Terahertz Time-domain Spectroscopy as a Novel Tool for Crystallographic Analysis in Cellulose: Cellulose I to Cellulose Ii, Tracing the Structural Changes Under Chemical Treatment

Han Wang
Nagoya University: Nagoya Daigaku https://orcid.org/0000-0001-5071-2020
Hiroki Kataoka
Nagoya University: Nagoya Daigaku
Satoru Tsuchikawa
Nagoya University: Nagoya Daigaku
Tetsuya Inagaki (inatetsu@agr.nagoya-u.ac.jp)
Nagoya University: Nagoya Daigaku

Research Article

Keywords: THz-TDS, XRD, Cellulose I, Cellulose II, Crystallinity

Posted Date: September 23rd, 2021

DOI: https://doi.org/10.21203/rs.3.rs-864353/v1

Version of Record: A version of this preprint was published at Cellulose on March 12th, 2022. See the published version at https://doi.org/10.1007/s10570-022-04493-x.

Terahertz time-domain spectroscopy as a novel tool for crystallographic analysis in cellulose: cellulose I to cellulose II, tracing the structural changes under chemical treatment

Han Wang†, Hiroki Kataoka†, Satoru Tsuchikawa†, Tetsuya Inagaki†*

†Graduate School of Bioagricultural Sciences, Nagoya University, Nagoya, Aichi, 464-8601, Japan.

*E-mail: inatetsu@agr.nagoya-u.ac.jp

ABSTRACT

Terahertz time-domain spectroscopy (THz-TDS) has expanded possibilities in cellulose crystallography research, as THz radiation detects most intermolecular vibrations and responds to the phonons of crystalline lattices. In this study, we traced the transformation of the cellulose crystalline lattice from cellulose I to cellulose II by THz-TDS and X-ray powder diffraction. Cellulose II was obtained by treating cellulose I with NaOH of different concentrations (0 wt%–20 wt%, at 2 wt% intervals). The THz absorption coefficient spectra of cellulose II showed three characteristic peaks (at 1.32 THz, 1.76 THz, and 2.77 THz). The THz absorption coefficient spectra of cellulose II treated with 20-wt% NaOH and cellulose I without NaOH treatment were fitted by a seventh-order Fourier series. Thus, the THz absorption coefficient spectra of samples treated with NaOH of other concentrations could be considered a combination of these two fitted profiles of cellulose I and cellulose II, multiplied by different coefficients. Furthermore, the coefficients could reflect the relative contents of cellulose I and cellulose II in the samples.

Keywords: THz-TDS, XRD, Cellulose I, Cellulose II, Crystallinity

Declarations.

Funding: JSPS KAKENHI Grant Number 21H02255 and 16H02559

Conflicts of interest/Competing interests: None

Availability of data and material: None

Code availability: Matlab 2018b

Compliance with Ethical Standards: This study following Compliance with Ethical Standards; this study does not involve human participants, animals, and potential conflicts of interest.

Authors' contributions: Han Wang and Tetsuya Inagaki conceived and designed the experiments. Han Wang and Hiroki Kataoka performed the experiments and analyzed the data. Han Wang wrote the paper, Tetsuya Inagaki and Satoru Tsuchikawa gave final approval of the manuscript.

Introduction

Cellulose, a polymer material where β (1 → 4)-D-glucose units aggregate into a highly ordered, chain-like structure, widely exists in nature (Updegraff 1969). The hydroxyl groups in each glucose unit can further interact to form intra- and intermolecular hydrogen bonds. This hydrogen bonding stabilizes the cellulose structure and makes it an important material that can support the structure of plants, algae, and some bacteria. The hydrogen bonds in natural cellulose can be modified by some physical and chemical treatments to change their connection. Typically, naturally occurring cellulose I can be converted into cellulose II by mercerization (treatment with aqueous NaOH) or regeneration. A previous study reported that the crystalline lattice of cellulose I will change into that of cellulose II (Zugenmaier 2001) by treating the former with aqueous NaOH (about 10 wt%), and such transformation is irreversible (Kroon-Batenburg and Kroon 1997). Cellulose II has the most stable crystalline structure among all cellulose polymorphs. The chains in cellulose II are arranged antiparallelly (Kolpak and Blackwell 1976). By contrast, the cell unit of cellulose II is the same as that of cellulose I_{β}, which is a relatively stable monoclinic unit (Kolpak and Blackwell 1976; Debzi et al. 1991). The concertation of NaOH and the type of cellulose used are key factors to be considered in attempting a complete transformation of cellulose.

