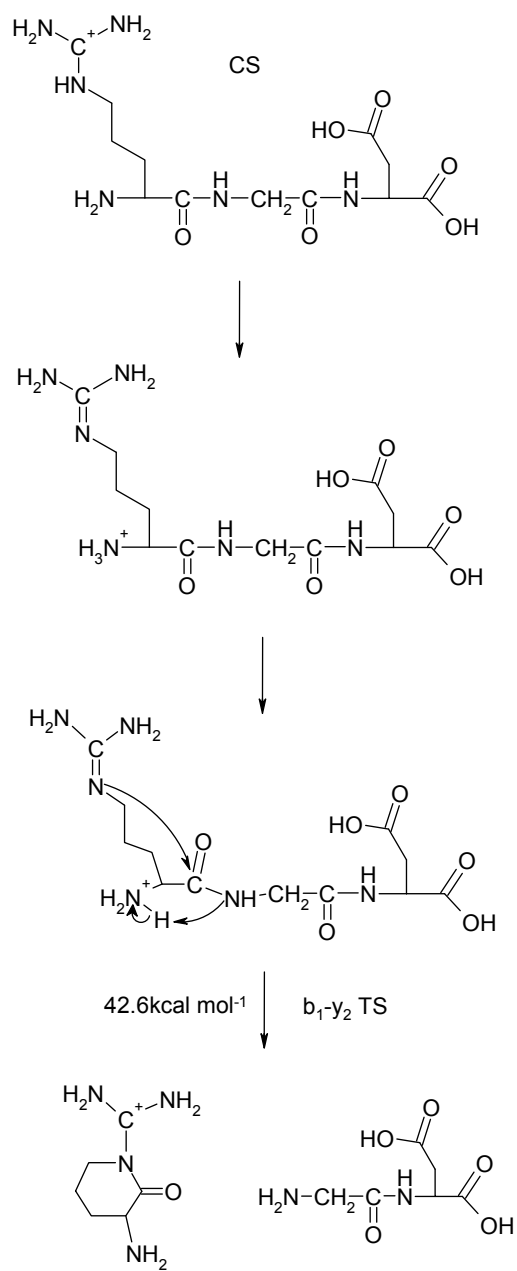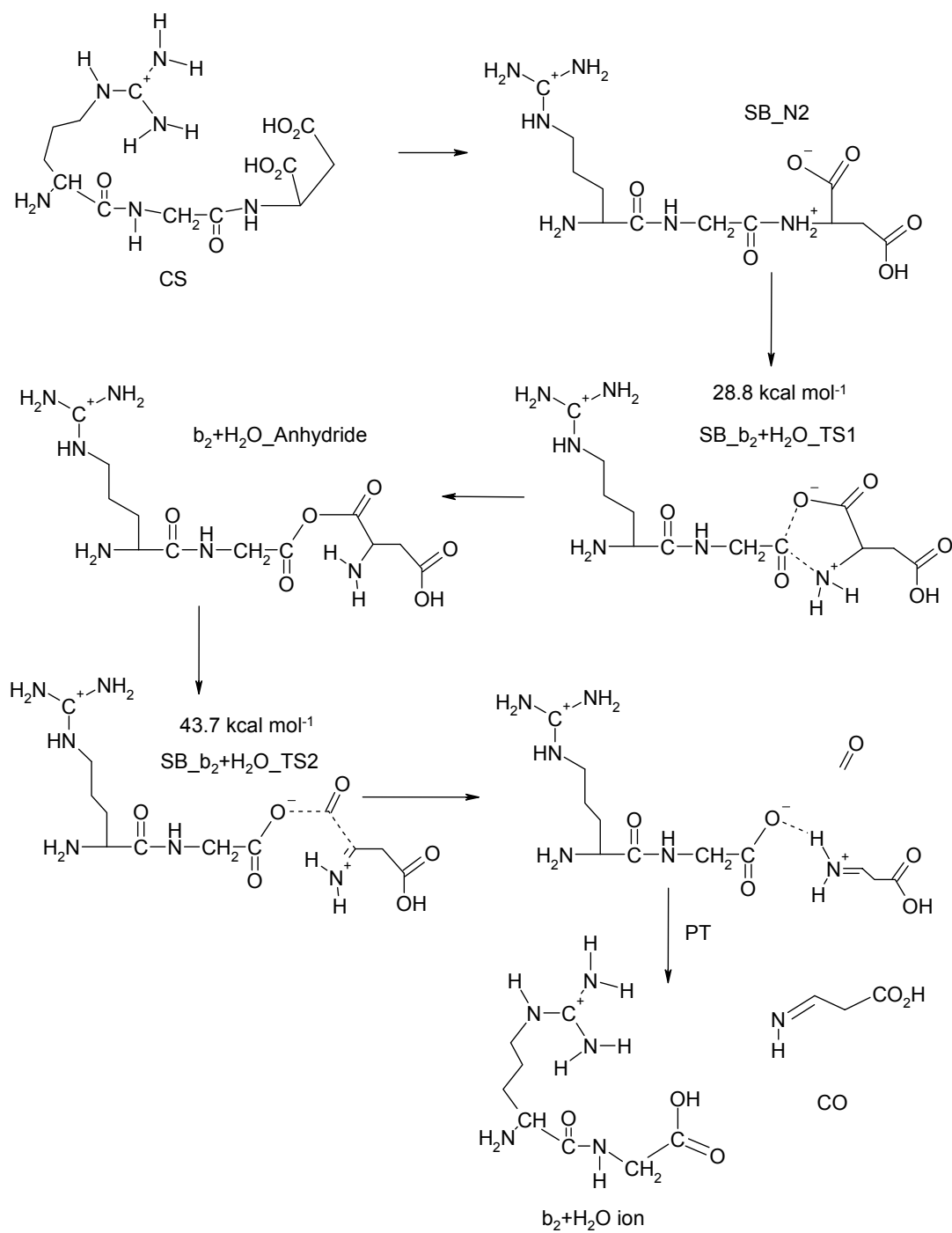
3.7 References

1) Steen, H.; Mann, M. *Nature Rev. Mol. Cell Biol.* **2004**, 5**,** 699–711.

2) Hunt, D. F.; Yates, J. R., III; Shabanowitz, J.; Winston, S.; Hauer, C. R. *Proc. Natl. Acad. Sci. U.S.A.* **1986***, 83*, 6233-6237.

3) Biemann, K.; Scoble, H. A. *Science* **1987**, *237*, 992-998.

4) Biemann, K. *Biomed. Environ. Mass Spectrom.* 1988, 16, 99.

5) Roepstorff, P.; Fohlmann, J. *J. Biomed. Mass Spectrom.* 1984, 11, 601.

6) Anderson, N.L.; Anderson, N.G *Mol. Cell. Prot.*, 2002, 1, 845-867.

7) Khatun, J.; Ramkissoon, K.; Giddings, M.G.; *Anal. Chem.* **2007**, 79, 3032-3040.

8) Eng, J. K.; McCormack, A. L.; Yates, J. R., III. *J. Am. Soc. Mass Spectrom.* 1984, 5, 976.

9) Perkins, D. N.; Pappin, D. J. C.; Creasy, D. M.; Cottrell, J. S. *Electrophoresis* 1999, 20, 3551.

10) Paizs, B.; Suhai, S. *Mass Spectrom. Rev.*, **2005**, 24, 508-548.,

11) Paizs, B.; Suhai, S. Rapid Commun. Mass Spectrom. 2002, 16, 1699-1702.

12) Wysocki, V. H.; Tsaprailis, G.; Smith, L. L.; Breci, L. A. *J. Mass Spectrom.,* **2000**, 35, 1399-1406.

13) Dongre, A. R.; Jones, J. L.; Somogyi, A.; Wysocki, V. H. *J. Am. Chem. Soc.* **1996**, *118*, 8365-8374.

14) Somogyi, Á.; Wysocki, V.H.; Mayer, I *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 704-717.

15) Biemann, K.; Martin, S. A. *Mass Spectrom. Rev.* **1987**, *6*, 1.

16)   Poulter, L.; Taylor, L. C. E. *Int. J. Mass Spectrom. Ion Processes* **1989**, *91*, 183.

17)   Burlet, O.; Yang, C. Y.; Gaskell, S. J. *J. Am. Soc. Mass Spectrom.* **1992**, *3*, 337.

18)   Tang, X.; Boyd, R. K. *Rapid Commun. Mass Spectrom.* **1992**, *6*, 651.

19)   Burlet, O.; Orkiszewski, R. S.; Ballard, K. D.; Gaskell, S. J. *Rapid Commun. Mass Spectrom.* **1992**, *6*, 658.

20)   McCormack, A. L.; Somogyi, A.; Dongre, A. R.; Wysocki, V. H. *Anal. Chem.* **1993**, *65*, 2859.

