Proceedings Paper:
Towler, CNA, Kendrick, J and Burnett, AD orcid.org/0000-0003-2175-1893 (2021) Understanding the Effect of Dispersion Corrections on the Calculated Spectra of α-Lactose Monohydrate. In: 2020 45th International Conference on Infrared, Millimeter, and Terahertz Waves (IRMMW-THz). 2020 45th International Conference on Infrared, Millimeter and Terahertz Waves (IRMMW-THz), 08-13 Nov 2020, Buffalo, NY, USA. IEEE , pp. 585-586. ISBN 9781728166209

https://doi.org/10.1109/irmmw-thz46771.2020.9370997

Understanding the Effect of Dispersion Corrections on the Calculated Spectra of α-Lactose Monohydrate.

Calum N. A. Towler[1], John Kendrick[1], and Andrew D. Burnett*[1]
[1]School of Chemistry, University of Leeds, Leeds, LS2 9JT UK
* Email: A.D.Burnett@leeds.ac.uk

2020 45th International Conference on Infrared, Millimeter and Terahertz Waves (IRMMW-THz) | 978-1-7281-6620-9/20/$31.00 ©2020 IEEE | DOI: 10.1109/IRMMW-THZ46771.2020.9370997

***Abstract*—THz spectra were recorded for a polycrystalline α-lactose monohydrate sample at 4 K. Solid-state DFT was then used to aid the interpretation of the spectrum. In particular, we investigate how the inclusion of a number of empirical dispersion corrections to the DFT calculations influences the correlation between calculation and experiment.**

I. INTRODUCTION

WITH the rapid development of terahertz time-domain spectroscopy (THz-TDS) in the last decade, the next challenge is the accurate interpretation and analysis of the spectra of crystalline organic molecules. This is often complicated, owing to the size of many of the molecules of interest which results in many IR active phonon modes. Phonon modes typically involve complex motion of the entire molecule and are extremely sensitive to their surrounding environment, hence making interpretation of the spectra without the aid of calculation impossible. The experimental spectrum was taken using a custom-built THz time domain spectrometer described elsewhere [1]. The sample was cooled to 4 K to maximise correlation between experimental and calculated spectra as at standard Density Functional Theory (DFT) calculations do not incorporate temperature effects.

Owing to the crystalline and periodic nature of most systems of interest, solid-state DFT calculations are used to interpret the spectra to understand the origin of all the phonon modes contributing to the spectra. The original crystal structure that was used to begin the calculations was obtained from the Cambridge Crystallographic Database [2]. All the calculated spectra presented here use VASP [3] to calculate an optimized structure and Born charges (which are required to calculate infrared intensities) while the Hessian matrix and phonon frequencies are calculated using Phonopy [4]. For each calculation the PBE functional [5] and the PAW pseudopotential [6] was used. The Monkhorst-Pack [7] method was utilised for the description of k-points. The calculated spectra are then generated by post processing the VASP and Phonopy outputs using PDielec which is a DFT post-processing tool that calculates the permittivity of the effective medium measured while also including sampling effects such as particle shape and mie-scattering when calculating the THz spectra [8].

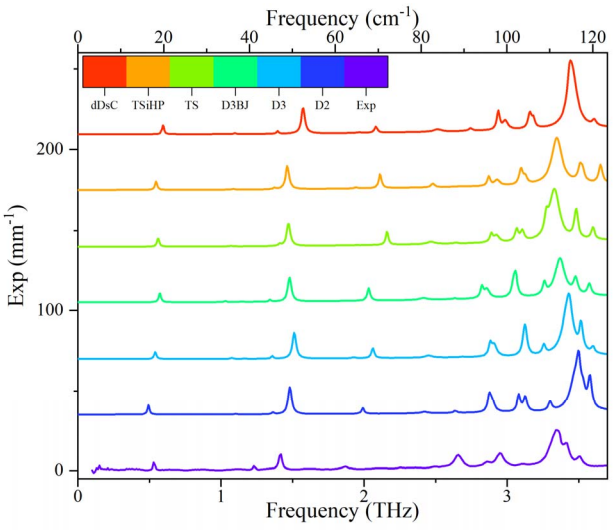

Figure 1. The experimental THz spectra of a 10% aLM sample (mixed with PTFE) recorded at 4 K and the calculated THz spectra with each dispersion correction.

As DFT often underestimates the inter-molecular forces that hold many molecular crystals together, an empirical dispersion correction is often introduced to improve the correlation between calculation and experiment. Here we investigate how the inclusion of these dispersion corrections into the calculation influences the calculated geometry of the system and the resultant THz spectra of α-lactose monohydrate (aLM). A number of dispersion corrections [9–14] have been included in this study.

This study focuses on understanding the THz spectra of powdered crystalline aLM. This material is an ideal THz spectral standard [15] particularly owing to an incredibly low

Name	a / Å	Δa %	b / Å	Δb %	c / Å	Δc %	β / °	Δβ %
Exp	4.783	-	21.54	-	7.760	-	105.91	-
D2	4.643	-2.92	21.42	-0.563	7.713	-0.164	105.74	-0.163
D3	4.691	-1.97	21.57	+0.124	7.742	-0.466	105.42	-0.466
D3BJ	4.705	-1.65	21.58	+0.188	7.740	-0.284	105.61	-0.284
dDsC	4.593	-4.03	21.33	-0.094	7.680	-0.319	105.57	-0.319
TS	4.661	-2.64	21.58	+0.209	7.737	-0.701	105.17	-0.701
TSiHP	4.670	-2.43	21.64	+0.501	7.748	-0.590	105.29	-0.590

Table 1 – The unit cell parameters for each structure and the percentage change from the experimental structure. Only the β angle is shown as the other angles were not varied owing to symmetry constraints.

