Pulsed Terahertz Spectroscopy of DNA, Bovine Serum Albumin and Collagen between 0.1 and 2.0 THz

A. G. Markelz[a,c*], A. Roitberg[b] and E. J. Heilweil[a]

a.)Optical Technology Division, Bldg. 221/Rm.B208, NIST, Gaithersburg, MD. 20899

b.) Biotechnology Division, Bldg. 222/Rm.A353, NIST, Gaithersburg, MD. 20899

c.) Materials Physics Research, 1D-410, Lucent Technologies, Murray Hill, NJ 07974

Abstract

We report the first use of pulsed terahertz spectroscopy to examine low frequency collective vibrational modes of biomolecules. Broadband absorption increasing with frequency was observed for lyophilized powder samples of calf thymus DNA, bovine serum albumin and collagen in the 0.06 THz - 2.00 THz, (2 cm^{-1} - 67 cm^{-1}) frequency range suggesting that a large number of the low frequency collective modes for these systems are IR active. Transmission measurements at room temperature showed increasing FIR absorption with hydration and denaturing.

Introduction

There is an increasing interest in studying low frequency collective vibrational modes occurring in proteins, DNA and oligonucleotides since these modes may provide information about a biomolecule's conformational state.[1] For example, low frequency modes are associated with collective motion of the tertiary subunits moving with respect to one another, or coherent movement of a portion of a structural subunit. Similar motions are associated with conformational movements that occur during ligand binding and are critical to protein function.[1] For DNA structures, Prohofsky and collaborators have predicted helix, base twisting, and librational modes in the $20\text{-}100 \text{ cm}^{-1}$ range.[2-4] Similarly, protein collective vibrational modes have been calculated to lie in the far-infrared (FIR) for bovine pancreatic trypsin inhibitor (BPTI) and human lysozyme.[5,6] In the past these vibrational modes have been studied with Raman and traditional FTIR transmission techniques.[7-9] While Raman studies have been extensive, there is some disagreement in band identification from various groups, possibly due to the application of complex Lorentzian line fitting procedures for extracting weak contributions to the main elastic scattering peak.[8,10]

FIR transmission studies have been limited due to the difficulty in accessing this frequency range, especially below 50 cm^{-1}. In addition, deconvolution of FIR transmission data for the real and imaginary parts of the dielectric function has required an iterative fitting process with a guess of the number of oscillators to include in the fit.[9,11,12] Finally, absorptions initially assigned below 100 cm^{-1} by earlier investigators

using FTIR techniques were later ascribed to etalon effects.[11,12] Here we use sensitive pulsed terahertz spectroscopy (PTS) to measure absorption down to 2 cm^{-1} and use sample preparation and data collection techniques to eliminate etalon effects in the absorbance measurements.

With reported signal to noise ratios of 10,000:1, PTS has become a preferable method for performing low frequency spectroscopy over standard FIR techniques.[13] Since PTS is a direct field measurement, determination of the complex dielectric function is straight-forward, without the need for fitting routines. While not yet explored, PTS should enable the determination of the sequence of intermediate tertiary structures involved in a conformational change using time-resolved pump - PTS probe measurements. These studies would be similar to time-resolved circular dichroism measurements of secondary structure evolution.[14] Towards this end we have performed the first known absorption measurements on biomolecular samples using PTS below 70 cm^{-1}. These measurements yield broadband FIR absorption spectra for calf thymus DNA, bovine serum albumin (BSA), and collagen. The measured absorption increase with optical pathlength and is hydration and conformation dependent. These results demonstrate that PTS is a viable technique for future time-resolved FIR measurements of protein folding, DNA helical motion, and base pair hydrogen bonding dynamics.

Experimental

In previous FIR transmission studies using FTIR techniques, Powell et al. discussed difficulties in fabricating durable free standing thin films of sufficient thickness to produce an adequate optical density for frequencies below 40 cm^{-1}.[15] Adequate optical density was achieved in the current studies using pressed pellets of the commercially prepared lyophilized powders. Calf thymus DNA was obtained from SIGMA, Lot #46H8874, average molecular weight, MW, of ~ 12 MDa (1 MDa=10^6 gm/mole).[**] The supplied DNA was dissolved in NaCl prior to lyophilization, implying that the counterion is Na. BSA is a transporter blood protein of 550 residues, MW = 66 kDa. Its crystallographic structure consists of nine α-helices with a heart shaped tertiary conformation (PDB code for Human Serum Albumin is 1AU6, BSA is > 73 % homologous to HSA).[16] Lyophilized BSA was obtained from Sigma, lot #107H0649. Type I Collagen consists of a three polypeptide superhelix. Lyophilized collagen was obtained from Sigma, lot #107H7049, average MW = 0.36 MDa. Pellet density was determined by measuring the thickness and weight at 66% relative humidity (r.h.).

