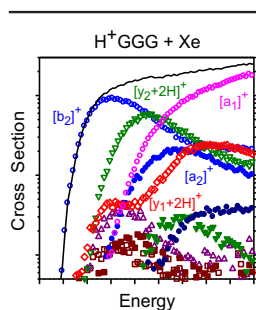


## RESEARCH ARTICLE

# Thermodynamics and Reaction Mechanisms of Decomposition of the Simplest Protonated Tripeptide, Triglycine: A Guided Ion Beam and Computational Study

Abhigya Mookherjee,<sup>1</sup> Michael J. Van Stipdonk,<sup>2</sup> P. B. Armentrout<sup>1</sup> <sup>1</sup>Department of Chemistry, University of Utah, 315 S.1400 E. Rm 2020, Salt Lake City, UT 84112, USA<sup>2</sup>Department of Chemistry and Biochemistry, Duquesne University, 600 Forbes Ave., Pittsburg, PA 15282, USA

**Abstract.** We present a thorough characterization of fragmentations observed in threshold collision-induced dissociation (TCID) experiments of protonated triglycine ( $\text{H}^+\text{GGG}$ ) with Xe using a guided ion beam tandem mass spectrometer (GIBMS). Kinetic energy-dependent cross-sections for 10 ionic products are observed and analyzed to provide 0 K barriers for six primary products:  $[\text{b}_2]^+$ ,  $[\text{y}_1 + 2\text{H}]^+$ ,  $[\text{b}_3]^+$ , CO loss,  $[\text{y}_2 + 2\text{H}]^+$ , and  $[\text{a}_1]^+$ ; three secondary products:  $[\text{a}_2]^+$ ,  $[\text{a}_3]^+$ , and  $[\text{y}_2 + 2\text{H} - \text{CO}]^+$ ; and two tertiary products: high energy  $[\text{y}_1 + 2\text{H}]^+$  and  $[\text{a}_2 - \text{CO}]^+$  after accounting for multiple ion-molecule collisions, internal energy of reactant ions, unimolecular decay rates, competition between channels, and sequential dissociations. Relaxed potential energy surface scans performed at the B3LYP-D3/6-311+G(d,p) level of theory are

used to identify transition states (TSs) and intermediates of the six primary and one secondary products. Geometry optimizations and single point energy calculations were performed at several levels of theory. These theoretical energies are compared with experimental energies and are found to give reasonably good agreement, in particular for the M06-2X level of theory. This good agreement between experiment and theory validates the reaction mechanisms explored computationally here and elsewhere and allows identification of the product structures formed at threshold energies. The present work presents the first measurement of absolute experimental threshold energies of important sequence ions and non-sequence ions:  $[\text{y}_1 + 2\text{H}]^+$ ,  $[\text{b}_3]^+$ , CO loss,  $[\text{a}_1]^+$ , and  $[\text{a}_3]^+$ , and refines those for  $[\text{b}_2]^+$  and  $[\text{y}_2 + 2\text{H}]^+$  previously measured.

**Keywords:** Protonated peptide, Glycylglycylglycine, Reaction mechanisms, Energetics, Mobile proton

Received: 29 January 2016/Revised: 20 December 2016/Accepted: 22 December 2016/Published Online: 14 February 2017

## Introduction

Tandem mass spectrometry (MS/MS) has been established as an important tool for studying gas-phase protonated peptide fragmentation in order to determine the amino acid sequence of peptides [1]. A typical collision-induced dissociation (CID) spectrum obtained at low collision energy is composed of multiple fragments from which the sequence of residues in a peptide can be elucidated. Effective sequencing requires rapid analysis of many CID spectra, which can be handled by automated algorithms [2–4]. A limitation of the existing sequencing algorithms is that spectra are matched

against a database of possible fragments by the mass to charge ratio without considering the relative propensities for fragmentation [5]. Also, in a process like de novo sequencing where the peptide's amino acid sequence is derived from a tandem MS/MS spectrum without the aid of a sequence database, non-sequence ions formed by losses of neutrals from primary fragments are often treated as noise by the algorithms barring some cases (e.g., peptides with serine).

Overall, limited knowledge of the underlying fragmentations may result in over-reliance on automated data interpretation, and may also prevent the critical assessment of the automated search results. Thus, sequence identification would benefit from understanding the fragmentation processes of protonated peptides, which in turn requires a thorough knowledge of the dissociation mechanisms. At present, such an understanding is incomplete even for a small peptide like triglycine made up of the “simplest” residues. Ultimately, being able to predict dissociation pathways and their probabilities will help the

**Electronic supplementary material** The online version of this article (doi:10.1007/s13361-016-1590-y) contains supplementary material, which is available to authorized users.

Correspondence to: P. B. Armentrout; e-mail: armentrout@chem.utah.edu

development of more efficient sequencing algorithms that make use of mass spectrometry data.

Gas-phase fragmentation of protonated peptides at low collision energies occurs via charge-directed mechanisms induced by migration of the proton added to the neutral peptide [5–10]. This leads to amide bond cleavage, which produces  $[b_n]^+$  (containing the N-terminus) and  $[y_n + 2H]^+$  (containing the C-terminus) ions, known as sequence ions because they provide information on the primary structure of the peptide. The sequence  $[b_n]^+$  and  $[y_n + 2H]^+$  ions can further dissociate to give smaller  $[b_m]^+$  ions,  $[y_m + 2H]^+$ , and  $[a_n]^+$  ions, internal fragments, or immonium ions. Non-sequence product ions are formed by the loss of small neutral molecules like water, carbon monoxide, and ammonia. The nomenclature used for peptide fragment ions in this study is adopted from the “all-explicit” nomenclature proposed by Siu and co-workers [11], a modification of that proposed by Roepstorff and Fohlmann [12], and Biemann [13].

The system of interest here, protonated triglycine,  $H^+GGG$ , is one of the most extensively studied model peptides. The earliest quantitative experimental study on the decomposition of  $H^+GGG$  was done by Klassen and Kebarle (KK) [14], who measured the CID of small protonated molecules and peptides as a function of kinetic energy in a triple quadrupole mass spectrometer. Reid, Simpson, and O’Hair [15] reexamined this system and performed MS/MS experiments that were not kinetic energy-resolved. In both studies, possible mechanisms for product formation were suggested, but no detailed mechanisms were explicitly developed for the  $H^+GGG$  system. Aribi et al. updated these results by investigating the fragmentations of  $H^+GGG$  experimentally using threshold collision-induced dissociation (TCID) and extracting threshold energies for three processes from energy-resolved CID spectra [16]. They also used density functional theory (DFT) at the B3LYP/6-31++G(d,p) level to elucidate the fragmentation mechanisms, building on previous computational work by Rodriguez et al. (RCSCHS), who had explored the “proton migration and tautomerism” in  $H^+GGG$  [17]. Studies undertaken by Paizs and Suhai (PS) took a close theoretical look at the  $H^+GGG$  system, albeit with a relatively small basis set, B3LYP/6-31G(d) [18, 19]. They identified key conformers (intermediates) formed as a result of protonation at different sites and the transition states (TSs) connecting the intermediates involved in the cis-trans isomerization of the peptide bond. By calculating unimolecular rate constants for TSs using Rice-Ramsperger-Kassel-Marcus (RRKM) theory [20, 21], the authors concluded that cis-trans isomerization of the N-terminal peptide bond in  $H^+GGG$  is energetically feasible but is a time-consuming process. Such cis-trans isomerization is crucial only if  $[y_1 + 2H]^+$  ions are formed via the “diketopiperazine” (DKP) pathway. PS also investigated the mechanisms for fragmentation of  $H^+GGG$  to  $[b_2]^+$  and  $[y_1 + 2H]^+$  ions via both the oxazolone and DKP pathways [19]. Although formation of  $[y_1 + 2H]^+$  was found to be loose (limited by product energies) along the oxazolone pathway, the DKP pathway was found to be limited by a tight TS (TTS). The DKP pathway is inconsistent with the experimental results obtained

by Morgan and Bursey [22], who saw a linear correlation between the relative intensities of  $[y_1 + 2H]^+$  and  $[b_2]^+$  and the proton affinities of the C-terminal amino acid residue (glycine in this case), suggesting that  $[b_2]^+$  and  $[y_1 + 2H]^+$  products share common pathways. Likewise, Wesdemiotis and co-workers [23] observed a very small kinetic energy release for formation of  $[y_1 + 2H]^+$  ions, consistent with a loose TS. Fragmentation of C-terminally amidated  $H^+GGG$ ,  $H^+GGGNH_2$ , was recently examined experimentally by MS/MS, infrared multiple photon dissociation (IRMPD), and chemical dynamics simulations [24]. Detailed reaction mechanisms were not characterized, although possible pathways were suggested.

The present study extends these previous works by providing a more comprehensive examination of the reaction mechanisms, detailed quantitative data, and much more in-depth analyses of the data, thereby providing more extensive and quantitative energetic information. Here, we measure absolute cross-sections of the observed TCID reactions with a sensitivity of nearly three orders of magnitude using a guided ion beam tandem mass spectrometer (GIBMS). An important result of this study is the assortment of sequence and non-sequence ions that are obtained as products from  $H^+GGG$  fragmentation (seven more than previously identified). These experimental results are analyzed using molecular parameters (vibrational frequencies and rotational constants) determined in theoretical calculations to extract absolute experimental threshold energies of each fragmentation pathway. The experimental threshold energies are compared with those obtained from single point energy calculations performed at the B3LYP, B3LYP-D3 (B3LYP with the GD3BJ dispersion correction [25, 26]), MP2(full), and M06-2X levels of theory. Agreement between the experimental and theoretical energies verifies that the key steps involved in  $H^+GGG$  fragmentation have been properly identified and supports the mobile proton model [6, 8–10, 27], which allows an understanding of peptide dissociation chemistry. The present analysis of the kinetic energy-dependent cross-sections of  $H^+GGG$  fragmentations provides the first complete experimental determination of the threshold energies of sequence and non-sequence ions formed from decomposition of any tripeptide.

## Experimental and Computational Details

### *Experimental Procedures*

Cross-sections of  $H^+GGG$  colliding with Xe are measured using a GIBMS that has been described previously in detail [28, 29]. The  $H^+GGG$  ions were generated using an electrospray ionization source (ESI) [30] under conditions similar to those described earlier, such that the  $H^+GGG$  ions are characterized by a temperature of 300 K [31–36]. Details of the experimental methods and means of data analysis can be found in the Supporting Information (SI). All energies reported here are in the center-of-mass (CM) frame unless otherwise noted.

### Computational Details

Model structures and vibrational frequencies of reactants, products, TSs, and intermediates were obtained at B3LYP/6-311+G(d,p) and B3LYP-D3/6-311+G(d,p) levels of theory. Single-point energies corrected for zero-point energies were calculated at B3LYP, B3LYP-D3, and MP2(full) (where full refers to correlation of all electrons, simply referred to as MP2 below) levels of theory using the 6-311+G(2d,2p) basis set. Details of these procedures can be found in the SI. Geometries of key H<sup>+</sup>GGG conformers, TSs, and products were also optimized at M06-2X/6-311+G(d,p) with single-point energies calculated at B3LYP, MP2, and M06-2X levels of theory. In the following, basis sets are not explicitly given as they always correspond to L1/6-311+G(2d,2p)/L2/6-311+G(d,p), where L1 refers to the level of theory for single-point energies and L2 for geometry optimizations. Single-point energies determined at the B3LYP and MP2 levels using 6-311+G(2d,2p) basis sets compared accurately with experimental energies (with mean absolute deviations of about 10 kJ/mol) in studies of protonated glycine (H<sup>+</sup>G) [37], diglycine (H<sup>+</sup>GG) [38, 39], [b<sub>2</sub>]<sup>+</sup> (H<sup>+</sup>AMOX), and diketopiperazine (H<sup>+</sup>DKP) [40] and also illustrate the range of energies that can be obtained computationally.

### Nomenclature

Conformers of the H<sup>+</sup>GGG reactant are named using a nomenclature that identifies the site of protonation within square brackets, followed by designation of eight dihedral angles starting from the N terminus and going along the backbone of H<sup>+</sup>GGG to the hydroxyl group: ∠N<sup>1</sup>CCN<sup>2</sup>, ∠CCN<sup>2</sup>C, ∠CN<sup>2</sup>CC, ∠N<sup>2</sup>CCN<sup>3</sup>, ∠CCN<sup>3</sup>C, ∠CN<sup>3</sup>CC, ∠N<sup>3</sup>CCO<sup>4</sup>, ∠CCO<sup>4</sup>H, where the superscripts 1, 2, 3 denote the first, second, and third glycine residue, and superscript 4 indicates the hydroxyl oxygen, the point where the fourth peptide bond would form. The dihedral angles are labeled as c (cis) for angles <45°, g (gauche) for angles between 45° and 135°, and t (trans) for angles >135°. Thus the B3LYP ground structure (GS) is named [O<sup>1</sup>]-ctgtttt, where the site of protonation is the carbonyl oxygen of the first glycine residue, and the superscript t denotes the orientation of the proton by specifying that ∠CCO<sup>1</sup>H is trans. When the lone pair on the N-terminal amino group is trans to the C–C bond, ∠N<sup>1</sup>CCN<sup>2</sup> is denoted by a superscript t. In some cases, we assign subscripts + and – for gauche and trans dihedral angles having opposite signs and leading to distinct conformations. Product conformers are named in a similar manner, defining the protonation site and the relevant dihedral angles.

Transition states are indicated by TS followed by a description of the change in the protonation site, dihedral angle, or bond cleavage. Thus, the TS for a proton transfer is named TS[O<sup>1</sup>-N<sup>2</sup>] indicating transfer of a proton from the carbonyl oxygen O<sup>1</sup> (the lower energy conformer) to the amide nitrogen N<sup>2</sup>. A TS involving a dihedral angle change is designated within parentheses, e.g., (cg), which indicates that the dihedral angle changes from cis (the lower energy conformer) to gauche. A TS involving cleavage of a bond indicates the bond broken by ~ inside curly brackets, e.g., {OC~N<sup>2</sup>}. Although

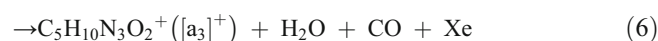
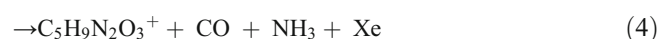
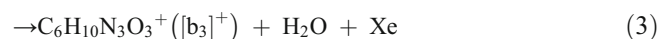
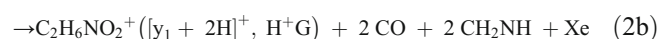
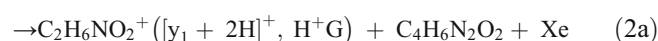
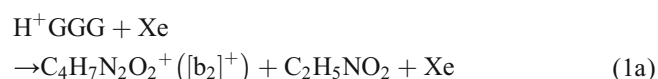
more complicated than simply numbering species, this nomenclature has proven to be useful in distinguishing conformations, while depicting connections between intermediates, and allows better visualization of the change taking place [37, 39–41].

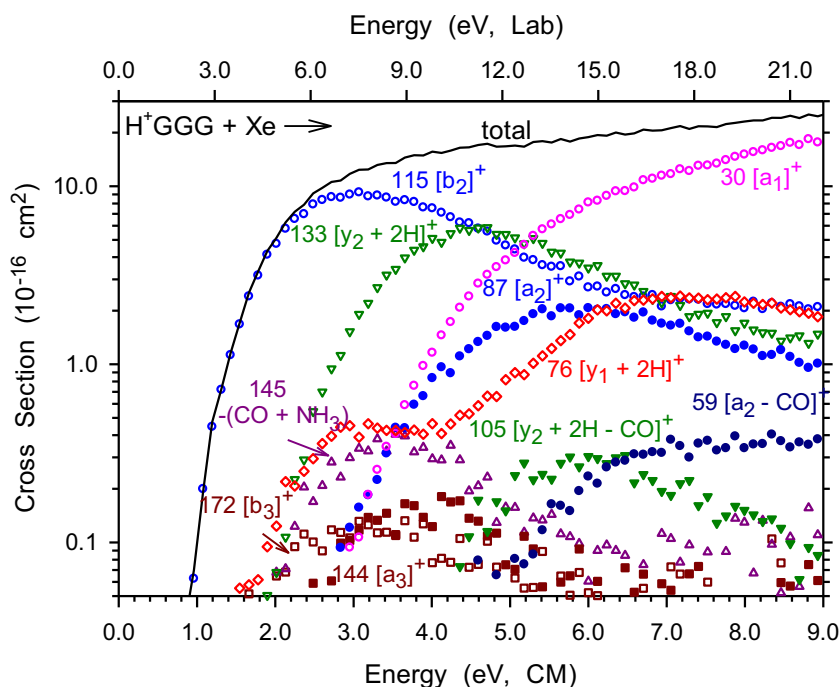
Energies for all species are given in terms of the range of values obtained computationally from the single point energy calculations corrected for zero point energies at the B3LYP, B3LYP-D3, MP2, and M06-2X levels of theory relative to their respective GSs (i.e., [O<sup>1</sup>]-ctgtttt for DFT and [N<sup>1</sup>]-ttgtgtt for MP2), as discussed in the SI. Complete tables of all computed values can be found in SI: Supplementary Tables S1 for H<sup>+</sup>GGG conformers, S3 for GGG conformers, S4 for intermediates, and S5 for TSs.

## Results

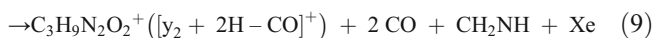
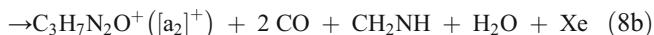
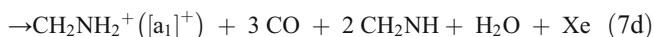
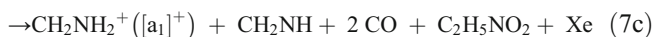
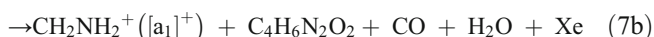
### Cross-Sections for Collision-Induced Dissociation

Figure 1 shows the experimental kinetic energy-dependent cross-sections obtained for the interaction of H<sup>+</sup>GGG with Xe. Ten ionic products are observed for H<sup>+</sup>GGG (*m/z* 190) fragmentation. On the basis of the theoretical results, these are assigned to reactions 1a–10.





