J|A|C|S ARTICLES

Published on Web 10/06/2010

Cyclization and Rearrangement Reactions of a_n Fragment Ions of Protonated Peptides

Benjamin J. Bythell,[†] Philippe Maître,*[,‡] and Béla Paizs*[,†]

Computational Proteomics Group, German Cancer Research Center, Im Neuenheimer Feld 580, 69120 Heidelberg, Germany, and Laboratoire de Chimie Physique, Université Paris-Sud 11, UMR8000 CNRS, Faculté des Sciences, Bât. 350, 91405 Orsay Cedex, France

Received February 23, 2010; E-mail: B.Paizs@dkfz.de; Philippe.Maitre@u-psud.fr

Abstract: a_n ions are frequently formed in collision-induced dissociation (CID) of protonated peptides in tandem mass spectrometry (MS/MS) based sequencing experiments. These ions have generally been assumed to exist as immonium derivatives ($-HN^+{=}CHR$). Using a quadrupole ion trap mass spectrometer, MS/MS experiments have been performed and the structure of a_n ions formed from oligoglycines was probed by infrared spectroscopy. The structure and isomerization reactions of the same ions were studied using density functional theory. Overall, theory and infrared spectroscopy provide compelling evidence that a_n ions undergo cyclization and/or rearrangement reactions, and the resulting structure(s) observed under our experimental conditions depends on the size (n). The a_2 ion (GG sequence) undergoes cyclization to form a 5-membered ring isomer. The a_3 ion (GGG sequence) undergoes cyclization initiated by nucleophilic attack of the carbonyl oxygen of the N-terminal glycine residue on the carbon center of the C-terminal immonium group forming a 7-membered ring isomer. The barrier to this reaction is comparatively low at 10.5 kcal mol^{-1}, and the resulting cyclic isomer (-5.4 kcal mol^{-1}) is more energetically favorable than the linear form. The a_4 ion with the GGGG sequence undergoes head-to-tail cyclization via nucleophilic attack of the N-terminal amino group on the carbon center of the C-terminal immonium ion, forming an 11-membered macroring which contains a secondary amine and three trans amide bonds. Then an intermolecular proton transfer isomerizes the initially formed secondary amine moiety ($-CH_2-NH_2^+-CH_2-NH-CO-$) to form a new $-CH_2-NH-CH_2-NH_2^+-CO-$ form. This structure is readily cleaved at the $-CH_2-NH_2^+-$ bond, leading to opening of the macrocycle and formation of a rearranged linear isomer with the $H_2C{=}NH^+-CH_2-$ moiety at the N terminus and the $-CO-NH_2$ amide bond at the C terminus. This rearranged linear structure is much more energetically favorable (-14.0 kcal mol^{-1}) than the initially formed imine-protonated linear a_4 ion structure. Furthermore, the barriers to these cyclization and ring-opening reactions are low (8−11 kcal mol^{-1}), allowing facile formation of the rearranged linear species in the mass spectrometer. This finding is not limited to 'simple' glycine-containing systems, as evidenced by the IRMPD spectrum of the a_4 ion generated from protonated AAAAA, which shows a stronger tendency toward formation of the energetically favorable (-12.3 kcal mol^{-1}) rearranged linear structure with the $MeHC{=}NH^+-CHMe-$ moiety at the N terminus and the $-CO-NH_2$ amide bond at the C terminus. Our results indicate that one needs to consider a complex variety of cyclization and rearrangement reactions in order to decipher the structure and fragmentation pathways of peptide a_n ions. The implications this potentially has for peptide sequencing are also discussed.

Introduction

Peptide sequencing in proteomics is primarily achieved by tandem mass spectrometry (MS/MS) of protonated peptides. In these experiments peptide ions are excited by collisions with inert gases (collision-induced dissociation (CID)) to induce fragmentation, and the resulting product ion spectra are used to decipher the sequences. The product ion spectra of protonated peptides are usually dominated by sequence informative *b*, *a*, and *y* fragments,[1] which are formed in a complex reaction cascade.[2,3] To facilitate rapid processing of the large number of spectra routinely produced in high-throughput proteomics experiments, various bioinformatics tools have been developed.[4] These programs utilize fragmentation models[2] to generate theoretical spectra for candidate sequences and various mathematical measures to assess similarity between these theoretical spectra and the experimental MS/MS spectra. One of the basic assumptions inherent to the current implementation of this strategy is that peptide ions or their fragments dissociate on

[†] German Cancer Research Center.
[‡] Université Paris-Sud 11.

(1) (a) Roepstorff, P.; Fohlmann, J. *J. Biomed. Mass Spectrom.* **1984**, *11*, 601. (b) Biemann, K. *Biomed. Environ. Mass Spectrom.* **1988**, *16*, 99.
(2) Paizs, B.; Suhai, S. *Mass Spectrom. Rev.* **2005**, *24*, 508.
(3) Paizs, B.; van Stipdonk, M. *J. Am. Soc. Mass Spectrom.* **2008**, *19*, 1717. papers in this focus issue on peptide fragmentation.
(4) (a) Steen, H.; Mann, M. *Nat. Rev. Mol. Cell. Biol.* **2004**, *5*, 699. (b) Nesvizhskii, A. E.; Vitek, O.; Aebersold, R. *Nat. Methods* **2007**, *4*, 787.

 10.1021/ja101556g

direct fragmentation pathways,[5] which do not introduce rearrangements of the original sequences. For example, current fragmentation models consider primary dissociations like PEPTIDE → PEPTID + E, PEPTIDE → PEPT + IDE, etc. (where PEPTIDE is a hypothetical heptapeptide and the ionizing proton is not shown for simplicity), secondary processes like PEPTID → PEPTI + D are also considered to be possible, but these algorithms completely exclude the possibility of fragmentation reactions like PEPTID → PEPID + T or PEPTID → PEPD + TI which produce *nondirect* sequence ions.[5]

Recent studies[5] indicate that the dissociation chemistry of *b* ions cannot be universally described by considering only *direct* fragmentation chemistries. Small *b* ions terminated by the 5-membered oxazolone ring[6,7] undergo sequential degradation[6a,b,8] at their C termini because the oxazolone ring is more labile than the backbone amide bonds ('oxazolone' rule,[9] PEPTID → PEPTI + D type reaction leading to *direct* sequence ions; here PEPTID refers to an oxazolone-terminated *b* ion). Larger *b* ions, however, often fragment by elimination of formally internal amino acid residues[5] in PEPTID → PEPID + T type reactions forming *nondirect* sequence[5] ions. Simultaneous cleavage of two amide bonds and subsequent reassociation of the terminal fragments (for example, PEPTID → PEP + T + ID → PEPID + T type reactions) is unlikely to occur. Instead, the initially formed oxazolone *b* isomer undergoes head-to-tail cyclization[5,9,10] to form a macrocyclic isomer of the *b* ion. Intermolecular proton transfer pathways can then result in protonation of various amide nitrogens, enabling opening of the cyclic backbone reforming linear isomers terminated by the oxazolone ring. This chemistry in all but the case where the original sequence is regenerated leads to linear *b* isomers with scrambled sequences which are linear permutations of the original sequence. Fragmentation of these 'scrambled' linear *b* ions leads to elimination of initially internal amino acid residues (for example, PEPTID → cyclo-(PEPTID) → IDPEPT → IDPEP + T). This 'scrambling' chemistry of *b* ions recently attracted significant research attention and was investigated in numerous studies involving MS/MS,[5,10,11] infrared spectroscopy,[9,11e] ion mobility,[11d] hydrogen/deuterium exchange,[7d,11e,f] and theoretical techniques.[5,9]

a_n ions also carry useful information about peptide sequence. A recent large-scale statistical study of validated protonated peptide spectra indicated that these ions are *more* prevalent than their corresponding b_n ions[12] in many cases. In addition to providing confirmation of the presence of the corresponding b_n ions, consecutive a_n ions enable amino acid strings to be read out and the corresponding sequences to be determined. a_n ions can be formed either directly from protonated peptides[13] or from the corresponding b_n ions via elimination of CO[2,6a,b,8,13b,f] (28 mass-unit mass difference for singly charged ions). The latter reaction was previously studied by measuring the accompanying kinetic energy releases.[6a,b] These indicate that loss of CO occurs through an energetically high barrier, and the product energy level is below that of the transition structure (TS) involved. The experimentally recorded energetics were subsequently independently supported by quantum chemical calculations.[8a]

a_n ions are initially formed as immonium derivatives ($-HN^+{=}CHR_n$).[2,13a,b] a_1 ions as well as other immonium ions formed from internal amino acid residues are just simple protonated imine derivatives ($H_2N^+{=}CHR_i$). These ions are often formed in CID of protonated peptides, resulting in useful information on the peptide's amino acid composition.[13c−e] Larger imine a_n ions ($n \geq 2$) have more than one potential protonation site: the N-terminal amino group and also the imine group itself, $-C(O)-N{=}CHR_n$, where the C-terminal imine group is bound to a carbonyl group. Unlike normal amide bonds, which are partially conjugated, the C(O)−N bond adjacent to the imine group is a single bond, allowing free rotation around

(5) (a) Harrison, A. G.; Young, A. B.; Bleiholder, B.; Suhai, S.; Paizs, B. *J. Am. Chem. Soc.* **2006**, *128*, 10364. (b) Bleiholder, C.; Osburn, S.; Williams, T. D.; Suhai, S.; Van Stipdonk, M.; Harrison, A. G.; Paizs, B. *J. Am. Chem. Soc.* **2008**, *130*, 17774. (c) Bythell, B. J.; Knapp-Mohammady, M.; Paizs, B.; Harrison, A. G. *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 1352.

(6) (a) Yalcin, T.; Khouw, C.; Csizmadia, I. G.; Peterson, M. R.; Harrison, A. G. *J. Am. Soc. Mass Spectrom.* **1995**, *6*, 1165. (b) Yalcin, T.; Csizmadia, I. G.; Peterson, M. B.; Harrison, A. G. *J. Am. Soc. Mass Spectrom.* **1996**, *7*, 233. (c) Paizs, B.; Lendvay, G.; Vékey, K.; Suhai, S. *Rapid Commun. Mass Spectrom.* **1999**, *13*, 525. (d) Rodriquez, C. F.; Cunje, A.; Shoeib, T.; Chu, I. K.; Hopkinson, A. C.; Siu, K. W. M. *J. Am. Chem. Soc.* **2001**, *123*, 3006. (e) Polfer, N. C.; Oomens, J.; Suhai, S.; Paizs, B. *J. Am. Chem. Soc.* **2005**, *127*, 17154. (f) Polfer, N. C.; Oomens, J.; Suhai, S.; Paizs, B. *J. Am. Chem. Soc.* **2007**, *129*, 5887. (g) Chen, X.; Turecek, F. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 1941.

(7) (a) Bythell, B. J.; Erlekam, U.; Paizs, B.; Maître, P. J. *ChemPhysChem* **2009**, *10*, 883. (b) Yoon, S. H.; Chamot-Rooke, J.; Perkins, B. R.; Hilderbrand, A. E.; Poutsma, J. C.; Wysocki, V. H. *J. Am. Chem. Soc.* **2008**, *130*, 17644. (c) Oomens, J.; Young, S.; Molesworth, S.; van Stipdonk, M. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 334. (d) Bythell, B.; Somogyi, A.; Paizs, B. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 618.

(8) (a) Paizs, B.; Szlavik, Z.; Lendvay, G.; Vekey, K.; Suhai, S. *Rapid Commun. Mass Spectrom.* **2000**, *14*, 746. (b) Allen, J. M.; Racine, A. H.; Berman, A. M.; Johnson, J. S.; Bythell, B. J.; Paizs, B.; Glish, G. L. *J. Am. Soc. Mass Spectrom.* **2008**, *19*, 1764.

