ACS Partner Journal

Journal of the American Society for Mass Spectrometry

pubs.acs.org/jasms

Article

Photochemical and Collision-Induced Cross-Linking in Stereochemically Distinct Scaffolds of Peptides and Nitrile Imines in Gas-Phase Ions

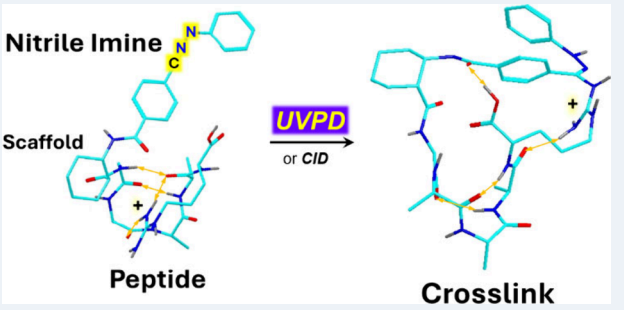
Hongyi Zhu, Marianna Nytka, Tuan Ngoc Kim Vu, Karel Lemr,* and František Tureček*

Cite This: *J. Am. Soc. Mass Spectrom.* 2024, 35, 3070−3088

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Intramolecular cross-linking between peptides and nitrile-imine intermediates was studied in stereochemically distinct conjugates in which the reacting components were mounted on *cis*-1,2-cyclohexane and *trans*-1,4-cyclohexane scaffolds that we call 1,2-*s*-peptides and 1,4-*s*-peptides, respectively. The nitrile-imine intermediates were generated by N_2 loss from 2,5-diaryltetrazole tags upon UV-photodissociation at 213 and 250 nm or by collision-induced dissociation, and further interrogated by CID and UVPD-MS^3. Peptide fragment ion series originating from linear structures and macrocyclic cross-links were distinguished and used to quantify the cross-linking yields. The yields in MS^2 varied between 27% for AAAG conjugates to 78% for GAAAK conjugates, depending on the peptide sequence. The CID-MS^3 yields were in a 57−97% range, depending on the peptide sequence. Structures of 1,2-*s*-peptide and 1,4-*s*-peptide ions as well as several of their nitrile-imine intermediates and cross-links were investigated by high-resolution cyclic ion mobility in combination with Born−Oppenheimer molecular dynamics and density functional theory calculations. Matches between the experimental and calculated collision cross sections and ion relative Gibbs energies were used to assign peptide structures. Peptide conjugates C-terminated with Gly and Lys residues underwent cross-linking by the carboxyl group, as established by MS^3 sequencing and corroborated by carboxyl blocking experiments that lowered the cross-linking yields. Peptide conjugates C-terminated with Arg also cross-linked via the side-chain guanidine group. A notable feature of the 1,4-*s*-peptide ions was the participation of low-energy twist-boat cyclohexane conformers that was enforced by strong hydrogen bonds between the peptide and nitrile imine.



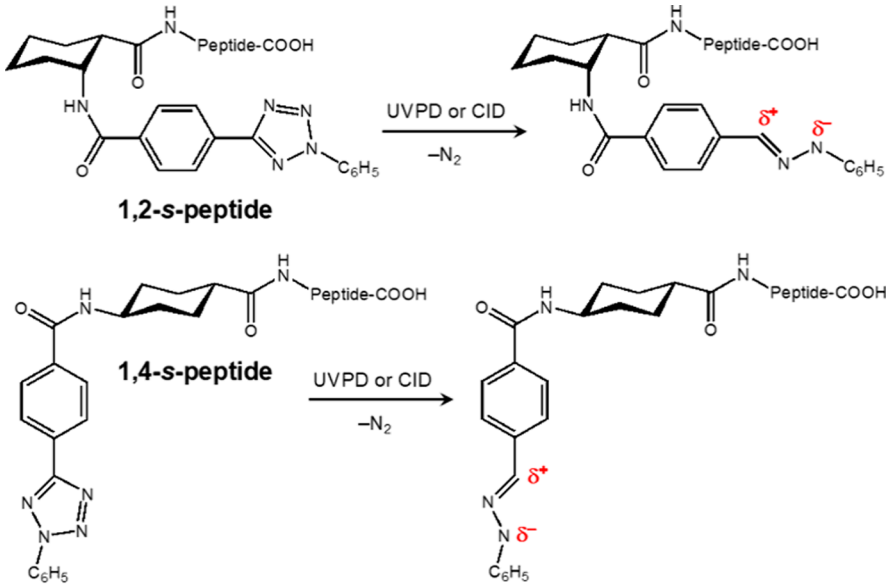
INTRODUCTION

The development of new reagents has been in the forefront of cross-linking studies aimed at the determination of biomolecular structure by mass spectrometry.[1] In addition to photoreactive tags such as diazirines[2−4] and benzophenones[5−7] that can be introduced into selected amino acid residues, various bifunctional cross-linkers have been introduced[8] and used[9] that targeted lysine residues, and new variations have been proposed that were cleavable by collision-induced dissociation (CID) in tandem mass spectrometry analysis.[10−14] Cycloaddition reactions, such as the azide−alkyne click reaction, have expanded the portfolio of cross-linking strategies in solution.[15,16] The tag approach has the advantage that at least one site of attachment is known so that the detection of cross-links indicates the residues that are sterically accessible to the tagged ones.

We have been pursuing a different approach in which cross-linking was achieved by photodissociation of gas-phase peptide ions that were site-specifically tagged with diazirine groups.[17] The main aim of these studies has been to elucidate ion structures and determine peptide−peptide[18] and peptide−oligonucleotide[19] noncovalent interactions in gas-phase complexes. Gas-phase cross-linking offers some advantages over classical solution studies. First, the reactive intermediates are distinguished and selected by mass which allows one to quantify the cross-linking yields.[20] Second, the cross-link sites can be elucidated by tandem mass spectrometry *de novo* sequencing.[21] However, similar to cross-linking in solution, gas-phase cross-linking is a low resolution method, the chief limitation being the incomplete and peptide-dependent sequence coverage. Recently, we have reported on a discovery that nitrile imines undergo photoinduced cross-linking to peptides in conjugates that were tagged with a 2,5-diaryltetrazole group.[22] The diaryltetrazole is conveniently attached to a suitable nucleophile in the peptide, such as the lysine ε-amine,[23] or cysteine SH.[22] Photodissociation at 250 nm breaks up the tetrazole ring[24] expelling N_2 and creating the reactive

Received: July 24, 2024
Revised: September 27, 2024
Accepted: October 16, 2024
Published: October 24, 2024

Scheme 1. Peptide-Nitrile-Imine Conjugates with Stereochemically Distinct *cis*-1,2- and *trans*-1,4-Cyclohexane Scaffolds



nitrile imine. The associated change of mass allows one to track the conversion and quantify the reaction yield.[22] Our previous investigation of nitrile-imine cross-linking reactions has revealed that they proceeded as a proton-catalyzed [3 + 2] addition to the peptide amide group that was followed by ring opening and proton migration.[22] In an effort to further explore nitrile-imine cross-linking, we now investigate conjugates in which peptide chains and the 2,5-diaryltetrazole moiety are mounted on cyclohexane scaffolds.[25] This design allows us to exploit the cyclohexane stereochemistry of *cis*-1,2- and *trans*-1,4-substituents to position the reacting moieties in a well-defined steric arrangement and study its effect on cross-linking yields and attachment sites (Scheme 1).

The flexible nature of the *cis*-1,2-cyclohexane scaffold was expected to provide unimpeded access of the incipient nitrile imine to the peptide backbone amide and side-chain groups. In contrast, the *trans*-1,4-cyclohexane scaffold was presumed to exert steric hindrance to the peptide-nitrile-imine interaction in a fashion analogous to peptides containing 1,4-*cis*- and *trans*-cycloornithine residues.[26] Here, we explore these potential stereochemical effects on cross-linking with peptides containing different functional groups, such as carboxyl, methoxycarbonyl, amine, imidazole, and guanidine that may compete in their reactions with the nitrile imine group.[27] The paper is organized as follows. In the first part we present and discuss photodissociation and collision-induced dissociation (CID) tandem mass spectra of peptide-tetrazole conjugates with the aim of identifying the fragment ions and establishing cross-linking yields. This is followed by structure analysis of the conjugates and dissociation products by high-resolution cyclic ion mobility mass spectrometry in combination with Born–Oppenheimer molecular dynamics (BOMD) and density functional theory (DFT) calculations of ion structures to obtain theoretical collision cross sections (CCS_{calc}) for comparison with experimental data.

EXPERIMENTAL SECTION

Materials and Methods. *cis*-1,2-Cyclohexane-diaryltetrazole-peptide conjugates were synthesized by sequential coupling of 5-phenyl-(2-phenyl-4-carboxyl)-tetrazole and peptide *N*-termini to *cis*-2-(Fmoc-amino)-cyclohexanecarboxylic acid (Sigma-Aldrich, St. Louis, MO). *trans*-1,4-Cyclohexane-diaryltetrazole-peptide conjugates were synthesized analogously using *trans*-4-(Fmoc-amino)-cyclohexanecarboxylic acid (Sigma-Aldrich, St. Louis, MO). The synthetic procedures are described in the Supporting Information. Ions were produced by electrospray ionization from 50:50 methanol–water solutions that contained 1% of acetic acid. UVPD spectra at 250 nm of mass-selected ions were measured on a Bruker amaZon 3D ion trap (Bruker Daltonik, Bremen, Germany) that was modified for photodissociation studies.[28] The light beam was provided by an EKSPLA NL301G (Altos Photonics, Bozeman, MT, USA) Nd:YAG laser (20 Hz) and directed into a PG142 unit consisting of an optical parametric oscillator that was coupled to a second-harmonic generator. The typical laser pulse energies at 250 nm were 1.8–2.0 mJ per pulse. The 5 mm diameter light beam from the OPO was focused into the 1 mm aperture in the ion trap ring electrode.[28] UVPD spectra at 213 nm were obtained on an Orbitrap Ascend Tribrid instrument (ThermoFisher, San Jose, CA) using 15–25 ms stage time (25–63 laser pulses per scan) at 1.5 ± 0.2 μJ/pulse. Spectra were obtained by scanning the ion trap or in the high-resolution mode in the Orbitrap at 100 000 resolving power. All ion assignments were corroborated by elemental compositions that were based on accurate m/z. Ion mobility measurements were carried out on a SELECT SERIES Cyclic Ion Mobility Spectrometer (c-IMS) (Waters Corp., Wilmslow, U.K.)[29] with direct infusion into a normal flow electrospray ion source at a flow rate of 5 μL/min. Each sample was measured six times in a positive mode. Two solutions of calibrant were measured once. Each calibrant ion was isolated in the quadrupole and its arrival time was determined. Typically, ion mobility separation over several cycles (n) was used at the total ion path of $n \times 98$ cm. Details

https://doi.org/10.1021/jasms.4c00317
J. Am. Soc. Mass Spectrom. 2024, 35, 3070–3088

of all measurement parameters (Table S1, S2) and calibration for collision cross section (CCS) determination (Table S3, S4) are given in the Supporting Information.

Calculations. Structures of the conjugates and their dissociation products were generated by Born−Oppenheimer molecular dynamics (BOMD) calculations that were run as 20 ps trajectories with 1 fs steps, using the Berendsen thermostat[30] at 510, 610, and 810 K. Trajectory analysis confirmed that the systems were reaching a near-stationary state after ca. 16000 steps (16 ps), as determined by low root-mean-square deviations of atomic Cartesian coordinates, so no longer trajectory runs were necessary because they generated only duplicate structures. Initial protonation sites were chosen at different amide groups for the nonbasic AAAG conjugates to be treated by BOMD. Since BOMD includes both electron and nuclear motion, the different protonation sites coalesced by proton transfer to the most basic position. The initial protonation sites in AAAHG, GAAAK, and GAAAR conjugates were at the basic residues. The trajectory calculations were performed with PM6-D3H4[31] which complements the semiempirical Hamiltonian with dispersion and hydrogen-bonding interactions, using the Cuby4 platform[32] and MOPAC.[33] Two hundred snapshots were selected at regular 100 fs intervals from the BOMD trajectories, and the structures were fully gradient-optimized with PM6-D3H4. Several selected low energy structures were further optimized with B3LYP[34] with the 6-31+G(d,p) basis set. These calculations were augmented by including empirical dispersion corrections (GD3-BJ)[35,36] and were used to provide harmonic frequencies. Several low-energy structures were reoptimized with M06-2X[37] using the 6-31+G(d,p) basis set. The M06-2X/6-31+G(d,p) optimized geometries were used to calculate single-point energies which were carried out with M06-2X and the def2qzvpp[38] basis set (<5000 basis functions). It should be noted that ion thermochemistry obtained at this level of theory is approximate within an estimated 5−10 kJ mol^{-1},[37,39] and thus no efforts were made to derive the ion isomer equilibrium populations from the calculated Gibbs free energies ($\Delta G_{g,310}$). Another set of M06-2X/6-31++G(d,p) single-point calculations were used to obtain charge densities according to Merz, Singh, and Kollman (MK).[40,41] All the standard DFT calculations were run using Gaussian 16 (Revision B.01) that was licensed from Gaussian, Inc. (Wallingford, CT). Collision cross sections in nitrogen were calculated by the modified ion trajectory method ($MobCal_{MPI}$)[42,43] using the MK charge densities. Standard van der Waals parameters for different atom types were obtained from the MMFF94 data set.[44]

RESULTS AND DISCUSSION

***cis*-1,2-Cyclohexane-diaryltetrazole-AAAG (1,2-*s*-AAAG) and *trans*-1,4-Cyclohexane-diaryltetrazole-AAAG (1,4-*s*-AAAG) Scaffolds.** Photodissociation of 1,2-*s*-$AAAG^+$ (m/z 662) resulted in backbone cleavages and loss of N_2 that was followed by dissociations of the primarily produced nitrile imine ion $\mathbf{1}^+$ (m/z 634) as shown in the UVPD-MS^2 spectrum (Figure 1a). The loss of N_2, as well as all other fragment ion assignments, were based on accurate mass measurements (Δm = 28.0060 for N_2, Table S5−S13, Supporting Information). The dissociations upon UVPD showed strong dependence on the photon energy, which was E_{photon} = 4.96 and 5.82 eV, respectively, for the 250 and 213 nm laser beams. This was reflected by the more extensive dissociations of the $(1,2\text{-}s\text{-AAAG}-N_2)^+$ ion ($\mathbf{1}^+$) when

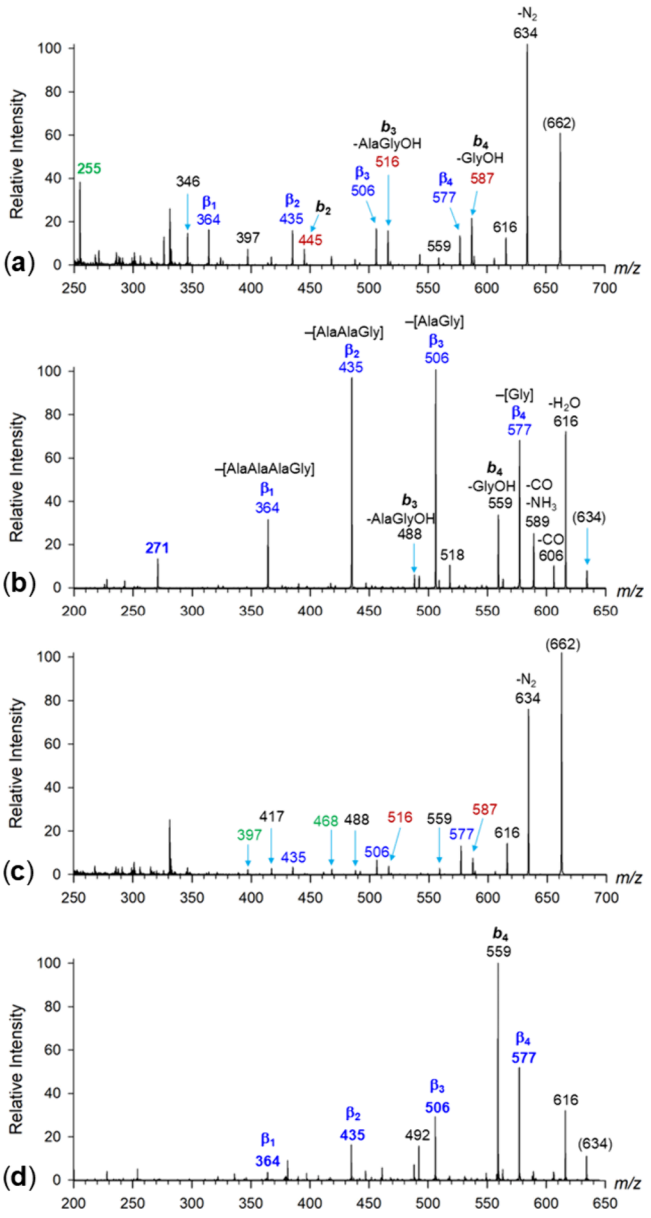


Figure 1. (a) UVPD-MS^2 of 1,2-*s*-$AAAG^+$ (m/z 662) obtained at 250 nm. (b) CID-MS^3 of $(1,2\text{-}s\text{-AAAG}-N_2)^+$ ($\mathbf{1}^+$) from 250 nm UVPD. (c) UVPD-MS^2 of 1,4-*s*-$AAAG^+$ (m/z 662) obtained at 250 nm. (d) CID-MS^3 of $(1,4\text{-}s\text{-AAAG}-N_2)^+$ ($\mathbf{2}^+$) from 250 nm UVPD. Fragment ion color-coding is as follows: Dark red: b_n and $[y_n + 2H]^+$ from the precursor ion; green: b_n after loss of N_2 and C_6H_5N; blue: $\boldsymbol{\beta}_n$ ions by loss of truncated residues from cross-links after loss of N_2. The aminocyclohexane carboxylic acid is counted as the first residue in b_n and $\boldsymbol{\beta}_n$ ions.

