Dario Barreiro-Lage, Paola Bolognesi et al.
Photofragmentation specificity of photoionized cyclic amino acids (diketopiperazines) as precursors of peptide building blocks
Photofragmentation specificity of photoionized cyclic amino acids (diketopiperazines) as precursors of peptide building blocks†

Dario Barreiro-Lage,†a,c Jacopo Chiarinelli,†b Paola Bolognesi,†b,⁎d Robert Richter,†d Henning Zettergren,†d,⁎ Mark H. Stockett,†d Sergio Díaz-Tendero†a,e and Lorenzo Avaldi†b

The photoionisation and photofragmentation of the two cyclic dipeptides cyclo(alanyl-glycine) cGA and cyclo(glycyl-glycine) cGG, have been studied combining experiments and simulations. State selected fragments from the ionized molecules are detected using photo-electron photo-ion coincidence (PEPICO) measurements and specific fragmentation paths are identified and characterized via the use of ion-neutral coincidence maps. The simulations, performed using Quantum Chemistry methods, allow us to infer the fragmentation mechanisms of the ionized and excited molecules. We show that ring opening is followed by emission of the neutral fragments CO and HNCO. In the case of cGG the emission of neutral CO leads to a metastable structure that breaks producing small cationic fragments. The studied cyclic dipeptides evolve under ionizing radiation generating different small aziridin moieties and oxazolidinones. These two species are key reactants to elongate producing peptide chains. The corresponding mechanisms have been computed and show that the reaction requires very low energy and may occur in the presence of ionizing radiation.

1. Introduction

The response and evolution of molecules of biological interest interacting with ionising radiation is fundamental not only for biology, where the damage by ionizing radiation strongly depends on the structural and chemical properties of the biomolecular constituents of life matter and can be related to processes initiated at the atomic and molecular level, but also in biotechnological applications, such as biosensors and molecular electronics. Moreover in astrochemistry and astrobiology, key information on the origin of life is provided by the understanding of the chemistry of relatively simple molecules in environments resembling primordial conditions, either on Earth or extraterrestrial objects where external agents may have triggered amino acid condensation, for example, in aqueous or fluctuating environments.

In the particular case of amino acids interacting with highly-charged and energetic ions in the gas phase3–5 the ionized and excited molecules present competitive decay channels: in addition to the so-called Coulomb explosion, leading to direct bond cleavages(s) and emission of two or several charged fragments, ultrafast isomerization has also been observed, with hydrogen migration being the most common alternative pathway. Interestingly, the interaction of clusters of amino acids with z particles in the gas phase leads to the formation of peptide bonds, thus being a possible source of prebiotic molecules.6 Gas phase peptide bond formation has also been observed in CID-MS/MS experiments of electrosprayed clusters of serine-serine dipeptides7 and the following photoexcitation of proton bound peptide complexes.8

A dipeptide can exist in two different structures, a linear or a cyclic one, depending on the formation of one or two peptide bonds between the carboxylic and amino groups of the two amino acids and the elimination of one or two water molecules.
The common feature of any cyclo dipeptide built on α amino acids is the 2,5-diketopiperazine (DKP) six member ring, which results in a peculiar behavior in the photoionization and photofragmentation of the cyclo dipeptide.\(^9,10\) At variance with the fragmentation of amino acids, DKP derivatives have attracted less attention. Zhang et al.\(^{11}\) reported the VUV photoinduced fragmentation of cyclo(alanyl-alanine), cyclo(glycyl-glycine) and cyclo(glycyl-valine) up to 11 eV, while the fragmentation of a series of both protonated and deprotonated cyclo-dipeptides has been studied by Guo et al.\(^{12,13}\) We have recently shown\(^9\) that the VUV irradiation of cyclo(alanyl-alanine) (cAA), besides being an excellent source of CO and HNCO prebiotic species can also produce oxazolidinone intermediates. In the same study, we proposed a model where the interaction among the oxazolidinone intermediate and other neutral and charged fragments released in the molecular decomposition of cAA may lead either to the reconstruction of the cyclic cationic dipeptide or to the formation of longer linear peptide chains via almost barrierless processes, thus proposing a new mechanism for the formation of more complex biomolecules as oligo-peptides.