Multiple reflection etalon effects were eliminated in absolute absorbance measurements at < 5% r.h. by mixing the lyophilized biomolecular powder samples with 200 mg of polyethylene (PET) powder such that the final pellet thickness was > 7.5 mm. Given the index of refraction of PET is 1.5, the etalon spacing due to multiple reflections in the pellets would be 0.4 cm^{-1} which is slightly less than the 0.5 cm^{-1} spectral resolution of these measurements. Each sample transmission spectrum was normalized to a 200 mg pure PET pellet (PET has low FIR loss) and all pellets were pressed at 6.9x10^7 Pa (10^4 PSI) at room temperature.

We used two PTS systems based on photoconductive switches embedded in Hertzian dipole antennas.[17] By using two different spectrometers for these measurements we demonstrate below that the reported spectra are independent of specific system details. The first PTS system is similar to that described elsewhere[18] however here the

optical system includes two parabolic mirrors for focusing the terahertz light onto the sample and recollecting the transmitted light for detection. The spot size measured by scanning a pinhole at the sample position was 2 mm FWHM. The main source of irreproducibility was beam pointing drift of the ultrafast 800 nm Ti:Sapphire laser used for generating and detecting the THz pulses. To compensate for this drift, short repeated measurements were made of the sample and the reference pellets. The sample holder consisted of two 6 mm diameter apertures for the sample and the reference. A measurement consisted of three scans of the transmitted THz electric field for the sample then three scans for the reference sample. A total of 10 consecutive sample (reference) measurements were averaged to determine a transmission spectrum. The reproducibility of the relative transmission $T_s = E_{t,s}/E_{t,r}$ was determined by the ratio of an average of five consecutive measurements to the average of the next five consecutive measurements. A typical power spectrum, $P(\nu) \sim E(\nu)^2$, and a reproducibility test are shown in Fig. 1.

For hydration and conformation dependence measurements, self-supporting disks were pressed from the lyophilized powders (without PET) so that the PET filler would not interfere with the hydration of the samples. However, the spectra from these thinner pellets (about 0.2 mm thickness) were contaminated with etalon effects as discussed below. Humidity dependence measurements were achieved by mounting the samples in a controlled humidity cell 90 minutes prior to scanning interferograms. The controlled humidity cell consisted of an adapted commercially available spectroscopic flow cell with 0.5 cm thick polyethylene windows. This thickness was chosen to eliminate etalon spectral oscillations from the windows. One window was machined with two displaced 10 mm diameter 1.1 mm deep recesses, one to accommodate the 7 mm diameter pressed pellets, and the second for positional reproducibility measurements. Samples were held in place with a stainless steel washer (10 mm OD, 5 mm ID) which defined the optical aperture. Controlled relative humidity in the cell was achieved by circulating air with a peristaltic pump above a saturated aqueous salt solution reservoir and measuring the r.h. with a meter after the cell. The accuracy of the r.h. meter was ~1% and the lowest r.h. that could be accurately measured was 5 %.

Denatured BSA samples were made by dissolving the lyophilized powder into a phosphate buffer solution (pH=7.07) which was then heated to 363 K for 20 minutes. At these temperatures the conformational change is irreversible.[19] The resultant opaque gel was dried in a fume hood, crushed to a powder and subsequently pressed into pellets for transmission measurements.

The PTS system used for hydration and conformation measurements is discussed in detail elsewhere.[20] In this study the receiver antenna was similar to the one used in Ref. 20 but fabricated on low temperature grown GaAs. Typical power spectra were similar to that shown in Fig. 1. Interferogram data are taken to provide 0.5 cm^{-1} spectral resolution and 200 consecutive spectral scans were averaged per measurement. All measurements were made at room temperature, and both PTS systems were enclosed in dry nitrogen purged boxes to diminish FIR absorption due to ambient humidity.

