Chirality-dependent structuration of protonated or sodiated polyphenylalanines: IRMPD and ion mobility studies†

Valeria Lepere, Katia Le Barbu-Debus, Carine Clavaguéra, Debora Scuderi, Giovanni Piani, Anne-Laure Simon, Fabien Chirot, Luke MacAleese, Philippe Dugourd and Anne Zehnacker*

Ion mobility experiments are combined with Infra-Red Multiple Photon Dissociation (IRMPD) spectroscopy and quantum chemical calculations for assessing the role of chirality in the structure of protonated and sodiated di- or tetra-peptides. Sodiated systems show a strong chirality dependence of the competition between Na⁺…O and Na⁺…π interactions. Chirality effects are more subtle in protonated systems and manifest themselves by differences in the secondary interactions such hydrogen bonds between neutral groups or those involving the aromatic rings.

Introduction

Chirality has ubiquitous effects in chemistry, both at the microscopic and macroscopic levels. The desire of studying these effects without the perturbation brought by the environment has prompted several gas-phase approaches. Isolated weakly bonded diastereomers resulting from the non-covalent association of two chiral molecules have been the subject of numerous studies. They include spectroscopic studies of neutral complexes in supersonic-jet conditions and that of ionic species by mass spectrometry, with methods ranging from collision-induced dissociation, to the kinetic method or its variant, ion-molecule reactions, or ion mobility. A few studies coupling spectroscopy and mass spectrometry have also been reported. In contrast, spectroscopic studies of the differences between diastereomers isolated in the gas phase are scarce, despite their interest. Indeed, diastereomers often possess different macroscopic chemical and biological properties, which stem from both different intra and intermolecular interactions. For example, the effect of ephedrine on blood pressure modification is twice that of pseudo ephedrine. The two molecules also show microscopic differences, both in the number of populated conformers in jet-cooled conditions and in the strength of their intramolecular hydrogen bond. While all the naturally-occurring aminoacids are L, post translational modifications can invert chirality, which has possible consequences on their structure and function. For example, reversal of chirality of a single amino-acid in the Trp-cage dication leads to structural modifications of its gas-phase structure detected by Electron Capture Dissociation (ECD). Also in isolated peptides of more limited size, inverting the chirality of a residue may modify the intramolecular interactions hence the structure. This has been observed in capped neutral dipeptides showing a γ turn, which exists in two opposite handednesses whose relative stability is dictated by the chirality of the residues. However, the chirality-induced structural changes seem to be limited in small peptides. While the reversal of chirality of one amino-acid unit in long protonated poly-alanine strands is sufficient to disrupt the helical shape adopted in the peptide when all the residues are of the same chirality, no difference has been observed between protonated or neutral LL and LD diphenylalanines. In contrast, a marked chirality effect has been observed in the binding pattern between the same dipeptide and Li⁺ or Na⁺. The observed binding motifs involve coordination of the alkali

PCCP
to the N-terminus aromatic ring, to the amide oxygen, and to one of the termini oxygens. While coordination to either of the termini has been proposed for LL, interaction with the carboxyl oxygen has been suggested to be energetically preferred in LD. The aim of this work is to extend the study of Dunbar et al. on diphenylalanines to tetraphenylalanines, either isotactic LLLL or syndiotactic LDLD. The question raised is whether chirality-induced structural differences also appear in the protonated forms when the sequence is longer. The aromatic character of the phenylalanine residue is expected to modify the competition between the strong electrostatic interaction between the cation and the peptide on the one hand, and the internal stabilization of the peptide on the other hand. Indeed, the aromatic rings can compete with C=O or N as a solvation site; they will increase stabilizing dispersion interactions within the peptide, and may intervene through the formation of NH···π hydrogen bonds which have been shown to play an important role in chiral recognition. Reversal of the chirality of one or several residues can modify the subtle balance between these different intra and intermolecular forces.

To this end, we combine two technical approaches, namely Infra-Red Multiple Photon Dissociation (IRMPD) and Ion Mobility Spectrometry (IMS). IRMPD is a very powerful technique for the structural characterization of ions in the gas phase. It has been applied to protonated peptides such as Gly or Ala, for which several binding sites have been suggested. The interaction between alkali and Gly, Ala or Phe has been the subject of numerous experimental and theoretical studies, with the aim of understanding the binding motif of the cation, either fully solvated by oxygen atoms or interacting also with one of the nitrogens. However, IRMPD mainly probes the hydrogen bond network at play in these systems. In the polyphenylalanines studied here, the position of the phenyl rings is not directly probed by the IR. A method sensitive to the global shape of the molecule is therefore highly desirable. IMS provides information on the 3D structure of charged biomolecules in the gas phase. Since the pioneering works of Bowers, Jarrold and Clemmer, numerous studies have been undertaken on systems, ranging from peptides and DNA strands to isolated proteins and complexes. These two methods (IRMPD and IMS) nicely complement each other as global information on the molecular shape extracted from ion mobility is combined with local information about the hydrogen bond network inferred from the IRMPD results. Such a dual approach was proposed already to probe the charge-solvated structure of guanidinium-carbonylpyrrole derivatives or to study the structural properties of metal–peptide complexes.

