Identifying Raman modes of Sb$_2$Se$_3$ and their symmetries using angle-resolved polarised Raman spectra†

Nicole Fleck,†a Theodore D. C. Hobson,†a Christopher N. Savory,@b John Buckeridge,a Tim D. Veal,b Maria R. Correia,d David O. Scanlon,b Ken Durosea and Frank Jäckela

The physical properties of antimony selenide (Sb$_2$Se$_3$) are highly anisotropic. Angle-resolved polarised Raman spectroscopy was employed to characterise oriented crystals and used in conjunction with group theory structural analysis to assign vibrational symmetries to the peaks observed in the Raman spectra. The phonon energies were corroborated via density functional theory (DFT) calculations. Furthermore, a straightforward method is proposed to verify the desirable (001) plane orientation of film growth via minimisation of the 155 cm$^{-1}$ peak in the Raman spectrum.

Introduction

Antimony selenide (Sb$_2$Se$_3$) is an anisotropic semiconductor with a room temperature band gap of 1.18 eV.¹ It is currently a highly attractive candidate for low-cost, large-volume photovoltaic electricity generation.² Indeed, single junction power conversion efficiencies have reached 9.2%³ from 2.1%⁴ in only 6 years. Pushing the solar cell efficiency to over 10% would make Sb$_2$Se$_3$ viable for industrial scale up. It is also of interest for water splitting, as a superconductor and as a thermoelectric material.⁵–⁷ Nevertheless the state of knowledge of its fundamental properties has been slow to develop.

Sb$_2$Se$_3$ has an orthorhombic structure under atmospheric pressure and belongs to the Pbnm space group and $D_{2h}$ point group (Schoenflies notation) with lattice parameters $a = d_{100} = 11.6311$ Å, $b = d_{010} = 11.7808$ Å, $c = d_{001} = 3.9767$ Å.⁸ The structure is comprised of 1D nanoribbons coordinated via van der Waals interactions.⁹ The three main crystal planes are shown in Fig. 1. The Pbnm space group nomenclature is used in this paper, meaning the [001] direction is parallel to the ribbon growth direction. Both Pbnm and Pnma space groups can be found in the discussion of Sb$_2$Se$_3$ which may lead to confusion ($Pnma$ convention aligns the [010] direction with the ribbon axis).

Raman spectroscopy is a fast and usually non-destructive technique which can be used to characterise Sb$_2$Se$_3$ and its impurity phases.¹¹ While gaps in the literature remain, some recent analysis of Sb$_2$Se$_3$ via Raman spectroscopy has furthered the current understanding of this material. In 2018, Shongalova et al. resolved an issue of misinterpretation of one of the peaks in the analysis of previous Raman spectra.¹¹ While much of the
literature indexed the peak at 250 cm$^{-1}$ as characteristic of Sb$_2$Se$_3$, it was found to be an antimony oxide ($\gamma$-Sb$_2$O$_3$) peak. This occurred mainly due to sample degradation under high laser power during measurement but these peaks are indistinguishable from Sb$_2$O$_3$ or elemental selenium impurities already present in the sample from the growth/synthesis. The following year, the process of comprehensively indexing the Raman active vibrational modes of Sb$_2$Se$_3$ was begun by Vidal-Fuentes et al. including a low temperature study. However, due to the small size of the crystallite analysed, its orientation could not be determined and the symmetries of the vibrations could not be assigned fully.

Raman mode symmetries describe the nature of the vibrations in a material. Sb$_2$Se$_3$ is known to be isostructural to Sb$_2$S$_3$ which has 60 zone centre phonons of symmetry $\Gamma = 10A_g + 5B_{1g} + 10B_{2g} + 5B_{3g} + 5A_u + 10B_{1u} + 5B_{2u} + 10B_{3u}$. Both A and B modes are non-degenerate with A modes showing symmetric behaviour and B modes being anti-symmetric. The subscript g indicates the mode is symmetric under inversion while u indicates it is anti-symmetric. In Sb$_2$Se$_3$, the $A_m$ modes are silent and the $B_m$ modes are IR-active. The set of Raman active modes have the following symmetries predicted by group theory: $\Gamma = 10A_g + 5B_{1g} + 10B_{2g} + 5B_{3g}$.

Angle-resolved polarised Raman spectroscopy has been used in the literature to assign vibrational symmetries to phonon modes, for example in ZnO, NbSe$_3$, MoTe$_2$ and BaTiO$_3$ crystals. For anisotropic crystals, birefringence has an impact on the scattering of polarised light described by Kranert et al. However, this effect is negligible for highly absorbing materials such as Sb$_2$Se$_3$, hence the more straightforward approach described in the results continues to hold.