Tracing the transformation of the crystalline structure is essential for a better understanding of the processes of industry and biosynthesis where cellulose is involved, such as viscose rayon manufacturing (O'sullivan 1997; Brown 2004). Several techniques have been used to investigate the transformation of cellulose I to cellulose II. Given that the crystalline structure of cellulose can be detected by X-ray, the most widely applied technique is X-ray diffraction (XRD); Fourier-transform infrared spectroscopy (FTIR), Raman spectroscopy, nuclear magnetic resonance (NMR), and near-infrared spectroscopy (NIR) have also been used in such investigation (Langan et al. 2001, 2005; Schenzel and Fischer 2001; Dinand et al. 2002; Oh et al. 2005; Schenzel et al. 2009; Kafle et al. 2014). After the NaOH treatment of cellulose, both the crystalline lattice structure and the crystallinity change. FTIR and NMR are used to observe the crystallinity of cellulose and elucidate the hydrogen bonding that changes during the crystalline lattice

transformation. Halonen et al. reported on the use of solid-state cross polarization/magic angle spinning carbon 13 nuclear magnetic resonance (CP/MAS 13C NMR) spectroscopy (Halonen et al. 2013). Oh et al. observed that multiple bands of the FTIR spectra of cellulose I shift after NaOH treatment, and the absorbance ratios correlate with the crystallinity obtained from XRD patterns (Oh et al. 2005).

Terahertz (THz) radiation lies in the frequency of 0.1–10 THz, which corresponds to the wavelengths of 3–0.03 mm (which lie at the interval between microwaves and infrared radiation). Rapid progress has been made in THz technology in the past two decades, and one of the earliest commercial applications was THz time-domain spectroscopy (THz-TDS). THz-TDS has been used in many fields because THz radiation detects the vibrations of many biomolecules and hydrogen bonds and directly responds to the phonons in crystal lattices. These fields include pharmaceutical polymorphs (Strachan et al. 2005; Pickwell and Wallace 2006; Zeitler et al. 2007; Xie et al. 2014); quantitative characterization of paper (Reid and Fedosejevs 2006; Trafela et al. 2013; Peccianti et al. 2017); and detection of the vibrational modes of water isotopes, DNA, and protein (Plusquellic et al. 2007; Markelz 2008; Born et al. 2009; Yada et al. 2009). Some case studies have confirmed the possibility of using THz-TDS for cellulose research, such as studies determining the crystallinity of cellulose I (Vieira and Pasquini 2014; Wang et al. 2021) and distinguishing cellulose I allomorphs (Wang et al. 2020).

As a follow-up to a previous study (Wang et al. 2020, 2021), the current work further explores the possibility of using THz-TDS in the research field of cellulose crystallography. We traced the transformation of the crystalline lattice of cellulose I to that of cellulose II by using THz-TDS. The cellulose II sample was obtained by treating cellulose I with NaOH of different concentrations. Treatment of cellulose I with NaOH of concentrations exceeding 10% changed the crystalline lattice structure into that of cellulose II, as confirmed by other methods (such as XRD); the transformation was also reflected in the THz absorption coefficient spectra. The THz absorption coefficient spectra of cellulose I showed two absorption peaks; the absorption peak at 3.04 THz was shared by the allomorphs of I_α and I_β, and the absorption peaks at 2.13 THz and 2.38 THz were characteristic absorption peaks of I_β and I_α, respectively (Wang et al. 2020, 2021).

