DOI: 10.1002/anie.201100702

Polypeptide Structures

Cold-Ion Spectroscopy Reveals the Intrinsic Structure of a Decapeptide**

*Natalia S. Nagornova, Matteo Guglielmi, Manuel Doemer, Ivano Tavernelli, Ursula Rothlisberger, Thomas R. Rizzo, and Oleg V. Boyarkin**

In memory of Irina Boyarkine

The three-dimensional (3D) structures of proteins and peptides in vivo largely determine their biological function. In vitro these native structures and their heterogeneity reflect a fine balance between noncovalent intramolecular interactions and those with the surrounding solvent molecules. Decoupling intra- and intermolecular interactions and revealing the intrinsic structures of biomolecules is crucial for understanding protein–peptide (–protein, –membrane) binding processes and protein folding, and can assist in silico drug design. Here we demonstrate the use of conformer-selective, cold-ion infrared spectroscopy and experimentally constrained calculations to solve the 3D structure of a natural antibiotic, gramicidin S (GS), isolated in the gas phase. It is the largest molecule for which the gas-phase structure has been accurately determined.

This benchmark decapeptide (cyclo-VOLFPVOLFP, where "O" designates ornithine and Phe is the D rather than the L enantiomer) has been studied in the condensed phase for decades owing to its practical importance.[1–8] GS exhibits strong antimicrobial activity, which is based on its binding to microbial membranes,[5,6] but it is toxic to human red blood cells. Rational design of GS analogues with improved pharmacological activity requires a better understanding of the GS structure and its interactions with solvent molecules and phospholipids of the cell membranes. The structure of the isolated peptide may serve as an additional starting point to model these interactions and help elucidate the mechanism of its antimicrobial activity.

While isolation of solvent-free biomolecules in the gas phase removes the intermolecular interactions, the decreased concentration of gas-phase samples requires sensitive structure-selective techniques. Ion-mobility techniques can separate different conformational isomers by their collisional cross-section,[9–11] but their accuracy in solving structures is limited by the low number of experimentally derived structural constraints and should be verified by complementary techniques. In recent years precise structures of several amino acids and small peptides in the gas phase have been determined using infrared spectroscopy.[12–18] This approach relies on measuring a "fingerprint" of vibrational transitions (frequencies, intensities, and linewidths) that serves as a benchmark for structural calculations. Unambiguous identification of calculated structures for a large molecule challenges experiments to provide a detailed fingerprint for each observed conformer, since this is exactly what theory calculates. This requires achieving vibrational resolution and conformational selectivity in the IR spectra, which becomes problematic for large species at room temperature. Theory typically employs classical molecular dynamics simulations to sample a large conformational space to identify candidate structures. Subsequently, a few of the lowest energy structures are optimized by ab initio theory to find the most stable conformer. The biggest challenge in these calculations is in narrowing the conformational search among the thousands of structures identified by molecular dynamics prior to optimizing their structures at higher levels of theory and calculating their spectra.

Our experiment combines electrospray ionization mass spectrometry, cryogenic cooling, and laser spectroscopy (see the Supporting Information for the details).[19] Cooling sample molecules to sufficiently low temperatures (≈ 10 K) allows vibrational resolution in the UV and IR spectra of GS.[20] High resolution in the UV spectrum enables the use of IR/UV double-resonance detection[12,13,21–23] for conformer-selective measurements of IR spectra. We recently demonstrated application of this approach for spectroscopy of GS in the 6 μm region.[20] Herein we extend it over a significant spectral range covering all the light-atom stretching vibrations, and we use some special techniques, such as ^{15}N isotopic substitution and complexation of the peptide with a crown ether to help assign the vibrational bands. Structural constraints derived from both our spectroscopic and mass spectrometric data guide the conformational search to find the most stable calculated structures of isolated, doubly protonated GS. By comparing the unique fingerprint provided by our highly resolved, conformation-specific infrared spectrum with the theoretically derived vibrational spectrum

