Resolving the Origins of Crystalline Anharmonicity Using Terahertz Time-Domain Spectroscopy and \textit{ab initio} Simulations

Michael T. Ruggiero and J. Axel Zeitler*

\textit{Department of Chemical Engineering and Biotechnology, University of Cambridge, New Museums Site, Pembroke Street, Cambridge, CB2 3RA, United Kingdom}

E-mail: jaz22@cam.ac.uk

Phone: +44 (0) 1223 334783
Abstract

Anharmonicity has been shown to be an important piece of the fundamental framework that dictates numerous observable phenomena. In particular, anharmonicity is the driving force of vibrational relaxation processes, mechanisms that are integral to the proper function of numerous chemical processes. However, elucidating its origins has proven difficult due to experimental and theoretical challenges, specifically related to separating the anharmonic contributions from other unrelated effects. While no one technique is particularly suited for providing a complete picture of anharmonicity, by combining multiple complementary methods such a characterisation can be made. In this study the role of individual atomic interactions on the anharmonic properties of crystalline purine, the building block of many DNA and RNA nucleobases, is studied by experimental terahertz time-domain spectroscopy and first-principles density functional theory (DFT) and ab initio molecular dynamics simulations (AIMD). In particular, the detailed vibrational information provided by the DFT calculations is used to interpret the atomic origins of anharmonic-related effects as determined by the AIMD calculations, which are in good agreement with the experimental data. The results highlight that anharmonicity is especially pronounced in the intermolecular interactions, particularly along the amine hydrogen bond coordinate, and yields valuable insight into what is similarly observed complex biosystems and crystalline solids.

Introduction

Molecular vibrations, including their relaxation dynamics, play a vital role in chemical processes where they are often responsible for transporting energy throughout a system, for example during catalysis, structural reorientation, and biomolecular ligand binding events.\textsuperscript{1–6} Vibrational relaxation can be incredibly complex, with motions propagating through hundreds, to even thousands, of atoms within very short time periods (< 1 ps).\textsuperscript{7,8} Despite this, the fundamental origins of these relaxation pathways can be traced to the anharmonicity of individual vibrations, a property that allows for coupling between two (or more) modes.\textsuperscript{9–12}
Such is the case in DNA and RNA, where it has been suggested that vibrational relaxation, originating from anharmonicity in the amine hydrogen bonds of the bases, is critical for dissipating the energy from electronic absorption events, important for maintaining structural and chemical integrity.\textsuperscript{13–17} Recent work has shown that relaxation in Watson-Crick bases is greatly enhanced through intermolecular interactions, indicating that both the molecular and bulk potential energy landscapes must be well understood in order to fully comprehend anharmonic coupling in DNA and RNA.\textsuperscript{16,17} But while a great deal is known about these interactions from bulk-phase experiments, the contributions from individual atoms or functional groups is difficult to achieve with high accuracy.\textsuperscript{18–23} On the other hand, crystalline materials offer the advantage of having a well-defined, regular structure, enabling the extraction of very precise atomic-level details regarding the forces present.\textsuperscript{24–26} Moreover, the dynamics observed in complex biomolecules are often similarity present in crystalline solids and thus provide an excellent opportunity for detailed study.\textsuperscript{27,28} And although biomolecular systems like DNA are not typically static nor do they exhibit long-range order, there does exist a significant amount of localised order on the molecular level, \textit{i.e.} base pairing and stacking,\textsuperscript{29–31} which is precisely where anharmonicity effects are believed to originate.

Crystalline purine (Figure 1) contains many of the same interaction types that are found in nature,\textsuperscript{32} specifically the same amine hydrogen bonding pattern that is believed to be the source of anharmonic-related effects in DNA, and is an excellent model system to explore how individual anharmonic interatomic potentials can influence bulk properties. Yet even the use of crystalline materials does not simplify the characterisation of anharmonicity, neither experimentally nor theoretically, especially in systems that contain entangled intramolecular and intermolecular interactions.\textsuperscript{33,34} Low-frequency terahertz time-domain spectroscopy (THz-TDS) is an experimental vibrational technique that is capable of probing both the intramolecular and intermolecular forces simultaneously, but decoupling the individual atomic contributions is not feasible without additional information, typically obtained using theoretical simulations.\textsuperscript{11,35–37} Density functional theory (DFT) calculations can yield valuable
insight regarding vibrational modes by providing normal mode vectors and force constants, but such calculations are typically performed within the harmonic approximation and without the inclusion of nuclear motion (the zero-Kelvin limit), neglecting anharmonicity and its related thermal effects entirely.\textsuperscript{38–40} Conversely, molecular dynamics (MD) simulations can explore the entire potential energy surface without harmonic restraints while taking temperature into account, but they do not explicitly determine the vibrational transitions nor mode-types, limiting their overall utility for interpreting experimental vibrational data.\textsuperscript{41–43}

