THE JOURNAL OF PHYSICAL CHEMISTRY A

Article

pubs.acs.org/JPCA

Conformational Study of Z-Glu-OH and Z-Arg-OH: Dispersion Interactions versus Conventional Hydrogen Bonding

Sander Jaeqx,[†] Weina Du,[‡] Evert Jan Meijer,[‡] Jos Oomens,[†,§,∥] and Anouk M. Rijs*[,∥]

[†]FOM Institute for Plasma Physics Rijnhuizen, Edisonbaan 14, 3439 MN Nieuwegein, The Netherlands

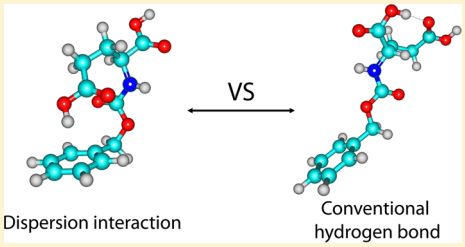
[‡]van't Hoff Institute for Molecular Sciences and Amsterdam Center for Multiscale Modeling, University of Amsterdam, Science Park 904, 1098 XH Amsterdam, The Netherlands

[§]van't Hoff Institute for Molecular Sciences, University of Amsterdam, Science Park 904, 1098 XH Amsterdam, The Netherlands

[∥]Radboud University Nijmegen, Institute for Molecules and Materials, FELIX Facility, Toernooiveld 7, 6525 ED Nijmegen, The Netherlands

Supporting Information

ABSTRACT: The gas-phase conformational preferences of the model dipeptides Z-Glu-OH and Z-Arg-OH have been studied in the low-temperature environment of a supersonic jet. IR-UV ion-dip spectra obtained using the free electron laser FELIX provide conformation-specific IR spectra, which in combination with density functional theory (DFT) allow us to determine the conformational structures of the peptides. Molecular dynamics modeling using simulated annealing generates a variety of low-energy structures, for which geometry optimization and frequency calculations are then performed using the B3LYP functional with the 6-311+G(d,p) basis set. By comparing experimental and theoretical IR spectra, three conformations for Z-Glu-OH and two for Z-Arg-OH have been identified. For three of the five structures, the dispersion interaction provides an important contribution to the stabilization, emphasizing the importance of these forces in small peptides. Therefore, dispersion-corrected DFT functionals (M05-2X and B97D) have also been employed in our theoretical analysis. Second-order Møller−Plesset perturbation theory (MP2) has been used as benchmark for the relative energies of the different conformational structures. Finally, we address the ongoing debate on the gas-phase structure of arginine by elucidating whether isolated arginine is canonical, tautomeric, or zwitterionic.



INTRODUCTION

Proteins are at the heart of the biochemical factory of an organism's cells, performing numerous tasks including the catalysis of biochemical reactions, transport of oxygen, transmitting nerve impulses, and acting as biomolecular motors. The operating mechanisms of proteins are not only determined by the sequence of amino acid (AA) residues from which they are constructed, but also by the three-dimensional structure, i.e., the conformation or folding structure, that is adopted by the protein. The structure of a protein is to a large extent determined by its backbone intramolecular hydrogen bonds (HB's), but also HB, electrostatic, and van der Waals interactions between the backbone, the amino acid side chains, and the environment are of influence.[1,2]

For a detailed understanding of these intramolecular interactions, the structural conformations of small benchmark peptides have been spectroscopically studied under isolated conditions, focusing mainly on neutral amino acids,[3−5] dipeptides,[3,6,7] and tripeptides.[3,8] To specifically mimic the behavior of the peptide bonds and to avoid interactions with the basic and acidic termini of the peptide, capped peptides with an acetyl group on the N-terminus and an amide or methylamide group on the C-terminus are often studied.[9−14] Secondary structural motifs such as helices,[9,15,16] β-turns,[11] and γ-turns[17,18] have been unambiguously identified in the gas phase based on their IR spectral signatures. Larger systems with four,[19,20] five,[21,22] and even systems up to 15 amino acids[23] have been studied as well, although the structural assignment becomes more challenging with increasing peptide size due to spectral congestion. Moreover, many studies on complexes of AAs with solvent molecules,[3,24−26] in a protonated[15,24,27] or deprotonated form,[27] or AA complexes with metal-cations[24,28,29] have been performed to understand the structure to function relation in peptides and proteins.

Gas-phase infrared (IR) spectroscopy in combination with quantum-chemical calculations has proven to be a very powerful tool to study the conformations of cold isolated biomolecules. These gas-phase measurements enable us to study the intrinsic properties of biomolecules in the absence of external perturbations, e.g., solvent molecules. Moreover, complexes of solvent molecules and metal-ions with AA's can be made selectively, to study the effect on the conformations and interactions of AA's.[25,26,30] The IR frequencies of the amide

Special Issue: Peter B. Armentrout Festschrift

Received: May 31, 2012
Revised: October 24, 2012
Published: October 24, 2012

dx.doi.org/10.1021/jp3053339 | J. Phys. Chem. A 2013, 117, 1216−1227

bond C=O stretch (Amide I) and NH bend (Amide II) modes undergo a diagnostic shift upon hydrogen bonding and thus provide us with a direct view on the hydrogen bond network. Together with the rest of the fingerprint region (900−1400 cm^{-1}) these modes typically reveal much of the peptide's secondary structure.

In this paper, we focus on the influence of some of the most acidic and the most basic AAs, glutamic acid (Glu) and arginine (Arg), on the peptide's conformational structure.[1] The protonation/deprotonation chemistry of these two AA's in their neutral form has been the subject of numerous publications. The most stable gas-phase structure of neutral Arg has been the subject of a long debate. It has been suggested that a zwitterionic form of Arg may be the preferred gas-phase structure, in which the C-terminus is deprotonated and the guanidine side chain is protonated (Figure 1d). For the

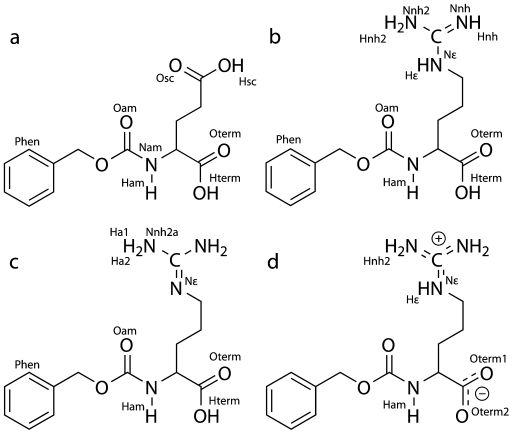


Figure 1. Chemical structures of (a) Z-Glu-OH, (b) canonical Z-Arg-OH, (c) tautomerized structure Z-Arg-OH, and (d) zwitterionic Z-Arg-OH.

nonzwitterionic form of Arg, two different tautomers can exist (Figure 1b,c), of which the relative stabilities differ by only ∼1−2 kcal/mol.[31−35] Upon the basis of the observation of carbonyl stretch vibrations in cavity ring-down spectroscopy experiments, Chapo et al. determined neutral gas-phase Arg to possess a nonzwitterionic structure,[36] although a detailed structural assignment of gas-phase Arg describing the different possible conformations has not been reported to date. Protonated and metal-ion cationized Arg have been structurally characterized using ion spectroscopy.[24,28]

For glutamic acid, various theoretical and experimental studies on its proton affinity[37] and gas-phase acidity have been reported; however, experimental studies addressing the conformational preferences of gas-phase Glu are largely lacking.[38−40]

Here, a conformational study is performed on two N-capped amino acids, Z-Glu-OH and Z-Arg-OH (Z = benzyloxycarbonyl, see Figure 1). The Z-cap is incorporated as a UV-chromophore, enabling us to perform conformation-specific IR-UV spectroscopy. The IR absorption spectra of three conformers of Z-Glu-OH and two conformers of Z-Arg-OH are recorded in the mid-infrared range (900−1850 cm^{-1}). By comparing these experimental spectra with theoretical spectra of low energy conformations, we identify the most abundant gas-phase conformation. Various levels of theory have been used to elucidate the influence of dispersion interactions in these peptides.

EXPERIMENTAL SECTION

Experimental Methods. Z-Glu-OH and Z-Arg-OH were obtained from Bachem and Fluka, respectively, and were used without further purification. The setup used for the experiments is described in more detail elsewhere.[41] For all experiments, the sample was mixed with graphite powder and applied onto a solid graphite bar. Laser desorption was used to bring the sample in the gas phase. A pulsed IR laser (Polaris Pulsed Nd: YAG Laser System from New Wave research) with a pulse energy of about 1.5 mJ and wavelength of 1064 nm was used to desorb the sample molecules from the graphite substrate as intact neutral molecules. The gas-phase sample molecules are taken along in the supersonic molecular beam, produced by a pulsed valve with a backing pressure of 3 bar of argon. The temperature of the molecules is estimated to be on the order of 10−20 K.[42]

The molecular beam is skimmed, and the sample molecules enter a differentially pumped reflector time-of-flight (TOF) mass spectrometer. Here, the molecules interact with a UV beam produced by a pulsed Nd: YAG laser (either Innolas GmbH spitlight 1200 or Quanta-Ray Lab Series) coupled to a frequency doubled dye laser from Radiant Dyes (laser dye: coumarin 153). The UV laser is operated at 10 Hz and typical pulse energies are on the order of 1−2 mJ. The sample molecules are two-photon ionized via 1 + 1 REMPI scheme.[43] Different conformations possess a slightly different electronically excited state energy, and thus appear as different peaks in the UV excitation spectrum.