21)   Tang, X.-J.; Thibault, P.; Boyd, R. K. *Anal. Chem.* **1993**, *65*, 2824.

22)   Johnson, R. S.; Krylov, D.; Walsh, K. A. *J. Mass Spectrom.* **1995**, *30*, 386.

23)   Nold, M. J.; Wesdemiotis, C.; Yalcin, T.; Harrison, A. G. *Int. J. Mass Spectrom. Ion Processes* **1997**, *164*, 137.

24)   Harrison, A. G.; Yalcin, T. *Int. J. Mass Spectrom.* **1997**, *165*, 339.

25)   Summerfield, S. G.; Whiting, A.; Gaskell, S. J. *Int. J. Mass Spectrom. Ion Processes* **1997**, *162*, 149.

26)   Harrison, A. G.; Tu, Y.-P. *J. Mass Spectrom.* **1998**, *33*, 532.

27)   Jorgensen, T. J. D.; Gardsvoll, H.; Ploug, M.; Roepstorff, P. *J. Am. Chem. Soc.* **2005**, *127*, 2785.

28)   Csonka, I. P.; Paizs, B.; Lendvay, G.; Suhai, S. *Rapid Commun. Mass Spectrom.* **2000**, *14*, 417-431.

29)   Paizs, B.; Csonka, I. P.; Lendvay, G.; Suhai, S. *Rapid Commun. Mass Spectrom.* **2001**, *15*, 637-650.

30) Csonka, I. P.; Paizs, B.; Lendvay, G.; Suhai, S. *Rapid Commun. Mass Spectrom.* **2001**, *15*, 1457.

31) Tsaprailis, G.; Nair, H.; Somogyi, A.; Wysocki, V. H.; Zhong, W.; Futrell, J. H.; Summerfield, S. G.; Gaskell, S. J. *J. Am. Chem. Soc.* **1999**, *121*, 5142.

32) Cooper, T.; Talaty, E.; Grove, J.; Suhai, S.; Paizs, B.; Van Stipdonk, M. *J. Am. Soc. Mass Spectrom.* **2006**, **17**, 1654-1664.

33) Bythell, B.J.; Barofsky, D. F.; Pingitore, F.; Polce, M. J.; Wang, P.; Wesdemiotis, C.; Paizs, B. *J. Am. Soc. Mass Spectrom.*, **2007**, 7, 1291-1303.

34) Harrison, A. G.; Young, A. B.; Bleiholder, C.; Suhai, S.; Paizs, B. *J. Am. Chem. Soc.,* **2006**, 128 , 10364 -10365.

35) Harrison, A. G.; Young, A. B. *J. Am. Soc. Mass Spectrom.* **2004**, *15*, 1810.

36) She, Y-M.; Krokhin, O.; Spicer, V.; Loboda, A.; Garland, G.; Ens, W.; Standing**,** K. G.; Westmore, J. B. *J. Am. Soc. Mass. Spectrom.*, **2007**, 18, 1024-1037.

37) Thorne, G. C.; Ballard, K.D.; Gaskell, S. J. *J. Am. Soc. Mass. Spectrom.*, **1990**, 1, 249-257.

38) Ballard, K.D.; Gaskell, S. J. *J. Am. Chem. Soc.*, **1992**, 114, 64-71.

39) March, R. E. *J. Mass Spectrom.* **1997**, 32, 351-369.

40) March, R. E. *Rapid Commun. Mass Spectrom.* **1998**, 12, 1543-1554.

41) Medzihradszky, K. F.; Campbell, J. M.; Baldwin, M. A.; Falick, Arnold M.; Juhasz, P.; Vestal, M. L.; Burlingame, A. L. *Anal. Chem.*, **2000**, 72(3), 552-558.

42) Juhasz, P.; Campbell, J. M.; Vestal, M. L. MALDI-TOF/TOF technology for peptide sequencing and protein identification in *Mass Spectrometry and Hyphenated Techniques in Neuropeptide Research*, Edited by Silberring, J.; Ekman, R., **2002**, 375-413, Wiley-Interscience, 1st Edition.,

43) Bienvenut, W. V.; Deon, C.; Pasquarello, C.; Campbell, J.M.; Sanchez, J.; Vestal, M.L.; Hochstrasser, D.F. *Proteomics*, **2002**, 2(7), 868-876.

44) McLuckey, S. A.; Van Berkel, G. J.; Goeringer, D. E.; Glish, G. L. *Anal. Chem*. **1994**, 66, 689A-696A.

45) Polfer, N.; Oomens, J.; Suhai, S.; Paizs, B. *J. Am. Chem. Soc.,* 129, 5887 - 5897, 2007.

46) Gaussian 03, Revision C.02, Gaussian, Inc., Wallingford CT, **2004**.

47) Baer, T.; Hase, W. L. *Unimolecular Reaction Dynamics*, **1996**, Oxford University Press, Oxford.

48) Beyer, T.; Swinehart, D.R.; *ACM Commun.* **1973**, 16, 379.

49) Lifshitz, C. *Eur. J. Mass Spectrom.*, **2002**, 8, 85-98.

50) Yalcin, T.; Khouw, C.; Csizmadia, I. G.; Peterson, M. R.; Harrison, A. G.*,* **1995**, 6, 1165-1174.

51) Yalcin, T.; Csizmadia, I. G.; Peterson, M. R.; Harrison, A. G., *J. Am. Soc. Mass.Spectrom.*, **1996**, 7, 233-242.

52) Paizs, B.; Suhai, S. *Rapid Commun. Mass Spectrom.*, **2001**, 15, 2307-2323.

53) Polfer, N. C.; Oomens, J.; Suhai, S.; Paizs, B. *J. Am. Chem. Soc.* **2005**, *127*, 17154-17155.

54)   Balta, B.; Aviyente, V.; Lifshitz, C. *J. Am. Soc. Mass Spectrom*., **2003**, 14,
        1192-1203.

55)   Paizs, B.; Suhai. S.; Hargittai, B.; Hruby, V.J.; Somogyi, A. Int. J. Mass
        Spectrom. 219, 203–232, 2002.

56)   Fang, S.; Takao, T.; Satomi, Y.; Mo, W.; Shimonishi, Y. *J. Am. Soc. Mass.*
        Spectrom., 2000, 11, 345-351.

57)   Farrugia. J.M.; O'Hair, R.A.J. *Int. J. Mass Spectrom.*, 2003, 222, 229–242.

58)   Hiserodt, R. D.; Brown, S.M.; Swijter, D.F.H.; Hawkins, N.; Mussinan, C. J.
        J. Am. Soc. Mass. Spectrom., 2007, 18, 1414-1422.

59)   Laerdahl, J. K.; Uggerud, E. *Int. J. Mass Spectrom*, **2002**, 214,277-314.

60)   Gritsenko, O. V.; Ensing, B.; Schipper, P. R. T.; Baerends, E. J.; *J. Phys.*
        Chem. A. **2000**, *104,* 8558-8565.

3.8 Supplementary Information For Chapter 3: *J. Am. Chem. Soc.* **2008.**

Novel Salt-Bridge Stabilized Peptide Fragmentation Mechanisms in Protonated RGD

Benjamin J. Bythell and Douglas F. Barofsky

Department of Chemistry, Oregon State University, Corvallis, Oregon, USA

Béla Paizs

German Cancer Research Center, Department of Molecular Biophysics, Heidelberg, Germany