It is clear that while none of these three methods are capable of completely describing anharmonicity entirely they each offer a unique perspective regarding its origins, and by combining all three a cohesive picture surrounding the vibrational dynamics of materials can be achieved. Here, this multi-faceted approach combining experimental THz-TDS with theoretical solid-state DFT and \textit{ab initio} MD (AIMD) simulations is applied in order to uncover the fundamental origins of anharmonicity in crystalline purine, ultimately contributing to the overall understanding of anharmonic effects in related crystalline and biomolecular systems.

**Methods**

**Terahertz Time-Domain Spectroscopy**

Samples were prepared to THz-TDS experiments by first mixing as-received purine (98%, Sigma-Aldrich, Poole, UK) with high-density polyethylene (HDPE) to an approximate 10% w/w concentration and subsequently grinding the mixture with a mortar and pestle in order to homogenise the sample and reduce particle size. The finely ground sample-HDPE mixture (\(\sim 300\) mg) was then placed into a 13 mm hydraulic press die and pressed under 2 tons of force, yielding 3 mm thick pellets of 13 mm diameter. A corresponding pellet containing only HDPE was used as a blank reference for all measurements.

The THz-TDS measurements were performed at 300 K and 80 K using a commercial
Figure 1: Unit cell structure of purine.\textsuperscript{32} An example of the hydrogen bonding pattern is shown for one chain found in the bulk
Terapulse 4000 spectrometer (TeraView Ltd., Cambridge, UK) equipped with a liquid nitrogen cryostat, providing a usable bandwidth of $0.5 - 2.75 \text{ THz}$. The terahertz spectrometer was housed in a continuously-purged nitrogen chamber to reduce absorption of atmospheric water, which is a strong absorber of terahertz radiation. The sample pellets were placed into the liquid nitrogen cryostat and held under vacuum for the duration of the experiment. At both temperatures 2000 terahertz time-domain waveforms were acquired and averaged for both the sample and blank pellets, and were subsequently Fourier transformed yielding frequency-domain transmission spectra. The sample spectra were then divided by the HDPE blank spectra, producing the presented terahertz absorption spectra.

**Solid-State Density Functional Theory**

Solid-state DFT simulations were performed with a developmental version of the CRYSTAL14 software package, which incorporates periodic boundary conditions into the model to reproduce the translational symmetry of the crystal. The Perdew-Burke-Ernzerhof (PBE) density functional, combined with the D3 dispersion correction, was coupled with the double-$\zeta$ def2-SVP basis set for all calculations. Initially, the atomic positions and lattice parameters (taken from the experimental single-crystal structure) were fully optimised with no constraints other than the space group symmetry of the solid, which has proven to produce excellent results for many organic and inorganic crystals. In the case of the simulated room temperature structure, the system was constrained to maintain the room-temperature volume, and all atomic positions and lattice vectors were allowed to relax within the that constraint. Upon complete optimisation, a vibrational analysis was performed by calculation of the second-derivative of the potential energy surface via a numerical finite difference scheme. Vibrational normal mode eigenvectors and eigenvalues were calculated within the harmonic approximation through diagonalisation of the mass-weighted Hessian (force constant) matrix, and IR intensities were calculated via the Berry Phase method.
**ab initio** Molecular Dynamics

AIMD simulations were performed using the CP2K software package,\(^5\,^2\,^5^3\) again incorporating periodic boundary conditions. In order to be comparable to the solid-state DFT calculations using CRystal14 the parameters for the AIMD simulations were kept as similar as possible. The PBE density functional,\(^4\,^5\) including the D3 correction by Grimme,\(^4\,^6\,^4\,^7\) was coupled with the double-\(\zeta\) DZVP basis set\(^5\,^4\) and the Goedecker-Teter-Hutter (GTH) pseudopotentials.\(^5\,^5\) Prior to performing the AIMD simulations, the atomic positions were fully optimised using the CP2K static-DFT engine, as well as the lattice constants in the case of the low-temperature structure. While the full optimisation of both atomic positions and lattice constants generates an effective 0 K structure, it is reasonable to use this structure to use for low-temperature (80 K) AIMD simulations. Following optimisation, the AIMD trajectories were calculated following a similar procedure used by Thomas *et al.*\(^5\,^6\) A 0.5 fs timestep was used for all simulations, and the chosen temperature maintained using the Nosé-Hoover chain thermostat.\(^5\,^7\,^5\,^8\) A 30 ps trajectory was used for determining the the vibrational spectra, which was taken following an equilibrium period of 3 ps (determined by monitoring the total energy and temperature stability of the system). In order to calculate the IR intensities the molecular dipole moments were required, which were determined every 2.5 fs through a localisation scheme using maximally localised Wannier functions and generated using the following relation

\[
\mu = -2e \sum r_i + e \sum Z_j R_j
\]

where \(\mu\) is the molecular dipole moment, \(r_i\) is the Wannier function centre, and \(Z_j\) and \(R_j\) are the nuclear charge and position, respectively.

