CHEMPHYSCHEM

# Folding Structures of Isolated Peptides as Revealed by Gas-Phase Mid-Infrared Spectroscopy

Joost M. Bakker,*[a] Christian Plützer,[b] Isabel Hünig,[b] Thomas Häber,[b] Isabelle Compagnon,[a] Gert von Helden,[c] Gerard Meijer,[c] and Karl Kleinermanns*[b]

*To understand the intrinsic properties of peptides, which are determined by factors such as intramolecular hydrogen bonding, van der Waals bonding and electrostatic interactions, the conformational landscape of isolated protein building blocks in the gas phase was investigated. Here, we present IR-UV double-resonance spectra of jet-cooled, uncapped peptides containing a tryptophan (Trp) UV chromophore in the 1000–2000 $cm^{-1}$ spectral range. In the series Trp, Trp-Gly and Trp-Gly-Gly (where Gly stands for glycine), the number of detected conformers was found to decrease from six (Snoek et al., PCCP, **2001**, 3, 1819) to four and two, respectively, which indicates a trend to relaxation to a global minimum. Density functional theory calculations reveal that the O–H in-plane bending vibration, together with the N–H in-plane bending and the peptide C=O stretching vibrations, is a sensitive probe to hydrogen bonding and, thus, to the folding of the peptide backbone in these structures. This enables the identification of spectroscopic fingerprints for the various conformational structures. By comparing the experimentally observed IR spectra with the calculated spectra, a unique conformational assignment can be made in most cases. The IR-UV spectrum of a Trp-containing nonapeptide (Trp-Ala-Gly-Gly-Asp-Ala-Ser-Gly-Glu) was recorded as well and, although the IR spectrum is less well-resolved (and it probably results from different isomers), groups of amide I (peptide C=O stretching) and amide II (N–H in-plane bending) bands can still be recognised, in agreement with predictions at the AM1 level.*

## 1. Introduction

The folding of proteins to a specific tertiary and quaternary structure determines their physiological activity. Important contributions to the free energy of α-helix and β-sheet formation are the enthalpy decrease by hydrogen bonding and the entropy increase due to hydrophobic interactions between nonpolar side chains and surrounding water molecules. The latter counterbalances the entropy decrease during the transition from a random conformation to the folded, native form. The relative importance of these factors for a specific folding pathway is largely unknown.

The spectroscopy of peptides, which are isolated in the gas phase and cooled to a few Kelvin in a supersonic expansion, allows the study of peptide folding, essentially without entropic effects due to the solvent or, alternatively, with a controlled change of free energy by forming peptide–water clusters with a known number of water molecules. Gas-phase measurements can thus contribute to a better understanding of the role of the solvent in the peptide-folding process. These measurements have the added advantage of a sufficiently high spectral resolution to obtain structural information and to provide stringent tests for molecular modelling computations that aim to simulate protein folding and explore the pathways towards a global folding minimum. Especially the use of IR-UV double-resonance techniques allows for a direct assignment of experimentally observed IR spectra to structures obtained from high-level ab initio calculations.

Recently, gas-phase spectroscopic measurements have been successfully applied to study the conformational properties and preferences of isolated or small clusters of amino acids,[1–8] protected amino acids[9–13] and dipeptides,[14] all of which can be vaporised by thermal heating. Molecular species that are more difficult to bring intact into the gas phase can be vaporised using laser desorption. This technique has previously been applied to study the gas-phase structure of protected tripepti-

[a] *Dr. J. M. Bakker,[+] Dr. I. Compagnon*
*FOM Institute for Plasma Physics Rijnhuizen, Edisonbaan 14*
*3439 MN Nieuwegein (The Netherlands)*
*E-mail: jmbakker@fhi-berlin.mpg.de*

[b] *Dr. C. Plützer,[++] Dr. I. Hünig,[+++] Dr. T. Häber, Prof. K. Kleinermanns*
*Institut für Physikalische Chemie, Heinrich-Heine Universität Düsseldorf*
*40225 Düsseldorf (Germany)*
*E-mail: kleinermanns@uni-duesseldorf.de*

[c] *Dr. G. von Helden, Prof. G. Meijer*
*Fritz-Haber-Institut der Max-Planck-Gesellschaft*
*Faradayweg 4–6, 14195 Berlin (Germany)*

[+] *Current address:*
*Fritz-Haber-Institut der Max-Planck-Gesellschaft, Faradayweg 4–6*
*14195 Berlin (Germany)*
*Fax: (+49) 30-8413-5603*

[++] *Current address:*
*JILA, University of Colorado, Boulder, CO 80309–0440 (USA)*

[+++] *Current address:*
*Physical & Theoretical Chemistry Laboratory, South Parks Road*
*Oxford OX1 3QZ (UK)*

  DOI: 10.1002/cphc.200400345 

des[15] and of unmodified "natural" di- or tripeptides.[16] In this last work, IR-UV double-resonance spectra of the peptides Gly-Trp, Trp-Gly and Trp-Gly-Gly were reported in the region of the N–H and O–H stretching vibrations (2800–3700 $cm^{-1}$). In UV-UV hole-burning experiments, two conformers were observed for Gly-Trp, four conformers for Trp-Gly and two conformers for Trp-Gly-Gly.[16] IR spectra were obtained for all these conformers, except for Trp-Gly, where the IR spectrum of only one conformer was measured. By comparing the vibrational spectra measured in the spectral region around 3 μm with spectra of bare tryptophan,[6] N-acetyl tryptophan methyl amide,[11] as well as to density functional theory (DFT) calculations, the observed spectral features could be assigned to the peptide N–H, the indole N–H and the carboxyl O–H stretching vibrations. Unfortunately, comparison of the experimental spectra to calculated spectra did not allow for an unambiguous assignment of the experimental spectrum to a single conformer in all cases.

Here, we present the results of a study in which the experiments on Trp containing peptides are extended into the mid-IR spectral range, which was recently shown to contain extensive conformer-sensitive spectral information.[8,13,14,17,18]

The IR absorption spectra reported here cover the spectral region of the O–H and N–H in-plane bending and the C=O stretching vibrations, thereby giving complementary information which enables us to make definite structural assignments in most cases. It will be shown that the O–H in-plane bending, the peptide C=O stretching and, to a lesser extent, the N–H in-plane bending vibrations are sensitive probes for structural assignments in the spectral region between 1000 and 2000 $cm^{-1}$, providing fingerprints for stretched and folded conformations, as well as for the type of hydrogen bond formed by the carboxylic O–H group. Additionally, we report newly obtained IR spectra of two conformers of Trp-Gly in the N–H and O–H stretching region. Finally, we present the first IR spectrum of a nonapeptide, the delta-sleep inducing peptide (DSIP) Trp-Ala-Gly-Gly-Asp-Ala-Ser-Gly-Glu,[19] in the region from 1000 to 2000 $cm^{-1}$. Groups of C=O stretching (amide I), N–H in-plane bending (amide II) and O–H in-plane bending vibrations are clearly visible in the spectrum.

