THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

Intermolecular Vibrations in Hydrophobic Amino Acid Crystals: Experiments and Calculations

Michael R. C. Williams, Daniel J. Aschaffenburg, Benjamin K. Ofori-Okai, and Charles A. Schmuttenmaer*

Department of Chemistry, Yale University, P.O. Box 208107, 225 Prospect Street, New Haven, Connecticut 06520-8107, United States

Supporting Information

ABSTRACT: Intermolecular vibrations of amino acid crystals occur in the THz, or far-infrared, region of the electromagnetic spectrum. We have measured the THz and Raman spectra of DL-leucine as well as two polymorphs of DL-valine, the spectroscopic properties of which have not previously been compared. Theoretical modeling of intermolecular vibrations in hydrophobic amino acids is challenging because the van der Waals interactions between molecules are not accounted for in standard density functional theory. Therefore, to calculate the vibrational modes, we used a recently developed approach that includes these nonlocal electron correlation forces. We discuss methods for comparing results from different theoretical models using metrics other than calculated vibrational frequency and intensity, and we also report a new approach enabling concise comparison of vibrational modes that involve complicated mixtures of inter- and intramolecular displacements.

INTRODUCTION

Hydrophobic amino acids interact with their chemical environment through a combination of electrostatic, hydrogen bonding, dipole, induced dipole, and dispersion forces, each of which have their own characteristic energy scale and distance dependence. Spectroscopy of the intermolecular vibrations in crystalline hydrophobic amino acids provides a direct measurement of the dynamics resulting from the interplay of all of these forces. The connections between intermolecular forces and the specific motions of amino acid molecules are interesting in their own right and also relevant to understanding the behavior of proteins and other important biomolecules.

For example, the vibrations of amino acid functional groups in dipeptide nanotubes may facilitate transport of molecules through the nanotube structures in a manner analogous to that of pores in biological membranes.[1] Additionally, the temperature threshold at which amino acid functional-group dynamics become anharmonic has been shown to be a property of the specific type of functional group involved, with implications for the interactions between particular functional groups and their chemical environments.[2] Some of these interactions can be observed as perturbations of dynamics that occur on a picosecond time scale, making spectroscopy in the THz frequency range an ideal probe. Recent research has shown that the motion of liquid water molecules at these frequencies is affected by the presence of dissolved amino acids in a manner that depends on the hydrophobicity of the solvated amino acid.[3−5]

IR-active intermolecular dynamics of crystalline amino acids occur in the 0.1−5 THz frequency range and may be measured directly using THz time-domain spectroscopy (THz-TDS).[6,7] The high sensitivity of this technique to the molecular interactions in a crystal lattice has found many uses, such as identifying polymorphs in pharmaceutical settings,[8] quantifying the ratio of substances in mixtures,[9] distinguishing between very similar cocrystalline structures,[10] and observing solid−solid phase transitions.[11] Differences in the lattice dynamics of racemic and enantiopure crystals of the same amino acid are readily observable using THz-TDS.[12,13] Polypeptides have also been studied using this method.[14,15] Other researchers have used THz-TDS to study the interactions between hydrophobic peptides and cocrystallized water.[16] Some THz-TDS experiments have recorded spectra as some external parameter is varied, such as the rotation angle of single crystals,[17] isotopic substitution,[18] or temperature.[19,20] Recent advances in THz-TDS polarimetry will begin to enable broadband measurement of vibrational circular dichroism at these frequencies.[21,22]

In addition to THz-TDS, the Raman shifts of intermolecular vibrations can be observed over a similar spectral range, although this requires measurements close to the Rayleigh line. Raman spectra of solid L-valine have been measured as a function of temperature between 17 and 300 K.[23] Measurements of low-frequency vibrations in L-valine at temperatures from 300 to 423 K have been made using both Raman and inelastic neutron scattering spectroscopy.[24] The pressure dependence of DL-valine Raman spectra has also been measured.[25] Raman spectroscopy has also been used to study the low-frequency vibrations of several amino acids (including L- and DL-valine) whose crystal lattice has been perturbed by exposure to γ radiation or the addition of metal dopants.[26]

Received: July 8, 2013
Revised: August 8, 2013
Published: August 9, 2013

Rather than relying on external perturbations, we chose to compare the intermolecular vibrations of similar systems that differ from one another in a limited number of structural aspects. The three systems studied are all racemates of either valine or leucine, whose chemical structures are illustrated in Figure 1. There are two polymorphs of solid-state DL-valine, the

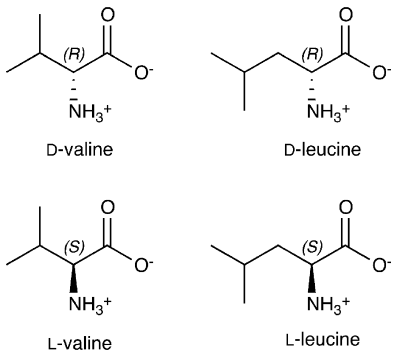

Figure 1. Chemical structures of leucine and valine. They are zwitterionic in the solid state, as depicted.

monoclinic and triclinic forms. Only one polymorph of DL-leucine is known to exist; it has a triclinic unit cell with an overall lattice structure that shares many features with the triclinic DL-valine polymorph. The crystallographic properties of these systems are listed in Table 1.

Several aliphatic amino acids, including valine and leucine, crystallize in layered structures with alternating regions of hydrophilic and hydrophobic interactions.[32] This pattern is illustrated in Figure 2A, where the regions of interactions between hydrophobic −R groups are highlighted in pale red and regions of hydrophilic interactions are highlighted in light blue. In other literature, these regions have sometimes been described as either layers or sheets. In the current work, however, we use the term "layer" to refer to a slab comprised of entire molecules as shown in Figure 2B where the alternating layers of molecules are highlighted in purple and orange. This is in contrast to slabs involving either hydrophobic or hydrophilic interaction regions such as illustrated in Figure 2A.

Two hydrogen bond patterns have been observed in racemic hydrophobic amino acid crystals and are described as the LD−LD and L1−D1 configurations.[33] In the LD−LD case, each layer of molecules contains both the L- and D-enantiomers. In structures with L1−D1 layers, only one enantiomer is present in each layer. Computational models indicate that the hydrogen bond network in LD−LD crystals is energetically preferred and that the L1−D1 layer structure is only observed where −R group steric considerations make the hydrogen bond network associated with L1−D1 layers the best available option.[32−34] This is the case for the −R groups of both valine and leucine

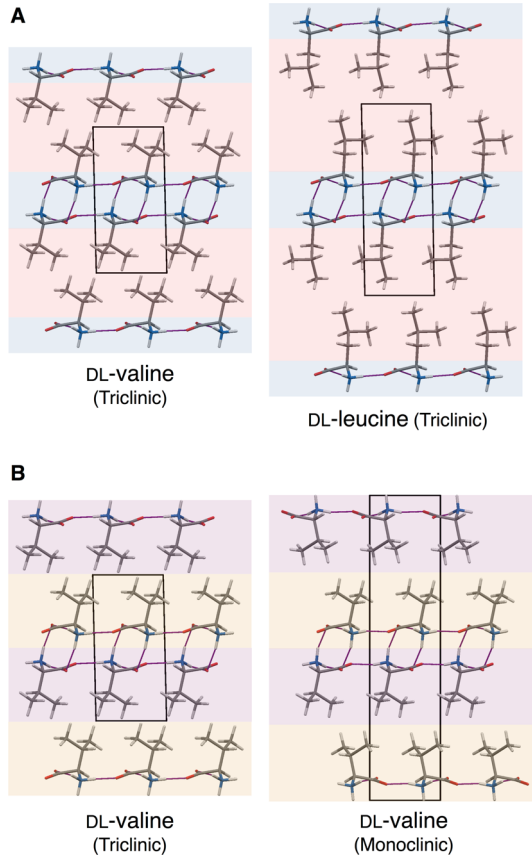

Figure 2. (A) Crystal structures of DL-leucine and the triclinic polymorph of DL-valine share the same hydrogen-bonding configuration and alternating regions of hydrophilic (highlighted in light blue) and hydrophobic (highlighted in pale red) interactions. (B) The triclinic and monoclinic DL-valine polymorph both consist of bilayers of the L-enantiomer (highlighted in purple) and D-enantiomer (highlighted in orange). However, the monoclinic polymorph has a screw axis symmetry element (oriented vertically in this perspective), resulting in a 180° rotation for every other bilayer.

(isopropyl and isobutyl, respectively), and as a consequence, each of the three systems discussed in the current work have molecules organized in the L1−D1 configuration. This can be seen in Figure 2B, where the layers highlighted in purple are made of L-valine molecules and those highlighted in orange are D-valine only.

The geometrical characteristics of the L1−D1 layers and their relationship to the unit cell parameters of these crystals are illustrated in Figure 3. Important atoms are labeled in Figure 3A. Each amino acid molecule has one ammonium group that acts as a donor of three hydrogen bonds. Two of these hydrogen bonds are within a layer of molecules, and the third

Table 1. Crystallographic Unit Cell Dimensions

	crystallographic properties			cell vector lengths (Å)			cell vector angles (deg)				
compound	lattice system	space group	Z	**a**	**b**	**c**	α	β	γ	temp.	CSD ID[a]
DL-leucine	triclinic	$P\bar{1}$	2	5.15	5.40	13.99	86.6	96.8	111.5	R.T.	DLLEUC03[27]
DL-valine	triclinic	$P\bar{1}$	2	5.23	5.42	10.83	90.8	92.3	110.0	100 K	VALIDL03[28]
	monoclinic	$P2_1/c$	4	5.21	5.41	22.10	90	90	109.2	R.T.	VALIDL[29]

[a]All crystal structures were retrieved from the Cambridge Structural Database.[30,31] The labels given to the unit cell vectors of the DL-leucine and monoclinic DL-valine structures have been redefined here to match the labeling system of the triclinic DL-valine structure for consistency and ease of comparison.

dx.doi.org/10.1021/jp406730a | J. Phys. Chem. B 2013, 117, 10444−10461

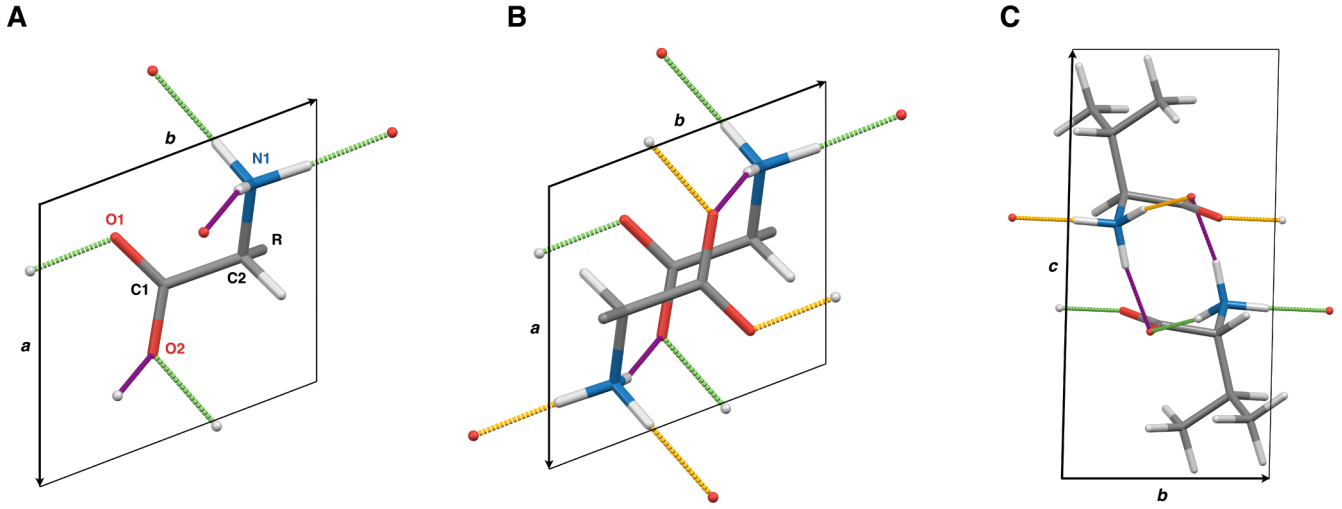

Figure 3. Unit cell of triclinic DL-valine used as an example to illustrate important geometric features shared by the three hydrophobic amino acid crystals studied. Molecules in these crystals are found in layers that are parallel to the **a**−**b** plane of the unit cell. Unit cell vector **c** is roughly perpendicular to this plane. Hydrogen bonds are depicted in green when they are located within the bottom layer of molecules in this drawing and purple when they connect to the layer above. Intralayer hydrogen bonds in the top layer of molecules are indicated in yellow-orange. (A) Only the molecule in the bottom layer is drawn. (B) Both molecules in the unit cell are now depicted. (C) The unit cell rotated 90° out of the page from the previous perspective; the view is now along cell vector **a**, and the entire −R group is illustrated.

joins the two layers. The layers are parallel to the **a**−**b** plane. Figure 3 shows the hydrogen bonds of the top layer of molecules in yellow-orange and the hydrogen bonds within the bottom layer in green. Bonds between layers are shown in purple. There are two oxygen atoms in each molecule. The atom labeled O1 accepts a single, intralayer hydrogen bond. Atom O2 accepts two hydrogen bonds; one is within the layer, and the other is between layers. The **a** and **b** unit cell vectors are largely determined by the hydrogen bond network, as can be seen in Figure 3B. On the other hand, the magnitude and orientation of vector **c** are primarily indicative of the optimal stacking orientation of the hydrophobic portions of the molecules, which are illustrated in Figure 3C.

The choice of molecular crystals that share the same hydrogen bond structure allows strategic comparisons to be made. Inspection of Figure 2A suggests that, by comparing the intermolecular vibrations of DL-leucine and the triclinic form of DL-valine, it may be possible to isolate effects that are due to the difference in −R groups. The comparison between the intermolecular vibrations of triclinic and monoclinic DL-valine allows for a somewhat more subtle difference to be investigated. As can be seen in Figure 2B, the same −R group (and in the same conformation) is present in both systems, and the hydrogen bond network is the same for both polymorphs. As a result, intralayer hydrophobic forces are probably very similar in both polymorphs, as well as the intra- and interlayer hydrophilic forces. This leaves the interlayer hydrophobic forces as the primary difference between these two systems.

EXPERIMENTAL METHODS

Recrystallization and Powder XRD. Both monoclinic and triclinic polymorphs of DL-valine are reported in the Cambridge Crystallographic database.[30,31] In a previous work,[20] we studied DL-valine which had been recrystallized by slow evaporation; this material was identified as the triclinic polymorph. However, it is also possible to work with the monoclinic polymorph at ambient pressure and a variety of temperatures. For the current work, we determined which polymorph is formed under a variety of crystallization scenarios using a Bruker-AXS D8 focus diffractometer to measure powder X-ray diffraction (XRD) spectra. The results are summarized below, and powder X-ray diffraction (XRD) spectra are found in the Supporting Information (SI).

1. Stock DL-valine (obtained from Fluka) is the monoclinic polymorph (Figure S1A).
2. Fast recrystallization of the stock DL-valine from aqueous solution (50 g/L) in an oven at 80 °C yields the monoclinic polymorph (Figure S1B).
3. Slow recrystallization of stock DL-valine from a dilute solution at room temperature yields the triclinic polymorph (Figure S1C). Concentrations between 2.0 g/L and 7.5 g/L were tested, all of which resulted in the formation of the triclinic polymorph when allowed to recrystallize slowly.
4. A solution created by mixing equivalent quantities (25 g/L each) of D-valine and L-valine (both obtained from Alfa-Aesar) always leads to the formation of the triclinic polymorph (Figure S1D), even when the recrystallization is carried out rapidly at 80 °C.

