Supplementary Information of:

Gas-phase Infrared Spectroscopy of neutral peptides:

Insights from the far-IR and THz domain

Sjors Bakels[1], Marie-Pierre Gaigeot[2], and Anouk M. Rijs[1, *]

[1] Radboud University, Institute for Molecules and Materials, FELIX Laboratory, Toernooiveld 7-c, 6525 ED Nijmegen, the Netherlands, *Email : a.rijs@science.ru.nl

[2] LAMBE CNRS UMR8587, Université d'Evry val d'Essonne, Blvd F. Mitterrand, Bât Maupertuis, 91025 Evry, France

Contents

Section 4

All experimental far-IR spectra (black) of the measured dipeptides are presented in Fig. SI.1 together with their assigned calculated DFT-MD spectra (in color). Data is from ref.[1-3]

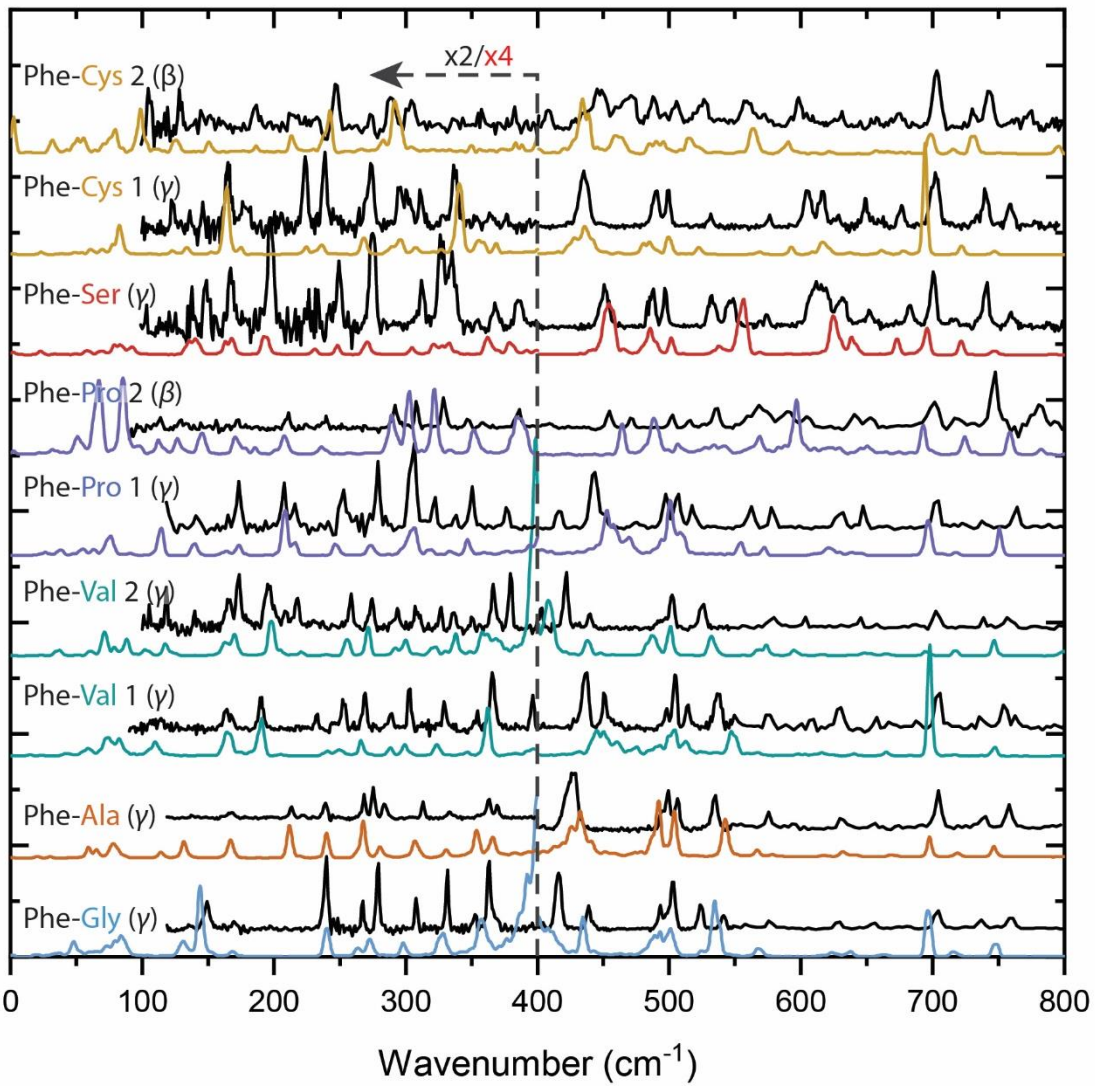

Figure S1: Overview of all experimental (black) far-IR spectra and their assigned DFT-MD spectra (colored traces) of the dipeptides Phe-AA, with AA=Gly, Ala, Val, Pro, Ser and Cys, from bottom to top. The symbol between brackets indicates the type of turn the dipeptide has adapted. All experimental and calculated spectra left to 400 cm^{-1} are multiplied by 2 and 4 times, respectively.

In order to compare the DFT-MD calculations as performed before[2] to static calculations with the same basis set (BLYP-D3 as compared to BLYP in the reference), we performed calculations on 6 dipeptides which are presented in Fig. SI.2. The agreement between the static DFT with the BLYP-D3 functional is rather inconsistent. For example, a surprisingly good agreement is found for Phe-Ala, while for Phe-Pro 2 (and others) the agreement is poor.