(9) Erlekam, U.; Bythell, B. J.; Scuderi, D.; Van Stipdonk, M.; Paizs, B.; Maitre, P. *J. Am. Chem. Soc.* **2009**, *131*, 11503.

(10) (a) Tang, X.-J.; Thibault, P.; Boyd, R. K. *Anal. Chem.* **1993**, *65*, 2824. (b) Tang, X.-J.; Boyd, R. K. *Rapid Commun. Mass Spectrom.* **1994**, *8*, 678.

(11) (a) Mouls, L.; Aubagnac, J. L.; Martinez, J.; Enjalbal, C. *J. Proteome Res.* **2007**, *6*, 1378. (b) Harrison, A. G. *J. Am. Soc. Mass Spectrom.* **2008**, *19*, 1776. (c) Molesworth, S.; Osburn, S.; Van Stipdonk, M. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 2174. (d) Riba-Garcia, I.; Giles, K.; Bateman, R. H.; Gaskell, S. J. *J. Am. Soc. Mass Spectrom.* **2008**, *19*, 609. (e) Chen, X.; Yu, L.; Steill, J. D.; Oomens, J.; Polfer, N. C. *J. Am. Chem. Soc.* **2009**, *131*, 18272. (f) Fattahi, A.; Zekavat, B.; Solouki, T. *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 358–369.

(12) Savitski, M. M.; Falth, M.; Eva Fung, Y. M.; Adams, C. M.; Zubarev, R. A. *J. Am. Soc. Mass Spectrom.* **2008**, *19*, 1755.

(13) (a) Vachet, R. W.; Ray, K. L.; Glish, G. L. *J. Am. Soc. Mass Spectrom.* **1998**, *9*, 341. (b) Ambihapathy, K.; Yalcin, T.; Leung, H.-W.; Harrison, A. G. *J. Mass Spectrom.* **1997**, *32*, 209. (c) Biemann, K. *Biomed. Environ. Mass Spectrom.* **1988**, *16*, 99. (d) Papayannopoulos, I. A. *Mass. Spectrom. Rev.* **1995**, *14*, 49. (e) Hohmann, L. J.; Eng, J. K.; Gemmill, A.; Klimek, J.; Vitek, O.; Reid, G. E.; Martin, D. B. *Anal. Chem.* **2008**, *80*, 5595. (f) Bythell, B. J.; Dain, R. P.; Curtice, S. S.; Oomens, J.; Steill, J. D.; Groenewold, G. S.; Paizs, B.; Van Stipdonk, M. *J. Phys. Chem. A* **2010**, *114*, 5076.

(14) (a) El Aribi, H.; Rodriquez, C. F.; Almeida, D. R. P.; Ling, Y.; Mak, W. N. V.; Hopkinson, A. C.; Siu, K. W. M. *J. Am. Chem. Soc.* **2003**, *125*, 9229. (b) El Aribi, H.; Orlova, G.; Rodriquez, C. F.; Almeida, D. R. P.; Hopkinson, A. C.; Siu, K. W. M. *J. Phys. Chem. B* **2004**, *108*, 18743. (c) Harrison, A. G.; Young, A. B.; Schnölzer, M.; Paizs, B. *Rapid Commun. Mass Spectrom.* **2004**, *18*, 1635. (d) Verkerk, U. H.; Siu, C.-K.; Steill, J. D.; El Aribi, H.; Zhao, J.; Rodriquez, C. F.; Oomens, J.; Hopkinson, A. C.; Michael Siu, K. W. M. *J. Phys. Chem. Lett.* **2010**, *1*, 868.

(15) (a) Vachet, R. W.; Bishop, B. M.; Erickson, B. W.; Glish, G. L. *J. Am. Chem. Soc.* **1997**, *119*, 5481. (b) Cooper, T.; Talaty, E.; Grove, J.; Suhai, S.; Paizs, B.; Van Stipdonk, M. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 1654. (c) Bythell, B.; Molesworth, S.; Osburn, S.; Cooper, T.; Paizs, B.; Van Stipdonk, M. *J. Am. Soc. Mass Spectrom.* **2008**, *19*, 1788. (d) Polfer, N. C.; Bohrer, B. C.; Plasencia, M. D.; Paizs, B.; Clemmer, D. E. *J. Phys. Chem. A* **2008**, *112*, 1286.

(16) (a) Prazeres, R.; Glotin, F.; Insa, C.; Jaroszynski, D. A.; Ortega, J. M. *Eur. Phys. J. D* **1998**, *3*, 87. (b) Mac Aleese, L.; Simon, A.; McMahon, T. B.; Ortega, J. M.; Scuderi, D.; Lemaire, J.; Maître, P. *Int. J. Mass. Spectrom.* **2006**, *249/250*, 14. (c) Lemaire, J.; Boissel, P.; Heninger, M.; Mauclaire, G.; Bellec, G.; Mestdagh, H.; Simon, A.; Caer, S. L.; Ortega, J. M.; Glotin, F.; Maitre, P. *Phys. Rev. Lett.* **2002**, *89*, 27300.

the C−N axis.[6f,8b] As a direct consequence, linear a_2 ions (formally $H_2N-CHR_1-CO-NH^+=CHR_2$) can undergo head-to-tail cyclization to form a 5-membered ring isomer containing a cis amide bond and a secondary amine.[14a] This reaction is initiated by nucleophilic attack of the N-terminal amino group on the carbon of the immonium group. a_2 ions most often fragment by elimination of CO, yielding the corresponding $H_2N^+=CHR_1$ or $H_2N^+=CHR_2$ fragments or a combination of the two.[14b,c] The proton affinities of the two neutral imines determine which is the more likely to be formed and detected. In a very recent paper[14d] Siu and co-workers showed that the a_2 ion (GG sequence) predominantly has a cyclic structure in a Fourier transform ion cyclotron resonance mass spectrometer after nozzle-skimmer dissociation.

a_3 peaks are rarely observed in the product ion spectra of protonated peptides[8b,12] despite both smaller (a_1 and a_2) and larger (a_n, $n > 3$) members of the a_n series being often detected. This phenomenon was explained[8b] by the reduced kinetic stability of the a_3 ion compared to that of a_2 and a_4. The a_4 ion of Leu-enkaphalin (YGGFL) has been extensively studied in recent years by several groups. Glish, Vachet, and co-workers studied[15a] the fragmentation reactions of this ion and discovered that elimination of ammonia leads to a rearranged sequence where the initially C-terminal C_α and the corresponding side chain ($-NH^+=CHR$) are transferred to the N terminus, leaving the initially penultimate amino acid residue exposed at the new C terminus in an oxazolone group. The detailed mechanism of this rearrangement reaction was recently studied[5b,15b,c] using theory and stable isotope labeling. The structure of the a_4 fragment of Leu-enkephalin was probed by IR spectroscopy, ion mobility spectrometry (IMS), and theoretical studies. Using IR spectroscopy and theory Polfer et al.[6f] demonstrated that this ion forms both linear and macrocyclic isomers. For the former, protonation can conceivably occur at the N-terminal amino and C-terminal imine groups; however, both experiment and theory indicate the presence of the amino-protonated form. The macrocyclic a_4 isomer can be formed[5b] from the imine-protonated linear structure by nucleophilic attack of the N-terminal amino nitrogen on the carbon center of the immonium group. This macrocyclic form is as energetically favorable as the best linear structures for the a_4 fragment of YGGFL.[6f] A recent computational study[5b] of the a_5 ions with the YAGFL, AGFLY, and LYAGF sequences indicates the macrocyclic form is energetically more favored than the best linear structures for each of the investigated sequence isomers. Furthermore, the barrier to cyclization for the a_5 ion with the YAGFL sequence is 9.6 kcal mol^{-1} as compared to the imine-protonated linear species, thereby enabling facile formation of the macrocyclic isomer under normal low-energy CID conditions.[5b]

Additionally, the structure of the a_4 fragment of Leu-enkaphalin was investigated further by using ion mobility spectrometry (IMS) in the Clemmer group.[15d] Three major features were observed in the ion mobility spectrum; the one with the lowest collision cross-section was assigned as the most compact macrocyclic structure, and the peak associated with the highest cross-section was assigned as the N-terminal-protonated linear form. However, the third feature which represents by far the most abundant isomer(s) could not be assigned to any of the known a_n isomers. This clearly indicates that further work is necessary to fully understand the structure and reactivity of a_n fragments formed in CID of protonated peptides. In an effort to address these issues, we investigated the structure and reactivity of the a_n ($n = 2-4$) ions formed from protonated oligoglycines and the a_4 ion from protonated AAAAA by means of gas-phase 'action' IR spectroscopy combined with modeling and density functional theory calculations considering a larger pool of structures than ever before.

Experiments and Computations

Mass Spectrometry and IR-MPD Spectroscopy. GGG, GGGG, GGGGG, and AAAAA (Sigma Aldrich, Hamburg, Germany) were dissolved in $CH_3OH:H_2O$ = 1:1 with 2% acetic acid in a concentration range of 50−80 μmol/L and sprayed with conventional ESI conditions into a 3D quadrupole ion trap mass spectrometer (Bruker Esquire 3000+). Infrared spectroscopy was carried out using the free electron laser (FEL) at CLIO[16a] coupled to the ion trap instrument.[16b] Following mass selection, the isolated singly charged precursor ions were subjected to CID (MS^n) to form a_2 ions from protonated GGG, a_3 ions from protonated GGGG, and a_4 ions from protonated GGGGG and AAAAA. Helium was used as the collision gas. The a_n ions were then mass selected and irradiated with the FEL IR beam for 200 ms. It is also important to stress that thermalization of the mass-selected a_n ions prior to their irradiation is ensured by multiple collisions with helium.[16b] Upon resonant vibrational excitation, dissociation of a_2 ions was monitored via the a_2-CO peak. Dissociation of a_3 ion was monitored via the a_3*, a_3*-CO, b_2, and a_2 peaks. Dissociation of a_4 of protonated GGGGG was monitored via the a_3*, b_3, a_3, b_2, a_2, and m/z 75 peak, while the a_4 of protonated AAAAA was monitored via the a_4°, a_4°-CO, a_4*$^\circ$-CO, a_4*$^\circ$-2CO, a_3°-CO, and m/z 171 peaks. The abundances of these multiple photofragments and their corresponding a_n ion precursors were recorded as a function of the IR wavelength in order to derive the IR action spectra where the IRMPD efficiency[16c] is plotted against the photon energy.

Computational Details. A recently developed conformational search engine[2,5,6e,f,9,17] devised to deal with protonated peptides was used to scan the potential energy surfaces (PESs) of the a_n ($n = 2-4$) ions composed from glycine residues and the a_4 ion of protonated AAAAA. These calculations began with molecular dynamics simulations using the Discover program (Biosym Technologies, San Diego, CA) in conjunction with the AMBER force field,[18] modified in house in order to enable the study of structures with oxygen- and nitrogen-protonated amide bonds and protonated or neutral imine groups. During the dynamics calculations we used simulated annealing techniques to produce candidate structures for further refinement, applying full geometry optimization using the AMBER force field. These optimized structures were analyzed by a conformer family search program developed in Heidelberg. This program groups optimized structures into families for which the most important characteristic torsion angles of the molecule are similar. The most stable species in the families were then fully optimized at the PM3, HF/3-21G, B3LYP/6-31G(d), and finally B3LYP/6-31+G(d,p) levels. The conformer families were regenerated at each level, and only structurally nondegenerate conformers were recomputed at the next level to prevent wasting computer time (i.e., only one of N identical structures is recomputed at the next level). Such series of calculations were performed for the various isomers discussed in this paper. Transition structures (TSs) for forming/opening the macrocyclic isomers and proton transfer reactions were probed in a similar fashion, and all optimized

(17) (a) Wyttenbach, T.; Paizs, B.; Barran, P.; Breci, L.; Liu, D.; Suhai, S.; Wysocki, V. H.; Bowers, M. T. *J. Am. Chem. Soc.* **2003**, *125*, 13768. (b) Paizs, B.; Suhai, S.; Hargittai, B.; Hruby, V. J.; Somogyi, A. *Int. J. Mass Spectrom.* **2002**, *219*, 203. (c) Paizs, B.; Suhai, S. *J. Am. Soc. Mass Spectrom.* **2004**, *15*, 103. (d) Paizs, B.; Suhai, S. *Rapid Commun. Mass Spectrom.* **2002**, *16*, 375. (e) Bythell, B. J.; Barofsky, D. F.; Pingitore, F.; Polce, M. J.; Wang, P.; Wesdemiotis, C.; Paizs, B. *J. Am. Soc. Mass Spectrom.* **2007**, *7*, 1291. (f) Bythell, B. J.; Suhai, S.; Somogyi, A.; Paizs, B. *J. Am. Chem. Soc.* **2009**, *131*, 14057.