[*] N. S. Nagornova, Prof. T. R. Rizzo, Dr. O. V. Boyarkin
Laboratoire de Chimie Physique Moléculaire
École Polytechnique Fédérale de Lausanne
1015 Lausanne (Switzerland)
Fax: (+41) 21-693-5170
E-mail: oleg.boiarkin@epfl.ch
Homepage: http://lcpm.epfl.ch/
Dr. M. Guglielmi, M. Doemer, Dr. I. Tavernelli, Prof. U. Rothlisberger
Laboratoire de Chimie et Biochimie Computationelle
École Polytechnique Fédérale de Lausanne
1015 Lausanne (Switzerland)

[**] We thank EPFL and FNS (grant no. 200020-120065) for their generous support of this work and Valenta Pharm Co. (Moscow) for providing us with samples of GS. O.B. thanks Irina Boyarkine for a lifetime of support of his research.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201100702.

we can assign one of these candidate structures to the predominant conformer of GS that we produce in the gas phase.

In the gas phase at low temperature $[GS+2H]^{2+}$ adopts three different conformations, one of which is significantly more abundant than the other two.[20] Figure 1a shows an infrared spectrum of this main conformer cooled to approximately 12 K measured by photofragment-detected IR/UV double resonance. The spectrum, which covers the NH, CH, and C=O stretching and NH bending bands provide a set of nearly 30 spectroscopic reference frequencies for selecting a 3D structure of doubly protonated GS from the calculated possibilities. In several regions of the spectrum the resolved, closely spaced peaks impose stringent requirements on the accuracy of calculated vibrational frequencies. Simply matching the calculated and observed frequencies is necessary but not sufficient for identifying the proper structure, however. The assignment of the peaks to specific vibrational modes provides the true link between experiment and theory.

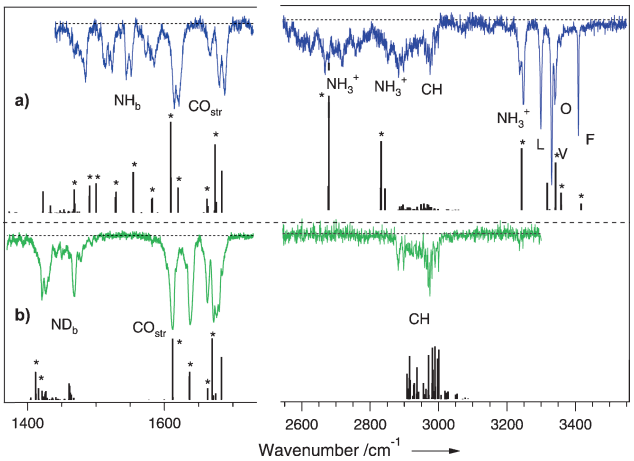

Figure 1. Infrared spectra of the most abundant conformer of a) $[GS+2H]^{2+}$ and b) its deuterated analogue (N–H→N–D) measured by IR/UV double-resonance photofragment spectroscopy, together with the corresponding calculated vibrational spectra for the most stable calculated structure of these species. The calculated frequencies are scaled by a factor of 0.961 in (a) and by a factor of 0.941 in (b). In (a) the frequencies of NH/CH stretching vibrations are additionally shifted by the term $\Delta\nu_i = -20\sqrt{\delta\nu_i}$, where $\delta\nu_i$ is the width of the *i*-th peak. Asterisks label the most intense (nearly) doubly degenerate calculated transitions.

