J|A|C|S
ARTICLES

Published on Web 12/04/2008

# Terahertz Spectra and Normal Mode Analysis of the Crystalline VA Class Dipeptide Nanotubes

Hailiang Zhang,[†] Karen Siegrist,[‡,§] David F. Plusquellic,[‡] and Susan K. Gregurick*[,†,¶]

*Department of Chemistry and Biochemistry, University of Maryland, Baltimore County, Baltimore, Maryland 21250, and Biophysics Group, Physics Laboratory, National Institute of Standards and Technology, Gaithersburg, Maryland 20899-8443*

Received July 17, 2008; E-mail: greguric@umbc.edu

***Abstract:*** Terahertz (THz) vibrational modes are characterized by nonlocal, collective molecular motions which are relevant to conformational changes and molecular functions in biological systems. We have investigated the THz spectra of a set of small bionanotubes which can serve as very simple models of membrane pores, and have examined the character of the THz modes which can impact transport processes. In this work, THz spectra of the crystalline VA class dipeptide nanotubes were calculated at both the harmonic and vibrational self-consistent field (VSCF) level using the CHARMM22 force field with periodic boundary conditions. Comparison of the calculated THz spectra against the experimental spectra revealed that the VSCF corrections generally improved the predictions in the low-frequency region. The improvements were especially manifested in the overall blue-shifts of the VSCF frequencies relative to the harmonic values, and blue shifts were attributed to the overall positive coupling strengths in all systems. Closer examination of the motions in the most significantly coupled normal mode pairs leads us to propose that, when two similar side-chain squeezing modes are coupled, the rapidly increased van der Waals interactions can lead to a stiffening of the effective potential, which in turn leads to the observed blue-shifts. However, we also noted that when the side-chain atoms become unphysically proximate and the van der Waals repulsion becomes too large, the VSCF calculations tend to deviate in the high frequency region and for the system of L-isoleucyl-L-valine. In addition, normal-mode analysis revealed a series of channel-breathing motions in all systems except L-valyl-L-alanine. We show that the inner products of the backbone vibrations between these channel-breathing motions divided the remaining VA class dipeptide systems into two subgroups. It is suggested that these modes may facilitate a pathway for the guest molecule absorption, substitution and removal in the VA class dipeptide nanotubes. Normal mode analysis also demonstrated that the THz motions may contribute to the pore permeability either directly by changing the pore size, or indirectly by affecting the solvent−host effective potentials.

## 1. Introduction

Terahertz (THz) spectroscopy, which investigates the low-frequency region of the vibrational spectrum, is used here to study vibrational spectra below 3 THz (∼100 $cm^{-1}$). In biological systems, THz absorbance normally originates from the collective motions associated with large-scale conformational changes, and thus THz radiation is a natural tool for probing the mechanisms important to biological function. Recent work in this rapidly growing field has proven its potential for unraveling picosecond dynamics of biomolecular interactions,[1−4] identifying conformations,[5] and imaging biological processes.[6] In much of the recent work directed at developing theoretical understanding of the lowest-frequency vibrational modes of biomolecules, crystalline solids have played an important role.[7−9] The low-frequency vibrational modes or phonon modes of crystalline systems can give rise to sparse, narrow, well-resolved spectral features, providing spectra ideally suited for development and validation of theoretical models. Progress has been demonstrated in reproducing the experimentally observed

[†] University of Maryland.
[‡] National Institute of Standards and Technology.
[§] Present Address: Electro-optical and Infrared Systems and Technologies Group, Johns Hopkins Applied Physics Laboratory, 11100 Johns Hopkins Rd., Laurel, MD 20723.
[¶] Present Address: Office of Biological and Environmental Research, U.S. Department of Energy, Germantown, MD 20874.

(1) Ebbinghaus, S.; Kim, S. J.; Heyden, M.; Yu, X.; Heugen, U.; Greubele, M.; Leitner, D. M.; Havenith, M. *Proc. Nat. Acad. Sci. U.S.A.* **2008**, *104*, 20749–20752.
(2) Heugen, U.; Schwaab, G.; Brundermann, E.; Heyden, M.; Yu, X.; Leitner, D. M.; Havenith, M. *Proc. Nat. Acad. Sci. U.S.A.* **2006**, *103*, 12301–12306.
(3) Heyden, M.; Brundermann, E.; Heugen, U.; Niehues, G.; Leitner, D. M.; Havenith, M. *J. Am. Chem. Soc.* **2008**, *130*, 5773–5779.
(4) Ebbinghaus, S.; Kim, S. J.; Heyden, M.; Yu, X.; Greubele, M.; Leitner, D. M.; Havenith, M. *J. Am. Chem. Soc.* **2008**, *130*, 2374–2375.
(5) Whitmire, S. E.; Wolpert, D.; Markelz, A. G.; Hillebrecht, J. R.; Galan, J.; Birge, R. R. *Biophys. J.* **2003**, *85*, 1269–1277.
(6) Masson, J. B.; Sauviat, M.-P.; Martin, J.-L.; Gallot, G. *Proc. Nat. Acad. Sci. U.S.A.* **2006**, *103*, 4808–4812.
(7) Jepsen, P. U.; Clark, S. J. *Chem. Phys. Lett.* **2007**, *442*, 275–280.
(8) Allis, D. G.; Fedor, A. M.; Korter, T. M.; Bjarnason, J. E.; Brown, E. R. *Chem. Phys. Lett.* **2007**, *440*, 203–209.
(9) Siegrist, K.; Bucher, C. R.; Mandelbaum, I.; Hight-Walker, A. R.; Balu, R.; Gregurick, S. K.; Plusquellic, D. F. *J. Am. Chem. Soc.* **2006**, *128*, 5764–5775.


**10.1021/ja805581n CCC: $40.75** 

vibrational spectra for biomolecular crystalline systems, especially at *ab initio* levels of theory. Contributions from several groups have shown reasonable agreement between computational and experimental results, and in particular, these efforts have clarified the necessity of imposing periodic boundary conditions to achieve sufficiently realistic models of strongly hydrogen-bonded crystalline systems[7−10] at very low frequencies. For modeling complex biological systems, however, it is ultimately desirable to develop reliable models at the level of empirical force field calculations. Furthermore, in our own previous work, the radical impact of small changes in hydrogen bonding on low-frequency spectra of trialanine[9] suggested that similarities in the character of THz modes might exist for systems having similar hydrogen-bonding networks. Consequently, we have used empirical force field models to investigate a set of well-characterized dipeptide nanotubes which share very similar hydrogen-bonding networks.

The most straightforward method to model the collective motions of biomolecular systems is by resorting to the harmonic approximation. However, the collective motions found at THz frequencies are generally dominated by soft noncovalent interactions such as van der Waals and hydrogen-bonding interactions, where the anharmonic effects, intermode couplings and vibrational energy resonance (VER)[11,12] are not necessarily negligible. In this situation, methodologies that incorporate the anharmonic effect and intermode coupling are required. Vibrational self-consistent field (VSCF) is one such methodology that has been successfully applied to various systems, such as BPTI with rigid-body constraints,[13] $(Ar)_{13}$ clusters,[14] isolated peptide−water complexes,[15] glucose,[16] *N*-methylacetamide,[17] $H_2SO_4 \cdot H_2O$ systems,[18,19] intermediates in the PYP photocycle,[20] and peptides by employing the predicted important coupling (PIC) criterion.[21] More recently, VSCF methods have been applied to various crystalline systems by employing periodic boundary conditions. These systems include HY in the rare-gas matrix where Y is an electronegative group,[22] and iodine in an argon matrix.[23] However, due to the high computational demands for evaluation of the intermode couplings, the VSCF methods are typically restricted to small or moderate crystalline systems. The present research utilized the VSCF method to model the THz spectra of a larger system, the crystalline VA class dipeptide nanotubes. The calculated spectra were compared against the experimental THz spectra in crystalline phase. In addition, a normal-mode analysis was carried out in order to investigate the dynamical factors that contribute to the structural flexibility and pore permeability of the VA class dipeptides. In the present VSCF calculation, up to 3300 atoms without constraints were included with periodic boundary conditions; to the best of our knowledge, this is the first time a condensed phase VSCF calculation on the large biomolecules has been performed.

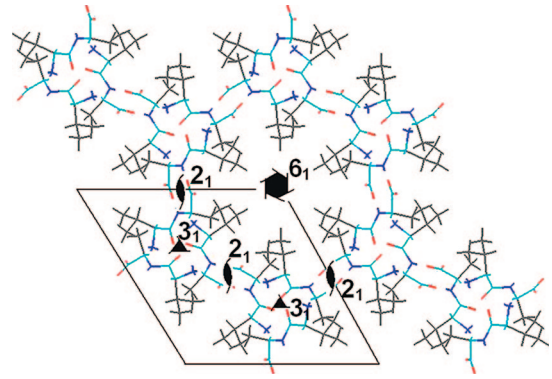


***Figure 1.*** Crystal structure of L-valyl-L-alanine viewed along the *c*-axis. The rhombus encompasses a unit cell. The side chains are shown in gray. The symmetric operations are shown in black.