Optimum Sb$_2$Se$_3$ solar cell device performance is currently achieved by growing the ribbons perpendicularly to the contacting layers i.e. preferential growth in the [001] direction. Angle-resolved polarised Raman spectroscopy has been used to determine the orientation and variation thereof across polycrystalline, thin film absorber, CuInSe$_2$. The same approach could be used to identify the orientation of Sb$_2$Se$_3$ thin films and even identify the variation from grain to grain which may significantly affect device performance. Vidal-Fuentes et al. also proposed this, however it has not been realised to date. In this way rapid determination of the dominant orientation of films of Sb$_2$Se$_3$ would streamline the fabrication and characterisation process.

In this study, a combined experimental, group theory and density functional theory (DFT) approach is presented to comprehensively measure and identify the Raman bands in Sb$_2$Se$_3$. This study differs from earlier work in that the experiments are for a set of oriented (100), (010) and (001) crystallographic surfaces which has enabled the symmetry of the vibrational modes to be assigned reliably for the first time. Furthermore, DFT calculations were used to determine the Raman active phonon mode energies of Sb$_2$Se$_3$ and these corroborated the experimental results as well as suggesting additional vibrational modes. The results are expected to have relevance to both the basic science and technological communities.

Experimental and computational details

Sb$_2$Se$_3$ crystals were prepared via the Bridgman technique using a single-zone vertical furnace. The details can be found in the ESI (ESI Section I†). Crystals were grown up to 4 mm in diameter and 1 cm in length. The crystals demonstrated bright cleavage planes extending over their full width and displayed narrow X-ray diffraction (XRD) peaks (see Table S1†). The crystals could be easily manually cleaved to expose the (010) plane and reveal reflective facets. This is consistent with the expectation that the weakest van der Waals interactions lie along the [010] axis. Straight, parallel lines observed using optical microscopy on the surface of the cleaved planes (see Fig. S1†) result from cleavage steps and are expected to run parallel to the 1D nanoribbons ([001] axis), as this is the axis with the most strongly bound layers. These were used as guides to orient the crystal optically on a goniometer and cut to reveal the (100) and (001) planes without the need for Laue equipment.

The cleaved (010) plane and cut (001) plane can be seen in Fig. 2a. Cutting the (001) plane was found to be most challenging as it was mechanically less stable than the other orientations. Crystals of this orientation were therefore mounted in epoxy to prevent unwanted cleaving during polish/storage. The surfaces of the cut (100) and (001) plane crystals were polished with 0.3 μm alumina particles in de-ionised water. The cleaved (010) faces did not require polishing. To verify the surface orientations of the crystals, XRD ($\theta$–$\theta$ scan) was performed using a Rigaku Smartlab diffractometer in parallel beam geometry using a monochromated Cu-Kα$_1$ X-ray source.

Angle-resolved polarised Raman spectroscopy (backscattering geometry) was used to assign the vibrational symmetries to the Raman modes. The setup is shown in Fig. 2b. Linearly polarised light impinges on the sample. The scattered light is...
passed through a polariser oriented either parallel (solid arrow) or perpendicular (dashed arrow) to the polarisation of the incident beam. The geometry can be described in Porto notation where Z(XX)Z implies the photons are incident in the Z-direction and directly backscattered light (Z) is detected. The labels in parenthesis indicate the incident beam polarisation is in the X-direction and X-polarised light is detected (parallel polariser geometry). The crossed polariser geometry, Z(ZY)Z was used also.

The crystals were rotated in the polarisation plane being studied while the Raman peak intensities were monitored. The sample angle (θ) was varied between 0° and 180° in 10° increments and a spectrum was acquired at each yielding 19 spectra in each dataset. A dataset was taken for each of the three crystal planes in both parallel and perpendicular polariser configurations. A total of 114 spectra were recorded.

Room temperature Raman spectroscopy was carried out using a Renishaw in-Via microscope with a 532 nm laser. The maximum laser power (at the sample) was limited to 0.1 mW with maximum power density of 3.2 kW cm⁻² to prevent sample damage. No degradation effects (peak at 250 cm⁻¹ or higher) were observed in any of the spectra acquired. The power density at which oxidation occurs varies across studies and was observed to vary even for different crystal planes. Analysis of this behaviour requires further work and is beyond the scope of this study but may be dependent on the wavelength of the illuminating light.