The THz absorption coefficient spectra of cellulose II showed a different profile of cellulose I, where the absorption peaks were at 1.32 THz, 1.76 THz, and 2.77 THz; the peak intensity of cellulose II was relatively small compared with that of cellulose I. The THz absorption coefficient spectra of cellulose II and cellulose I were fitted by a seventh-order Fourier series. The THz absorption coefficient spectra of the samples treated with NaOH of other concentrations could be considered a combination of the profiles of cellulose I and cellulose II; only the multiplied coefficients were different, and these coefficients could reflect the change in the relative content of the two components. Thus, the THz absorption coefficient spectra can be used to investigate the transformation of the crystalline lattice during the chemical treatment and the relative content of cellulose I and cellulose II of the mixed cellulose samples.

Experimental

Sample preparation

For observing the transformation from cellulose I to cellulose II, microcrystalline cellulose (MCC) powders (cellulose I, EMD Millipore 1.02331.0500) were treated with NaOH of different concentrations (0 wt%, 2 wt%, 4 wt%, 6 wt%, 8 wt%, 10 wt%, 12 wt%, 14 wt%, 16 wt%, 18 wt%, and 20 wt%) for 30 min at room temperature. After the reaction, the sample was washed with acetic acid and distilled water up to pH 7 and filtered. Then, the washed cellulose powders were dried in a desiccator containing P_2O_5. The NaOH, acetic acid, and phosphorus pentoxide used here were all purchased from KISHIDA CHEMICAL Co., Ltd. In this study, MCC samples without NaOH treatment were considered as cellulose I standard samples, while MCC samples treated with 20% NaOH were considered as standard cellulose II samples.

The above powders (treated with NaOH of different concentrations) were collected at a mass of 0.075 g by an electronic balance (± 0.0001 g). All the powders were compressed into tablets with a diameter of 14 mm and a thickness of approximately 0.35 mm using a compact heating press (IMC-180C, Imoto Machinery Co., Ltd.). Three tablets were prepared for each powder to ensure

reproducibility of the experiment. The thickness of these tablet samples was measured using a micrometer (± 0.001 mm).

XRD and THz-TDS measurement

XRD measurement of all the tablet samples was performed with Cu-Kα radiation (λ = 0.1542 nm) using an X-ray diffractometer (Ultima IV, Rigaku) at a voltage of 40 kV and a current of 40 mA. Diffractograms were recorded from 5° to 40°. The scan speed was set to 5° min^{-1}, and the sampling step was 0.05°. The background diffractogram was obtained from an empty sample holder.

The THz transmission spectra of all the tablet samples were measured by using a Tera Prospector (Nippo Precision Co., Ltd.), and the reference signals were obtained by measurement of air before and after sample measurement. The THz beam was horizontally polarized with a bandwidth of about 0.1 THz to 4.00 THz, and the spectral resolution was 0.02 THz, which corresponded to the inverse of the temporal scan range (50 ps). The diameter of the THz beam spot on the sample was around 3 mm. Each measurement was recorded by averaging 100 scans to improve the signal-to-noise ratio. For reproducibility, all measurements were conducted thrice. To avoid the influence of the THz absorption of water vapor on the measurement, we placed the whole THz optical system in an almost-closed acrylic box, which was filled with dry air until all the THz measurements were completed to ensure stability of humidity. All samples were placed in the box for 24 h before measurement to balance the ambient humidity.

Results and Discussion

XRD pattern and THz absorption spectrum analysis of cellulose treated with NaOH

XRD analysis: transformation of cellulose

The original XRD pattern was cut out with a scattering range of 10° to 30°, which included the main crystalline peaks of cellulose I and cellulose II. Before further peak deconvolution,

background subtraction and baseline correction were performed on all original XRD patterns. The background pattern was obtained as mentioned in the experimental section, and the baseline was fitted as a first-order polynomial after background subtraction. The processing is shown in Fig. 1 (a).