## Experimental Section

To study the IR absorption spectra of gas-phase peptides, a mixture of the solid peptide and graphite powder was deposited onto the surface of a bar of solid graphite ($50\times15\times2$ $mm^3$). H-Gly-Trp-OH (Bachem AG, 261 amu), H-Trp-Gly-OH (Sigma, 261 amu), H-Trp-Gly-Gly-OH (Bachem AG, 318 amu) and DSIP (Bachem AG, 849 amu) were used without further purification. The bar of solid graphite was placed directly underneath the orifice of a pulsed valve. The pulsed valve (0.5 mm nozzle diameter) operated at a 10 Hz repetition rate and at an argon backing pressure of up to 3.5 bar, with pulse durations of about 50 μs. Directly after opening the nozzle, sample molecules were desorbed from the graphite matrix by a pulsed Nd:YAG laser (Thales Diva-2, 1064 nm, 5 ns pulse duration, $<1$ mJ per pulse), entrained in the supersonic expansion and cooled to rotational temperatures of about 10 K. Because of the rapid cooling in the expansion, different conformers, separated by energy barriers of typically a few hundred $cm^{-1}$,[20] could be frozen out in local minima and could not relax to a single global minimum. About 4 cm downstream from the nozzle, the expansion was skimmed, and the molecular beam entered a differentially pumped Wiley–McLaren type linear time-of-flight (TOF) mass spectrometer. The UV and IR lasers were spatially overlapped at the crossing point with the molecular beam. Ions produced in this region were accelerated and detected using a micro-channel-plate (MCP) detector. The MCP signal was amplified and fed into a 10 bit, 100 mega-samples $s^{-1}$ digital oscilloscope (LeCroy), which was connected to a PC. Two digital delay/pulse generators were used to synchronise the molecular beam and the various laser sources.

For the measurement of IR absorption spectra of jet-cooled peptides, IR ion-dip spectroscopy was used.[21–26] The ions were produced from the vibrational ground state in the electronic ground state $S_0$ using a two-photon ionisation scheme. Ground-state molecules were excited to the first electronically excited singlet state $S_1$ using a frequency-doubled, Nd:YAG pumped dye laser (Rhodamine 575, 8 ns pulse duration, spectral bandwidth $\approx 0.4$ $cm^{-1}$). A second photon from the same excitation laser pulse then ionised the molecules. Typically, a few μs before the UV excitation laser is fired, the IR laser interacts with the molecules. If the IR laser is resonant with a vibrational transition, population is transferred from the vibrational ground state in $S_0$ to an excited vibrational state, which leads to a depletion of ground-state molecules and a dip in the ion signal. By measuring the ion yield, while scanning the frequency of the IR laser, the IR ion-dip spectrum is obtained.

IR radiation was produced at the free-electron laser for infrared experiments (FELIX) user facility at the FOM institute for Plasma Physics Rijnhuizen.[27,28] The frequency range that could be covered extended from 40 to 2000 $cm^{-1}$, but only measurements in the range from 1000 to 2000 $cm^{-1}$ are presented in this paper.[29] The temporal output of this pulsed-laser system consisted of a few μs long burst of micropulses, forming a so-called macropulse. The micropulse spacing within the macropulse was set to 1 ns, and the micropulse duration to about 100 optical cycles, which resulted in a spectral bandwidth of around 0.5 % (fwhm) of the central frequency. Typically, energies of up to 100 mJ were reached in the macropulse. In this experiment, the UV detection laser was pulsed at a 10 Hz repetition rate, while FELIX was operated at 5 Hz. By independently recording both the IR-on and the IR-off signal, a "normalised" ion-dip spectrum could be obtained, which was insensitive to long-term drifts in source conditions and/or in UV laser power. For the newly obtained IR-UV spectra of conformers of Trp-Gly in the N–H and O–H stretching region, the IR radiation was generated by difference frequency mixing in a $LiNbO_3$ crystal (2800–4000 $cm^{-1}$).[30]

Quantum chemical calculations were performed using the Gaussian 98 program package.[31] The procedure of obtaining and selecting the relevant conformers, based on structural considerations, has been described previously.[16] Selected conformers of Gly-Trp, Trp-Gly and Trp-Gly-Gly were submitted to a full geometry optimisation at the B3LYP/6–31G(d,p) level, from which the energies, as given in the figures, were derived. For the optimised structures, harmonic frequencies were calculated at the same level. For DSIP, prototypical structures of a fully stretched and a helical conformer were optimised using the semiempirical AM1 method. Harmonic frequencies were calculated at the same level.

Finally, a short note about nomenclature: all experimentally observed conformers are labelled with lower-case letters, calculated

  

structures with capital letters. The labelling that is used is consistent with previously published work.[16]

## 2. Results

### 2.1. Gly-Trp

UV-UV double-resonance spectroscopic studies have revealed the presence of two different conformers under jet-cooling conditions. Strong resonances at 34715 and 34900 $cm^{-1}$ are associated with conformers **a** and **b**, respectively.[16] Figure 1

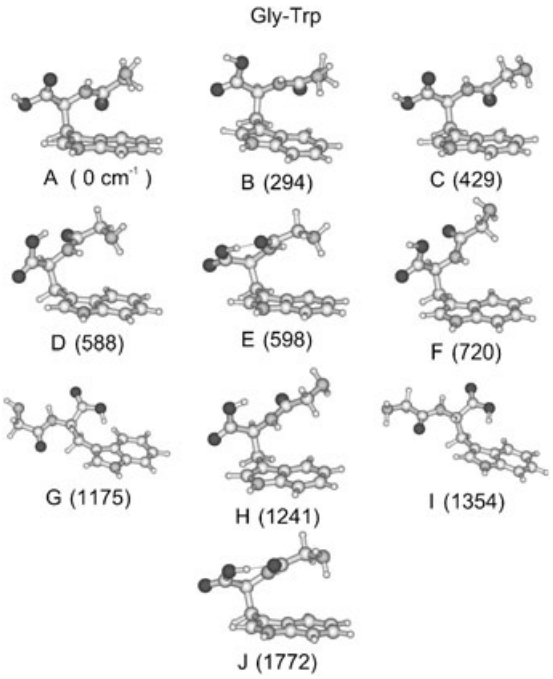


***Figure 1.*** *Calculated structures of Gly-Trp at the B3LYP/6–31G(d,p) level of theory.[16] Relative electronic energies [in $cm^{-1}$] are given in parentheses. Structures E and J exhibit strong COOH···O=C hydrogen bonds, whereas structures D, G, H and I form weak hydrogen bonds either with the π-system of the indole ring (G, I) or with the nitrogen end of the peptide group (D, H).*

displays the optimised geometries of the ten calculated structures of Gly-Trp with their relative electronic energies, neglecting zero-point corrections.[16] Our earlier studies in the O–H and N–H stretching regions showed that the experimental spectrum of conformer **a** belongs to a hydrogen-bonded structure, which involves a hydrogen bond between the carboxyl O–H and the peptide C=O groups, and conformer **a** was unambiguously assigned to structure E. Conformer **b** could not be unambiguously assigned based only on the frequencies of the N–H and O–H stretching vibrations, so that both structures A and B remained possible.

