PCCP

View Online
Dynamic Article Links

Cite this: *Phys. Chem. Chem. Phys.*, 2011, **13**, 11719–11730

www.rsc.org/pccp

PAPER

Terahertz spectroscopy of enantiopure and racemic polycrystalline valine†

Michael R. C. Williams, Alan B. True, Artur F. Izmaylov, Timothy A. French,‡ Konstanze Schroeck and Charles A. Schmuttenmaer*

Received 2nd March 2011, Accepted 28th April 2011
DOI: 10.1039/c1cp20594c

Experimental and computational THz (or far-infrared) spectra of polycrystalline valine samples are reported. The experimental spectra have been measured using THz time-domain spectroscopy. Spectra of the pure enantiomers, both D and L, as well as the DL racemate have been taken at room temperature and low temperature (78 K). The spectra of the pure D and L enantiomers are essentially identical, and they are markedly different from the DL racemate. In addition, a temperature-dependent study of L-valine was undertaken in which the absorption maxima were found to red shift as a function of increasing temperature. The vibrational absorption spectra (frequencies and intensities) were calculated using the harmonic approximation with the Perdew–Burke–Ernzerhof (PBE) functional, localized atomic orbital basis sets, and periodic boundary conditions. The calculated and experimental spectra are in good qualitative agreement. A general method of quantifying the degree to which a calculated mode is intermolecular *versus* intramolecular is demonstrated, with the intermolecular motions further separated into translational *versus* rotational/librational motion. This allows straightforward comparison of spectra calculated using different basis sets or other constraints.

1. Introduction

Terahertz (THz) spectroscopy has become an increasingly popular method to study organic molecular crystals such as amino acid crystals,[1–3] pharmaceuticals,[4] explosives,[5] and even macromolecules of biological interest.[3,6–9] Large-scale motions of biomolecules and proteins occur in the frequency range of 0.1–3 THz (3–100 cm^{-1}), and there has been much effort to learn more about solid state librations and protein dynamics such as folding and conformational changes.[10–12] There has been some success measuring infrared modes associated with secondary protein structures, but many of the modes predicted in the terahertz range have yet to be positively identified.[13–17] This is due to the fact that vibrational spectra of proteins, even when crystallized, are dominated by broad absorption backgrounds rather than sharp features. Fortunately, there is much to be learned from more tractable systems, such as organic molecular crystals.

Intermolecular interactions dominate many aspects of biology: examples include DNA base pairing, secondary and tertiary protein structures, and interactions between proteins. Several groups have carried out THz studies of DNA ranging from its hydration and conformation to the characteristics of the various nucleosides and bases.[6,7,9] Whitmire and coworkers found that the THz absorption for wild-type bacteriorhodopsin is dependent upon the conformation of the protein, yet did not resolve any vibrational modes.[11] Kutteruf *et al.* showed that different sequences of amino acids in di- and tripeptides have independent THz spectra unrelated to the solid-state monomer absorption.[6]

There have also been many studies of the far-IR (or THz) spectra of amino acid crystals in the last 30 years.[18–21] More recently, Yamaguchi,[2] *et al.* reported the differing THz spectra of enantiopure and racemic polycrystalline alanine and King,[22] *et al.* reported the different THz spectra of L- and DL-serine as well as density functional theory (DFT) calculations on those systems. Here, the same type of behavior is observed for valine samples.

In addition to THz or far-IR spectroscopy, Raman scattering and inelastic neutron scattering (INS) methods have been employed in the study of amino acid crystals, such as alanine.[23] Raman scattering, which provides information derived from

Yale University, Department of Chemistry, PO Box 208107, 225 Prospect St., New Haven, CT 06520-8107, USA. E-mail: charles.schmuttenmaer@yale.edu

† Electronic supplementary information (ESI) available: (1) Table S1: unit cell parameters for all instances of L- and DL-valine (without any salts or co-solvents) in the CSD. (2) Powder XRD spectra for L-valine and DL-valine. (3) Plots of the quantified mode character for all systems studied. (4) Movie files of vibrations. See DOI: 10.1039/c1cp20594c

‡ Current address: Department of Chemistry and Chemical Biology, Harvard University, 1 Oxford Street, Cambridge, MA 02138, USA.

variations in polarizability as a function of atomic displacements, is a valuable complement to THz spectroscopy, which provides a direct measurement of the coupling between incident electromagnetic radiation and the oscillating dipole moments of collective atomic motions. INS (unlike THz and Raman spectroscopy) can be used to obtain full phonon dispersion curves rather than being limited to the measurement of modes for which $\boldsymbol{k}$ is zero. Unfortunately, INS requires a substantial, dedicated infrastructure that is not always available. In addition, while it is common practice in INS experiments on molecular crystals to use perdeuterated samples, there are examples of cases where the temperature dependence of the dynamics[24] and even the structure[25] of the perdeuterated material varies substantially from the naturally occurring isotopologue.

Amino acid crystals, and organic molecular crystals in general, can be studied computationally at several levels. *Ab initio* methods using full periodic boundary conditions (PBC) are ideal, and should be used whenever possible. There are a few examples using the Car–Parrinello Molecular Dynamics (CPMD) approach which carries out molecular dynamics simulations with density functional methods.[1,26] Calculations that have very good agreement with experimental THz spectra of organic molecular crystals have been performed using density functional theory (with periodic boundary conditions) as implemented in the Dmol3 package[5] as well as the CASTEP program.[27] In both cases, high cut-off energies (up to 1200 eV) for the plane wave basis used were necessary, with a correspondingly high computational cost. Generally, plane wave bases are best suited for metals, and one needs relatively large cut-off energies for semiconductors and insulators in order to describe localized electrons. Since crystalline amino acids are insulators with band gaps of 4–5 eV, we employed the GAUSSIAN[28] program in the current work, which uses localized Gaussian atomic orbitals as basis sets. Other DFT software that uses an atom-centered approach is available, such as the well-known SIESTA project. Unlike many of these packages (which use pseudopotential methods), the method employed here using GAUSSIAN is an all-electron approach. It is also possible to gain a certain amount of insight utilizing empirical methods as employed in CHARMM,[29] however, one should keep in mind that parameterizations available in this program were created to reproduce protein properties in solution or gas phase, rather than in crystals.

The crystallographic atomic coordinates for all three forms of valine (D, L, and DL) are found in the Cambridge Structural Database (CSD);[30] see Table S1 in the ESI.† This offers a chance to compare the experimental and calculated spectra, both frequencies and intensities, using the known atomic coordinates as a starting point in the calculations. In addition, the measured powder X-ray diffraction (XRD) spectra can be matched to those calculated from the known coordinates to verify that the experimental samples have the same morphologies as used in the calculations.

Fig. 1 presents the molecular structure of valine, which is zwitterionic in its crystalline state. Valine is a small hydrophobic amino acid, having an isopropyl side chain. Fig. 2 shows the crystal structures of DL-valine and L-valine.

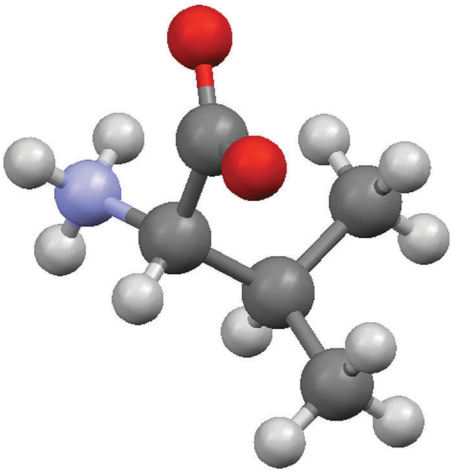

Fig. 1 Molecular structure of the zwitterionic form of L-valine. Hydrogen is white, carbon is gray, nitrogen is blue, and oxygen is red.

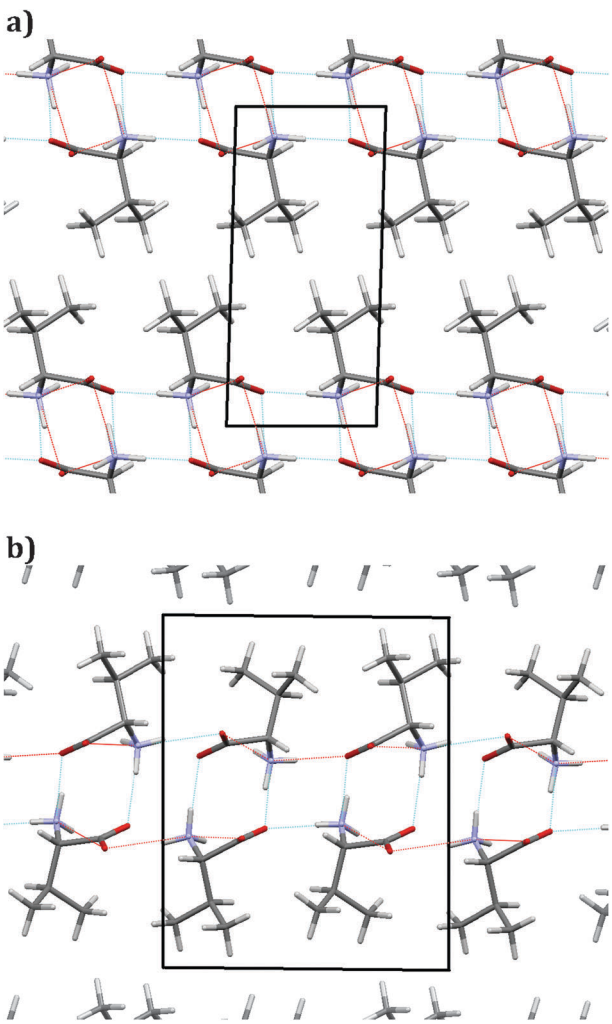

Fig. 2 Part (a): DL-valine crystal structure looking down the *a*-axis. Part (b): L-valine crystal structure looking down the *b*-axis. In both parts, hydrogen-bonding interactions are shown with dotted lines.

2. Methods

2.1 Computational details

All calculations for this work were carried out using the Perdew–Burke–Ernzerhof (PBE) density functional[31] within the development version of the Gaussian program. It is only recently that analytical frequencies and IR intensities of periodic systems have been implemented in Gaussian.[32,33]

The space group, atomic coordinates, and lattice parameters for the D, L, and DL crystals of valine were obtained from the Cambridge Structural Database (CSD),[30] and used as the starting configuration for subsequent geometry optimizations. The powder X-ray spectra of CSD entries were compared with experimental powder X-ray spectra of our material to verify polymorph identities. In the case of multiple CSD entries for the same polymorph, the lowest temperature coordinates and lattice parameters were preferred due to the fact that the computational methods employed include no correction for non-zero temperature. The specific entries in the CSD that were used are LVALIN01[34] for L-valine, and VALIDL03[35] for DL-valine, and were chosen because their calculated powder patterns are in agreement with our measured ones.