SI3.1 PES Structures of protonated RGD

CS   0.0kcal mol^{-1}

SB_D   12.6kcal mol^{-1}

SB_Cterm   14.1kcal mol^{-1}

SB_N2   17.7 kcal mol^{-1}

NG   12.5kcal mol^{-1}
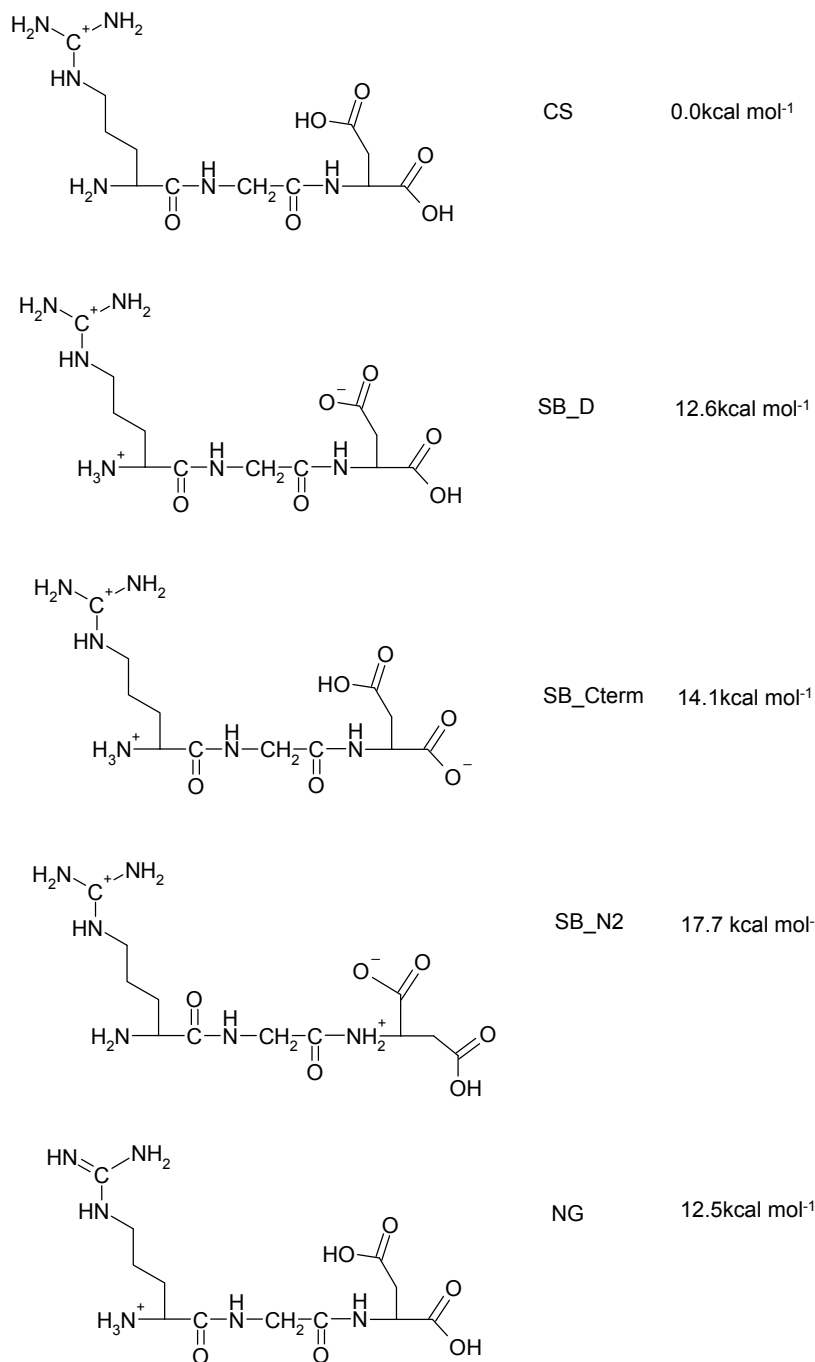
SI Figure 3.1 Structures on the alternate water loss PFP

RC_H$_2$O_loss_2

6.5 kcal mol^{-1}

H$_2$O_loss_2 TS

50.1 kcal mol^{-1}

H$_2$O_loss_2_PBD

47.7 kcal mol^{-1}
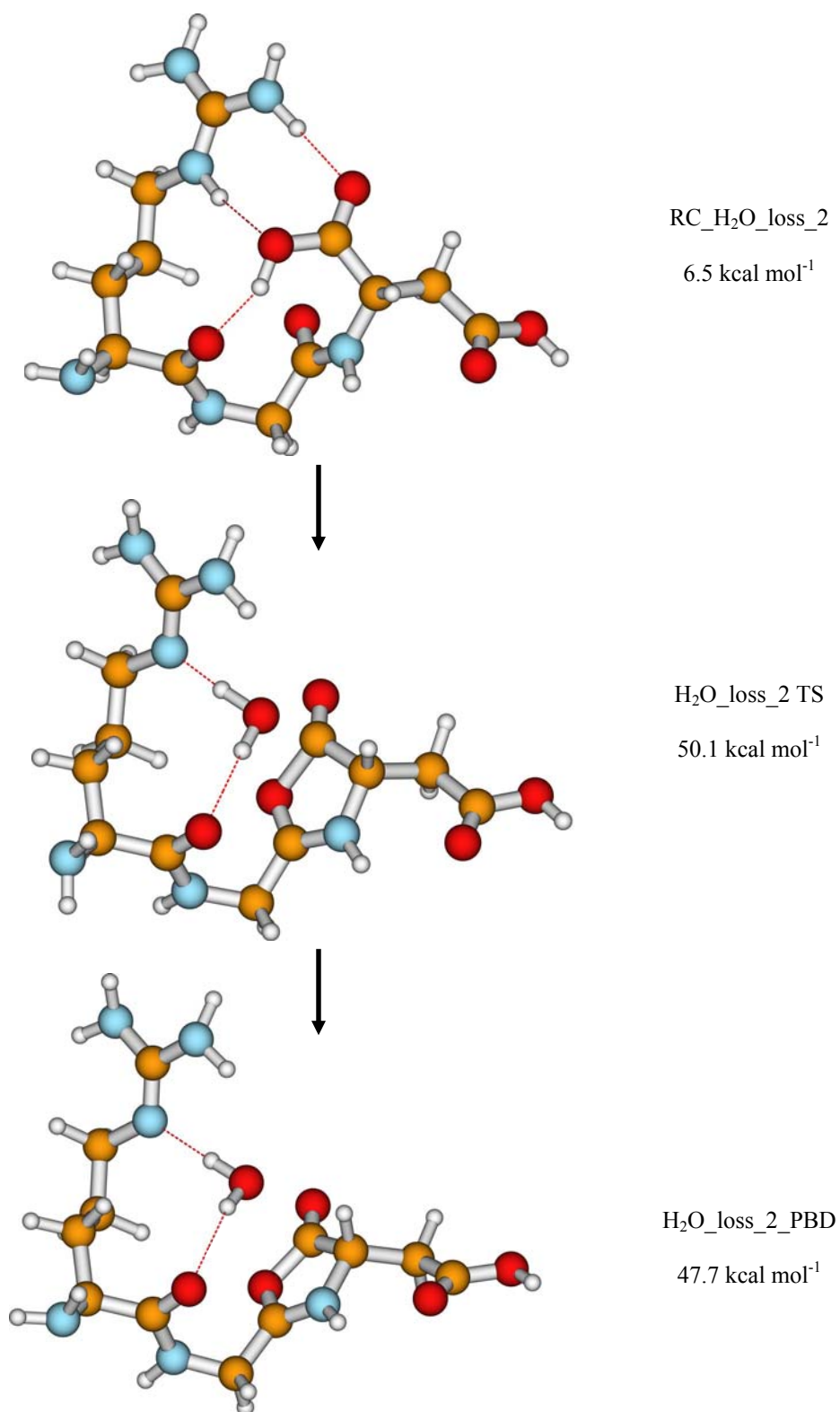
SI Figure 3.2 The most stable structures of the protonated RGD structure families with the preformed NH_3 moiety with relative energies.