**Figure 1.** Cross-sections for collision-induced dissociation of  $\text{H}^+\text{GGG}$  with xenon at 0.2 mTorr as a function of kinetic energy in the center-of-mass (CM) frame in the lower x-axis and in the laboratory frame (Lab) in the upper x-axis. Products are identified by their mass to charge ratio and the all explicit nomenclature for fragment ions. Open symbols show the primary products and closed symbols show the products formed by sequential dissociation. The solid line shows the total cross-section



Importantly, Figure 1 shows that the total cross-section increases smoothly with increasing energy, consistent with the sequential nature of the coupled reactions 1a–10 as outlined in the following.

The lowest energy process arises from the loss of neutral glycine to form the  $[\text{b}_2]^+$  ion ( $m/z$  115) in reaction 1a. Previous work in our laboratory and others has identified this species as protonated 2-aminomethyl-5-oxazolone ( $\text{H}^+\text{AMOX}$ ) [16, 19, 40, 42–47] as originally suggested by Harrison and co-workers [48]. The next largest cross-section is associated with formation of  $[\text{y}_2 + 2\text{H}]^+$  ( $m/z$  133, protonated diglycine,  $\text{H}^+\text{GG}$ ) in reaction 5. At slightly lower energies than  $[\text{y}_2 + 2\text{H}]^+$  formation, the  $[\text{y}_1 + 2\text{H}]^+$  ion ( $m/z$  76, protonated glycine,  $\text{H}^+\text{G}$ ) is formed in reaction 2a, which directly competes with reaction 1a because both channels differ only in which fragment keeps the extra proton. After rising to about  $0.5 \text{ \AA}^2$ , the  $[\text{y}_1 + 2\text{H}]^+$  cross-section rises again near 4 eV, indicating the contribution of a new pathway to this product, reaction 2b. Formation of the  $[\text{b}_3]^+$  ion ( $m/z$  172) has a similar apparent threshold as  $[\text{y}_2 + 2\text{H}]^+$  and corresponds to loss of water from  $\text{H}^+\text{GGG}$  to form protonated 2-glycylaminomethyl-5-oxazolone ( $\text{H}^+\text{GAMOX}$ ) in reaction 3. A similar threshold is also found for the  $m/z$  145 product ion formed in reaction 4 by the combined losses of  $\text{CO} + \text{NH}_3$ . Although both reactions 4 and 5 include the loss of  $\text{CO}$ , no primary ion at the appropriate mass ( $m/z$  162) was observed, despite carefully looking for it. This can occur if  $\text{CO}$  loss involves a TTS such that loss of additional fragments from  $m/z$  162 can occur readily at the energy of the TTS. This result is also consistent with similar low-energy and millisecond timescale ion trap/CID experiments of Bythell et al. for fragmentation of  $\text{H}^+\text{AGG}$  [49]; however, they did observe the loss of  $\text{CO}$  in MALDI/TOF-

TOF, metastable ion (MI), and CID experiments occurring at microsecond timescales and higher energies. The next product observed here is formed in reaction 6, where CO is lost from the  $[b_3]^+$  ion to form  $[a_3]^+$  ( $m/z$  144). Formation of  $[a_1]^+$  ( $m/z$  30) in reaction 7 has a similar apparent threshold as  $[a_3]^+$  but is clearly much more efficient, dominating the products seen at high energies. This can be attributed to the multiple pathways that are capable of yielding  $[a_1]^+$  adding to its overall product intensity, discussed below. At about the same apparent threshold as  $[a_1]^+$ , formation of  $[a_2]^+$  ( $m/z$  87) in reaction 8 is observed, and corresponds to loss of CO from  $[b_2]^+$ . Two minor products,  $[y_2 + 2H - CO]^+$  ( $m/z$  105) and  $[a_2 - CO]^+$  ( $m/z$  59,  $H^+(CH_2NH_2)_2$ ) are observed at higher energies and correspond to decarbonylation of  $[y_2 + 2H]^+$  in reaction 9 and  $[a_2]^+$  in reaction 10.

Cross-sections of five of the six primary products ( $[b_2]^+$ ,  $[y_2 + 2H]^+$ ,  $[y_1 + 2H]^+$ ,  $[b_3]^+$ , and  $-(CO + NH_3)$ ) start to decline after an initial rise to a maximum cross-section. The  $[b_2]^+$  ion in reaction 1a can lose CO to form  $[a_2]^+$ , studied previously in our laboratory and by others [14, 16, 40, 43–46, 48, 50–55]. This is consistent with the decline of  $[b_2]^+$  and rise of  $[a_2]^+$  at about 3 eV, Figure 1. The  $[a_2]^+$  cross-section declines at higher energies because it decomposes by additional CO loss to form  $m/z$  59 ( $[a_2 - CO]^+$ ), which then dissociates to  $[a_1]^+$  in reaction 7c, as shown previously [40, 56]. The decline in the  $[y_2 + 2H]^+$  ( $H^+GG$ ) cross-section at higher energies is associated with loss of  $H_2O$  to form  $[b_2]^+$ , reaction 1b, and with loss of CO, reaction 9, as shown in previous work [38]. The  $[y_2 + 2H - CO]^+$  ( $m/z$  105) product ion rapidly dissociates to  $[y_1 + 2H]^+$  ( $H^+G$ ) by reaction 2b, which accounts for the second feature in the  $[y_1 + 2H]^+$  cross-section starting at  $\sim 4$  eV. These relative energetics are consistent with a study of  $H^+GG$  (here the  $[y_2 + 2H]^+$  product ion) decomposition where formation of  $m/z$  76 begins at  $\sim 1.4$  eV (which should be added to the threshold for  $[y_2 + 2H]^+$  formation) [38]. The primary  $[y_1 + 2H]^+$  product cross-section declines above 3 eV because it can dissociate by loss of  $CO + H_2O$  to form  $[a_1]^+$  in reaction 7b, consistent with direct studies of the decomposition of  $H^+G$  [37]. The  $[b_3]^+$  primary product dissociates by CO loss to form  $[a_3]^+$  in reaction 6, consistent with the shapes and magnitudes of these two cross-sections. The  $[a_3]^+$  product can dissociate further to lose  $CH_2NH$  forming the  $[b_2]^+$  ion (reaction 1b), a process that has been examined and documented in the past [51, 57, 58], or lose  $AMOX$  to form  $[a_1]^+$  ions in reaction 7b. Neither of these pathways is evident in the  $[b_2]^+$  and  $[a_1]^+$  cross-sections as these products are also formed via other more probable routes. The  $m/z$  145 product ion formed by loss of  $CO + NH_3$  must also dissociate at higher energies, but because the magnitude of its cross-section is relatively low, no subsequent product channels can be assigned definitively.

Note that there are five pathways leading to formation of  $[a_1]^+$  ( $CH_2NH_2^+$ ): a primary decarbonylation of  $H^+GG$  (7a) competitive with reaction 5, a secondary product formed via  $[y_1 + 2H]^+$  (7b), a quaternary product formed via  $[b_2]^+$  and  $[a_2]^+$  (7c), a tertiary product formed via  $[y_2 + 2H]^+$  (7c), and a quaternary product formed via  $[y_2 + 2H]^+$  and  $[y_1 + 2H]^+$  (7d) [37, 38, 40]. A sixth potential process in which the  $[b_2]^+$

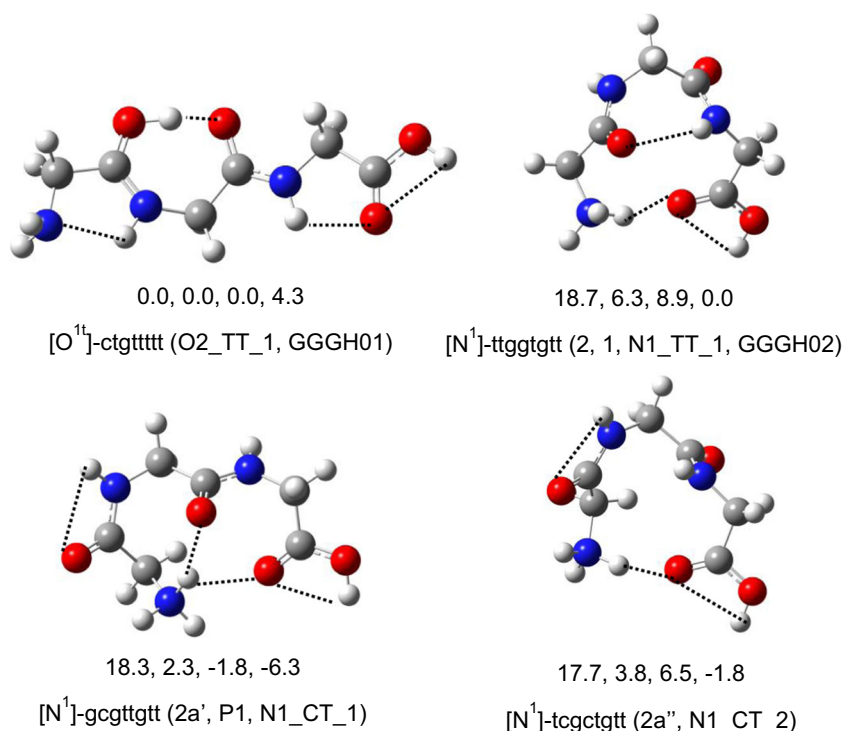
ion decomposes directly to form  $[a_1]^+$  along with  $C_3H_3NO_2$ , originally proposed by Siu and co-workers [52], was found to be unimportant in our previous analysis of the  $[b_2]^+$  ion decomposition [40]. Further support for this conclusion comes from a recent study that examined the direct fragmentation of  $[b_2]^+$  to  $[a_1]^+$  for a series of protonated tripeptides and concluded that this dissociation is insignificant for  $[b_2]^+$  ions with a glycyglycine sequence [59]. Notably, the magnitude of the  $[a_2 - CO]^+$ ,  $H^+(CH_2NH_2)_2$ , product cross-section in the current and previous studies [40] are comparable,  $\sim 0.4 \text{ \AA}^2$  at about 8 eV in the CM frame, although these cross-sections have slightly different shapes, which can be attributed to the different energy dependences of their precursor  $[a_2]^+$  ion.

As mentioned above, the decomposition of  $H^+GGG$  as a function of kinetic energy has been investigated previously by KK in a triple quadrupole mass spectrometer [14]. They observed the ionic products:  $[b_2]^+$ ,  $[y_1 + 2H]^+$ ,  $[y_2 + 2H]^+$ ,  $[a_1]^+$ , and  $[a_2]^+$ , the dominant products observed in the current results (those with cross-sections exceeding  $1 \text{ \AA}^2$ ). Similar to our results,  $[b_2]^+$  was found to be the energetically favored process, with  $[y_2 + 2H]^+$  next,  $[y_1 + 2H]^+$  showed bimodal behavior, and  $[a_1]^+$  was the dominant product at high energies. Relative intensities of  $[y_2 + 2H]^+$ ,  $[y_1 + 2H]^+$ ,  $[a_1]^+$ , and  $[a_2]^+$  are similar to those observed here, but KK's data differ mainly in the relative  $[b_2]^+$  intensity, which is smaller than observed here, and in a pronounced low energy tail to the  $[a_1]^+$  cross-section. In their MS/MS spectra of  $H^+GGG$ , Reid et al. observed the formation of  $[b_2]^+$ ,  $[b_3]^+$ ,  $[y_1 + 2H]^+$ , combined loss of  $CO + NH_3$ , and  $[a_2]^+$  products with the  $[b_2]^+$  product ion being the most intense [15]. No energy-resolved data were acquired. Aribi et al. [16] examined the kinetic energy dependence of  $H^+GGG$  fragmentation to form  $[b_2]^+$  followed by its decarbonylation to  $[a_2]^+$ , and formation of  $[a_1]^+$  ions from both  $[b_2]^+$  and  $[a_2]^+$ . (None of the other products observed here were included and only the  $[b_2]^+$  cross-section was shown.) These cross-sections were analyzed to obtain experimental thresholds.

### $H^+GGG$ Ground Structure

B3LYP//B3LYP, B3LYP-D3//B3LYP-D3, B3LYP//M06-2X, and M06-2X//M06-2X calculations suggest that protonation occurs most favorably at the carbonyl oxygen  $O^1$  of  $H^+GGG$ , Figure 2. Thus, these levels of theory indicate  $[O^{1t}]$ -ctgttttt is the ground structure (GS), in concert with the B3LYP results of PS (for their **O2\_TT\_1**) [18, 19] and Wu and McMahon (WM) (for their **GGGH01**) [60]. In this conformer, the proton on  $O^1$  interacts with the carbonyl oxygen of the second glycine residue, leading to a stabilizing  $O^1H \cdots O^2$  hydrogen bond of 1.42 Å. RSCSHS [17] found that the  $[O^{1c}]$ -ttttgttt conformer is the GS, which is calculated here to be 16–20 kJ/mol above  $[O^{1t}]$ -ctgttttt, Supplementary Table S1.

In contrast, calculations done at MP2//B3LYP and MP2//M06-2X find  $N^1$  is the most favorable site for protonation, suggesting that  $[N^1]$ -gcgtgttt (called **2a'** by Wu and Lebrilla [61], **P1** by RSCSHS, and **N1\_CT\_1** by PS) is the lowest energy species, with  $[O^{1t}]$ -ctgttttt being 2–6 kJ/mol higher,



**Figure 2.** H<sup>+</sup>GGG conformers, [O<sup>1</sup>]-ctgttttt, [N<sup>1</sup>]-gcgttggt, [N<sup>1</sup>]-ttggtggt, and [N<sup>1</sup>]-tcgctggt with their relative energies at B3LYP//B3LYP, B3LYP-D3//B3LYP-D3, M06-2X//M06-2X, and MP2//B3LYP levels of the theory

Supplementary Table S1. We also located two more [N<sup>1</sup>] conformers with similar energies: [N<sup>1</sup>]-ttggtggt (**2** by Zhang et al. [62, 63], **1** by RCSCHS, N1\_TT\_1 by PS, GGGH02 by WM), and [N<sup>1</sup>]-tcgctggt (**2a''** by Wu and Lebrilla, N1\_CT\_2 by PS). Both [N<sup>1</sup>]-gcgttggt and [N<sup>1</sup>]-tcgctggt conformers are characterized by a cis N-terminal peptide bond (second dihedral angle), hence the CT designation of PS, whereas [O<sup>1</sup>]-ctgttttt and [N<sup>1</sup>]-ttggtggt have trans N-terminal peptide bonds. For overlapping species and levels of theory, our results for these species are comparable to the calculations of WM, PS, and RCSCHS, Supplementary Table S1; however, only PS considered all four of these conformers.

WM found that a combination spectrum of their lowest energy isomers, [O<sup>1</sup>]-ctgttttt and [N<sup>1</sup>]-ttggtggt, compared well with their experimentally obtained IRMPD spectrum, thus concluding that both species are present in their experiments. Here, we find that both [N<sup>1</sup>]-gcgttggt and [N<sup>1</sup>]-tcgctggt are *lower* in energy than [N<sup>1</sup>]-ttggtggt at all levels of theory, and neither was considered by WM. The predicted spectra for these four conformers are compared with the experimental IRMPD spectrum of WM in Supplementary Figure S2. The discussion in the SI suggests that [O<sup>1</sup>]-ctgttttt and [N<sup>1</sup>]-ttggtggt are indeed the dominant species, in agreement with WM's conclusions. The two lower energy [N<sup>1</sup>] conformers are either not present or have low abundance. These results indicate that B3LYP relative energies for the [O<sup>1</sup>] and [N<sup>1</sup>] conformers are likely to be inaccurate, and either the B3LYP-D3, MP2, and M06-2X relative energies are also flawed or the two low-energy [N<sup>1</sup>] isomers having a cis N-terminal peptide bond are inaccessible to an ESI source. The latter hypothesis is explored more fully in

the SI, where we conclude that the cis species are probably not formed in solution or in the ESI source because the neutral GGG molecule does not have ready access to the cis peptide bond conformers. Additionally, the rate-limiting TS for cis-trans isomerization in the protonated molecule, TS[N<sup>1</sup>]-g(ct)gttttt, lies 61–80 kJ/mol above the GS (Supplementary Figure S4 and Supplementary Table S5), which is high enough to prevent the cis-trans isomerization from occurring in solution or an ESI source. As a consequence, Figure 2 and Supplementary Tables S1, S4, and S5 list MP2 energies of all species relative to the [N<sup>1</sup>]-ttggtggt isomer and M06-2X energies relative to [O<sup>1</sup>]-ctgttttt, as these are the lowest energy conformer according to each of these levels that is accessible experimentally. For B3LYP and B3LYP-D3, energies are listed relative to their [O<sup>1</sup>]-ctgttttt ground structure.

In analogy to this work, both [N<sup>1</sup>] and [O<sup>1</sup>] structures were found in the IRMPD study of H<sup>+</sup>GGGNH<sub>2</sub> [24]. The N-terminal protonated H<sup>+</sup>GGGNH<sub>2</sub> was calculated to be the lowest energy H<sup>+</sup>GGGNH<sub>2</sub> at DFT levels of theory. Although the authors identified two [N<sup>1</sup>] structures with folded conformations as major contributors to the experimental spectra of H<sup>+</sup>GGGNH<sub>2</sub>, an [O<sup>1</sup>] structure with a linear backbone (similar to [O<sup>1</sup>]-ctgttttt) was required to account for some features of their IRMPD spectra.