(18) Case, D. A.; et al. *AMBER 99*; University of California: San Francisco, 1999.

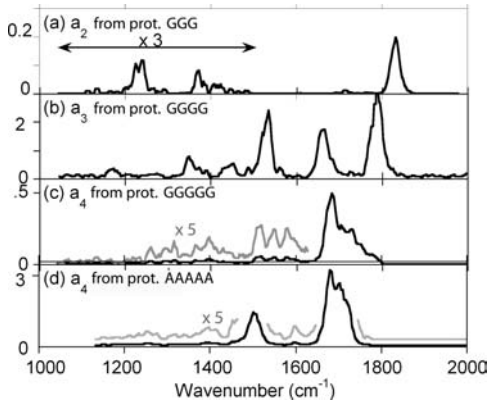


Figure 1. Experimental infrared spectra of (a) the a_2 ion of protonated GGG, (b) the a_3 ion of protonated GGGG, (c) the a_4 ion of protonated GGGGG, and (d) the a_4 ion of protonated AAAAA.

structures were examined by vibrational analysis and then submitted to intrinsic reaction coordinate (IRC) calculations to determine which minima they connect. The total energies of the various optimized structures are presented in Tables S1−4 (Supporting Information). Zero-point energy corrected relative energies were computed at the B3LYP/6-31+G(d,p) level. The theoretical IR spectra were determined using harmonic frequencies scaled by a factor of 0.98. The calculated stick spectra were convoluted assuming a Lorentizan profile with a 20 cm^{-1} full-width at half-maximum. The Gaussian set of programs[19] was used for all ab initio and DFT calculations. Orientation-averaged projection cross-sections were obtained for selected structures of the a_4 ions of protonated GGGGG and AAAAA using the Sigma program[20] developed by the Bowers group.

Results and Discussion

Figure 1 displays the experimental IR spectra of the investigated a_2, a_3, and a_4 ions, composed exclusively of glycine residues as well as the spectrum of the a_4 ions composed of alanine residues. In the 1000−2000 cm^{-1} so-called "fingerprint" range explored here, IR bands are characteristic of different functional groups of the molecular ions. This is particularly important in the high-frequency range of CO stretches, which have been shown to be highly diagnostic for structure identification of the b_n ions.[6,7] In this respect, it is important to notice that none of the spectra contain peaks above 1850 cm^{-1}.

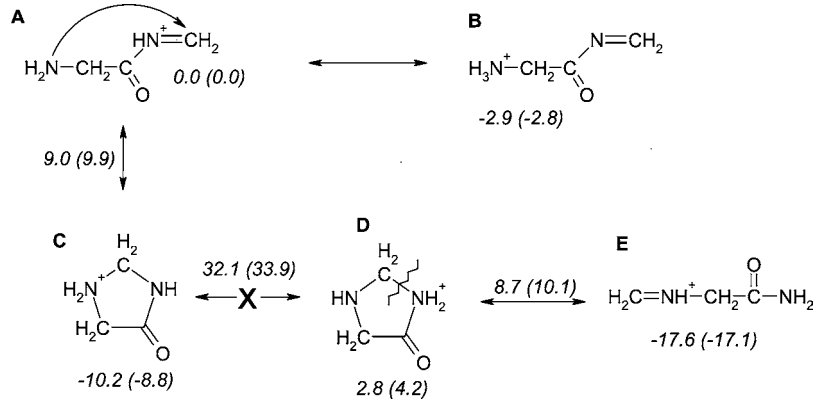
A first look at these spectra clearly indicates the presence of substantially different structures for the above a_n fragments. Overall, the IRMPD process was found to be more efficient for the a_3 ions. Anecdotally, this finding ties in with the general instability of the a_3 ions reported in mass spectrometers.[8b,12] The IRMPD effiency is an order of magnitude smaller for the a_2 ions. In this case, the corresponding spectrum (Figure 1a) has a dominant peak at ∼1830 cm^{-1} and other much less prominent features in the 1200−1500 cm^{-1} range. In sharp contrast, the a_3 ion spectrum displays several well-resolved features (Figure 1b). This is likely to be in part related to the fact that the IRMPD process is more facile in this case, thereby allowing for a more sensitive spectroscopic probe. The a_4 (composed of glycine residues) spectrum has a broad band with a maximum at ∼1680 cm^{-1} and a shoulder on its blue side at 1730 cm^{-1} which in turn has a shoulder that extends to ∼1785 cm^{-1}. This group of bands is red shifted as compared to the major peak for the a_2 ion and the largest peak of the a_3 ion. Analogous to the glycyl form(s), the spectrum of the a_4 ions of protonated AAAAA has a broad band with a maximum at ∼1680 cm^{-1} and a shoulder on its blue side at ∼1730 cm^{-1} terminating at ∼1740 cm^{-1}. In contrast to the glycl form(s), the alanyl a_4 spectrum has a large and broad absorption at ∼1510 cm^{-1} with a lower frequency shoulder at ∼1480 cm^{-1} and a smaller absorption at ∼1600 cm^{-1}. Neither the glycyl a_3 ion nor either of the a_4 spectra show peaks above 1800 cm^{-1}. In the following we will assign each individual spectrum with the help of detailed modeling and density functional theory calculations. In doing this both linear and various cyclic isomers (some of which are newly conceived possibilities) will be considered in various protonated forms along with cyclization and ring-opening reactions connecting these major isomers.

1. Structure and IR Spectroscopy of the a_2 Ion of Protonated GGG. The isomeric structures and reactions considered for the a_2 ion of protonated GGG are presented in Scheme 1. The linear isomer $[H_2N{-}CH_2{-}CO{-}N{=}CH_2 + H]^+$ has two competitive protonation sites: the C-terminal imine nitrogen and the N-terminal amino group (Figure 2A and 2B, respectively, Scheme 1 and Figure S1, Supporting Information). The former is generated directly when the a_2 ion is formed from the b_2 ion,[8] so this will be our reference structure for the relative energies. Thus, negative relative energies will represent effective stabilization of the initially formed structure by structural reorganization (i.e., proton transfer(s) and/or cyclization). Our calculations indicate the N-terminal amino-protonated linear structure (B in Figures 2 and S1, Supporting Information, and in Scheme 1) is 2.9 kcal mol^{-1} more energetically favored (Table S1, Supporting Information) than the initially formed imine-protonated form (A, Figure 2).

The imine-protonated form (A in Figures 2 and Scheme 1) can undergo cyclization initiated by nucleophilic attack of the N-terminal amino nitrogen on the carbon of the protonated imine.[14a] The barrier to this reaction (A ↔ C in Scheme 1) was calculated by Siu and co-workers[14d] at 8.9 kcal mol^{-1} (measured relative to the imine-protonated from, A) at the B3LYP/6-311++G(d,p) level of theory; our B3LYP/6-31+G(d,p) value is almost identical at 9.0 kcal mol^{-1}. The resulting 5-membered ring (structure C in Figures 2 and S1, Supporting Information) that incorporates a cis amide bond and a protonated secondary amine is 10.2 kcal mol^{-1} energetically more favored (Table S1, Supporting Information) than the C-terminal imine-protonated linear species. This 5-membered ring isomer can in principle undergo a proton transfer to form the amide nitrogen-protonated form (D in Figures 2 and S1) that is at 2.8 kcal mol^{-1} relative energy. However, the barrier to this proton transfer reaction (TS C ↔ D in Figure S1, Supporting Information) is high at 32.1 kcal mol^{-1} relative energy, which makes formation of the protonated amide nitrogen species unlikely. Should such species be formed, these can undergo ring opening at the $-CH_2-NH_2^+-$ bond (TS D ↔ E in Figure S1, Supporting Information), forming a rearranged $H_2C{=}NH^+{-}CH_2{-}CO{-}NH_2$ linear isomer (E in Figures 2 and S1, Supporting Information) that features imine and amide groups at its N and C termini, respectively. This ring-opening TS requires at least 8.7 kcal mol^{-1} relative energy (TS (D ↔ E)) to initiate. This reaction, in principle, generates a rearranged linear a_2 isomer (E in Figures 2 and S1, Supporting Information) that is the energetically most favorable

(19) Frisch, M. J.; et al. *Gaussian 03*, revision C.02; Gaussian Inc.: Pittsburgh, PA, 2004.
(20) (a) von Helden, G.; Hsu, M.-T.; Gotts, N.; Bowers, M. T. *J. Phys. Chem.* **1993**, *97*, 8182–8192. (b) Wyttenbach, T.; von Helden, G.; Batka, J. J.; Carlat, D.; Bowers, M. T. *J. Am. Soc. Mass Spectrom.* **1997**, *8*, 275–282. (c) Wyttenbach, T.; Witt, M.; Bowers, M. T. *J. Am. Chem. Soc.* **2000**, *122*, 3458–3464.

Scheme 1. Various a_2 Ion Structures of Protonated GGG and Their Reactions[a]



[a] (A) C-terminal imine-protonated linear form, (B) N-terminal amino-protonated linear form, (C) 5-membered ring isomer protonated at the secondary amine, (D) 5-membered ring isomer protonated at the amide nitrogen, and (E) rearranged linear form protonated at the N-terminal imine group. Relative ZPE-corrected energies (Gibbs free energies at 298 K) in kcal mol^{-1} are given in italics.

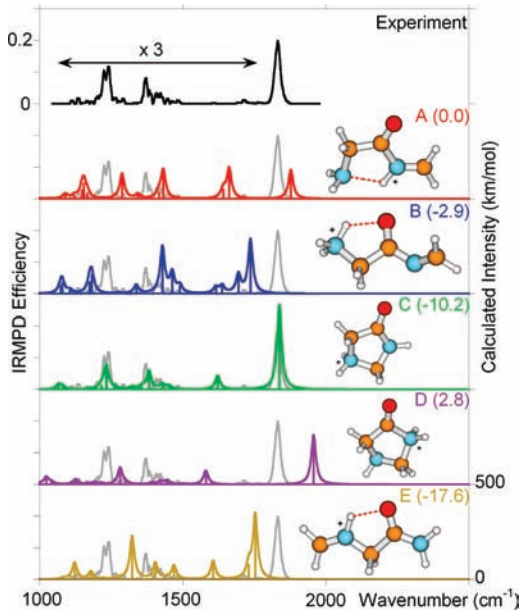


Figure 2. Experimental and calculated infrared spectra of the a_2 ion of protonated GGG. Theoretical spectra are presented for the C-terminal imine-protonated linear form (A), the N-terminal amino-protonated linear form (B), the 5-membered ring isomer protonated at the secondary imine (C), the 5-membered ring isomer protonated at the amide nitrogen (D), and the rearranged linear form protonated at the N-terminal imine (E). The computed relative energies (kcal mol^{-1}) are presented in parentheses for each isomer.

structure on the PES of the a_2 ion at −17.6 kcal mol^{-1} relative energy (Table S1, Supporting Information).