AMM_CS_A
21.9 kcal mol^{-1}

AMM_CS_B
26.5 kcal mol^{-1}

AMM_SB_D_A
35.4 kcal mol^{-1}

AMM_SB_D_B
28.4 kcal mol^{-1}

AMM_SB_Cterm_A
39.1 kcal mol^{-1}

AMM_SB_Cterm_B

36.7 kcal mol^{-1}
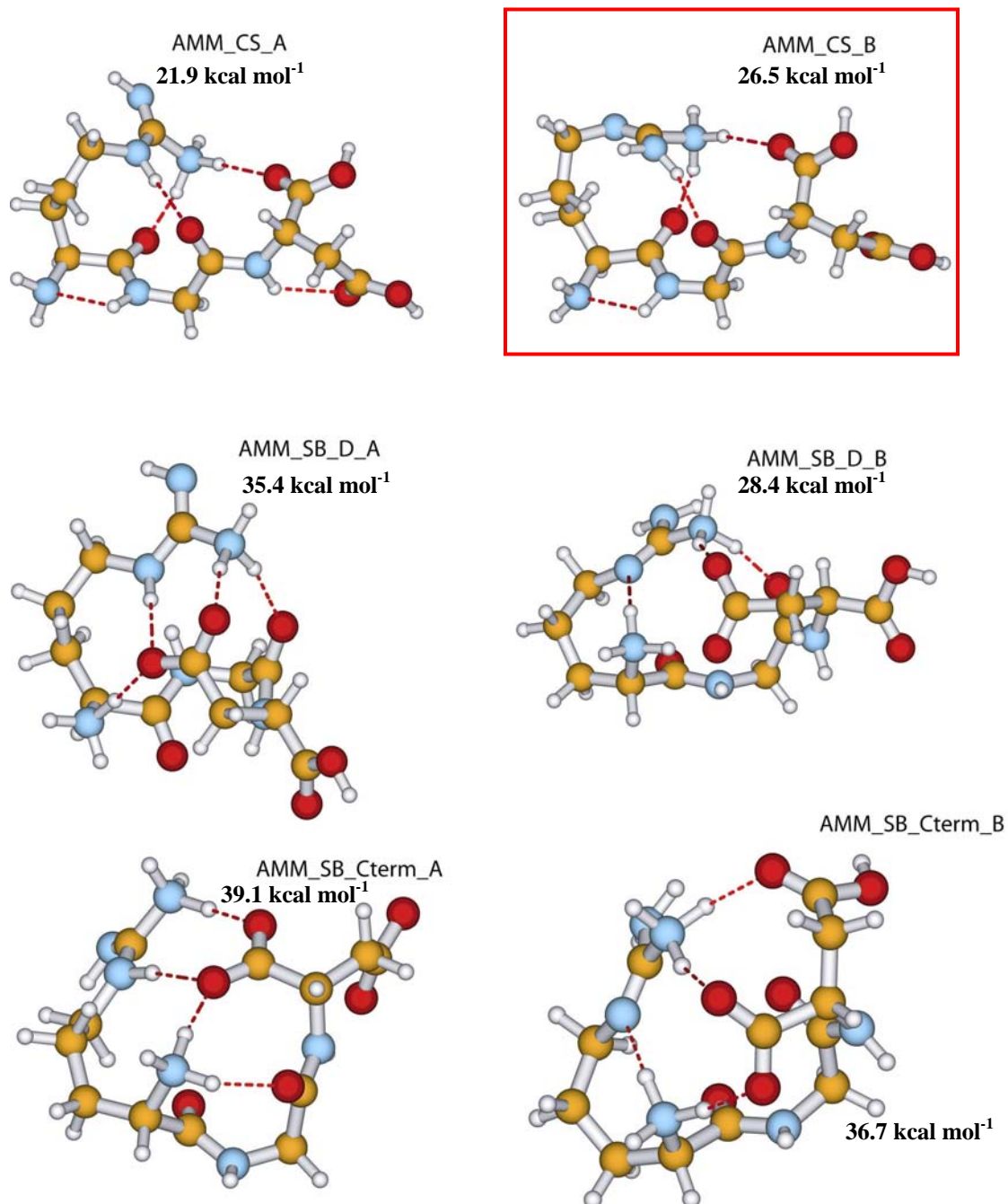
SI Figure 3.3 Structures on the S$_N$2 guanidine loss PFP

RC_59_loss_S$_N$2

7.0 kcal mol^{-1}

59_loss_S$_N$2 TS

44.9 kcal mol^{-1}

59_loss_S$_N$2_PBD

21.9 kcal mol^{-1}
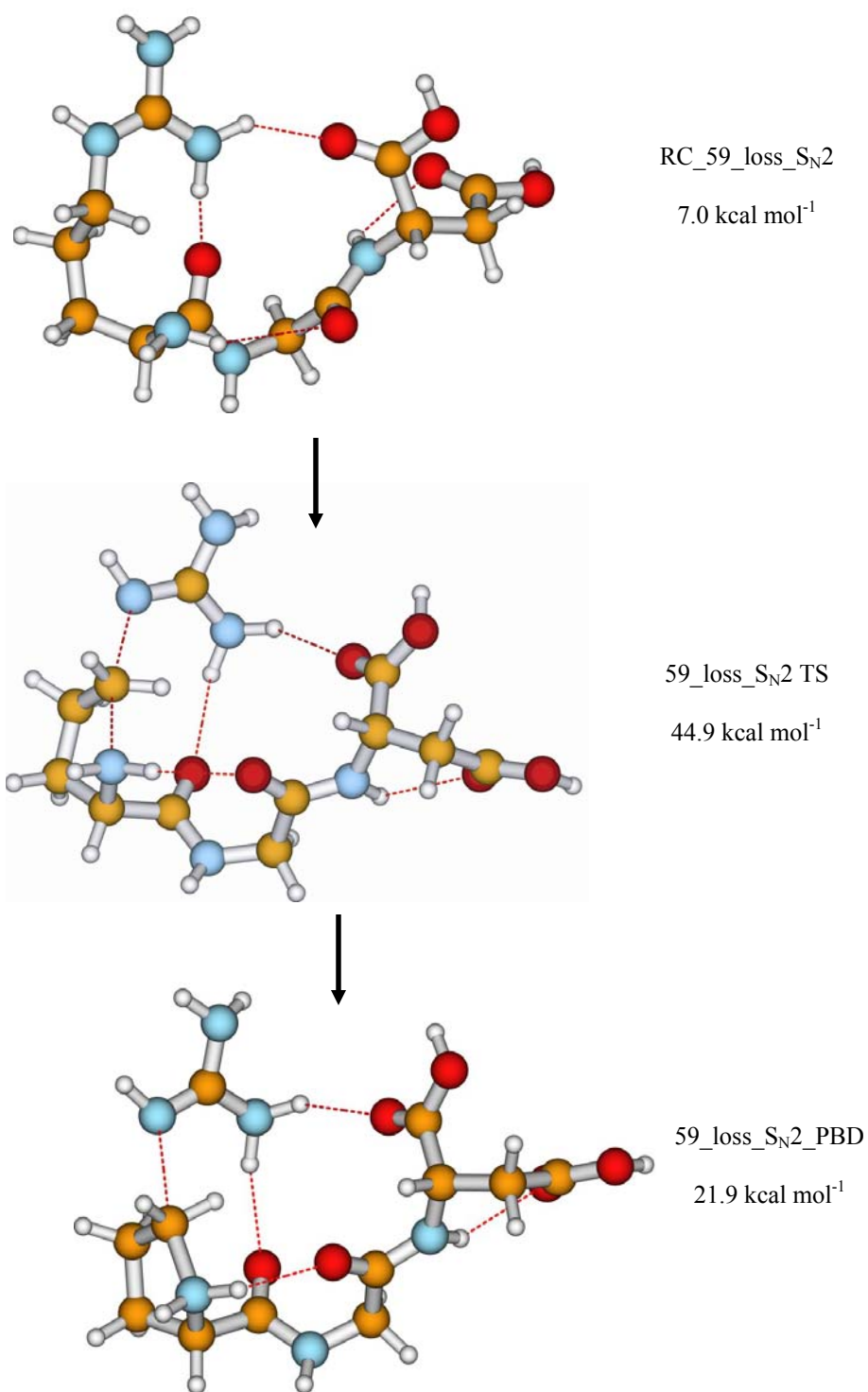
Full citation for Reference 46:

Gaussian 03, Revision C.02, Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, Jr., J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; and Pople, J. A.; Gaussian, Inc., Wallingford CT, **2004**.

Chapter 4

4. Conclusion

Researchers today rely heavily on automated protonated peptide sequencing algorithms. These algorithms generally incorporate fragmentation models which are overly simplistic and ignore all fragment intensity information. In so doing, much of the gas-phase chemistry is ignored. This approach can be successful if protonated peptides fragment in a regular, simplistic manner producing sequence (b, a and y) ions almost exclusively. It is increasingly clear however, that things are not always so simple. Other fragment ion types are formed reasonably regularly [Glish & vachet, Chapter 2, stipdonk, Harrison, polfer] and that these can have significant intensities. These problems are not widely described by the This added complexity need not be all bad news. Using a more complicated model which incorporated

Protonated peptide fragmentation is more complicated than is generally appreciated. Reactions like those in Chapter 2, where fragments held together in proton-bound dimers and undergo transitions, including association reactions and subsequent dissociation of the reorganized parent ion.

The PIC model has had recent success in explaining that fragments held together in proton-bound dimers can undergo various transitions, including association reactions and subsequent dissociation of the reorganized parent ion [32, 33].

Combined tandem MS experiments offer a powerful tool for probing the complex gas-phase chemistry of protonated peptides. Accompanying computational and labeling studies can provide important mechanistic details not readily observed from tandem mass spectra alone.

Similar reactions to the formation of $[AGG + H - CO - NH_3]^+$ peak could easily occur for other protonated peptides and consequently lead to unexpected fragment ions in MS/MS spectra. This, in turn, could lead to erroneous peptide and protein identification using current bioinformatics tools. As such, the post-cleavage phase of peptide fragmentation warrants further study so that the PIC model may be improved and implemented into software to enhance the effectiveness of peptide sequencing algorithms.

Bibliography

1) Borman, S. *Chem. Eng. News*, May 26, **1998**.

2) Nobel website http://nobelprize.org/

3) Barber, M.; Bordoli, R.A.; Sedgewick, R. D.; Tyler, A. N., *J. Chem. Soc., Chem. Commun.* **1981**, 7, 325-327.

4) Yamashita, M.; Fenn, J. B.; *J. Phys. Chem*. **1984**, 88, 4451-4459.