These experiments suggest that microscopic seed crystals of the monoclinic form of DL-valine persist in solution until a suitable level of dilution is reached. When the racemate was prepared by dissolving equal quantities of the pure enantiomers in water, the triclinic polymorph was obtained at all temperatures and concentrations investigated. This implies that recrystallizing a sample of triclinic DL-valine into the monoclinic polymorph would require adjusting an experimental parameter other than those investigated here, perhaps the identity of the solvent. The fact that commercially available DL-valine arrives in the monoclinic form must be due to the specific process(es) through which the material is synthesized, extracted, or purified. There are many examples in the literature of measurements involving solid DL-valine where the polymorph that was used is not specified. Our results indicate that research involving crystalline or even aqueous DL-valine should consider the polymorphism issue with care.

dx.doi.org/10.1021/jp406730a | J. Phys. Chem. B 2013, 117, 10444−10461

THz Time-Domain Spectroscopy. Recrystallized materials were pulverized using a ball mill for 3 min to obtain fine polycrystalline powders with average crystallite size significantly less than 50 μm. For the THz measurements, these powders were pressed into thin pellets with a 13 mm diameter die. Under 3.5 kbar of pressure, approximately 100 mg of powder yielded a pellet about 400 μm thick. The samples were mounted in a cryostat, and the propagation of THz light through the pellet and through a clear aperture was measured alternately. Reference and sample spectra were each obtained at room temperature and 80 K.

A detailed description of the THz spectrometer and the principles of its operation are available elsewhere.[20] The current measurements were made using a Ti:sapphire laser (KM Laboratories Griffin) as the source of ultrafast (~50 fs) optical pulses with a central wavelength of 800 nm. Half of this beam was used for THz light generation with an interdigitated photoconductive antenna (Batop iPCA-21-05-1000-800-h). A square wave (±10 V at 30 kHz) was used to bias the emitter as well as to provide a reference signal for lock-in detection. The other half of the optical beam was directed to an identical photoconductive antenna for THz detection. The electric field of the THz pulse was mapped over a 30 ps window in the time domain. In the frequency domain, these measurements have a resolution of approximately 0.035 THz. The amino acid pellets were not thick enough to allow truncation of THz pulses reflected inside the sample, and thus the absorption coefficients were determined by using a method that accounts for the reflections.[35] Representative time-domain reference and sample spectra are included in the SI.

Raman Spectroscopy. Raman measurements were made using an argon-ion laser (Spectra-Physics Stabilite 2017) operating at 488 nm. Samples consisted of polycrystalline material that was pulverized in the same manner as for the XRD and THz measurements. Scattered light was passed through a SPEX 1403 Ramalog double spectrometer and an attached SPEX 1442U third monochromator accessory. In this configuration, it was possible to measure spectral features with Raman shifts less than 0.5 THz (~15 cm^{-1}) from the Rayleigh line. Known laser plasma lines were used for final calibration of the frequency axis of the spectra.[36] These lines have been numerically removed from the reported measurements, but the unmodified spectra are included in the SI.

COMPUTATIONAL METHODS

Using theory to interpret the vibrational spectra of these systems requires an accurate model of the multidimensional potential energy surface. The need to include both hydrophobic and hydrophilic interactions makes the task particularly demanding. It is now *de rigueur* to model molecular crystals and their low-frequency vibrations using density functional theory (DFT) with infinite periodic boundary conditions, although this method is computationally expensive for systems with large numbers of atoms. It is sometimes possible to investigate noncovalent interactions in molecular crystals by calculating the properties of increasingly large clusters of their constituent molecules.[37] Periodic-boundary DFT calculations have previously been carried out on solid-state hydrophobic amino acids and polypeptides,[38−41] but those studies did not focus on low-frequency intermolecular vibrations. Accurately modeling these vibrations requires that the calculated optimized structure have extremely small residual forces, although it is possible to use DFT to calculate THz absorption spectra that closely match experiments.[20,42−46]

A particular challenge when modeling intermolecular vibrations of hydrophobic amino acids is including the effects of van der Waals forces that result from long-range electron correlation. Standard DFT methods do not account for these dispersion interactions. One approach to including them is DFT-D, which adds an empirically determined scalar correction term to standard DFT calculations. This method has recently been used to model THz spectra of molecular crystals with good agreement.[47] A second approach, which is available in the SIESTA DFT software package[48,49] that was used for our calculations, is to employ an exchange-correlation functional that accounts for van der Waals forces within the framework of DFT. This functional, known as vdW-DF, was first implemented for sheets and slabs,[50−52] but it has more recently been modified for use with systems of any geometry.[53]

Unlike many other methods of incorporating dispersion interactions in DFT, vdW-DF is intended to be nonempirical. It uses a standard generalized gradient approximation (GGA) functional to calculate electron exchange energy, and the local part of the correlation energy is calculated using the local density approximation (LDA). The nonlocal correlation energy is calculated through a function that depends on the electron densities, the density gradients, and the spatial separation of pairs of points in the system being modeled. This allows a self-consistent calculation of the nonlocal correlation energy as a function of the electron density. The overall vdW-DF functional is therefore an approximation of the true exchange-correlation functional with—in principle—no need for corrections from outside of density functional theory.[54] In weakly bound systems that are not well-modeled with standard DFT methods, the vdW-DF functional has been shown to have a level of accuracy on par not only with DFT-D but ab initio methods such as MP2 and CCSD(T) as well.[55] While the computational cost of MP2, CCSD(T), and other post-Hartree−Fock methods increases very rapidly for larger systems, vdW-DF scales as $O(N^3)$, which is the same scaling order as most traditional DFT calculations.[54] The version of the SIESTA DFT software used in our calculations has an efficient implementation of vdW-DF, and the increase in computational cost relative to a standard DFT calculation is minimal.[56] A comparison of a large variety of methods incorporating van der Waals interactions in DFT has recently been published.[57]

The original version of vdW-DF uses the revPBE (revised Perdew−Burke−Ernzerhof) functional to calculate the exchange energy of a system. In practice, however, the use of other exchange functionals has been shown to increase the accuracy of calculations on a variety of systems.[58] Therefore, we chose a variation of vdW-DF advocated by Klimeš et al.,[58] which uses an underlying exchange functional known as optB88. While the term vdW-DF might technically refer only to the original revPBE-based version, for convenience we use the term to refer to the optB88 variant as well. In addition, to be consistent with the terminology used by the developers of vdW-DF, we refer to the quantum-mechanical, long-range electron correlation as van der Waals interactions, but they are more precisely described as London dispersion interactions.

Calculations were carried out using version 3.0 (revision 367) of SIESTA.[48] The Atomic Simulation Environment (ASE)[59] package of Python programming modules was used for writing scripts for pre- and postprocessing the DFT calculations. SIESTA uses numerical atomic orbitals as a basis

set, and we used a double-ζ, double-polarized (DZDP) basis for all calculations. In these calculations, the wave functions and electron density are projected onto a real-space grid. A cutoff energy of 1000 Ry was chosen for this grid, corresponding to a separation between grid points of roughly 0.05 Å and a total number of points on the order of 1×10^6 for the systems being modeled.

The initial coordinates used in the calculations were obtained from the Cambridge Crystallographic Database.[30,31] The specific database entries that were used are listed in Table 1. The first stage of each calculation was the optimization of these atomic positions to find the configuration with minimum interatomic forces. In all of our calculations, the unit cell geometry was optimized concurrently with the atomic positions. We compiled and ran SIESTA on clusters of Intel Xeon 5400 or 5500 series processors at the Yale High Performance Computing Center. A typical geometry optimization calculation ran on 64 or 128 CPU cores and required roughly 24−48 h of wall time (actual time elapsed) to locate the optimal geometry using the modified Broyden's quasi-Newton−Raphson algorithm.[60] The convergence conditions were: (1) the maximum force experienced by any atom is less than 0.002 eV/Å and (2) the maximum unit cell stress tensor element is less than 1.0 MPa (0.01 kbar, ∼10 atm) in magnitude. Example input files are included in the SI.

Calculation of Vibrational Mode Frequencies and IR Intensities. In the current work, the harmonic vibrational modes of each system were determined using the finite-difference method of calculating the force-constant matrix.[61] Each atom in the unit cell was individually displaced from its equilibrium position in the positive and negative direction along each Cartesian axis by 0.02 Å while leaving the other atoms fixed. Repeating the calculations with displacements between 0.01 Å and 0.04 Å yielded essentially identical results.

In addition to determining the force constants, the macroscopic polarization of the system was calculated using the Berry phase approach.[62] This allows for the calculation of the Born effective charge tensor for each atom[61]

$$Z^*_{ij,\tau} = V\frac{\partial P_i}{\partial r_{j,\tau}} \tag{1}$$

where P_i is the macroscopic polarization in the i direction when atom τ undergoes displacement r along the j axis in a system with unit cell volume V. (The asterisk in this notation is to indicate that the quantity is an effective charge, not that it is a complex conjugate.) The Born effective charge tensor equivalently represents[63]

$$Z^*_{ij,\tau} = \frac{\partial F_{j,\tau}}{\partial E_i} \tag{2}$$

where F_j is the force experienced in the j direction by atom τ in the presence of E_i, a macroscopic electric field in the i direction.

The connection between the Born effective charge tensors and the IR activity of a given mode is as follows. If the normal-mode coordinate of vibrational mode m is denoted Q_m, then the IR activity of the mode, I_m, can be expressed as[61]

$$I_m \propto \sum_{i=1}^{3}\left(\frac{\partial P_i}{\partial Q_m}\right)^2 \tag{3}$$

where i again refers to one of the three Cartesian directions. The change in polarization as displacement occurs along this normal-mode coordinate is[64]

$$\frac{\partial P_i}{\partial Q_m} = \sum_{k=1}^{3N}\frac{\partial P_i}{\partial r_k}X_{k,m} \tag{4}$$

where $X_{k,m}$ represents the normal mode m in terms of each of its $3N$ Cartesian displacement coordinates r_k. This can equivalently be expressed as

$$\frac{\partial P_i}{\partial Q_m} = \sum_{j=1}^{3}\sum_{\tau=1}^{N}\frac{\partial P_i}{\partial r_{j,\tau}}X_{j\tau,m} \tag{5}$$

where j is a Cartesian direction and τ is one of the atoms in the unit cell. (The Cartesian displacements of the normal mode have been reshaped from the $3N \times 1$ array $X_{k,m}$ to the $N \times 3$ array $X_{j\tau,m}$.) Combining eq 5 with eq 3, and then employing the definition of the Born effective charge tensor (eq 1), yields the relative IR absorption intensities[61]

$$I_m \propto \sum_{i=1}^{3}\left(\sum_{j=1}^{3}\sum_{\tau=1}^{N}Z^*_{ij,\tau}X_{j\tau,m}\right)^2 \tag{6}$$

For computational efficiency, the calculation of the force constants and polarizations was divided into N jobs executed in parallel (where N is the number of atoms per unit cell); each job calculated the force constants and Born effective charge tensor for a particular atom. For DL-leucine and the triclinic form of DL-valine (44 and 38 atoms per cell, respectively) each of these calculations were typically carried out on eight processors and required roughly 3 h of wall time. In the case of monoclinic DL-valine (76 atoms per cell), each job typically ran on 32 processors and also required approximately 3 h of wall time. The total computational cost (optimization and IR-activity steps) of the calculations was approximately 0.5 CPU-years for the triclinic systems and 1.5 CPU-years for the larger monoclinic system. Each calculation was carried out several times to check for convergence of various parameters.

RESULTS

Spectroscopic Measurements. The THz absorption spectra are reported in Figure 4. The ordinate is the Naperian power absorption coefficient, which is conventionally reported on a per-centimeter basis. The features observed are listed in Table 2. The frequency, fwhm, and amplitude of peaks were obtained using multipeak fitting with Lorentzian lineshapes. When fitting THz-TDS spectra, an extra Lorentzian was included to model the strong absorptions that are partly visible at the upper limit of the usable frequency range of the spectrometer. This approach also accounted for the rising absorption present in the room temperature spectra.

At 80 K (thick lines in Figure 4), both DL-valine polymorphs have an absorption peak at 1.73 THz. In the case of the triclinic polymorph, the frequency of the peak has blueshifted slightly from a room temperature value of 1.70 THz, but in the monoclinic system, the feature seems to have actually redshifted from 1.75 THz. Monoclinic DL-valine has an additional peak at 0.85 THz in the room temperature spectrum that blueshifts to 0.90 THz at 80 K. The room temperature absorption spectrum of DL-leucine has peaks at 1.37 and 1.87 THz as well as a broader absorption at approximately 2.25 THz. When the sample is cooled to 80 K, all of these features

dx.doi.org/10.1021/jp406730a | J. Phys. Chem. B 2013, 117, 10444−10461

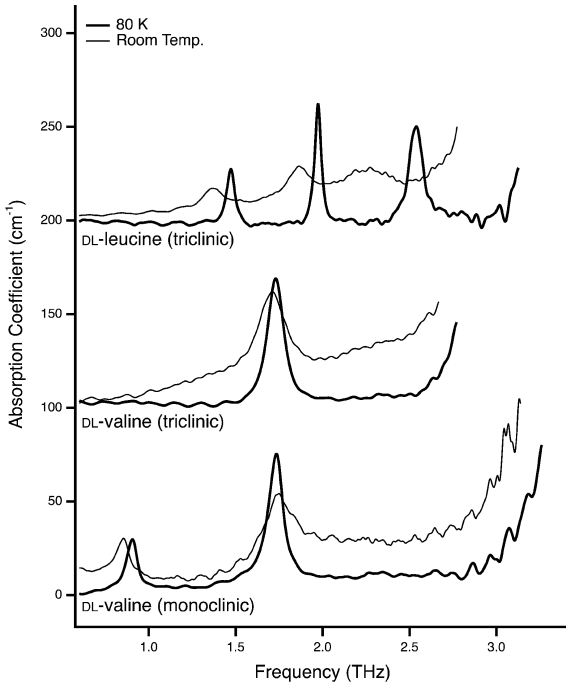

Figure 4. THz absorption spectra of DL-leucine and both DL-valine polymorphs. The DL-leucine and triclinic DL-valine spectra are vertically offset (1 THz ~ 33 cm^{-1}).

blueshift and narrow. In all of the THz spectra, there are indications of very strong absorptions that occur at frequencies just above the upper limit of the usable experimental bandwidth. As the temperature is reduced, these absorptions all appear to blueshift, sharpen, or both.

The Raman spectra of DL-leucine and both DL-valine polymorphs were measured at room temperature over the range 0.5−7.0 THz (~15−235 cm^{-1}) and are shown in Figure 5. The spectra of the two DL-valine polymorphs are remarkably similar aside from the extra peak (at 1.19 THz) in the spectrum of the monoclinic system. The three lowest-frequency modes in the DL-leucine spectrum appear in a similar pattern to those in the triclinic valine spectrum, only they are shifted to lower frequencies. Observed features are listed in Table 2. Each Raman spectrum was fit using a set of Lorentzian functions, with excellent agreement. A Lorentzian function centered at 0 THz was included to account for the Rayleigh scattering signal. An example of the fitting procedure is available in the SI.

Optimized Unit Cell Geometries. An accurate DFT model should reproduce the unit cell parameters of the system being studied. However, many DFT studies of hydrophobic amino acids find that the unit cell vector normal to the layers of molecules is substantially longer than the crystallographic value, which reflects the absence of van der Waals interactions in those models. In the current work, the calculated unit cell of each of the three systems agrees extremely well with the crystallographic data. The **a** and **b** vectors determined in each vdW-DF calculation are within a few hundredths of an angstrom of the experimental values (Table 3). The **c** vector of each of the three systems is actually shorter than the crystallographic value. However, this makes sense given that the calculations are a model of the system at 0 K. In the Supporting Information of our previous work[20] we noted that for L-valine unit cell, which also has a layer structure, most of the decrease in unit cell size at low temperature occurs along the vector normal to these layers, which would be analogous to the **c** vector of the systems here.

It is worth pointing out that symmetry was *not* enforced in any of our calculations. For example, the monoclinic form of DL-valine was not constrained to have a unit cell with two 90° angles, but the calculation nonetheless locates an optimal geometry that retains the experimental symmetry.