The IR spectra were produced from the AIMD trajectories through Fourier transformation of the time-dependant dipole moment autocorrelation function, as previously described.\(^5\,^6\)
\[ A(\omega) \propto \left\langle \dot{\mu}(\tau)\dot{\mu}(t+\tau) \right\rangle_\tau \exp^{-i\omega t} \, dt \]

where \( A \) is the absorption cross section, \( \omega \) is frequency, \( t \) is time, and \( \dot{\mu} \) is the time derivative of the dipole moment leading to the dipole-velocity autocorrelation function. The post-processing of the AIMD trajectories and generation of the IR spectra was performed using the TRAVIS program.\(^{56,59,60}\)

**Results and Discussion**

**Terahertz Time-Domain Spectroscopy**

THz-TDS is a powerful technique for sensing the physical manifestation of potential energy anharmonicity because the low-frequency (0.3 - 3.0 THz) vibrations that can be measured are comprised of both intramolecular and intermolecular motions, unlike in mid-infrared experiments that can only directly probe intramolecular bonds.\(^{61-63}\) Furthermore, terahertz transitions have energies that are well below the thermal frequency at ambient conditions \( (\nu_{300K} \approx 6.25 \text{ THz}) \), meaning that vibrational excited states, which are very sensitive to anharmonicity, are sufficiently populated.\(^{64,65}\) This implies that the vibrational modes that are highly anharmonic will show a pronounced thermal dependence in their peak position and lineshape. This effect is illustrated in the terahertz spectra of purine which is presented in Figure 2.

The 300 K spectrum of purine contains a number of broad resonances that overlap with one another, while the 80 K spectrum has significantly sharper features that are more well-defined. Although it is relatively common for the terahertz spectra of organic crystals to blue-shift and sharpen upon cooling,\(^{66-68}\) the thermal effects observed here are much more pronounced than in most other systems,\(^{69,70}\) with an average blue shift of 0.10 THz and an average change in full-width half-maxima (FWHM) of 0.12 THz for \( \Delta T = 220 \text{ K} \). Interest-
Figure 2: Experimental THz-TDS data for crystalline purine taken at 300 K (red) and 80 K (blue).
ingly, the spectral features at 0.74 THz and 1.46 THz (at 80 K) do not shift significantly \( \Delta \nu = 0.01 \) THz in both cases), and when those modes are not included in the analysis the average shift between the 300 K and 80 K spectra is even larger \( \Delta \nu = 0.14 \) THz, or \( \sim 9\% \).

**Density Functional Theory Vibrational Analyses**

To explore the atomic origins of the observed vibrational properties, solid-state DFT simulations were performed with the CRYSTAL14 software package. The results of the DFT simulations are shown in Figure 3, and it is clear that the calculation significantly over-estimates the frequencies when compared to the 80 K experimental spectrum (more representative of the effective 0 K simulation), a common consequence of the harmonic approximation. Therefore a frequency position scalar of 0.85 was applied in order to come to a good agreement with the experimental data, while IR intensities are not scaled at all. The intensity of the calculated feature at 2.384 THz at first appears to be highly overestimated, but comparison with the 4 K spectrum previously reported by Shen et al. shows that between 80 K and 4 K the experimental intensity almost doubles, with the 4 K intensity in good relative agreement with the theory. That being said, the intensities are slightly over-estimated in some cases, which is possibly caused by the harmonic constraint of the vibrational eigenvectors. It is important to note that the effect of basis set superposition error was found to be minimal on the vibrational calculation, with a simulation using the larger triple-\( \zeta \) 6-311G(2d,2p) basis set still requiring a scalar of 0.87 to come into agreement with experiment (see Supporting Information).

