J|A|C|S
ARTICLES
Published on Web 04/06/2006

# High-Resolution Terahertz Spectroscopy of Crystalline Trialanine: Extreme Sensitivity to $\beta$-Sheet Structure and Cocrystallized Water

Karen Siegrist,[†] Christine R. Bucher,[†] Idan Mandelbaum,[†] Angela R. Hight Walker,[†] Radhakrishnan Balu,[‡] Susan K. Gregurick,[‡] and David F. Plusquellic*,[†]

*Contribution from the Optical Technology Division, National Institute of Standards and Technology, Gaithersburg, Maryland 20899-8443, and Department of Chemistry and Biochemistry, University of Maryland, Baltimore, Maryland 21250*

Received December 14, 2005; E-mail: david.plusquellic@nist.gov

***Abstract:*** High-resolution terahertz absorption spectra (0.06−3 THz) have been obtained at 4.2 K for three crystalline forms of trialanine [$H_2^+$-(Ala)$_3$-$O^-$]. The crystal structures differ in their $\beta$-sheet forms (parallel vs antiparallel) and in their water composition (hydrated vs dehydrated antiparallel $\beta$-sheet). The spectra are nearly vibrationally resolved, with little absorption below 1 THz. In sharp contrast to observations made in the mid-IR region, the spectral patterns of all three forms are qualitatively different, illustrating the extreme sensitivity to changes in the *inter*molecular hydrogen-bonding networks that stabilize peptide crystals. Predictions obtained from a classical force field model (CHARMM) and density functional theory (DFT/PW91) for periodic solids are compared with the X-ray structural data and the terahertz absorption spectra. In general, the results for the parallel $\beta$-sheet are in better agreement with experiment than those for the antiparallel $\beta$-sheet. For all three structures, however, most hydrogen bond distances are underestimated at both levels of theory, and the predicted absorption features are significantly red-shifted for the two antiparallel $\beta$-sheet structures. Moreover, the nuclear motions predicted at the two levels of theory are qualitatively different. These results indicate that the PW91 functional is not sufficient to treat the weak intersheet hydrogen bonding present in the different $\beta$-sheet forms and strongly suggest the need for improved force field models that include three-atom hydrogen-bonding terms for periodic solids.

## I. Introduction

Terahertz (THz) vibrational spectroscopy of biomolecular systems is widely recognized for its sensitivity to force constants influencing the global nuclear motions that extend over a large fraction of the biomolecular framework. Previous studies of polypeptides,[1,2] proteins,[3−6] DNA,[7,8] sugars,[9−11] and other biomolecules[12,13] have examined the dependence of THz absorption on such factors as the degree of hydration and amorphous versus crystalline forms. However, as a result of the conformational flexibility and the variety of environmental interactions that exist in solutions and amorphous solids, the THz vibrational spectra reported for these systems are often unresolved. In contrast, the THz spectra of low-defect crystalline solids can be very sharp, with line widths $<1$ cm$^{-1}$.[9−13] The discrete absorption features of crystals characterize highly degenerate vibrational modes and reflect, to varying degrees, lattice translations, intermolecular librations, and intramolecular vibrations. Extension of such studies to crystalline peptides offers a rich but challenging area of study in the THz region. The extensive hydrogen-bonding networks that extend along the peptide backbone as well as between the zwitterionic end groups ($NH_3^+$ and $CO_2^-$) make the preparation of low-defect single crystals a slow and difficult process. Furthermore, such networks greatly impact the phonon frequencies of the lattice

[†] National Institute of Standards and Technology.
[‡] University of Maryland.

(1) Kutteruf, M. R.; Brown, C. M.; Iwaki, L. K.; Campbell, M. B.; Korter, T. M.; Heilweil, E. J. *Chem. Phys. Lett.* **2003**, *375*, 337.
(2) Yamamoto, K.; Tominaga, K.; Sasakawa, H.; Tamura, A.; Murakami, H.; Ohtake, H.; Sarukura, N. *Bull. Chem. Soc. Jpn.* **2002**, *75*, 1083.
(3) Walther, M.; Fischer, B.; Schall, M.; Helm, H.; Uhd Jepsen, P. *Chem. Phys. Lett.* **2000**, *332*, 389.
(4) Zhang, C.; Tarhan, E.; Ramdas, A. K.; Weiner, A. M.; Durbin, S. M. *J. Phys. Chem. B* **2004**, *108*, 10077.
(5) Plusquellic, D. F.; Korter, T. M.; Fraser, G. T.; Lavrich, R. J.; Benck, E. C.; Bucher, C. R.; Domench, J.; Hight Walker, A. R. In *Terahertz Sensing Technology. Volume 2: Emerging Scientific Applications and Novel Device Concepts*; Wollard, D. L., Loerop, W. R., Shur, M. S., Eds.; World Scientific: Hackensack, NJ, 2003: pp 385−404.
(6) (a) Chen, J.-Y.; Knab, J. R.; Cerne, J.; Markelz, A. G. *Phys. Rev. E Rapid* **2005**, *72*, 040901. (b) Knab, J. R.; Chen, J.-Y.; Markelz, A. G. *Biophysics* **2006**, *90*, 2576.
(7) Markelz, A. G.; Roitberg, A.; Heilweil, E. J. *Chem. Phys. Lett.* **2000**, *320*, 42.
(8) (a) Bykhovskaia, M.; Gelmont, B.; Globus, T.; Wollard, D. L.; Samuels, A. C.; Duong, T. H.; Zakrzewska, K. *Theor. Chem. Acc.* **2001**, *106*, 22. (b) Globus, T.; Bykhovskaia, M.; Wollard, D.; Gelmont, B. *J. Phys. D* **2003**, *36*, 1314.
(9) Nishizawa, J.; Suto, K.; Sasaki, T.; Tanabe, T.; Kimura, T. *J. Phys. D* **2003**, *36*, 2958.
(10) Korter, T. M.; Balu, R.; Campbell, M. B.; Beard, M. C.; Gregurick, S. K.; Heilweil, E. J. *Chem. Phys. Lett.* **2005**, *418*, 65.
(11) Walther, M.; Fischer, B. M.; Jepson, P. U. *Chem. Phys.* **2003**, *288*, 261.
(12) Korter, T. M.; Plusquellic, D. F. *Chem. Phys. Lett.* **2004**, *385* 45.
(13) (a) Fischer, B.; Hoffmann, M.; Helm, H.; Modjesch, G.; Uhd Jepsen, P. *Semicond. Sci. Technol.* **2005**, *20*, S246. (b) Rutz, F.; Wilk, R.; Kleine-Ostmann, T.; Grunenberg, J.; Koch, M. *Proceedings of the Joint 29th International Conference on Infrared and Millimeter Waves and 12th International Conference on Terahertz Electronics*; IEEE: Piscataway, NJ, 2004, p 737.


**10.1021/ja058176u CCC: $33.50** 

and the lowest intramolecular vibrational frequencies of the embedded subunits.

Because of the conformational flexibility exhibited by large biomolecules and the need for high-quality force fields to describe hydrogen-bonding networks, definitive interpretations of individual resonances in the THz region have yet to be reported. Furthermore, while high-level ab initio calculations have provided the necessary accuracy to interpret local-mode vibrational frequencies in the mid-infrared region and the shifts associated with hydrogen bonding for gas-phase peptide mimetics,[14] similar calculations for periodic solids are currently out of reach, except for the smallest systems. Even at modest levels of quantum theory (Hartree−Fock or density functional theory), the predictive accuracy in the THz region is further compromised because (i) the crystal lattice vibrational modes have substantial amounts of intermolecular or phonon character, requiring application of periodic boundary conditions for a minimalistic description of the system, (ii) the lowest frequency modes invariably involve bending and torsional degrees of freedom along "soft" potential energy surfaces, where mechanical anharmonicity is expected to be most important, and (iii) shallow potential energy surfaces are intrinsically more sensitive to the numerical accuracy of the method than the surfaces associated with local modes in the mid-IR region.

In this study, we have examined a set of peptide crystal structures that contain the same primary sequence, trialanine, but exist in different $\beta$-sheet forms. Unlike many peptide crystals, trialanine is known to exist as both parallel (*p*-$Ala_3$) and antiparallel (*ap*-$Ala_3$) $\beta$-sheets, depending on the solvent composition during crystallization. Furthermore, the *ap*-sheet of trialanine cocrystallizes with water, which (as we discovered) is easily removed under vacuum, offering a further opportunity to examine the effect of solvation on the THz vibrational spectra. A principal focus of this study is to examine the sensitivity of THz vibrational features to the character of well-defined hydrogen-bonding networks that exist in peptide crystals.

A second goal of this work is to explore the predictive quality of classical force field and density functional theory (DFT) models. Since the vibrational motions probed in the THz region represent the incipient motions for large-scale conformational changes, these studies provide some quantitative measure of the validity of the force fields used to model protein folding dynamics. Moreover, the crystalline forms of trialanine are just small enough to treat at the quantum level of theory using DFT. The results presented in this work are used to assess the accuracy of these theories for predictions of THz spectra and crystal structure and to elucidate the areas where further refinements might be needed.

## II. Experimental Section

A schematic diagram of the continuous-wave THz spectrometer is given in Figure 1. A full description of its performance for high-resolution THz laser studies is given elsewhere.[5,12,15] Briefly, the system consists of a low-temperature-grown GaAs photomixer[16] driven at the difference frequency of two near-infrared lasers. The two lasers include a fixed-frequency diode laser operating near 850 nm ($\Delta\nu_{fwhm} \approx 0.0001$ $cm^{-1}$) and a newly constructed standing-wave Ti:sapphire (Ti:Sapp) laser having a resolution of $\Delta\nu_{fwhm} \approx 0.04$ $cm^{-1}$. The Ti:Sapp laser is broadly tunable and gives more than an order-of-magnitude improvement in resolution over that of the ring laser discussed previously.[5,12] Similar to grating-tuned lasers reported elsewhere,[17] the laser is seeded by feedback from an external grating-tuned cavity through a 4% output coupler. This configuration improves the scan repeatablility of the spectrometer to better than $\pm 0.03$ $cm^{-1}$, as determined from repeated scans of water vapor in the THz region. The scan repeatability is important to minimize "noise" induced by small shifts in the standing wave interference patterns (with modulation depths that exceed 50% of the transmitted power) when background ratios are performed.