In this context, a question arises whether the different amino acids in cyclo dipeptides, i.e. different side chains attached to the DKP skeleton, may affect the behaviour of the irradiated molecules and their potential as prebiotic building blocks. To advance the understanding of this key open question, in this work we have studied the photoionisation and photofragmentation of the two cyclic dipeptides cyclo(alanyl-glycine) cGA and cyclo(glycyl-glycine) cGG, combining experiments and simulations and compared the results with those previously obtained for cyclo(alanyl-alanine) cAA.\(^9,10\) From the experimental side photoelectron spectroscopy and mass spectrometry combined with the photoelectron–photoion coincidence (PEPICO) and ion-neutral experiments have been used to investigate the state-selective fragmentation of the dipeptides as a function of the molecular internal energy. Since their introduction in 1967\(^{14}\) PEPICO experiments have been proved to be the most suited approach to study dissociation dynamics providing branching ratios towards the different channels, absolute decay rates and energy partitioning among the products.\(^{15,16}\) Ion-neutral coincidence experiments,\(^{10}\) on the other hand, allow us to identify the correlated pair of charged and neutral fragments released in the two-body fragmentation of a cation and from the map of the ion-neutral coincidence events sequential fragmentation paths can be proposed. Then \emph{ab initio} molecular dynamics simulations and potential energy surface exploration on the one hand allow us to interpret the PEPICO and ion-neutral coincidence results inferring the fragmentation mechanisms in the ionized and excited molecules and on the other predict the reactivity of the newly formed species. This methodology, recently detailed in ref. 10 for the case of cAA, has allowed us to investigate the potential capability of DKP species to somehow ‘survive’ VUV irradiation, either reconstituting the cationic dipeptide structure or even giving rise to the formation of more complex species.

The experimental methods and details of the simulations are presented in Section 2, while the results are presented and discussed in Section 3. Finally Section 4 is devoted to some conclusions.

## 2. Methods

### 2.1 Experimental

The photoemission and electron–ion coincidence measurements were performed at the CIPO beamline\(^37\) of the Elettra synchrotron radiation facility, using a monochromatic photon beam of 60 eV. The end-station is equipped with a 150 mm radius hemispherical electron energy analyser (VG 220i) and a time-of-flight (TOF) ion mass spectrometer mounted opposite to each other at the magic angle, i.e. approximately 55° with respect to the linearly polarized photon beam, in order to cancel out angular effects in the photoionisation cross section of the different molecular orbitals. All details of the experimental apparatus and procedures have been presented and discussed in ref. 10 and references therein, and therefore will be only briefly summarised here. The hemispherical analyser and TOF spectrometers have been operated independently for photoelectron (PE) spectroscopy or photoion mass spectrometry (MS) measurements, or ‘in conjunction’ for photoelectron–photoion coincidence (PEPICO) experiments. The 2D position sensitive detector (PSD) in the electron analyzer\(^{17,18}\) provides an acquisition window and an energy resolution of about 10% and 2% of the pass energy, respectively. The PE spectra reported in Fig. 1 have been measured with a pass energy of 10 eV, an overall energy resolution of about 200 meV and an acquisition time of about 10 s per point.

For MS measurements, the repeller and extractor electrodes of the TOF spectrometer are polarized with antisymmetric voltages up to 370 V cm\(^{-1}\) and operated in pulsed mode, triggered by a 1 kHz pulse generator. In the PEPICO mode the extraction is triggered by either the signal due to the detection of a photoelectron or by the one from a pulse generator running at 100 Hz, a frequency comparable with the photoelectron counting rate during PEPICO acquisitions. The latter is used for measurement of the so-called ‘random mass spectrum’ generated by uncorrelated particles. The random mass spectrum has to be evaluated and subtracted from the measured electron–ion coincidence spectrum after appropriate normalization to their respective number of triggers. From these background subtracted PEPICO spectra, the next step in the data analysis consists of integrating the yield of each fragment over the region of mass-to-charge ratio \(m/z\) ± 0.5.

Finally, these yields are reported versus the binding energy, \(\text{BE} = hv - \text{KE}\), see Fig. 2, where \(hv\) is the photon energy and KE is the electron kinetic energy. The PEPICO measurements have been performed at a fixed photon energy of 60 eV, with an analyser pass energy of 30 eV and by scanning the KE over the range of interest in steps of 2.7 eV to provide sufficient overlap between two adjacent KE settings in the electron analyser. The entire KE scan is then repeated several times to improve the statistics. The stability of the apparatus is checked by PE measurements performed before and after each PEPICO acquisition.

