Available online at www.sciencedirect.com
SCIENCE DIRECT®
Chemical Physics Letters 375 (2003) 337–343
CHEMICAL PHYSICS LETTERS
www.elsevier.com/locate/cplett

Terahertz spectroscopy of short-chain polypeptides

M.R. Kutteruf [1], C.M. Brown [2], L.K. Iwaki [3], M.B. Campbell [4], T.M. Korter [5], E.J. Heilweil *

Optical Technology Division and NIST Center for Neutron Research, National Institute of Standards and Technology, Gaithersburg, MD 20899-8443, USA

Received 31 January 2003
Published online 10 June 2003

Abstract

Terahertz (THz) absorption and neutron vibrational spectra for solid-phase short-chain peptide sequences at 77 and 298 K are reported. We demonstrate that a high degree of spectral and structural information exists for these systems in the 1–15 THz (33–500 cm^{-1}) spectral range. The density and uniqueness of distinct spectral features for each system suggests that sequence-dependent structural information can potentially be extracted from these spectra using quantum mechanical solid-state modeling and theories.
Published by Elsevier Science B.V.

1. Introduction

A complete understanding of protein function and behavior in living systems hinges on a detailed picture of polypeptide secondary and tertiary structures and dynamics on a range of timescales [1–3]. In an effort to develop methodologies to directly monitor complex macromolecule dynamics in real time, we [4] and others [5,6] suggested that an atomic level picture of concerted motions of polypeptides and DNAs may be accessible through accurate measurement of low-frequency vibrational spectra.

These vibrations are expected to occur in the THz frequency regime and may be observed using Raman, low-energy neutron and infrared spectroscopies, and other techniques. In this Letter, we present a series of THz infrared and neutron vibrational densities of states spectra for short-chain solid-state polypeptides that demonstrate a high degree of complexity and structural information in the 1–15 THz (33–500 cm^{-1}) range. Starting with X-ray crystal structures, it is also suggested that

* Corresponding author. Fax: +301-869-5700.
E-mail address: edwin.heilweil@nist.gov (E.J. Heilweil).
[1] NIST Summer Undergraduate Research Fellow, Box C-1394 Bryn Mawr College, Bryn Mawr, PA 19010, USA.
[2] Department of Materials and Nuclear Engineering, University of Maryland, College Park, MD 20742 and NIST Center for Neutron Research, 100 Bureau Drive, Mailstop 8562, NIST, Gaithersburg, MD 20899, USA.
[3] Present address: 2128 Plaza Court, CEO Lasers, St. Charles, MO 63303, USA.
[4] SPARTA, Inc., 1911 N. Ft. Meyer Drive, Suite 1100, Arlington, VA 22209, USA.
[5] Optical Technology Division, NIST Guest Researcher.

0009-2614/03/$ - see front matter. Published by Elsevier Science B.V.
doi:10.1016/S0009-2614(03)00856-X

application of modern ab initio structure and frequency calculations will enable assignment of these spectral features to both internal and external vibrational motions of these complex hydrogen-bonded molecular solids.

While it would be of highest priority to obtain low-frequency vibrational spectra for peptides and proteins in aqueous-phase environments that mimic their natural states, this is not immediately feasible with infrared spectroscopy because absorption by peptides or other biomolecules is masked by stronger solvent absorption in the THz spectral region [7]. The present study was performed to show that pure solid samples of low molecular weight protein fragments can indeed show sharp spectral features that are uniquely determined from molecular symmetries and structure. Spectral data presented and discussed here are believed to be the first of their kind and demonstrate that this expectation is borne out.

2. Experimental

Solid lyophilized samples of various short-chain peptide sequences were used as received from ICN Biomedicals [8]. Powders requiring low-temperature storage were maintained at 273 K until use to prevent decomposition or exposure to atmospheric water. Samples were rapidly prepared by weighing 2–10 mg of each solid and homogenizing the material in a mortar and pestle. This procedure ensured particle sizes sufficiently smaller than THz wavelengths to reduce baseline offsets at higher frequencies arising from non-resonant light scattering. Each sample was thoroughly mixed with approximately 100 mg of spectrophotometric grade high density polyethylene powder (Sigma–Aldrich) and pressed as a pellet in a 13 mm diameter vacuum die at the lowest possible pressures (ca. 200 psi or 1.4×10^6 Pa) to minimize decomposition from transient heating. The 1.5 mm thick pellets have sufficient path length to eliminate etaloning artifacts (3 cm^{-1} period) in the spectra. Sample pellets were mounted in an aluminum sample holder fixed in position with a teflon securing ring for 298 K measurements. Pellets were also encapsulated in a brass fixture that adapts to the copper cold finger of a vacuum cryostat (Janis Research Company, Inc. model ST-100) [8] fitted with 3 mm thick high-density polyethylene windows for broadband 77 K infrared studies.

