

The Gas-Phase Dipeptide Analogue Acetyl-phenylalanyl-amide: A Model for the Study of Side Chain/Backbone Interactions in Proteins

Wutharath Chin,[†] Michel Mons,^{*,†} Jean-Pierre Dognon,[†] Reinard Mirasol,[‡] Gregory Chass,^{†,‡} Iliana Dimicoli,[†] François Piuze,[†] Patrick Butz,[§] Benjamin Tardivel,[†] Isabelle Compagnon,^{||} Gert von Helden,[¶] and Gerard Meijer[¶]

Laboratoire Francis Perrin (URA CNRS 2453), Service des Photons, Atomes et Molécules, Centre d'Etudes de Saclay, Bât. 522, 91191 Gif-sur-Yvette Cedex, France, Global Institute of Computational Molecular and Materials Sciences (GIOCOMMS), 158-2 Major Street, Toronto, Ontario M5S 2L2, Canada, PTCL, South Parks Road, University of Oxford, OX3 3QZ, United Kingdom, FOM Institute for Plasma Physics "Rijnhuizen", Edisonbaan 14, 3439 MN Nieuwegein, The Netherlands, and Fritz-Haber-Institut der Max-Planck-Gesellschaft, Faradayweg 4-6, 14195 Berlin, Germany

Received: May 7, 2004; In Final Form: November 15, 2004

The issue of the influence of the side chain/backbone interaction on the local conformational preferences of a phenylalanine residue in a peptide chain is addressed. A synergetic approach is used, which combines gas-phase UV spectroscopy as well as gas-phase IR/UV double-resonance experiments with DFT and post Hartree–Fock calculations. *N*-Acetyl-Phe-amide was chosen as a model system for which three different conformers were observed. The most stable conformer has been identified as an extended β_L conformation of the peptide backbone. It is stabilized by a weak but significant NH– π interaction bridging the aromatic ring on the residue (i) with the NH group on residue (i+1), with the aromatic side chain being in an anti conformation. This stable conformation corresponds to the common NH(i+1)–aromatic(i) interaction encountered in proteins for the three aromatic residues (phenylalanine, tyrosine, and tryptophan), which illustrates the relevance of gas-phase investigations to structural biology issues. The two other less abundant conformers have been assigned to two γ -folded backbone conformations that differ by the orientation of the side chain. In all cases, the IR data provided spectroscopic fingerprints of these interactions. Finally, the strong conformational dependence of the fluorescence yield found for *N*-acetyl-Phe-amide illustrates the role of the environment on the excited-state dynamics of these species, which is often exploited by biochemists to monitor protein structural changes from tryptophan lifetime measurements.

1. Introduction

Amide–aromatic interactions have been early recognized in biology as significant interactions that are able to stabilize specific backbone secondary structures in proteins.^{1–3} Recent data mining investigations⁴ have shown that the Ar(i) ← NH(i+1) interaction acting between an aromatic residue *i* and its neighbor residue *i*+1 (Figure 1) is the most abundant amide–aromatic interaction. Correlatively, those residues, that exhibit such an interaction, have a marked preference for an anti (trans) orientation of the aromatic side chain relative to the peptide backbone. Quantum chemists have investigated the role of such an amide–aromatic interaction by studying model molecular complexes. In particular, the binding and the most favorable geometries of an amide molecule relative to an aromatic ring, the so-called *parallel* and *pointing* conformations, have been studied.^{5,6} In model systems such as benzene–formamide, interaction energies on the order of 4 kcal/mol were found which are mainly due to dispersion and high-order electrostatic interactions such as quadrupole (benzene)–dipole (amide) interactions.

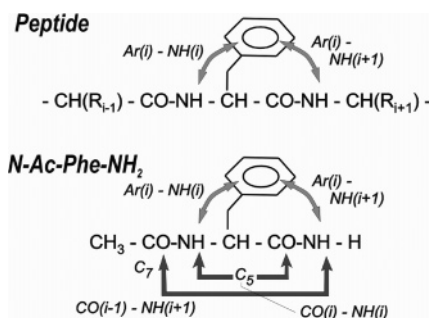


Figure 1. Upper panel: possible side chain/backbone local interactions for a phenylalanine residue within a peptide chain. Lower panel: schematic view of the several interactions controlling the conformational preferences of *N*-Ac-Phe-NH₂.

Gas-phase studies on aromatic amino acids with protected N- and C-termini provide an original approach to understand the role of local backbone/backbone and backbone/side chain interactions that goes beyond investigations of benzene–amide molecular complexes. In protected amino acids (Figure 1), among the nine possible conformations (according to Ramachandran terminology⁷ based on the φ and ψ dihedral angles), folded γ conformations, which are similar to the central part of the γ -turns encountered in biology,⁸ are the only remarkable H-bonded structures expected. These γ -turns consist of three

[†] Laboratoire Francis Perrin.

[‡] Global Institute of Computational Molecular and Materials Sciences.

[§] University of Oxford.

^{||} FOM Institute for Plasma Physics.

[¶] Fritz-Haber-Institut der Max-Planck-Gesellschaft.

amino acid residues, where a H-bond bridges the backbone CO of residue ($i-1$) with the NH group of residue ($i+1$) via a seven-atom ring (referred to as C_7). Recently, Lavrich et al. have provided spectroscopic evidence for the formation of such a C_7 structure for the *N*-Ac-Ala-NH-Me dipeptide analogue in the gas phase.⁹ From the theoretical side, numerous ab initio calculations have shown that such C_7 H-bonded conformations are the most stable forms in the gas phase for residues such as glycine (Gly) or alanine (Ala).^{10–12} In particular, the extended backbone structures (labeled β_L) that correspond to the β -strands in proteins are more than 1.5 kcal/mol less stable, mainly because of the weakness of the CO(i)–NH(i) close contact attractive interaction (referred to as a C_5 interaction; Figure 1). The issue of the backbone/aromatic side chain interaction was also addressed theoretically,^{13–17} but so far experimental studies remain sparse. A few groups investigated modified amino acids with either tryptophan (Trp) or phenylalanine (Phe) as an aromatic side chain chromophore using UV and IR spectroscopy.^{18–23} These studies focused on the conformational landscape of dipeptide analogues and provided evidence for both folded (γ) and extended (β_L) backbone conformations.

The goal of the present study is to investigate the role of the backbone/side chain interactions on the local conformational preferences of the peptide backbone. The small protected amino acid *N*-acetyl-phenylalanyl-amide (Figure 1), short: *N*-Ac-Phe-NH₂ or NAPA, is used as a model system. The study is conducted in the gas phase to focus on intrinsic properties, free of any solvent effect. Two complementary approaches are used: quantum calculations and optical spectroscopic methods.