The ion-neutral coincidence experiments have been performed at the CNR-ISFM in Rome, using a Wiley-McLaren TOF mass spectrometer. 

---

N atom at a flight time longer than the free flight at a constant speed and may or may not be detected emission of neutral fragments in cGA and cGG, see Fig. 3. daughter ion, D acceleration regions of the TOF spectrometer, then the lighter ion’. Briefly, when a non-stationary singly charged ion, the ‘parent ion’, is suddenly accelerated and will reach the TOF detector at a shorter time than the Pion nominal flight time. The correlated neutral fragment, Nfrag, will continue its free flight at a constant speed and may or may not be detected at a flight time longer than the Pion nominal flight time.

The distribution of the arrival times of these (Pion, Nfrag) pairs produces the ‘traces’ in Fig. 3 and allows us to determine the corresponding Pion \rightarrow D_{ion} + N_{frag} reaction paths. In our set-up, considering the geometry of the TOF and the applied voltages, the residence time in the extraction/acceleration regions is in the range of 790–850 ns for cGG+ and 840–900 ns for cGA+. So, for fragmentation events happening within these time limits it has been possible to correlate the fragmentation of a parent ion to its ion-neutral products. Each trace originates at the TOF1 of Pion, corresponding to ‘late’ fragmentation events, and asymptotically converges to the TOF1 of D_{ion}, corresponding to ‘early’ fragmentation events. ‘Early’ and ‘late’ times here are in relation to the time spent in the extraction/acceleration regions of the spectrometer, and refer respectively to prompt fragmentation, i.e. ‘early’, and fragmentation happening just before entering the free fly tube, i.e. ‘late’. Considering that ‘early’ fragmentation events have a lower detection efficiency for N_{frag} due to the still low KE and directionality of motion acquired by Pion at the time of fragmentation, each trace naturally fades out towards its convergence limit at TOF1 of D_{ion}. Therefore, only an ion optics simulation of the process allows one to reliably extract the reaction path, Pion \rightarrow D_{ion} + N_{frag}. A fit of the intensity distribution of the coincidence events along the traces might in principle provide relevant information on the kinetics of the fragmentation. However, the unknown detection efficiency for neutral species in the set-up and the unknown ion’s internal energy, due to the detection of electrons with unselected kinetic energy prevent the use of these data to extract thermochemical and kinetics information on the photofragmentation. It is also important to clarify that the different traces obtained in ion-neutral coincidence spectra correspond to individual events, and thus do not assure sequentiality. In this work, simulations have been performed using a custom Igor Pro procedure based on calculation of the electric field in the apparatus obtained by the SIMION 8.0 program.

The background pressure in the vacuum chamber was about 2 \times 10^{-8} mbar. Each one of the two samples, cGG (C4H6N2O2, mass m = 114 amu, CAS 106-57-0) and cGA (C5H8N2O2, mass m = 128 amu, CAS 4526-77-6) is in the form of powder at standard ambient temperature and pressure. The samples have been introduced in the vacuum chamber in a crucible and sublimated at a temperature of 80 °C and 90 °C for cGG and cGA, respectively.

### 2.2 Computational details

Quantum chemistry calculations were performed for the simulation of the photoelectron spectra and the fragmentation of all the diketopiperazines. Photoelectron spectra were estimated calculating ionisation potentials from valence shells using two different methods: (i) the equation of motion coupled cluster method including single and double excitation (EOM-CCSD) in combination with the Karlsruhe basis-set def2-QZVPP and (ii) the electron propagation theory (EPT) of correlated electron quasiparticle method within the Outer Valence Greens Function propagator, OVGF and the 6-311G** basis set, considering an NVE microcanonical-ensemble. For the study of the fragmentation of
the ionized DKPs we have used a combination of molecular dynamics simulation and exploration of the potential energy surfaces. This strategy has been previously employed with success for ionized molecules and molecular clusters.\textsuperscript{4–6,10,37–45} \textit{Ab initio} molecular dynamics allow following the fragmentation pathways for the first hundreds of femtoseconds after ionisation. We have used the atom-centred density matrix propagation (ADMP) formalism\textsuperscript{46–48} at the DFT level of theory, in particular using the B3LYP functional\textsuperscript{49–52} with the 6-31++G(d,p) basis set.\textsuperscript{53,54} All simulations were performed with a time step of $\Delta t = 0.1$ fs, a fictitious mass of $m = 0.1$ a.u., to ensure adiabaticity during the propagation, and a maximum propagation time of $t_{\text{max}} = 500$ fs. To mimic experimental conditions, a typical excess/excitation internal energy of $E_{\text{exc}} = 10$ and 15 eV, was used to consider states prepared upon ionisation, where the excess of excitation energy is randomly redistributed on the nuclear degrees of freedom. We considered a total of 100 trajectories for each value of excitation energy, and performed statistics over them (see the ESI). For the calculation of the energy barriers in the pathways leading to the observed exit channels, further exploration of the potential energy surface was carried out at the DFT B3LYP/6-311++G(d,p) level of theory. Relative energies in the PES were corrected with the zero point energy (ZPE). The combination of \textit{ab initio} molecular dynamics simulations and the exploration of the PES (see also the ESI) allowed us to obtain the fragmentation mechanisms that explain the moieties observed in the PEPICO experiments. All calculations were carried out using the Gaussian16 code.\textsuperscript{55}