5) Karas, M.; Bachmann, D.; Bahr, U.; Hillenkamp, F. *Int. J. Mass Spectrom. Ion Proc.*, **1987**, 78, 53-68.

6) Karas, M.; Hillenkamp, F. *Anal. Chem.*, **1988**, 60, 2299-2301.

7) Tanaka, K. H.; Wake, H.; Ido, Y.; Akita, S.; Yoshida, Y.; Yoshida, I. *Rapid Commun. Mass Spectrom*. **1988**, 8, 2.

8) Barber, M.; Bordoli, R. S.; Elliot, G. J.; Sedgewick, R. D.; Tyler, A. N. *Anal. Chem*., **1981**, 54, 645A.

9) Spengler, B.; Bahr, U.; Karas, M.; Hillenkamp, F. *Anal. Instrum*. **1988**; 17, 173.

10) Vestal, M. L.; Juhasz, P. ; Martin, S.A. *Rapid Commun. Mass Spectrom*. 1995; 9**,** 1044- .

11) Brown, R.S.; Lennon, J.J. *Anal. Chem*. **1995**; 67: 1998-.

12) Karas, M. *J. Mass Spectrom*. 2000, **35**, 1–12.

13) Fenn, J.B.; Mann, M.; Meng, C.K.; Wong, S.F.; Whitehouse, C.M.; *Mass Spectrom. Rev*., **1990**, 9, 37-70.

14) Cooks, R.G.; Glish, G.; McLuckey, S.A.; Kaiser, R.E. *Chem. Eng. News*, March 25, **1991**.

15) Polce, M. J.; Cordero, M. M.; Wesdemiotis, C.; Bott, P. A. *Int. J. Mass Spectrom. Ion Proc.,* **1992**, 113, 35-58.

16) Stephens, W.E. *Phys. Rev.* 1946, **69**, 691.

17) Cotter, R.J., *Anal. Chem*., **1999**, 445A-451A.

18) Wiley, W.C.; McLaren, I.H. *Rev. Sci. Instrum*. **1955**, 26, 1150-1157.

19) Mamyrin, B.A.; Karataev, V.I.; Shmikk, D.V.; Zagulin, V.A. *Sov. Phys. JETP*, **1973**, 37, 45.

20) Cotter, R.J., Time-of-Flight Mass Spectrometry, **1997**, ACS, Washington.

21) Juhasz, P.; Campbell, J. M.; Vestal, M. L. MALDI-TOF/TOF technology for peptide sequencing and protein identification in *Mass Spectrometry and*

Hyphenated Techniques in Neuropeptide Research, Edited by Silberring, J.; Ekman, R., **2002**, 375-413, Wiley-Interscience, 1st Edition

22) Vestal, M. L.; Juhasz, P. *J. Am. Soc. Mass Spectrom*., **1998**, 9, 892-.

23) Cotter, R. J.; Gardner, B. D.; Iltchenko, S.; English, R. D. *Anal. Chem*. **2004,** 76, 1976-1981.

24) Jennings, K. R. *Int. J. Mass Spectrom. Ion Phys*. **1968**, 2, 227.

25) McLafferty, F. W.; Schuddemage, H. D. R. *J. Am. Chem. Soc*., **1969**, 91, 1866.

26) Martin, S. A.; Johnson, R. S.; Costello, C. E.; Biemann, K., The Structure Determination of Peptides by Tandem Mass Spectrometry. In *Analysis of Peptides and Proteins by MassSpectrometry*; McNeal, C. J., Ed.; John Wiley and Sons: Chichester, England, **1988**, 135-150.

27) Roepstorff, P.; Fohlman, J. *Biomed. Mass Spectrom.* **1984**, 11(11), 601.

28) a)Biemann, K. *Methods Enzymol*., 1990, 193, 455-479. b) 351-360.

29) Khatun, J.; Ramkissoon, K.; Giddings, M.G.; *Anal. Chem*. **2007**, 79, 3032-3040.

30) Paizs, B.; Suhai, S. *Mass Spectrom. Rev.*, **2005**, 24, 508-548.

31) Bythell, B. J.; Barofsky, D. F., Pingitore, F.; Polce, M. J., Wang, P., Wesdemiotis, C.; Paizs, B. *J. Am. Soc. Mass Spectrom.*, **2007**, 7, 1291-1303.

32) Dongre, A. R.; Jones, J. L.; Somogyi, A.; Wysocki, V. H., *J. Am. Chem. Soc.* **1996**, *118*, 8365-8374.

33) Wysocki, V. H.; Tsaprailis, G.; Smith, L. L.; Breci, L. A. *J. Mass Spectrom.,* **2000**, 35, 1399-1406.

34) Burlet, O.; Yang, C.Y.; Gaskell, S. J. *J. Am. Soc. Mass Spectrom.*, **1992**, 3, 337–344.

35) Somogyi, Á.; Wysocki, V.H.; Mayer, I. *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 704-717.

36) Tsaprailis, G.; Nair, H.; Somogyi, A.; Wysocki, V. H.; Zhong, W.; Futrell. J. H.; Summerfield, S. G.; Gaskell, S. J. *J. Am. Chem. Soc.*, **1999**, 121, 5142–5154.

37) Tsang CW, Harrison AG. *J. Am. Chem. Soc.*, **1978**, 98, 1301–1308.

38) Mueller, D.R.; Eckersley, M.; Richter, W. *Org. Mass Spectrom.* **1988**, 23, 217–222.

39) Johnson, R. S.; Krylov, D.; Walsh, K. A. *J. Mass Spectrom.,* **1995**, 30, 386–387.

40) Harrison, A.G.; Yalcin, T. *Int. J. Mass Spectrom. Ion Processes,* **1997**, 165, 339–347.

41) Csonka, I. P.; Paizs, B.; Lendvay, G.; Suhai, S. *Rapid Commun. Mass Spectrom.*, **2000**, 14, 417–431.

42) Csonka, I. P.; Paizs, B.; Lendvay, G.; Suhai, S. *Rapid Commun. Mass Spectrom.*, **2001**. 15, 1457–1472.

43) Paizs, B.; Suhai, S. *Rapid Commun. Mass Spectrom.*, **2001**, 15, 651–663.

44) Huang, Y.; Wysocki, V.H.; Tabb, D. L.; Yates, J. R. *Int. J. Mass Spectrom. Ion Processes,* 2002, 219, 233-244.

45) Kapp, E. A.; Schutz, F.; Reid, G. E.; Eddes, J. S.; Moritz, R. L.; O'Hair, R. A. J.; Speed, T. P.; Simpson, R. J. *Anal. Chem,* 2003, 75, 6251–6264.

46) Huang, Y.; Triscari, J. M.; Pasa-Tolic, L.; Anderson, G. A.; Lipton, M. S.; Smith, R. D.; Wysocki, V. H. *J. Am. Chem. Soc.,* 2004, 126, 3034-3035.

47) Huang, Y.; Triscari, J. M.; Pasa-Tolic, L.; Anderson, G. A.; Lipton, M. S.; Smith, R. D.; Wysocki, V. H. *Anal. Chem.,* 2005, 77, 5800-5813.

48) Tabb, D. L.; Huang, Y.; Wysocki, V. H.; Yates, J. R., III, *Anal. Chem.* **2004**, 76, 1243-1248.

49) Zhang, Z. *Anal. Chem*., **2004**, 76, 6374-6383.

50) Harrison, A.G. *Mass Spec. Rev*., **1997**, 16, 201-217.

51) Harrison, A. G. *J. Mass Spectrom.*, **1999**, 34, 577-589.

52) Paizs, B.; Suhai, S. *Rapid Commun. Mass Spectrom.* **2002**, 16, 1699-1702.

53) Harrison, A. G.; Young, A. B.; Bleiholder, B.; Suhai, S.; Paizs, B. *J. Am. Chem. Soc.* **2006**, *128*, 10364-10365.

54) Cooper, T.; Talaty, E.; Grove, J.; Suhai, S.; Paizs, B.; Van Stipdonk, M. *J. Am. Soc. Mass Spectrom.* **2006**,

55) http://nobelprize.org/nobel_prizes/chemistry/laureates/1998/press.html   Nobel Prize website.

56) Pople, J.A. Nobel Lecture http://nobelprize.org

57) Wiliams, I.H. *Chem. Soc. Rev*., **1993**, 22, 277-283.

58) Koch, W.; Holthausen, M.C.; A Chemist's Guide to Density Functional Theory, Second Edition, 2001 Wiley-VCH.

59) Jensen, F. Introduction to Computational Chemistry, Second Edition, 2007 John Wiley & Sons, Ltd.

60) Markus, R.A.; Rice, O.K.; *J. Phys. Colloid Chem.* **1951**, 55, 894-908.

