Assessing the Performance of Density Functional Theory Methods on the Prediction of Low-Frequency Vibrational Spectra

Peter A. Banks[1] · **Zihui Song**[1] · **Michael T. Ruggiero**[1]

Abstract

The low-frequency (terahertz) dynamics of condensed phase materials provide valuable insight into numerous bulk phenomena. However, the assignment and inter-pretation of experimental results require computational methods due to the complex mode types that depend on weak intermolecular forces. Solid-state density functional theory has been used in this regard with great success, yet the selection of specific computational parameters, namely the chosen basis set and density functional, has a profound influence on the accuracy of predicted spectra. In this work, the role of these two parameters is investigated in a series of organic molecular crystals, in order to assess the ability of various methods to reproduce intermolecular forces, and sub-sequently experimental terahertz spectra. Specifically, naphthalene, oxalic acid, and thymine were chosen based on the varied intermolecular interactions present in each material. The results highlight that unconstrained geometry optimizations can be used as an initial proxy for the accuracy of interatomic forces, with errors in the calculated geometries indicative of subsequent errors in the calculated low-frequency vibra-tional spectra, providing a powerful metric for the validation of theoretical results. Finally, the origins of the observed shortcomings are analyzed, providing a basic framework for further studies on related materials.

Keywords Density functional theory · Molecular dynamics · Vibrational spectroscopy · Low frequency · Crystalline disorder

---

Peter A. Banks and Zihui Song contributed equally to this work.

✉ Michael T. Ruggiero
Michael.Ruggiero@uvm.edu

1   Department of Chemistry, University of Vermont, 82 University Place, Burlington, VT 05405, USA

1 Introduction

Terahertz time-domain spectroscopy is a powerful method for the evaluation of crystalline phonons, which are increasingly being shown to be related to the bulk properties of the solids [1–4]. Low-frequency vibrations in molecular crystals often involve motions of entire molecules, for example hindered translations, making the vibrational dynamics dependent upon inherently weak long-range forces [1, 3, 5–9]. This ultimately results in each molecular crystal, including crystals consisting of the same molecule but with different bulk packing (e.g., polymorphism or disorder), exhibiting a unique terahertz spectrum—both in mode type and frequency [3, 10–13]. But while terahertz time-domain spectroscopy is increasingly becoming more widely utilized, the assignment of the low-frequency vibrational spectra is not straightforward, and remains a challenge for the interpretation of experimental results. This is because unlike the well-defined functional group-specific transitions probed by mid-infrared spectroscopy (e.g., FTIR and Raman), terahertz vibrations are much more complex and depend on many factors, making assignment difficult [1, 2, 4, 14–18].

Computational methods have been successfully used to assign the predicted vibrational spectra with success. However, the simulation of low-frequency vibrations presents a greater challenge compared with mid-infrared motions, due to the incredibly weak forces that dictate terahertz motions. Thus, the employed theoretical model must be capable of reproducing the weakest forces present within a material, for example van der Waals and related intermolecular forces. A number of methods have been used to predict terahertz spectra, but solid-state density functional theory (ss-DFT) has proven to be an excellent and robust method for predicting low-frequency vibrational modes with an accuracy orders of magnitude better than "gold standard" of chemical accuracy [19].

However, while ss-DFT is an appropriate method for describing terahertz dynamics, there still are a number of parameters—such as the employed basis set, density functional, and dispersion correction—that ultimately influences the calculation of crystalline forces, which in turn drive vibrational dynamics. For example, the range-corrected ωB97-x functional accurately predicts unit cell parameters and vibrational transition frequencies for π-stacked ionic pyridinium molecular crystals, due to its ability to more accurately capture charge-separated forces, while other functionals fail to reproduce the same forces [20]. Thus, to explore the dependence of the calculated forces in molecular crystals on ss-DFT parameters, three molecular crystalline systems—naphthalene, anhydrous oxalic acid, and thymine—were selected based on the varied and distinct intermolecular forces present in each solid. The ability of a wide range of basis sets and density functionals to reproduce specific forces is assessed, both with respect to the experimental low-frequency vibrational spectra, and with the reproduction of the experimentally observed bulk geometry, which we show can act as a measure of the accuracy of the force calculation. We also perform cross-correlation analyses to showcase the similarities and differences between the various methods, and highlight the successes—and failures—of particular methods with respect to the handling of specific force types. Overall, this work provides

a measure of validating theoretical results, yielding a framework for further studies utilizing computational methods for the assignment of low-frequency vibrational spectra.

2 Methods

2.1 Experimental

Thymine (99%, Sigma-Aldrich) and naphthalene (99%, Acros Organics) are available commercially, and were used as received without any further purification. Anhydrous oxalic acid (98%, Acros Organics) was pre-mixed with polytetrafluoroethylene (PTFE) to an approximate 3% w/w concentration, and placed in a dessicator overnight prior to sample preparation to minimize hydration. Pellets for THz-TDS measurements were prepared by grinding with PTFE to an approximate 3% w/w concentration. The mixture was then pressed under about 225.32 MPa (32163.74 Psi) of pressure in a 13-mm-diameter die (Specac), furnishing 3-mm-thick samples. All THz-TDS were acquired using a Teraflash spectrometer (Toptica Photonics AG). A fiber-coupled emitter and receiver were used, and two pairs of off-axis parabolic mirrors served to collimate and focus the terahertz radiation, respectively. The entire spectrometer was enclosed and continuously purged with dry nitrogen gas to minimize absorption of atmospheric water. Cryogenic (50 K) spectra were collected using a closed-cycle helium cryostat (Cryocool Industries), with the samples placed under vacuum in a sample chamber containing TPX windows. For each sample, 20,000 time-domain waveforms were collected and averaged, and the resultant waveform underwent a Fourier transform to produce a terahertz power spectrum with a spectral resolution 30 GHz. This process was repeated in quintuplicate for each sample. The reported spectra are divisions of the recorded sample terahertz power spectra by PTFE blank spectra. The experimental peak positions were determined by fitting the spectra with Lorentzian functions.

2.2 Theoretical

Solid-state DFT simulations were performed using the fully periodic CRYSTAL17 software package [21, 22]. All structures underwent geometry optimization, and the calculations were initiated using the experimental atomic positions and lattice vectors retrieved from Cambridge Crystallographic Data Centre (CCDC). Prior to any vibrational analyses, all atoms and lattice vectors were allowed to fully relax with no constraints other than the space group symmetry of the solid. Frequency calculations were executed by using the optimized coordinates and lattice parameters to yield the vibrational modes and IR intensities. Eigenvalues and eigenvectors were calculated numerically through the harmonic approximation [23, 24], which has previously proven effective for vibrational normal modes [1–5, 10, 11, 25–27], and IR intensities were calculated from the dipole moment derivatives ($\frac{\partial \mu}{\partial Q}$) which were determined using the Berry phase method [28, 29]. Reciprocal space sampling was performed using the Monkhorst–Pack scheme, with a k-point mesh in the first Brillouin Zone

(program keyword SHRINK: X X, where X X = 8 8, 6 6, and 8 8 for naphthalene, oxalic acid, and thymine, respectively), corresponding to a commensurate grid of k-points of 170, 64, and 170, once Bravais symmetry was taken into account for these three crystals, respectively. The tolerances for Coulomb and exchange integral cutoffs were set to $\Delta E < 10^{-8}$ hartree (program keyword TOLINTEG: 8 8 8 8 16). The energy convergence criterion for geometric optimizations was set to $\Delta E < 10^{-8}$ hartree (program keyword TOLDEE: 8). The energy convergence criterion for frequency calculations was set to a more stringent value from the geometry optimization, $\Delta E < 10^{-11}$ hartree [30].