Thanks to progress in computer technology, quantum calculations can nowadays be performed on small model peptides, allowing us to explore their conformational landscape theoretically.^{10–17,19,23–31} When investigating weakly polar interactions such as backbone/side chain interaction, it is necessary to include electron correlation; consequently, only a few computational studies have been devoted to this issue. The goal of the accompanying paper³² is to give an accurate theoretical description of the here studied peptide.

On the experimental side, gas-phase spectroscopy on biological molecules made significant progress over the past decade. The combined use of laser-desorption with supersonic expansion methods^{33–35} leads to an efficient cooling that nevertheless enables conformational relaxation toward the lowest energy structures of the system. Resonant two-photon ionization (R2PI) spectroscopy allows then for a sensitive detection and in some cases it even allows to distinguish the UV spectral signatures of several conformers populated in the jet.^{36–39} The use of double-resonance techniques, such as gas-phase resonant ion-dip infrared (RID-IR) spectroscopy,³⁷ can then lead to conformer specific IR spectra and in some cases to a complete characterization of the H-bonding. Following such a strategy, Gerhards and co-workers published very recently a study of *N*-Ac-Phe-NHMe (short: NAPMA), a dipeptide model system that is very similar to that studied in the present experiment.²³ Zwier and co-workers investigated another dipeptide model system, *N*-Ac-Trp-NHMe (short: NATMA) as well as its nonmethylated derivative (*N*-Ac-Trp-NH₂, short: NATA).¹⁹ In both cases, detailed assignments of the observed conformers in terms of intramolecular H-bonding are proposed; the NH– π interaction, however, was not discussed thoroughly.

The present paper presents experimental data on NAPA together with a comparison to theoretical results. The experimental strategy combines both UV and IR spectroscopy. UV spectroscopy of the Phe chromophore is sensitive enough to

the phenyl environment to spectrally distinguish between gas-phase conformers, and the specific UV signature of each conformer allows us to estimate their abundance in the supersonic expansion. Conformer selective UV excitation enables us then to record IR spectra of each conformer using the RID-IR technique. Two spectral regions have been investigated: the amide A (NH stretch: 2.8–3.2 μ m) and the amide I and II regions (CO stretch/NH bend: 5–8 μ m). The first region is very sensitive to the presence of H-bonds and leads to a straightforward, qualitative diagnosis about the H-bonding of each conformer observed, as it directly gives the number of intramolecular H-bonds in the molecule. The combination of the two spectral regions then enables us to distinguish among several possible backbone arrangements. To derive information in particular concerning the location of the H-bonds in the molecule, the amino acids used in the study are amidated on the C terminus, even though methyl-amidated peptides have a closer resemblance to a peptidic linkage. We shall see that for such amidated peptides, IR spectroscopy brings additional information about the involvement of the amino group of the C-end amide in the H-bonding.

The experimental gas-phase conformational abundances and the measured IR spectra will be compared to the relevant theoretical results that are presented in detail in the accompanying paper.³² In addition, the NAPA data will be compared to theoretical results on simpler related systems that do not have an aromatic side chain (*N*-Ac-Gly-NH₂ and *N*-Ac-Ala-NH₂) and the results will be discussed in terms of aromatic/backbone interaction effects.

2. Experimental Section

The R2PI and RID-IR spectroscopy experiments were carried out on two molecular beam/laser desorption setups located in Saclay^{35,40} and in Nieuwegein.³⁹ NAPA was purchased from Sigma and used without further purification. In Saclay, the peptide powder was mixed and pressed with spectroscopic-grade graphite powder to produce a pellet, which is fixed below the pulsed gas nozzle at a distance of ca. 2 mm. The second harmonic of a Nd:YAG laser was sent on the pellet surface from the end of a 550- μ m diameter transport optical fiber located above it at ca. 3 mm. In Rijnhuizen, the sample powder is deposited on a graphite surface and desorbed by the mildly focused output of the fundamental output of a Nd:YAG laser.³⁹ Argon was used as carrier gas (typically 3–5 bar backing pressure).

In Saclay, the infrared radiation (2.8–3.2 μ m) was generated as the idler of a Nd:YAG-pumped LiNbO₃ OPO (Euroscan/BM Industries). A line width of 1 cm^{-1} was achieved by an intracavity IR Pérot-Fabry plate. The spectral scan of the OPO, controlled by a LabView-operating computer, was carried out by tilting the plate, simultaneously with the LiNbO₃ crystal. A typical energy of 1–3 mJ per pulse could be obtained, except in the 3470–3500 cm^{-1} energy range where the OPO emission vanishes because of an intense IR absorption in the crystal. The IR light was mildly focused ($f = 500$ mm) onto the supersonic jet between the extraction plates of a time-of-flight mass spectrometer. Peptide molecules were photoionized via an R2PI scheme using the output of an excimer-pumped (Lambda-Physik EMG 200) dye laser (Lambda-Physik FL 2002E) after a time delay of 100 ns.

In Nieuwegein, mid-IR radiation was generated by the free-electron laser FELIX. Its output consists of microsecond-long bursts (macropulses) of light that come at a repetition rate of 5 Hz. Each macropulse consists of a train of picosecond-long

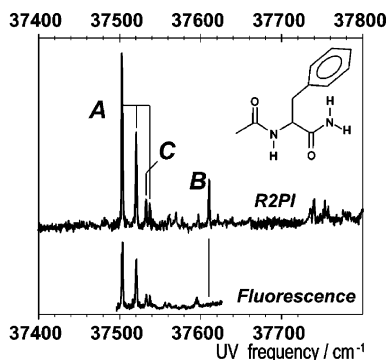


Figure 2. Mass-selective one-color R2PI spectrum (top) and excitation fluorescence spectrum (bottom) of the model peptide *N*-Ac-Phe-NH₂ in the origin region of the $S_1 \leftarrow S_0$ transition of the phenyl chromophore. Conformational features are identified and labeled according to the isomers deduced from the resonant ion-dip IR experiments.

micropulses, with a spacing of one nanosecond between micropulses. In the present experiment, FELIX was optimized for an intense output in the 5–8- μm region (typically 50 mJ per pulse); the line width employed was typically 1% fwhm, corresponding to $\sim 15 \text{ cm}^{-1}$ in the region scanned. The unfocused IR beam crossed the molecular jet and is aligned counterpropagating to the focused UV photoionization beam of a Nd:YAG-pumped dye laser (Spectra Physics GCR-100/PDL3), operating at 10 Hz.