generated by UVPD at 213 nm (Figure S1, Supporting Information) We note that the UV−vis absorption spectrum of 2-(4-carboxyphenyl)-5-phenyl-tetrazole in acetonitrile solution showed strong absorption at both wavelengths due to the 2,5-diaryltetrazole chromophore (Figure S2, Supporting Information), while absorption at 250 nm in gas-phase ions has also been established by UVPD action spectra.[24] UVPD produced three series of fragment ions whose relative intensities depended on the photon energy. Peptide backbone cleavage without tetrazole N_2 loss produced b_n type ions at m/z 587, 516, 445, 374, and $[y_n + 2H]^+$ ions at m/z 289, and 271 that were particularly abundant at 213 nm (red-annotated ions in

https://doi.org/10.1021/jasms.4c00317
J. Am. Soc. Mass Spectrom. 2024, 35, 3070−3088

Scheme 2. Proposed Formation of β_n Sequence Fragment Ions from Crosslinked $(1,2\text{-}s\text{-AAAG}-N_2)^+$

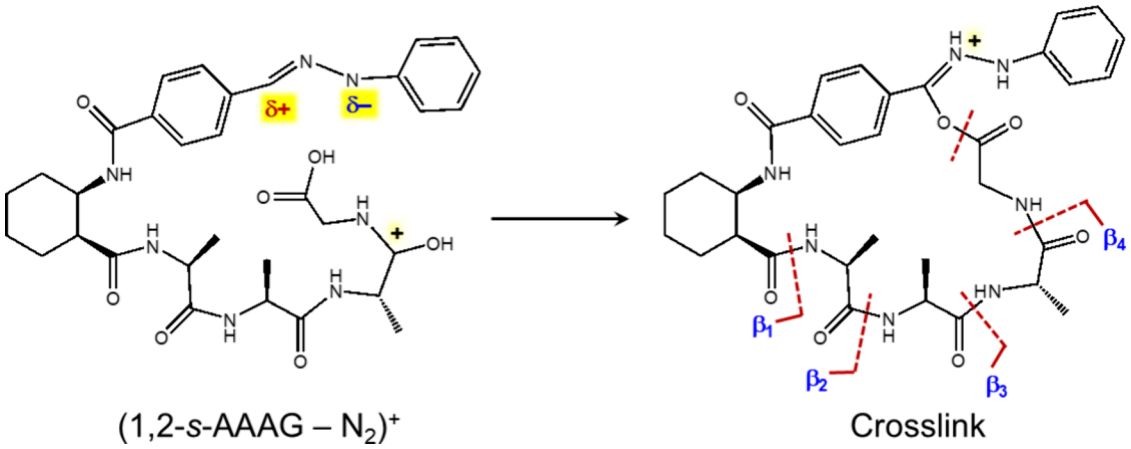


Figure 1a and S1). We note that backbone cleavages leading to *b*-type ions were predominant in the CID-MS^2 spectrum of 1,2-*s*-$AAAG^+$ (Figure S3, Table S6) that did not show loss of N_2. A second ion series originated from $\mathbf{1}^+$ (m/z 634) by loss of C_6H_5N, possibly phenylnitrene or an isomer, giving rise to fragment ions at m/z 543 and their sequence *b*-type ions at m/z 468, 397, 326, and 255 (green-annotated ions in Figure 1a). This ion series was prominent in the 213 nm UVPD spectrum (Figure S1) while being much less abundant at 250 nm (Figure 1a). The third ion series (blue-annotated ions in Figure 1a and S1) also originated from $\mathbf{1}^+$ but involved internal residue losses forming m/z 577 (loss of [$HNCH_2CO$], denoted as [Gly]), m/z 506 (loss of [$HNCH(CH_3)CONHCH_2CO$], denoted as [AlaGly]), m/z 435 (loss of [AlaAlaGly]), and m/z 364 (loss of [AlaAlaAlaGly]). In this notation, we use square brackets to distinguish the neutral internal losses, [Amino acid residues], from those of standard amino acids and peptides leading to *b* and $[y_n + 2H]^+$ fragment ions.[45] Since losses of internal amino acid residues are characteristic of cyclic peptides,[46−48] their presence in the UVPD-MS^2 spectra allowed us to assign internal cross-links between the nitrile imine and peptide groups. The dramatic difference between the UVPD and CID-MS^2 spectra was due to the different modes of ion excitation that affected the competing dissociations. In particular, absorption of a 250 or 213 nm photon by the 2,5-diaryltetrazole chromophore produced excited electronic states that could dissociate by N_2 expulsion before the electronic energy was dissipated within the ion, resulting in vibrational excitation. In contrast, slow heating by collisions drove ground-electronic state dissociations where the peptide bond cleavages prevailed. This was despite that previous energy analysis of tetrazole ring dissociations has indicated a low threshold energy for the N_2 loss,[17,22] as also addressed by calculations described later in the paper.

In the next step, we selected ions $\mathbf{1}^+$ (m/z 634) and investigated their CID-MS^3 spectra (Figure 1b, Table S7). The main fragment ion series was represented by internal fragments that were formed by consecutive loss of [Gly] and [Ala] residues; these ions are denoted as $\boldsymbol{\beta}_n$ to distinguish them from the more common b_n sequence ions. Out of the latter, the b_2, b_3, and b_4 were found in the spectrum at relatively low intensities, representing structures with an open peptide chain in a minor population of ions $\mathbf{1}^+$. The main fraction, represented by the blue-annotated $\boldsymbol{\beta}_n$ ions, were due to cross-links. The Figure 1b spectrum showed consecutive losses of internal fragments that all started with [Gly]. In contrast, an initial loss of [Ala] (m/z 563) or [AlaAla] (m/z 492) from $AAAG^+$ to give ions in which the [Gly] residue would be retained was very weak. To mark proton migrations in the cross-links, we used full H/D exchange in 1,2-*s*-$AAAG^+$ to produce ion $[D_7]$-*s*-$AAAG^+$ (m/z 669) that was photolyzed at 250 nm, and the $[D_7]$-$\mathbf{1}^+$ ion formed (m/z 641) was investigated by CID-MS^3 (Figure S4, Supporting Information). The observed deuterium retention in the $\boldsymbol{\beta}_n$ fragment ions was consistent with the expected neutral losses of $[[D_1]\text{Gly}]$ and $[[D_1]\text{Gly}([D_1]\text{Ala})_n]$ (n = 1−3). It should be noted that the fragment ions underwent back exchange while stored in the ion trap, resulting in a distribution of deuterated species. For example, the abundant $\boldsymbol{\beta}_4$ ion of an expected $[D_6]$ content showed a distribution of $[D_6]$, $[D_5]$, and $[D_4]$ species, and likewise for the other $\boldsymbol{\beta}_n$ fragment ions (Figure S4).

The sequential loss of [Gly] and [Ala] from $\mathbf{1}^+$ may be interpreted as resulting from cross-linking to the nitrile imine by the Gly COOH group, which was associated with proton transfer and anchored one of carboxyl oxygen atoms in the nitrile imine moiety. This is shown in Scheme 2 that indicates rupture of the cross-linked O—C═O ester bond accompanied by backbone amide dissociations. These two types of dissociations can occur competitively, leading to the same result.

We probed this hypothesis by blocking the carboxyl OH by converting it to a methyl ester in 1,2-*s*-$AAAG\text{-}OCH_3^+$ and obtained the CID and UVPD-MS^2 spectra (Table S8,S9). CID-MS^3 of the methyl ester ion, $\mathbf{1}\text{-}OCH_3^+$, m/z 648, that was generated by UVPD of 1,2-*s*-$AAAG\text{-}OCH_3^+$ (Figure S5, Supporting Information), showed a markedly diminished formation of $\boldsymbol{\beta}_n$ ions while the spectrum was dominated by b_1-b_4 ions from cleavages of the presumably linear peptide chain (Figure S5, Table S10, Supporting Information).

The isomeric 1,4-*s*-$AAAG^+$ ion was photolyzed at 250 and 213 nm with the results shown in Figure 1c and S6, respectively. In addition to the loss of N_2 from the tetrazole moiety (ion $\mathbf{2}^+$, m/z 634), UVPD-MS^2 showed a series of internal $\boldsymbol{\beta}_n$ fragment ions from consecutive dissociations of $\mathbf{2}^+$ (blue-annotated m/z 577, 506, 435, and 364) that were analogous to those from $\mathbf{1}^+$. For ion elemental compositions cf. Table S11 (Supporting Information). Peptide chain dissociations without nitrogen loss (red-annotated ions) and those starting from nitrile imine fragments (green-annotated ions) were also observed. Similar to 1,2-*s*-$AAAG^+$, CID-MS^2 of 1,4-*s*-

https://doi.org/10.1021/jasms.4c00317
J. Am. Soc. Mass Spectrom. 2024, 35, 3070−3088

$AAAG^+$ gave a straight b_n ion series while no loss of N_2 was observed (Figure S7, Table S12). CID-MS^3 of ion $\mathbf{2}^+$ showed a facile loss of GlyOH (m/z 559) which could be assigned to represent the fraction of ions with the straight peptide chain (Figure 1d). However, most of the CID-MS^3 products from $\mathbf{2}^+$ were $\boldsymbol{\beta}_n$ internal fragments by loss of [Gly] (m/z 577) in combination with $[(Ala)_n]$ (m/z 506, 435). In addition, loss of [AlaAla] (m/z 492) was observed, and cleavage of the Ala_1 N–C_α bond formed the complementary fragment ions at m/z 381, 254, and 228. According to their accurate m/z (Table S13, Supporting Information), these ions involved the formation of $HC(CH_3)CO$-Ala-Ala-$NHCH_2CO$ neutral or ion fragments that indicated cyclization to the nitrile imine by the Gly carboxyl group.

Note on $\boldsymbol{\beta}_n$ versus (b_i + H_2O) Ions. The above-discussed $\boldsymbol{\beta}_n$ ions are formally analogous to (b_i + H_2O) fragment ions that have been reported by Gaskell and co-workers to be formed by C-terminal carboxyl participation in amide bond cleavage, resulting in OH group and proton migration.[49−51] The two types of ions differ in structure, with $\boldsymbol{\beta}_n$ ions having the carboxyl oxygen incorporated in the nitrile-imine moiety as a result of cross-linking, whereas the (b_i + H_2O) are C-terminally truncated peptide ions. Since the ions are isomeric, their distinction was of concern with regard to peptide-nitrile imine cross-linking. The formation of (b_i + H_2O) fragment ions has been reported to be sequence dependent, and where present in the spectra, they accompanied as minor satellites the more common b_i ions.[50] Perusal of the CID-MS^2 spectra of 1,2-*s*-$AAAG^+$ (Figure S3) and 1,4-*s*-$AAAG^+$ (Figure S7) revealed no (b_i + H_2O) ions accompanying the b_n ion series, indicating that this peptide sequence in the scaffolds did not undergo backbone rearrangements with OH migration. In addition, we investigated CID-MS^2 spectrum of a 1,2-*s*-$GAAAK^+$ analogue in which the 2,5-diaryltetrazole moiety was replaced by a benzamide group (1,2-*b*-$GAAAK^+$, m/z 646, Figure S8, Supporting Information) to exclude N_2 loss. The spectrum showed series of b_n and complementary $[y_n + 2H]^+$ fragment ions, but no (b_i + H_2O) ions were formed. From these results, we inferred that the peptide sequences used in the conjugates were not conducive to forming (b_i + H_2O) ions, thus affirming the $\boldsymbol{\beta}_n$ ion series assignment as cross-link markers. For the discussion of b_4 fragment ions see Figure S9 (Supporting Information).

We utilized the combined intensities of the identified straight-chain and cyclized ions to estimate the cross-linking yield. The fraction of cyclized ions was represented by the sum of the $\boldsymbol{\beta}_n$ ion intensities (m/z 577, 506, and 435) along with that of the loss of water ion (m/z 616) that was absent in the CID spectra of linear peptide ions and may indicate a cyclized structure. The fraction of noncyclized ions was represented by the sum of the ion intensities corresponding to the b_n and (b_n − C_6H_5N) ions. The data are summarized in Table 1. The more extensive fragmentation on UVPD at 213 nm was reflected in the ($M-N_2$) ion survivor fractions that were 0.9–5% when compared with the 250 nm data (38–60%). The cross-linked ion fractions, 27% and 38% for 1,2-*s*-AAAG and 1,4-*s*-AAAG, respectively, at 213 nm were lower compared to the 250 nm data that gave 65% and 77%, respectively. The high fraction of cross-links from 1,4-*s*-AAAG was somewhat unexpected because the scaffold's *trans*-1,4-stereochemistry was designed to place the peptide and tetrazole moieties at the remote opposite sides of the cyclohexane ring and thus hamper their mutual steric accessibility. This issue is addressed later in the paper when we discuss the ion structures. The effect of blocking the Gly carboxyl as a methyl ester, 1,2-*s*-AAAG-OCH_3^+, was very dramatic, resulting in nearly complete suppression of cross-linked fragment ions on UVPD-MS^2. Cross-linking yields were also determined for CID-MS^3 of nitrile-imine ions generated by UVPD at 213 and 250 nm (Table 1). The 250 nm data for $\mathbf{1}^+$ and $\mathbf{2}^+$ showed high fractions of cross-links, 89% and 84%, respectively. The observed increase of the cross-link fractions from UVPD-MS^2 to CID-MS^3 was understandable, as the linear peptide ions were likely to be depleted on UVPD because of their facile backbone dissociation, leaving a proportionately larger fraction of cyclic cross-links to be probed in the next step by CID-MS^3. This effect was particularly large for the methyl esters where the cross-links did not dissociate on UVPD and were then revealed by CID-MS^3. Nonetheless, the fraction of cross-linked methyl esters was notable lower than that of the carboxyl-terminated peptide conjugates.