Results

In Figure 2 we show the measured absorbance, $A = - \log (E_{t,s}/E_{t,r})$, for several biomolecular samples, where $E_{t,s}$ is the THz field transmitted through the sample and $E_{t,r}$ is the THz field transmitted through the reference. From these measurements we wished to confirm linear absorption and also directly determine ε, the molar absorptivity from $A = \varepsilon c l$ where c is the molar concentration and l is the pathlength. For DNA samples: #1 $cl = 1.3 \times 10^{-6}$ M·cm and $\varepsilon(20$ cm$^{-1}) = 1.9 \times 10^5$ and #2 $cl = 2.6 \times 10^{-6}$ M·cm and $\varepsilon(20$ cm$^{-1}) = 2.0 \times 10^5$. Similarly for the BSA samples, sample #1 had $cl = 0.655 \times 10^{-3}$ M·cm and $\varepsilon(20$ cm$^{-1}) = 0.90 \times 10^3$, sample #2 had $cl = 1.122 \times 10^{-6}$ M·cm and $\varepsilon(20$ cm$^{-1}) = 0.95 \times 10^3$ and sample #3 had $cl = 2.806 \times 10^{-3}$ M·cm and $\varepsilon(20$ cm$^{-1}) = 0.96 \times 10^3$. For the collagen samples, sample #1 had $cl = .206 \times 10^{-3}$ M·cm and $\varepsilon(20$ cm$^{-1}) = 4.1 \times 10^3$ and sample #2 had $cl = .514 \times 10^{-3}$ M·cm and $\varepsilon(20$ cm$^{-1}) = 3.8 \times 10^3$. Thus for all samples examined we observe good agreement with Beer's law behavior. The absorbance to higher frequencies is strong and broadband for all samples. For both DNA and BSA the absorbance increases nearly linearly with frequency. For collagen, however, the absorbance increases more rapidly with increasing frequency. Also, in Fig. 2 we show two spectra from pressed pellets of pure DNA and BSA. In order to compare the pure pellet data to the PET mixed pellet data we normalize the absorbance of the pure pellets, A_{norm}, using $A_{norm} = A_p * c_m l_m / c_p l_p$ where A_p is the absorbance of the pure pellet and c_m (c_p) and l_m (l_p) are the concentration and length of the mixed (pure) pellet. This normalization depends only on the measured pellet parameters c and l and is independent of the actual spectra. We note that using such a simplistic normalization assumes the absorbance does not depend on density and that PET is not interacting with the sample. Surprisingly the data nearly overlay for the entire frequency range suggesting that these assumptions are at least in part valid. One readily observes the similarity of the spectral shapes for these samples compared to those pressed in transparent PET.

Figure 3(a) shows the hydration dependence of the DNA spectra. The spectra are clearly different than those in Fig. 2(a) for < 5% r.h. The low frequency absorption below 15 cm^{-1} diminishes from that in Fig. 2. Similarly for BSA, as shown in Figure 3(b), The entire transmission spectrum shifts down in frequency as the native BSA is hydrated, as was seen in the DNA samples. The denatured BSA at <5% r.h. has still lower transmission at the low frequencies.

We also investigate whether PTS can be used to measure the level of hydration for biomolecular samples by utilizing the direct field measurement of the real part of the refractive index. To determine the real part of the index directly from the transmission measurements we note that the field transmittance, $t = E(\nu)/E_{ref}(\nu)$ ($E(\nu)$ is the field spectrum transmitted through the sample and $E_{ref}(\nu)$ is the field spectrum for an empty aperture), is related to the complex index $N = n + i\kappa$ by:

$$t = \frac{4N}{(1+N)^2} e^{i2\pi(N-1)d\nu} = |t| e^{i\varphi}$$

(1.)

where d is the sample thickness and ν is the frequency in cm^{-1}. For $n \sim 1.5$, $d \sim 200$ μm and $\nu = 25$ cm^{-1}, the phase shift ϕ in the exponential factor would be 1.57 rad. If κ was as large as n/3, or T = 0.04, the phase shift given by the Fresnel prefactor is only -0.07 rad, i.e. for the sample thicknesses used in these studies the phase shift due to the

interfaces is negligible for $\kappa < n$, or $T > 0.10$. This condition is satisfied for the entire frequency range for DNA samples and for $\nu < 30$ cm^{-1} for BSA samples. Thus the index and extinction coefficient are given by:

$$n = \frac{\phi}{2\pi d\nu} + 1 \qquad (2\ a.)$$

$$\kappa = -\frac{1}{2\pi d\nu} \ln \frac{4n}{|t|(1+n)^2} \qquad (2\ b.)$$

Our measurements of $\varepsilon = (n + i\kappa)^2$ as a function of sample density give good agreement with Clausius Mossotti behavior, giving polarizability magnitudes at $\nu = 25$ cm^{-1} and r.h.<5 % of: $|\gamma_{mol}|$= 1900 nm^3 for calf thymus DNA; $|\gamma_{mol}|$ = 9.6 nm^3 for BSA; and $|\gamma_{mol}|$= 88 nm^3 for calf skin collagen. The frequency dependent real part of the refractive index (n) and absorption coefficient ($\alpha = 4\pi\kappa\nu$) are shown in the inset of Figure 4 for DNA #3 described in Fig. 2 and Fig 3.