In order to evaluate the role of basis set and functional on the calculated vibrational frequency results, a large number of configurations were tested. Both double- and triple-ζ basis sets were evaluated, with varying degrees of additional polarization, including the Pople basis sets 6-31G [31, 32] and 6-311G [33] from no polarization to two polarization functions, i.e., 6-31G(d,p), 6-31G(2d, 2p), 6-311G(d,p), and 6-311G(2d,2p). The original Ahlrich's VTZ and VTZP [34, 35] basis sets were compared against recently developed def2- basis sets [36], including def2-SV and def2-TZV, as well as with additional polarization functions (def2-SVP and def2-TZVP). Finally, the solid-state optimized pob-TZVP [37] basis set was also utilized. The summary of basis sets selected is shown in Table 1.

A wide range of density functionals were also used and evaluated, including the generalized gradient approximation (GGA) and hybrid-GGA functionals, where the latter incorporates a finite percentage of exact Hartree-Fock exchange. For the GGA class, the Perdew-Burke-Ernzerhof (PBE), [38] PBE revised for solids (PBEsol) [39], and the Beck-Lee-Yang-Parr (BLYP) [40] density functional were used, where BLYP is a combination of the Becke B88 and Lee-Yang-Parr (LYP) functional for exchange and correction, respectively. Hybrid functionals were represented by PBE0 [41], PBE0-13 [42], and B3LYP [43] functionals, which are modifications of PBE and BLYP functionals, with 25%, 33.33%, and 20% exact exchange, respectively. London dispersion forces were accounted for using the Grimme DFT-D3 correction, [44] with the Becke-Johnson damping function, and was included for all the functionals tested, using the previously published dispersion coefficients [45, 46].

Table 1   Summary of basis sets used for predicting terahertz spectra

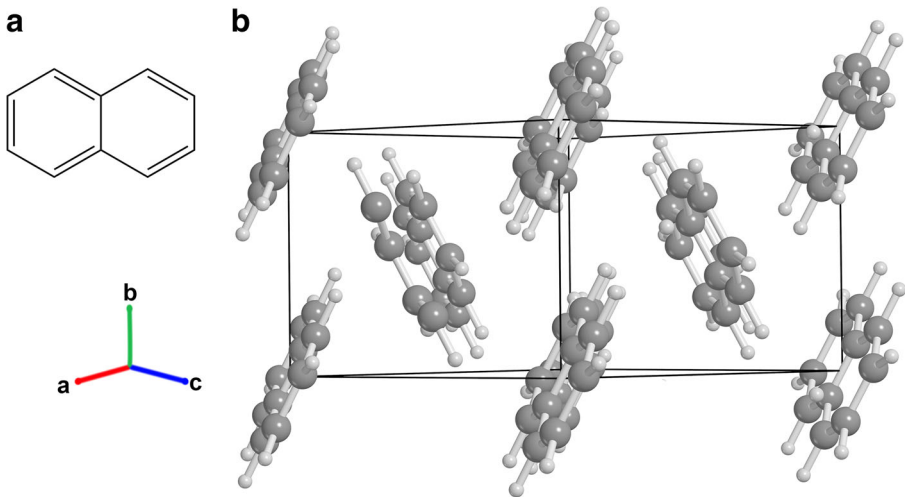|	No polarization	One polarization	Two polarization
Double - ζ (DZ)	6-31G	6-31G (d, p)	6-31G (2d, 2p)
	def2-SV	def2-SVP	—
Triple - ζ (TZ)	6-311G	6-311G (d, p)	6-311G (2d, 2p)
	def2-TZV	def2-TZVP	—
	VTZ	VTZP & pob-TZVP	—

**Fig. 1** **a** Molecular structure and **b** crystalline unit cell of naphthalene

3 Results and Discussion

3.1 Naphthalene

Naphthalene, $C_{10}H_8$, crystallizes in the monoclinic $P2_1/c$ space group, with 2 molecules in the unit cell (Fig. 1). The lattice parameters, obtained from previously published single crystal X-ray diffraction measurements (CSD reference code: NAPHTA15, 100 K) are $a = 7.824$Å, $b = 5.934$ Å, $c = 8.099$ Å, $\alpha = \gamma = 90\,°$, and $\beta = 114.441°$ [47]. The molecules are arranged in an offset π-stacking system, bonding externally through weak van der Waals interactions.

The THz-TDS spectrum of naphthalene below 3 THz has been previously published by King and Korter, where a sharp absorption feature was observed at 2.29 ± 0.01 THz at low temperature [30]. The cryogenic THz-TDS spectrum of naphthalene was reacquired for this study using an instrument with a slightly increased bandwidth, which uncovers a second absorption feature at 3.21 ± 0.02 THz. The experimental spectrum also exhibits a rising background absorption that suddenly increases significantly at the edge of the instrumental bandwidth. This is likely due to a phonon mode around 5 THz, which is predicted by the DFT simulations as shown in Figures S3 and S4 in the ESI.

In order to assign the experimental spectrum, a large number of vibrational simulations were performed using ss-DFT, and the results for the PBE functional are presented in Fig. 2a (other functionals are shown in Figure S3). The simulated spectra were generated by convolving Lorentzian functions with an empirically determined full-width at half-maximum (FWHM) of 0.12 THz. The results highlight that the spectra, while generally similar, experience significant variability. An initial assignment performed by comparing the simulated spectra to the experimentally observed data yields two simulations, PBE/VTZP and PBE/pob-TZVP, that appear
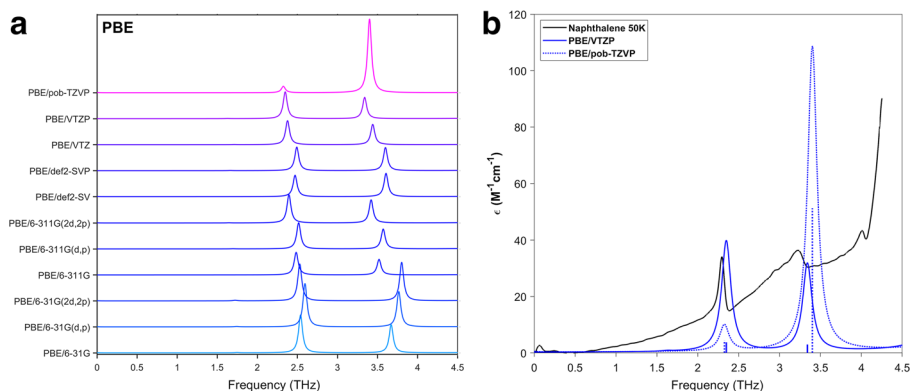
Fig. 2 **a** Naphthalene-simulated terahertz spectra with the PBE functional method. **b** Experimental (black) and predicted (blue solid) terahertz spectrum from the PBE/TZVP method and (blue dash) from PBE/pob-TZVP method. The peak positions of the experimental spectra were generated by Lorentzian line shape fitting, with FWHM of 0.12 THz

to be in agreement with the experimental spectrum. However, close inspection of the predicted mode types (Figures S5 and S6 in the ESI) highlights significant differences between the two methods, which makes a definitive assignment difficult and necessitates deeper investigation.

Further analysis of the various simulations enables direct comparisons between the methods to be obtained. In order to explore the variability between the simulation methods, the calculated results can be internally compared against each other by examining the cross-correlation between the theoretical predicted frequencies, as shown in Fig. 3. The cross-correlation analysis was modified based on a previously reported method [14]. For each normal mode, the predicted transition frequencies were normalized:

$$\nu_i^{\textbf{norm}} = \frac{\nu_i - \overline{\nu}}{\sqrt{n}\sigma(\nu)}$$

where ν_i is the frequency of the ith IR-active normal mode, $\overline{\nu}$ and $\sigma(\nu)$ are the mean and standard deviation of the predicted IR-active modes below 6 THz, respectively, and n is total number of IR-active modes. Upon normalization, the cross-correlation for each simulation was calculated using the standard MATLAB package. While the general range of the correlation coefficient is from 0 to 1, the correlation coefficients are very closely grouped, indicating that overall the results are in fact correlated, as expected. However, within this cross-correlation, minor deviations in the correlation coefficient are meaningful, as they provide insight into how similar—or dissimilar—the various methods are. Within the dataset, the pob-TZVP basis set stands out as being the most poorly correlated compared with the other simulation methods, regardless of the utilized functional. It is important to note that our analysis is focused primarily on the determination of the vibrational frequencies, and the predicted intensities are not taken into account. This is intentional, as the accurate prediction of intensities is dependent on the accurate electronic structure determination, which is
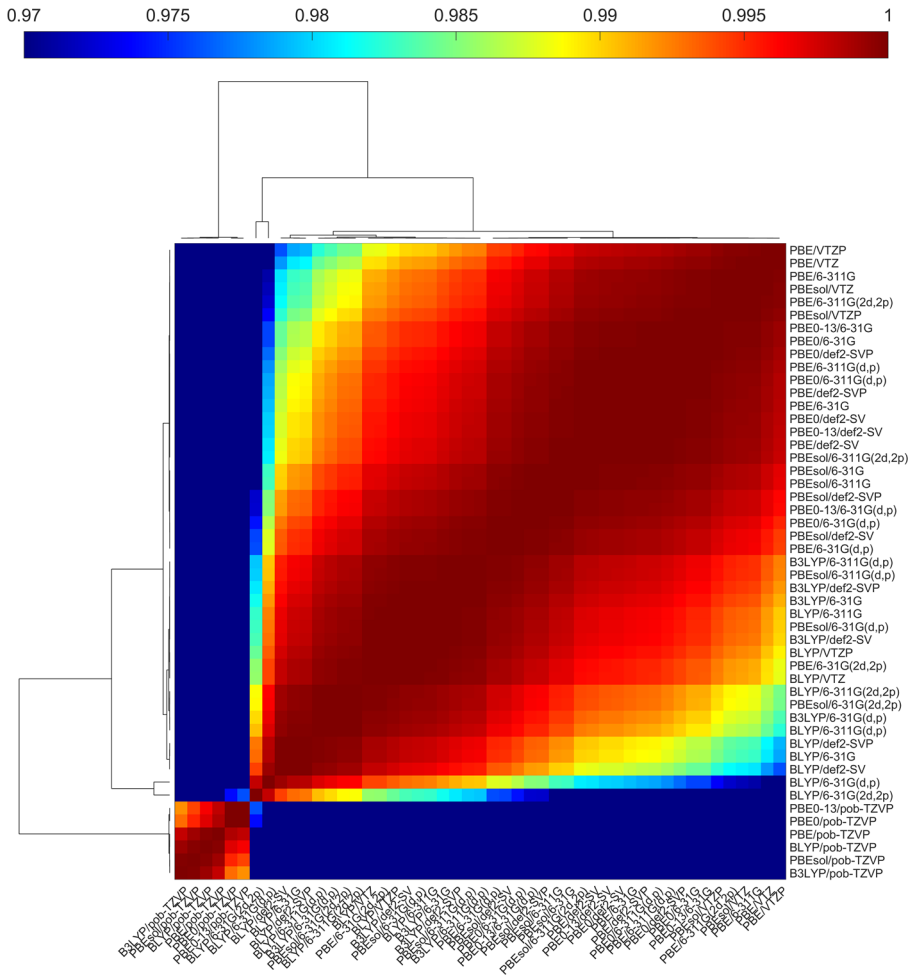
Fig. 3 Cross-correlation heat map of calculated IR-active vibrational transition frequencies predicted from simulations of naphthalene below 6 THz. The methods have been clustered based on correlation, with self-correlation along the diagonal

strongly influenced by a large number of factors besides those related to the prediction of interatomic forces. Furthermore, direct comparisons between experimental and theoretical intensities and peak shapes require empirical data, for example empirical FWHM and the assumption that all peaks follow Lorentzian behavior. Therefore, we limit our analysis to the vibrational transitions.

As mentioned before, two methods appear to be in good agreement with experimental spectrum. However, the cross-correlation analysis indicates that PBE/VTZP and PBE/pob-TZVP have little agreement between them, despite the apparent agreement of their simulated terahertz spectra. While internal consistency between
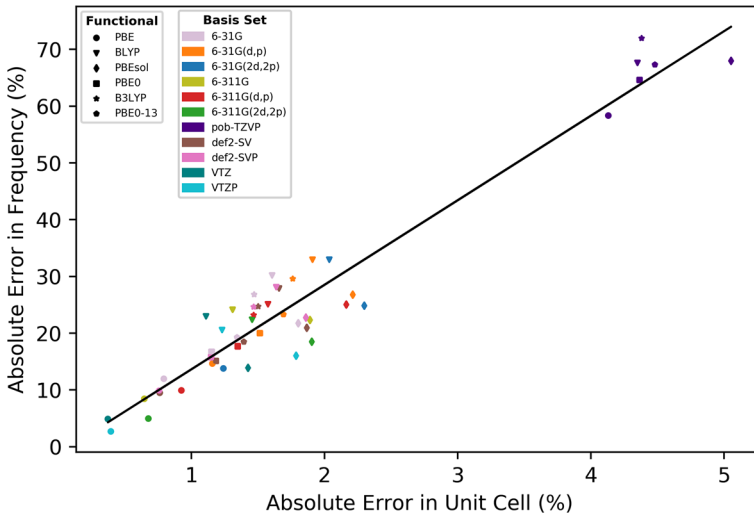
Fig. 4 Relationship between the average absolute error in lattice parameter and vibrational frequency. The black solid line is a least-squared fit

methods alone is not sufficient to describe the accuracy of a simulation, the lack of correlation of the pob-TZVP calculations provides strong evidence of suspect results.

In order to determine the origins of the contrasting results, additional comparisons with experimental results are required. Due to the dependence of vibrational simulations on accurate forces, unconstrained geometry optimizations can be used to provide insight into the accuracy of the interatomic forces. Therefore, the average absolute error in frequency (in the terahertz range) was compared with the average absolute error in unit cell dimension (Fig. 4). The resulting errors in the unit cell parameters and the average absolute error for each methods employed are tabulated and plotted in the Supporting Information. A linear relationship is illustrated, highlighting how poorly calculated forces manifest in both the geometry and frequency calculations. Such a trend suggests that PBE/VTZP is the best method to assign the experimentally observed spectrum, due to smallest errors for both the predicted vibrational frequencies and predicted unit cell parameters. Clearly, simulations employing the pob-TZVP basis set exhibit a high relative error in predicted structural parameters, specifically a significant contraction of all lattice vectors compared with the X-ray structure. Such a contraction indicates that the strength of the intermolecular forces is overestimated, which manifests in a blue shift of the predicted vibrational features. Specifically, the mode that appears to be in agreement with the experimental spectrum actually corresponds to the second experimentally observed absorption peak, where the intensity is reduced because the blue shift results in a lower magnitude displacement that ultimately reduces the value of the transition dipole moment derivative. Thus, PBE/pob-TZVP actually is not a good method, and highlights that a low-frequency assignment should not be compared with the experimentally observed data strictly based upon proximity of calculated vibrational frequencies, but also the structural data from the geometric optimization should also be considered. This is a

critical finding, as it is readily apparent that theoretical results can be misleading if not evaluated in the context of other experimental data. Without question, additional effects such as temperature, anharmonicity, and vibrational relaxation can play an important role in the final experimental frequencies; however, as shown here, it is possible to remove potential pitfalls through evaluation of structural information.

With a suitable method found, the two experimental absorption features could be assigned. The experimental mode at 2.29 THz corresponds to an inter-sheet π-stacking translation motion along the b-unit cell axis, and the latter mode at 3.21 THz corresponds to a π-stacking translation motion along the diagonal direction of a- and c-unit cell axes with simultaneous molecular bending, as shown in Figures S5 and S6.

Upon assignment, it is clear that contraction of the unit cell would significantly alter the predicted frequencies of these modes, highlighting the link between forces, geometry, vibrational frequency, and vibrational mode type.