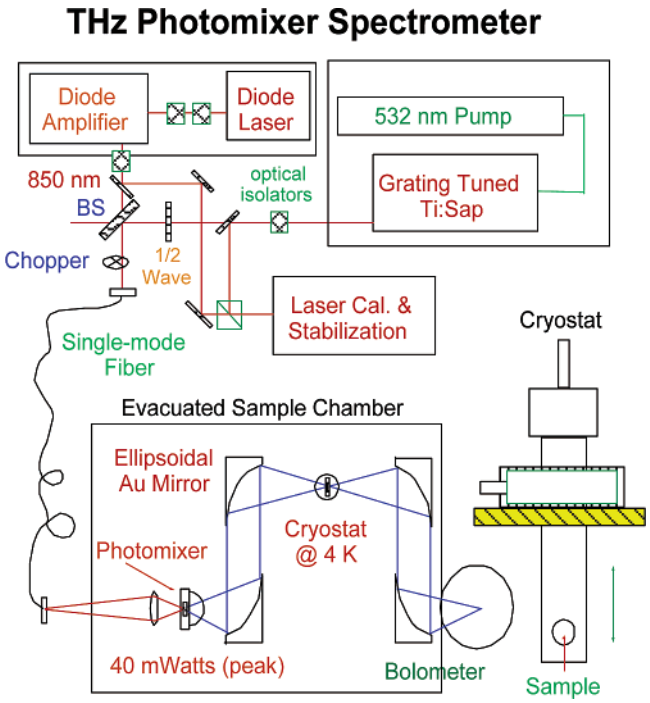


***Figure 1.*** Schematic of the high-resolution THz instrument. See text for details.

The Ti:Sapp beam is combined with the output from a fixed-frequency diode laser and focused by an aspherical lens onto the photomixer. A maximum power of $\sim 1$ $\mu$W of THz radiation is obtained from 0.06 to 3 THz, with output power decreasing as $\omega^{-4}$ beyond the peak value at 0.6 THz. The focused THz beam passes through the cryogenically cooled sample at 4.2 K and is detected by a liquid-helium-cooled silicon-composite bolometer. Power detection sensitivity of the bolometer is <1 nW up to 3 THz in a 400 Hz band-pass (noise equivalent power of the bolometer is 1 $pW/Hz^{1/2}$).

Tri-L-alanine was obtained both from Aldrich and from Bachemat in 98% purity and handled in different ways to produce the crystalline forms of the parallel and antiparallel $\beta$-sheets. Following the recipe of Hempel, Camerman, and Camerman,[18] the parallel $\beta$-sheet was recrystallized from aqueous solutions at dimethylformamide (DMF) concentrations of 50, 33, and 25% to produce platelet-type crystals. To obtain good crystals of the hemihydrated antiparallel $\beta$-sheet, which is characterized by long, needle-shaped crystals, the concentration of DMF in water was varied over a range from 0 to 18 vol %. After recrystallization, the peptide was mixed with photometric-grade polyethylene powder either by grinding with a mortar and pestle or by

(14) (a) Lavrich, R. J.; Plusquellic, D. F.; Suenram, R. D.; Fraser, G. T.; Hight Walker, A. R.; Tubergen, M. J. *J. Chem. Phys.* **2003**, *118*, 1253. (b) Chin, W.; Piuzzi, F.; Dognon, J. P.; Dimicoli, L.; Tardivel, B.; Mons, M. *J. Am. Chem. Soc.* **2005**, *127*, 11900.
(15) Siegrist, K.; Plusquellic, D. F. *Rev. Sci. Instrum.*, in preparation.
(16) (a) McIntosh, K. A.; Brown, E. R.; Nichols, K. B.; McMahon, O. B.; DiNatale, W. F.; Lyszczarz, T. M. *Appl. Phys. Lett.* **1995**, *67*, 3844. (b) Verghese, S.; McIntosh, K. A.; Brown, E. R. *Appl. Phys. Lett.* **1997**, *71*, 2743. (c) Brown, E. R. *Appl. Phys. Lett.* **1999**, *75*, 769. (d) Duffy, S. M.; Verghese, S.; McIntosh, K. A.; Jackson, A.; Gossard, A. C.; Matsuura, S. *IEEE Trans. Microwave Theory Tech.* **2001**, *49*, 1032.
(17) (a) German, K. R. *Appl. Opt.* **1981**, *20*, 3168. (b) Vieira, N. D., Jr.; Mollenauer, L. F. *IEEE J. Quantum Electron.* **1985**, *QE-21*, 195.
(18) Hempel, A.; Camerman, N.; Camerman, A. *Biopolymers* **1991**, *31*, 187.

shaking in a container with stainless steel bearings, at concentrations between 5 and 20%. A vacuum die was used to press 0.2 g of the mixture at pressures up to 69 MPa when coarse grain (~150 $\mu$m) polyethylene was used and 20 MPa when fine grade polyethylene (~4 $\mu$m) was used. Over the range of conditions used to pulverize the crystalline samples and press the pellets, we have found no evidence for pressure-dependent frequency shifts at the level of $\pm 0.1$ cm$^{-1}$. Typically, pellet disks were ~2 mm thick and 1.25 cm in diameter. Pure polyethylene (PE) disks were similarly pressed for background scans.

Two disks, one mixed sample and one pure PE, were secured by screw-in rings in the vertically stacked compartments of a brass sample holder. The brass holder, attached to a cryogenically cooled arm, was lowered into the chamber through a home-built "load-lock" system, permitting sample replacement without loss of vacuum in the main chamber. Baseline power spectra (with no sample in place) were obtained, and the position of the cryostat was adjusted to optimize the THz power detected at the bolometer. Final adjustments were made to the cryostat position after cooling the sample to 4.2 K. Background and sample spectra were obtained by raising or lowering the arm assembly by about 1 cm, without appreciably altering the incident angle of the THz beam on the sample, thereby minimizing phase shifts in the interference patterns.

The Ti:Sapp laser frequency was scanned at 0.04 cm$^{-1}$ resolution, beginning at the diode laser's frequency near 11 860 cm$^{-1}$ to a difference frequency of 100 cm$^{-1}$. The combined diode and Ti:Sapp beams were chopped at 400 Hz prior to entering the vacuum chamber containing the photomixer assembly. Both the amplified bolometer signal and a voltage proportional to the photomixer current were sampled using lock-in amplifiers. These signals were oversampled at an interval 5 times smaller than the 30 ms time constant of the bolometer to eliminate line shape distortion. The Ti:Sapp frequency is measured with a wavemeter every 0.5 s to an accuracy of 0.02 cm$^{-1}$. The absolute accuracy was confirmed by comparing water vapor lines measured at atmospheric pressure ($\Delta\nu_{\mathrm{fwhm}} \approx 0.2$ cm$^{-1}$) with predictions from the HITRAN database.[19] Typically, 10 or more scans were performed consecutively, with each scan taking approximately 8 min. Temperature equilibrium was confirmed from comparisons of consecutive scans.

Each scan was first scaled for changes in the lock-in input sensitivity, normalized to changes in the AC component of the photocurrent, and linearized according to the wavemeter scale. The data were then averaged and normalized to the transmitted power through a PE blank. Finally, the absorption spectra were obtained using a base 10 log scale.

FTIR spectra of the recrystallized samples were taken from 400 to 4000 cm$^{-1}$ at room temperature with instrument resolution set at 1 cm$^{-1}$. A 1% mixture of trialanine crystals in KBr was prepared, and a total 0.2 g of the mixture was pressed into disks using the same procedure described above to produce 1 mm thick disks. No effort was made to completely remove the residual water and $CO_2$ lines near 2350 cm$^{-1}$.

## III. Results

**A. Sample Characterization.** Crystalline trialanine may exist in one of two different $\beta$-sheet forms, depending on the solvent composition during recrystallization.[18] The X-ray crystal structures shown in Figure 2a/b and c/d differ in two important ways. The crystal structure in the top panels is that of a parallel $\beta$-sheet ($p$-Ala$_3$),[18] shown along the crystallographic $b$ (panel a) and $c$ axes (panel b). The antiparallel $\beta$-sheet ($ap$-Ala$_3$-H$_2$O)[20] structure is shown in the lower panels along the $b$ (panel c) and $c$ axes (panel d). Unlike $p$-Ala$_3$, water is cocrystallized in $ap$-Ala$_3$-H$_2$O in the 2:1 mole-to-mole ratio of Ala:H$_2$O. The side view shown in Figure 2d illustrates how water serves to bridge the

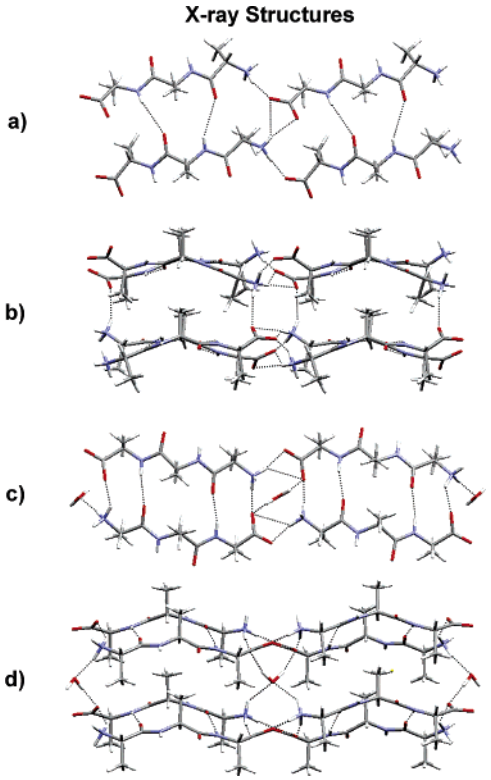


***Figure 2.*** Crystal structures of the parallel ($p$-Ala$_3$) and the antiparallel hemi-hydrated ($ap$-Ala$_3$-H$_2$O) $\beta$-sheet forms of trialanine when viewed along the (a) $b$-axis and (b) $c$-axis of $p$-Ala$_3$ and the (c) $b$-axis and (d) $c$-axis of $ap$-Ala$_3$-H$_2$O.