Additionally, Wyttenbach, Bushnell, and Bowers [64, 65] measured experimental collision cross-sections (CCS) of protonated Gly<sub>n</sub> (n = 1–6) oligopeptides, with their values for n = 1–3 listed in Supplementary Table S2. Their ion mobility experiments used a matrix laser desorption ionization (MALDI) source and were conducted at 300 and 80 K (where

the resolution at the colder temperature is twice that at room temperature). The experimental CCS measurements found no evidence of multiple structures for any of the systems, which is expected for the smaller  $H^+G$  and  $H^+GG$  systems, but is inconsistent with the IRMPD studies for  $H^+GGG$ . This disparity can be attributed to the use of different ion sources for the two sets of experiments. As discussed in detail in the SI, the CCSs can be compared with those computed by two different programs [65, 66]. Good agreement between experiment and theory for  $H^+G$  and  $H^+GG$  is found, but not for the  $H^+GGG$  system, such that an unequivocal assessment of which conformer is present cannot be made. However, the low-energy  $[N^1]$  structures located here are most consistent with the experimental CCS results as discussed in detail in the SI, which would suggest that MP2 theory provides the most reliable relative energetics.

Given that both  $[N^1]$ -ttggtgtt and  $[O^1]$ -ctgttttt isomers of  $H^+GGG$  appear to be present in the results of WM, we have explored the pathway connecting these low-energy isomers, described in detail in SI, Supplementary Figure S5. Overall, the rate-limiting TS connecting the  $[N^1]$ -ttggtgtt and  $[O^1]$ -ctgttttt isomers is  $TS[O^1c]-(tc)tggtttt$ , which lies 77–82 kJ/mol above the GS. Such energies are probably inaccessible in an ESI source such that interconversion of  $[O^1]$ -ctgttttt and  $[N^1]$ -ttggtgtt are unlikely, and thus must be formed independently in the source. However, this TS energy is sufficiently low that interconversion of these isomers is energetically possible in the TCID experiments, as this energy is well below the thresholds for the decomposition reactions discussed below,  $\geq 171 \pm 6$  kJ/mol.

### $[b_2]^+/[y_1 + 2H]^+$ Formation

The reaction mechanism for the formation of  $[b_2]^+$  and  $[y_1 + 2H]^+$  ions has been studied previously [17, 19, 49]. At a higher level of theory, we reproduce the pathway found by PS, as shown in Supplementary Figure S6 starting from  $[O^1]$ -ctgttttt, the DFT GS. As noted above, rearrangement of this species to the MP2 GS,  $[N^1]$ -ttggtgtt, is possible upon activation such that the rate-limiting TS (in all mechanisms explored here) is the same for both isomers. Overall, formation of the  $[b_2]^+$  and  $[y_1 + 2H]^+$  product ions is limited by the energies of the separated products (i.e., they have loose TSs). This conclusion is consistent with the large cross-sections observed for these two products in Figure 1. Formation of  $[b_2]^+ + G$  is calculated to lie 153–172 (170 by PS) kJ/mol above the GS and that of  $[y_1 + 2H]^+ (H^+G) + AMOx$  is 182–200 (188 by PS) kJ/mol above the GS, and 20–30 kJ/mol above  $[b_2]^+ + G$ . RCSCHS reported a pathway for the formation of the  $[b_2]^+$  ion with a rate-limiting TS above the pathway elucidated here by  $\sim 30$  kJ/mol. Importantly, RCSCHS used relative free energies to assess the rate-limiting steps; however, comparison of relative enthalpies is needed for TCID measurements, which determines bond enthalpies at 0 K. Entropy cannot be used to overcome barriers; it only determines the relative probability of the process (i.e., the Arrhenius pre-exponential factor in the rate coefficient).

### $[y_2 + 2H]^+/[a_1]^+$ Formation

Formation of the  $[y_2 + 2H]^+$  and  $[a_1]^+$  product ions from  $H^+GGG$  has been examined in the past [14, 19]. Here we have reexamined these mechanisms at higher levels of theory and report the entire mechanism in SI, Supplementary Figure S8. Briefly the formation of  $[y_2 + 2H]^+$  and  $[a_1]^+$  product ions must involve transfer of the mobile proton to the  $N^2$  amide nitrogen and the subsequent loss of CO. We find that the rate-limiting TS,  $TS[N^2]-t_tgttttt\{N^1C-CO^1-N^2C\}$  for the decarbonylation reaction lies 164–175 kJ/mol above the GS and leads to the formation of complexes of  $CH_2NH_2^+$ , GG, and CO (eight variants were found), which lie 62–95 kJ/mol below the TS and 77–102 kJ/mol above the GS, Supplementary Tables S4 and S5. In addition to this pathway, we found another two pathways with similar rate-limiting TSs that lie 2–7 kJ/mol higher in energy, Supplementary Table S5. One of these reproduces the pathway and rate-limiting TS of PS,  $TS[N^2]-cggttttt\{N^1C-CO^1-N^2C\}$  (**Y2\_A1Y1\_TS**), and the third pathway occurs via rate-limiting  $TS[N^2]-ctgttttt\{N^1C-CO^1-N^2C\}$ . These three TSs differ from each other only in the spatial arrangement of the incipient immonium ion; however,  $TS[N^2]-t_tgttttt\{N^1C-CO^1-N^2C\}$  is less constrained than the other two TSs, which means it should dominate the observed reaction as it is entropically and slightly enthalpically favored. The TSs presented here are analogous forms of the  $TS[N^2]-cggttt\{C^{\alpha 1}-OC-N^2\}$  obtained for loss of CO from  $H^+GG$  found by Armentrout and Heaton [39] and from  $H^+AGG$  by Bythell et al. [49]. Additionally, we tried to locate a pathway that does not involve sequential losses of CO and  $CH_2NH$  or GG to form  $[y_2 + 2H]^+$  and  $[a_1]^+$  ions, but were unable to do so.

Once over one of the rate-limiting  $TS[N^2]\{N^1C-CO^1-N^2C\}$ , several complexes of the  $CH_2NH_2^+$  ion bound to diglycine and CO can be formed. We located nine of these, including one reported by PS, as detailed in the SI. Similarly, once CO is lost, the  $(CH_2NH)H^+(GG)$  complexes can explore different conformations (six were located, including two found by PS, as detailed in the SI). The various  $CH_2NH(H^+)GG + CO$  isomers are well below the three  $TS[N^2]\{N^1C-O^1C-CN^2\}$  rate-limiting TSs, Supplementary Tables S4 and S5. Any of these  $CH_2NH(H^+)GG$  species can dissociate to form either the  $[y_2 + 2H]^+ (H^+GG) + CH_2NH$  or  $[a_1]^+ (CH_2NH_2^+) + GG$  products at higher energies. These subsequent dissociations are limited by the separated product energies, consistent with the large cross-sections observed for these channels, Figure 1. Formation of  $[y_2 + 2H]^+ + CH_2NH + CO$  is calculated to lie 1–36 kJ/mol above the rate-limiting TSs, and  $[a_1]^+ + GG + CO$  lies 62–94 kJ/mol above the rate-limiting TSs (52–67 kJ/mol above the competing  $[y_2 + 2H]^+$  channel). Dissociation to yield  $[y_2 + 2H]^+$  and  $[a_1]^+$  ions is a competitive process because the proton is shared between the two molecules at their favored protonation sites.

PS also examined the pathway yielding a three-membered ring aziridinone product,  $c-CH_2NHCO$  [19], originally postulated by Cordero et al. [67] and suggested for  $H^+GGG$  by KK [14] and explored for dipeptides by PS [5]. This pathway was found to be limited by TSs that exceed those outlined here by

~40 kJ/mol, in agreement with the conclusions of PS. Thus the aziridinone pathway was not explored here.

### $[b_3]^+$ Formation (Loss of Water from $H^+GGG$ )

Mechanisms for the loss of water from  $H^+GGG$  have been previously explored by O'Hair et al. [68] and Reid et al. [15]. These studies demonstrated that water loss occurred from the C-terminus and resulted in an oxazolone product from  $H^+GGG-OH$ , although no atomistic details were provided. Other notable efforts examining water loss in peptides include pioneering work by Ballard and Gaskell with bigger peptides like bradykinin [69], leucine-enkephalin, and *des*-Arg<sup>1</sup>-substance P (PKPQQFFGLM-NH<sub>2</sub>) [70],  $H^+AGG$  [49], dipeptides of glycine and cysteine [71, 72], and bigger peptides of glycine [15, 73, 74]. Here we have reexamined the mechanism in detail in SI and Supplementary Figure S9, at higher levels of theory than previously utilized. Briefly, loss of water occurs here by transfer of a proton to the C-terminal hydroxyl group, O<sup>4</sup>H, and subsequent ring formation. This occurs via the tight TS[O<sup>1</sup>-O<sup>4</sup>]-ct(cg)t{C~O<sup>4</sup>H}, which lies 107–136 kJ/mol above the GS and forms a complex of protonated 2-glycylaminomethyl-5-oxazolone,  $H^+GAMOX[N^3]$  and water. This complex can form the final products by losing H<sub>2</sub>O via a loose TS, 107–116 kJ/mol above the GS and 4–22 kJ/mol below the TTS. Because this channel is limited by a TTS, it does not compete effectively with the pathways having loose TSs, explaining its small cross-section in Figure 1.

The TTS located here is similar to TS[O<sub>1</sub>-O<sub>3</sub>]-ctggt identified by Armentrout and Heaton for the formation of  $[b_2]^+$  by loss of water from  $H^+GG$  [39] and to the TS found for loss of water from  $H^+AGG$  [49]. The former study also located an alternative pathway involving TS[N<sub>1</sub>-O<sub>3</sub>]-ttgct, which can occur analogously for  $H^+GGG$  by transfer of a proton from the N<sup>2</sup> nitrogen to the terminal hydroxyl group via TS[N<sup>2</sup>-O<sup>4</sup>]- (tg)tg(tg){C~O<sup>4</sup>H}; however, this pathway exceeds TS[O<sup>1</sup>-O<sup>4</sup>]-ct(cg)t{C~O<sup>4</sup>H} by 78–97 kJ/mol (194–232 kJ/mol above the GS), Supplementary Table S5. The constrained head-to-tail cyclic geometry of this TS raises its energy and makes it entropically less favored as well. Thus, details of this route were not investigated further here. We also investigated the loss of water from [N<sup>1</sup>]-ttggtgt. This reaction pathway is limited by TS[N<sup>1</sup>-O<sup>4</sup>]- (tg)tg(tg){C~O<sup>4</sup>H} that lies 16–26 kJ/mol higher than TS[O<sup>1</sup>-O<sup>4</sup>]-ct(cg)t{C~O<sup>4</sup>H}, Supplementary Table S5 and, therefore, was not explored further.

Bythell et al. [73] used tandem mass spectrometry, CID, and IRMPD to study dehydration from protonated tetraglycine,  $H^+GGGG$ , and characterized the resultant product ion structure. They concluded that dehydration of  $H^+GGGG$  occurs by elimination of a non-C-terminal oxygen atom resulting in an oxazole ring near the N-terminus and thus does not correspond to the  $[b_4]^+$  ion, which would otherwise require elimination from the C-terminal free acid end. Later, Verkerk et al. [75] re-examined the dehydration product of  $H^+GGGG$  and reported a lower-energy protonated imidazolone product formed by the proton transfer from N-terminal amide oxygen. A similar reaction mechanism was also suggested by Spezia et al. for loss of

water from  $H^+GGGNH_2$  [24], which has no hydroxyl group. Comparison of the experimental IRMPD spectrum with that calculated for the lowest energy water loss product led the authors to conclude that water loss occurs from the N-terminal residue. The possibility of such mechanisms was explored here but we were unable to find pathways lower in energy than the one in which water loss occurs from the C-terminus. Details of these alternative pathways can be found in SI. Notably, Spezia et al. [24] reported the formation of a  $[b_3]^+$  ion with an oxazolone structure similar to the present study, but this product is formed by the loss of ammonia from  $H^+GGGNH_2$ .

In related experiments and computations on the  $[b_3]^+$  ion generated from a protonated pentapeptide,  $H^+AAAAR$ , Allen et al. also found the oxazolone structure to be the lowest energy structure [57], which agrees well with the present results. These authors also located a nine-membered ring structure for their  $[b_3]^+$  ion, which lies 90 kJ/mol higher than the oxazolone structure. Similarly, we located a nine-membered ring structure, 2-hydroxy-5,8-dioxo-1,4,7-triazonane<sup>+</sup>[O] for our  $[b_3]^+$  ion, which is 69–85 kJ/mol higher in energy than  $H^+GAMOX[N^3]$ , Supplementary Figure S9 and Supplementary Table S4. Because a lower energy pathway to form  $H^+GAMOX[N^3]$  was already found, we did not explore the pathway for the much higher energy cyclic  $[b_3]^+$  product ion.

### Further Dissociation of $[b_3]^+$ to $[a_3]^+$

Decarbonylation of  $[b_3]^+$  to  $[a_3]^+$  is rarely observed [57, 58, 76], but is found here and for larger peptides [48, 51]. The lowest energy pathways for this process in the presence and absence of the water product are shown in Supplementary Figure S10. Although the former is lower in energy, the rate-limiting tight TS( $H^+GAMOX[N^3]$ -ctg(tg){C~CO<sup>3</sup>})(H<sub>2</sub>O<sub>HN3,CO1-HO</sub>) along the pathway that includes water lies well above the loose TS associated with the  $[b_3]^+ + H_2O$  products. Thus, the  $H^+GAMOX(H_2O)$  intermediate should lose water rapidly, as demonstrated by calculations of the rate coefficients for this competition where the probability of reaching the TTS is four orders of magnitude less probable than water loss. Thus, decomposition of  $[b_3]^+$  to  $[a_3]^+$  probably occurs after water loss. Overall the formation of the  $[a_3]^+$  (GGNHCH<sub>2</sub><sup>+</sup>[N<sup>3</sup>]-ctggt) + H<sub>2</sub>O + CO products is limited by the TTS, TS( $H^+GAMOX[N^3]$ -ctg(tg){C~CO<sup>3</sup>}), which is 223–256 kJ/mol above the GS. Here, the small magnitude of the  $[a_3]^+$  cross-section in Figure 1 is determined primarily by the cross-section magnitude of its  $[b_3]^+$  precursor.

In previous work, Allen et al. studied decarbonylation of the  $[b_3]^+$  ion from  $H^+AAAAR$  [57] and explored the mechanism for this process in the absence of water. They found that the most stable  $[a_3]^+$  structure is a protonated linear imine structure formed by ring cleavage of  $[b_3]^+$  and is limited by a TTS. Both of these findings are similar to the results of the present study.

In an IRMPD study, Bythell, Maître, and Paizs (BMP) [51] compared the calculated IR spectra of the  $[a_3]^+$  structures with their experimental spectrum for  $[a_3]^+$  (from  $H^+GGGG$ ) and found that neither of the acyclic  $[a_3]^+$  structures gave a good



match. Rather, a seven-membered ring structure (**L** by BMP), calculated to be  $\sim 23$  kJ/mol below  $\text{GGNHCH}_2^+[\text{N}^3]\text{-ctgg}$  (**A**), agreed well with the IRMPD spectrum. Consequently BMP concluded that the majority of the  $[\text{a}_3]^+$  population is composed of the cyclic species (**L**) under the experimental conditions, which is kinetically the “most favored product” starting from  $\text{GGNHCH}_2^+[\text{N}^3]\text{-ctgg}$  (**A**) and formed by passing over a TTS 44 kJ/mol higher than **A**. This result motivated us to investigate cyclization of the  $[\text{a}_3]^+$  ion formed from  $\text{H}^+\text{GGG}$ , which potentially can occur before or after losing water, although, as noted above, the water is probably lost upon formation of the  $[\text{b}_3]^+$  precursor (Supplementary Figure S10). Details of these pathways and the different  $[\text{a}_3]^+$  structures can be found in Supplementary Figures S11 and S12 of the SI. Overall, the rate-limiting TS for cyclization of  $[\text{a}_3]^+$  in the absence of water,  $\text{TS}(\text{GGNHCH}_2^+[\text{N}^3]\text{-ctgg}(\text{tg}))$ , lies 3–13 kJ/mol above the rate-limiting  $\text{TS}(\text{H}^+\text{GAMOX}[\text{N}^3]\text{-ctg}(\text{cg})\{\text{C}\sim\text{CO}^3\})$  for decarbonylation of the  $[\text{b}_3]^+$  ion, Supplementary Table S5. Thus, the acyclic  $[\text{a}_3]^+$  species  $\text{GGNHCH}_2^+[\text{N}^3]\text{-ctgg}$  should be formed at the experimental threshold, but cyclization can occur with little additional energy needed (although probably entropically disfavored). These theoretical results are not inconsistent with the observation of BMP that their  $[\text{a}_3]^+$  ion was dominated by the cyclic structure for two possible reasons: (1) this  $[\text{a}_3]^+$  ion was formed by decomposition of  $\text{H}^+\text{GGGG}$  (rather than  $\text{H}^+\text{GGG}$ ), and (2), the source conditions needed to yield adequate amounts of  $[\text{a}_3]^+$  for spectroscopic study could provide additional energy needed for the cyclization.

### Loss of CO and NH<sub>3</sub>

The lowest energy pathway for the loss of CO and NH<sub>3</sub> from  $\text{H}^+\text{GGG}$  begins after the loss of CO shown in Supplementary Figure S8 and parallels the mechanism elucidated previously for  $\text{H}^+\text{GG}$  [39]. The path in Supplementary Figure S13 starts at the  $(\text{CH}_2\text{NH}_2^+)(\text{GG}_{\text{NH}\cdot\text{OC}_{\text{NH}\cdot\text{N}}})$  complex, but any low-energy variant of the  $(\text{CH}_2\text{NH})\text{H}^+(\text{GG})$  complex can be used to initiate the loss of ammonia. A key finding of this pathway is covalent bond formation between the carbon of the immonium ion and the terminal nitrogen of GG. Depending on the level of theory, the products formed,  $\text{CH}_2\text{GG}^+[\text{N}^2]\text{-ttttt} + \text{NH}_3 + \text{CO}$ , lie above the rate-limiting TS for the initial CO loss (by  $\sim 10$  kJ/mol at B3LYP-D3, Supplementary Figure S13, and 8 kJ/mol at M06-2X), and below this rate-limiting step by  $\sim 17$  kJ/mol at the B3LYP and MP2 levels of theory, Supplementary Tables S4 and S5. Importantly, the products of this pathway lie 24–28 kJ/mol below the  $[\text{y}_2 + 2\text{H}]^+ + \text{CH}_2\text{NH} + \text{CO}$  products, and thus are energetically favored although entropically disfavored because the latter products require no rearrangement from the starting complex, Supplementary Figure S8. As will be shown in the analysis of the data below, results from B3LYP and MP2 levels of theory appear to more reliably reproduce the experimental results.