All calculations were initially carried out using the 6-31G** basis. The basis set dependence of the calculations is analyzed in detail below. The number of $\boldsymbol{k}$-points included in each system was determined automatically by Gaussian on the basis of estimated band gap and the size of the unit cell. For L-valine, the system with the largest unit cell, 120 $\boldsymbol{k}$-points were included, while the DL-valine calculation used 300 $\boldsymbol{k}$-points. Geometry optimizations were performed using internal coordinates with the convergence criterion of 2.0×10^{-6} Hartree per Bohr (approximately 1×10^{-4} eV Å^{-1}) in maximum force. The unit cell vectors were optimized along with the atomic positions.

Although fully analytic procedures for harmonic frequencies and IR intensities are available, in order to take advantage of parallel performance we have employed a semi-analytical procedure, which numerically differentiates analytical forces and dipoles. The displacements (±0.001 Å) for each atom from its equilibrium position in each Cartesian direction were calculated in parallel (6 total displacements per atom). Also, note that in the current work, the proper definition of periodic dipole moment has been used.[36] This definition, compared to a simple generalization of an isolated molecule's dipole, involves a Berry phase contribution that is absolutely crucial in order for a calculated periodic dipole to reproduce the result from the corresponding oligomeric limit.

2.2 Quantifying intermolecular *vs.* intramolecular motion

In order to quantify the fractional intermolecular and intramolecular contributions to each normal mode, we have employed a technique that is more rigorous than that used by Jepsen and Clark[27] and also allows the intermolecular contribution to be further broken down into translational and rotational motions. It is not limited to any particular computational method; it only requires access to the vibrational eigenvectors of the system of interest.

The equilibrium position of each atom is $r_{0,i}$, and that for the distorted molecule is $r_{1,i}$, where $r_{1,i}$ is obtained from $r_{1,i} = r_{0,i} + \delta_i$, where δ_i is the displacement vector of that atom for the normal mode in question. The values of δ_i are scaled such that the atom that undergoes the maximum displacement moves by 0.01 Å. The total root-mean-squared (rms) mass-weighted atomic displacement for each molecule in the unit cell is calculated using

$$\mathrm{Total}_{\mathrm{rms}} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} m_i \delta_i^2} \qquad (1)$$

and their average is taken. N is the number of atoms in each molecule and m_i is the mass of each atom. The contribution of the translational motion to the total displacement is obtained from

$$\mathrm{Trans}_{\mathrm{rms}} = \mathrm{Total}_{\mathrm{rms}} - \sqrt{\frac{1}{N}\sum_{i=1}^{N} m_i (r_{0,\mathrm{com},i} - r_{1,\mathrm{com},i})^2} \qquad (2)$$

where $r_{0,\mathrm{com},i}$ and $r_{1,\mathrm{com},i}$ are the positions of each atom in the equilibrium and distorted configurations, respectively, using the molecular center of mass as the origin. This amounts to translating the distorted configuration to have the same center of mass as the equilibrium configuration, and calculating the rms mass-weighted differences in atomic positions.

The intramolecular contribution to each mode is obtained by minimizing the difference in the distorted structure relative to the equilibrium structure. This is achieved by minimizing the quantity[37,38]

$$\sum_{i=1}^{N} m_i (r_{0,\mathrm{com},i} - \boldsymbol{U} r_{1,\mathrm{com},i})^2, \qquad (3)$$

where $\boldsymbol{U}$ is the rotation matrix given by[39]

$$\boldsymbol{U} = \begin{pmatrix} \cos\phi\cos\psi - \sin\phi\cos\theta\sin\psi & \sin\phi\cos\psi + \cos\phi\cos\theta\sin\psi & \sin\theta\sin\psi \\ -\cos\phi\sin\psi - \sin\phi\cos\theta\cos\psi & -\sin\phi\sin\psi + \cos\phi\cos\theta\cos\psi & \sin\theta\cos\psi \\ \sin\phi\sin\theta & -\cos\phi\sin\theta & \cos\theta \end{pmatrix} \qquad (4)$$

where ϕ, θ, and ψ are the Euler angles. A nonlinear least squares routine (Levenberg–Marquardt steepest decent[40]) is employed to iteratively adjust the Euler angles until convergence is obtained.

At this point, the rms difference between the equilibrium structure and the translated/rotated distorted structure yields the intramolecular contribution to the mode:

$$\mathrm{Intra}_{\mathrm{rms}} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} m_i (r_{0,\mathrm{com},i} - \boldsymbol{U}_{min} r_{1,\mathrm{com},i})^2}, \qquad (5)$$

where $\boldsymbol{U}_{\mathrm{min}}$ is the rotation matrix when using the optimized angles ϕ_{min}, θ_{min}, and ψ_{min}.

In summary, the total, translational, and intramolecular rms mass-weighted displacements are calculated directly.

From these, the intermolecular and rotational contributions are determined using $\text{Inter}_{rms} = \text{Total}_{rms} - \text{Intra}_{rms}$, and $\text{Rot}_{rms} = \text{Inter}_{rms} - \text{Trans}_{rms}$.

2.3 Experimental procedures

The D- and L-valine were purchased from Alfa-Aesar, and the DL-valine was purchased from Fluka. The D- and L-valine were used without further purification; the DL-valine was recrystallized by slow evaporation from dilute solution (approximately 1 g per liter of water). Samples were pulverized into a fine, uniform powder using a ball mill to minimize Mie scattering and crystallite anisotropy. Approximately 100 mg of each powder was then pressed between nylon discs in a 13 mm diameter die. Free standing pellets with a thickness of 0.25–0.50 mm were obtained using a pressure of 3.5 kbar. The pellets were placed in a cryostat (Janis) with Mylar windows, and spectra were taken under vacuum at both 298 K and 78 K, as well as at intermediate temperatures for the D- and L-valine samples. The cryostat sample holder had a clear aperture adjacent to the amino acid pellet, and by translating the assembly a separate reference spectrum was obtained at each temperature for which data are reported, minimizing the effect of any drift of the system over time.

The THz apparatus used has been described elsewhere.[41] A Ti:Sapphire laser (KM Labs Griffin) with a repetition rate of 80 MHz provided 50 fs pulses with an energy of ~3 nJ at a wavelength of 800 nm. The beam was split, with 50% of the power producing THz radiation *via* a GaAs photoconductive switch modulated with a 30 kHz square wave bias (±10 V) to allow for detection using a lock-in amplifier. The remainder of the near-IR pulse was directed to the detector switch (low-temperature grown GaAs). The region of the spectrometer through which the THz beam propagated was purged with dry nitrogen to reduce the absorption by atmospheric water vapor. Data points were collected in the time domain over a 50 ps window, corresponding to a resolution of 20 GHz in the frequency domain.

Powder XRD spectra for the samples were measured with a Bruker-AXS D8 Focus diffractometer (1.54 Å Cu-Kα radiation, a step size of 0.010°, a step time of 2 s, and a detector slit width of 0.1 mm).

3. Results and discussion

3.1 Structural confirmation

Given the well-known issue of polymorphism in organic molecular crystals, we measured the powder XRD spectra in order to confirm that our polycrystalline samples are the same polymorph as reported in the CSD.[30] This is necessary for two reasons: (1) when calculating the THz spectra, we want to be certain that we are using the same structure on which we are making measurements, and (2) when different labs compare results, they should verify that they are working with the same polymorph, and this is a convenient way to do so. The THz spectra of different polymorphs of the same species can vary dramatically due to the intermolecular nature of the low frequency motions.[42]

Fig. 3 compares three powder XRD spectra for L-valine, but the other samples behave similarly (see the ESI†). In particular,

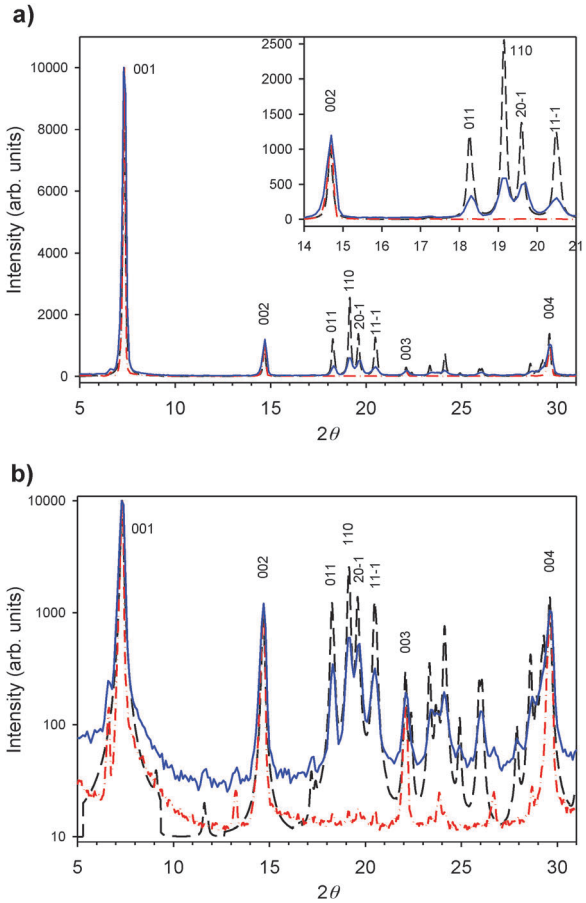

Fig. 3 Part (a) is X-ray powder diffraction spectra of L-valine. Dashed black line shows calculated spectrum, dash-dotted red line is for polycrystalline sample crushed with mortar and pestle, and solid blue line is after 2 minutes in the ball-mill. All spectra are normalized to 10 000. Part (b) is the same data as in part (a), except plotted on a semi-logarithmic scale. An offset of 10 has been added to each data set to keep range within 3 orders of magnitude.

the measured scattering angles are in excellent agreement with that calculated from the 3-d atomic coordinates, but only the intensities of (00*l*) peaks are in good agreement with the calculated intensities. L- and DL-valine form plate-like crystallites because of strong in-plane hydrogen bonding, yielding layers that are held together by van der Waals interactions. For L-valine, the hydrogen bonding network is in the ***a*–*b*** plane. Thus, there is preferential diffraction along the ***c***-axis. However, when the sample is pulverized in a ball mill for 2 minutes, the intensities of all the peaks are in much better agreement with the calculated ones (Fig. 3a, inset) because the sample is more isotropic. Fig. 3b shows the same data on a semi-logarithmic scale. This allows the peak positions to be compared even when the intensities are not in good agreement.

When possible, polymorphs were identified by comparing room temperature XRD spectra to room temperature crystal structures. In the case of DL-valine, of the three entries in the CSD for the triclinic polymorph, only the two low temperature crystal structures (VALIDL02 and VALIDL03) include atomic coordinates. The VALIDL01 entry is at room temperature,

although only the cell vectors are reported. The initial agreement of our room temperature powder XRD data with spectra calculated using data from either of the low temperature CSD entries was poor. However, after simply scaling the VALIDL03 unit cell (but not the atomic positions) to match that of VALIDL01 (a factor of 1.02), the calculated powder XRD pattern was in very good agreement with our data; this is illustrated in the ESI.†

3.2 Terahertz spectra and computational results

At room temperature, D- and L-valine display broad absorption across the 1.5–2.0 THz range, and an isolated peak at 2.25 THz (Fig. 4a and b). The broad feature is comprised of several peaks, at least three of which are well-resolved at 78 K. The racemic DL-valine sample has a single strong absorption feature at 1.70 THz. Unlike the D and L samples, the peak in DL-valine does not split into multiple peaks at low temperature, although it does sharpen (Fig. 4c).