For the IR-UV double resonance spectra, the IR and UV beams were spatially overlapped, but the IR pulse preceded the UV pulse by ∼200 ns. The IR radiation is produced by the Free Electron Laser for Infrared eXperiments (FELIX) at the FOM institute Rijnhuizen.[44] The frequency of the UV probe pulse is fixed on a transition found in the REMPI experiment to select the conformation of interest, producing a constant ion signal. Whenever the IR hole-burn laser excites a transition that shares the same ground state as the probe laser (and thus is the same conformation), a dip in the ion signal is observed since the ground state is depleted by the IR laser. An IR absorption spectrum of a single conformer can thus be constructed by taking the logarithm of the ion signal without the IR pulse divided by the ion signal with IR pulse. To correct for long-term UV power drifts and changing source conditions, alternating IR-on and IR-off signals are measured by operating the IR laser at 5 Hz and the UV laser at 10 Hz.

To probe the number of conformations present in the experiment, a technique similar to that described above is employed, where the IR laser is, however, fixed on a vibrational transition of a specific conformer, while the UV laser is tuned in frequency. In this way, peaks in the REMPI spectrum belonging to the same conformation are removed, while peaks originating from another conformation are not affected.

Theoretical Methods. Both Z-Glu-OH and Z-Arg-OH are very flexible molecules with many degrees of freedom. As a consequence, the potential energy surface (PES) is very complex with numerous local minima. An elaborate conformational search was performed to probe the PES and find the local/global minima that corresponded to the structures observed in the experiment. Structural assignment is based on comparison between theory and experiment. Briefly, the computational strategy used consisted of input structures for higher level calculations being generated by force field

dx.doi.org/10.1021/jp3053339 | J. Phys. Chem. A 2013, 117, 1216−1227

calculations. Since the structures are ultimately optimized using DFT, the lack of flexible molecular charge in the force field was assumed to have minor influence on the final results.

The conformational search was performed by applying the simulated annealing (SA) approach using the GROMACS4 package[45] and the amber99sb forcefield.[46] The maximum temperature used in the simulations was 1300 K, and the simulation lengths were 20 ns with time steps of 2 fs. The temperature was lowered exponentially to 5 K in 20 ps. The structures at 5 K were stored. 1000 structures were generated, from which about 50 structures were selected, based on their relative energy and on chemical intuition, to be optimized at the B3LYP/6-31G(d,p) level of theory using the Gaussian09 program package.[47] About 25 "promising" structures were reoptimized at the B3LYP/6-311+G(d,p) level of theory, and their harmonic vibrational frequencies were computed at the same level. To correct for anharmonicity, a scaling factor of 0.9845 was used. The theoretical spectra were convoluted with a Gaussian line shape function with a full width at half-maximum of 1% of the center frequency to approximately match the observed experimental line width. Additional low temperature SA simulations were performed on promising structures that showed good agreement with experiment in the Amide I and Amide II region, but poor agreement in the fingerprint region, so that small conformational changes are induced, while keeping the hydrogen bond network unchanged.

Neutral Glu, neutral Arg, Arg with proton transfer from N_ε to N_{nh} (see Figure 1), as well as the Z-cap are not present in the amber99sb forcefield, and thus had to be implemented manually. The parametrization was done by first building the specific residue in Chemcraft,[48] with an alanine residue on both the N- and the C-terminus to simulate a peptide environment. The atomic charges were determined by the AM1-BCC charge method implemented in AmberTools. Parameters for new bond lengths, bond angles, dihedrals and impropers were copied from existing analogues in the Ambersb99 force field. The new parameters were derived in a manner consistent with how the rest of the force field was originally derived.[49]

Although B3LYP is frequently used for geometry optimizations and frequency calculations of biomolecules, a known deficiency is that dispersion interactions are not taken into account.[50] In addition, computing zero point energies remains problematic.[51] Therefore, structure optimizations and frequency calculations are also performed with the DFT-D functionals M05-2X and B97D, which make use of an empirical correction for the dispersion energy.[52] Furthermore, structure optimization is also performed with the second order Møller–Plesset perturbation theory (MP2), which is assumed to give the most accurate energy because of the inclusion of electron correlation.[53,54] These additional calculations all employ the 6-311+G(d,p) basis set. For the DFT methods, not only the electronic energies are determined, but in addition the zero-point energy (ZPE)-corrected energies and Gibbs free energies at 300 K are evaluated.

RESULTS AND DISCUSSION

Z-Glu-OH. *Experimental Results: UV and IR spectroscopy.* The top trace of Figure 2 shows the UV excitation spectrum of Z-Glu-OH recorded in the UV region from 37500 to 37800 cm^{-1}. Numerous sharp transitions are observed over a range of 250 cm^{-1}. The number of conformations present in the gas phase is usually determined by UV–UV hole burning spectroscopy (HBS). However, due to low depletion signals,

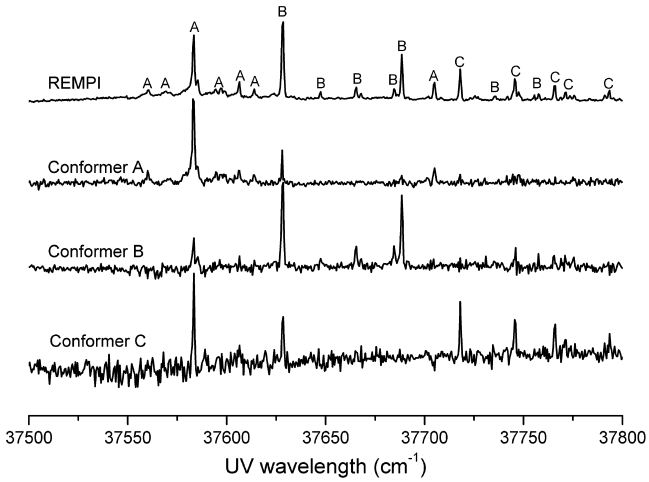


Figure 2. REMPI spectrum of Z-Glu-OH (top trace) and IR-UV HB spectra for three conformations of Z-Glu-OH.

this technique could not be applied for Z-Glu-OH. Therefore, IR-UV ion-dip spectra were recorded for five different UV wavelengths corresponding to the main peaks in the REMPI spectrum. Whenever the IR-UV spectra were identical, one can assume that they originate from the same conformation. Three distinct IR absorption spectra were found, indicating that at least three conformations (A, B and C) coexist under our experimental conditions. To check for additional conformations, another version of the IR-UV IDS technique was used. Here, the IR laser was fixed on a peak in an IR absorption spectrum of one of the conformers, where the other conformers showed no or little absorption. By scanning the UV laser, peaks in the REMPI spectrum originating from the same conformation are depleted. The IR laser was fixed on 1164 cm^{-1} for conformation A, for conformation B on 1530 cm^{-1}, and for conformation C on 1016 cm^{-1}, as indicated in Figure 3. These hole-burn spectra are presented in the three lower traces

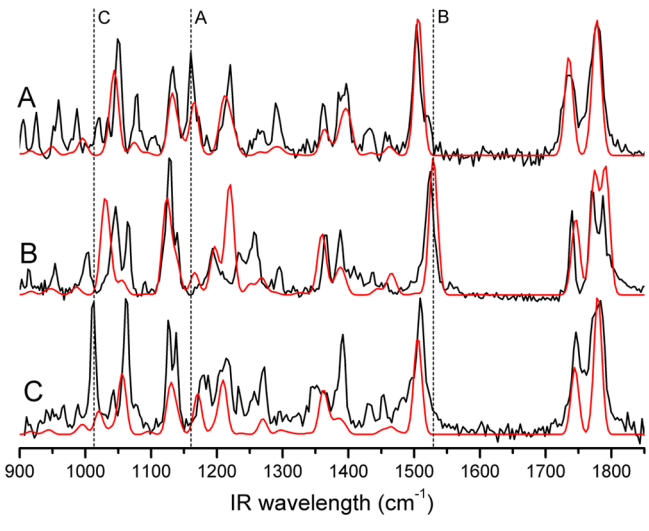


Figure 3. IR-UV ion dip spectra of three Z-Glu-OH conformations (black), and the theoretical spectra of the assigned structures (red). The vertical lines indicate the IR frequencies used for IR-UV HBS experiments. UV wavelengths used are 37583.5 cm^{-1} for conformation A, 37628.5 cm^{-1} for conformation B and 37718 cm^{-1} for conformation C. Conformation A is assigned to G5, conformation B to G1, and conformation C to G2.

dx.doi.org/10.1021/jp3053339 | J. Phys. Chem. A 2013, 117, 1216–1227

of Figure 2. Overlapping IR absorptions and the finite spectral width of the IR laser caused a minor depletion of other conformers besides the selected conformer. However, by comparing intensities and the fraction of the signal depleted, all of the peaks found in the REMPI spectrum could be assigned to one of the three conformations.

The three conformations identified have their fundamental $S_1 \leftarrow S_0$ transition at 37583.5 cm^{-1} (conformer A), 37628.5 cm^{-1} (conformer B), and 37718 cm^{-1} (conformer C). They show a vibrational progression, conformation A and B even to up to 125 cm^{-1} above the fundamental $S_1 \leftarrow S_0$ transition. Hot bands are found on the red side of the fundamental transition, from which a vibrational temperature of 15 K was determined for the molecules.

Figure 3 presents the IR absorption spectra of the three conformations of Z-Glu-OH in the region between 900 cm^{-1} and 1850 cm^{-1}. The IR absorption spectra are recorded for conformers A, B and C with the ionization laser parked on the peaks at 37583.5 cm^{-1}, 37628.5 cm^{-1} and 37718 cm^{-1}, respectively. In all three spectra, the peaks between 1600 and 1800 cm^{-1} can be assigned to the C═O stretching modes, with the carboxylic acid C═O stretch vibrations at about 1780 cm^{-1} and the Amide I (C═O stretch vibration) around 1740 cm^{-1}. Additionally, in the measured wavelength region, the Amide II (NH bend- and C−N stretch vibration) band can be recognized between 1500 and 1520 cm^{-1}, and many other conformationally diagnostic modes are observed in the fingerprint region (900−1400 cm^{-1}).