3. Results and discussion

Photoelectron spectroscopy is a powerful tool to probe the relationship between the electronic structures of the neutral molecule and its cationic states reached in the photoionization event. This relationship depends on the measurable ionization energy and the change in the molecular geometry following...
ionization. Quantum chemistry methods calculate the minimum energy geometries of both initial neutral and final singly charged states and their results provide the assignment of the photoelectron spectrum. The experimental photoemission spectra of cGG and cGA measured in this work together with cAA from ref. 10 are shown in Fig. 1. The structure of the three molecules (see the insets in Fig. 1) is made by the central DKP ring and side chains made by the combination of one or two H atoms and methyl groups, depending on the constituent amino acids. Thus the three spectra share some common features, like the two bands at BE < 11 eV, a gap up to about 12 eV and then several broad features up to and above BE 20 eV. Both cGG and cGA spectra display a single feature just above 12 eV, which merges in a broader band (partially masked by the water contamination at 12.62 eV in the case of cAA). The cGG spectrum is consistent with the one previously reported in the literature. The energy of the first feature in the photoelectron spectra increases from 9.53 ± 0.05 eV in cAA,10 to 9.61 ± 0.05 eV in cGA and finally 9.88 ± 0.05 eV in cGG, mirroring the variation in the ionisation potential, IP, of the constituent amino acids, the IP being about 150 meV larger in glycine than in alanine.57 Furthermore, the ionization energies of the two amino acids57 are higher than the ones for the corresponding cyclodipeptides. This is at variance with the findings in cyclodipeptides with aromatic chains,58 where the orbitals of the DKP ring provide molecular stabilization with respect to the aromatic rings of the side chains, which are the frontier orbitals of the molecule.

Consistent with the experimental observations, the theoretical analysis of the PE spectra of the three diketopiperazine derivatives shows that the four lowest direct ionisation values (from HOMO to HOMO−3) require less than 11 eV of excitation energy, with a gap between this first group and the HOMO−4 above 12 eV. As the main structural differences among the three cyclic amino acids are the addition of one or two methyl group(s), it is reasonable to attribute this gap to their common features. Indeed, analyzing the orbital participation in the HOMO−4, we observe that in all three cases it corresponds to a combination of π orbitals in the carbonyl groups C=O of the DKPs, while the first four orbitals (HOMO to HOMO−3) are formed via the participation of the lone-pairs of the heteroatoms with some contribution from the bonding orbitals located in the ring.

Both theoretical methods employed to simulate the photoelectron spectra (OVGF and EOM-CCSD) provide similar results, and agree with the experimental results reproducing fairly well the measured bands (and gaps). In the energy region near the second ionization threshold the theory slightly overestimates the binding energies corresponding to ionization from inner valence electrons.

