THE JOURNAL OF PHYSICAL CHEMISTRY B

ARTICLE

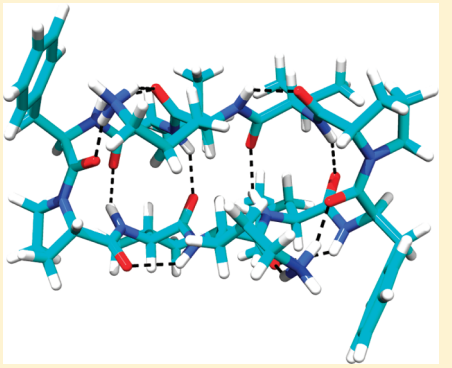
pubs.acs.org/JPCB

Structures and IR Spectra of the Gramicidin S Peptide: Pushing the Quest for Low-Energy Conformations

Kaustubh Joshi, David Semrouni, Gilles Ohanessian,* and Carine Clavaguéra*

Laboratoire des Mécanismes Réactionnels, Department of Chemistry, Ecole Polytechnique, CNRS, 91128 Palaiseau Cedex, France

ABSTRACT: An extensive molecular modeling study was carried out on the doubly protonated cyclic decapeptide Gramicidin S following several recent gas-phase experiments. Our computational strategy includes replica-exchange molecular dynamics simulations with the new generation force field AMOEBA for exploration and density functional calculations using several functionals for refinement of structures and computation of IR spectra. This procedure yields low-energy structures of which three are proposed to correspond to the three conformers detected in low-temperature IR experiments. The most stable structure has C_2 symmetry and four strong β-sheet interactions between Orn and Val residues. Furthermore, all the other peptidic N−H bonds are involved in seven-membered C_7 motifs. The computed IR spectra of the three conformers are in good agreement with the experimental ones in the 1400−2000 cm^{-1} range. In the 3000−3600 cm^{-1} region, the computed spectrum is also in good agreement with experiment for the main conformer, and predictions are made of structure-specific signatures for the other two conformers. The accuracy of several density functionals is discussed in detail. These results point out that efficient potential energy surface explorations coupled to appropriate density functional theory (DFT) calculations are able to reveal the structures of molecules as large and flexible as decapeptides.



1. INTRODUCTION

The peptide antibiotic Gramicidin S (GS) was discovered in the 1940s.[1] It is a cyclodecapeptide with the (Pro-Val-Orn-Leu-D-Phe)$_2$ sequence using two singular amino acids, that is, Orn and D-Phe.[2] Its crystal structure consists of an antiparallel β-sheet stabilized by four intramolecular hydrogen bonds between leucine and valine residues. Because of its cyclic constraints, GS is well documented as a model for a β-sheet motif, and furthermore, the two ornithines with basic side chains may play an important role in governing its structure.[3] GS has been the subject of numerous experimental studies using a variety of techniques in condensed phase as a benchmark for a β-sheet immersed in various crystalline and liquid media.[4−6] Unraveling the factors which determine the structures of GS in the condensed phase may be greatly facilitated by a detailed knowledge of its low-energy conformers in the absence of surrounding medium. To this end, several experimental and computational studies of gaseous, doubly protonated GS have recently been carried out.[7−9] Two different techniques have emerged in the past decade as able to provide vibrational signatures of small biological molecules in the gas phase. Infrared multiple photon dissociation (IRMPD) spectroscopy is based on the fragmentation of molecular ions when irradiated with photons of variable wavelengths. It is most often performed at room temperature. Despite the complexity of GS, this technique has been successfully used recently by measuring the mid-IR spectrum (1250−1750 cm^{-1} range) at 300 K in the gas phase.[7] The results highlighted dominant interactions of the charged side chains with the backbone that govern the lower frequency part of the amide I peak. This study provided a general agreement between the calculated and the experimental spectra; lack of a perfect agreement was partly attributed to the possible presence of several conformers. The second technique that has recently been used is IR-UV double resonance spectroscopy. Compared to IRMPD, it has the double advantage that it is carried out on very cold ions leading to significantly higher resolution and that exploiting resolved UV transitions allows monitoring infrared spectra for conformers or isomers one at a time while IRMPD spectra are those of a mixture if any. Nagornova et al. obtained such spectra for cooled GS in the 1400−1750 cm^{-1} range.[8] Transitions recorded in the electronic spectrum were assigned to three different conformers A, B, and C. The IR depletion spectra distinctly confirmed the presence of three different conformers suggesting that conformer B is similar to conformer A with a clear differentiation for conformer C.

On the basis of the high-resolution spectrum of the most abundant conformer in the fingerprint region, it was stressed[8] that none of the structures obtained by Kupser et al.[7] or by Nagornova et al.[8] led to a computed spectrum which was compatible with experiment thus setting up a challenge to theoretical chemists.

We take up this challenge of obtaining reliable structures and spectra for GS by combining several approaches. First, a realistic exploration of the potential energy surface is based on replica-exchange

Received: July 25, 2011
Revised: November 14, 2011
Published: November 16, 2011

molecular dynamics (REMD) coupled to the AMOEBA force field which includes polarization effects and multipolar distributions for electrostatics. Second, the structures of low-energy conformations thus obtained are refined, and infrared spectra are computed at the quantum chemical level. Because of the size of Gramicidin S, only density functional theory (DFT) is tractable, and we use M06 and B3LYP-D functionals so as to include both Hartree−Fock exchange and dispersion effects. The spectra are compared with recent experimental data[8] in the 1400−2000 cm^{-1} region and were initially meant to predict spectra in the N−H stretch mode range. While this work was near completion, an additional study of GS appeared[9] describing both well-resolved IR spectrum in the 2600−3500 cm^{-1} range and a computed spectrum which provided an excellent match between experimental and computed vibrational bands. Close comparison with the present results indicated both structures to be very similar. We describe below our own strategy, we compare our results to those previously published, we provide a discussion of the current status of DFT in this regard, and we describe structures and spectra which match experimental results for the minor conformers.