We present here a study of the influence of chirality on the structure of protonated (H+) or sodiated (Na+ in what follows) isotactic or syndiotactic di- or tetraphenylalanines, obtained by combining the results of IRMPD spectroscopy, ion mobility and quantum chemical calculations. Fig. 1 schematically shows the studied molecules, with the chiral centers denoted by *. The peptides L-Phe–L-Phe, L-Phe–D-Phe, L-Phe–l-Phe–l-Phe and l-Phe–d-Phe–l-Phe–d-Phe will be denoted LL or LD and LLLL or LDLD, respectively.

Experimental and theoretical methods

LL or LD diphenylalanines and LLLL or LDLD tetraphenylalanines were purchased from GeneCust-Luxembourg and used without further purification. Protonated or sodiated ions were produced by Electrospray Ionization (ESI) of 500 μM peptide solutions prepared by dissolving the polypeptides in a H2O/MeOH 50:50 mixture.

I. IRMPD spectroscopy

IRMPD spectra were obtained through MS2 experiments by means of two set-ups. The first one was a modified Paul ion trap (Bruker, Esquire 3000+) equipped with a Diamond entrance window and coupled to a Free Electron Laser (FEL) at the Centre Laser Infrarouge d’Orsay (CLIO). The FEL was focused by a 350 mm focal length ZnSe lens. IR spectra were obtained by mass selecting the precursor ions in a 2 Da window, with 300 ms irradiation time. Mass spectra were recorded after six accumulations, and the photon energy was scanned by steps of 4–5 cm−1. The second set up involved a 7T Fourier transform ion cyclotron resonance (FT-ICR) hybrid mass spectrometer (Bruker, Apex Qe). The CLIO FEL IR beam was mildly focused by a 2000 mm Ag-protected spherical concave mirror, with a typical beam waist at the centre of the FT-ICR of the order of 1 mm diameter. The CLIO FEL operated at 25 Hz with 8 μs long bunch pulses containing 0.5–3 ps micropulses separated by 16 ns. The spectral bandwidth (full width at half maximum FWHM) was about 7 cm−1 with pulse energy of 1600 to 900 mW from 900 to 2000 cm−1. The wavelength calibration was ensured by simultaneously recording the spectrum of polystyrene and that of the studied system.

Infrared spectra were obtained by monitoring the fragmentation efficiency $F = -\ln(P(F + P))$ as a function of the IR wavelength, with $F$ being the sum of the abundances of the fragment ions produced by IRMPD and $P$ that of the parent ion.
II. Ion mobility

Mobility measurements were performed using a homemade 1 m long drift cell coupled to a commercial quadrupole time-of-flight (micro-qTOF, Bruker-Daltonics, Bremen, Germany, mass resolution 10 000), as previously described in details.52

Electrosprayed ions were introduced through a heated capillary (around 250 °C) and accumulated in an hourglass-shaped ion funnel. Ion packets were periodically injected in the drift tube, across which they travelled, driven by a uniform electric field operated between 8.3 and 4.7 V cm⁻¹. The helium pressure in the tube was maintained at 12 Torr. A second ion funnel was used to focus the diffuse ion packet at the exit of the drift tube and to guide them to the transfer region of the mass spectrometer through a 0.7 mm diameter aperture. Ions were then conveyed to the orthogonal extraction region of the time-of-flight mass analyzer which is operated at 10 kHz. Mass spectra could eventually be recorded as a function of the drift time of the ions in the tube. Under the experimental conditions, this drift time was inversely proportional to the drift field, and proportional to the orientationally-averaged momentum transfer cross section for ion drift gas collisions.64 Absolute values for these collision cross sections could then be extracted from drift time measurements at different drift voltages across the mobility cell. Following this procedure, the uncertainty on the measured cross sections was estimated to be no more than 3%.

III. Theoretical methods

The di- and tetra-peptides studied here are shown in Fig. 1. The atoms are numbered according to the residue they are contained in, starting from the N terminus. The ionic intra-molecular hydrogen bonds involved in the protonated peptides will be named Cn according to the number of atoms contained in the ring. Additional interaction involving neutral hydrogen bonds will be called C′n. Complexes of the oligopeptides with sodium will be named after the nature of the interaction with the alkali, i.e. an interaction between sodium and the carbonyl oxygen, the nitrogen, or the aromatic ring will be called O, N, π, respectively. For example, a structure with Na⁺ interacting with two carbonyl oxygens and an aromatic cloud will be called OOπ.