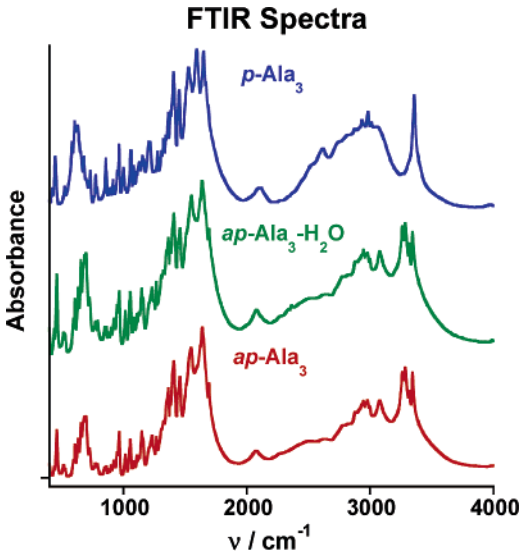


***Figure 3.*** FTIR spectra of the $p$-Ala$_3$ (top trace), $ap$-Ala$_3$-H$_2$O (middle trace), and $ap$-Ala$_3$ (lower trace) $\beta$-sheet forms of trialanine. Distinguishing features are easily seen with regard to the different $\beta$-sheet forms in the top two traces. However, little change occurs in this region upon the removal of water.

sheets via a tetrahedral network of hydrogen bonds to the zwitterionic end groups of the trialanine molecules. The water molecules in the structure are not equivalent: the water at the center is more strongly bound than those at the edges.

The mid-infrared FTIR spectra of $p$-Ala$_3$ and $ap$-Ala$_3$-H$_2$O obtained at room temperature are shown as the upper and middle traces of Figure 3, respectively. The $p$-Ala$_3$ crystals were grown from a 25% DMF/75% water solution and evaporated under

(19) Rothman, L. S.; et al. *J. Quant. Spectrosc. Radiat. Transfer* **2003**, *82*, 5.
(20) Fawcett, J. K.; Camerman, N.; Camerman, A. *Acta Crystallogr. B* **1975**, *31*, 658.

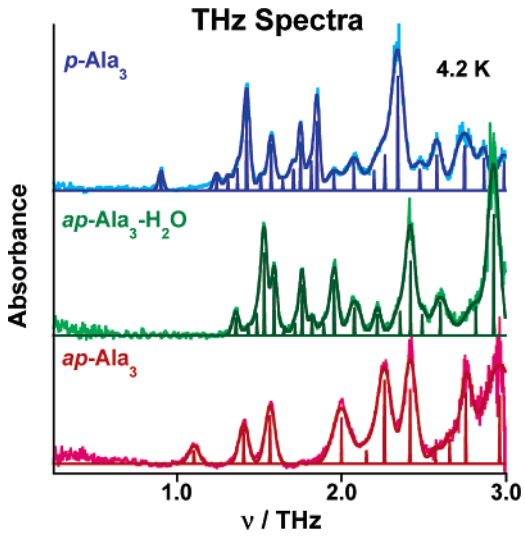


***Figure 4.*** THz spectra of the *p*-$Ala_3$ (top trace), *ap*-$Ala_3$-$H_2O$ (middle trace), and *ap*-$Ala_3$ (lower trace) $\beta$-sheet forms of trialanine. Distinguishing spectral features are seen in each of these spectra. The best-fit simulations are superimposed on the experimental data, and the line shape parameters are summarized in Table 1.

mild vacuum conditions using desiccant over a 1 week period. The *ap*-$Ala_3$-$H_2O$ crystals were recrystallized from approximately 200 mg of trialanine in 150 mL of a 15% DMF/85% water solution. The slow evaporation of solvent required about 4 weeks. The spectra in Figure 3 are in excellent agreement with the previous studies, where many of the local-mode features have been assigned.[21] While overall the FTIR spectra are quite similar, subtle but distinguishing features sensitive to the different $\beta$-sheet forms are seen in the bend and stretch regions. These features, as well as the different crystal shapes, have served to verify the $\beta$-sheet forms of our samples.

**B. Sensitivity of Terahertz Spectra to $\beta$-Sheet Structure and Cocrystallized Water.** The THz spectra of *p*-$Ala_3$ and *ap*-$Ala_3$-$H_2O$ are shown in the top and middle panels of Figure 4, respectively. They were obtained at 4.2 K over a range from 2 to 100 $cm^{-1}$. The sample concentrations were 20 and 12 wt % *p*-$Ala_3$ and *ap*-$Ala_3$-$H_2O$ in polyethylene, respectively. The spectrum of the *ap*-$Ala_3$-$H_2O$ sheet is similar to that previously reported,[22] except the features shown here are sharper and better resolved. While little absorption occurs below 1 THz in both samples, the THz spectrum of *p*-$Ala_3$ is qualitatively different from that of *ap*-$Ala_3$-$H_2O$.

To aid in further comparisons, the individual resonances in each spectrum were identified using nonlinear least-squares line shape fits where, for each line, the center frequency, $\bar{\nu}_e$, width, $\Delta\bar{\nu}_e$, and peak absorption intensity, $\kappa_e$, were iteratively varied.[12] The minimum number of features that account for the observed THz absorption intensity are summarized in Table 1. The best-fit line shapes are superimposed on the experimental data in Figure 4. We identify 22 and 19 features for *p*-$Ala_3$ and *ap*-$Ala_3$-$H_2O$ samples, respectively. The observed spectra are best fit using Gaussian line shape functions, indicating that all THz resonances are inhomogeneously broadened. Since the conformational structures of $Ala_3$ are similar in the two crystals (cf. Figure 2),[18] these spectral changes reflect the differences in the force fields associated with the vastly different network of hydrogen bonds that characterize the two $\beta$-sheet structures.

(21) Qian, W.; Bandekar, J.; Krimm, S. *Biopolymers* **1991**, *31*, 193.
(22) Shotts W. J.; Sievers, A. J. *Biopolymers* **1974**, *13*, 1593.

***Table 1.*** Center Frequencies, $\bar{\nu}_e$, Peak Absorption Coefficients, $\kappa_e$, and Gaussian Line Widths, $\Delta\bar{\nu}$, Obtained from Nonlinear Least-Squares Fits[a] of the Three $\beta$-Sheet Crystalline Forms of $Ala_3$ at 4.2 K

| | *p*-$Ala_3$ | | | *ap*-$Ala_3$-$H_2O$ | | | *ap*-$Ala_3$ | | |
|---|---|---|---|---|---|---|---|---|---|
| | $\bar{\nu}_e$ | $\kappa_e$ | $\Delta\bar{\nu}$ | $\bar{\nu}_e$ | $\kappa_e$ | $\Delta\bar{\nu}$ | $\bar{\nu}_e$ | $\kappa_e$ | $\Delta\bar{\nu}$ |
| 1 | 30.5 | 0.30 | 1.2 | 45.3 | 0.20 | 1.6 | 36.8 | 0.14 | 2.4 |
| 2 | 41.4 | 0.26 | 1.5[b] | 47.8 | 0.83 | 1.5[b] | 46.9 | 0.30 | 2.3 |
| 3 | 43.7 | 0.23 | 1.6[b] | 49.5 | 0.21 | 1.5[b] | 52.2 | 0.43 | 2.3 |
| 4 | 45.6 | 0.42 | 1.5[b] | 51.0 | 0.87 | 1.5 | 66.7 | 0.42 | 4.2 |
| 5 | 47.5 | 1.61 | 1.6 | 53.0 | 0.57 | 1.5 | 71.7 | 0.14 | 3.4[b] |
| 6 | 50.3 | 0.24 | 1.5[b] | 54.9 | 0.06 | 1.4[b] | 75.4 | 0.72 | 3.4 |
| 7 | 52.5 | 0.81 | 1.7 | 57.3 | 0.13 | 1.5[b] | 80.1 | 0.79 | 3.0 |
| 8 | 54.8 | 0.18 | 2.0[b] | 58.7 | 0.53 | 1.3 | 85.8 | 0.12 | 2.5[b] |
| 9 | 56.9 | 0.44 | 1.9 | 60.7 | 0.16 | 1.5 | 88.6 | 0.20 | 2.5[b] |
| 10 | 58.4 | 0.85 | 1.1 | 63.1 | 0.12 | 2.0[a] | 91.9 | 0.59 | 2.8 |
| 11 | 60.5 | 0.49 | 2.5 | 65.2 | 0.55 | 1.8 | 98.7 | 0.74 | 7.8 |
| 12 | 61.7 | 1.15 | 1.2 | 69.3 | 0.28 | 2.9 | | | |
| 13 | 65.2 | 0.30 | 3.1 | 73.9 | 0.23 | 1.9 | | | |
| 14 | 69.2 | 0.48 | 3.3 | 78.6 | 0.25 | 3.1 | | | |
| 15 | 73.3 | 0.34 | 3.0[b] | 80.7 | 0.83 | 2.2 | | | |
| 16 | 75.5 | 0.59 | 3.0[b] | 83.1 | 0.18 | 1.9 | | | |
| 17 | 78.1 | 0.20 | 2.7 | 86.7 | 0.31 | 3.7 | | | |
| 18 | 82.6 | 0.43 | 3.8 | 93.9 | 0.27 | 5.0[b] | | | |
| 19 | 86.0 | 0.70 | 2.3 | 97.5 | 1.34 | 2.6 | | | |
| 20 | 91.6 | 0.86 | 4.5 | | | | | | |
| 21 | 95.6 | 0.53 | 2.0 | | | | | | |
| 22 | 99.6 | 0.53 | 3.4 | | | | | | |

[a] All values in $cm^{-1}$. Uncertainties are $\pm 0.1$ $cm^{-1}$ or type B, $k = 2$ errors as determined from repeated scans. [b] Fixed in the fit at a reasonable value based on the trend.

The THz spectrum shown in the lower part of Figure 4 was also obtained from the sample pellet containing the original *ap*-sheet crystals. However, sample scans were repeated at 4.2 K after a day under vacuum at room temperature. Since the spectrum is so remarkably different from that of *ap*-$Ala_3$-$H_2O$, the initial expectation was that the sample had decomposed under these conditions. This was proved false upon reexamination of the mid-IR spectrum. The FTIR spectrum shown in the lower part of Figure 3 is nearly identical to that obtained for *ap*-$Ala_3$-$H_2O$. The same sharp features are seen to exist in both spectra, although the relative intensities over broader spectral regions near 800 and 3600 $cm^{-1}$ are somewhat different.