2. COMPUTATIONAL STRATEGY

Exploration of the potential energy surface of a complex molecule such as GS cannot be carried out at the quantum chemical level. While any force field may be used as a structure generator, reliable energetics is important to sort out plausible structures to use at more accurate levels. In the case of gas-phase amino acids, force fields whose electrostatic energy component is only based on point charges have been shown to reproduce only half of a set of conformations as established by accurate quantum chemical calculations, while the new generation force field AMOEBA reaches 80% or more of reproduction thanks to high-level electrostatics and inclusion of polarization.[10−12] In the GS peptide studied herein, the presence of the charged Orn side chain and of strong β-sheet interactions suggests a special importance of electrostatics and polarization in both structures and energetics. For this purpose, we have used the AMOEBA force field developed by Ponder et al.[13]

The potential energy surface of doubly protonated GS was explored using the coupling of REMD and AMOEBA that was found to be successful in a previous study of the IR spectrum of $G_8−Na^+$.[11] The REMD simulations have been performed with 16 replicas allocated according to a geometric progression of temperatures in the range 200−600 K. Each trajectory was propagated with a time step of 1 fs. Exchanges of random pairs of adjacent configurations were attempted every 10 ps, and the configurations from all replicas were saved every 5 ps for further local optimization with the force field. The total simulation time for each replica was 5 ns, and the REMD simulations were repeated four times with all replicas starting each time from the most stable structure from the preceding REMD simulation. More than 30 000 structures were generated in this manner. Looking at the instantaneous index of the replica initially set to the lowest temperature (200 K), we find a large number of successful exchanges with other replicas up to the highest temperature of 600 K. This ensures that our temperature schedule allowed for an efficient communication between replicas, and hence, a much improved sampling with respect to conventional MD. rmsd analysis was carried out to identify distinct structures out of the large set obtained from REMD simulations. In a first step, a clustering analysis was used to distinguish the structures on the basis of their C_α atoms superposition with a 0.1 Å rmsd threshold and to create families. These families were further refined to sort the unique structures considering all nonhydrogen atoms using an rmsd analysis with a threshold of 0.1 Å. These two steps were first applied to the subset of lowest energy structures in a 1 kcal·mol^{-1} window. To generate several unique structures that remain tractable for full geometry reoptimization at the quantum mechanical level, this procedure was repeated twice to the next 1 kcal·mol^{-1} energy windows yielding a total analysis within the lowest 3 kcal·mol^{-1} window. The lowest potential energy configurations obtained in each family localized by this two-step clustering procedure were further subjected to electronic structure calculations at the DFT level to determine the best candidates for comparison to experiment.

The description of noncovalent interactions in large-size peptides in terms of structures, of identification of low-energy conformations, and of computation of infrared spectra remains a challenge to theoretical chemistry. Indeed, standard density functional theory (DFT) fails to describe dispersion interactions[14] which are expected to be important in densely packed molecules.[15] In recent years, two main approaches have been proposed to overcome this problem. The M06 series of functionals developed by Zhao and Truhlar has been parametrized to take into account various physical effects such as noncovalent interactions.[16] They have been shown to provide a significant improvement for the general description of chemical properties by DFT.[15] In parallel, the DFT-D approach has been developed by adding an empirical dispersion term to previously existing functionals with functional-specific parameters.[17] It has been proved to be an efficient way to describe dispersion effects in biomolecules.[18,19] Both types of functionals have been used in the present work, namely, M06 and B3LYP-D associated with 6-31G(d,p) and def2-SVP basis set, respectively. We have found previously that MP2 calculations can produce IR spectra in very good agreement with experimental data for gas-phase peptides;[11] however, this level of calculation is intractable because of the size of GS.

Recent functionals are not all able to reproduce the secondary structures and infrared spectra of gas-phase peptides.[20] In a first step, the protonated dipeptide H^+TyrAla was used to benchmark the performance of M06 and B3LYP-D functionals to evaluate the stretching frequencies of both peptidic and ammonium N−H bonds. Calculations were performed for the lowest energy structure, named conformer A in ref 21. The optimized structures are very similar at both levels with an rmsd smaller than 0.05 Å. One ammonium N−H is free, and the peptidic N−H is involved in a C_5 arrangement with the amide carbonyl leading to two experimental peaks at 3350 and 3370 cm^{-1}, respectively. The other two ammonium N−H bonds are involved in a π interaction with the aromatic ring of Tyr and in a C_5 interaction with the peptidic C═O with an experimental broad peak in the 3100−3200 cm^{-1} range and a maximum at 3180 cm^{-1}. With M06 vibrational frequencies scaled by 0.95, the three N−H frequencies are at 3359, 3356, and 3162 cm^{-1}, which are in good agreement with experiments.[21] B3LYP-D results give similar agreement, however, only when using two different scaling factors for peptidic and ammonium N−H bonds. In a second step, the performance of B97-D, B3LYP, B3LYP-D, and M06 functionals was evaluated on conformer A of Gramicidin S. From a common starting geometry, the C_2-symmetric structure with all

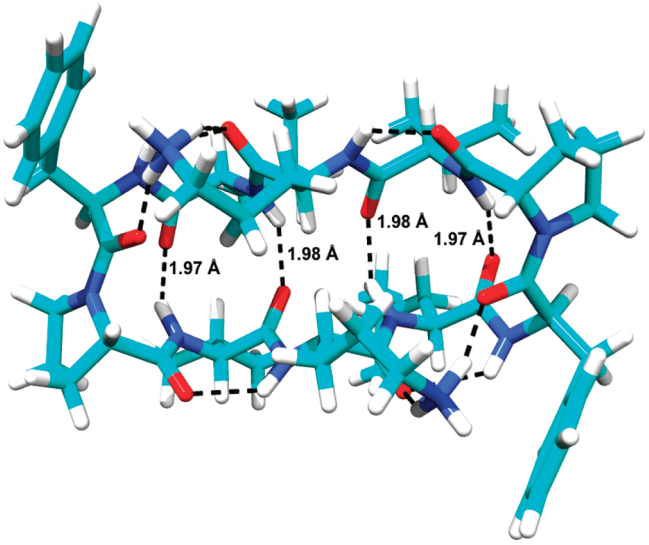


Figure 1. C_2-symmetric structure proposed for conformer A.

peptidic N−H bonds being hydrogen bound was only located using B3LYP-D and M06 leading to the conclusion that a treatment of Hartree−Fock exchange and dispersion interactions is essential. Consequently, geometry optimizations and infrared spectra calculations of other GS conformers were performed at the DFT level using the M06 functional and the 6-31G(d,p) basis set. Harmonic vibrational frequencies were scaled by a factor of 0.95 above 3000 cm^{-1} and 0.94 below 2000 cm^{-1} to take account of the potential energy surface anharmonicities. The M06 functional was preferred to B3LYP-D to avoid the use of two different scaling factors for peptidic and ammonium N−H stretch modes. Inclusion of anharmonic corrections in the vibrational frequency calculations using second-order perturbation theory or variational self-consistent field (SCF) methods has been recently applied to small biological molecules;[22−24] however, these methodologies are impractical for a decapeptide such as GS. The calculations described in this study were performed using the Tinker,[25] Turbomole v6.1,[26] and Gaussian 09[27] packages.