Calculated Vibrational Frequencies and IR Intensities. The computational model accurately reproduces the number of low-frequency IR-active peaks and their relative intensities in each of the three systems, as illustrated in Figure 5A−C. It is not practically feasible to calculate the Raman activity of the vibrational modes using the methods at hand. However, each of

Table 2. Spectral Features Observed in the THz-TDS (IR Absorption) and Raman Scattering Measurements

	DL-valine (monoclinic)			DL-valine (triclinic)			DL-leucine (triclinic)		
measurement	frequency[a]	fwhm	amplitude[b]	frequency	fwhm	amplitude	frequency	fwhm	amplitude
IR absorption (at 80 K)	0.90	0.07	27.5				1.47	0.05	29.5
	1.73	0.11	69.5	1.73	0.10	68.2	1.97	0.04	62.0
							2.54	0.09	50.6
IR absorption (room temp.)	0.85	0.09	22.7				1.37	0.15	17.3
	1.75	0.31	37.9	1.70	0.24	46.0	1.87	0.20	29.4
							2.25	0.31	28.0
Raman scattering (room temp.)	1.19	0.11	49.2				1.13	0.15	100.0
	1.37	0.10	100.0	1.38	0.18	100.0	1.39	0.21	48.1
	1.78	0.13	32.4	1.79	0.22	35.7	1.64	0.24	77.5
	2.02	0.11	29.4	2.09	0.21	34.7	2.39	0.54	45.1
	3.00	0.19	23.8	3.13	0.34	64.1	2.66	0.24	34.1
	3.39	0.49	41.1	3.48	0.44	22.7	3.69	0.53	37.6
	4.07	0.57	23.9	4.00	0.55	42.1	4.11	0.43	10.8
	5.22	0.42	15.5	5.20	0.44	22.1	4.99	0.43	13.8
	6.35	0.39	9.7	6.38	0.38	13.2	6.34	0.58	12.9

[a]Frequency and fwhm of peaks are in units of THz. [b]The IR power absorption coefficient has units of cm^{-1} (Naperian absorption per cm, not to be confused with wavenumbers), while the Raman scattering intensities are scaled relative to the strongest peak in each spectrum.

dx.doi.org/10.1021/jp406730a | J. Phys. Chem. B 2013, 117, 10444−10461

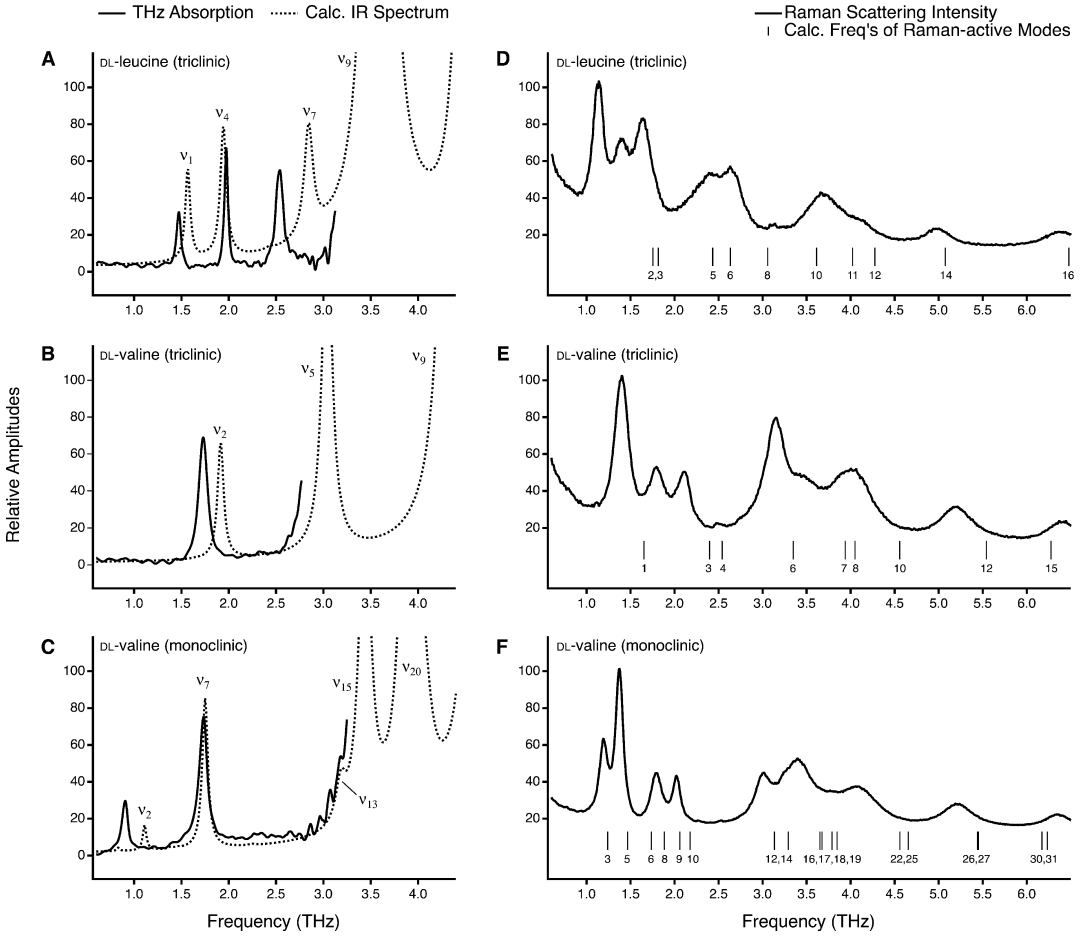

Figure 5. (A–C) Calculated IR absorption spectra (dashed lines) compared to the THz-TDS measurements (solid lines) at 80 K. (D–F) Room-temperature Raman spectra (solid lines) and calculated vibrational frequencies for Raman active modes (thin bars, beneath which are the mode numbers). All calculated frequencies have been scaled by a factor of 0.9.

Table 3. Comparison of Calculated and Experimental (in Bold) Unit Cell Dimensions[a]

		cell vector lengths (Å)			cell vector angles (deg)				
compound	DFT model and software	a	b	c	α	β	γ	vol. (Å^3)	temp.
DL-leucine (triclinic)	vdW-DF (SIESTA)	5.07	5.43	13.54	88.71	95.56	112.85	341.9	0 K
		5.15	**5.40**	**13.99**	**86.58**	**96.76**	**111.47**	**359.6**	**R.T.**
DL-valine (triclinic)	PBE (GAUSSIAN)[b]	5.31	5.46	11.66	90.97	93.03	110.10	317.2	0 K
	PBE (SIESTA)	5.41	5.46	11.69	92.17	92.03	110.17	323.7	0 K
	vdW-DF (SIESTA)	5.20	5.45	10.59	90.67	91.55	110.74	280.6	0 K
		5.23	**5.42**	**10.83**	**90.83**	**92.29**	**110.01**	**287.8**	**100 K**
DL-valine (monoclinic)	vdW-DF (SIESTA)	5.17	5.45	21.48	90.00	90.00	110.37	567.8	0 K
		5.21	**5.41**	**22.1**	**90**	**90**	**109.2**	**588.3**	**R.T.**

[a]The Cambridge Spectral Database (CSD) IDs of the experimental structures are given in Table 1. To distinguish among DFT calculations, some specifics of the particular model and software used are included in parentheses. [b]The results of this calculation were reported in our previous work[20] and are included here for comparison.

the crystal systems studied is centrosymmetric, meaning that all vibrational modes are expected to be either IR-active or Raman-active but not both. Therefore, in Figure 5D–F, the frequencies of the Raman-active modes are indicated with thin black bars below each spectrum.

The calculated vibrational frequencies are all higher than observed, so they have been scaled by a factor of 0.9 in Figure 5 and elsewhere. Linear scaling factors are often used to account for overestimation of vibrational frequencies in DFT, including in modeled THz spectra.[65] The origin of the overestimation is usually presumed to be anharmonicity that is not accounted for by the calculation. However, although a frequency-scaling factor may improve the apparent agreement with experimental spectra, the anharmonicity and temperature dependence of intermolecular vibrations in molecular crystals is mode-specific. There is no reason to expect that a single scaling factor will be appropriate for all vibrations. Instead, analyzing the underlying motion of the calculated vibrational modes must be a

dx.doi.org/10.1021/jp406730a | J. Phys. Chem. B 2013, 117, 10444–10461

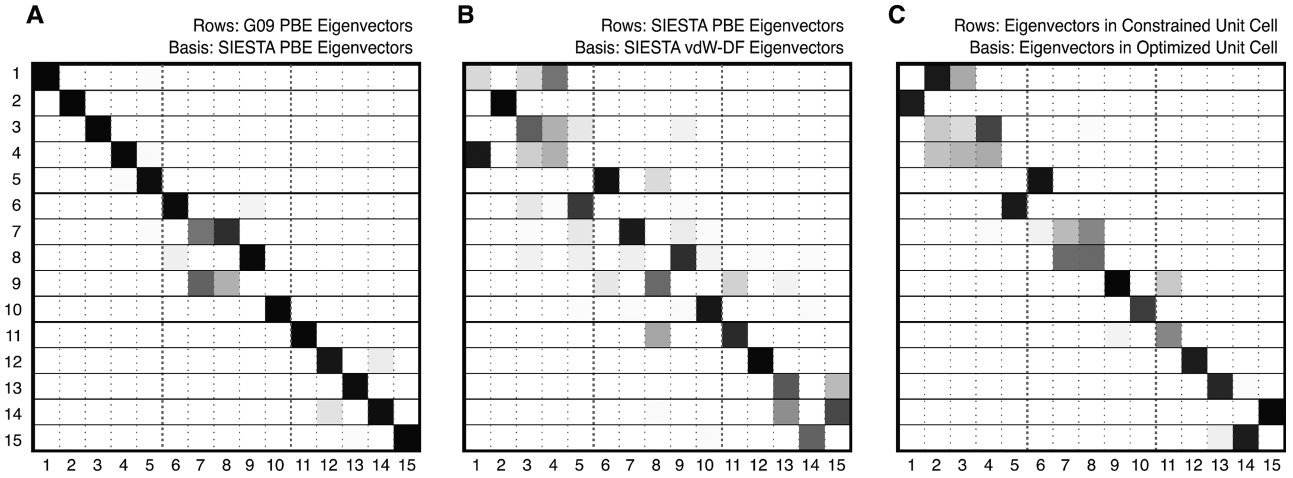

Figure 6. By projecting the set of eigenvectors from one calculation into a basis comprised of the set of eigenvectors from a second calculation of the same system (triclinic DL-valine in these examples), one can quickly obtain a quantitative measure of the salient differences between motions predicted by different models. This is more informative than a comparison of, for instance, calculated absorption spectra alone. In each figure above, every row is an eigenvector from one calculation represented in the normalized basis set of modes from a second calculation. (A) Similar vibrational modes are predicted by two different DFT software packages when the same exchange-correlation functional (PBE) is used, despite dissimilarities in the implementation of the theory. (B) A computational model (vdW-DF) that accounts for van der Waals dispersion forces predicts some of the same low-frequency modes as one that lacks these forces (PBE); however, several of the modes differ in terms of their actual motion, not just the calculated vibrational frequency or IR absorption strength. (C) The decision to constrain the unit cell geometry to the crystallographic values rather than allow them to be optimized as part of the calculation can result in changes to the character of the predicted vibrations, not merely their frequencies.

prerequisite to evaluating the accuracy of a calculation and making peak assignments. The first step in this process is to obtain useful descriptions of these modes, which becomes difficult when they involve several dozen atoms and have a combination of intra- and intermolecular character. Methods of addressing this problem are described below.

Eigenvector Projections. Since there exists a tremendous number of DFT variations and other computational methods, it is very important to have the ability to compare results from different calculations. However, it is not always clear what the most informative comparison may be. Factors to consider include similarity of unit cell parameters, relative deviations of optimized atomic positions from crystallographic values, comparison of vibrational frequencies and intensities, and so on. We propose a new method of comparison, one that is quantitative, but also visual and concise.

Since the goal is to model vibrational modes, the most relevant objects to compare are the dynamical matrices obtained from the various calculations. While the common practice of comparing the normal-mode frequencies of two or more calculations reveals some of this information, these scalar quantities do not provide insight into the actual character of the underlying motions. The greater part of the dynamical information is contained in the eigenvectors. One way to use this information in a comparison is to project the eigenvectors of one calculation into the basis of eigenvectors from a second calculation. To some extent, the resulting matrix is similar to the Duschinsky rotation matrix $\mathbf{J}$ in the expression $\mathbf{Q} = \mathbf{J}\mathbf{Q}' + \mathbf{K}$,[66] where $\mathbf{Q}$ and $\mathbf{Q}'$ are the normal modes of a system in two different electronic states and $\mathbf{K}$ is a displacement vector. In this case, $\mathbf{J}$ is the transformation matrix that describes the linear combinations of vibrations needed to connect $\mathbf{Q}$ and $\mathbf{Q}'$.

In the present context of comparing eigenvectors from two different calculations, we consider a matrix in which each row is a normal mode of one calculation expressed in the basis of the normal modes of the other calculation. Each of the grids in Figure 6 represents the "$\mathbf{J}$ matrix" that connects two different sets of eigenvectors. They are arrays of row vectors that represent one set of eigenvectors as a series of coefficients of eigenvectors from a different calculation. The rows are normalized such that the sum of squares of the elements equals 1 and each square on the grid is shaded such that white corresponds to 0 and black to 1. Any difference in equilibrium coordinates of the atoms in each calculation is minimized by optimally rotating one set of coordinates on to the other. Therefore, we do not explicitly consider any $\mathbf{K}$ vector in this procedure. Although the current application is simply a concise method of comparing results from different calculations, it might also be possible to purposely vary some calculation parameter so that $\mathbf{J}$ has a well-defined physical interpretation.

As an example of the method, in Figure 6A the eigenvectors of the triclinic DL-valine calculated in our previous work[20] using Gaussian software and the PBE exchange−correlation functional are represented in the basis of eigenvectors determined with a new calculation with SIESTA, also using the PBE functional (i.e., no van der Waals forces are included in either calculation). Diagonal elements indicate that nearly identical modes appear in both calculations and in the same frequency order. Elements that are off diagonal yet black indicate that the same mode occurs in both calculations but the frequency order differs (for example, mode 8 in the Gaussian calculation corresponds to mode 9 in the SIESTA one).