Infrared absorption spectra in the 0.2–12 THz range were obtained with a modified Nicolet Magna 550 Fourier Transform infrared spectrometer (FTIR) using a silicon-coated broadband beam-splitter and DTGS room temperature detector fitted with a high density polyethylene window. The dry-air purged, globar source FTIR has sufficient sensitivity to generate high quality spectra in the 1.2–12 THz range for both room temperature pellets and samples placed in a low-temperature cryostat (with two polyethylene windows) placed in the sample compartment. All measurements made using the FTIR (after ~1 h of sample compartment purging to eliminate water vapor interference) were obtained at 4 cm^{-1} spectral resolution and averaging 64 interferometric scans. In standard fashion, spectra were converted to optical density (OD) units after ratioing raw sample transmission spectra (T_{sample}) to that of a pressed 100 mg polyethylene blank (T_{PE}) disk ($OD = -\log_{10}(T_{sample}/T_{PE})$ obtained under identical acquisition conditions. Reproducibility of THz spectral features were checked by measuring several pellets containing varying amounts of each peptide.

For the neutron work, approximately 3 g of the powders were placed inside aluminum sample holders (ca. 1 cm diameter) and press-sealed with indium wire. The Fermi-Chopper Spectrometer (FCS) [9] and Filter-Analyzer Neutron Spectrometer (FANS) [10] were used to collect data over the energy transfer ranges of 0–15 THz (0–500 cm^{-1}) and from 5.8 to 39 THz (194–1300 cm^{-1}), respectively. FCS is located on a cold neutron guide at the National Institute of Standards and Technology Center for Neutron Research (NCNR) and collects data primarily in neutron energy gain, resulting in a characteristic degradation of resolution with increasing energy transfer. FANS is a thermal neutron spectrometer working best with samples at low temperature (15 K in our experiments) with instrumental resolution better than the FCS over the stated energy transfer range. The IR and neutron spectral intensities shown in this

paper were selected from several single runs and were not averaged. Since the absolute IR intensity is concentration dependent and difficult to quantify, the uncertainty in infrared optical densities was estimated to be less than 10% by comparing multiple spectra for the same material in different pellets (Type B uncertainty, $k = 1$).

3. Results and discussion

We show characteristic terahertz infrared absorption spectra in Figs. 1–4 for several different groupings of peptides. First, representative spectra for several dipeptides at room temperature are shown in Fig. 1. These include Gly–Ala, Ala–Ala, Pro–Gly and Leu–Gly (see caption for amino acid nomenclature) to show the differences and complexity of these spectra in the 1–15 THz range. Similarly, we have obtained uniquely identifying spectra of a more extended range of solid di-peptides (and single amino acids) that are not shown here. Varying pellet concentrations confirmed that these features originate from the molecular system

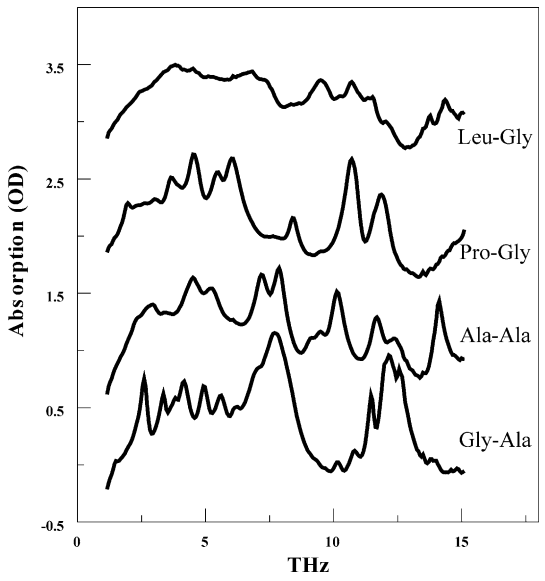

Fig. 1. Representative THz spectra of several solid dipeptides at 298 K in polyethylene matrix pellets (Gly, glycine; Ala, alanine; Pro, proline; Leu, leucine). The three upper spectra were shifted vertically by 1.0 optical density (OD) unit for clarity.