The reason for the change in the THz spectrum became apparent after sample weight measurements were made before and after pumping on the sample at room temperature. The weight loss of 4.8% is equivalent to the complete removal of water from the sample. The loss attributed to cocrystallized water was later confirmed by reversing the process. THz spectra identical to those of the *ap*-$Ala_3$-$H_2O$ sample were obtained following rehydration at room temperature, further confirming the chemical integrity of the *ap*-sheet. Eleven features were identified in the spectrum of dehydrated *ap*-$Ala_3$. The best-fit line shapes are superimposed on the experimental data in Figure 4, and the parameters are given in Table 1. *This single result illustrates the enormous impact that the hydrogen-bonded solvent has on the global nuclear motions probed in the THz region.* The lack of any distinct spectral signature from water in the mid-IR region is likely a result of line broadening in the OH stretch and bend regions, which may account for the diffuse

absorption differences noted above. Similar line broadening effects have been observed in the gas phase.[23]

## IV. Discussion

The three different forms of trialanine in Figure 4 serve as a benchmark system to investigate the lowest-frequency vibrational modes of a peptide crystal and to examine the impact the hydrogen-bonding network has on these vibrational modes. Two other aspects of this system further facilitate a rigorous comparison with theory. The X-ray crystal structures of the *p*-$Ala_3$ and *ap*-$Ala_3$-$H_2O$ sheets are known to high accuracy, and the 4.2 K absorption spectra are devoid of sequence level structure, i.e., the absorption resonances represent the fundamental frequencies from the zero-point levels. Therefore, these structures and spectra provide two points of comparison with theoretical models and the necessary data to systematically investigate strengths and weaknesses of the theories.

The theoretical challenge becomes apparent from the scope of the calculation required to adequately model these systems. The crystal structure of *p*-$Ala_3$ is monoclinic and belongs to space group *P*21. The unit cell contains four $Ala_3$ and has 132 atoms including 56 heavy atoms. The *ap*-$Ala_3$-$H_2O$ crystal is also monoclinic but has *C*2 space group symmetry. There are eight $Ala_3$ and four $H_2O$ in the unit cell, or 276 atoms and 116 heavy atoms. For the results presented here, all calculations were performed on a minimum of one unit cell and began with the crystallographic coordinates reported for *p*-$Ala_3$ and *ap*-$Ala_3$-$H_2O$.[18,20] Positions of hydrogen atoms were used as reported although not explicitly measured in the X-ray studies. The initial geometry and unit cell symmetry of *ap*-$Ala_3$ were those of *ap*-$Ala_3$-$H_2O$ with the waters removed. In the *P*21 and *C*2 space groups, only two of the $Ala_3$ and two of the four waters of *ap*-$Ala_3$-$H_2O$ sheet are symmetry unique. When possible, the space group symmetry was exploited to minimize computational cost.

Geometry optimizations and harmonic frequency calculations were performed on periodic systems using the CHARMM[24] and DMol[3] software[25] packages. Two theoretical approaches that proved to be unsuccessful deserve some mention first.

**A. Isolated or Periodic-like System Calculations.** Initial attempts to model the THz absorption spectra were performed on antiparallel dimers of $Ala_3$ using density functional theory [B3LYP/6-31G(d)] available in the Gaussian program suite.[26] Although theoretical calculations based on isolated systems have been successful for non-hydrogen-bonded crystal structures in the condensed phase,[27] results of efforts on hydrogen-bonded structures of $Ala_3$ proved to be problematic. As an example of the intrinsic problems with this approach, the two predicted absorption spectra of the dimer are shown in the upper panel of Figure 5. The differences in these spectra arise from subtle but important distinctions in the optimized geometries shown in the inset; the proton donor/acceptor roles of the $NH_2$ and COOH end groups are reversed. Ignoring the obvious problem that the calculation fails to reproduce the zwitterionic forms

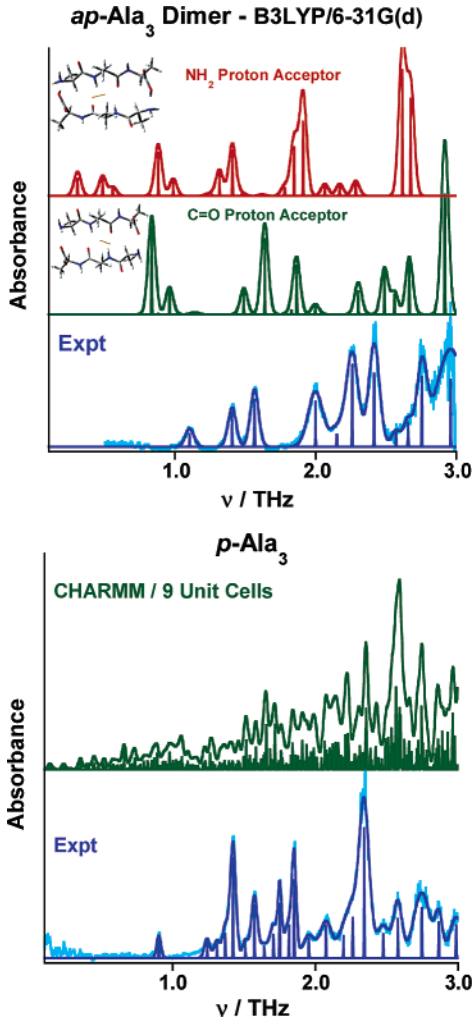


***Figure 5.*** DFT calculations of two different forms of the antiparallel sheet dimer of $Ala_3$, shown together with the observed THz spectrum of *ap*-$Ala_3$ crystal in the upper panel. In the upper part, the end groups of the dimer are linked via the COH- - -$NH_2$ hydrogen bond, while in the second structure, the roles of the proton donor and acceptor are reversed to give HNH- - -O═C hydrogen bonds. The predicted spectra are also shown and illustrate the impact of the dipole moment orientations on the THz frequencies and intensities. The CHARMM results for the periodic-like system of *p*-$Ala_3$ consisting of nine unit cells is shown in the lower panel. The predicted spectrum is severely fragmented as a result of the variety of dangling end groups orientations.

known to exist in the crystal, it is clear that maintaining the correct orientations of the end groups is of paramount importance for modeling periodic systems containing hydrogen-bonded networks. Such constraints are not possible in isolated systems.

In a second attempt to mimic the crystal environment and minimize problems with the ill-defined end group orientations, calculations were performed on "crystal-like" systems containing multiple unit cells using the CHARMM program suite. Systems ranging in size from 3 to 12 unit cells (more than 1500 atoms) were examined. The calculated spectrum for 9 unit cells of *p*-$Ala_3$ is illustrated in the lower panel of Figure 5. Similar results were obtained for the other systems examined. The spectral splittings from distorted edge group orientations lead to a near-complete "wash out" of the spectrum. It is clear that much larger systems would need to be examined using this approach to minimize the broken degeneracies associated with

(23) Florio, G. M.; Zwier, T. S.; Myshakin, E. M.; Jordan, K. D.; Sibert, E. L. *J. Chem. Phys.* **2003**, *118*, 1735 and references therein.
(24) (a) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187. (b) MacKerell, A. D.; Banavali, N.; Foloppe, N. *Biopolymers* **2000**, *56*, 257.
(25) (a) Delley, B. *J. Chem. Phys.* **1990**, *92*, 508. (b) Delley, B. *J. Chem. Phys.* **2000**, *113*, 7756.
(26) Frisch, M. J.; et al. *Gaussian 03*, revision B.04; Gaussian, Inc.: Wallingford, CT, 2003.
(27) Heilweil, E. J., private communication.

***Table 2.*** Results of CHARMM Calculations for the Periodic Systems of *p*-$Ala_3$, *ap*-$Ala_3$-$H_2O$, and *ap*-$Ala_3$

| lattice | $\Delta a, \Delta b, \Delta c, \Delta\beta$ | energy (kcal/mol) unit cell | $\Delta E^b$ | average[c] (Å) | max[c] (Å) |
|---|---|---|---|---|---|
| | | *p*-$Ala_3$ | | | |
| X-ray[a] | | −479.232 | − | 0.114/0.184 | 0.326/0.480 |
| Opt | 0.072, 0.041, 0.142, 0.53 | −480.242 | 1.010 | 0.125/0.189 | 0.352/0.500 |
| | | *ap*-$Ala_3$-$H_2O$ | | | |
| X-ray[a] | | −535.640 | − | 0.296/0.304 | 1.159/1.159 |
| $H_2O$ only | | | | 0.268/0.292 | 0.356/0.419 |
| Opt | 0.549, −0.208, 0.507, 0.79 | −540.090 | 4.450 | 0.517/0.538 | 0.856/0.952 |
| $H_2O$ only | | | | 0.406/0.466 | 0.556/0.816 |
| | | *ap*-$Ala_3$ | | | |
| 8 $Ala_3$[e] | | −886.19 | − | | |
| 16 $Ala_3$[f] | | −899.73[d] | 13.54 | | |

[a] Differences specified relative to X-ray lattice parameters: $a = 11.849$ Å, $b = 10.004$ Å, $c = 9.862$ Å, $\beta = 101.30°$ for *p*-$Ala_3$ and $a = 18.513$ Å, $b = 5.330$ Å, $c = 24.775$ Å, $\beta = 98.64°$ for *ap*-$Ala_3$-$H_2O$. [b] Differences are specified relative to the previous value. [c] Average and maximum values of distances between X-ray and calculated atom coordinates for the heavy atoms/all atoms of the unit cell. [d] Total energy of the 16-$Ala_3$ unit cell divided by 2. [e] Optimized cell parameters are $a = 18.092$ Å, $b = 5.153$ Å, $c = 24.634$ Å, $\beta = 99.18°$. [f] Optimized cell parameters are $a = 18.058$ Å, $b = 10.297$ Å, $c = 24.458$ Å, $\beta = 99.33°$.

the dangling end groups. The advantage of this approach is that three-atom hydrogen-bonding terms may be included in CHARMM in contrast to the period system calculations described below.