3. RESULTS AND DISCUSSION

3.1. Structures and Energetics. The C_2-symmetric structure of GS shown in Figure 1 is the lowest energy structure found in our study at the M06 level. It involves a large number of hydrogen-bonded interactions which are detailed in Table 1. All the peptidic N−H bonds are involved in hydrogen-bond interactions: the Leu and Val N−H bonds build the β-sheet structure in strong interaction with Val and Leu C═Os, respectively (1.97 and 1.98 Å); furthermore, both Phe N−H bonds interact with Orn C═Os in seven-membered cyclic motifs (2.39 Å), and both Orn N−H bonds are involved in seven-membered interactions with Pro C═Os (2.09 Å). As discussed in the next section, this conformer is likely to give rise to the experimental spectrum labeled as A.[8] A set of other conformers lies in an energy range of 7 kJ·mol^{-1}, two of which can be associated with the experimental spectra B and C (Figure 2).[8] Their hydrogen-bonding patterns are detailed in Table 1 highlighting similarities and differences. The structure proposed for conformer B is slightly asymmetric strongly resembling that of conformer A and is only 3 kJ·mol^{-1} higher in energy. One of the β-sheet interactions is stretched by 0.25−2.138 Å relative to A. The other β-sheet and ammonium N−H···O═C distances are all within 0.03 Å of their values in A. There are probably several such structures surrounding the symmetric conformer on the potential energy surface. The structure proposed for conformer C is 7 kJ·mol^{-1} higher in energy than conformer A. Only three β-sheet interactions are present instead of four for conformers A and B (see Figures 1 and 2). The other differences come from the absence of two seven-membered ring interactions, one (Orn) N−H···O═C (Pro) and one (Leu) N−H···O═C (Val). The stability of the structure is ensured by stronger hydrogen bonds between the other Leu N−H and Val C═O.

Table 1. Distances (in Å) for the Hydrogen-Bonded Interactions in Conformers A, B, and C at the M06 Level

interaction	A	B	C
Leu:C═O···H−N:Val	1.965	1.936	2.021
Leu:N−H···O═C:Val	1.983	2.138	1.934
Leu:N−H···O═C:Val	1.983	1.975	2.711
Leu:C═O···H−N:Val	1.965	1.992	2.071
NH_3^+···O═C Phe	1.723	1.751	1.720
NH_3^+···O═C Orn	1.779	1.760	1.770
NH_3^+···O═C Phe	1.723	1.723	1.724
NH_3^+···O═C Orn	1.779	1.764	1.808
C_7 Phe:N−H···O═C:Orn	2.391	2.588	2.431
C_7 Orn:N−H···O═C:Pro	2.088	2.216	3.917
C_7 Phe:N−H···O═C:Orn	2.391	2.315	2.725
C_7 Orn:N−H···O═C:Pro	2.088	2.074	2.135
Leu:N−H···O═C:Val(2)	2.479	2.576	2.115
Leu:N−H···O═C:Val(2)	2.479	2.238	3.980

The structure obtained recently by Nagornova et al.[9] was reoptimized at the M06/6-31G(d,p) level, starting from their coordinates, and was found to be 8.3 kJ/mol higher in energy although structurally very similar to our C_2 lowest-energy conformer. Slight variations in geometries can be discerned leading to the conclusion that there exists a potential energy basin encompassing a manifold of conformations with very small structural and energetic differences. The importance of the choice of density functional will be discussed in the next section. Geometry optimizations were also performed using the two lowest-energy structures found by Kupser et al.[7] and labeled A and C. At the M06/6-31G(d,p) level, these structures are higher in energy than our conformer A by 13.4 and 15.7 kJ/mol, respectively. Their structure A is also a C_2-symmetric structure with four β-sheet interactions. However, it is different than our conformer A because C_7 hydrogen bonds between Phe and Orn and between Orn and Pro are missing leading to a significantly higher energy. A similar structure was obtained by Nagornova et al.[8] (described in the supporting information of ref 8) with four hydrogen bonds of β-sheet type, however, with shorter interactions of ammoniums with carbonyl groups and thus with longer C_7 hydrogen bonds between peptidic N−H and C═O groups. As stated by these authors, its computed spectrum does not provide a satisfactory match to the experimental spectrum. In addition, structure C of Kupser et al.[7] has three β-sheet interactions as in our conformer C; however, they are quite different in the orientations of the side chains of Orn and Phe with the Orn NH_3^+ remote from the Phe rings interacting more strongly with Val, Orn, and Phe C═Os.

dx.doi.org/10.1021/jp207102v |J. Phys. Chem. B 2012, 116, 483–490

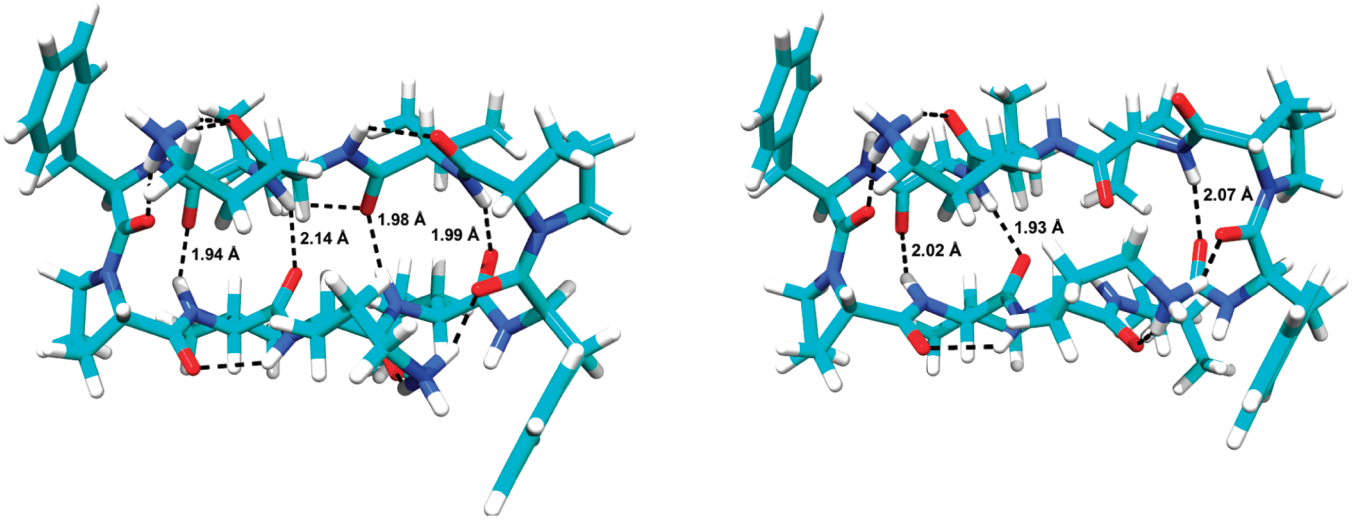


Figure 2. Structures proposed for conformers B (left side) and C (right side).