We also examined another pathway involving the direct transfer of a proton from  $\text{N}^2$  to  $\text{N}^1$  but found its rate-limiting

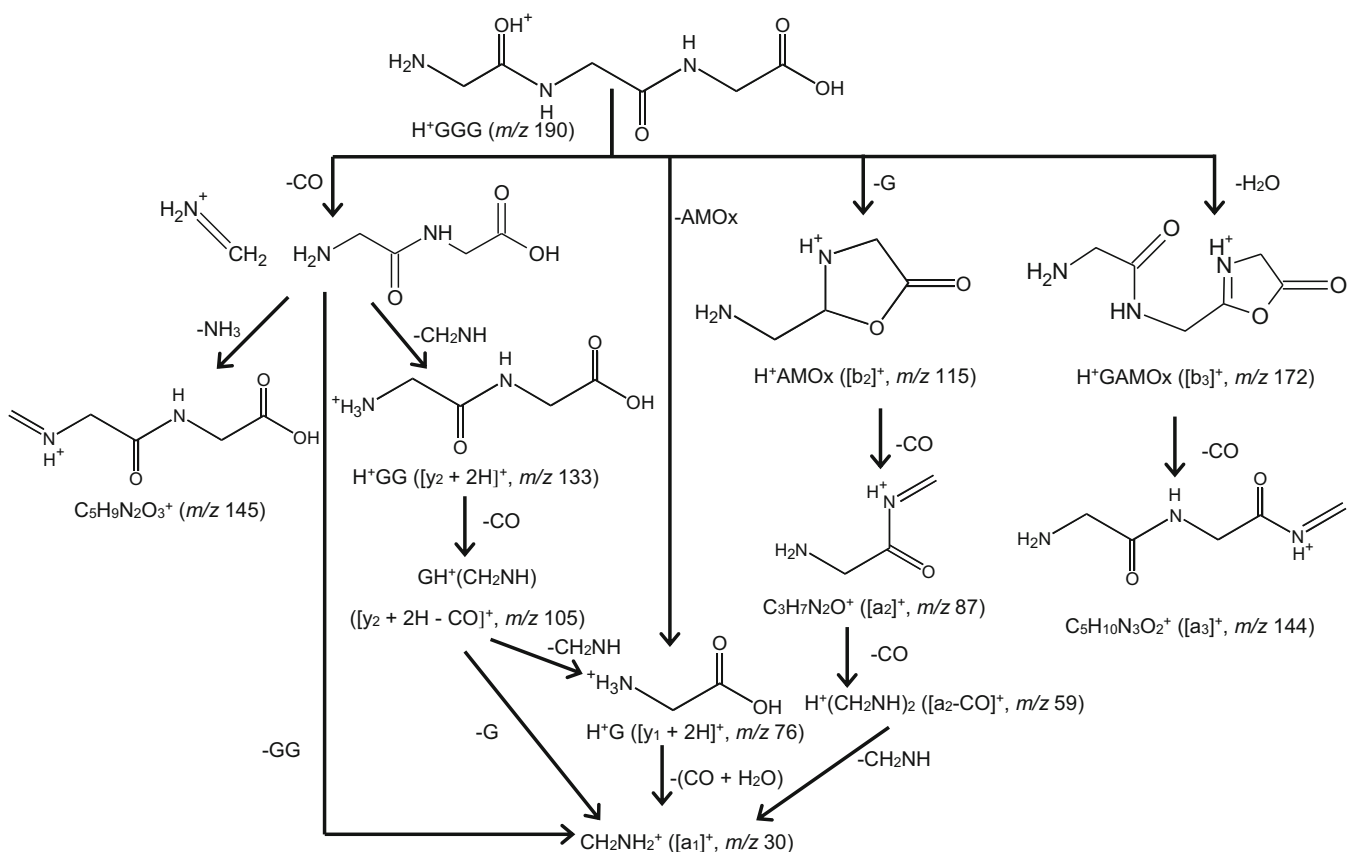
TS,  $\text{TS}(\text{H}^+\text{AMGG}[\text{N}^2\text{-N}^1]\text{-}(\text{tg})\text{gtttt})$  (Supplementary Figure S13), to be higher in energy, Supplementary Table S5. These two pathways are analogous to the low-energy mechanisms found by Armentrout and Heaton for loss of CO and NH<sub>3</sub> from  $\text{H}^+\text{GG}$ , although the former pathway is the rate-limiting TS for the loss of CO + NH<sub>3</sub> from  $\text{H}^+\text{GG}$  [39]. Bythell et al. [49] discovered a similar low-energy pathway for CO + NH<sub>3</sub> loss from  $\text{H}^+\text{AGG}$  involving proton transfer from O<sup>2</sup> to N<sup>1</sup> and limited by the loss of CO (161 kJ/mol above GS  $\text{H}^+\text{AGG}$ ) at the B3LYP/6-31+G(d,p)//B3LYP/6-31G(d) level of theory.

An alternative pathway for the loss of CO + NH<sub>3</sub> from  $\text{H}^+\text{GGG}$  involves nucleophilic displacement ( $\text{S}_{\text{N}}2$  mechanism) of ammonia by the  $\text{CH}_2\text{NH}$  moiety in the  $\text{CH}_2\text{NH}(\text{H}^+\text{GG})$  complex formed after expulsion of CO. This mechanism results in rate-limiting  $\text{TS}[\text{N}^2]\text{-ttttt}\{\text{CC}\sim\text{N}^2\text{H}\}$ , which is 385–433 kJ/mol above the ground reactants, Supplementary Table S5. Further details of this pathway were not explored because Supplementary Figure S13 provides much lower energy pathways for deamidation following the initial decarbonylation. Similar results were also obtained by Bythell et al. for loss of CO + NH<sub>3</sub> from  $\text{H}^+\text{AGG}$  [49] and by Armentrout and Heaton for the  $\text{H}^+\text{GG}$  [39].

### Analysis of Primary Dissociation Channels

Scheme 1 outlines the relationships between reactions 1–10 on the basis of the reaction mechanisms calculated above and previous work [39, 40]. The primary dissociations from  $\text{H}^+\text{GGG}$  are the formation of  $[\text{b}_2]^+$  in reaction 1a,  $[\text{y}_1 + 2\text{H}]^+$  in reaction 2a, loss of water to form  $[\text{b}_3]^+$  in reaction 3, and loss of CO, which rapidly leads to formation of  $[\text{y}_2 + 2\text{H}]^+$  in reaction 5,  $[\text{a}_1]^+$  in reaction 7a, and loss of NH<sub>3</sub> in reaction 4. The rate-limiting TSs for  $[\text{b}_2]^+$ ,  $[\text{y}_1 + 2\text{H}]^+$ ,  $[\text{y}_2 + 2\text{H}]^+$ , and  $[\text{a}_1]^+$  correspond to the asymptotic energies of their respective products (loose TSs), whereas TTSs limit the loss of water forming  $[\text{b}_3]^+$ ,  $\text{TS}[\text{O}^{\text{t}}\text{-O}^{\text{t}}]\text{-ct}(\text{cg})\text{t}\{\text{C}\sim\text{O}^{\text{t}}\text{H}\}$ , and loss of (CO + NH<sub>3</sub>),  $\text{TS}[\text{N}^2]\text{-t}_4\text{gtttt}\{\text{N}^1\text{C}\sim\text{CO}^1\sim\text{N}^2\text{C}\}$ , as detailed above. As noted, these TS characterizations are consistent with the relative magnitudes observed in Figure 1, namely, TTSs generally lead to small cross-sections and loose TSs lead to larger ones. Direct competition between  $[\text{b}_2]^+$  and  $[\text{y}_1 + 2\text{H}]^+$ , where the former is clearly energetically favored, explains why the cross-section of the latter product is inhibited. Note that the primary product for CO loss ( $m/z$  162) was not observed, which, given the mechanisms elucidated above, implies that this product decomposes rapidly when the thresholds for its subsequent decompositions are reached, primarily formation of  $[\text{y}_2 + 2\text{H}]^+$  and ammonia loss, and  $[\text{a}_1]^+$  at somewhat higher energies (all formed via loose PSL TSs according to theory). Because loss of CO over a TTS competes with the loose TS for the dominant  $[\text{b}_2]^+$  production at a similar energy, its cross-section apparently never reaches an appreciable level until after it can begin to dissociate further.

These complex inter-related processes are analyzed by first summing the cross-sections of the major sequential channels



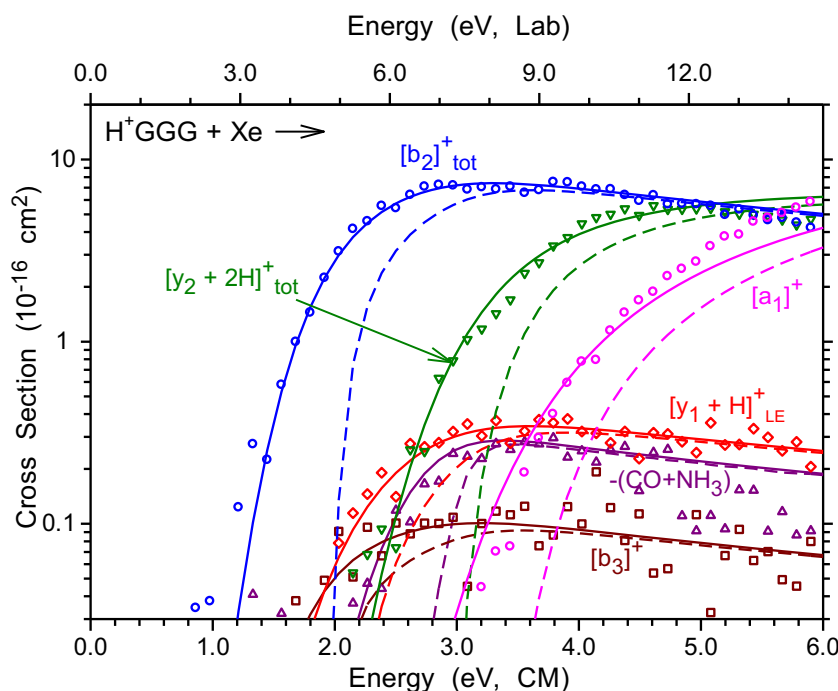
Scheme 1. Schematic reaction mechanism for  $\text{H}^+\text{GGG}$  decomposition

with their precursor channels to form composite channels:  $m/z$  115 +  $m/z$  87 ( $[\text{b}_2]^+ + [\text{a}_2]^+ = [\text{b}_2]_{\text{tot}}^+$ ), and  $m/z$  133 +  $m/z$  105 + high-energy (HE) portion (beyond 4 eV) of  $m/z$  76 ( $[\text{y}_2 + 2\text{H}]^+ + ([\text{y}_2 + 2\text{H} - \text{CO}]^+) + [\text{y}_1 + 2\text{H}]_{\text{HE}}^+ = [\text{y}_2 + 2\text{H}]_{\text{tot}}^+$ ). The cross-section for  $[\text{y}_1 + 2\text{H}]_{\text{HE}}^+$  is isolated by subtracting the model for the low-energy (LE) portion from the  $[\text{y}_1 + 2\text{H}]_{\text{tot}}^+$  product cross-section. Channels  $[\text{b}_3]^+$  and  $[\text{a}_3]^+$  should also be summed, but the latter product is sufficiently small that it was not observed at lower pressures (as discussed further below). The cross-sections of the composite channels change smoothly with increasing energy, which is consistent with the assigned sequential dissociations, Figure 3. The cross-sections of the six primary channels,  $[\text{b}_2]_{\text{tot}}^+$ ,  $[\text{y}_1 + 2\text{H}]_{\text{LE}}^+$ ,  $[\text{b}_3]^+$ , loss of CO +  $\text{NH}_3$ ,  $[\text{y}_2 + 2\text{H}]_{\text{tot}}^+$ , and  $[\text{a}_1]^+$  are modeled competitively using Equation S1, as shown in Figure 3 with the optimized parameters listed in Table 1. This model provides the threshold energies for reactions 1a ( $[\text{b}_2]^+$ ), 2a ( $[\text{y}_1 + 2\text{H}]_{\text{LE}}^+$ ), 3 ( $[\text{b}_3]^+$ ), 4 (loss of CO +  $\text{NH}_3$ ), 5 ( $[\text{y}_2 + 2\text{H}]_{\text{tot}}^+$ ), and 7a ( $[\text{a}_1]^+$ ). Cross-sections for  $[\text{b}_2]_{\text{tot}}^+$ ,  $[\text{y}_1 + 2\text{H}]_{\text{LE}}^+$  and  $[\text{y}_2 + 2\text{H}]_{\text{tot}}^+$  can be modeled without scaling their cross-sections (i.e., using common  $\sigma_{0,j}$  values) or molecular parameters, suggesting that the statistical factors associated with these channels are sufficient to accurately describe their relative cross-section magnitudes over an extended energy range.

It is important to point out that the modeling of the ammonia loss,  $[\text{y}_2 + 2\text{H}]^+$ , and  $[\text{a}_1]^+$  cross-sections is not

straightforward because theory shows that all three channels result from sequential decompositions from a product ion formed by CO loss. The challenge in modeling these three cross-sections as sequential decompositions is that the cross-section for the primary product ion (CO loss) was not observed. Our search for the best model for reproducing these three cross-sections led us to analyze these product cross-sections in three ways. These models are named after the TSs for  $\text{NH}_3$  loss,  $[\text{y}_2 + 2\text{H}]^+$  formation, and  $[\text{a}_1]^+$  formation with subscripts designating their treatment as primary (P) or secondary (S) dissociations, namely  $\text{T}_\text{P}\text{L}_\text{P}\text{L}_\text{P}$  (tight-loose-loose),  $\text{T}_\text{P}\text{L}_\text{S}\text{L}_\text{S}$  (tight-loose-loose), and  $\text{L}_\text{S}\text{L}_\text{S}\text{L}_\text{S}$  (loose-loose-loose). In the latter model, the TS for initial CO loss is assumed to lie below the  $\text{NH}_3$  loss products enough that its molecular parameters are inconsequential. Details of the analysis of the  $\text{T}_\text{P}\text{L}_\text{S}\text{L}_\text{S}$  (which requires extensive frequency scaling) and  $\text{L}_\text{S}\text{L}_\text{S}\text{L}_\text{S}$  (which does not reproduce the data) models can be found in the SI, because that discussion indicates that the best analysis of these three products is obtained by the  $\text{T}_\text{P}\text{L}_\text{P}\text{L}_\text{P}$  approach.

In the  $\text{T}_\text{P}\text{L}_\text{P}\text{L}_\text{P}$  model, shown in Figure 3, all three product channels are modeled as primary decompositions from  $\text{H}^+\text{GGG}$ . Here, ammonia loss is analyzed as a primary product limited by the TTS for CO loss, which explains why the CO loss product is not observed. This is consistent with theory at



**Figure 3.** Competitive modeling of the six primary products: composite cross-sections of  $[b_2]^+_{\text{tot}}$ ,  $[y_2 + 2H]^+_{\text{tot}}$ ,  $[y_1 + 2H]^+_{\text{LE}}$ ,  $[b_3]^+$ , loss of  $(\text{CO} + \text{NH}_3)$ , and  $[a_1]^+$  from  $\text{H}^+\text{GGG}$ . Symbols show zero-pressure extrapolated data for the indicated process between  $\text{H}^+\text{GGG}$  and  $\text{Xe}$  as a function of collision energy in the center-of-mass frame (lower x-axis) and in the laboratory frame (upper x-axis). Solid lines show the best fit to the data using the model in Equation S1 convoluted over the neutral and ion kinetic and internal energy distributions. Dashed lines show the model cross-sections in the absence of experimental kinetic energy broadening for reactants with internal energies at 0 K

the B3LYP and MP2 levels. Alternative TSs for CO loss are energetically similar to  $\text{TS}[\text{N}^2]\text{-t,tgttttt}\{\text{N}^1\text{C}\sim\text{CO}^1\sim\text{N}^2\text{C}\}$  that was used to model reaction 4. Using different TSs in the analysis does not change the thresholds for the primary channels significantly; thus it is possible that the experimental threshold is a weighted average of the various TSs having similar energies leading to the same CO loss products.