3.2 Oxalic Acid

While naphthalene is a good model system for the evaluation of simple organic materials involving only π-stacking and van der Waals forces, it does not display other non-covalent interactions common to organic molecular crystals, for example hydrogen bonding. Thus, anhydrous oxalic acid, a simple organic crystal that predominantly features hydrogen bonding interactions, was selected as a second test system. Anhydrous oxalic acid crystallizes in the orthorhombic *Pbca* space group (Fig. 5). There are 4 molecules in the unit cell, with corresponding lattice parameters of $a = 6.493$ Å, $b = 6.060$ Å, $c = 7.803$ Å, and $\alpha = \beta = \gamma = 90°$ (CSD reference code: OXALAC05, 130 K) [48]. Four vibrational modes are observed below 4.25 THz at 50 K, occurring at 2.10 ± 0.04 THz, 2.99 ± 0.03 THz, 3.83 ± 0.02 THz, and 3.99 ± 0.01 THz, respectively. Similarly to naphthalene, a series of simulations were performed using a variety of differing basis sets and functionals, yielding the
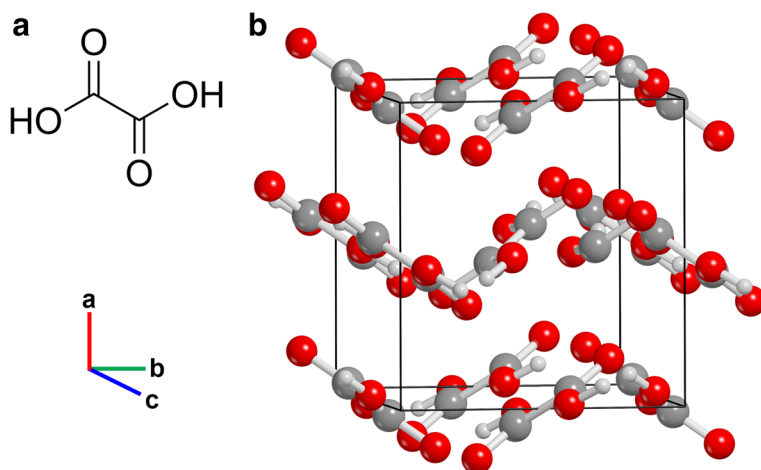

Fig. 5 **a** Molecular structure and **b** crystalline unit cell of anhydrous oxalic acid

spectra displayed in Fig. 6. As with naphthalene, a range of conflicting spectra are produced, with some containing features that correlate well with the experimentally observed spectrum, while others offer little agreement. The agreement of each predicted spectrum to the experimentally observed spectrum is a crucial aspect in the evaluation of simulation accuracy, but the relative differences and similarities of the simulations also provide strong insight into the specific theoretical considerations that allow for such accuracy. Thus, further analysis of these methods was performed by comparing the relative differences of terahertz active modes, specifically with respect to the calculated vibrational transition frequency.

These discrepancies are illustrated in the cross-correlation analysis shown in Fig. 7. In contrast to naphthalene, the simulations involving basis sets without additional polarization functions are distinct, while larger triple-ζ are more closely correlated. The employed functional, however, does not present significant deviations between the predicted spectra, with both hybrid and GGA functionals capable of strong and weak correlations, based upon the size of the employed basis set. However, PBEsol is an exception to this trend, and consistently presents poor correlation with respect to the other theoretical models—a surprising finding given that this functional was developed specifically for solids [39].

Crystalline anhydrous oxalic acid is heavily dominated by hydrogen bonding, a force that ss-DFT is generally capable of accurately capturing [5, 16]. Due to the strong electrostatics that dictate hydrogen bonds, and that it is the dominant contribution to the binding forces of anhydrous oxalic acid, the choice of functional might not be as critical. However, in order to capture the perturbations to electron density introduced by the strong electrostatics, additional flexibility to the atomic wavefunctions must be introduced, in this case through additional polarization functions. Thus, the inclusion of polarization functions is a critical parameter for the accuracy of the simulation, as the hydrogen bonding is more appropriately accounted for [16]. The
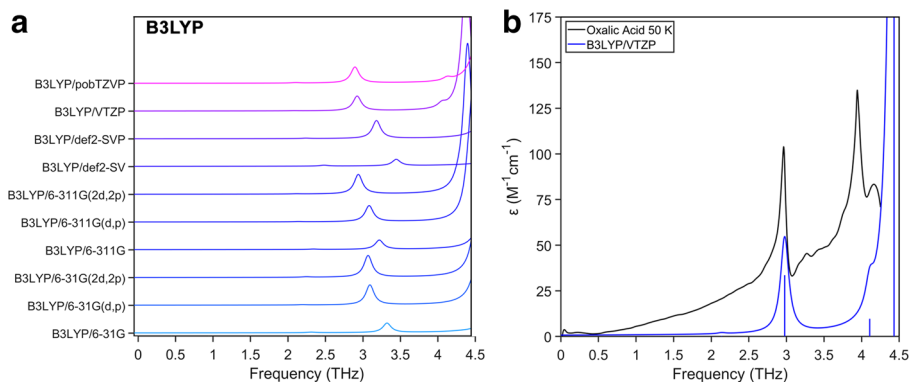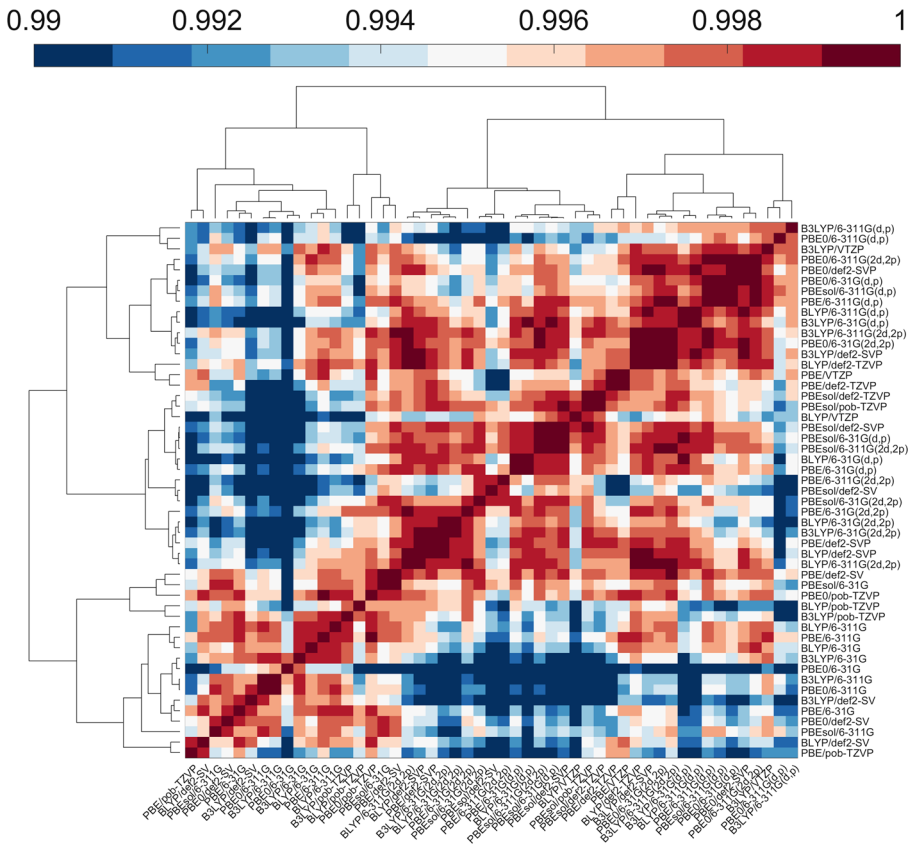

Fig. 6 **a** Results of anhydrous oxalic acid calculations performed with the B3LYP functional, with basis sets labeled for each spectrum. **b** Experimental 50 K anhydrous oxalic acid spectrum overlaid with calculated spectrum produced by the B3LYP/VTZP method. The spectra have been generated with Lorentzian fits around the predicted vibrational transition frequencies, with FWHM values of 0.15 THz

effects of additional polarization functions, size of basis set, and selection of functional are best interpreted in the evaluation of relative errors of each simulation, both with respect to their structural and spectral reproduction of the experimentally observed data.