For a given sample density the dependence of the real part of the index on relative humidity was measured, shown in Fig. 4 for $\nu = 25$ cm^{-1}, where absorption is small enough that the condition of $\kappa<n$ is fulfilled. The error bars shown for the DNA n values are the $\pm 1\sigma$ error limits for a total of 5 measurements for each point, where a measurement consists of removal of the sample from the purged PTS system and taking the measurement again at least 24 hours later. The BSA values are for single measurements. The real part of the DNA refractive index increases as a function of r.h. and then levels off to a value of 1.5. The highest r.h. reached was 77%, and no condensation occurred on the samples or the windows. The index of the BSA samples has a small decreasing dependence on relative humidity over the range measured.

Discussion

We first address the possible structural forms of the molecules in lyophilized powder pellet form. Crystallographic measurements on thin films of calf thymus DNA with Na counterion give the helical B form for r.h. > 92%, helical A form for 92% > r.h. > 45%, and a disordered helical A form for r.h.< 45%.[7] The triple helix structure of collagen remains intact down to the lowest measured hydration of 20 % r.h.[21] The main effect of dehydration for the DNA and collagen samples is the diminishing inter-strand coupling as H_2O is removed. Measurements on recombinant human serum albumin, HSA, show a reversible increase in β-sheet content from r.h. < 16% and decrease of α-helix content from r.h. 58% to 30% upon lyophilization.[22] Given that HSA and BSA are highly homologous, we expect similar structural content change for the BSA samples.[16]

While it has been shown that weight change occurs immediately upon exposure to new humidity conditions, structural equilibrium is not achieved for several days.[7] Thus with regard to the net hydration achieved in the optical measurements, we can only comment on trends in the FIR spectral data and not make quantitative hydration correlations. As shown in Fig. 4, the real part of the refractive index for DNA samples increases with hydration and then reaches a constant value, whereas for BSA no such increase occurs. We are not aware of any other *direct* measurement of the refractive index as a function of r.h. at these frequencies. Microwave refractive index measurements on DNA thin films at 5 cm^{-1} showed a similar rise with hydration.[12] This behavior can be

explained by assuming that the net measured dielectric response is an average from both the DNA and the water of hydration. Given that the refractive index for dry DNA and pure water are 1.3 and 2.3 respectively at 25 cm^{-1},[23] one expects that the average index would increase with water content. The lack of hydration dependence for BSA may result from a lower total number of water molecules absorbed per unit surface area for the globular BSA structure compared to DNA, which has many backbone phosphate and sugar sites available for hydration.

As DNA is hydrated the absorption features in the FIR transmission spectra appear to shift to lower frequencies. We compare the overall behavior of the increasing FIR absorption with increasing hydration to the well established 25 cm^{-1} Raman-active mode measured for DNA thin films.[10] Measurements by Weidlich et al. showed that as the r.h. is reduced the 25 cm^{-1} mode shifts up in frequency and broadens until at 0 % r.h. the mode is nearly indiscernible, parallel to the behavior seen in Figure 3(a) . Shifting of the 25 cm^{-1} Raman mode was adequately accounted for by mass loading by absorbed water.[7,10]

The limited FIR polypeptide spectroscopy performed with FTIR under low humidity conditions showed broad featureless absorptions down to 20 cm^{-1}.[24] As is surmised for DNA, the FIR absorption shift arising from the hydration of BSA may be due to mass loading by water. The increased overall low frequency FIR absorption for denatured material maybe a consequence of increased hydrogen bonding within the structural network as the protein sample denatures to a gel.

Conclusions

We have demonstrated the use of PTS for measuring low frequency absorption spectra and refractive indices of biomolecular pressed-pellet samples as a function of relative humidity. All three biomolecules studied have broadband FIR absorption suggesting that a large density of low frequency collective modes are indeed IR active. This gives rise to the possibility of preferentially exciting a particular mode via initiating a conformational change associated with that mode. Future experimental plans include spectral studies of lyophilized solid samples of short-chain DNA oligomers, and application of this technique to obtain time-resolved THz absorption spectra after photo-initiation of a biopolymer conformational change.

* Corresponding author. Present address: amarkelz@acsu.buffalo.edu, Physics Dept., SUNY, Buffalo, NY 14260.

** Certain commercial materials are identified in this paper in order to adequately specify the experimental procedure. In no case does such identification imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the materials or equipment are necessarily the best available for the purpose.