Modeling vibrations in the THz spectral region is challenging because determining their inherently small force constants requires that the optimized structure be very highly converged. Residual forces between atoms that might have an insignificant effect on the calculation of mid-IR frequency vibrations can be large on the energy scale of the very low frequency modes of interest here. Furthermore, the energy separation between calculated THz frequency modes can be quite small (a few cm^{-1}, or tens of GHz). As a consequence, it is not advisable to make assignments solely on the basis of proximity to an observed spectral feature as well as the calculated and measured intensity unless the computational approach being used is accurate to this level of detail, especially if the calculated frequencies are harmonic ones, since it is well known that these low frequency modes are highly anharmonic. One should first analyze the computational model in greater depth to determine if a putative assignment is plausible at the given level of theory and/or other limitations. The dependence of calculated THz spectra on the choice of exchange and correlation functional has been reported,[5,43] as well as comparisons of THz spectra calculated by DFT and by classical models.[1] In this work we consider the effects on calculated THz spectra of two other computational factors: (1) basis set dependence and (2) the choice of either allowing the unit cell parameters of the crystal to be optimized along with the atomic positions or fixing these values to those obtained by crystallography.

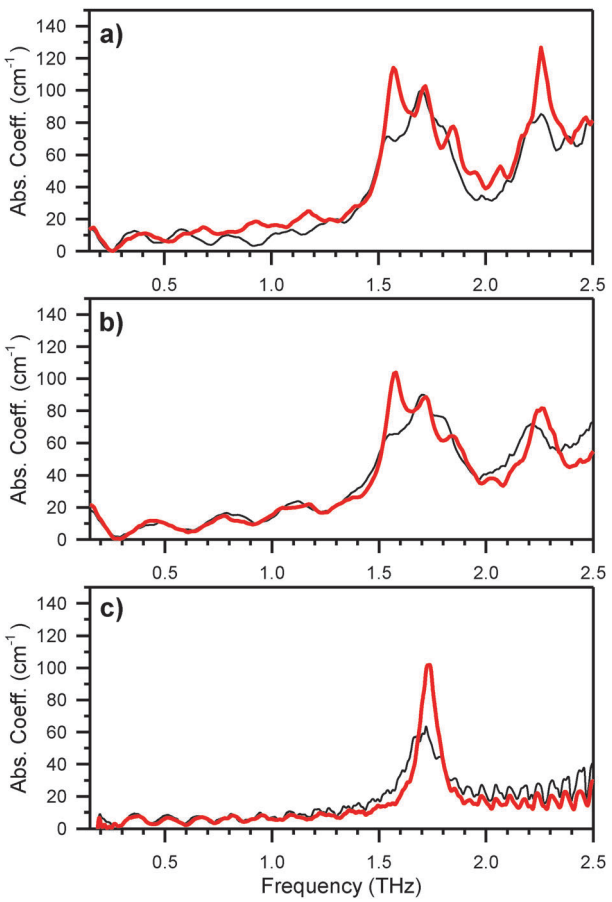

Fig. 4 The THz absorption spectra of polycrystalline D-valine (part a) and L-valine (part b) are identical, and they are very different from the DL-valine racemate (part c). Spectra acquired at room temperature (thin black lines) show a red shift in the absorption features relative to the spectra taken at 78 K (thick red lines) for all three species.

3.2.1 Basis set dependence. While the employment of a larger basis set can generally be expected to increase the accuracy of a model (within the limitations of the method, theory, or implementation thereof), it is not obvious how the calculated vibrational absorption spectra of complicated systems should vary as the number of basis functions increases. Whether the underlying modes themselves vary significantly in actual character or whether the variation is limited to calculated intensities and frequencies is an important question as well. For systems such as those being investigated, in which the intermolecular and intramolecular degrees of freedom are substantially mixed, these concerns are especially relevant.

The basis set dependence of the DL-valine calculations was investigated in terms of the calculated vibrational frequency and intensity, as well as the character of the mode's motion quantified using the method described above. The DL-valine system was selected for this exercise because it has half as many atoms per unit cell as crystalline L- and D-valine (two molecules *versus* four, with 19 atoms per molecule), and the computational cost is therefore substantially lower for calculations on DL-valine. It is reasonable to expect that the basis set dependence of L- and D-valine is similar to that of DL-valine (in terms of overall convergence), given the similarity in the constituents and intermolecular interactions of these systems. The DL-valine system's geometry was optimized from the original crystallographic coordinates for each basis set to ensure the independence of the different calculations.

The assessment of a computational model's convergence with respect to parameters such as frequency, intensity, or vibrational character is much more time consuming than evaluating the convergence based on single-point energy calculations. However, it is not clear how to correlate the degree of convergence of the energy at a particular geometry with an expected degree of convergence of the subsequently calculated vibrations.

When molecules have high symmetry, it is convenient to refer to a vibrational mode by its symmetry label, which also has the advantage of universality among researchers. However, when describing the low-frequency vibrations of solid-state systems composed of molecules of C_1 symmetry and further complicated by intermolecular coupling, it is more difficult to find a common vocabulary. Our approach of reporting quantitative summaries of mode character (such as in Fig. 5) is intended to improve the ease of comparing computational results on these systems. When used to compare the THz spectra of two similar calculations, the quantitative descriptions of the modes can provide additional insight beyond the notion that one calculated spectrum "agrees" better than another. Quantified mode character essentially provides a middle ground between exhaustive numerical tables and inevitably vague qualitative descriptions.

In Fig. 5 we report the calculated frequencies and intensities of vibrational modes over the range of 0.25 to 4.0 THz using four different basis sets as well as a quantitative description of the nature of the motion for the ten lowest frequency vibrations from each calculation. The smallest basis set considered was 3-21G. The second smallest basis set, 6-31G, differs from 3-21G in that the core and valence orbitals are composed of a larger number of primitive Gaussian functions. This increases the flexibility of electron distribution for each valence orbital in terms of spatial extent (but not shape). The 6-31G* basis set adds polarization functions to the valence orbitals of non-hydrogen atoms, the mixing of which with the atomic orbitals increases flexibility of electron distribution in terms of shape, *i.e.*, allows valence orbitals to be distorted relative to purely atom-like orbitals as a result of the chemical environment. Finally, the 6-31G** basis set differs from the 6-31G* basis set in that polarization functions are also added to hydrogen atoms.

Table 1 provides the frequencies of the first 12 modes calculated with each of the four basis sets. In addition, the mode number of the corresponding vibration in the 6-31G** calculation is reported, as well as the ratio of this frequency to the vibrational frequency of the corresponding mode in the 6-31G** calculation. Putative mode assignments made to spectral features are discussed in a separate section below. Note that the experimental THz spectrum of DL-valine in Fig. 5 is plotted over a larger range than in Fig. 4c. The strength of the absorption feature appearing at approximately 2.75 THz almost certainly exceeds the dynamic range of our spectrometer at that frequency. The dynamic range of this instrument, which is at a maximum between 1 and 2 THz, diminishes at higher frequencies.[44] Therefore, while the

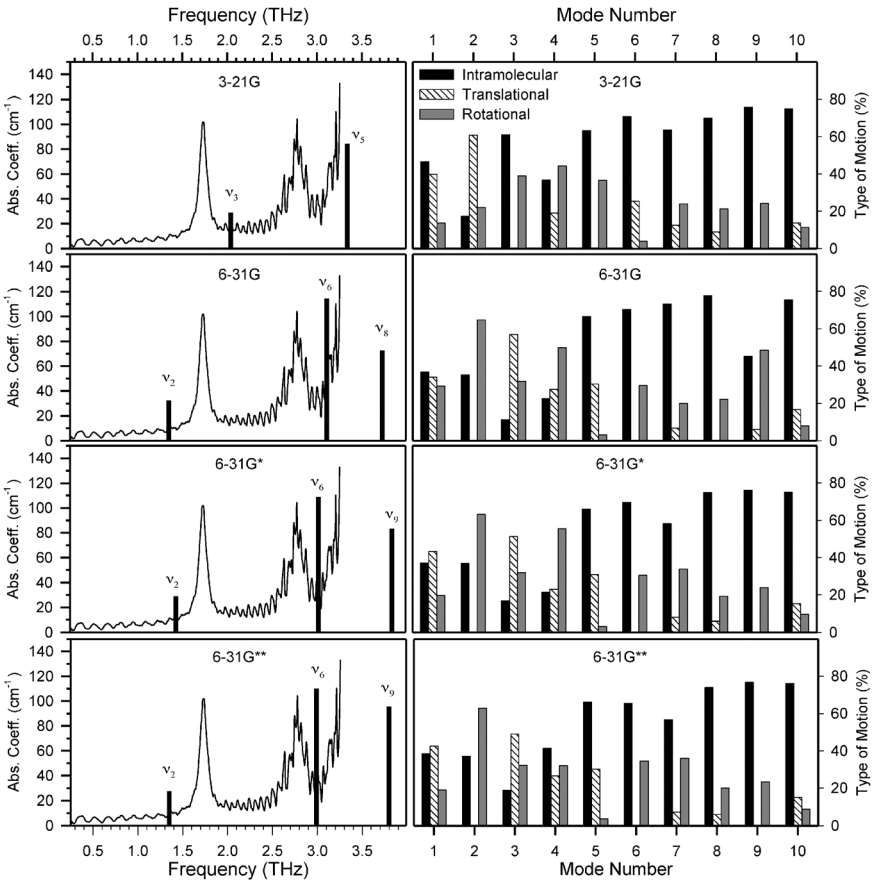

Fig. 5 The frequencies and intensities of infrared active DL-valine vibrational modes were calculated with four different basis sets (black stick spectra in the series of plots on the left). The experimental 78 K THz absorption spectrum is provided for comparison. The series of plots on the right show the quantified character of the vibrational motion for the ten lowest frequency modes from each calculation. Labels above the stick spectra refer to the mode number as calculated using that particular basis set. All of the infrared active modes below 4.0 THz happen to be a mix of librational (hindered rotational) motion with substantial intramolecular displacement.