On the basis of the frequencies of the C═O stretch vibrations of the side chain, the C-terminus and the amide C═O, we conclude that there is no strong hydrogen bond involving any of these C═O moieties in each of the three conformations. Such HB's would cause a noticeable red shift of 10 to 50 cm^{-1},[55] which is not observed. However, conformation B shows a split C═O stretch peak with the two components separated by 16 cm^{-1} at about 1780 cm^{-1}. This relatively small splitting indicates a slightly different HB environment for the two carboxylic acid groups. The splitting is possibly due to hydrogen bonding of the hydroxyl group of one of the carboxylic acid moieties, shifting the C═O stretch vibration slightly to the red as a consequence of the so-called "spill over effect" upon hydrogen bonding.[56] In addition, the NH bend vibration in conformer B is shifted to the blue by 20 cm^{-1} with respect to the amide NH bending modes in conformations A and C, indicating that the NH group is involved in an HB.[57]

In addition to the Amide I and Amide II region, the fingerprint region from 900 to 1400 cm^{-1} shows a wealth of resolved peaks, which can further aid in the structural assignment.[55] In this region, mostly vibrations involving the phenyl ring and the CH_2 rocking and wagging modes are expected. These vibrations are often strongly coupled to each other. Other vibrational modes found in this region are COH bend vibrations and single-bond CO, CN, and CC stretching modes.

Calculations. Z-Glu-OH possesses multiple hydrogen donor and acceptor sites, so that it can adopt various hydrogen-bonding conformations. Hydrogen bond acceptors include the side-chain carboxylic acid C═O, the terminal carboxylic acid C═O and the phenyl ring of the Z-cap. The amide NH, side-chain carboxylic acid OH and the terminal carboxylic acid OH group act as hydrogen bond donors. Figure 4 presents the 12 most stable structures of Z-Glu-OH. Rotamers or similar structures have been excluded. The intramolecular interactions

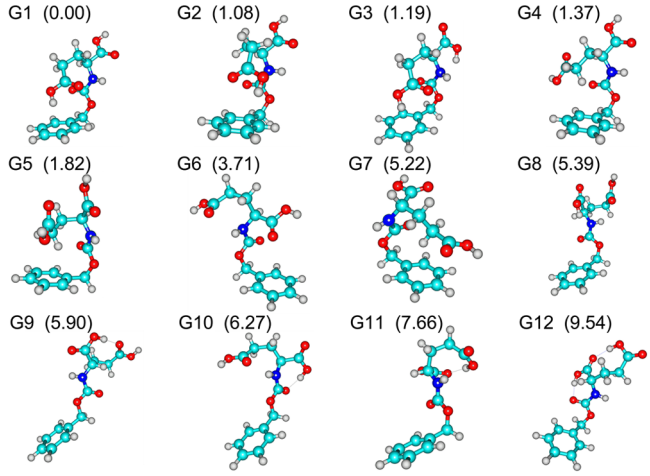


Figure 4. Lowest-energy structures of Z-Glu-OH. Relative energies at the MP2-level are given in parentheses in kcal/mol.

of each of the structures are listed in Table 1, and the atom labels can be found in Figure 1.

Table 1. Intramolecular Interactions for Optimized Structures of Z-Glu-OH[a]

structure	interactions
G1	$H_{am} - O_{term} + H_{sc} - phen$
G2	$H_{am} - O_{term} + H_{sc} - phen$
G3	$H_{am} - O_{sc} + H_{sc} - phen + H_{term} - N_{am}$
G4	$H_{am} - O_{term}$
G5	$H_{am} - O_{term}$
G6	$H_{am} - O_{sc}$
G7	
G8	$H_{am} - O_{sc} + H_{term} - N_{am}$
G9	$H_{am} - O_{term} + H_{term} - O_{sc}$
G10	$H_{am} - O_{sc} + H_{term} - O_{am}$
G11	$H_{term} - O_{sc} + H_{am} - O_{am}$
G12	$H_{sc} - O_{term} + H_{term} - O_{am}$

[a]For atom labels, see Figure 1.

The computed relative energies for the different conformers using the different theoretical methods are listed in Table 2. Some striking differences between the energies of the various structures at the different levels of theory are noticed. For the pure DFT method (B3LYP), structures with strong intramolecular hydrogen bonds (between amide or carboxylic acid C═O groups and carboxylic acid OH or amide NH groups) are lowest in energy. As expected, structures with no hydrogen bonds are higher in energy. In addition, London dispersion interactions between the side chain carboxylic acid OH group and the phenyl group can provide substantial stabilization, i.e., of the same order of magnitude as a normal hydrogen bond, as was shown by Vondrásek et al.[58] and Kumar et al.[59] However, it is well-known that DFT does not take dispersion interactions into account, so that structures with such interactions (G1, G2, and G3 of Figure 4) become further stabilized when dispersion-corrected DFT-D methods (M05-2X and B97D) are used instead of B3LYP. Hydrogen bond interactions are treated equally in DFT and DFT-D.

Structures with strong HB's between C═O groups and OH or NH groups are very stable using DFT; however, their relative energies increase significantly when MP2 is used. The

dx.doi.org/10.1021/jp3053339 | J. Phys. Chem. A 2013, 117, 1216−1227

Table 2. Relative Electronic Energies (E_{elec}), ZPE-Corrected Energies (ZPE) and Gibbs Free Energies at 300 K (ΔG) for the Lowest-Energy Structures of Z-Glu-OH[a]

	MP2	B3LYP			B97D			M05-2X		
struc	E_{elec}	E_{elec}	ZPE	ΔG	E_{elec}	ZPE	ΔG	E_{elec}	ZPE	ΔG
G1	0.00	1.30	1.57	0.53	0.00	0.00	0.00	0.00	0.00	0.00
G2	1.08	2.16	2.41	1.26	1.59	1.44	1.34	1.70	1.84	1.94
G3	1.19	2.53	2.45	2.11	0.60	0.87	1.39	1.46	1.89	1.84
G4	1.37	0.46	0.63	−0.30	1.56	1.53	0.74	0.83	1.07	0.39
G5	1.82	0.99	1.21	0.11	2.68	2.56	1.64	2.23	2.49	1.87
G6	3.71	1.43	1.65	0.07	3.34	3.25	1.60	3.49	3.60	1.91
G7	5.22	4.79	5.16	3.62	5.96	5.60	4.52	5.27	5.32	5.92
G8	5.39	3.77	3.40	4.05	4.09	4.40	3.79	4.49	5.15	4.73
G9	5.90	0.00	0.00	0.00	2.96	2.87	1.75	2.20	2.43	2.77
G10	6.27	0.73	0.79	0.16	3.32	3.01	1.48	3.58	3.58	3.69
G11	7.66	5.98	5.66	6.91	5.90	6.24	5.71	6.18	6.88	6.63
G12	9.54	4.80	4.26	6.31	5.61	5.98	5.92	6.53	7.20	5.82

[a]Energies are given in kcal/mol. All calculations employed the 6-311+G(d,p) basis set.

lowest-energy structure (G1) according to the B97D and M05-2X methods agrees with that based on the MP2 energy calculation. However, these DFT-D methods also seem to overestimate the stabilization energy due to hydrogen bonding, as is observed for structures G4, G9, G10, G11, and G12. Although B3LYP is a good method to perform structure optimization, correct energy values are a known problem for DFT calculations.[51]

As observed in the IR spectra of Figure 3, the frequencies of the C═O stretch and NH bend vibration indicate that there are no strong intramolecular interactions involving one of the C═O groups. Although interactions between the π-cloud of the aromatic ring and OH-group of the carboxylic acid moiety may introduce small shifts in the C═O stretch position,[60] structures exhibiting strong interactions with one of the C═O groups can be excluded from the structural assignment. Such strong interactions are observed in structures G3, G8, G9, G10, G11 and G12.

Conformation B shows a splitting between the two carboxylic acid C═O stretch vibrations, indicating a different hydrogen bonding environment for the side-chain and terminal COOH groups. In Figure 5a, the experimental IR spectrum of conformation B in the C═O stretch region is compared with calculated spectra for structures G1 and G6, which both show a splitting between the two carboxylic acid C═O stretch vibrations. The frequencies of the carboxylic C═O stretch vibrations match better for structure G1 than for structure G6.

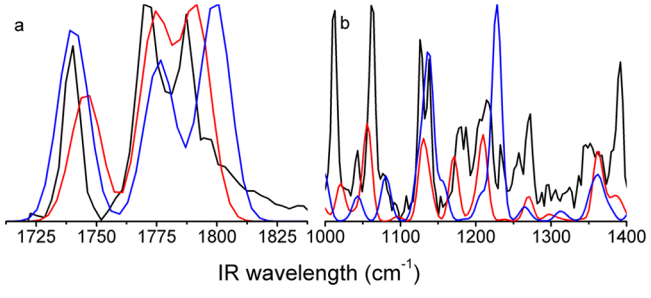


Figure 5. (a) Experimental IR spectrum of conformation B (black) and computed spectra for structures G1 (red) and G6 (blue) in the C═O stretch range. (b) Experimental IR spectrum of conformation C (black) and computed spectra for structures G2 (red) and G7 (blue) in the fingerprint region.

