



**HAL**  
open science

## THz to far-infrared spectra of the known crystal polymorphs of Phenylalanine

Thomas Niehaus, Emilien Prost, Vincent Loriot, Franck Lépine, Luc Bergé,  
Stefan Skupin

► **To cite this version:**

Thomas Niehaus, Emilien Prost, Vincent Loriot, Franck Lépine, Luc Bergé, et al.. THz to far-infrared spectra of the known crystal polymorphs of Phenylalanine. *Physical Chemistry Chemical Physics*, 2024, 26 (9), pp.7329-7334. 10.1039/D3CP05805K . hal-04400026v1

**HAL Id: hal-04400026**

**<https://hal.science/hal-04400026v1>**

Submitted on 17 Jan 2024 (v1), last revised 1 Mar 2024 (v2)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Cite this: DOI: 00.0000/xxxxxxxxxx

# THz to far-infrared spectra of the known crystal polymorphs of Phenylalanine

Thomas A. Niehaus,<sup>a‡</sup> Emilien Prost,<sup>a‡</sup> Vincent Lorient,<sup>a</sup> Franck Lépine,<sup>a</sup> Luc Bergé,<sup>b</sup> and Stefan Skupin<sup>a</sup>

Received Date  
Accepted Date

DOI: 00.0000/xxxxxxxxxx

There is renewed interest in the structure of the essential amino acid Phenylalanine in the solid state. Three new polymorphs were found in the years 2012 to 2014. Here, we investigate the structure, stability, and energetical ordering of these phases using first-principles simulations at the level of Density Functional Theory incorporating van-der-Waals interactions. Two of the distinct crystal forms are found to be structurally similar and energetically very close after vibrational free energy corrections have been taken into account. Infrared absorption spectra are likewise calculated and compared to experimental measurements. By combining measurements obtained with a commercial Fourier Transform Infra-Red spectrometer and a homemade air-photonics-based THz Time Domain spectrometer, we could carry out this comparison in the vibrational frequency region from 1 to 40 THz. The excellent agreement of the line positions and the established energy ranking allow us to identify the most stable polymorph of Phenylalanine.

## 1 Introduction

Over the past years, there have been significant advances in X-ray diffraction and subsequent interpretation of observed reflections in terms of the geometrical structure of the specimen<sup>1</sup>. This is accompanied by an improved ability to obtain single crystals of sufficient size even for materials that are considered challenging to crystallize<sup>2</sup>. One example is the essential amino acid Phenylalanine (Phe). Beginning with the measurements of Khawas<sup>3</sup>, a full three-dimensional crystal structure of the D enantiomer was obtained only in 1990 by Weissbuch<sup>4</sup>. This was followed by the rapid discovery of different crystal structures, or polymorphs, of L-Phe. Williams found a new form free of crystal water (**F-II**), starting from hydrated L-Phe in 2013<sup>5</sup>. Shortly thereafter, in 2014, Moussou reported a new monoclinic polymorph (**F-III**)<sup>6</sup>, which was structurally similar to the Weissbuch structure but had lower symmetry. The space group assignment was already questioned earlier through Density Functional Theory calculations<sup>7</sup>. Also in 2014, Ihlefeldt was able to refine the Weissbuch structure, giving rise to the variant (**F-I**) of the L-enantiomer and proposed a new, previously unknown polymorph (**F-IV**)<sup>8</sup>.

Given this rich structural diversity, it is essential to establish independent and alternative ways of structure determination and

characterization. Infrared (IR) absorption spectroscopy can differentiate between phases, and provides information about subtle geometrical changes, such as modifications of hydrogen-bonding patterns. L-Phe has been well studied in the gas phase<sup>9–11</sup>. Still, in the solid state, only the study by Pan and co-workers using THz Time Domain Spectroscopy (THz-TDS) is available, which provided spectra for a limited frequency window from 0.5 to 4.5 THz<sup>12</sup>.

In this work, first-principles Density Functional Theory (DFT) calculations are carried out for all known polymorphs of L-Phe, which provides previously unavailable information on the energetical ordering and stability of the various phases. We also produced simulated THz to far-IR spectra to supply vibrational fingerprints of the crystal variants. These spectra allow direct comparison to experimental data obtained using a commercial Fourier Transform Infra-Red (FTIR) spectrometer and homemade broadband THz-TDS using two-color plasma generation<sup>13,14</sup> coupled with Air-Biased Coherent Detection (ABCD)<sup>15,16</sup>. Using air as a generation and detection medium for THz-TDS allows extended bandwidths while avoiding the inherent phonon absorption appearing in solid-state devices.

## 2 Methods

### 2.1 DFT simulations

Initial crystal structures as published in<sup>8</sup> were retrieved from the Cambridge Structural Database (CSD)<sup>17</sup>. Polymorph **F-I** at 105 K is available under the database identifier QQAUJ05, **F-I** at 298 K as QQAUJ06, **F-IV** as QQAUJ07, and the L-Phenylalanine

<sup>a</sup> Univ Lyon, Université Claude Bernard Lyon 1, CNRS, Institut Lumière Matière, F-69622 Villeurbanne, France, Tel: +33 472 431 571; E-mail: thomas.niehaus@univ-lyon1.fr

<sup>b</sup> Centre Lasers Intenses et Applications, Université de Bordeaux-CNRS-CEA, 33405 Talence Cedex, France

‡ These authors contributed equally to this work.