The peak position values were not constrained to a given range during the fitting. Constraining the parameters would force consistency across the datasets. The positions were kept as shared parameters within a dataset but free to vary between datasets. The resulting values were compared across datasets to determine the accuracy and reliability of the fitting procedure.

Group theory calculations were performed to evaluate the sample angle dependence of the intensities of the Raman active modes. These dependences were calculated using a matrix method for both the parallel and perpendicular polariser geometry. The full description can be found in ESI Section II and the outcomes are described in the results.

Periodic DFT calculations provided theoretical prediction of the Raman vibrational energies using scalar relativistic potentials within the Vienna Ab initio Simulation Package (VASP). The Perdew–Burke–Ernzerhof generalised gradient approximation (GGA) exchange-correlation functional revised for solids (PBEsol) was used for both structural relaxation and phonon supercell calculations. Interactions between the valence and core electrons were described via the projector augmented wave method. The Phonopy package was used to generate twenty 2 × 4 × 2 (320 atom) finite displacement supercells from the PBEsol relaxed structure to obtain the phonon frequencies of Sb₂Se₃. In all calculations, the total energy was converged to within 1 × 10⁻⁸ eV and the forces per atom were converged to within 1 × 10⁻⁵ eV Å⁻¹. A plane wave cut off of 500 eV was used throughout and a Γ-centred k-point mesh of 2 × 2 × 2 was used for the supercell calculations.

Previous computational studies on Sb₂Se₃ have either used the PBE or HSE06 functionals (with or without dispersion corrections) to describe its structural properties. PBEsol optimises the structure to within 0.5% of experimental values for the a and c lattice parameters while underestimating the b parameter by just over 2%. While HSE06 gives an optimised geometry to within 2% of experiment, PBEsol is much more cost effective for the large supercells required to converge the phonon frequencies. PBEsol has also shown improved performance over PBE in the calculation of vibrational properties in other chalcogenide materials. While electronic structure calculations of Sb₂Se₃ have been conducted by numerous groups and calculation of the phonon dispersion has also been reported, tabulation of predicted phonon mode energies via DFT calculations and comparison to experimental data has not been published to date.

Results

Fig. 3 shows XRD patterns of the cut (100) and (001) and cleaved (010) crystal planes. In all cases, the first structure-factor-allowed reflection is present followed by the higher order peaks. This confirms that the use of the cleaving method to identify the (010) plane and cutting relative to the [001] axis is a reliable means of orienting Sb₂Se₃ on its orthorhombic faces. This method may also be successful for other van der Waals crystals.

A summary of the results of group theory analysis is presented in Table 1. The constants A–F represent Raman tensor elements. The Raman peak intensity variation, with respect to sample rotation (θ), of each vibrational mode (Aₐ, B₁g, B₂g and B₃g) was determined. This peak intensity variation depends on the crystal plane being analysed (θ or θ₁) and whether the polarisation was parallel or perpendicular to the incident laser polarisation (l or l₁). In the parallel polariser geometry, all Aₐ mode intensities in each of the three crystal planes have a periodicity of 180°. The B₃g mode intensities will repeat after
due to their $2\theta$ dependence. In perpendicular polariser geometry, both $A_g$ and $B_g$ modes have $90^\circ$ periodicity, however, $A_g$ modes show sine-like behaviour and $B_g$ modes show cosine-like behaviour. $A_g$ and $B_g$ modes can therefore easily be distinguished in both polariser orientations. In this way, knowing the crystal plane as well as the expected peak intensity variation of the individual Raman peaks allows their symmetries to be identified.

Fig. 4 shows an example (the (100) plane, parallel polariser) of the angle-resolved datasets collected and the fitting procedure used for analysis. The black line indicates the raw data while the coloured peaks show the fitted Lorentzian lineshapes. Similar plots for the (010) and (001) planes are shown in Fig. S2† and a more detailed example fit of a single Raman spectrum can be found in Fig. S3.† For each peak in the figure, the peak intensity ($I$) vs. sample angle ($\theta$) was extracted and plotted. This data was compared and fitted to the expected $I-\theta$ behaviour from Table 1. Fig. 5 shows the data points for the 185 cm$^{-1}$ peak $I-\theta$ plot. The solid lines are fits to the data points of the relevant functions from Table 1. For example, the 185 cm$^{-1}$ peak has maximum intensity at 0° rotation, minimum at 90° and returns to its maximum at 180° rotation of the crystal. The 180° period is characteristic of an $A_g$ symmetry vibration. Therefore the function fitted to the data was for an $A_g$ mode in the (100) plane with parallel polariser geometry. The tensor elements were used as fitting parameters. Due to the reliable and consistent fit of the expected $A_g$ function for all crystal orientations and both polariser geometries the symmetry can be confirmed as $A_g$. These plots were prepared for each peak in each of the six datasets in order to identify the mode symmetries. The remaining plots can be found in the Fig. S5–S7.† These plots show excellent consistency in peak position with only small variations (within the expected error of 1 cm$^{-1}$) in values.
obtained from the different planes and polariser configurations. This further validates the rigorous analysis procedure.