Figure 2 shows the IR-UV double-resonance spectra of conformers **a** and **b** (top traces) of Gly-Trp in the 1000–1850 $cm^{-1}$ region. The experimental spectra are obtained by tuning the UV excitation laser to the transition frequencies mentioned above. Conformer **a** shows prominent bands at 1781, 1659, 1512, and 1400 $cm^{-1}$, which can tentatively be assigned to the

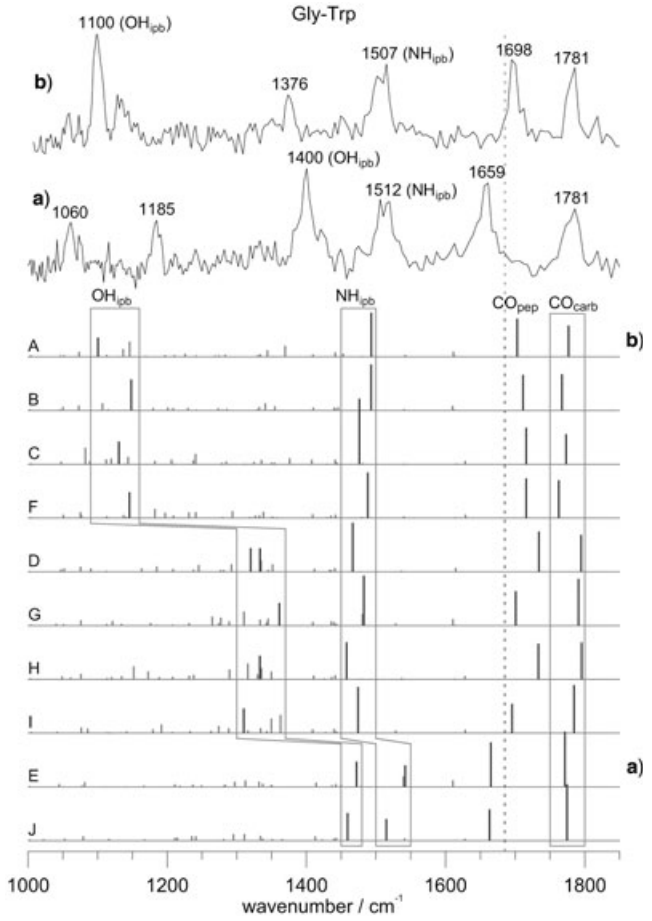


***Figure 2.*** *IR-UV double-resonance spectra (top 2 traces) and calculated spectra of Gly- Trp in the spectral region from 1000 to 2000 $cm^{-1}$. Harmonic frequencies of the calculated spectra are scaled by an empirical factor of 0.964. The assigned calculated structures are indicated by the appropriate conformer labels.*

carboxyl C=O stretching (from now on referred to as $CO_{carb}$), the peptide C=O stretching ($CO_{pep}$) vibrations, the peptide N–H in-plane bending ($NH_{ipb}$) and the carboxylic O–H in-plane bending ($OH_{ipb}$) vibrations. Conformer **b** differs mostly in the position of the $CO_{pep}$ band, which is shifted to higher wavenumbers (at 1698 $cm^{-1}$), and in the substantially lowered frequency of the $OH_{ipb}$ vibration, assuming this is the mode found at ≈1100 $cm^{-1}$.

Below the experimental spectra, the theoretical spectra of the ten calculated structures are shown. The calculated harmonic frequencies are scaled by a factor of 0.964, which is the empirical scaling factor required for the most prominent tryptophan monomer to match the experimental and calculated carboxylic C=O stretching frequencies.[8, 32] In the following, we will focus on the $OH_{ipb}$, $NH_{ipb}$, $CO_{pep}$ and $CO_{carb}$ vibrations, which are indicated in boxes in the figures.

The calculated spectra for Gly-Trp structures can be arranged in three groups with respect to the carboxylic O–H in-plane bending vibration $OH_{ipb}$: 1) The O–H group is not involved in a hydrogen bond and the $OH_{ipb}$ transition frequency is below 1200 $cm^{-1}$ (structures A, B, C and F). 2) A weak hydrogen bond is formed between the O–H group and the π-system of the

  

indole ring (structures G and I) or between the O–H group and the nitrogen end of the peptide group (structures D and H). 3) The O–H group forms a strong hydrogen bond with a lone electron pair of the peptidic carbonyl oxygen (structures E and J). With increasing hydrogen-bond strength, the $OH_{ipb}$ transition frequency shifts to higher values, covering a frequency range of about 400 $cm^{-1}$. Depending on the type of hydrogen bond, all transition frequencies fall into a certain wavenumber range, which allows for an immediate classification of the interaction. Additionally, there is a direct correlation between the red-shift of the carboxylic O–H stretch[16] and the blue-shift of the $OH_{ipb}$ vibration.

A second reliable probe of hydrogen bonding is the peptide C=O stretching vibration. If a lone pair of the oxygen atom is involved in a hydrogen bond, the C=O distance will increase and the vibrational frequency will red-shift by about 40–50 $cm^{-1}$. Although the spectroscopic shift is much smaller than for $OH_{ipb}$, we can draw a frequency limit around 1685 $cm^{-1}$ (indicated with a dashed vertical line in Figure 2): C=O groups with vibrational frequencies below this limit are involved in a strong hydrogen bond with the carboxylic O–H group (structures E and J). This is consistent with the observed blue-shift in the $OH_{ipb}$ frequency. Peptidic C=O groups that either are free or establish only weak interactions with the terminal amino group exhibit $CO_{pep}$ frequencies higher than 1685 $cm^{-1}$. As will be shown later, the frequency limit around 1685 $cm^{-1}$ is a general feature of the systems studied here (using the chosen theoretical method).

The peptidic N–H in-plane bending vibration $NH_{ipb}$ is less sensitive to structural changes, because the N–H group forms weaker hydrogen bonds than the O–H group. Except for cases where a strong O–H···O=C peptide bond is formed (structures E and J), the $NH_{ipb}$ band is hardly shifted at all and bears little structural information. In contrast to the O–H group, there is no direct correlation between the N–H stretching and the N–H in-plane bending frequency shifts.

Due to structural constraints, hydrogen bonds involving the carboxylic C=O group are not feasible, and the calculated $CO_{carb}$ frequencies fall into a narrow wavenumber range between 1750 and 1800 $cm^{-1}$. The effects of *syn*- and *anti*-configuration of the carboxyl group are too small here to be of significance in the structural assignment.

Based on the agreement between the experimental and theoretical spectra, the previous assignment of conformer **a** to structure E is confirmed. Although structure J shows a similar spectral fingerprint in this region, this structure could be ruled out previously by comparing experimental and calculated spectra in the O–H and N–H stretching regions.[16] However, the cause of the relative strengths of the resonances observed at 1060 and 1185 $cm^{-1}$ is unclear.

Conformer **b** could not be unambiguously assigned previously, and both structures A and B remained possible. Although the spectral differences in the $OH_{ipb}$ vibrations are small, we tend to assign conformer **b** to structure A, because the experimental $OH_{ipb}$ band at 1100 $cm^{-1}$ lines up best with the theoretical predictions. This example nicely illustrates how the complementary information of both spectral regions helps to discriminate between different structures, and how it enables the assignment of experimental IR spectra of "natural" peptides to a single conformer. It also shows that the structures observed in molecular-beam experiments are not necessarily those with the lowest calculated energies.