On the other hand, the amide C═O stretch matches better with structure G6. Comparing the energies of the two structures (Table 2), one notices that although both structures are almost isoenergetic at the B3LYP level of theory, structure G1 is much lower in energy when dispersion interactions are included (B97D, M05-2X and MP2). We therefore assign conformation B to structure G1.

The frequencies of the amide NH bend and C═O stretch vibrations of conformation A, conformation C, and structures G2, G4, G5, and G7 are listed in Table 3. From an

Table 3. Experimentally Observed and Calculated Frequencies in the Amide Region of Several Conformations of Z-Glu-OH[a]

		C═O stretch	
	NH bend	amide	carboxylic acid
conformation A	1503.7	1735.3	1776.9
conformation C	1509.2	1745.9	1783.6
structure G2	1504	1743.8	1778
structure G4	1514.1	1736.7	1777.8
structure G5	1505	1734	1778.5
structure G7	1516.6	1743.1	1780.5

[a]Frequencies are given in cm^{-1}.

experimental-versus-theoretical comparison of these three frequencies, conformation A can be assigned to structure G5. The largest deviation between theory and experiment for the NH bend and C═O stretch vibrations is only 1.6 cm^{-1}. Moreover, the remainder of the fingerprint region also shows a good match between theory and experiment, as can be seen in the top trace in Figure 3. While the computed frequencies for G4 are close, Figure SI-2 (Supporting Information) shows that structure G5 provides a sensitively better match to the experimental frequencies. However, structures G4 and G5 are very similar (see Figure 4), differing only in the orientation of the carboxylic acid group of the side chain. It is possible that the two conformations are simultaneously excited, despite the conformation selectivity of the IR-UV ion dip method. This may occur when the electronic $S_1 \leftarrow S_0$ transitions of G4 and G5 are very close.[61] The small shoulder on the blue side of the Amide II peak near 1500 cm^{-1} and the shoulder on the red side of the peak at 1221 cm^{-1} indeed suggest the coexistence of

dx.doi.org/10.1021/jp3053339 | J. Phys. Chem. A 2013, 117, 1216−1227

both G4 and G5 structures in the gas phase. Moreover, both structures have comparable energies with G5 only 1 kcal/mol higher. We suspect that both conformations are present in the gas phase, and that the observed spectrum of conformer A originates from G5 with a small contribution of G4.

The vibrational frequencies presented in Table 3 also assist in the structural assignment of conformation C. The red-shifted amide C═O stretch vibration of structure G4 excludes this structure as the assignment for conformer C. However, based on the NH bend and C═O stretch frequencies of the remaining structures G2 and G7, it is not clear to which structure conformation C should be assigned; the spectrum toward longer wavelengths in the fingerprint region provides further clues for the structural assignment of this conformer. In Figure 5b, the spectrum of conformation C in the 1000−1400 cm^{-1} range is compared with the calculated spectra of structures G2 and G7. The overall agreement is clearly better for structure G2. In addition, the calculated energy of structure G2 is about 4 kcal/mol lower compared with that of structure G7 at the MP2 and DFT-D levels of theory. Hence, we assign conformation C to structure G2.

Discussion. Figure 3 shows the IR spectra of the three conformations of Z-Glu-OH together with the calculated IR spectra of the assigned structures. In the Amide I and Amide II regions of the spectrum, the agreement between theory and experiment is excellent. However, in the fingerprint region some discrepancies are observed, especially when dispersion interactions are present, since this type of interaction is neglected in B3LYP. Moreover, agreement in terms of relative intensities is not nearly as good as in terms of frequencies.

Figure SI-1 of the Supporting Information shows the calculated spectra of the three conformations of Z-Glu-OH using the DFT functional B3LYP and the DFT-D functionals B97D and M05-2X. One notices that, although the calculated energies are more accurate for B97D than for B3LYP, the observed frequencies, especially those of the C═O stretch vibrations, are poorly reproduced by B97D. The M05-2X functional reproduces the experimental frequencies in the Amide I and Amide II region reasonably well, although not as accurately as B3LYP. For structures exhibiting dispersion interactions, the agreement between the M05-2X calculated and experimental frequencies is better in the fingerprint region; see, for example, conformers B and C of Z-Glu-OH. If dispersion interactions are not present, the B3LYP functional outperforms M05-2X in the fingerprint region (see conformer A of Z-Glu-OH).

On the basis of the relative stabilities in Table 2, it is not surprising that structures G1 and G2 are observed, i.e., they have the lowest energies. However, one may wonder why structure G3 is not observed, while the higher-energy structures G4 and G5 are experimentally observed. Perhaps its relatively high Gibbs free energy (at 300 K) prevents structure G3 from being present in the molecular beam. As was shown by Gloaguen et al.,[62] Řeha et al.,[63] and Shubert et al.,[64] the relative conformational abundances measured in the experiment are related to the relative conformational abundances at the temperatures prior to expansion cooling. During the rapid cooling in the supersonic expansion, it is assumed that the system is unable to isomerize and thus access conformations that were high in free energy at 300 K. Another explanation can be found by comparing structures G3 and G1, which are very similar and differ only in the orientation of the C-terminal COOH group. In structure G1, the C═O moiety has a C_5-interaction with the amide NH, while in structure G3 the amide NH interacts with the side chain C═O and the C-terminal OH group interacts with the lone pair of the amide nitrogen. If the barrier between these two structures is low, structure G3 will not become populated because it will easily be converted into G1.

Z-Arg-OH. *Experimental Results: REMPI.* The mass spectrum of Z-Arg-OH obtained upon UV irradiation at 37574 cm^{-1} is shown in Figure 6. The peak at m/z = 308

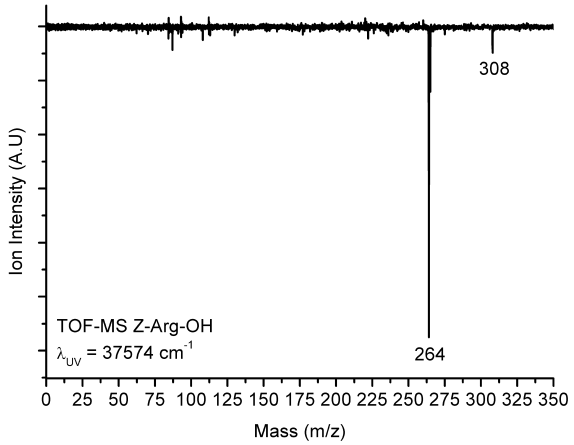


Figure 6. Time-of-flight mass spectrum of Z-Arg-OH at 37574 cm^{-1}.

corresponds to the Z-Arg-OH parent ion. In addition to this peak, a major fragment is observed at m/z = 264, which suggests either loss of a CO_2 group or a $C(NH_2)_2$ group. In the tautomeric structure, as well as in the proton transfer structure, the $C(NH_2)_2$ group is present. An exact mass determination on the m/z 264 fragment produced by collision induced dissociation (CID) of protonated Z-Arg-OH in a high-resolution FTICR mass spectrometer suggests that the fragment originates from loss of CO_2.

A REMPI spectrum recorded on the TOF mass channel of the parent m/z 308 ion as well as on the m/z 264 fragment ion of Z-Arg-OH is shown in Figure 7. Unlike the REMPI spectrum of Z-Glu-OH (Figure 2), the UV excitation spectrum of mass 308 is not as well resolved. Instead, a broad spectrum is observed with some sharp resonances superimposed on it. This

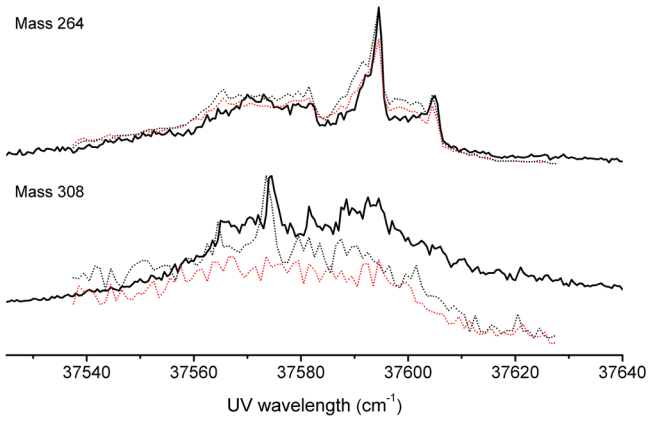


Figure 7. REMPI spectrum of Z-Arg-OH probing mass 264 (top trace) and mass 308 (bottom trace) in solid. Dotted lines indicate IR-UV HBS spectra, where the red line is preceded by IR irradiation at 1140 cm^{-1}.

dx.doi.org/10.1021/jp3053339 | J. Phys. Chem. A 2013, 117, 1216−1227

broad spectrum may result from incomplete cooling of the low energy vibrational modes. The combination of a sharp transition in the REMPI spectrum of m/z 264 and an intense hot band also suggests incomplete cooling. However, the broadening may also be due to a short lifetime of the electronically excited state.