Acknowledgments

References

[1] R. Nossal and H. Lecar, *Molecular and Cell Biophysics*, 1st ed. (Addison-Wesley, Redwood City, CA, 1991).

[2] W. Zhuang, Y. Feng, and E. W. Prohofsky, *Phys. Rev. A* **41,** 7033 (1990).

[3] L. Young, V. V. Prabhu, E. W. Prohofsky, and G. S. Edwards, *Phys. Rev. A* **41,** 7020 (1990).

[4] Y. Feng, W. Zhuang, and E. W. Prohofsky, *Phys. Rev. A* **43,** 1049 (1991).

[5] A. Roitberg, R. B. Gerber, R. Elber, and M. A. Ratner, *Science* **268,** 1319 (1995).

[6] S. Hayward and N. Go, *Annu. Rev. Phys. Chem.* **46,** 223 (1995).

[7] S. M. Lindsay, S. A. Lee, J. W. Powell, T. Weidlich, C. Demarco, G. D. Lewen, N. J. Tao, and A. Rupprecht, *Biopolymers* **27,** 1015-1043 (1988).

[8] V. Lisy, P. Miskovsky, B. Brutovsky, and L. Chinsky, *J. Biomol. Struct. Dyn.* **14,** 517 (1997).

[9] J. W. Powell, W. L. Peticolas, and L. Genzel, *J. Mol. Struct.* **247,** 119 (1991).

[10] T. Weidlich, S. M. Lindsay, Q. Rui, A. Rupprecht, W. L. Peticolas, and G. A. Thomas, *J. Biomol. Struct. and Dyn.* **8,** 139 (1990).

[11] T. Weidlich, J. W. Powell, L. Genzel, and A. Rupprecht, *Biopolymers* **30,** 477-480 (1990).

[12] A. Wittlin, L. Genzel, F. Kremer, S. Haseler, and A. Poglitsch, *Phys. Rev. A* **34,** 493 (1986).

[13] R. A. Cheville and D. Grischkowsky, *Opt. Lett.* **20,** 1646 (1995).

[14] E. F. Chen, M. J. Wood, A. L. Fink, and D. S. Kliger, *Biochem.* **37,** 5589-5598 (1998).

[15] J. W. Powell, G. S. Edwards, L. Genzel, F. Kremer, A. Wittlin, W. Kubasek, and W. Peticolas, *Phys. Rev. A* **35,** 3929 (1987).

[16] D. C. Carter and J. X. Ho, *Adv. Protein Chem.* **45,** 153 (1994).

[17] N. Katzenellenbogen and D. Grischkowsky, in *Ultra-Wideband, Short-Pulse Electromagnetics*, edited by Betroni (Plenum Press, New York, 1992).

[18] Y. Cai, I. Brener, A. J. Lopata, J. Wynn, L. Pfeiffer, J. B. Stark, W. Q, X. C. Zhang, and J. F. Federici, *Appl. Phys. Lett.* **73,** 444 (1998).

[19] *Advances in Infrared and Raman Spectroscopy*, Vol. 1, edited by B. G. Frushour and J. L. Koenig (Heyden, 1975).

[20] A. G. Markelz and E. J. Heilweil, *Appl. Phys. Lett.* **72,** 2229 (1998).

[21] R. D. B. Fraser and T. P. MacRae, *Conformation in fibrous proteins* (Academic Press, New York, 1973).

[22] K. Griebenow and A. M. Klibanov, *Proc. Natl. Acad. Sci.* **92,** 10969 (1995).

[23] J. T. Kindt and C. A. Schmuttenmaer, *J. Chem. Phys.* **100,** 10373 (1996).