### 2.2. Trp-Gly

In previous REMPI and UV-UV hole-burning experiments, four different conformers of Trp-Gly were observed in the molecular beam.[16] In those experiments, the IR spectrum of conformer **a** was reported as well. In Figure 3, this spectrum—measured

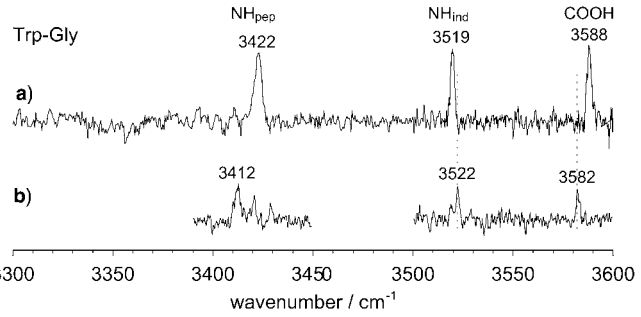


***Figure 3.*** *IR spectra of conformers* ***a*** *and* ***b*** *of Trp-Gly in the 3 μm region.*

with the UV probe laser set at 34613 $cm^{-1}$—is depicted once more in the top trace. The three observed IR bands are assigned to the free (not hydrogen bonded) carboxyl O–H stretching, the indole N–H stretching and the peptide N–H stretching vibrations at 3588, 3519 and 3422 $cm^{-1}$, respectively. These frequencies indicate a stretched, unfolded structure of the peptide.

The bottom trace in Figure 3 shows the newly obtained IR spectrum for conformer **b** in the N–H and O–H stretching regions. This is measured with the UV probe laser tuned to a frequency of 34761 $cm^{-1}$. In the additional material, also a spectrum of conformer **d** (detected with the UV laser at 34946 $cm^{-1}$) is presented.[33] As mid-IR spectra are only recorded for conformers **a** and **b**, only their 3 μm spectra are displayed in Figure 3. Conformers **b** and **d** both have a vibrational spectrum in this region that is very similar to that of conformer **a**, and it is thus likely that they all have very similar structures.

The structures of the nine calculated structures of Trp-Gly are depicted in Figure 4.[16] By comparing the IR spectra with calculated spectra in the region of the N–H and O–H stretching vibrations, conformer **a** was previously assigned to structures B or C, but it proved not to be possible to distinguish between these two structures.[16] Structures B and C differ only in the orientation of the carboxyl group, which is flipped by 180°. As all three observed conformers (**a**, **b** and **d**) exhibit very similar spectra in the 3 μm range, and as only two calculated spectra match this pattern, it is not unlikely that other structures may exist, which are not evaluated in the conformational

  

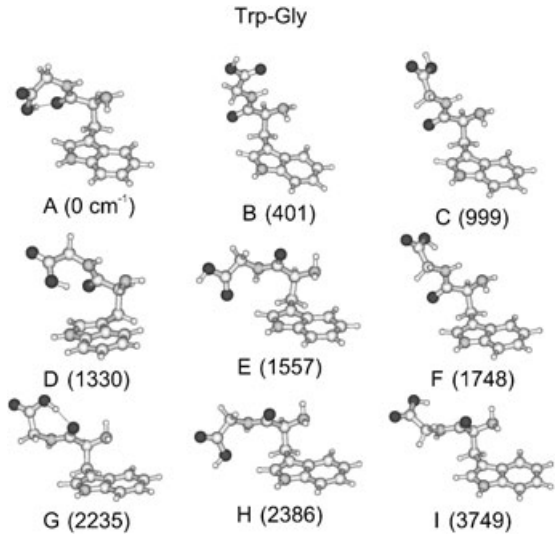


***Figure 4.*** *Calculated structures of Trp-Gly with their relative electronic energies [in cm$^{-1}$] in parentheses. Structures A and G exhibit strong COOH···O=C hydrogen bonds, whereas structures D, F and I form weak hydrogen bonds either with the $\pi$-system of the indole ring (D) or with the nitrogen end of the peptide group (F, I).*

search procedure. One characteristic in which the "missed" structure could differ from structures B and C is the orientation of the indole ring, which has only little influence on the vibrational spectrum if there is little interaction between the peptide backbone and the ring system. It is well-known however that DFT calculations do not take dispersions interactions properly into account and higher-level calculations are necessary here.

In Figure 5, the IR spectra of conformers **a** and **b** are shown in the spectral region from 1000 to 2000 cm$^{-1}$. The experimental spectra are obtained by positioning the UV excitation laser at 34613 (a) and 34761 cm$^{-1}$ (b). Similar to what happens in the spectral region of the N–H and O–H stretching vibrations, the spectra in the mid-IR region are also very similar. Four prominent peaks, which are assigned (from higher to lower wavenumbers) to the carboxyl and peptide C=O stretching vibrations, to the peptide N–H in-plane bending vibration, and to the carboxyl O–H in-plane bending vibration are observed in both spectra.

Below the experimental spectra, calculated stick spectra of nine different Trp-Gly structures are displayed. Harmonic frequencies are scaled again by a factor of 0.964. As in the case of Gly-Trp, the vibrations in the calculated spectra can be divided into three groups, which are characterised by their $OH_{ipb}$ frequency and by the type of intramolecular hydrogen bond. If the peptide C=O group is involved in a strong hydrogen bond (structures A and G), the $CO_{pep}$ frequency is red-shifted to below 1685 cm$^{-1}$, as also seen in the Gly-Trp system. $NH_{ipb}$ and $CO_{carb}$ are, again, largely insensitive to structural changes, and can thus not be used as structure-sensitive probes. There is one feature that does deviate from the described spectral pattern: in structure B, the $OH_{ipb}$ vibration is coupled to the $CH_2$ wagging vibration, but the O–H group remains non-hydrogen-bonded.

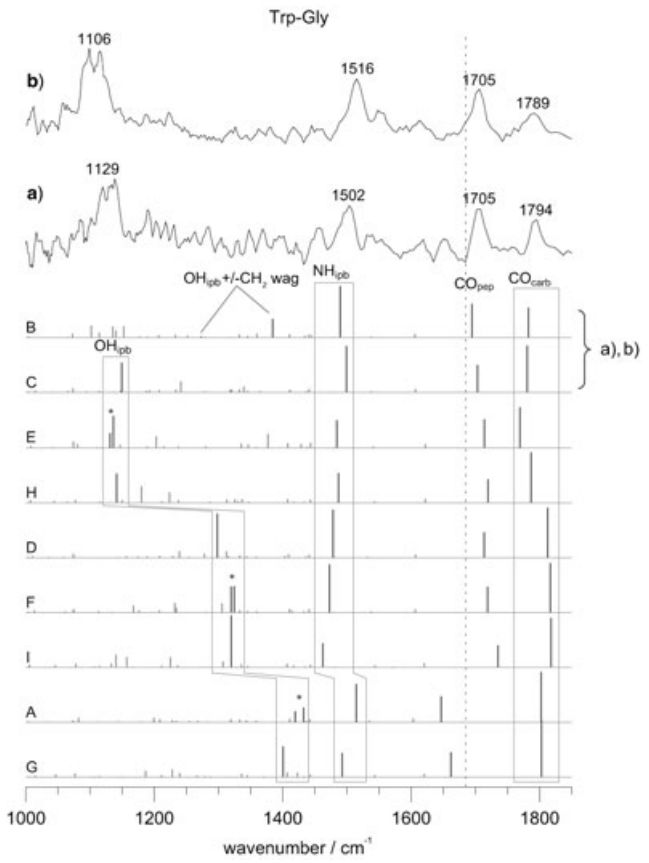


***Figure 5.*** *IR spectra of conformers **a** and **b** of Trp-Gly (top 2 traces), together with calculated spectra of Trp-Gly structures in the spectral region from 1000 to 2000 cm$^{-1}$. The frequencies of the calculated stick spectra are scaled by 0.964. The assigned calculated structures are indicated by the appropriate conformer labels. Peaks marked with an asterisk correspond to bands involving a strong coupling between the O–H in-plane bending and the $CH_2$ wagging vibration.*