The loose TSs for  $[y_2 + 2H]^+$  and  $[a_1]^+$  products involve dissociation to three molecular species (for which there are no appropriate statistical approaches yet developed). Therefore, when treated as “primary” dissociation channels, we use complexes of  $(\text{CO} + \text{GG})$  and  $(\text{CO} + \text{CH}_2\text{NH})$ , respectively, as the “neutral” that is formed along with the primary ion to approximate the molecular parameters of the neutral products,  $\text{CO} + \text{GG}$  and  $\text{CO} + \text{CH}_2\text{NH}$ . Geometry optimizations were performed on different conformers of the putative  $(\text{CO} + \text{GG})$  and  $(\text{CO} + \text{CH}_2\text{NH})$  complexes to locate the lowest energy conformer, which was used in the data analysis.  $(\text{GG}\text{-ctttt})_{(\text{OH})\text{OC}}$ , where the CO oxygen is H-bonded to the hydroxyl hydrogen of GG, was used to approximate the molecular parameters of  $\text{CO} + \text{GG}$ . Similarly,  $(\text{CH}_2\text{NH})_{(\text{N})\text{CO}}$ , where the CO carbon is loosely coordinated to the  $\text{CH}_2\text{NH}$  nitrogen, was used to approximate the molecular parameters of  $\text{CO} + \text{CH}_2\text{NH}$ . The  $\text{T}_p\text{L}_p\text{L}_p$  model fits the  $[a_1]^+$  cross-section well only up to  $\sim 4.5$  eV without the use of any cross-section scaling factor. The  $[a_1]^+$  cross-section can be

accurately reproduced over the entire energy range shown if the PSL TS frequencies below  $900\text{ cm}^{-1}$  are loosened (multiplying by  $0.813 \pm 0.004$ ) and the threshold is shifted up by  $0.19 \pm 0.09$  eV, Table 1. The latter shift results because the kinetic shift is larger for the loose TS frequencies (compare  $\Delta S_{1000}^{\ddagger}$  value of 243 with 207 J/mol K when TS frequencies were not scaled). The need for a scaling factor to reproduce the  $[a_1]^+$  cross-section over an extended range of energy indicates that  $[a_1]^+$  is formed via the PSL TS at threshold energies and by different pathways at higher energies. We believe that the  $[a_1]^+$  threshold energy obtained without the use of a scaling factor is our best threshold energy value for this product because it is not biased by the contribution of additional channels.

In order to accurately reproduce reaction 3 ( $[b_3]^+$ ) in the competitive modeling as shown in Figure 3, the TTS frequencies below  $900\text{ cm}^{-1}$  needed to be scaled by  $0.786 \pm 0.003$ . The need for this modest scaling of the vibrational frequencies indicates that the TTS is looser than suggested by theory. We also modeled the  $[b_3]^+$  cross-section using a switching TS in which the rate is determined by a low-energy TTS for  $\text{H}_2\text{O}$  loss coupled with a higher energy PSL TS. This model and its fitting parameters can be found in the SI. Here the PSL TS energy could be as high as 1.64 eV without affecting the reproduction of the  $[b_3]^+$  and other primary product cross-sections. At higher PSL TS energies, the model misses the threshold region of the  $[b_3]^+$  cross-section, Supplementary

**Table 1.** Fitting Parameters of Equations S1, S3a, and S3b, Threshold Energies at 0 K, and Entropies of Activation at 1000 K for the Decomposition of H<sup>+</sup>GGG

Reaction	TS <sup>a</sup>	Competition <sup>b</sup>	$\sigma_0$	n	E <sub>0</sub> (eV)	$\Delta S_{1000}^{\ddagger}$ (J/K mol)
1a [b <sub>2</sub> ] <sup>+</sup> (H <sup>+</sup> AMOX) + G	PSL	none	20.1 ± 2.0	1.1 ± 0.1	1.76 ± 0.06 <sup>c</sup>	92 ± 5
		[y <sub>1</sub> + 2H] <sup>+</sup> , [y <sub>2</sub> + 2H] <sup>+</sup> , [a <sub>1</sub> ] <sup>+</sup> , [b <sub>3</sub> ] <sup>+</sup> , CO loss	20.4 ± 1.7	1.0 ± 0.1	1.77 ± 0.06	92 ± 5
		[y <sub>1</sub> + 2H] <sup>+</sup> , [y <sub>2</sub> + 2H] <sup>+</sup> , [a <sub>1</sub> ] <sup>+</sup> , [b <sub>3</sub> ] <sup>+</sup> , CO loss, + seq [a <sub>2</sub> ] <sup>+</sup>	20.3 ± 1.2	1.0 ± 0.1	1.76 ± 0.06	92 ± 5
		[y <sub>1</sub> + 2H] <sup>+</sup> , [y <sub>2</sub> + 2H] <sup>+</sup> , [a <sub>1</sub> ] <sup>+</sup> + seq [y <sub>2</sub> + 2H - CO] <sup>+</sup>	20.6 ± 1.4	1.0 ± 0.1	1.78 ± 0.06	92 ± 5
2a [y <sub>1</sub> + 2H] <sup>+</sup> (H <sup>+</sup> G) + AMOX	PSL	[b <sub>2</sub> ] <sup>+</sup> , [y <sub>2</sub> + 2H] <sup>+</sup> , [a <sub>1</sub> ] <sup>+</sup> , [b <sub>3</sub> ] <sup>+</sup> , CO loss	20.4 ± 1.7	1.0 ± 0.1	1.90 ± 0.07	81 ± 5
		[b <sub>2</sub> ] <sup>+</sup> , [y <sub>2</sub> + 2H] <sup>+</sup> , [a <sub>1</sub> ] <sup>+</sup> , [b <sub>3</sub> ] <sup>+</sup> , CO loss, + seq [a <sub>2</sub> ] <sup>+</sup>	21.1 ± 1.6	1.0 ± 0.1	1.90 ± 0.06	81 ± 5
		[b <sub>2</sub> ] <sup>+</sup> , [y <sub>2</sub> + 2H] <sup>+</sup> , [a <sub>1</sub> ] <sup>+</sup> + seq [y <sub>2</sub> + 2H - CO] <sup>+</sup>	20.6 ± 1.4	1.0 ± 0.1	1.91 ± 0.06	81 ± 5
		none	10.7 ± 1.0	1.5 ± 0.1	2.64 ± 0.06 <sup>c</sup>	175 ± 5
5 [y <sub>2</sub> + 2H] <sup>+</sup> (H <sup>+</sup> GG) + CO + CH <sub>2</sub> NH	PSL	[b <sub>2</sub> ] <sup>+</sup> , [y <sub>1</sub> + 2H], [a <sub>1</sub> ] <sup>+</sup> , [b <sub>3</sub> ] <sup>+</sup> , CO loss	20.4 ± 1.7	1.0 ± 0.1	2.28 ± 0.06	176 ± 5
		[b <sub>2</sub> ] <sup>+</sup> , [y <sub>1</sub> + 2H] <sup>+</sup> , [a <sub>1</sub> ] <sup>+</sup> , [b <sub>3</sub> ] <sup>+</sup> , CO loss + seq [a <sub>2</sub> ] <sup>+</sup>	21.1 ± 1.6	1.0 ± 0.1	2.29 ± 0.05	176 ± 5
		[b <sub>2</sub> ] <sup>+</sup> , [y <sub>1</sub> + 2H] <sup>+</sup> , [a <sub>1</sub> ] <sup>+</sup> + seq [y <sub>2</sub> + 2H - CO] <sup>+</sup>	20.6 ± 1.4	1.0 ± 0.1	2.28 ± 0.06	176 ± 5
		[b <sub>2</sub> ] <sup>+</sup> , [y <sub>1</sub> + 2H] <sup>+</sup> , [y <sub>2</sub> + 2H] <sup>+</sup> , [b <sub>3</sub> ] <sup>+</sup> , CO loss	20.3 ± 1.7	1.0 ± 0.1	2.75 ± 0.06 <sup>d</sup>	243 ± 5
7a [a <sub>1</sub> ] <sup>+</sup> (CH <sub>2</sub> NH <sub>2</sub> <sup>+</sup> ) CO + GG	PSL	[b <sub>2</sub> ] <sup>+</sup> , [y <sub>1</sub> + 2H] <sup>+</sup> , [y <sub>2</sub> + 2H] <sup>+</sup> , [b <sub>3</sub> ] <sup>+</sup> , CO loss	20.2 ± 2.9	1.0 ± 0.1	2.56 ± 0.07 <sup>c</sup>	207 ± 5
		[b <sub>2</sub> ] <sup>+</sup> , [y <sub>1</sub> + 2H] <sup>+</sup> , [y <sub>2</sub> + 2H] <sup>+</sup> , [b <sub>3</sub> ] <sup>+</sup> , CO loss	21.1 ± 1.7	1.0 ± 0.1	2.76 ± 0.05	243 ± 5
		[b <sub>2</sub> ] <sup>+</sup> , [y <sub>1</sub> + 2H] <sup>+</sup> , [y <sub>2</sub> + 2H] <sup>+</sup> , [b <sub>3</sub> ] <sup>+</sup> , CO loss + seq [a <sub>2</sub> ] <sup>+</sup>	20.5 ± 1.5	1.0 ± 0.1	2.76 ± 0.06	243 ± 5
		[b <sub>2</sub> ] <sup>+</sup> , [y <sub>1</sub> + 2H] <sup>+</sup> , [y <sub>2</sub> + 2H] <sup>+</sup> + seq [y <sub>2</sub> + 2H - CO] <sup>+</sup>	20.1 ± 1.7	1.0 ± 0.1	1.51 ± 0.07 <sup>e,g</sup>	64 ± 1
3 [b <sub>3</sub> ] <sup>+</sup> (H <sup>+</sup> GAMOX) + H <sub>2</sub> O	TS[O <sup>1t</sup> -O <sup>4</sup> ]	[b <sub>2</sub> ] <sup>+</sup> , [y <sub>1</sub> + 2H] <sup>+</sup> , [y <sub>2</sub> + 2H] <sup>+</sup> , [a <sub>1</sub> ] <sup>+</sup> , CO loss	20.4 ± 1.3	1.0 ± 0.1	1.52 ± 0.05	65 ± 9
		[b <sub>2</sub> ] <sup>+</sup> , [y <sub>1</sub> + 2H] <sup>+</sup> , [y <sub>2</sub> + 2H] <sup>+</sup> , [a <sub>1</sub> ] <sup>+</sup> , CO loss + seq [a <sub>2</sub> ] <sup>+</sup>	20.0 ± 1.7	1.0 ± 0.1	1.70 ± 0.06	107 ± 1
4 (CO+NH <sub>3</sub> ) loss	TS[N <sup>2</sup> ]	[b <sub>2</sub> ] <sup>+</sup> , [y <sub>1</sub> + 2H] <sup>+</sup> , [y <sub>2</sub> + 2H] <sup>+</sup> , [a <sub>1</sub> ] <sup>+</sup> , [b <sub>3</sub> ] <sup>+</sup>	20.9 ± 1.5	1.0 ± 0.1	1.70 ± 0.05	106 ± 2
		[b <sub>2</sub> ] <sup>+</sup> , [y <sub>1</sub> + 2H] <sup>+</sup> , [y <sub>2</sub> + 2H] <sup>+</sup> , [a <sub>1</sub> ] <sup>+</sup> , [b <sub>3</sub> ] <sup>+</sup> + seq [a <sub>2</sub> ] <sup>+</sup>	15.3 ± 1.1	1.1 ± 0.1	2.22 ± 0.15 <sup>i</sup>	32 ± 1
6 [a <sub>3</sub> ] <sup>+</sup> ([b <sub>3</sub> ] <sup>+</sup> - CO) <sup>h</sup>	TS[O <sup>1t</sup> -O <sup>4</sup> ] → TS(H <sup>+</sup> GAMOX)	[b <sub>2</sub> ] <sup>+</sup> , [y <sub>1</sub> + 2H] <sup>+</sup> , [y <sub>2</sub> + 2H] <sup>+</sup> , [a <sub>1</sub> ] <sup>+</sup> , CO loss, [b <sub>3</sub> ] <sup>+</sup>				

<sup>a</sup>PSL = phase space limit. TS[O<sup>1t</sup>-O<sup>4</sup>] = TS[O<sup>1t</sup>-O<sup>4</sup>]-ct(cg)t{C~O<sup>4</sup>H}, TS[N<sup>2</sup>] = TS([N<sup>2</sup>]-t<sub>g</sub>ttttt{N<sup>1</sup>C~CO<sup>1</sup>~N<sup>2</sup>C}), TS(H<sup>+</sup>GAMOX) = TS(H<sup>+</sup>GAMOX[N<sup>3</sup>])-ctg(cg){C~CO<sup>3</sup>}. Arrow indicates a sequential process with the indicated TSs for the initial and sequential steps

<sup>b</sup>Competition with listed ions is included in the modeling where + seq indicates a sequential reaction

<sup>c</sup>E<sub>0</sub> without RRKM = 2.04 ± 0.09 eV for b<sub>2</sub> and 3.05 ± 0.08 eV for [y<sub>2</sub> + 2H]<sup>+</sup>

<sup>d</sup>Value when PSL TS frequencies below 900 cm<sup>-1</sup> of [a<sub>1</sub>]<sup>+</sup> product in T<sub>p</sub>L<sub>p</sub>L<sub>p</sub> model are scaled by 0.813 ± 0.004

<sup>e</sup>Value when no scaling factor used for PSL TS frequencies of [a<sub>1</sub>]<sup>+</sup> product in T<sub>p</sub>L<sub>p</sub>L<sub>p</sub> model. See text for details

<sup>f</sup>Value when TS frequencies below 900 cm<sup>-1</sup> of [b<sub>3</sub>]<sup>+</sup> product are scaled by 0.786 ± 0.003

<sup>g</sup>E<sub>0</sub> = 1.31 ± 0.06 when [b<sub>3</sub>]<sup>+</sup> product is analyzed with switching TS, details in text

<sup>h</sup>Values from analysis of high-pressure data rather than zero pressure extrapolated results

<sup>i</sup>E<sub>0</sub> = 2.32 ± 0.12 eV when corrected to zero pressure conditions

Figure S16. In comparison to the TTS model, the switching TS model does not require scaling of the cross-section or TS frequencies, such that the kinetic shift is larger, thereby reducing the [b<sub>3</sub>]<sup>+</sup> threshold energy by 0.20 ± 0.08 eV, Table 1 and Supplementary Table S6, which now agrees better with theory (as discussed below). Nevertheless, the model obtained using the TTS model reproduces the [b<sub>3</sub>]<sup>+</sup> cross-section better than the switching TS model, such that our best determination of the [b<sub>3</sub>]<sup>+</sup> threshold energy is an average value of the threshold energies obtained from these two approaches, 1.41 ± 0.09 eV, Table 1. This uncertainty includes a contribution from analyzing the [b<sub>3</sub>]<sup>+</sup> channel without the [a<sub>3</sub>]<sup>+</sup> product added.

We also analyzed the water loss [b<sub>3</sub>]<sup>+</sup> product using the high energy alternative TS, TS[N<sup>1</sup>-O<sup>4</sup>]-t<sub>g</sub>tg{C~O<sup>4</sup>H} (16–26 kJ/mol above TS[O<sup>1t</sup>-O<sup>4</sup>]-ct(cg)t{C~O<sup>4</sup>H}), but this TS is entropically less favorable, such that the contribution from this TS should be minimal. Similarly, another high-energy TS, TS[N<sup>2</sup>-O<sup>4</sup>]-t<sub>g</sub>t<sub>g</sub>(t<sub>g</sub>){C~O<sup>4</sup>H}, is much higher in energy than TS[O<sup>1t</sup>-O<sup>4</sup>]-ct(cg)t{C~O<sup>4</sup>H} and also more constrained (discussed above), such that it will not contribute significantly to the [b<sub>3</sub>]<sup>+</sup> channel.

All of these competitive products were also analyzed assuming that the reaction begins at the MP2 GS of H<sup>+</sup>GGG, [N<sup>1</sup>]-t<sub>g</sub>tg<sub>g</sub>t<sub>g</sub>t<sub>g</sub>. The uncertainties in the threshold energies of all products examined here include deviations in the threshold

energies associated with use of molecular parameters for either the  $[O^1]$  or  $[N^1]$  conformer in the data analysis.

Overall, Figure 3 shows that the competition between all six primary products can be modeled nicely over extended ranges of energy ( $>5$  eV) and two and a half orders of magnitude. Note that in agreement with its large cross-section at high energies,  $[a_1]^+$  is entropically the most favored process among the six competitive channels as reflected in its large  $\Delta S_{1000}^{\ddagger}$  value, Table 1.

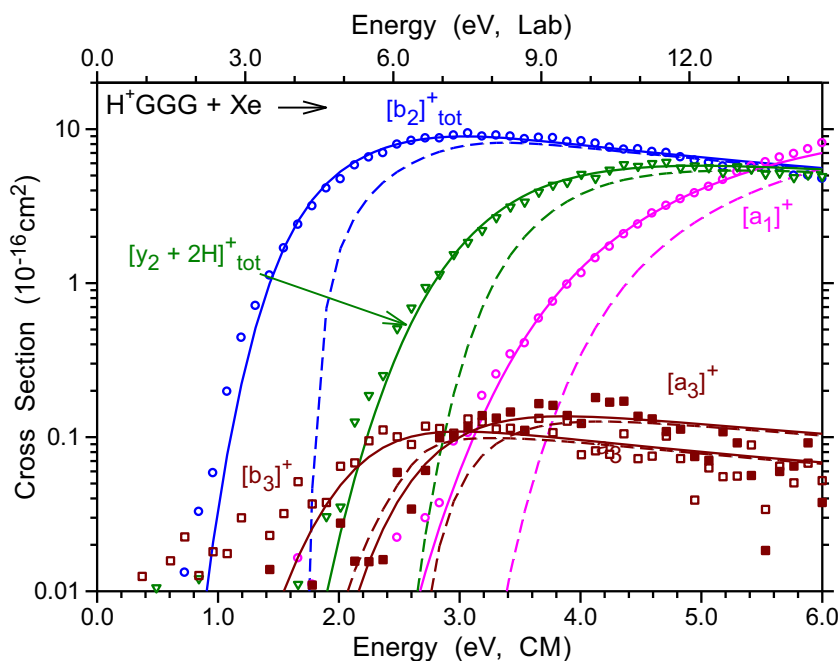
### Analysis of Secondary Dissociation Channels

Once the fitting parameters for the six primary channels are determined, sequential dissociation channels  $[a_3]^+$  (from  $[b_3]^+$ ),  $[a_2]^+$  (from  $[b_2]^+$ ), and  $[y_2 + 2H - CO]^+$  (from  $[y_2 + 2H]^+$ ) in reactions 6, 8a, and 9 are analyzed. In all of these analyses, threshold energies of the competing  $[b_2]^+$ ,  $[y_1 + 2H]^+$ ,  $[y_2 + 2H]^+$ , and  $[a_1]^+$  channels are the same as those obtained when the sequential channels are not included. Other secondary dissociations include reactions 1b ( $[y_2 + 2H]^+ - H_2O \rightarrow [b_2]^+$ ) and 7b ( $[y_1 + 2H]^+ - H_2O - CO \rightarrow [a_1]^+$ ) that are not included in the current analysis because they are both inefficient processes that will contribute little to either product channel making it difficult to determine their energetics. In both cases, comprehensive studies have been done previously in our lab to yield threshold energies for these processes [37, 38].