The gray elements near the middle of the grid in Figure 6A indicate that a description of modes 7 or 9 from the Gaussian calculation requires a linear combination of eigenvectors 7 and 8 from the SIESTA calculation. The fact that these modes can be well-described using only two modes from the basis-set eigenvectors indicates that the great majority of the underlying motion of these modes occur in both calculations. In general, Figure 6A reveals that, despite the fact that SIESTA employs pseudopotentials instead of explicit core electrons as well as a

dx.doi.org/10.1021/jp406730a | J. Phys. Chem. B 2013, 117, 10444−10461

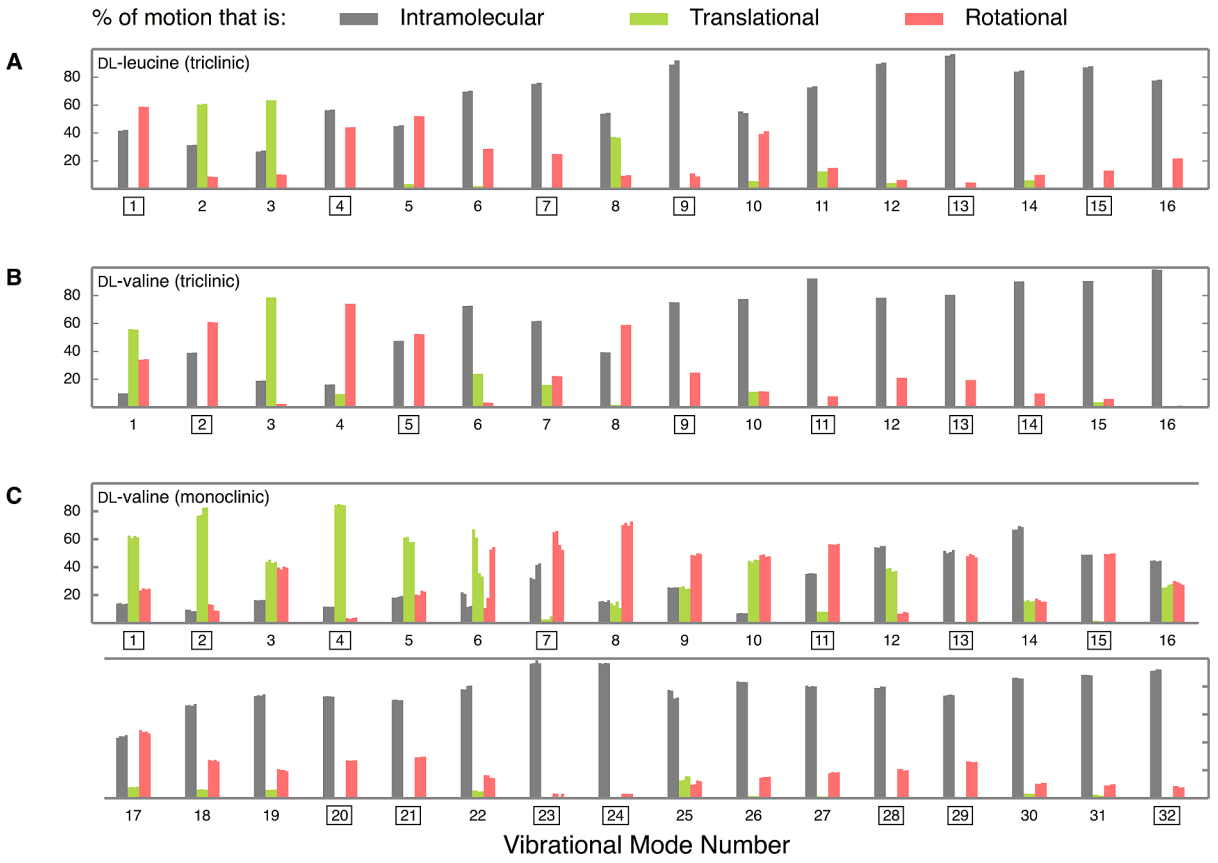

Figure 7. Character of the lowest-frequency vibrational modes in solid-state DL-leucine and the two DL-valine polymorphs quantified in terms of translational, rotational, and intramolecular motion. IR-active vibrations are indicated with a box around the mode number. For a given mode, the motion of each molecule in the unit cell is indicated with its own set of bars: four for monoclinic DL-valine and two for the triclinic systems. Since monoclinic DL-valine has twice as many molecules per unit cell, it also has approximately twice as many vibrational modes over the same frequency range.

different kind and number of basis functions, the two sets of eigenvectors obtained are quite similar.

Having established that SIESTA largely reproduces the Gaussian results when both use the PBE functional, we next compare the results from the SIESTA PBE calculation with the SIESTA vdW-DF calculation. As can be seen in Figure 6B, some of the same modes appear in both calculations: Modes 2, 10, and 12 for example. But other vibrations (such as PBE modes 1 and 3) can only be described in the vdW-DF basis by using a linear combination of many eigenvectors. This illustrates a much more important difference than would be revealed by comparing the two calculated IR spectra or vibrational frequencies. While some modes occurring in both calculations have merely shifted in terms of relative frequency, other modes are unique to the particular computational model.

In a previous paper[20] we showed that a calculation using a unit cell constrained to the crystallographic values, which are often measured at temperatures far from 0 K, yields vibrations that do not exist in the set of modes from a calculation using an optimized unit cell (atomic positions within the unit cell are optimized in both cases). An unoptimized unit cell calculation effectively represents the system under some arbitrary, anisotropic external stress; the structure is not at equilibrium within the physics of the theory being employed. The stress experienced by the unoptimized unit cell depends on several factors, such as how much the experimental unit cell vectors differ from what their values would be at 0 K or the inadequacies of the theoretical model. Even a small amount of stress will be significant in a situation where many vibrational modes occur with frequencies that differ by only a few cm^{-1} (tens of GHz). Figure 6C underscores this point by projecting the eigenvectors from a fixed-cell calculation into the basis of eigenvectors from the relaxed cell calculation (both sets of eigenvectors are those obtained using Gaussian in our previous work).[20] The eigenvectors from all calculations mentioned are included in the SI.

DISCUSSION

Descriptions of Intermolecular Vibrations. The collective motions of crystalline molecules are often referred to as lattice modes or external vibrations, but in the broadest terms they are phonons. A phonon is a quasi-particle with a wavevector $\boldsymbol{k}$ that describes the relative phase of atomic displacements in adjacent unit cells. The vibrations we have modeled correspond to the subset of phonons with $\boldsymbol{k} = 0$; in other words, the displacements in all unit cells are in phase, which is a requirement for most optical measurements. However, we use the term "intermolecular vibration" because it highlights the molecular aspect of the discussion and also makes it easier to contrast motions that primarily involve intramolecular displacements. The molecules in the systems that we studied are not rigid, and thus there is no sharp distinction between vibrations that are purely intermolecular and purely intramolecular. In spite of this, we find the terminology of rigid motion to be useful even in describing

dx.doi.org/10.1021/jp406730a | J. Phys. Chem. B 2013, 117, 10444−10461

vibrations that have substantial intramolecular components. This somewhat surprising observation is discussed in detail below.

In the rigid molecule approximation, intermolecular vibrational modes take the form of librations (hindered rotations) and center-of-mass translations. Librational modes can be thought of in terms of a molecule in a potential well of the form $V(\phi)$, where ϕ is the angle of rotation around a particular axis. The specific form of this potential is determined by the intermolecular forces involved.[67] For naphthalene crystals under the rigid molecule approximation, it has been estimated that molecules librate through approximately 4° and have center-of-mass displacements of roughly 0.02 Å at room temperature.[68]

Factor Group Analysis. The basic selection rules for intermolecular vibrational modes in molecular crystals can be determined using factor group analysis.[69,70] The space group of DL-leucine and the triclinic form of DL-valine is $P\bar{1}$ (C_i^1 in Schoenflies notation; space group number 2), and both of these systems have two C_1 molecules located at sites that are related by inversion. This configuration leads to $3A_g + 3A_u$ translational modes, and there are another $3A_g + 3A_u$ torsional modes derived from the rotational degrees of freedom. The three A_u translations correspond to displacements of both molecules in the same direction simultaneously. These are acoustic phonon modes that have zero frequency at $\boldsymbol{k} = 0$. Therefore, the remaining relevant intermolecular vibrations are $3A_u + 6A_g$. The $3A_u$ modes derive from the rotational degrees of freedom, and there are $3A_g$ translational and $3A_g$ rotational modes. A_u modes are IR-active and A_g modes Raman-active.

The space group of the monoclinic DL-valine system is $P2_1/c$ (C_{2h}^5 in Schoenflies notation; space group number 14). There are four sites that are filled using symmetry operations on a single C_1 molecule. This configuration results in translational representations $3A_g + 3A_u + 3B_g + 3B_u$. The rotational degrees of freedom also have representations $3A_g + 3A_u + 3B_g + 3B_u$. In this case, the acoustic modes are $A_u + 2B_u$, leaving $6A_g + 5A_u + 6B_g + 4B_u$. Since the $P2_1/c$ system has an inversion center, the A_u and B_u modes are IR-active, and the A_g and B_g modes are Raman-active. (The "B" indicates that the screw-axis translation operation has a character of −1 instead of 1, which is its value for the A_g and A_u modes.)

While factor group analysis of these systems provides general predictions about their IR and Raman spectra, these alone are not enough to ensure that peak assignments can be made rigorously. For example, although factor group analysis dictates that a certain number of intermolecular modes will be IR or Raman-active, there is no reason that some of the lower-frequency intramolecular modes cannot occur in the same spectral range as these modes. This is further complicated by the fact that molecules are not actually rigid and the resulting vibrational modes are a mixture of inter- and intramolecular motions. To address the difficulty of making peak assignments, as well as to find quantitative descriptions of the calculated vibrational motions, we have developed the methods described below to analyze calculated results as well as to effectively compare results obtained from different models.

Quantifying the Nature of Complicated Vibrations. In addition to comparing results from different theoretical models as described above, there exists a need to succinctly characterize aspects of the vibrational modes beyond their frequencies and intensities. In Figure 7, the motion of the calculated low-frequency vibrations of DL-valine and DL-leucine are depicted in the manner described in our previous work.[20] The IR-active vibrations are indicated with a box around the mode number, and the other modes are Raman-active based on the symmetry of the systems. In both DL-leucine and triclinic DL-valine, the lowest-frequency IR-active modes all have significant intramolecular character (gray bars), and the intermolecular character is entirely librational (orange bars). The intermolecular character of the low-frequency Raman-active modes consists of translational (green bars) as well as librational displacements.

For DL-leucine (Figure 7A), the modes that are IR-active have no translational component, as expected. Any translational motion that is antisymmetric with respect to inversion results in a net center-of-mass displacement for the unit cell. In the simplest picture, the three intermolecular IR-active modes would be the antisymmetric rigid librations. However, even the lowest-frequency vibration (mode 1) has ∼40% intramolecular character, and the intramolecular component is even larger for the next IR-active vibrations, modes 4 and 7.

Based on factor group symmetry, one would expect DL-leucine to have three translational and three librational Raman-active modes. As can be seen in Figure 7A, DL-leucine modes 2 and 3 have very similar overall character, the largest component of which is translational. Figure 5D indicates that the calculated frequencies of these vibrations are also quite close (1.75 and 1.82 THz). The third vibration with notable translational motion (mode 8) occurs at higher frequency than the first two (3.06 THz) and has a much larger intramolecular component. This leaves the three expected librational Raman-active modes to consider. There are, in fact, three vibrations that fit this description to a certain extent: modes 5, 6, and 10. Interestingly, mode 6 has more intramolecular character than mode 10, despite occurring at significantly lower frequency.

Triclinic DL-valine, having a unit cell analogous to DL-leucine, might be expected to have roughly the same set of vibrations. However, this is only partly evident in Figure 7B. The IR-active modes again have no translational character. Mode 2 of triclinic DL-valine has a ratio of intramolecular and rotational motion that is very similar to mode 1 of DL-leucine. This is slightly less true for triclinic DL-valine mode 5 and DL-leucine mode 4. Mode 9 of triclinic DL-valine looks very similar in character to DL-leucine mode 7, although the IR activities of these two modes (Figure 5A,B) are very different.

Like DL-leucine, the two lowest-frequency Raman-active modes of triclinic DL-valine are predominantly translational. However, mode 1 of triclinic DL-valine has a substantial rotational component, while mode 3 is almost entirely translational. The third translational mode expected from factor group theory is not apparent in Figure 7B. Instead, modes 4, 6, 7, and 10 all have modest translational components. A similar situation exists for the librational modes: two vibrations (modes 4 and 8) have large rotational components, but there is no obvious third Raman-active mode that could be considered substantially librational.

In addition to comparing the behavior of the slightly different molecules in the two triclinic systems, it is also informative to compare the character of modes in the two DL-valine polymorphs (Figure 7B,C). Because monoclinic DL-valine has four molecules in the unit cell instead of two, it is possible for that system to have antisymmetric translational displacements that are not acoustic (i.e., do not have a net center-of-mass displacement). This can be seen in Figure 7C, where modes 1, 2, and 4 are predominantly translational but have IR-active

symmetry. However, only mode 2 has significant calculated IR intensity; this agrees with the experimental THz absorption spectrum in Figure 5C.

In modes 6 and 7 of monoclinic DL-valine, not every molecule in the unit cell appears to have exactly the same type of motion, unlike all of the other vibrations depicted in Figure 7. This is most likely a result of their nearly identical calculated frequencies (1.74 and 1.75 THz) and the fact that no symmetry was enforced during the calculation of the force constant matrix. Ignoring this slight mixing, mode 7 in the monoclinic system has a combination of intramolecular and librational character that is similar to that of mode 2 of triclinic DL-valine. These modes also have very similar calculated frequencies (and appear to coincide with absorptions in the experimental spectra in Figure 5B,C). Modes 13 and 15 of monoclinic DL-valine both appear to be similar in character to triclinic DL-valine mode 5, although the calculated IR activity of mode 15 (monoclinic) is a much closer match.

New Method of Quantifying Vibrational Character. The quantified character of the vibrations shown in Figure 7 allows for general comparisons to be made among modes, but deeper analysis requires additional details. For example, it is possible for two modes (such as triclinic DL-valine mode 9 and DL-leucine mode 7) to have similar percentages of the different types of motion but actually involve different degrees of freedom. This is especially true given the large variety of intramolecular motions that are possible.

To make comparisons between different molecular systems or between results of different calculations, it is useful to have information that is as specific as possible without being impractical or esoteric. In what follows, we describe an approach that allows the presentation of detailed vibrational character that is not only quantitatively rich but also concise and intuitive. There are other methods of visualizing various kinds of information about interactions in molecular crystals,[71,72] and there also exists an efficient method for coarse-grained characterization of intermolecular vibrations[73,74] and force constants.[75] An advantage to our approach is that it allows one to analyze the vibrations that simultaneously involve a large number of inter- and intramolecular degrees of freedom. Additionally, our approach can be applied regardless of the method or model used to obtain the vibrational eigenvectors. Since the discussion that follows will involve specific intramolecular coordinates, the atom numbers used to define these coordinates are labeled in Figure 8.

In the new quantification procedure, the character of each mode is illustrated in a horizontal bar with several segments, each corresponding to different types of motion. For example, in Figure 9, the specific properties of the calculated low-frequency vibrations of triclinic DL-valine are depicted. We refer to one of these bars as a mode's "vibrational character ID strip," or more briefly, its "ID strip". Calculated separately for each molecule in the unit cell, the ID strip of a particular mode consists of:

1. The x, y, and z components of any center of mass translational displacement, labeled T_x, T_y, and T_z.
2. The projection of the axis of rotation (for librational displacements) onto the three principal axes of the molecule, denoted R_A, R_B, and R_C.
3. The relative displacements of atoms along selected, important internal coordinates. In the current case, these are limited to relevant bond angles and dihedral angles (torsions).

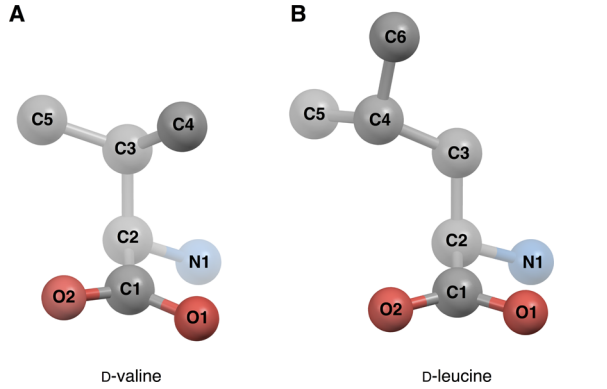

Figure 8. Atom numbers for (A) valine and (B) leucine molecules used to identify intramolecular coordinates (bond angles and dihedral angles) in the text and in Figures 9–11.

This procedure can be carried out without the quantification method used to generate Figure 7, but the percentages of translational, librational, and intramolecular character are useful quantities for scaling purposes.

For a particular vibration, the process of generating its ID strip begins by displacing all atoms in the unit cell along the normal mode coordinates. The translational parts of the ID strip are then calculated by finding the vector connecting the equilibrium and displaced center of mass for each molecule. The three Cartesian components of this vector are scaled such that the largest component equals one. The vector is then multiplied by the scalar "percent translational character" calculated previously. The sign of each component of the translational displacement is indicated with red for negative and blue for positive, and the opacity of that portion of the mode ID strip is proportional to the magnitude of the particular displacement.