**B. CHARMM Calculations on Periodic Systems.** To eliminate problems associated with finite systems sizes, periodic system calculations were performed using the CHARMM 22 empirical force field.[24] The CHARMM/TIP3P parameter set was used to model the cocrystallized water in the *ap*-$Ala_3$-$H_2O$ $\beta$-sheet. Periodic structures were created using features of the software and consisted of a set of image cells surrounding a core block (typically one unit cell). The number of image cells generated depended on the core atom cutoff radius for including pairwise nonbonding interactions. Changes in energy for radii larger than 15 and 25 Å for *p*-$Ala_3$ and *ap*-$Ala_3$-$H_2O$, respectively, were negligible and used for all subsequent calculations. We note that three-atom terms to describe hydrogen bonds between core and image cells are not permitted in CHARMM. Therefore, all intermolecular hydrogen bonds (including intracell bonds) are modeled with two-atom terms. While this limitation is rather severe for the peptide crystal calculations considered here, it is also noted that incorporating such interactions presents an extreme theoretical challenge. Finally, it was also important to ensure convergence in the electrostatic energy sums for periodic systems using the Ewald approximation.[28] The Gaussian width of 0.34 was used for transitioning from real to **k**-space electrostatic sums, and the number of grid points in Fourier space was set to ~2 times the unit cell dimensions.

The normal-mode frequencies were obtained from the eigenvalues of the mass-weighted Hessian matrix[29] derived using a two-point finite difference method. The atomic coordinate displacements used were ±0.01 Å. The intensity of the vibrational fundamental, $v' = v'' + 1$, of each normal mode, $\mathbf{Q}_i$, is proportional to the square of the transition dipole moment derivative, $\delta\boldsymbol{\mu}_i/\delta\mathbf{Q}_i$, and numerically evaluated according to

$$I_i \propto \left|\left\langle v'\left|\frac{\partial\boldsymbol{\mu}_i}{\partial\mathbf{Q}_i}\right|v''\right\rangle\right|^2 \cong \left|\sum_j^{3N}\frac{e_j\, q_{i,j} - e_j^\circ\, q_{i,j}^\circ}{\Delta Q_i}\right|^2 \cong \left|\sum_j^{3N}\frac{e_j^\circ(q_{i,j} - q_{i,j}^\circ)}{\Delta Q_i}\right|^2 \quad (1)$$

where the sum is over all 3*N* Cartesian coordinate displacements, $q_j - q_j^\circ$, derived from the eigenvectors of the Hessian by dividing by square-root of the atomic masses, and $\Delta Q_i$ represents the root-mean-square displacement of atoms in mode $\mathbf{Q}_i$, arbitrarily set to 1 in the final approximation of eq 1. Since CHARMM assigns a static partial charge, $e_j^\circ$, according to atom type regardless of bond length, the electrical anharmonicity is by definition 0, making the latter approximation necessary. Equation 1 yields the integrated intensity; therefore, for comparisons with experiment, we have used the same Gaussian width in each line shape function. We note that these frequency calculations are strictly at the harmonic level. Improvements that include mechanical anharmonicity are currently underway.

Two checks were performed to confirm that the optimized structures of the unit cell conformed to their space group symmetries. First, the atomic coordinates were compared between the symmetry-unique and symmetry-equivalent $Ala_3$ monomers after transformation (at most a $C_2$ rotation) and translation to the center-of-mass of each monomer. The average coordinate differences were always near zero, confirming that the space group symmetry was maintained following optimization. It is also important to note that the proper orientation of the core block prior to optimization was critical to arrive at this result (especially for the waters of *ap*-$Ala_3$-$H_2O$). The use of factional coordinates followed by the internal conversion to orthogonal Cartesian coordinates proved to be the most reliable way to guarantee conformity with the space group symmetry. Second, calculations were repeated after increasing the core block size by 2-fold or more along each crystallographic axis. The scaling of the computed energies with size was nearly exact to within the convergence criteria of $10^{-6}$ kcal/mol. The predicted spectra were also identical after normalizing the intensities to the number of unit cells making up the core block.

Calculations were performed with and without optimizing the four lattice parameters, *a*, *b*, *c*, and $\beta$. The results from these

(28) (a) Ewald, P. *Ann. Phys*. **1921**, *64*, 253. (b) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577. (c) Allen, M. P.; Tildesley, D. J. *Computer Simulations of Liquids*; Oxford Press: New York, 1989.
(29) Wilson, E. B., Jr.; Decius, J. C.; Cross, P. C. In *Molecular Vibrations: The Theory of Infrared and Raman Vibrational Spectra*; McGraw-Hill: New York, 1955.

***Table 3.*** Results of DMol Calculations for the Periodic Systems of *p*-$Ala_3$, *ap*-$Ala_3$-$H_2O$, and *ap*-$Ala_3$[a]

| functional | condition | energy (hartrees) unit cell | $\Delta E$ ($\times 10^{-3}$) | average[g] (Å) | max[g] (Å) |
|---|---|---|---|---|---|
| | | *p*-$Ala_3$ | | | |
| **PW91** | $10^{-6}$ Ha[b] | −3274.0999338 | −[f] | 0.057/0.138 | 0.157/0.462 |
| **PW91** | $10^{-5}$ Ha[b] | −3274.0998969 | −0.037 | 0.064/0.142 | 0.188/0.476 |
| **PW91** | **k**-point[c] | −3274.1020013 | 2.067 | 0.057/0.139 | 0.154/0.457 |
| **PBE** | $10^{-6}$ Ha[b] | −3271.3874757 | −2712 | 0.066/0.142 | 0.187/0.463 |
| **PW91**[d] | $R_c = 3.7$[e] | −3274.0999338 | −[f] | 0.057/0.138 | 0.157/0.462 |
| **PW91**[d] | $R_c = 4.7$[e] | −3274.1088925 | 8.959 | 0.064/0.143 | 0.193/0.469 |
| PW91[d] | $R_c = 5.7$[e] | −3274.1096619 | 0.769 | 0.067/0.145 | 0.215/0.469 |
| PW91[d] | $R_c = 6.7$[e] | −3274.1101423 | 0.480 | 0.066/0.144 | 0.201/0.470 |
| PW91[d] | $R_c = 7.7$[e] | −3274.1101343 | −0.0079 | 0.066/0.144 | 0.198/0.471 |
| | | *ap*-$Ala_3$-$H_2O$ | | | |
| **PW91**[d] | $R_c = 3.7$[e] | −6854.2214485 | −[f] | 0.209/0.286 | 0.599/0.754 |
| | $H_2O$ only | | | 0.132/0.166 | 0.191/0.223 |
| PW91[d] | $R_c = 4.7$[e] | −6854.2439329 | 22.48 | 0.207/0.280 | 0.560/0.715 |
| | $H_2O$ only | | | 0.140/0.173 | 0.207/0.236 |
| | | *ap*-$Ala_3$ | | | |
| PW91[d,h] | $R_c = 3.7$[e] | −6548.3043171 | −[f] | | |
| **PW91**[d,h] | $R_c = 4.7$[e] | −6548.3314366 | 27.12 | | |

[a] For *p*-$Ala_3$, energies are given for different functionals and convergence criteria and for a range of atomic cutoff radii, $R_c$. Predicted spectra for the entries in bold font are shown in the figures. [b] Convergence criteria: $10^{-5}$ hartrees for energy, $2 \times 10^{-3}$ Ha/Å for maximum gradient, and $5 \times 10^{-3}$ Å for maximum displacement; and $10^{-6}$ hartrees for energy, $2 \times 10^{-4}$ Ha/Å for maximum gradient, and $5 \times 10^{-4}$ Å for maximum displacement. [c] **k**-point mesh used for energy was $2 \times 2 \times 2$. [d] Tighter convergence criteria used as given in footnote *a*. [e] Numerical atomic radius cutoff radius in Å. [f] Energy differences are specified relative to the first value. For the different $R_c$ values, differences are specified relative to the previous value. [g] Average and maximum values of distances between X-ray and calculated atom coordinates for heavy atoms/all atoms of the unit cell. Differences are specified relative to X-ray cell parameters: $a = 11.849$ Å, $b = 10.004$ Å, $c = 9.862$ Å, $\beta = 101.30°$ for *p*-$Ala_3$ and $a = 18.513$ Å, $b = 5.330$ Å, $c = 24.775$ Å, $\beta = 98.64°$ for *ap*-$Ala_3$-$H_2O$. [h] Unit cell parameters used for *ap*-$Ala_3$ were fixed at those of *ap*-$Ala_3$-$H_2O$.

optimizations are given in Table 2. For both the *p*-$Ala_3$ and *ap*-$Ala_3$-$H_2O$ crystal structures, the average distances between the calculated and experimental atom positions increased when lattice parameters were permitted to vary. For the *p*-$Ala_3$ sheet, the average distance increased moderately by ∼0.01 Å/heavy atom. The total energy decreased by ∼1 kcal/mol relative to the energy of the fixed lattice, and the four lattice parameters decreased by at most 1.5%. The situation was much worse for *ap*-$Ala_3$-$H_2O$. The cell dimensions changed by 2−4% for the full relaxed lattice (∼4.5 kcal/mol lower in energy), and the average distance increased to ∼0.5 Å/heavy atom versus the fixed lattice value of ∼0.3 Å/heavy atom. Given these trends, the lattice parameters were held fixed at the X-ray values for the CHARMM calculations discussed below.

**C. DFT Calculations on Periodic Systems.** We have used the DMol[3] program suite[25] to perform DFT calculations on periodic systems of trialanine. All calculations included all electrons and made use of a double numerical plus polarization (DNP) atomic orbital basis set (equivalent to a double-ζ) and two different functionals described below. Since there is currently no computationally efficient means to optimize the crystal cell parameters, the lengths, *a*, *b*, and *c*, and angle, *β*, of the monoclinic cells were held fixed at the values reported from the X-ray studies.

The harmonic frequencies and normal-mode displacements were calculated at the Γ-point by a two-point numerical differentiation of the forces, requiring six energy calculations for each atom in the unit cell. Of the 132/268/276 atoms of *p*-$Ala_3$, *ap*-$Ala_3$, and *ap*-$Ala_3$-$H_2O$, only 66, 66, and 72 atoms are symmetry unique. Significant savings in computational cost (2- to 8-fold) were realized by generating the energy derivatives of the symmetry-equivalent atoms from those of the symmetry-unique atoms after transformation according to the operations of the space group. These energy derivatives were then used to construct the full second-derivative (Hessian) matrix of the unit cell. The normal-mode frequencies and displacements were obtained following diagonalization of the mass-weighted Hessian. Numerical errors propagated using this procedure were found to be negligible based on comparisons of results that included all 132 atoms of *p*-$Ala_3$.