Both R2PI spectra and RID-IR spectra were measured. In both cases, the mass-selected ion signal was acquired on a LeCroy digital oscilloscope, averaged over 100 laser shots, transferred to a PC, and measured as a function of laser frequency. In Nieuwegein, a homemade synchronization interface between FELIX and the UV laser enables to average simultaneously “IR on” and “IR off” signals on the oscilloscope. In Saclay, this is done sequentially, using a PC controlled shutter.

The setup in Saclay is also fitted with a fluorescence detector, described elsewhere,^{40,41} which includes as a filter a broad bandwidth monochromator followed by photomultiplier (Hamamatsu). The fluorescence excitation spectrum of NAPA was recorded under the same jet conditions as its R2PI spectrum.

3. Results

3.1. UV Spectroscopy. The frequency positions of the bands of the chromophore (the phenyl ring) in the UV spectrum are sensitive to its environment. However, it is difficult to extract structural information from these data. The Franck–Condon (FC) activity on the other hand is more interesting, as it provides information on the structural changes between ground and excited electronic states and, hence, on ring/backbone interactions.

The R2PI spectrum of mass-selected NAPA has been obtained in the region of the origin of the Phe chromophore $S_1 \leftarrow S_0$ transition. Because of possible pollution of the NAPA mass channel by fragmented peptide hydrates or peptide dimers, the composition of the species responsible for the weak signals was carefully checked by monitoring the arrival times after the opening of the pulsed valve.^{42,43}

The UV spectrum of NAPA (Figure 2) is rich and, with the help of the IR/UV measurements, three spectral systems labeled A, B, and C can be distinguished. The most intense system A is composed of a low-frequency (18 cm^{-1}) vibrational progression accompanied by much weaker bands. The two other conformers B and C are less populated than the A conformer

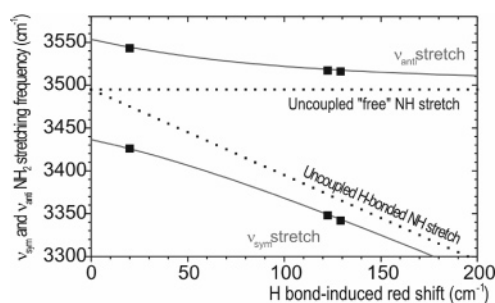


Figure 3. Evolution of the NH stretch frequencies of a H-bonded terminal amide group of NAPA as a function of the H-bond strength (expressed in terms of red-shift) as predicted from a phenomenological vibrational coupling model.⁴⁶ When being uncoupled and unperturbed, the two modes are degenerate. They split and one component shifts to the red when the corresponding H atom is involved in H-bonding. Vibrational coupling between the two modes causes splitting into a “symmetric” and an “antisymmetric” component. The parameters of the model (unperturbed frequency and vibrational coupling intensity) have been adjusted to fit the NH stretches of the three conformers of *N*-Ac-Phe-NH₂. The abscissa provides the strength of the perturbation on the bonded NH.

and are each characterized by a single weak UV band to the blue of the A origin.

Fluorescence excitation spectra carried out under the same conditions exhibit basically the same shape as the R2PI spectrum, apart from the fact that the UV band of species B is missing. Fluorescence decay measurements yield similar lifetimes for the A ($62 \pm 2 \text{ ns}$) and C ($42 \pm 2 \text{ ns}$) bands, values comparable to the lifetime of toluene on its origin band (76 ns ⁴⁴). The absence of a significant fluorescence signal for the B species indicates a dramatic conformation-dependent quenching effect in which the fluorescence quantum yield of B is lower than that of A or C by at least one order of magnitude.

3.2. IR Spectroscopy and First Structural Assignment. *The Amide A Region.* Structural information can be obtained from IR spectroscopy because of the sensitivity of band positions to the environment of the NH or CO chromophore. This is especially the case in the amide A region where the frequencies of NH stretch modes typically range from 3420 to 3550 cm^{-1} ,⁴⁵ depending upon the residue and its environment. With the exception of NH₂ amino groups, NH oscillators on different residues are usually not strongly coupled. A red-shift of an amide A mode will be the fingerprint of an intramolecular H-bond: the stronger the bond, the larger the shift as a general rule. H-bonding to a medium acceptor like O, for example, will cause 100 – 200 cm^{-1} red-shifts.⁴⁵

The use of peptides that are amidated (instead of methylamidated) on the C terminus allows us to obtain additional information. When the NH₂ group is not involved in H-bonding, the NH stretches are quasi degenerate and vibrational coupling will cause a splitting of about 120 cm^{-1} into a symmetric and an antisymmetric component.⁴⁵ When one of the H-atoms is involved in a H-bond, the original symmetry and degeneracy are broken, leading to an increased splitting of the modes. This can be described by a simple model of a system of two levels interacting by vibrational coupling.⁴⁶ The splitting is then controlled by the strength of the H-bond and the vibrational coupling constant, assumed to be independent of the H-bond. As a result the originally “symmetric” low-frequency component is the most perturbed (Figure 3); its frequency is red-shifted and this red-shift increases with the H-bond strength. Correlated, the frequency of the “antisymmetric” component decreases only slightly as the H-bond strength increases, approaching the

TABLE 1: Experimental UV Origins and IR Frequencies of the Three Conformers Observed for *N*-Ac-Phe-NH₂ Assigned by Comparison with the Calculated IR NH and CO Stretch Frequencies (cm⁻¹) and Intensities (Given between Brackets in km/mol) of the Relevant Most Stable Conformations^a

species	UV	IR								assignment see section 4
	frequency/shift (cm ⁻¹)	stretches				bends				
		NH _{Phe}	sym. NH ₂	anti. NH ₂	CO _{Ac}	CO _{Phe}	NH _{Phe}	sym. NH ₂	anti. NH ₂	
		experimental								
conf. A	37503	3434	3426	3543	1694	1725	1492	1592	1393	β_L (a)
conf. B	+107	3438	3342	3516	1682	1718	1503	1589		γ_L (g+)
conf. C	+30	not obs.	3348	3517	1680	1719	1494	1580		γ_L (g-)
		theoretical								
β_L (a)		3427 (49)	3422 (80)	3546 (63)	1695 (358)	1725 (267)	1495 (268)	1601 (71)	1384 (158)	
γ_L (g+)		3443 (96)	3342 (134)	3517 (81)	1685 (210)	1725 (408)	1504 (191)	1595 (374)		
γ_L (g-)		3470 (28)	3348 (127)	3512 (75)	1687 (188)	1735 (416)	1504 (238)	1597 (185)		
γ_L (a)		3466 (25)	3373 (84)	3513 (66)	1690 (183)	1735 (379)	1498 (216)	1601 (125)		
γ_D (g-)		3484 (26)	3337 (166)	3517 (93)	1690 (128)	1724 (460)	1520 (149)	1592 (222)		