From the structures that had a matching spectrum in the spectral region of the N–H and O–H stretching vibrations (structures B and C[16]), structure C matches the experimental spectra in the range between 1000 and 2000 cm$^{-1}$ considerably better. Although both calculated spectra match the C=O stretching and the N–H in-plane bending vibrations, the observed O–H in-plane bending vibration is better reproduced for structure C. Also, the strong band predicted for structure B around 1400 cm$^{-1}$ is not observed in the experimental spectra. However, as mentioned before, three conformers with very similar experimental spectra are found. Therefore, we cannot completely rule out structure B, which has several bands of moderate strength around 1150 cm$^{-1}$ which could give rise, for example, to the band structure observed around 1106 cm$^{-1}$ for conformer **b**.

### 2.3. Trp-Gly-Gly

In the previous REMPI and UV-UV hole-burning spectra of Trp-Gly-Gly, only two conformers were observed.[16] For each of these conformers, the IR spectra in the 3 μm region were recorded as well. The most striking difference between these two IR spectra is that both the free (not hydrogen bonded)

  

indole N–H and the free carboxyl O–H stretching vibrations are missing in the spectrum of one conformer (**a**), while both vibrational bands are present in the spectrum of the other conformer (**b**). It can be therefore concluded that, in conformer **a**, the carboxyl O–H group forms a hydrogen bond within the peptide backbone, while the backbone folds back to the indole ring, forming a hydrogen bond between the indole N–H group and the carboxylic C=O group. This structural motif is what is observed for the two lowest-energy structures (A and B) of Trp-Gly-Gly, as shown in Figure 6. In contrast, the presence of free carboxyl O–H and indole N–H stretching vibrations in the IR spectrum of conformer **b** suggests a stretched backbone conformation, which is similar to that of structures D and G in Figure 6. However, based solely on the N–H and O–H stretching spectra, no definite assignment could be made.

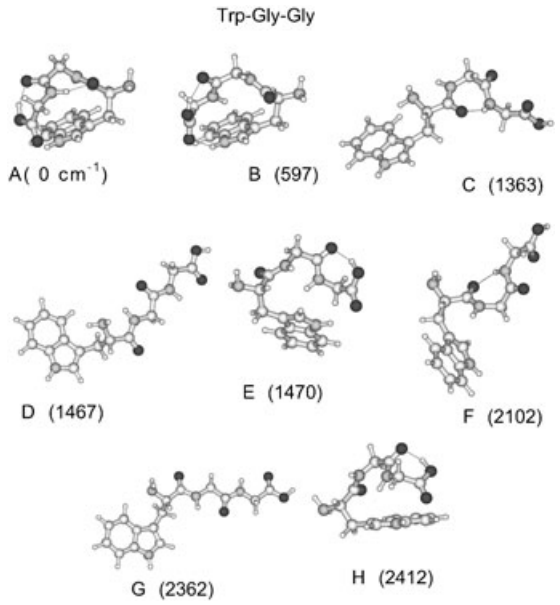


***Figure 6.*** *Calculated structures of Trp-Gly-Gly at the B3LYP/6–31G(d,p) level of theory. Relative electronic energies [in cm$^{-1}$] are given in parentheses. Structures A, B, E and H exhibit strong COOH···O=C hydrogen bonds. Structures A and B are further stabilised by NH···O=C bonds.*

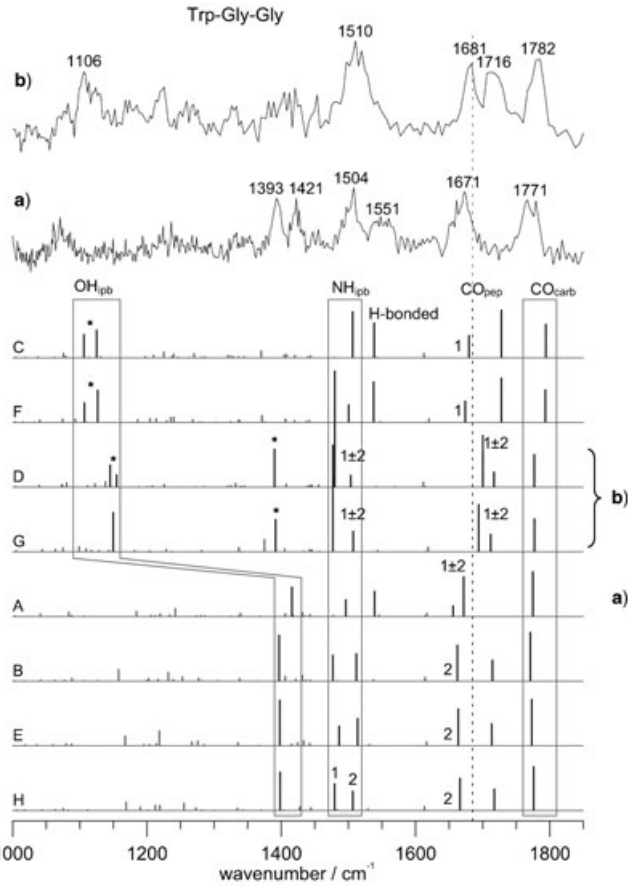


***Figure 7.*** *IR absorption spectra (top 2 traces) and calculated spectra of Trp-Gly-Gly in the spectral region from 1000 to 2000 cm$^{-1}$. The calculated harmonic frequencies are scaled by 0.964. The assigned calculated structures are indicated by the appropriate conformer labels. Peaks marked with an asterisk correspond to coupled vibrations involving the O–H in-plane bending and either the $CH_2$ wagging or the C–N stretching vibrations. The labels 1 and 2 classify vibrations of the peptide groups according to which peptide group they are localised in, counted from the N-terminus. A combination $1 \pm 2$ indicates a coupling between such vibrations, resulting in one symmetric and one asymmetric component, with low and high intensity, respectively.*