The experimentally observed terahertz spectrum of anhydrous oxalic acid (Fig. 6) is in agreement with the previously published spectrum from King and Korter [16]. The B3LYP/VTZP method produces an optimized geometry most closely resembling the experimentally observed unit cell parameters, offering an absolute percent error with respect to unit cell axes of 0.79% (Fig. 8), as well as predicted vibrational transition frequencies in close agreement with the experimentally observed spectrum, and is thus used to assign the experimentally observed spectrum, as it offers small improvements over previous studies [16]. The spectra presented in Fig. 6 have been generated using an empirical FWHM of 0.15 THz. This method predicts vibrational transition frequencies occurring at 2.138 THz, 2.977 THz, 4.108 THz, and 4.431 THz (Fig. 6), corresponding to an asymmetric sliding in the direction of the *c*-unit cell axis of layers of molecules stacked along the *a*-unit cell axis, an intramolecular torsional motion about the C–C bond coupled with a slight translation along the *a*-unit cell axis, a pinching-like movement along the *a*-unit cell axis of layers of molecules, and a stretching/contracting oscillation along the *b*-unit cell axis within layers of molecules. Animations and graphical schematics of these motions are available in the ESI. Again, a correlative effect can be seen between the relative accuracy of the predicted vibrational transition frequencies and predicted unit cell parameters. Consequently, the theoretical terahertz spectra that most closely resemble the experimentally observed spectrum are produced from methods producing a small ($< 2\%$) error in the unit cell axes. The largest contributions to overall absolute average errors of unit cell axes are largely in the *b*-unit cell axis (Figure S2), which constitutes the direction of hydrogen bonding within layers of oxalic acid molecules. Polarization functions, in the case of anhydrous oxalic acid, greatly increase the accuracy of the theoretical methods, both with respect to the experimentally observed spectrum, and convergence to similarly accurate methods. Thus, the simulations in which this axis is not accurately predicted, commonly those employing basis sets without additional polarization functions, present an inaccurate depiction of the forces present in the molecular crystal, which yield theoretical spectra in poor agreement with the experimental spectrum. This trend can be seen in both Figs. 7 and 8, with methods including polarization functions most closely correlated, with lower errors with respect to lattice parameters, as well as predicted vibrational transition frequencies. These results emphasize that systems with extensive hydrogen bonding must include polarization functions in their respective theoretical model to reach agreement with the experimentally observed data.

3.3 Thymine

The aforementioned solids are excellent models for assessing basic forces such as π-stacking and hydrogen bonding, respectively. However, many organic molecular crystals are complex systems, displaying a large array of weak intermolecular forces that are often entangled, each of which contribute to the complex low-frequency

Fig. 7 Cross-correlation heat map of calculated IR-active vibrational transition frequencies predicted from simulations of anhydrous oxalic acid. The methods have been sorted to maximize correlation, with self-correlation along the diagonal

vibrations. Thus, thymine, one of the four DNA bases, was chosen as a system in which a complex mixture of such intermolecular forces is present. Thymine crystallizes in the monoclinic $P2_1/c$ space group, with the unit cell containing four molecules (Fig. 9), and lattice parameters of $a = 12.889$ Å, $b = 6.852$ Å, $c = 6.784$ Å, $\alpha = \gamma = 90°$, and $\beta = 104.92°$ (CSD reference code: THYMIN01 [49], 293 K).

Thymine exhibits four well-resolved spectral features below 4.25 THz at 50 K, occurring at 1.27 ± 0.01 THz, 2.10 ± 0.01 THz, 2.29 ± 0.03 THz, and 2.91 ± 0.02 THz, respectively (Fig. 10). The experimentally measured terahertz spectrum is in agreement with the cryogenic spectrum of thymine collected by Fischer et al. [50] The theoretical spectra presented in Fig. 10 have been generated using an empirical FWHM of 0.15 THz. The hydrogen bonding, π-stacking, and close intermolecular packing present in the crystalline structure of thymine (Fig. 9) create significant difficulty in the accurate simulation of its corresponding terahertz motions, resulting in drastic variations in the predicted terahertz spectra based upon the utilized model. A number of these simulations demanded rigorous treatments
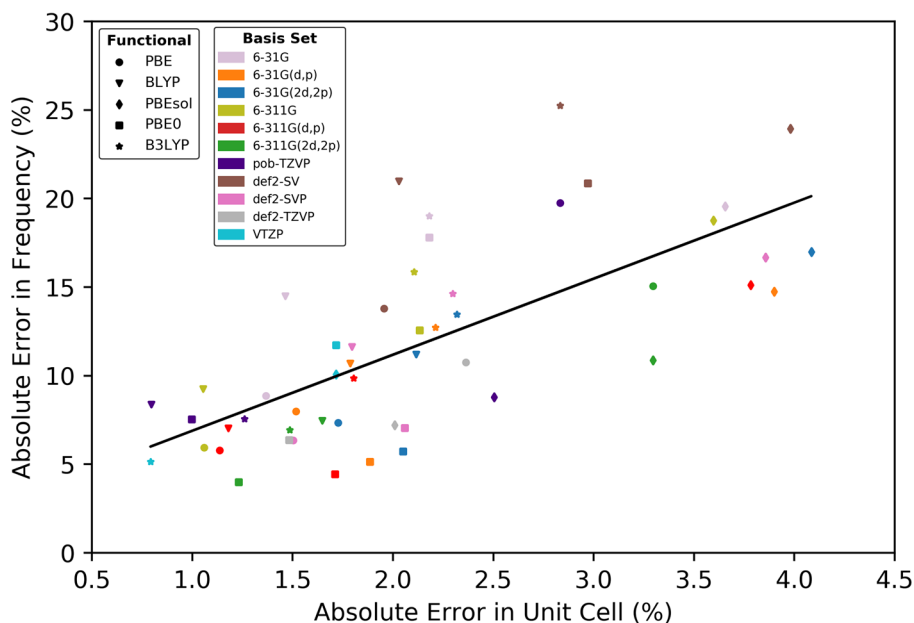
Fig. 8 Average absolute error of predicted unit cell axes against average absolute error in predicted vibrational transition frequencies of anhydrous oxalic acid below 4.25 THz. A linear fit to the data is represented by a solid black line

of the exchange and correlation integrals, which ultimately resulted in a failure to converge, highlighting the difficulty in the simulation of the system as a whole. Consequently, the number of total calculations performed upon thymine is slightly reduced from those of oxalic acid and naphthalene. Of the successful simulations, a number of methods greatly underestimated the unit cell volume, leading to large errors in the predicted lattice vectors, in some cases on the order of angstroms. It is important to note that the experimental thymine structure used in this comparison was collected at ambient conditions. Consequently, the errors in lattice vectors are significantly larger than those of oxalic acid and napthalene, whose experimental structures determined at low temperatures. Even larger triple-ζ basis sets produced such errors, for example the calculation performed using the PBEsol/6-311G(d,p) reduced the a-unit cell vector to a value of 7.957 Å, an absolute percent error of this axis of 38.27%, with an average absolute error of unit cell axes of 21.88% (Fig. 11). For these reasons, the assignment of the experimentally observed spectrum must be conducted carefully.