Table 1. Crosslinking Yields of Peptide-Nitrile-Imine Scaffolds

Compound/ion	Ion activation	% ($M-N_2$)[a]	% Cross-links[a]
UVPD-MS^2			
1,2-*s*-AAAG	250 nm	46	65
1,2-*s*-AAAG	213 nm	0.9	27
1,4-*s*-AAAG	250 nm	60	77
1,4-*s*-AAAG	213 nm	5	38
1,2-*s*-AAAG-OCH_3	250 nm	38	<0.1
1,2-*s*-AAAG-OCH_3	213 nm	2.7	<0.1
1,2-*s*-AAAHG	213 nm	42	65
1,4-*s*-AAAHG	213 nm	47	60
1,2-*s*-AAAHG-OCH_3	213 nm	38	<0.1
1,2-*s*-GAAAK	250 nm	77	28
1,2-*s*-GAAAK	213 nm	32	66
1,4-*s*-GAAAK	213 nm	38	78
1,2-*s*-GAAAK-OCH_3	213 nm	48	15
1,2-*s*-GAAAR	213 nm	52	48
1,2-*s*-GAAAR-OCH_3	213 nm	47	40
1,4-*s*-GAAAR	213 nm	72	58
CID-MS^3			
$\mathbf{1}^+$ (from 1,2-*s*-AAAG at 250 nm)			89
$\mathbf{1}^+$ (from 1,2-*s*-AAAG at 213 nm)			74
$\mathbf{2}^+$ (from 1,4-*s*-AAAG at 250 nm)			84
$\mathbf{2}^+$ (from 1,4-*s*-AAAG at 213 nm)			57
$\mathbf{1}^+$-OCH_3 (from 1,2-*s*-AAAG-OCH_3 at 250 nm)			28
$\mathbf{1}^+$-OCH_3 (from 1,2-*s*-AAAG-OCH_3 at 213 nm)			37
$\mathbf{3}^+$ (from 1,2-*s*-AAAHG at 250 nm)			87
$\mathbf{3}^+$ (from 1,2-*s*-AAAHG at 213 nm)			91
$\mathbf{4}^+$ (from 1,4-*s*-AAAHG at 213 nm)			91
3-OCH_3^+ (from 1,2-*s*-AAAHG-OCH_3 at 213 nm)			2.1
$\mathbf{5}^+$ (from 1,2-*s*-GAAAK at 250 nm)			94
$\mathbf{5}^+$ (from 1,2-*s*-GAAAK at 213 nm)			93
5-OCH_3^+ (from 1,2-*s*-GAAAK-OCH_3 at 213 nm)			22
$\mathbf{6}^+$ (from 1,4-*s*-GAAAK at 213 nm)			97
$\mathbf{7}^+$ (from 1,2-*s*-GAAAR at 213 nm)			74
$\mathbf{7}^+$ (from CID of 1,2-*s*-GAAAR)			74
7-OCH_3^+ (from 1,2-*s*-GAAAR-OCH_3 at 213 nm)			74
7-OCH_3^+ (from CID of 1,2-*s*-GAAAR-OCH_3)			78
$\mathbf{8}^+$ (from 1,4-*s*-GAAAR at 213 nm)			67
$\mathbf{8}^+$ (from CID of 1,4-*s*-GAAAR)			62

[a]Relative to the sum of all identified fragment ion intensities.

https://doi.org/10.1021/jasms.4c00317
J. Am. Soc. Mass Spectrom. 2024, 35, 3070−3088

***cis*-1,2-Cyclohexane-diaryltetrazole-AAAHG (1,2-*s*-AAAHG) and *trans*-1,4-Cyclohexane-diaryltetrazole-AAAHG (1,4-*s*-AAAHG) scaffolds.** To utilize nitrile-imine formation by both UVPD and CID, we investigated scaffold conjugates that were furnished with the basic His, Lys, and Arg amino acid residues. In this way, the peptide backbone fragmentation was suppressed by proton sequestration at the basic site, allowing us to generate nitrile-imine intermediates by CID-MS^2 and study their properties by CID and UVPD-MS^3, as well as by ion mobility. Both CID and photodissociation of the 1,2-*s*-$AAAHG^+$ ions gave abundant fractions of the denitrogenated intermediates $\mathbf{3}^+$ (m/z 771, Figure 2a, S10a,b Table S14,S15). Backbone dissociations were chiefly observed for 213 nm UVPD (Figure S10a) where they primarily led to two ion series, both originating from $\mathbf{3}^+$. In contrast, competitive peptide chain dissociations in tetrazole conjugates, i.e. without N_2 loss, were only minor as indicated by the red-annotated ions. The main series of blue-annotated ions in the CID-MS^3 spectra corresponded to $\boldsymbol{\beta}_n$ fragments by loss of [Gly] (m/z 714), and continuing by sequential losses of [His] and [Ala] at m/z 577, 506, 435, and 364 (Figure 2b). In addition, the presence of the basic His residue steered protonation to complementary $\boldsymbol{\gamma}_n$ ions at m/z 266, 337, and 408 that were also indicative of cyclized structures. The $\boldsymbol{\gamma}_n$ ions formally corresponded to $[y_n + 2H - H_2O)^+$, but because of the absence of $[y_n + 2H)^+$ ions, they were better represented as a separate ion series. The cross-linking yields at 213 nm for 1,2-*s*-$AAAHG^+$ and 1,4-*s*-$AAAHG^+$ were similar at 65 and 60%, respectively (Table 1). A significant drop of yield was recorded for 1,2-*s*-$AAAHG\text{-}OCH_3^+$ (<0.1%), indicating the critical role of the free carboxyl in binding to the nitrile imine.

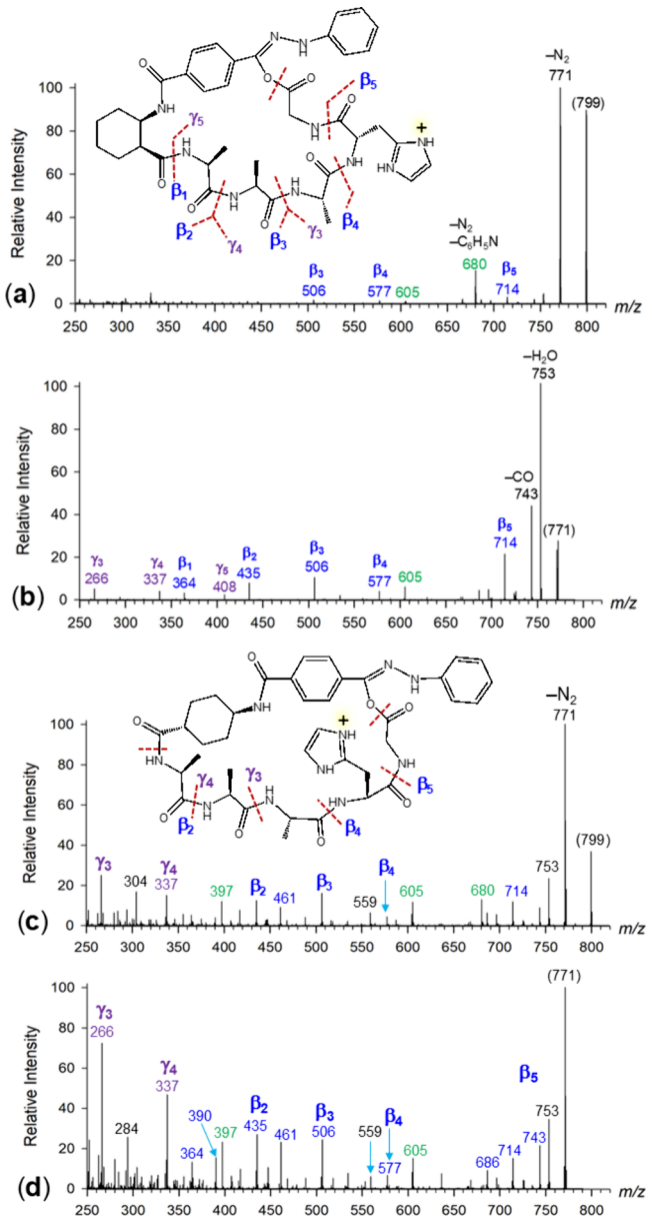


Figure 2. (a) UVPD-MS^2 at 250 nm of 1,2-*s*-$AAAHG^+$ (m/z 799). (b) CID-MS^3 of $(1,2\text{-}s\text{-}AAAHG - N_2)^+$ ($\mathbf{3}^+$, m/z 771). (c) UVPD-MS^2 of 1,4-*s*-$AAAHG^+$. (d) CID-MS^3 of $(1,4\text{-}s\text{-}AAAHG - N_2)^+$ ($\mathbf{4}^+$, m/z 771). The fragment ion series are annotated as in Figure 1.

We utilized the complementary formation of the $(1,2\text{-}s\text{-}AAAHG - N_2)^+$ ions ($\mathbf{3}^+$, m/z 771) by CID and UVPD to investigate their further dissociations in the UVPD(213 nm)-CID-MS^3 and CID-UVPD(213 nm)-MS^3 sequences (Figure 2b, S11, Table S16−S18). CID-MS^3 of $\mathbf{3}^+$ showed prominent peaks by loss of water and CO at m/z 753 and 743, respectively (Figure 2b). Although these losses do not explicitly indicate cyclization, they were not observed for the 1,2-*s*-$AAAHG^+$ ions and arguably can be counted to originate from cyclized structures. The presence of linear, noncyclized, structures was indicated by the m/z 605 ion which was formed by the combined loss of GlyOH and C_6H_5N. Most fragment ions were of the $\boldsymbol{\beta}_n$ type, m/z 714, 577, 506, 435, and 364, that were due to loss of internal residues and thus represented cyclized structures. UVPD of $\mathbf{3}^+$ showed similar results (Figure S11), although because of the higher excitation energy the sequence ion series reached to lower-mass ions by more extensive dissociation of light-absorbing intermediate fragment ions.

Photodissociation of 1,4-*s*-$AAAHG^+$ (m/z 799) at 213 nm gave an intense $(1,4\text{-}s\text{-}AAAHG - N_2)^+$ ion ($\mathbf{4}^+$, m/z 771) along with a series of backbone fragment ions (Figure 2c, Table S19). These consisted of the $\boldsymbol{\beta}_n$ series of sequential losses of [Gly], [His], and [Ala] internal fragments (blue-annotated ions), indicating cross-linking. The other main series was initiated by loss of C_6H_5N from the nitrile imine (m/z 680), followed by standard backbone cleavages yielding b_n type fragment ions (green-annotated series). Loss of N_2 was also induced by CID that was followed by backbone cleavages (Table S20). CID-MS^3 (Figure 2d, Table S21) and UVPD-MS^3 of CID-generated $\mathbf{4}^+$ (m/z 771, Table S22), as well as CID-MS^3 of UVPD-generated $\mathbf{4}^+$ (Table S23) showed mainly the $\boldsymbol{\beta}_n$ ions that were complemented by the $\boldsymbol{\gamma}_3$ and $\boldsymbol{\gamma}_4$ ions at m/z 266 and 337, respectively. Both the sequential fragmentation starting from the C-terminal Gly, and the absence of direct loss of $[Ala]_n$ from $\mathbf{4}^+$ indicated that cross-linking to the nitrile imine was chiefly realized by the Gly residue. The linkage and peptide chain dissociations could be depicted by structures and ester and amide bond cleavages analogous to those shown in Scheme 2 for 1,2-*s*-AAAG (Figure 2a,c). The cross-linking yields upon CID of $\mathbf{3}^+$ and $\mathbf{4}^+$ that were generated at 213 nm were comparable at 91% (Table 1). Ions $\mathbf{3}^+$ produced by UVPD at 250 nm gave a similar yield of cross-links (87%) upon CID. In contrast, carboxyl methylation, as in $\mathbf{3}^+\text{-}OCH_3$, caused a large drop in the cross-linking yield (2.1%, Table 1) which was consistent with the yield decrease upon UVPD of 1,2-*s*-$AAAHG\text{-}OCH_3^+$ methyl ester. Con-

https://doi.org/10.1021/jasms.4c00317
J. Am. Soc. Mass Spectrom. 2024, 35, 3070−3088

versely, this large effect of blocking the carboxyl indicated that the His side-chain imidazole ring did not react with the nitrile imine to affect the cross-linking yield. The fragment ion assignment was corroborated by accurate mass measurements (Table S24−S27, Supporting Information).

Peptide Scaffolds Containing Lysine and Arginine Residues, 1,2-*s*-GAAAK, 1,2-*s*-GAAAR 1,4-*s*-GAAAK, and 1,4-*s*-GAAAK. The above-discussed results from the 1,2-*s*-AAAHG and 1,4-*s*-AAAHG scaffolds indicated that cross-linking chiefly involved the C-terminal residue. To elucidate the effect of basic side-chain groups, we synthesized scaffolds of the 1,2-*s*-GAAAK, 1,2-*s*-GAAAR 1,4-*s*-GAAAK, and 1,4-*s*-GAAAR types, where the charge-carrying side chain could provide a proton to catalyze cross-linking and engage as a nucleophile in attacking the nitrile-imine dipole. UVPD-CID-MS^3 and CID-UVPD-MS^3 of these scaffold ions were studied at 213 nm, including accurate mass measurements for fragment ion identification.

1,2-*s*-GAAAK. UVPD at 213 nm of 1,2-*s*-$GAAAK^+$ (m/z 790) resulted in loss of N_2 followed by backbone dissociations of the $(1,2\text{-}s\text{-}GAAAK{-}N_2)^+$ ion ($\mathbf{5}^+$, m/z 762, Figure 3a, Table S28). The dissociations produced two main series of fragment ions. The sequential loss of [Lys], [Ala], and [Gly] gave rise to ions at m/z 634, 563, 492, 421, and 364 (blue annotated ions in Figure 3a). These represented the cross-linked fraction of $\mathbf{5}^+$ (93%, Table 1) where dissociation started from the C-terminal Lys residue while the carboxyl OH was retained in the $\boldsymbol{\beta}_n$ ions. The other main series (green annotated ions) started from $\mathbf{5}^+$ by loss of C_6H_5N (m/z 671) that was followed by regular backbone cleavages forming the b_n ions at m/z 525, 454, 383, and 312. CID-MS^2 of 1,2-*s*- $GAAAK^+$ gave very similar results (Figure S12a, Table S29, Supporting Information). CID-MS^3 of $\mathbf{5}^+$ generated by UVPD (Figure 3b) and CID (Figure S12b) gave very similar spectra that were dominated by the $\boldsymbol{\beta}_n$ ions. In addition, CID-MS^3 led to a loss of $C_6H_8N_2$ (m/z 654) most likely phenylhydrazine, that has been observed for other peptide-nitrile imine conjugates (Table S30, S31).[22] Ions $\mathbf{5}^+$ generated by UVPD at both 250 and 213 nm showed large fractions of cross-links upon CID analysis (94 and 93%, respectively, Table 1). In contrast to the CID spectrum, UVPD of $\mathbf{5}^+$ resulted in the loss of C_6H_5N followed by backbone cleavages forming the b_n ions (Figure S12c). The Figure 3b and S12c data apparently gave conflicting results, as the $\boldsymbol{\beta}_n$ ions represented the cross-linked fraction whereas the b_n ions were deemed to be indicative of the straight chain isomers. The different dissociations on CID and UVPD can be explained by the very different modes of ion excitation. CID under multiple collision conditions follows the slow heating internal energy build up,[52] promoting low activation energy reactions. This is depicted in Scheme 3 by proton transfer onto the phenylhydrazone moiety in cross-linked $\mathbf{5}^+$, resulting in the weakening and cleavage of the lysine CO−O bond and triggering the formation of the $\boldsymbol{\beta}_n$ ion series. In contrast, photoexcitation chiefly affects the aromatic ring chromophore, resulting in a photochemical cleavage of the C−OCO bond that restores the open-chain lysine residue. The internal energy supplied by the photon (562 kJ mol^{-1} at 213 nm) then drives standard peptide amide dissociations forming the b_n ions. That the lysine carboxyl was important in cross-linking was supported by the results of UVPD of the related methyl ester, 1,2-*s*-GAAAK-OCH_3^+. UVPD of the latter ion (m/z 804) and UVPD-CID-MS^3 as well as CID-CID-MS^3 of the denitrogenated intermediate ($\mathbf{5}$-OCH_3^+, m/z 776) showed