Table 1 The frequencies of DL-valine vibrational modes calculated with four different basis sets are listed in the order (ascending frequency) that they appear in the particular calculation. Frequency values in bold type indicate that the calculated mode is IR active. For modes calculated using a basis set smaller than 6-31G**, the mode number of the corresponding vibration in the 6-31G** calculation is reported, as well as the fraction of its calculated frequency relative to the analogous 6-31G** mode. For example, the third lowest frequency mode in the 3-21G calculation is IR active, analogous to mode number two in the 6-31G** calculation, and occurs at a frequency 1.51 times higher (2.04 THz/1.35 THz)

	Mode frequency/THz				Analogous 6-31G** mode number			Fraction of analogous 6-31G** mode frequency		
Mode	3-21G	6-31G	6-31G*	6-31G**	3-21G	6-31G	6-31G*	3-21G	6-31G	6-31G*
1	1.45	1.09	1.09	1.07	1	1	1	1.36	1.02	1.02
2	1.97	**1.35**	**1.43**	**1.35**	3	2	2	1.29	1.00	1.06
3	**2.04**	1.62	1.53	1.53	2	3	3	1.51	1.06	1.00
4	2.76	2.20	2.02	1.98	4	4	4	1.39	1.11	1.02
5	**3.34**	2.76	2.70	2.66	6	5	5	1.12	1.04	1.02
6	3.51	**3.11**	**3.02**	**2.99**	5	6	6	1.32	1.04	1.01
7	4.00	3.57	3.55	3.52	7	8	7	1.14	0.97	1.01
8	4.32	**3.73**	3.73	3.67	8	9	8	1.18	0.98	1.02
9	**4.47**	3.82	**3.84**	**3.80**	9	7	9	1.16	1.09	1.01
10	4.99	4.56	4.58	4.44	10	10	10	1.12	1.03	1.03
11	**5.13**	**4.83**	**4.66**	**4.60**	11	11	11	1.12	1.05	1.01
12	6.43	5.77	5.77	5.59	12	12	12	1.15	1.03	1.03

measurement of this feature's position is included in Fig. 5 for comparison with the calculations, its intensity should not be considered accurate.

For the purpose of calculating THz absorption spectra, it is of interest to compare the effect of adding basis functions that increase only the flexibility of orbital extent to the effect of adding polarization functions. While plane-wave based computational approaches to modeling periodic systems have a systematic and smooth approach to basis set convergence (simply increasing the cut-off energy and thereby decreasing the wavelength of the plane waves), the basis set dependence of atomic orbital based models is more idiosyncratic and therefore worthy of investigation.

Two IR-active modes are calculated in the 0.25 to 4 THz region of the spectrum when using the 3-21G basis set (Fig. 5). The motion of both modes, $\nu_3^{(3\text{-}21G)}$ at 2.03 THz and $\nu_5^{(3\text{-}21G)}$ at 3.32 THz, are primarily intramolecular, with essentially all of the intermolecular components of the character being rotational in nature.

The results of the calculation carried out using the 6-31G basis set are also presented in Fig. 5. The lowest frequency calculated mode that is IR-active is $\nu_2^{(6\text{-}31G)}$ (1.34 THz), with $\nu_6^{(6\text{-}31G)}$ and $\nu_8^{(6\text{-}31G)}$ being the other two IR-active modes below 4.0 THz. As was the case for both IR-active modes in the 3-21G calculation, the three IR-active modes in the 6-31G calculation all have significant intramolecular character and all have no translational contribution to the intermolecular portion of the atomic displacements. In other words, the three modes calculated to be IR-active in the region below 4.0 THz are all librational modes with varying amounts of displacement occurring along intramolecular degrees of freedom.

Despite occurring at different frequencies and in a different order, the modes appearing in the set of low frequency vibrations using the 3-21G basis are largely conserved in the set of vibrational modes calculated using the larger 6-31G basis. For example, $\nu_5^{(3\text{-}21G)}$ is analogous to $\nu_6^{(6\text{-}31G)}$ in terms of the character of their motions. Likewise, $\nu_3^{(3\text{-}21G)}$ and $\nu_2^{(6\text{-}31G)}$ display roughly the same motion, although $\nu_3^{(3\text{-}21G)}$ is somewhat more intramolecular. The mode $\nu_9^{(3\text{-}21G)}$, which occurs at 4.32 THz, is similarly analogous to $\nu_8^{(6\text{-}31G)}$ (and is also IR-active). Even though a comparison of the quantified character of two modes is not sufficient in itself to determine if the vibrations are analogous or not, it is a very useful starting point and guide for inspecting animations of the modes. In addition, the quantified character can highlight important details that one might miss simply watching a movie of the vibrations. Animations of all modes discussed are available in the ESI.†

While increasing the size of the basis set from 3-21G to 6-31G results in an obvious change in the calculated THz spectrum, the addition of polarization functions to the second row elements in the system (*i.e.*, changing the basis set from 6-31G to 6-31G*) results in a much smaller change in the frequencies and intensities of IR-active vibrational modes, as seen in Fig. 5. The modes from the 6-31G* calculation largely conserve the character of those from the 6-31G calculation, and the three IR-active modes below 4.0 THz again all consist of intramolecular motions combined with semi-rigid torsion.

Again, there are some differences in the order of vibrational modes. For instance, the third-lowest frequency calculated IR-active mode is $\nu_9^{(6\text{-}31G^*)}$ as compared to $\nu_8^{(6\text{-}31G)}$ previously. This change in ordering is due less to a change in the character or frequency of this mode than changes in nearby IR-inactive modes. The modes that are most affected by the addition of polarization functions in the 6-31G* basis set are $\nu_1^{(6\text{-}31G^*)}$ and $\nu_7^{(6\text{-}31G^*)}$ (formerly $\nu_1^{(6\text{-}31G)}$ and $\nu_9^{(6\text{-}31G)}$, respectively).

Finally, there is very little change in the calculated vibrations of crystalline DL-valine when the basis set size is increased from 6-31G* to 6-31G** except for $\nu_4^{(6\text{-}31G^*)}$, which becomes more intramolecular (less of a rigid torsion) as $\nu_4^{(6\text{-}31G^{**})}$.

3.2.2 Fixed *versus* relaxed unit cell models. Not all quantum chemistry software includes the option of optimizing unit cell parameters in calculations of periodic systems. In addition, some researchers choose to use a unit cell fixed at crystallographic values because their computational model finds an optimized unit cell for the system that is very different than experimental values. This approach is not optimal and leads to results that are not self-consistent: although one may obtain the minimum energy configuration of the system with fixed

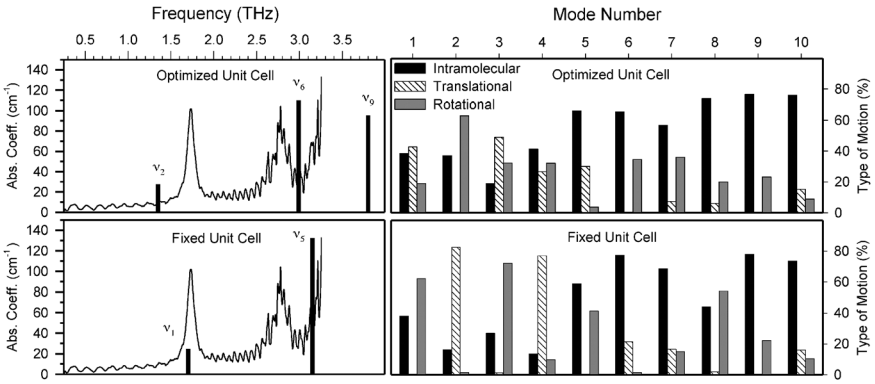

Fig. 6 The frequency, intensity, and quantified character of DL-valine vibrational modes obtained from a calculation in which the unit cell geometry was optimized as well as the atomic coordinates (our standard procedure) are compared to those calculated with the unit cell fixed at experimental crystallographic values. While the order of infrared active modes is different, the two calculated spectra are quite similar. However, the apparent similarity is largely due to the fact that the vibrational modes most affected by the implicit stress in the fixed cell calculation happen to be infrared inactive. That the calculations are in fact quite different is illustrated by the appearance of predominantly translational modes in the fixed cell calculation whose character is not found among the modes using the optimized cell calculation.

unit cell parameters, the optimized positions and subsequently calculated vibrations will not correspond to those of the system modeled at equilibrium. Instead, the calculated vibrations will be those of the molecular crystal under a complicated and indeterminate amount of stress. If one is willing to accept the fact that the calculated atomic positions within the unit cell differ from those measured experimentally (vibrational spectra are only calculated once the atomic positions have been optimized), then there should be no requirement to fix the unit cell parameters at experimental values. Furthermore, most calculations are implicitly carried out at 0 K, while diffraction data are always obtained at a higher temperature, sometimes much higher.

To illustrate this point, the vibrational frequencies and IR intensities of DL-valine were calculated with the 6-31G** basis keeping the unit cell fixed at crystallographic values but optimizing atomic positions, and then compared to the same 6-31G** level calculation in which the unit cell had been optimized as well as the atomic positions (the latter calculation has been discussed above). A comparison of these results is presented in Fig. 6 along with quantified descriptions of the motion of calculated low-frequency modes. Upon first impression, the THz absorption spectrum calculated using the fixed unit cell parameters appears to be very similar to the spectrum calculated using optimized unit cell geometry. In addition, the IR-active modes in both calculations are closely analogous in character: largely intramolecular with the entire intermolecular contribution coming from rotational motion. Inspection of the actual motion of these modes (animations available in the ESI†) confirms that $\nu_1^{(fix)}$ is almost identical to $\nu_2^{(opt)}$ and likewise for $\nu_5^{(fix)}$ and $\nu_6^{(opt)}$. The mode analogous to $\nu_9^{(opt)}$ also exists in the fixed cell calculation ($\nu_9^{(fix)}$) although it occurs at 4.37 THz instead of 3.80 THz. Some of the IR-inactive modes are also conserved between calculations: the motion of $\nu_6^{(fix)}$ corresponds closely with that of $\nu_5^{(opt)}$, as does $\nu_7^{(fix)}$ with $\nu_8^{(opt)}$, and $\nu_{10}^{(fix)}$ with $\nu_{10}^{(opt)}$. Table 2 lists the frequencies of all the modes calculated using the fixed cell parameters and identifies the corresponding mode (if any) in the optimized cell calculation.

Table 2 Comparison of the frequencies of DL-valine vibrations calculated with the unit cell parameters either fixed at crystallographic values or optimized along with atomic coordinates (both calculations carried out with 6-31G** basis sets). While modes with predominantly torsional intermolecular motion appear in both calculations, several modes with substantial translational character appear in the fixed cell calculation for which it is difficult to identify a corresponding vibration among the optimized cell modes

	Mode frequency/THz			
Mode	Optimized cell	Fixed cell	Analogous mode number in optimized cell	Fraction of analogous mode frequency
1	1.07	**1.70**	2	1.26
2	**1.35**	1.74	—	—
3	1.53	2.18	—	—
4	1.98	2.37	—	—
5	2.66	**3.15**	6	1.05
6	**2.99**	3.17	5	1.19
7	3.52	3.93	8	1.09
8	3.67	4.05	—	—
9	**3.80**	**4.37**	9	1.15
10	4.44	4.59	10	1.03
11	**4.60**	**4.88**	11	1.06
12	5.59	5.94	12	1.06

Despite the similarities in the vibrations discussed thus far, there are several modes in the fixed cell calculation that have no analog in the optimized cell geometry calculation. Specifically, modes $\nu_2^{(fix)}$, $\nu_3^{(fix)}$, $\nu_4^{(fix)}$ and $\nu_8^{(fix)}$ are vibrations whose motion is not found in the set of vibrations using the optimized unit cell. In this sense, although the IR-active vibrations in the THz region are similar in the fixed and optimized cell calculations, the systems represented by the two calculations are quite different. The presence of vibrational modes that are completely absent from the optimized cell calculation suggests that the fixed cell calculation is reflecting something akin to a high-pressure solid phase of DL-valine, albeit an idiosyncratic state in which stress has not been applied uniformly.