Fig. 6 shows the DFT-calculated dispersion of the phonon bands along the high symmetry path within the Brillouin zone. Due to the fact that Raman spectroscopy probes only zone-centre phonons the corresponding theoretical values can be noted from the intersection of the phonon bands with the $\Gamma$ point. 30 phonon energies were predicted in this way which is consistent with the expected $10A_g + 5B_{1g} + 10B_{2g} + 5B_{3g}$ modes of Sb$_2$Se$_3$. The peaks determined using DFT were found to be shifted from the experimental positions in a systematic manner ($-11$ cm$^{-1}$) which has been discussed previously for phonon energies calculated using the PBEsol functional with a similar shift found ($-10$ cm$^{-1}$). Such disagreement can result from a combination of differences in computational lattice parameters versus experiment. This in turn may be a consequence of the athermal limit and inaccuracies in the DFT approach in modelling the interatomic forces precisely. For a material such as Sb$_2$Se$_3$, where strong and weak interactions are present in different directions, modelling the interatomic forces is challenging. Nevertheless a consistent linear shift of the DFT computed peaks upwards by $11$ cm$^{-1}$ resulted in excellent agreement with experimental results as seen in Table 2.

A complete description of the determined vibrational modes of Sb$_2$Se$_3$ is presented in Table 2. The experimental peaks, their symmetries, theoretically predicted modes and previously reported peak positions$^{35}$ are compared and show very good agreement. Seven $A_g$ modes and seven $B_g$ modes were assigned experimentally. Several assignments deviate from those seen by Vidal-Fuentes and are further discussed below. Vidal-Fuentes et al. could not resolve the peaks found at 212, 187 and $156$ cm$^{-1}$ with a randomly oriented crystal. For the same reason, the $B_g$ symmetries could previously not be assigned. All other peaks seen by Vidal-Fuentes were corroborated by the experimental peak positions determined here. In addition, the partial symmetry assignments Vidal-Fuentes report are consistent with the symmetries determined in this study.

While Vidal-Fuentes report a single peak at $213$ cm$^{-1}$, the experimental results here from the (001) plane indicated two closely spaced peaks (213 and $212$ cm$^{-1}$). The $B_{1g}$ modes show large relative intensity in the perpendicular polariser geometry for this plane (see Fig. S7†). Due to intensity differences, the $213$ cm$^{-1}$ $A_g$ mode cannot be resolved in the perpendicular polariser case and the $212$ cm$^{-1}$ $B_{1g}$ mode is of too low intensity to be seen in the parallel polariser geometry. The existence of two peaks in this region is also consistent with the closely spaced modes predicted via DFT. Data from the (010) plane suggests two $A_g$ peaks in this region, however this second peak of $A_g$ symmetry was not observed in any other plane. It has been included for completeness but could not be confidently labelled (Fig. S6†).

For the (001) plane, the global fit for the $A_g$ mode around $185$ cm$^{-1}$ was shifted to $187$ cm$^{-1}$ for both the parallel and perpendicular polariser case (see Fig. S7†) over numerous repetitions of the experiment. This is a strong indication of the existence of another mode and is again corroborated by the theoretical calculations. This also holds true for the mode at $155$ cm$^{-1}$. The peak at $155$ cm$^{-1}$ is clearly visible in the (010) plane spectra (see Fig. S6†) and a $156$ cm$^{-1}$ feature is visible in the (100) plane (Fig. S5†). The latter could be an artefact from the collection cone angle of the objective lens which does not uniquely collect light scattered normal to the sample surface. However the consistency of the results with theory and literature suggest this is not the case.