monohydrate **F-W** as GOFWOP01. We also considered **F-II** published in<sup>5</sup> available as QQQUAUJ04, and **F-III** published in<sup>6</sup> and available as QQQUAUJ03. The experimental cell parameters are summarized in Tab. 1. Furthermore, the crystal structure proposed by King et al.<sup>7</sup> (**F-K**) was taken from the Supplemental Information for that article. Density functional theory simulations under periodic boundary conditions were then carried out using the QUANTUM ESPRESSO code<sup>18,19</sup> employing plane waves as the basis set. We used the Perdew-Burke-Ernzerhof (PBE) exchange-correlation functional together with the PAW pseudopotentials available in the library `pslibrary.1.0.0`. The cutoff values for the wave functions and the density were chosen as 45 a.u. and 360 a.u., respectively, and a convergence threshold of  $0.5 \times 10^{-10}$  a.u. was taken for the determination of the self-consistent field. Since van-der-Waals interactions are assumed to play an essential role in these partially  $\pi$ -stacked molecular crystals, we apply the Grimme dispersion model D3<sup>20</sup> with default parameters. In a benchmark of different dispersion models, this choice was recently shown to provide accurate vibrational spectra in molecular crystals also in the low THz regime<sup>21</sup>. The total energy computed by QUANTUM ESPRESSO contains contributions from the kinetic energy of the electrons, the external (electron-ion), Hartree, exchange-correlation, and van-der-Waals energy, as well as the ion-ion repulsion<sup>18,19</sup>. Since we are interested only in the energy differences between the different polymorphs, lattice energies are not evaluated. The Brillouin zone was sampled with a Monkhorst-Pack set of  $2 \times 2 \times 1$  (**F-I**, **F-K**),  $2 \times 3 \times 2$  (**F-II**, **F-IV**, **F-W**), and  $2 \times 1 \times 2$  (**F-III**). The atomic positions and unit cell parameters were then optimized without any constraints using a tight convergence threshold of  $1 \times 10^{-4}$  a.u.. The resulting structures maintained the initial space group in all cases. As expected, optimizing **F-I** at 298 K and 105 K leads to identical structures, as the DFT simulations correspond to 0 K. The published structure for **F-IV** included three different sets of atomic positions, reflecting the sidechain disorder<sup>8</sup>. We optimized all three structures and obtained slightly different final geometries. Only the lowest energy conformer is discussed further. For each polymorph, the vibration modes and the THz to far-IR absorption spectrum at the  $\Gamma$  point were obtained using density functional perturbation theory implemented in QUANTUM ESPRESSO. After application of the acoustic sum rule, we found no imaginary frequencies, which indicates that all final structures are indeed true minima on the potential energy surface. The spectral intensities have been Gaussian broadened with a linewidth of  $5 \text{ cm}^{-1}$  to facilitate comparison with the experimental data. Having obtained the vibrational frequencies  $\omega_i$  for mode  $i$ , we computed zero point energy corrected total energies  $E_{ZPE} = E_{tot} + \frac{1}{2} \sum_i \hbar \omega_i$ . Likewise, we determined vibrational free energy corrections  $E_{vib} = E_{tot} + F_{vib}$ , with  $F_{vib} = -k_B T \sum_i \ln Q_{vib}^i$ ,  $k_B$  being the Boltzmann constant,  $T$  the temperature and  $Q_{vib}^i$  the partition function for mode  $i$  of a quantum harmonic oscillator. In the evaluation of the free energy corrections, the phonon dispersion should ideally be accounted for by performing a Brillouin zone integration<sup>22</sup>. Our results are evaluated at the  $\Gamma$  point only. Note, that we also do not account for the thermal expansion of the crystal. As shown in Ref.<sup>23</sup>, neglecting this effect may lead to errors up to 1 kJ/mol in absolute terms.

Since we are interested in free energy differences of polymorphs that should feature very similar expansions, we disregard it in the following.

Table 1 Experimental cell parameters of L-Phe.

Form	I Ref. <sup>8</sup>	II Ref. <sup>5</sup>	III Ref. <sup>6</sup>	IV Ref. <sup>8</sup>
space group	P2 <sub>1</sub>	P2 <sub>1</sub>	P2 <sub>1</sub>	C2
a [Å]	8.7829(4)	12.063(11)	6.0010(5)	9.6806(7)
b [Å]	5.9985(3)	5.412(5)	30.8020(17)	5.2362(4)
c [Å]	31.0175(13)	13.676(13)	8.7980(4)	15.8474(12)
$\beta$ [°]	96.9220(12)	99.5976(26)	90.120(4)	96.291(3)
V [Å <sup>3</sup> ]	1622.12(13)	880.3(24)	1626.24(17)	798.46(10)
Z'	4	2	4	1
T [K]	105	294	100	100

## 2.2 Experimental setups

The FTIR measurements were performed using a commercial Bruker Invenio-S spectrometer. This machine uses a polychromatic light source that works in the range of  $400 \text{ cm}^{-1}$  to  $4000 \text{ cm}^{-1}$ . The light goes through a Michelson-type interferometer before being directed through the sample and finally detected. Data acquisition is performed by scanning the distance between the two arms of the interferometer and recording the resulting interferogram. For each position of the interferometer, interference leads to a change in the spectral profile of the beam. It is possible to extract this information by taking the Fourier transform of the interferogram. Given that this interferogram measures the light intensity versus a position corresponding to a time delay between the two arms, we directly obtain the spectral profile in reciprocal space.

