Proceedings Paper:
Burnett, AD orcid.org/0000-0003-2175-1893 and Kendrick, J (2022) Understanding The THz Spectrum Of L-cysteine. In: Proceedings of the 2022 47th International Conference on Infrared, Millimeter and Terahertz Waves (IRMMW-THz). 2022 47th International Conference on Infrared, Millimeter and Terahertz Waves (IRMMW-THz), 28 Aug - 02 Sep 2022, Delft, Netherlands. IEEE . ISBN 978-1-7281-9428-8

https://doi.org/10.1109/IRMMW-THz50927.2022.9896024

Understanding the THz spectrum of L-cysteine.

Andrew D. Burnett[1], John Kendrick[1]
[1]School of Chemistry, University of Leeds, Leeds, West Yorks, LS2 9JT, UK

***Abstract*—In this article we look at understanding the THz spectra of L-cysteine. L-cysteine exhibits a high degree of polymorphism with Form I being the most stable. This polymorph however shows disorder above 30 K via rotation around the C-C-S-H torsion angle. We aim to understand how this disorder influences the THz spectrum using a combination of static DFT phonon calculation of a number of different supercells along with molecular dynamics simulations of appropriate structures.**

I. Introduction

THE amino acid L-cysteine is one of the building blocks required for the synthesis of proteins and is particularly important as it forms disulfide bridges which are structurally important in a larger number of proteins. L-cysteine is known to crystallise into two different polymorphs at ambient pressure. Form I [1] is orthorhombic ($P2_12_12_1$) with one molecule in the asymmetric unit. At temperatures below 30 K [2] all molecules are arranged with a S-H...S hydrogen bonding pattern (see Fig 1a) with a C-C-S-H torsion angle (τ) of 77.6°, however at higher temperatures disorder arises allowing the torsional angle to rotate to −85.4° forming the S–H...O hydrogen bonding pattern seen in Fig 1b. Therefore at room temperature, a powder sample of Form I is disordered with a roughly equal amount of both S–H...S and S–H...O bonding patterns, which may be dispersed randomly throughout the crystal or they may be localised into domains. Form II [3] is monoclinic ($P2_1$) with two independent molecules in the asymmetric unit. The two molecules differ in the orientation of S–H bond, one of which forms a hydrogen bond with sulfur and the other with oxygen meaning both hydrogen bonding motifs (Fig 1) are observed.

In order to understand how the hydrogen bonding pattern influences the calculated spectra static simulations of Form I and II were performed along with a second Form I structure with the S–H...O bonding motifs. Additional static calculations were then performed on supercells containing up to 32 molecules where the population of S–H...S and S–H...O hydrogen-bonding motifs were randomly distributed with a roughly equal split across the supercell [2]. Finally these large supercells were used as the starting point for molecular dynamics simulations for additional spectral calculations.

Static Density Functional Theory (DFT) calculations were performed with the VASP package [4] using the Perdew-Burke-Ernzerhof (PBE) functional [5] and the Projector Augmented Wave (PAW) pseudo-potentials [6] distributed with VASP 5.4.1. Dispersion corrections were included using the method described by Grimme [7] with the damping functions suggested by Becke and Johnson [8] referred to as GD3/BJ. A plane-wave cutoff of 600 eV and a reciprocal space k-point resolution of 0.2 $Å^{-1}$ were chosen, The wavefunction was converged so that the energy changed by less than 1×10^{-6} eV and the atom positions and unit cell dimensions were optimised so that all

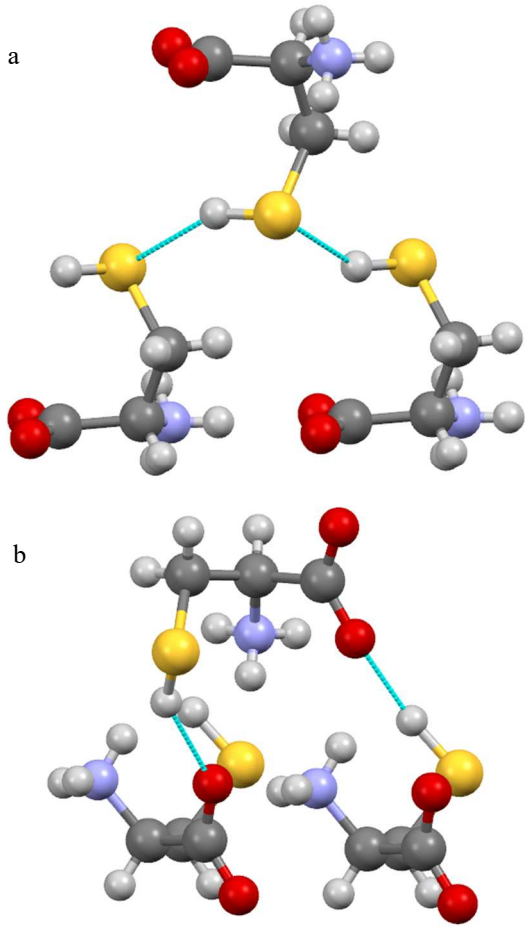

Fig. 1. (a) shows the S-H...S Hydrogen Bonding Pattern while (b) shows the S-H...O Hydrogen Bonding Pattern