61) Steinfeld, J.I.; Fransico, J.S.; Hase, W.L.; Chemical Kinetics and Dynamics, Second Edition, 1999.

62) Baer, T.; Hase, W.L.; Unimolecular Reaction Dynamics, Oxford, 1996.

63) Cramer, C.J. Essentials of Computational Chemistry, Second Edition, 2004, John Wiley & Sons, Ltd.

64) Lewars, E. Computational Chemistry, 2003, Kluwer Academic Publishers.

65) Born, M.; Oppenheimer, J.R.; *Ann. Physik.*, **1927**, 84, 457-484.

66) Eckart, C. *Physical Rev.*, **1935**, 46, 383-387.

67) Simons, J. Energetic Principles of Chemical Reactions, Jones and Bartlett (Boston), 1983.

68) K. Fukui. *Acc. Chem. Res.*, **1981**, 14, 363-368.

69) Gonzalez, C.; Schlegel, H. B.; *J. Chem. Phys.* **1989**, 90, 2154-2161.

70) Gonzalez, C.; Schlegel, H. B.; *J. Phys. Chem.* **1990**, 94, 5523-5527.

71) Paizs, B.; Lendvay, G.; Vékey, K.; Suhai, S. *Rapid Commun. Mass Spectrom.*, **1999**, 13, 525-533.

72) Rodriquez, C.F.; Cunje, A.; Shoeib, T.; Chu, I. K.; Hopkinson, A.C.; Siu, K. W. M. *J. Am. Chem. Soc.,* **2001**, 123, 3006-3012.

73) Balta, B.; Aviyente, V.; Lifshitz, C. *J. Am. Soc. Mass Spectrom*., **2003**, 14, 1192-1203.

74) Hohenberg, P.; Kohn, W., *Phys. Re*v., **1964**, 136, B864-B871.

75) Kohn, W.; Sham, L.J., *Phys. Re*v., **1965**, 140, A1133-A1138.

76) Parr, R.G.; Yang, W.; Density-Functional Theory of the Electronic Structure of Molecules, Oxford University Press, New York, 1989.

77) Stephens, P.J.; Devlin, J.F.; Chabalowski, C.F.; Frish, M.J.; *J. Phys. Chem*. **1994**, 98, 11623-11627.

78) Barone, V.; Orlandini, L.; Adamo, C. *Chem. Phys. Lett.* **1994**, 231, 295-*300*.

79) Mijoule, C.; Latajka, Z.; Borgis, D. *Chem. Phys. Lett.* **1993**, 208, 364-368.

80) Becke, A.D., J. Chem. Phys., **1988**, 88, 1053-1062.

81) Lee, C.; Yang, W.; Parr, R.G.; *Phys. Rev. B*, **1988**, 37, 785-789.

82) Lozynski, M.; Rusinska-Roszak, D.; Mack, H.-G. *J. Phys. Chem. A,* **1998**, 102, 2899-2903.

83) Adamo, C.; Barone, V.; in "Recent Advances in DFT Methods. Part II," Chong, D.P., Ed., World Scientific, Singapore, 1997.

84) Sim, F.; St.-Amant, A.; Papai, I.; Salahub, D. R. *J. Am. Chem. Soc*., **1992**, 114, 4391- 4400.

85) Ghosh, A. *Curr. Opin. Chem. Biol.* **2003**, 7, 110-112.

86) Paizs, B; Suhai, S., *J. Comput. Chem*., **1998**, 19, 575-584.

87) Rice, O.K.; Ramsperger, H.C.; *J. Am. Chem. Soc*. 1927, 49, 1617- 1629.

88) Rice, O.K.; Ramsperger, H.C.; *J. A.m Chem. Soc*. 1928, 50, 617-620.

89) Kassel, L.S. *J. Phys. Chem.* 1928, 32, 225-242.

90) Gilbert, R.G.; Smith, S.C.; Theory of Unimolecular and Recombination Reactions, 1990, Blackwell Scientific Publications.

91) Rosenstock, H.M.; Wallenstein, M.B.; Wahrhaftig, A.L.; Eyring, H.; *Proc. Nat. Acad. Sci.* 38, 1952, 667-678.

92) Beyer, T.; Swinehart, D.R.; *ACM Commun.* **1973**, 16, 379.

93) Simons, J. An Introduction to Theoretical Chemistry, 2003, Cambridge University Press.

94) Ochterski, J.W. Vibrational Analysis in Gaussian, 1999, available at http://www.gaussian.com/

95) Wang, P.; Kish, M.M.; Wesdemiotis, C. Encycl. Mass Spectrom. 2005, 2, 139-151.

96) Ballard, K.D.; Gaskell, S. J. *J. Am. Soc. Mass. Spectrom*., **1993**, 4, 477-481.

97) Cordero, M. M.; Houser, J. J.; Wesdemiotis, C. *Anal. Chem*. **1993**, 65, 1594-1601.

98) Yalcin, T.; Khouw, C.; Csizmadia, I. G.; Peterson, M. R.; Harrison, A. G. *J. Am. Soc. Mass. Spectrom.,* **1995**, 6(12), 1165-1174.

99) Yalcin, T.; Csizmadia, I. G.; Peterson, M. R.; Harrison, A. G. *J. Am. Soc. Mass. Spectrom*., **1996**, 7, 233-242.

100)   Dongré, A. R.; Somogyi, A.; Wysocki, V. H. *J. Mass Spectrom.* **1996**, 31(4), 339-350.

101)   Ambihapathy, K.; Yalcin, T; Leung, Hei-Wun; Harrison, *J. Mass Spectrom.* **1997**, 32, 209-215.

102)   Nold, M.J.; Wesdemiotis, C.; Yalcin, T.; Harrison, A. G.; *Int. J. Mass Spectrom. Ion Proc.* **1997**, 164, 137-153.

103)   Klassen, J.S.; Kebarle, P. *J. Am. Chem. Soc.* **1997**, 164, 6552-6563.

104)   Reid, G. E.; Simpson, R. J.; O'Hair, R. A. J.   *J. Am. Soc. Mass Spectrom.*, **1998**, 9(9), 945-956.

105)   O'Hair, R. A. J.; Styles, M. L.; Reid, G. E.   *J. Am. Soc. Mass Spectrom,* **1998**, 9(12), 1275-1284.

106)   Reid, G.E.; Simpson, R. J.; O'Hair, R. A. J.   *Int. J. Mass Spectrom. Ion Proc.*, **1999**, 190/191, 209-230.

107)   Nold, M. J.; Cerda, B. A.; Wesdemiotis, C. *J. Am. Soc. Mass Spectrom.,* **1999**, 10, 1-8.

108)   Paizs, B.; Szlavik, Z.; Lendvay, G.; Vékey, K.; Suhai, S. Formation of a_2^+ ions of protonated peptides. An ab initio study. *Rapid Commun. Mass Spectrom.*, **2000**, 14, 746-755.

109)   Harrison, A. G.; Csizmadia, I. G.; Tang, T.-H..; Tu, Y.-P. *J. Mass Spectrom.,* **2000**, 35, 683-688.

110)   Polce, M. J.; Ren, D; Wesdemiotis, C. *J. Mass Spectrom.,* **2000**, 35(12), 1391-1398.

111)   Laskin, J.; Denisov, E.; Futrell, J. H *J. Am. Chem. Soc.,* **2000**, 122, 9703-9714.

112)   Paizs, B.; Suhai, S. *Rapid Commun. Mass Spectrom.*, **2001**, 15, 651-663.

113)   Paizs, B.; Suhai, S. *Rapid Commun. Mass Spectrom*., **2002**, 16, 375-389.

114)   Paizs, B.; Suhai, S.; Harrison, A. G. *J. Am. Soc. Mass Spectrom.,* **2003**, 14, 1454-1469.

115)   Balta, B.; Aviyente, V.; Lifshitz, C. *J. Am. Soc. Mass Spectrom*., **2003**, 14, 1192-1203.

116)   Pingitore, F.; Polce, M. J.; Wang, P.; Wesdemiotis, C.; Paizs, B.   *J. Am. Soc. Mass. Spectrom*., **2004**, 15(7), 1025-1038.

117)   Harrison, A. G. Linear free energy correlations in mass spectrometry. *J. Mass Spectrom.*, **1999**, 34, 577-589.

118)   Cooper, T.; Talaty, E.; Grove, J.; Suhai, S.; Paizs, B.; Van Stipdonk, M. *J. Am. Soc. Mass Spectrom.* **2006**, **17**, 1654-1664.

119)   Kinser, R. D.; Ridge, D. P.; Hvistendahl, G.; Rasmussen, B.; Uggerud, E. *Chem.-Eur.J*., **1996**, 2, 1143-1149.