In Figure 7, the IR spectra of Trp-Gly-Gly in the range between 1000 and 2000 cm$^{-1}$ are shown. The experimental spectra are obtained by tuning the UV detection laser to 34651 and 34870 cm$^{-1}$ for conformers **a** and **b**, respectively.[16] The highest frequency modes in both IR spectra are assigned to the C=O stretching vibrations of the carboxyl group. The resonances of conformer **b** at 1716 and 1681 cm$^{-1}$ are tentatively assigned to free or only weakly interacting C=O stretching vibrations, and the band at 1671 cm$^{-1}$ of conformer **a** to a hydrogen-bond-shifted C=O stretching vibration. The bands at 1510 cm$^{-1}$ (conformer **b**) and 1504 cm$^{-1}$ (conformer **a**) can safely be assigned to N–H in-plane bending vibrations. Conformer **b** shows a band at 1106 cm$^{-1}$, which is most likely associated with the O–H in-plane bending vibration (OH$_{ipb}$). Due to the closeness of its absorption frequency to 1100 cm$^{-1}$, the carboxyl O–H group is assumed not to be hydrogen bonded. In contrast, the O–H in-plane bending vibration of conformer **a** is found at either 1393 or 1421 cm$^{-1}$, which indicates a strong hydrogen bond with a carbonyl group. This is consistent with the red-shifted C=O stretching vibration observed at 1671 cm$^{-1}$. These assignments are also consistent with the conclusions drawn earlier from the spectra in the region of the N–H and O–H stretching vibrations.

In the lower part of Figure 7, calculated IR absorption spectra of eight different structures of Trp-Gly-Gly are displayed. Harmonic frequencies are again scaled by a factor of 0.964. To distinguish between the vibrations of different peptide-bond groups, the peptide C=O stretching and N–H in-plane bending vibrations are numbered starting from the N-terminal amino acid, for example, CO(1) is the C=O stretching vibration of the peptide group closest to the tryptophan subunit. In case of NH$_{ipb}$, the lowest frequency band always corresponds to NH$_{ipb}$(1), unless stated otherwise. For structures D and G, both peptide C=O stretching and NH$_{ipb}$ vibrations are coupled, which leads to the appearance of a strong and a weak compo-

  

nent for each vibration. Bands marked with an asterisk correspond to strongly coupled vibrations.

In the case of Gly-Trp and Trp-Gly, structural constraints permit only O–H···O=$C_{pep}$ bonds. Due to a longer and more flexible backbone, Trp-Gly-Gly has the possibility of forming weaker N–$H_{pep}$···O=C interactions, as can be seen in structures C, F and A. This interaction is reflected in the H-bonded $NH_{ipb}$ vibration frequencies, which are blue-shifted by about 50 $cm^{-1}$ from that of the free $NH_{ipb}$ vibration. At the same time, the frequency of the C=O stretching vibration of the hydrogen acceptor group is slightly red-shifted, to around 1680 $cm^{-1}$. The C=O red-shift observed for the N–$H_{pep}$···O=C hydrogen bonds is somewhat smaller than that observed for the stronger O–H···O=C bonds, such as those found in structures B, E and H. In structure A, both types of hydrogen bonds occur at the same time, which forces the C–O stretching vibrations to appear below 1685 $cm^{-1}$. Structures A, B, E and H also exhibit hydrogen bonds between the indole N–H and the carboxyl C=O groups, which are not observed for Gly-Trp and Trp-Gly. However, no significant red-shift of the $CO_{carb}$ band upon hydrogen-bond formation is observed because the N–$H_{ind}$···O=$C_{carb}$ interaction is only a weak one.

Based on the spectra measured between 1000 and 2000 $cm^{-1}$, conformer **a** is assigned to structure A. The spectrum of structure A is the only calculated spectrum that exhibits no peptide C=O stretching band above 1685 $cm^{-1}$. Due to the coupling between the peptide C=O vibrations in structure A, only one C=$O_{pep}$ band is predicted with appreciable intensity, which agrees well with the experimental results. The fairly accurately predicted bands at 1504 and 1551 $cm^{-1}$ can be assigned to the N–H in-plane bending vibrations at peptide bonds 1 and 2. The higher-frequency N–H in-plane bending vibration results from a hydrogen bond between N–H(2) and C=O(1). Structure B can be ruled out as a possibility for conformer **a**, because its calculated spectrum does not match the experimental spectrum in the C=O stretching region. Once again, only the combination of spectra in the region of the N–H and O–H stretching vibrations with those in the region of the C=O stretching and the N–H and O–H in-plane bending vibrations allows the assignment of the experimental IR spectrum to a single structure.

Unfortunately, we cannot distinguish between structures D and G for conformer **b**, because their calculated spectra are almost identical. It has to be concluded that, as in the case of Trp-Gly, stretched conformations are too similar to be distinguishable by IR spectroscopy alone.

### 2.4. The Nonapeptide DSIP

To further explore the possibilities of structural identification of small gas-phase peptides by means of their IR absorption properties, the IR absorption spectrum of the nonapeptide Trp-Ala-Gly-Gly-Asp-Ala-Ser-Gly-Glu (mass 849 amu) is measured. This peptide, which is also known as the delta-sleep inducing peptide[19] exhibits a rather broad and structureless REMPI spectrum between 34700 and 35000 $cm^{-1}$ (not shown). Figure 8 shows the IR spectrum of DSIP measured with the UV

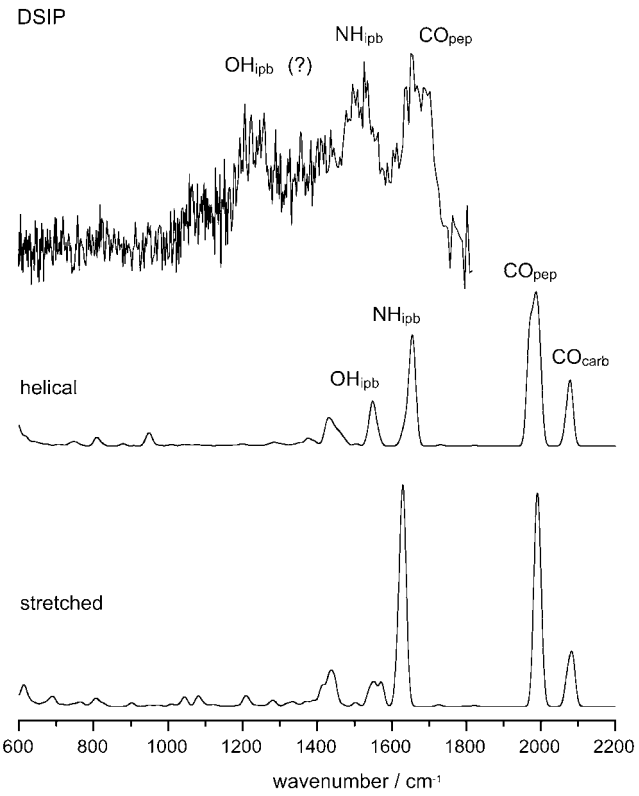


***Figure 8.*** *Gas-phase IR spectrum (top trace) of the delta sleep-inducing peptide (DSIP) Trp-Ala-Glu-Gly-Asp-Ala-Ser-Gly-Glu in the spectral region from 600 to 2000 $cm^{-1}$. Below, spectra calculated at the AM1 level are displayed for a stretched and for a helical structure. The stick spectra are convoluted with a Gaussian profile (fwhm = 15 $cm^{-1}$). Harmonic frequencies are presented unscaled.*

detection laser frequency tuned to 34890 $cm^{-1}$. For the measurement of the IR spectrum, a rather high IR pulse energy of up to 100 mJ is used, which can lead to contributions from IR multiphoton dissociation of the neutral peptide to the depletion of the ion signal.