978-1-7281-6620-9/20/$31.00 ©2020 IEEE

	Correlation	Frequency Scaling	Frequency Shift	Average RMS Displacement	Minimum Mapping	Fingerprint Comparison
D2	0.8213	0.96	0	0.0741	0.14394	0.4029
D3	0.8787	1.01	-3.4	0.05964*	0.12132*	0.3904*
D3BJ	0.8475	1.02	-3	0.08199	0.16924	0.401
dDsC	0.8782	0.99	-1.8	0.07811	0.14932	0.4302
TS	0.9021*	1.04	-3.6	0.09177	0.18429	0.4039
TSiHP	0.8899	1.03	-3.2	0.08826	0.18373	0.4002

Table 2 – The different numerical values used to compare the produced structure and spectra. The * indicates the best correction in each column. Frequency Scaling and Frequency Shift are for information purposes and are not graded.

frequency mode (17 cm^{-1}/0.52 GHz) that is very sharp, even at room temperature, but also has a complex spectra which has been heavily studied previously [16].

II. Results

In order to understand the influence of the dispersion corrections on the structure of aLM Table 1 shows the experimental and optimised unit cell parameters for calculations using the various dispersion corrections. While all calculations seem to represent the overall structure well, generally a compression along the *a* and *c*-axis along with a decrease in the β-angle for all calculations is observed. However, the changes in all calculations are relatively small and a closer look shows that most molecular changes are owing to shifts in hydrogen positions, which is unsurprising given that these are difficult to locate experimentally in the X-ray measured structure. Determining what the 'best structure' is is difficult from this comparison so we also used three numerical methods to compare the optimised structures to the experimental structure and these are summarised in Table 2. Pymatgen [17] was used to calculate the average root mean squared (RMS) displacement along with a minimum mapping value which corresponds to the minimal species mapping required to align the two structures. Having a value as close to 0 as possible represents the smallest shift.

Finally, a python package called Matminer [18] was used to create a 'fingerprint' of each structure which were compared. With this parameter a value of 0 means the structures are identical but any number < 0.9 can be considered similar. Using these three numerical comparisons the D3 correction clearly shows the closest structure to experiment.

Fig. 1 shows the THz spectra of aLM measured at 4 K along with the calculated spectra for each calculation that incorporated a different dispersion correction. What is clear is that all calculations show a reasonably similar spectral shape. A closer look at the calculated modes using each method show that the motion involved for each mode, in each calculation, is the same and the frequency order of the modes is the same for all calculations, although the exact position and relative intensity of each mode can change substantially between calculations. As an example, the experimental mode at 0.52 THz varies between 0.492 and 0.594 THz across the various calculations. Spectra were compared using a cross-correlation coefficient (a value of 1 being identical) while the frequency scaling and frequency shift required to obtain the best cross-correlation between experiment and calculation was also determined. These comparison metrics have been discussed in detail elsewhere [8]. Using these methods, the TS dispersion correction seems to provide the best overall correlation between experiment and calculation, although all methods generally overestimate the frequency of the modes.

III. Summary

Here we have shown that solid-state DFT calculations can be used to interpret the THz spectra of aLM. A dispersion correction is needed to better model inter-molecular forces. This work has shown that the choice of dispersion correction is still however, unclear and needs to be chosen carefully. For aLM, D3 provides a good choice for structural optimisation while the TS correction provides the best correlation with experimental THz spectra. It is now important to extend this study to incorporate temperature effects, anhamonicity and phonon lifetimes into the DFT calculations in order to understand the origins of the discrepancies between experiment and calculation still further.

References

1 Burnett, A. D. *et al*. *Materials Today*. 2008, **11**(3), pp. 18–26.
2 Groom, C. R *et al. Acta Crystallogr B* 2016, **72**(2), pp. 171–179.
3 Kresse, G. *et al. Phys. Rev. B*. 1993, **47**(1), pp. 558–561.
4 Togo, A. *et al.* Science. *Scr. Mater.* 2015, **108** pp. 1–5.
5 Csonka, G. I *et al*. *Phys. Rev. B*. 2009, **79**(15), 155107.
6 Kresse, G *et al*. *Phys. Rev. B*. 1999, **59**(3), pp. 1758-1775.
7 Monkhorst, H. J. *et al. Physical Review B*. 1976, **13**(12), pp. 5188–5192.
8 Burnett, A. D. *et al. J. Infrared. Millim. THz*. 2019, pp.1–32.
9 Grimme, S. *J. Comput. Chem*. 2006, **27**(15), pp. 1787–1799.
10 Grimme, S. *et al*. *J. Chem. Phys*. 2010, **132**(15), 154104.
11 Becke, A. D. *et al*. *J. Chem. Phys*. 2005, **123**(15), 154101.
12 Tkatchenko, A. *et al. Phys. Rev. Lett.* 2009, **102**(7), 073005.
13 Bučko, T *et al. J. Chem. 2013, 9(10), pp. 4293-4299.*
14 Steinmann, S. N *et al. J. Chem. Phys*. 2011, **134**(4), 044117.
15 Withayachumnankul, W *et al*. *Opt. Express*. 2008, **16**(10), pp. 7382-7396.
16 Allis, D. G *et al*. *Chem. Phys. Lett*. 2007, **440**(4–6), pp. 203–209.
17 S. Ong *et al. Comp. Mats. Sci.* 2013, **68**, pp. 314-319
18 L. Ward *et al. Comp. Mats. Sci.* 2018, **152**, pp. 60-69