Once a shift of $+11$ cm$^{-1}$ was applied, the DFT calculations predicted all of the experimentally observed modes to a high degree of accuracy. Some modes which were not observed experimentally were also predicted. In contrast to the calculations, no modes around $140$ cm$^{-1}$ were seen in the Raman spectra which is consistent with the experimental data from Vidal-Fuentes et al. This discrepancy may be due to the low intensity of these vibrational modes. The remaining predicted modes show very good agreement with those experimentally observed.
The 110–140 cm$^{-1}$ region was difficult to analyse in the (100) and (001) planes as there are many overlapping modes. The broad nature and low intensity of the features in this region prohibit accurate global fits to be calculated. Individual spectra in each crystal orientation show a clear peak at 117 cm$^{-1}$ (see ESI Fig. S2–S4†). This mode was assigned as Ag as only these modes are expected to be visible in spectra for all three crystal orientations. However, this assignment should be considered tentative on the present evidence, even though it is consistent with both Vidal-Fuentes and DFT predictions. Nevertheless, the additional peaks predicted by DFT in this region are expected to be present although they cannot be resolved in the data measured at room temperature.

The peak at 100 cm$^{-1}$ lies at the edge of the Raman spectrometer filter cut off and thus the position may be artificially high. This mode showed strong Ag vibrational symmetry in all parallel polariser configurations and was therefore included. The cut off using the perpendicular polariser case was slightly higher so that the mode at 100 cm$^{-1}$ was no longer visible and could not be analysed. The filter cut-off generally limited the analysis of the vibrations at wavenumbers below 100 cm$^{-1}$.

### Table 2
Experimental Raman peak positions and symmetries of Sb$_2$Se$_3$ and comprehensive comparison to DFT calculations and literature (Lit.) at room temperature (RT) and 10 K. All values in the table have units of cm$^{-1}$. DFT predicted peaks shifted upwards by 11 cm$^{-1}$.

<table>
<thead>
<tr>
<th>Exp.</th>
<th>Symmetry</th>
<th>FWHM</th>
<th>DFT</th>
<th>Lit. (RT)</th>
<th>Symmetry</th>
<th>FWHM</th>
<th>Lit. (10 K)</th>
</tr>
</thead>
<tbody>
<tr>
<td>213</td>
<td>A$_g^1$</td>
<td>6.4</td>
<td>213</td>
<td>212.8</td>
<td>A$_g$</td>
<td>6.0</td>
<td>215.6</td>
</tr>
<tr>
<td>212</td>
<td>B$_{2g}^1$</td>
<td>5.3</td>
<td>211</td>
<td>207.8</td>
<td>B$_{2g}$</td>
<td>7.3</td>
<td>209.0/203.7</td>
</tr>
<tr>
<td>205</td>
<td>B$_{2g}^2$</td>
<td>11.5</td>
<td>205</td>
<td>191.5</td>
<td>A$_g$</td>
<td>10.6</td>
<td>195.8</td>
</tr>
<tr>
<td>192</td>
<td>A$_g^3$</td>
<td>11.8</td>
<td>187</td>
<td>115.5</td>
<td>A$_g$</td>
<td>11.5</td>
<td>189.1</td>
</tr>
<tr>
<td>187</td>
<td>A$_g^4$</td>
<td>10.6</td>
<td>185</td>
<td>184.7</td>
<td>A$_g$</td>
<td>11.5</td>
<td>189.1</td>
</tr>
<tr>
<td>185</td>
<td>A$_g^5$</td>
<td>15.1</td>
<td>156</td>
<td>111.4</td>
<td>A$_g$</td>
<td>8.7</td>
<td>119.0/114.7</td>
</tr>
<tr>
<td>156</td>
<td>B$_{2g}^1$</td>
<td>13.9</td>
<td>155</td>
<td>154.4</td>
<td>B$_{2g}$</td>
<td>13.2</td>
<td>157.4</td>
</tr>
<tr>
<td>155</td>
<td>B$_{2g}^2$</td>
<td>143/142/132</td>
<td>131</td>
<td>132.2</td>
<td>B$_{2g}$</td>
<td>14.0</td>
<td>138.7</td>
</tr>
<tr>
<td>131</td>
<td>B$_{2g}^3$</td>
<td>9.9</td>
<td>124</td>
<td>125.7</td>
<td>B$_{2g}$</td>
<td>12.3</td>
<td>130.0</td>
</tr>
<tr>
<td>124</td>
<td>B$_{2g}^4$</td>
<td>111/107</td>
<td>117</td>
<td>118.7</td>
<td>A$_g$</td>
<td>8.7</td>
<td>119.0/114.7</td>
</tr>
<tr>
<td>117</td>
<td>A$_g^5$</td>
<td>7.5</td>
<td>111</td>
<td>109.3</td>
<td>B$_{2g}$</td>
<td>13.0</td>
<td>111.4</td>
</tr>
<tr>
<td>111</td>
<td>B$_{2g}^6$</td>
<td>8.3</td>
<td>101</td>
<td>100.1</td>
<td>A$_g$</td>
<td>3.8</td>
<td>103.3</td>
</tr>
<tr>
<td>101</td>
<td>A$_g^7$</td>
<td>80.1</td>
<td>82</td>
<td>83</td>
<td>A$_g$</td>
<td>2.9</td>
<td>83.4</td>
</tr>
<tr>
<td>82</td>
<td>A$_g^8$</td>
<td>72/69</td>
<td>68</td>
<td>61.6</td>
<td>A$_g$</td>
<td>2.4</td>
<td></td>
</tr>
<tr>
<td>58/56</td>
<td>54.4</td>
<td>55/49</td>
<td>55/49</td>
<td>54.4</td>
<td>B$_{2g}$</td>
<td>2.0</td>
<td></td>
</tr>
</tbody>
</table>