The rotational components of the vibration's ID strip are more complicated to calculate. For each molecule, the displaced coordinates of the distorted molecule are translated such that the center of mass is coincident with the center of mass of the equilibrium coordinates. Then the distorted molecule is rigidly rotated using Euler angles in a least-squares minimization routine that obtains the best overlap between the equilibrium and displaced coordinates.[20,76,77] The problem is similar to the general attitude orientation problem in astronomy, also known as Wahba's problem.[78,79]

If the Euler angles that optimally return the displaced coordinates to their equilibrium values are (ϕ, θ, ψ), then $(-\psi, -\theta, -\phi)$ are the Euler angles that correspond to the rotational part of the motion. These angles are used to construct the 3×3 rotation matrix, $\mathbf{M}$.[20] This matrix will have three eigenvectors with three corresponding eigenvalues. However, only one of the eigenvalues will be real; the other two will be complex. If the eigenvalue λ is a real number, then the corresponding eigenvector $\mathbf{u}$ is coincident with the axis about which $\mathbf{M}$ rotates the coordinates in Cartesian space. The amount that $\mathbf{M}$ rotates the molecule about $\mathbf{u}$ can be calculated using

$$\mathrm{tr}(\mathbf{M}) = 1 + 2\cos\vartheta \tag{7}$$

where $\mathrm{tr}(\mathbf{M})$ is the trace of $\mathbf{M}$, and ϑ is the amount of rotation.[80]

dx.doi.org/10.1021/jp406730a | J. Phys. Chem. B 2013, 117, 10444–10461

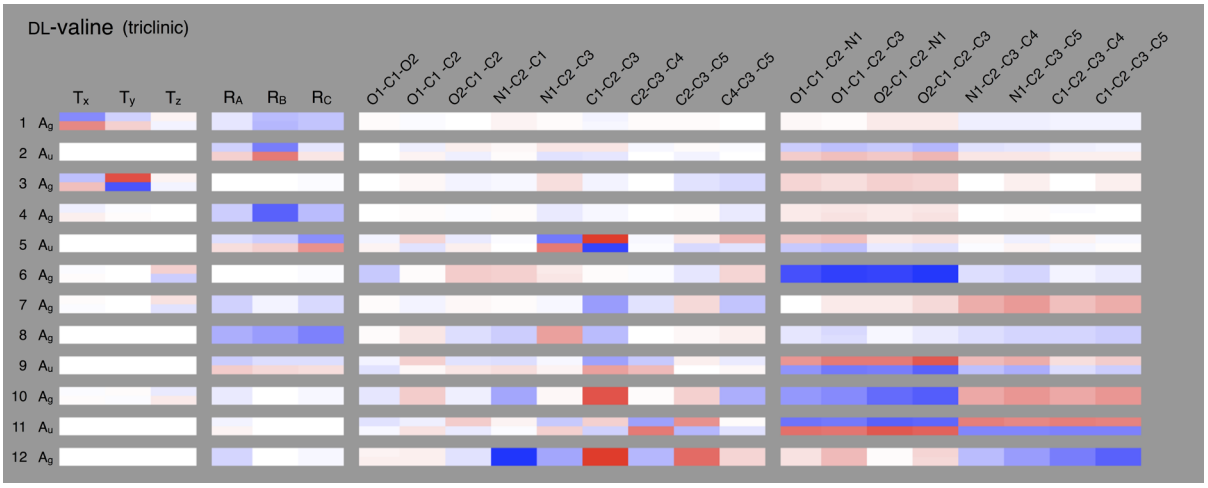

Figure 9. "Vibrational character ID strips" of the 12 lowest-frequency vibrational modes in triclinic DL-valine calculated using the method described in the text. For each mode, the motion of each molecule in the unit cell is depicted as a separate row within the ID strip. The colors red and blue indicate the sign of that particular motion for one molecule relative to the other. Each ID strip indicates the specific nature of motion in these modes, including combinations of inter- and intramolecular displacements. The translational components of vibrations are indicated in the Cartesian coordinates of the unit cell, while the librational components are represented as the projection of the rotation axis on the principle axes of the molecule (see Table 4). The final two categories of motion reported in each ID strip are the displacements along selected intramolecular coordinates, specifically the bond angles and torsions for the non-hydrogen atoms. The symmetry label for each vibration was determined numerically using the complete eigenvector for the mode. A_u modes are IR-active.

Table 4. Unit Cell Vectors and Moment of Inertia Tensors for Molecules in Their Equilibrium Geometry As Calculated by DFT (SIESTA, vdW-DF) Reported in Cartesian Coordinates

	unit cell vectors				principal rotation axes				
		coordinates (Å)				orientation			moment of inertia
compound	vector	x	y	z	axis	x	y	z	(amu·Å^2)
DL-leucine (triclinic)	**a**	5.07	0	0	R_A	−0.14	−0.45	0.88	224.29
	b	−2.11	5.00	0	R_B	0.74	−0.64	−0.2	523.32
	c	−1.31	−0.22	13.47	R_C	0.65	0.63	0.43	583.93
DL-valine (triclinic)	**a**	5.20	0	0	R_A	−0.05	−0.62	0.79	211.48
	b	−1.93	5.10	0	R_B	0.88	−0.4	−0.26	296.13
	c	−0.29	−0.24	10.59	R_C	0.48	0.68	0.56	356.93
DL-valine (monoclinic)	**a**	5.17	0	0	R_A	−0.04	−0.6	0.8	211.33
	b	−1.9	5.11	0	R_B	0.87	−0.41	−0.27	296.34
	c	0	0	21.48	R_C	0.49	0.68	0.54	356.04

Once the rotation axis **u** has been determined, it is transformed from $x-y-z$ coordinates into the basis of the principal axes of the molecule: R_A, R_B, R_C. The standard definition of the principal axes is used such that rotation about R_A has the smallest moment of inertia and R_C has the largest. The calculated principal rotation axes for the molecules in each of the three systems are reported in Table 4. The elements of **v**, the rotation axis in the principal axis coordinates, are scaled such that the largest of the three components is equal to 1. The elements of **v** are then multiplied by the "percent rotational character" determined for the mode. These scaled values are used in the ID strip. For the rotational components, the colors red or blue are assigned to each molecule in the unit cell in order to indicate its relative rotation direction (as observed from a fixed reference point).

The intramolecular displacements are quantified by finding the difference between the values of each particular internal coordinate of interest in the displaced and equilibrium geometries. The calculated amplitudes of the various intramolecular displacements are inherently scaled relative to one another, both within and between modes and therefore are not adjusted. The largest bond angle change among all vibrations represented in a particular figure determines the maximum opacity value for that category of motion and likewise for the dihedral angles. Changes in bond and dihedral angles can be either positive or negative, which is indicated by using red and blue to indicate the relative signs of displacements along various coordinates. As in the case of the rotation direction, the sign of dihedral angle changes are defined for an observer at a fixed external vantage point.

Character of Triclinic DL-Valine Vibrations. The ID strips of triclinic DL-valine modes (Figure 9) increase the information available from the more general character quantification procedure used to construct Figure 7. For example, although Figure 7B reveals that DL-valine modes 1 and 3 largely involve translational displacements, using Figure 9 one can quickly determine the specific orientation of these motions: translations parallel to the $x-y$ plane. This corresponds to

dx.doi.org/10.1021/jp406730a | J. Phys. Chem. B 2013, 117, 10444−10461

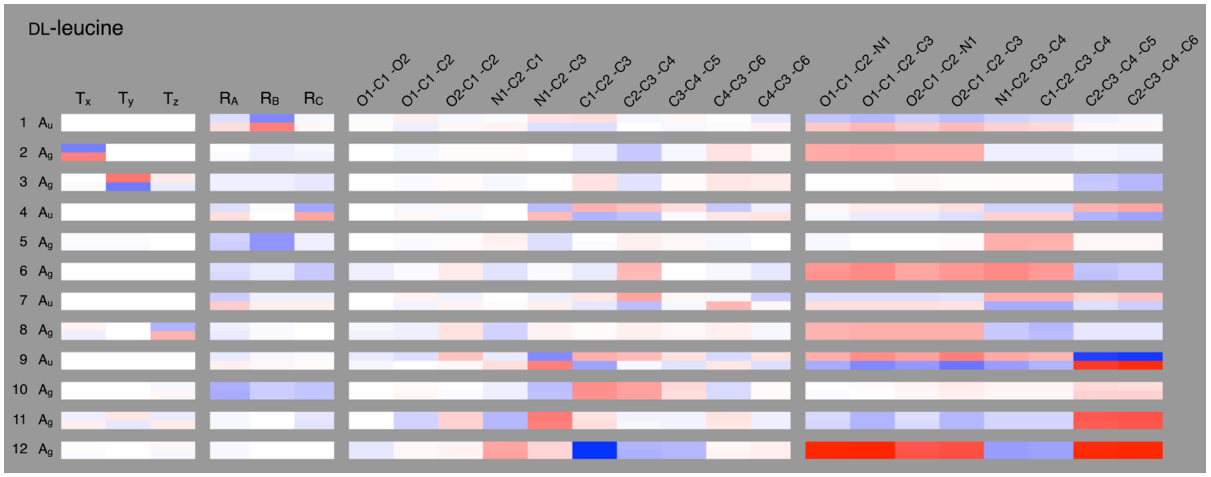

Figure 10. Character ID strips for the low-frequency vibrations in DL-leucine. A comparison to the ID strips for triclinic DL-valine vibrations (Figure 9) reveals that some of the modes have close analogues, although the intramolecular motion that accompanies specific translational or librational displacement can differ. A_u modes are IR-active.

motion parallel to the **a**−**b** plane of the unit cell. In mode 3, a molecule translating in the positive y direction simultaneously moves in the negative x direction with about one-third the amount of displacement. Using Table 4, which contains the Cartesian components of the unit cell vectors depicted in Figure 3, one can see that mode 3 consists of molecules translating essentially parallel to **b**. On the other hand, it is now apparent that the translational displacements of mode 1 are perpendicular to **b**.

Motion along the z-direction translational degree of freedom is divided mostly between modes 6 and 7, with a small amount in mode 10. This is probably due to the fact that these modes require pulling against the interlayer hydrogen bonds, raising the energy of this motion into the frequency region where it couples with a variety of intramolecular motions. Despite the fact that more than three modes have some translational character, the selection rule from group theory that translational modes in this system will not be IR-active remains valid.

Unlike the translational modes, there are expected to be librational modes of both A_g and A_u character; that is, both IR and Raman-active modes may have some librational character. In the simplest picture, these librations should exist in pairs where the rotation is around the same axis but the molecules rotate either symmetrically (A_g) or antisymmetrically (A_u). The ID strips indeed show that modes with similar but opposite parity motion exist, for example, modes 2 and 4 or modes 5 and 8 in Figure 9. The A_g mode in each pair appears at higher frequency because the symmetric rotations require greater distortion of the interlayer hydrogen bonds.

The lowest-frequency modes with librational character seem to involve rotation of molecules around an axis that is close to R_B, the axis with the second highest moment of inertia rather than R_C, the heaviest axis (see Table 4). This can be explained by the hydrogen bond geometry of the system. With reference to Figure 3, the R_C axis of each molecule is roughly perpendicular to the plane containing the interlayer hydrogen bonds (depicted in purple). As a consequence, rotation around this axis requires stretching of the interlayer hydrogen bonds. On the other hand, rotating the molecules around R_B allows the hydrogen bonds to remain relatively rigid.

While the librational motion around the heavier two principal axes seems to be mostly confined to the two pairs of modes mentioned above, a rotational component around R_A, the lightest principal axis, appears in a large number of vibrations (all modes but 3 and 6). There are no modes that are both strongly librational in nature and also have a rotation axis oriented near R_A. Instead, modes whose axis of rotation is closely aligned with R_A involve relatively little net rotation because it would require twisting of the intralayer hydrogen bond network (green and orange lines in Figure 3). The hydrophilic portion of the molecule is hindered by these hydrogen bonds, and so the rotational motion is mostly limited to torsions of the −R group. In other words, this particular rotational degree of freedom is coupled with a number of intramolecular degrees of freedom that involve large-scale torsional displacements.

Character of DL-Leucine Modes. It is also possible to compare the ID strips of two different systems. Several of the DL-leucine ID strips (Figure 10) seem to have analogues among the triclinic DL-valine ID strips. For example, DL-leucine mode 1 and triclinic DL-valine mode 2 each have no translational character and have rotational displacements almost exclusively around the R_B axis, and their intramolecular displacements are very similar. Some of the difference in the observed ratio of frequencies (1.73 THz for DL-valine mode 2 versus 1.45 THz for DL-leucine mode 1) can be explained by the fact that leucine has a larger moment of inertia about its R_B axis (296.34 amu·Å^2 for valine versus 523.32 amu·Å^2 for leucine, as seen in Table 4). However, this is not a full explanation because while the inverse square root of the ratio of moments of inertia is 1.33, the ratio of their frequencies is only 1.19.

In addition to the similarities between their respective lowest-frequency modes, DL-leucine mode 4 shares many characteristics with DL-valine mode 5. For both modes, the rotational component is mostly around R_C, the heaviest axis. The angles that change in both cases are ones that indicate that the −R group is rocking back and forth (N1−C2−C3 and C1−C2−C3). For DL-leucine, the −R group is somewhat less rigid overall, with the isopropyl group (C4, C5, and C6) at the end undergoing torsional motion, as can be seen in the last two segments of the dihedral angle portion of the ID strip for mode 4 of that system.

In the DL-leucine THz spectrum, a third peak is observed, which, based on the DFT calculation, is speculatively identified

dx.doi.org/10.1021/jp406730a | J. Phys. Chem. B 2013, 117, 10444−10461

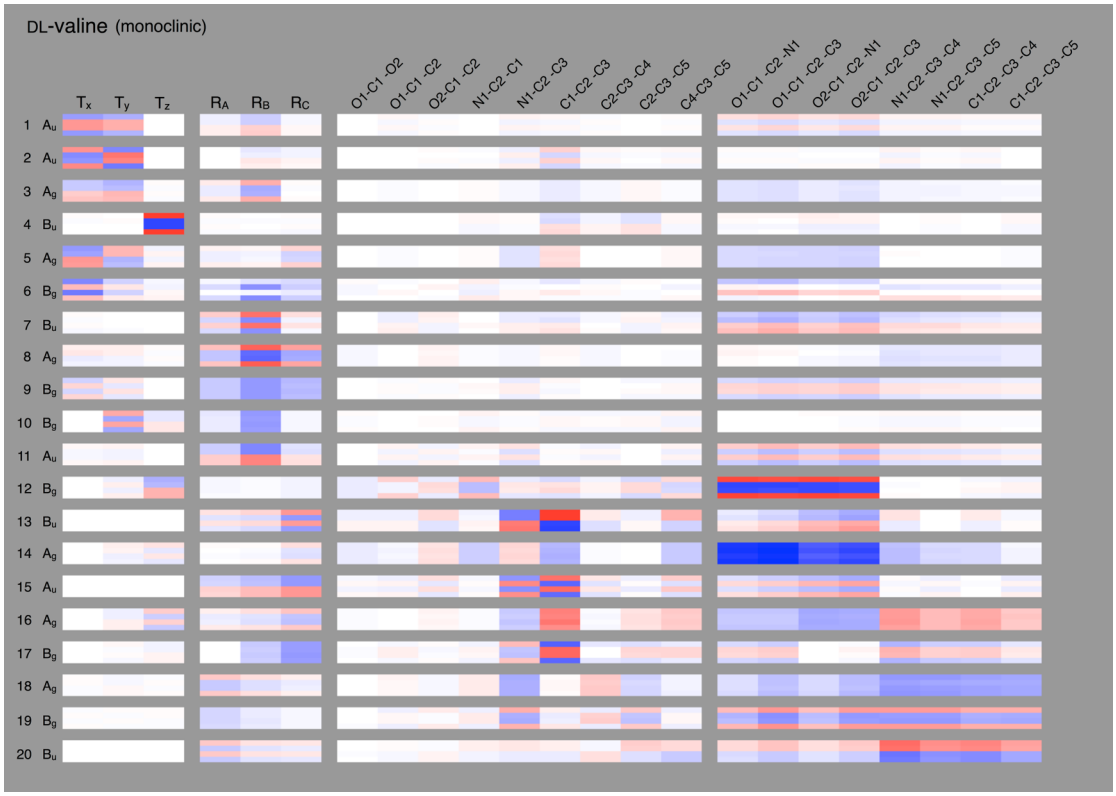

Figure 11. Character ID strips for the low-frequency vibrations in monoclinic DL-valine. Each ID strip has four rows, one for each molecule in the unit cell (compared to two for triclinic DL-valine and DL-leucine). The stacking order of the rows in each ID strip corresponds to the orientation of the molecules in monoclinic DL-valine as depicted in Figure 2B. The presence of four molecules in the unit cell means that it is possible for vibrational modes in this system to involve translational intermolecular displacements and also be IR active. For example, mode 2 is predicted to be IR-active, and it corresponds to a feature in the experimental THz absorption spectrum that is not present in the spectrum of triclinic DL-valine (Figure 5B,C). A_u and B_u modes are IR-allowed.

as mode 7. In Figure 7 it was seen that this mode and triclinic DL-valine mode 9 had almost identical proportions of rotational and intramolecular character. However, the ID strips show that, although the rotational portion of this mode is somewhat similar to that of mode 9 in the triclinic DL-valine calculation, the characteristic intramolecular aspect of this vibration is that torsion of the −R group (last four dihedral angles in the ID strip) changes out of phase with the torsion involving the H-bonding region (first four dihedral angles in the ID strip). This pattern of intramolecular motion is different from that of triclinic DL-valine mode 9 and DL-leucine mode 9, where the torsion in the H-bonding portion of the molecule happens in phase with the twist around the C2−C3 bond in the respective −R groups. Similar behavior is observed in DL-valine mode 11 (the fourth IR-active mode, rather than the third). The ID strips also give a possible explanation for the fact that triclinic DL-valine mode 9 has much greater calculated IR activity than DL-leucine mode 7: there is more displacement along intramolecular coordinates corresponding to twisting of the carboxylate group in the DL-valine mode, indicated by the greater relative opacity in the first four dihedral angles of its ID strip.