**VA Class Dipeptide Nanotubes.** Since Ghadiri and co-workers pioneered the study of cyclic D,L-peptide-based nanotubes with 8−12 residues,[24] bionanotubes formed by small peptide building blocks have been the subject of considerable research efforts.[25−27] Recently, bionanotubes constructed from even smaller components, i.e. dipeptides having two hydrophobic residues, were found to form microporous structures stable to the admission and removal of guest molecules such as water. In 2007, Görbitz reviewed and summarized the structural patterns of all the available hydrophobic dipeptide nanotubes.[28] It was found that for combinations of L-alanine (Ala), L-valine (Val) or L-isoleucine (Ile), including L-alanyl-L-isoleucine (Ala-Ile or AI), L-isoleucyl-L-alanine (Ile-Ala or IA), L-alanyl-L-valine (Ala-Val or AV), L-valyl-L-alanine (Val-Ala or VA), L-isoleucyl-L-valine (Ile-Val or IV), L-valyl-L-isoleucine (Val-Ile or VI) and L-valyl-L-valine (Val-Val or VV), the crystalline dipeptides share essentially the same structural pattern. These seven systems constitute a unique group of isostructural nanotubes with identical three-dimensional hydrogen-bond networks.[29] These were termed the VA class of dipeptide nanotubes, which was named after Val-Ala, the first example of a microporous dipeptide structure.[30] All the VA class dipeptide crystals have hexagonal symmetry and belong to $P6_1$ Hermann Mauguin space group (Figure 1), where the side chains cluster around the hexagonal axis $\mathbf{6_1}$, forming the inner surface of the nanotube channels. The framework of each channel constitutes a double helix that is linked by consecutive $-NH_3^+ \cdots {}^-OOC-$ head-to-tail hydrogen bonds. The interface between two adjacent channels at $\mathbf{2_1}$ axis are connected by hydrogen bonds from the

(10) Allis, D. G.; Korter, T. M. *ChemPhysChem* **2006**, *7*, 2398–2408.
(11) Fujisaki, H.; Bu, L.; Straub, J. E. In *Normal Mode Analysis: Theory and Applications to Biological and Chemical Systems*; Chapman and Hall/CRC Press: Boca Raton, FL, 2005.
(12) Moritsugu, K.; Miyashita, O.; Kidera, A. *Phys. Rev. Lett.* **2000**, *85*, 3970–3973.
(13) Roitberg, A.; Gerber, R. B.; Elber, R.; Ratner, M. A. *Science* **1995**, *268*, 1319–1322.
(14) Jung, J. O.; Gerber, R. B. *J. Chem. Phys.* **1996**, *105*, 10682.
(15) Gregurick, S. K.; Fredj, E.; Elber, R.; Gerber, R. B. *J. Phys. Chem. B* **1997**, *101*, 8595–8606.
(16) Gregurick, S. K.; Liu, J. H.-Y.; Brant, D. A.; Gerber, R. B. *J. Phys. Chem. B* **1999**, *103*, 3476–3488.
(17) Gregurick, S. K.; Chaban, G. M.; Gerber, R. B. *J. Phys. Chem. A* **2002**, *106*, 8696–8707.
(18) Miller, Y.; Chaban, G. M.; Gerber, R. B. *J. Phys. Chem. A* **2005**, *109*, 6565–6574.
(19) Miller, Y.; Chaban, G. M.; Zhou, J.; Asmis, K. R.; Neumark, D. M.; Gerber, R. B. *J. Chem. Phys.* **2007**, *127*, 094305.
(20) Adesokan, A. A.; Pan, D.; Fredj, E.; Mathies, R. A.; Gerber, R. B. *J. Am. Chem. Soc.* **2007**, *129*, 4584–4594.
(21) Pele1, L.; Gerber, R. B. *J. Chem. Phys.* **2008**, *128*, 165105.
(22) Gerber, R. B. *Annu. Rev. Phys. Chem.* **2004**, *55*, 55–78.
(23) Bihary, Z.; Gerber, R. B.; Apkarian, V. A. *J. Chem. Phys.* **2001**, *115*, 2695.
(24) Ghadiri, M. R.; Granja, J. R.; Milligan, R. A.; McRee, D. E.; Khazanovich, N. *Nature* **1993**, *366*, 324–327.
(25) Hartgerink, J. D.; Clark, T. D.; Ghadiri, M. R. *Chem. Eur. J.* **1998**, *4*, 1367–1372.
(26) Zhang, S.; Marini, D.; Hwang, W.; Santoso, S. *Curr. Opin. Chem. Biol.* **2002**, *6*, 865–871.
(27) Gao, X.; Matsui, H. *Adv. Mater.* **2005**, *17*, 2037.
(28) Görbitz, C. H. *Chem. Eur. J.* **2007**, *13*, 1022–1031.
(29) örbitz, C. H. *New J. Chem.* **2003**, *27*, 1789–1793.
(30) Görbitz, C. H.; Gundersen, E. *Acta Crystallogr., Sect. C* **1996**, *52*, 1764–1767.

second hydrogen of $-NH_3^+$ to amide >C=O, as well as from amide >NH and $>C^{\alpha}H$ to $^-OOC-$. The junction between three channels at $\mathbf{3_1}$ axis are connected by hydrogen bonds from the third hydrogen of $-NH_3^+$ to amide >C=O.

Depending on the size of side chains, the van der Waals diameters of VA class nanotubes range from ~3 to ~6 Å, and the channel shapes vary significantly from the narrow, flattened channel of Ala-Ile to a partially obstructed, zigzag channel of Ile-Val.[28] Despite this, all the VA class dipeptide systems share an extraordinary property that certain solvent guest molecules can be absorbed and removed from the channels without impairing the structure of hydrogen-bonded networks.[28] This was illustrated in a striking manner for Ala-Val by substitution of acetonitrile from crystallization with methanol and further substitution of methanol with 2-propanol.[28,31] Interestingly, the retroanalogue dipeptide Val-Ala was found to be impermeable to the larger guest alcohol molecules although its pore is comparable in size to that of Ala-Val. These phenomena indicate that VA class dipeptide nanotubes are rather flexible and stable, but of varying permeability.

The sorption behavior study of this class may serve as a simple model for an important biological process, which is molecular transport through a cell membrane. It is well-known that membrane transport is facilitated by specialized proteins, e.g. ion channels or water pores (aquaporin), which share common functional characteristics such as high specificity and transport rates.[32,33] The dynamical nature of transport phenomena cannot be understood from static crystal structures alone; therefore, in the past decade, transport mechanisms have been investigated using computational methods.[32−37] For example, Allen et al. showed that water molecules can permeate small hydrophobic pores by forming highly correlated single chains threading down the nanotube,[34] de Groot and Grubmuller attributed high rates of water transport through the hydrophobic pore of Aquaporin-1 to the collective motions of the intrapore water,[35] and Hummer et al. reported pulse-like transmission of water through hydrophobic carbon nanotubes.[36] The coordinated motions described by these results are exactly the type of collective behavior characterizing vibrations in the THz region, wherein the motion of intracrystalline guest molecules is inherently coupled to the motions of the surrounding crystal structure. We suggest that investigation of the character of THz vibrations of the VA class dipeptide nanotubes, which share the peptide helical framework, can aid in comprehending the transport dynamics of more complex structures such as ion channels and aquaporins. The present research implemented the THz calculations of all seven of the VA class dipeptide systems and investigated the structural flexibility and pore permeability by employing normal-mode analysis.

**VSCF Methodology.** Vibrational self-consistent field (VSCF) theory is one of the most important methodologies for vibrational calculations that incorporate anharmonic and intermode coupling corrections.[38−40] From the Born−Oppenheimer approximation, which separates the motion of the nuclei and the electrons, the exact nuclear Schrödinger equation is:

$$\left[-\frac{\hbar^2}{2}\sum_{j=1}^{N}\frac{\partial^2}{\partial Q_j^2}+V(Q_1,...,Q_N)\right]\Psi(Q_1,...,Q_N)=E\Psi(Q_1,...,Q_N) \tag{1}$$

where $N$ is the total number of modes in the system, $Q$ represents the normal coordinate, and $V(Q_1,...,Q_N)$ is the full potential energy. In most cases, especially for the floppy systems, due to anharmonicity and coupling between different normal modes, the full potential energy generally is not separable, and thus eq 1 is not analytically solvable. VSCF methodology utilizes the Hartree approximation, wherein the wave function is assumed to be factorable in the following manner:

$$\Psi(Q_1,...,Q_N)=\prod_{k=1}^{N}\psi_k(Q_k) \tag{2}$$

Then, eq 2 leads to $N$ single VSCF Schrödinger equations that can be easily solved in the following form:

$$\left[-\frac{\hbar^2}{2}\frac{\partial^2}{\partial Q_k^2}+V_k^n(Q_k)-\varepsilon_k\right]\psi_k^n(Q_k)=0 \tag{3}$$

where $V_k^n(Q_k)$ is the effective potential of mode $k$ that takes the form of:

$$V_k^n(Q_k)=\left\langle\prod_{l\neq k}^{N}\psi_l^n(Q_l)\middle|V(Q_1,...,Q_N)\middle|\prod_{l\neq k}^{N}\psi_l^n(Q_l)\right\rangle \tag{4}$$

VSCF calculation generally starts from the wave functions based on the harmonic approximation, and the effective potentials, energies and the wave functions are iteratively solved in a self-consistent fashion. The VSCF approximation for the total energy is given by:

$$E_{total}^n=\sum_{k=1}^{N}\varepsilon_k^n-(N-1)\left\langle\prod_{k=1}^{N}\psi_k^n\middle|V(Q_1,...,Q_N)\middle|\prod_{k=1}^{N}\psi_k^n\right\rangle \tag{5}$$

The accuracy of VSCF calculation can be improved by applying second-order perturbation theory, assuming that there are no degenerate excited states:

$$E_n^{cc}=E_n^{VSCF}+\sum_{m\neq n}\frac{\left|\left\langle\prod_{k=1}^{N}\psi_k^n(Q_k)\middle|\Delta V\middle|\prod_{k=1}^{N}\psi_k^n(Q_k)\right\rangle\right|^2}{E_n^0-E_m^0} \tag{6}$$

where the perturbation, $\Delta V$, is the correlation between each normal mode:

$$\Delta V(Q_1,...,Q_N)=V(Q_1,...,Q_N)-\sum_{k=1}^{N}V_k^n(Q_k) \tag{7}$$

Numerically, the most time-consuming step in the VSCF process is the multidimensional integral performed in evaluating the effective potential of eq 4. For computational manageability, the three-body and higher-order interactions are neglected in this approach, and it is assumed that only pairwise interactions contribute significantly to the correlations between all normal modes. Finally, one has:

(31) Görbitz, C. H. *Acta Crystallogr., Sect. B* **2002**, *58*, 849–854.
(32) Fujiyoshi, Y.; Mitsuoka, K.; de Groot, B. L.; Philippsen, A.; Grubmüller, H.; Agre, P.; Engel, A. *Curr. Opin. Struct. Biol.* **2002**, *12*, 509–515.
(33) Beckstein, O.; Sansom, M. S. P. *Phys. Biol.* **2004**, *1*, 42–52.
(34) Allen, T. W.; Kuyucak, S.; Chung, S.-H. *J. Chem. Phys.* **1999**, *111*, 7985–7999.
(35) De Groot, B. L.; Grubmuller, H. *Science* **2001**, *294*, 2353–2357.
(36) Hummer, G.; Rasaiah, J. C.; Noworyta, J. P. *Nature* **2001**, *414*, 188–190.
(37) Pomes, R.; Roux, B. *Biophys. J.* **1998**, *75*, 33–40.
(38) Bowman, J. M. *J. Chem. Phys.* **1978**, *68*, 608.
(39) Gerber, R. B.; Ratner, M. A. *Chem. Phys. Lett.* **1979**, *68*, 195–198.
(40) Gerber, R. B.; Ratner, M. A. *Adv. Chem. Phys.* **1988**, *70*, 97–132.

$$V(Q_1, ..., Q_N) \approx \sum_{i=1}^{N} V_i(Q_i) + \sum_{i<j}^{N} V_{ij}(Q_i, Q_j) \quad (8)$$

where:

$$V_i(Q_i) = V(0, ...Q_i, ..., 0) \quad (9)$$

represents the diagonal potential of mode $i$, and:

$$V_{ij}(Q_i, Q_j) = V(0, ...Q_i, ..., Q_j, ..., 0) - V(0, ...Q_i, ..., 0) - V(0, ...Q_j, ..., 0) \quad (10)$$

represents the pairwise coupling potential between modes $i$ and $j$. In this way, evaluation of the $N$-dimensional integral in eq 4 is reduced to $(N - 1)$ two-dimensional integrals, and consequently, the computational expense is significantly reduced.

## 2. Methods

**2.1. THz Experiments.** A brief synopsis of the operation of the continuous wave THz spectrometer diagrammed schematically in Figure 2 is given here and described in detail elsewhere.[41,42] In the present work, the spectrometer includes a tunable Ti:sapphire laser where frequency selection is achieved by coupling through an end mirror to an external grating-tuned cavity. The 14 cm wide, 1200 line/mm gold-coated grating is configured to scan >100 $cm^{-1}$ at a frequency resolution of 0.06 $cm^{-1}$ (1.6 GHz). The spectrometer employs an upgraded ErAs:GaAs photomixer,[43] which provides up to 5 times the THz output obtained from the preceding generation of photomixers.[9] Output from a fixed frequency diode laser near 11800 $cm^{-1}$ is combined with the output from the tunable Ti:sapphire laser and focused onto the photomixer, driving the antenna structure at the difference frequency which sweeps through 3 THz (100 $cm^{-1}$) as the grating tunes the Ti:sapphire laser. Maximum power is obtained near 0.6 THz (20 $cm^{-1}$) and THz output falls off with frequency as $\omega^{-4}$ ($\omega = 2\pi f$) beyond this maximum. The THz radiation, confined within a vacuum chamber to avoid water-vapor absorption, is collimated by a silicon lens, and then focused by off-axis parabolic mirrors.

The sample assembly, fixed to the coldfinger of a cryostat in the vacuum chamber, is placed in the THz beam, and transmitted power is detected by a helium-cooled silicon composite bolometer having a power detection sensitivity of <1 nW up to 3 THz in a 400 Hz bandpass. The noise equivalent power (NEP) of the bolometer is 1 $pW/Hz^{1/2}$. For normalization purposes, a voltage proportional to the photomixer photocurrent is recorded simultaneously with the detected THz power.

Dipeptides obtained from Bachem at >99% purity were diluted with fine grain polyethylene powder (average particle diameter 4 $\mu$m), and 100 mg of a 10−15% concentration mixture were pressed at $2 \times 10^7$ Pa (3000 psi) to form a disk ~750 $\mu$m thick and 13 mm in diameter. Some samples were recrystallized by evaporation from aqueous solution, or by vapor diffusion of solvents into aqueous solutions. It should be noted that, in light of the solvent-retaining characteristics of some members of the VA class, only solvents known to be removable under vacuum were utilized in recrystallization for this work. It is further noted that recrystallized samples produced the same spectra generated by the as-received samples, although linewidths may be broader or narrower, depending upon the crystal quality obtained in a particular trial. The single exception was L-alanyl-L-valine (Ala-Val). For Ala-Val, the samples recrystallized from aqueous solution generate a noticeable intensity change in one peak, as compared to spectra from as-received samples. All samples, both those used as-received and those prepared by recrystallization, were placed under vacuum (<$10^{-4}$ torr) for 5−7 days before spectra were recorded.

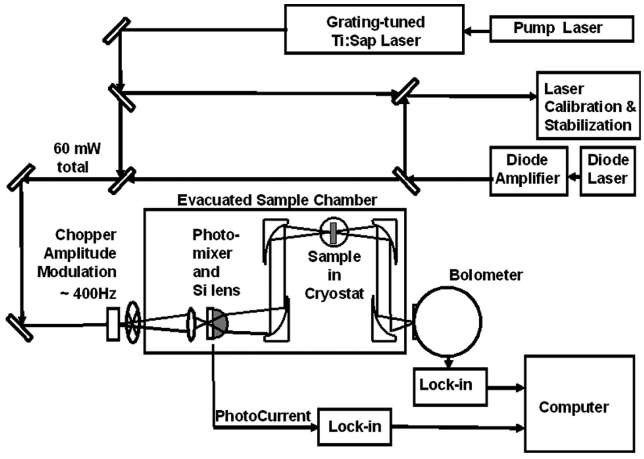


***Figure 2.*** Schematic of continuous wave Terahertz spectrometer.

The brass sample assembly contains two sample positions, which are accessed by lowering or raising the cryostat by 1 cm. This allows a background spectrum of the polyethylene blank occupying one position to be acquired with minimal change in the THz beam path and angle of incidence. This procedure is important for minimizing interference artifacts when transmission spectra are ratioed against the background spectrum of pure polyethylene. During acquisition of spectra, the helium-cooled sample assembly was held at 4.2 K in vacuum for the disk-shaped samples in the brass assembly.

The dipeptide absorption spectra have been fit using a nonlinear least-squares procedure to a minimal set of Gaussian line shape functions, characterized by center frequency, $\bar{\nu}_e$, peak absorption coefficient, $\kappa_e$, and width, $\Delta\bar{\nu}$ that account for the observed intensity. The parameter values from the fits that account for the observed THz absorption intensity are summarized in Table 1. The widths for a few poorly resolved or weak lines were fixed to widths commensurate with those of nearby features, and these are denoted by *b* in Table 1.

**2.2. Force Field.** In the unit cell of the VA class dipeptide systems, there are 192 atoms for Ala-Ile and Ile-Ala, 175 atoms for Ala-Val and Val-Ala, 228 atoms for Val-Ile and Ile-Val, and 210 for Val-Val; when periodic boundary conditions are applied, there are ~2500−3300 atoms in the image lists at a cutoff radius of 20 Å. Given such large system sizes, *ab initio* or hybrid *ab initio* potential energy surface calculations are computationally expensive. On the other hand, all-atom empirical force fields have been successfully utilized to model the vibration spectra of similar biological systems. For example, Korter and co-workers showed that the CHARMM22 force field is adequate and computationally inexpensive for modeling two complex hydrogen-bonded biological systems, solid serine and cysteine,[44] while Siegrist and co-workers[9] have demonstrated more limited success of this force field for different $\beta$-sheet forms of trialanine. Moreover, Gerber et al. showed that optimized potential for liquid simulations (OPLS) was even superior to *ab initio* calculations in some intermediate frequency regions for the glycine systems.[45] Therefore, in the present VSCF calculations, we utilized the all-atom force field, CHARMM22.[46,47] This work will serve as a benchmark test of the CHARMM22 force field for the potential energy surfaces of the crystalline dipeptide systems.

(41) 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*; Emerging Scientific Applications and Novel Device Concepts, Vol. 2; World Scientific: NJ, 2003.
(42) Korter, T. M.; Plusquellic, D. F. *Chem. Phys. Lett.* **2004**, *385*, 45–51.
(43) Bjarnason, J. E.; Brown, E. R. *Appl. Phys. Lett.* **2005**, *87*, 134105.
(44) Korter, T. M.; Balu, R.; Campbell, M. B.; Beard, M. C.; Gregurick, S. K.; Heilweil, E. J. *Chem. Phys. Lett.* **2006**, *418*, 65–70.
(45) Gerber, R. B.; Chaban, G. M.; Gregurick, S. K.; Braucer, B. *Biopolymers* **2003**, *68*, 370–382.
(46) MacKerell, A. D., Jr. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.
(47) MacKerell,, A. D., Jr.; Brooks, B.; Brooks, C. L., III.; Nilsson, L.; Roux, B.; Won, Y.; Karplus, M. In *The Encyclopedia of Computational Chemistry*; John Wiley & Sons: Chichester, U.K., 1998; Vol. 1.