The application of a frequency scalar yields a comparable spectrum to experiment, but upon closer inspection it is found to move the lowest absorption feature (experimentally at 0.74 THz) well out of agreement, indicating that particular mode is well-modelled by the harmonic approximation. This helps to confirm the hypothesis that the lack of any significant shifting for that feature is a result of its harmonic nature. Such a disparity is meaningful, since the atomic displacement vectors produced by the DFT calculations can be
Figure 3: DFT predicted vibrational spectra for both the 80 K and 300 K (inset) measurements, with lineshapes convolved using Lorentzians with linewidths corresponding to the average FWHM of the experimental spectra (99 and 36 GHz for the 300 K and 80 K spectra, respectively). The top panel shows the unscaled DFT spectra, while the bottom panel is a result of scaling the frequencies by 0.85 and in the case of the 300 K spectrum an intensity scalar of 0.5.
analysed to investigate if any correlation exists between the observed spectral changes and the vibrational mode-types (and participating atoms).

All of the predicted vibrations of purine below 3 THz are rotations of entire molecules, with the exception of a translational mode at 2.384 THz (full mode assignment provided in the Supporting Information). Visualisation of the vibrational normal modes highlights that some motions involve very little overall displacement of the amine-hydrogen bond, such as the transition at 0.557 THz which is a rotation about that coordinate, while others are very significant, such as the antisymmetric rotation of two hydrogen bonded purines in the 1.413 THz vibration. Comparing the individual shift in peak positions between the 300 K and 80 K spectra with the amount of distortion to the amine \( N - H \) and amine hydrogen bonds results in a very strong correlation, with greater distortion related to the amount of observed spectral shifting.

Because the interatomic potentials in chemical bonds are relatively stiff it is uncommon for temperature changes to have a large effect on internal molecular structure. Contrastingly, intermolecular potentials are much softer, meaning that in addition to low-frequency vibrational anharmonicity there can also be temperature-dependant structural changes that must be accounted for when interpreting experimental data. In order to probe if the differences in the THz-TDS spectra are a result of structural anisotropy (different structures at 300 K than 80 K), the room-temperature structure was optimised with the lattice parameters fixed to their experimental values, followed by a vibrational calculation. The results, shown in Figure 3, again require a positional scalar of 0.85, but now the intensities must also be scaled by a factor of 0.5 as well. This is not unexpected, as the integrated intensity of the experimental 300 K spectrum is half that of the 80 K spectrum. This is most likely caused by increased vibrational motion in transitions from thermally populated excited states, which are expected to have larger amplitude displacements that in turn would result in reduced IR intensities given the proportionality of \( I \propto \frac{\delta \mu}{\delta Q} \) (where \( \mu \) is the dipole moment and \( Q \) is the vibrational displacement vector). This can be illustrated by calculated the potential energy
Table 1: Experimental and scaled predicted (×0.85) frequencies and corresponding temperature shifts (THz), the change in amine $N-H$ bond length, amine hydrogen bond length (pm), and the change in the amine hydrogen bond angle (degrees) after displacement along each normal mode eigenvector.

<table>
<thead>
<tr>
<th>Experiment</th>
<th>DFT</th>
<th>Bond Changes After Eigenvector Displacement</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\nu_{300K}$</td>
<td>$\nu_{80K}$</td>
<td>$\Delta \nu$</td>
</tr>
<tr>
<td>0.73</td>
<td>0.74</td>
<td>0.01</td>
</tr>
<tr>
<td>-</td>
<td>1.13</td>
<td>-</td>
</tr>
<tr>
<td>1.24</td>
<td>1.41</td>
<td>0.17</td>
</tr>
<tr>
<td>1.45</td>
<td>1.46</td>
<td>0.01</td>
</tr>
<tr>
<td>1.56</td>
<td>1.68</td>
<td>0.12</td>
</tr>
<tr>
<td>1.77</td>
<td>1.90</td>
<td>0.13</td>
</tr>
<tr>
<td>2.21</td>
<td>2.34</td>
<td>0.13</td>
</tr>
</tbody>
</table>

curve of the terahertz normal modes, which show that at 300 K ($E = k_BT$) the potentials are almost exactly twice as wide as they are in the ground state ($E = \frac{1}{2} \hbar \omega$). This broadening trend in turn leads to a more spatially distributed probability density, ultimately yielding increased displacement that would be manifested in a lower terahertz absorption intensity.