For all the studied systems, exploration of potential energy surface was performed combining two different force fields (OPLS-2005 and MMFFs) with the advanced conformational search implemented in the MacroModel program available in the Schrödinger package.65 At the end of the conformational searching step, all the peptide structures with energy below 20 kJ mol⁻¹ were optimized within the frame of the DFT theory at the RI-B97D/TZVPP level. The most promising conformers in terms of energetics and spectroscopy were further optimized using the dispersion-corrected functional B3LYP-D3 associated to the TZVPP basis set. Based on previous studies,66,67 this method was chosen because it combines a good frequency description68 and inclusion of empirical dispersion corrections.69,70 Final relative energies have been confirmed at the RI-CC2/TZVPP level on the B3LYP-D3 geometries. The calculations were performed with the Turbomole 6.671 and the Gaussian packages.72 The vibrational spectra were simulated by convoluting the harmonic frequencies obtained at the same level of calculation by a Lorentzian line shape (FWHM 20 cm⁻¹). For comparison with IMS measurements, collision cross sections were also calculated for the selected conformers using the trajectory method described by Mesleh et al.73

Experimental results

I. Mass and CID spectra of polyphenylalanines

The mass spectra of protonated or sodiated di- and tetra-phenylalanines were recorded prior to spectroscopic studies (Fig. 2). The same peaks with similar intensities appear in the mass spectrum of LLH⁺ and LDH⁺. The protonated monomer appears at m/z 313.4, the dimer at m/z 625.8 and the trimer at m/z 938.2. Fragments appear at m/z 120.4 and m/z 166.4. The mass spectra of LLLLH⁺ and LDLDH⁺ differ much more from each other than those of the dipeptides do. They are dominated by the protonated monomer (m/z 607.7), and dimer (m/z 1214.4). The trimer (m/z 1821.1) as well as a

![Fig. 2](image-url) Mass spectra of an acidified (a) dipeptides and (b) tetrapeptides solution.
doubly-charged trimer (m/z 911.1) and pentamer (m/z 1517.8) strongly contribute to the spectrum of LDLDH+, but are hardly seen in LLLLH+. The heterochiral system shows therefore much more clustering propensity. In contrast with that of LDLDH+, the mass spectrum of LLLLH+ shows extensive fragmentation.

Complexation with sodium impurity is also observed as a weak contribution at m/z 335.4 for diphenylalanines and m/z 629.7 for tetraphenylalanines.

MS2 Collision-Induced Dissociation (CID) spectra of LLH+ and LDH+ show several fragments. The main fragmentation path corresponds to carboxyl loss (m/z 268.4), y1 fragment (m/z 166.4) and a1 fragment (m/z 120.4). The peaks at m/z 295.4 and m/z 222.4 which correspond to the loss of water (m/z 295.4) and of the radical residue (Ph-CH2*) (m/z 222.4) are about one order of magnitude less intense. LLNa+ and LDNa+ show the same fragments as LLH+ and LDH+, except the loss of water, with however very different relative intensities. y1 is dominant while the intensity of the other fragments is one to two orders of magnitude smaller. The influence of the chirality of the residues manifests itself by a smaller efficiency for the carboxyl loss in LLNa+ than LDNa+.

The MS2 CID spectra of LLLLH+ and LDLDH+ show the same fragments, though LLLLH+ shows higher fragmentation efficiency than LDLDH+. The spectra are dominated by the y2 and b2 fragments which correspond to the dissociation into two dipeptide moieties. The dominant channel is b2/y2 followed by b3/y3 then a2. In contrast to what is observed for the protonated dipeptides, loss of water or carboxyl group is negligible. The relative intensity of the fragments is similar in LLLLH+ and LDLDH+ except the y1 fragment which is more intense in LLLLH+. CID of sodiated tetraphenylalanines mainly leads to the loss of sodium, which is not observed in our experimental set up. The other fragments amount to 1% of the parent intensity (y1) or less (b3, y2, a1). Loss of water is only present for LLLLNa+.

II. IRMPD spectra of protonated LL and LD

The IRMPD spectra of protonated LL and LD recorded in the fingerprint region are shown in Fig. 3. They are very similar and present several bands. The band around 1150 cm⁻¹ is the signature of the carboxylic OH bend devoid of hydrogen-bonding interaction. The band around 1400 cm⁻¹ corresponds to the NH3+ umbrella motion, and that around 1530 cm⁻¹ to the NH3+ scissoring motion. Last, the higher-energy features around 1700 and 1760 cm⁻¹ are typical of the carbonyl and carboxyl CO stretching modes. The only difference between the LL and LD spectra is the higher intensity of the band at 1530 cm⁻¹ in the former. The bands are listed in Tables S1 and S2 in the ESI†.

These results can be compared to those of Dunbar et al., who reported the IRMPD spectrum of protonated LL and LD recorded with a FT-ICR ion trap coupled with the FELIX IR free-electron laser. The two spectra are identical except for a band at 1100 cm⁻¹ which is not present in the spectra reported here for either of the systems. The differences observed between the two set-ups might arise from different trapping conditions, leading to different temperatures and conformer populations. They might also come from different irradiation conditions, such as laser power, or spectral bandwidth, which influence in a different way the absorption of modes with different anharmonicity.

III. IRMPD spectra of sodiated LL and LD

The IRMPD spectra of sodiated LL and LD are presented in Fig. 4. The bands are listed in Tables S3 and S4 in the ESI†. Strong bands appear around 1700 cm⁻¹, which are characteristic of the CO stretching modes. The features around 1540 cm⁻¹ correspond to the NH bending modes and that around 1160 cm⁻¹ to the carboxylic OH bend. Differences in the position of the bands appear for the two systems of different chirality. While the band at 1163 cm⁻¹ in LLNa+ is blue shifted by 8 cm⁻¹ in LDNa+, the bands at 1537 and 1754 for LLNa+ are moved towards the red at 1526 and 1742 for LDNa+. Finally, the major difference is a shoulder at 1780 cm⁻¹ observed for LLNa+, which is not present in LDNa+. This observation is in agreement with the studies of Dunbar et al. and can be explained in terms of the coexistence of two conformers. As mentioned for the protonated species, the intensity ratio and bandwidths observed here slightly differ from those previously reported, in particular the intensity ratio between the 1742 and 1780 cm⁻¹ bands observed in LLNa+. This observation suggests that the population distribution between the isomers is not the same in the two experiments.