The photofragmentation of the cyclic dipeptides has been studied by analyzing the PEPICO spectra. This type of experiments provide state-selected mass spectra, where a fragmentation pattern can be correlated to specific molecular orbitals, within the experimental energy resolution. Furthermore, the first onset of the PEPICO yields provide a rough indication of the appearance energy of the fragment itself. As commonly observed in other polyatomic molecules, higher excitation energies, E_{exc} = BE − IP, are needed in order to release a larger number of smaller fragments due to increased bond breakage. On the other hand, the signal of the intact molecular ions for the three cyclodipeptides, m/z 142, 128 and 114 for cAA, cGA and cGG respectively, are observed only in the lower excitation energy region. In all cases, while only a couple of fragmentation channels exists in the region BE < 11 eV, in the energy range above 12 eV the molecules rapidly become unstable showing a
large amount of fragmentation products. Here, we find a relevant common feature in the three DKPs: the direct release at relatively low energies of neutral CO and HNCO, i.e. neutral molecules that have been regarded as heavily involved in the abiotic synthesis of prebiotic species (see e.g. ref. 59 and 60 and references therein). This behaviour was previously observed for cAA. As has already been discussed in relation with the photoelectron spectra (Fig. 1), at lower energies electronic density is localized in the bonds forming the DKP ring. Therefore their charge depletion may lead to the prompt opening of the ring. A wavefunction analysis performed using the Quantum Theory Atoms in Molecules of the most stable configuration for the three singly charged diketopiperazines, shows how the CO–CHCH₃ bond, for cAA and cGA, or CO–CH₂ bond, for cGG, are the most labile and consequently they will be the most likely ones to break upon ionization. The opening of the ring allows the direct release of neutral CO and HNCO moieties, which would lead to the m/z 114 (cAA), 100 (cGA) and 86 (cGG) fragments in the case of CO release, and to m/z 99 (cAA), 85 (cGA) and 71 (cGG) for the HNCO release. Note that it was observed how in the mass spectrum of cAA (Fig. 1 and 2 of ref. 9) the intensity of the charged fragment after CO release is much smaller than that for the HNCO-loss. A similar situation arises in cGA, where the equivalent fragment [cGA-CO]+ at m/z 100 is not observed within the sensitivity of the present experiment. As it will be shown below, this is the consequence of the higher stability of the charged product generated after the direct release of HNCO; on the other hand, less stable species are reached after the CO release and, therefore, less energy is needed for their further fragmentation.

Despite these common fragmentation aspects in the loss of CO and HNCO, we observe certain counter-intuitive trends that make the evolution of DKPs after ionization something far from trivial or general. In order to obtain insight into these differences, a combined experimental-theoretical methodology has been employed. Some of these results are reported in Table 1, together with the assignment of all relevant peaks in the PEPICO spectra. Experimentally, ion-neutral coincidence experiments have been performed to investigate the delayed emission of neutral fragments. Within the sensitivity of this technique as implemented in our apparatus, it is possible to track the fragmentation process Ion → Frag + Nfrag provided the fragmentation occurs during the residence time of the parent ion in the extraction/acceleration region of the spectrometer. Then, following the same approach as for cAA, Quantum Chemistry calculations have been performed to explore the potential energy surface of the two cyclic dipeptides cGA and cGG, and to determine all relevant fragmentation mechanisms. This provides the energy of the optimised structures of each fragment as well as the energy barriers connecting them along fragmentation patterns. Fig. 3 shows the ion-neutral coincidence

---

Table 1: The main fragments observed in the PEPICO measurement of cGA and cGG, Fig. 2, are reported with their proposed assignment, first onset in PEPICO experiments (estimated uncertainties 0.2–0.4 eV) and theoretically predicted appearance energy, AEth values.

<table>
<thead>
<tr>
<th>Assignment</th>
<th>m/z</th>
<th>PEPICO (eV)</th>
<th>AEth (eV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>cGA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M⁺</td>
<td>128</td>
<td>9.1³</td>
<td>9.30ᵇ</td>
</tr>
<tr>
<td>[M-CO]⁺</td>
<td>100</td>
<td>—</td>
<td>10.21/10.65</td>
</tr>
<tr>
<td>[M-HNCO⁺]</td>
<td>85</td>
<td>9.5</td>
<td>9.79</td>
</tr>
<tr>
<td>[H₂C₄NO⁺]</td>
<td>57</td>
<td>11.9</td>
<td>—</td>
</tr>
<tr>
<td>[H₂C₄N⁺⁻]</td>
<td>44</td>
<td>12.1</td>
<td>—</td>
</tr>
<tr>
<td>[HNCO⁺⁻]</td>
<td>—</td>
<td>43³</td>
<td>12.1</td>
</tr>
<tr>
<td>[H₂NCH₂⁺⁻]</td>
<td>30</td>
<td>11.9</td>
<td>30³</td>
</tr>
<tr>
<td>[CO⁺⁻]</td>
<td>28</td>
<td>13.9</td>
<td>28³</td>
</tr>
<tr>
<td>cGG</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M⁺</td>
<td>114</td>
<td>9.2³</td>
<td>9.40ᵇ</td>
</tr>
<tr>
<td>[M-CO]⁺</td>
<td>100</td>
<td>—</td>
<td>10.21/10.65</td>
</tr>
<tr>
<td>[H₂C₄NO⁺]</td>
<td>57</td>
<td>11.9</td>
<td>—</td>
</tr>
<tr>
<td>[H₂C₄N⁺⁻]</td>
<td>44</td>
<td>12.1</td>
<td>—</td>
</tr>
<tr>
<td>[HNCO⁺⁻]</td>
<td>—</td>
<td>43³</td>
<td>12.1</td>
</tr>
<tr>
<td>[H₂NCH₂⁺⁻]</td>
<td>30</td>
<td>11.9</td>
<td>30³</td>
</tr>
<tr>
<td>[CO⁺⁻]</td>
<td>28</td>
<td>13.9</td>
<td>28³</td>
</tr>
</tbody>
</table>