^a UV origins of the secondary conformers are given relative to main species A. Harmonic calculated frequencies, taken from ref 32, have been scaled by 0.960, 0.980, and 0.970 in the Amide A, I, and II regions respectively (see section 4). The experimental precisions are ± 10 cm⁻¹ for the absolute UV frequencies (conformer A), ± 1 cm⁻¹ for the UV shifts, ± 2 and ± 5 cm⁻¹ for the absolute IR frequencies in the amide A and I/II regions, respectively.

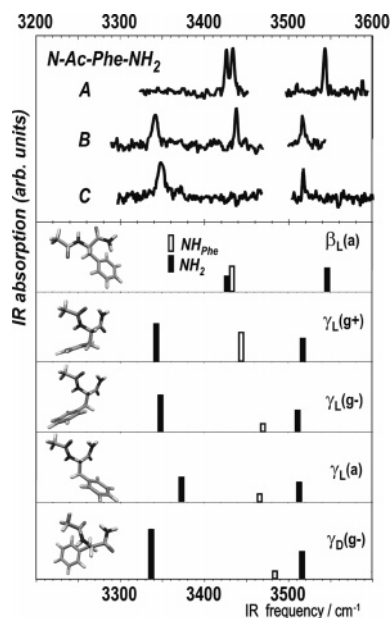


Figure 4. Top panel: resonant ion-dip IR spectra of the three conformers of *N*-Ac-Phe-NH₂, in the spectral region of the NH stretches (amide A); bottom panel: calculated stick spectra of the five lowest energy conformations of NAPA, as obtained from DFT B3LYP/6-31+G(d) calculations together with the corresponding DFT optimized structures.³² The frequency positions in the stick spectra are scaled by a factor of 0.960.

asymptotic frequency of a NH stretch free from vibrational coupling. These properties provide us therefore with an efficient indicator of the involvement of the terminal amino group in a H-bond.

The IR spectrum of the main conformer A exhibits three IR bands, all located in the 3420–3550 cm⁻¹ spectral region, corresponding to free or nearly free NH stretch modes (Figure 4 top panel, Table 1). The energy difference between the bands at extreme positions is 117 cm⁻¹, close to the typical splitting of the two NH components of the amino group in unsubstituted amides.⁴⁵ These bands are therefore assigned to a free (or nearly free, see below) terminal amino group of the NAPA molecule.

The third band in the center, corresponding to the NH moiety on Phe, is at 3434 cm⁻¹, slightly red-shifted compared to a free NH stretch, but not compatible with a classical H-bond.

The IR spectra of species B and C (Figure 4 top panel, Table 1) differ from that of A by the presence of a red-shifted broad band in the 3340–3350 cm⁻¹ energy range and by having a much weaker red-shift of the blue most band (3520 cm⁻¹ region). The splitting between the outermost lines in the spectrum of structure B and C of 174 and 169 cm⁻¹, respectively, together with the overall red-shift is indicative for a terminal amino group engaged in significant intramolecular H-bonding. In B, a third band is located around 3438 cm⁻¹. For C, a mode corresponding to the third NH oscillator is not observed and most likely located in the gap of the OPO tuning range.

When one compares the above results to the H-bonding analysis in Figure 1, one can conclude that the two minor conformers B and C have their terminal NH₂ engaged in a H-bond in a folded C₇ backbone with a γ turnlike structure. The two conformers could differ by the type of turn (inverse or direct) or by the orientation of their side chain. In contrast, the major conformer A does not have any strong NH \rightarrow CO H-bond and consequently no γ -like structure.

The Amide I and II Regions. The 5–8- μ m region contains the CO stretch and NH in-plane bending modes. The NH bend vibrations are often tightly coupled with CH bends and C–N stretches, and the assignment of the spectra in this region is generally more complicated. However, because these couplings are often conformation-dependent, the spectral patterns observed turn out to be useful information for assignment.

The 5–8- μ m IR spectrum of conformer A (Figure 5, Table 1) exhibits four main bands, two in the 1620–1700 cm⁻¹ range (amide I) and two in the 1350–1600 cm⁻¹ region (amide II). The IR spectra of species B and C (Figure 5, Table 1) are very similar and resemble that of A, especially in the amide I region, apart from a change in the band intensity pattern. In the amide II region, one band (~ 1500 cm⁻¹) is present in the spectrum of all three species whereas the second most intense in A, at 1393 cm⁻¹, does not show up in the spectra of B and C.

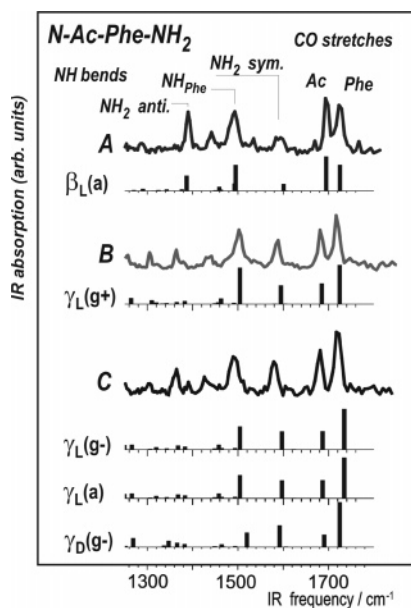


Figure 5. Resonant ion-dip IR spectra of the three conformers of *N*-Ac-Phe-NH₂, in the amide I (CO and acetyl stretches) and II (NH bends) spectral regions together with the stick spectra of the five lowest energy conformations of NAPA, as obtained from DFT B3LYP/6-31+G(d) calculations.³² The frequencies in the stick spectra are scaled by a factor of 0.980 and 0.970 in the amide I and II regions, respectively.