forces were less than 5×10^{-3} eV/Å. Full geometry optimisations of the unit-cell and the molecular geometries, were performed. The optimisations were terminated when the internal coordinate changes were less than 0.02 (Bohr or radians), the gradients were less than 0.0004 (Hartree/Bohr or Hartree/radian) and the energy change was less than 1.0×10^{-5} Hartree. Phonopy [9] using VASP as the DFT package was used to calculate the phonon frequencies while the Born charges and the high frequency permittivity of the cells were calculated using VASP.

Ab Initio molecular dynamics calculations were performed with CP2K package [10]. All calculations used the PBE exchange-correlation functional [5] along with the GD3/BJ dispersion correction [6,7]. The Molopt double zeta valence plus polarisation basis sets [11] were used with the GTH PBE pseudopotentials. [12] The cutoff and relative cutoff parameters for all calculations were 450 and 50 Rydberg respectively.

Molecular dynamics calculations were performed using both the NPT and NVT ensembles. The Nose thermostat was used to constrain the temperature with a time constant of 100 fs and the pressure was constrained with a Nose-Hoover-Chain with the same time constant. The time step for the simulations was 0.5 fs. After equilibration of the NPT calculations the average cell dimensions were determined from a simulation of at least 3 ps. The average cell dimensions from the NPT calculations were then used in the NVT ensemble simulations. NVT equilibration was at least 2 ps before a production run of at least 30 ps

II. Results

Fig. 2 shows the calculated THz spectra from static DFT calculations of single unit cells of Form I with S–H...S and S–H...O bonding patterns where the spectra were generated using the post-processing tool PDielec [13,14]. Clearly both spectra show significant differences but neither correlate particularly well with experimental spectra, even those recorded at low temperature [15]. The effect of hydrogen-bonding motif on the spectra was further explored using a number of supercells with randomly distributed motifs. The black trace in Fig. 3 shows the calculated spectra from a 2x1x4 supercell of Form I which is typical of all the large supercells calculations performed. Also shown in red is the experimental spectra recorded at room temperature presented previously in [15]. Correlation between calculation and experiment is reasonable above ~150 cm^{-1} with modes generally calculated at higher frequencies to experiment. This is likely a combination of the lack of anharmonic or temperature effects included in the static calculations. However, at frequencies below 100 cm^{-1} the calculation shows a large number of modes that show poor correlation with experiment.

Finally the blue trace in Fig. 3 is calculated using the post processing tool TRAVIS [16] from the NVT MD simulation of the 2x1x4 supercell. Above ~150 cm^{-1} the static and dynamic calculations are similar with the dynamic calculation showing modes shifted to lower frequencies because of the inclusion of both temperature and anharmonicity in the calculation. Below 100 cm^{-1} the blue trace shows 2 broad features which correlate well with experiment. Interestingly in MD runs we clearly see some of the dihedral angles flip between two states during the simulation and it appears to be these dynamics that allow for the improved correlation between experiment and calculation below100 cm^{-1}.

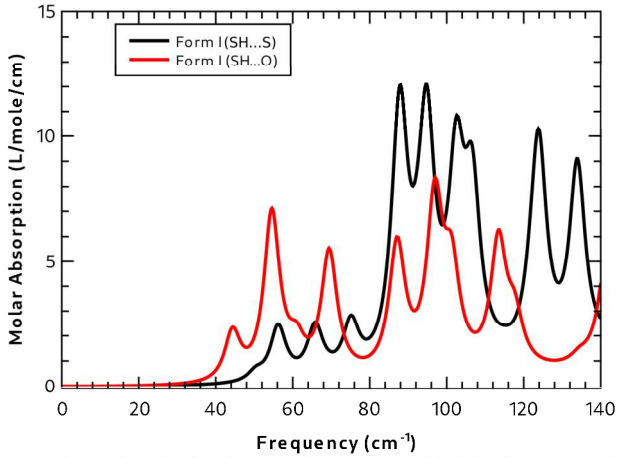

Fig. 2. Shows the calculated spectra for Form I. In black is the spectra with the S-H...S bonding pattern which is the most stable structure below 30 K while in red is the structure with only the S-H...O bonding motif.