The XRD patterns after background subtraction and baseline correction could be considered composites of the cellulose I and cellulose II profiles, that is, patterns having the amorphous intensity curves and crystalline peaks of cellulose I and cellulose II, respectively. All fitting processes in peak deconvolution had to be performed under the same conditions. Hence, all patterns, even that of the cellulose I without NaOH treatment, were considered composites of the cellulose I and cellulose II profiles; that is, the patterns had both the amorphous and crystalline peaks of cellulose I and II, respectively. However, to ensure fitting accuracy, we allowed the intensity of the crystalline peaks to be zero during the curve-fitting process. Hence, for the cellulose I without NaOH treatment, the crystalline peaks of cellulose II would be almost zero during the fitting process. The three main crystalline peaks of cellulose I had Miller indices of ($1\bar{1}$ 0), (110), and (200); for cellulose II, the Miller indices were ($1\bar{1}0$), (110), and (020). In Fig. 1 (b), the crystalline peaks of cellulose I are denoted as peak1, peak2, and peak3, and those of cellulose II are denoted as peak4, peak5, and peak6.

Deconvolution was conducted on the six crystalline peaks and two amorphous intensity curves of all patterns via curve fitting using a pseudo-Voigt profile, which has been used in many other research cases. The profile is a linear combination of a Gaussian curve and a Lorentzian curve, as shown in the following equations:

$$I_G(2\theta) = I_{max} \cdot exp\left\{-4\,ln(2) \cdot \left(\frac{2\theta - 2\theta_{max}}{\beta}\right)^2\right\} \quad (1)$$

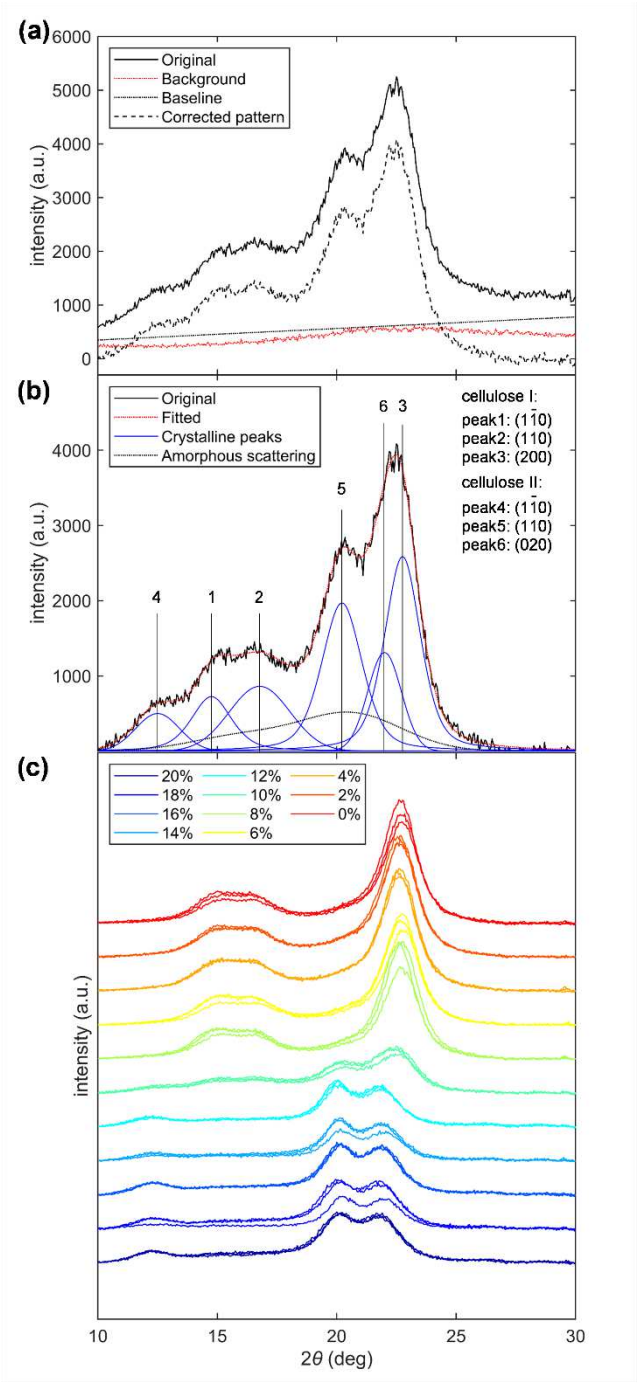
$$I_L(2\theta) = I_{max}\left\{1 + 4\left(\frac{2\theta - 2\theta_{max}}{\beta}\right)^2\right\}^{-1} \quad (2)$$