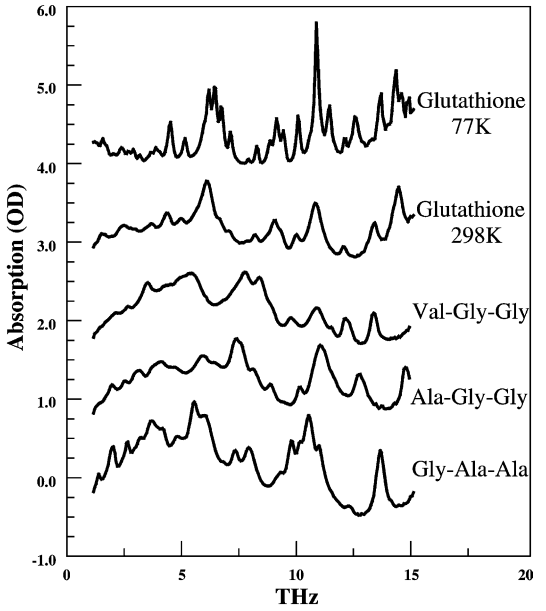

Fig. 2. THz spectra obtained for selected tripeptides at 298 K in polyethylene matrix pellets (Gly, glycine; Val, valine; Ala, alanine) and glutathione (Glu–Cys–Gly) at 298 and 77 K for comparison. The upper four spectra were shifted vertically by 1.0 optical density (OD) unit for clarity.

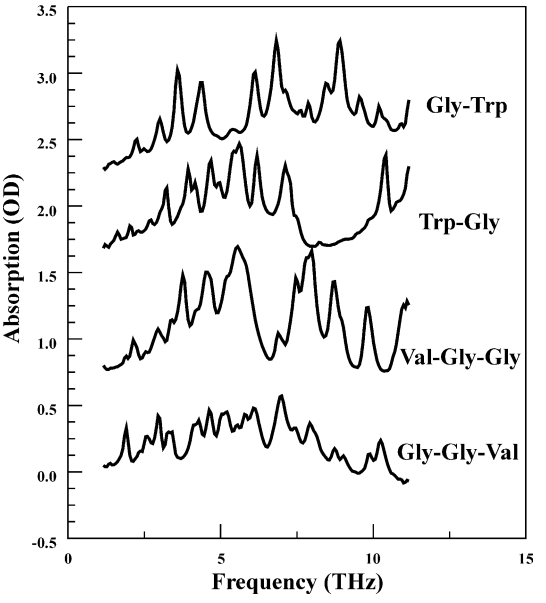

Fig. 3. THz spectra for Trp–Gly, Gly–Trp, Gly–Gly–Val, and Val–Gly–Gly (Trp, tryptophan; Gly, glycine; Val, valine) obtained at 77 K exhibiting a high degree of spectral content and sequence dependent crystalline mode specificity.

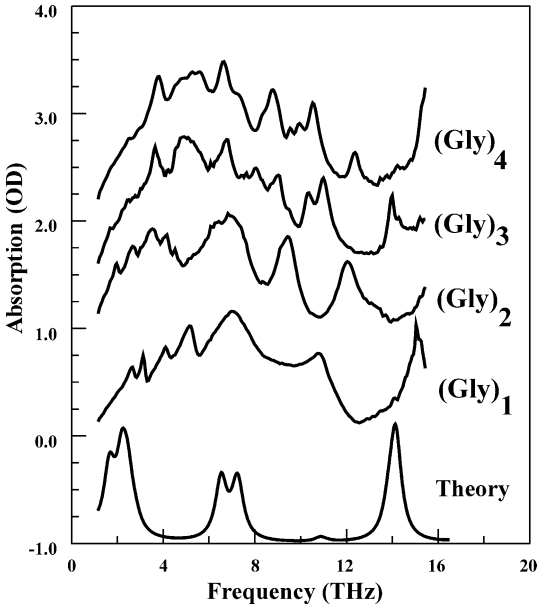

Fig. 4. THz absorption spectra of $(Gly)_n$ ($n = 1 - 4$) at 298 K exhibiting increasing mode density to lower frequencies as the chain length increases. Spectra are vertically displaced by 0.6 OD unit to eliminate overlap. A calculation of predicted frequencies and infrared absorption intensities for the four molecules in the unit cell of glycine $(Gly)_1$ is shown for comparison as the lower trace (see text).