TABLE 2: Energetics of the Lowest Conformations of the *N*-Ac-Phe-NH₂, *N*-Ac-Gly-NH₂, and *N*-Ac-Ala-NH₂ Model Dipeptides Taken from Ref 32, B3LYP/6-31+G(d) and MP2/6-31+G(d) Relative Energies (ΔE_{B3LYP} and ΔE_{MP2}), Zero-Point Vibrational Relative Energies (ΔZPE) at the B3LYP/6-31+G(d) Level of Theory (scaling factor 0.960) and ZPE-corrected B3LYP and MP2 Relative Energies ($\Delta E_{\text{B3LYP}} + \Delta ZPE$ and $\Delta E_{\text{MP2}} + \Delta ZPE$) (in kcal/mol) for All Three Peptides, with the β_L Conformation as the Energy Reference

	ΔE_{B3LYP}	ΔE_{MP2}	ΔZPE	$\Delta E_{\text{B3LYP}} + \Delta ZPE$	$\Delta E_{\text{MP2}} + \Delta ZPE$
<i>N</i> -Ac-Phe-NH ₂					
β_L (a)	0.00	0.00	0.00	0.00	0.00
γ_L (g+)	0.18	0.05	0.51	0.69	0.56
γ_L (g-)	0.35	0.72	0.21	0.56	0.93
γ_L (a)	0.50	1.06	0.25	0.75	1.31
γ_D (g-)	1.74	1.20	0.50	2.24	1.70
<i>N</i> -Ac-Gly-NH ₂					
γ_{LD}	-0.92	-2.23	0.84	-0.08	-1.39
β_L	0.00	0.00	0.00	0.00	0.00
<i>N</i> -Ac-Ala-NH ₂					
γ_L	-1.19	-1.72	0.45	-0.74	-1.27
β_L	0.00	0.00	0.00	0.00	0.00

4. Refined Assignment

The IR data collected for the three species observed are now compared to the results from calculations (Tables 1 and 2) reported in the accompanying paper³² to derive a more precise structure assignment.

The most stable forms calculated that are expected to be populated under jet cooling conditions (Figure 4) are listed below, labeled according to the Ramachandran terminology for the backbone conformations as well as to the classical (a, g+, and g-, for anti, gauche+, and gauche-) nomenclature for the side chain orientation ($C(O)-C^{\alpha}-C^{\beta}-C^{\gamma}$ side chain dihedral angle): (i) the extended β_L (a) conformer related to the β strandlike conformation in proteins, (ii) the three-folded γ_L conformations, related to the so-called inverse γ -turns in proteins (also referred to as the C_{7eq} conformation in the literature,

because of the equatorial position of the residue compared to the backbone fold), and, (iii) the folded γ_D (g-) form, related to the direct γ -turns in proteins (referred to as the C_{7ax} conformation, because of the axial position of Phe compared to the backbone fold). This last form is about 1 kcal/mol higher in energy than the four other forms (Table 2), but it has nevertheless been considered in our analysis as it represents the other low-energy topological basin (γ_D) able to account for intramolecular H-bonding. The other forms and basins that are either even higher in energy³² or not able to account for H-bonding (or both) have not been considered.

Harmonic frequencies and IR absorption intensities (DFT B3LYP/6-31+G(d)) have been obtained for each conformation and are compared to the experiment.³² To match the experiment, scaling factors had to be introduced. In agreement with what is anticipated from qualitative spectroscopic arguments alone, the amide A data of the A conformer can only be explained by the β_L (a) extended conformation. The data for the B conformer, in particular the NH band at 3438 cm⁻¹, can only be reproduced by the γ_L (g+) species. Starting from this assignment, different scaling factors are used for amide A, I, and II spectral regions. Figures 4 and 5 show that a fair agreement is found between experimental and theoretical frequencies for A and B in these three regions. Discrepancies are less than 10 cm⁻¹ in all cases.

This good agreement suggests using theoretical spectra together with the above-mentioned scaling factors to assign the structure of the minor conformer C. Figure 4 show that the best agreement between experiment and theory is with either the γ_L (g-) or a γ_D (g-) conformation. At this stage, the amide I and II data appear as a decisive source of information, since the band intensity pattern of the γ_D (g-) form (Figure 5) does not fit the experiment in contrast to γ_L (g-) or γ_L (a). We can thus assign the C form to the γ_L (g-) conformation of NAPA.

5. Discussion

5.1. Energetics of the NAPA Forms. The most abundant form of NAPA in the experiment is obviously the extended β_L (a) form, followed by the γ_L (g+) and γ_L (g-) forms. Despite the numerous minima in the potential energy landscape of NAPA,³² only these three conformations are observed under jet conditions, which illustrates the efficiency of the structural relaxation processes during the expansion. The coexistence of the three structures should be considered as a consequence of very similar energies for these three species together with relatively low isomerization barriers (which can be crossed easily and for a significant period of time during the expansion).

The comparison with relative energies obtained from quantum chemistry calculations³² provides a stringent test of the methods and level of theory used (Table 2). DFT B3LYP/6-31+G(d) and MP2/6-31+G(d) calculations corrected for the zero-point vibrational energy (ZPE) seem to provide results that are in fair agreement with experimental abundances. Both methods show the same trends: the β_L (a) conformation is the most stable one, followed by the three γ_L conformations, which lie within a 1 kcal/mol energy range, starting typically 0.5 kcal/mol higher in energy (Table 2). The lowest energy γ_D structure is significantly higher in energy, certainly higher than the lowest γ_L conformation (Table 2) as well as the other β_L conformations (g+ and g-).³² The differences in relative stabilities and ordering within the γ_L family when comparing the two types of calculations give an indication of the order of magnitude of the precision achieved by the theoretical methods.

Comparison with calculated energies of related species, which do not exhibit an aromatic/side chain interaction such as *N*-Ac-

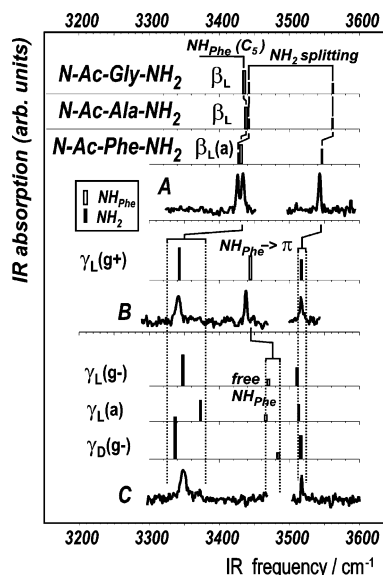


Figure 6. Comparison of the calculated IR stick spectra (amide A region) of the most stable conformers of *N*-Ac-Phe-NH₂, obtained from DFT B3LYP/6-31+G(d) (scaled) calculations, with calculated spectra of the extended backbone β_L conformations of *N*-Ac-Gly-NH₂, *N*-Ac-Ala-NH₂.³² Experimental resonant ion-dip IR data of the three conformers of *N*-Ac-Phe-NH₂ are also given for comparison.