$$I_{PV}(2\theta) = \mu I_L + (\mu - 1) I_G \quad (3)$$

where $I_G(2\theta)$ and $I_L(2\theta)$ are the Gaussian and Lorentzian curves, respectively (Wada et al. 1997); I_{max} is the peak intensity; $2\theta_{max}$ is the peak position; and β is the full width at half maximum (FWHM).

The positions (2θ) of the crystalline peaks of the cellulose I profile were fixed in ranges of 15° ± 0.5°, 16.5° ± 0.5°, and 22.75° ± 0.25°, which corresponded to Miller indices of $(1\bar{1}0)$, (110), and (200), respectively. For the cellulose II profile, the crystalline peaks were fixed in ranges of 12° ± 0.5°, 20° ± 0.5°, and 21.5° ± 0.5°, which corresponded to Miller indices of $(1\bar{1}0)$, (110), and (020), respectively. The amorphous intensity curves of cellulose I and cellulose II were fixed at 20.6° and 16°, respectively (Oh et al. 2005; French 2014). The other parameters, namely, the FWHM, peak intensity, and coefficient μ, were all determined via curve fitting. Fig. 1 (b) shows the deconvolution process of the 10% NaOH–treated cellulose sample, where all the crystalline peaks and amorphous intensity curves of cellulose I and cellulose II were fitted as a pseudo-Voigt profile.

The baseline-corrected XRD patterns of the cellulose samples are arranged in Fig. 1 (c), with the color gradient expressing the different concentrations of NaOH during the treatment. The changes in the XRD patterns after the NaOH treatment can be easily identified; under treatment with 0% to 8% NaOH, typical cellulose I patterns can be observed. With the increase in NaOH concentration, the pattern treated with 10% NaOH showed a superimposition of the cellulose I and cellulose II patterns, and the samples treated with 12% to 20% NaOH showed patterns dominated by the cellulose II pattern.



152

153 Fig. 1 Curve-fitting process of a 10% NaOH–treated cellulose sample. (a) Background subtraction and

154 baseline correction, (b) deconvolution of peaks with a pseudo-Voigt profile (the amorphous scattering

155 in the figure is the sum of the amorphous intensity curves of cellulose I and cellulose II), (c)

156 baseline-corrected XRD patterns of all cellulose samples (the gradient colors express the different

157 concentrations of NaOH during the treatment).

158 *THz analysis: transformation of cellulose*

159 The measured THz time-domain signal was Fourier transformed into the frequency domain, and
160 the absorption coefficient α was calculated using the following equations:

161 $$n = -\frac{\varphi c}{2\pi v L} + 1 \quad (4)$$

162 $$\alpha = -\frac{2}{L}\ln\left[R\frac{(n+1)^2}{4n}\right] \quad (5)$$

163 where φ is the phase difference between the reference and measured samples ($\varphi_{reference} - \varphi_{sample}$), R
164 is the ratio of the amplitude in the frequency domain of the measured samples to the reference, c is
165 the speed of light (3×10^8 m/s), v is the frequency, and L is the thickness of the samples (Reid and
166 Fedosejevs 2006).