The ***a***, ***b***, and ***c*** vectors of the crystallographic unit cell generally do not all change by the same percentage upon optimization. This is illustrated in Table 3, which compares the optimized unit cell geometry in each calculation to the

Table 3 Calculated unit cell parameters for DL-valine and L-valine are reported. Entries in bold are the experimental values for the crystallographic structure used as a starting point for each calculation (these structures are referred to using their Cambridge Crystallographic Database identifier). In the case of DL-valine, the calculations were performed using a variety of basis sets and the unit cell parameters are listed for each of these. The cell vectors calculated using the minimal 3-21G basis set are closest to experimental values, but this is only a consequence of basis set superposition error and its resultant artificial forces

	Temp./K	*a*	*b*	*c*	*α*	*β*	*γ*
VALIDL03	**100**	**5.233**	**5.415**	**10.830**	**90.83**	**92.29**	**110.01**
3-21G		5.158578	5.398365	11.163325	91.400760	91.534037	110.242593
6-31G		5.348202	5.477154	11.630725	91.013071	93.460178	109.401412
6-31G*		5.321987	5.491873	11.658328	91.111177	92.812791	110.241482
6-31G**		5.314786	5.462077	11.659320	90.967448	93.028513	110.102025
LVALIN01	**120**	**9.682**	**5.247**	**11.930**	**90**	**90.57**	**90**
6-31G**		9.744634	5.289243	12.485422	90.000684	90.654787	90.004880

lowest temperature crystallographic values available. In every DL-valine calculation, as well as the L-valine calculation, the largest discrepancy is in the value for ***c***, which for both species is the vector normal to the alternating hydrophobic and hydrophilic layers. This has been observed in other DFT studies of L-valine,[45] and is almost certainly due to inadequate treatment of van der Waals interactions by the model. While fixing the cell at smaller dimensions than predicted by the DFT calculation might appear to compensate for the missing van der Waals forces, it is not expected to lead to more accurate calculated vibrations. An increase in other types of interaction is in no way equivalent to the actual van der Waals interaction potential.

Experimental unit cell values reflect the physics of the real world, but they are completely adventitious with regard to the physics of any less than perfect model. If one's model does not reproduce the unit cell geometry of a given system to one's satisfaction, there is no reason to assume that enforcing crystallographic parameters will increase the accuracy of the model or any subsequent calculation based on it. Even if the calculated vibrational spectrum using a fixed unit cell is more similar to experiment than that obtained with an optimized cell, it is not logical to assert that the former calculation is necessarily more accurate than the latter. The inaccuracies of a model applied self-consistently can be more illuminating than the accuracies of a model applied arbitrarily.

It is worth remarking that the modes appearing in both the optimized and fixed cell calculations are those that have no translational character. This is consistent with the observation that translational intermolecular modes are generally the most strongly affected by changes in unit cell size due to temperature or pressure variations.[46] In the case of DL-valine, it happens that these modes are not IR-active, which results in similar calculated spectra. This fortuity actually obscures the differences between the two calculations. Other systems are expected to have low-frequency modes that are both IR-active and translational in nature (see L-valine calculations below), in which case calculated spectra of the fixed system will most likely differ from those of a fully optimized system.

3.3 L-Valine

The experimental and calculated (6-31G** basis) THz spectra of L-valine are presented in Fig. 7, along with a quantitative description of the type of motion involved for each mode. While all of the IR-active modes calculated in the THz region for DL-valine are of a similar type (mixed intramolecular and librational, with no translational character), the IR-active modes calculated for L-valine have a variety of different types of motion. For example, ν_4 is predominantly intramolecular, while ν_5, ν_6, and ν_7 have significant translational intermolecular character and ν_8 is a mix of intramolecular and librational motion.

Temperature dependence of valine spectra. Fig. 8 displays the temperature-dependent absorption spectrum of L-valine between 1.1 THz and 2.5 THz over a range of temperatures from 78 K to 298 K. The Raman spectrum of L-valine between 1 THz and 4 THz was reported by Lima and coworkers.[47] At room temperature they report peaks at 1.4 THz and 1.6 THz. At low temperature they report one Raman active peak at 1.5 THz (having migrated from 1.4 THz at room temperature).

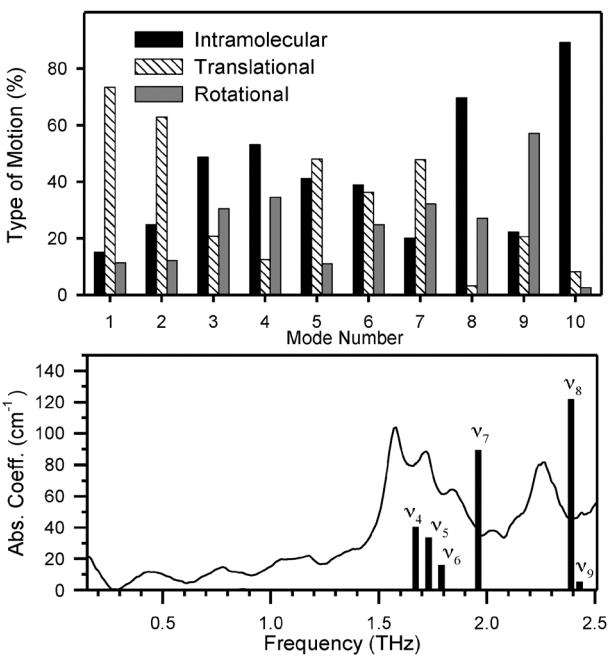

Fig. 7 The frequencies and intensities of infrared active L-valine vibrational modes calculated with the 6-31G** basis set (black stick spectra in the lower plot). The experimental 78 K THz absorption spectrum is provided for comparison. The upper plot shows the quantified character of the vibrational motion for the ten lowest frequency modes calculated. Labels above the stick spectra refer to the mode number as calculated using that particular basis set.

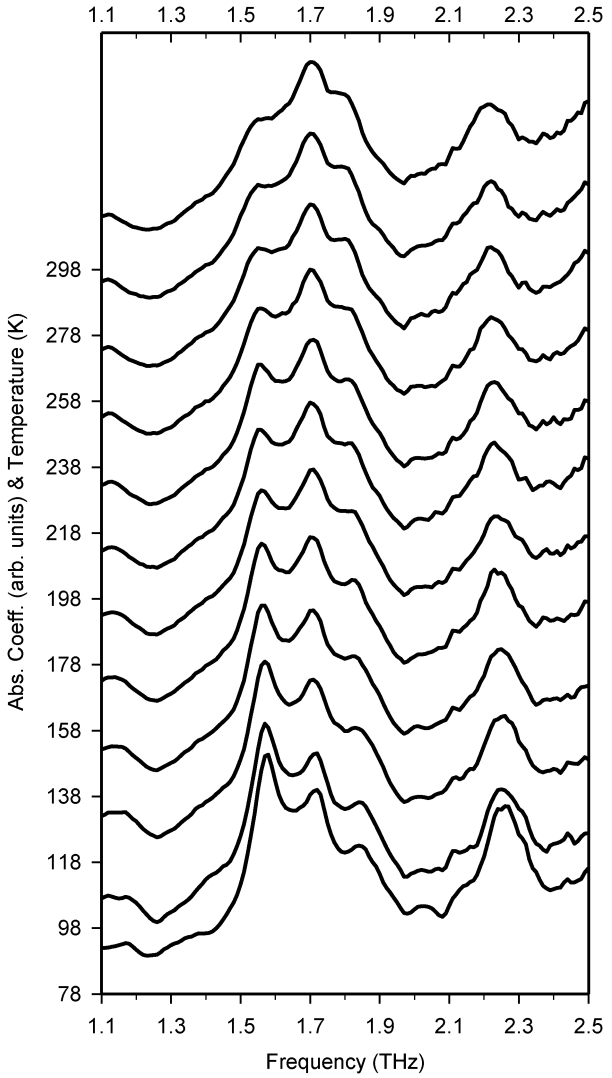

Fig. 8 Temperature dependence of L-valine THz spectra. The absorption coefficients (in cm^{-1}) have been multiplied by 0.7, and the spectra have been offset by the temperature at which the spectrum was collected (from 78 K to 298 K in steps of 20 K). The temperature dependence of D-valine is identical.

The THz spectrum is somewhat richer, and has four discernable peaks at 78 K.

All observed peaks in L-valine red shift with increasing temperature. In Fig. 9 the position of the four resolved features is reported as a function of temperature. These values were extracted from our data by fitting Voigt functions to the pertinent regions of the spectrum. It was assumed that three peaks overlapped between 1.5 and 1.9 THz and that the feature at 2.25 THz was a single peak. The baseline was also a fitted parameter and was assumed to be linear. The features at 1.5, 1.85 and 2.25 THz red shift as the temperature rises. The frequency of these three peaks has a similar temperature dependence when considered on a percentage basis, while that for the peak at approximately 1.7 THz is much smaller.

In general, the phenomenon of temperature-dependent vibrational frequency is explained by anharmonicity. In the solid state, a certain amount of care must be taken in describing the nature of this anharmonicity in relation to any particular mode. In systems with more than one mode, concepts such as a diatomic Morse oscillator (or any other one-dimensional model) are not sufficient to describe the temperature dependence of spectra due to modes coupling to each other. Another approximate description of the frequency shift consists in evaluating harmonic frequencies for the unit cell parameters taken from experiment for different temperatures. This approach simply evades treating the anharmonicity that is the essential cause of the change in the unit cell parameters with temperature. The average positions of atoms in a totally harmonic crystal have no temperature dependence, and hence there would be no change in volume as a function of temperature.[48] A detailed account of the anharmonicity in the potential energy surface along the displacement coordinate of normal modes as well as the coupling between modes is a challenging problem that will be addressed in future work.

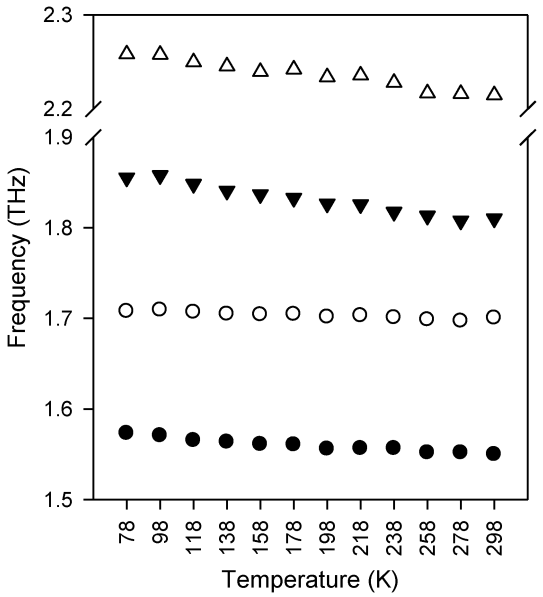

Fig. 9 The positions of IR absorption peaks in the THz spectrum of L-valine depend on temperature to an extent that varies from mode to mode. These data points were obtained by fitting Voigt functions (with a linear baseline) to each spectrum in Fig. 8 with the assumption that the overlapping features between 1.5 and 1.9 THz consisted of three peaks while the feature at 2.25 THz corresponded to a single peak.