Ala-NH₂ and *N*-Ac-Gly-NH₂, is quite informative (Table 2). For these molecules, the folded conformations (γ_L and γ , respectively) are more stable than the $\beta_L(a)$ conformation by typically 1 kcal/mol, in agreement with the experimentally detected conformer on *N*-Ac-Ala-NHMe.⁹ This marked contrast with the NAPA case, together with the lack of stability of the $\beta_L(g+)$ and $\beta_L(g-)$ in NAPA,³² demonstrates that the increase in stability of the NAPA $\beta_L(a)$ structure is essentially caused by the side chain. Of particular importance is the occurrence of Ar(i)-NH(i+1) interactions, which are only possible for an anti conformation of the side chain. The corresponding stabilization energy, estimated from the change in relative energy between the β_L and γ_L structures when going from *N*-Ac-Ala-NH₂ to NAPA, is about 2 kcal/mol.

In the same spirit, the difference in energies of the three γ_L conformations of NAPA of about 1 kcal/mol stems from differences in γ_L backbone/side chain Ar(i)-NH(i) interactions (which turn out to be weaker than the Ar(i)-NH(i+1) interaction). The $\gamma_L(g+)$ conformation exhibits the strongest Ar(i)-NH(i) interaction, in contrast to $\gamma_L(a)$, where this interaction cannot occur (Figure 4).

Finally, one will notice that the absence of the γ_D conformer in the jet illustrates the efficiency of isomerization toward the γ_L or β_L topological basins during the expansion.

5.2. Structure and Spectroscopy of the Most Stable NAPA Conformers. *The $\beta_L(a)$ Conformation (A).* The calculated β_L conformation exhibits an extended backbone structure, with two close contacts, corresponding to the NH(i+1)-C^γ(Phe) (NH → π) and the NH(i)-CO(i) C₅ interactions (DFT interatomic distances of 2.64 and 2.16 Å, respectively). The latter, although being characterized by a rather short H-O distance, cannot be considered as a H-bond, in particular because of its nonlinear character (NH-O angle of 107°). IR spectroscopy does exhibit the fingerprint of these weak interactions. Figure 6 shows that the stretching frequency of the NH group engaged in a C₅ bond is in the 3430 cm⁻¹ region for Gly, Ala, or Phe, independent of the residue considered and thus red-shifted by ~50 cm⁻¹ from the free NH region (3460–3485 cm⁻¹), as obtained in the calculations of the γ conformations (Figure 4). Similarly, the

simultaneous red-shift of the two components of the terminal amino group, when going from the Gly or Ala residue to Phe (Figure 6), corresponds to the spectroscopic fingerprint of a weak perturbation of the terminal amino group, namely, here by the interaction of the trans NH of the amide with the C^γ carbon of the phenyl ring. This effect on the frequencies is accounted for by the vibrational coupling model presented in Figure 3 assuming a H-bond strength of about 20 cm⁻¹. Such a red-shift would be expected for a methylated amide group on the C terminus¹⁹ where vibrational coupling is absent, clearly showing the weak but significant effect of the interaction. In the mid-IR range, the amide II bands are the bands that are most specific to the $\beta_L(a)$ conformation, showing an intense band at 1390 cm⁻¹, mainly because of the in-plane bend of the *cis*-NH of the terminal amide group, being strongly coupled to several CH bends on the backbone.

Both the C₅ as well as the NH(i+1) → π interactions contribute to the overall stability of the conformation. In Gly and Ala, the C₅ interaction does not provide enough stabilization for the β_L conformation to make it the most stable conformation. However, the situation is different in Phe, where there is a significant NH(i+1)- π interaction.

The γ_L Conformations (B and C). In the γ_L structures, a C₇ H-bond between the NH(i+1)-CO(i-1) groups is present, with H-O DFT interaction distances of 2.00, 2.03, and 2.16 Å for the g+, g-, and side chain orientations, respectively, and with a more linear geometry than the β_L structure (NH-O angle of 145, 143, and 138°, respectively). The coupling model of Figure 3 yields a H-bond strength of ~125 cm⁻¹, illustrating the increased interaction compared to both the C₅ interaction (~50 cm⁻¹) or the NH(i+1) → π bond (~20 cm⁻¹) for the β_L structure. The width of the band, about 10 cm⁻¹, is also indicative of a stronger H-bonding as well as of a larger coupling of the vibrational mode to the rest of the molecule.

Among the three γ conformations, only the most stable one, $\gamma_L(g+)$, exhibits a NH(i)-C^γ(Phe) interaction. Experimentally, this leads to a red-shift of the frequency of ~30 cm⁻¹ relative to the free NH mode for conformer B, and the DFT-optimized structure yields a H-C^γ distance of 2.46 Å. In the $\gamma_L(a)$ conformation, the anti orientation of the Phe side chain clearly forbids such an interaction. In the $\gamma_L(g-)$ conformation, the H-C^γ DFT (B3LYP/6-31+G(d)) distance of 2.83 Å suggests a much less intense interaction than for $\gamma_L(g+)$. This is confirmed by both the absence of a red-shifted NH_{Phe} stretch band in the IR spectrum of conformer C (Figure 4) and the relative ordering of the MP2 energies within the γ_L manifold (Table 2), which indicates that the energetic contribution of the NH(i)-C^γ(Phe) interaction is about 0.8 kcal/mol.

The γ_D Conformations. The γ_D conformation considered, $\gamma_D(g-)$, exhibits the C₇ H-bond between the NH(i+1) and the CO(i-1) group, with a short H-O interaction distance of 1.94 Å (DFT B3LYP/6-31+G(d)) and a linearity similar to the γ_L geometries. The axial geometry of the residue relative to the C₇ ring causes a significant destabilization, which explains the higher energy of this conformation with respect to the γ_L basin. No other interaction is noticed, as illustrated by the calculated amide A IR spectra.