The dissociation of  $[b_3]^+$  ( $m/z$  172) to form  $[a_3]^+$  ( $m/z$  144) in reaction 6 occurs by the loss of CO. Here the precursor  $[b_3]^+$  ion has low intensity such that the intensity of the  $[a_3]^+$  product ion is equally weak, and is also obscured by the nearby  $m/z$  145 product ion, which is more intense. Because of these

difficulties, zero pressure extrapolated cross-sections for the  $[a_3]^+$  product ion were of insufficient quality to model with any confidence. Therefore, we analyze the  $[a_3]^+$  product cross-section acquired at a high pressure ( $\sim 0.2$  mTorr) of xenon in order to obtain accurate relative threshold information. Calculations discussed above indicate there are two pathways leading to  $[a_3]^+ + H_2O + CO$ , but the pathway associated with loss of water followed by decarbonylation of  $[b_3]^+$  is almost certainly operative, such that formation of  $[a_3]^+$  is limited by the tight TS( $H^+GAMOX[N^3]$ -ctg(cg)) $\{C\sim CO^3\}$ , Supplementary Figure S10. Thus, we analyzed the  $[a_3]^+$  product as a sequential loss of water followed by CO loss ( $[b_3]^+ \rightarrow [a_3]^+$ ) while still including the competition with the other five primary channels, Figure 4. Fitting parameters of the  $[a_3]^+$  product obtained from the analysis of high-pressure data are listed in Table 1. This analysis indicates that the threshold for reaction 6 is  $0.81 \pm 0.09$  eV above that for reaction 3. When this difference is added to the threshold energy for formation of  $[b_3]^+$  determined using a zero-pressure extrapolated data set, we obtain a threshold for  $[a_3]^+$  production of  $2.32 \pm 0.12$  eV.

The sequential dissociation of the primary  $[b_2]^+$  ion formed in reaction 1a to  $[a_2]^+$  in reaction 8a occurs via the tight TS( $H^+AMOX[N^2]$ -c $\{OC\sim O\}$ ) as previously calculated by Armentrout and Clark [40]. The subsequent dissociation of  $[b_2]^+$  to  $[a_2]^+$  is modeled while still including competition from the other five primary product ions, Supplementary Figure S17. Here, the  $[a_2]^+$  ion required a slight scaling of its cross-section, which differs from Armentrout and Clark's [40] analysis of  $[a_2]^+$  where no scaling was required. The sequential



**Figure 4.** Sequential modeling of the  $[a_3]^+$  ion formed by CO loss from  $[b_3]^+$  along with competitive modeling of the remaining primary ions. The  $[y_1 + 2H]^+_{LE}$  and loss of CO +  $NH_3$  products are not shown for purposes of clarity. Symbols show high-pressure data for the indicated process between  $H^+GGG$  and Xe as a function of collision energy in the center-of-mass frame (lower x-axis) and in the laboratory frame (upper x-axis). Solid lines show the best fit to the data using the models in Equations S1, S3a, and S3b convoluted over the neutral and ion kinetic and internal energy distributions. Dashed lines show the model cross-sections in the absence of experimental kinetic energy broadening for reactants with internal energies at 0 K

dissociation model reproduces the  $[a_2]^+$  cross-section up to  $\sim 4.5$  eV. Analysis of the  $[a_2]^+$  cross-section beyond 4.5 eV does not reproduce the threshold region and requires extensive scaling of its cross-section to capture its magnitude. Above 4.5 eV, the  $[a_2]^+$  ion can also be made from a  $[b_2]^+$  product ion that is generated by dissociation of the primary  $[y_2 + 2H]^+$  product ion in reaction 8b, as shown directly in previous studies of  $H^+GG$  decomposition [38]. These previous studies indicate that this pathway should start at  $\sim 4.5$  eV ( $\sim 2.5$  eV above the apparent  $[y_2 + 2H]^+$  threshold), but is probably not a major contributor to the  $[a_2]^+$  cross-section above 4.5 eV, although the energy distributions involved in the sequential pathways from  $H^+GGG$  might alter this conclusion. Nevertheless, analysis of the  $[a_2]^+$  data as a sequential dissociation from  $H^+GGG$  via  $[b_2]^+$  formed in reaction 1a beyond 4.5 eV should not be accurate and explains the need for significant scaling of  $[a_2]^+$  cross-section at these higher energies. Supplementary Table S6 lists the parameters for this sequential fit. We find the threshold energy of the  $[a_2]^+$  product to be  $3.39 \pm 0.06$  eV when formed by reaction 8a, which lies  $1.62 \pm 0.08$  eV above the threshold for  $[b_2]^+$  (whose average threshold energy is  $1.77 \pm 0.06$  eV, Table 1). The  $1.62 \pm 0.08$  eV value can be favorably compared with a threshold energy of  $1.57 \pm 0.12$  eV found for the same process in a previous study [40].

Armentrout and Heaton [38] also found that the  $[b_2]^+$  product ion can dissociate to  $[a_1]^+$  via  $[a_2]^+$ , reaction 7c. Here, we explored this possibility by summing the  $[a_2]^+$  and  $[a_1]^+$  cross-sections and fitting this composite channel as a sequential dissociation from  $[b_2]^+$  where formation of  $[a_2]^+$  is limited by the tight  $TS(H^+AMOX[N^2]-c\{OC\sim O\})$ . We also analyzed the  $[a_1]^+$  cross-section as a sequential dissociation from  $[b_2]^+$ , limited by a PSL TS. These analyses show that the threshold energy of  $[a_1]^+$  (from  $[b_2]^+$ ) exceeds that of  $[a_2]^+$  (from  $[b_2]^+$ ), thus indicating that  $[b_2]^+$  forms  $[a_2]^+$  first, which then dissociates to  $[a_1]^+$ , consistent with previous results [40]. However, the threshold energies obtained for  $[a_1]^+$  and  $[a_2]^+$  differ by only 0.20 eV, which is much smaller than the difference of 1.36 eV [40] obtained previously for dissociation of  $[b_2]^+$ . This comparison simply confirms that  $[b_2]^+$  contributes to the  $[a_1]^+$  product at higher energies and that other processes dominate the onset for production of  $[a_1]^+$ . This result is consistent with a recent study conducted by Bythell and Harrison where the direct  $[b_2]^+ \rightarrow [a_1]^+$  pathway was found to be insignificant for a  $[b_2]^+$  ion having a glycyglycine sequence [59].

Lastly, we analyze the secondary dissociation of  $H^+GGG$  in reaction 9, decarbonylation of the primary  $[y_2 + 2H]^+_{tot}$  product formed in reaction 5. As mentioned above, the direct study of  $H^+GG$  ( $[y_2 + 2H]^+$ ) revealed that the ( $[y_2 + 2H - CO]^+$ ) product ion formed after CO loss can undergo further fragmentation to form the  $[y_1 + 2H]^+$  product ion. As in that study, the composite cross-section was modeled as a sequential loss of CO from  $[y_2 + 2H]^+$  limited by tight  $TS[N^2]-ttgtt\{C^{\alpha 1}\sim OC\sim N^2\}$  [39], while still including competition with the major primary channels:  $[b_2]^+$ ,  $[y_1 + 2H]^+$ , and  $[a_1]^+$ . In this data analysis, we used a complex of CO and  $CH_2NH$  to approximate the molecular

parameters of the “neutral” formed after CO loss from the primary ion  $[y_2 + 2H]^+$  because this dissociation involves three molecular species. This assumption was not needed in the previous study because CO was the only neutral formed in that system. Our sequential model requires no scaling of the cross-section of reaction 9 similar to the previous analysis of this product [38], and reproduces the data well in the threshold region and up to  $\sim 5$  eV, Supplementary Figure S18. The threshold energy of the  $[y_2 + 2H - CO]^+$  product from our analysis is  $3.91 \pm 0.13$  eV, which lies  $1.63 \pm 0.14$  eV above the  $[y_2 + 2H]^+$  ( $H^+GG$ ) threshold and agrees well with the energy obtained previously for the dissociation of  $H^+GG$ ,  $1.62 \pm 0.05$  eV [38]. The result of this sequential model, along with its optimized parameters, is listed in Supplementary Table S6. Analysis of the composite cross-section of reaction 9 beyond 5 eV (1.4 eV above the apparent threshold for  $[y_2 + 2H]^+$ ) cannot be performed accurately because the  $[y_2 + 2H - CO]^+$  product decomposes to  $[y_1 + 2H]^+_{HE}$  starting  $\sim 1.5$  eV above the onset for  $[y_2 + 2H]^+$ , as shown by previous studies of  $H^+GG$ .

### Analysis of Tertiary Dissociation Channels

Decomposition of  $H^+GGG$  also produces third-order dissociation products:  $[y_1 + 2H]^+_{HE}$  in reaction 2b,  $[a_1]^+$  in reaction 7c,  $[a_2]^+$  in reaction 8b, and  $H^+(CH_2NH)_2$  in reaction 10. None of these reactions can be modeled directly in the present system because our data analysis program CRUNCH cannot handle such high order reactions for reasons noted above. However, formation of  $[a_1]^+$  from  $H^+GG$  ( $[y_2 + 2H]^+ \rightarrow [y_2 + 2H - CO]^+ \rightarrow [a_1]^+$ ) and  $[a_2]^+$  from  $H^+GG$  ( $[y_2 + 2H]^+ \rightarrow [b_2]^+ \rightarrow [a_2]^+$ ) have been analyzed previously in our lab yielding their threshold energies [38]. These previous studies on  $H^+GG$  decomposition also indicate that  $[y_1 + 2H]^+_{HE}$  formed in reaction 2b follows the loss of CO from  $[y_2 + 2H]^+_{tot}$  [38, 39]. Beyond 5 eV,  $[y_1 + 2H]^+_{HE}$  is the only contributor to the composite cross-section for reaction 9 ( $[y_2 + 2H - CO]^+ + [y_1 + 2H]^+_{HE}$ ) such that the threshold energy of reaction 9 beyond 5 eV is the threshold energy of the  $[y_1 + 2H]^+_{HE}$  product. Analysis of the composite cross-section for reaction 9 as a sequential dissociation from  $[y_2 + 2H]^+_{tot}$  requires slight scaling of the cross-section and yields a threshold energy of  $4.28 \pm 0.11$  eV, Supplementary Figure S19. This is  $2.00 \pm 0.13$  eV above the threshold energy of  $[y_2 + 2H]^+$  ( $H^+GG$ ), which agrees with the previously measured threshold energy of  $1.98 \pm 0.07$  eV [38] for the same process. A detailed description of this analysis that accounts for the differences in this process when formed from  $H^+GGG$  versus from  $H^+GG$  can be found in the SI. Optimized parameters of this model are listed in Supplementary Table S6.

Decomposition of the  $[a_2]^+$  product to form  $H^+(CH_2NH)_2$  in reaction 10 was observed in a previous study of the decomposition of the  $[b_2]^+$  ion [40] with a threshold energy of  $<2.50 \pm 0.31$  eV [40]. This work showed that reaction 10 is limited by tight  $TS(C_3H_7N_2O^+[N^2]-ct\{C\sim OC\sim N\})$ . Here, we find the threshold energy for formation of  $[b_2]^+$  is  $1.77 \pm 0.06$  eV, such that a fair estimate of the threshold energy for  $H^+(CH_2NH)_2$  formation is the sum of these two energies,  $<4.27 \pm 0.32$  eV. This value is somewhat lower than the apparent threshold of

~4.9 eV for the  $\text{H}^+(\text{CH}_2\text{NH})_2$  product, Figure 1, but this would be expected because of the competition inherent in this more complicated system.

### Comparison with Previous Experiments and Theory

In their study, KK [14] obtained the threshold energies for dissociation of  $\text{H}^+\text{GGG}$  into the two primary channels,  $[\text{b}_2]^+$  and  $[\text{y}_2 + 2\text{H}]^+$ , using an early version of CRUNCH. They obtained threshold energies of 2.36 and 3.17 eV, respectively, although neither value includes competitive or kinetic shifts. An estimate for kinetic shifts on the basis of their results for the  $\text{H}^+\text{GG}$  system yielded a more reasonable threshold energy value of <1.65 eV for  $[\text{y}_2 + 2\text{H}]^+$ . If we independently analyze the present data for  $[\text{b}_2]^+$  and  $[\text{y}_2 + 2\text{H}]^+$  without considering competitive or kinetic shifts, we obtain threshold energies of  $2.04 \pm 0.09$  and  $3.05 \pm 0.08$  eV, respectively, which are similar to the values of KK. When competitive and kinetic shifts are included, our threshold for the  $[\text{y}_2 + 2\text{H}]^+$  channel is  $2.28 \pm 0.06$  eV, well above the threshold estimated by KK, which simply points out the limitation of their estimate.

Aribi et al. [16] obtained a threshold energy of 1.37 (+0.10/−0.12) eV for the  $[\text{b}_2]^+$  ion including kinetic shifts and 2.81 eV without including kinetic shifts. These values differ wildly from our average values of  $1.77 \pm 0.06$  and  $2.04 \pm 0.09$  eV, respectively. As discussed in detail by Armentrout and Clark in their study of the  $[\text{b}_2]^+$  ion formed from  $\text{H}^+\text{GGG}$  [40], the  $[\text{b}_2]^+$  cross-section reported in the earlier study is shifted to higher energies and has an inflated cross-section, explaining why the thresholds without kinetic shifts differ. Once kinetic shifts are included, the difference between these threshold energies can be attributed to the choice of the rate-limiting TS used for data analysis. Aribi et al. used a tight TS [16], which should lead to a larger kinetic shift (1.44 eV) and lower threshold energy. In

contrast, we find the  $[\text{b}_2]^+$  product to be limited by a loose PSL TS, Supplementary Figure S6, such that the kinetic shift is modest (0.27 eV).

### Comparison with Theoretical Results

Figures 3, 4, and Supplementary Figures S14–S19 show that the experimental cross-sections for reactions 1a, 2a, 2b, 3–7a, 8a, and 9 are reproduced well by Equations S1, S3a, and S3b, over a wide range of energies (up to 4–6 eV) and over two orders of magnitude in cross-section. One additional means of evaluating the success of this analysis is to compare the experimental threshold energies obtained from the modeling with those obtained from calculations. Comparison of the experimental threshold energies of reactions 1a, 2a, 2b, 3–7a, 8a, 9, and 10 with 0 K values for the rate-limiting TSs for each channel calculated at the B3LYP, B3LYP-D3, MP2, and M06-2X levels of theory is shown in Table 2 and Figure 5.

In general, agreement between theory and experiment is reasonable with bigger differences between theory and experiment for the higher energy sequential dissociations. Thus, we first compare the mean absolute deviations (MADs) from experimental threshold values excluding these products. All four levels of theory yield between two and four threshold values that agree with experiment within the uncertainties. The MADs for the six primary products are similar for all four levels of theory, 6–14 kJ/mol. These MADs compare well with those of ~10 kJ/mol found previously in our laboratory for the decompositions of  $\text{H}^+\text{GG}$ ,  $\text{H}^+\text{G}$ , and  $\text{H}^+\text{Asn}$  systems [37, 38, 77]. At least two theoretical values fall within our range of experimental uncertainty for all primary products except reaction 5. For reaction 5 ( $[\text{y}_2 + 2\text{H}]^+$ ), theory underestimates the threshold energy by  $13\text{--}46 \pm 6$  kJ/mol. As noted above, our modeling is speculative for this product channel because theory indicates its

**Table 2.** Comparison of Experimental and Theoretical Reaction Energies at 0 K for Decomposition of  $\text{H}^+\text{GGG}$  in kJ/mol

Reaction	Transition state <sup>a</sup>	Experiment <sup>b</sup>		Theory <sup>c</sup>			
		this work	literature <sup>d</sup>	B3LYP	B3LYP-D3	MP2	M06-2X
$[\text{b}_2]^+$ ( $\text{H}^+\text{AMOX} + \text{G}$ )	PSL	$171 \pm 6$		153	172	169	171
$[\text{y}_1 + 2\text{H}]^+_{\text{LE}}$ ( $\text{H}^+\text{G} + \text{AMOX}$ )	PSL	$183 \pm 7$		182	200	189	196
$[\text{b}_3]^+$ ( $\text{H}^+\text{GAMOX} + \text{H}_2\text{O}$ )	$\text{TS}[\text{O}^{1\text{L}}\text{-O}^4]$	$136 \pm 9$		132	120	111	136
CO loss	$\text{TS}[\text{N}^2]$	$164 \pm 6$		164	164	172	175
$[\text{y}_2 + 2\text{H}]^+$ ( $\text{H}^+\text{GG} + \text{CO} + \text{CH}_2\text{NH}$ )	PSL	$220 \pm 6$		174	200	186	207
$[\text{a}_1]^+$ ( $\text{CH}_2\text{NH}_2 + \text{CO} + \text{GG}$ )	PSL	$247 \pm 7$		241	258	242	259
$[\text{a}_3]^+$ ( $\text{H}_2\text{O} + \text{CO}$ loss)	$[\text{b}_3]^+ \rightarrow \text{TS}(\text{H}^+\text{GAMOX})$	$224 \pm 12$		237	243	228	256
$[\text{a}_2]^+$ ( $\text{G} + \text{CO}$ loss)	$[\text{b}_2]^+ \rightarrow \text{TS}(\text{H}^+\text{AMOX})$	$327 \pm 6$	$322 \pm 13$	285	304	295	309
$[\text{y}_2 + 2\text{H} - \text{CO}]^+$	$[\text{y}_2 + 2\text{H}]^+ \rightarrow \text{TS}([\text{N}^2]\text{-ttgtt})$	$377 \pm 13$	$377 \pm 15$	328	352	353	366
$[\text{y}_1 + 2\text{H}]^+_{\text{HE}}$	$[\text{y}_2 + 2\text{H}]^+ \rightarrow \text{TS}([\text{N}^2]\text{-ttgtt})$	$413 \pm 11$	$411 \pm 11$	353	404	376	408
$[\text{a}_2 - \text{CO}]^+_{\text{e}}$	$[\text{a}_2]^+ \rightarrow \text{TS}(\text{C}_3\text{H}_7\text{N}_2\text{O}^+[\text{N}^2])$		$<412 \pm 31$	345	378	378	393
MAD <sup>f</sup>				$13 \pm 18, 29$	$11 \pm 9, 17$	$14 \pm 12, 21$	$6 \pm 6, 13$

<sup>a</sup>PSL = phase space limit.  $\text{TS}[\text{O}^{1\text{L}}\text{-O}^4] = \text{TS}[\text{O}^{1\text{L}}\text{-O}^4]\text{-ct}(\text{cg})\text{t}\{\text{C}\text{-O}^4\text{H}\}$ .  $\text{TS}[\text{N}^2] = \text{TS}[\text{N}^2]\text{-ttgtttt}\{\text{N}^1\text{C}\text{-CO}^1\text{-N}^2\text{C}\}$ .  $\text{TS}(\text{H}^+\text{GAMOX}) = \text{TS}(\text{H}^+\text{GAMOX}[\text{N}^3])\text{-ctg}(\text{cg})\{\text{C}\text{-CO}_3\}$ .  $\text{TS}(\text{H}^+\text{AMOX}) = \text{TS}(\text{H}^+\text{AMOX}[\text{N}^2]\text{-c}\{\text{C}\text{-C}\})$  [40].  $\text{TS}[\text{N}^2]\text{-ttgtt} = \text{TS}[\text{N}^2]\text{-ttgtt}\{\text{C}^{\text{a}1}\text{-OC}\text{-N}^2\}$  [39].  $\text{TS}(\text{C}_3\text{H}_7\text{N}_2\text{O}^+[\text{N}^2]) = \text{TS}(\text{C}_3\text{H}_7\text{N}_2\text{O}^+[\text{N}^2])\text{-ct}\{\text{C}\text{-OC}\text{-N}\}$  [40].