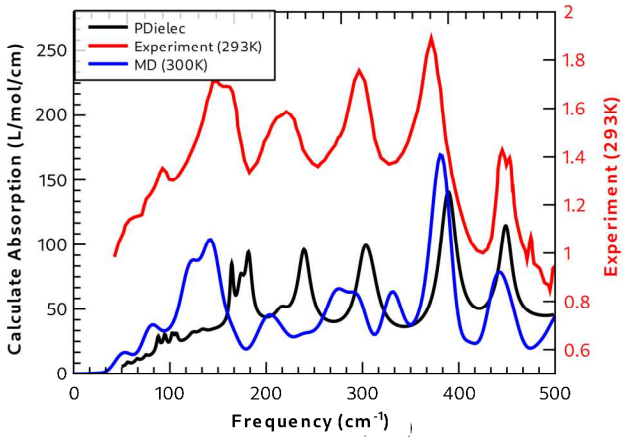

Fig. 3. shows the experimental spectra (red) of cysteine recorded at 293 K taken from [15] along with the spectra calculated from a static DFT calculation from a disordered supercell (black) and a MD simulation at 300 K (blue)

References

[1] K.A. Kerr, J.P. Ashmore, "Structure and conformation of orthorhombic l-cysteine," *Acta Crystallographica Section B,* vol. 29, pp. 2124–2127, 1973.

[2] S.A. Moggach, S.J. Clark, S. Parsons, "l-Cysteine-I at 30 K," *Acta Crystallographica Section E,* vol. 61, pp. o2739–o2742, 2005.

[3] M.M. Harding, H.A. Long, "The crystal and molecular structure of l-cysteine," *Acta Crystallographica Section B,* vol. 24, pp. 1096–1102, 1968.

[4] J. Hafner, "Ab-initio simulations of materials using VASP: Density-functional theory and beyond," *Journal of Computational Chemistry,* vol. 29, pp. 2044–2078, 2008.

[5] J.P. Perdew, K. Burke, M. Ernzerhof, "Generalized Gradient Approximation Made Simple," *Physical Review Letters,* vol. 77, pp. 3865–3868, 1996.

[6] Blöchl, P. E. "Projector augmented-wave method," *Physical Review B,* vol. 50, 17953, 1994.

[7] S. Grimme, S. Ehrlich, L. Goerigk, "Effect of the damping function in dispersion corrected density functional theory," *Journal of Computational Chemistry*, vol. 32, pp. 1456–1465, 2011.

[8] A.D. Becke, E.R. Johnson, "A density-functional model of the dispersion interaction," *The Journal of Chemical Physics*, vol. 123, 154101, 2005.

[9] A. Togo, I. Tanaka, "First principles phonon calculations in materials science," *Scr. Mater,* vol. 108, pp. 1–5, 2015.

[10] T. D. Kühne, *et al*. "CP2K: An electronic structure and molecular dynamics software package -Quickstep: Efficient and accurate electronic structure calculations," *Journal of Chemical Physics*, vol. 152, 194103, 2020.

[11] J. VandeVondele, J. Hutter, "Gaussian basis sets for accurate calculations on molecular systems in gas and condensed phases," *Journal of Chemical Physics* vol 127, 114105, 2007.

[12] S. Goedecker, M. Teter, "Separable dual-space Gaussian pseudopotentials," *Physical Review B - Condensed Matter and Materials Physics*, vol. 54, 1753, 1996.

[13] J. Kendrick, A.D. Burnett, "PDielec: The calculation of infrared and terahertz absorption for powdered crystals," *Journal of Computational Chemistry*, vol. 37, pp. 1491–1504, 2016.

[14] J. Kendrick, A.D. Burnett, "Exploring the Reliability of DFT Calculations of the Infrared and Terahertz Spectra of Sodium Peroxodisulfate," *Journal of Infrared, Millimeter, and Terahertz Waves*, vol. 41, pp. 382-413, 2020.

[15] T. Korter, R. Balu, M. Campbell, M. Beard, S. Gregurick, E. Heilweil, "Terahertz spectroscopy of solid serine and cysteine," *Chemical Physics Letters* vol. 418, pp. 65–70, 2006.

[16] M. Brehm, B. Kirchner, "TRAVIS - A Free Analyzer and Visualizer for Monte Carlo and Molecular Dynamics Trajectories," *Journal of Chemical Information and Modeling*, vol. 51, pp. 2007–2023, 2011.