The sodiated systems do not show any band at 1485 cm⁻¹, as expected when no NH3+ group is present in the system. This observation proves unambiguously that formation of a zwitterion can be excluded.

IV. IRMPD spectra of protonated LLLL and LDLD

The IRMPD spectra of protonated tetraphenylalanines are presented in Fig. 5 and the position of the bands is listed in Tables S5 and S6 (ESI†). Five different regions appear for LLLLH+.
The presence of three distinct CO stretching bands between 1600 and 1800 cm\(^{-1}\) suggests the existence of CO groups involved in three different interactions. Two strong peaks appear at 1440 and 1500 cm\(^{-1}\) in the NH\(_2^+\) umbrella and NH\(^+_2\) scissoring region. Last, the region of the OH bend shows a single feature at 1141 cm\(^{-1}\) band. The LDLDH\(^+\) spectrum is less structured. A congested band appears around 1700 cm\(^{-1}\) for the CO stretch, suggesting that the carbonyls are in similar environments. The NH bend region shows a single band around 1550 cm\(^{-1}\). Besides much less structured features in the NH bend and CO stretch regions in LDLDH\(^+\), the main difference relative to LLLLH\(^+\) is the presence of a band at 1241 cm\(^{-1}\) and the absence of that at 1141 cm\(^{-1}\). It has to be stressed that the spectra of protonated tetrapeptides are much more sensitive to chirality than those of dipeptides.

V. IRMPD spectra of sodiated LLLL and LDLD

The spectra of the sodiated tetraphenylalanines are presented in Fig. 6 and the bands are listed in Tables S7 and S8 (ESI†). They are dominated by two intense and broad bands around 1680 cm\(^{-1}\), characteristic of the CO stretches, and around 1500 cm\(^{-1}\), typical of the NH bending modes. The weak band at 1750 cm\(^{-1}\) corresponds to the carboxylic CO stretch and is abnormally weak, as frequently observed for this sort of systems.\(^{75}\) This might be due to two reasons. First, the laser power strongly decreases in this range (it falls from ~1.6 W at 900 cm\(^{-1}\) to 0.7 W at 1800 cm\(^{-1}\)). Second, the main dissociation channel is the loss of Na\(^+\), which is not observed in the experimental set-up used. This might affect the relative intensities of the band. The LLLLNa\(^+\) and LDLDNa\(^+\) only differ by the relative intensity of the 1680 cm\(^{-1}\) peak. As in the case of diphenylalanines, the absence of a band around 1450 cm\(^{-1}\) indicates that no zwitterion is formed.
Table 1  Relative free energies (kcal mol$^{-1}$) at the B3LYP-D3/TZVPP level together with experimental and calculated (at the B3LYP-D3/TZVPP geometries) collision cross sections ($\AA^2$) of the studied systems

<table>
<thead>
<tr>
<th></th>
<th>$\Delta G$ (kcal mol$^{-1}$)</th>
<th>Calculated cross section ($\AA^2$)</th>
<th>Experimental cross section ($\AA^2$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LLH$^+$</td>
<td>0.0</td>
<td>115</td>
<td>119</td>
</tr>
<tr>
<td>LLcalc1</td>
<td>3.2</td>
<td>113</td>
<td>116</td>
</tr>
<tr>
<td>LDH$^+$</td>
<td>0.0</td>
<td>111</td>
<td>124</td>
</tr>
<tr>
<td>LDcalc1</td>
<td>3.2</td>
<td>112</td>
<td>116</td>
</tr>
<tr>
<td>LDNa</td>
<td>0.0</td>
<td>124</td>
<td></td>
</tr>
<tr>
<td>LDLH$^+$</td>
<td>0.3</td>
<td>118</td>
<td>116</td>
</tr>
<tr>
<td>LLLLH$^+$</td>
<td>0.0</td>
<td>178</td>
<td></td>
</tr>
<tr>
<td>LLLLcalc1</td>
<td>0.0</td>
<td>167</td>
<td></td>
</tr>
<tr>
<td>LLLLcalc2</td>
<td>0.7</td>
<td>162</td>
<td>183</td>
</tr>
<tr>
<td>LLLLNa</td>
<td>0.0</td>
<td>174</td>
<td>188</td>
</tr>
<tr>
<td>LDLDNa$^+$</td>
<td>0.0</td>
<td>184</td>
<td></td>
</tr>
</tbody>
</table>

### VI. Ion mobility measurements

The experimental collision cross sections (Table 1) deduced from the IMS measurements for LLH$^+$, LDH$^+$, and LLNa$^+$ (119, 116, and 116 $\pm$ 3 $\AA^2$, respectively) are very similar considering the experimental error bars. In contrast, the cross section of 124 $\pm$ 3 $\AA^2$ measured for LDNa$^+$ is significantly higher.