The band structure observed around 1675 $cm^{-1}$ is attributed to the C=O stretching vibrations of the peptide bonds (amide I bands), and the bands around 1500 $cm^{-1}$ are attributed to the N–H in-plane bending modes (amide II bands). The absorption structure around 1200 $cm^{-1}$ is most probably due to O–H in-plane bending vibrations.

Below the experimental data, two calculated spectra are shown as examples of a helical structure of DSIP (middle trace) and of a stretched structure (lower trace). Both structures, depicted in Figure 9, are submitted to geometry optimisation at the semiemperical AM1 level. The helical structure is more than 0.25 eV more stable than the fully stretched conformation. It must be stressed that one should be cautious with results obtained using semiempirical standard methods. However, the application of ab initio methods is not feasible for systems with the size of DSIP. Although AM1 is known to give a rather poor description of hydrogen-bonded systems in the calculation of harmonic frequencies,[34] calculated IR spectra are shown for both structures in Figure 8. The stick spectrum is

  

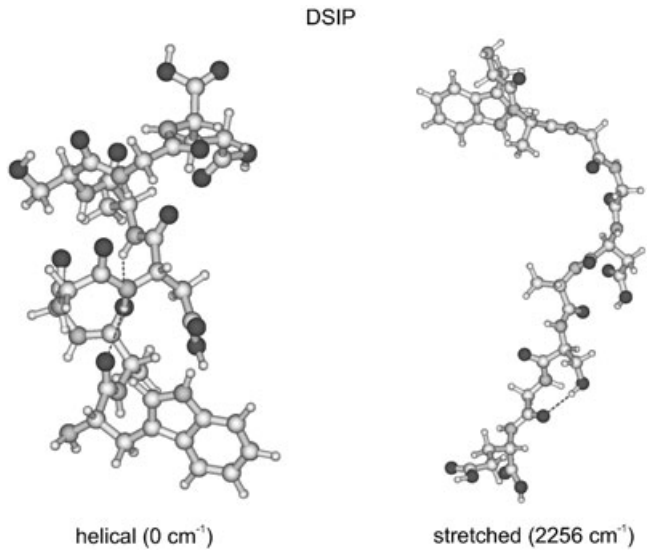


***Figure 9.*** *Calculated structures of DSIP using the semiempirical AM1 method. Relative electronic energies [in $cm^{-1}$] are given in parentheses.*

convoluted with a Gaussian profile of 15 $cm^{-1}$ full-width-at half-maximum (fwhm). Although theoretical methods usually overestimate harmonic frequencies by up to 15%, we have chosen to present unscaled frequencies here.

In the calculated spectra, groups of carboxylic C=O stretching, peptidic C=O stretching (amide I), N–H in-plane bending (amide II) and O–H in-plane bending vibrations are discernible. However, based on the simulations, it is not possible to decide between a stretched or a folded conformation. The experimental band shapes are broader than the width of the IR laser. However, at this point, it cannot be ruled out that a simultaneous probing of several conformers takes place, since the UV spectrum is broad and structureless. Additionally, the temperature of the probed molecules may be higher than expected. Although the experiments are carried out under similar conditions as for the di- and tripeptides, the cooling properties of larger laser-desorbed peptides may be inferior, resulting in a broadening of the spectral lines. Further investigations at improved spectral resolution throughout the IR spectral region are planned to elucidate the structure of DSIP in the gas phase.

## 3. Conclusions and Summary

The O–H in-plane bending, peptidic C=O stretching and, in part, the N–H in-plane bending vibrations appear to be reliable and sensitive fingerprints for hydrogen bonding and peptide folding. The IR spectra between 1000 and 2000 $cm^{-1}$ thus give complementary information to the spectra in the region of the N–H and O–H stretching vibrations and can confirm the hydrogen-bonding character of the carboxylic O–H group, even if the O–H stretching vibration is not observed directly.

By comparing experimental and calculated spectra, it is demonstrated that the O–H in-plane bending mode is an especially sensitive probe for hydrogen-bonding interactions, and that it can reveal information about the type of hydrogen bond formed. The $OH_{ipb}$ frequencies can be classified into the following spectral regions: 1) 1100–1175 $cm^{-1}$: free (non-bonded) carboxyl O–H group. 2) 1250–1350 $cm^{-1}$: weak hydrogen bond between the O–H group and the $\pi$-system of the indole ring (or with the peptide group). 3) 1400–1500 $cm^{-1}$: strong hydrogen bond between the carboxyl O–H group and a carbonyl group. In the harmonic approximation at the B3LYP/6–31G(d,p) level, the frequency shift is mainly a combination of a change in the reduced mass and the force constant, where their ratio remains roughly constant within each group of vibrations. No systematic change in reduced mass and force constant is observed, except in the presence of strong hydrogen bonds where the reduced mass is substantially lowered. Unfortunately, no conformer with a weak O–H···$\pi$ hydrogen bond is observed experimentally, with which our predictions for group (2) frequencies could be tested. Currently, we are investigating di-peptides containing serine, which have a $CH_2OH$ side chain in addition to the carboxylic OH group. These peptides should form a good test for our model predictions in the $OH_{ipb}$ region.

The peptide C=O stretching vibration allows for a clear distinction between free and strongly hydrogen-bonded C=O groups, where an electron lone pair on the oxygen atom acts as proton acceptor. For all systems studied here, the frequency of a hydrogen-bonded C=O stretching vibration is shifted to below 1685 $cm^{-1}$. In contrast to the $OH_{ipb}$ vibrations, the red-shift is dominated by a decrease of the force constant. Thus, a strongly red-shifted O–H stretching vibration, a blue-shifted O–H in-plane bending vibration and a C=O stretching frequency below 1685 $cm^{-1}$ form the fingerprint of a $COOH\cdots O{=}C_{pep}$ hydrogen bond.

In contrast to the peptide C=O group, the frequency of the carboxyl C=O stretching vibration is largely insensitive to hydrogen bonding and bears little structural information. For example, the $CO_{carb}$ band in Trp-Gly-Gly is shifted from 1782 to 1771 $cm^{-1}$ upon hydrogen bonding with the indole N–H group. One should keep in mind, though, that N–H groups form weaker hydrogen bonds than O–H groups, and that the latter cannot be present due to structural constraints. Therefore, the $CO_{carb}$ band is not expected to shift significantly for the systems studied here. IR-UV spectra of Trp-Ser,[35] where a hydrogen bond between the $CH_2OH$ and the carboxylic C=O groups is possible, suggest a red-shift of about 30 $cm^{-1}$, which makes this peptide an ideal candidate for the observation of $CO_{carb}$ shifts. Effects due to *syn*- and *anti*-configurations of the carboxyl group are small and masked by changes in the electron density upon hydrogen-bond formation of the carboxyl C=O or O–H groups.