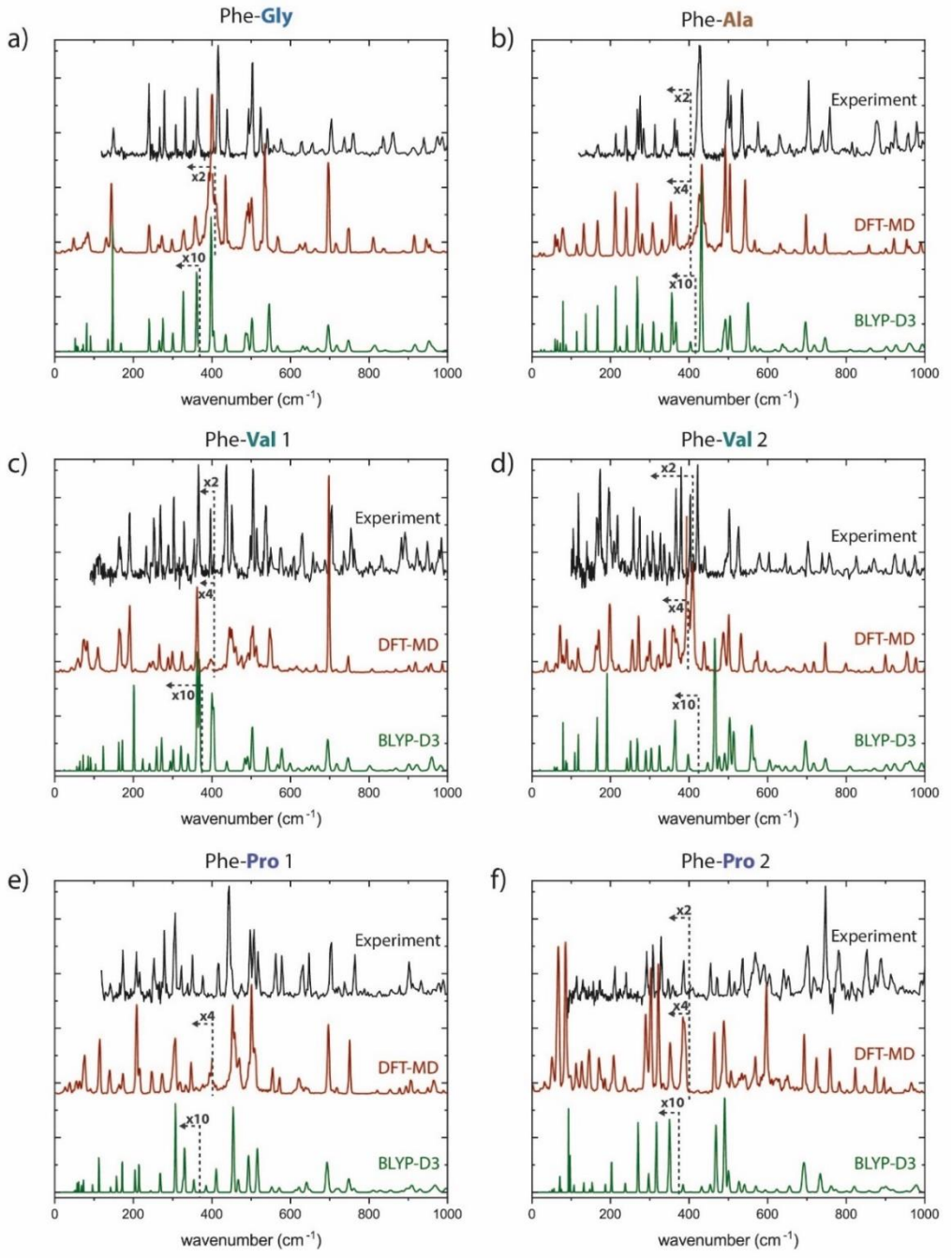

Figure S2: Experimental (black), DFT-MD (red) and static BLYP-D3 calculations as performed for 6 Phe-AA peptides: a) Phe-Gly, b) Phe-Ala, c) Phe-Val 1, d) Phe-Val-2, e) Phe-Pro 1 and f) Phe-Pro 2.

Section 5

Z-Ala REMPI spectra

The REMPI spectra of the monomer (m/z 223.2) and dimer of Z-Ala are presented in Fig. SI.3. The spectrum of the monomer shows one dominant peak at 37589 cm^{-1}, at which the experiments were conducted, and some other, less intense, peaks, either arising from vibrational progression or other conformations. The broad feature around 37570 cm^{-1} is a hot band. The spectrum of the dimer (Fig. SI.3.b), shows broader features, but with a clear maximum at 37420 cm^{-1}, and also either vibrational progression or conformational variety.

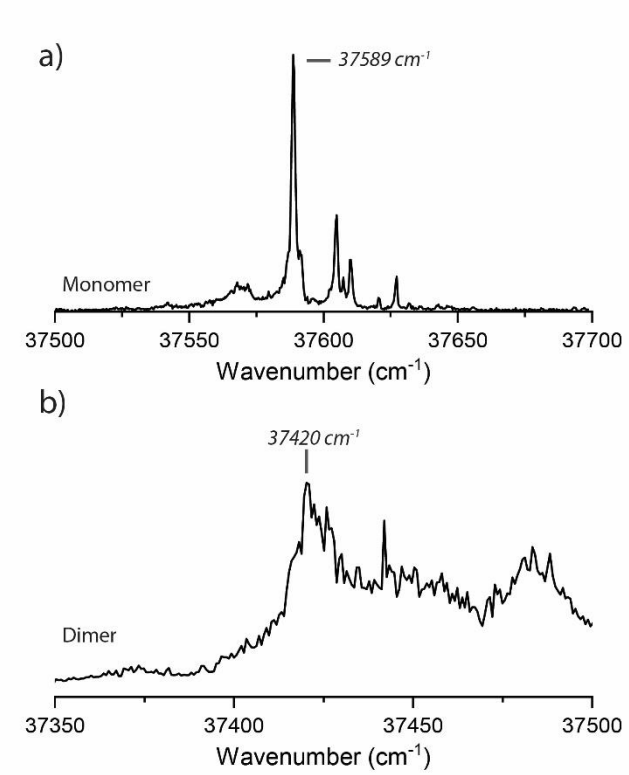

Fig. S3: REMPI spectra of the a) dimer of Z-alanine-OH, and b) the monomer of Z-alanine-OH. The peaks that are used to measure the far-IR spectra are indicated with the wavenumber.

Z-Ala dimer conformational search

An extensive conformational search was performed for the dimer of Z-Ala, with over 50 structures geometry optimized and frequency calculated. All were ordered in structural families, based on their hydrogen bond pattern: A C14 means for example that 14 atoms are enclosed in the ring that is formed by the dimerization of the two strands. In Fig.SI.4 the infrared spectra of the lowest energy structures of each structural family are plotted to show the distinctions in the spectra, with their relative energies in kJ/mol between brackets. These spectra resemble the general features of all spectra of the same family. No match could be found for any of the lower energy structural families, such as the anti-parallel C14 and C10 structures, based mainly on the amide I and II region, still the most used and studied region for secondary structure. Only the non-hydrogen bonded structural family showed good overlap with the experiment in these regions, and, together with a reasonable overlap in the other regions, we therefore assigned the dimer to a non-hydrogen bonded structure, consisting of two not perturbed monomer strands.