It is interesting to note that the experimental temperature dependence of DL-leucine mode 1 seen in Figure 4 is much larger than DL-valine mode 2 even though their ID strips are so similar. Mode 1 of DL-leucine blueshifts by 7% upon cooling from room temperature to 80 K, while DL-valine mode 2 blueshifts by less than 2%. The temperature dependence of a vibrational mode in a solid is a complicated property. In a quasi-harmonic approximation,[81] one contribution to the temperature dependence is the anharmonicity of the vibrational potential, meaning that the observed central frequency will depend on the relative occupation of higher vibrational states. The second contribution to the temperature dependence is ascribed to the change in unit cell geometry as a function of temperature, which causes the vibrational potential of the mode to change. Strictly speaking, these two quantities are not fully separable because the change in unit cell geometry with temperature is itself caused by anharmonic interactions.

Because triclinic DL-valine mode 2 does not have strong temperature dependence, one could speculate that the vibration is neither particularly anharmonic, nor is it especially sensitive to changes in the unit cell. Therefore, assuming that DL-leucine mode 1 is as similar to DL-valine mode 2 as the calculations indicate, the larger temperature dependence of the DL-leucine mode suggests that either the motion of the −R group of leucine introduces anharmonicity to the overall vibrational potential of the mode or that the larger −R group somehow makes the mode more sensitive to changes in the unit cell geometry. The relatively symmetric and Lorentzian line shape of the peaks in Figures 4 and 5 supports the second possibility. Although it is true that the lower frequency of DL-leucine mode 1 means that it would have a greater change in average occupation level with temperature than the corresponding triclinic DL-valine vibration (mode 2), the two other DL-leucine vibrations observed in the THz spectra both have higher frequencies than DL-valine mode 2 yet still exhibit very strong temperature dependence. It seems more likely that, compared

dx.doi.org/10.1021/jp406730a | J. Phys. Chem. B 2013, 117, 10444−10461

to valine, the −R group of leucine has more flexibility and undergoes larger conformational changes as the unit cell geometry varies, which would have effects on both the vibrational potentials as well as the moment of inertia tensor of each molecule. In other words, at a given temperature, the potentials of the observed DL-leucine vibrations are relatively harmonic, but conformational changes in the −R group as a function of the unit cell geometry affect the steepness of those potentials.

Character of Monoclinic DL-Valine Vibrations. Both DL-valine polymorphs have an IR absorption peak that occurs at 1.73 THz (at 80 K). In the calculated spectra, this feature seems to correspond to mode 2 in triclinic DL-valine and mode 7 in monoclinic DL-valine. The character of this mode, as illustrated by its ID strip in Figure 11, is very similar to that of mode 2 in the triclinic DL-valine calculation (Figure 9). The intermolecular character of the mode is primarily rotational (around the R_B axis), and the intramolecular displacements are primarily torsional. The symmetry of mode 7 is B_u, while mode 2 of the triclinic polymorph has symmetry A_u. If one were to observe the four molecules rotating in one unit cell of the monoclinic polymorph, the directions of their respective rotations would be the same as that of the four molecules in two unit cells of the triclinic polymorph stacked on top of each other.

In addition to the IR absorption feature that it shares with the triclinic polymorph, the monoclinic form of DL-valine has an additional absorption peak at roughly 0.90 THz. The calculated spectrum of monoclinic DL-valine has a similar feature, mode 2, which consists primarily of rigid translations of the molecules in the x−y (**a**−**b**) plane. This mode involves the two pairs of hydrogen-bonded molecules translating out of phase, while each set of interlayer hydrogen bonds remains rigid. The triclinic system has only one pair of molecules, and when they translate together in a way that leaves the hydrogen bonds rigid, the mode corresponds to an acoustic phonon. The molecules remain more rigid in this mode than any other IR-active mode in any of the three systems studied here. In addition, the fact that the motion of the mode is limited to the **a**−**b** plane suggests that its behavior might be a useful source of information about interlayer shear forces.

Raman-Active Modes. Using character ID strips is also a useful way to begin analysis of the low-frequency vibrations in the Raman spectra. In Figure 5, the calculated frequencies of Raman-active modes are indicated with small black bars. The modes in the monoclinic DL-valine spectrum occur in pairs, especially above 3 THz. As seen in the ID strips (Figure 11), these pairs often correspond to one A_g and an analogous B_g mode (see for example modes 12 and 14, 18 and 19, and the intramolecular portions of modes 16 and 17). The correspondence between these pairs of modes in the monoclinic calculation and an analogue in the triclinic DL-valine calculation is also often quite good. For example, modes 12 and 14 of the monoclinic polymorph are very similar to mode 6 of the triclinic polymorph.

However, in monoclinic DL-valine, below about 3 THz the A_g and B_g members of a pair begin to differ from one another both in frequency and to some extent in character as well. For example, the most similar B_g counterpart to mode 5 is mode 9, not mode 6. Modes 8 and 10 as well as 3 and 6 are other examples of A_g and B_g pairs that are dissimilar in frequency. For a given pair of modes, the one with B_g symmetry usually occurs at a higher frequency, if only slightly. The only exception is found in modes 12 and 14, whose intramolecular component is essentially torsion of the carboxylate groups. In mode 14, however, the intramolecular motion also involves a twisting of the −R group.

Compared to the triclinic polymorph, the experimental Raman spectrum of monoclinic DL-valine has only one extra peak in the frequency range below ∼3 THz despite the fact that the calculations predict three additional modes that are Raman active. However, not all modes that have Raman-active symmetry will necessarily have strong intensity. This is analogous to the fact that modes 1 and 4 in monoclinic DL-valine have IR-active symmetry but negligible calculated intensity, which agrees with the observed THz absorption spectrum.

CONCLUSIONS

Vibrations that involve both inter- and intramolecular degrees of freedom become more difficult to describe as the size of the system increases. We have investigated several means of addressing this problem, including new ways of quantifying the character of calculated vibrational modes, presenting detailed information in compact and visual forms, and comparing the results of different calculations.

One of the more surprising observations is that terminology used to describe the intermolecular motions of rigid molecules remains useful for vibrational modes with substantial intramolecular displacements. While one might surmise that the applicability of rigid models would be determined *a priori* by the identity of the molecules in the system, our results indicate that the situation is not all-or-nothing: molecules may behave rigidly when undergoing motion involving certain intermolecular degrees of freedom but not others. In fact, to a large extent, the relative rigidity of different modes in these systems is determined by interactions between molecules, especially hydrogen bonding.

Vibrations in the frequency range below approximately 3 THz (100 cm^{-1}) are especially interesting because they often exhibit enough of a particular intermolecular motion to allow comparisons to be made between similar modes in different systems. In other words, for many of the lowest-frequency modes, it is possible to invoke the idealized rigid motion (e.g., a translation in a particular direction) as a starting point and then consider exactly how that particular motion is perturbed in different systems, including which intramolecular degrees of freedom become involved. This type of analysis would not be possible if, for instance, the intermolecular degrees of freedom were distributed among a very large number of vibrations or if, on the other hand, the motions were completely rigid. Rather than being an anonymous profusion of indistinguishable bath modes, low-frequency vibrations with a mixture of inter- and intramolecular components can be both uniquely identified and quantitatively described; with appropriate analysis, they represent a rich source of information.

ASSOCIATED CONTENT

Supporting Information

Powder XRD spectra used to identify polymorphs, representative time-domain THz measurements, detailed treatment of plasma lines in Raman spectra, example DFT calculation input files, calculated vibrational frequencies and eigenvectors, and table of values used in character ID strips (Figures 9−11). This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*E-mail: charles.schmuttenmaer@yale.edu.

Present Address

B.K.O.-O.: Massachusetts Institute of Technology, Department of Chemistry, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, United States.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

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 no. CNS 08-21132 that partially funded acquisition of the facilities. Funding for this work was also provided through NSF grant no. CHE-0911593.

REFERENCES

(1) Zhang, H.; Siegrist, K.; Plusquellic, D. F.; Gregurick, S. K. Terahertz Spectra and Normal Mode Analysis of the Crystalline VA Class Dipeptide Nanotubes. *J. Am. Chem. Soc.* **2008**, *130*, 17846−17857.

(2) Schiró, G.; Caronna, C.; Natali, F.; Cupane, A. Direct Evidence of the Amino Acid Side Chain and Backbone Contributions to Protein Anharmonicity. *J. Am. Chem. Soc.* **2010**, *132*, 1371−1376.

(3) Heugen, U.; Schwaab, G.; Bründermann, E.; Heyden, M.; Yu, X.; Leitner, D. M.; Havenith, M. Solute-Induced Retardation of Water Dynamics Probed Directly by Terahertz Spectroscopy. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 12301−12306.

(4) Born, B.; Weingärtner, H.; Bründermann, E.; Havenith, M. Solvation Dynamics of Model Peptides Probed by Terahertz Spectroscopy. Observation of the Onset of Collective Network Motions. *J. Am. Chem. Soc.* **2009**, *131*, 3752−3755.

(5) Niehues, G.; Heyden, M.; Schmidt, D. A.; Havenith, M. Exploring Hydrophobicity by THz Absorption Spectroscopy of Solvated Amino Acids. *Faraday Discuss.* **2011**, *150*, 193−207.

(6) Grischkowsky, D.; van Exter, M.; Fattinger, C. Terahertz Time-Domain Spectroscopy of Water Vapor. *Opt. Lett.* **1989**, *14*, 1128−1130.

(7) Kindt, J. T.; Schmuttenmaer, C. A. Far-Infrared Dielectric Properties of Polar Liquids Probed by Femtosecond Terahertz Pulse Spectroscopy. *J. Phys. Chem.* **1996**, *100*, 10373−10379.

(8) Ikeda, Y.; Ishihara, Y.; Moriwaki, T.; Kato, E.; Terada, K. A Novel Analytical Method for Pharmaceutical Polymorphs by Terahertz Spectroscopy and the Optimization of Crystal Form at the Discovery Stage. *Chem. Pharm. Bull.* **2010**, *58*, 76−81.

(9) Watanabe, Y.; Kawase, K.; Ikari, T.; Ito, H.; Ishikawa, Y.; Minamide, H. Component Analysis of Chemical Mixtures using Terahertz Spectroscopic Imaging. *Opt. Commun.* **2004**, *234*, 125−129.

(10) Parrott, E. P. J.; Zeitler, J. A.; Friščić, T.; Pepper, M.; Jones, W.; Day, G. M.; Gladden, L. F. Testing the Sensitivity of Terahertz Spectroscopy to Changes in Molecular and Supramolecular Structure: A Study of Structurally Similar Cocrystals. *Cryst. Growth Des.* **2009**, *9*, 1452−1460.

(11) Franz, M.; Fischer, B. M.; Walther, M. Probing Structure and Phase-Transitions in Molecular Crystals by Terahertz Time-Domain Spectroscopy. *J. Mol. Struct.* **2011**, *1006*, 34−40.

(12) Yamaguchi, M.; Miyamaru, F.; Yamamoto, K.; Tani, M.; Hangyo, M. Terahertz Absorption Spectra of L-, D-, and DL-Alanine and their Application to Determination of Enantiometric Composition. *Appl. Phys. Lett.* **2005**, *86*, 053903.

(13) King, M. D.; Hakey, P. M.; Korter, T. M. Discrimination of Chiral Solids: A Terahertz Spectroscopic Investigation of L- and DL-Serine. *J. Phys. Chem. A* **2010**, *114*, 2945−2953.

(14) Kutteruf, M. R.; Brown, C. M.; Iwaki, L. K.; Campbell, M. B.; Korter, T. M.; Heilweil, E. J. Terahertz Spectroscopy of Short-Chain Polypeptides. *Chem. Phys. Lett.* **2003**, *375*, 337−343.

(15) Ponseca, C. S.; Kambara, O.; Kawaguchi, S.; Yamamoto, K.; Tominaga, K. Low-Frequency Spectra of Amino Acids and Short-Chain Peptides Studied by Terahertz Time-Domain Spectroscopy. *J. Infrared Millim. Te.* **2010**, *31*, 799−809.

(16) Ahmed, Z.; Chou, S. G.; Siegrist, K.; Plusquellic, D. F. State-Resolved THz Spectroscopy and Dynamics of Crystalline Peptide−Water Systems. *Faraday Discuss.* **2011**, *150*, 175−192.

(17) Rungsawang, R.; Ueno, Y.; Tomita, I.; Ajito, K. Angle-Dependent Terahertz Time-Domain Spectroscopy of Amino Acid Single Crystals. *J. Phys. Chem. B* **2006**, *110*, 21259−21263.

(18) Hufnagle, D. C.; Taulbee-Combs, A. R.; Spendel, W. U.; Pacey, G. E. Collective Mode Frequency Shifts in L-Serine and a Series of Isotopologues in the Terahertz Regime. *Talanta* **2012**, *88*, 61−65.

(19) Walther, M.; Plochocka, P.; Fischer, B.; Helm, H.; Uhd Jepsen, P. Collective Vibrational Modes in Biological Molecules Investigated by Terahertz Time-Domain Spectroscopy. *Biopolymers* **2002**, *67*, 310−313.

(20) Williams, M. R. C.; True, A. B.; Izmaylov, A. F.; French, T. A.; Schroeck, K.; Schmuttenmaer, C. A. Terahertz Spectroscopy of Enantiopure and Racemic Polycrystalline Valine. *Phys. Chem. Chem. Phys.* **2011**, *13*, 11719−11730.

(21) Aschaffenburg, D. J.; Williams, M. R. C.; Talbayev, D.; Santavicca, D. F.; Prober, D. E.; Schmuttenmaer, C. A. Efficient Measurement of Broadband Terahertz Optical Activity. *Appl. Phys. Lett.* **2012**, *100*, 241114.

(22) Morris, C. M.; Aguilar, R. V.; Stier, A. V.; Armitage, N. P. Polarization Modulation Time-Domain Terahertz Polarimetry. *Opt. Express* **2012**, *20*, 12303−12317.

(23) Lima, J. A.; Freire, P. T. C.; Lima, R. J. C.; Moreno, A. J. D.; Mendes Filho, J.; Melo, F. E. A. Raman Scattering of L-Valine Crystals. *J. Raman Spectrosc.* **2005**, *36*, 1076−1081.

(24) da Silva, J. H.; Lima, J. A., Jr.; Freire, P. T. C.; Lemos, V.; Mendes Filho, J.; Melo, F. E. A.; Pizani, P. S.; Fischer, J.; Klemke, B.; Kemner, E.; Bordallo, H. N. Raman Spectroscopy and Inelastic Neutron Scattering Study of Crystalline L-Valine. *J. Phys.: Condens. Matter* **2009**, *21*, 415404.

(25) Murli, C.; Vasanthi, R.; Sharma, S. M. Raman Spectroscopic Investigations of DL-Serine and DL-Valine Under Pressure. *Chem. Phys.* **2006**, *331*, 77−84.