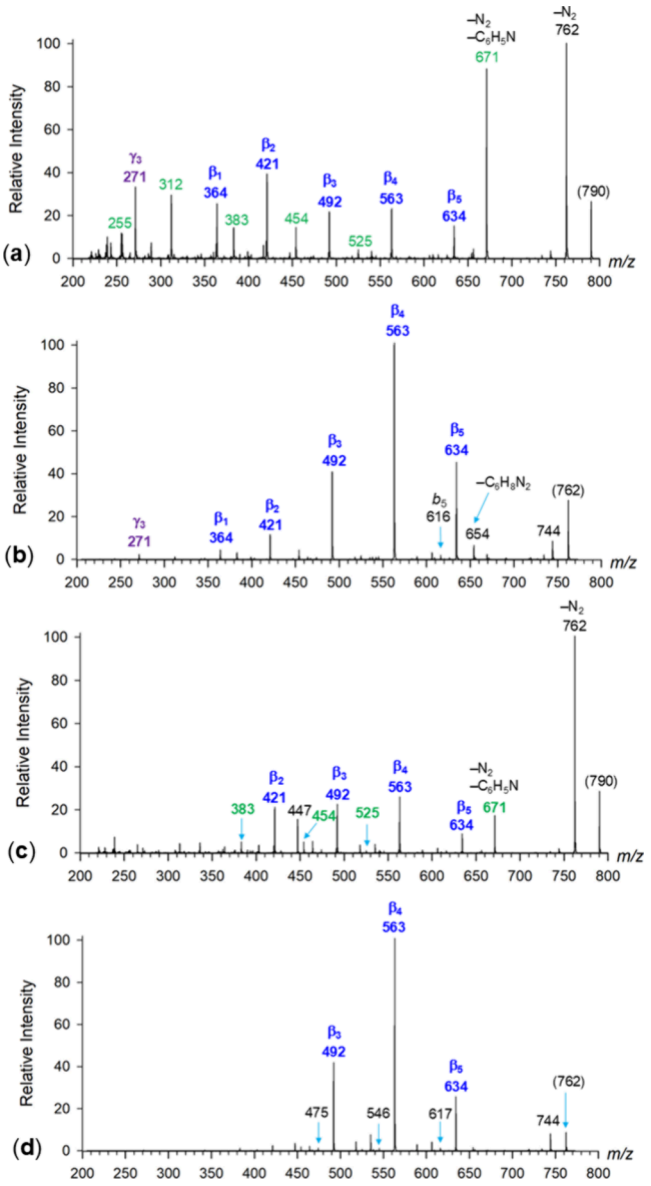


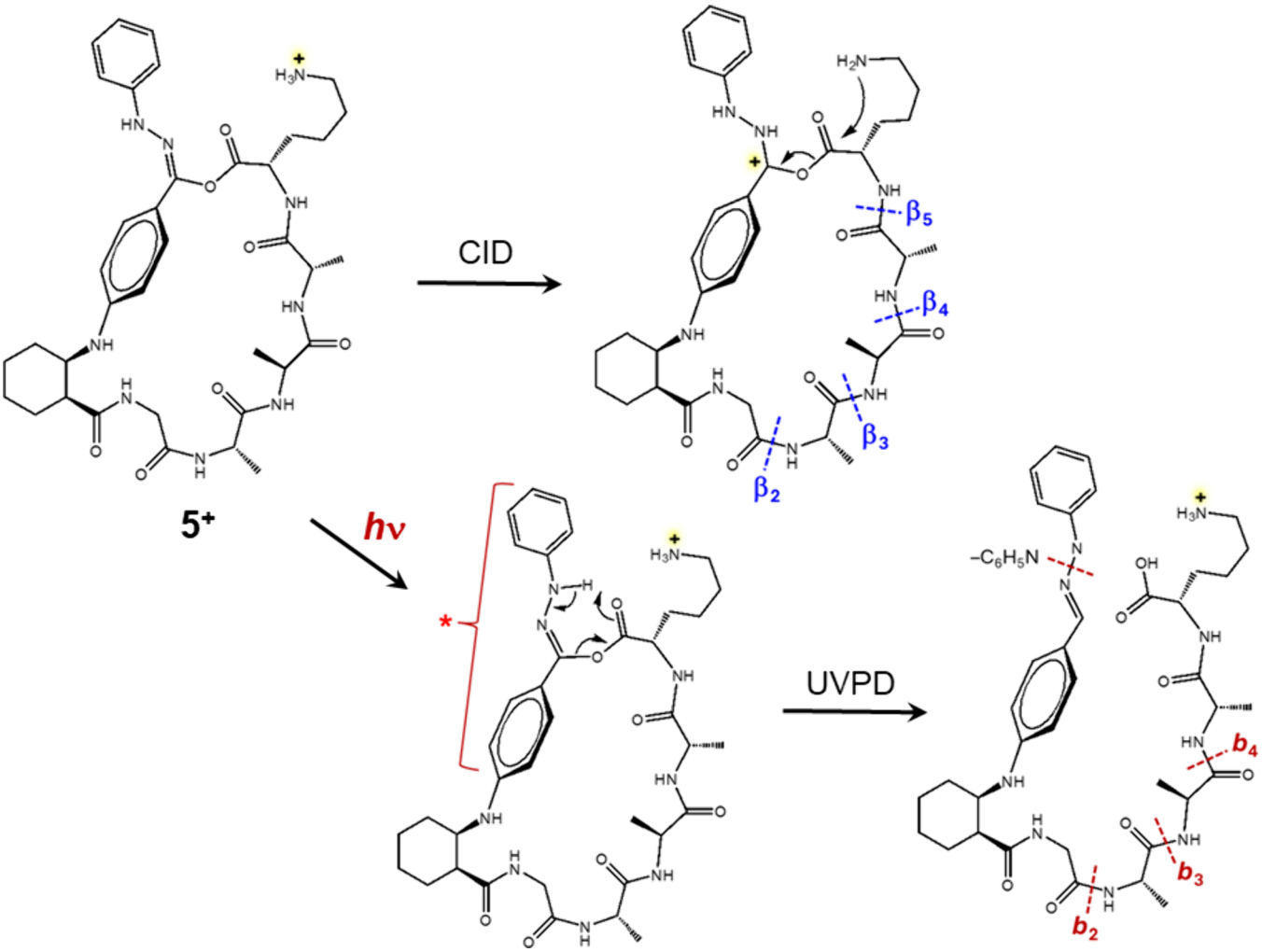
Figure 3. (a) UVPD-MS^2 of 1,2-*s*-$GAAAK^+$ (m/z 790) at 213 nm. (b) UVPD-CID-MS^3 of $\mathbf{5}^+$ (m/z 762). (c) UVPD-MS^2 of 1,4-*s*-$GAAAK^+$ (m/z 790) at 213 nm. (d) UVPD-CID-MS^3 of $\mathbf{6}^+$ (m/z 762). The fragment ion series are annotated as in Figure 1.

chiefly the b_n ion series after loss of C_6H_5N, whereas the $\boldsymbol{\beta}_n$ ions, albeit present, were only minor (Figure S13a,b, Table S32−S35, Supporting Information). The loss of $C_6H_5NHNH_2$ from $\mathbf{5}$-OCH_3^+ was a major dissociation (m/z 668, Figure S13b), indicating a facile proton transfer onto the nitrile imine group.

The important role of the carboxyl in cross-linking was also emphasized by comparing the yields for 1,2-*s*-GAAAK-OCH_3^+ and 1,2-*s*-$GAAAK^+$ (Table 1). Carboxyl blocking by methylation caused a significantly lower cross-linking yields in both UVPD-MS^2 and CID-MS^3 of $\mathbf{5}$-OCH_3^+ (15 and 22%, respectively). Interestingly, the drop for 1,2-*s*-GAAAK-OCH_3^+ was not as severe as for 1,2-*s*-AAAG-OCH_3^+ and 1,2-*s*-AAAHG-OCH_3^+ (*vide supra*), indicating that a minor fraction of 1,2-*s*-GAAAK-OCH_3^+ might use the Lys amine group in reacting with the nitrile imine.

https://doi.org/10.1021/jasms.4c00317
J. Am. Soc. Mass Spectrom. 2024, 35, 3070−3088

Scheme 3. Different Dissociations of Crosslinked Intermediate 5^+ upon Collision Activation and 213 nm Photon Absorption



1,4-*s*-GAAAK. The *trans*-1,4-cyclohexane conjugate 1,4-*s*-GAAAK was investigated by a combination of sequential excitations by UVPD and CID (Table S36–S39). The UVPD-MS^2 spectrum of 1,4-*s*-GAAAK$^+$ (m/z 790) showed a series of $\boldsymbol{\beta}_n$ ions originating from the denitrogenated intermediate $\mathbf{6}^+$ (m/z 762, Figure 3c). CID-MS^3 of $\mathbf{6}^+$ (m/z 762, Figure 3d) was representative of the ion structure and dissociations. The spectrum showed a series of dominant $\boldsymbol{\beta}_n$ ions at m/z 634, 563, and 492 by sequential loss of [Lys] and [Ala] that resembled the analogous ion series produced from the 1,2-*s*-GAAAK$^+$ ion $\mathbf{5}^+$. A minor C_6H_5N loss on UVPD was observed (m/z 671), but its subsequent backbone cleavage to form b_n ions was weak (m/z 525, 454, 383, Figure 3d). CID-MS^2 of 1,4-*s*-GAAAK$^+$ was similar to the UVPD spectrum, showing loss of N_2 and $\boldsymbol{\beta}_n$ ions (Figure S14a). Both UVPD-MS^2 of 1,4-*s*-GAAAK$^+$ and CID-MS^3 of $\mathbf{6}^+$ showed substantial cross-linking yields, 78 and 97%, respectively (Table 1). Further UVPD at 213 nm of $\mathbf{6}^+$ resulted in deep dissociation, forming the $C_{14}H_9N_2O$ and $C_{14}H_{11}N_2O_2$ ions that contained the diarylnitrile imine and its oxygenated form, respectively (Figure S14b). Prominent $\boldsymbol{\beta}_n$ and (b_n − C_6H_5N) ion series were also observed We interpreted the UVPD and CID-MS^n data as being indicative of prevalent cross-linking of Lys to the transient nitrile imine.

1,2-*s*-GAAAR. UVPD at 213 nm of 1,2-*s*-GAAAR$^+$ (m/z 818) showed some new features that were specific for Arg (Figure 4a, Table S40). In addition to the $\boldsymbol{\beta}_n$ ion series (m/z 563, 492, 421, and 364) and combined loss of N_2 and C_6H_5N (m/z 699), there was a prominent ion by loss of $C_7H_7NO_2$ (m/z 653). CID-MS^3 of the denitrogenated ion $\mathbf{7}^+$ (m/z 790, Table S41) showed a similar fragmentation pattern that was dominated by the $\boldsymbol{\beta}_n$ ion series and the m/z 762 (loss of CO) and m/z 653 ions (Figure 4b). Ion activation by the CID-MS^2 and CID-CID-MS^3 sequences gave similar results (Figure S15a,b, Table S42,S43, Supporting Information). The loss of $C_7H_7NO_2$ pointed to carboxyl cross-linking to the nitrile imine phenyl group that was unprecedented with other peptide sequences. This was further studied by CID-MS^4 of the m/z 762 (loss of CO) ion that gave a complete $\boldsymbol{\beta}_n$ ion series but no m/z 653 (Figure S15c), indicating that the $C_7H_7NO_2$ neutral fragment was eliminated in a single step as a distinct moiety containing the nitrile imine phenyl linked to the carboxyl group and including two hydrogen atoms transferred from the peptide chain. An interesting feature of the Figure S15c spectrum was the elimination of an arginine $C_5H_{12}N_4$ neutral fragment, m/z 762.4029 → m/z 634.2971, starting the $\boldsymbol{\beta}_4$ sequence which clearly indicated cross-linking by the arginine C-terminus. This was further investigated by blocking the free carboxyl as a methyl ester in 1,2-*s*-GAAAR-OCH_3 (Table S44, S45). CID-MS^3 of the denitrogenated ion (7-OCH_3^+, m/z 804, Figure S16, Table S47) displayed three major fragment ion series. The classical b_n ions (m/z 346, 403, 474, and 545) indicated a fraction of non-cross-linked nitrile imine isomers

https://doi.org/10.1021/jasms.4c00317
J. Am. Soc. Mass Spectrom. 2024, 35, 3070−3088

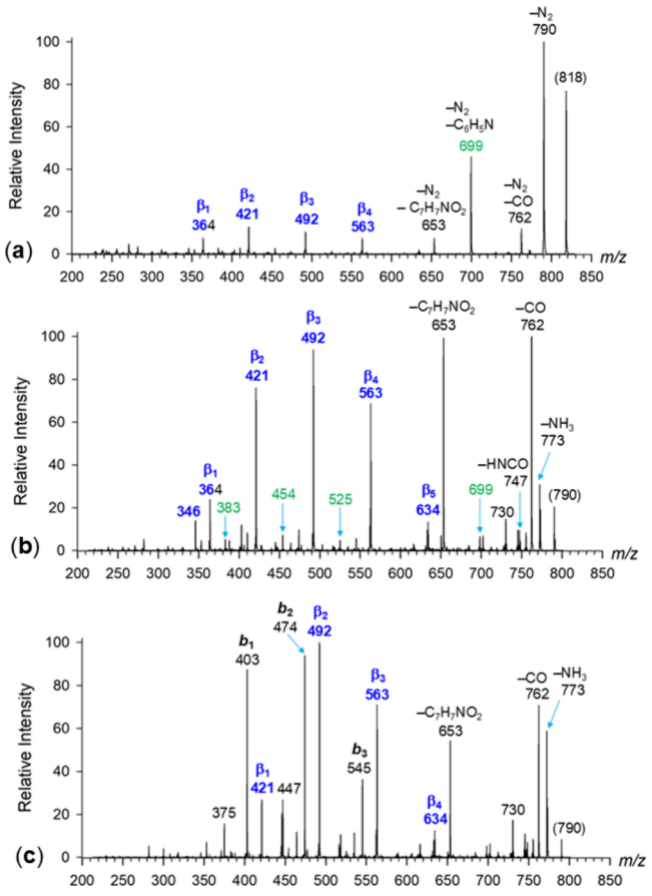


Figure 4. (a) UVPD at 213 nm of 1,2-*s*-GAAAR$^+$ (m/z 818); (b) CID-MS^3 of **7**$^+$ (m/z 790); (c) CID-MS^3 of **8**$^+$ (m/z 790).

that were analogous to those in the spectrum of the 1,2-*s*-GAAAK derivative **5**-OCH$_3^+$ (Figure S13b, Table S35). In contrast to **5**-OCH$_3^+$, the Arg methyl ester (**7**-OCH$_3^+$) showed a major loss of ammonia but virtually no loss of C_6H_5N, which had a major effect on the Figure S16 spectrum. The ammonia loss was followed by internal fragmentations leading to a sequential loss of two [Ala] units. This indicated cyclic peptide structures, although it was not clear if they were present in the **7**-OCH$_3^+$ ion population or produced upon CID by the ammonia loss. Finally, the CID-MS^3 spectrum of **7**-OCH$_3^+$ displayed a series of fragment ions starting with the elimination of $C_7H_{13}N_3O_2$ (m/z 633) from the Arg-OCH$_3$ residue and followed by sequential losses of [Ala] and [Gly] units (m/z 562, 491, 420, and 363, Figure S16), forming a new $\boldsymbol{\beta}_n$ ion series. Note that these fragment ions had different formulas (Table S46, Supporting Information) than those in the Figure 4b spectrum, which indicated cross-linking by an NH$_2$ group of the Arg side-chain, followed by N−C bond cleavage leaving the guanidine NH$_2$ group incorporated in the $\boldsymbol{\beta}_n$ ions. Guanidine cross-linking was also evident from the small effect that carboxyl blocking had on the cross-linking yields of 1,2-*s*-GAAAR-OCH$_3$ and ion **7**-OCH$_3^+$. The Table 1 data showed that the cross-linking yield upon UVPD dropped by only 8% upon carboxyl methylation in 1,2-s-GAAAR-OCH$_3$ and virtually did not change in CID of **7**-OCH$_3^+$, amounting to 74−78% for ions generated by UVPD or CID-MS^2.

1,4-*s*-GAAAR. The GAAAR peptide sequence was investigated as to its propensity for nitrile- imine cross-linking when mounted on the *trans*-1,4-cyclohexane scaffold. Both UVPD and CID-MS^2 of 1,4-*s*-GAAAR$^+$ (m/z 818, Figure S17a,b, Table S48, S49, Supporting Information) resulted in loss of N_2, and the resulting ions **8**$^+$ (m/z 790) were subjected to further investigations. CID-MS^3 of **8**$^+$ (Figure 4c, S17c, Table S50, S51) showed losses of NH_3, CO, and $C_7H_7NO_2$ at m/z 773, 762, and 653, respectively, that were analogous to the dissociations of **7**$^+$. In addition, we observed a $\boldsymbol{\beta}_n$ ion series (m/z 634, 563, 492, and 421) that was indicative of cross-links, and a b_n ion series (m/z 545, 474, 403) that were formed from the fraction of linear chain isomers. Compared with the sterically more favorable **7**$^+$, ion **8**$^+$ showed a smaller fraction of cross-links, as expressed by the 67% yield (Table 1). The loss of $C_7H_7NO_2$ (m/z 653) was indicative of a proton and carboxyl transfer to the nitrile imine C_6H_5N moiety and was investigated by CID-MS^4, as detailed in the Supporting Information (Figure S18a,b).

It follows from the above spectra and data interpretation, that chemical bond-forming reactions resulting in peptide-nitrile-imine cross-linking occurred in gas phase ions with both the *cis*-1,2-cyclohexane and across the *trans*-1,4-cyclohexane scaffolds. The latter was not quite expected from the point of view of organic stereochemistry where *trans*-1,4-substituents on cyclohexane are considered remote and their mutual approach is hindered by the ring geometry. Similar arguments have been used to discuss the stereochemistry of water elimination from various cyclohexanol ions in the gas phase,[53−56] and backbone dissociations of *cis*- and *trans*-1,4-cycloornithine peptide isomers that showed dramatic stereochemical effects.[26] The cross-links in the present 1,2-*s*-peptide and 1,4-*s*-peptide ions chiefly involved the C-terminal residues which can be related to steric constraints. These were determined by the nucleophile reaching the nitrile-imine group within the stiff, rod-like, diaryl moiety. The nature of the cross-links was gleaned from the major $\boldsymbol{\beta}_n$ ion series that preserved an oxygen atom from the carboxyl group. This was observed for C-terminal Gly and Lys and also for peptides terminated with HisGly. This indicated that, despite their steric availability, the side-chain nucleophilic imidazole and amine group in His and Lys, respectively, did not compete with the carboxyl for reacting with the nitrile imine. In contrast, the guanidine group of Arg displayed specific reactivity toward the nitrile imine that was evidenced by NH$_2$ transfer. Since the Arg guanidine is by far the most basic group in 1,2-*s*-GAAAR and 1,4-*s*-GAAAR, it is likely to be protonated in the gas-phase ions, and its reactions as a nucleophile must be preceded by proton transfer. The possible mechanism for the Arg-guanidine cross-link formation will be addressed later in the paper along with fully optimized ion structures. To gain more insight into the pertinent ion structures we carried out ion mobility measurements and attempted to relate the experimental collision cross sections (CCS_{IMS}) to theoretical values (CCS_{calc}) that we calculated for several low-energy tautomers and conformers of the ions in question.