***Table 1.*** Experimental Center Frequencies[a], $\bar{\nu}_e$, Peak Absorption Coeffecients, $\kappa_e$, and Gaussian Line Widths[a], $\Delta\bar{\nu}$, Obtained from Nonlinear Least-Squares Fits of the VA Class Dipeptide Nanotubes

| | AI | | | IA | | | AV | | | VA | | | IV | | | VI | | | VV | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $\bar{\nu}_e$ | $\kappa_e$ | $\Delta\bar{\nu}$ | $\bar{\nu}_e$ | $\kappa_e$ | $\Delta\bar{\nu}$ | $\bar{\nu}_e$ | $\kappa_e$ | $\Delta\bar{\nu}$ | $\bar{\nu}_e$ | $\kappa_e$ | $\Delta\bar{\nu}$ | $\bar{\nu}_e$ | $\kappa_e$ | $\Delta\bar{\nu}$ | $\bar{\nu}_e$ | $\kappa_e$ | $\Delta\bar{\nu}$ | $\bar{\nu}_e$ | $\kappa_e$ | $\Delta\bar{\nu}$ |
| 1 | 40.4 | 0.09 | 3.5[b] | 35.7 | 0.18 | 3.5[b] | 42.4 | 0.09 | 3.5[b] | 38.8 | 0.14 | 4.0[b] | 44.1 | 0.10 | 2.0[b] | 44.9 | 0.08 | 2.5[b] | 39.3 | 0.14 | 3.0[b] |
| 2 | 44.0 | 0.49 | 3.3 | 41.6 | 0.16 | 3.5[b] | 46.2 | 0.12 | 3.5[b] | 43.5 | 0.34 | 4.0[b] | 46.1 | 0.38 | 2.0[b] | 48.0 | 0.07 | 2.5[b] | 43.1 | 0.74 | 3.0 |
| 3 | 46.8 | 0.08 | 3.0[b] | 45.0 | 0.18 | 3.5[b] | 50.0 | 0.25 | 3.5[b] | 46.9 | 1.20 | 3.4 | 48.1 | 1.13 | 2.2 | 50.7 | 0.32 | 2.8 | 46.4 | 0.20 | 3.0[b] |
| 4 | 49.9 | 0.15 | 3.0 | 49.1 | 1.13 | 4.3 | 52.2 | 0.42 | 2.2 | 49.5 | 0.32 | 4.0[b] | 50.5 | 0.15 | 2.0[b] | 53.1 | 0.11 | 2.5[b] | 49.9 | 0.15 | 3.0[b] |
| 5 | 53.9 | 0.09 | 3.5[b] | 53.0 | 0.28 | 3.5[b] | 56.9 | 0.17 | 4.0[b] | 53.2 | 0.10 | 4.0[b] | 52.9 | 0.12 | 2.5[b] | 55.4 | 0.13 | 2.5[b] | 53.5 | 0.10 | 3.0[b] |
| 6 | 58.7 | 0.21 | 3.7 | 56.6 | 0.16 | 3.5[b] | 59.9 | 0.42 | 2.8 | 57.4 | 0.35 | 3.8 | 55.6 | 0.11 | 2.5[b] | 56.9 | 0.20 | 2.0 | 57.5 | 0.12 | 3.3[b] |
| 7 | 61.9 | 0.65 | 3.1 | 60.6 | 0.14 | 4.0[b] | 64.2 | 0.30 | 2.4 | 59.7 | 0.05 | 4.0[b] | 58.1 | 0.48 | 3.1 | 59.1 | 0.26 | 2.8 | 60.8 | 0.13 | 3.3[b] |
| 8 | 64.1 | 0.12 | 4.0[b] | 64.8 | 0.13 | 4.0[b] | 67.2 | 0.49 | 3.3 | 63.3 | 0.45 | 4.2 | 61.1 | 0.40 | 2.8 | 60.3 | 0.06 | 3.0[b] | 64.1 | 0.61 | 3.1 |
| 9 | 71.5 | 0.96 | 3.5[b] | 69.3 | 0.42 | 4.2 | 71.2 | 0.12 | 3.5[b] | 68.3 | 0.09 | 4.5[b] | 62.9 | 0.23 | 2.2 | 63.2 | 0.05 | 3.0[b] | 68.1 | 0.16 | 3.3[b] |
| 10 | 77.5 | 0.08 | 4.0[b] | 74.6 | 0.08 | 3.5[b] | 74.9 | 0.25 | 2.4 | 74.8 | 0.07 | 4.5[b] | 64.9 | 0.18 | 2.5[b] | 66.8 | 0.08 | 3.0[b] | 70.9 | 0.13 | 3.0[b] |
| 11 | 80.3 | 0.24 | 3.9 | 77.9 | 0.30 | 3.0 | 79.5 | 0.10 | 3.0[b] | 78.7 | 0.14 | 4.0[b] | 67.1 | 0.20 | 2.5[b] | 69.6 | 0.39 | 2.5 | 73.7 | 0.54 | 3.3 |
| 12 | 85.3 | 0.05 | 4.0[b] | 80.9 | 0.13 | 4.0[b] | 83.0 | 0.26 | 2.9 | 81.9 | 0.12 | 4.5[b] | 69.6 | 0.67 | 2.6 | 72.9 | 0.21 | 2.6 | 77.0 | 0.19 | 3.5[b] |
| 13 | 89.8 | 0.15 | 4.0[b] | 85.7 | 0.52 | 4.6 | 86.4 | 0.11 | 3.0[b] | 86.0 | 0.34 | 5.1 | 71.4 | 0.19 | 2.5[b] | 76.3 | 0.21 | 2.5 | 80.6 | 0.14 | 3.5[b] |
| 14 | 93.3 | 0.69 | 0.4 | 90.3 | 0.16 | 4.0[b] | 89.4 | 0.29 | 4.0[b] | 90.8 | 0.15 | 4.5[b] | 77.4 | 0.04 | 3.5[b] | 80.0 | 0.04 | 3.0[b] | 85.1 | 0.14 | 3.5[b] |
| 15 | 98.2 | 0.63 | 0.4 | 93.5 | 0.16 | 4.0[b] | 95.8 | 0.35 | 4.8 | 94.9 | 0.25 | 4.5[b] | 81.1 | 0.88 | 3.7 | 84.3 | 0.14 | 2.8 | 88.7 | 0.12 | 3.5[b] |
| 16 | | | | 97.7 | 0.19 | 4.0[b] | 99.8 | 0.27 | 6.0[b] | 99.4 | 0.68 | 5.0[b] | 85.3 | 0.43 | 3.0 | 89.2 | 0.08 | 3.5[b] | 92.8 | 0.15 | 3.8[b] |
| 17 | | | | | | | | | | | | | 88.7 | 0.36 | 3.0[b] | 92.3 | 0.25 | 3.4 | 98.1 | 0.61 | 4.9 |
| 18 | | | | | | | | | | | | | 91.7 | 1.12 | 3.0 | 97.2 | 0.27 | 3.9 | | | |
| 19 | | | | | | | | | | | | | 93.7 | 0.46 | 2.0[b] | | | | | | |
| 20 | | | | | | | | | | | | | 96.4 | 0.78 | 4.2 | | | | | | |

[a] All values in $cm^{-1}$. [b] Fixed in the fit at a reasonable value based on the trend.

**2.3. Energy Minimization.** The starting geometries of Ala-Ile, Ile-Ala, Ile-Val, Val-Ile and Val-Val were obtained from the crystal structures deposited in the Cambridge Crystallographic Data Centre, whereas those of Ala-Val and Val-Ala were obtained from the International Union of Crystallography. It is known that, for VA class dipeptide systems, solvent guest molecules can be removed by drying with the peptide host intact.[28] Therefore, to obtain initial geometries for completely dry crystal structures, the solvent molecules in the starting geometry files (in CIF format) of certain systems, such as the acetonitrile molecules in Ala-Val and the water molecules in Val-Ile, Ile-Val and Val-Val, were simply removed without modifying the dipeptide host structures. The equivalent molecules in a single unit cell were generated based on the symmetric operations and unit cell parameters provided by the CIF files. Periodic boundary conditions were created using the CRYSTAL facility in the CHARMM-c32b2 package[48] with a cutoff radius of 20 Å. The energy minimizations start from 1000 steps of steepest descent (SD) algorithm, followed by 20,000 steps of adopted basis set Newton−Raphson (ABNR) algorithm. The lattice parameters were also optimized in the ABNR algorithm; it is noted that the crystal lattice symmetry was maintained during the lattice parameter optimization process. During the energy minimization, the electrostatic interactions were computed by invoking the Particle Mesh Ewald summation strategy,[49,50] in which the Gaussian width of 0.34 was used for transitioning from the real space to **k**-space electrostatic sums, and the number of the grid points in Fourier space was set to ∼2 times each unit cell dimension.

**2.4. Vibration Calculation.** After the energies have been successfully minimized, the mass-weighted Hessian matrices[51] at the global/local minimum were diagonalized by using finite differences. Similar to Day's method, only the wavevector $\mathbf{k} = 0$ was considered.[52] The eigenvectors generated by the Hessian diagonalization, which correspond to each normal mode, served as the starting point for the VSCF process. In the present VSCF calculation, the second-order perturbation-based correlation correction (CC-VSCF) was not implemented for the following reasons. (1) Based on eq 6, for large systems, CC-VSCF involves the integrations over a large number of excited states. Given the system sizes of the present crystalline VA class dipeptides, this CC-VSCF calculation is computationally intractable. (2) The Gerber group pointed out that the importance of the correlation corrections decreases with cluster size, and the accuracy of the VSCF increases correspondingly.[53] This group found that the average correlation-correction percentage decreased gradually from ∼2% to ∼0.7% with the water cluster increasing from 2 to 5 in size.[53] They therefore proposed that, for very large systems, the VSCF calculations should suffice, and correlation corrections are unnecessary. Considering this, the present VSCF calculations on the crystalline VA class dipeptides are carried out without correlation corrections.

The VSCF wave functions and energies were numerically calculated by iteratively solving the Schrödinger eq 3 using Yang and Peet's collocation method.[54,55] To validate the reliability of the collocation subroutine in numerically solving the nucleic Schrödinger equations, we tested a series of harmonic oscillators. Notably, we found that given a fixed number of mesh points, the convergence of the collocation calculation is generally sensitive to the input coordinate range $\Delta Q$. For a harmonic oscillator with a specific force constant $k$, when the input $\Delta Q$ is beyond a certain range, the output eigenenergies and wave functions begin to deviate from the analytical solutions. Therefore, we predefined a series of harmonic oscillators with various force constants $k$, and collected their corresponding valid $\Delta Q$ values. We then obtained a polynomial interpolation on these ($k$, $\Delta Q$) data, which served as a guide to determine the input $\Delta Q$ value for a normal mode based on its specific $k$ value.