As in the case of the carboxyl C=O stretching vibration, the peptide N–H in-plane bending vibration is largely insensitive to structural changes. Significant blue-shifts of about 30–40 $cm^{-1}$ are only observed if an N–H···O=C bond is formed. Indole N–H in-plane bending vibrations fall in the same wavenumber region but have negligible absorption strengths. The insensitivity of the $CO_{carb}$ and $NH_{ipb}$ bands make them ideally

  

suited for day-to-day calibration, and also as starter bands when investigating new systems.

In summary, IR spectra are measured of isolated, chemically unmodified peptides, whose conformational landscape can be probed free of interaction with solvent molecules. Both stretched and folded conformers were observed. The O–H in-plane bending and the peptide C=O stretching vibrations are the most sensitive probes for structural assignments in the spectral region between 1000 and 2000 $cm^{-1}$, providing fingerprints for stretched and folded conformations, as well as for the type of hydrogen bond formed by the carboxylic O–H group. The approach also makes the generation of complexes with water molecules possible, and should therefore allow the study of the influence of the solvent in detail, one water molecule at a time.[7,25,36] Spectroscopic measurements of isolated gas-phase peptides at very low temperatures lead to the most direct data sets to calibrate and improve existing ab initio, semiempirical or force-field-based calculations of peptide conformers and protein folding.

## Acknowledgements

*We thank M. Gerhards, M. Schmitt, K. Seefeld and M. Nispel for fruitful discussions. 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 skilful assistance by the FELIX staff. This work has been supported by the Deutsche Forschungsgemeinschaft (Kl 531/24). T. Häber thanks the Deutsche Forschungsgemeinschaft for a postdoctoral research fellowship.*

**Keywords:** conformers · density functional calculations · IR spectroscopy · lasers · peptides

[1] T. R. Rizzo, Y. D. Park, L. A. Peteanu, D. H. Levy, *J. Chem. Phys.* **1986**, *84*, 2534.
[2] L. A. Philips, S. P. Webb, S. J. Martinez, G. R. Fleming, D. H. Levy, *J. Am. Chem. Soc.* **1988**, *110*, 1352.
[3] S. J. Martinez, J. C. Alfano, D. H. Levy, *J. Mol. Spectrosc.* **1992**, *156*, 421.
[4] S. J. Martinez, J. C. Alfano, D. H. Levy, *J. Mol. Spectrosc.* **1993**, *158*, 82.
[5] L. C. Snoek, E. G. Robertson, R. T. Kroemer, J. P. Simons, *Chem. Phys. Lett.* **2000**, *321*, 49.
[6] L. C. Snoek, R. T. Kroemer, M. R. Hockridge, J. P. Simons, *Phys. Chem. Chem. Phys.* **2001**, *3*, 1819.
[7] L. C. Snoek, R. T. Kroemer, J. P. Simons, *Phys. Chem. Chem. Phys.* **2002**, *4*, 2130.
[8] J. M. Bakker, L. Mac Aleese, G. Meijer, G. von Helden, *Phys. Rev. Lett.* **2003**, *21*, 203003 .
[9] M. Gerhards, C. Unterberg, *Phys. Chem. Chem. Phys.* **2002**, *4*, 1760.
[10] M. Gerhards, C. Unterberg, A. Gerlach, *Phys. Chem. Chem. Phys.* **2002**, *4*, 5563 .
[11] B. C. Dian, A. Longarte, S. Mercier, D. A. Evans, D. J. Wales, T. S. Zwier, *J. Chem. Phys.* **2002**, *117*, 10688 .
[12] B. C. Dian, A. Longarte, T. S. Zwier, *J. Chem. Phys.* **2003**, *118*, 2696.
[13] M. Gerhards, C. Unterberg, A. Gerlach, A. Jansen, *Phys. Chem. Chem. Phys.* **2004**, *6*.
[14] C. Unterberg, A. Gerlach, T. Schrader, M. Gerhards, *J. Chem. Phys.* **2003**, *118*, 8296.
[15] W. Chin, M. Mons, J.-P. Dognon, F. Piuzzi, B. Tardivel, I. Dimicoli, *Phys. Chem. Chem. Phys.* **2004**, *6*, 2700.
[16] I. Hünig, K. Kleinermanns, *Phys. Chem. Chem. Phys.* **2004**, *6*, 2650.
[17] R. A. Jockush, R. T. Kroemer, F. Talbot, J. Simons, L. Snoek, P. Carcabal, M. Havenith, J. M. Bakker, I. Compagnon, G. Meijer, G. von Helden, *J. Am. Chem. Soc.* **2004**, *126*, 5709.
[18] J. M. Bakker, I. Compagnon, G. Meijer, G. von Helden, M. Kabelác, P. Hobza, M. S. de Vries, *Phys. Chem. Chem. Phys.* **2004**, *6*, 2810.
[19] G. Schoenenberger, P. Maier, H. Tobler, M. Monnier, *Pflügers Arch.* **1977**, *369*, 99.
[20] B. C. Dian, J. R. Clarkson, T. S. Zwier, *Science* **2004**, *303*, 1169.
[21] R. N. Pribble, A. W. Garrett, K. Haber, T. S. Zwier, *J. Chem. Phys.* **1995**, *103*, 531.
[22] C. Riehn, C. Lahmann, B. Wasserman, B. Brutschy, *Chem. Phys. Lett.* **1992**, *197*, 3197.
[23] S. Tanabe, T. Ebata, M. Fuji, N. Mikami, *Chem. Phys. Lett.* **1993**, *215*, 347.
[24] R. N. Pribble, F. C. Hagemeister, T. S. Zwier, *J. Chem. Phys.* **1997**, *106*, 2145.
[25] C. Janzen, D. Spangenberg, W. Roth, K. Kleinermanns, *J. Chem. Phys.* **1999**, *110*, 9898.
[26] J. M. Bakker, L. Mac Aleese, G. von Helden, G. Meijer, *J. Chem. Phys.* **2003**, *119*, 11180.
[27] D. Oepts, A. F. G. van der Meer, P. van Amersfoort, *Infrared Phys. Technol.* **1995**, *36*, 297.
[28] http://www.rijnh.nl/felix.
[29] Spectra covering the range from 500–1000 $cm^{-1}$ are available as additional material at http://www-public.rz.uni-duesseldorf.de/~pc1/kleinermanns.
[30] C. Jacoby, W. Roth, M. Schmitt, C. Janzen, D. Spangenberg, K. Kleinermanns, *J. Phys. Chem. A* **1998**, *102*, 4471.
[31] Gaussian 98 (Revision A.7), M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. Montgomery, R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P. M. W. Gill, B. G. Johnson, W. Chen, M. W. Wong, J. L. Andres, M. Head-Gordon, E. S. Replogle, J. A. Pople, Gaussian, Inc., Pittsburgh, PA, **1998**..
[32] J. M. Bakker, Ph. D. thesis, Radboud Universiteit Nijmegen **2004**.
[33] http://www-public.rz.uni-duesseldorf.de/~pc1/kleinermanns.
[34] F. Jensen, *Introduction to Computational Chemistry*, John Wiley & Sons Ltd., **1999**.
[35] T. Häber, K. Seefeld, K. Kleinermanns, unpublished results.
[36] J. P. Simons, *Comptes Rendus Chimie* **2003**, *6*, 17.

Received: July 23, 2004
Revised: October 28, 2004

  