Ion Mobility, Collision Cross Sections, and Ion Structures. The UVPD and CID results pointed to extensive cross-linking to the transiently generated nitrile imine group by the peptide chains in all peptide sequences we investigated. The sequential nature of the internal peptide chain fragments indicated that cross-linking occurred close to the peptide C-terminus regardless of the type of the amino acid residues. Nitrile-imine conjugates of pentapeptides have been shown to undergo unprecedented reactions with the peptide amide bonds, resulting in cyclization.[22] This raised the question of the

https://doi.org/10.1021/jasms.4c00317
J. Am. Soc. Mass Spectrom. 2024, 35, 3070−3088

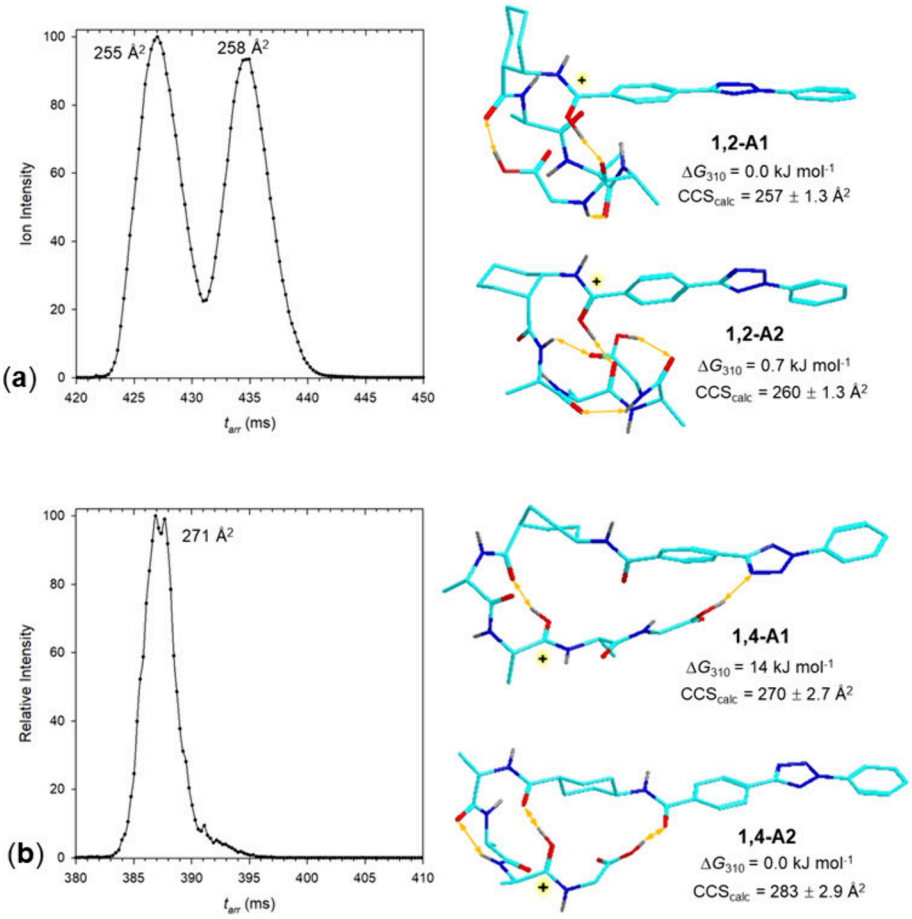


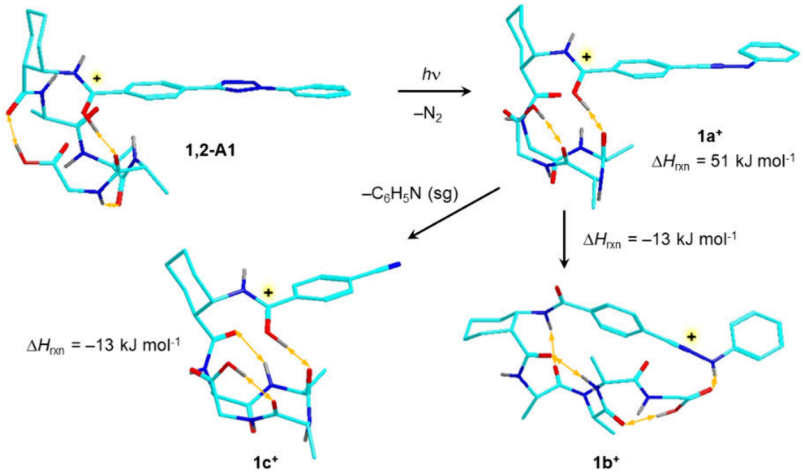
Figure 5. Arrival time distributions (ATD) of (a) 1,2-*s*-AAAG$^+$ and (b) 1,4-*s*-AAAG$^+$ after (a) 20 and (b) 15 passes. M06-2X/6-31+G(d,p) optimized structures, M06-2X/def2qzvpp relative Gibbs energies, and CCS_{calc} of low energy (a) 1,2-*s*-AAAG$^+$ and (b) 1,4-*s*-AAAG$^+$ ions. Atom color coding is as follows: cyan = C, blue = N, red = O, gray = H. Only exchangeable NH, OH hydrogens are shown to avoid clutter. Major hydrogen bonds are indicated by ochre double-headed arrows.

composition of the cross-links in the present scaffold conjugates, as well as the conformation of the precursor ions. Addressing these questions was thought to help elucidate the nature and mechanisms of the cyclization reactions. To investigate the composition of the precursor ion conformers and cross-linking products, we used cyclic ion mobility mass spectrometry along with the determination of collision cross sections (CCS_{IMS}) that were compared and possibly matched with theoretical CCS_{calc} that were obtained for fully optimized ion structures and charge distributions. We note that although the optimized structures are local energy minima at 0 K whereas the ions in ion mobility are thermal, a previous ab initio-molecular dynamics study of peptide ions has indicated very small (ca. 1%) changes of CCS_{calc} when going from 0 to 298 K.[57]

1,2-*s*-AAAG$^+$ and 1,4-*s*-AAAG$^+$. The arrival time distribution (ATD) of 1,2-*s*-AAAG$^+$ obtained after 20 passes (19.6 m path length) showed two peaks with very similar CCS_{IMS} = 255 and 258 Å^2, respectively (Figure 5a). Compared with that, the ATD of 1,4-*s*-AAAG$^+$ (15 passes) gave one peak of CCS_{IMS} = 271 Å^2 (Figure 5b) with a minor tail toward longer arrival times which indicated that the ion was produced as one component or an unresolved mixture of several components of very close CCS. Using a combination of BOMD of several protonation isomers and M06-2X/6-31+G(d,p) geometry optimization of several low-energy structures, we were able to identify the lowest-energy ions and calculate their CCS_{calc}. Figure 5a shows two lowest-energy isomers of 1,2-*s*-AAAG$^+$, **1,2-A1** and **1,2-A2**, that had virtually identical relative energies, and can be considered energetically equivalent at the DFT level of theory. Upon BOMD and gradient optimization of initial structures with several different amide protonation sites, all low-energy structures collapsed to a single protonation isomer being protonated at the diaryltetrazole carboxamide carbonyl. The protonated amides developed strong hydrogen bonds to the Ala amide groups, which were Ala-2 in **1,2-A2** and Ala-3 in **1,2-A1** (Figure 5a). The structures also differed in the orientation of the *cis*-1,2-diaryltetrazole carboxamide and peptide groups on the cyclohexane ring. The peptide chain was equatorial in **1,2-A1** and axial in **1,2-A2**. The hydrogen bonding patterns determined the compact peptide conformations whereas the rigid 2,5-diaryltetrazole moiety stuck out of the structures. This indicated that the incipient nitrile imine was not in the close proximity of the peptide groups, and cross-linking must have involved peptide conformational change. The CCS_{calc} of **1,2-A1** and **1,2-A2**, 257 and 260 Å^2, respectively, matched within 1% the experimental CCS_{IMS} of the peaks detected in the Figure 5a ion mobilogram. It should be noted that the MobCal-MPI CCS calculations that were based on averaging 10 iterations gave 0.5−1.1% root-mean-square deviations which should be considered when comparing the calculated

https://doi.org/10.1021/jasms.4c00317
J. Am. Soc. Mass Spectrom. 2024, 35, 3070−3088

Scheme 4. Dissociation and Crosslinking Reaction Energies of 1,2-*s*-AAAG$^+$



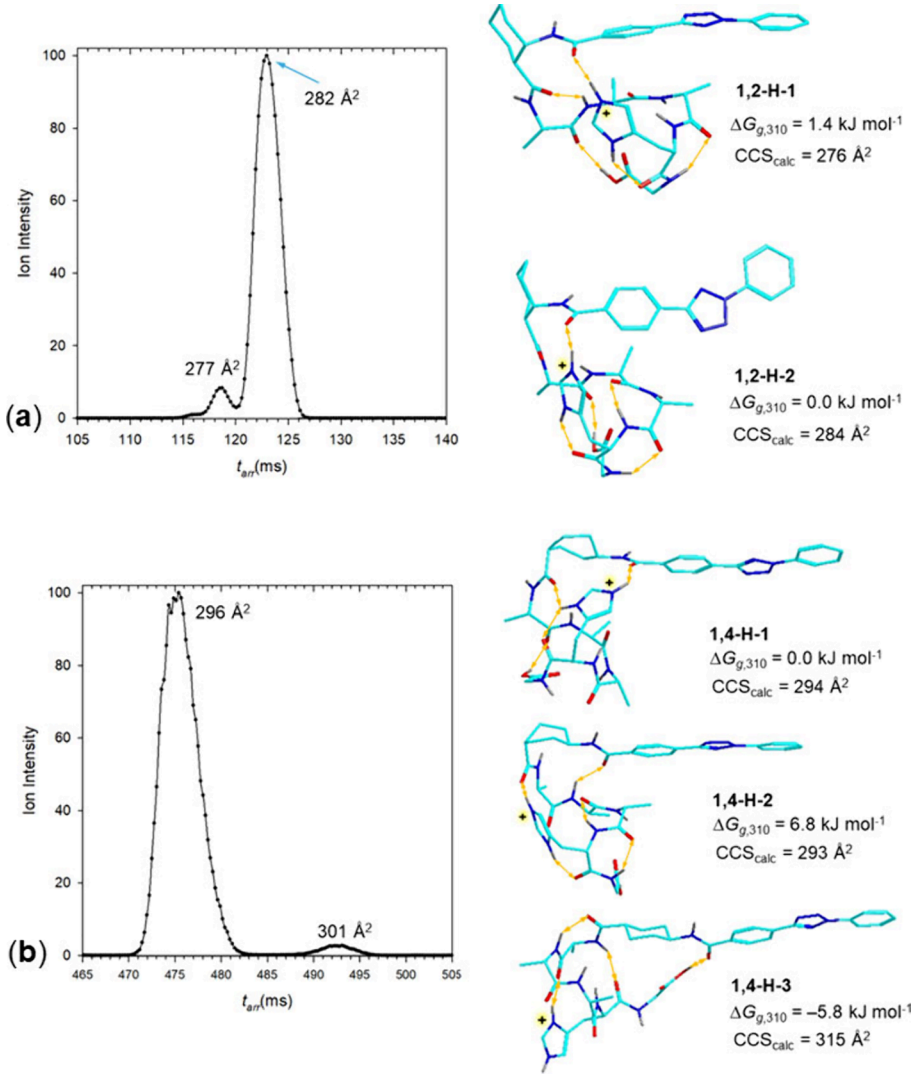


Figure 6. Arrival time distributions of (a) 1,2-*s*-AAAHG$^+$ and (b) 1,4-*s*-AAAHG$^+$ after (a) 4 and (b) 15 passes. M06-2X/6-31+G(d,p) optimized structures, M06-2X/def2qzvpp relative Gibbs energies, and CCS_{calc} of low energy (a) 1,2-*s*-AAAHG$^+$ and (b) 1,4-*s*-AAAHG$^+$ ions. Structure description as in Figure 5.

and experimental CCS. The CCS_{calc} of other low-energy conformers were 2−4% larger and provided a less satisfactory fit with the CCS_{IMS} (Figure S19, Supporting Information).

Several low-energy optimized structures of 1,4-*s*-AAAG$^+$ were obtained that all showed a uniform protonation site that was on the Ala-2 amide (Figure 5b). The structures fell

https://doi.org/10.1021/jasms.4c00317
J. Am. Soc. Mass Spectrom. 2024, 35, 3070−3088

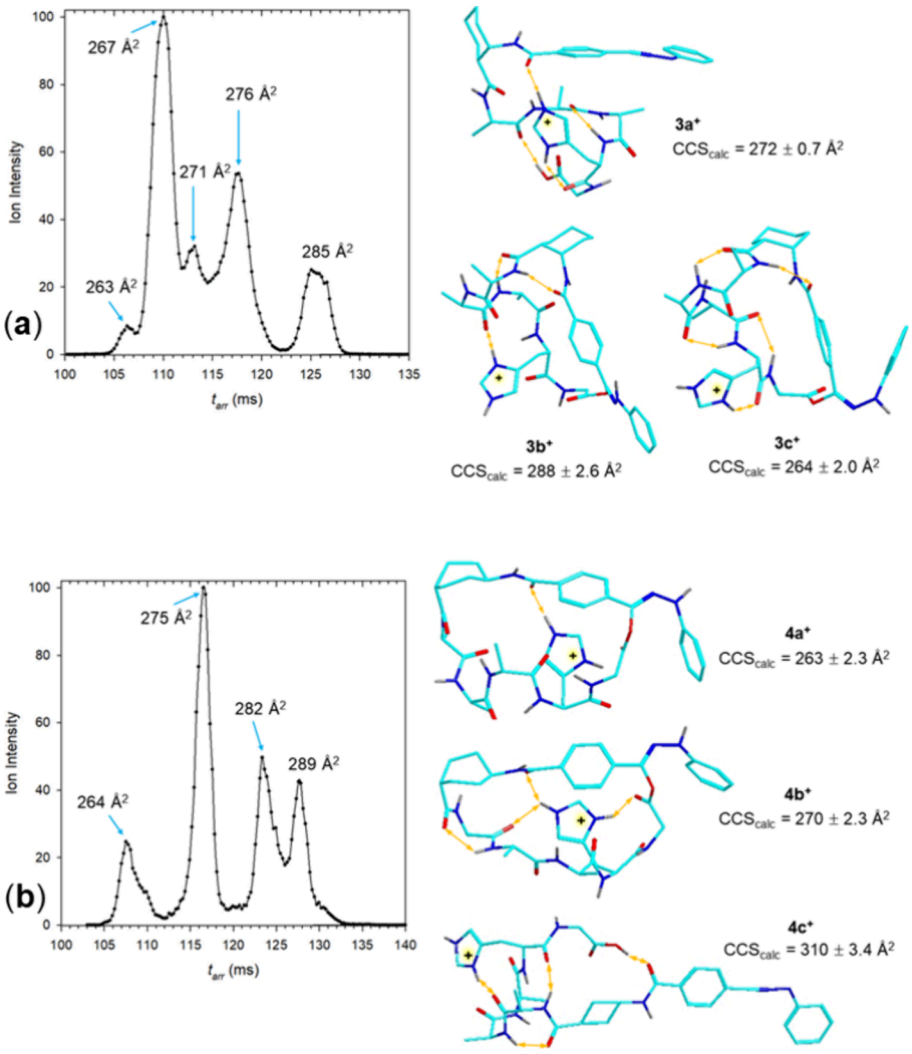


Figure 7. Arrival time distributions of (a) $(1,2\text{-}s\text{-AAAHG}-N_2)^+$ ions $\mathbf{3}^+$ and (b) $(1,4\text{-}s\text{-AAAHG}^+ - N_2)^+$ ions $\mathbf{4}^+$ after 4 passes. M06-2X/6-31+G(d,p) optimized structures and CCS_{calc} of low energy isomers $\mathbf{3a}^+$-$\mathbf{3c}^+$ and $\mathbf{4a}^+$-$\mathbf{4c}^+$. Structure description as in Figure 5.

into two groups; one with the carboxyl group reaching to the tetrazole ring, as represented by **1,4-A1**, and the other with a folded peptide sequence as represented by **1,4-A2**. The two groups differed in their CCS_{calc} which ranged between 262 and 270 Å^2 for the **1,4-A1** types and 282−286 Å^2 for the **1,4-A2** types. The gas-phase structures favored the **1,4-A2** group by Gibbs energy, which was largely mitigated by solvation in water that favored **1,4-A1** (Figure 5b). The structures of the **1,4-A1** type gave a closer agreement between the CCS_{IMS} and CCS_{calc} and thus we considered them representative of the 1,4-*s*-$AAAG^+$ ions produced by electrospray. Loss of N_2 from 1,2-*s*-$AAAG^+$ could not be studied by CID and therefore we did not have the relevant CCS_{IMS} of the nitrile imine and cross-linked products. Calculations suggested that the loss of N_2 from the tetrazole ring in 1,2-*s*-$AAAG^+$ was only 51 kJ mol^{-1} endothermic leaving a substantial excitation energy in nitrile imine ion $\mathbf{1a}^+$ (Scheme 4). From the energy balance one gets $E_{exc}(\mathbf{1}^+) \leq (E_{photon} - 51) = 479$ and 562 kJ mol^{-1} at 250 and 213 nm, respectively, where the $\leq$ sign reflects the unknown translational energy of the departing N_2 molecule. The energy excess can drive proton migration onto the nitrile imine group that has been considered essential for facilitating peptide cyclization.[22] Alternatively, the N−N bond in excited $\mathbf{1a}^+$ can dissociate, driving the observed loss of C_6H_5N and subsequent backbone dissociations.