5.3. Comparison with Similar Dipeptide Analogues. The present results can be compared to those obtained by the same technique on the analogous system NAPMA and on model dipeptides in which Phe is replaced by another more extended aromatic side chain, such as tryptophan. For NAPMA, Gerhards and co-workers found experimentally four conformations, of which three of them were assigned to $\beta_L(a)$, $\gamma_L(g+)$, and

$\gamma_L(g^-$ or a).²³ This study, in qualitative agreement with the present observation of the $\beta_L(a)$, $\gamma_L(g^+)$, and $\gamma_L(g^-)$ conformations on NAPA, emphasizes the negligible role of the methyl capping group on the C terminal amide on the relative conformational stabilities of the system. In both NATA and NATMA, Zwier and co-workers found that β_L and γ_L conformations also do compete: both structures are observed simultaneously in the jet and exhibit very similar DFT energies.¹⁹ A careful analysis of the reported experimental data and DFT calculations indicates that, as in the NAPA case, IR spectroscopy of these molecules in the amide A region exhibits the spectral fingerprints, specific of the weak $NH-\pi$ interactions, identified in the present work. The convergence of the NAPA, NAPMA, and NATA or NATMA data therefore provides a consistent set of data that provides evidence for the role of backbone/aromatic side chain interactions in the stabilization of extended backbone structures.

5.4. Conformer-Selective Fluorescence Yield. The dramatic dependence of the fluorescence quantum yield upon the structure requires us to examine the issue of fluorescence quenching. We showed that the conformer for which the low fluorescence yield is observed is in a $\gamma_L(g^+)$ conformation, which contains a $NH(i) \rightarrow \pi$ interaction. This suggests that this interaction is responsible for the fluorescence quenching. Such issues have been widely addressed in the condensed phase. In particular, the existence of conformers has, for instance, been evoked to explain biexponential fluorescence behaviors of the tyrosine (Tyr) and tryptophan amino acid derivatives and dipeptides.⁴⁷ Charge transfer with the amide group as an electron acceptor has also been proposed as the mechanism responsible for the quenching. In such a case, however, solvation is thought to be the fundamental process for stabilizing charge-transfer states, and the charged terminus (NH_3^+ or COO^-) when one of the termini is not protected is also expected to play a role as it influences the electron density. In the present gas-phase experiments, none of these effects can be effective and other origins have to be considered. One should also notice that the quantum yields in the condensed phase are rather high for Trp and Tyr (~ 0.2) but much lower for Phe (< 0.04),⁴⁸ showing that the processes responsible for quenching in Phe are probably different. Dramatic phenomena in fluorescence were reported by Zwier and co-workers for the hydrogen-bonded complexes benzene-HCl and toluene-HCl.⁴⁹ Two-color two-photon ionization experiments, in which the delay between the $S_1 \leftarrow S_0$ laser and the ionization laser was continuously scanned, gave evidence that the hydrogen-bonding interactions lead to enhanced intersystem crossing to the triplet state. As the present dipeptide has a similar interaction as the π -type H-bond in toluene-HCl, the same interpretation in terms of enhanced intersystem crossing seems to be plausible here.

6. Conclusion

The conformational preferences on a model dipeptide in the gas phase deduced from the comparison of IR experimental and theoretical considerations enable us to provide chemical physics grounds, on the basis of characterization of local interactions, for the observed conformational distributions of residues in proteins in biological systems. The most abundant forms observed in biological systems correspond to the most stable forms observed in the modified gas-phase peptides. This suggests that simple energetic reasoning, free of any entropic, solvent, or environment effects, may be significant and relevant to biological systems.

The IR/UV double-resonance laser techniques provide spectroscopists with important information about the interactions

that play a role in the system, including weakly polar ones such as $NH \rightarrow \pi$ interactions. The experimental information collected, which was interpreted with the help of theory, also enabled us to validate the energetic side of these calculations, providing a good example of synergy between experiment and theory. This encouraging result suggests that the study of larger model peptide systems, which is already on its way,^{23,50,51} should provide a fruitful approach for our understanding of the competition between local conformational preferences of each residue and the more global secondary structures in the peptide chain, like β -turns.

Acknowledgment. We gratefully acknowledge the support by the Stichting voor Fundamenteel Onderzoek der Materie (FOM) in providing the required beam time on FELIX and highly appreciate the skillful assistance by the FELIX staff. This work was supported by the European Community, Research Infrastructure Action under the FP6 "Structuring the European Research Area" Program through the Integrated Infrastructure Initiative "Integrating Activity on Synchrotron and Free Electron Laser Science".