We find that the three most stable conformers identified in our study have computed spectra which bear strong enough similarities with the A, B, and C UV/IR spectra that the assignment of a structure can be safely made to each spectrum (see below). It turns out that our energy ordering matches the intensity ordering in the UV spectrum. This may be partly coincidental because differences in absorption cross sections may obscure the link between intensity and populations. Still, the IR spectra of A and B are so similar that structures and absorption cross sections are also expected to be similar. Thus, it is likely that A is more abundant than B in the experimental conditions used, which is in good agreement with our energy ordering.

3.2. The C_2-Symmetric Conformer A. *Spectral Assignments.* The superposition of computed and experimental spectra of conformer A is provided in Figure 3 together with spectra assignments. As indicated in Figure 3, there are three broad spectral ranges: 1610−1700 cm^{-1} includes mainly the amide I bands, ammonium-dominated bands appear in the 1520−1660 cm^{-1} region, while amide II bands lie in the 1450−1500 cm^{-1} range. There is significant overlap between the amide I and the ammonium groups of bands including some strong mode coupling between ammonium scissors and the stretches of those of the C═Os with which they interact directly (those of Phe and Orn). The three peaks at 1689, 1683, and 1682 cm^{-1} are attributed to the coupled stretchings of Val, Leu, and Pro C═O corresponding to the experimental bands with maxima at 1689 and 1681 cm^{-1}. Small peaks are due to coupled C═O stretchings of Val and Leu (1681 cm^{-1}) and to pure C═O stretching of Pro (1670 cm^{-1}). The 1669 cm^{-1} band has contributions from C═O stretch modes in Pro and Val and may be correlated to the experimental band around 1670 cm^{-1}. Peaks at 1653 cm^{-1} are attributed to Orn C═O stretches coupled with NH_3^+ scissors. The 1619 cm^{-1} peaks correspond to Phe C═O stretches coupled with NH_3^+ scissor in good agreement with the experimental spectrum.[8] It matches well with the most intense experimental feature although the latter appears as a doublet. Peaks at 1573−1574 cm^{-1} and 1519−1520 cm^{-1} are all attributed to NH_3^+ scissorings. They are in good agreement with two of the three experimental features in this region; however, there is no match to the experimental doublet at 1540−1550 cm^{-1}.

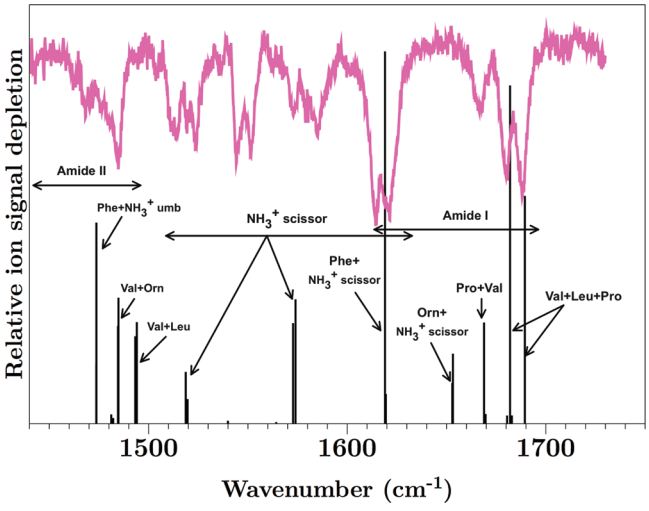


Figure 3. Experimental (pink lines, ref 8) and computed M06 (black lines) IR spectra for conformer A in the 1450−1750 cm^{-1} region.

A possible explanation for the small mismatches between experimental and computational results is that our structure may underestimate the Orn C═O···ammonium interactions. Would this interaction be slightly stronger, the amide I band computed at 1653 cm^{-1} would be red-shifted and would become the second component of the experimental doublet at 1610−1630 cm^{-1}, while the ammonium scissoring computed at 1520 cm^{-1} would be blue-shifted and would become a component of the experimental doublet at 1530−1540 cm^{-1}.

In the amide II region, the peaks at 1494 and 1493 cm^{-1} are attributed to the coupling of Val and Leu modes while those at 1485 cm^{-1} correspond to Val and Orn coupling. The Phe amide II is predicted at 1474 cm^{-1} with a coupling with the umbrella mode of the NH_3^+. The amide II group of peaks appears to stand slightly to the blue compared to experiment. However, the general good agreement between experimental and computed spectra lends confidence to the overall structural identification of this conformer.

In Table 2, the experimental values from ref 9 and the scaled harmonic frequencies are listed in the 2900−3600 cm^{-1} range, a

dx.doi.org/10.1021/jp207102v |J. Phys. Chem. B 2012, 116, 483–490

Table 2. Experimental Frequencies (in cm^{-1}) from Ref 9 and Mode Assignments and Computed M06 Vibrational Frequencies (in cm^{-1}) for A, B, and C in the N−H Stretch Region

assignments	A exptl[9]	A	B	C
N−H Phe	3409	3395	3420; 3401	3414; 3400
N−H Val	3331	3372	3368 (Val + Orn asym); 3365 (Val + Orn sym)	3350; 3347
N−H Orn	3340	3332	3354 (Val); 3328 (Orn)	3438; 3341
N−H Leu	3299	3296	3323; 3291	3356; 3251
$NH_3^+ \cdots \pi$ Phe	3237−3249	3237	3242; 3238	3334; 3243
$NH_3^+ \cdots$ C═O Orn	∼2880	3048	3037; 3029	3094; 3079
$NH_3^+ \cdots$ C═O Phe	∼2670	2945	2978; 2959	2995; 2955