The experimental IR spectrum of the a_2 ion is shown in Figure 2 along with theoretical spectra computed for structures A−E. The experimental spectrum is dominated by a strong IRMPD feature centered at ∼1830 cm^{-1}. Other features were observed at 1233 and 1370 cm^{-1} but are much weaker. This is in contrast with the recently published[14d] IRMPD spectrum of the a_2 ion of GGG, which has the same features but the intensities of the bands observed near 1230 and 1370 cm^{-1} are larger than in the present case. We recorded multiple scans in the requisite wavenumber range, but the IRMPD efficiency was always very low in this region (Figure 1). The second difference between the IRMPD spectrum of the a_2 ion of GGG of Verkerk et al.[14d] and the present one is associated with the shape of the IRMPD band. In the present case, the strong feature is relatively sharp, and the observed bandwidth (fwhm ∼25 cm^{-1}) is the width expected when mass-selected ions are thermalized through multiple collisions with helium.[16b] On the other hand, broader features were observed by Verkerk et al..[14d] which may indicate that the internal energy of the ions is larger than in our experiment. This would also be consistent with the fact that the IRMPD efficiency in the 1000−1700 cm^{-1} range in our case is much lower than that observed by Verkerk et al.[14d] The internal energy of the ion is likely to be a critical issue for the IR-induced fragmentation of a small ion such as the a_2 ion for which intramolecular vibrational energy redistribution is likely to be limited by the low density of vibrational states after absorption of the first photons.[21] Another clear difference between the two experiments is the pressure regime in the two instruments. The pressure is on the order of 10^5 times lower in the Verkerk et al.[14d] experiment in the FT-ICR than in our experiments, meaning that collisional relaxation/cooling of the ions via collisions with the background gas will be correspondingly more frequent in our experiment. However, it should be explicitly noted that it is unlikely that significant energy damping will occur in the ion trap experiment. This is because the CLIO free electron laser delivers ∼8 μs long trains of picosecond pulses which are spaced by 16 ns. Multiple IRMPD experiments performed with both an FTICR (low-pressure regime) or an ion-trap (high-pressure regime) have shown that the IRMPD efficiency scales linearly with the irradiation time, thereby indicating that the whole IRMPD process (i.e., multiple photon absorption and fragmentation) occurs within the time scale of a macropulse. Thus, the collision frequency which should be on the order of 10^4 Hz in the ion trap is too low to enable collisional relaxation to compete with the IR heating of the ions. This type of process however is likely to be a relevant problem when using lower fluence lasers in the ion trap, i.e., OPO/OPA lasers delivering nanosecond pulses at a low repetition rate (typically 10 Hz).

Overall, as also observed by Verkerk et al.,[14d] the computed IR absorption spectrum of the C structure closely fits the experimental spectrum. The structural assignment is driven by assessing the C=O stretch vibrations which appear at characteristic frequencies for the isomers studied. In our case, only one sharp band was observed in the CO stretch region, suggesting that essentially one isomer is populated under our

(21) Oomens, J.; Sartakov, B. G.; Meijer, G.; Von Helden, G. *Int. J. Mass Spectrom.* **2006**, *254* (1−2), 1–19.

experimental conditions. The presence of the linear isomers A and B can be excluded since no experimental absorption is observed at 1876 and 1730 cm^{-1}, respectively. Similarly, the presence of the amide nitrogen-protonated cyclic D and rearranged linear E isomers can be excluded since no experimental absorption is observed at 1760 and 1970 cm^{-1}, respectively. On the other hand, the large absorption at 1830 cm^{-1} is nicely matched by the major peak belonging to the C═O stretch at 1837 cm^{-1} of the theoretical spectrum of the secondary amine-protonated cyclic C isomer. Furthermore, the weak features observed at 1233 and 1370 cm^{-1} nicely match the positions of IR active bands of structure C. Consequently, either all or the vast majority of the a_2 ion population studied under our applied experimental conditions is composed of structure C.

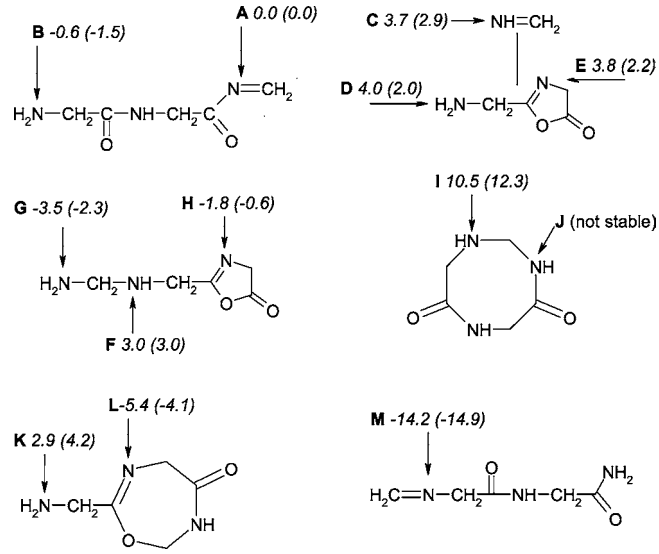
As already stated above, the IRMPD spectrum of the a_2 ion of GGG recently published by Verkerk et al.[14d] is slightly different from the one presented here. As in the present case, the strongest absorption was observed near 1830 cm^{-1}. Nevertheless, while this main feature has a relatively sharp profile in the present case, Verkerk et al.[14d] observed a broader and asymmetric profile with a shoulder at 1727 cm^{-1} on the red side of the strong band reported at 1828 cm^{-1}. Since none of the individual structures has an IR cross-section fitting with the broad and asymmetric band,[14d] it was proposed that the experimental IRMPD spectrum of the a_2 ion of GGG recently published by Verkerk et al.[14d] was the result of the combination of the IRMPD spectrum of several structures, isomer C being the most abundant (55%).

The above analysis of our experimental data suggests that isomer C is predominantly formed under our experimental conditions, which is in apparent contradiction with the energetics computed for structures A−E. These suggest that the rearranged linear form (structure E) is much more favored than the cyclic C structure. However, this contradiction can be resolved by considering the barrier which must be overcome in order to form structure E (Figure 2). The C → D proton transfer reaction has a threshold energy of 32.1 kcal mol^{-1}, so this reaction is very unlikely to occur, thereby preventing formation of structures D and E. It is worth noting that the majority of the a_2 ions appear to have enough energy to form structure C from B via crossing the cyclization TS at 9.0 kcal mol^{-1}. On the other hand, the TS of the proton transfer reaction to form D from C requires at least 32.1 kcal mol^{-1} relative energy, and the vast majority of the a_2 ion populations does not have enough energy to overcome this barrier.

2. Structure and IR Spectroscopy of the a_3 Fragment of Protonated GGGG. The isomeric structures and reactions considered for the a_3 ion of protonated GGGG are presented in Chart 1 and Scheme 2, respectively. The linear isomer of protonated $H_2N-CH_2-CO-NH-CH_2-CO-N=CH_2$ has two competitive sites of protonation: the N-terminal amino and the C-terminal imine groups. Similar to the discussion on the a_2 ion above, the initially formed imine-protonated species (A in Figures 3 and S2) will be used as the reference structure when comparing the relative energies of the various a_3 isomers. Our theoretical data indicate the N-terminal amino-protonated linear structure (B in Figures 3 and S2) is 0.6 kcal mol^{-1} more energetically favorable (Table S2, Supporting Information) than the imine-protonated form A.

Structure A can potentially undergo a variety of reactions (Scheme 2). Nucleophilic attack of the N-terminal carbonyl oxygen[15b] on the C-terminal carbonyl carbon (TS at 14.2 kcal mol^{-1} relative energy, Table S2, Supporting Information) leads

Chart 1. Various a_3 (GGG sequence) Ion Structures[a]

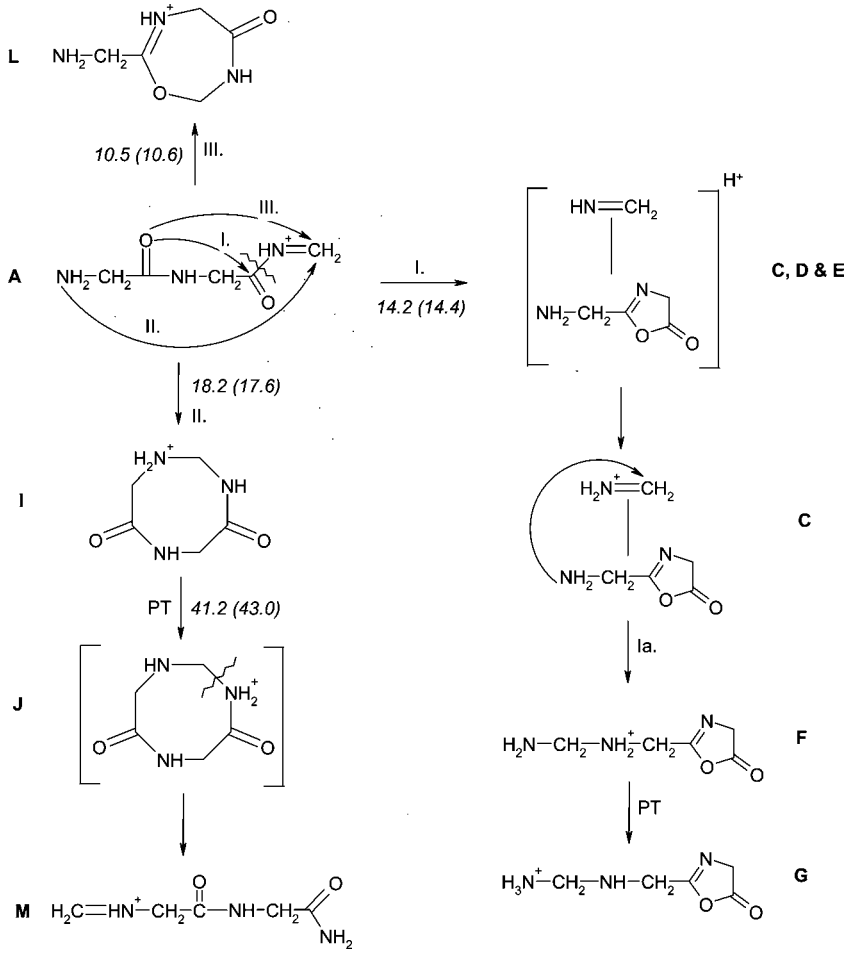


[a] (A and B) Linear form protonated at the imine and N-terminal amino groups, respectively; (C, D, and E) proton-bound dimer of HN═CH_2 and 2-(aminomethyl)-5-oxazolone protonated at the imine nitrogen, amino group of the oxazolone, and ring nitrogen of the oxazolone, respectively; (F, G, and H) oxazolone-terminated re-associated linear form protonated at the secondary amine, N-terminal amino, and ring nitrogen, respectively; (I and J) macrocyclic (8-membered ring) isomer protonated at the secondary amine or amide nitrogen, respectively; (K and L) 7-membered ring isomer protonated at the N-terminal amine or ring imine nitrogen; and (M) rearranged linear form protonated at the N-terminal imine group. The respective protonation sites are indicated by arrows. Relative ZPE-corrected energies (Gibbs free energies at 298 K) in kcal mol^{-1} are given in italics.

to formation of the 5-membered oxazolone ring by concerted cleavage of the $CO-NH^+$ bond of the protonated imine (Reaction I, Scheme 2). This reaction produces a proton-bound dimer (PBD) of HN═CH_2 and 2-(aminomethyl)-5-oxazolone which can exist in three competitive protonation forms: the imine nitrogen (structure C in Chart 2) and the amino and ring nitrogens of the oxazolone (structures D and E in Chart 2, respectively). These PBD structures are energetically less favored than the linear A species (relative energies at 3.7, 4.0, and 3.8 kcal mol^{-1}, respectively, Chart 1, Table S2, Supporting Information). The PBD structure C can undergo a reassociation reaction[15b] where the amino nitrogen (N terminus) of the oxazolone attacks the carbon center of the protonated imine (reaction Ia, Scheme 2). This leads to a new a_3 isomer (F, Scheme 2) that contains the $H_2N-CH_2-NH_2^+-$ group at the N terminus and an oxazolone group at the C terminus. Two other competitive protonation sites (Chart 1) are available for this isomer: the N-terminal primary amine and the oxazolone ring nitrogen (structures G and H, respectively, Scheme 2). Our calculations indicate (Table S2, Supporting Information) that the N-terminal amino and oxazolone nitrogen-protonated structures (G and H) are energetically more favorable than the imine-protonated linear reference structure (A, Scheme 2, Chart 1).