Due to the strong correlative effect demonstrated in the previously aforementioned solids between the error in the optimized unit cell axes and error in the resultant predicted vibrational transition frequencies, the experimental spectrum is in reasonable agreement with the theoretical spectrum produced with the PBE/def2-TZVP method. This method accurately predicts vibrational transition frequencies occurring at 1.316 THz, 1.966 THz, and 2.382 THz. These modes correspond to a butterfly like motion of hydrogen-bonded dimers along the c-axis, an asymmetric rocking
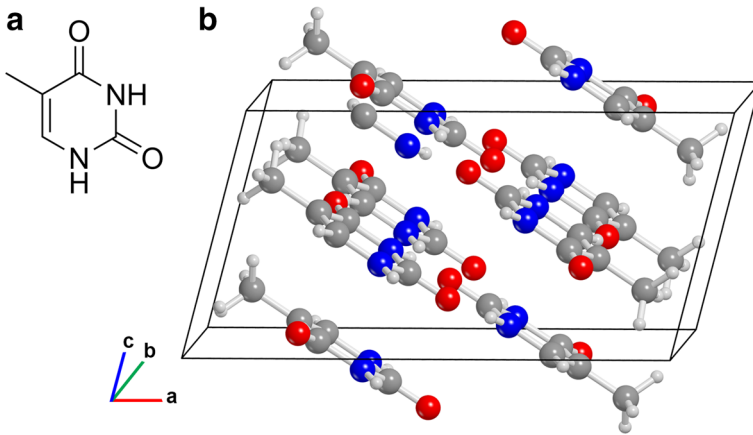
Fig. 9 Molecular structure **a** and crystalline unit cell **b** of thymine

of hydrogen-bonded dimers coupled with a slight methyl group rotation, and an asymmetric translation along the $\pi-$stacking direction of hydrogen-bonded dimers, coupled with a methyl group rotation, respectively. These predictions are in strong agreement with the experimentally observed spectrum. The largest discrepancies between the calculated vibrational transition frequencies from this method and the experimentally observed spectrum fall between 2.5 and 3.5 THz, where three vibrational transition frequencies are predicted, contrasting with the single broad feature centered at 2.91 THz. However, it is likely that this feature contains several unresolved features at 50 K, and for this reason, the analyses of all thymine simulations are based solely upon the first three modes to retain uniformity in the analysis. It is important to note that the predicted features' mode types were analyzed, along with the mode symmetry, in order to ensure direct comparisons between the methods
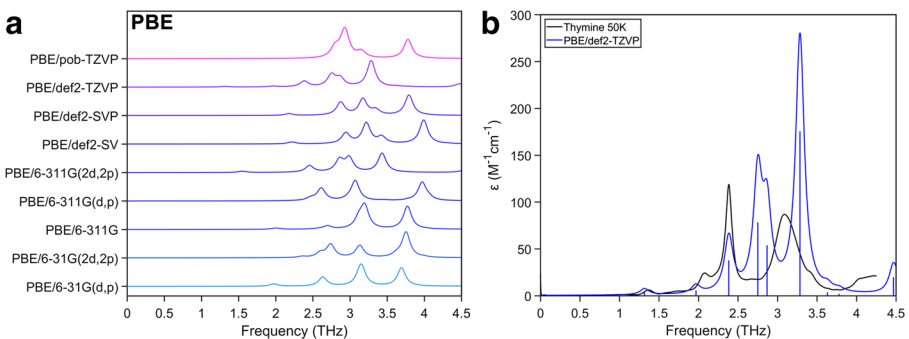
Fig. 10 **a** Results of thymine calculations performed with the PBE functional, with basis sets labeled for each spectrum. **b** Experimental 50 K thymine spectrum overlaid with calculated spectrum produced by the PBE/def2-TZVP method. The spectra have been generated with Lorentzian fits to the predicted vibrational transition frequencies, with FWHM values of 0.15 THz
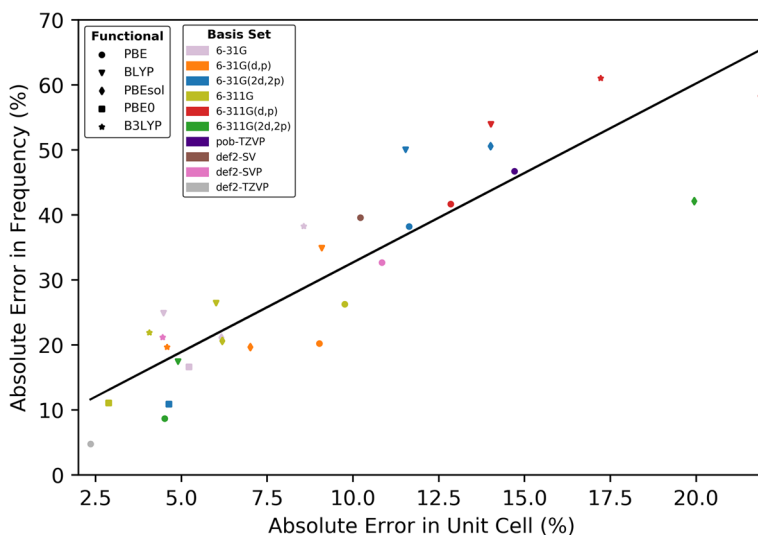
Fig. 11 Average absolute error of predicted unit cell axes against average absolute error in predicted vibrational transition frequencies of thymine below 4.25 THz. A linear fit to the data is represented by a solid black line

were appropriately made. Under this premise, the simulations most closely resembling the experimentally observed unit cell parameters produce the least deviation with respect to the experimentally observed terahertz spectrum, in agreement with the findings of oxalic acid and naphthalene. The relative similarities and differences of the methods employed in the simulation of phonons present in the structure of crystalline thymine are visualized in Fig. 12. Here, a combination of hybrid functionals (PBE0 and B3LYP) and large basis sets with added polarization functions clearly perform the best, indicating the rigor required to accurately capture the complex forces present in this sample. This trend is readily observed in Figs. 11 and 12, with methods employing hybrid functionals and large basis sets most closely correlated, and theoretical models that employ both of these considerations producing lattice parameters and predicted vibrational transition frequencies that most closely resemble the experimentally observed data. This suggests that systems presenting an elaborate interplay of differing intermolecular forces require especially rigorous calculations.

4 Conclusions

Solid-state density functional theory is an effective method for the prediction of terahertz spectra of molecular crystals, but the choice of theoretical method is central to the successful reproduction of the experimentally observed data. The results of this work have highlighted the large variability in the computed terahertz spectra of solids when using some of the more popular DFT methodologies. While each of the studied systems was different, including different interatomic interactions and
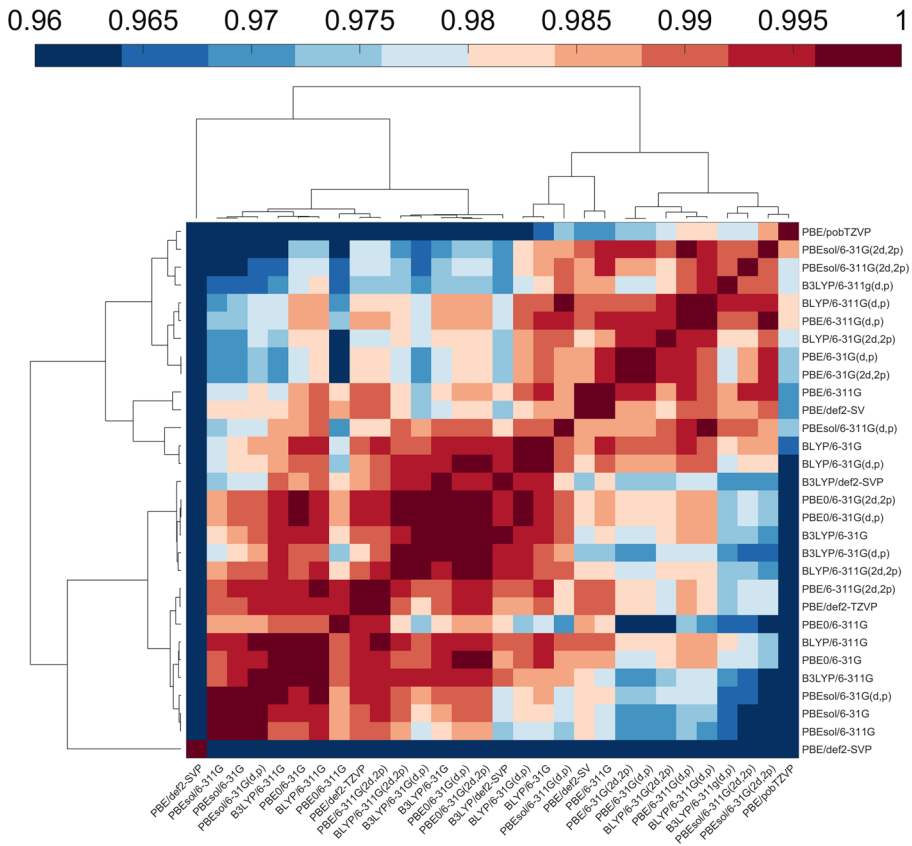
Fig. 12 Cross-correlation heat map of calculated IR-active vibrational transition frequencies predicted from simulations of thymine. The methods have been sorted to maximize correlation, with self-correlation along the diagonal