Regarding the tetraphenylalanines, the experimental cross sections reported in Table 1 display a clear difference between the protonated systems, for which the cross section is 178 $\pm$ 3 $\AA^2$, and the sodiated systems, which were found to show significantly higher cross sections (183 and 188 $\pm$ 3 $\AA^2$ for LLLLNa$^+$ and LDLDNa$^+$ respectively). The protonated systems then appear to be more compact than the corresponding sodiated species. However, no marked difference as a function of the residue chirality could be detected for the sodiated tetrapeptides.

### Calculation results and attribution

1. Calculated structures of diphenylalanines

The assignment has been made on the basis of a good agreement between the experimental spectrum and that simulated from the calculated harmonic frequencies and between the measured and calculated values for the collision cross sections. Among the calculated structures that satisfy these conditions, we have kept those that are the most stable. The selected diphenylalanine structures are shown in Fig. 7. Table 1 presents their calculated cross sections and relative free energies obtained at the B3LYP-D3/TZVPP level. The calculated frequencies are listed in Tables S1–S8 in the ESI.

1) Protonated systems. The calculations yield six LLH$^+$ structures with Gibbs free energy below 2 kcal mol$^{-1}$ for LL and five for LDH$^+$. They all display a C5 interaction between N(1)H$^+$ and CO(1). The two stable LL dipeptides which meet the requirements on both vibrational spectroscopy (Fig. 3) and ion mobility are LLcalc1 and LLcalc2. Both structures display a similar hydrogen bond pattern, with a strong ionic C5 interaction between N(1)H$^+$ and CO(1) and a weak C5 neutral hydrogen bond between the peptide N(2)H and the carboxylic CO(2). The structures involving an interaction between N(1)H$^+$ and the carboxylic CO(2) are much higher in energy and have been discarded, as well as the structure where the NH$^+$···OC hydrogen bond is replaced by NH$^+$···aromatic ring interactions only.

While the C5 neutral hydrogen bond between the peptide N(2)H and the carboxylic CO(2) is identical in LLcalc1 and LLcalc2 (2.40 vs. 2.42 Å, respectively), it is not the case for the strong interaction between N(1)H and CO(3). In LLcalc1, there is a compromise between the N(3)H$^+$···O(1) and the N(3)H$^+$···aromatic ring interaction. The resulting tradeoff leads to a quite large N(3)H$^+$···O(1) distance (2.03 Å), while the two other NH occupy equivalent positions and bridge the two aromatic rings which are parallel to each other. In contrast in LLcalc2, N(3)H$^+$ optimizes its interaction with CO(1p), resulting in a closer N(3)H$^+$···O(1) distance (1.73 Å) and a larger distance between the aromatic rings which adopt a “V” shape. Although both satisfactorily match the experimental data, we tend to favour LLcalc1 for energetic reasons as it is lies 3.2 kcal mol$^{-1}$ lower in energy, at the B3LYP-D3 level.

The cross section of the selected protonated LDcalc structure amounts to 110 $\AA^2$. This structure shows a very strong ionic N(1)H$^+$···O(1) bond (1.73 Å) and no neutral C5 N(2)H···O(2) interaction. The two aromatic rings are perpendicular to each other; it allows one of the ammonium N(1)H to interact with one of them and thereby compensates for the lack of neutral C5.

The main difference between LL and LD (absence of neutral C5 and perpendicular aromatic rings in LD) does not show clear vibrational signature. However, it manifests itself by a difference in drift time. Indeed, the theoretical results fairly well reproduce the fact that the collision cross section of LL (119 $\AA^2$) is slightly larger than that of LD (116 $\AA^2$).

2) Sodiated systems. The calculations yield three LLNa$^+$ structures with Gibbs free energy below 2 kcal mol$^{-1}$, which show either OOπ or Oππ interaction. The three calculated LD Na$^+$ structures are all of OOπ nature. The vibrational spectra of the sodiated dipeptides have been reported by Dunbar et al. together with calculations supporting the existence of a very stable OOπ form, common to the two diastereomers. The results presented here (Fig. 4) parallel the latter results, as the very stable LLNacalc1 possesses such a structure and has a satisfactory cross section of 115 $\AA^2$. In LLNacalc2, the interaction between the alkali and the CO groups is strong as indicated by a short distance, 2.21 and 2.27 Å for the amide and carboxyl CO, respectively. The Na$^+$···π distance is short too, 2.90 Å. The two aromatic rings are in an extended geometry and do not interact with each other. The structure displays a C5 neutral hydrogen bond between the peptide N(2)H and the N(3) terminal. The second stable form compatible with the experimental results, namely LLNacalc2$^+$, is an Oππ structure in which the alkali interacts with the amide CO(1) (distance of 2.16 Å) and the two aromatic rings which are therefore kept parallel to each other. The peptide N(2)H is involved in a bifurcated C5 neutral hydrogen bond involving both N(1) and the carboxylic CO(2). This structure is...
different from the NO\textpi structure proposed by Dunbar \textit{et al.} Indeed, both NO\textpi and O\textpi structures account well for the high-energy shoulder of the band due to the CO stretch. However, the NO\textpi structure is calculated much higher in energy here (2.7 kcal mol\textsuperscript{-1} at the B3LYP-D3/TZVPP level). Moreover, its calculated cross section amounts to 119 Å\textsuperscript{2}, which is larger than expected from the measurements.