spectral region where the strong sensitivity of N−H stretching frequencies to N−H environment may provide highly structure specific signatures.[9,28−30] Because of its symmetric structure, four doubly degenerate bands are calculated for the peptidic N−H stretch vibrations of conformer A. As described in the previous section, the eight peptidic N−H bonds are involved in strong hydrogen bond interactions, either as β-sheets or as seven-membered cycles, resulting in red shifts such that no band appears above 3400 cm^{-1}. These general features are in good agreement with the experimental spectrum (Table 2). Experiments carried out with ^{15}N-labeled GS at the Leu and Val residues[9] allowed identification of the Leu and Val N−H bands as those at 3299 and 3331 cm^{-1}. Thus, the Phe and Orn N−H stretch bands must complete the set of bands between 3250 and 3400 cm^{-1} as it is likely that the doublet near 3240−3250 cm^{-1} arises from ammonium N−H motions. A complete band assignment based on computations has been proposed in ref 9. Our results match very well with both experimental and computation ones except for the Val N−H stretching frequency which is overestimated by 41 cm^{-1}. This is due to the functional used rather than to a significant structural difference as discussed in detail below. In addition to the β-sheet interactions, the Leu N−H bonds interact also with the other Val C═O (interactions labeled (2) in Tables 1 and 4) leading to larger red shifts for Leu than for Val N−H stretches. The N−H stretchings of each NH_3^+ group in the Orn side chains appear at lower frequencies since all three N−H stretchings are involved in strong interactions: one cation−π-interaction with the phenyl ring of Phe and hydrogen bonds with Orn C═O and Phe C═O. Frequency matching for the most strongly bound ammonium N−H is not as good as in ref 9 because we use a single scaling factor in this entire region, while additional, mode-specific scaling factors were used in ref 9.

The IRMPD study of Kupser et al.[7] yielded a spectrum in the 1100−1750 cm^{-1} range. It mainly consists of a broad split amide I band at 1580−1720 cm^{-1} and a broad amide II band at 1430−1580 cm^{-1}. Ammonium bands are likely to contribute to both features. It is apparent that the split character of the amide I band reflects the superposition of the better resolved five bands in the UV/IR spectrum.[9] Band assignment was done[7] on the basis of the discussion of five structures of which the two most stable have been briefly described above. The amide I massif computed for the lowest energy structure has the same ordering of residue assignment as ours as a function of wavelength. This assignment was described for deuterated species in which ammonium bands are strongly red-shifted so that a full comparison with our results is difficult in the nondeuterated case. It appears, however, that by disregarding precise frequency values, qualitative band assignments are similar for conformers A of ref 7 and of the present study. As mentioned in the previous section, there exist significant structural differences between the two structures in terms of C_7 hydrogen bonds involving Phe/Orn and Orn/Pro pairs. While these interactions hardly have spectacular effects in the fingerprint range especially with the low resolution of IRMPD spectroscopy, they lead to distinct differences in N−H stretching frequencies. Indeed, the absence of C_7 interactions results in Phe and Orn N−H stretches lying above 3400 cm^{-1}. Moreover, Val N−H streching modes are blue-shifted by 20 cm^{-1} because of weaker β-sheet interactions. It is thus clear that the two structures have significant structural differences and that conformer A from ref 7 is incompatible with the experimental spectrum of ref 9.

Table 3. Differences (in cm^{-1}) between Computed Vibrational Frequencies Using M06 and B3LYP Functionals and Experimental N−H Stretching Frequencies for the Conformer A (relative energies between structures at the same level of computation in kJ/mol)

mode	M06	M06	B3LYP[a]	B3LYP[b]
N−H Phe	−14	9	−9	−7
N−H Orn	−8	8	−5	15
N−H Val	41	36	−1	6
N−H Leu	−3	30	29	14
N−H (π Phe) NH_3^+	−6	2	−1	−4
mean unsigned error	14	17	9	9
relative energy	0.0	+8.3	+7.2	0.0

[a] Optimization starting from the M06 structure. [b] Optimization starting from the structure available in ref 9.

As a conclusion, the room-temperature IRMPD spectrum recorded in the fingerprint region[7] does not appear to provide specific enough information to make structure-spectrum correlations for peptides of this size. While IRMPD experiments carried out in the fingerprint region have been highly successful at deciphering several structural features of small peptides,[11,31−38] experiments on proteins have revealed limitations[39,40] because of spectral congestion.[41,42]

Role of the Functional on the Structure and the Spectrum. As we obtained a structure that can be attributed to conformer A with minor geometric differences in comparison with the one obtained by Nagornova et al.,[9] an analysis of the influence of the functional was carried out on geometries, energies, and spectra. To this aim, the B3LYP structure of ref 9 was reoptimized with the M06 functional, and our own M06 structure was reoptimized with the B3LYP functional. Table 3 presents the deviations between calculated and experimental frequencies in the 3200−3400 cm^{-1}

Table 4. Distances (in Å) for the Hydrogen-Bonded Interactions in Conformer A at B3LYP-D/SVP, M06/6-31G(d,p), and B3LYP/6-31G(d,p) Levels

interaction	B3LYP-D[a]	M06[a]	M06[b]	B3LYP[a]	B3LYP[b]
Leu:C═O···H−N:Val	1.91	1.97	1.92	1.97	2.01
Leu:N−H···O═C:Val	1.95	1.98	2.10	2.30	2.22
C_7 Phe:N−H···O═C:Orn	2.28	2.39	2.45	2.40	2.41
C_7 Orn:N−H···O═C:Pro	2.04	2.09	2.07	2.18	2.23
Leu:N−H···O═C:Val(2)	2.38	2.48	2.29	2.36	2.32

[a] Optimization starting from the M06 structure. [b] Optimization starting from the structure available in ref 9.