To extend the experimentally accessible frequency range below  $400 \text{ cm}^{-1}$ , we use a homemade air-photonics-based THz Time Domain Spectroscopy (THz-TDS) system relying on two-color plasma THz generation. Ultrashort optical pulses of 30 fs at 800 nm produced by a Ti:Sa amplifier are focused in the laboratory atmosphere. A BBO crystal, placed between the focusing lens and the geometrical focus, produces a second color by type I second harmonic generation. An ultra-thin half-waveplate at 800 nm placed after the BBO crystal is used to align the polarization of both colors. Near the geometrical focus, an air plasma is generated by the two-color laser pulse. This plasma emits THz radiation, collected and steered by off-axis parabolic mirrors. The THz field is first focused on the sample and then refocused on the THz Air-Biased Coherent Detection (THz-ABCD) detection system. This detection scheme overlaps the THz field with an IR probe pulse and a high-voltage bias field. Then, second harmonic (SH) radiation is generated through four-wave mixing in air. The SH signal produced is proportional to the THz and bias electric field, and by taking successive measurements for opposite signs of the bias field, one can extract the THz electric field. The THz field temporal profile is then reconstructed by scanning the delay between the THz pulse and the IR probe pulse. Fourier transform gives access to the THz spectrum. Our setup allows us to perform THz-TDS measurements between 4 THz and 30 THz. More details on the specific implementation of our THz-TDS system can be found in<sup>24</sup>.

## 2.3 Sample preparation and measurements

High-purity D-Phenylalanine powder ( $\geq 98\%$ , Sigma-Aldrich) was used in our experiments. The powder grain size is of the order of the wavelengths ( $\sim 10 - 300 \mu\text{m}$ ) in our frequency window, thus diffusing the THz radiation. Therefore, sample preparation was required for the powder to be used with our apparatus. We chose to deposit a thin film of Phe on a silicon wafer substrate by physical vapor deposition. The Phe powder was heated to  $260^\circ\text{C}$  to evaporate under vacuum while the substrate was kept at room temperature, allowing the molecules to condensate on the substrate. We obtained Phe films of about  $10 \mu\text{m}$  in thickness.

The silicon wafer used as the substrate induces a Fabry-Pérot effect because of multiple Fresnel reflections at its interfaces. This has been filtered out with THz-TDS by restricting the time domain to delays below the first back-and-forth reflection in the substrate. For FTIR, however, it led to oscillations in the spectra that were removed by Fourier filtering. This filtering led to a floating baseline in the absorption signal. Therefore, the discussion is focused on the positions of the absorption and relative amplitudes unaffected by these artifacts. Finally, reference measurements employing a silicon wafer without deposition were performed to extract the pure D-Phe transmission spectra from our THz-TDS and FTIR measurements.

## 3 Results and discussion

### 3.1 Calculated structure and energetics

First, we will briefly discuss the relaxed crystal structures shown in Fig. 1. All polymorphs are characterized by hydrogen-bonded bilayers and hydrophobic regions given by the phenyl sidechains. **F-I** and **F-III** are structurally highly similar and contain four molecules *A* to *D* in the asymmetric unit. Monomers differ in the dihedral angle of the phenyl rings (**F-I**: *A*  $93.5^\circ$ , *B*  $59.7^\circ$ , *C*  $57.3^\circ$ , *D*  $96.0^\circ$ ) and (**F-III**: *A*  $94.0^\circ$ , *B*  $59.7^\circ$ , *C*  $57.4^\circ$ , *D*  $95.8^\circ$ ), with *A* being similar to *D* and *B* similar to *C*. As visible in Fig. 1, the two crystal forms can only be differentiated by their monomer arrangement in the hydrophobic region. Transforming **F-I** into **F-III** would require a shift of every other bilayer along the *a* axis. Cuppen and co-workers investigated this phase transition in detail<sup>25</sup>, but found no experimental evidence of its occurrence in the temperature range from 100 K to ambient temperature. Interestingly, though, they reported that polymorph **F-I** could no longer be crystallized at room temperature after growth of **F-III**. This indicates that **F-III** is the kinetically preferred form. Further, Polymorph **F-II** was found by Williams<sup>5</sup> under dry conditions in a study of L-Phe hydrates. The open structure with a smaller density than **F-I/F-III** can host additional water molecules. The recent polymorph **F-IV** is the most compact and features a different stacking with respect to **F-I/F-III**<sup>8</sup>. The DFT calculations can reproduce the experimental lattice parameters satisfactorily. Table 2 shows that the deviation does not exceed 2% for **F-I** to **F-III**, with larger differences for **F-IV**. As reported in Ref.<sup>8</sup>, the latter structure exhibits a large side chain disorder, making structural refinement more difficult. Given that for the studied molecular crystals rather small changes in sidechain stacking can lead to substantial changes in the lattice parameters, van-der-Waals in-

teractions seem well described by the Grimme dispersion model employed here. The optimized crystal structures have been deposited on the Zenodo repository<sup>26</sup>.