[24] L. Genzel, L. Santo, and S. C. Shen, in *Spectroscopy of Biological Molecules*, edited by C. Sandory and T. Theophanides (D. Reidel, Boston, 1984), pp. 609-619.
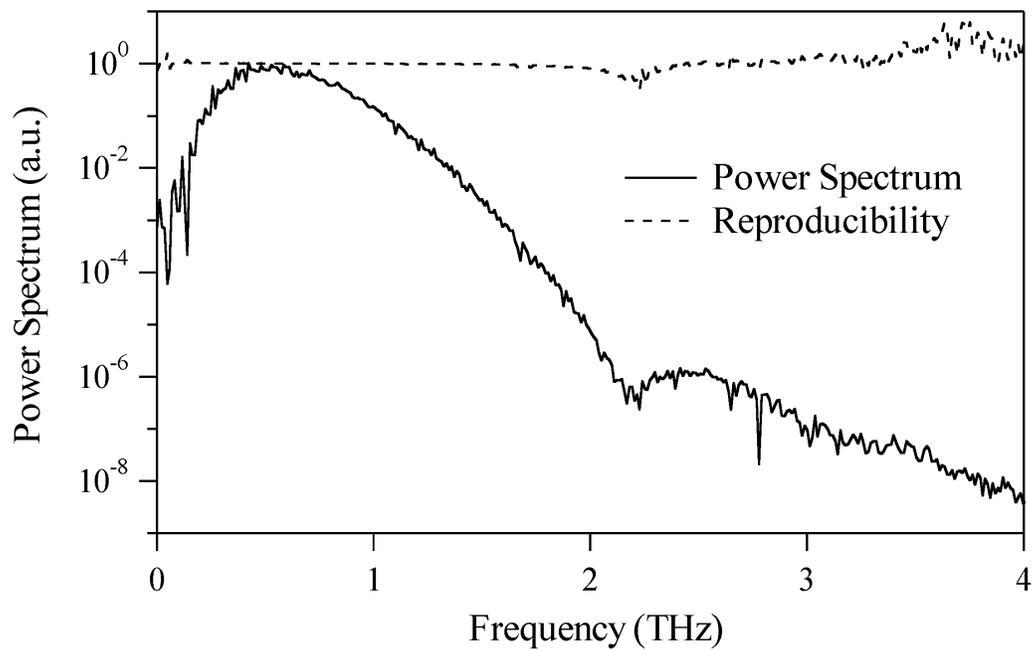
Figure 1. Representative power spectrum and baseline reproducibility for PTS systems used for these measurements.
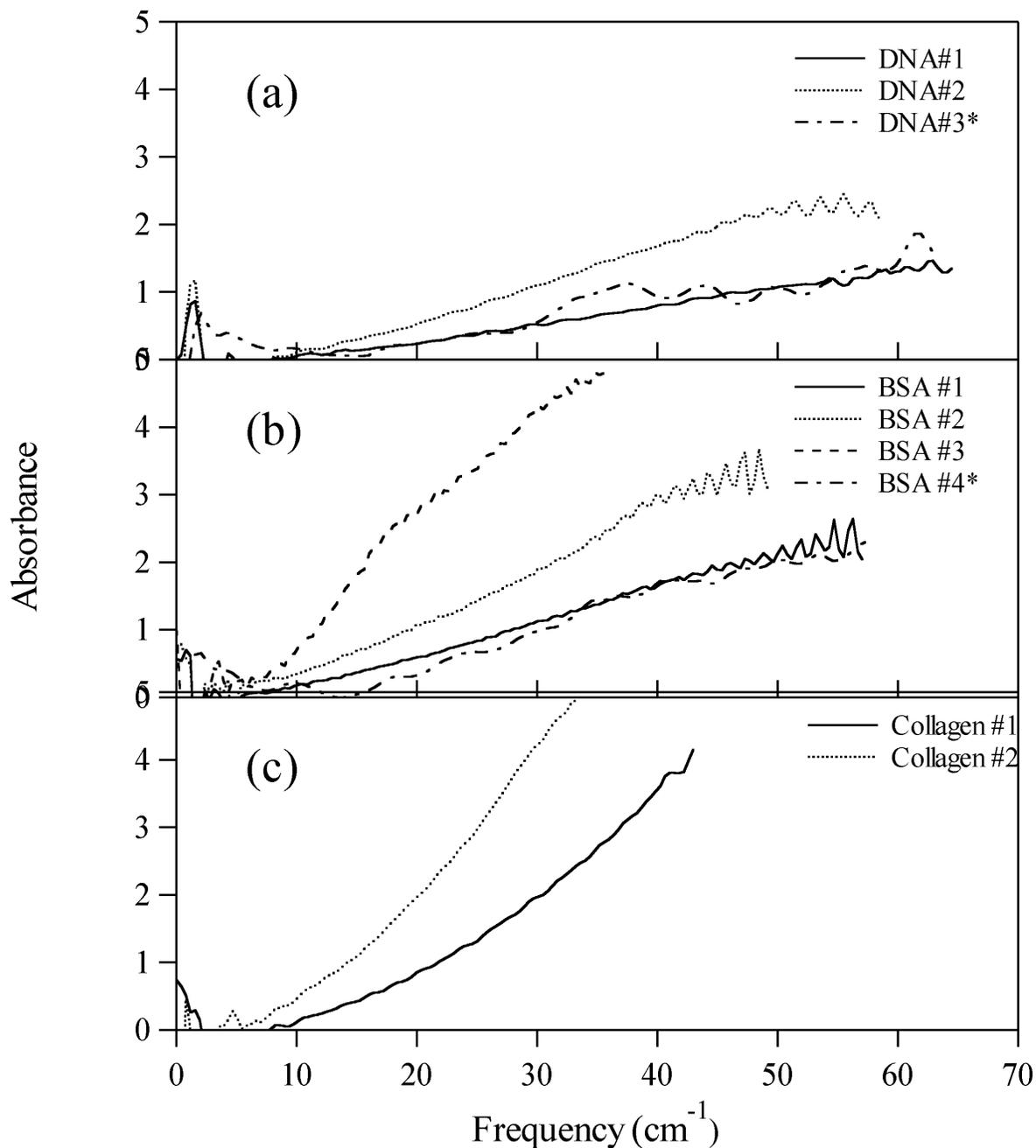
Markelz, Roitberg and Heilweil