The sodiated LD dipeptide stands out with the largest experimental cross section among all the dipeptides studied. Among the calculated sodiated LD dipeptides which reproduce the experimental IRMPD spectrum (Fig. 4), the calculated cross section of LD\textit{calc} \textsuperscript{−} shown in Fig. 7, of OO\textpi nature, is the most compatible with the experiment. Its hydrogen bond pattern is similar to that observed in LLNa\textit{calc} \textsuperscript{−}, with a distance between the alkali and the amide or carboxyl oxygens of 2.20 or 2.30 Å, respectively. The effect of chirality is that the LD\textit{calc} \textsuperscript{−} differs from LLNa\textit{calc} \textsuperscript{−} by the outwards orientation of the aromatic ring, which explains the larger experimental and calculated cross section.

The results obtained here align with those previously reported; only one form of sodiated LD is observed in our experimental conditions while two forms coexist for LL. Moreover, the differences in experimental frequencies are qualitatively reproduced by the calculations; the NH band is red shifted and the OH bends blue shifted in LDNa\textsuperscript{+}.

II. Calculated structures of tetraphenylalanines

The assignment of the observed tetrapeptides to one or several of the calculated forms has been done following the same line as for the dipeptides. The selected structures are shown in Fig. 8.

(1) Protonated systems. The calculations yield two LLLLH\textsuperscript{+} and seven LDLDH\textsuperscript{+} structures with Gibbs free energy below 2 kcal mol\textsuperscript{-1}. Among the calculated structures, only two LDLDH\textsuperscript{+} and one LLLLH\textsuperscript{+} match the experimental spectra. They are further sorted out by their energy and mobility. The LLLL\textit{calc} structure shown in Fig. 8 presents satisfactory agreement with both the vibrational spectroscopy (Fig. 5) and ion mobility results, and is the most stable at the B3LYP-D3/TZVPP level. This structure involves a strong C11 interaction between N\textsubscript{(3)}H\textsuperscript{+} and the C terminal CO\textsubscript{(3)}, with a N\textsubscript{(3)}H\textsuperscript{+}−O\textsubscript{(3)} distance as short as 1.66 Å. The two other N\textsubscript{(3)}H\textsuperscript{+} are involved in interactions with the aromatic rings located on the C and N termini (z\textsubscript{1} and z\textsubscript{4} positions), with distances between N\textsubscript{(3)}H\textsuperscript{+} and the aromatic rings in z\textsubscript{1} and z\textsubscript{4} of 2.4 and 2.6 Å, respectively. The latter larger distance is due to the fact that the N\textsubscript{(3)}H\textsuperscript{+} interacting with the aromatic ring in z\textsubscript{4} position is also involved in a C5 ring with CO\textsubscript{(1)}. These N\textsubscript{(3)}H\textsuperscript{+}−π interactions keep the two aromatic rings parallel, and make the structure rather compact, with a calculated cross section of 167 Å\textsuperscript{2}. The carboxylic acid group is directed outwards and only interacts with the rest of the molecule via its CO\textsubscript{(4)} involved as an acceptor in a neutral C′5 interaction with N\textsubscript{(4)}H. The fact that the carboxylic group does not act as a hydrogen bond donor manifests itself by the value of the C−OH bend (1173 cm\textsuperscript{-1}). Lastly, the other amide carbonyl CO\textsubscript{(3)} and the peptidic NH\textsubscript{(1)} are free.

LDLD\textit{calc} \textsuperscript{1} is the most stable LDLD structure; its collision cross section is also the most satisfactory. Its N\textsubscript{(3)}H\textsuperscript{+} is involved in a similar interaction pattern as in LLLL\textit{calc} \textsuperscript{1}, namely a C11 and two NH−π. However, the rings involved in the interactions are not the same as in LLLL\textit{calc} \textsuperscript{1}, but are in z\textsubscript{1} and z\textsubscript{3} positions. LDLD\textit{calc} \textsuperscript{1} strongly differs from LLLL\textit{calc} \textsuperscript{1} by the environment of the carboxylic acid. A C′11 interaction takes place between the carboxylic carbonyl CO\textsubscript{(0)} and the amide N\textsubscript{(3)}H. This interaction is accompanied by a weak C\textsubscript{(4)}OH−π interaction with the aromatic ring in z\textsubscript{2} position, with an OH−π distance of 2.58 Å. The spectroscopic fingerprint of this interaction is a blue shift of the C−OH bending mode which is calculated at 1199 cm\textsuperscript{-1}, as observed in sodiated polyglycines. None of the CO or NH are free as two C′7 interactions take place between C\textsubscript{(3)}O−HN\textsubscript{(4)} and C\textsubscript{(4)}O−HN\textsubscript{(3)}.
In LDLD_{calc2}, the N(1)H$_3^+$ ammonium interacts with C(3)O and the aromatic ring in $a_3$, like in LDLD_{calc1}, but there is no interaction with another aromatic ring. The same C'11 and C'7 interactions are observed. LLLL and LDLD have the same experimental cross section, which should be reproduced by the calculation. In this respect, the LDLD_{calc1} structure fits the experimental cross section better than LDLD_{calc2}, although both of them are satisfactory in terms of spectroscopy.