³ Adiabatic value. ^= Vertical value.
spectra of cAA and cGG (see ref. 10 for cAA). In these maps, the direct loss of CO and HNCO by the molecular ion (paths labeled A and B, respectively) are clearly visible as delayed fragmentation processes for both cGA and cGG. This supports the theoretical hypothesis of the relevance of the direct formation of HNCO and CO in the photodegradation of DKPs as proposed by the analysis of the AIMD. The corresponding theoretical results for energetics of these direct loss pathways are shown in the potential energy surface in Fig. 4, together with cAA reported in ref. 10 for comparison (for further PES exploration of the three DKPs see Fig. S3–S5 in the ESI†). Following the initial vertical ionization (VIP), adiabatic relaxation occurs, culminating in the singly charged form of the DKPs. We can identify this point as the adiabatic ionization potential (AIP), and compare it with the estimated appearance energy’s experimental value obtained in the PEPICO spectra. In all cases the mechanism starts with the ring opening after cleavage of the CO–CHCH₃ or CO–CH₂ bonds, followed by the release of a (Dion, Nfrag) pair in a further step. Lower energy barriers are needed to produce HNCO, while the release of CO implies overcoming transition states ~0.5 eV higher in energy.

A careful exploration of the ion-neutral coincidence maps for cGG, shows a fragmentation pathway that directly connects fragment 86⁺ with 30⁺. Considering the large difference in energy barriers needed to produce a very stable fragment CH₂NH₂⁺ (30⁺) (see Fig. 4) is the reason behind this experimental observation of the dominant 86⁺ → 30⁺ channel.

In Fig. 3 we see that for cGA, CO loss is followed by NH loss, but this sequential process does not occur for cGG, being the first step, CO loss leading to 86⁺, the only one observed. Most probably the low energy barriers needed to produce a very stable fragment CH₂NH₂⁺ (30⁺) (see Fig. 4) is the reason behind this experimental observation of the dominant 86⁺ → 30⁺ channel.

After the assignment of the main fragmentation paths resulting from the ionization of cGA and cGG, we now discuss the possible role of these molecules as prebiotic species. In a previous work, 42 we have shown theoretically how small neutral fragments obtained after the fragmentation of singly-charged cAA dipeptide, combined with a reactive charged moiety (5-oxazolidinones) that can also be obtained from isomerisation cAA⁺, may act as a neutral ‘seed’ and reactive agent, to support a mechanism of peptide chain elongation. Fig. 5 reports a schematic representation of the aziridine derivatives that may act as a highly reactive key intermediates of the mechanism for elongating peptide chains. In the first column, the three DKPs are displayed, with alanine units in red and those corresponding to glycine in blue. The second column features a representative peptide chain, highlighting these same units. The third column illustrates how cyclisation of each unit results in the formation of aziridine
derivatives. Then, if one considers cAA to be composed of two alanine unit cells (see the first structure in Fig. 5a) that have lost OH and H at the carboxylic and amino ends, respectively, a fragment of mass 71 amu is the ‘residue’, i.e. the unit cell that will be replicated in a peptide chain (see the second structure in Fig. 5a). That is, if we split symmetrically cAA in two identical fragments, each half would be a fragment with mass 71 amu. In a similar way, cGA and cGG, could be divided into their ‘unit cells’, where the one of glycine has a mass of 57 amu (see Fig. 5b and c).