Figure 2. Absorbance of biomolecular pressed pellet samples at r.h.<5%: (a) DNA #1 and #2 DNA lyophilized powder mixed with PET whereas #3 is the absorbance for a pure DNA lyophilized powder pellet renormalized to the equivalent molecular path length to DNA #1, see text. (b) BSA #1, #2 and #3 are BSA lyophilized powder mixed with PET whereas #4 is pure BSA lyophilized powder pellet renormalized to the equivalent molecular path length to BSA #1, see text. (c) Collagen #1 and #2 are Collagen lyophilized powder mixed with PET. Sample preparation described in the text. The data is truncated to remove values laying outside the dynamic range of the PTS system, defined by the ratio of the noise floor intensity to the reference transmitted intensity.
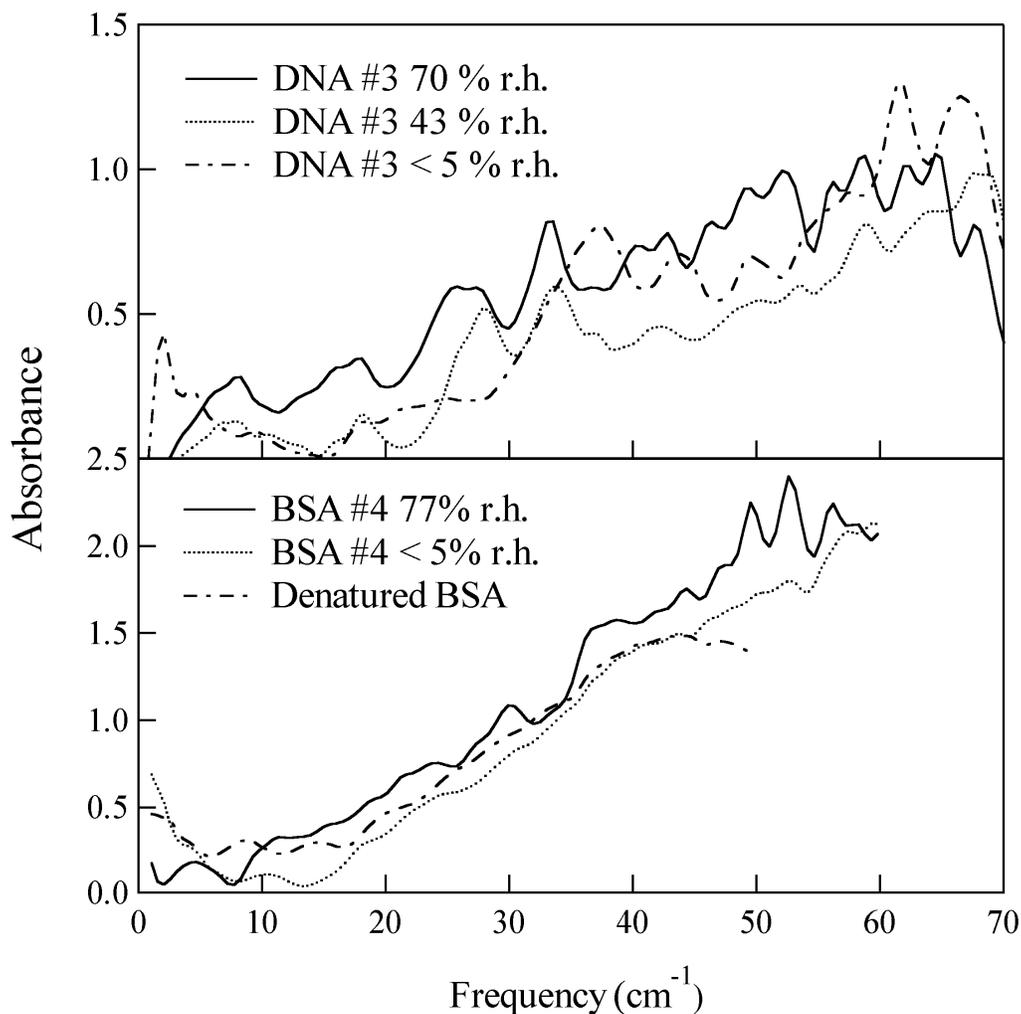