3.4 Mode assignments

The space group of the DL-valine polymorph studied here is $P\bar{1}$. The unit cell is triclinic and contains a pair of valine enantiomers each belonging to the C_1 point group and whose positions are related by an inversion center. If one were to assume that molecules are rigid, factor group analysis[49] of this crystalline system predicts 12 intermolecular modes, six of A_g symmetry and six of A_u symmetry. Three of the A_u vibrations are acoustic modes (appearing in our calculations as the three modes at zero frequency in which all atoms translate in the same direction) and the remaining nine modes ($6\ A_g + 3\ A_u$) are IR-allowed.

The results of our DL-valine calculations indicate that the internal and external degrees of freedom of the molecules in

these crystals are not independent, even at the lowest vibrational frequencies. In this case, a rigid molecule approximation is not appropriate. However, although we cannot assign uniquely identifying symmetry labels to the vibrations, it is still possible to describe their motion in quantitative terms. In particular, the calculated DL-valine modes with non-negligible IR activity in the THz region all have motions consisting of libration of the molecular units combined with displacements along intramolecular degrees of freedom. Considering the fact that this motion was calculated consistently using a variety of basis sets and even in the fixed cell calculation, it is proposed that the features appearing in the THz spectrum correspond to vibrations of this type. Given the fairly large separation between their frequencies, it is reasonable to assign mode $\nu_2^{(6\text{-}31G^{**})}$ to the peak in the DL-valine spectrum at 1.70 THz and mode $\nu_6^{(6\text{-}31G^{**})}$ to the peak at approximately 2.80 THz.

The L-valine polymorph belongs to the space group $P2_1$ and contains four molecules of C_1 symmetry. There are two unique sites in the monoclinic unit cell, and a pair of molecules occupying one of each type is related to the other pair of molecules in the cell by a two-fold screw axis. Rigid-molecule factor group analysis of this system predicts 24 external modes represented by 12A + 12B. The acoustic modes are A + 2B, leaving 11A + 10B rigid intermolecular modes, all of which are IR-allowed. As in the case of DL-valine, the L-valine crystals have THz frequency vibrations whose internal and external degrees of freedom are mixed, suggesting that a rigid molecule model is not an accurate description of the system. However, the result of factor group analysis that L-valine should have a larger number of potentially observable intermolecular vibrational modes than DL-valine seems to be reflected in both the calculated and experimental spectra.

While the calculated spectrum of L-valine appears to reproduce some of the key features of the experimental one, a direct, one-to-one assignment of features is not justified. For example, it might seem obvious to assign ν_8 to the feature at 2.25 THz. However, this would imply that either the calculated intensity of ν_7 is particularly inaccurate, or that its calculated frequency is wrong and that perhaps ν_7 should correspond to the strong feature at 1.55 THz. However, if one is willing to assume that the calculated frequency of ν_7 (1.97 THz) differs that much from the mode it represents, there is no reason not to assign ν_7 instead of ν_8 to the peak at 2.25 THz. Nonetheless, the fact that the many peaks visible in the THz spectrum of L-valine represent a variety of different types of motion suggests that this system will be very interesting for further study of the relationship between temperature dependence and mode character.

4. Conclusions

Terahertz spectra have been measured for polycrystalline samples comprised of the D and L enantiomers of valine, as well as the 50/50 racemic mixture. The spectra of enantiopure and racemic samples differ strongly. Powder XRD data were compared with calculated powder patterns (based on the known atomic coordinates of the solids) to verify the morphology of the sample. DFT calculations utilizing periodic boundary conditions were performed using the GAUSSIAN program, and a method to derive the relative proportions of vibrational and rotational characters in calculated intermolecular modes was demonstrated. It was shown that this quantitative approach to describing specific modes substantially aids comparison between calculations carried out with slightly different parameters and will be useful when making comparisons between calculations carried out using different methods or by different researchers. This is due to the fact that the underlying character of the low-frequency modes is less sensitive to variation in the computational model than are the calculated frequencies and infrared intensities.

Future studies will include the investigation of methods that can aid in the identification of particular modes based on the nature of their motion as well as their sensitivity to temperature change. In addition, THz spectra from liquid helium temperature upwards will be acquired to provide detailed data on temperature dependence. Developing a model of the mode-specific temperature dependence of low-frequency vibrations in molecular crystals that is both accurate and tractable is challenging. Future computational work will investigate the possibility of rigorously building this model up from 0 K DFT calculations, rather than making comparisons between systems where the lattice vectors have been artificially fixed in order to approximate the unit cell volume at finite temperatures (the latter approach, as discussed above, resulting in optimized structures with hidden strain). Finally, in order to calculate the dynamical properties of hydrophobic molecular crystals more accurately, it will be necessary to employ computational methods that explicitly account for van der Waals interactions. It is hoped that the experimental spectra and rigorous calculations presented here will provide a useful basis of comparison for the next generation of computational work.

Acknowledgements

We acknowledge the National Science Foundation (CHE-0911593) and AstraZeneca for partial support of this work. This work was supported in part by the facilities and staff of the Yale University Faculty of Arts and Sciences High Performance Computing Center and by the National Science Foundation under grant #CNS 08-21132 that partially funded acquisition of the facilities. KS thanks the German Academic Exchange Service (DAAD) for funding, TAF acknowledges the Dan David Foundation for additional funding as well. We thank Drs Chris Incarvito, J. Michael McBride, and Jon Parr for technical assistance and/or helpful discussions.

References

1 T. M. Korter, R. Balu, M. B. Campbell, M. C. Beard, S. K. Gregurick and E. J. Heilweil, *Chem. Phys. Lett.*, 2006, **418**, 65.
2 M. Yamaguchi, F. Miyamaru, K. Yamamoto, M. Tani and M. Hangyo, *Appl. Phys. Lett.*, 2005, **86**, 053903.
3 K. Yamamoto, K. Tominaga, H. Sasakawa, A. Tamura, H. Murakami and H. Ohtake, *et al.*, *Biophys. J.*, 2005, **89**, L22.
4 C. J. Strachan, P. F. Taday, D. A. Newnham, K. C. Gordon, J. A. Zeitler and M. Pepper, *et al.*, *J. Pharm. Sci.*, 2005, **94**, 837.
5 D. G. Allis, D. A. Prokhorova and T. M. Korter, *J. Phys. Chem. A*, 2006, **110**, 1951.

6 M. R. Kutteruf, C. M. Brown, L. K. Iwaki, M. B. Campbell, T. M. Korter and E. J. Heilweil, *Chem. Phys. Lett.*, 2003, **375**, 337.
7 A. G. Markelz, A. Roitberg and E. J. Heilweil, *Chem. Phys. Lett.*, 2000, **320**, 42.
8 Y. C. Shen, P. C. Upadhya, E. H. Linfield and A. G. Davies, *Appl. Phys. Lett.*, 2003, **82**, 2350.
9 Y. C. Shen, P. C. Upadhya, E. H. Linfield and A. G. Davies, *Vib. Spectrosc.*, 2004, **35**, 111.
10 C. F. Zhang, E. Tarhan, A. K. Ramdas, A. M. Weiner and S. M. Durbin, *J. Phys. Chem. B*, 2004, **108**, 10077.
11 S. E. Whitmire, D. Wolpert, A. G. Markelz, J. R. Hillebrecht, J. Galan and R. R. Birge, *Biophys. J.*, 2003, **85**, 1269.
12 A. Markelz, S. Whitmire, J. Hillebrecht and R. Birge, *Phys. Med. Biol.*, 2002, **47**, 3797.
13 G. V. Papamokos and I. N. Demetropoulos, *J. Phys. Chem. A*, 2004, **108**, 8160.
14 K. Itoh and H. Katabuch, *Biopolymers*, 1973, **12**, 921.
15 K. Itoh, T. Shimanou and H. Katabuch, *Nature (London), New Biol.*, 1972, **239**, 42.
16 K. Itoh and T. Shimanou, *Biopolymers*, 1970, **9**, 383.
17 K. Itoh, T. Nakahara, T. Shimanou, M. Oya, K. Uno and Y. Iwakura, *Biopolymers*, 1968, **6**, 1759.
18 W. N. Wang, W. M. Yue, H. T. Yan, C. L. Zhang and G. Z. Zhao, *Chin. Sci. Bull.*, 2005, **50**, 1561.
19 A. Matei, N. Drichko, B. Gompf and M. Dressel, *Chem. Phys.*, 2005, **316**, 61.
20 P. F. Taday, I. V. Bradley and D. D. Arnone, *J. Biol. Phys.*, 2003, **29**, 109.
21 B. Yu, F. Zeng, Y. Yang, Q. Xing, A. Chechin and X. Xin, *et al.*, *Biophys. J.*, 2004, **86**, 1649.
22 M. D. King, P. M. Hakey and T. M. Korter, *J. Phys. Chem. A*, 2010, **114**, 2945.
23 A. M. Micu, D. Durand, M. Quilichini, M. J. Field and J. C. Smith, *J. Phys. Chem.*, 1995, **99**, 5645.
24 M. Walther, B. M. Fischer and P. U. Jepsen, *Chem. Phys.*, 2003, **288**, 261.
25 J. M. de Souza, P. T. C. Freire, H. N. Bordallo and D. N. Argyriou, *J. Phys. Chem. B*, 2007, **111**, 5034.
26 R. Car and M. Parrinello, *Phys. Rev. Lett.*, 1985, **55**, 2471.
27 P. U. Jepsen and S. J. Clark, *Chem. Phys. Lett.*, 2007, **442**, 275.
28 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb and J. R. Cheeseman, *et al.*, *Gaussian Development Version, Revision G.03*, Wallingford, CT, 2008.
29 B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan and M. Karplus, *J. Comput. Chem.*, 1983, **4**, 187.
30 F. H. Allen, *Acta Crystallogr., Sect. B: Struct. Sci.*, 2002, **58**, 380.
31 J. P. Perdew, K. Burke and M. Ernzerhof, *Phys. Rev. Lett.*, 1996, **77**, 3865.
32 A. F. Izmaylov and G. E. Scuseria, *J. Chem. Phys.*, 2007, **127**, 144106.
33 A. F. Izmaylov and G. E. Scuseria, *Phys. Rev. B: Condens. Matter*, 2008, **77**, 165131.
34 B. Dalhus and C. H. Gorbitz, *Acta Chem. Scand.*, 1996, **50**, 544.
35 R. Flaig, T. Koritsanszky, B. Dittrich, A. Wagner and P. Luger, *J. Am. Chem. Soc.*, 2002, **124**, 3407.
36 K. N. Kudin, R. Car and R. Resta, *J. Chem. Phys.*, 2007, **126**, 234101.
37 G. R. Kneller, *J. Chem. Phys.*, 2008, **128**, 194101.
38 K. N. Kudin and A. Y. Dymarsky, *J. Chem. Phys.*, 2005, **122**, 224105.
39 P. F. Bernath, *Spectra of Atoms and Molecules*, Oxford University Press, New York, 1995.
40 W. H. Press, S. A. Teukolsky, W. T. Vetterling and B. P. Flannery, *Numerical Recipes in Fortran*, Cambridge University Press, Cambridge, 2nd edn, 1992.
41 M. C. Beard, G. M. Turner and C. A. Schmuttenmaer, *J. Phys. Chem. B*, 2002, **106**, 7146.
42 A. B. True, K. Schroeck, T. French and C. A. Schmuttenmaer, *J. Infrared, Millimeter, Terahertz Waves*, 2010, DOI: 10.1007/s10762-010-9645-0.
43 P. M. Hakey, D. G. Allis, W. Ouellette and T. M. Korter, *J. Phys. Chem. A*, 2009, **113**, 5119.
44 P. U. Jepsen and B. M. Fischer, *Opt. Lett.*, 2005, **30**, 29.
45 P. R. Tulip and S. J. Clark, *Phys. Rev. B: Condens. Matter*, 2005, **71**, 195117.
46 A. Schauer, *Can. J. Phys.*, 1964, **42**, 1857.
47 J. A. Lima, P. T. C. Freire, R. J. C. Lima, A. J. D. Moreno, J. Mendes and F. E. A. Melo, *J. Raman Spectrosc.*, 2005, **36**, 1076.
48 N. W. Ashcroft and D. N. Mermin, *Solid State Physics*, Saunders College, Philadelphia, 1976.
49 D. L. Rousseau, R. P. Bauman and S. P. S. Porto, *J. Raman Spectrosc.*, 1981, **10**, 253.