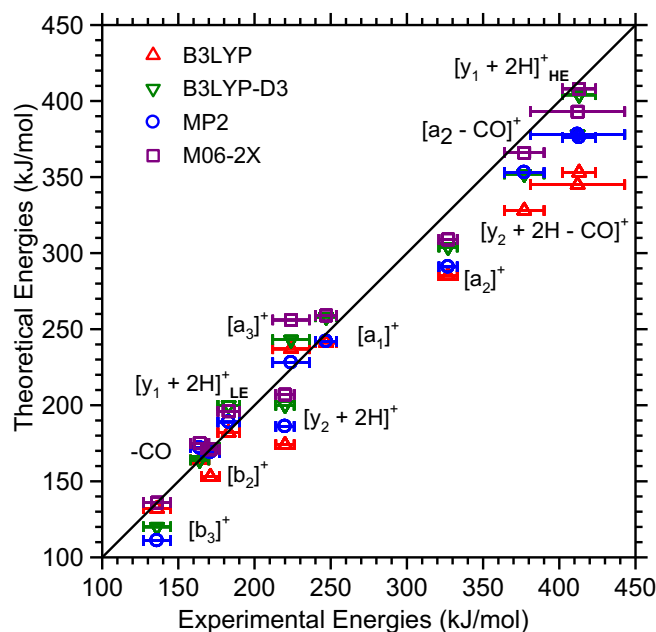
<sup>b</sup>Experimental values are taken from Table 1

<sup>c</sup>Computational results from Tables S3 and S4

<sup>d</sup>Values from references [38, 40]

<sup>e</sup>Threshold energy of this product was calculated using the measured energy of  $\text{TS}(\text{C}_3\text{H}_7\text{N}_2\text{O}^+[\text{N}^2])\text{-ct}\{\text{C}\text{-OC}\text{-N}\}$  [40] and that of the  $[\text{b}_2]^+$  product from this study

<sup>f</sup>Mean absolute deviation (MAD) from experimental threshold energies of products  $[\text{b}_2]^+$ ,  $[\text{y}_1 + 2\text{H}]^+$ ,  $[\text{b}_3]^+$ , CO loss,  $[\text{y}_2 + 2\text{H}]^+$ , and  $[\text{a}_1]^+$ . Italicized values are MADs from experimental threshold energies of all products investigated here



**Figure 5.** Comparison of experimental threshold energies for the products shown modeled using Equations S1 and S3 with B3LYP (red triangles), B3LYP-D3 (green triangles), MP2 (blue circles), and M06-2X (purple squares) energies taken from Table 1. The diagonal line indicates perfect agreement between theory and experiment

formation involves dissociation to three molecular species (one ion and two neutrals), whereas our best modeling ( $T_{PLP}$ ) assumes  $[y_2 + 2H]^+$  is formed as a primary product, i.e., dissociation occurs to one ion and one neutral (a loose complex of the two neutral species). This approach may over-estimate the true threshold energy because competition between the other primary products and the TTS for CO loss is not adequately represented.

For the secondary products,  $[a_3]^+$ ,  $[a_2]^+$ , and  $[y_2 + 2H - CO]^+$ , and tertiary products,  $[a_2 - CO]^+$  and  $[y_1 + 2H]_{HE}^+$ , at least one theoretical value is within experimental uncertainty. If all 11 fragmentations from  $H^+GGG$  are considered, the MADs increase to 13–29 kJ/mol, Table 2. Clearly, M06-2X level of theory gives the best agreement with experiment when the primary and when all products are considered, although the deviations are reasonable for all levels of theory. Overall, the good agreement between experiment and theory, especially for the primary products, validates the reaction mechanisms described above.

We can also compare the energy difference between the TS energy for CO loss and the sequential products  $[y_2 + 2H]^+$  and  $[a_1]^+$  that follow this putative primary product. Calculations suggest that the TS energy for CO loss is 6–36 kJ/mol below the  $[y_2 + 2H]^+$  PSL TS and 63–94 kJ/mol below the  $[a_1]^+$  PSL TS. Modeling the experimental data suggests that these differences are <57 kJ/mol for  $[y_2 + 2H]^+$  and <102 kJ/mol for the  $[a_1]^+$  product, consistent with theory.

We can also favorably compare the energy measured here for the subsequent decomposition of the  $[b_2]^+$  ion in reaction 8a,  $156 \pm 8$  kJ/mol, with that measured previously for the direct dissociation of  $[b_2]^+$ :  $151 \pm 12$  kJ/mol [40]. Similarly, the subsequent dissociation of the  $[y_2]^+$  ion to lose CO in reaction 9 is measured here to cost  $157 \pm 14$  kJ/mol, which can be compared with the CO loss threshold energy of  $157 \pm$

7 kJ/mol for the same fragmentation of  $H^+GG$  measured previously in our laboratory. Further, the threshold energy of reaction 2b involving the sequential dissociation of  $[y_2 + 2H]^+$  to  $[y_1 + 2H]_{HE}^+$  is found to be  $193 \pm 11$  kJ/mol above that of  $[y_2 + 2H]^+$ , which is consistent with  $191 \pm 7$  kJ/mol measured previously [38].

## Conclusion

In the present study, we have examined the conformers of  $H^+GGG$  using B3LYP, B3LYP-D3, MP2, and M06-2X levels of theory. The lowest energy  $[O^{1t}]$ -ctgttttt structure predicted by B3LYP and B3LYP-D3 agrees with the ground structure (GS) obtained previously by PS [18] and WM [60]. MP2 and M06-2X suggest that  $[N^1]$ -gcgttgtt is the lowest energy structure and is lower than the  $[N^1]$ -ttgttgtt GS previously suggested by the MP2 calculations of WM. Comparison of the calculated spectra for  $[O^{1t}]$ -ctgttttt,  $[N^1]$ -gcgttgtt, and the low-lying  $[N^1]$ -ttgttgtt and  $[N^1]$ -tcgctgtt conformers with the IRMPD spectrum obtained by WM indicate that  $[O^{1t}]$ -ctgttttt and  $[N^1]$ -ttgttgtt are the dominant species formed under experimental conditions, in agreement with the previous conclusions of WM (who did not consider the other two  $[N^1]$  conformers). Examination of the neutral GGG and pathways for cis-trans isomerization of  $H^+GGG$  indicate that the  $[N^1]$ -gcgttgtt and  $[N^1]$ -tcgctgtt conformers, both of which have a cis peptide bond between the first and second residues, are unlikely to be formed in solution or in the electrospray ionization process, explaining why they are not obvious species in the IRMPD spectrum. The presence of both  $[O^{1t}]$ -ctgttttt and  $[N^1]$ -ttgttgtt is consistent with MP2 populations calculated at 298 K, whereas B3LYP, B3LYP-D3, and M06-2X energies would suggest that only the



[O<sup>1+</sup>] species should be present. We also find that interconversion of [N<sup>1+</sup>]-ttggtgt and [O<sup>1+</sup>]-ctgtttt requires surmounting an appreciable barrier, indicating that both species are probably formed directly in solution or the electrospray ionization process. Interconversion can occur after collisional activation, however, such that both isomers share common rate-limiting TSs for decomposition.

The [b<sub>2</sub>]<sup>+</sup> (H<sup>+</sup>AMOX) and [y<sub>1</sub> + 2H]<sup>+</sup> (H<sup>+</sup>G) product ions are formed competitively via the “oxazolone” pathway, comparable to the results of PS, and are limited by the separated product energies and, thus, their transition states are loose, Supplementary Figure S6. CO loss from H<sup>+</sup>G<sub>3</sub> has a tight TS and can occur through three similar pathways, with the lowest energy pathway lying 164–175 kJ/mol above the GS (slightly lower in energy than the pathway found by PS). Loss of CO from H<sup>+</sup>G<sub>3</sub> forms a (CH<sub>2</sub>NH<sub>2</sub><sup>+</sup>)(GG) complex, which dissociates competitively to form ionic products [y<sub>2</sub> + 2H]<sup>+</sup> (H<sup>+</sup>GG) and [a<sub>1</sub>]<sup>+</sup> (CH<sub>2</sub>NH<sub>2</sub><sup>+</sup>) via loose rate-limiting TSs, Supplementary Figure S8. At somewhat lower energies, the (CH<sub>2</sub>NH<sub>2</sub><sup>+</sup>)(GG) complex can also rearrange and lose a molecule of ammonia, Supplementary Figure S13, which is limited by the initial loss of CO at 164–175 kJ/mol at the B3LYP and MP2 levels, a result that is largely confirmed by the analysis of the present data. The loss of water from H<sup>+</sup>G<sub>3</sub> has a tight TS, Supplementary Figure S9, and results in the formation of the [b<sub>3</sub>]<sup>+</sup> product, H<sup>+</sup>GAMOX, which has an oxazolone structure similar to [b<sub>2</sub>]<sup>+</sup> with an additional glycyl group. Further cleavage of the oxazolone ring of the [b<sub>3</sub>]<sup>+</sup> product can release CO and form [a<sub>3</sub>]<sup>+</sup>, Supplementary Figure S10. Because loss of water from H<sup>+</sup>GAMOX(H<sub>2</sub>O) is both energetically and entropically more favorable than the decarbonylation, Supplementary Figure S10, these processes are likely to occur sequentially in the experiment. Loss of CO from [b<sub>3</sub>]<sup>+</sup> is limited by a tight TS that forms an acyclic ion, [a<sub>3</sub>]<sup>+</sup> (GGNHCH<sub>2</sub><sup>+</sup>); however, cyclization of GGNHCH<sub>2</sub><sup>+</sup> can take place to form a seven-membered ring conformer with little additional energy (although the extra step makes it entropically less likely to occur).

The kinetic energy dependence of the collision-induced dissociation of H<sup>+</sup>G<sub>3</sub> with Xe is examined in a guided ion beam tandem mass spectrometer. Kinetic energy-dependent cross-sections for six primary products: [b<sub>2</sub>]<sup>+</sup>, [y<sub>1</sub> + 2H]<sup>+</sup><sub>LE</sub>, [b<sub>3</sub>]<sup>+</sup>, CO loss, [y<sub>2</sub> + 2H]<sup>+</sup>, and [a<sub>1</sub>]<sup>+</sup>; three secondary products: [a<sub>2</sub>]<sup>+</sup>, [a<sub>3</sub>]<sup>+</sup>, and [y<sub>2</sub> + 2H – CO]<sup>+</sup>; and two tertiary products: [y<sub>1</sub> + 2H]<sup>+</sup><sub>HE</sub> and [a<sub>2</sub> – CO]<sup>+</sup> are observed and analyzed. Threshold energies at 0 K for all of these processes are determined after consideration of the effects of multiple collisions with Xe, internal energy, lifetime effects, and competition [78–80]. Notably, this work entails the first simultaneous analysis of six product channels, with reasonable results obtained for their threshold energies. Experimental results are compared with detailed quantum chemical calculations discussed above. Thresholds for reactions 2a ([y<sub>1</sub> + 2H]<sup>+</sup><sub>LE</sub>), 3 ([b<sub>3</sub>]<sup>+</sup>), 4 (CO loss), 6 ([a<sub>3</sub>]<sup>+</sup>), and 7a ([a<sub>1</sub>]<sup>+</sup>) constitute the first experimental measurements, and those for reactions 1a ([b<sub>2</sub>]<sup>+</sup>) and 5 ([y<sub>2</sub> + 2H]<sup>+</sup>) have been refined to provide more accurate thermodynamics for these two products. We find that the threshold

energy values of all reactions (except reaction 5) lie within experimental uncertainty of at least one of the four theoretical values. This good agreement validates the reactions mechanisms explored theoretically and also allows identification of the structures of the products formed at threshold. The discrepancy between the experimental and theoretical threshold energies of the [y<sub>2</sub> + 2H]<sup>+</sup> primary product in reaction 5 is attributed to its more speculative data analysis. Comparison of the threshold energies of primary products with theory indicates that the four levels of theory give MADs of 6–14 kJ/mol, with M06-2X giving the best agreement. Threshold energies of reactions 2b ([y<sub>1</sub> + 2H]<sup>+</sup><sub>HE</sub>), 8a ([a<sub>2</sub>]<sup>+</sup>), 9 ([y<sub>2</sub> + 2H – CO]<sup>+</sup>), and 10 ([a<sub>2</sub> – CO]<sup>+</sup>) are consistent with those measured previously in our laboratory for the same products [38, 40].

Good agreement between experimental and theoretical energies also indicates that the reactions begin at the ground conformation (or those conformations populated within a thermal equilibrium). Formation of the products observed in the decomposition of H<sup>+</sup>G<sub>3</sub> requires that the extra proton is able to move. Reaction mechanisms in the present study clearly show that the extra proton is mobile once the ground state reactant is collisionally activated. Thus, these results provide a detailed exposition of the utility and scope of the mobile proton model [6, 27] in understanding fragmentations of the simplest tripeptide. Overall, these results hopefully can be used to better predict mass spectra and, thus, improve our understanding of the fragmentations of larger peptides and also improve the algorithms used to predict sequences [81].

As noted above, the ability to obtain accurate thermodynamic information allows a comparison with theoretical results, which permits the identification of the structures of products. Thus, TCID methods provide a similar component of structural identification as spectroscopic approaches, while additionally yielding valuable thermodynamic information. Notably, such structures need not always agree depending on whether the ion sources and conditions used to generate adequate product ion intensity for spectroscopic interrogation allow subsequent isomerization. In this regard, it is also interesting to contemplate how large a peptide can be studied by TCID. Clearly, larger peptides will exhibit increasing kinetic shifts resulting from increased degrees of freedom, which will limit the accuracy of the kinetic modeling needed to extract useful threshold information. Further, slower processes (those occurring via TTSSs) may not be observed within the experimental time-frame available. Given the success of the present study in comprehensively examining a tripeptide, we conservatively speculate that accurate studies of hexapeptides should be feasible in the future.

Another important finding of this work is that the data analysis does not demand that we have exact knowledge of the rate-limiting TSs involved in these reactions. In fact, the magnitude of the experimental cross-section of a particular product generally indicates the nature of the rate-limiting TS involved in that reaction. For instance, the large cross-section of [b<sub>2</sub>]<sup>+</sup> indicates that it is dictated by a loose TS, whereas the small cross-section

for the  $[b_3]^+$  product suggests that it is formed via a tight TS. Thus, the data determines whether a particular product is formed via a tight or loose TS, such that adjustment of the relative tightness or looseness of the TS (in the form of  $\Delta S^\ddagger$ ) should enable accurate analysis of the data. Consequently, for more complex systems, a generic set of vibrational frequencies and rotational constants (which can be modified to fine-tune  $\Delta S^\ddagger$ ) as input should suffice for data analysis.

## Acknowledgements

The authors acknowledge support for this work by the National Science Foundation, grant CHE-1359769, and grants of computational time from the Center for High Performance Computing at the University of Utah. The authors thank Dr. Thomas Wytenbach and Dr. Mike Bowers, University of Santa Barbara (UCSB), for sharing their ion mobility data and calculating the collision cross-sections of our calculated protonated triglycine conformers.