Figure 3. (a) Absorbance hydration dependence for sample DNA #3 also shown in Fig. 2(a). (b) Absorbance hydration dependence for BSA #4 shown in Fig. 2(b), and transmittance of a denatured BSA sample with r.h. < 5%. The data is truncated to remove values laying outside the dynamic range of the PTS system, defined by the ratio of the noise floor intensity to the reference transmitted intensity.
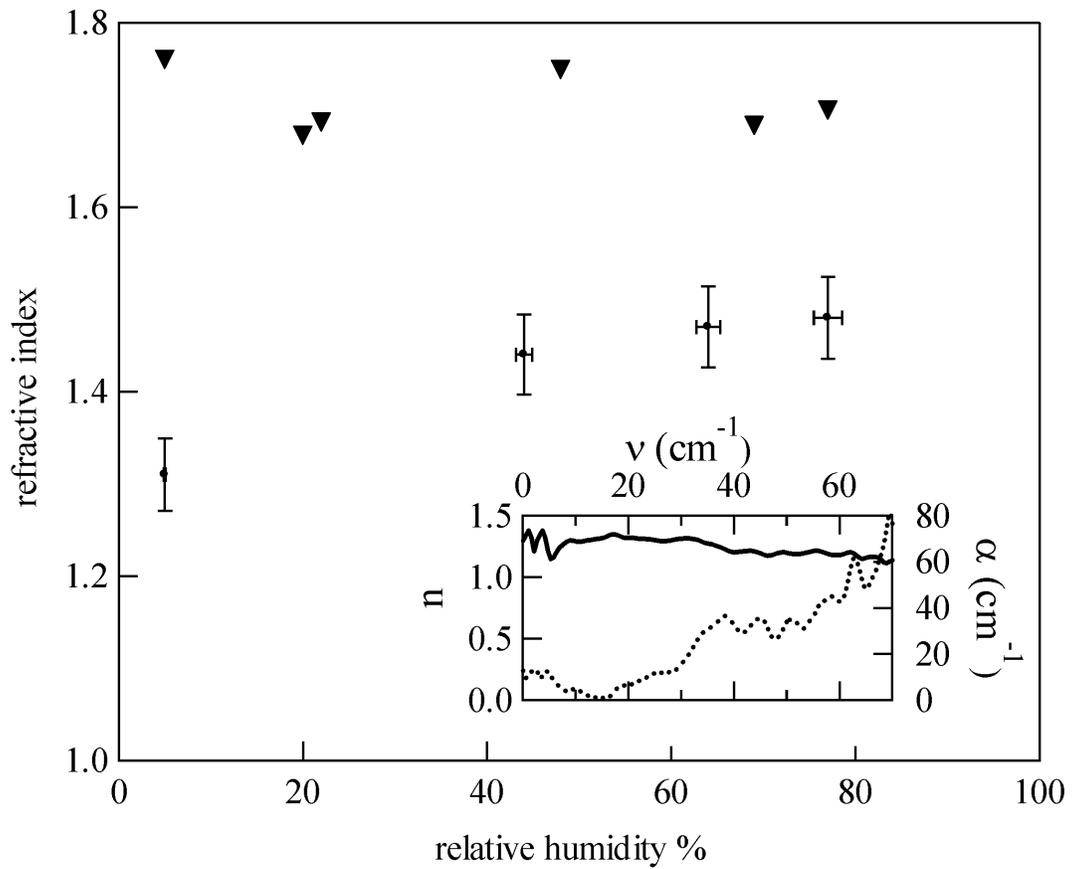
Figure 4. Real part of the refractive index as a function of r.h. at $\nu = 25$ cm^{-1}. Density of the DNA (●) and BSA (▼) samples are 0.48 gm/cm^3 and 1.04 gm/cm^3, respectively. Inset: the frequency dependent real part of the index, (n, solid line) and the absorption coefficient (a, dashed line) for DNA #3 shown in Fig. 2(a) and 3(a), derived from the PTS data using equations (2 a) and (2 b).