1,2-*s*-AAAHG$^+$ and 1,4-*s*-AAAHG$^+$. Ion mobility of the 1,2-$AAAHG^+$ ions (m/z 799, 4 passes) showed two resolved peaks with CCS_{IMS} = 277 and 282 Å^2 for the minor (7%) and major (93%) components, respectively (Figure 6a). These were closely matched by the CCS_{calc} of the two lowest Gibbs energy structures **1,2-H-1** and **1,2-H-2** of 276 and 284 Å^2 (Figure 6a). Both these low-energy structures were protonated on His and showed coiled peptide chains with similar hydrogen bonding patterns, reflecting their similar CCS_{calc} that were within 3%. Ion mobility of the 1,4-$AAAHG^+$ ions (m/z 799, 15 passes) also showed two peaks of CCS_{IMS} 296 and 301 Å^2 for the major (98%) and minor (2%) components (Figure 6b). The calculated lowest-energy ion structures fell into two groups. One group, represented by **1,4-H-1** and **1,4-H-2**, showed the peptide hydrogen bonded to the benzoyl C=O which resulted in compact structures. The CCS_{calc} of **1,4-H-1** closely matched the CCS_{IMS} of the major peak in the Figure 6b mobilogram. The other group, represented by **1,4-H-3**, had more extended structures, resulting in larger CCS_{calc}.

Ion mobility separation of the CID-denitrogenated ions $\mathbf{3}^+$ and $\mathbf{4}^+$ (m/z 771) showed five fully or partially resolved peaks with CCS_{IMS} ranging between 263 and 289 Å^2 (Figure 7a,b).

https://doi.org/10.1021/jasms.4c00317
J. Am. Soc. Mass Spectrom. 2024, 35, 3070−3088

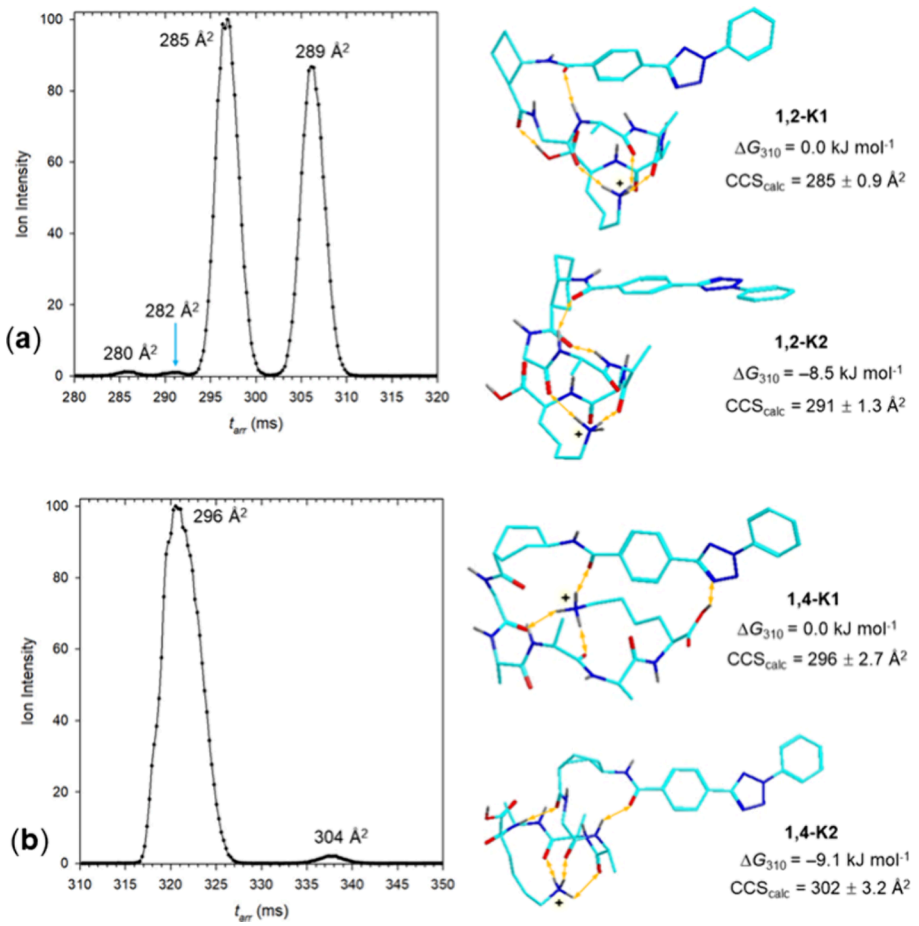


Figure 8. Arrival time distributions and optimized structures of (a) 1,2-*s*-GAAAK$^+$ (10 passes) and (b) 1,4-*s*-GAAAK$^+$ (14 passes). Structure description as in Figure 5.

This made the structure assignment difficult because the CCS_{calc} of conformers of several different ion structures showed substantial overlap. In addition, the phenyl group conformation in both the nitrile imines and cross-links had a large effect on the CCS_{calc}, causing overlap between conformers of different structure types. This is illustrated in Figure 7a for low energy conformers of the nitrile imine (**3a**$^+$) and carboxyl-cross-linked ions **3b**$^+$ and **3c**$^+$. In this case, a match between the CCS_{calc} and CCS_{IMS} cannot be taken as an unequivocal structure assignment, although the presence of the linear (nitrile imine) and carboxyl cross-linked structures was strongly indicated by the UVPD-MS^2 and CID-MS^3 spectra (Figure 2a,b). With **4**$^+$ (Figure 7b), the composite peak with CCS_{IMS} at 264 Å^2 was resolved after slicing and subsequent separation in 10 passes (Figure S20, Supporting Information). The major peak at CCS_{IMS} = 275 Å^2 was not split upon slicing and additional IMS cycles.

The CCS_{calc} for cross-linked structures **4a**$^+$ and **4b**$^+$ had close matches with peaks in the mobilogram whereas the CCS_{calc} for the nitrile imines (**4c**$^+$) had not. An interesting feature of the (1,4-*s*-AAAHG$^+$ − N_2)$^+$ cross-links was that the carboxyl bond to the nitrile imine was achieved only after flipping the cyclohexane ring to the twist-boat conformation. The **4c**$^+$ → **4b**$^+$ cyclization was calculated to be exergonic, $\Delta G_{g,\,310}$ = −59 kJ mol^{-1}, which compensated for the energy gain due to cyclohexane ring inversion. A further stabilization of the twist-boat conformation could be due to the strong hydrogen bond between the protonated histidine imidazole and the benzoyl carbamide in **4a**$^+$ and **4b**$^+$ (Figure 7b).

1,2-*s*-GAAAK$^+$ and 1,4-*s*-GAAAK$^+$. The presence of the Lys residue in these conjugates steered protonation to the basic side-chain ε-amino group in all low-energy isomers. Internal solvation of the ε-NH_3^+ group then governed the peptide chain folding, as shown for the representative 1,2-*s*-GAAAK$^+$ structures **1,2-K1** and **1,2-K2** (Figure 8a). Both **1,2-K1** and **1,2-K2** showed peptide hydrogen bonds to the benzamide carbonyl and chiefly differed in the hydrogen bonding of the C-terminal carboxyl and orientation of the phenyltetrazole moiety. These low-energy conformers had CCS_{calc} that closely matched those of the major peaks in the ion mobilogram. We did not find adequate CCS matches for the two minor peaks in the Figure 8a mobilogram that corresponded to more compact structures at CCS_{IMS} = 280 and 282 Å^2.

The low-energy conformers of 1,4-*s*-GAAAK$^+$ (**1,4-K1** and **1,4-K2**) had cyclohexane twist-boat conformations that allowed the ions to achieve efficient hydrogen bonding between the peptide moiety and benzamide carbonyl, in particular regarding internal solvation of the ε-NH_3^+ group by the benzamide in **1,4-K1** (Figure 8b). We note that the typical Gibbs energy difference between the cyclohexane twist-boat and chair conformations has been reported as 22−24 kJ mol^{-1},[58,59] which can be readily overcome by the binding energy of a strong intramolecular hydrogen bond. The **1,4-K1** ion of CCS_{calc} = 296 Å^2, which was more compact and had a

https://doi.org/10.1021/jasms.4c00317
J. Am. Soc. Mass Spectrom. 2024, 35, 3070−3088

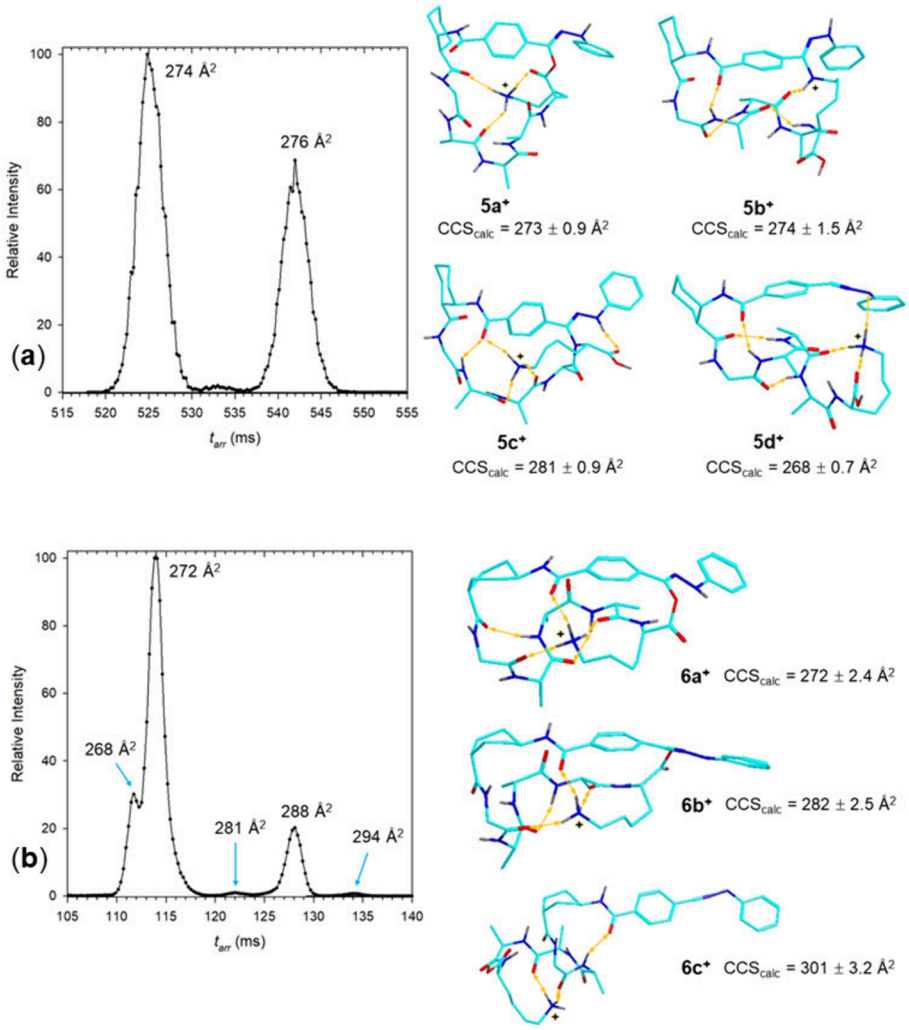


Figure 9. Arrival time distributions of (a) $(1,2\text{-}s\text{-GAAAK}-N_2)^+$ ions ($\mathbf{5}^+$) and (b) $(1,4\text{-}s\text{-GAAAK}^+ - N_2)^+$ ions ($\mathbf{6}^+$) after (a) 20 and (b) 4 passes. M06-2X/6-31+G(d,p) optimized structures and CCS_{calc} of low energy isomers $\mathbf{5a}^+$-$\mathbf{5d}^+$ and $\mathbf{6a}^+$-$\mathbf{6c}^+$. Structure description as in Figure 5.

lower Gibbs energy, also developed a hydrogen bond between the carboxyl and the tetrazole ring. Whether or not that may be conducive to carboxyl cross-linking in the incipient nitrile imine will be discussed later in the paper. The CCS_{calc} of **1,4-K1** and **1,4-K2** showed the best match with the CCS_{IMS} of the two resolved peaks, 296 Å^2 and 304 Å^2, in the Figure 8b ion mobilogram.

Ion mobility separation of the CID-denitrogenated ions $(1,2\text{-}s\text{-GAAAK}-N_2)^+$ ($\mathbf{5}^+$, Figure 9a) showed two major peaks of CCS_{IMS} = 274 and 276 Å^2, respectively, along with a very minor peak at CCS_{IMS} = 275 Å^2. We considered several types of ion structures, both nitrile imines and cross-linked ones, that were evaluated by their relative energy and CCS_{calc}. The representative lowest-energy conformers in each group are shown in Figure 9a. The carboxyl-cross-linked isomer $\mathbf{5a}^+$ of CCS_{calc} = 273 Å^2, that was the global energy minimum among the $(1,2\text{-}s\text{-GAAAK}-N_2)^+$ ions, showed a close match with the CCS_{IMS} of the first major peak in the mobilogram. A low-energy isomer cross-linked by the Lys NH_2 ($\mathbf{5b}^+$) of CCS_{calc} = 274 Å^2 was close to the second major peak in the $(1,2\text{-}s\text{-GAAAK}-N_2)^+$ mobilogram. Considering the CID-MS^3 spectrum of $\mathbf{5}^+$ (Figure 4b) that indicated carboxyl cross-linking, the presence of structures of the $\mathbf{5a}^+$ type in the stable ion population was possible and was consistent with the matching CCS. The presence of Lys-NH_2 cross-linked ions was also possible, as indicated by the matching CCS. Interestingly, no close CCS match was found for the nitrile-imine structures, such as $\mathbf{5d}^+$ of CCS_{calc} = 268 Å^2 (Figure 9a). The lack of residual nitrile-imine structures was consistent with the Figure 4b spectrum and the 93−94% cross-linking yield for $\mathbf{5}^+$ (Table 1).

Ion mobility of CID-denitrogenated ions $(1,4\text{-}s\text{-GAAAK}-N_2)^+$ ($\mathbf{6}^+$, Figure 9b) showed at least five fully and partially resolved components. The most abundant peak at CCS_{IMS} = 272 Å^2 found a close match with the carboxyl-cross-linked isomer $\mathbf{6a}^+$ of CCS_{calc} = 272 Å^2 (Figure 9b). Another carboxyl-cross-linked conformer ($\mathbf{6b}^+$ of CCS_{calc} = 282 Å^2) could fit one of the peaks with CCS_{IMS} = 281 or 288 Å^2, although this assignment was uncertain. The nitrile-imine isomers were more extended and had larger CCS_{calc}, e.g. for $\mathbf{6c}^+$, that could be accounted for by the weak peak at CCS_{IMS} = 294 Å^2 in the mobilogram. The assignment of structure $\mathbf{6a}^+$ was consistent with the CID-MS^3 spectrum of $\mathbf{6}^+$ (Figure 3b) and the high cross-linking yield for this scaffold (97%, Table 1). A salient feature of structures $\mathbf{6a}^+$−$\mathbf{6c}^+$ was that they all had the cyclohexane ring in a twist-boat conformation that allowed for cross-linking in $\mathbf{6a}^+$ and $\mathbf{6b}^+$ but also favored a peptide hydrogen bond to the benzamide carbonyl in $\mathbf{6c}^+$.

https://doi.org/10.1021/jasms.4c00317
J. Am. Soc. Mass Spectrom. 2024, 35, 3070−3088

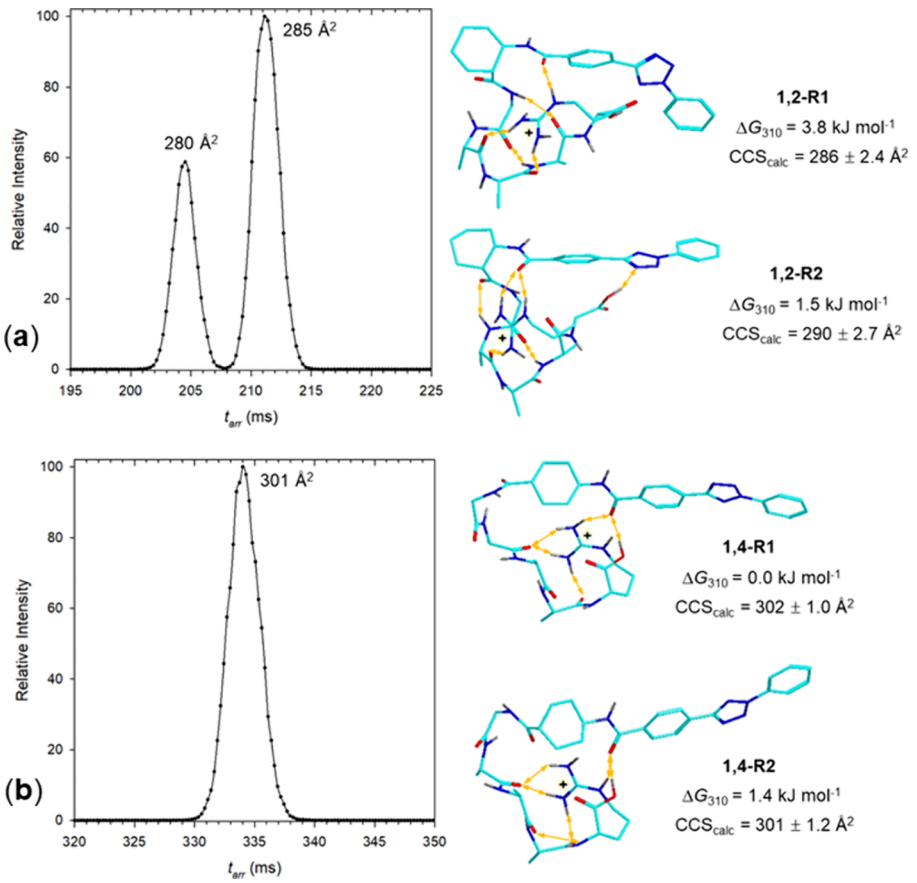


Figure 10. Arrival time distributions of (a) (1,2-*s*-GAAAR)$^+$ and (b) (1,4-*s*-GAAAR)$^+$ ions after (a) 7 and (b) 10 passes. M06-2X/6-31+G(d,p) optimized structures and CCS_{calc} of low energy isomers. Structure description as in Figure 5.