Turning now to the energetical ordering of the different phases indicated in Tab. 3, one observes that the high structural similarity of **F-I** and **F-III** translates into a very small energy difference of  $\Delta E_{tot} = 0.186 \text{ kJ/mol}$ . Taking into account the zero point energy (ZPE) and vibrational entropy corrections, the difference even reduces to  $\Delta E_{vib} = 0.064 \text{ kJ/mol}$  at  $T = 300 \text{ K}$ . This indicates that both phases should be present in L-Phe samples in equal amounts unless special care is taken during the crystallization. Our results are in line with the work of Cuppen et al.<sup>25</sup> that also suggests a higher stability of **F-III** than **F-I**. Given the very small energy difference and the inevitable approximations in our DFT approach, a final verdict on the thermodynamic stability of both phases cannot be given at this point. We briefly comment on phase **F-K** found by King et al.<sup>7</sup> via DFT simulations. Although these authors refer to **F-K** as the "correct crystal structure of L-phenylalanine," we find it to be higher in energy ( $\Delta E_{tot} = 0.296 \text{ kJ/mol}$ ) and so far it was not detected in crystallography. Therefore, we do not discuss it further. The experimentally confirmed phases **F-II** and **F-IV** are even higher in energy. In the case of **F-II**, this could be due to the more porous structure leading to reduced dispersion interactions between the phenyl rings. The high energy of **F-IV** is surprising, as this conformer was speculated to be the lowest energy phase in Ref.<sup>8</sup> because of its high density. So far, however, the large structural disorder of this phase was not taken into account (see section 2.1). The three experimentally available structures give rise to configurational entropy ( $S_{conf} = -k_B \sum_i p_i \ln(p_i)$  with  $p_i \propto \exp(-E_{tot,i}/k_B T)$ ), which we can estimate from the DFT energies for the respective structures. The corresponding free energy contribution of  $-2.273 \text{ kJ/mol}$  at  $T = 300 \text{ K}$  is sizable and stabilizes **F-IV**, but does not change the general energy ranking. Furthermore, the experimentally derived occupancies for the three structures (0.471:0.260:0.267) do not match well with the probabilities  $p_i$  obtained from the DFT calculations (0.501:0.079:0.420). A larger deviation of theory and experiment for phase **F-IV** was already mentioned in the context of the lattice parameters and deserves further study.

Table 2 Percent deviation  $[100(x_{cal} - x_{exp})/x_{exp}]$  of calculated lattice parameters from the experimental ones listed in Tab. 1

Form	I	II	III	IV
a	0.455	-1.848	0.791	4.658
b	0.861	0.341	1.105	-1.000
c	1.173	0.777	0.333	1.390
$\beta$	0.132	0.867	0.036	1.481
V	2.486	-1.008	2.245	4.730

### 3.2 THz to far-IR spectra

We compare the simulated THz to far-IR absorption spectra of **F-I** to **F-IV** with the THz-TDS results from Pan *et al.*<sup>12</sup> and our measurements in Fig. 2. Although our experimental sample consisted of the enantiomer D-Phe, we compare it here to calculations for L-Phe. In the framework of this study, both enantiomers are equivalent because conventional infrared spectroscopy involv-

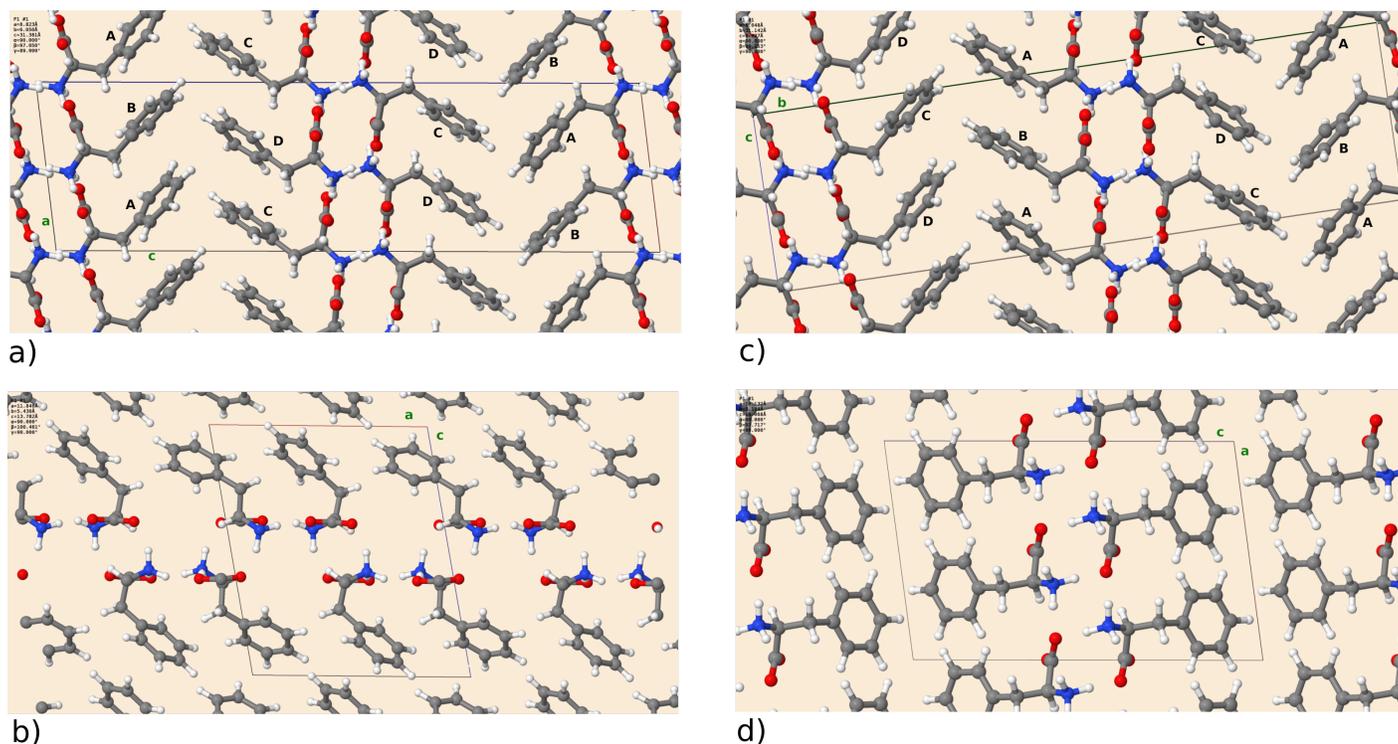


Fig. 1 DFT optimized crystal structures of L-Phe for polymorph a) F-I, b) F-II, c) F-III, and d) F-IV. The unit cell is depicted as a box, and the crystallographic directions (a,b,c) are marked by green letters. F-I and F-III have four distinct L-Phe units, which are labeled A to D.