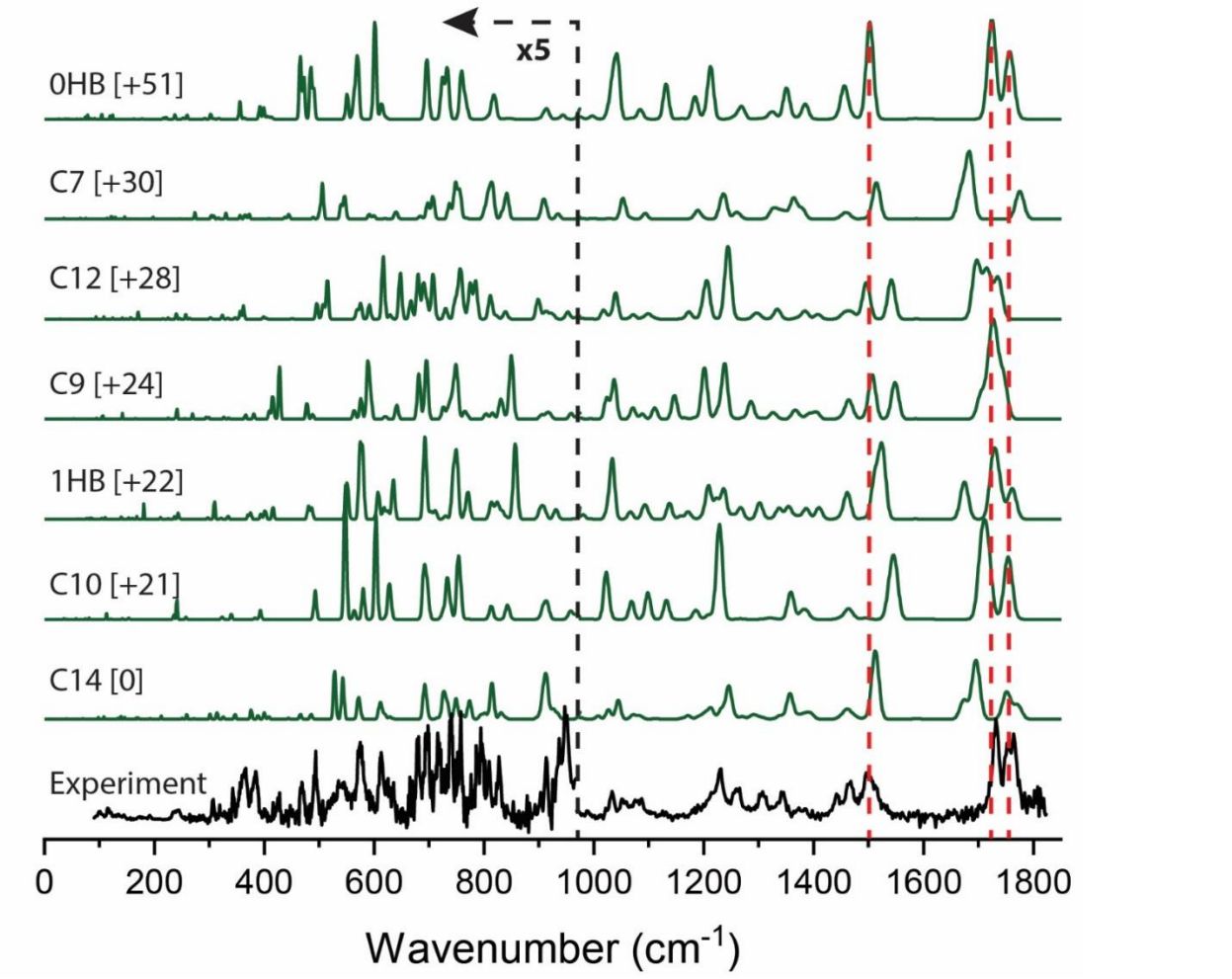

*Fig. S4: Experimental spectrum (black), and spectra of the lowest energy structures of each structural family (green) of the Z-alanine dimer. The spectra are scaled with a scaling factor of 0.976, and were calculated at the B3LYP-D3/6-311+G** level of theory. The red lines indicate the position of the bands in the amide I and II bands of the 0HB structure.*