The intensity of each mode was evaluated through the dipole derivatives based on the normal mode motions, which are numerically evaluated according to:

$$I_i \propto \left| \left\langle v' \left| \frac{\partial \mu_i}{\partial \mathbf{Q}_i} \right| v'' \right\rangle \right|^2 \tag{11}$$

A Lorentzian function was used to describe each line shape, which was uniformly broadened at a temperature of 4 K. The previously

(48) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187–217.
(49) Ewald, P. *Ann. Phys.* **1921**, *64*, 253–287.
(50) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577.
(51) 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.
(52) Day, G. M.; Zeitler, J. A.; Jones, W.; Rades, T.; Taday, P. F. *J. Phys.Chem. B* **2006**, *110*, 447–456.
(53) Jung, J. O.; Gerber, R. B. *J. Chem. Phys.* **1996**, *105*, 10332–10348.
(54) Yang, W.; Peet, A. C. *Chem. Phys. Lett.* **1988**, *153*, 98–104.
(55) Yang, W.; Peet, A. C. *J. Chem. Phys.* **1989**, *90*, 1746–1751.

developed vector field representation tool[56] was used to visualize the motion of each normal mode by the visual molecular dynamics the program VMD,[57] with an expansion of vector field calculation from the α-carbon to all atom levels.

The VSCF subroutine components were implemented in the CHARMM-c32b2 package[48] using FORTRAN77 language. The program was compiled and implemented under x86_64 Linux Cluster on 3.0 GHz Intel Xeon "Irwindale" Processors from University at Buffalo, Center for Computational Research. The VSCF calculations of the VA class dipeptide systems took up to 4 CPU hours running on 4 nodes with 2 processors per node.

**2.5. Numerical Comparison of THz Spectra.** To numerically compare the calculated spectra against the corresponding experimental spectra, we employed a sophisticated method similar to that used by Hageman et al. for spectra comparison.[58] Readers are referred to ref 58 for a more detailed description of this method. Briefly, this method takes into account the small frequency shift in the peak position and utilizes the cross correlation function to compare two shifted spectra. The cross correlation function, also known as sliding dot product, takes the form of:

$$Cr_{fg}(r) = \frac{\sum_{x=0}^{x=k} (f(x) - \langle f \rangle) \cdot (g(x+r) - \langle g \rangle)}{\sqrt{\sum_{x=0}^{x=k} (f(x) - \langle f \rangle)^2 \sum_{x=0}^{x=k} (g(x) - \langle g \rangle)^2}} \quad (12)$$

where $\langle f \rangle$ and $\langle g \rangle$ are the average intensity value of spectrum $f$ and $g$, respectively, and $r$ is the frequency shift of spectrum $g$ with respective to spectrum $f$. A penalty is given to the larger shift by introducing a weight function:

$$w(r) = 1 - \frac{|r|}{l} \quad (13)$$

where $l$ is the maximum frequency shift performed, which is 0.2 THz for the present calculation. The averaged cross correlation function is obtained by integrating the weighted cross correlation function over the spectrum shift:

$$Cr_{fg}^{wt} = \frac{\sum_{r=-l}^{r=l} Cr_{fg}(r) w(r)}{\sqrt{\sum_{r=-l}^{r=l} Cr_{ff}(r) w(r) \cdot \sum_{r=-l}^{r=l} Cr_{gg}(r) w(r)}} \quad (14)$$

The final evaluation result of the comparison is defined as:

$$F = 100 \cdot (1 - |Cr_{fg}^{wt}|) \quad (15)$$

For two identical spectra $F$ is 0, and for two distinctly different spectra $F$ is 100.

## 3. Results and Discussion

**3.1. Energy Minimization.** Reaching the global/local energy minimum is a prerequisite for Hessian diagonalization and further vibration calculations; moreover, the inaccuracy of X-ray-determined hydrogen atom positions also makes the structural optimization necessary. For the VA class dipeptide systems under study, all of the ABNR energy minimizations converged below 0.000001 kcal/mol tolerance after about 500−1000 cycles, indicating that the global/local minimums had been reached.

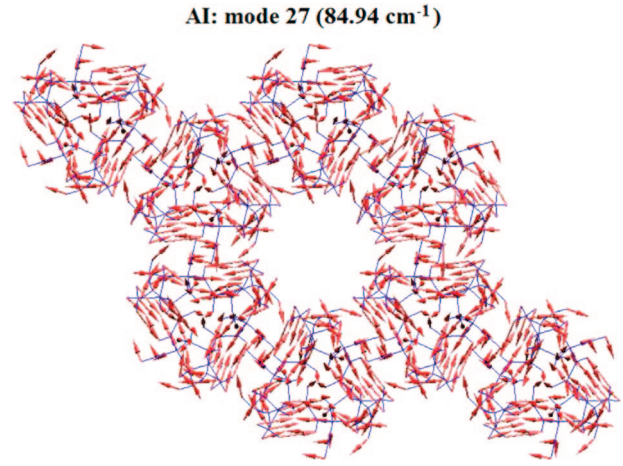


***Figure 3.*** Vector field representation of a channel-breathing motion of the IA based on the vibrational calculations using CHARMM22 force field with periodic boundary conditions. The dipeptide molecules are shown in blue. The arrows representing atomic vibrations are colored based on the vibrational amplitudes. The mode number and the corresponding VSCF frequency are shown on the top of the plot.

**3.2. Normal Mode Analysis.** After energy minimization, the mass-weighted Hessian matrix was diagonalized to obtain the harmonic frequency eigenvalues. Table S of the Supporting Information lists the first 40 harmonic frequencies (the first column of each system) of the VA class dipeptides, from which one finds that all of the first three eigenfrequencies range from −0.068559 $cm^{-1}$ to 0.078171 $cm^{-1}$. This confirms that the structures have been sufficiently minimized for the crystalline systems under periodic boundary conditions. The eigenvectors upon Hessian diagonalization correspond to the normal modes. To facilitate normal-mode analysis, we generated the three-dimensional vibration animations using PyMOL™ v0.99 for the normal modes of all the VA class dipeptides in the THz region. The complete package including all the animations is available at http://userpages.umbc.edu/~zhangh1/.

Normal-mode analysis revealed that almost all of the vibration modes in the THz region involved the hydrogen bonds of the nanotube frame, inducing motions of side chains constituting the channel inner surface. From the three-dimensional animations, we found these side-chain motions frequently gave rise to channel wall deformation (e.g., side-chain wagging or channel squeezing), but the channel volumes were not notably changed. However, for all the VA class dipeptide systems except Val-Ala itself, 1−3 normal modes were found between (80−120) $cm^{-1}$ which demonstrated a remarkable channel-breathing motion that enlarged the channel capacity significantly, characterized by the synchronous vibrations of side chains toward the channel centers ($\mathbf{6_1}$ axis). A selected example of this breathing motion is shown in Figure 3, whereas the complete list is shown in Figure S1 (Supporting Information). The animations of these motions are also available at http://userpages.umbc.edu/~zhangh1/.

The dipeptide of Val-Ala itself was the only exception in which the obvious channel-breathing mode was *not* found in THz region. It is notable that, although Val-Ala and Ala-Val are composed of the same side-chain types, these two structures have extremely different permeability character-

(56) Balu, R.; Zhang, H.; Zukowski, E.; Chen, J.-Y.; Markelz, A. G.; Gregurick, S. K. *Biophys. J.* **2008**, *94*, 3217–3226.
(57) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33–38.
(58) Hageman, J. A.; Wehrens, R.; de Gelder, R.; Meerts, W. L.; Buydens, L. M. C. *J. Chem. Phys.* **2000**, *113*, 7955–7962.

***Table 2.*** Eigenvector Inner Products of Backbone Vibrations[a] between the Channel-Breathing Modes of VA Class Dipeptide Nanotubes[b]

| | | AI | | AV | | IA | | IV | | VI | | VV | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| dipeptide | mode number | 27 | 31 | 27 | 37 | 25 | 37 | 22 | 36 | 25 | 26 | 37 | 23 |
| AI | 27 | 1.00 | – | **0.57**[c] | **0.51**[c] | 0.20 | **−0.44**[c] | 0.00 | 0.01 | 0.06 | 0.29 | −0.13 | 0.04 |
| | 31 | – | 1.00 | 0.01 | **−0.58**[c] | 0.25 | 0.02 | 0.15 | −0.04 | 0.17 | 0.08 | 0.06 | −0.19 |
| AV | 27 | **0.57**[c] | 0.01 | 1.00 | – | **0.60**[c] | −0.30 | −0.25 | 0.21 | 0.11 | 0.26 | −0.07 | 0.31 |
| | 37 | **0.51**[c] | **−0.58**[c] | – | 1.00 | −0.02 | −0.01 | −0.22 | 0.17 | −0.10 | 0.17 | −0.06 | 0.17 |
| IA | 25 | 0.20 | 0.25 | **0.60**[c] | −0.02 | 1.00 | – | −0.29 | 0.15 | 0.01 | 0.27 | 0.10 | 0.25 |
| | 37 | **−0.44**[c] | 0.02 | −0.30 | −0.01 | – | 1.00 | −0.25 | 0.19 | −0.15 | 0.16 | 0.27 | 0.04 |
| IV | 22 | 0.00 | 0.15 | −0.25 | −0.22 | −0.29 | −0.25 | 1.00 | – | −0.01 | 0.01 | −0.12 | **−0.65**[c] |
| | 36 | 0.01 | −0.04 | 0.21 | 0.17 | 0.15 | 0.19 | – | 1.00 | 0.04 | 0.31 | 0.02 | 0.00 |
| VI | 25 | 0.06 | 0.17 | 0.11 | −0.10 | 0.01 | −0.15 | −0.01 | 0.04 | 1.00 | – | – | 0.11 |
| | 26 | 0.29 | 0.08 | 0.26 | 0.17 | 0.27 | 0.16 | 0.01 | 0.31 | – | 1.00 | – | −0.01 |
| | 37 | −0.13 | 0.06 | −0.07 | −0.06 | 0.10 | 0.27 | −0.12 | 0.02 | – | – | 1.00 | 0.08 |
| VV | 23 | 0.04 | −0.19 | 0.31 | 0.17 | 0.25 | 0.04 | **−0.65**[c] | 0.00 | 0.11 | −0.01 | 0.08 | 1.00 |

[a] Based on the vibrational calculations using CHARMM22 force field; see text for details. [b] The lines divide VA class dipeptides into two subgroups based on the distribution of large inner product values. [c] The inner product values greater than 0.4 shown in bold.

istics. For example, Görbitz' et al. found that, although both dipeptides can accommodate the linear chains of acetonitrile molecules, Ala-Val can flexibly accommodate larger guest alcohols, whereas Val-Ala did not.[28,31,59] It can further be noted that in Görbitz' experiments, the only crystal that shows any sign of deterioration upon solvent substitutions or removal was also Val-Ala.[28,59] Therefore, we suggest that the channel-breathing modes can enlarge the channel capacity and thus facilitate the pathway of guest molecule absorption, substitution and removal, and the lack of this THz mode in Val-Ala may result in the exclusion of large alcohol guest molecules. This flexibility, or the lack thereof, may also explain the tendency of Val-Ala for structural deterioration upon environmental perturbation.