Table 3 Relative energies in kJ/mol with respect to the lowest energy structure.  $\Delta E_{ZPE}$  contains the vibrational zero point energy and  $\Delta E_{vib}$  contains entropy corrections at  $T = 300$  K.

	$\Delta E_{tot}$	$\Delta E_{ZPE}$	$\Delta E_{vib}$
F-I	0.186	0.144	0.064
F-II	4.452	4.442	6.796
F-III	0.000	0.000	0.000
F-IV	6.400	6.542	7.706
F-K	0.296	N/A	N/A

ing linearly polarized radiation is not sensitive to chirality. The first observation that can be made is that F-I and F-III are not only structurally and energetically close, but their vibrational spectrum is also nearly identical. This holds not only in the depicted region up to 40 THz but also for higher frequencies (see Fig. 4). The general agreement between the experimental results and the simulations is highest for F-I/F-III, which confirms that they are indeed the most stable polymorphs and the main constituents of the sample. In Fig. 2, the low energy region of the spectrum measured by our THz-TDS spectroscopy setup is not well described because of incomplete spatial overlap between the THz field and the probe pulse in the THz-ABCD. This suppresses the THz signal at the lowest frequencies<sup>27</sup> up to approximately 150  $\text{cm}^{-1}$ . Therefore, we compare to the experimental results of Pan et al.<sup>12</sup> in this frequency range. There are two main absorption peaks at 89  $\text{cm}^{-1}$  and 140  $\text{cm}^{-1}$ . Using the eigenvectors of the dynamical matrix for the F-III simulations, we can assign these modes to wagging (90  $\text{cm}^{-1}$ ) and twisting (146  $\text{cm}^{-1}$ ) of phenyl rings, respectively. We further identify the F-III peak at 218  $\text{cm}^{-1}$  as a hydrogen bond stretching motion  $\nu(\text{O-NH}_3)$ . Experimentally, this

peak appears around 200  $\text{cm}^{-1}$ . In the range 250-400  $\text{cm}^{-1}$ , F-III features three important rocking modes, at 308  $\text{cm}^{-1}$  ( $\rho(\text{CH}_2)$ ), 371  $\text{cm}^{-1}$  ( $\rho(\text{NH}_3)$ ) and 388  $\text{cm}^{-1}$  ( $\rho(\text{NH}_3)$ ). The intensity of the latter two modes is overestimated in the simulations compared to the FTIR measurements, where a broad feature is found at 366  $\text{cm}^{-1}$ . The forms F-II and F-IV show large deviations from the F-I/F-III spectrum, which is due to the different hydrogen-bonded network of the hydrophilic bilayer in these structures. Further important peaks in the range up to 800  $\text{cm}^{-1}$  are found at 461  $\text{cm}^{-1}$ , 512  $\text{cm}^{-1}$  and 683  $\text{cm}^{-1}$ , all due to out-of-plane motions of the phenyl rings, which fairly match the corresponding experimental absorption features at 470  $\text{cm}^{-1}$ , 526  $\text{cm}^{-1}$  and 700  $\text{cm}^{-1}$ . The intensity of the mode at 683  $\text{cm}^{-1}$  seems to be overestimated in the simulations. We attribute this fact to the underestimated dispersion of this mode. In fact, the four molecules in the unit cell lead to nearly identical normal modes (682  $\text{cm}^{-1}$  to 685  $\text{cm}^{-1}$ ) with a strong absorption intensity, giving rise to one large peak. We speculate that the FTIR excitations at 683 and 700  $\text{cm}^{-1}$  derive from the same mode, indicating a much larger dispersion and leading to well separated peaks with smaller absorption.

Above this frequency range, notable absorption can be found at 850  $\text{cm}^{-1}$  (wagging of the phenyl hydrogens, calculation at 830  $\text{cm}^{-1}$ ), 1076  $\text{cm}^{-1}$  (phenyl stretching, calculation at 1049  $\text{cm}^{-1}$ ), 1155  $\text{cm}^{-1}$  (wagging of the amino groups, calculation at 1117  $\text{cm}^{-1}$ ) and strong absorption around 1309  $\text{cm}^{-1}$  (C-O stretch, calculation at 1263  $\text{cm}^{-1}$ ). A detailed list of all calculated and measured modes is given in Tab. 4. We find that FTIR and THz-TDS spectroscopy provide identical peak positions (except for the mode at 366  $\text{cm}^{-1}$ ), while the spectral intensity