Graph Theory for delocalized torsional modes

Another motion that we illustrate is in the ~200 cm^{-1} spectral domain, as shown in Fig. SI.5.a. The graph for the g^+ monomer in Fig. SI.5.b illustrates a highly delocalized torsional 200 cm^{-1} mode, where two motions dominate, i.e. the torsion around the C-terminal *–C*11*–O*2– backbone bond and the bending around the 'central' backbone CNC. Each of these two motions is mechanically coupled to several torsions and bending motions, these coupled motions however do not systematically participate to the final IR activity of the mode (there are no values on the associated vertices). Note in the graphs the systematic out-of-phase couplings to the C-O torsion, and the mixing of in-/out-of-phase couplings to the CNC bending. Nicely, the 200 cm^{-1} IR mode of g^+ shows the mechanical coupling of the backbone motions to the bending of the side-chain, with the χ_1 (*C*7*–C*8*–C*11 in the graph) torsion coupled to both the dominant CO torsion and CNC bending. The zero-value on the graph edge associated to this motion shows that it does not participate to the final IR activity. The same motions are responsible for the 192 cm^{-1} mode of g^- on the right side of Fig. SI.5.b, but one can see that the graph of this mode is simpler than the one for g^+, with only the two dominant CO torsion and CNC backbone bending motions (similarities in the red rectangles) dominating the whole mode and providing its final IR intensity.

The IR mode at 197 cm^{-1} in the dimer in Fig. SI.5.c is composed of the two CO torsions from each monomer strand, however uncoupled to each other as the disconnected graphs show. The CO torsion on the g^+ strand is mechanically coupled to the CNC bending on the same strand, the latter not contributing to the final IR activity. Added to the overall IR activity of the dimer mode, the torsion around one of the intermolecular hydrogen bonds is a large contributor, as revealed by the largest value on the *–O*1 ⋯ *H* vertex.

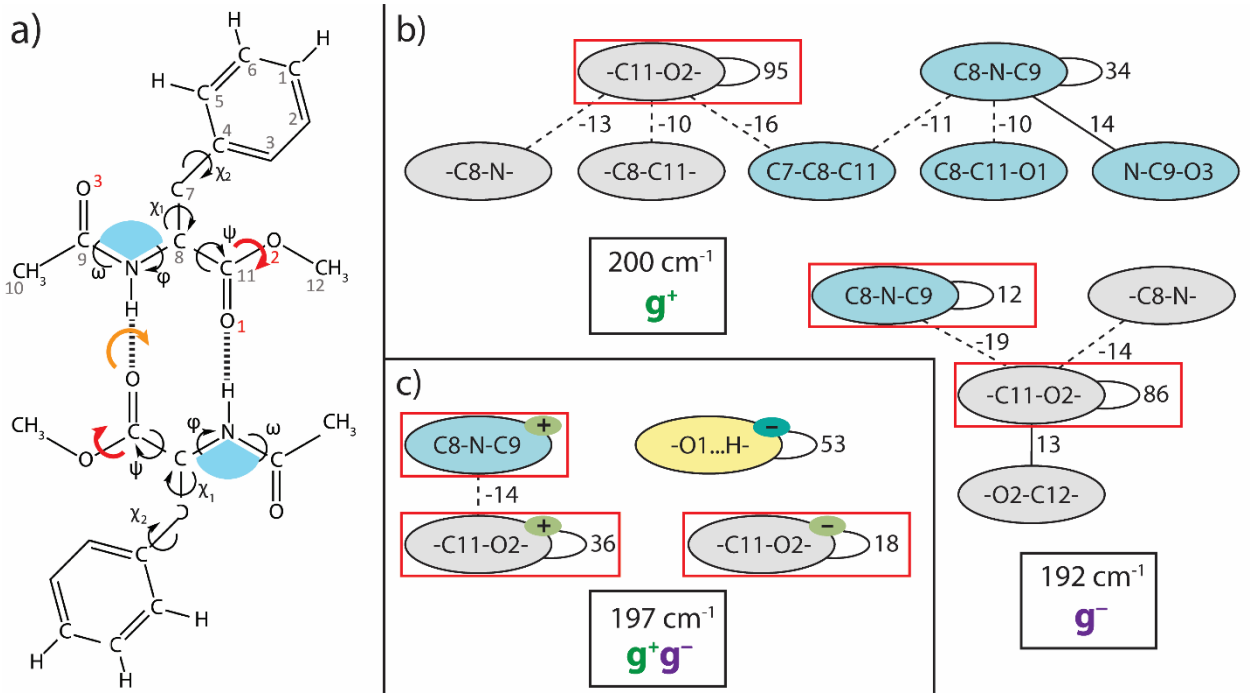

Figure S5: Graphs for the modes as shown in a) at b) 200/192 cm^{-1} respectively for the g^+ and g^- monomers, and c) 197 cm^{-1} for the g^+-g^- dimer. See Fig. 12.d in the main text for the nomenclatures. The colors in a) correspond to the modes in the graphs in b) and c).