The current software release of DMol[3] does not support the calculation of the atomic polar tensors ($\delta\mathbf{E}/\delta\boldsymbol{\mu}\delta\mathbf{Q}$) for periodic systems needed to predict the infrared intensities. Therefore, the intensity for each normal mode was calculated using the last approximation given in eq 1. The static charges were set to the Mulliken atomic charges of the optimized structure. Intensity calculations that include the electrical anharmonicity given in eq 1 are currently underway. All vibrational lines were convoluted with a Gaussian line shape function of fixed width.

Two different functionals were tested for their accuracy in predictions of the crystal structure and THz spectrum of the *p*-$Ala_3$ sheet. The exchange-correlation functional based on the generalized gradient approximation (GGA) in the parametrization of Perdew and Wang (PW91)[30] was selected for its known reliability to accurately model hydrogen-bonded systems. The second functional tested was PBE.[31] The average distances between the predicted atom positions and the X-ray structural data are given in Table 3 for the heavy atoms only and when the hydrogen atoms are included. Compared to that of the PBE functional, the optimized PW91 geometry is in better agreement with the X-ray data by nearly 0.01 Å/heavy atom. The THz spectra predicted using the PBE and PW91 functionals are shown as traces A and E of Figure 6, respectively. Clearly, the

(30) Perdew, J. P.; Wang, Y. *Phys. Rev. B* **1992**, *45*, 13244.
(31) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865.

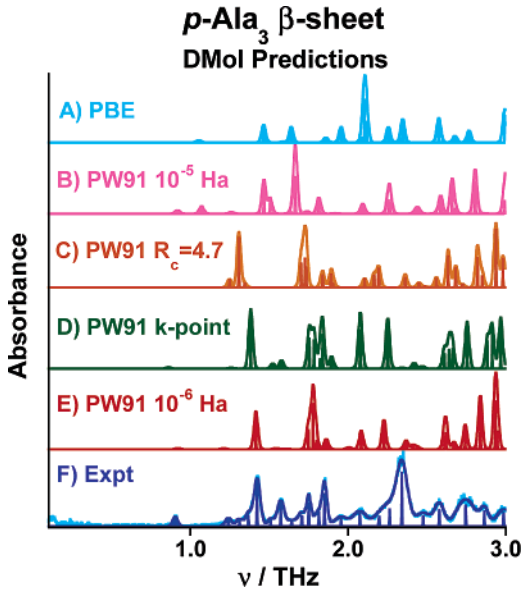


***Figure 6.*** THz spectral predictions obtained from DFT calculations using DMol[3] and (A) the PBE functional, (B) the PW91 functional with the moderate (default) optimization tolerances, (C) the PW91 functional with an atomic orbital cutoff radius of $R_c = 4.7$ Å, (D) the PW91 functional and a $2 \times 2 \times 2$ **k**-point mesh, (E) the PW91 functional with a 10-fold increase in the optimization criteria, and (F) the experimental spectrum of crystalline *p*-Ala$_3$. See text for details.

spectral predictions of the latter are also in better agreement with the experimental THz spectrum.

Given the numerical nature of the theoretical approach implemented in DMol[3], a few of the numerical parameters were explored to ensure the adequate convergence of the structure and total energy. Most of these tests were performed on *p*-Ala$_3$ because of its smaller size. First, the sensitivity of the THz frequencies to the geometry convergence criteria is illustrated in Figure 6. Trace B is the predicted spectrum obtained from the minimum energy structure obtained under the following convergence criteria: $10^{-5}$ hartrees (Ha) for energy, $2 \times 10^{-3}$ Ha/Å for maximum gradient, and $5 \times 10^{-3}$ Å for maximum displacement. The spectrum obtained after decreasing each of these criteria by 10-fold is shown as trace E. The later results are in better agreement with experiment, as are the geometries (see Table 3). It is clear from these results that the force constants are remarkably sensitive to small changes in the equilibrium geometry. Therefore, the tighter tolerances were used in all subsequent calculations. Second, the atomic orbital cutoff radius was varied over the range from 3.7 to 7.7 Å at intervals of 1 Å; the results are summarized in Table 3. In general, the energies improve with increasing $R_c$ values up to 6.7 Å. The largest change of ~9 mHa occurs at $R_c = 4.7$ Å. However, the discrepancies with the X-ray geometry increased for the larger cutoff values and the spectral predictions for $R_c = 4.7$ Å (trace C of Figure 6) are somewhat red-shifted relative to experiment. Therefore, the default value of $R_c = 3.7$ Å was used, except where otherwise noted.

Finally, calculations were also performed on an unshifted $2 \times 2 \times 2$ **k**-point mesh over the Brillouin zone and compared with the Γ-point calculation (**k** = 0 in reciprocal space). The results are given in Table 3 and the predicted spectrum is shown as trace D in Figure 6. The energy is only slightly lower (≈2 mHa) than the energy obtained for the Γ-point. Furthermore,

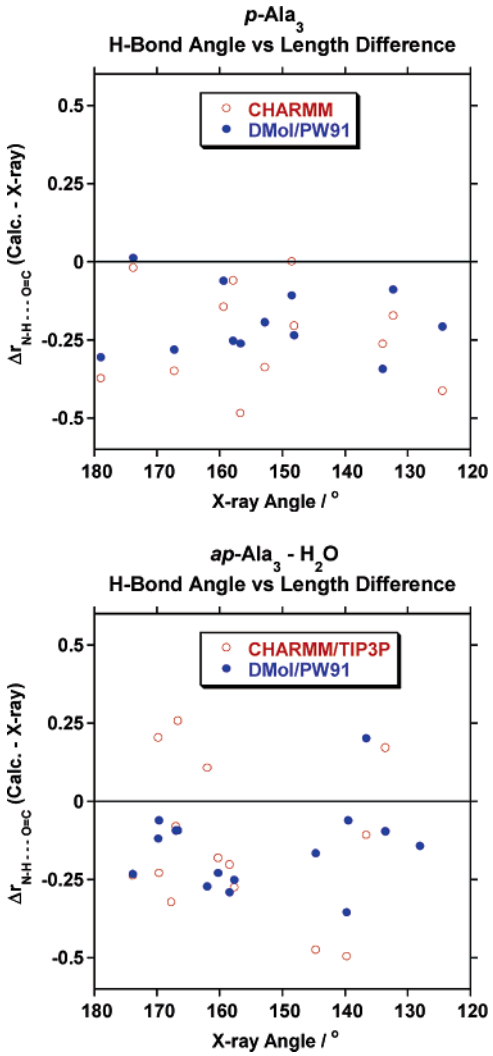


***Figure 7.*** Differences in the [O- - -(H)—N] bond lengths obtained from CHARMM (open circles) and DMol[3] (filled circles) relative to the X-ray data. These are shown versus the X-ray [O- - -H—N] angles for the 12 hydrogen bonds of the *p*-Ala$_3$ and the 16 hydrogen bonds of *ap*-Ala$_3$-H$_2$O crystal structures.

the atomic coordinate error and the THz spectrum are also very similar to the Γ−point results. Because of the larger computational cost of the **k**-point calculations, the later approximation was used throughout this work.

**D. Structural Predictions from CHARMM and DMol[3].** The crystal structures from the DFT calculations are generally better than those predicted by CHARMM. From Tables 2 and 3, the average error in the atomic coordinates is 0.057 Å/heavy atom for *p*-Ala$_3$ sheet, which is nearly half that of the best CHARMM structure. For *ap*-Ala$_3$-H$_2$O, the average error at the DFT level increases by nearly 3-fold to 0.20 Å/heavy atom relative to that of the parallel sheet, but it is still more accurate than the CHARMM structure by more than 50%.

As mentioned above, a major limitation of CHARMM for periodic systems is that all intermolecular interactions are described by two-atom terms. Given the absence of an angular dependence of these interaction energies, it is interesting to compare the structural properties of the hydrogen bonds predicted by CHARMM with the X-ray data and the DFT results. The separations between O and N atoms of the hydrogen bonds relative to the crystallographic data are shown in Figure

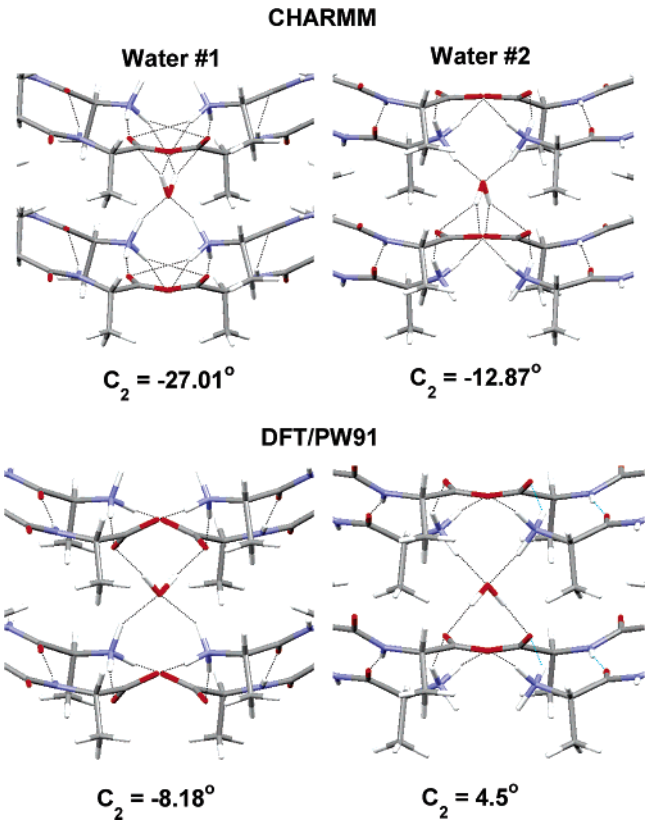


***Figure 8.*** Detailed views of the two nonequivalent water molecules in the optimized structures predicted by CHARMM (top) and DMol[3] (bottom). The twist angles are given relative to the X-ray crystal data.