$^a$ Tentative symmetry assignment.
Further work could be done with a Raman spectrometer configured to measure below 100 cm\(^{-1}\) using the same analysis technique to identify and characterise the low wavenumber vibrations.

While the orientation of the ribbons is difficult to unequivocally deduce from the Raman spectra, it is clear that the device relevant (001) plane shows little to no Raman scattering intensity around 155 cm\(^{-1}\). Minimisation of this peak may provide a fast approach to testing the alignment of the ribbons in thin films and crystals. In the crossed polariser geometry, altering growth parameters to maximise the 205 cm\(^{-1}\) peak intensity (see Fig. S7†) can also ensure growth in the [001] direction. Raman spectroscopy already provides a fast, non-destructive approach to material characterisation and this could be further exploited.

**Conclusions**

van der Waals crystals such as Sb\(_2\)Se\(_3\) may easily cleave, revealing cleavage steps. These may be used in optical orientation of crystals for cutting without the need for X-ray methods. Angle-resolved polarised Raman spectroscopy of individual crystal planes was used to identify the Raman modes of Sb\(_2\)Se\(_3\) and assign vibrational symmetries to each peak in the spectrum. A total of ten A\(_g\), five B\(_{1g}\), ten B\(_{2g}\) and five B\(_{3g}\) modes are expected from group theory and DFT calculations. Experimentally, seven A\(_g\), two B\(_{1g}\), three B\(_{2g}\) and two B\(_{3g}\) modes were identified. Results are consistent with those of Vidal-Fuentes,\(^{15}\) but additional peaks were identified and concretesymmetry assignments made through the use of oriented crystals in this work. DFT calculations supported and verified the phonon mode energies as well as predicting some low intensity and lower wavenumber modes. Some low energy modes were not observable experimentally, but are expected to be measured in further studies. A methodology was reported to quickly identify the technologically important (001) orientation using the 155 cm\(^{-1}\) Raman peak. This study both strengthens the fundamental understanding of Sb\(_2\)Se\(_3\) as a material and paves the way for rapid determination of film orientation and grain to grain variation of growth.

**Conflicts of interest**

There are no conflicts to declare.

**Acknowledgements**

Many thanks to Laurence Hardwick for use of the Raman spectrometer. This work was supported by the Engineering and Physical Sciences Research Council (EPSRC) grant number EP/N000020/1 (N. Fleck and T. D. C. Hobson) and funding from the University of Liverpool (N. Fleck). The authors acknowledge the use of the UCL Legion, Grace and Myriad High Performance Computing Facilities (Legion@UCL, Grace@UCL and Myriad@UCL) and associated support services in the completion of this work. We are also grateful to the ARCHER UK National Supercomputing Service, provided via our membership of the UK’s HEC Materials Chemistry Consortium, funded by EPSRC (EP/L000202, EP/R029431), and the UK Materials and Molecular Modelling Hub for computational resources, MMM Hub, which is partially funded by EPSRC (EP/P020194/1). The XRD facility was used supported by the EPSRC (EP/P001513/1). M. R. Correa thanks the funding provided from the i3N, UIDB/50025/2020 and UIDP/50025/2020 projects, financed by national funds through the FCT/MEC. The underlying data in this paper is available from https://doi.org/10.17638/datacat.liverpool.ac.uk/1062.

**Notes and references**