Similar to Z-Glu-OH, the number of distinct conformations present is determined by recording IR absorption spectra with the UV laser tuned to each of the dominant REMPI peaks as well as to a few wavelengths in the broad part of the UV excitation spectrum. Such IR-UV ion dip spectra were recorded for both the 308 and 264 mass channels. From this analysis, two different conformations were identified, as shown in Figure 8. The IR spectrum measured with the UV excitation laser at

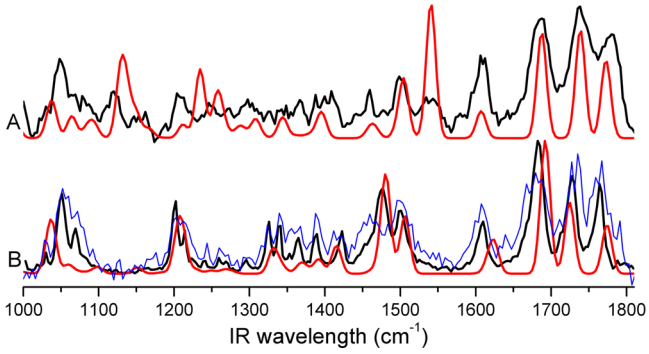


Figure 8. IR-UV ion dip spectra of two Z-Arg-OH conformations (black), and the theoretical spectra of the assigned structures (red). The top spectrum is of mass 308, the bottom trace of mass 264. The blue line in the bottom trace is measured at mass 308 showing that they belong to a single conformation. UV wavelengths used are 37574 cm^{-1} for conformation A and 37594 cm^{-1} for conformation B. Conformation A is assigned to AC3, and conformation B is assigned to AT3.

37597 cm^{-1} for mass 308 is identical to the IR spectrum extracted from mass channel 264 at the same UV excitation wavelength, as seen in the lower trace of Figure 8.

As can be seen in Figure 8, conformation A exhibits an absorption at 1140 cm^{-1}, while conformation B has no absorption at this frequency. We used the IR-UV double resonance technique, where we fixed the IR laser at 1140 cm^{-1} and scanned the UV laser to check for additional conformations. The spectra are shown in Figure 7. The peak at 37574 cm^{-1} imposed on the broad background for mass 308 disappears upon IR irradiation, leaving the broad component of the spectrum practically unchanged. In the spectrum extracted from mass channel 264, the IR laser also does not have a strong influence.

The IR spectra, in combination with the IR-UV IDS spectra, confirm the presence of two dominant conformations of Z-Arg-OH in the gas phase. In mass channel 308, conformation A is found, with its fundamental $S_1 \leftarrow S_0$ transition at 37574 cm^{-1}. Two small peaks, at 37601 cm^{-1} and 37621 cm^{-1}, are also associated with conformation A, and are due to vibrational progression. The broad spectrum in mass channel 308 is attributed to conformation B. We conclude that conformation B readily dissociates and the entire spectrum in mass channel 264 is attributed to this conformation.

IR-UV Ion-Dip Spectra. IR absorption spectra are obtained from the two most abundant conformers present in the gas phase for Z-Arg-OH via IR-UV ion-dip spectroscopy (Figure 8). Both conformations show an absorption band around 1770 cm^{-1}, which is associated with the carboxylic acid C═O stretch vibration. Proton transfer from the carboxylic acid C-terminus to the guanidine group has been proposed to occur in the gas phase.[31] However, from the observation of the carboxylic acid C═O stretch, it is clear that such a zwitterionic structure is not formed under the present molecular beam conditions.

Further examination of the two spectra reveals that conformation A exhibits an absorption band around 1550 cm^{-1}, which is not present in conformation B. Peaks in this spectral range are associated with the NH bend vibration of the backbone and the guanidine side chain. A possible explanation is tautomerization of the guanidine side chain of Arg for conformer B, as presented in Figure 1c. Such tautomerization has been suggested to occur in the gas phase.[31−33,35] Another possible explanation for the absence of this peak in conformer B is that it has shifted to lower energies and now overlaps with the peak at 1500 cm^{-1}.

In the IR spectrum of conformation B, peaks are observed around 1610 cm^{-1}, 1685 cm^{-1}, 1725 cm^{-1} and 1765 cm^{-1}. These vibrations are associated with NH_2 scissor modes, C═N stretch modes, and the C═O stretch modes of the amide group and the C-terminus, respectively. Both the COOH- and NH_2- moieties are observed, indicating that the Z-Arg-OH molecule is intact when it is probed by the IR radiation and that dissociation likely occurs only upon UV excitation.

Conformational Search. A conformational search was performed for Z-Arg-OH in its canonical, tautomeric, and zwitterionic structure. Similar to Z-Glu-OH, Z-Arg-OH has multiple hydrogen bond donor and acceptor sites. Due to the flexibility of the Arg side chain, hydrogen bonds can be formed easily. Therefore, low energy structures tend to have more hydrogen bond interactions compared to Z-Glu-OH. The PES of Arg is known to exhibit many local minima at relatively low energies resulting in a complicated and relatively shallow PES,[33] which makes the conformational search more difficult; it is not impossible that low-energy structures are missed. The most stable structures found for Z-Arg-OH are presented in Figure 9, where AC, AT, and AZ refer to the canonical, tautomeric, and zwitterionic forms of Arg, respectively. The intramolecular interactions are listed in Table 4.

The electronic energies, ZPE-corrected energies, and Gibbs free energies at 300 K are given in Table 5. Z-Arg-OH shows the same general trends as Z-Glu-OH in the sense that the lowest energy structures at the MP2 level all exhibit dispersion interactions, emphasizing the importance of such interactions. The effect of the neglect of dispersion interactions by the B3LYP functional becomes evident when structures are reoptimized using MP2 or one of the DFT-D functionals. The distance between the phenyl group and a nearby hydrogen bond donor decreases substantially, and the structure changes noticeably. If, on the other hand, no dispersion interactions are present, the changes in the structure are quite small.

The lowest energy structure found for the zwitterionic form is 3.72 kcal/mol higher than the lowest energy structure found for the canonical form. In addition, several zwitterionic structures converged to canonical or tautomeric structures upon optimization at the DFT level. This suggests that zwitterionic Z-Arg-OH is not stable in the gas phase, in agreement with the fact that it has not been observed under our experimental conditions, as shown below. The energy difference between the lowest energy tautomeric structure and the

dx.doi.org/10.1021/jp3053339 | J. Phys. Chem. A 2013, 117, 1216−1227

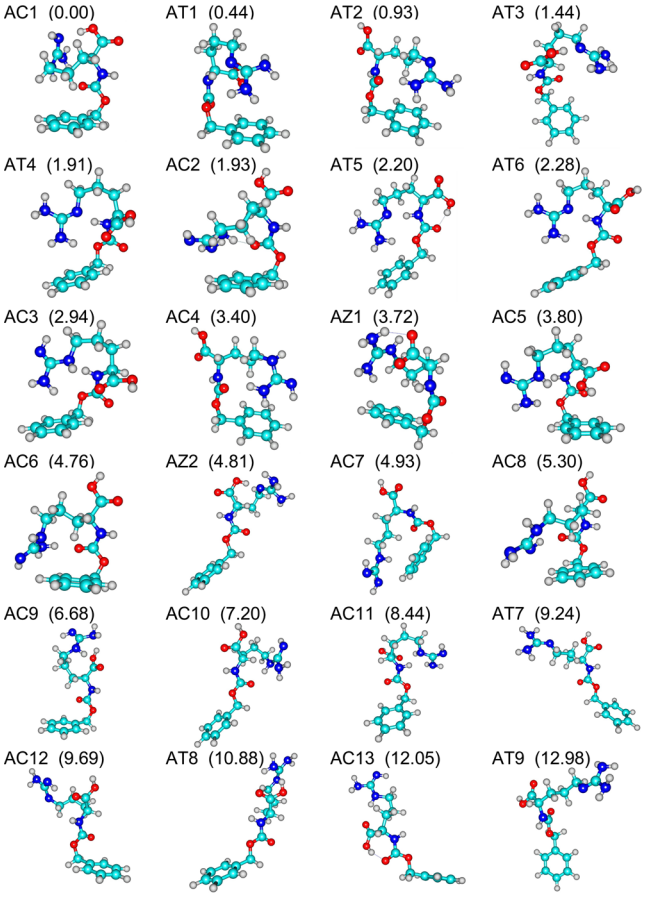


Figure 9. Lowest-energy structures of Z-Arg-OH, calculated at the MP2 level. Energies are given in parentheses in kcal/mol. The AC, AT, and AZ nomenclature refers to Arg in its canonical, tautomeric, and zwitterionic form, respectively.

lowest energy canonical structure is only 0.44 kcal/mol, so that one may expect to observe both forms in the experiment.

Structural Assignment. The structural assignment is predominantly based on the Amide I and II range of the spectrum above 1500 cm^{-1}, since the modes in this range are most sensitive to dispersion and hydrogen bond interactions. In Table 6 the experimental IR frequencies of conformation A and B as well as the computed frequencies of all canonical structures of Z-Arg-OH are listed. For conformer A, the guanidine NH bend vibration is located at 1540 cm^{-1}, making structures AC6, AC7, AC8, AC12, and AC13 unlikely due to their strongly red-shifted NH bend vibration. Structures AC1, AC2, and AC10 can be excluded based on the frequency of the amide C═O stretch vibration. In these structures, the amide C═O group is involved in a strong HB, causing its frequency to shift substantially to the red, which is not observed in the experimental spectrum of conformer A. Structure AC4 is excluded because its NH_2 scissor vibration is shifted to the blue by about 25 cm^{-1}. The red-shift of the carboxylic acid C═O stretch vibration excludes structures AC9 and AC11.