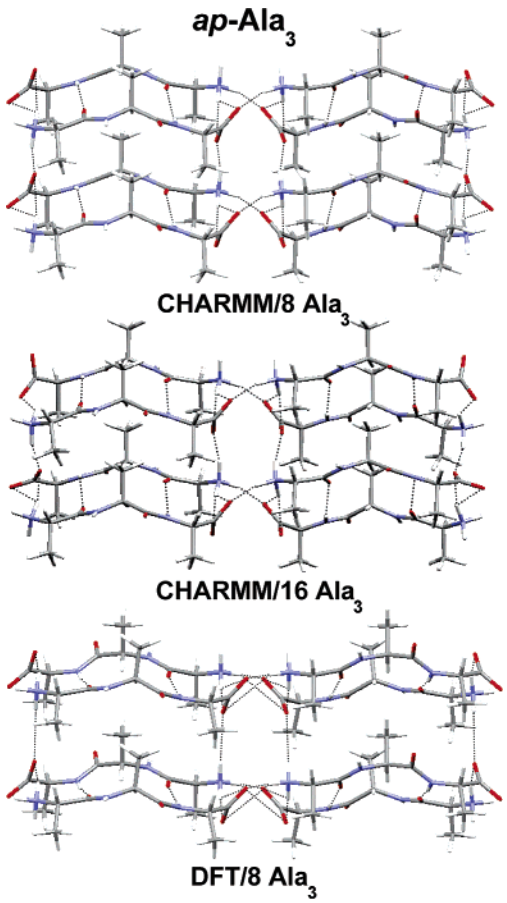


***Figure 9.*** Optimized *ap*-$Ala_3$ structures from CHARMM for 8-$Ala_3$ (top) and 16-$Ala_3$ unit cells (middle). The optimized structure from DMol[3] for the 8-$Ala_3$ unit cell is shown at the bottom. The principal difference between the two CHARMM structures is the rotation of two $CO_2$ groups near the outer edges of the structure in the top sheet.

7 as a function of the [C═]O- - -H—N angles, where a linear hydrogen bond corresponds to an angle of 180°. For nearly all of the 12 hydrogen bonds of the *p*-$Ala_3$ $\beta$-sheet shown in the top panel, the heavy atom separations at both levels of theory are smaller than experiment by up to 0.5 Å. Although the differences predicted by CHARMM are generally larger than those from DMol[3], there is at best only a small increase in $\Delta r$ with increasing nonlinearity, indicating that the absence of an angular-dependent energy term is not a dominant source of error in the force field calculations. For the 16 hydrogen bonds of the *ap*-$Ala_3$-$H_2O$ structure shown in the lower panel, the errors from CHARMM are larger than those of *p*-$Ala_3$ and also increase slightly with the degree of nonlinearity. The smaller DFT errors relative to CHARMM may be attributed to better structural predictions for water, as will be discussed next.

The geometries of the cocrystallized water in *ap*-$Ala_3$-$H_2O$ are of particular interest in these comparisons. Expanded portions of the structures calculated by CHARMM and DMol[3] are shown in the upper and lower panels of Figure 8, respectively. Problems associated with the TIP3P water model are immediately apparent in the top panels, where the $sp^3$ tetrahedral arrangement of hydrogen bonds is rather poorly reproduced. The two waters are twisted about their $C_2$ axes by −27° and −13°, and the oxygen atoms of water are displaced by an average distance of 0.27 Å relative to the X-ray data (see Table 3). This rotation appears to be energetically driven by the formation of two nonlinear hydrogen bonds with each proton donor of water and perhaps further facilitated by the absence of the $sp^3$-orientated lone pairs on the oxygen atom. In contrast, quantum theory is capable of accounting for the polarizibility and directional preferences of the proton donors and lone pairs of water; consequently, the predicted geometries are in much better agreement with the crystal data. The average structural error is roughly half that of CHARMM, and the angle errors are only −8° and +4°.

The crystal structure of the dehydrated form of the antiparallel $\beta$-sheet, *ap*-$Ala_3$, is not known. However, given the reversibility of the hydration process, large changes in structure are very unlikely. Therefore, initial calculations were performed beginning with the X-ray crystal structure of *ap*-$Ala_3$-$H_2O$ with the waters removed. The optimized structure from CHARMM is shown in the upper panel of Figure 9. The cell geometry maintains the *C*2 space group symmetry of the *ap*-$Ala_3$-$H_2O$ $\beta$-sheet.

Repeating the CHARMM calculations on structures doubled in size along the *b*-axis (i.e., a core block containing 16 $Ala_3$) resulted in the optimized geometry shown in the middle panel of Figure 9. The $CO_2$ groups in every other sheet and row are rotated by >20° relative to the eight-$Ala_3$ unit cell. Furthermore, the total energy of this larger cell was found to be more stable by ∼27 kcal/mol when compared to twice the energy of the smaller cell. Further doubling of the core block size did not produce any new structures.

Initial optimizations performed using DMol[3] on the unit cell containing eight $Ala_3$ always produced two imaginary frequencies for $R_c$ = 3.7 Å. However, a fully optimized geometry of

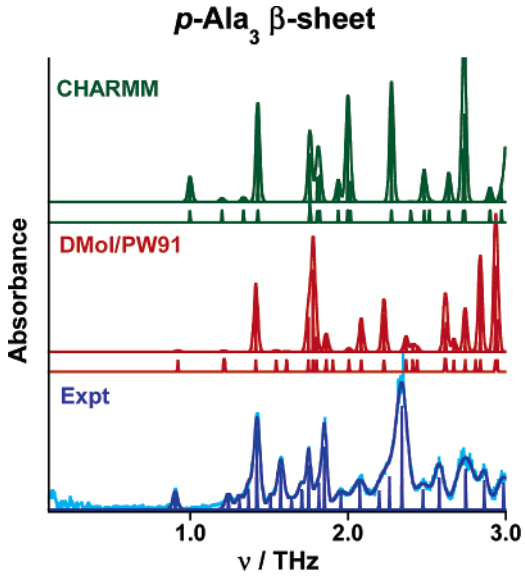


***Figure 10.*** THz spectra calculated from CHARMM (top) and DFT/PW91 (middle) compared to the observed spectrum of *p*-$Ala_3$ (bottom).

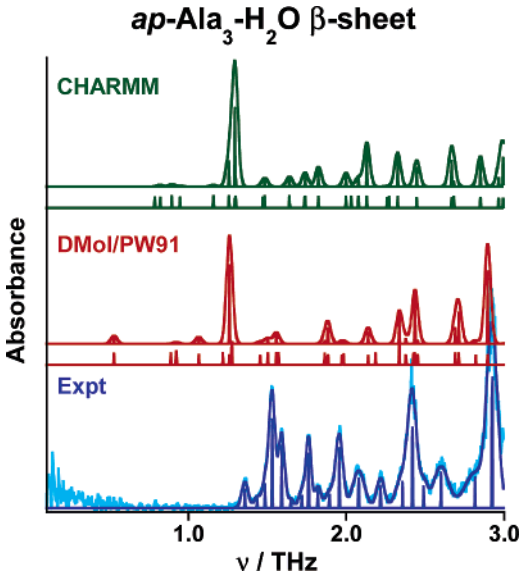


***Figure 11.*** THz spectra calculated from CHARMM (top) and DFT/PW91 (middle) compared to the observed spectrum of *ap*-$Ala_3$-$H_2O$ (bottom).

any periodic system should never lead to a transition-state structure, even if the unit cell dimensions are incorrect. Rather, this result indicates that the removal of water destabilizes the structure and softens the potential energy surfaces, making it more difficult to numerically locate the energy minimum. After increasing the cutoff radius to $R_c = 4.7$ Å, the equilibrium structure was successfully located and shown in the lower panel of Figure 9. Like the CHARMM structure, the cell geometry conforms to the *C*2 space group. Unfortunately, DFT calculations on the larger cell containing 528 atoms are not currently possible with DMol$^3$, which is limited to systems having less than 500 atoms.

**E. Spectral Predictions from CHARMM and DMol$^3$.** The THz spectral predictions from CHARMM and DMol$^3$ are shown together with the experimental data of the *p*-$Ala_3$ $\beta$-sheet in Figure 10. Given the rather crude approximations used to calculate intensities, the positions of all modes that are infrared-active by symmetry are shown beneath each simulation. Below 1.5 THz, the spectral predictions at both levels of theory are in fair agreement with experiment where, for example, the predicted locations of two transitions are near the experimental features at 0.9 and 1.4 THz. However, other weaker transitions in the CHARMM spectrum are missing near 2.0 THz, and the DMol$^3$ intensities are only poorly estimated in this region.

The spectral predictions for *ap*-$Ala_3$-$H_2O$ are shown in Figure 11. Despite the significant structural differences noted above for water, the CHARMM and DMol$^3$ spectra are again similar with regard to the overall intensity distribution and total number of features below 3 THz. However, at both levels of theory, the spectra are significantly red-shifted by as much as 0.25 THz relative to experiment, and neither level of theory adequately reproduces the observed intensity pattern.

The theoretical THz spectra of *ap*-$Ala_3$ for the three structures in Figure 9 are illustrated in Figure 12. In all three cases, the features appear to be significantly red-shifted relative to experiment. Of the two CHARMM predictions, the spectrum of the eight-$Ala_3$ unit cell shares more in common with the experimental data than that calculated for the lower-energy cell with 16 $Ala_3$. Furthermore, the calculated spectrum from DMol$^3$

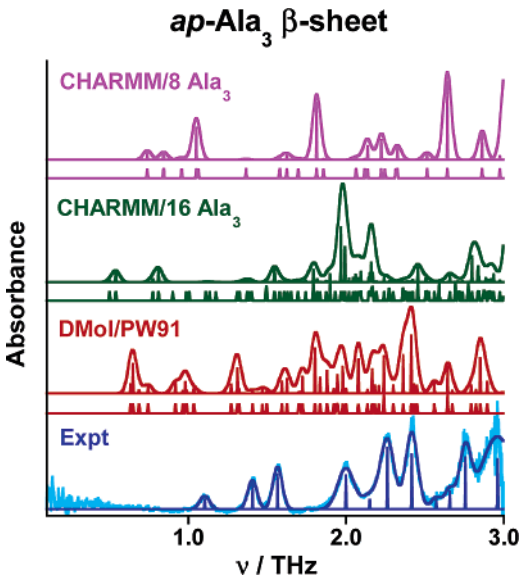


***Figure 12.*** THz spectra calculated from CHARMM for unit cells consisting of 8 (top) and 16 (middle) $Ala_3$ and from DMol$^3$ for an 8-$Ala_3$ unit cell. The observed spectrum of *ap*-$Ala_3$ is shown as the lower trace.

is more like that of the smaller CHARMM structure and perhaps best accounts for many of the observed features. These arguments would suggest that the changes in the end group orientations that occur in the larger CHARMM cell may not be correct.