Alternatively, cyclization initiated by nucleophilic attack (reaction II in Scheme 2) of the N-terminal amino nitrogen of structure A on the carbon of the protonated imine can lead to an 8-membered ring ('macrocyclic') isomer. This 8-membered ring accommodates two amide bonds, the first of which must exist in the trans amide bond configuration, while the other, which is generated by ring formation, can be either cis or trans. The presence of the trans amide bond introduces significant ring strain, and the resulting structure (I in Chart 1, Scheme 2) is

Scheme 2. Various a_3 Ion Structures (GGG sequence) Derived from Rearrangement Reactions Starting from Structure A Which Is Generated Directly From the b_3 Ion[a]



[a] Nucleophilic attack (reaction I) of the N-terminal carbonyl oxygen at the C-terminal carbonyl results in cleavage of the CO−NH^+ bond and formation of the PBD of $HN{=}CH_2$ and 2-(aminomethyl)-5-oxazolone (structures C, D, and E). Nucleophilic attack (reaction Ia) of the N-terminal amino nitrogen of the oxazolone on the carbon center of the protonated imine in C leads to F from which G and H can be formed by intramolecular proton transfer. Nucleophilic attack (reaction II) of the N-terminal amino nitrogen on the carbon of the protonated imine group can lead to the macrocyclic (8-membered ring) isomer. After proton transfer to the just-formed amide bond, ring opening occurs spontaneously to form the rearranged linear structure M. Nucleophilic attack (reaction III) of the N-terminal carbonyl oxygen at the C-terminal carbon leads to the 7-membered ring isomer L. Relative ZPE-corrected energies (Gibbs free energies at 298 K) in kcal mol^{-1} are given in italics.

not energetically favorable (relative energy at 10.5 kcal mol^{-1}). Furthermore, the cyclization TS to form structure I is at 18.2 kcal mol^{-1} relative energy, which is much less energetically favorable than the corresponding cyclization TS for the a_2 ion (Scheme 1). The protonation energetics of the nitrogen of the amide bond introduced by ring formation (structure J in Chart 1) were also tested. Despite the many attempts made, no such stable species was found on the a_3 ion PES. Typically, the geometry optimizations of such structures lead to opening of the ring at the $-CH_2-NH_2^+-$ bond and formation of a rearranged linear structure ($H_2C{=}NH^+-CH_2-CO-NH-CH_2-CO-NH_2$, M in Chart 1) that features imine and amide groups at its N and C termini, respectively. The J ↔ I proton transfer TS was located at 41.2 kcal mol^{-1} relative energy (Scheme 2, Figure S2, Supporting Information). The amide nitrogen-protonated species on the product wing of this proton transfer TS is not itself stable, and IRC calculations indicate immediate reorganization of the structure to form isomer M after leaving the TS region of the potential energy surface. Structure M is the energetically most favorable candidate structure on the PES of the a_3 ion with a relative energy of −14.2 kcal mol^{-1} (Chart 1, Table S2, Supporting Information).

A third type of reaction is possible in principle. This reaction is initiated by nucleophilic attack (reaction III in Scheme 2) of the N-terminal carbonyl oxygen on the carbon center of the protonated imine of structure A. This generates an isomer with a 7-membered ring which contains a cis amide bond and a protonated imine group (from the first amide group). This class of structure has two competitive protonation sites: the ring imine nitrogen formed directly in the reaction and the N-terminal amino group (structures L and K, respectively). Our calculations indicate the former is the more energetically favorable (relative energy at −5.4 kcal mol^{-1}). In fact, structure L is the second most energetically favorable species on the PES of the a_3 ion; only the rearranged linear structure M (see preceding text) is more so. Furthermore, isomer L can be readily formed from structure A by crossing a comparatively low barrier of 10.5 kcal mol^{-1} height (Scheme 2, Table S2 and Figure S2, Supporting Information).

The experimental IR spectrum of the a_3 ion is shown in Figure 3 along with theoretical spectra computed for structures A, B, C, G, I, L, and M. Theoretical spectra for all other structures were computed and compared to the experimentally recorded spectrum, but none of these showed a good fit. For brevity, only

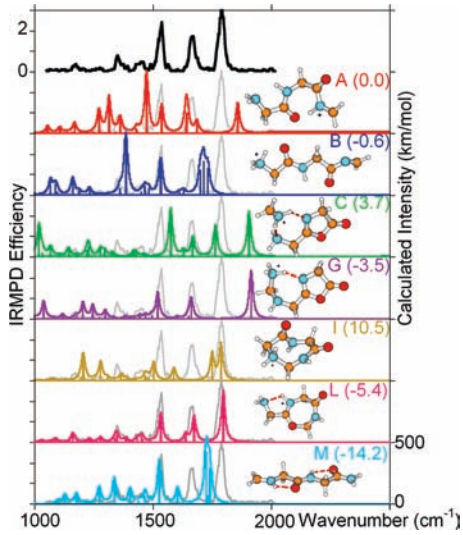


Figure 3. Experimental and calculated infrared spectra of the a_3 ion of protonated GGGG. Theoretical spectra are presented for the C-terminal imine-protonated linear form (structure A); the N-terminal amino-protonated linear form (structure B); the complex of $H_2^+N{=}CH_2$ and 2-(aminomethyl)-5-oxazolone (structure C); the reassociated, oxazolone-terminated linear form protonated at the N-terminal amine (structure G); the macrocyclic (8-membered ring) isomer protonated at the secondary amine (structure I); the 7-membered ring isomer protonated at the ring imine (structure L); and the rearranged linear form protonated at the N-terminal imine (structure M). The computed relative energies (kcal mol^{-1}) are presented in parentheses for each isomer.

the energetically most favored species are discussed in the C−D−E, F−G−H, and K−L families. The experimental spectrum contains three major peaks in the higher frequency region, and three less intense features are also observed at ∼1170, 1350, and 1445 cm^{-1}. The three most intense features are observed at ∼1530, ∼1665, and ∼1790 cm^{-1}, and their bandwidth (fwhm) is on the order of ∼25−30 cm^{-1}, as in the case of the a_2 ion.

The C═O stretching vibration which appears in the higher frequency range explored here is highly diagnostic since all the calculated spectra are characterized by a strong C═O absorption. Inspection of Figure 3 also reveals that the calculated IR absorption spectrum of the 7-membered ring isomer L fits the experimental spectrum for the three most intense bands observed at ∼1530, ∼1665, and ∼1790 cm^{-1}. An assignment of these three main experimental features can thus be proposed. The band observed at ∼1790 cm^{-1} can be assigned to the C═O stretching mode, which has a calculated frequency at 1798 cm^{-1} for structure L. The bands at ∼1530 and ∼1665 cm^{-1} can be assigned to the N−H bending and C═N stretching modes of the C═N$^+$H motif (calculated frequencies are 1533 and 1674 cm^{-1}).

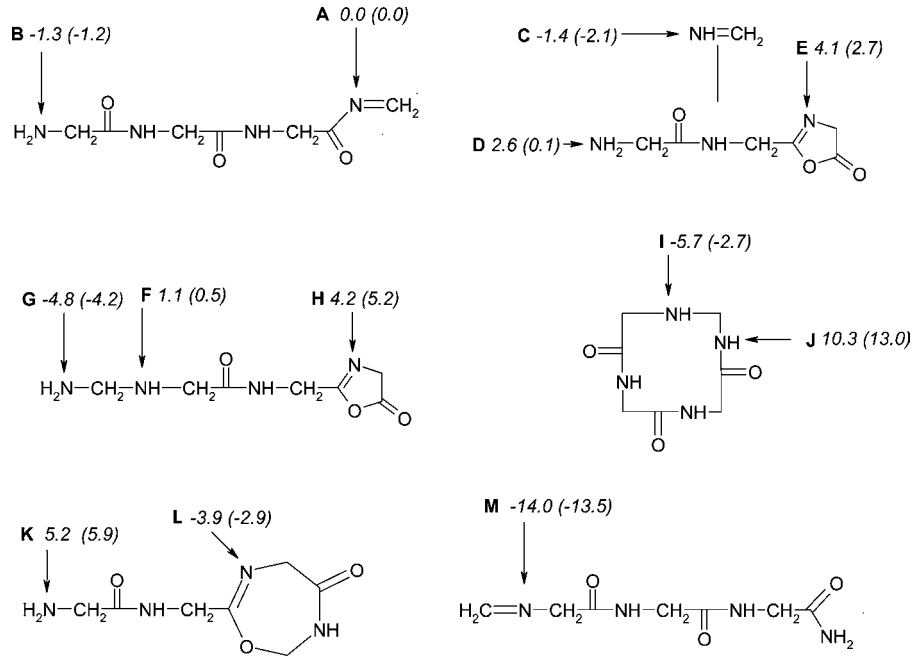
Although it is difficult to definitively exclude the presence of small populations of other structures, the relatively sharp character of the three most intense IRMPD bands strongly suggests that the vast majority of the a_3 ion population studied under the applied experimental conditions is composed from structure L. The linear isomers A and B (Chart 1) can be excluded for two reasons. First, their calculated spectra significantly differ with the experimental spectrum in the C═X (X = O, N) stretching region where structure A has a band at 1858 cm^{-1} and structure B three closely spaced, strongly IR active stretches, giving rise to a strong and broad IR cross-section peaking at 1714 cm^{-1}. Second, the predicted spectra of these two linear structures have a strong absorption in the lower frequency range associated with the N$^+$H bending mode (predicted at 1472 and 1387 cm^{-1} for A and B, respectively), whereas no strong IRMPD feature was detected in this spectral range. The theoretical spectra of the proton-bound dimer structures (C, D, and E) and also of structure G show strong absorption at wavenumbers above 1800 cm^{-1} corresponding to the characteristic oxazolone C═O stretching mode.[7a−c] Since no peak is observed in this range experimentally, the presence of these isomers can be excluded. The presence of the secondary amine-protonated cyclic structure I, which is not energetically favorable, can also be excluded for spectroscopic reasons considering in particular the spectral range above 1500 cm^{-1}. While three strong features are observed experimentally, the spectrum of structure I has only a broad band due to two nearly degenerate C═O stretching modes (1750 and 1787 cm^{-1}). Finally, and most importantly, the rearranged linear M isomers, which are predicted to be the a_3 ion structure of lowest energy on the PES, can also be excluded since the two C═O stretching frequencies are predicted to be very close (1725 and 1741 cm^{-1}), giving rise to a broad absorption feature, while two relatively sharp bands were observed in this spectral range at ∼1665 and ∼1790 cm^{-1}.

On the basis of the experimental and energetic considerations, it is clear that the vast majority of the a_3 ion population studied under the applied experimental conditions is composed from structure L. This observation indicates that the energetically most favored rearranged linear M structure, which is much more so than structure L (relative energies at −14.2 versus −5.4 kcal mol^{-1}), is not formed in the mass spectrometer. One can rationalize this by considering the relatively high energy of the macrocyclic structure I (10.5 kcal mol^{-1}) and the corresponding ring-closing TS (18.2 kcal mol^{-1}). Additionally, the very high energy barrier (TS at 41.2 kcal mol^{-1} relative energy) to the proton transfer reaction necessary to form the unstable amide nitrogen-protonated J structure necessary to form structure M is likely to be rate limiting. These combined effects prevent formation of structure M (reaction II of Scheme 2). The observation of structure L is thus consistent with the theoretical predictions since it is lower in energy than structure A and is the kinetically most favored product starting from structure A, with a single transition state (A ↔ L) at 10.5 kcal/mol.