Two structures then remain as possible structural assignments for conformation A, being structures AC3 and AC5. Based only on the frequencies of the vibrations above 1500 cm^{-1}, it is not possible to distinguish between structures AC3 and AC5. Additionally, the fingerprint region does not allow us to assign either AC3 or AC5 to conformation A. Therefore, the energetics of the two structures are examined. The electronic

Table 4. Intramolecular Interactions for Structures of Canonical Z-Arg-OH[a]

structure	interactions
AC1	$H_{term} - N_{nh} + H_{nh2} - O_{am} + H_{am} - O_{term}$
AC2	$H_{am} - O_{term} + H_{\varepsilon} - O_{am}$
AC3	$H_{nh2} - phen + H_{\varepsilon} - O_{term} + H_{am} - N_{\varepsilon}$
AC4	$H_{nh2} - phen + H_{nh2} - O_{am} + H_{am} - O_{term}$
AC5	$H_{am} - N_{\varepsilon} + H_{term} - phen$
AC6	$H_{am} - O_{term} + N_{nh2} - O_{am}$
AC7	$H_{am} - O_{term}$
AC8	$H_{am} - O_{term} + N_{nh} - O_{am}$
AC9	$H_{am} - O_{term} + H_{\varepsilon} - O_{term} + H_{term} - N_{nh2}$
AC10	$H_{am} - O_{term} + H_{\varepsilon} - O_{am}$
AC11	$H_{e} - O_{term} + H_{am} - N_{\varepsilon}$
AC12	$H_{am} - O_{term} + H_{nh} - O_{term}$
AC13	$H_{term} - O_{am} + H_{\varepsilon} - O_{term}$
AT1	$H_{a1} - phen + H_{am} - N_{\varepsilon}$
AT2	$H_{a1} - phen + H_{am} - O_{term} + H_{a2} - O_{am}$
AT3	$H_{am} - O_{term} + H_{term} - N_{\varepsilon}$
AT4	$H_{a1} - phen + H_{am} - N_{\varepsilon}$
AT5	$H_{a1} - phen + H_{am} - N_{\varepsilon} + H_{term} - O_{am} + H_{a1} - O_{es}$
AT6	$H_{a1} - phen + H_{am} - N_{\varepsilon} + H_{a1} - O_{es}$
AT7	$H_{am} - O_{term}$
AT8	$H_{am} - O_{term} + H_{term} - N_{nh2a}$
AT9	
AZ1	$O_{term1} - H_{nh2} + O_{term1} - H_{\varepsilon} + H_{am} - O_{term2}$
AZ2	$O_{term1} - H_{nh2} + O_{term1} - H_{\varepsilon} + H_{am} - O_{term2}$

[a]For atom labels, see Figure 1.

energy of structure AC3 at the MP2 level is lower than that of structure AC5, although only by 0.86 kcal/mol. However, the Gibbs free energy at 300 K at the DFT-D levels of theory clearly favors structure AC3 as the assignment for conformer A. It should be noted, however, that structure AC3 is quite high in relative energy, (about 5−7 kcal/mol at the DFT or DFT-D levels and 2.94 kcal/mol at the MP2 level. In Figure SI-3, conformation A is compared with computed spectra for several structures (partly of lower relative energy), clearly showing that despite their lower energy, these structures do not match with the experimental spectrum of conformation A. Therefore, although slightly higher in energy, we conclude that conformation A is due to structure AC3.

In the IR spectrum of conformer B, the vibration around 1540 cm^{-1} due to the guanidine NH bending mode is clearly missing. If conformation B possesses the canonical Z-Arg-OH structure, the frequency of this guanidine NH-bending mode is apparently strongly red-shifted to at least near 1500 cm^{-1}. The only structures having this vibration near 1500 cm^{-1} are AC6, AC7, AC8, AC12, and AC13. However, these structures exhibit frequencies for the other modes in the diagnostic spectral range above 1500 cm^{-1}, in particular, the C-terminal C═O stretch, which deviate strongly from the experimental values for conformation B. It is therefore concluded that conformation B does not match with any of the canonical structures, and conformation B is likely to be assigned to one of the tautomerized structures.

In Table 7, the experimental frequencies in the diagnostic Amide I and II frequency range above 1500 cm^{-1} are compared to computed frequencies for tautomerized structures ATx (Figure 1c). The tautomerized structures possess two NH_2 moieties exhibiting two NH_2 scissor vibrations with closely spaced vibrational frequencies. However, only one band is

dx.doi.org/10.1021/jp3053339 | J. Phys. Chem. A 2013, 117, 1216−1227

Table 5. Electronic Energies (E_{elec}), ZPE-Corrected Energies (ZPE) and Gibbs Free Energies at 300 K (ΔG) for the lowest-Energy Structures of Z-Arg-OH[a]

	MP2	B3LYP			B97D			M05-2X		
struc	E_{elec}	E_{elec}	ZPE	ΔG	E_{elec}	ZPE	ΔG	E_{elec}	ZPE	ΔG
AC1	0.00	3.22	4.20	5.35	0.00	0.69	2.21	0.09	1.48	1.87
AT1	0.44	6.85	6.42	5.89	4.13	3.97	6.00	3.85	4.60	5.55
AT2	0.93	4.91	5.14	5.47	3.97	3.88	4.75	2.04	2.92	3.28
AT3	1.44	0.00	0.00	0.00	0.42	0.00	0.00	0.00	0.00	0.00
AT4	1.91	5.34	5.19	3.45	4.79	3.98	3.95	4.97	5.00	3.83
AC2	1.93	5.11	5.45	5.02	4.93	4.86	5.45	3.83	4.72	4.74
AT5	2.20	2.39	2.35	2.33	2.71	1.70	0.98	2.40	2.27	1.07
AT6	2.28	3.45	3.10	2.10	4.04	2.86	1.90	3.92	3.79	1.81
AC3	2.94	7.16	7.63	6.84	4.71	4.89	5.99	4.55	5.65	5.37
AC4	3.40	6.94	7.33	7.73	4.99	5.00	5.72	3.54	4.56	4.89
AZ1	3.72	3.59	3.62	3.65	3.05	3.08	3.45	1.59	2.93	2.14
AC5	3.80	8.58	9.09	8.13	7.10	7.65	9.82	6.72	7.77	6.95
AC6	4.76	6.52	7.13	6.84	7.07	7.42	8.37	5.37	6.60	6.90
AZ2	4.81	4.93	3.98	3.19	7.10	6.23	5.14	6.01	4.47	2.03
AC7	4.93	6.99	6.89	4.87	7.27	6.86	6.25	5.89	6.11	3.99
AC8	5.30	7.43	7.47	6.39	7.75	7.60	7.12	5.74	6.60	5.48
AC9	6.68	7.21	7.95	8.02	6.31	6.74	7.20	4.82	5.99	4.96
AC10	7.20	6.59	6.87	6.25	7.52	7.47	6.61	5.60	6.43	4.99
AC11	8.44	8.38	8.71	7.96	8.68	8.37	7.43	8.00	5.65	5.37
AT7	9.24	6.37	5.86	4.01	10.91	9.66	7.45	8.27	8.07	6.47
AC12	9.69	9.07	9.29	8.23	11.46	11.19	10.56	8.34	9.55	8.59
AT8	10.88	9.93	10.08	9.72	10.21	10.06	9.26	9.72	10.45	8.50
AC13	12.05	8.37	8.60	7.99	10.91	10.46	9.55	8.88	9.41	8.99
AT9	12.98	11.52	10.90	9.11	15.29	14.02	12.05	12.42	12.14	8.67

[a]Energies are given in kcal/mol. All calculations employed the 6-311+G(d,p) basis set.

Table 6. Experimentally Observed and Calculated Frequencies in the Amide Region for Several Conformations Z-Arg-OH[a]

	C=O stretch				NH bend	
	carboxylic acid	amide	C=N	NH_2 scissor	side chain	amide
CONF A	1779	1738	1685	1608	1540	1499
CONF B	1764	1727	1683	1609		1499
AC1	1778	1706	1651	1602	1542	1515
AC2	1778	1716	1676	1605	1545	1526
AC3	1773	1739	1688	1607	1541	1505
AC4	1780	1729	1669	1632	1451	1521
AC5	1786	1743	1690	1606	1544	1492
AC6	1779	1714	1682	1630	1443	1520
AC7	1781	1740	1682	1599	1506	1516
AC8	1782	1725	1686	1596	1497	1520
AC9	1744	1733	1688	1609	1541	1515
AC10	1775	1709	1673	1605	1556	1516
AC11	1762	1734	1682	1613	1542	1509
AC12	1777	1734	1682	1601	1516	1501
AC13	1780	1685	1674	1604	1508	1517

[a]Frequencies are given in cm^{-1}.

Table 7. Experimentally Observed Frequencies for Conformer B of Z-Arg-OH Compared to Calculated Frequencies for Several Conformations in the Amide Region[a]

	C=O stretch			NH_2 scissor		
structure	carboxylic acid	amide	C=N	1	2	NH bend
CONF B	1764	1727	1683	1609		1499
AT1	1771	1733	1693	1629	1616	1506
AT2	1777	1732	1698	1637	1602	1524
AT3	1774	1725	1692	1624	1617	1506
AT4	1794	1733	1708	1617	1608	1549
AT5	1789	1687	1707	1637	1627	1607
AT6	1794	1746	1709	1630	1622	1600
AT7	1775	1738	1711	1619	1605	1514
AT8	1757	1738	1723	1631	1608	1518
AT9	1787	1748	1711	1618	1605	1523

[a]Frequencies are given in cm^{-1}.

observed in the experimental spectrum, as the spacing between the NH_2 modes is not large enough to be resolved with FELIX.