Chapter 6

Larger alanine containing peptides

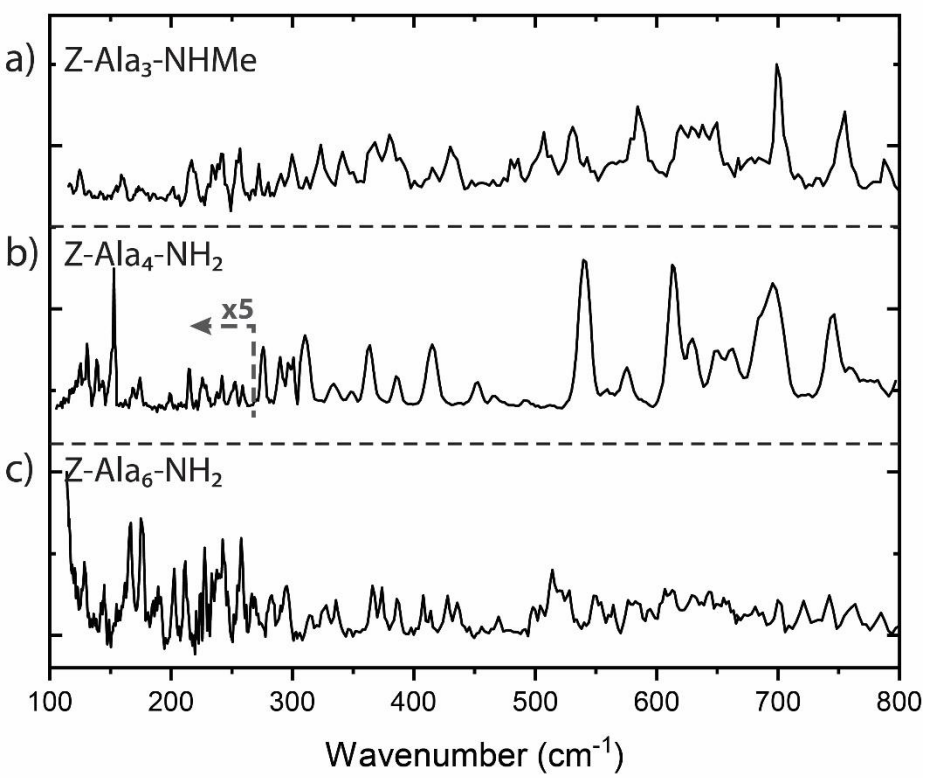

Figure S6: Experimental spectra of a) Z-Ala$_3$-NHMe, b) Z-Ala$_4$-NH$_3$ and c) Z-Ala$_6$-NH$_2$, taken using the IR-UV Ion-Dip method as explained in section 2.1.2.[4]

References

(1) Mahe, J.; Bakker, D. J.; Jaeqx, S.; Rijs, A. M.; Gaigeot, M. P. Mapping Gas Phase Dipeptide Motions in the Far-Infrared and Terahertz Domain. *Phys. Chem. Chem. Phys.* **2017,** *19*, 13778-13787.

(2) Jaeqx, S.; Oomens, J.; Cimas, A.; Gaigeot, M. P.; Rijs, A. M. Gas-Phase Peptide Structures Unraveled by Far-IR Spectroscopy: Combining IR-UV Ion-Dip Experiments with Born-Oppenheimer Molecular Dynamics Simulations. *Angew. Chem. Int. Ed.* **2014,** *53*, 3663-3666.

(3) Mahe, J.; Jaeqx, S.; Rijs, A. M.; Gaigeot, M. P. Can Far-IR Action Spectroscopy Combined with BOMD Simulations be Conformation Selective? *Phys. Chem. Chem. Phys.* **2015,** *17*, 25905-25914.

(4) Jaeqx, S. Ph.D. Dissertation, Radboud University, 2014.