120)   March, R. E. *J. Mass Spectrom*. **1997**, 32, 351-369.

121)   March, R. E. *Rapid Commun. Mass Spectrom*. **1998**, 12, 1543-1554.

122)   Medzihradszky, K. F.; Campbell, J. M.; Baldwin, M. A.; Falick, Arnold M.; Juhasz, P.; Vestal, M. L.; Burlingame, A. L. *Anal. Chem.*, **2000**, 72(3), 552-558.

123)   Bienvenut, W. V.; Deon, C.; Pasquarello, C.; Campbell, J.M.; Sanchez, J.; Vestal, M.L.; Hochstrasser, D.F. *Proteomics*, **2002**, 2(7), 868-876.

124)   McLuckey, S. A.; Van Berkel, G. J.; Goeringer, D. E.; Glish, G. L. *Anal. Chem*. **1994**, 66,689A-696A.

125)   Gaussian 03, Revision C.02, Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, Jr., J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; and Pople, J. A.; Gaussian, Inc., Wallingford CT, **2004**.

126)   El Aribi, H.; Orlova, G.; Rodriquez, C.F.; Almeida, D. R. P.; Hopkinson, A.C.; Siu, K. W. M. *J. Phys. Chem. B,* 2004, 108(48),18743 -18749.

127)   Pingitore, F.; Wesdemiotis, C.; Paizs, B. *Proceedings of the 52nd ASMS Conference on Mass Spectrometry and Allied Topics*, Nashville, Tennessee, May 23-27, **2004**.

128)   Lifshitz, C. *Eur. J. Mass Spectrom.*, **2002**, 8, 85-98.

129)   Paizs, B.; Suhai, S. *Rapid Commun. Mass Spectrom.*, **2001**, 15, 2307-2323.

130)   Polfer, N. C.; Oomens, J.; Suhai, S.; Paizs, B. *J. Am. Chem. Soc.* **2005**, *127*, 17154-17155.

131)   Bleiholder, C.; Suhai, S.; Paizs, B. *J. Am. Soc. Mass Spectrom.* **2006**, 17, 1275-1281.

132)   Lifshitz, C. *Chem. Soc. Rev.*, **2001**, 30, 186-192.

133)   Paizs, B.; Csonka, I. P.; Lendvay, G.; Suhai, S. *Rapid Commun. Mass Spectrom.* 2001, 15, 637-650.

134)   Laerdahl, J. K.; Uggerud, E. *Int. J. Mass Spectrom*, **2002**, 214, 277-314.

135)   Gritsenko, O. V.; Ensing, B.; Schipper, P. R. T.; Baerends, E. J.; *J. Phys. Chem. A.* **2000**, *104,* 8558-8565.

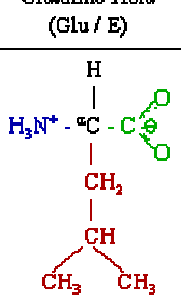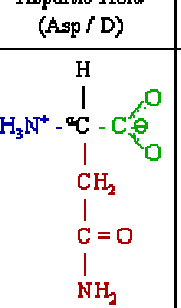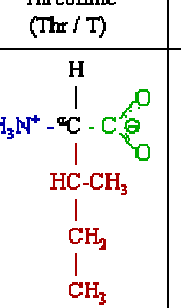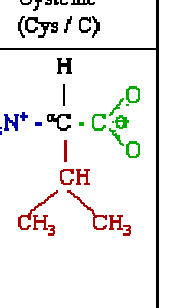136)   Tu, Y. -P.; Harrison, A. G. *J. Am. Soc. Mass Spectrom.,* **1998**, 9(5), 454-462.
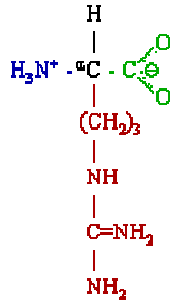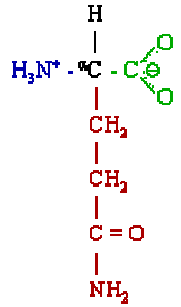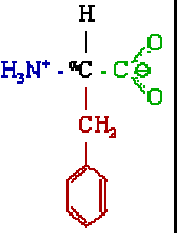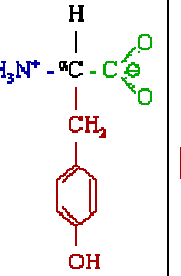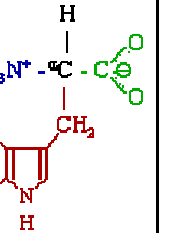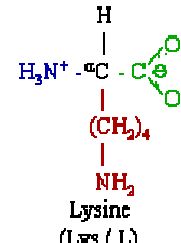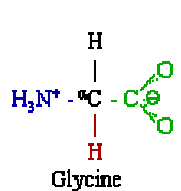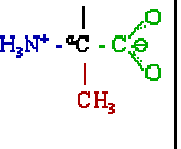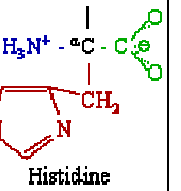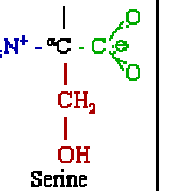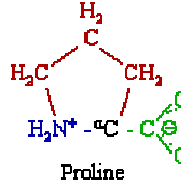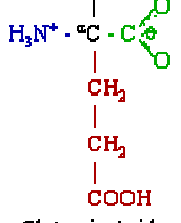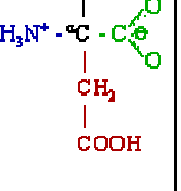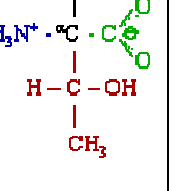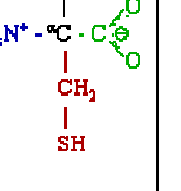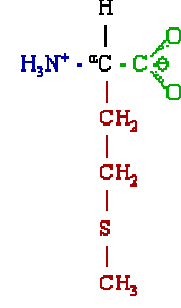
137)   McLafferty, F. W.; Turecek, F. *Interpretation of Mass Spectra*, 4th ed. University Science Books: Mill Valley, CA, **1993,** 51-84.

138)   Steen, H.; Mann, M. *Nature Rev. Mol. Cell Biol.* **2004**, 5**,** 699–711.

139)   Hunt, D. F.; Yates, J. R., III; Shabanowitz, J.; Winston, S.; Hauer, C. R. *Proc. Natl. Acad. Sci. U.S.A.* **1986***, 83*, 6233-6237.

140)   Biemann, K.; Scoble, H. A. *Science* **1987**, *237*, 992-998.

141)   Anderson, N.L.; Anderson, N.G *Mol. Cell. Prot*., 2002, 1, 845-867.

142)   Khatun, J.; Ramkissoon, K.; Giddings, M.G.; *Anal. Chem.* **2007**, 79, 3032-3040.

143)   Eng, J. K.; McCormack, A. L.; Yates, J. R., III. *J. Am. Soc. Mass Spectrom*. 1984, 5, 976.

144)   Perkins, D. N.; Pappin, D. J. C.; Creasy, D. M.; Cottrell, J. S. *Electrophoresis* 1999, 20, 3551.

145)   Biemann, K.; Martin, S. A. *Mass Spectrom. Rev.* **1987**, *6*, 1.

146)   Poulter, L.; Taylor, L. C. E. *Int. J. Mass Spectrom. Ion Processes* **1989**, *91*, 183.

147)   Tang, X.; Boyd, R. K. *Rapid Commun. Mass Spectrom.* **1992**, *6*, 651.

148)   Burlet, O.; Orkiszewski, R. S.; Ballard, K. D.; Gaskell, S. J. *Rapid Commun. Mass Spectrom.* **1992**, *6*, 658.

149)   McCormack, A. L.; Somogyi, A.; Dongre, A. R.; Wysocki, V. H. *Anal. Chem.* **1993**, *65*, 2859.

150)   Tang, X.-J.; Thibault, P.; Boyd, R. K. *Anal. Chem.* **1993**, *65*, 2824.

151)   Summerfield, S. G.; Whiting, A.; Gaskell, S. J. *Int. J. Mass Spectrom. Ion Processes* **1997**, *162*, 149.

152)   Harrison, A. G.; Tu, Y.-P. *J. Mass Spectrom.* **1998**, *33*, 532.

153)   Jorgensen, T. J. D.; Gardsvoll, H.; Ploug, M.; Roepstorff, P. *J. Am. Chem. Soc.* **2005**, *127*, 2785.

154)   Bythell, B.J.; Barofsky, D. F.; Pingitore, F.; Polce, M. J.; Wang, P.; Wesdemiotis, C.; Paizs, B. *J. Am. Soc. Mass Spectrom*., **2007**, 7, 1291-1303.