To verify whether these channel-breathing modes originate from similar hydrogen-bond vibrations of the structural framework, we compared the framework eigenvectors between the breathing modes of different dipeptide systems. Since all the VA class nanotubes share the identical framework topology, the obvious method of comparison is to numerically evaluate the overlap of the peptide framework component of the vibrational modes. The inner product of the backbone components of the normal-mode eigenvectors is a straightforward measure of mode similarity, since it measures the overlap of the backbone structural motions, where an overlap close to 1 corresponds to nearly identical backbone motions and nearly 0 overlap corresponds to extremely dissimilar, or orthogonal, motions. Since the CHARMM mass-weighted eigenvectors are normalized over the full crystal structure, the inner product in our calculations takes the form of:

$$dp_{r,s} = \frac{\sum_{i}^{backbone} \left[(q_i^r\sqrt{m_i})(q_i^s\sqrt{m_i})\right]}{\sqrt{\left[\sum_{i}^{backbone} (q_i^r\sqrt{m_i})^2\right]\left[\sum_{i}^{backbone} (q_i^s\sqrt{m_i})^2\right]}} \qquad (16)$$

where $m_i$ is the mass of the $i$th backbone atom, and $q_i^r$ and $q_i^s$ represent the $i$th backbone eigenvector component of the two normal modes, $r$ and $s$, respectively. The inner products between the framework components of the breathing modes are shown in Table 2, from which one finds that the distribution of large inner product values (shown in bold) roughly divides the VA class dipeptides into two subgroups: **I**: AI, AV, IA and **II**: IV, VI, VV. Within the same subgroup, the channel-breathing modes share similarities of characteristic motion (as shown by the larger overlap values), whereas those from different subgroups show less similarity (as shown by the smaller overlap values). Note that this division method is rather arbitrary, and the border of subclass **II** is actually not as clear as that of subclass **I** based on these mode overlap values. Despite this, our findings by and large agree with Görbitz's division of the VA class dipeptide nanotubes based on the torsion angles of backbone bonds, in which three subgroups, **I**: AV, AI; **II**: IA, VA; **III**: IV, VI, VV, were distinguished.[29]

The channel-breathing and deforming motions also yield insight into the transport mechanisms of biological channels. This will be illustrated from three aspects. (1) It was suggested that the geometry of a pore (i.e., radius, shape and length) can have a great influence on pore permeability to ions or water.[33] Therefore, for the present systems, we expect the THz motions that affect the channel sizes or shapes play roles in the pore permeability. For example, the channel-breathing motions affect the physical diameter of the pore, and we suggest these motions bear on the ability of pores to admit or exclude solvent molecules and facilitate the pore transport. (2) Previous research found that the critical radius for permeability depends also on potentials between the solvent and the pore molecules. For example, in a molecular dynamics simulation of a water pore model consisting of concentric rings of methane-like pseudoatoms, Beckstein et al. observed the positional fluctuations of the wall atoms, and they hypothesized that the wall fluctuations change the effective potential between the wall atom and water molecules thus influencing water transport.[33] For the present systems, we similarly suggest the channel wall deformation motions play roles in solvent transport through the hydrophobic pore. The localized side-chain fluctuations induced by these modes can transiently expose the hydrogen-bond framework to the pore solvent, thus affecting the effective potential between solvent and the pore and impacting the critical radius for solvent permeability. (3) Finally, the collective motions of intrapore water have been shown to be important to water transport through the hydrophobic pore.[35,36] For example, de Groot and Grubmuller attribute high rates of water transport through the hydrophobic pore of Aquaporin-1 to collective motions of the intrapore water, [35] while the work of Pomes et al. implied that collective motions were needed to sustain proton conduction along intrapore water chains.[37] For the present systems, the concerted side-chain motions

(59) Görbitz, C. H. *CrystEngComm* **2005**, *7*, 670–673.

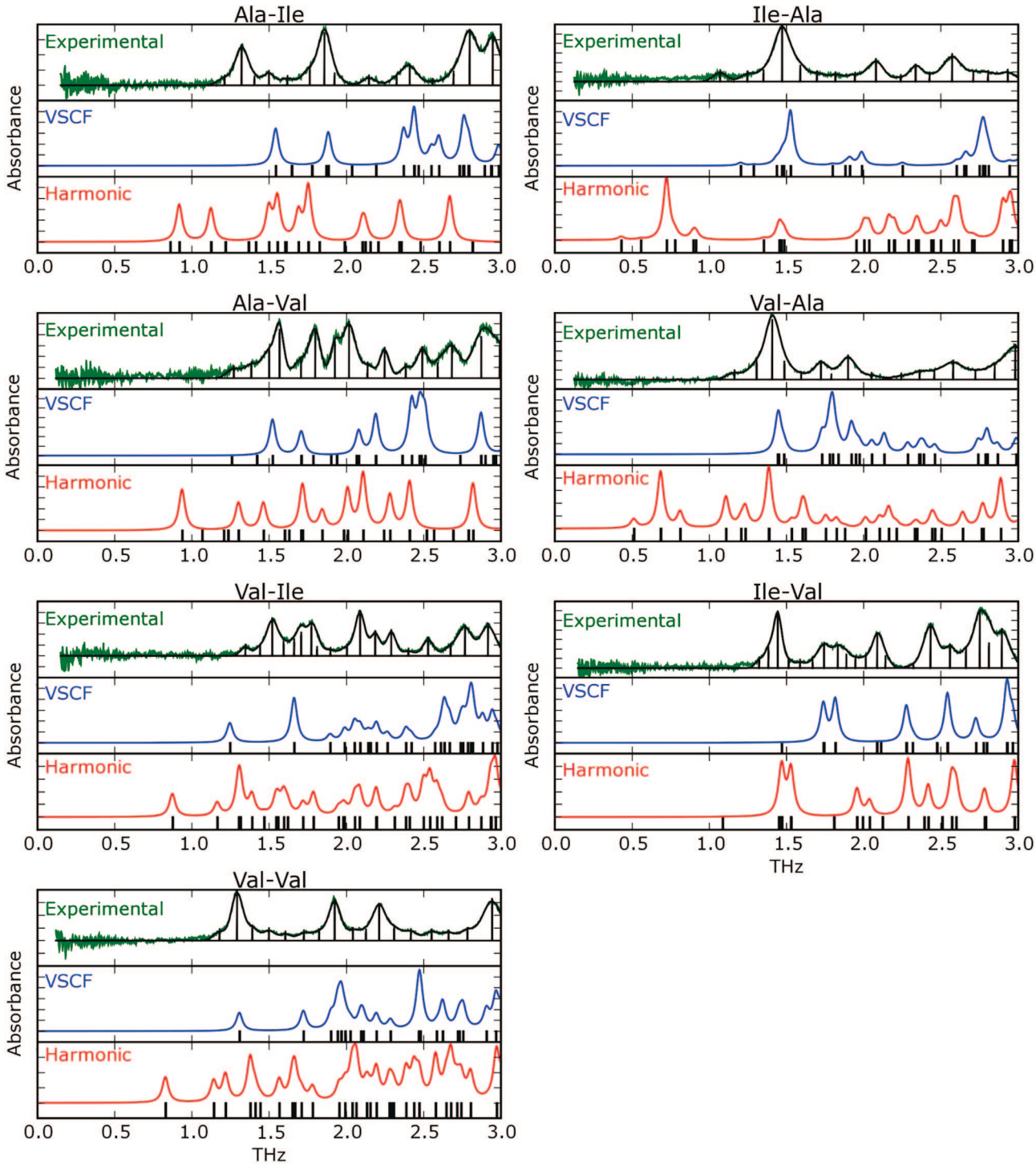


***Figure 4.*** Calculated THz spectra of VA class dipeptide nanotubes compared to the experimental spectra. In each subplot, from top to bottom are respectively: "Experimental"—the experimental spectra of crystalline pressed with polyethylene with best-fit lines and distinguishing spectral features superimposed, "VSCF", the calculated spectra at VSCF level using CHARMM22 with periodic boundary conditions, and "Harmonic", the calculated spectra at harmonic level using CHARMM22 with periodic boundary conditions. The intensities of the calculated spectra were evaluated using the dipole derivatives, and the line shapes were described by the Lorentzian functions. The calculated mode frequencies of each system are indicated by the stick representations below each calculated spectrum.

of the channel-breathing modes may be coupled to the motions of intrapore solvent and drive solvent transmission through the hydrophobic pore.