packing arrangements, in general, larger basis sets and hybrid functionals tended to produce more accurate results. Each of the three studied systems was best modeled by triple-zeta basis sets, namely VTZP for naphthalene and oxalic acid, and def2-TZVP for thymine. However, as demonstrated in the case of pob-TZVP, size is not the only factor that matters, but the quality of the basis set construction can play a critical role on the computed results. In terms of polarization functions, they were found to be of little importance for naphthalene, but were critical in oxalic acid and thymine, likely due to the nature of the intermolecular interactions (hydrogen bonds) promoting a more anisotropic electron distribution. Thus, it is critical that for future work, these factors are adequately checked and benchmarked for each individual system, in order to utilize the most appropriate method. In terms of functional, it was rather surprising that the PBEsol functional, which was developed specifically for solid applications, performed poorly in each of the studied materials. Overall, these results highlight the importance of accurately modeling interatomic forces for the simulation of terahertz spectra, and how bulk structural optimizations can be used as

a proxy for the evaluation of such forces. Thus, experimental terahertz spectra should be assigned using theoretical parameters that led to a sufficient representation of the three-dimensional structure. This study highlights the correlation between the two, and suggests minimizing error in structural parameters before undergoing vibrational frequency calculations. While each material is different, the results of this work have found that errors in the lattice vectors above 2% tend to result in poor representations of the terahertz spectra, and this should be taken into account when performing these analyses. This method provides the most potential for success, and provides a framework for future studies using THz-TDS.

Drawing further comparisons across the studied systems, the PBEsol density functional presented generally inaccurate results. Similarly, the pob-TZVP basis set demonstrated mostly inaccurate results across the three studied systems. These results highlight the importance of accurately modeling interatomic forces for the simulation of terahertz spectra, and how bulk structural optimizations can be used as a proxy for the evaluation of the forces. However, when calculating the error in lattice parameters, the theoretical results should be compared with experimental structural data collected at low temperatures for a better representation of the delicate forces governing terahertz spectra.

References

1. Schweicher, G.; D'Avino, G.; Ruggiero, M. T.; Harkin, D. J.; Broch, K.; Venkateshvaran, D.; Liu, G.; Richard, A.; Ruzié, C.; Armstrong, J.; Kennedy, A. R.; Shankland, K.; Takimiya, K.; Geerts, Y. H.; Zeitler, J. A.; Fratini, S.; Sirringhaus, H. "Chasing the "Killer" Phonon Mode for the Rational Design of Low–Disorder, High–Mobility Molecular Semiconductors". *Adv. Mater.*, **31** (43), 2019, 1902407.
2. Zhang, W.; Maul, J.; Vulpe, D.; Moghadam, P. Z.; Fairen-Jimenez, D.; Mittleman, D. M.; Axel Zeitler, J.; Erba, A.; Ruggiero, M. T. "Probing the Mechanochemistry of Metal-Organic Frameworks with Low-Frequency Vibrational Spectroscopy". *J. Phys. Chem. C*, **122** (48), 2018, 27442–27450.
3. Ruggiero, M. T.; Sutton, J. J.; Fraser-Miller, S. J.; Zaczek, A. J.; Korter, T. M.; Gordon, K. C.; Zeitler, J. A. "Revisiting the Thermodynamic Stability of Indomethacin Polymorphs with Low-Frequency Vibrational Spectroscopy and Quantum Mechanical Simulations". *Cryst. Growth Des.*, **18**, 2018, 6513–6520.
4. Ruggiero, M. T.; Sibik, J.; Orlando, R.; Zeitler, J. A.; Korter, T. M. "Measuring the Elasticity of Poly-L-Proline Helices with Terahertz Spectroscopy". *Angew. Chem. Int. Ed.*, **55** (24), 2016, 6877–6881.
5. Delaney, S. P.; Korter, T. M. "Terahertz Spectroscopy and Computational Investigation of the Flufenamic Acid/Nicotinamide Cocrystal". *J. Phys. Chem. A*, **119**, 2015, 3269–3276.
6. Allis, D.; Fedor, A.; Korter, T.; Bjarnason, J.; Brown, E. "Assignment of the lowest-lying THz absorption signatures in biotin and lactose monohydrate by solid-state density functional theory". *Chem. Phys. Lett.*, **440** (4-6), 2007, 203–209.
7. Neu, J.; Stone, E. A.; Spies, J. A.; Storch, G.; Hatano, A. S.; Mercado, B. Q.; Miller, S. J.; Schmuttenmaer, C. A. "Terahertz Spectroscopy of Tetrameric Peptides". *J. Phys. Chem. Lett.*, **10**, 2019, 2624–2628.
8. Tan, N. Y.; Ruggiero, M. T.; Orellana-Tavra, C.; Tian, T.; Bond, A. D.; Korter, T. M.; Fairen-Jimenez, D.; Axel Zeitler, J. "Investigation of the terahertz vibrational modes of ZIF-8 and ZIF-90 with terahertz time-domain spectroscopy". *Chem. Commun.*, **51** (89), 2015, 16037–16040.
9. Parrott, E. P.; Zeitler, J. A. "Terahertz time-domain and low-frequency Raman spectroscopy of organic materials". *Appl. Spectrosc.*, **69**, 2015, 1–25.
10. Ruggiero, M. T.; Zeitler, J. A.; Korter, T. M. "Concomitant polymorphism and the martensitic-like transformation of an organic crystal". *Phys. Chem. Chem. Phys*, **19**, 2017, 28502–28506.

11. Ruggiero, M. T.; Sibik, J.; Zeitler, J. A.; Korter, T. M. "Examination of L-Glutamic Acid Polymorphs by Solid-State Density Functional Theory and Terahertz Spectroscopy". *J. Phys. Chem. A*, **120**, 2016, 7490–7495.

12. Tanno, T.; Asari, J.; Yodokawa, S.; Kurabayashi, T. "Terahertz spectroscopic study on order-disorder phase transition of nonadecane". *Chem. Phys.*, **461**, 2015, 25–28.

13. Li, R.; Zeitler, J. A.; Tomerini, D.; Parrott, E. P.; Gladden, L. F.; Day, G. M. "A study into the effect of subtle structural details and disorder on the terahertz spectrum of crystalline benzoic acid". *Phys. Chem. Chem. Phys.*, **12**, 2010, 5329–5340.

14. Kendrick, J.; Burnett, A. D. "Exploring the Reliability of DFT Calculations of the Infrared and Terahertz Spectra of Sodium Peroxodisulfate". *J. Infrared Millim. Terahertz Waves (In Press)*.

15. Jepsen, Peter Uhd, Clark, S. J. "Precise ab-initio prediction of terahertz vibrational modes in crystalline systems". *Chem. Phys. Lett*, **442**, 2007, 257–280.

16. King, M. D.; Korter, T. M. "Effect of Waters of Crystallization on Terahertz Spectra: Anhydrous Oxalic Acid and Its Dihydrate". *J. Phys. Chem. A*, **114** (26), 2010, 7127–7138.

17. Delaney, S. P.; Korter, T. M. "Terahertz Spectroscopy and Computational Investigation of the Flufenamic Acid/Nicotinamide Cocrystal". *J. Phys. Chem. A.*, **119**, 2015, 3269–3276.

18. Ruggiero, M. T.; Gooch, J.; Zubieta, J.; Korter, T. M. "Evaluation of Range-Corrected Density Functionals for the Simulation of Pyridinium-Containing Molecular Crystals". *J. Phys. Chem. A*, **120** (6), 2016, 939–947.

19. Ruggiero, M. T.; Zeitler, J. A.; Erba, A. "Intermolecular anharmonicity in molecular crystals: interplay between experimental low-frequency dynamics and quantum quasi-harmonic simulations of solid purine". *Chem. Commun.*, **53**, 2017, 3781–3784.