(26) Gorelik, V. S.; Zlobina, L. I. Influence of Perturbations on the Lattice Spectra of Crystalline Amino Acids. *J. Russ. Laser Res.* **1996**, *17*, 119−132.

(27) Iwama, S.; Horiguchi, M.; Sato, H.; Uchida, Y.; Takahashi, H.; Tsue, H.; Tamura, R. Observation of the Preferential Enrichment Phenomenon for Essential α-Amino Acids with a Racemic Crystal Structure. *Cryst. Growth Des.* **2010**, *10*, 2668−2675.

(28) Flaig, R.; Koritsanszky, T.; Dittrich, B.; Wagner, A.; Luger, P. Intra- and Intermolecular Topological Properties of Amino Acids: A Comparative Study of Experimental and Theoretical Results. *J. Am. Chem. Soc.* **2002**, *124*, 3407−3417.

(29) Mallikarjunan, M.; Rao, S. T. Crystal Structure of DL-Valine. *Acta Crystallogr.* **1969**, *25*, 296−303.

(30) Thomas, I. R.; Bruno, I. J.; Cole, J. C.; Macrae, C. F.; Pidcock, E.; Wood, P. A. WebCSD: The Online Portal to the Cambridge Structural Database. *J. Appl. Crystallogr.* **2010**, *43*, 362−366.

(31) Allen, F. H. The Cambridge Structural Database: a Quarter of a Million Crystal Structures and Rising. *Acta Crystallogr., Sect. B: Struct. Sci.* **2002**, *58*, 380−388.

(32) Dalhus, B.; Henrik Görbitz, C. Crystal Structures of Hydrophobic Amino Acids: Interaction Energies of Hydrogen-Bonded Layers Revealed by Ab Initio Calculations. *J. Mol. Struct.: THEOCHEM* **2004**, *675*, 47−52.

(33) Görbitz, C. H.; Vestli, K.; Orlando, R. A Solution to the Observed Z′ = 2 Preference in the Crystal Structures of Hydrophobic Amino Acids. *Acta Crystallogr., Sect. B: Struct. Sci.* **2009**, *65*, 393−400.

(34) Day, G. M.; Cooper, T. G. Crystal Packing Predictions of the Alpha-Amino Acids: Methods Assessment and Structural Observations. *CrystEngComm* **2010**, *12*, 2443−2453.
(35) Duvillaret, L.; Garet, F.; Coutaz, J. L. A Reliable Method for Extraction of Material Parameters in Terahertz Time-Domain Spectroscopy. *IEEE J. Sel. Top. Quantum Electron.* **1996**, *2*, 739−746.
(36) Craig, N. C.; Levin, I. W. Calibrating Raman Spectrometers with Plasma Lines from the Argon Ion Laser. *Appl. Spectrosc.* **1979**, *33*, 475−476.
(37) Kleine-Ostmann, T.; Wilk, R.; Rutz, F.; Koch, M.; Niemann, H.; Güttler, B.; Brandhorst, K.; Grunenberg, J. Probing Noncovalent Interactions in Biomolecular Crystals with Terahertz Spectroscopy. *ChemPhysChem* **2008**, *9*, 544−547.
(38) Tulip, P. R.; Clark, S. J. Dielectric and Vibrational Properties of Amino Acids. *J. Chem. Phys.* **2004**, *121*, 5201−5210.
(39) Tulip, P. R.; Clark, S. J. Structural and Electronic Properties of L-Amino Acids. *Phys. Rev. B* **2005**, *71*, 195117.
(40) Tulip, P. R.; Clark, S. J. Lattice Dynamical and Dielectric Properties of L-Amino Acids. *Phys. Rev. B* **2006**, *74*, 064301.
(41) Tulip, P. R.; Bates, S. P. First Principles Determination of Structural, Electronic and Lattice Dynamical Properties of a Model Dipeptide Molecular Crystal. *Mol. Phys.* **2009**, *107*, 2201−2212.
(42) Jepsen, P. U.; Clark, S. J. Precise Ab-Initio Prediction of Terahertz Vibrational Modes in Crystalline Systems. *Chem. Phys. Lett.* **2007**, *442*, 275−280.
(43) King, M. D.; Buchanan, W. D.; Korter, T. M. Application of London-Type Dispersion Corrections to the Solid-State Density Functional Theory Simulation of the Terahertz Spectra of Crystalline Pharmaceuticals. *Phys. Chem. Chem. Phys.* **2011**, *13*, 4250−4259.
(44) King, M. D.; Buchanan, W. D.; Korter, T. M. Identification and Quantification of Polymorphism in the Pharmaceutical Compound Diclofenac Acid by Terahertz Spectroscopy and Solid-State Density Functional Theory. *Anal. Chem.* **2011**, *83*, 3786−3792.
(45) King, M. D.; Davis, E. A.; Smith, T. M.; Korter, T. M. Importance of Accurate Spectral Simulations for the Analysis of Terahertz Spectra: Citric Acid Anhydrate and Monohydrate. *J. Phys. Chem. A* **2011**, *115*, 11039−11044.
(46) King, M. D.; Ouellette, W.; Korter, T. M. Noncovalent Interactions in Paired DNA Nucleobases Investigated by Terahertz Spectroscopy and Solid-State Density Functional Theory. *J. Phys. Chem. A* **2011**, *115*, 9467−9478.
(47) King, M. D.; Korter, T. M. Application of London-Type Dispersion Corrections in Solid-State Density Functional Theory for Predicting the Temperature-Dependence of Crystal Structures and Terahertz Spectra. *Cryst. Growth Des.* **2011**, *11*, 2006−2010.
(48) Soler, J. M.; Artacho, E.; Gale, J. D.; García, A.; Junquera, J.; Ordejón, P.; Sánchez-Portal, D. The SIESTA Method for Ab Initio Order-N Materials Simulation. *J. Phys.: Condens. Matter* **2002**, *14*, 2745−2779.
(49) Artacho, E.; Anglada, E.; Diéguez, O.; Gale, J. D.; García, A.; Junquera, J.; Martin, R. M.; Ordejon, P.; Pruneda, J. M.; Sánchez-Portal, D.; Soler, J. M. The SIESTA Method; Developments and Applicability. *J. Phys.: Condens. Matter* **2008**, *20*, 064208.
(50) Rydberg, H.; Lundqvist, B. I.; Langreth, D. C.; Dion, M. Tractable Nonlocal Correlation Density Functionals for Flat Surfaces and Slabs. *Phys. Rev. B* **2000**, *62*, 6997−7006.
(51) Rydberg, H.; Dion, M.; Jacobson, N.; Schröder, E.; Hyldgaard, P.; Simak, S.; Langreth, D.; Lundqvist, B. Van der Waals Density Functional for Layered Structures. *Phys. Rev. Lett.* **2003**, *91*, 126402.
(52) Langreth, D. C.; Dion, M.; Rydberg, H.; Schröder, E.; Hyldgaard, P.; Lundqvist, B. I. Van der Waals Density Functional Theory with Applications. *Int. J. Quantum Chem.* **2004**, *101*, 599−610.
(53) Dion, M.; Rydberg, H.; Schröder, E.; Langreth, D. C.; Lundqvist, B. I. Van der Waals Density Functional for General Geometries. *Phys. Rev. Lett.* **2004**, *92*, 246401.
(54) Thonhauser, T.; Cooper, V. R.; Li, S.; Puzder, A.; Hyldgaard, P.; Langreth, D. C. Van der Waals Density Functional: Self-Consistent Potential and the Nature of the van der Waals Bond. *Phys. Rev. B* **2007**, *76*, 125112.
(55) Kelkkanen, A. K.; Lundqvist, B. I.; Nørskov, J. K. Density Functional for van der Waals Forces Accounts for Hydrogen Bond in Benchmark Set of Water Hexamers. *J. Chem. Phys.* **2009**, *131*, 046102.
(56) Román-Pérez, G.; Soler, J. Efficient Implementation of a van der Waals Density Functional: Application to Double-Wall Carbon Nanotubes. *Phys. Rev. Lett.* **2009**, *103*, 096102.
(57) Burns, L. A.; Vázquez-Mayagoitia, Á.; Sumpter, B. G.; Sherrill, C. D. Density-Functional Approaches to Noncovalent Interactions: A Comparison of Dispersion Corrections (DFT-D), Exchange-Hole Dipole Moment (XDM) Theory, and Specialized Functionals. *J. Chem. Phys.* **2011**, *134*, 084107.
(58) Klimeš, J.; Bowler, D. R.; Michaelides, A. Chemical Accuracy for the van der Waals Density Functional. *J. Phys.: Condens. Matter* **2010**, *22*, 022201.
(59) Bahn, S. R.; Jacobsen, K. W. An Object-Oriented Scripting Interface to a Legacy Electronic Structure Code. *Comput. Sci. Eng.* **2002**, *4*, 56−66.
(60) Johnson, D. D. Modified Broyden's Method for Accelerating Convergence in Self-Consistent Calculations. *Phys. Rev. B* **1988**, *38*, 12807−12813.
(61) Fernández-Torre, D.; Escribano, R.; Archer, T.; Pruneda, J. M.; Artacho, E. First-Principles Infrared Spectrum of Nitric Acid and Nitric Acid Monohydrate Crystals. *J. Phys. Chem. A* **2004**, *108*, 10535−10541.
(62) King-Smith, R.; Vanderbilt, D. Theory of Polarization of Crystalline Solids. *Phys. Rev. B* **1993**, *47*, 1651−1654.
(63) Gonze, X.; Lee, C. Dynamical Matrices, Born Effective Charges, Dielectric Permittivity Tensors, and Interatomic Force Constants from Density-Functional Perturbation Theory. *Phys. Rev. B* **1997**, *55*, 10355.
(64) Porezag, D.; Pederson, M. Infrared Intensities and Raman-Scattering Activities within Density-Functional theory. *Phys. Rev. B* **1996**, *54*, 7830−7836.
(65) King, M. D.; Buchanan, W. D.; Korter, T. M. Investigating the Anharmonicity of Lattice Vibrations in Water-Containing Molecular Crystals through the Terahertz Spectroscopy of L-Serine Monohydrate. *J. Phys. Chem. A* **2010**, *114*, 9570−9578.
(66) Yacovitch, T. I.; Kim, J. B.; Garand, E.; van der Poll, D. G.; Neumark, D. M. Slow Photoelectron Velocity-Map Imaging Spectroscopy of the N-Methylvinoxide Anion. *J. Chem. Phys.* **2011**, *134*, 134307.
(67) Press, W. *Single-Particle Rotations in Molecular Crystals*; Springer-Verlag: Berlin, 1981.
(68) Kitaigorodsky, A. I. *Molecular Crystals and Molecules*; Academic Press: New York, 1973.
(69) Fateley, W. G.; McDevitt, N. T.; Bentley, F. F. Infrared and Raman Selection Rules for Lattice Vibrations: The Correlation Method. *Appl. Spectrosc.* **1971**, *25*, 155−173.
(70) Rousseau, D. L.; Bauman, R. P.; Porto, S. P. S. Normal Mode Determination in Crystals. *J. Raman Spectrosc.* **1981**, *10*, 253−290.
(71) McKinnon, J. J.; Spackman, M. A.; Mitchell, A. S. Novel Tools for Visualizing and Exploring Intermolecular Interactions in Molecular Crystals. *Acta Crystallogr., Sect. B: Struct. Sci.* **2004**, *60*, 627−668.
(72) Allis, D. G.; Fedor, A. M.; Korter, T. M.; Bjarnason, J. E.; Brown, E. R. Assignment of the Lowest-Lying THz Absorption Signatures in Biotin and Lactose Monohydrate by Solid-State Density Functional Theory. *Chem. Phys. Lett.* **2007**, *440*, 203−209.
(73) Houjou, H. Coarse Graining of Intermolecular Vibrations by a Karhunen-Loève Transformation of Atomic Displacement Vectors. *J. Chem. Theory Comput.* **2009**, *5*, 1814−1821.
(74) Houjou, H. Evaluation of Coupling Terms Between Intra- and Intermolecular Vibrations in Coarse-Grained Normal-Mode Analysis: Does a Stronger Acid Make a Stiffer Hydrogen Bond? *J. Chem. Phys.* **2011**, *135*, 154111.
(75) Houjou, H.; Koga, R. Explicit Representation of Anisotropic Force Constants for Simulating Intermolecular Vibrations of Multiply Hydrogen-Bonded Systems. *J. Phys. Chem. A* **2008**, *112*, 11256−11262.

(76) Kneller, G. R. Eckart Axis Conditions, Gauss' Principle of Least Constraint, and the Optimal Superposition of Molecular Structures. *J. Chem. Phys.* **2008**, *128*, 194101.
(77) Kudin, K. N.; Dymarsky, A. Y. Eckart Axis Conditions and the Minimization of the Root-Mean-Square Deviation: Two Closely Related Problems. *J. Chem. Phys.* **2005**, *122*, 224105.
(78) Markley, F. L. Attitude Determination using Vector Observations and the Singular Value Decomposition. *J. Astronaut. Sci.* **1988**, *36*, 245−258.
(79) Markley, F. L. Attitude Determination using Vector Observations: A Fast Optimal Matrix Algorithm. *J. Astronaut. Sci.* **1993**, *41*, 261−280.
(80) Giacovazzo, C.; Monaco, H. L.; Artioli, G.; Viterbo, D.; Ferraris, G.; Gilli, G.; Zanotti, G.; Catti, M. *Fundamentals of Crystallography*, 2nd ed.; Oxford University Press: Oxford, 2002.
(81) Leibfried, G.; Ludwig, W. Theory of Anharmonic Effects in Crystals. In *Solid State Physics: Advances in Research and Applications*, Seitz, F., Turnbull, D., Eds.; Academic Press: New York, 1961; pp 275−444.

Supporting Information for:

Intermolecular Vibrations in Hydrophobic Amino Acid Crystals: Experiments and Calculations

Michael R. C. Williams, Daniel J. Aschaffenburg, Benjamin Ofori-Okai[†], Charles A. Schmuttenmaer[*]

Yale University, Department of Chemistry, PO Box 208107, 225 Prospect Street, New Haven, CT 06520-8107, USA

[*]Author to whom correspondence should be addressed: charles.schmuttemaer@yale.edu

[†] Present address: MIT Department of Chemistry, 77 Massachusetts Avenue, Cambridge MA 02139

This document contains the following materials to supplement the main text: Powder XRD spectra used to identify polymorphs, representative time-domain THz measurements, detailed treatment of plasma lines in Raman spectra, example DFT calculation input files, calculated vibrational frequencies and eigenvectors, table of values used in character ID strips.

I. Powder X-ray Diffraction Measurements of DL-leucine and Both DL-valine Polymorphs

The measured powder X-ray diffraction spectra (thick lines, offset vertically) for DL-valine recrystallized under various conditions are compared to calculated powder patterns (thin lines) are shown in Figure S1. The experimental and calculated powder XRD spectra of DL-leucine are reported in Figure S2. All calculated powder patterns were computed from full crystal structures using the Mercury program.[1] The crystal structures were obtained from the Cambridge Crystallographic Database, and the specific entries that we used are listed in Table 1 of the main article.

The experimental spectra in Figure S1A,B are plotted with the calculated powder pattern for monoclinic DL-valine, while those in Figure S1C,D are plotted with the calculated DL-valine triclinic powder pattern. The crystal structure of triclinic DL-valine was measured at 100 K. Therefore, in order to be able compare the calculated powder pattern to our room-temperature powder X-ray measurements, the unit cell was scaled to its value at room temperature. For more details, see the supporting information section of Reference 2.