## References

- Papayannopoulos, I.: The interpretation of collision-induced dissociation tandem mass spectra of peptides. *Mass Spectrom. Rev.* **14**, 49–73 (1995)
- Mann, M., Wilm, M.: Error-tolerant identification of peptides in sequence databases by peptide sequence tags. *Anal. Chem.* **66**, 4390–4399 (1994)
- Eng, J.K., McCormack, A.L., Yates, J.R.: An approach to correlate tandem mass spectral data of peptides with acid sequences in a protein database. *J. Am. Soc. Mass Spectrom.* **5**, 976–989 (1994)
- Perkins, D.N., Pappin, D.J.C., Creasy, D.M., Cottrell, J.S.: Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **20**, 3551–3567 (1999)
- Paizs, B., Suhai, S.: Fragmentation pathways of protonated peptides. *Mass Spectrom. Rev.* **24**, 508–548 (2005)
- Dongré, A.R., Jones, J.L., Somogyi, Á., 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.* **118**, 8365–8374 (1996)
- Wysocki, V.H., Tsaprailis, G., Smith, L.L., Brechi, L.A.: Mobile and localized protons: a framework for understanding peptide dissociation. *J. Mass Spectrom.* **35**, 1399–1406 (2000)
- Burlet, O., Orkiszewski, R.S., Ballard, K.D., Gaskell, S.J., Bertrand, M.J.: Charge promotion of low-energy fragmentations of peptide ions. *Rapid Commun. Mass Spectrom.* **6**, 658–662 (1992)
- Burlet, O., Yang, C.Y., Gaskell, S.J.: Influence of cysteine to cysteic acid oxidation on the collision-activated decomposition of protonated peptides: evidence for intraionic interactions. *J. Am. Soc. Mass Spectrom.* **3**, 337–344 (1992)
- Cox, K.A., Gaskell, S.J., Morris, M., Whiting, A.: Role of the site of protonation in the low-energy decompositions of gas-phase peptide ions. *J. Am. Soc. Mass Spectrom.* **7**, 522–531 (1996)
- Chu, I.K., Siu, C.-K., Lau, J.K.-C., Tang, W.K., Mu, X., Lai, C.K., Guo, X., Wang, X., Li, N., Xia, Y., Kong, X., Oh, H.B., Ryzhov, V., Tureček, F., Hopkinson, A.C., Siu, K.W.M.: Proposed Nomenclature for Peptide Ion Fragmentation. *Int. J. Mass Spectrom.* **390**, 24–27 (2015)
- Roepstorff, P., Fohlman, J.: Proposal for a common nomenclature for sequence ions in mass spectra of peptides. *Biomed. Mass Spectrom.* **11**, 601 (1984)
- Biemann, K.: Contributions of mass spectrometry to peptide and protein structure. *Biomed. Environ. Mass Spectrom.* **16**, 99–111 (1988)
- 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.* **119**, 6552–6563 (1997)
- 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.* **190/191**, 209–230 (1999)
- Aribi, H.E., Rodriguez, C.F., Almeida, D.R.P., Ling, Y., Mak, W.W.N., Hopkinson, A.C., Siu, K.W.M.: Elucidation of fragmentation mechanisms of protonated peptide ions and their products: a case study on glycyglycylglycine using density functional theory and threshold collision-induced dissociation. *J. Am. Chem. Soc.* **125**, 9229–9236 (2003)
- Rodriguez, 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.* **123**, 3006–3012 (2001)
- Paizs, B., Suhai, S.: Combined quantum chemical and RRKM modeling of the main fragmentation pathways of protonated GGG. I. Cis-trans isomerization around protonated amide bonds. *Rapid Commun. Mass Spectrom.* **15**, 2307–2323 (2001)
- Paizs, B., Suhai, S.: Combined quantum chemical and RRKM modeling of the main fragmentation pathways of protonated GGG. II. Formation of  $b_2$ ,  $y_1$ , and  $y_2$  ions. *Rapid Commun. Mass Spectrom.* **16**, 375–389 (2002)
- Robinson, P.J., Holbrook, K.A.: *Unimolecular Reactions*. Wiley Interscience, New York (1972)
- Gilbert, R.G., Smith, S.C.: *Theory of Unimolecular and Recombination Reactions*. Blackwell Scientific, London (1990)
- Morgan, D., Bursey, M.: A linear free-energy correlation in the low-energy tandem mass spectra of protonated tripeptides Gly–Gly–Xxx. *Org. Mass Spectrom.* **29**, 354–359 (1995)
- Polce, M.J., Ren, D., Wesdemiotis, C.: Dissociation of the peptide bond in protonated peptides. *J. Mass Spectrom.* **35**, 1391–1398 (2000)
- Spezia, R., Martens, J., Oomens, J., Song, K.: Collision-induced dissociation pathways of protonated  $\text{Gly}_2\text{NH}_2$  and  $\text{Gly}_3\text{NH}_2$  in the short time-scale limit by chemical dynamics and ion spectroscopy. *Int. J. Mass Spectrom.* **388**, 40–52 (2015)
- Grimme, S., Antony, J., Ehrlich, S., Krieg, H.: A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H–Pu. *J. Chem. Phys.* **132**, 154104–154119 (2010)
- Grimme, S., Ehrlich, S., Goerigk, L.: Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* **32**, 1456–1465 (2011)
- Somogyi, Á., Wysocki, V.H., Mayer, I.: The effect of protonation site on bond strengths in simple peptides: application of ab initio and modified neglect of differential overlap bond orders and modified neglect of differential overlap energy partitioning. *J. Am. Soc. Mass Spectrom.* **5**, 704–717 (1994)
- Ervin, K.M., Armentrout, P.B.: Translational energy dependence of  $\text{Ar}^+ + \text{XY} \rightarrow \text{ArX}^+ + \text{Y}$  ( $\text{XY} = \text{H}_2, \text{D}_2, \text{HD}$ ) from thermal to 30 eV c.m. *J. Chem. Phys.* **83**, 166–189 (1985)
- Muntean, F., Armentrout, P.B.: Guided ion beam study of collision-induced dissociation dynamics: integral and differential cross-sections. *J. Chem. Phys.* **115**, 1213–1228 (2001)
- Moision, R.M., Armentrout, P.B.: An electrospray source for thermochemical investigation with the guided ion beam mass spectrometer. *J. Am. Soc. Mass Spectrom.* **18**, 1124–1134 (2007)
- Moision, R.M., Armentrout, P.B.: An electrospray ionization source for thermochemical investigation with the guided ion beam mass spectrometer. *J. Am. Soc. Mass Spectrom.* **18**, 1124–1134 (2007)
- Carl, D.R., Moision, R.M., Armentrout, P.B.: Binding energies for the inner hydration shells of  $\text{Ca}^{2+}$ : an experimental and theoretical investigation of  $\text{Ca}^{2+}(\text{H}_2\text{O})_x$  complexes ( $x = 5–9$ ). *Int. J. Mass Spectrom.* **265**, 308–325 (2007)
- Heaton, A.L., Moision, R.M., Armentrout, P.B.: Experimental and theoretical studies of sodium cation interactions with the acidic amino acids and their amide derivatives. *J. Phys. Chem. A* **112**, 3319–3327 (2008)
- Heaton, A.L., Armentrout, P.B.: Experimental and theoretical studies of potassium cation interactions with the acidic amino acids and their amide derivatives. *J. Phys. Chem. B* **112**, 12056–12065 (2008)
- Ye, S.J., Clark, A.A., Armentrout, P.B.: An experimental and theoretical investigation of alkali metal cation interactions with hydroxyl side chain amino acids. *J. Phys. Chem. B* **112**, 10291–10302 (2008)
- Ye, S.J., Armentrout, P.B.: Absolute thermodynamic measurements of alkali metal cation interactions with a simple dipeptide and tripeptide. *J. Phys. Chem. A* **112**, 3587–3596 (2008)
- Armentrout, P.B., Heaton, A.L., Ye, S.J.: Thermodynamics and mechanisms for decomposition of protonated glycine and its protonated dimer. *J. Phys. Chem. A* **115**, 11144–11155 (2011)
- Armentrout, P.B., Heaton, A.L.: Thermodynamics and mechanisms of protonated diglycine decomposition: a guided ion beam study. *J. Am. Soc. Mass Spectrom.* **23**, 632–643 (2012)

39. Armentrout, P.B., Heaton, A.L.: Thermodynamics and mechanisms of protonated diglycine decomposition: a computational study. *J. Am. Soc. Mass Spectrom.* **23**, 621–631 (2012)
40. Armentrout, P.B., Clark, A.A.: The simplest  $b_2^+$  ion: determining its structure from its energetics by a direct comparison of the threshold collision-induced dissociation of protonated oxazolone and diketopiperazine. *Int. J. Mass Spectrom.* **316/318**, 182–191 (2012)
41. Mookherjee, A., Armentrout, P.B.: Theoretical investigation and reinterpretation of the decomposition of lithiated proline and N-methyl proline. *Int. J. Mass Spectrom.* **370**, 16–28 (2014)
42. Wang, D., Gulyuz, K., Stedwell, C.N., Polfer, N.C.: Diagnostic NH and OH vibrations for oxazolone and diketopiperazine structures:  $b_2$  from protonated triglycine. *J. Am. Soc. Mass Spectrom.* **22**, 1197–1203 (2011)
43. Oomens, J., Young, S., Molesworth, S., Stipdonk, M.J.V.: Spectroscopic evidence for an oxazolone structure of the  $b_2$  fragment ion from protonated tri-alanine. *J. Am. Soc. Mass Spectrom.* **20**, 334–339 (2009)
44. Yoon, S.H., Chamot-Rooke, J., Perkins, B.R., Hilderbrand, A.E., Poutsma, J.C., Wysocki, V.H.: IRMPD spectroscopy shows that AGG forms an oxazolone  $b_2^+$  ion. *J. Am. Chem. Soc.* **130**, 17644–17645 (2008)
45. Polfer, N.C., Oomens, J., Suhai, S., Paizs, B.: Infrared spectroscopy and theoretical studies on gas-phase protonated leu-enkephalin and its fragments: direct experimental evidence for the mobile proton. *J. Am. Chem. Soc.* **129**, 5887–5897 (2007)
46. Bythell, B.J., Erkelam, U., Paizs, B., Maitre, P.: Infrared spectroscopy of fragments from doubly protonated tryptic peptides. *Chem. Phys. Chem.* **10**, 883–885 (2009)
47. Chen, X., Yu, L., Steill, J.D., Oomens, J., Polfer, N.C.: Effect of peptide fragment size on the propensity of cyclization in collision-induced dissociation: oligoglycine  $b_2$ – $b_8$ . *J. Am. Chem. Soc.* **131**, 18272–18282 (2009)
48. 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.* **6**, 1165–1174 (1995)
49. Bythell, B.J., Barofsky, D.F., Pingitore, F., Polce, M.J., Wang, P., Wesdemiotis, C., Paizs, B.: Backbone cleavages and sequential loss of carbon monoxide and ammonia from protonated AGG: a combined tandem mass spectrometry, isotope labeling, and theoretical study. *J. Am. Soc. Mass Spectrom.* **18**, 1291–1303 (2007)
50. Verkerk, U.H., Siu, C., Steill, J.D., Aribi, H.E., Zhao, J., Rodriguez, C.F., Oomens, J., Hopkinson, A.C., Siu, K.W.M.:  $a_2$  Ion derived from triglycine: an  $N_1$ -protonated 4-imidazolidinone. *J. Phys. Chem. Lett.* **1**, 868–872 (2010)
51. Bythell, B.J., Maitre, P., Paizs, B.: Cyclization and rearrangement reactions of  $a_n$  fragment ions of protonated peptides. *J. Am. Chem. Soc.* **132**, 14766–14779 (2010)
52. Aribi, H.E., Orlova, G., Rodriguez, 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* **108**, 18743–18749 (2004)
53. Yalcin, T., Csizmadia, I.G., Peterson, M.R., Harrison, A.G.: The structure and fragmentation of  $B_n$  ( $n \geq 3$ ) ions in peptide spectra. *J. Am. Soc. Mass Spectrom.* **7**, 233–242 (1996)
54. Ambihapathy, K., Yalcin, T., Leung, H.-W., Harrison, A.G.: Pathways to immonium ions in the fragmentation of protonated peptides. *J. Mass Spectrom.* **32**, 209–215 (1997)
55. Harrison, A.G.: To b or not to b: the ongoing saga of peptide b ions. *Mass Spectrom. Rev.* **28**, 649–654 (2009)
56. Verkerk, U.H., Zhao, J., Lau, J.K.-C., Lam, T.-W., Hao, Q., Steill, J.D., Siu, C.-K., Oomens, J., Hopkinson, A.C., Siu, K.W.M.: Structures of the  $a_2$  Ions of Ala-Ala-Ala and Phe-Phe-Phe. *Int. J. Mass Spectrom.* **330/332**, 254–261 (2012)
57. Allen, J.M., Racine, A.H., Berman, A.M., Johnson, J.S., Bythell, B.J., Paizs, B., Glish, G.L.: Why are  $a_3$  ions rarely observed? *J. Am. Soc. Mass Spectrom.* **19**, 1764–1770 (2008)
58. Savitski, M.M., Falth, M., Fung, Y.M.E., Adams, C.M., Zubarev, R.A.: Bifurcating fragmentation behavior of gas-phase tryptic peptide dications in collisional activation. *J. Am. Soc. Mass Spectrom.* **19**, 1755–1763 (2008)
59. Bythell, B.J., Harrison, A.G.: Formation of  $a_1$  ions directly from oxazolone  $b_2$  ions: an energy-resolved and computational study. *J. Am. Soc. Mass Spectrom.* **26**, 774–781 (2015)
60. Wu, R., McMahon, T.B.: Infrared multiple photon dissociation spectroscopy as structural confirmation for GlyGlyGlyH<sup>+</sup> and AlaAlaAlaH<sup>+</sup> in the gas phase. Evidence for amide oxygen as the protonation site. *J. Am. Chem. Soc.* **129**, 11312–11313 (2007)
61. Wu, J., Lebrilla, C.B.: Gas-phase basicities and sites of protonation of glycine oligomers (Gly<sub>n</sub>; n = 1–5). *J. Am. Chem. Soc.* **115**, 3270–3275 (1993)
62. Zhang, K., Cassady, C.J., Chung-Phillips, A.: Ab initio studies of neutral and protonated triglycines: comparison of calculated and experimental gas-phase basicity. *J. Am. Chem. Soc.* **116**, 11512–11521 (1994)
63. Zhang, K., Zimmerman, D.M., Chung-Phillips, A., Cassady, C.J.: Experimental and ab initio studies of the gas-phase basicities of polyglycines. *J. Am. Chem. Soc.* **115**, 10812–10822 (1993)
64. Wytenbach, T., Bushnell, J.E., Bowers, M.T.: Salt bridge structures in the absence of solvent? The case for the oligoglycines. *J. Am. Chem. Soc.* **120**, 5098–5103 (1998)
65. Wytenbach, T., Helden, G.V., Batka, J.J.J., Carlat, D., Bowers, M.T.: Effect of the long-range potential on ion mobility measurements. *J. Am. Soc. Mass Spectrom.* **8**, 275–282 (1997)
66. Bleiholder, C., Wytenbach, T., Bowers, M.T.: A novel projection approximation algorithm for the fast and accurate computation of molecular collision cross-sections (I): Method. *Int. J. Mass Spectrom.* **308**, 1–10 (2011)
67. Cordero, M.M., Houser, J.J., Wesdemiotis, C.: The neutral products formed during backbone fragmentations of protonated peptides in tandem mass spectrometry. *AC* **65**, 1594–1601 (1993)
68. 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.* **9**, 1275–1284 (1998)
69. Ballard, K.D., Gaskell, S.J.: Dehydration of peptide  $[M + H]^+$  ions in the gas phase. *J. Am. Soc. Mass Spectrom.* **4**, 477–481 (1993)
70. Ballard, K.D., Gaskell, S.J.: Sequential mass spectrometry applied to the study of the formation of “internal” fragment ions of protonated peptides. *Int. J. Mass Spectrom. Ion Processes* **111**, 173–189 (1991)
71. Balta, B., Aviyente, V., Lifshitz, C.: Elimination of water from the carboxyl group of GlyGlyH<sup>+</sup>. *J. Am. Soc. Mass Spectrom.* **14**, 1192–1203 (2003)
72. 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.* **9**, 945–956 (1998)
73. Bythell, B.J., Dain, R.P., Curtice, S.S., Oomens, J., Steill, J.D., Groenewold, G.S., Paizs, B., Stipdonk, M.J.V.: Structure of  $[M + H - H_2O]^+$  from protonated triglycine revealed by tandem mass spectrometry and IRMPD spectroscopy. *J. Phys. Chem. A* **114**, 5076–5082 (2010)
74. Lau, J.K.C., Zhao, J., Siu, K.W.M., Hopkinson, A.C.: Elimination of water from the backbone of protonated tetraglycine. *Int. J. Mass Spectrom.* **316/318**, 268–272 (2012)
75. Verkerk, U.H., Zhao, J., Stipdonk, M.J.V., Bythell, B.J., Oomens, J., Hopkinson, A.C., Siu, K.W.M.: Structure of the  $[M + H - H_2O]^+$  ion from tetraglycine: a revisit by means of density functional theory and isotope labeling. *J. Phys. Chem. A* **115**, 6683–6687 (2011)
76. Cooper, T., Talaty, E., Grove, J., Stipdonk, M.J.V., Suhai, S., Paizs, B.: Isotope labeling and theoretical study of the formation of  $a_3^*$  ions from protonated tetraglycine. *J. Am. Soc. Mass Spectrom.* **17**, 1654–1664 (2006)
77. Heaton, A.L., Armentrout, P.B.: Thermodynamics and mechanism of protonated asparagine decomposition. *J. Am. Soc. Mass Spectrom.* **20**, 852–866 (2009)
78. Rodgers, M.T., Armentrout, P.B.: Statistical modeling of competitive threshold collision-induced dissociation. *J. Chem. Phys.* **109**, 1787–1800 (1998)
79. Stein, S.E., Rabinovich, B.S.: On the use of exact state counting methods in RRKM rate calculations. *Chem. Phys. Lett.* **49**, 183–188 (1977)
80. Rodgers, M.T., Ervin, K.M., Armentrout, P.B.: Statistical modeling of collision-induced dissociation thresholds. *J. Chem. Phys.* **106**, 4499–4508 (1997)
81. Wysocki, V.H., Cheng, G., Zhang, Q., Herrmann, K.A., Beardsley, R.L., Hilderbrand, A.E.: Peptide Fragmentation Overview. In *Principles of Mass Spectrometry Applied to Biomolecules*. Wiley, Hoboken, New Jersey. 277–300 (2006)