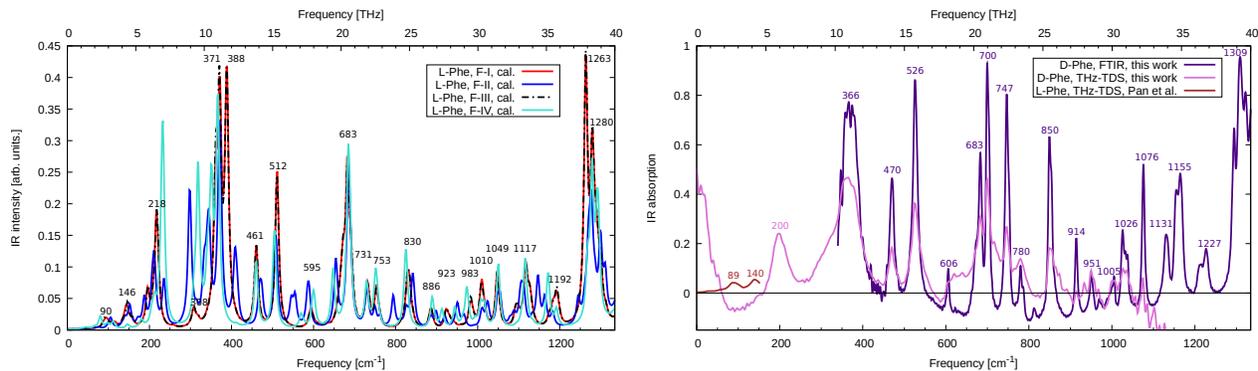


Fig. 2 (Left, theory) Calculated infrared absorption spectrum for L-Phe polymorphs **F-I** to **F-IV**. Spectra are normalized to account for the different numbers of monomers in the unit cell. (Right, experiment) FTIR and THz-TDS infrared absorption spectrum for D-Phe together with the results for L-Phe from Pan et al.<sup>12</sup>.

is more reliably given at the FTIR level. This holds in particular above  $900\text{ cm}^{-1}$  where the bandwidth limit of the THz-TDS is reached. The mean absolute error of the calculated frequencies for **F-III** with respect to the FTIR data is  $24\text{ cm}^{-1}$ . Scaling the frequencies by a factor of 1.03 to account for systematic errors in the DFT calculation reduces the error to only  $6\text{ cm}^{-1}$ . Such scaling is routinely performed for simulations in the gas phase<sup>28</sup> but is less common for molecular crystals, though it works quite well in the present example.

Simulations of the monohydrate **F-W** (see Fig. 3 and 4) provide a complex spectrum significantly different from the experimental results and simulations. This is also true for **F-II**, structurally similar to **F-W**, but without the additional water molecule. This indicates that the studied samples contain pure D-Phe, as expected.

## 4 Conclusions

Extensive DFT simulations of all known crystal polymorphs of Phenylalanine have been carried out. Experimentally known lattice parameters are well reproduced, and van-der-Waals interaction seems well described by the Grimme dispersion model. An energetical ordering of the different phases is established, where the polymorphs (**F-I**), Ref.<sup>8</sup>, and (**F-III**), Ref.<sup>6</sup>, feature the lowest energies. They are found to be very close structurally and energetically, and their simulated vibrational spectra are nearly identical. They best agree with our experimental spectra obtained with the FTIR and THz-TDS methods in the frequency range of 1 to 40 THz from our thin-film deposition sample. This confirms that (**F-I**) and (**F-III**) are indeed the most stable polymorphs of Phenylalanine. Our work demonstrates how advanced numerical and experimental techniques can be used in concert to reveal the structure of complex organic molecules in the solid state. We believe that detailed structural information on the amino acid Phenylalanine obtained in our study will be helpful for further research.

## Acknowledgements

We acknowledge the Agence Innovation Defense (AID) and Agence Nationale de la Recherche (ANR-19-ASMA-0007-ALTESSE2) for providing financial support and thank GENCI for

computational resources under projects DARI A0110810637 and A0130810637. We thank Eric Constant and Isabelle Compagnon for stimulating discussions.

## Notes and references

- 1 R. Fornari and M. Roth, *MRS bulletin*, 2009, **34**, 239–244.
- 2 Z. K. Nagy and R. D. Braatz, *Annual review of chemical and biomolecular engineering*, 2012, **3**, 55–75.
- 3 B. Khawas, *Acta Crystallographica Section B: Structural Crystallography and Crystal Chemistry*, 1970, **26**, 1919–1922.
- 4 I. Weissbuch, F. Frolow, L. Addadi, M. Lahav and L. Leiserowitz, *Journal of the American Chemical Society*, 1990, **112**, 7718–7724.
- 5 P. A. Williams, C. E. Hughes, A. B. Buanz, S. Gaisford and K. D. Harris, *The Journal of Physical Chemistry C*, 2013, **117**, 12136–12145.
- 6 E. Mossou, S. C. Teixeira, E. P. Mitchell, S. A. Mason, L. Adler-Abramovich, E. Gazit and V. T. Forsyth, *Acta Crystallographica Section C: Structural Chemistry*, 2014, **70**, 326–331.
- 7 M. D. King, T. N. Blanton and T. M. Korter, *Physical Chemistry Chemical Physics*, 2012, **14**, 1113–1116.
- 8 F. S. Ihlefeldt, F. B. Pettersen, A. Von Bonin, M. Zawadzka and C. H. Görbitz, *Angewandte Chemie International Edition*, 2014, **53**, 13600–13604.
- 9 L. Snoek, E. Robertson, R. Kroemer and J. Simons, *Chemical Physics Letters*, 2000, **321**, 49–56.
- 10 G. Von Helden, I. Compagnon, M. Blom, M. Frankowski, U. Erlekam, J. Oomens, B. Brauer, R. Gerber and G. Meijer, *Physical Chemistry Chemical Physics*, 2008, **10**, 1248–1256.
- 11 B. Ravikumar, R. Rajaram and V. Ramakrishnan, *Journal of Raman Spectroscopy: An International Journal for Original Work in all Aspects of Raman Spectroscopy, Including Higher Order Processes, and also Brillouin and Rayleigh Scattering*, 2006, **37**, 597–605.
- 12 T. Pan, S. Li, T. Zou, Z. Yu, B. Zhang, C. Wang, J. Zhang, M. He and H. Zhao, *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 2017, **178**, 19–23.
- 13 M. Kress, T. Löffler, S. Eden, M. Thomson and H. G. Roskos,