Based on only the frequencies listed in Table 7, all structures except AT1 and AT3 can be excluded as possible structures since they have at least one frequency deviating strongly from the experimental value in the region above 1500 cm^{-1}. Although AT1 lies 1.00 kcal/mol lower in energy than AT3, the experimental frequencies in the 1500−1900 cm^{-1} region match better with the calculated frequencies for AT3. To confirm our tentative assignment of structure AT3, the fingerprint region of the spectrum is investigated in Figure 10, in which the experimental IR spectrum is compared to the computed spectra for structures AT1 and AT3 in the 1000−1600 cm^{-1} range. It is clear that the computed spectrum for AT3 agrees well with the experimental spectrum, particularly for the strong doublet band near 1500 cm^{-1}, the set of closely spaced weak bands between 1300 and 1450 cm^{-1}, and the sharp feature near 1200 cm^{-1}. The predicted frequency for the doublet feature between 1000 and 1100 cm^{-1} appears to be off by about 20 cm^{-1}. The computed spectrum for AT1 matches

dx.doi.org/10.1021/jp3053339 | J. Phys. Chem. A 2013, 117, 1216−1227

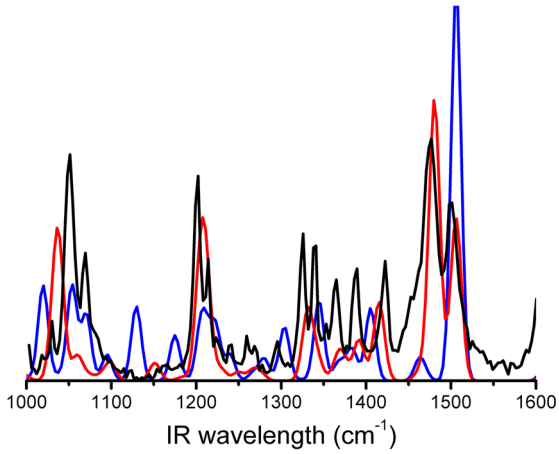


Figure 10. Experimental IR spectrum of conformation B of Z-Arg-OH (black) and computed spectra for structures AT1 (blue) and AT3 (red) in the fingerprint region.

not nearly as well with the experimental frequencies, and hence conformation B is assigned to structure AT3.

DISCUSSION

For Z-Arg-OH, the structures computed to possess the lowest energy in the gas phase were not observed, for which we can only provide speculative explanations. Possibly, there are deficiencies in the computational methods employed to establish the energetic ordering. For instance, close examination of the MP2-optimized structure AC1 reveals that there appears to be some interaction between the H atoms of the $-CH_2-$ group of the Arg side chain. The MP2 calculation may overestimate the stabilization energy due to this dispersion interaction, which has been reported previously.[65] From an experimental point of view, another possibility is that the electronically excited state of the lowest-energy structure is very short-lived. For Trp-Gly, Shemesh et al. proposed an excited state deactivation by an electron-driven excited-state proton-transfer process,[66] where radiationless excited-state decay to the ground state occurs via a conical intersection. In Trp-Gly, a strong HB is formed between the amide oxygen and the C-terminal OH-group. In Z-Arg-OH, structure AC2 has a similar strong HB interaction, i.e., a HB between the amide O_{am} and H_{ε}. It is thus possible that the excited state lifetime of structure AC2 is too short for detection by REMPI. Yet another possible explanation may be found in the process of expansion cooling that occurs in the molecular beam, in which molecules may become trapped in local minima, resulting in a nonthermal distribution of conformers. While it is not possible to pinpoint exactly which of the above effects gives rise to the current observations, we note that it is not uncommon to not observe the lowest energy conformation in this type of molecular-beam experiments. For instance, studies by the groups of de Vries,[3] Kleinermanns,[67] and Miller[68] report similar findings for other biomolecular systems.

Although the agreement between theory and experiment for frequencies above 1500 cm^{-1} is good, the agreement between theory and experiment is usually somewhat less favorable in the fingerprint region, certainly when dispersion interactions are present. The fact that vibrations in the fingerprint region are often strongly delocalized and perhaps more anharmonic and coupled may make them more difficult to calculate accurately. However, for conformation A of Z-Glu-OH, which we assigned to a structure without dispersion interactions, the agreement between the experimental and theoretical spectrum is good even in the fingerprint region. Moreover, comparing MP2- and B3LYP-optimized structures generally shows that structural differences are much larger when dispersion interactions are present. The intramolecular HB interactions, which are most important in determining the frequencies of the vibrations above 1500 cm^{-1}, are still retained, which may explain why the amide region is generally better reproduced than the fingerprint region when the calculations are performed with B3LYP.

The agreement between theoretical and experimental spectra is compared for the two conformations of Z-Arg-OH using the B3LYP, B97D, and M05-2X functionals (see Figure SI-4 of the Supporting Information). As for Z-Glu-OH, while providing somewhat worse relative energy values (using the MP2 values as the standard), the B3LYP calculations result in a much better frequency match than the B97D calculations. However, when dispersion interactions are present, the M05-2X reproduces the experimental spectrum more closely (see conformer B of Z-Arg-OH).

CONCLUSIONS

We have performed a conformational analysis of the gas phase dipeptide models Z-Glu-OH and Z-Arg-OH. Three conformations were experimentally found for Z-Glu-OH, two of them exhibiting a dispersion interaction between the side chain H_{term} and the phenyl ring. These structures correspond to the two lowest energy conformations found. The third conformation identified in the experiment is structure G5, which does not exhibit dispersion interaction. Structures G3 and G4, which are lower in energy than structure G5, are not observed. While structure G3 is likely not observed due to its high Gibbs free energy at 300 K, it is not clear why structure G4 is not observed.

For Z-Arg-OH, two conformations are found experimentally: one with a canonical Arg side chain structure, and one with a tautomerized guanidine structure. Zwitterionic structures were not observed, likely on account of their high relative energies. The assigned structures are not the lowest-energy structures found using the simulated annealing approach. Possibly, this is due to a very short electronically excited state lifetime, making the structure unobservable in the REMPI spectrum. Overestimation of dispersion interactions by MP2 may also explain why the experimentally observed structures do not correspond to the structures calculated to be lowest in energy.

Three of the five assigned structures exhibit dispersion interactions, emphasizing the importance of these interactions in peptides. Conventional DFT ignores these interactions, which is a major drawback of this frequently used and efficient computational method. Frequency calculations in the fingerprint region have to be interpreted with care when dispersion interactions are present. However, for the structurally very diagnostic Amide I and II region of the spectrum, B3LYP frequency calculations are very useful in the structural assignment of small biomolecules.

In addition, the present study provides further evidence for the deficiencies of B3LYP-based relative energy calculations for small biomolecules when dispersion interactions are present. This means that B3LYP energy calculations should be handled with care in the case of tyrosine, phenylalanine, and tryptophane, as well as when other aromatic moieties are incorporated in the studied biomolecular system. Energy calculations performed with DFT-D functionals perform

dx.doi.org/10.1021/jp3053339 | J. Phys. Chem. A 2013, 117, 1216−1227

much better, since dispersion interactions are included. However, HB interaction energies still appear to be overestimated.

In the critically diagnostic amide region of the spectrum, frequency calculations performed with the B3LYP functional reproduce the experimental spectrum more closely than the DFT-D functionals B97D and M05-2X. However, for conformations stabilized by dispersion interactions, the fingerprint region is better reproduced by M05-2X as compared to B3LYP.

■ ASSOCIATED CONTENT

Supporting Information

A comparison between the frequency calculations for the B3LYP, B97D, and M05-2X functionals for the assigned structures. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: a.rijs@science.ru.nl.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We gratefully acknowledge the expert assistance of the FELIX staff, in particular Drs. A.F.G. van der Meer and B. Redlich. This work is part of the research program of FOM, which is financially supported by the Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO). J.O. thanks the Stichting Physica for support.