The differences between the DFT and experimental frequency positions, as well as the deviation from equilibrium after displacement along the normal mode eigenvector for the amine $N-H$ bond, amine hydrogen bond, and hydrogen bond angle are provided in Table 1. The calculated frequency differences between the two absorptions that showed the smallest shifts experimentally (0.74 and 1.46 THz) are in good agreement, further indicating that they are more harmonic than the other modes. Furthermore, for the experimental feature with the largest shift (1.41 THz) the DFT predictions are also in good agreement, indicating that this particular mode is strongly dependant on the intermolecular structural anharmonic potential rather than intrinsic vibrational anharmonicity. However for the remaining modes the calculated frequency differences make up no more than half of what is observed, indicating that vibrational anharmonicity plays a significant role in the low-frequency vibrations.
**ab initio** Molecular Dynamics Vibrational Analyses

The assignment and interpretation of the experimental THz-TDS spectra using static DFT calculations suggests that a combination of anharmonic factors contribute to the overall vibrational properties of bulk purine. Even though such a conclusion is valuable, additional analysis is required to fully understand the interactions. Classical MD simulations have had an overwhelming amount of success in describing anharmonic and related temperature effects in a plethora of systems, but since they require the explicit input of bond force constants any vibrational analyses will be biased towards those parameters.\textsuperscript{77–79} To avoid this, fully \textit{ab initio} MD simulations, a method that updates the forces via a first-principles DFT calculation at every step of motion, have become a very powerful tool for determining material properties that are highly sensitive on the accurate description of electronic structure. Recently, AIMD has been used to simulate the mid-IR and Raman spectra of small organic molecules \textit{via} the Fourier transform of the velocity autocorrelation function, yielding highly accurate vibrational spectra, inclusive of anharmonic effects.\textsuperscript{56,60,80,81}

AIMD simulations were performed at 300 K and 80 K using CP2K, again incorporating periodic boundary conditions. In order to be comparable to the solid-state DFT calculations the parameters for the AIMD simulations were kept as similar as possible. The results of the AIMD simulations are presented in Figure 4 and are in excellent agreement with the experimental spectra. It is important to note that unlike in the solid-state DFT vibrational analysis the AIMD spectra have not been scaled in any way, and accurately reproduce experimental line-widths, intensities, and frequency positions. Moreover, the inclusion of anharmonicity into the model properly accounts for the internal vibrational anharmonicity, allowing the exploration of thermal effects in addition to those caused by structural changes. For example, it is evident that the minor feature at 0.74 THz is clearly unaffected by either structural or temperature change, which is why it is one of the only features that is well modelled by DFT prior to scaling.

The interplay between structural and intrinsic anharmonicity is an important compo-
Figure 4: Experimental 300 K (top, red) and 80 K (bottom, blue) THz-TDS spectra of purine and corresponding simulated AIMD spectra (black). The dotted curves represent the calculations performed on the reciprocal structures, with the top showing the effective 0 K structure simulated at 300 K, and the bottom showing the 300 K structure simulated at 80 K.
nent to consider. To evaluate this further, AIMD calculations were performed on the 300 and effective 0 K structures as before, except that the temperatures of the simulations were inverted, i.e. the cryogenic structure was evaluated at 300 K and vice versa. Overall, the spectra show a similar behaviour to experiment, with the reciprocal 80 K simulation exhibiting sharper features than when performed at 300 K. Of significant interest is the spectral feature occurring at 2.34 THz in the 80 K experimental spectrum, since the solid-state DFT simulations were not able to fully reproduce its observed blue-shifting. The inverted AIMD calculations highlight that for this particular mode the vibrational potential energy surface is highly anharmonic, since the simulations of both the 300 and low-temperature structures at a temperature of 80 K predict the same energy for that transition, while it shifts significantly in the 300 K simulations ($\Delta \nu \approx 0.16$ THz).

Conclusion

By combining static DFT calculations with AIMD simulations it is possible to interpret the observed low-frequency vibrational dynamics in crystalline purine. The results confirm the long-held belief that vibrational relaxation in systems containing amine hydrogen bonds has its origin in the anharmonicity of that coordinate, and that the covalent bond and intermolecular potentials are both key to understanding the experimental results. The analysis of the THz-TDS data and DFT-predicted normal modes highlights that the intermolecular forces are strongly anisotropic, and external motions (which are thermally active at ambient temperatures) must be taken into consideration when characterising anharmonic-related properties. Ultimately, the combination of these three techniques have led to a more in-depth understanding of the anharmonicity in purine, paving the way for application to more complex crystalline and biomolecular systems.
Acknowledgement

The authors thank the UK Engineering and Physical Sciences Research Council for funding (EP/N022769/1). Additional data related to this publication are available at the University of Cambridge data repository (https://www.repository.cam.ac.uk/handle/XXX).

Supporting Information Available

Experimental methods, calculated spectrum of purine using the 6-311g(2d,2p) basis set, and vibrational mode assignment details. This material is available free of charge via the Internet at http://pubs.acs.org.

This material is available free of charge via the Internet at http://pubs.acs.org/.

References


(15) Pecourt, J. M. L.; Peon, J.; Kohler, B. Ultrafast Internal Conversion of Electronically


Graphical TOC Entry