In conclusion, the LLLL and LDLD systems share in common a strong C11 interaction between N$_3^+$H$_3^+$ and OC(3), accompanied by NH$_3^+$/C1/C1/C1$_p$ interactions. They differ however by the nature of the aromatic rings involved in the NH$_3^+$/C1/C1/C1$_p$ interactions and by the neutral interactions, LLLL showing exclusively short-distance interactions like C'5 while LDLD shows long distance structuration like C'7 or C'11. The differences in experimental spectroscopy of protonated LLLL and LDLD is qualitatively reproduced by the calculations, in particular the more structured aspect of the NH bend region and the presence of an intense OH bend band in LLLL.

(2) Sodiated systems. The calculations yield four LLLLNa$^+$ and seven LDLDNa$^+$ structures with Gibbs free energy below 2 kcal mol$^{-1}$. The low-energy calculated structures of LLLLNa$^+$ all display OO$\pi$ structures. Those with an OO$\pi$ pattern are higher in energy by at least 1.6 kcal mol$^{-1}$. Among the OO$\pi$ structures, the LLLLNa$^+$ shown in Fig. 8 gives the best agreement with the experimental vibrational spectrum (Fig. 6) and is the most stable. The sodium interacts with two carbonyl groups, namely CO(1) and CO(3), with distances as short as 2.18 Å. The carboxylic acid is directed outwards as does not interact with the sodium. The cation also interacts strongly with the terminal aromatic rings, as indicated by short distances of 2.9–3.1 Å between Na and the benzene in $a_1$ or $a_3$. In contrast, the two other benzene rings in $a_2$ and $a_4$ are dangling; they interact much less with Na$^+$ (distances of ~5 Å) but adopt a parallel geometry which allows optimising dispersion. In addition, two neutral C'5 interactions are observed from NH(4) to NH(1) and from N(4)H to CO(4). The carboxyl, the N$_3^+$H$_3^+$ as well as the peptidic CO(2) groups are free.

All the LDLDNa$^+$ calculated structures in a 2 kcal mol$^{-1}$ window display an OO$\pi$ structure, with the sodium interacting with the four CO as well as the aromatic ring in $a_4$. The LDLDNa$^+$ structure compatible with both measured collision cross section and vibrational spectrum is shown in Fig. 8, it is the second most stable of all calculated forms. The Na$^+$/O(1)$_p$, Na$^+$/O(2)$_p$, Na$^+$/O(3)$_p$, and Na$^+$/O(4)$_p$ distances amount to 2.28, 2.20, and 2.56, and 2.39 Å, respectively; they are longer than in LLLLNa$^+$. In contrast with LLLLNa$^+$, only one aromatic ring interacts with the cation, with a distance between the phenyl and the cation of 3.08 Å. Among the neutral C$\pi$ interactions observed in LLLLNa$^+$, only that from NH(2) to NH(1) exists in LDLDNa$^+$.}

**Discussion**

We shall first compare the protonated systems studied here to the protonated phenylalanine monomer. The spectroscopy of protonated phenylalanine has been studied both in room temperature and cryogenic ion traps. Two conformers have been evidenced, which show different IR signatures as well as different dynamics in the first electronic excited state. They both display a C5 NH$_3^+$/O hydrogen-bond and an NH$_3^+$/... interaction with the aromatic ring and only differ in the orientation of the amino-acid backbone relative to the aromatic ring. This pattern is retained in both protonated dipeptides. In the protonated tetrapeptides, the aromatic ring in $a_4$ competes with remote aromatic rings for interacting with the charged NH$_3^+$. The resulting trade-off is a bifurcated interaction...
involving Phe(21) and Phe(24) for LLLL_{calc}. In the LDLDH_{calc} structures, the remote aromatic ring involved in this competition is Phe(21). A bifurcated interaction takes place between Phe(21) and Phe(24) for LDLDH_{calc} while NH\_2\_1 interacts only with Phe(21) in LDLDH_{calc2}.

Another interesting aspect is to assess the role of the aromatic ring in the structure of the protonated peptides. The IRMPD spectrum of protonated Ala\_2 has been interpreted thanks to Car-Parrinello molecular dynamics at 300 K.\textsuperscript{45} The results show that protonated Ala\_2 spends most of the time in a trans configuration with \textit{NH} \cdots \cdot O (C5) and \textit{NH} \cdots \cdot \textit{OOC} (C5) interactions, which is similar to that observed in protonated LL. In LD however, no C5 interaction takes place. Ion mobility\textsuperscript{78} as well as vibrational spectroscopy of photochemically-produced protonated Ala\_4\_79 have suggested that it adopts a folded structure in which the protonated N terminal is involved in a C8 interaction possibly accompanied by a C11 interaction with the carboxylic acid. The terapeptides studied here do not show this long-distance structuration but retain the \textit{NH} \cdots \cdot O (C5) pattern observed in protonated dialanine or diphenylalanine, probably because the long-distance structuration is prevented by the bulkiness of the aromatic ring and the \textit{NH} \cdots \cdot \pi interactions. The two protonated tetraphenylalanines show differences in their neutral interactions (C5 for LLLL and C7 or C11 for LDLD). It is to be stressed here that the strong C5 ionic interactions is common to all the studied protonated peptides and that the sensitivity to chirality depends on much weaker neutral interactions.