References and Notes

- (1) Burley, S. K.; Petsko, G. A. *FEBS Lett.* **1986**, *203*, 139.
- (2) Levitt, M.; Perutz, M. F. *J. Mol. Biol.* **1988**, *201*, 751.
- (3) Steiner, T. *Angew. Chem., Int. Ed.* **2002**, *41*, 48.
- (4) Toth, G.; Watts, C. R.; Murphy, R. F.; Lovas, S. *Proteins* **2001**, *43*, 373.
- (5) Duan, G.; Smith, V. H.; Weaver, D. F. *J. Phys. Chem. A* **2000**, *104*, 4521.
- (6) Cheng, J. G.; Kang, C. M.; Zhu, W. L.; Luo, X. M.; Puah, C. M.; Chen, K. X.; Shen, J. H.; Jiang, H. L. *J. Org. Chem.* **2003**, *68*, 7490.
- (7) Branden, C.; Tooze, J. *Introduction to Protein Structure*; Garland Publishing: New York, 1999.
- (8) Chou, K. *Anal. Biochem.* **2000**, *286*, 1.
- (9) Lavrich, R. J.; Plusquellic, D. F.; Suenram, R. D.; Fraser, G. T.; Walker, A. R. H.; Tubergen, M. J. *J. Chem. Phys.* **2003**, *118*, 1253.
- (10) Perczel, A.; Csizmadia, I. G. *Int. Rev. Phys. Chem.* **1995**, *14*, 127.
- (11) Vargas, R.; Garza, J.; Hay, B. P.; Dixon, D. A. *J. Phys. Chem. A* **2002**, *106*, 3213.
- (12) Endredi, G.; Perczel, A.; Farkas, O.; McAllister, M. A.; Csonka, G. I.; Ladik, J.; Csizmadia, I. G. *J. Mol. Struct. (THEOCHEM)* **1997**, *391*, 15.
- (13) Farkas, O.; McAllister, M. A.; Ma, J. H.; Perczel, A.; Hollosi, M.; Csizmadia, I. G. *J. Mol. Struct. (THEOCHEM)* **1996**, *369*, 105.
- (14) Perczel, A.; Farkas, O.; Csaszar, A. G.; Csizmadia, I. G. *Can. J. Chem.-Rev. Can. Chim.* **1997**, *75*, 1120.
- (15) Jakli, I.; Perczel, A.; Farkas, O.; Hollosi, M.; Csizmadia, I. G. *J. Mol. Struct. (THEOCHEM)* **1998**, *455*, 303.
- (16) Hudaky, P.; Jakli, I.; Csaszar, A. G.; Perczel, A. *J. Comput. Chem.* **2001**, *22*, 732.
- (17) Chasse, G. A.; Rodriguez, A. M.; Mak, M. L.; Deretey, E.; Perczel, A.; Sosa, C. P.; Enriz, R. D.; Csizmadia, I. G. *J. Mol. Struct. (THEOCHEM)* **2001**, *537*, 319.
- (18) Dian, B. C.; Longarte, A.; Zwier, T. S. *Science* **2002**, *296*, 2369.
- (19) Dian, B. C.; Longarte, A.; Mercier, S.; Evans, D. A.; Wales, D. J.; Zwier, T. S. *J. Chem. Phys.* **2002**, *117*, 10688.
- (20) Evans, D. A.; Wales, D. J.; Dian, B. C.; Zwier, T. S. *J. Chem. Phys.* **2004**, *120*, 148.
- (21) Dian, B. C.; Longarte, A.; Winter, P. R.; Zwier, T. S. *J. Chem. Phys.* **2004**, *120*, 133.
- (22) Gerhards, M.; Unterberg, C. *Phys. Chem. Chem. Phys.* **2002**, *4*, 1760.
- (23) Gerhards, M.; Unterberg, C.; Gerlach, A.; Jansen, A. *Phys. Chem. Chem. Phys.* **2004**, *6*, 2682.
- (24) Perczel, A.; Jakli, I.; McAllister, M. A.; Csizmadia, I. G. *Chem.-Eur. J.* **2003**, *9*, 2551.
- (25) Berg, M. A.; Chasse, G. A.; Deretey, E.; Fuzery, A. K.; Fung, B. M.; Fung, D. Y. K.; Henry-Riyad, H.; Lin, A. C.; Mak, M. L.; Mantas, A.; Patel, M.; Repyakh, I. V.; Staikova, M.; Salpietro, S. J.; Tang, T. H.; Vank, J. C.; Perczel, A.; Csonka, G. I.; Farkas, O.; Torday, L. L.; Szekely, Z.; Csizmadia, I. G. *J. Mol. Struct. (THEOCHEM)* **2000**, *500*, 5.
- (26) Improta, R.; Benzi, C.; Barone, V. *J. Am. Chem. Soc.* **2001**, *123*, 12568.
- (27) Benzi, C.; Improta, R.; Scalmani, G.; Barone, V. *J. Comput. Chem.* **2002**, *23*, 341.

- (28) Chass, G. A.; Lovas, S.; Murphy, R. F.; Csizmadia, I. G. *Eur. Phys. J. D* **2002**, *20*, 481.
- (29) Langella, E.; Rega, N.; Improta, R.; Crescenzi, O.; Barone, V. *J. Comput. Chem.* **2002**, *23*, 650.
- (30) Improta, R.; Barone, V.; Kudin, K. N.; Scuseria, G. E. *J. Am. Chem. Soc.* **2001**, *123*, 3311.
- (31) Czinki, E.; Csaszar, A. G.; Perczel, A. *Chem.—Eur. J.* **2003**, *9*, 1182.
- (32) Chass, G. A.; Mirasol, R.; Setiadi, D. H.; Tang, T.-H.; Chin, W.; Mons, M.; Dimicoli, I.; Dognon, J.-P.; Viskolcz, B.; Lovas, S.; Penke, B.; Csizmadia, I. G. *J. Phys. Chem. A*, **2005**, *109*, 5289.
- (33) Cable, J. R.; Tubergen, M. J.; Levy, D. H. *J. Am. Chem. Soc.* **1989**, *111*, 9032.
- (34) Cohen, R.; Brauer, B.; Nir, E.; Grace, L.; de Vries, M. S. *J. Phys. Chem. A* **2000**, *104*, 6351.
- (35) Piuze, F.; Dimicoli, I.; Mons, M.; Tardivel, B.; Zhao, Q. *Chem. Phys. Lett.* **2000**, *320*, 282.
- (36) Robertson, E. G.; Simons, J. P. *Phys. Chem. Chem. Phys.* **2001**, *3*, 1.
- (37) Zwier, T. S. *J. Phys. Chem. A* **2001**, *105*, 8827.
- (38) Weinkauff, R.; Schermann, J. P.; de Vries, M. S.; Kleinermanns, K. *Eur. Phys. J. D* **2002**, *20*, 309.
- (39) Bakker, J. M.; Aleese, L. M.; Meijer, G.; von Helden, G. *Phys. Rev. Lett.* **2003**, *91*, 203003.
- (40) Piuze, F.; Mons, M.; Dimicoli, I.; Tardivel, B.; Zhao, Q. *Chem. Phys.* **2001**, *270*, 205.
- (41) Piuze, F.; Dimicoli, I.; Mons, M.; Millié, P.; Brenner, V.; Zhao, Q.; Soep, B.; Tramer, A. *Chem. Phys.* **2002**, *275*, 123.
- (42) Li, S.; Bernstein, E. R. *J. Chem. Phys.* **1992**, *97*, 804.
- (43) Yao, J.; Im, H. S.; Foltin, M.; Bernstein, E. R. *J. Phys. Chem. A* **2000**, *104*, 6197.
- (44) Jacon, M.; Lardeux, C.; Lopez-Delgado, R.; Tramer, A. *Chem. Phys.* **1977**, *24*, 145.
- (45) Robertson, E. G.; Hockridge, M. R.; Jelfs, P. D.; Simons, J. P. *Phys. Chem. Chem. Phys.* **2001**, *3*, 786.
- (46) Nakanaga, T.; Ito, F. *J. Phys. Chem. A* **1999**, *103*, 5440.
- (47) Beechem, J. M.; Brand, L. *Annu. Rev. Biochem.* **1985**, *24*, 43.
- (48) *Handbook of Biochemistry and Molecular Biology*, 3rd ed.; CRC Press: Cleveland, OH, 1976.
- (49) Gotch, A. J.; Zwier, T. S. *J. Chem. Phys.* **1990**, *93*, 6977.
- (50) Chin, W.; Mons, M.; Dognon, J.-P.; Piuze, F.; Tardivel, B.; Dimicoli, I. *Phys. Chem. Chem. Phys.* **2004**, *6*, 2700.
- (51) Chin, W.; Dognon, J.-P.; Piuze, F.; Tardivel, B.; Dimicoli, I.; Mons, M. *J. Am. Chem. Soc.* **2005**, *127*, 707.