Supplementary Information for: Terahertz Spectroscopy of Enantiopure and Racemic Polycrystalline Valine

Michael R. C. Williams, Alan B. True, Artur F. Izmaylov, Timothy A. French,[a] Konstanze Schroeck, and Charles A. Schmuttenmaer*
Yale University, Department of Chemistry, PO Box 208107, 225 Prospect St., New Haven, CT 06520-8107

[a]Current address: Department of Chemistry and Chemical Biology, Harvard University, 1 Oxford Street, Cambridge, MA 02138

*Email: charles.schmuttenmaer@yale.edu

March 3, 2011

There are four parts of this Supplementary Information. 1. Unit cell parameters for all instances of valine (without any salts or co-solvents) from the Cambridge Structural Database (CSD).[2] 2. Powder XRD spectra. 3. Plots of the quantified type of motion for the systems studied. 4. Movie files for vibrational modes discussed in the text.

I. Information from the CSD

Sample	Cambridge Database Identifier	Temp (K)	Space Group	Cell Lengths (Å)			Cell Angles (°)		
				a	*b*	*c*	α	β	γ
D-valine	**AHEJEC03[1]**	**293**	**$P2_1$**	**9.673**	**5.252**	**12.043**	**90**	**90.75**	**90**
	AHEJEC02	270	$P2_1$	9.67	5.268	12.026	90	90.72	90
	AHEJEC	223	$P2_1$	9.666	5.257	11.984	90	90.66	90
	AHEJEC01	173	$P2_1$	9.661	5.246	11.941	90	90.59	90
L-valine	**LVALIN[3]**	**295**	**$P2_1$**	**9.710**	**5.270**	**12.060**	**90**	**90.80**	**90**
	LVALIN01	120	$P2_1$	9.682	5.247	11.930	90	90.57	90
	LVALIN02	270	$P2_1$	9.674	5.266	12.020	90	90.72	90
DL-valine	VALIDL02	120	P-1	5.222	5.406	10.838	90.89	92.34	110.02
	VALIDL01	295	-	5.430	11.050	5.250	92.40	109.40	91.00
	VALIDL03[5]	**100**	**P-1**	**5.233**	**5.415**	**10.830**	**90.83**	**92.29**	**110.01**
	VALIDL	295	$P2_1/c$	5.210	22.100	5.410	90	109.20	90

Table S1. The cell parameters of valine found in the CSD.

X-ray powder diffraction measurements are used to confirm the morphology of the samples. Table S1 provides all available entries in the CSD for valine (without any salts or co-solvents). Our samples are found to have the structures shown in bold type. The cell parameters and space groups confirm that the crystals of pure D- and L- enantiomers have the same structure as each other, and the racemate has a completely different crystal structure.

II. Powder XRD Spectra

The atomic coordinates from the CSD are used to calculate x-ray powder patterns using Mercury 1.4.1 software.[4] The experimental and calculated results are then compared to identify the samples based on the main features in the diffraction patterns (see Figure 3 in main text, and Figures S1 and S2). The intensities of the powder patterns are not as important for identification as is the scattering angle, 2θ. Scattering intensities are highly dependent on sample preparation (*e.g.*, anisotropy) and experimental parameters. Pulverizing the samples in a ball mill improves the data considerably by providing more uniform crystallite sizes and eliminating any anisotropies arising from unusual crystal shapes such as needles or plates.[6] Pulverizing the samples for two minutes is sufficient, and longer times causes the peak widths to increase. This could be due the particles becoming so small that their diffraction is broader than the instrument resolution, or because the samples become somewhat amorphous due to heating or mechanical distortion due to thrashing during the pulverization process. The unit cell parameters of the amino acids studied in this work, including any known polymorphs, are presented in Table S1.

Our powder XRD spectra were obtained at room temperature, and unfortunately the only room temperature CSD entry for the triclinic polymorph of DL-valine (VALIDL01) did not include atomic positions, only unit cell parameters. However, by scaling the unit cell dimensions (but not atomic coordinates) of the VALIDL03 structure to match that of VALIDL01, it was possible to identify our sample material as the triclinic form of DL-valine. The calculated diffraction pattern of the approximated room temperature VALIDL03 structure is in good agreement with our data, while the calculated diffraction pattern of the monoclinic polymorph (VALIDL), is not. Figures S1 and S2 illustrate the process of identifying the DL-valine polymorph.

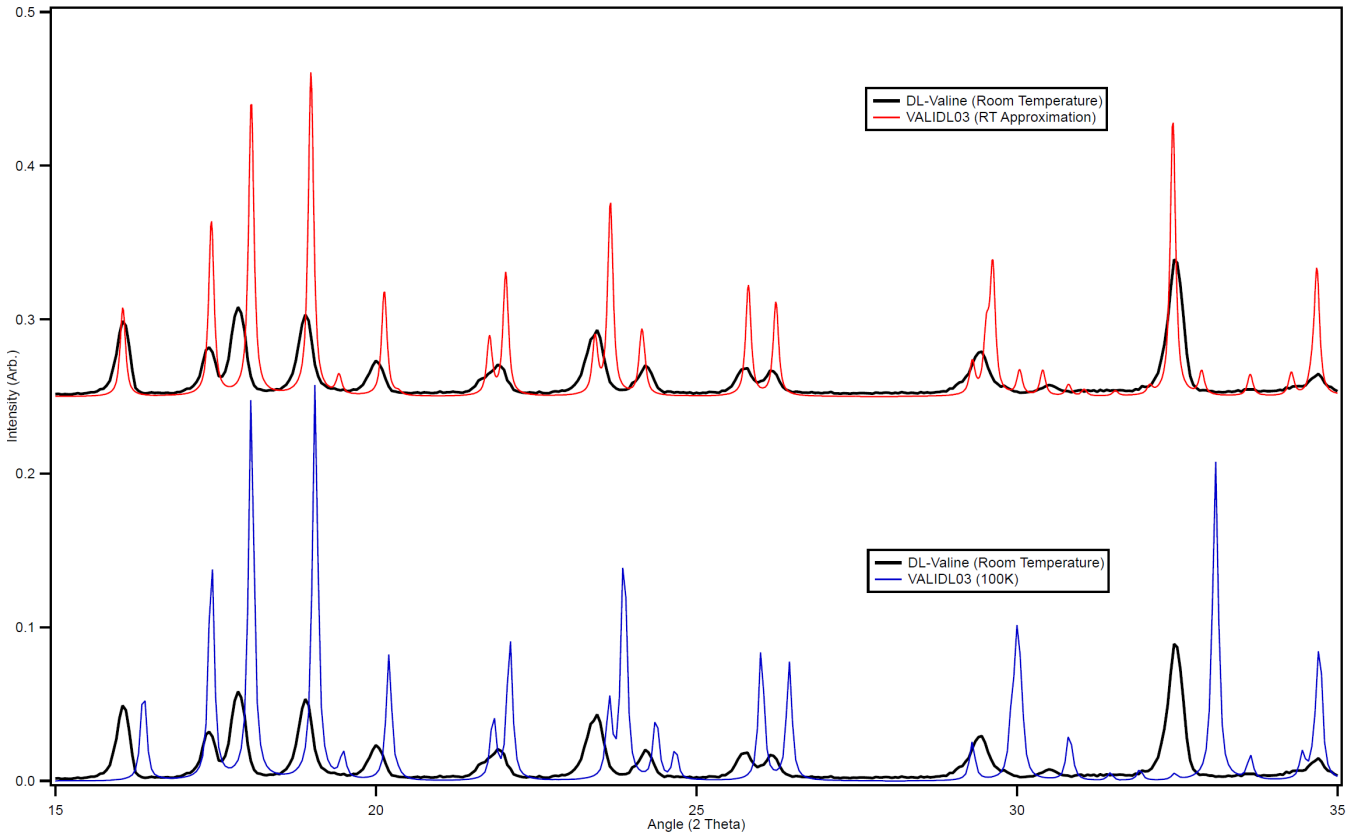

Figure S1. Low temperature crystallographic coordinates of the triclinic form of DL-valine (VALIDL03) do not initally result in a calculated diffraction pattern in accord with our room temperature powder XRD data (bottom). However, after simply scaling the unit cell of VALIDL03 to room temperature values, the calculated pattern is in

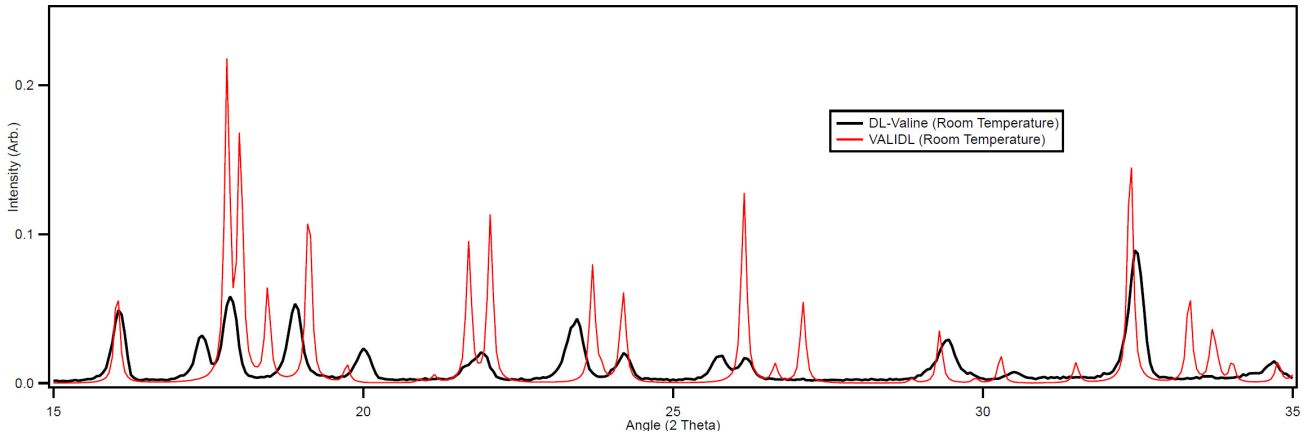

Figure S2. The calculated powder pattern of the monoclinic form of DL-valine (VALIDL) agrees poorly with our data. Unlike VALIDL03, the VALIDL coordinates were obtained at room temperature, and are therefore compared directly with the experimental powder XRD pattern.