We use several different methods to assign the vibrational bands in Figure 1a. Isotopic labeling of the two Val and two Leu residues by ^{15}N should shift the amide NH stretching vibrations to lower frequencies by approximately 8 cm^{-1} (in a harmonic oscillator approximation). In the IR spectrum of Figure 2b, we indeed observe a −8.5 cm^{-1} shift of two peaks in the isotopically substituted molecule, allowing us to unambiguously assign these peaks to four NH stretching vibrations of these residues. The fact that the isotopic substitution of four residues results in the shift of only two peaks implies that the NH stretches in each pair of identical residues have nearly degenerate frequencies, suggesting symmetrically equivalent positions of the identical residues. This observation supports our earlier suggestion that the $[GS+2H]^{2+}$ structure should be highly symmetric (C_2).[20] Replacement of all amide hydrogens by deuterium atoms also helps in the assignment of vibrational bands in that it shifts all ND stretching vibrations below roughly 2500 cm^{-1}, leaving only CH stretches in the 3 μm region (Figure 1b). It also shifts ND bending vibrations to lower wavenumbers, allowing us to distinguish amide bending bands from the C=O stretching bands.

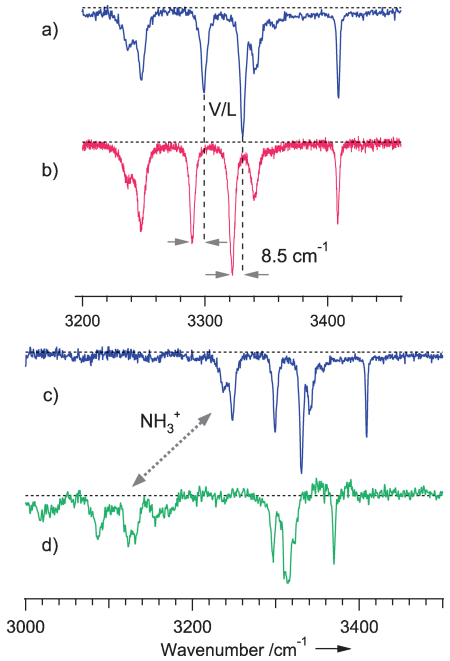

Figure 2. Portions of the infrared spectra of the most abundant conformer of a,c) $[GS+2H]^{2+}$, b) its isotopologues ($^{15}N\leftarrow^{14}N$ in Leu and Val), and d) $[GS+2H]^{2+}$/[18]crown-6 complex.

We assign the two peaks around 3240 cm^{-1} to the NH_3^+ stretching vibrations based on the general expectation that the frequencies of such charged groups shift strongly to lower energy relative to amide NH stretching bands because of stronger hydrogen bonding. We verified this assignment by complexing $[GS+2H]^{2+}$ with two crown ether molecules ([18]crown-6), which form particularly strong hydrogen bonds with the ammonium groups.[24,25] As shown in Figure 2c and d, the complexation leads to an additional shift of these two peaks by 120–150 cm^{-1}.

The UV-induced photofragment mass spectrum of $[GS+2H]^{2+}$ (Figure S1 in the Supporting Information) provides additional information that directly constrains our structural search prior to calculation of the infrared spectrum. The two most abundant fragments that result from photoexcitation of the Phe chromophores correspond to the loss of neutral -CH_2NH_2 and -$CH_2CH_2NH_2$ from the ornithine side chains, although these channels are only negligible ones in collisional-induced dissociation[7] and in infrared multiphoton dissociation (IRMPD; Figure S1 in the Supporting Information). The observed nonstatistical dissociation[26] suggests an initial transfer of photoexcitation energy directly from the

Phe chromophores to the amines of the Orn side chain and implies a certain proximity and coupling of the two groups. This conclusion, along with the symmetry inferred from the vibrational spectra, drastically narrows the initial conformational search for suitable candidate structures.