**3.3. THz Spectra.** Based on the normal modes generated by Hessian diagonalization, the VSCF calculations incorporating the anharmonic effects and intermode couplings were carried out for the VA class dipeptide systems according to eqs 3−10. The calculated THz spectra of the crystalline VA class dipeptide systems below 3 THz (~100 $cm^{-1}$) at the VSCF level, as well as at the harmonic level, are shown in Figure 4. The experimental THz spectra of all the VA class dipetides from the crystalline dipeptide systems pressed into polyethylene pellets are also shown for comparison. From Figure 4, one intuitively finds that, in most cases, the VSCF calculations demonstrated overall improvement compared to the harmonic calculations, especially in the lower-frequency region between 0 THz to 2 THz. This can be numerically

***Table 3.*** *F* Values of Numerical Comparisons between the Calculated and Experimental Terahertz Spectra of the VA Class Dipeptide Systems from Different Frequency Regions. A Smaller *F* Value Implies a Larger Similarity

| | terahertz regions | | | |
|---|---|---|---|---|
| | 0−2 THz | | 2−3 THz | |
| dipeptide systems | harm. vs exp. | VSCF vs exp. | harm. vs exp. | VSCF vs exp. |
| AI | 68.22 | 38.49 | 41.26 | 76.04 |
| IA | 96.93 | 7.85 | 83.43 | 87.49 |
| AV | 28.73 | 20.51 | 74.60 | 60.96 |
| VA | 43.43 | 40.04 | 46.71 | 88.14 |
| VI | 32.82 | 25.52 | 77.97 | 47.72 |
| IV | 27.60 | 43.57 | 70.48 | 84.60 |
| VV | 31.07 | 28.18 | 38.59 | 82.55 |

confirmed by evaluating the *F* values between calculated and experimental spectra based on the cross-correlation function introduced in section 2.5. Table 3 summarized the *F* values evaluated over both lower-frequency (0−2 THz) and higher-frequency (2−3 THz) regions. One finds that at lower frequency (0−2 THz), the VSCF calculation demonstrated significant improvement compared to the harmonic calculation for all systems except IV; however, at higher frequency (2−3 THz), there is no obvious VSCF improvement. The following section will respectively discuss these VSCF improvements and discrepancies.

**3.3.1. VSCF Improvements.** The VSCF improvements in the lower-frequency region (0−2 THz) is generally reflected in the overall *blue-shifts* from the corresponding harmonic frequencies, which can be found from both Figure 4 and Table S of the Supporting Information. Similar blue-shifts had been previously reported in the VSCF versus harmonic-level calculations of an isolated di-L-serine−$H_2O$ complex, where these blue-shifts were attributed to the anharmonic and mode couplings.[15] Therefore, the present research will systemically investigate the anharmonic mode couplings of the VA class dipeptide systems. The anharmonic mode couplings have considerable relevance to the molecular physical properties. For example, these couplings can affect the vibrational frequencies, and govern the flow of vibrational energy between different modes in dynamical processes.[18] To quantify the magnitudes of anharmonic couplings, we introduced a *coupling strength* term that is very similar to the average absolute value of the coupling potential defined by Miller et al.[18] The coupling strength between mode *i* and *j* is defined as:

$$cs_{ij} = \left\langle \psi^2(Q_i) \middle| V_{ij}(Q_i, Q_j) \middle| \psi^2(Q_j) \right\rangle \qquad (17)$$

where $V_{ij}(Q_i,Q_j)$ is the pair-wise coupling potential from eq 10, and $\psi(Q_i)$ and $\psi(Q_j)$ are the ground-state wave functions. Note that this coupling strength term slightly differs from Miller et al.'s definition in not evaluating the absolute value of the coupling potential. The coupling strength contours for all the present dipeptide systems are shown in Figure S2 (Supporting Information). From this figure, one finds that despite the diverse coupling strength distribution patterns, the coupling strength values are almost exclusively *positive*. As a result, according to eqs 8 and 4, the effective potentials will be stiffened in comparison to the diagonal potentials for the VA class dipeptide systems.

Then, one may ask, why are the coupling strengths all positive values? To answer this question, we selectively investigated the most significantly coupled mode pairs characterized by the largest coupling strength value within each dipeptide system. A selected

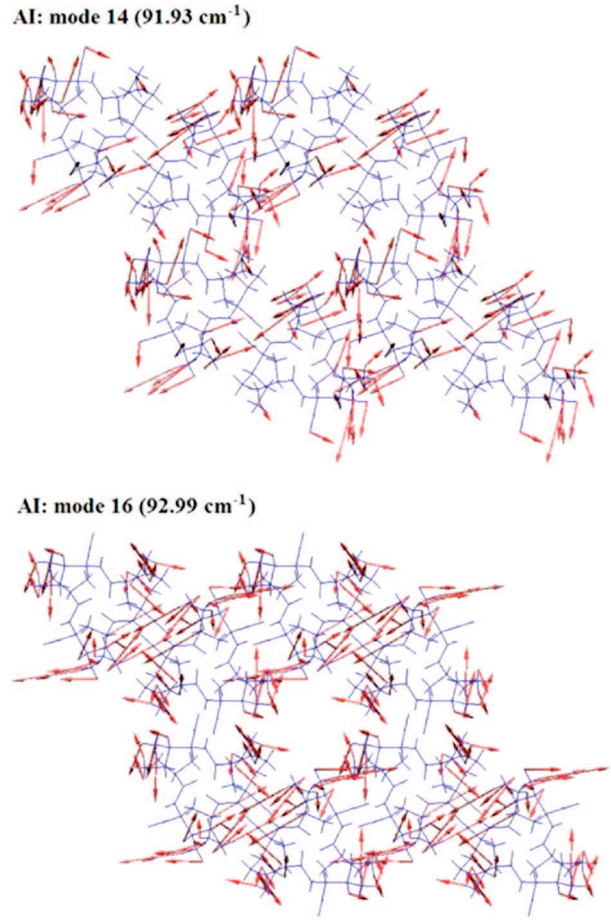


***Figure 5.*** Vector field representations of the most significantly coupled mode pair of AI based on the vibrational calculations using the CHARMM22 force field with periodic boundary conditions. The dipeptide molecules are shown in blue. The arrows representing atomic vibrations are colored based on the vibrational amplitudes. The weak vibrations are not shown for the sake of clarity. The mode number and the corresponding VSCF frequency are shown on the top of each plot.

example of these mode pairs is demonstrated in Figure 5, whereas the complete list is shown in Figure S3 (Supporting Information). By comparing each significantly coupled mode pair from the same system, we found that their overall vibrational patterns are quite similar. Moreover, unlike the channel-breathing motions, these significantly coupled modes mostly involve the *asynchronous* wagging of the side chains that leads to *squeezing* of the channel inner surfaces. Since the VA class dipeptide side chains are all hydrophobic, squeezing of the inner surface constituted by these side chains will inevitably experience certain van der Waals repulsions. Remarkably, when two similar squeezing motions, i.e. each mode pair of the same system in Figure S3 (Supporting Information), are coupled or overlapped, the channel inner surface will be squeezed to a larger extent beyond which the van der Waals repulsions increase very *sharply* (of order $(1/r)^{1/12}$ in the Lennard-Jones potential form). However, according to eq 10, the coupling potential between two modes is defined by the *linear* subtraction of $V(\Delta x_1,\Delta x_2) - V(\Delta x_1,0) - V(\Delta x_2,0)$. Therefore, the *sharply* increased van der Waals repulsions will easily overcome this *linear* subtraction, resulting in the *positive* coupling potentials and consequently the *positive* coupling strengths of eq 17. Based on the normal-mode analysis, these side-chain squeezing or wagging motions are quite prevalent in THz region, from which it follows that the coupling between these motions leads to stiffening of the

effective potential surfaces from eq 4 and thus blue-shifting of the VSCF-corrected frequencies from the harmonic frequencies.

In contrast, the channel-breathing modes between 80 to 120 $cm^{-1}$ generally involve the *synchronous* vibrations of side chains toward the channel centers ($\mathbf{6_1}$ axis) as illustrated in Figure S1 (Supporting Information); in other words, they generally do not introduce the inner surface squeezing or side-chain wagging that may experience strong van der Waals repulsions. In this case, they are normally weakly coupled modes. This can be confirmed by examining the coupling strength contour. In Figure S2, these channel-breathing motions are highlighted by the vertical white lines, from which one finds that they are generally located in the shallow area of the contour, which corresponds to the small coupling strengths.

**3.3.2. VSCF Discrepancies.** As discussed above, the rapidly increased van der Waals interaction, due to couplings between similar side-chain squeezing modes, leads to the blue-shifts that improve VSCF calculations. However, care must be taken during a VSCF calculation to avoid the side-chain atoms being too proximate and the subsequent van der Waals repulsions being too large. There are two reasons for this: (1) It is well-known that the theoretical repulsion force field should depend exponentially on the distance; however, for ease of computing, the Lennard-Jones formula employs the $(1/r)^{12}$ term to describe the repulsion, which could be inaccurate at very short distances. Consequently, the classical force field-based VSCF calculation may also be inaccurate in this situation. (2) In a previous study of $H_2SO_4 \cdot H_2O$ system, Miller et al. pointed out that when a treatment of rectilinear normal modes leads to unphysically large anharmonic coupling, VSCF usually gives poor results.[18] The proximate side-chain atoms and rather strong repulsions are physically very unrealistic, and thus the VSCF calculations that can lead to these situations are expected to be inapplicable.

VSCF under these circumstances may yield insight into the discrepancies for the IV system and for the higher frequency region (2−3 THz): (1) The VSCF inapplicability due to proximate distance is more likely to be found when the side chains are very crowded. In the present VA class dipeptide nanotubes, IV and VI have the most bulky sides chains; compared to VI, IV has the larger pore size (3.9 Å vs 3.7 Å) and more empty space (35.6% vs 34.8%).[28,60] Therefore, the bulky side chains of IV are confined in an even smaller space and hence are most crowded. In this case, the physically unrealistic proximity and repulsion between side-chain atoms are very likely, and the VSCF calculation of IV is expected to deteriorate. (2) For the present dipeptides nanotubes, this VSCF inapplicability is also more likely to be found for the modes involving side-chain wagging or squeezing that can lead to strong van der Waals repulsions. Due to the strong van der Waals repulsions, the anharmonicity of this mode type is expected to be very large and the potential energy surface resembles a square well. This is very similar to the shearing and torsional modes in water clusters as well as $H_2SO_4 \cdot H_2O$ system reported by Gerber's group.[18,53] In this case they found that, due to the VSCF inapplicability, these modes tended to be dramatically blue-shifted. Similarly, we suggest their counterparts in the present system will likewise be considerably blue-shifted to the higher-frequency region, leading to the deterioration of VSCF spectra in the higher-frequency region (2−3 THz).