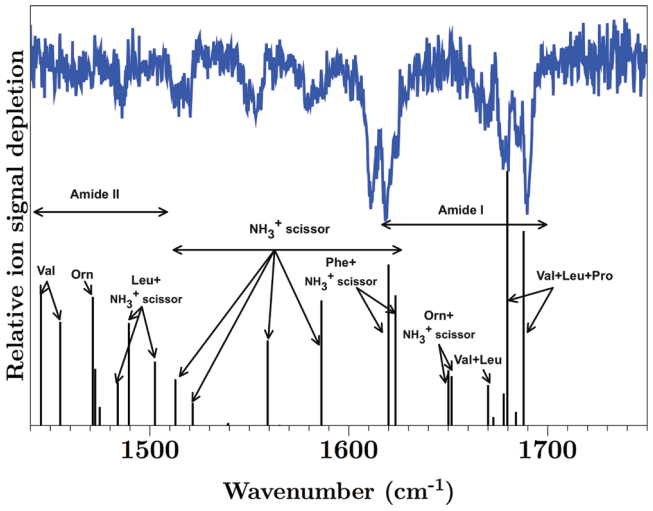


Figure 4. Experimental (blue lines, ref 8) and computed M06 (black lines) IR spectra for conformer B in the 1450−1750 cm^{-1} region.

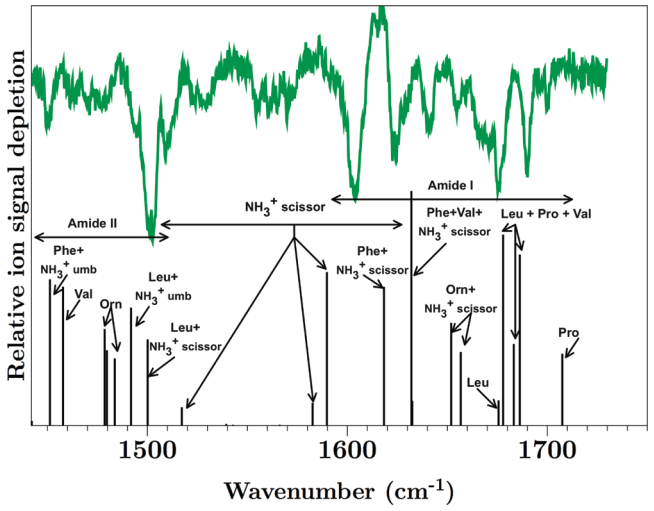


Figure 5. Experimental (green lines, ref 8) and computed M06 (black lines) IR spectra for conformer C in the 1450−1750 cm^{-1} region.

range using M06 and B3LYP functionals. In Table 4, the β-sheet interaction and C_7-motif distances are provided for these structures.

The two structures are characterized by the same network of hydrogen bonds with differences that are chemically insignificant when using the same density functional for both. Yet, they correspond to two different minima on the potential energy surface. The small structural differences (Table 4) lead to differences in N−H stretching frequencies of at most 33 cm^{-1} and most often much less. The largest error arises from the shortest β-sheet interactions: that involving the Val N−H for M06 structures and that involving the Leu N−H for B3LYP structures. Their overall agreement with experiment is the same with mean unsigned errors of 14 versus 17 cm^{-1} at the M06 level and of 9 cm^{-1} for both structures at the B3LYP level (see Table 3). Their relative energies are reversed at the M06 and B3LYP levels. However, such differences remain within the expected uncertainties with the limited basis sets used.

In addition to the functional, the basis set, the starting structure, and the geometry optimizer influence the final results. B3LYP/SVP optimization failed to locate the correct symmetric structure as four C_7 motifs were lost, while B3LYP/6-31G(d,p) was successful. Adding the dispersion correction to yield B3LYP-D/SVP corrected this error. With any method, lack of enforcement of C_2 symmetry when switching functionals led to nonsymmetrical structures that were higher in energy. These results indicate that final refinement of structures is particularly challenging for GS even when C_2 symmetry is enforced. However, within the latter manifold, limited structural differences lead to limited frequency variations and to an overall agreement with experiment.

3.3. Spectral Assignments for Conformer B. Superposition of experimental and calculated spectra of conformer B is given in Figure 4. Amide I bands span the same range as for conformer A with essentially identical attributions. The loss of symmetry leads to a splitting of amide I modes especially for Phe residues. On the basis of the discussion for conformer A, it is, however, unlikely that the experimental doublet at 1610−1630 cm^{-1} arises from the Phe amide I bands only. Contribution of Orn amide I may be invoked here as well. The shapes of the multiplets in the 1650−1690 cm^{-1} range are very similar for A and B, which is in agreement with limited splittings. For the same reason, the intense modes of NH_3^+ scissoring at 1573 and 1520 cm^{-1} for conformer A appear at 1586 and 1559 cm^{-1} and for conformer B appear at 1521 and 1513 cm^{-1}, respectively, which is in better agreement with experiment than what was found above for conformer A. Finally, the amide II region is predicted to significantly spread more than for conformer A. It is, however, difficult to compare to experiment in this region because of the very small signal-to-noise ration below 1500 cm^{-1}.

The predicted frequencies in the 3000−3600 cm^{-1} region are listed in Table 2. The lack of symmetry of conformer B leads to eight distinct bands corresponding to the individual peptidic N−H in the 3290−3420 cm^{-1} range. As for conformer A, all these N−H bonds are involved in hydrogen bonds (see Table 1) albeit with a slightly larger distribution of strengths and therefore

dx.doi.org/10.1021/jp207102v |J. Phys. Chem. B 2012, 116, 483–490

also of frequencies inducing some distinct differences with the spectrum of conformer A. There are no major differences between the ammonium N−H stretching frequencies for conformers A and B except for their systematic splitting which should make their spectra distinguishable in this spectral region as well.

3.4. Conformer C. Superposition of experimental and calculated spectra of the proposed structure for conformer C is given in Figure 5. Comparison with Figures 3 and 4 reveals that the structure of conformer C is likely to be significantly different from those of A and B. This agrees with the features described for C in section 3.1 such as the presence of only three β-sheet interactions, the loss of two C_7-ring interactions, and a stabilization thanks to stronger β-sheet interaction between Leu N−H and Val C═O.

Consequently, some bands appear to be at distinct positions when compared with those of A and B in the amide I range. An additional band at 1700 cm^{-1} is well matched by a pure, free Pro C═O amide I computed at 1708 cm^{-1}. The other Pro C═O is coupled with the Leu and Val C═O stretches yielding the three peaks between 1678 and 1683 cm^{-1} similarly as for conformers A and B, which is in good agreement with the experimental bands in the 1670−1690 cm^{-1} range. The small peak at 1676 cm^{-1} is attributed to a pure Leu C═O stretch mode. As for conformers A and B, the peaks at 1652 and 1657 cm^{-1} correspond to Orn amide I modes coupled with NH_3^+ scissor with an experimental counterpart in the present case. The 1590−1650 cm^{-1} portion of the experimental spectrum displays some clear differences with that of conformer A. Three bands near 1600, 1620, and 1640 cm^{-1} can be seen instead of a doublet around 1610−1630 cm^{-1}. The new band near 1640 cm^{-1} is attributed to the two peaks computed at 1632 cm^{-1} which correspond to the coupling of Phe and Val C═O stretch modes with NH_3^+ scissor. Indeed, this is correlated to a red shift of the C═O stretch frequency of the Val residue which is involved in a stronger β-sheet interaction in conformer C than in A. The band near 1620 cm^{-1} is the same for A and C: Phe amide I mixed with ammonium scissoring. A stronger mixing leads to a blue shift of its companion band dominated by ammonium scissoring with admixture of Phe (and Orn) amide I modes to 1600 cm^{-1}.