- Opt. Lett.*, 2004, **29**, 1120–1122.
- 14 K. Y. Kim, J. H. Glowina, A. J. Taylor and G. Rodriguez, *Opt. Express*, 2007, **15**, 4577–4584.
- 15 N. Karpowicz, J. Dai, X. Lu, Y. Chen, M. Yamaguchi, H. Zhao, X.-C. Zhang, L. Zhang, C. Zhang, M. Price-Gallagher, C. Fletcher, O. Mamer, A. Lesimple and K. Johnson, *Appl. Phys. Lett.*, 2008, **92**, 011131.
- 16 I.-C. Ho, X. Guo and X.-C. Zhang, *Opt. Express*, 2010, **18**, 2872–2883.
- 17 C. R. Groom, I. J. Bruno, M. P. Lightfoot and S. C. Ward, *Acta Crystallographica Section B: Structural Science, Crystal Engineering and Materials*, 2016, **72**, 171–179.
- 18 P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo *et al.*, *Journal of physics: Condensed matter*, 2009, **21**, 395502.
- 19 P. Giannozzi, O. Andreussi, T. Brumme, O. Bunau, M. B. Nardelli, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, M. Cococcioni *et al.*, *Journal of physics: Condensed matter*, 2017, **29**, 465901.
- 20 S. Grimme, J. Antony, S. Ehrlich and H. Krieg, *The Journal of chemical physics*, 2010, **132**, 154104.
- 21 N. Bedoya-Martínez, A. Giunchi, T. Salzillo, E. Venuti, R. G. Della Valle and E. Zojer, *Journal of chemical theory and computation*, 2018, **14**, 4380–4390.
- 22 J. Nyman and G. M. Day, *CrystEngComm*, 2015, **17**, 5154–5165.
- 23 Y. N. Heit and G. J. Beran, *Acta Crystallographica Section B: Structural Science, Crystal Engineering and Materials*, 2016, **72**, 514–529.
- 24 E. Prost, V. Lorient, E. Constant, I. Compagnon, L. Bergé, F. Lépine and S. Skupin, *Review of Scientific Instruments*, 2022, **93**, 033004.
- 25 H. M. Cuppen, M. M. Smets, A. M. Krieger, J. A. van den Ende, H. Meekes, E. R. van Eck and C. H. Görbitz, *Crystal Growth & Design*, 2019, **19**, 1709–1719.
- 26 T. Niehaus, *DFT optimized crystal structures for different polymorphs (F-I to F-IV, F-W) of L-Phenylalanine in Crystallographic Information File (cif) format*, 2023, <http://dx.doi.org/10.5281/zenodo.10057575>.
- 27 Prost, Emilien, Lorient, Vincent, Constant, Eric, Compagnon, Isabelle, Bergé, Luc, Lépine, Franck and Skupin, Stefan, *Eur. Phys. J. Spec. Top.*, 2022, 1–10.
- 28 J. P. Merrick, D. Moran and L. Radom, *The Journal of Physical Chemistry A*, 2007, **111**, 11683–11700.