**3.4. Vibrational Degeneracy.** Due to the intrinsic symmetry of the crystalline VA class dipeptide systems, degenerate vibrational modes are prevalent as indicated by Table S of the Supporting Information. In some cases, the VSCF calculations reduced the number of degenerated states but did not eliminate them entirely. For example, modes 25, 26 and 27 of IA share very close harmonic frequencies, and after VSCF corrections, mode 25 and 26 still retain the degeneracy, whereas mode 27 is split out (Table S, Supporting Information). Figure 6 demonstrated the vector field representations of these three modes, from which one finds that mode 26 is actually nearly identical to mode 25 after a $C_3^1$ rotation, but mode 27, a breathing motion as discussed in section 3.2, distinguishes the others under any circumstances. Here we consider the symmetrically identical modes 25 and 26 as an "E" symmetry pair, whereas the distinct mode 27 is considered to have "A" symmetry. This implies that the VSCF calculation can split the degenerate "A" and "E" symmetry modes; however, it does not split the "E" states. This is not surprising since, by applying periodic boundary conditions and maintaining the lattice symmetry, the present VSCF calculation should not discriminate between these symmetrically identical modes that are just aligned along different symmetry axes.

Nevertheless, the coupling between the E mode pair can be included by applying the configuration interaction (CI) calculations, which is generally implemented in CC-VSCF when degenerate excited states are involved.[53] As previously discussed, CC-VSCF was not implemented in the present calculation, and hence the degenerate states interaction was not considered. Moreover, we interestingly found that most E modes in the present VA class dipeptide systems are intrinsically rather asymmetric; in other words, although these E modes can be overlapped after certain symmetric operation, the inner product between them at the *original* positions is quite small. For example, the inner product between structure A and C in Figure 6 is nearly 1.0, but that between structure A and B is only −0.000004. Subsequently, the coupling strength, or the correlation, between the E mode pair is also quite small, and hence the degenerate E mode couplings are neglected in the present calculation.

## 4. Conclusion

The terahertz (THz) spectra of the VA class dipeptide systems, including AI, IA, AV, VA, IV, VI and VV, were calculated using the CHARMM22 force field with periodic boundary conditions. The calculations were carried out at both the harmonic and vibrational self-consistent field (VSCF) levels; to the best of our knowledge, this is the first time that condensed-phase VSCF calculation of biomolecules has been performed. The calculated THz spectra were compared against the experimental THz spectra of crystalline solids in polyethylene pellets. We find that in general the VSCF level calculations demonstrated significant improvement, except for IV and the higher-frequency region (2−3 THz). The improvements were mostly reflected in the overall blue-shifts of the VSCF frequencies from the harmonic values.

To investigate the nature of these blue-shifts, the intermode coupling strength contours for all the VA class dipeptide systems were investigated. We find that all of the coupling

(60) Soldatov, D. V.; Moudrakovski, I. L.; Grachev, E. V.; Ripmeester, J. A. *J. Am. Chem. Soc.* **2006**, *128*, 6737–6744.

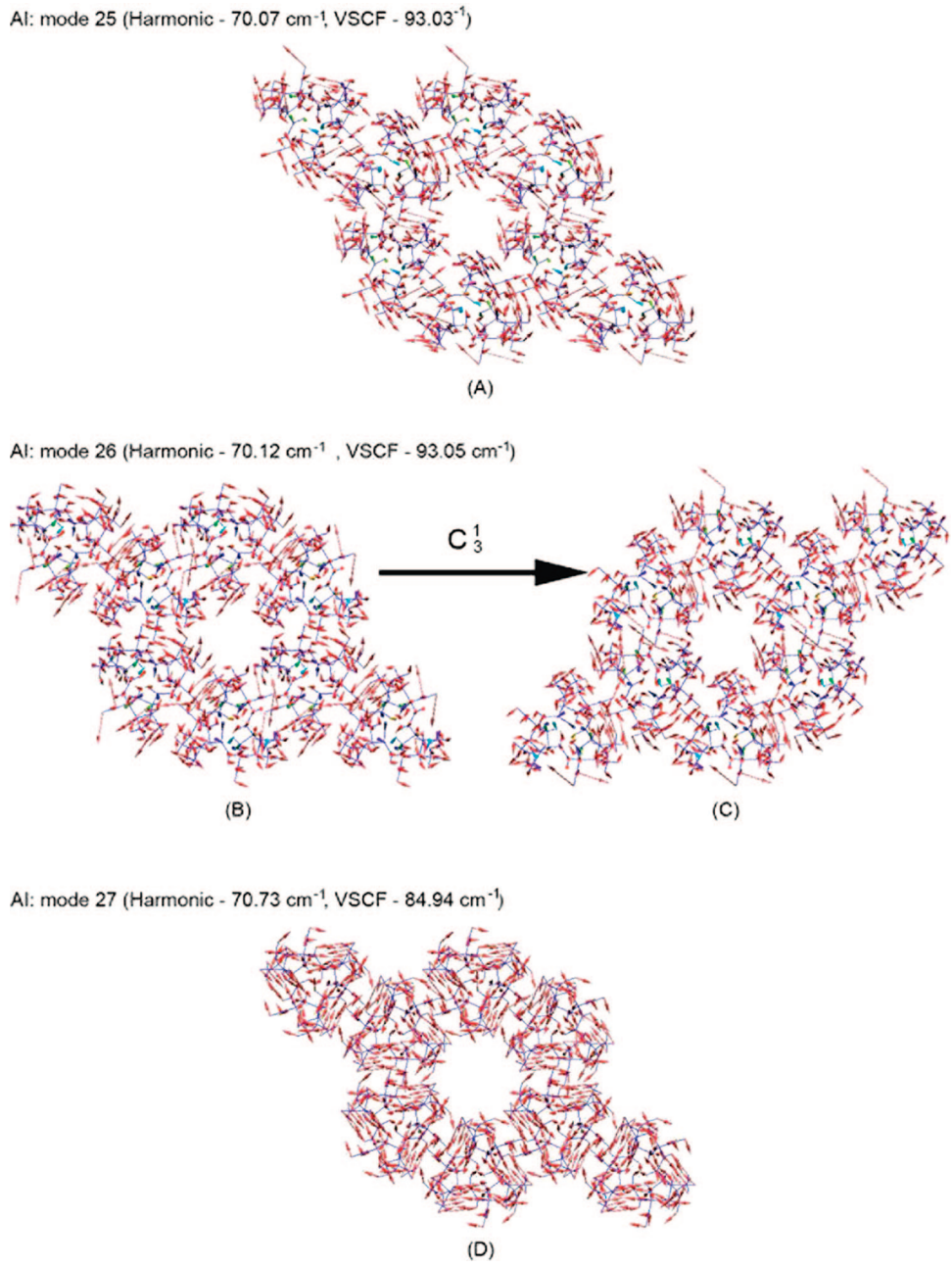


***Figure 6.*** Vector field representations of the degenerated normal modes of AI based on the vibrational calculations using CHARMM22 force field with periodic boundary conditions. The original structure of mode 26 is clockwise-rotated 120° to compare with mode 25. The dipeptide molecules are shown in blue. The arrows representing atomic vibrations are colored based on the vibrational amplitudes. The mode number and the corresponding VSCF frequency are shown on the top of each plot.

strengths are almost exclusively positive. By further investigating the most significantly coupled mode pair of each dipeptide system, we propose that when two similar side-chain squeezing modes are coupled, the sharply increased van der Waals interactions will lead to a stiffening of the effective potential, which eventually causes the blue-shifts in the VSCF frequencies. We also noted that VSCF tended to deteriorate when the side-chain atoms are too proximate and the subsequent van der Waals repulsions are too large, which accounts for the inaccuracies of IV and the high-frequency regions. We suggest that this physically unrealistic situation should be avoided in VSCF calculations, in which case higher-order correlations, such as CC-VSCF, or *ab initio* force field may be needed.

In addition, normal-mode analysis reveals a series of channel-breathing motions ranging from 80 $cm^{-1}$ to 120 $cm^{-1}$, which are proposed to facilitate the pathway of the guest molecules absorption, substitution and removal in the VA class dipeptides. Based on the inner products of the backbone vibrations between these channel-breathing motions, the VA class dipeptide systems were divided into two subgroups, **I**: AI, AV, IA and **II**: IV, VI, VV. The normal-mode analysis validated Görbitz's experimental findings with respect to the extraordinary flexibility and stability of the VA class dipeptide nanotubes. Normal mode analysis also demonstrated that THz motions may contribute to the pore permeability either directly by changing the pore size, or indirectly by affecting the solvent host effective potentials.

**Acknowledgment.** University at Buffalo, Center for Computational Research is acknowledged for providing the x86_64 Linux Cluster on 3.0 GHz Intel Xeon "Irwindale" Processors to compile the programs and implement the vibrational calculations.

**Supporting Information Available:** Complete ref 46, table of the calculated frequencies for the crystalline VA class dipeptide nanotubes at the harmonic and VSCF level using the CHARMM22 force field with periodic boundary conditions, the vector field representations of the channel-breathing motions and the most significantly coupled mode pairs, and the coupling strength contour map of all the VA class dipeptide systems. Request for all programs as well as the ($k$, $\Delta Q$) data to reproduce these calculations can be made to zhangh1@umbc.edu. This material is available free of charge via the Internet at http://pubs.acs.org.

JA805581N