3. Structure and IR Spectroscopy of the a_4 Fragment of Protonated GGGGG. The isomeric structures and reactions considered for the a_4 ion of protonated GGGGG are presented in Chart 2 and Scheme 3, respectively. The $[H_2N{-}CH_2{-}CO{-}(NH{-}CH_2{-}CO)_2{-}N{=}CH_2 + H]^+$ linear isomer has two competitive protonation sites: the N-terminal amino and the C-terminal imine nitrogens. Once again we will use the initially formed imine-protonated species (structure A in Figures 4 and S3, Supporting Information) as the reference structure. Our theoretical data indicate the N-terminal amino-protonated linear structure (B in Chart 2) is 1.3 kcal mol^{-1} more energetically favored (Table S3, Supporting Information) than the imine-protonated A form.

Similar to the case of the a_3 ion, the imine-protonated form of the a_4 ion can undergo a variety of reactions (Scheme 3). Nucleophilic attack of the N-terminally adjacent carbonyl oxygen[15b] on the C-terminal carbonyl leads to formation of the 5-membered oxazolone ring, cleavage of the CO−NH$^+$ bond, and expulsion of the C-terminal imine (reaction I, Scheme 3). The barrier to this reaction is 15.1 kcal mol^{-1} (Scheme 3, Table S3, Supporting Information). The resulting postreaction complex

Chart 2. Various Structures for the a_4 Ion of Protonated GGGGG[a]



[a] (A and B) Linear form protonated at the imine and N-terminal amino groups, respectively; (C, D, and E) proton-bound dimer of $NH{=}CH_2$ and 2-(glycylaminomethyl)-5-oxazolone protonated at the imine nitrogen, amino group of the oxazolone, and ring nitrogen of the oxazolone, respectively; (F, G, and H) oxazolone-terminated re-associated linear form protonated at the secondary amine, N-terminal amino, and ring nitrogen, respectively; (I and J) macrocyclic (11-membered ring) isomer protonated at the secondary amine and amide nitrogen, respectively; (K and L) 7-membered ring isomer protonated at the N-terminal amine or ring imine nitrogen; and (M) rearranged linear form protonated at the N-terminal imine group. The respective protonation sites are indicated by arrows. Relative ZPE-corrected energies (Gibbs free energies at 298 K) in kcal mol^{-1} are given in italics.

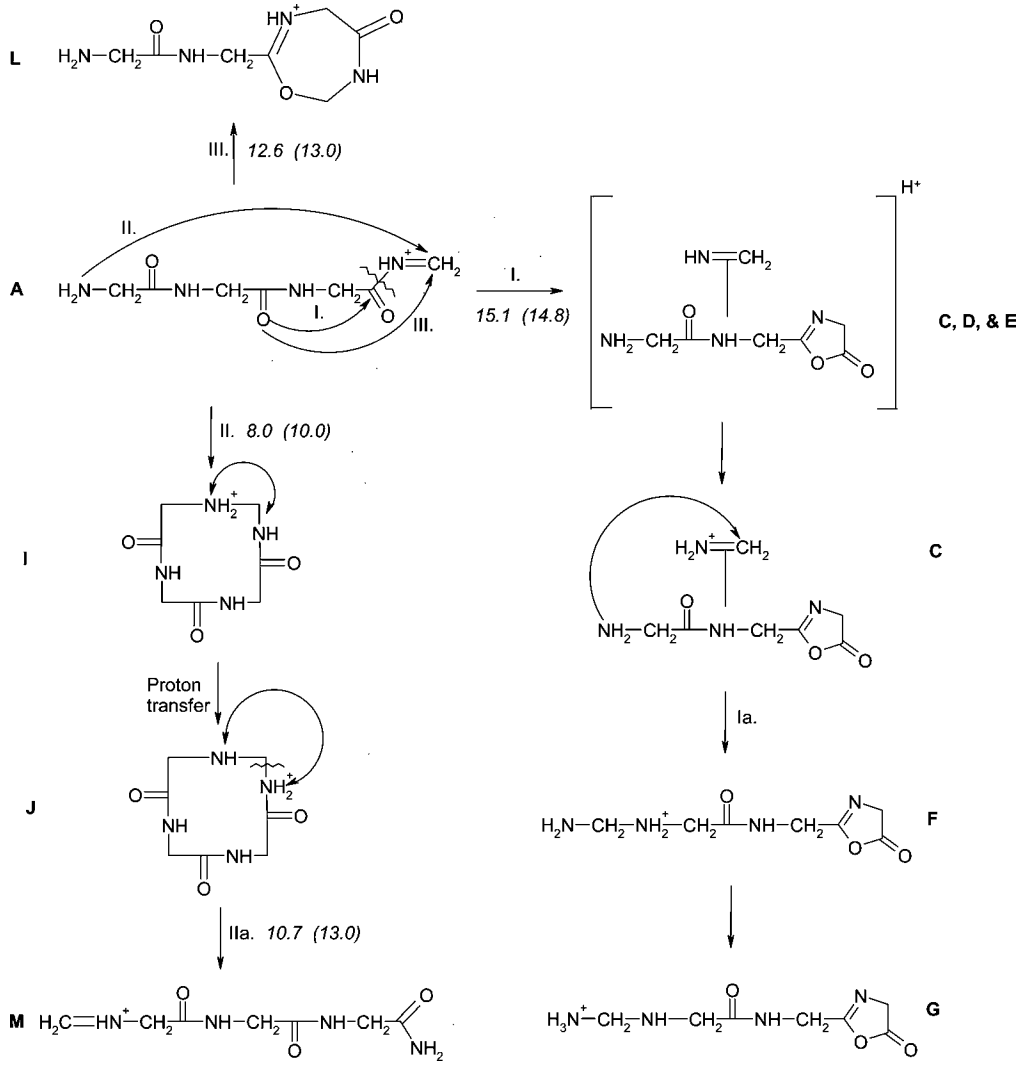
is a proton-bound dimer (PBD) of $HN{=}CH_2$ and 2-(glycylaminomethyl)-5-oxazolone, which has three competitive protonation sites: the imine nitrogen (structure C in Chart 3) as well as the N-terminal amino and ring nitrogens of 2-(glycylaminomethyl)-5-oxazolone (structures D and E in Chart 3, respectively). Of these proton-bound dimers, only structure C is energetically more favorable (relative energy at -1.4 kcal mol^{-1}) than the reference imine-protonated linear structure A. Structure C can undergo a reassociation reaction[15b] where the amino nitrogen of the oxazolone attacks the carbon center of the protonated imine (reaction Ia, Scheme 3). This leads to a 'reassociated' isomer which is terminated on the N-terminal side by the $H_2N{-}CH_2{-}NH_2^+{-}$ group and at the C-terminal side by the oxazolone group. Among the corresponding F, G, and H structures (protonation sites at the backbone secondary amine, the N-terminal primary amine, and the oxazolone ring nitrogen, respectively, Chart 2), the N-terminal amine-protonated form G is significantly more energetically favorable (relative energy at -4.8 kcal mol^{-1}) than the other (F and H) structures.

Cyclization initiated by nucleophilic attack[5b,6f] (reaction II in Scheme 3) of the N-terminal amino nitrogen on the carbon of the protonated imine leads to an 11-membered ring ('macrocyclic') isomer. This macroring accommodates a secondary amine and three amide bonds, two of which must be in the normal trans configuration while the third generated by ring formation can be either cis or trans. This macroring is much less strained than the corresponding 8-membered macroring isomer of the a_3 ion and the secondary imine-protonated form (structure I in Chart 2) is energetically favorable (relative energy at -5.7 kcal mol^{-1}). Furthermore, the barrier to ring formation is low at 8.0 kcal mol^{-1}, allowing facile generation of structure I from A via TS A $\leftrightarrow$ I (Figure S3, Supporting Information). The protonation energetics of the adjacent nitrogen of the amide bond (structure J in Chart 2) were also tested. While several stable species were located (relative energy at 10.3 kcal mol^{-1}) on the PES of the a_4 ion, many of the initial geometries probed did not produce macrocyclic isomers. Instead, these structures opened up at the $-CH_2{-}NH_2^+{-}$ bond during the course of geometry optimization, producing structure type M (Chart 2). The lowest energy J $\leftrightarrow$ M ring-opening TS we found is at 10.7 kcal mol^{-1} relative energy (Figure S3, Supporting Information). This reaction leads to the rearranged linear structure (M in Chart 2, $H_2C{=}NH^+{-}CH_2{-}CO{-}(NH{-}CH_2{-}CO)_2{-}NH_2$) that features imine and amide groups at its N and C termini, respectively (i.e., the imine position is reversed from the initially formed structure A, where the imine is at the C terminus). Structure M is the energetically most favored structure on the PES of the a_4 ion with a relative energy at -14.0 kcal mol^{-1} (Table S3, Supporting Information).

In section 2, the experimental spectrum of the a_3 ion was assigned to an isomer containing a 7-membered ring introduced by nucleophilic attack of the N-terminal adjacent glycine *oxygen* on the carbon center of the protonated imine of structure A. An analogous reaction (reaction III in Scheme 3) is possible for the a_4 ion too, and the corresponding ring imine-protonated isomer (structure L in Chart 2) is again energetically favorable (relative energy at -3.9 kcal mol^{-1}). We also tested another possibility where the oxygen of the N-terminal glycine residue attacks the carbon center of the protonated imine of structure A, forming a 10-membered ring isomer. Our calculations (data not shown) indicate such isomers are energetically much less favorable than the other cyclic forms, so these will not be discussed in detail here (additionally, the spectra were poor fits to the experimental data).

The experimental IR spectrum of the a_4 ion is shown in Figure 4 along with theoretical spectra computed for structures A, B,

Scheme 3. Various Structures and Their Reactions for the a_4 Ion of Protonated GGGGG[a]



[a] Nucleophilic attack (reaction I) of the N-terminal carbonyl oxygen at the C-terminal carbonyl results in cleavage of the CO−NH$^+$= bond and formation of the PBD of NH=CH$_2$ and 2-(glycylaminomethyl)-5-oxazolone (structures C, D, and E). Nucleophilic attack (reaction Ia) of the N-terminal amino nitrogen of the oxazolone on the carbon center of the protonated imine in C leads to F from which G and H can be formed by intramolecular proton transfer. Nucleophilic attack (reaction II) of the N-terminal amino nitrogen on the carbon of the protonated imine group can lead to the macrocyclic (11-membered ring) isomer. After proton transfer to the just-formed amide bond, ring opening occurs, forming the rearranged linear structure M. Nucleophilic attack (reaction III) of the N-terminal carbonyl oxygen at the C-terminal carbon leads to the 7-membered ring isomer L. Relative energies (Gibbs free energies at 298 K) in kcal mol^{-1} are given in italics.

C, G, I, L, and M. Theoretical spectra for all other structures were computed, but for brevity only the energetically most favored species are discussed in the C−D−E, F−G−H, I−J, and K−L families. The experimental spectrum essentially consists of a broad, asymmetric band from 1650 to 1800 cm^{-1}. On the red side of this band, one can clearly distinguish a sharp contribution with a maximum at ∼1685 cm^{-1} and a shoulder at ∼1730 cm^{-1} can also be seen on the blue side of the broad experimental band. In the lower frequency range three well-resolved peaks (at ∼1515, ∼1550, and ∼1585 cm^{-1}) can be distinguished, although the IRMPD signal is low below 1650 cm^{-1}. On the other hand, it is remarkable that no IRMPD signal could be observed when the laser frequency was scanned in the 1800−2000 cm^{-1} region. One can thus safely conclude that a_4 ions with oxazolone-containing structures such as C (or D and E) and G (or F and H) are not present in significant quantities under our experimental conditions since a signal associated with a strongly IR active C=O stretching mode should give rise to a band above 1800 cm^{-1}.