■ REFERENCES

(1) Stryer, L.; Tymoczko, J. L.; Berg, J. M. *Biochemistry*, 5th ed. (International ed.); Freeman: New York, 2002.
(2) Steiner, T.; Koellner, G. *J. Mol. Biol.* **2001**, *305*, 535−557.
(3) Abo-Riziq, A.; Grace, L.; Crews, B.; Callahan, M. P.; van Mourik, T.; de Vries, M. S. *J. Phys. Chem. A* **2011**, *115*, 6077−6087.
(4) von Helden, G.; Compagnon, I.; Blom, M. N.; Frankowski, M.; Erlekam, U.; Oomens, J.; Brauer, B.; Gerber, R. B.; Meijer, G. *Phys. Chem. Chem. Phys.* **2008**, *10*, 1248−1256.
(5) Linder, R.; Seefeld, K.; Vavra, A.; Kleinermanns, K. *Chem. Phys. Lett.* **2008**, *453*, 1−6.
(6) Häber, T.; Seefeld, K.; Kleinermanns, K. *J. Phys. Chem. A* **2007**, *111*, 3038−3046.
(7) Häber, T.; Seefeld, K.; Engler, G.; Grimme, S.; Kleinermanns, K. *Phys. Chem. Chem. Phys.* **2008**, *10*, 2844−2851.
(8) Valdes, H.; Spiwok, V.; Rezac, J.; Reha, D.; Abo-Riziq, A. G.; de Vries, M. S.; Hobza, P. *Chem.—Eur. J.* **2008**, *14*, 4886−4898.
(9) Brenner, V.; Piuzzi, F.; Dimicoli, I.; Tardivel, B.; Mons, M. *J. Phys. Chem. A* **2007**, *111*, 7347−7354.
(10) Chin, W.; Piuzzi, F.; Dimicoli, I.; Mons, M. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1033−1048.
(11) Cocinero, E. J.; Stanca-Kaposta, E. C.; Gamblin, D. P.; Davis, B. G.; Simons, J. P. *J. Am. Chem. Soc.* **2009**, *131*, 1282−1287.
(12) Compagnon, I.; Oomens, J.; Bakker, J.; Meijer, G.; von Helden, G. *Phys. Chem. Chem. Phys.* **2005**, *7*, 13−15.
(13) James, W. H.; Baquero, E. E.; Choi, S. H.; Gellman, S. H.; Zwier, T. S. *J. Phys. Chem. A* **2010**, *114*, 1581−1591.
(14) James, W. H.; Buchanan, E. G.; Muller, C. W.; Dean, J. C.; Kosenkov, D.; Slipchenko, L. V.; Guo, L.; Reidenbach, A. G.; Gellman, S. H.; Zwier, T. S. *J. Phys. Chem. A* **2011**, *115*, 13783−13798.
(15) Rossi, M.; Blum, V.; Kupser, P.; von Helden, G.; Bierau, F.; Pagel, K.; Meijer, G.; Scheffler, M. *J. Phys. Chem. Lett.* **2010**, *1*, 3465−3470.
(16) Stearns, J. A.; Seaiby, C.; Boyarkin, O. V.; Rizzo, T. R. *Phys. Chem. Chem. Phys.* **2009**, *11*, 125−132.
(17) Gloaguen, E.; Pagliarulo, F.; Brenner, V.; Chin, W.; Piuzzi, F.; Tardivel, B.; Mons, M. *Phys. Chem. Chem. Phys.* **2007**, *9*, 4491−4497.
(18) Chin, W.; Piuzzi, F.; Dognon, J. P.; Dimicoli, I.; Mons, M. *J. Chem. Phys.* **2005**, *123*, 084301.
(19) Gloaguen, E.; Pollet, R.; Piuzzi, F.; Tardivel, B.; Mons, M. *Phys. Chem. Chem. Phys.* **2009**, *11*, 11385−11388.
(20) Fricke, H.; Schafer, G.; Schrader, T.; Gerhards, M. *Phys. Chem. Chem. Phys.* **2007**, *9*, 4592−4597.
(21) Rijs, A. M.; Ohanessian, G.; Oomens, J.; Meijer, G.; von Helden, G.; Compagnon, I. *Angew. Chem., Int. Ed.* **2010**, *49*, 2332−2335.
(22) Abo-Riziq, A.; Bushnell, J. E.; Crews, B.; Callahan, M.; Grace, L.; de Vries, M. S. *Chem. Phys. Lett.* **2006**, *431*, 227−230.
(23) Rijs, A. M.; Kabelac, M.; Abo-Riziq, A.; Hobza, P.; de Vries, M. S. *ChemPhysChem* **2011**, *12*, 1816−1821.
(24) Bush, M. F.; Prell, J. S.; Saykally, R. J.; Williams, E. R. *J. Am. Chem. Soc.* **2007**, *129*, 13544−13553.
(25) Fricke, H.; Schwing, K.; Gerlach, A.; Unterberg, C.; Gerhards, M. *Phys. Chem. Chem. Phys.* **2010**, *12*, 3511−3521.
(26) Biswal, H. S.; Loquais, Y.; Tardivel, B.; Gloaguen, E.; Mons, M. *J. Am. Chem. Soc.* **2011**, *133*, 3931−3942.
(27) Oomens, J.; Steill, J. D.; Redlich, B. *J. Am. Chem. Soc.* **2009**, *131*, 4310−4319.
(28) Bush, M. F.; O'Brien, J. T.; Prell, J. S.; Saykally, R. J.; Williams, E. R. *J. Am. Chem. Soc.* **2007**, *129*, 1612−1622.
(29) Heaton, A. L.; Moision, R. M.; Armentrout, P. B. *J. Phys. Chem. A* **2008**, *112*, 3319−3327.
(30) Prell, J. S.; Demireva, M.; Oomens, J.; Williams, E. R. *J. Am. Chem. Soc.* **2009**, *131*, 1232−1242.
(31) Julian, R. R.; Jarrold, M. F. *J. Phys. Chem. A* **2004**, *108*, 10861−10864.
(32) Ling, S. L.; Yu, W. B.; Huang, Z. J.; Lin, Z. J.; Haranczyk, M.; Gutowski, M. *J. Phys. Chem. A* **2006**, *110*, 12282−12291.
(33) Rak, J.; Skurski, P.; Simons, J.; Gutowski, M. *J. Am. Chem. Soc.* **2001**, *123*, 11695−11707.
(34) Schlund, S.; Muller, R.; Grassmann, C.; Engels, B. *J. Comput. Chem.* **2007**, *29*, 407−415.
(35) Skurski, P.; Gutowski, M.; Barrios, R.; Simons, J. *Chem. Phys. Lett.* **2001**, *337*, 143−150.
(36) Chapo, C. J.; Paul, J. B.; Provencal, R. A.; Roth, K.; Saykally, R. J. *J. Am. Chem. Soc.* **1998**, *120*, 12956−12957.
(37) Jones, C. M.; Bernier, M.; Carson, E.; Colyer, K. E.; Metz, R.; Pawlow, A.; Wischow, E. D.; Webb, I.; Andriole, E. J.; Poutsma, J. C. *Int. J. Mass. Spectrom.* **2007**, *267*, 54−62.
(38) Bouchoux, G.; Bimbong, R. N. B.; Nacer, F. *J. Phys. Chem. A* **2009**, *113*, 6666−6676.
(39) Meng, L.; Lin, Z. *Comput. Theor. Chem.* **2011**, *976*, 42−50.
(40) Sun, W.; Kinsel, G. R.; Marynick, D. S. *J. Phys. Chem. A* **1999**, *103*, 4113−4117.
(41) Rijs, A. M.; Kay, E. R.; Leigh, D. A.; Buma, W. J. *J. Phys. Chem. A* **2011**, *115*, 9669−9675.
(42) Johnston, M. V. *Trends Anal. Chem.* **1984**, *3*, 58−61.
(43) Johnson, P. M.; Otis, C. E. *Annu. Rev. Phys. Chem.* **1981**, *32*, 139−157.
(44) Oepts, D.; van der Meer, A. F. G.; van Amersfoort, P. W. *Infrared Phys. Technol.* **1995**, *36*, 297−308.
(45) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. *J. Chem. Theory Comp.* **2008**, *4*, 435−447.
(46) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. *Proteins: Struct., Funct., Bioinf.* **2006**, *65*, 712−725.
(47) Frisch, M. J.;Trucks, G. W.;Schlegel, H. B.;Scuseria, G. E.;Cheeseman, M. A. R. J. R.;Scalmani, G.;Barone, V.;Mennucci, B.;Petersson, G. A.;Nakatsuji, H., et al. *Gaussian 09*, revision A.1; Gaussian, Inc.: Wallingford, CT, 2009.
(48) www.chemcraftprog.com.
(49) Howard, A.; Ross, B. *AMBER User Manual, Appendix C: Parameter Development.*

(50) Jensen, F. *Introduction to Computational Chemistry*, 2nd ed.; John Wiley & Sons Ltd: Chichester, U.K., 2007.
(51) Schreiner, P. R. *Angew. Chem., Int. Ed.* **2007**, *46*, 4217−4219.
(52) Morgado, C.; Vincent, M. A.; Hillier, I. H.; Shan, X. *Phys. Chem. Chem. Phys.* **2007**, *9*, 448−451.
(53) Chin, W.; Dognon, J.-P.; Piuzzi, F. o.; Tardivel, B.; Dimicoli, I.; Mons, M. *J. Am. Chem. Soc.* **2005**, *127*, 707−712.
(54) Compagnon, I.; Oomens, J.; Meijer, G.; von Helden, G. *J. Am. Chem. Soc.* **2006**, *128*, 3592−3597.
(55) Rijs, A. M.; Compagnon, I.; Oomens, J.; Hannam, J. S.; Leigh, D. A.; Buma, W. J. *J. Am. Chem. Soc.* **2009**, *131*, 2428−2429.
(56) Herrebout, W.; Clou, K.; Desseyn, H. O.; Blaton, N. *Spectrochim. Acta, Part A* **2003**, *59*, 47−59.
(57) Herrebout, W. A.; Clou, K.; Desseyn, H. O. *J. Phys. Chem. A* **2001**, *105*, 4865−4881.
(58) Vondrášek, J.; Bendová, L.; Klusák, V.; Hobza, P. *J. Am. Chem. Soc.* **2005**, *127*, 2615−2619.
(59) Kumar, S.; Pande, V.; Das, A. *J. Phys. Chem. A* **2012**, *116*, 1368−1374.
(60) Gerlach, A.; Unterberg, C.; Fricke, H.; Gerhards, M. *Mol. Phys.* **2005**, *103*, 1521−1529.
(61) Leon, I.; Cocinero, E. J.; Millan, J.; Jaeqx, S.; Rijs, A. M.; Lesarri, A.; Castano, F.; Fernandez, J. A. *Phys. Chem. Chem. Phys.* **2012**, *14*, 4398−4409.
(62) Gloaguen, E.; de Courcy, B.; Piquemal, J. P.; Pilme, J.; Parisel, O.; Pollet, R.; Biswal, H. S.; Piuzzi, F.; Tardivel, B.; Broquier, M.; Mons, M. *J. Am. Chem. Soc.* **2010**, *132*, 11860−11863.
(63) Řeha, D.; Valdés, H.; Vondrášek, J.; Hobza, P.; Abu-Riziq, A.; Crews, B.; de Vries, M. S. *Chem.—Eur. J.* **2005**, *11*, 6803−6817.
(64) Shubert, V. A.; Zwier, T. S. *J. Phys. Chem. A* **2007**, *111*, 13283−13286.
(65) van Mourik, T. *J. Chem. Theory Comput.* **2008**, *4*, 1610−1619.
(66) Shemesh, D.; Hättig, C.; Domcke, W. *Chem. Phys. Lett.* **2009**, *482*, 38−43.
(67) Hunig, I.; Kleinermanns, K. *Phys. Chem. Chem. Phys.* **2004**, *6*, 2650−2658.
(68) Choi, M. Y.; Miller, R. E. *J. Am. Chem. Soc.* **2006**, *128*, 7320−7328.