rather than artifacts arising from pellet inhomogeneities. Subtle changes in the spectra were observed if temperature-sensitive samples were allowed to age in air at room temperature over several days. This potentially arises from sample degradation or polymerization (e.g., self-cyclization or intermolecular hydrolysis) of the peptides with time and elevated temperature [11].

Fig. 2 shows room temperature spectra for several representative linear tripeptide species. In general, the addition of another amino acid to a dipeptide chain adds additional absorption features arising from increased mode and spectral density. Each spectrum is uniquely identified and other tripeptide solids (not shown) exhibit similar but characteristically different and equally complex spectral structure. Fig. 2 also includes spectra for solid glutathione (sequence Glu–Cys–Gly) at 77 and 298 K. This demonstrates that a mixed tripeptide species exhibits multiple features in the 1–15 THz range that are persistent and weakly temperature dependent (features much sharper and better resolved at the lower temperature).

Terahertz absorption spectra for Gly–Trp, Trp–Gly, Gly–Gly–Val and Val–Gly–Gly at 77 K are presented in Fig. 3 for comparison. One readily observes that each species yields characteristic and highly structured spectra in the 1–15 THz range. The lower temperature again enhances spectral absorption features for these di- and tri-peptides. The observed spectral features are apparently also sequence dependent (dependent on N-terminus versus C-terminus ordering), indicating that infrared absorption spectra in this frequency range are potentially sensitive to amino-acid chain sequence, peptide composition, intermolecular hydrogen-bonding and crystal structure packing.

Fig. 4 demonstrates that for the homologous series of glycine, $(Gly)_n$ $n = 1 - 4$, being the simplest amino acid (a proton side chain), distinct and new absorption features are produced as the chain length increases. Weak absorption features below 4 THz are observed for monomeric glycine and $(Gly)_2$. These disappear for the longer-chain homologs $(Gly)_3$ and $(Gly)_4$. Higher frequency absorptions seem to shift to lower frequency as the chain length increases and the density of features increases. As also observed for the dipeptides and tripeptides (see Figs. 1 and 2), there appears to be an underlying very broad absorption between 1 THz to about 8 THz. This may arise from intermolecular hydrogen bonding between neighboring constituents, crystallite surface adsorbed water, or water incorporated within the solid samples.

We also show in Fig. 4 a spectral calculation using GAUSSIAN 98 [12] for four glycine molecules situated in the X-ray structure unit cell. This calculation was performed using B3LYP theory and the 6-31G(d) basis set [13]. The structure of the unit cell was fixed at the X-ray determined coordinates and no structural optimization was performed. There seems to be some agreement between the more intense features (at 7, 11 and 15 THz) for the experimental THz absorption spectrum and theory. While similar features are predicted for isolated 'gas phase' molecule calculations at high frequencies (for only intramolecular

glycine modes), the unit cell result is superior to any free molecule glycine calculation because lower frequency intermolecular modes contribute state density below ca. 5 THz. Furthermore, if one were to remove the broad underlying absorption from the experiment (between 1 and 8 THz), similar feature positions and intensities result. One problem with this simple calculation is that it yields several imaginary low frequency values (most likely because the starting structure is not at a global minimum, the X-ray structure is constrained or perhaps in a transition state configuration). Such calculations are also not designed to model the entire crystalline structure, but it does appear to serve as a useful starting point to obtain frequency and intensity predictions for solid-state spectra. More advanced solid state theories with rigorous treatment of environmental and intermolecular effects are expected to provide better frequency and infrared intensity estimates (see below).