While the structural assignment is not as straightforward as in the cases of the a_2 and a_3 ions, one can safely conclude that the most abundant structure of the a_2, a_3, and a_4 ions formed upon CID conditions in an ion trap is different in each case. A cyclic, 5-membered ring structure was shown to be the most abundant structure for the a_2 ions. This structure is generated upon nucleophilic attack of the N-terminal amino nitrogen on the carbon of the protonated imine. A similar mechanism leads to an 11-membered macrocyclic structure in the case of the a_4 ion. As can be seen in Figure 4, the calculated IR absorption spectrum of the corresponding I structure does not closely match the observed spectrum for the a_4 ions. This is especially true in the 1650−1800 cm^{-1} spectral range where the broad experimental peak is observed. The three CO stretching modes of the 11-membered ring are predicted at 1690, 1722, and 1750 cm^{-1}, but the shape of the broad and asymmetric band observed for the a_4 ions is not well reproduced by the simulated IR cross-section for structure I. This is small at 1685 cm^{-1}, where the maximum of the IRMPD signal is observed (∼1685 cm^{-1}), and

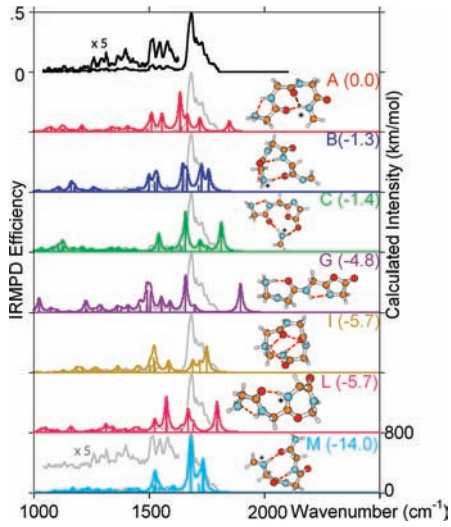


Figure 4. Experimental and calculated infrared spectra of the a_4 ion of protonated GGGGG. Theoretical spectra are presented for the C-terminal imine-protonated linear form (structure A); the N-terminal amino-protonated linear form (structure B); the complex of $NH_2^+{=}CH_2$ and 2-(glycylaminomethyl)-5-oxazolone (structure C); the reassociated, oxazolone-terminated linear form protonated at the N-terminal amine (structure G); the macrocyclic (11-membered ring) isomer protonated at the secondary amine (structure I); the 7-membered ring isomer protonated at the ring imine (structure L); and the rearranged linear form protonated at the N-terminal imine (structure M). The computed relative energies (kcal mol^{-1}) are presented in parentheses for each isomer.

it progressively increases up to its maximum at 1750 cm^{-1} corresponding to the most strongly IR active CO stretch, while the IRMPD signal was found to progressively decrease and finally disappear. One can thus safely conclude that structure I is not the most abundant a_4 isomer formed under our experimental conditions.

Structure L, which was found to have a calculated IR absorption very similar to the experimental spectrum of the a_3 ion in the 1000−2000 cm^{-1} spectral range, cannot be considered as being the predominantly formed structure for the a_4 ion. As in the case of the a_3 ions, theory predicts a strong IR active band (intracyclic C═O stretch) at ∼1795 cm^{-1} for the L isomer of a_4, while no IRMPD signal was experimentally present in this frequency range. Similarly, as in the case of the a_3 ion, structure L of the a_4 ion also has a quite intense IR absorption band at 1575 cm^{-1} corresponding to the N^+H bending mode, and since the IRMPD signal is also weak in this range, one can safely conclude that the population of a_4 ions with structure L should be marginal or nonexistent.

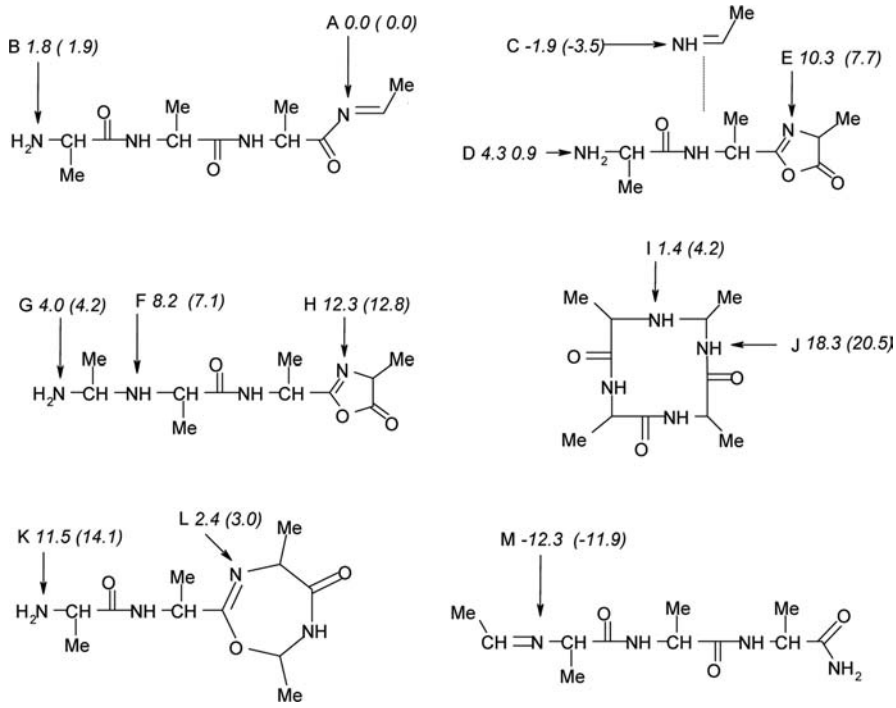
The comparison of the broad and strong experimental band with each individual calculated IR absorption spectrum suggests that the shape of this band nicely matches with the IR cross-section calculated for structure M. It is particularly interesting to notice that the sharp maximum of the experimental signal at ∼1685 cm^{-1} nicely matches with the calculated frequency (1680 cm^{-1}) of the strongest IR band of structure M. This absorption feature is associated with a normal mode of the terminal amide involving the C═O and CN stretches as well as the NH_2 scissoring. Moreover, structure M has two amide C═O stretching modes contributing to the IR cross-section between 1685 and 1800 cm^{-1}, and the calculated frequency of the most strongly IR active C═O stretch is 1733 cm^{-1}, which is very close to the position of the observed shoulder (∼1730 cm^{-1}) of the broad experimental feature.

In contrast to the cases of the a_2 and a_3 ions, it is more difficult to rule out the possibility that isomers other than M are formed with a small abundance in the case of the a_4 ions. As stated above, while the shape of the simulated IR cross-section for the linear B or macrocyclic I structures does not match the broad and asymmetric IRMPD band observed between 1650 and 1800 cm^{-1}, these two structures have multiple IR absorption bands in this spectral range and, most importantly, no predicted IR absorption band above 1800 cm^{-1}. Hence, while structure M is likely to be the predominant structure for a_4 ions, one could also conceive that a smaller fraction of the a_4 ions are formed with a macrocyclic structure such as I and/or an amino-protonated imine B structure. Furthermore, the multiple weak features observed below 1650 cm^{-1} are difficult to assign conclusively considering that structure M is the unique structure which has only one moderately IR active band (at 1527 cm^{-1}) in the 1500−1600 cm^{-1} spectral range, while three features were observed experimentally at ∼1515, ∼1550, and ∼1585 cm^{-1}. The band predicted at 1527 cm^{-1} for structure B corresponds to an amide NH bending mode and could account for one of these features. On the other hand, the two other features could be tentatively assigned to the NH bending mode (1522 cm^{-1}) and NH_2 scissoring mode (1585 cm^{-1}) of the 11-membered ring I structure.

The preceding discussion leads to the possibility of a multicomponent gas-phase mixture of a_4 ion structures. A mixture of a_4 ion structures would also be consistent with the earlier ion mobility data for the a_4 ions of protonated YGGFL,[15d] which showed three peaks, the largest of which was unassignable at the time. From the present combination of spectroscopic and theoretical data, one could conclude that rearranged linear structure M could be the most abundant structure, while macrocyclic structures such as I as well as amino-protonated imine B isomer could also be present as minor fractions. To probe this hypothesis we determined theoretical collision cross-sections (Figure S4, Supporting Information) for the most stable B, I, and M structures of a_4 of protonated GGGGG. As expected the macrocyclic I structures bear the lowest computed cross-sections (72−74 $Å^2$), and these are likely to account for the experimental IMS peak with the highest mobility. The cross-sections for the M structures spread a relatively wide range (78−91 $Å^2$); the two energetically most stable isomers feature collision cross-sections differing by more than 10 $Å^2$. The existence of these two energetically competitive M structures is fully consistent with the observed IMS features for the a_4 ions of protonated YGGFL;[15d] the two lowest mobility peaks could reasonably be explained by these M structures. The calculated cross-sections for the B isomers (79−81 $Å^2$) are similar to that of the lowest energy M structure, so these species can in principle also contribute to the IMS features of lesser mobility. It is worth noting here that the experimental IMS data was obtained for the a_4 ions of protonated YGGFL,[15d] and our theoretical data was computed for the a_4 ions of protonated GGGGG, so perfect agreement between the two data sets should not necessarily be expected. It is however clear that the main experimental IMS features are easily explained based on our theoretical data by assuming that M structures are populated in the mass spectrometer.

The assignment of our IRMPD data above is in line with the computed energetics for the a_4 isomers studied; M is the energetically most favored isomer. Formation of M is possible because of the low-energy A ↔ I and J ↔ M ring-closing and -opening transition structures and the relatively favorable amide

Chart 3. Various a_4 Ion Structures for the a_4 Ion of Protonated AAAAA[a]



[a] (A and B) Linear form protonated at the imine and N-terminal amino groups, respectively; (C, D, and E) proton-bound dimer of HN═CHMe and 2-(alanylaminoethyl)-(4-methyl)5-oxazolone protonated at the imine nitrogen, amino group of the oxazolone, and ring nitrogen of the oxazolone, respectively; (F, G, and H) oxazolone-terminated re-associated linear form protonated at the secondary amine, N-terminal amino, and ring nitrogen, respectively; (I and J) macrocyclic (11-membered ring) isomer protonated at the secondary amine and amide nitrogen, respectively; (K and L) 7-membered ring isomer protonated at the N-terminal amine and ring imine nitrogen; and (M) rearranged linear form protonated at the N-terminal imine group. The respective protonation sites are indicated by arrows. Relative ZPE-corrected energies (Gibbs free energies at 298 K) in kcal mol^{-1} are given in italics.

nitrogen-protonated macrocyclic structures. However, it should be noted that the unfavorable energetics of similar reactions and/or amide nitrogen-protonated macrocyclic intermediates effectively hindered formation of the rearranged linear structures for the a_2 and a_3 ions (E in Scheme 1 and M in Scheme 2, respectively). For the a_4 ion the ring-closing and -opening A ↔ I and J ↔ M TSs are at 8−11 kcal mol^{-1} relative energy and the majority of the a_4 ions appear to have enough energy remaining following their generation to cross these barriers, enabling formation of the most energetically favorable structure M. This type of process has precedence, as evidenced by formation of O-protonated cyclic b_5 ion (and larger) structures in mass spectrometers.[5,9] It should also be explicitly noted that structure M has a rearranged sequence, so further fragmentation of this structure may be detrimental if it occurs, leading to formation of *nondirect* sequence ions.[5]

4. Structure and IR Spectroscopy of the a_4 Fragment of Protonated AAAAA. Here we briefly compare the findings of the IRMPD action spectroscopy of the a_4 ion of protonated GGGGG to a larger a_4 ion generated from protonated AAAAA. In so doing we assess the generality of the findings presented in the preceding text, i.e., whether the complex rearrangement chemistries demonstrated for the a_4 ion of protonated GGGGG are present for the analogous ion generated from protonated AAAAA. Again, we calculated a large number of possible structures in order to offer the most objective comparison with the recorded spectrum. The relative energies of the structures calculated as well as the free energies (at 298 K) are summarized in Chart 3 and Table S4 (Supporting Information). The experimental spectrum of the a_4 ion generated from protonated AAAAA and the corresponding computed spectra for structures A, B, C, G, I, L, and M are shown in Figure 5.