Peaks at 1583 and 1517 cm^{-1} are also attributed mainly to NH_3^+ scissors and may be associated with the experimental bands in the 1570−1550 cm^{-1} region. In the amide II region, the Leu modes are predicted at 1500 and 1491 cm^{-1} coupled with NH_3^+ modes matching with the intense experimental peak at 1500 cm^{-1}. The new band near 1450 cm^{-1} is attributed to Phe amide II coupled to NH^{3+} motions. Amide II peaks for Val and Orn are predicted to lie in-between the former two; however, the low signal-to-noise ratio of the experimental spectrum does not allow for a meaningful comparison. Except for the low-intensity features at 1550−1570 cm^{-1}, experimental bands are matched by computed peaks with maximum discrepancies of 15 cm^{-1}.

As for conformer B in the 3000−3600 cm^{-1} region, eight distinct peptidic N−H stretching frequencies are predicted for conformer C (see Table 2). One Orn N−H is free giving a characteristic peak at 3438 cm^{-1}. The other Phe, Val, and Orn N−H are involved in similar interactions as for conformer B. One N−H Leu is involved in a very strong interaction with Val C═O resulting in a red shift of the vibrational N−H stretch frequency below 3300 cm^{-1}. Ammonium N−H bonds are predicted to be involved in the same types of interaction as in conformer B; however, frequencies are extremely sensitive to small structural differences leading to some strong shifts which should make the three conformers easy to distinguish in this region.

4. CONCLUSIONS

As stated by Nagornova et al.,[8] obtaining reliable structures and spectra for peptides as large as GS is challenging for computational chemistry. We took up this challenge by combining an REMD/AMOEBA approach to obtain a reliable set of low-energy conformers and a DFT approach including both Hartree−Fock exchange and dispersion effects to obtain reliable structures and IR spectra. This strategy is quite different from the one taken by Nagornova et al.[9] in that it does not include experimentally determined constraints, which are available for the most abundant conformer only, to guide the conformational search. This allowed identifying several low-energy conformers rather than only the most symmetrical one. Our strategy is more similar to the one used by Kupser et al.[7] The main difference between the two is the use of a polarizable force field for the exploration of the potential energy surface with the REMD approach and for refinement of structures prior to switching to a quantum chemical level. Comparison of final relative energies and spectra suggests that the use of a new generation force field may be crucial to obtain the lowest energy structures. Finally, the use of density functionals including Hartree−Fock exchange appears to be essential for obtaining correct structures as well as accurate spectra. Inclusion of dispersion effects is anticipated to be important as well. The results obtained at the B3LYP-D level (not shown) are similar to the M06 ones lending confidence in our theoretical predictions. Yet, B3LYP which does not include dispersion appears to perform well. The body of DFT results described herein suggests that full mastering of the computation of accurate IR spectra warrants further assessment and improvement of density functionals.

Our results indicate that the gas-phase structures of GS are governed by a large number of hydrogen bonds. As is already well-known, the cyclic backbone can form up to four β-sheet interactions; we find that this maximum number is established in the two most stable conformers, while only three exist in conformer C. In addition, the charged Orn side chains allow for significant structural variations by forming variable numbers of C_7 hydrogen bonds. Release or formation of some of the latter has direct consequences on the lengths of the β-sheet interactions and, therefore, on their strengths and on the corresponding vibrational frequencies. It is this interplay which makes the potential energy surface particularly complex and challenging to explore. The resulting frequency shifts render the N−H stretching frequencies particularly structure-telling. Thus, spectra in the 3000−3600 cm^{-1} range should clearly distinguish the three conformers.

Overall, the present computational study demonstrates the performance of our modeling strategy to reveal the low-energy conformers of peptides such as Gramicidin S for which the secondary structure is governed by numerous noncovalent interactions.

AUTHOR INFORMATION

Corresponding Author

*E-mail: gilles.ohanessian@polytechnique.fr (G. O.); carine.clavaguera@dcmr.polytechnique.fr (C. C.).

■ ACKNOWLEDGMENT

This work was performed using HPC resources from GENCI-[CCRT/CINES/IDRIS] (Grant x2010085107). We thank Prof. T. Rizzo and his group for communicating experimental data prior to publication and for permission to reproduce their published data in Figures 3−5.