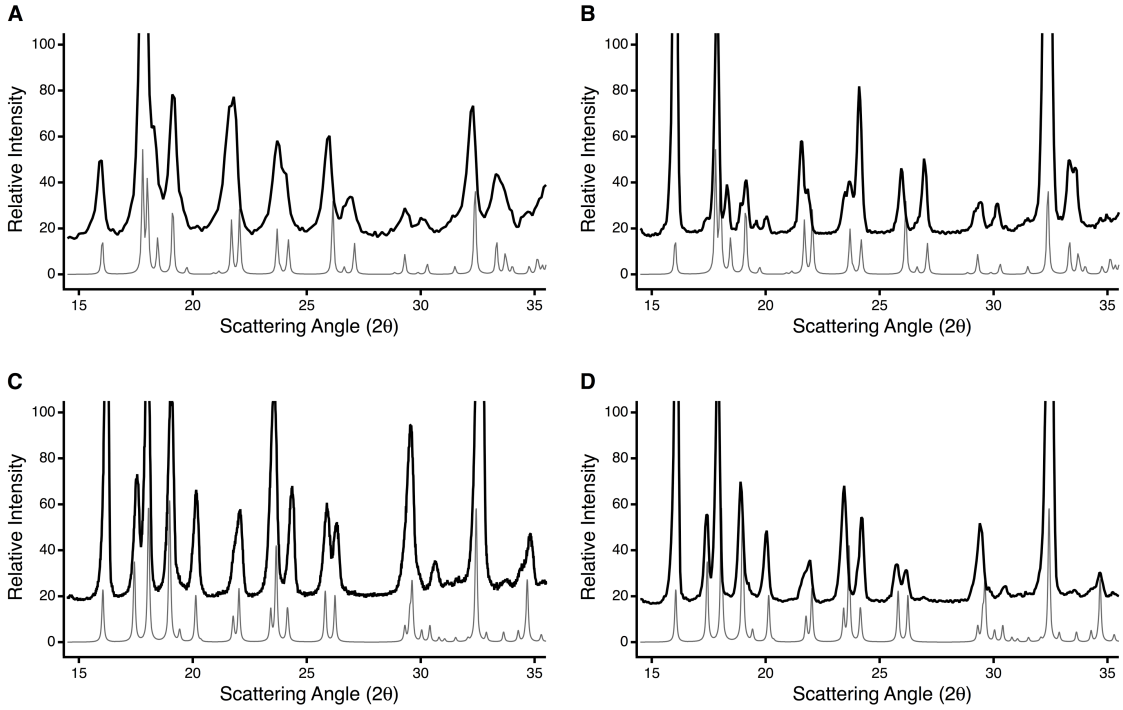

Figure S1. Experimental (thick lines) and calculated (thin lines) powder XRD spectra of two DL-valine polymorphs. (A) Stock DL-valine (Fluka) is a good match to the calculated powder pattern of the monoclinic DL-valine polymorph. (B) Fast recrystallization at 80 °C of a concentrated solution (50 g/L) of stock DL-valine is also identified as the monoclinic form. (C) Slow recrystallization of a dilute solution (2 g/L) of stock DL-valine results in the formation of the triclinic DL-valine polymorph. (D) When pure D-valine and L-valine (both from Alfa-Aesar) are dissolved in water at a concentration of 25 g/L each, the mixture crystallizes as triclinic DL-valine, even if recrystallized rapidly at 80 °C.

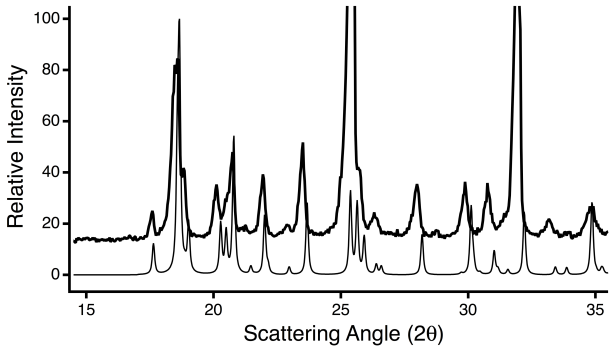

Figure S2. The experimental powder XRD spectrum of DL-leucine (thick line) is compared to the powder pattern calculated from the crystal structure (thin line). Only one polymorph of DL-leucine has been reported, and the measured spectrum is a good match to the pattern expected for the structure.

II. Representative THz Spectra

Examples of the measured time-domain spectrum of DL-leucine (at 80K) and a reference THz pulse are provided in Figure S3. The corresponding power spectra are depicted in Figure S4.

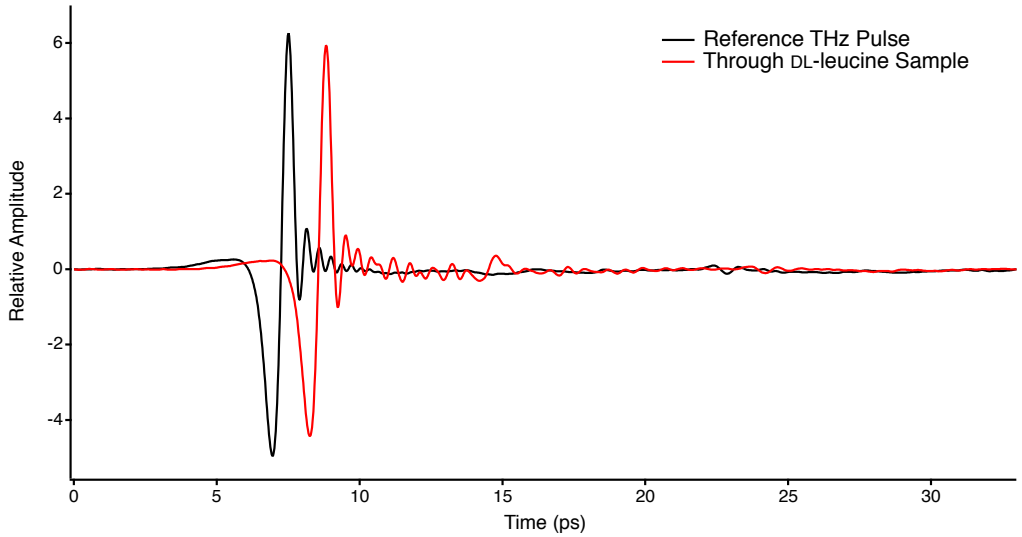

Figure S3. An example of the THz measurements in the time domain.

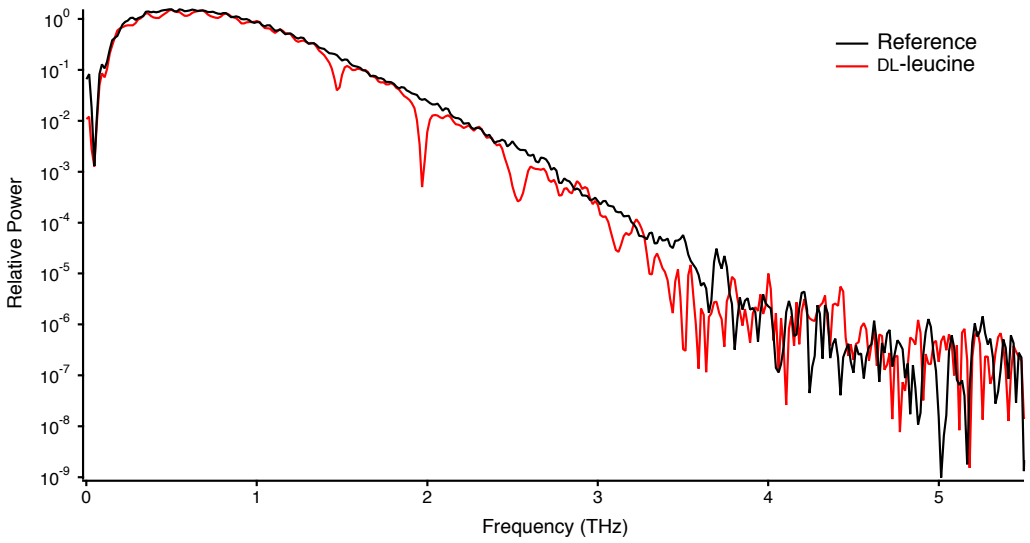

Figure S4. The power spectra of the time-domain measurements in Figure S3.

III. Plasma Lines and Peak-Fitting in Raman Spectra

The original Raman spectra, which include plasma lines from the argon ion laser, are reported in Figure S5. The frequencies at which these lines occur are known and have been reported in the literature.[3] A strong plasma line occurs at approximately 6.65 THz (221.5 cm^{-1}), and weaker ones are found at 1.15 THz (38.4 cm^{-1}) and 3.12 THz (104.0 cm^{-1}). The one at approximately 1.15 THz is troublesome since it gives the appearance of a shoulder on the lowest-frequency peak in the triclinic DL-valine spectrum. That this shoulder is spurious (and not analogous to the actual peak seen in the monoclinic DL-valine spectrum near this frequency) was determined by taking advantage of the fact that the intensity of plasma lines depends more strongly on the laser power than the Raman scattering signal, as clearly illustrated in Figure S6. To numerically remove plasma lines and also simultaneously determine the frequency and FWHM of Raman scattering peaks, a global fit function comprised of Lorentzian lineshapes was employed. In addition, a Lorentizian fixed at 0 THz was included to account for Rayleigh scattering. An example of the quality of the fit (in this case, for the triclinic DL-valine spectrum) can be seen in Figure S7.

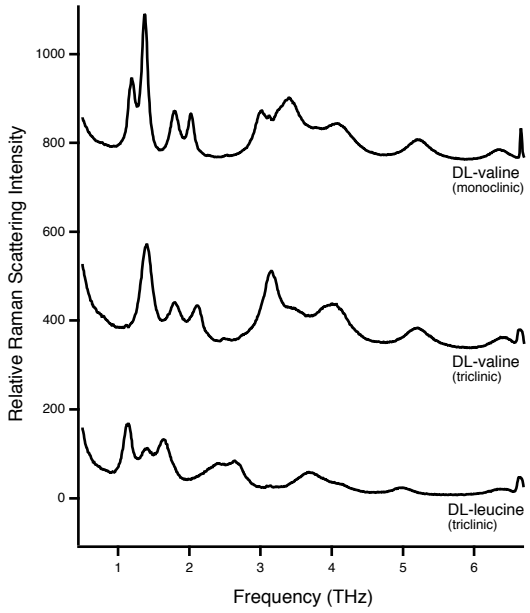

Figure S5. The original Raman spectra, including plasma lines.

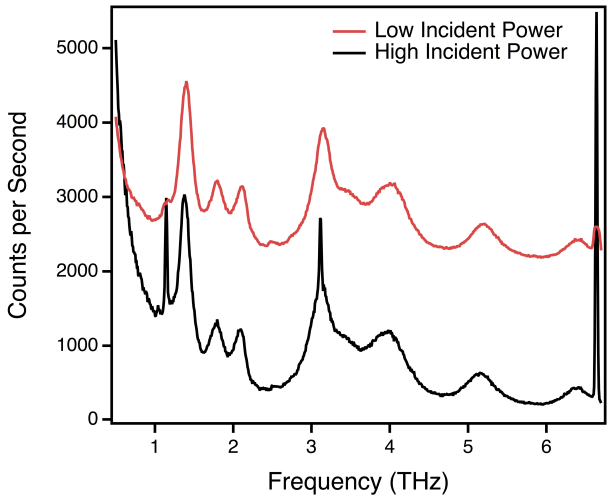

Figure S6. Triclinic DL-valine Raman spectrum measured at two different laser excitation powers. The apparent shoulder on the lowest-frequency peak in the spectrum is actually due to the presence of a plasma line at approximately 1.15 THz.

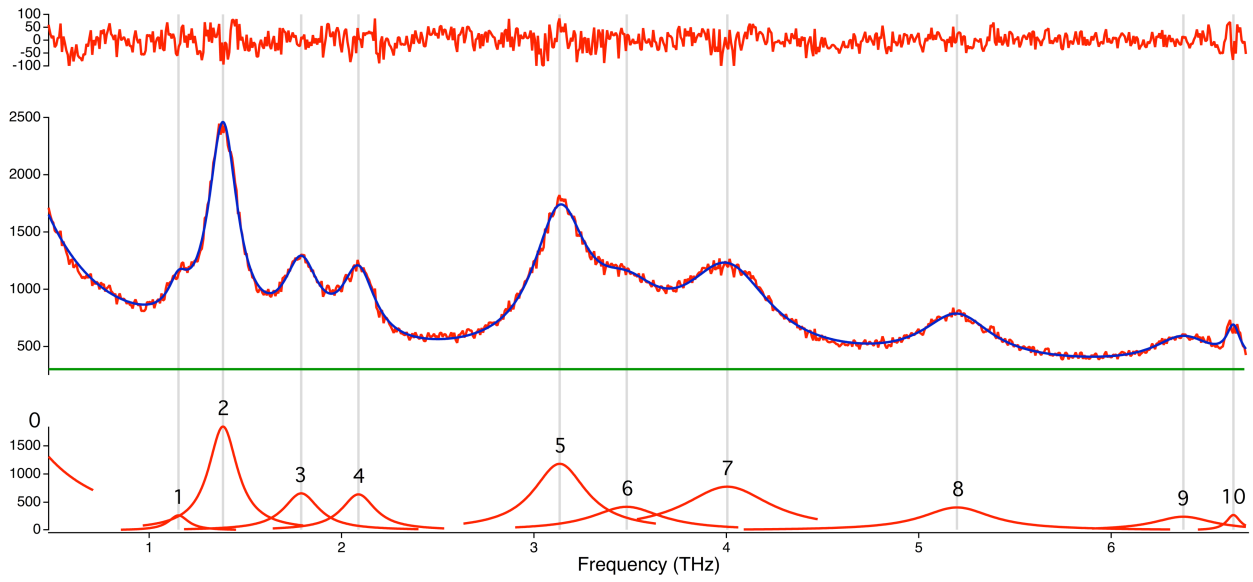

Figure S7. The positions and FWHM of peaks in the Raman spectra were obtained using a set of Lorentzian lineshapes as a model function. Prominent plasma lines (Lorentzian functions 1 & 10 above) were then subtracted from the spectra.

IV. Example of DFT Calculation Input

An example SIESTA input file ("triclinic_dlvaline_input.txt") used for our DFT calculations can be found within "calculation_files.zip."

V. Optimized Atomic Positions and Unit Cell Geometries

For the calculations performed using the vdW-DF functional (opt-B88 variant), the optimized structure can be found in the following files within "calculation_files.zip":

"dlleucine_coords.txt"

"monoclinic_dlvaline_coords.txt"

"triclinic_dlvaline_coords.txt"

VI. Calculated Vibrational Mode Frequencies and Eigenvectors

For each of the optimized structures above, the corresponding vibrational mode frequencies and eigenvectors can be found in the following files can be found in the following files within "calculation_files.zip":

"dlleucine_vectors.txt"

"monoclinic_dlvaline_vectors.txt"

"triclinic_dlvaline_vectors.txt"

Each eigenvector is reported such that every row corresponds to the three Cartesian components of the displacement for a particular atom. The order of the atoms matches that listed in the files containing the optimized structures. The vibrational frequencies given in these files are not scaled.

VII. Summary of Inter-/Intramolecular Vibrational Character

The calculated % intramolecular, % translational, and % rotational motion of low-frequency vibrations in each system (depicted in Figure 7) are tabulated in the spreadsheet file "mode_character.xls".

VIII. Quantitative Description of Vibrational Modes – Character ID Strips

The quantitative descriptions of the vibrational modes used to construct the character ID strips (Figures 9-11) are also tabulated within the "mode_character.xls" spreadsheet.

References

1. Macrae, C. F.; Bruno, I. J.; Chisholm, J. A.; Edgington, P. R.; McCabe, P.; Pidcock, E.; Rodriguez-Monge, L.; Taylor, R.; van de Streek, J.; Wood, P. A. Mercury CSD 2.0 - new features for the visualization and investigation of crystal structures. *J. Appl. Cryst.* **2008,** *41*, 466-470.
2. Williams, M. R. C.; True, A. B.; Izmaylov, A. F.; French, T. A.; Schroeck, K.; Schmuttenmaer, C. A. Terahertz Spectroscopy of Enantiopure and Racemic Polycrystalline Valine. *Phys. Chem. Chem. Phys.* **2011,** *13*, 11719-11730.
3. Craig, N. C.; Levin, I. W. Calibrating Raman Spectrometers with Plasma Lines from the Argon Ion Laser. *Appl. Spectrosc.* **1979,** *33*, 475-476.