III. Quantified Character of Motion for Low-Frequency Vibrational Modes in DL-Valine and L-Valine Crystal Systems

DL-valine 3-21G	Freq (THz)	Intramolecular (%)	Total Intermolecular (%)	Translational Intermolecular (%)	Rotational Intermolecular (%)
1	1.45	46.52	53.48	39.88	13.60
2	1.97	17.28	82.72	60.81	21.91
3	2.04	60.98	39.02	0.01	39.01
4	2.76	36.73	63.27	19.01	44.26
5	3.34	63.32	36.68	0.01	36.67
6	3.51	70.78	29.22	25.30	3.92
7	4.00	63.63	36.37	12.46	23.91
8	4.32	69.86	30.14	8.89	21.25
9	4.48	75.86	24.14	0.01	24.13
10	4.99	74.90	25.10	13.72	11.38
11	5.13	99.02	0.98	0.00	0.98
12	6.43	89.13	10.87	1.76	9.11
13	6.93	95.21	4.79	0.00	4.79
14	7.03	87.70	12.30	1.34	10.96
15	7.10	84.72	15.28	0.00	15.27
16	7.98	94.78	5.22	0.12	5.10
17	8.23	92.55	7.45	0.02	7.44
18	9.18	98.44	1.56	0.02	1.53
19	9.81	95.41	4.59	0.02	4.57
20	10.01	99.61	0.39	0.04	0.35

DL-valine 6-31G	Freq (THz)	Intramolecular (%)	Total Intermolecular (%)	Translational Intermolecular (%)	Rotational Intermolecular (%)
1	1.09	36.75	63.25	34.01	29.24
2	1.35	35.31	64.69	0.01	64.68
3	1.62	11.21	88.79	56.94	31.85
4	2.20	22.55	77.45	27.56	49.89
5	2.77	66.57	33.43	30.32	3.12
6	3.11	70.41	29.59	0.01	29.59
7	3.57	73.34	26.66	6.72	19.93
8	3.74	77.88	22.12	0.01	22.12
9	3.82	45.31	54.69	6.17	48.51
10	4.56	75.43	24.57	16.67	7.90
11	4.83	99.94	0.06	0.00	0.06
12	5.77	79.79	20.21	0.95	19.26
13	6.10	75.51	24.49	0.01	24.48
14	6.34	92.56	7.44	2.15	5.30
15	6.39	95.76	4.24	0.02	4.22
16	7.24	99.82	0.18	0.10	0.08
17	7.26	99.38	0.62	0.01	0.61
18	8.35	97.54	2.46	0.09	2.37
19	8.75	96.54	3.46	0.00	3.45
20	9.01	99.93	0.07	0.03	0.04

DL-valine 6-31G*	Freq (THz)	Intramolecular (%)	Total Intermolecular (%)	Translational Intermolecular (%)	Rotational Intermolecular (%)
1	1.09	37.10	62.90	43.26	19.64
2	1.44	36.86	63.14	0.00	63.14
3	1.54	16.89	83.11	51.28	31.83
4	2.02	21.38	78.62	23.09	55.53
5	2.70	65.90	34.10	30.91	3.18
6	3.02	69.53	30.47	0.01	30.47
7	3.55	58.17	41.83	8.03	33.80
8	3.74	74.80	25.20	5.89	19.31
9	3.84	76.03	23.97	0.02	23.95
10	4.59	75.11	24.89	15.33	9.56
11	4.66	99.77	0.23	0.01	0.22
12	5.77	81.74	18.26	1.19	17.06
13	6.04	75.61	24.39	0.01	24.39
14	6.24	91.77	8.23	2.32	5.91
15	6.38	96.27	3.73	0.00	3.72
16	7.24	99.31	0.69	0.16	0.53
17	7.28	99.41	0.59	0.02	0.57
18	8.28	98.01	1.99	0.08	1.91
19	8.71	96.73	3.27	0.01	3.26
20	9.31	99.80	0.20	0.04	0.16

DL-valine 6-31G**	Freq (THz)	Intramolecular (%)	Total Intermolecular (%)	Translational Intermolecular (%)	Rotational Intermolecular (%)
1	1.07	38.43	61.57	42.61	18.96
2	1.35	37.09	62.91	0.01	62.91
3	1.53	18.84	81.16	48.93	32.23
4	1.98	41.39	58.61	26.55	32.06
5	2.67	66.14	33.86	30.10	3.75
6	2.99	65.44	34.56	0.01	34.55
7	3.52	56.74	43.26	7.28	35.98
8	3.67	73.96	26.04	6.00	20.05
9	3.80	76.66	23.34	0.02	23.31
10	4.44	76.01	23.99	15.18	8.81
11	4.60	99.65	0.35	0.00	0.34
12	5.59	79.32	20.68	0.54	20.14
13	5.89	75.29	24.71	0.00	24.71
14	6.14	93.15	6.85	1.95	4.90
15	6.29	95.87	4.13	0.01	4.12
16	7.21	99.53	0.47	0.18	0.28
17	7.24	99.31	0.69	0.03	0.66
18	8.20	98.17	1.83	0.03	1.80
19	8.60	97.61	2.39	0.01	2.39
20	9.19	99.84	0.16	0.02	0.14

DL-valine 6-31G** Fixed Cell	Freq (THz)	Intramolecular (%)	Total Intermolecular (%)	Translational Intermolecular (%)	Rotational Intermolecular (%)
1	1.70	37.90	62.10	0.01	62.10
2	1.74	16.29	83.71	82.42	1.29
3	2.18	26.88	73.12	1.18	71.94
4	2.37	13.44	86.56	76.96	9.60
5	3.15	58.70	41.30	0.01	41.29
6	3.17	77.32	22.68	21.36	1.32
7	3.93	68.59	31.41	16.39	15.02
8	4.05	43.97	56.03	1.74	54.29
9	4.37	77.90	22.10	0.01	22.09
10	4.59	73.66	26.34	16.04	10.29
11	4.88	95.51	4.49	0.00	4.49
12	5.94	82.26	17.74	0.47	17.27
13	6.23	79.41	20.59	0.01	20.57
14	6.43	94.57	5.43	0.01	5.42
15	6.48	93.49	6.51	1.48	5.03
16	7.45	99.21	0.79	0.22	0.57
17	7.71	99.10	0.90	0.05	0.85
18	8.39	98.22	1.78	0.01	1.77
19	8.70	98.05	1.95	0.00	1.95
20	9.50	99.69	0.31	0.16	0.15

L-valine 6-31G**	Freq (THz)	Intramolecular (%)	Total Intermolecular (%)	Translational Intermolecular (%)	Rotational Intermolecular (%)
1	0.87	15.17	84.83	73.50	11.33
2	1.32	24.91	75.09	62.97	12.12
3	1.37	48.70	51.30	20.78	30.52
4	1.67	53.11	46.89	12.44	34.45
5	1.73	41.10	58.90	47.95	10.95
6	1.79	38.78	61.22	36.25	24.97
7	1.96	20.17	79.83	47.79	32.04
8	2.39	69.74	30.26	3.15	27.11
9	2.43	22.24	77.76	20.63	57.13
10	2.63	89.32	10.68	8.21	2.47
11	2.71	80.33	19.67	5.81	13.87
12	2.76	52.37	47.63	43.17	4.45
13	2.79	70.72	29.28	22.75	6.53
14	2.95	58.32	41.68	5.60	36.09
15	3.13	71.01	28.99	5.02	23.97
16	3.16	50.44	49.56	4.01	45.55
17	3.28	70.67	29.33	2.41	26.92
18	3.60	71.67	28.33	1.38	26.95
19	3.61	70.53	29.47	2.58	26.89
20	3.91	78.05	21.95	10.49	11.47
21	4.00	87.91	12.09	5.57	6.52
22	4.34	58.99	41.01	32.28	8.73
23	4.59	78.90	21.10	11.62	9.49
24	4.63	80.48	19.52	4.54	14.98
25	4.85	65.86	34.14	7.08	27.06
26	5.02	73.54	26.46	4.05	22.41
27	5.30	73.56	26.44	5.19	21.25
28	5.65	89.21	10.79	8.68	2.11
29	5.74	68.74	31.26	5.53	25.74
30	6.10	97.14	2.86	1.46	1.40
31	6.28	95.73	4.27	3.45	0.82
32	6.39	97.66	2.34	1.91	0.43
33	6.51	84.04	15.96	7.83	8.13
34	6.55	92.86	7.14	2.91	4.22
35	6.85	93.69	6.31	2.93	3.38
36	6.87	99.63	0.37	0.21	0.17
37	6.88	99.09	0.91	0.36	0.55
38	7.57	99.41	0.59	0.09	0.50
39	7.72	98.45	1.55	0.47	1.08
40	8.19	99.33	0.67	0.33	0.33

IV. Movie Files

Animations of vibrational modes discussed in the article are available here in the animated .gif format, which allows convenient viewing of several modes at a time. The files are named in the following manner:

{compound}_{basis set}_{mode number}_{unit cell vector}.gif

Where the unit cell vector refers to the direction along which one is viewing the molecules in a particular animation. For example, the file "dlval_631G_nu8_b.gif" is an animation of the eighth vibrational mode calculated for DL-valine using the 6-31G basis set and viewed along unit cell vector b. Note that since the * character is not allowed in file names, the other common designation for basis sets is used. Specifically, 6-31G(d) is the same as 6-31G*, and 6-31G(d,p) is the same as 6-31G**.

References

1. W. Q. Wang, Y. Gong, Z. M. Wang, and C. H. Yan, *Chin. J. Struct. Chem.*, 2003, **22,** 539.
2. F. H. Allen, *Acta Crystallogr., Sect. B: Struct. Sci.*, 2002, **58,** 380.
3. K. Torii and Y. Iitaka, *Acta Crystallogr., Sect. B: Struct. Sci.*, 1970, **B 26,** 1317.
4. C. F. Macrae, P. R. Edgington, P. McCabe, E. Pidcock, G. P. Shields, R. Taylor *et al.*, *J. Appl. Crystallogr.*, 2006, **39,** 453.
5. R. Flaig, T. Koritsanszky, B. Dittrich, A. Wagner, and P. Luger, *J. Am. Chem. Soc.*, 2002, **124,** 3407.
6. P. Yang, H. L. Wei, H. L. Huang, B. A. Baum, Y. X. Hu, G. W. Kattawar *et al.*, *Appl. Optics*, 2005, **44,** 5512.