This search employs extended molecular dynamics simulations with a minimal set of the above-mentioned experimentally determined constraints imposed as structural restraints to guide the exploration of configurational space. An initial pool of candidate structures was generated in this way through multiple simulated annealing runs in which the system was heated to high temperature (1500 K) to accelerate phase-space sampling and then slowly cooled down. From these confined conformational searches the four lowest-energy structures were selected and freely optimized using density functional theory as a starting point for the calculation of their harmonic vibrational frequencies (see the Supporting Information for the details of the calculations). The frequencies of the most stable structure (Table S2 in the Supporting Information), after scaling to account for vibrational anharmonicity, match well with the measured IR spectrum of the most abundant conformer (Figure S2 in the Supporting Information). The assignments of all the computed vibrational bands are in full agreement with our experimentally determined assignments. Only the frequencies of the ammonium NH stretching bands are not well predicted by theory. We do not expect a perfect reproduction of these strongly coupled bands, because their anharmonicities should be greater than those for weakly coupled NH/CH stretching modes. A refinement of the two scaling coefficients that is rooted in the physics of intramolecular vibrational coupling (see details in the Supporting Information) results in a better match between the calculated and measured frequencies.

A stringent test for the computed lowest-energy structure is to calculate the vibrational spectrum of the deuterated peptide. Deuteration does not change the structure but shifts the NH stretching bands to lower frequencies. The predicted spectrum of the deuterated species matches well with the experimental data (Figure 1b), reinforcing our confidence that the calculated structure is the correct one.

Figure 3 compares the structure of isolated $[GS+2H]^{2+}$ determined in this work with that of the crystallized, hydrated species measured by X-ray diffraction.[8] The nearly symmetrical (C_2) structure of the isolated peptide appears 40% less elongated and more compact. It exhibits a characteristic parallel alignment of the two Phe rings, each of which is in close proximity to an ammonium group of an Orn side chain. This difference largely results from solvation of the charged Orn side chains in the crystal that prevents their participation in cation–π hydrogen bonds with the Phe rings. In the isolated structure, the ammonium groups also form hydrogen bonds with the carbonyl oxygens of the Phe and Orn residues, which anchor them to the peptide backbone. Table S3 in the Supporting Information provides atomic coordinates of the calculated structure.

This work demonstrates that cold-ion spectroscopy, together with high-level theory, can be used to solve conformer-selective structures of isolated midsize peptides. Although isolated structures may not reflect the structures in vivo, in certain cases they should be helpful for understanding in vivo interactions. For instance, the structure of gramicidin S interacting with a membrane can differ from the structure determined in vitro by NMR or X-ray methods, making the intrinsic structure determined here a valuable starting point for modeling the biological activity of this antibiotic in vivo. To our knowledge gramicidin S is the largest molecule for which the accurate intrinsic structure has ever been determined.

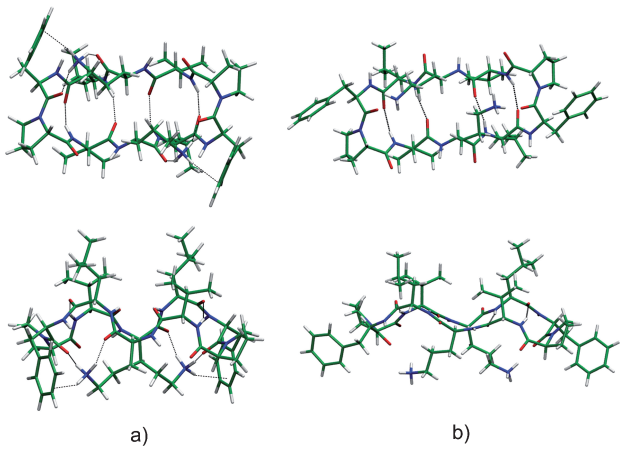

Figure 3. Two 3D views of the $[GS+2H]^{2+}$ structures a) determined in this work by cold-ion spectroscopy for the lowest-energy conformer of the isolated species, and b) solved by X-ray diffraction for crystallized species (reconstructed from the data of Ref. [8]).