We also obtained neutron spectra for the as-received diglycine peptide. The results of experimental investigations are shown in Fig. 5 where the neutron spectrum is displayed and compared to the THz absorption spectrum for this species. There is fairly good overlap of the three strongest neutron and IR absorption bands for diglycine occurring above 6 THz and at even higher frequencies where the FANS instrument yields high quality spectra (not shown). The frequency overlap of neutron scattering peaks with infrared absorptions near 7.5, 9 and 12 THz for the as-received material suggests that the corresponding vibrational modes arise from proton (hydrogen) motions of the glycine side-chain CH, backbone NH and end terminal groups. The lower frequency sharp neutron scattering and infrared absorption features observed for this species presumably arise from intermolecular crystalline phonon modes.

Deuteration of the labile OH and NH backbone and terminal group protons of diglycine, along with substitution of any possible crystallized water molecules with D_2O should simplify the neutron spectrum due to the very low scattering cross-section of deuterons compared to that of protons [13]. Since neutron spectra are almost

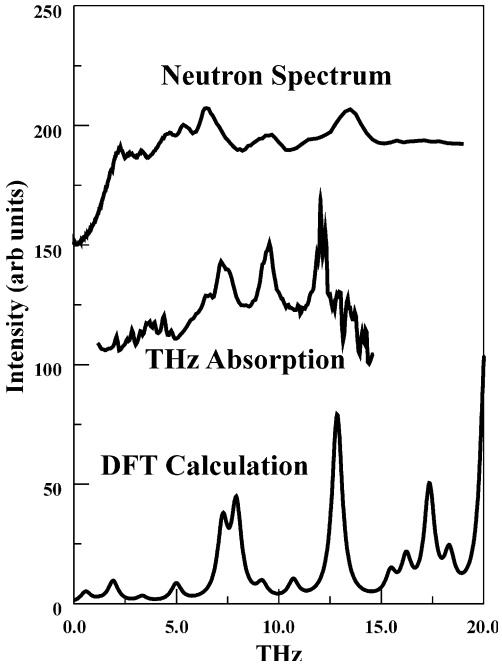

Fig. 5. THz absorption (298 K) and neutron scattering (15 K) spectra for solid protonated diglycine. Comparison to isolated molecule DFT calculation for the neutral planar structure shows similarities at the higher frequencies (intramolecular backbone motions) but neglects solid-state intermolecular interactions and hydrogen bonding that produces features at lower frequencies.

exclusively dominated by hydrogen atomic motion and the carbon-coordinated side-chain glycine protons are not expected to substitute upon deuteration, one would expect to observe neutron peaks that predominantly arise from side-chain motions. In addition, since neutrons are weakly interacting, one can often calculate using density functional theory (DFT), molecular mode frequencies and neutron scattering vibrational intensities for these relatively simple structures with more success than infrared intensities. Terahertz absorption, neutron spectroscopy and DFT calculations for diglycine and other deuterated species are being conducted and will be presented in a future publication.

Density functional theory structure and frequency calculations (using the GAUSSIAN 98 suite of programs utilizing the B3LYP/6-31+G(d,p) basis set) [14] were also tried to predict infrared

absorption spectra for isolated extended structures of the glycine monomer, dimer and trimer species. The resulting predicted IR absorption spectra for diglycine, obtained by convolving the calculated frequencies with a 20 cm^{-1} FWHM bandwidth function, is also displayed in Fig. 5. One observes that the predicted number and intensities of the higher frequency modes (>6 THz) overlap somewhat with the infrared absorption and neutron scattering spectra. This fair agreement reflects the possibility that these modes are predominantly internal molecular motions (since local environmental interactions are ignored for this gas phase calculation) that are not affected strongly by hydrogen-bonding interactions with the environment. The complex many-featured spectral absorptions below 6 THz are not modeled at all well, presumably again because all contributing intermolecular hydrogen-bonding and crystalline phonon modes are ignored by the current calculation. More of these types of calculations on related molecular systems are clearly needed to assess their applicability and usefulness for assigning molecular motions (modes) to the observed THz absorption and neutron spectral features.