1,2-*s*-GAAAR$^+$ and 1,4-*s*-GAAAR$^+$. IMS of the 1,2-*s*-GAAAR$^+$ ion was represented by two baseline-resolved peaks in a 36:64 intensity ratio for which we determined CCS_{IMS} = 280 and 285 Å^2, respectively (Figure 10a). Several 1,2-*s*-GAAAR$^+$ ion structures were obtained by BOMD and DFT geometry optimizations. All low-energy structures converged to a common peptide folding pattern shown by **1,2-R1** and **1,2-R2**, in which the charged Arg side-chain group was wrapped by the peptide amides and H-bonded to the benzamide carbonyl (Figure 10a). A similar H-bonding pattern was found for other low-energy conformers **1,2-R3**-**1,2-R6** (Figure S21, Supporting Information). Interestingly, the cyclohexane ring in these low-energy conformers appeared in either the chair (**1,2-R1**, **1,2-R3**) or twist-boat (**1,2-R2**, **1,2-R4**-**1,2-R6**) conformation and the ring showed multiple reversible flipping in the course of BOMD trajectories at 610 K. On the basis of their Gibbs energies and CCS_{calc}, structures **1,2-R1** and **1,2-R2**, CCS_{calc} = 286 and 290 Å^2, respectively, gave the closest fit with the experimental data. The agreement was somewhat inferior to that for the other sequences, although still within 2%. The other low-energy conformers had CCS_{calc} that were similar or slightly larger than those of **1,2-R1** and **1,2-R2** (Figure S21, Supporting Information).

IMS of the 1,4-*s*-GAAAR$^+$ ion gave a single peak at CCS_{IMS} = 301 Å^2 even after 10 passes (Figure 10b). Geometry optimization pointed to the lowest-energy structures **1,4-R1** and **1,4-R2** with CCS_{calc} = 302, and 301 Å^2, respectively, providing near perfect match with the CCS_{IMS}. The nearly isoenergetic structures **1,4-R1** and **1,4-R2** were rotamers differing only in the orientation of the phenyltetrazole group whereas their peptide conformations were virtually identical (Figure 10b). These were characterized by a tightly folded peptide chain that was maintained by multiple hydrogen bonds and capped by carboxyl hydrogen bonding to the benzamide carbonyl. The cyclohexane ring was in a twist-chair conformation allowing the compact overall conformation for the peptide conjugates, as in the other low-energy ions with the *trans*-1,4-cyclohexane scaffold.

IMS of the CID-denitrogenated (1,2-*s*-GAAAR−N_2)$^+$ ions showed four resolved peaks after 15 passes (Figure 11a). The major component had CCS_{IMS} = 273 Å^2 that was accompanied by more compact minor components at CCS_{IMS} = 269, 270, and 271 Å^2. These CCS_{IMS} were substantially smaller than the CCS_{calc} of nitrile imine conformers that were in the 295−296 Å^2 range (**7c**$^+$, **7d**$^+$, Figure S22, Supporting Information). This result, indicating the absence of non-cross-linked isomers in the (1,2-*s*-GAAAR−N_2)$^+$ ion population, was qualitatively consistent with the high cross-linking yield for (1,2-*s*-GAAAR−N_2)$^+$ (Table 1). We obtained two types of low-energy cross-linked structures, a carboxyl cross-link **7a**$^+$, and an arginine cross-link **7b**$^+$ of CCS_{calc} = 275 and 277 Å^2, respectively, that were close to the CCS_{IMS} of the main peak in the mobilogram. The minor components with smaller CCS_{IMS} could belong to slightly more compact conformers of **7a**$^+$ and **7b**$^+$ that, however, did not find matching structures among the other low-energy ions (**7e**$^+$-**7g**$^+$, Figure S22) that showed larger CCS_{calc}.

https://doi.org/10.1021/jasms.4c00317
J. Am. Soc. Mass Spectrom. 2024, 35, 3070−3088

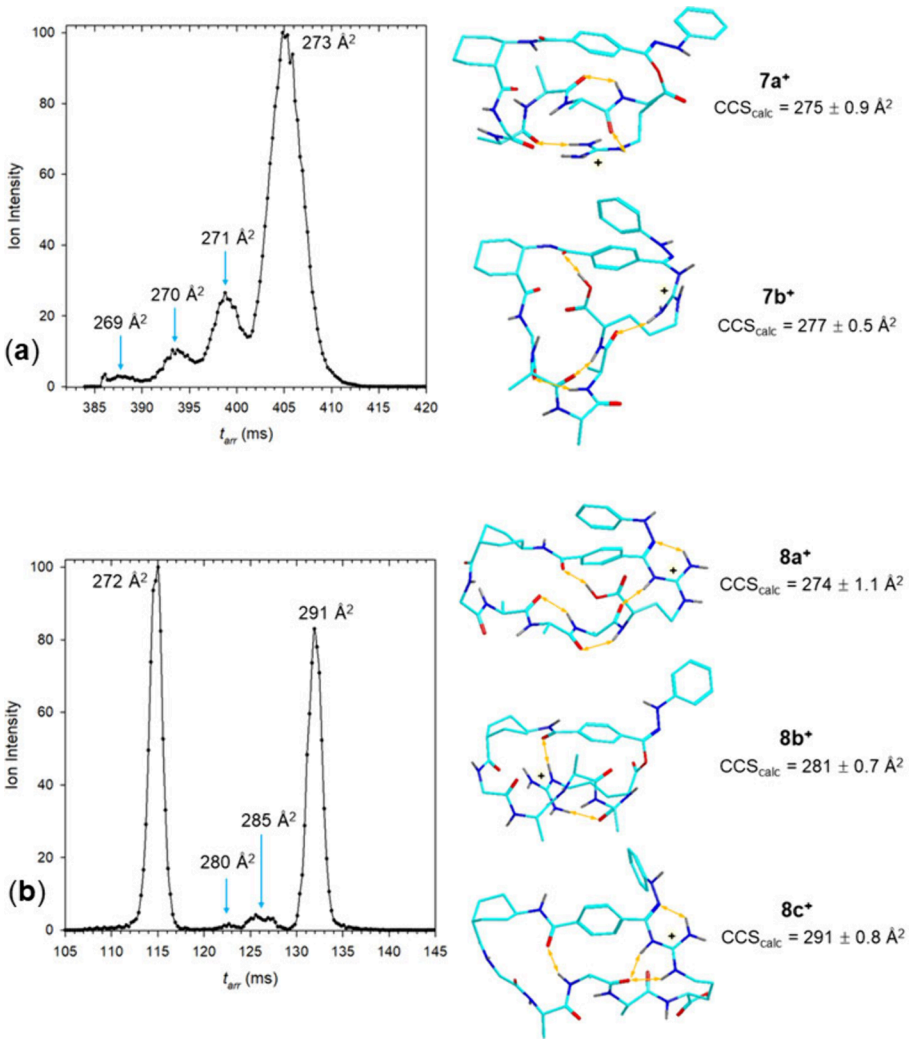


Figure 11. Arrival time distributions of (a) $(1,2\text{-}s\text{-GAAAR}-N_2)^+$ ions ($\mathbf{7}^+$) and (b) $(1,4\text{-}s\text{-GAAAR}^+ - N_2)^+$ ions ($\mathbf{8}^+$) after (a) 15 and (b) 4 passes. M06-2X/6-31+G(d,p) optimized structures and CCS$_{calc}$ of low energy isomers $\mathbf{7a}^+$, $\mathbf{7b}^+$ and $\mathbf{8a}^+$-$\mathbf{8c}^+$. Structure description as in Figure 5.

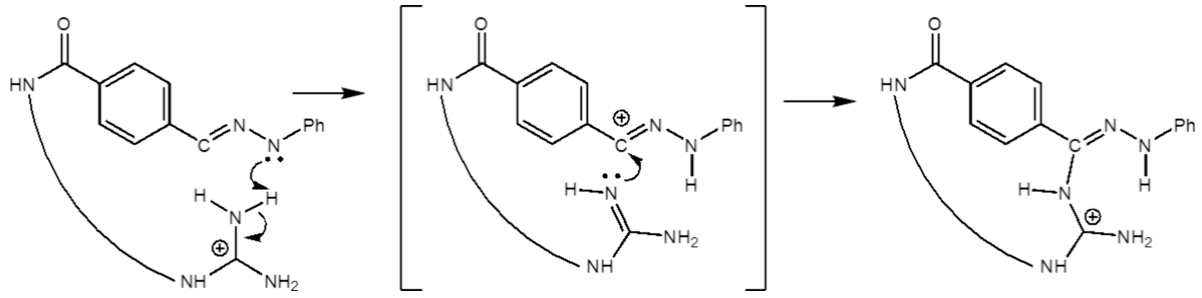
IMS of the CID-denitrogenated $(1,4\text{-}s\text{-GAAAR}-N_2)^+$ ions showed two major peaks at CCS$_{IMS}$ = 272 and 291 Å^2 that were accompanied by very minor peaks at CCS$_{IMS}$ = 280 and 285 Å^2 (Figure 11b). We found the best fit for the major IMS peaks with the Arg-cross-linked structures $\mathbf{8a}^+$ and $\mathbf{8c}^+$ of CCS$_{calc}$ = 274 and 291 Å^2, respectively. A carboxyl-cross-linked structure ($\mathbf{8b}^+$, CCS$_{calc}$ = 281 Å^2) could account for the minor peaks in the mobilogram. Ions $\mathbf{8a}^+$ and $\mathbf{8c}^+$ differed in their peptide conformations. In $\mathbf{8a}^+$, the compact structure of the macrocyclic ring was clinched by a hydrogen bond of the carboxyl to the benzamide carbonyl. This feature was absent in $\mathbf{8b}^+$ that had a bulkier macrocyclic ring with fewer hydrogen bonds and a free carboxyl group. In addition to $\mathbf{8a}^+$-$\mathbf{8c}^+$ we obtained several related ions that had CCS$_{calc}$ within the experimental range. These were represented by Arg-cross-linked structures $\mathbf{8d}^+$ and $\mathbf{8e}^+$ of CCS$_{calc}$ = 270 and 274 Å^2, respectively, carboxyl-cross-linked $\mathbf{8f}^+$ and $\mathbf{8g}^+$ with the respective CCS$_{calc}$ = 278 and 281 Å^2, and nitrile imine $\mathbf{8h}^+$ of CCS$_{calc}$ = 275 Å^2 (Figure S23, Supporting Information). All the cross-linked structures had the cyclohexane ring in a twist-boat conformation while differing in the conformation of the macrocyclic ring. Nitrile imine $\mathbf{8h}^+$ had a standard cyclohexane chair conformation (Figure S23). The close similarity of the CCS$_{calc}$ for the Arg-cross-linked ions $\mathbf{8a}^+$, $\mathbf{8d}^+$ and $\mathbf{8e}^+$ did not allow us to unambiguously assign the 272 Å^2 peak in the mobilogram to a single conformer. It appears, however, that the carboxyl-cross-linked ions $\mathbf{8b}^+$, $\mathbf{8f}^+$, and $\mathbf{8g}^+$ were only a small minority among the cross-links.

CONCLUSIONS

The experimental IMS data indicated that the *cis*-1,2-*s*-peptide scaffold ions were formed in the gas phase as mixtures of components that could be assigned to different but closely related conformers of the same protonation isomer. The cyclohexane ring in most of these ions had the standard chair conformation with the peptide chain assuming the equatorial position. The protonation site depended on the peptide composition, being at the benzamide carbonyl in $1,2\text{-}s\text{-AAAG}^+$, and at the basic residues in $1,2\text{-}s\text{-AAAHG}^+$, $1,2\text{-}s\text{-GAAAK}^+$, and $1,2\text{-}s\text{-GAAAR}^+$. The benzamide carbonyl also was a frequent acceptor of hydrogen bonds from the peptide charged groups. The *trans*-1,4-*s*-peptide scaffold ions displayed a tendency for one dominant conformation based on a cyclohexane twist-boat. This unusual phenomenon was clearly enforced by strong hydrogen bonding within the peptide and across the cyclohexane ring that compensated the energy needed for the chair-twist-boat conversion.[58,59] This was corroborated by the virtually equal Gibbs energies of the chair and twist-boat-

https://doi.org/10.1021/jasms.4c00317
J. Am. Soc. Mass Spectrom. 2024, 35, 3070−3088

Scheme 5. Proposed Mechanism for Arginine Crosslinking



like 1,4-*s*-GAAAR$^+$ conformers **1,4-R1** and **1,4-R2** (Figure 11b). Photodissociation and CID generated mixtures of isomers and conformers that were composed of cross-links and linear forms, presumably surviving nitrile imines. Photodissociation of 1,2-*s*-AAAG$^+$, 1,4-*s*-AAAG$^+$, 1,2-*s*-AAAHG$^+$, 1,4-*s*-AAAHG$^+$, 1,2-*s*-GAAAK$^+$, and 1,2-*s*-GAAAK$^+$ resulted in predominant cross-linking by the carboxyl group, forming diarylhydrazide enol esters as shown in Scheme 2 and 3. The role of the carboxyl was underscored by the sharp decrease of the cross-linking yields upon carboxyl methylation. We note that cross-linking in solution by a lysine-anchored nitrile imine to Glu side-chain carboxyl groups in proteins has been reported.[60,61] As discussed previously and corroborated by the absence of cross-linking in sodiated peptide conjugates,[22] the attack at the nitrile imine group was facilitated by proton transfer to the negative imine nitrogen. The tetrazole nitrogens were not protonated in the precursor ions, and the protonation sites were remote from the incipient nitrile imine. Hence, proton transfer must have involved conformational changes in the peptide moiety of the nitrile imine that were driven by the excitation energy provided by photon absorption or collisions. This was consistent with the modest threshold energies for the N_2 expulsion, e.g., 51 kJ mol^{-1} in the Scheme 4 reaction. The necessity for proton transfer and the associated conformational change implied that the cross-linking to nitrile imines in the peptide scaffolds probed the dynamics of the peptide conformations rather than the static structure of the precursor ion. This aspect was particularly salient in cross-linking of the *trans*-1,4-cyclohexane scaffolds that must have undergone a chair to twist-boat conversion. Cross-linking by Arg in 1,2-*s*-GAAAR$^+$ and 1,4-*s*-GAAAR$^+$ further emphasized the proton transfer steps. Carboxyl cross-linking was minor in these scaffolds, indicating that the carboxyl proton was not involved in the reaction. Cross-linking by the guanidinium group can be viewed as a two step process whereby an endothermic proton transfer from the guanidinium group to the imine nitrogen was followed by a highly exothermic formation of the C−N bond. From our calculations at the M06-2X/def2qzvpp level of theory and including zero-point corrections, the overall reaction was associated with the Gibbs energy change of $\Delta G_{310,rxn}$ = −88 and −72 kJ mol^{-1} for the cyclization in the (1,2-*s*-GAAAR−N_2)$^+$ and (1,4-*s*-GAAAR−N_2)$^+$ intermediates, respectively (Scheme 5). The cyclization can be classically viewed as a [3 + 2] addition of the polar guanidinium N−H bond across the nitrile imine dipolarophile that may proceed in two steps involving a proton-transfer intermediate, as shown in Scheme 4. It should be noted that analogous reactions have not been reported in the condensed phase and represent a new feature of gas-phase peptide ion chemistry.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/jasms.4c00317.