155)   Harrison, A. G.; Young, A. B.; Bleiholder, C.; Suhai, S.; Paizs, B. *J. Am. Chem. Soc.,* **2006**, 128 , 10364 -10365.

156)   Harrison, A. G.; Young, A. B. *J. Am. Soc. Mass Spectrom.* **2004**, *15*, 1810.

157)   She, Y-M.; Krokhin, O.; Spicer, V.; Loboda, A.; Garland, G.; Ens, W.; Standing**,** K. G.; Westmore, J. B. *J. Am. Soc. Mass. Spectrom*., **2007**, 18, 1024-1037.

158)   Thorne, G. C.; Ballard, K.D.; Gaskell, S. J. *J. Am. Soc. Mass. Spectrom*., **1990**, 1, 249-257.

159)   Ballard, K.D.; Gaskell, S. J. *J. Am. Chem. Soc*., **1992**, 114, 64-71.

160)   Polfer, N.; Oomens, J.; Suhai, S.; Paizs, B. *J. Am. Chem. Soc.,* 129, 5887-5897, 2007.

161)   Paizs, B.; Suhai. S.; Hargittai, B.; Hruby, V.J.; Somogyi, A. Int. J. Mass Spectrom. 219, 203–232, 2002.

162)   Fang, S.; Takao, T.; Satomi, Y.; Mo, W.; Shimonishi, Y. *J. Am. Soc. Mass. Spectrom*., 2000, 11, 345-351.

163)   Farrugia. J.M.; O'Hair, R.A.J. *Int. J. Mass Spectrom.*, 2003, 222, 229–242.

164)   Hiserodt, R. D.; Brown, S.M.; Swijter, D.F.H.; Hawkins, N.; Mussinan, C.

J. *J. Am. Soc. Mass. Spectrom.*, 2007, 18, 1414-1422.

APPENDICES

Appendix A:

The 20 naturally occurring amino acid residues. The single letter and three letter codes
are included. These are shown in zwitterions form, which is the natural form in solution.
From www.chemistrydaily.com/chemistry/Amino_acid.

Appendix B

Center of Mass Considerations in CAD

This actual amounts of energy transferred in collisions can be explained by looking at the processes in terms of center-of-mass frame (E_{COM}) where $E_{COM} = E_{LAB}[m_{gas}/(m_p+m_{gas})]$ for each collision (m_{gas} = mass of collision gas, m_p = mass of analyte ion). E_{COM} is the maximum amount of energy that can be transferred in a collision. This value is seldom achieved and varies as a function of collision partner amongst other things; more similarly sized collision partners are usually more efficient at transferring energy than very differently sized collision partners. The maximum possible total energy transfer is thus the number of collisions occurring multiplied by the E_{COM} for each collision.

For an ion of 100 m/z, colliding with a nitrogen, in a low energy instrument (E_{LAB} =30eV), E_{COM} = 6.6eV. In a MALDI-TOF/TOF, E_{LAB} = 1000eV and consequently, E_{COM} = 218eV. Assuming the transfer of energy is equally efficient in the two instruments, 33 collisions would have to occur in the low-energy instrument to transfer the same amounts of energy. The longer timescale of low-energy instrument experiments makes this number of collisions easily achievable. Thus the analyte ion may have higher internal energy in the low-energy instrument than the high-energy instrument.

Appendix C

Glossary of Abbreviations and Terms Used

B3LYP: hybrid density functional theory model; B = Becke exchange functional; LYP = Lee, Yang and Parr correlation functional; with 3 semiempirical coefficients used to determine the weighting of the various components.

CAD: collision-activated-dissociation

CID: collision-induced-dissociation

DFT: density functional theory

ESI: Electrospray Ionization

FAB: Fast-Atom-Bombardment

GB: gas-phase basicity $= -\Delta G^{\circ}$ for a protonation reaction, $(GB = PA + T\Delta S^{\circ})$.

FFR: field-free region

GM: global minimum

HP: Hartree-Fock

IRC: intrinsic reaction coordinate

IT: ion trap mass spectrometer

MALDI: Matrix-Assisted-Laser-Desorption/Ionization

MP2: 2nd order Møller-Plesset perturbation theory

PIC: Pathways In Competition fragmentation model

PFPs: peptide fragmentation pathways

PTMs: post translational modifications

PES: potential energy surface

PA: proton affinity, $= -\Delta H°$ for a protonation reaction (PA $=$ GB $-$ TΔS$°$)

PBD: proton-bound-dimer

RRKM: Rice-Ramsperger-Kassel-Markus

MS/MS or MS^n: tandem mass spectrometry

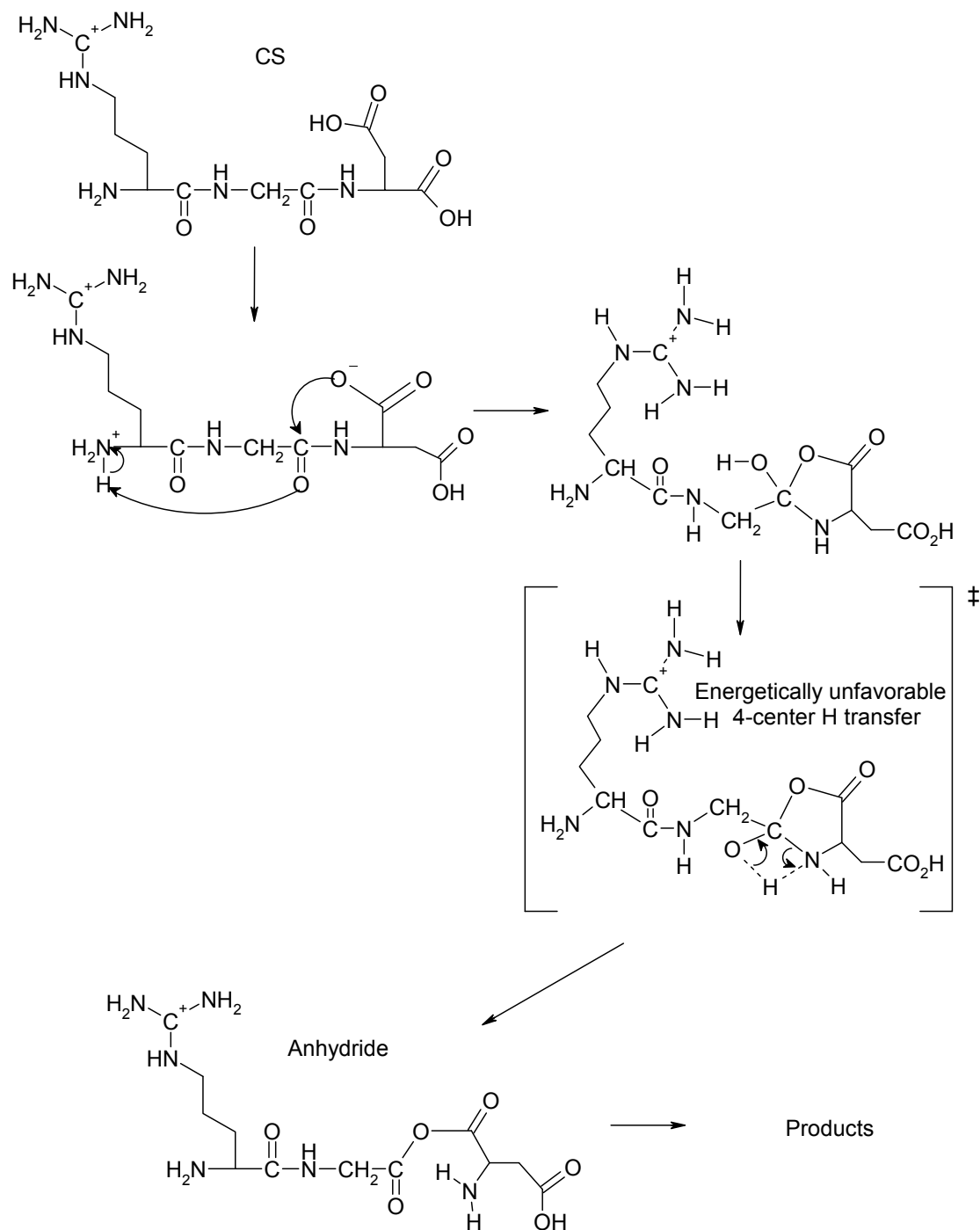TSs: transition states/structures

TOF: time-of-flight

TIS: timed ion selector device

ZPE: zero-point energy

Appendix D1

Proposed mechanism of anhydride formation Farrugia & O'Hair [57] & Mussiman et al [58].

Appendix D2
The b_2+H_2O PFP as proposed by Thorne et al. [37]