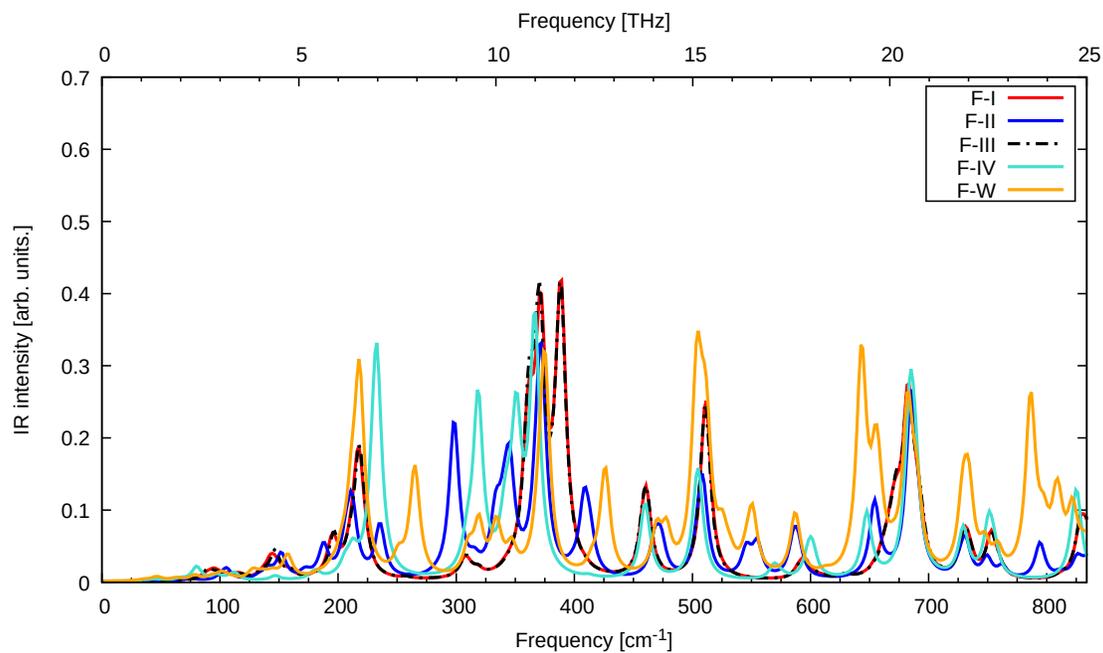


Fig. 3 Calculated absorption spectrum of all considered polymorphs **F-I** to **F-IV** together with the monohydrate **F-W** up to 25 THz. The intensity is normalized for the number of molecules in the simulation cell.

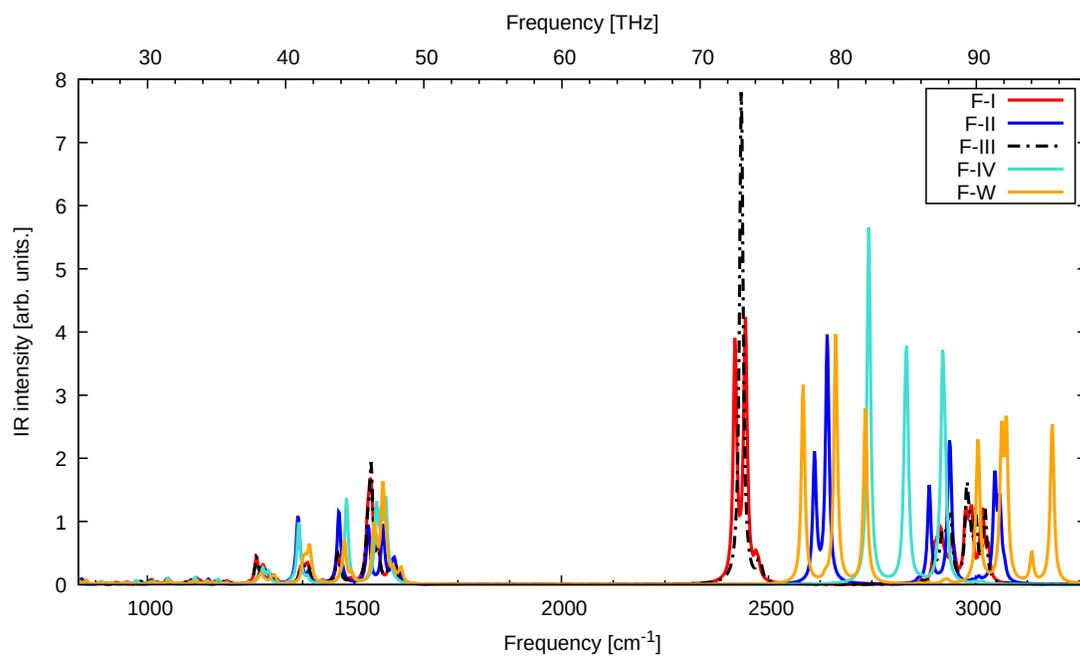


Fig. 4 Calculated absorption spectrum of all considered polymorphs **F-I** to **F-IV** together with the monohydrate **F-W** between 25 and 100 THz. The intensity is normalized for the number of molecules in the simulation cell.

theory (F-III)	experiment (THz-TDS)	experiment (FTIR)	mode description
90			$\omega$ (Phe)
146			$\tau$ (Phe)
158			$\rho$ (Phe)
197			$\rho$ (Phe)
218	200		$\nu$ (O-NH <sub>3</sub> )
308			$\rho$ (CH <sub>2</sub> )
371	370	366	$\rho$ (NH <sub>3</sub> )
388		374	$\rho$ (NH <sub>3</sub> )
461	470	470	$\gamma$ (Ph)
512	526	526	$\gamma$ (Ph)
595	606	606	$\tau$ (NH <sub>3</sub> )
683	700	700	$\gamma$ (Ph)
731	747	747	$\omega$ (H-Ph)
753	780	780	$\gamma$ (CO <sub>2</sub> )
830	853	850	$\omega$ (H-Ph)
886	914	914	$\tau$ (H-Ph)
923		951	$\delta$ (CH-NH <sub>3</sub> )
983		1005	$\omega$ (NH <sub>3</sub> )
1010		1026	$\nu$ (Ph)
1049		1076	$\nu$ (Ph)
1096		1131	$\nu$ (C-N)
1117		1155	$\omega$ (NH <sub>3</sub> )
1192		1227	$\nu$ (C-C)
1263		1309	$\nu$ (C-O)
1280		1322	$\nu$ (C-C)

Table 4 Vibrational frequencies of Phenylalanine up to 1300 cm<sup>-1</sup> from the DFT calculations for conformer **F-III** and the two experimental techniques (THz-TDS, FTIR) used in this article in units of wavenumbers (cm<sup>-1</sup>). The abbreviations in the approximate mode description stand for:  $\tau$ : twisting,  $\rho$ : rocking,  $\omega$ : wagging,  $\nu$ : stretching,  $\delta$ : scissoring,  $\gamma$ : out-of-plane motion, Ph: phenyl ring, Phe: Phenylalanine.