Received: January 27, 2011
Revised: March 7, 2011
Published online: May 6, 2011

Keywords: cold-ion spectroscopy · gramicidin S · molecular modeling · peptides · structure elucidation

[1] G. F. Gause, M. G. Brazhnikova, *Nature* **1944**, *154*, 703.
[2] D. T. Warner, *Nature* **1961**, *190*, 120.
[3] S. E. Hull, R. Karlsson, P. Main, M. M. Woolfson, E. J. Dodson, *Nature* **1978**, *275*, 206.
[4] Y. Xu, I. Sugár, N. Krishna, *J. Biomol. NMR* **1995**, *5*, 37.
[5] G. N. Tishchenko, V. I. Andrianov, B. K. Vainstein, M. M. Woolfson, E. Dodson, *Acta Crystallogr. Sect. D* **1997**, *53*, 151.
[6] E. J. Prenner, R. N. A. H. Lewis, R. N. McElhaney, *Biochim. Biophys. Acta Biomembr.* **1999**, *1462*, 201.
[7] E. Pittenauer, M. Zehl, O. Belgacem, E. Raptakis, R. Mistrik and G. Allmaier, *J. Mass Spectrom.* **2006**, *41*, 421.
[8] A. L. Llamas-Saiz, G. M. Grotenbreg, M. Overhand, M. J. van Raaij, *Acta Crystallogr. Sect. D* **2007**, *63*, 401.
[9] G. von Helden, T. Wyttenbach, M. T. Bowers, *Science* **1995**, *267*, 1483.
[10] K. B. Shelimov, D. E. Clemmer, R. R. Hudgins, M. F. Jarrold, *J. Am. Chem. Soc.* **1997**, *119*, 2240.
[11] R. R. Hudgins, M. F. Jarrold, *J. Am. Chem. Soc.* **1999**, *121*, 3494.
[12] J. R. Carney, T. S. Zwier, *J. Phys. Chem. A* **2000**, *104*, 8677.
[13] L. C. Snoek, E. G. Robertson, R. T. Kroemer, J. P. Simons, *Chem. Phys. Lett.* **2000**, *321*, 49.
[14] C. Unterberg, A. Jansen, M. Gerhards, *J. Chem. Phys.* **2000**, *113*, 7945.

[15] W. Chin, F. Piuzzi, J. P. Dognon, I. Dimicoli, B. Tardivel, M. Mons, *J. Am. Chem. Soc.* **2005**, *127*, 11900.
[16] M. S. de Vries, P. Hobza, *Annu. Rev. Phys. Chem.* **2007**, *58*, 585.
[17] J. A. Stearns, O. V. Boyarkin, T. R. Rizzo, *J. Am. Chem. Soc.* **2007**, *129*, 13820.
[18] T. R. Rizzo, J. A. Stearns, O. V. Boyarkin, *Int. Rev. Phys. Chem.* **2009**, *28*, 481.
[19] O. V. Boyarkin, S. R. Mercier, A. Kamariotis, T. R. Rizzo, *J. Am. Chem. Soc.* **2006**, *128*, 2816.
[20] N. S. Nagornova, T. R. Rizzo, O. V. Boyarkin, *J. Am. Chem. Soc.* **2010**, *132*, 4040.
[21] R. H. Page, Y. R. Shen, Y. T. Lee, *J. Chem. Phys.* **1988**, *88*, 5362.
[22] C. J. Gruenloh, J. R. Carney, F. C. Hagemeister, C. A. Arrington, T. S. Zwier, S. Y. Fredericks, J. T. Wood, K. D. Jordan, *J. Chem. Phys.* **1998**, *109*, 6601.
[23] C. Plützer, E. Nir, M. S. de Vries, K. Kleinermanns, *Phys. Chem. Chem. Phys.* **2001**, *3*, 5466.
[24] J. S. Brodbelt, *Int. J. Mass Spectrom.* **2000**, *200*, 57.
[25] R. R. Julian, J. L. Beauchamp, *Int. J. Mass Spectrom.* **2001**, *210/211*, 613.
[26] M. Guidi, U. J. Lorenz, G. Papadopoulos, O. V. Boyarkin, T. R. Rizzo, *J. Phys. Chem. A* **2009**, *113*, 797.