20. Ruggiero, M. T.; Gooch, J.; Zubieta, J.; Korter, T. M. "Evaluation of Range-Corrected Density Functionals for the Simulation of Pyridinium-Containing Molecular Crystals". *J. Phys. Chem. A*, **120** (6), 2016, 939–947.

21. Erba, A.; Baima, J.; Bush, I.; Orlando, R.; Dovesi, R. "Large-Scale Condensed Matter DFT Simulations: Performance and Capabilities of the CRYSTAL Code". *J. Chem. Theory Comput.*, **13** (10), 2017, 5019–5027.

22. Dovesi, R.; Erba, A.; Orlando, R.; Zicovich-Wilson, C. M.; Civalleri, B.; Maschio, L.; Rérat, M.; Casassa, S.; Baima, J.; Salustro, S.; Kirtman, B. "Quantum-mechanical condensed matter simulations with CRYSTAL". *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, **8** (4), 2018, e1360.

23. Pascale, F.; Zicovich-Wilson, C. M.; López Gejo, F.; Civalleri, B.; Orlando, R.; Dovesi, R. "The calculation of the vibrational frequencies of crystalline compounds and its implementation in the CRYSTAL code". *J. Comput. Chem.*, **25** (6), 2004, 888–897.

24. Zicovich-Wilson, C. M.; Pascale, F.; Roetti, C.; Saunders, V. R.; Orlando, R.; Dovesi, R. "Calculation of the vibration frequencies of α-quartz: The effect of Hamiltonian and basis set". *J. Comput. Chem.*, **25** (15), 2004, 1873–1881.

25. Jepsen, P. U.; Clark, S. J. "Precise ab-initio prediction of terahertz vibrational modes in crystalline systems". *Chem. Phys. Lett.*, **442** (4), 2007, 275 – 280.

26. Delaney, S. P.; Pan, D.; Galella, M.; Yin, S. X.; Korter, T. M. "Understanding the Origins of Conformational Disorder in the Crystalline Polymorphs of Irbesartan". *Cryst. Growth Des.*, **12** (10), 2012, 5017–5024.

27. Pascale, F.; Tosoni, S.; Zicovich-Wilson, C.; Ugliengo, P.; Orlando, R.; Dovesi, R. "Vibrational spectrum of brucite, $Mg(OH)_2$: a periodic ab initio quantum mechanical calculation including OH anharmonicity". *Chem. Phys. Lett.*, **396** (4), 2004, 308 – 315.

28. Noel, Y.; Zicovich-Wilson, C. M.; Civalleri, B.; D'Arco, P.; Dovesi, R. "Polarization properties of ZnO and BeO: An ab initio study through the Berry phase and Wannier functions approaches". *Phys. Rev. B*, **65**, 2001, 014111.

29. Dovesi, R.; Kirtman, B.; Maschio, L.; Maul, J.; Pascale, F.; Rérat, M. "Calculation of the Infrared Intensity of Crystalline Systems. A Comparison of Three Strategies Based on Berry Phase, Wannier Function, and Coupled–Perturbed Kohn–Sham Methods". *J. Phys. Chem. C*, **123**, 2018, 8336–8346.

30. King, M. D.; Korter, T. M. "Modified Corrections for London Forces in Solid-State Density Functional Theory Calculations of Structure and Lattice Dynamics of Molecular Crystals". *J. Phys. Chem. A*, **116** (25), 2012, 6927–6934.

31. Francl, M. M.; Pietro, W. J.; Hehre, W. J.; Binkley, J. S.; Gordon, M. S.; DeFrees, D. J.; Pople, J. A. "Self–consistent molecular orbital methods. A polarization–type basis set for second–row elements". *J. Chem. Phys.*, **77** (7), 1982, 3654–3665.

32. Hehre, W. J.; Ditchfield, R.; Pople, J. A. "Self–Consistent Molecular Orbital Methods. Further Extensions of Gaussian–Type Basis Sets for Use in Molecular Orbital Studies of Organic Molecules". *J. Chem. Phys.*, **56** (5), 1972, 2257–2261.

33. Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. "Self–consistent molecular orbital methods. A basis set for correlated wave functions". *J. Chem. Phys.*, **72** (1), 1980, 650–654.

34. Schäfer, A.; Horn, H.; Ahlrichs, R. "Fully optimized contracted Gaussian basis sets for atoms Li to Kr". *J. Chem. Phys.*, **97** (4), 1992, 2571–2577.

35. Schäfer, A.; Huber, C.; Ahlrichs, R. "Fully optimized contracted Gaussian basis sets of triple zeta valence quality for atoms Li to Kr". *J. Chem. Phys.*, **100** (8), 1994, 5829–5835.

36. Weigend, F.; Ahlrichs, R. "Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy". *Phys. Chem. Chem. Phys.*, **7**, 2005, 3297–3305.

37. Peintinger, M. F.; Oliveira, D. V.; Bredow, T. "Consistent Gaussian basis sets of triple-zeta valence with polarization quality for solid-state calculations". *J. Comput. Chem.*, **34** (6), 2013, 451–459.

38. Perdew, J. P.; Burke, K.; Ernzerhof, M. "Generalized Gradient Approximation Made Simple". *Phys. Rev. Lett.*, **77**, 1996, 3865–3868.

39. Perdew, J. P.; Ruzsinszky, A.; Csonka, G. I.; Vydrov, O. A.; Scuseria, G. E.; Constantin, L. A.; Zhou, X.; Burke, K. "Restoring the Density-Gradient Expansion for Exchange in Solids and Surfaces". *Phys. Rev. Lett.*, **100**, 2008, 136406.

40. Becke, A. D. "Density-functional exchange-energy approximation with correct asymptotic behavior". *Phys. Rev. A*, **38**, 1988, 3098–3100.

41. Adamo, C.; Barone, V. "Toward reliable density functional methods without adjustable parameters: The PBE0 model". *J. Chem. Phys.*, **110** (13), 1999, 6158–6170.

42. Guido, C. A.; Brémond, E.; Adamo, C.; Cortona, P. "Communication: One third: A new recipe for the PBE0 paradigm". *J. Chem. Phys.*, **138** (2), 2013, 021104.

43. Civalleri, B.; Zicovich-Wilson, C. M.; Valenzano, L.; Ugliengo, P. "B3LYP augmented with an empirical dispersion term (B3LYP-D*) as applied to molecular crystals". *CrystEngComm*, **10**, 2008, 405–410.

44. Grimme, S. "Density functional theory with London dispersion corrections". *WIREs Comput. Mol. Sci.*, **1** (2), 2011, 211–228.

45. Grimme, S.; Ehrlich, S.; Goerigk, L. "Effect of the damping function in dispersion corrected density functional theory". *J. Comput. Chem.*, **32** (7), 2011, 1456–1465.

46. Smith, D. G. A.; Burns, L. A.; Patkowski, K.; Sherrill, C. D. "Revised Damping Parameters for the D3 Dispersion Correction to Density Functional Theory". *J. Phys. Chem. Lett.*, **7** (12), 2016, 2197–2203.

47. Oddershede, J.; Larsen, S. "Charge Density Study of Naphthalene Based on X-ray Diffraction Data at Four Different Temperatures and Theoretical Calculations". *J. Phys. Chem. A*, **108** (6), 2004, 1057–1063.

48. Thalladi, V. R.; Nu, M.; Boese, R. "The Melting Point Alternation in R,ω-Alkanedicarboxylic Acids". *J. Am. Chem. Soc.*, **122** (38), 2000, 9227–9236.

49. Portalone, G.; Bencivenni, L.; Colapietro, M.; Pieretti, A.; Ramondo, F. "The Effect of Hydrogen Bonding on the Structures of Uracil and Some Methyl Derivatives Studied by Experiment and Theory". *Acta Chem. Scand.*, **53**, 1999, 57–68.

50. Fischer, B.; Hoffmann, M.; Helm, H.; Modjesch, G.; Jepsen, P. U. "Chemical recognition in terahertz time-domain spectroscopy and imaging". *Semicond. Sci. Techol.*, **20** (7), 2005, S246–S243.