The complexity and large number of observed IR absorption and neutron scattering features at low frequencies in the observed solid-state spectra suggests that intermolecular phonon modes must be taken into account in order to replicate the spectra. An additional complication is that strongly self-hydrogen-bonding species such as peptides may also form crystalline polymorphs with varying internal structures that will also contribute with different overall intensities to the THz spectra. In a collaborative effort to address these issues, solid-state DFT calculations are being performed by Dr. Mark Johnson at the Institut Laue Langevin in Grenoble, France for simple sugars such as α-D-glucose [13]. Starting with the known crystal structure for glucose, these calculations yield synthetic neutron spectra that compare extremely favorably in both frequency positions and intensities to the neutron spectra in the 5–20 THz regime. We envision successfully using this approach for the peptide systems examined in this work and will report these results in a more detailed publication.

4. Conclusions

Terahertz and neutron spectra for a variety of dipeptides and tripeptides in the solid state have been obtained and shown in this paper for comparison. Highly structured infrared absorption is typically observed for these species pressed into polyethylene pellets and studied at 298 K. Increased spectral structure arises at 77 K but without a large enhancement or increase in the number of absorption bands. The degree of complexity and structure of these THz spectra indicates that for small peptides, molecular solid-state structure, and sequence-dependent information may be contained in this region of the spectrum.

Neutron spectra were also obtained for fully protonated diglycine to compare to the THz spectra. Similar features and frequencies are observed from both spectroscopies indicating the higher energy THz spectrum is dominated by modes characterized primarily by side-chain hydrogen atom motion.

Quantum mechanical calculations for the isolated diglycine species suggest that higher frequency proton motions may be found in the THz and neutron spectrum but detailed structure at lower frequencies cannot be modeled without including solid-state, hydrogen-bonding and intermolecular interactions. Solid-state DFT calculations including these important sample characteristics are scheduled to be attempted for peptides and will be reported at a later date.

We are obtaining other small peptide homologs and longer chain length peptides to investigate whether THz spectral features persist as the molecular system becomes increasingly more complex and the opportunity for other crystalline polymorphs occurs. Additional studies of biologically active small protein sequences are also planned.

Acknowledgements

We gratefully thank Dr. Jacob Yeston for his thoughtful discussions and insights towards improving this manuscript. This work was supported through an internal NIST Competence Project entitled 'Advanced Terahertz Metrology' and STRS funding.

References

[1] R.H. Callender, R.B. Dyer, R. Gilmanshin, W.H. Woodruff, Annu. Rev. Phys. Chem. 49 (1998) 173.
[2] W.A. Eaton, V. Munoz, S.J. Hagen, G.S. Jas, L.J. Lapidus, E.R. Henry, J. Hofrichter, Ann. Rev. Biophys. Biomol. Struct. 29 (2000) 327.
[3] S. Takahashi, S.R. Yeh, T.K. Das, C.K. Chan, D.S. Gottfried, D.L. Rousseau, Nat. Struct. Biol. 4 (1997) 44.
[4] A.G. Markelz, A. Roitberg, E.J. Heilweil, Chem. Phys. Lett. 320 (2000) 42.
[5] R. Nossal, H. Lecar, Molecular and Cell Biophysics, first ed., Addison-Wesley, Redwood City, CA, 1991.
[6] K. Yamamoto, K. Tominaga, H. Sasakawa, A. Tamura, H. Murakami, H. Ohtake, N. Sarukura, Bull. Chem. Soc. Jpn. 75 (2002) 1083.
[7] J.T. Kindt, C.A. Schmuttenmaer, J. Phys. Chem. 100 (1996) 10373.
[8] Certain commercial equipment, instruments, or materials are identified in this paper to adequately specify the experimental procedure. In no case does identification imply recommendation or endorsement by NIST, nor does it imply that the materials or equipment identified are necessarily the best available for the purpose.
[9] J.R.D. Copley, T.J. Udovic, J. Res. Natl. Inst. Stand. Technol. 98 (1993) 71.
[10] J.R.D. Copley, D.A. Neumann, W.A. Kamitakahara, Can. J. Phys. 73 (1995) 763.
[11] A. Thomas, B. Roux, J.C. Smith, Biopolymers 33 (1993) 1249.
[12] M.J. Frisch et al., GAUSSIAN 98 (Revision A.11.3), Gaussian, Inc., Pittsburgh PA, 2001.
[13] B.S. Hudson, J. Phys. Chem. A 105 (2001) 3949.
[14] J. Yeston, T.M. Korter, M.R. Kutteruf, M.B. Campbell, C.M. Brown, M.R. Johnson, E.J. Heilweil, in preparation.