After ionization, photodegradation may lead to the formation of both ‘seeds’ in their cationic or neutral forms. The latter implies the formation of a three-member ring and therefore the formation of aziridine derivatives42 (see the third structure of each series in Fig. 5). The fragmentation leading to the formation of these residues, which in the case of cAA provided the necessary ‘seed’ for further reactivity, looks trivial since only two bonds must be broken splitting the target molecules in half. Nevertheless, the capricious nature of these apparently very similar diketopiperazines shows up again. While fragment 57+ (attributable to the glycine ‘unit cell’) is one of the most prominent signals for cGA, it is not observed in the PEPICO spectra of cGG. Moreover, fragment 71+ (attributable to the alanine ‘unit cell’), while observed in cAA, it has not been detected in cGA measurements. The absence of fragment 57+ in cGG has already been previously discussed when analyzing the ion-neutral coincidence spectrum. Even though the fragment is formed in an intermolecular bonded structure, low-in-energy barriers lead to a fast evolution towards smaller fragments, making the 57+ moiety a metastable fragment not observed in the experiments. The reason why fragment 71+ is not formed in the fragmentation after ionization of cGA is not trivial. After analyzing the potential energy surface of this diketopiperazine, all possible reaction pathways that could potentially lead to this fragment lead instead to different ones. Therefore, consistent with the experimental evidence that fragment 71+ was not observed neither in PEPICO nor in ion-neutral coincidence measurements, also theoretically we could not identify any path leading to 71+ in our calculations (see Fig. S3–S5 of the ESI†). At the heart of this significantly different and quite unpredictable behaviour among the three cyclic dipeptides, the electronic chargedistribution controlling bonding properties and fast molecular rearrangement following photoionisation are likely playing a role.

As alternative ‘reactive moieties’ for peptide elongation we have also identified the 2- and 5-oxazolidinone species. The former can be obtained through the recombination of charged and neutral aziridin ‘seeds’ via a barrierless reaction as previously discussed for 71 in ref. 9. These mechanisms were obtained theoretically through relaxed scans, where the distance between the atoms forming the new bond is shortened while the rest of the structure is allowed to relax into the configuration of minimum energy (see Fig. 6a). 2-Oxazolidinone derivatives, on
the other hand, can be formed for every singly-charged diketo-piperazine after the first bond-break (see Fig. 6a); isomerization of the original ionized cyclic structure leads to the corresponding 2-oxazolidinone in a single step with relatively low energy barriers, accessible for the excitation in these experiments (see Fig. S3–S5 of the ESI†).

With these two ingredients, ‘seeds’ and ‘reactive moieties’, we have theoretically studied the mechanisms of elongation producing peptide chains, considering all possible combinations of both reactants. Fig. 7 shows some of the relevant mechanisms obtained through the relaxed scans. From the scan, we have located the highest points and optimized the geometry of the first order transition states that connect reactants and products. We can see in Fig. 7 that, for the example of the cationic 5-oxazolidinone derivative obtained from cGG and the arizidin-2-one moiety (neutral fragment of mass 57 amu), the reaction takes place with a low energetic barrier (0.66 eV). This behaviour is observed in all the reactions considered. In particular, the process proceeds as follows: first, a new peptide bond is formed as the seed’s amino group bonds with the carboxylic carbon of the oxazolidinone derivative; this triggers the opening of the oxazolidinone ring, creating the new peptide chain. Then, a new oxazolidinone ring is created with the atoms of the newly formed peptide bond. Thus, the final fragment has two differentiated parts: a non-reactive peptide chain and a reactive oxazolidinone ring that could be involved...
in further capturing of new ‘seeds’, resulting again in the elongation of the peptide chain and reconstitution of the reactive oxazolidinone site (see the lowest in energy reactive channel in Fig. S3–S5 in the ESI†).

This theoretically proposed mechanism suggests a new possible route for peptide elongation in the gas phase. However, it doesn’t solve the question of homochirality in life. Oxazolidinones have been recognized as non-desired intermediates for peptide elongation as their tautomerization can lead to the loss of optical purity. Within the proposed mechanism one could argue that, even though tautomerization is possible, the energetic barriers for peptide elongation are much lower compared to those for tautomerization (see Fig. S6 in the ESI†). Additionally, the proposed mechanism retains the original chirality of its components. As in our experiments and calculations we have only considered ‘S’ enantiomers, all longer peptides show the same chirality. However, if the fragments had ‘R’ chirality, the same energetic barriers would take place for peptide elongation, and therefore, although conserving the optical purity, would not show the preference or an imbalance between R and S enantiomers.