Given the likeness of the spectral patterns predicted by CHARMM and DMol$^3$ for the *p*-$Ala_3$ and *ap*-$Ala_3$ $\beta$-sheet structures, one might also expect the nuclear motions associated with these modes to be similar in character. To test this assertion, the dot products were formed between the normalized eigenvectors, $\varphi$, of DMol$^3$ and CHARMM using

$$DP^2_{m,n} = \left| \sum_j^{3N} \phi^{DMol}_{j,m} \, \phi^{CHARMM}_{j,n} \right|^2 \times 100\% \qquad (2)$$

where the sum over all 3*N* eigenvectors of CHARMM for a given DMol$^3$ eigenvector equals 100%. The results for *p*-$Ala_3$

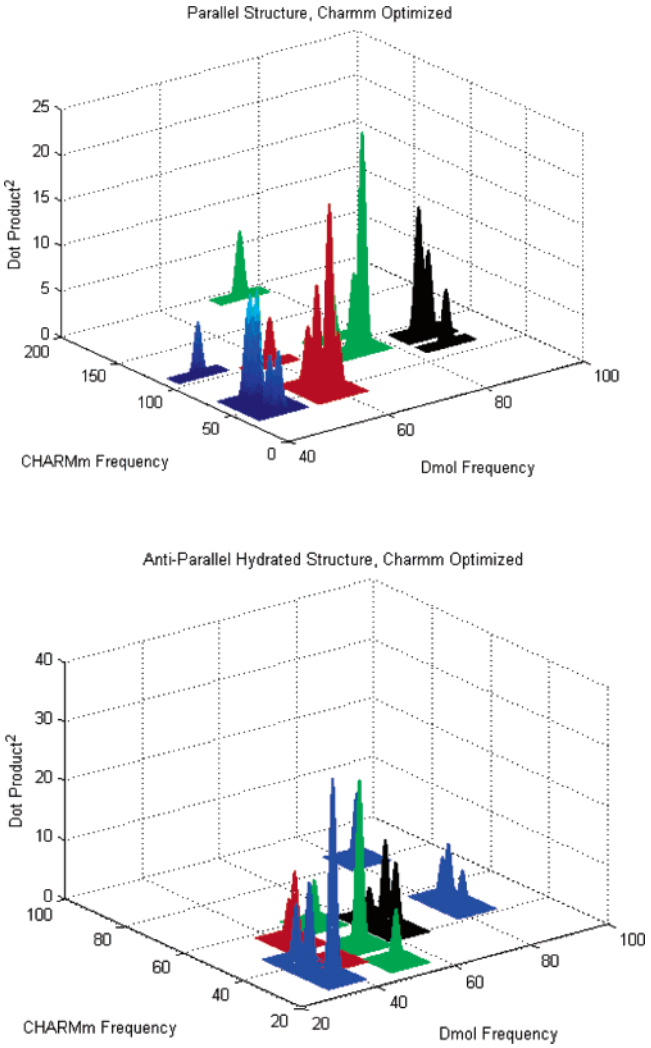


***Figure 13.*** Eigenvector dot product analysis applied to some of the prominent features present in the predicted spectra of *p*-$Ala_3$ and *ap*-$Ala_3$-$H_2O$. For each $DMol^3$ mode shown on the right, the percent composition of each CHARMM eigenvector is shown when these components exceed 5%. The plots illustrate that, in all cases, the nuclear motions predicted at these two levels of theory are highly mixed.

and *ap*-$Ala_3$-$H_2O$ are shown in the upper and lower panels of Figure 13 for a few of the most prominent features below 3 THz. Components less than 5% are not shown. Notice that, for each $DMol^3$ mode, significant mixing is found with three or more modes of CHARMM. Furthermore, the maximum component of any given CHARMM mode is less than 20%. These results indicate that the similarity exhibited in the spectral patterns at these two levels of theory is coincidental and that nuclear motions of lowest-frequency modes share very little in common for the two $\beta$-sheet structures.

We speculate that the disagreement between the calculated and experimentally measured spectra of the two $\beta$-sheet structures can be attributed to a breakdown in the harmonic approximation and/or a breakdown in the description of the intersheet peptide hydrogen-bonding interactions. Due to the size of the calculations, we have resorted to a harmonic analysis of the vibrational frequencies. However, previous ab initio calculation of the anharmonic vibrational spectrum of gas-phase peptides, which included hydrogen-bonded waters, found that for the soft, skeletal vibrations, deviations due to anharmonicity can be as large as 11%.[32] This can lead to an anharmonic shift in some of our calculated frequencies from 3 to 10 $cm^{-1}$ for both the DFT and empirical calculations. Nevertheless, anharmonic effects alone cannot account for the larger disagreements between the empirically calculated and experimental spectra, particularly in the cases of the antiparallel $\beta$-sheet samples.

Since the largest discrepancies with experiment are found for the antiparallel $\beta$-sheet forms, deficiencies in CHARMM and DFT may be linked to the strength of the intersheet binding. As can be seen from Figure 1, the parallel structure has four hydrogen bonds that cross-link two $Ala_3$ in one sheet with two $Ala_3$ in an adjacent sheet. In contrast, there are on average only two hydrogen bonds that bridge the antiparallel $\beta$-sheets, and for *ap*-$Ala_3$-$H_2O$, each of these is mediated by water. Therefore, the weaker intersheet binding of the antiparallel $\beta$-sheet structures likely softens the intermolecular surfaces, making estimations of the equilibrium structure and THz spectrum more susceptible to errors.

As mentioned earlier, a known deficiency in CHARMM is related to the nature of the empirical description of hydrogen bonding. Our current calculations utilized the CHARMM22 force field to describe both the internal forces of the peptide system and the intersheet interactions of the peptides and the peptide−water systems. It is well known that CHARMM22 does not explicitly include hydrogen-bonding terms per se in periodic system calculations but rather represents this interaction based on van der Waals and electrostatic terms, which have been parametrized to reproduce hydrogen-bonding geometries and propensities of smaller molecules. Recent studies on the hydrogen-bonding properties of *N*-methylacetimide with water and the statistical analysis of known protein structures suggests that hydrogen bonding should include both a radial and an orientational component.[33−36] We are currently investigating the effects of anharmonic coupling on the underlying spectroscopy and the influence of orientational degrees of freedom in hydrogen bonding on the energetics and the spectroscopy.

One final aspect of the experimental data may be related to the strengths of intersheet hydrogen-bonding networks of the three peptide crystals. The line widths reported in Table 1 are seen to increase with frequency, consistent with an inhomogeneous broadening mechanism.[12] Notice also that the widths increase across the series from *p*-$Ala_3$ to *ap*-$Ala_3$-$H_2O$ to *ap*-$Ala_3$. With the removal of water from the *ap*-sheet, so is lost the connectivity between sheets. Consequently, one or more of the $CO_2^-$ and $NH_3^+$ groups must rotate to some degree to reestablish some hydrogen bond connectivity between sheets, but at the expense of the stability of the intrasheet network. Therefore, the increasing inhomogeneous widths across this series may reflect weaker hydrogen-bonding networks and/or larger system disorder of the antiparallel sheet structures. Clearly, the shallow intermolecular potential surfaces of the $\beta$-sheet structures pose extreme theoretical challenges for ab initio methods and a demand for improved force field models that include hydrogen-bonding interactions for periodic systems.

(32) Gerber, R. B.; Brauer, B.; Gregurick, S. K.; Chaban, G. M. *Phys. Chem. Commun.* **2002**, *6*, 142.
(33) Chapman, M. S.; Fabiola, F.; Bertram R.; Korostelev, A. *Protein Sci.* **2002**, *11*, 1415.
(34) Baker, D.; Morozov, A. V.; Kortemme, T.; Tsemekhman, K. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 6946.
(35) Baker, D.; Kortemme, T.; Morozov, A. V. *J. Mol. Biol.* **2003**, *326*, 1239.
(36) Karplus, M.; Buck, M. *J. Phys. Chem.* **2001**, *105*, 11000.

## V. Conclusion

We have obtained vibrationally resolved THz spectra of three $\beta$-sheet forms of crystalline trialanine at 4.2 K, including one parallel and two antiparallel $\beta$-sheet forms that differ by the presence or absence of cocrystallized water. Both the *p*-$Ala_3$ and *ap*-$Ala_3$-$H_2O$ $\beta$-sheets have been well characterized in the mid-IR region and by X-ray crystallography, while the dehydrated *ap*-sheet has not been reported previously. The THz spectrum of the *p*-$Ala_3$ shares little in common with the spectra of either *ap*-sheet, reflecting differences in the strength and character of the hydrogen-bonding networks of these crystals. Furthermore, since water is only weakly (and reversibly) bound in the *ap*-sheet, we show for the first time the remarkable sensitivity of THz absorption features to dehydration, which stands in sharp contrast to observations made in the mid-IR region.

Such well-characterized periodic systems have permitted a detailed investigation of the quality of both classical (CHARMM) and quantum (DFT/PW91) theoretical models for predictions of crystal structure and THz spectra. While the results at both levels of theory are generally in better agreement with experiment for the *p*-$Ala_3$ $\beta$-sheet, most hydrogen bond distances are underestimated for both structures, and the normal-mode THz frequencies are significantly red-shifted relative to experiment for the antiparallel forms. Furthermore, the nuclear motions at these two levels of theory share little in common. Although the exact reasons for the discrepancies with experiment are not yet clear, possible sources of error include the neglect of vibrational anharmonicity and, at the classical level, the absence of an orientation dependence of the hydrogen-bonding interactions. It is further emphasized that the classical force field parameters and the DFT functionals[37] are based on best-fit descriptions of small molecules and hydrogen-bonded clusters. The competitive nature of the hydrogen-bonding interactions present in these crystal structures may require more emphasis on contributions from many-body effects.[38]

**Acknowledgment.** Financial support was provided to S.K.G. from UMBC in the form of a Special Research Initiative Award and the ADVANCE program, grant SBE-0244880 (NSF). I.M. and K.S. acknowledge the National Research Council for postdoctoral fellowship support.

**Supporting Information Available:** Complete refs 19 and 26. This material is available free of charge via the Internet at http://pubs.acs.org.

JA058176U

(37) Staroverov, V. N.; Scuseria, G. E. *J. Chem. Phys.* **2003**, *119*, 12129 and references therein.
(38) Guo, H.; Gresh, N.; Roques, B. P.; Salahub, D. R. *J. Phys. Chem. B* **2000**, *104*, 9746.