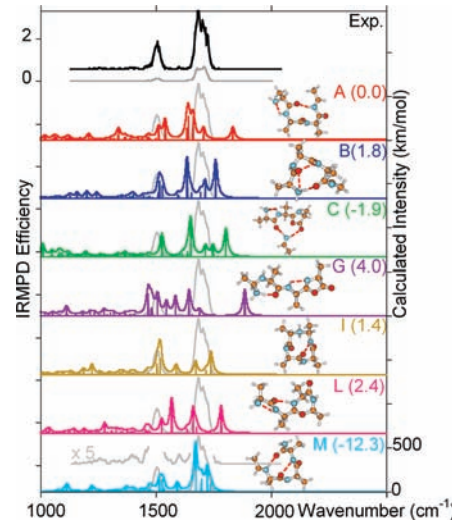


Figure 5. Experimental and calculated infrared spectra of the a_4 ion of protonated AAAAA. Theoretical spectra are presented for the C-terminal imine-protonated linear form (structure A); the N-terminal amino-protonated linear form (structure B); the complex of H_2N^+═CH_2 and 2-(alanylaminoethyl)-(4-methyl)5-oxazolone (structure C); the reassociated, oxazolone-terminated linear form protonated at the N-terminal amine (structure G); the macrocyclic (11-membered ring) isomer protonated at the secondary amine (structure I); the 7-membered ring isomer protonated at the ring imine (structure L); and the rearranged linear form protonated at the N-terminal imine (structure M). The computed relative energies (kcal mol^{-1}) are presented in parentheses for each isomer.

In the preceding spectrum of the a_4 ion of protonated GGGGG (Figure 4), structure M had the best match while contributions of structures B and I to the experimental spectrum could not be

ruled out. In Figure 5 the situation appears to be clearer and strongly in favor of the new, rearranged a_4 ion with structure M (Chart 3, $MeHC{=}NH^+{-}CHMe{-}CO{-}(NH{-}CHMe{-}CO)_2{-}NH_2$). Structure M is the energetically most favored structure on the PES of the a_4 ion with a relative energy at -12.3 kcal mol^{-1} (Chart 3, Table S4, Supporting Information). This structure provides the best match for the recorded spectrum. First, the large absorption at $\sim$1510 cm^{-1}, which is not present for the glycyl analogue (Figure 4), is reproduced in the theoretical spectrum of structure M by a linear combination of the amide N$-$H bends. The lower frequency shoulder at $\sim$1480 cm^{-1} is assigned as a linear combination of alanine methyl group C$-$H stretches. The small absorption at $\sim$1600 cm^{-1} corresponds nicely to the symmetric N$-$H bends in the newly created $-C(O)-NH_2$ group. The most prominent, diagnostic group of bands for structure M begins at $\sim$1680 cm^{-1} and stretches to $\sim$1740 cm^{-1}. This broad experimental absorption is composed of several closely spaced, large absorptions. First, there is one at $\sim$1680 cm^{-1}, corresponding to the strong C$=$O stretch of the of the C-terminal amine group ($-C(O)-NH_2$) which is coupled to the N$-$H stretches of this group. Second, there are two large amide C$=$O stretches corresponding to the internal amide groups ($\sim$1697 and $\sim$1723 cm^{-1}). Lastly, the N-terminally-protonated imine group has an absorption at $\sim$1737 cm$^{-1,}$ corresponding to the C$=$N stretch coupled to N$-$H and C$-$H bends.

In analogy to the findings for the glycyl analogue (Figure 4), structure I does not provide as good a match in terms of intensity of absorption but cannot be ruled out as a contributor to the experimental spectrum as the cyclic structure has predicted modes that also match the experimental spectrum in terms of position. It is thus conceivable that some a_4 ions of protonated AAAAA may have been trapped in this conformation too despite the higher relative energy (1.4 kcal mol^{-1}). In contrast, structure B, which was also a potential contributor to the experimental spectrum of the glycyl analogue (Figure 4), does not appear to contribute significantly to the alanyl form (Figure 5). This conclusion is based on the large absorptions predicted at 1640 and 1760 cm^{-1} for this structure which are not observed in the experimental spectrum. This follows from the higher relative energy of this type of structure generated from protonated AAAAA (1.8 kcal mol^{-1}, Figure 5), making structure B less energetically favorable than it is gylcyl analogue (-1.3 kcal mol^{-1}, Figure 4).

Our IRMPD and theoretical data suggest the presence of isomers I and M under the experimental conditions of our ion trap. To test the compatibility of this finding with the earlier ion mobility data for the a_4 ions of protonated YGGFL[15d] we calculated theoretical collision cross-sections (Figure S5, Supporting Information) for the most stable I and M structures of a_4 of protonated AAAAA. As expected the macrocyclic I structures are characterized by much lower cross-sections than the linear M structures. The cross-sections for the M structures spread a relatively wide range, and the energetically most stable isomers feature two classes of collision cross-sections differing by $\sim$10 Å^2. The existence of these two energetically competitive types of M structures is fully consistent with the observed IMS features; the two lowest mobility peaks could reasonably be explained by the corresponding M structures. This indicates that the experimentally observed IMS features could be rationalized by assuming the presence of only the I- and M-type a_4 isomers. To test this hypothesis we are currently performing detailed structural and cross-section calculations on the A$-$M isomers of the a_4 ions of protonated YGGFL. The findings of this work will be communicated separately.

Conclusions

(1) The present study demonstrates the importance of the synergism between gas-phase infrared spectroscopy and quantum chemical calculations for deciphering the complex reaction mechanisms that are involved in peptide fragmentation. The experimental infrared spectra of the a_n ($n = 2-4$) ions of oligoglycines show very distinct absorption patterns in the 1000$-$2000 cm^{-1} range, indicating that dissimilar isomers and/or protonation forms are formed under our experimental conditions. Exploration of the potential energy surface of these ions shows that three types of rearrangement reactions could be at play starting from the iminium a_n structure formed following the loss of CO from the corresponding b_n ion. The combination of the experimental infrared spectrum and theoretical spectrum for each low-lying a_n isomer, however, clearly suggests that the isomer associated with a low activation energy is observed in each case, e.g., fragmentation is kinetically controlled.

(2) Our combined IR spectroscopy and DFT study indicates that the a_2 ion with the GG sequence undergoes cyclization to form a 5-membered ring isomer. The corresponding 5-membered ring accommodates a secondary amine and cis amide bond. Protonation at the former is energetically more favorable, and proton transfer to form the amide nitrogen-protonated species is prevented due to the high energy barrier involved. This effectively eliminates the energetically favorable ring opening from the amide nitrogen-protonated species that could lead to a lower energy rearranged linear isomer. This is consistent with the findings of Verkerkk et al.[14d]

(3) The a_3 ion with the GGG sequence can in principle undergo various cyclization, proton transfer, and reassociation reactions. Our IR spectroscopy and DFT data indicate that the imine-protonated linear form undergoes cyclization that is initiated by nucleophilic attack of the carbonyl oxygen of the N-terminal glycine residue on the carbon center of the C-terminal immonium group, forming a 7-membered ring isomer. The barrier to this reaction is low at 10.5 kcal mol^{-1}, and the resulting cyclic isomer is energetically more favorable than the linear form. Cyclization to form the macrocyclic (8-membered ring) isomer does not occur since the resulting cyclic form which must accommodate a trans amide bond is energetically unfavorable due to the significant ring strain involved. Furthermore, formation of the amide nitrogen-protonated macrocyclic structure which could then open up to form a low-energy rearranged linear isomer is kinetically hindered.

(4) The a_4 ion with the GGGG sequence undergoes head-to-tail cyclization via nucleophilic attack of the N-terminal amino group on the carbon center of the C-terminal immonium ion, forming an 11-membered macroring which contains a secondary amine and three amide bonds. This macrocyclic isomer is energetically much more favorable than the initial imine-protonated linear isomer. An intermolecular proton transfer converts the initially formed $-CH_2-NH_2^+-CH_2-NH-CO-$ moiety to the $-CH_2-NH-CH_2-NH_2^+-CO-$ group which undergoes ring opening at the $-CH_2-NH_2^+-$ bond, forming a rearranged linear isomer with the $H_2C{=}N-CH_2-$ moiety at the N terminus and a $-CO-NH_2$ amide bond at the C terminus. This rearranged a_4 linear structure is the most energetically favorable isomer at -14 kcal mol^{-1} relative energy. The barriers to the above cyclization and ring-opening reactions

are low, allowing facile formation of the rearranged linear species. The energetic feasibility of such reactions is supported by prior work on cyclization reactions in *b* ions.[5,9] The presence of some component(s) of the B and/or I isomers in the ion trap cannot be excluded based on the available experimental data. Collision cross-section calculations indicate that the simultaneous presence of B-, I-, and M-type structures in the mass spectrometer is consistent with earlier ion mobility data for the a_4 ions of protonated YGGFL[15d] and could explain the trimodal IMS spectrum features.

(5) The experimental spectrum of the a_4 ion with the nominal AAAA sequence provides even stronger support for these rearrangement reactions as by far the best match was achieved using structure M (Figure 5, a linear isomer with the protonated $MeHC{=}N{-}CH_2{-}$ moiety at the N terminus and a $-CO-NH_2$ amide bond at the C terminus). This finding shows that the presence of larger side chains is not necessarily an insurmountable impediment to forming such structures. The presence in the experimental spectrum of some component(s) of isomer I (a cyclic structure) in the ion trap cannot be completely excluded; however, the N-terminally protonated imine structure B can. Cross-section calculations indicate that a mixture composed of only I- and M-type structures is fully consistent with earlier ion mobility data[15d] in this case. This supports our IR assignments.

(6) Our results indicate that one needs to consider both cyclization and rearrangement reactions in order to decipher the structure(s) and fragmentation pathways of peptide a_n ions. We propose new a_n isomers, the presence of some of which was supported by IR spectroscopy and DFT calculations. While it was previously suggested[5b] that a_n ions do not undergo 'scrambling' reactions, formation of the macrocyclic isomer and ring opening at the $-CH_2-NH_2^+-$ bond after proton transfer (structures I, J, and M for the a_4 ion) presented here necessitates revision of this 'fragmentation rule'. This clearly has implications for peptide sequence assignment based on the use of a_n ions. Our laboratories are currently studying these reactions and the new potential isomers to discover how general their occurrence is.

Acknowledgment. B.P. is grateful to the Deutsche Forschungsgemeinschaft for a Heisenberg fellowship and the Landesstiftung Baden-Württemberg (P-LS-Prot/57) for further financial support. B.B. thanks the DKFZ for a guest scientist fellowship. Financial support by the European Commission EPITOPES project (NEST program, project #15367) is gratefully acknowledged. The authors are grateful to J. M. Ortega, V. Steinmetz, and J. Lemaire for technical support.

Supporting Information Available: Total and relative energies of the species presented in the text (Tables S1−4), structures of the isomers discussed (Figures S1−4), and complete refs 18 and 19. This material is available free of charge via the Internet at http://pubs.acs.org.

JA101556G