Synthetic procedures, detailed ion mobility parameters, supplementary tables of high-resolution data, tandem mass spectra, auxiliary optimized ion structures and CCS_{calc} (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Karel Lemr − *Department of Analytical Chemistry, Faculty of Science, Palacký University, Olomouc 77900, Czech Republic; Institute of Microbiology of the Czech Academy of Sciences, Prague 14220, Czech Republic;* orcid.org/0000-0003-3158-0637; Phone: +420-585634415; Email: karel.lemr@upol.cz

František Tureček − *Department of Chemistry, Bagley Hall, University of Washington, Seattle, Washington 98195-1700, United States;* orcid.org/0000-0001-7321-7858; Phone: +1-206-685-2041l; Email: turecek@uw.edu

Authors

Hongyi Zhu − *Department of Chemistry, Bagley Hall, University of Washington, Seattle, Washington 98195-1700, United States*

Marianna Nytka − *Department of Analytical Chemistry, Faculty of Science, Palacký University, Olomouc 77900, Czech Republic;* orcid.org/0000-0002-9242-8295

Tuan Ngoc Kim Vu − *Department of Chemistry, Bagley Hall, University of Washington, Seattle, Washington 98195-1700, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/jasms.4c00317

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

Research at the University of Washington was supported by the Chemistry Division of the U.S. National Science Foundation, Grants CHE-1951518 and CHE-2347921. F.T. acknowledges support by the Klaus and Mary Ann Saegebarth Endowment. Research at Palacký University was supported by projects IGA PrF 2023 027 and IGA PrF 2024 026. Thanks are due to Mr. Jiahao Wan for the UV−vis absorption spectrum of 2-(4-carboxyphenyl)-5-phenyl tetrazole.

REFERENCES

(1) Piersimoni, L.; Kastritis, P. L.; Arlt, C.; Sinz, A. Cross-Linking Mass Spectrometry for Investigating Protein Conformations and Protein−Protein Interactions−A Method for All Seasons. *Chem. Rev.* **2022**, *122*, 7500−7531.

(2) Knowles, J. R. Photogenerated Reagents for Biological Receptor-Site Labeling. *Acc. Chem. Res.* **1972**, *5*, 155−160.

(3) Smith, R. A. G.; Knowles, J. R. Aryldiazirines: Potential Reagents for Photolabeling of Biological Receptor Sites. *J. Am. Chem. Soc.* **1973**, *95*, 5072−5073.

(4) Das, J. Aliphatic Diazirines as Photoaffinity Probes for Proteins: Recent Developments. *Chem. Rev.* **2011**, *111*, 4405−4417.

(5) Kauer, J. C.; Erickson-Viitanen, S.; Wolfe, H. R., Jr.; DeGrado, W. W. P-benzoyl-L-Phenylalanine, a New Photoreactive Amino Acid. Photolabeling of Calmodulin with a Synthetic Calmodulin-Binding Peptide. *J. Biol. Chem.* **1986**, *261*, 10695−10700.

(6) Dorman, G.; Prestwich, G. D. Benzophenone Photophores in Biochemistry. *Biochemistry* **1994**, *33*, 5661−5673.

(7) Lee, H. S.; Dimla, R. D.; Schultz, P. G. Protein-DNA Photo-Crosslinking with a Genetically Encoded Benzophenone-Containing Amino Acid. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 5222−5224.

(8) Lomant, A. J.; Fairbanks, G. Chemical Probes of Extended Biological Structures: Synthesis and Properties of the Cleavable Protein Cross-Linking Reagent [35S]Dithiobis(Succinimidyl Propionate). *J. Mol. Biol.* **1976**, *104*, 243−261.

(9) Merkley, E. D.; Rysavy, S.; Kahraman, A.; Hafen, R. P.; Daggett, V.; Adkins, J. N. Distance Restraints from Crosslinking Mass Spectrometry: Mining a Molecular Dynamics Simulation Database to Evaluate Lysine-Lysine Distances. *Protein Sci.* **2014**, *23*, 747−759.

(10) Tang, X.; Munske, G. R.; Siems, W. F.; Bruce, J. E. Mass Spectrometry Identifiable Cross-Linking Strategy for Studying Protein-Protein Interactions. *Anal. Chem.* **2005**, *77*, 311−318.

(11) Soderblom, E. J.; Goshe, M. B. Collision-Induced Dissociative Chemical Cross-Linking Reagents and Methodology: Applications to Protein Structural Characterization Using Tandem Mass Spectrometry Analysis. *Anal. Chem.* **2006**, *78*, 8059−8068.

(12) Kao, A.; Chiu, C.; Vellucci, D.; Yang, Y.; Patel, V. R.; Guan, S.; Randall, A.; Baldi, P.; Rychnovsky, S. D.; Huang, L. Development of A Novel Cross-Linking Strategy for Fast and Accurate Identification of Cross-Linked Peptides of Protein Complexes. *Mol. Cell. Proteomics* **2011**, *10*, No. M110.002170.

(13) Petrotchenko, E. V.; Serpa, J. J.; Borchers, C. H. An Isotopically Coded CID-Cleavable Biotinylated Cross-Linker for Structural Proteomics. *Mol. Cell. Proteomics* **2011**, *10*, No. M110.001420.

(14) Clifford-Nunn, B.; Showalter, H. D.; Andrews, P. C. Quaternary Diamines as Mass Spectrometry Cleavable Crosslinkers for Protein Interactions. *J. Am. Soc. Mass Spectrom.* **2012**, *23*, 201−212.

(15) Stadlmeier, M.; Runtsch, L. S.; Streshnev, F.; Wuhr, M.; Carell, T. A Click-Chemistry-Based Enrichable Crosslinker for Structural and Protein Interaction Analysis by Mass Spectrometry. *ChemBioChem.* **2020**, *21*, 103−107.

(16) Nury, C.; Redeker, V.; Dautrey, S.; Romieu, A.; van der Rest, G.; Renard, P. Y.; Melki, R.; Chamot-Rooke, J. A Novel Bio-Orthogonal Cross-Linker for Improved Protein/Protein Interaction Analysis. *Anal. Chem.* **2015**, *87*, 1853−1860.

(17) Tureček, F. Covalent Crosslinking in Gas-Phase Biomolecular Ions. An Account and Perspective. *Phys. Chem. Chem. Phys.* **2023**, *25*, 32292−32304.

(18) Liu, Y.; Ramey, Z.; Tureček, F. Non-Covalent Interactions of a Neuroprotective Peptide Revealed by Photodissociative Cross-Linking in the Gas Phase. *Chem.—Eur. J.* **2018**, *24*, 9259−9263.

(19) Liu, Y.; Tureček, F. Photodissociative Crosslinking of Diazirine-Tagged Peptides with DNA Dinucleotides in the Gas Phase. *J. Am. Soc. Mass Spectrom.* **2019**, *30*, 1992−2006.

(20) Shaffer, C. J.; Andrikopoulos, P. C.; Řezáč, J.; Rulíšek, L.; Tureček, F. Efficient Covalent Bond Formation in Gas-Phase Peptide-Peptide Ion Complexes with the Photoleucine Stapler. *J. Am. Soc. Mass Spectrom.* **2016**, *27*, 633−645.

(21) Nguyen, H. T. H.; Andrikopoulos, P. C.; Rulíšek, L.; Shaffer, C. J.; Tureček, F. Photodissociative Cross Linking of Noncovalent Peptide-Peptide Ion Complexes in the Gas Phase. *J. Am. Soc. Mass Spectrom.* **2018**, *29*, 1706−1720.

(22) Wan, J.; Nytka, M.; Vu, K.; Qian, H.; Lemr, K.; Turecek, F. Nitrile Imines as Peptide and Oligonucleotide Photocrosslinkers in Gas-Phase Ions. *J. Am. Soc. Mass Spectrom.* **2024**, *35*, 344−356.

(23) Song, W.; Wang, Y.; Qu, J.; Madden, M. M.; Lin, Q. A Photoinducible 1,3-Dipolar Cycloaddition Reaction for Rapid, Selective Modification of Tetrazole-Containing Proteins. *Angew. Chem., Int. Ed. Engl.* **2008**, *47*, 2832−2835.

(24) Marshall, D. L.; Menzel, J. P.; McKinnon, B. I.; Blinco, J. P.; Trevitt, A. J.; Barner-Kowollik, C.; Blanksby, S. J. Laser Photodissociation Action Spectroscopy for the Wavelength-Dependent Evaluation of Photoligation Reactions. *Anal. Chem.* **2021**, *93*, 8091−8098.

(25) Zhu, H.; Zima, V.; Ding, E.; Tureček, F. Carbene Crosslinking in Gas-Phase Peptide Ion Scaffolds. *J. Am. Soc. Mass Spectrom.* **2023**, *34*, 763−774.

(26) Marek, A.; Nguyen, H. T. H.; Brož, B.; Tureček, F. Stereospecific Control of Peptide Gas-Phase Ion Chemistry with *cis* and *trans* Cyclo Ornithine Residues. *J. Mass Spectrom.* **2018**, *53*, 124−137.

(27) Sharp, J. T. Nitrile Ylides and Nitrile Imines. In *Chemistry of Heterocyclic Compounds 59: Synthetic Applications of 1,3-Dipolar Cycloaddition Chemistry Toward Heterocycles and Natural Products*; Padwa, A., Pearson, W. H., Eds.; John Wiley & Sons: New York, 2002.

(28) Dang, A.; Korn, J. A.; Gladden, J.; Mozzone, B.; Tureček, F. UV-Vis Photodissociation Action Spectroscopy on Thermo LTQ-XL ETD and Bruker amaZon Ion Trap Mass Spectrometers: A Practical Guide. *J. Am. Soc. Mass Spectrom.* **2019**, *30*, 1558−1564.

(29) Giles, K.; Ujma, J.; Wildgoose, J.; Pringle, S.; Richardson, K.; Langridge, D.; Green, M. A. Cyclic Ion Mobility-Mass Spectrometry System. *Anal. Chem.* **2019**, *91*, 8564−8573.

(30) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984**, *81*, 3684−3690.

(31) Řezáč, J.; Fanfrlík, J.; Salahub, D.; Hobza, P. Semiempirical Quantum Chemical PM6Method Augmented by Dispersion and H Bonding Correction Terms Reliably Describes Various Types of Noncovalent Complexes. *J. Chem. Theory Comput.* **2009**, *5*, 1749−1760.

(32) Řezáč, J. Cuby: An Integrative Framework for Computational Chemistry. *J. Comput. Chem.* **2016**, *37*, 1230−1237.

(33) Stewart, J. J. P. *MOPAC 16*; Stewart Computational Chemistry: Colorado Springs, CO, 2016.

(34) Becke, A. D. Density-Functional Exchange-Energy Approximation with Correct Asymptotic Behavior. *Phys. Rev. A* **1988**, *38*, 3098−3100.

(35) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the Damping Function in Dispersion Corrected Density Functional Theory. *J. Comput. Chem.* **2011**, *32*, 1456−1465.

(36) Nickerson, C. J.; Bryenton, K. R.; Price, A. J. A.; Johnson, E. R. Comparison of Density-Functional Theory Dispersion Corrections for the DES15K Database. *J. Phys. Chem. A* **2023**, *127*, 8712−8722.

(37) Zhao, Y.; Truhlar, D. G. The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: Two New Functionals and Systematic Testing of Four M06-Class Functionals and 12 Other Functionals. *Theor. Chem. Acc.* **2008**, *120*, 215−241.

(38) Weigend, F. Accurate Coulomb-Fitting Basis Sets for H to Rn. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1057−1065.

(39) Gray, M.; Bowling, P. E.; Herbert, J. M. Comment on "Benchmarking Basis Sets for Density Functional Theory Thermochemistry Calculations: Why Unpolarized Basis Sets and the Polarized 6-311G Family Should Be Avoided. *J. Phys. Chem. A* **2024**, *128*, 7739−7745.

https://doi.org/10.1021/jasms.4c00317
J. Am. Soc. Mass Spectrom. 2024, 35, 3070−3088

(40) Singh, U. C.; Kollman, P. A. An Approach to Computing Electrostatic Charges for Molecules. *J. Comput. Chem.* **1984**, *5*, 129−145.
(41) Besler, B. H.; Merz, K. M., Jr.; Kollman, P. A. Atomic Charges Derived from Semiempirical Methods. *J. Comput. Chem.* **1990**, *11*, 431−439.
(42) Ieritano, C.; Crouse, J.; Campbell, J. L.; Hopkins, W. S. A Parallelized Molecular Collision Cross Section Package with Optimized Accuracy and Efficiency. *Analyst* **2019**, *144*, 1660−1670.
(43) Ieritano, C.; Hopkins, W. S. Assessing Collision Cross Section Calculations Usi ng MobCal-MPI with a Variety of Commonly Used Computational Methods. *Mater. Today Commun.* **2021**, *27*, No. 102226.
(44) Halgren, T. A. Merck olecular Force Field. I. Basis, Form, Scope, Parametrization, and Performance of MMFF94. *J. Comput. Chem.* **1996**, *17*, 490−519.
(45) Chu, I. K.; Siu, C.-K.; Lau, J. K.-C.; Tang, W. K.; Mu, X.; Lai, C. K.; Guo, X.; Wang, X.; Li, N.; Yao, Z.; 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.* **2015**, *390*, 24−27.
(46) Bleiholder, C.; Osburn, S.; Williams, T. D.; Suhai, S.; Van Stipdonk, M.; Harrison, A. G.; Paizs, B. Sequence-Scrambling Fragmentation Pathways of Protonated Peptides. *J. Am. Chem. Soc.* **2008**, *130*, 17774−17789.
(47) Novák, J.; Lemr, K.; Schug, K. A.; Havlíček, V. CycloBranch: De Novo Sequencing of Nonribosomal Peptides from Accurate Product Ion Mass Spectra. *J. Am. Soc. Mass Spectrom.* **2015**, *26*, 1780−1786.
(48) Townsend, C.; Furukawa, A.; Schwochert, J.; Pye, C. R.; Edmondson, Q.; Lokey, R. S. CyCLS: Accurate, Whole-Library Sequencing of Cyclic Peptides Using Tandem Mass Spectrometry. *Bioorg. Med. Chem.* **2018**, *26*, 1232−1238.
(49) 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 Process.* **1991**, *111*, 173−189.
(50) Thorne, G. C.; Ballard, K. D.; Gaskell, S. J. Metastable Decomposition of Peptide [M + H]+ Ions via Rearrangement Involving Loss of C-Terminal Amino Acid Residue. *J. Am. Soc. Mass Spectrom.* **1990**, *1*, 249−257.
(51) Ballard, K. D.; Gaskell, S. J. Intramolecular Oxygen-18 Isotopic Exchange in the Gas Phase Observed during the Tandem Mass Spectrometric Analysis of Peptides. *J. Am. Chem. Soc.* **1992**, *114*, 64−71.
(52) Goeringer, D. E.; McLuckey, S. A. Evolution of Ion Internal Energy during Collisional Excitation in the Paul Ion Trap: A Stochastic Approach. *J. Chem. Phys.* **1996**, *104*, 2214−2221.
(53) Green, M. M. Mass Spectrometrry − a Sensitive Probe of Molecular Structure. *Pure Appl. Chem.* **1978**, *50*, 185−196.
(54) Mandelbaum, A. Stereochemical Effects in Mass Spectrometry. *Mass Spectrom. Rev.* **1983**, *2*, 223−284.
(55) Turecek, F. Stereochemistry of Organic Ions in the Gas Phase: A Review. *Collect. Czech. Chem. Commun.* **1987**, *52*, 1928−1984.
(56) Splitter, J. S.; Turecek, F. *Applications of Mass Spectrometry to Organic Stereochemstry*; VCH Publishers: New York, 1994; pp 123−128.
(57) Pepin, R.; Petrone, A.; Laszlo, K. J.; Bush, M. F.; Li, X.; Tureček, F. Does Thermal Breathing Affect Collision Cross Sections of Gas Phase Peptide Ions? An Ab initio Molecular Dynamics Study. *J. Phys. Chem. Lett.* **2016**, *7*, 2765−2771.
(58) Squillacote, M.; Sheridan, R. S.; Chapman, O. L.; Anet, F. A. L. Spectroscopic Detection of the Twist-Boat Conformation of Cyclohexane. Direct Measurement of the Free Energy Difference between the Chair and the Twist-Boat. *J. Am. Chem. Soc.* **1975**, *97*, 3244−3246.
(59) Johnson, W. S.; Bauer, V. J.; Margrave, J. L.; Frisch, M. A.; Dreger, L. H.; Hubbard, W. N. The Energy Difference between the Chair and Boat Forms of Cyclohexane. The Twist Conformation of Cyclohexane. *J. Am. Chem. Soc.* **1961**, *83*, 606−614.
(60) Herner, A.; Marjanovic, J.; Lewandowski, T. M.; Marin, V.; Patterson, M.; Miesbauer, L.; Ready, D.; Williams, J.; Vasudevan, A.; Lin, Q. 2-Aryl-5-Carboxytetrazole as a New Photoaffinity Label for Drug Target Identification. *J. Am. Chem. Soc.* **2016**, *138*, 14609−14615.
(61) Tian, Y.; Jacinto, M. P.; Zeng, Y.; Yu, Z.; Qu, J.; Liu, W. R.; Lin, Q. Genetically Encoded 2-Aryl-5-Carboxytetrazoles for Site-Selective Protein Photo-Cross-Linking. *J. Am. Chem. Soc.* **2017**, *139*, 6078−6081.

https://doi.org/10.1021/jasms.4c00317
J. Am. Soc. Mass Spectrom. 2024, 35, 3070−3088