4. Conclusions

In summary, we have studied the ionization and subsequent fragmentation of two cyclic dipetides, cyclo(alanyl-glycine) cGA and cyclo(glycyl-glycine) cGG, induced by VUV radiation. In particular, we report results from three different techniques: (i) photoelectron spectroscopy, which allows us to characterize the electronic structure and ionization energies of different states; (ii) photoelectron–photoion coincidence measurement (PEPICO) spectroscopy, which allows identifying state-selected fragmentation channels; and (iii) ion-neutral coincidence maps, which allow providing unambiguous assignment of some reaction paths. In addition to these experiments, we have performed a thorough computational study using quantum chemistry simulations, where (i) ionization has been simulated using the equation of motion coupled cluster method (EOM-CCSD) and the electron propagation theory (EPT); (ii) fragmentation was studied by means of molecular dynamics with the atom-centred density matrix propagation (ADMP) formalism; and (iii) the potential energy surface was exploited using density functional theory (DFT), locating the transition states and the intermediates in the paths of the most relevant fragments observed experimentally, thus proposing the corresponding fragmentation mechanisms. These calculations provide insight into the different processes taking place in the photo-ionization induced fragmentation of these two molecules. We have compared our results to those previously obtained for the cyclo(alanyl-alanine) cAA.9,10

Starting from the study of the molecular rearrangement and fragmentation following VUV ionisation, and investigating the reactivity among some of the products, we also explored theoretically the potential of DKP species to survive radiation or even act as intermediates for the construction of oligopeptides, as already proposed in cAA.9 We have shown that in the three cationic diketopiperazines (DKPs), cGG, cGA and cAA, a similar evolution occurs at the beginning, just after photoionization, i.e. ring opening at the C–C bond that has been weakened to a greater extent in the cationic form with respect to the neutral canonical molecule. This ring opening can be followed either by the loss of neutral moieties, or by molecular rearrangement. In the first situation, neutral CO and HNCO are emitted with high probability, i.e. low energy barriers. In the second one, cationic oxazolidinone structures are formed.

Finally, we have explored the possibility of peptide bond formation using the fragmentation products as ‘reagents’, thus leading to the formation and elongation of a peptide chain. To this end, we have shown that two ‘reactive moieties’ are needed:
cationic oxazolidinones and neutral aziridin molecules, and both are products that can be obtained in the VUV irradiation of DKPs molecules. However, specific oxazolidinone and aziridin structures are obtained for each cyclic dipeptide. Thus, depending on which DKP is irradiated, different peptide chains might be formed via different mechanisms, making this process neither trivial nor general for cyclic dipeptides.

The joint experimental-theoretical approach here adopted is crucial to obtain a complete picture of the great variety of processes that occur in the photo-ionization of peptides, in the identification of the species produced, as well as in their reactivity. Although the three discussed molecules a priori look like very similar, specificity is observed in their chemical behaviour and in their response to photoionization.

Conflicts of interest
There are no conflicts to declare.

Acknowledgements
This article is based upon work from COST action CA18212 – Molecular Dynamics in the GAS phase (MD-GAS), supported by COST (European Cooperation in Science and Technology). The authors acknowledge the generous allocation of computer time at the Centro de Computación Científica at the Universidad Autónoma de Madrid (CCC-UAM). This work was partially supported by MICINN (Spanish Ministry of Science and Innovation) project PID2019-110091GB-I00 funded by MCIN/AEI/10.13039/501100011033, the “María de Maeztu” (CEX2018-000805-M) Program for Centers of Excellence in RD, MAECI Italy-Sweden project CTQ2016-76061-P. H. Z. acknowledges the FPI grant associated with MICINN project PRIN 20173B72NB project “Novel molecular tools for the exploration of the nanoworld”, and FPI grant 2018-105800-J-I00000 funded by Ministerio de Economía y Competitividad. This research was funded by the grant PGC2017-87752-C2-1-I00 and the Carl Trygger Foundation (Grant no. 17:436).

Notes and references
Paper

PCCP