Sodiated peptides raise the question of the competition between O and N as the interaction sites. For example, the IRMPD spectrum of Ala\_2Na\_+\_2 is compatible with a OO structure, possibly accompanied by an additional interaction with NH\_2.\textsuperscript{49} As in most of the cationised peptide studied, the carboxylic acid interacts with Na\_+ through its carboxyl. The structuring effect of Na\_+ is also evidenced in Ala\_2Na\_+ in which the charge is fully solvated by three carboxyls.\textsuperscript{49} Despite the fact that no experimental spectroscopic data is available for Ala\_2Na\_+, Na\_+ sits on top of the oxygen of three nearly parallel CO in its most stable calculated form, the structure being closed by the hydrogen bonding interaction between the two C and N termini. However, the structure in which Na\_+ interacts with the four CO is not much higher in energy.\textsuperscript{80}

The presence of an aromatic ring adds a third site. Indeed, cationised phenylalanine, PhNa\_+, shows a charge-solvated ON\_p structure.\textsuperscript{81} However, this structure is given up in the diphenylalanines studied here to the benefit of OOn or ONP\_p structures. It should be mentioned that the structure responsible for the blue shoulder of the CO stretch band has been assigned to an ONP\_p structure by Dunbar et al., which contrasts with our assignment to an OOn\_p form. The difference in energy between ONP\_p and OOn\_p structures might depend on the computational method used.

The IRMPD spectra of AlaPheNa\_+ and PheAlaNa\_+ have been explained in terms of the coexistence of OOn and ONP\_p forms involving the C terminus and the amide NH or the amide CO and NH\_2. OOn and ONP\_p are proposed for PhePheNa\_+ and PhePheK\_+.\textsuperscript{50} OON\_p and ONP\_p are observed for divalent ions such as Ni\_2.\textsuperscript{82} The presence of a second aromatic ring seems therefore to exclude interaction with N in natural aminoacids.

In the tetrapeptides as well, the important aspect is the competition between O and \pi for solvating the cation. This competition depends on chirality as OOn\_p is observed in LLLL in contrast with OOn\_p in LDLD, with a wrapped structure. The quantum chemical method should be carefully chosen to be able to reproduce the equilibrium between the physical effects involved in these manifold structures.

### Conclusion

IRMPD and ion mobility spectrometry are complementary methods for structural characterization of di- or tetra-phenylalanines composed of residues of identical or alternating chirality. While IRMPD is sensitive to the hydrogen bond network that stabilizes the structure, IMS gives information of the global shape of the molecule. The only difference between LDH\_+ and LLH\_+ was found to be the presence of a neutral C5 in the former, which has only limited consequences such as slightly different vibrational transition intensities and collision cross sections. This system confirms the previous hypothesis that the interaction responsible for chiral differentiation, here the neutral C5, may be much weaker than the main ionic interaction.\textsuperscript{22} Two isomers of LNNa\_+ are formed, namely OOn\_p and ONP\_p with different spectroscopic signatures. This conclusion differs from that of Dunbar et al. who suggested OOn\_p and ONP\_p conformers.\textsuperscript{39} In contrast, LDNa\_+ exists in one OOn\_p isomer only, with cross section much larger than that of LNNa\_+.

Although the binding pattern changes from OOn\_p in LLLLNa\_+ to OOn\_p in LDLDNa\_+, neither the spectroscopic signature nor the collision cross section is strongly affected by the chirality of the residues in the sodiated tetrapeptides. In contrast, LLLLH\_+ and LDLDH\_+ show different spectroscopic signatures explained in terms different secondary \textit{NH} \cdots \cdot \pi interactions.

The spectra simulated on the basis of harmonic frequencies lead to satisfactory agreement with the experimental values. Quantum chemical calculations have provided low-energy structures with a size compatible with IMS measurements. However, molecular dynamics calculations would bring interesting information on these highly flexible systems.\textsuperscript{83} Those resting on a well-adapted polarizable force field are especially promising for molecules of this size.\textsuperscript{84}

The structural differences induced by chirality should manifest itself in the aggregation properties. Conformer-specific studies at low temperature of these systems would bring interesting information, especially in the UV range. Specific photofragments as a function of the residues chirality could be searched for.\textsuperscript{77,85,86}

### Acknowledgements

The mass spectrometry platform SMAS of the Laboratoire de Chimie Physique (University Paris-Sud) is gratefully acknowledged. We thank Dr J. M. Ortega and the CLIO team for technical assistance. We acknowledge V. Steinmetz for his assistance in the experimental work. We acknowledge the use of the computing facility cluster GMPCS of the LUMAT federation (FR LUMAT 2764) and of the DI Univ. Paris-Sud.
References