■ REFERENCES

(1) Gause, G. F.; Brazhnikova, M. G. *Nature* **1944**, *154*, 703–703.
(2) Llamas-Saiz, A. L.; Grotenbreg, G. M.; Overhand, M.; Van Raaij, M. J. *Acta Crystallogr., Sect. D* **2007**, *D63*, 401–407.
(3) Crowfoot Hodgkin, D.; Oughton, B. M. *Biochem. J.* **1957**, *65*, 752–756.
(4) Xu, Y.; Sugár, I. P.; Krishna, N. R. *J. Biomol. NMR* **1995**, *5*, 37–48.
(5) Prenner, E. J.; Lewis, R. N. A. H.; McElhaney, R. N. *Biochim. Biophys. Acta Biomembr.* **1999**, *1462*, 201–221.
(6) Tishchenko, G. N.; Andrianov, V. I.; Vainstein, B. K.; Woolfson, M. M.; Dodson, E. *Acta Crystallogr., Sect. D* **1997**, *53*, 151–159.
(7) Kupser, P.; Pagel, K.; Oomens, J.; Polfer, N.; Koksch, B.; Meijer, G.; von Helden, G. *J. Am. Chem. Soc.* **2010**, *132*, 2085–2093.
(8) Nagornova, N. S; Rizzo, T. R.; Boyarkin, O. V. *J. Am. Chem. Soc.* **2010**, *132*, 4040–4041.
(9) Nagornova, N. S; Guglielmi, M.; Doemer, M.; Tavernelli, I.; Rothlisberger, U.; Rizzo, T. R.; Boyarkin, O. V. *Angew. Chem., Int. Ed.* **2011**, *50*, 5383–5386.
(10) Rasmussen, T. D.; Ren, P.; Ponder, J. W.; Jensen, F. *Int. J. Quantum Chem.* **2007**, *107*, 1390–1395.
(11) Semrouni, D.; Balaj, O. P.; Calvo, F.; Correia, C.; Clavaguera, C.; Ohanessian, G. *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 728–738.
(12) Jacquemin, D.; Perpete, E. A.; Michaux, C.; Frison, G. *J. Phys. Chem. B* **2011**, *115*, 3604–3613.
(13) Ponder, J. W.; Wu, C. J.; Ren, P. Y.; Pande, V. S.; Chodera, J. D.; Schnieders, M. J.; Haque, I.; Mobley, D. L.; Lambrecht, D. S.; DiStasio, R. A.; et al. *J. Phys. Chem. B* **2010**, *114*, 2549–2564.
(14) Riley, K. E.; Pitoňák, M.; Jurečka, P.; Hobza, P. *Chem. Rev.* **2010**, *110*, 5023–5063.
(15) Zhao, Y.; Truhlar, D. G. *Acc. Chem. Res.* **2008**, *41*, 157–167.
(16) Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120*, 215–241.
(17) Schwabe, T.; Grimme, S. *Acc. Chem. Res.* **2008**, *41*, 569–579.
(18) Bouteiller, Y.; Poully, J. C.; Desfrançois, C.; Gregoire, G. *J. Phys. Chem. A* **2009**, *113*, 6301–6307.
(19) Gloaguen, E.; Valdes, H.; Pangliarulo, F.; Pollet, R.; Tardivel, B.; Hobza, P.; Piuzzi, F.; Mons, M. *J. Phys. Chem. A* **2010**, *114*, 2973–2982.
(20) Semrouni, D.; Clavaguera, C.; Dognon, J. P.; Ohanessian, G. *Int. J. Mass Spectrom.* **2010**, *297*, 152–161.
(21) Stearns, J.; Guidi, M.; Boyarkin, O. V; Rizzo, T. R. *J. Chem. Phys.* **2007**, *127*, 154322.
(22) Cappelli, C.; Monti, S.; Scalmani, G.; Barone, V. *J. Chem. Theory Comput.* **2010**, *6*, 1660–1669.
(23) Carbonniere, P; Dargelos, A.; Ciofini, I.; Adamo, C.; Pouchan, C. *Phys. Chem. Chem. Phys.* **2009**, *11*, 4375–4384.
(24) Corral, I.; Lamsabhi, A. M.; Mó, O.; Yáñez, M. *Int. J. Quantum Chem.* **2011**ASAP.
(25) Ponder, J. W. TINKER, Software Tools for Molecular Design, version 5.1, 2010; http://dasher.wustl.edu/tinker/.
(26) Ahlrichs, R.; Bär, M.; Häser, M.; Horn, H.; Kölmel, C. *Chem. Phys. Lett.* **1989**, *162*, 165−169; for the current version, see http://www.turbomole.com.
(27) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; et al. *Gaussian 09*, revision B.01; Gaussian, Inc.: Wallingford, CT, 2009.
(28) Stearns, J.; Boyarkin, O. V; Rizzo, T. R. *J. Am. Chem. Soc.* **2010**, *129*, 13820–13821.
(29) Rizzo, T. R.; Stearns, J. A.; Boyarkin, O. V. *Int. Rev. Phys. Chem.* **2009**, *28*, 481–515.
(30) Stearns, J.; Seaiby, C.; Boyarkin, O. V; Rizzo, T. R. *Phys. Chem. Chem. Phys.* **2009**, *11*, 125–132.
(31) Prell, J. S.; O'Brien, J. T.; Steill, J. D.; Oomens, J.; Williams, E. R. *J. Am. Chem. Soc.* **2009**, *131*, 11442–11449.
(32) Dunbar, R. C.; Steill, J. D.; Oomens, J. *J. Am. Chem. Soc.* **2011**, *133*, 9376–9386.
(33) Dunbar, R. C.; Steill, J. D.; Polfer, N. C.; Oomens, J. *J. Phys. Chem. B* **2009**, *113*, 10552–10554.
(34) Polfer, N. C.; Oomens, J. *Mass Spectrom. Rev.* **2009**, *28*, 468–494.
(35) Polfer, N. C.; Paizs, B.; Snoek, L. C.; Compagnon, I.; Suhai, S.; Meijer, G.; von Helden, G.; Oomens, J. *J. Am. Chem. Soc.* **2005**, *127*, 8571–8579.
(36) Lucas, B.; Gregoire, G.; Lemaire, J.; Maitre, P.; Glotin, F.; Schermann, J. P.; Desfrançois, C. *Int. J. Mass Spectrom.* **2005**, *243*, 105–113.
(37) Gregoire, G.; Gaigeot, M. P.; Marinica, D. C.; Lemaire, J.; Schermann, J. P.; Desfrançois, C. *Phys. Chem. Chem. Phys.* **2007**, *9*, 3082–3097.
(38) Balaj, O. P.; Kapota, C.; Lemaire, J.; Ohanessian, G. *Int. J. Mass Spectrom.* **2008**, *269*, 196–209.
(39) Poully, J. C.; Lecomte, F.; Nieuwjaer, N.; Manil, B.; Schermann, J.-P.; Desfrançois, C.; Calvo, F.; Gregoire, G. *Phys. Chem. Chem. Phys.* **2010**, *12*, 3606–3615.
(40) Poully, J. C.; Lecomte, F.; Nieuwjaer, N.; Manil, B.; Schermann, J.-P.; Desfrançois, C.; Gregoire, G.; Ballivian, R.; Chirot, F.; Lemoine, J.; Calvo, F.; Antoine, R.; Dugourd, P. *Int. J. Mass Spectrom.* **2010**, *297*, 28–35.
(41) Oomens, J.; Polfer, N. C.; Moore, D. T.; Van der Meer, L.; Marshall, A. G.; Eyler, J. T.; Meijer, G.; von Helden, G. *Phys. Chem. Chem. Phys.* **2005**, *7*, 1345–1348.
(42) Fung, Y. M. E.; Besson, T.; Lemaire, J.; Maître, P.; Zubarev, R. A. *Angew. Chem., Int. Ed.* **2009**, *48*, 8340–8342.