167 The original THz absorption coefficient spectra from 0.2 THz to 3.5 THz were all corrected for
168 baseline fluctuations with a standard normal variate (SNV) algorithm and then smoothed by the
169 application of a Savitzky–Golay filter with a second-order polynomial and 21 smoothing points to
170 remove the noise, as shown in Fig. 2 (a), where the used color gradient is the same as that in Fig.
171 1. The THz absorption coefficient spectra of all the cellulose samples (including treated
172 with/without NaOH) showed a similar change trend as that of the XRD patterns. The absorption
173 characteristics shown in the THz absorption coefficient spectra of cellulose I and cellulose II
174 totally differed from those of other methods, such as FTIR spectra (which show a peak shift after
175 NaOH treatment of cellulose (Oh et al. 2005)). At NaOH concentrations below 10%, the THz
176 absorption coefficient spectra showed the typical characteristics of cellulose I observed in
177 previous studies: the absorption peak at 2.13 THz corresponded to the cellulose I_β type (Wang et
178 al. 2020), and the absorption peak at 3.04 THz correlated with the amount of cellulose I,
179 regardless of the I_α and I_β allomorphs (Wang et al. 2021). For the 20% NaOH–treated cellulose
180 samples, which were almost cellulose II, the spectra showed different characteristics of cellulose I;
181 two absorption peaks (at 1.32 THz and 2.77 THz) can be observed. Given that the absorption
182 peaks in the THz region of cellulose II were smaller than those of cellulose I, the Savitzky–Golay
183 second derivative of the THz absorption coefficient spectra was examined, as shown in Fig. 2 (b),
184 to distinguish the absorption peaks clearly. The peaks at 1.32 THz, 1.76 THz, and 2.77 THz could

185 be easily observed, whereas the peak at 1.76 THz was almost invisible in the THz absorption
186 coefficient spectra. As in the XRD patterns, the cellulose treated with 10% NaOH showed the
187 characteristics of both cellulose I and cellulose II. The integral intensity of the absorption
188 coefficient peaks in the THz region can be used to evaluate the crystallinity of cellulose I_β (Wang
189 et al. 2021).

190 Cellulose I_β has a monoclinic unit cell with two chains packed in parallel. After the NaOH
191 treatment of cellulose I_β, the allomorph transformed into cellulose II. Similar to cellulose I_β,
192 cellulose II has a monoclinic unit cell, but two chains are arranged antiparallelly. The hydrogen
193 bonds are important for the cellulose crystalline structure, and the hydrogen bonds of cellulose II
194 differ from those of cellulose I. The intramolecular O3-H···O5 hydrogen bond exists in both
195 cellulose I and cellulose II polymorphs, whereas the intramolecular O2-H···O6 hydrogen bond
196 only occurs in cellulose I (Langan et al. 2001, 2005; Nishiyama et al. 2002). The intermolecular
197 hydrogen bonds of cellulose I and cellulose II are also different; O6-H···O3 intermolecular
198 hydrogen bonds occur in cellulose I, whereas O6-H···O2 intermolecular hydrogen bonds occur in
199 cellulose II, as shown in Fig. 3. Unlike cellulose I, which has no intersheet hydrogen bonds,
200 cellulose II contains O6—H…O6, O2—H…O2 intersheet hydrogen bonds, which make cellulose
201 II more stable than cellulose I (Langan et al. 1999, 2001).

202 The peak positions of the THz absorption coefficient spectra, 2θ values, and d-spacing values
203 calculated from the XRD patterns of different types of cellulose are summarized in Table 1. The
204 d-spacing values were calculated as follows:

205 $$d = \frac{\lambda}{2\,sin\,\theta} \tag{6}$$

206 where λ is the wavelength of the X-ray radiation (0.1542 nm) and θ is the Bragg angle
207 (Zsigmondy and Scherrer 1912). As shown in Table 1, the change in the crystalline structures was
208 reflected in the observed d-spacing values, the 2θ of the XRD patterns, and the peak positions of
209 the THz absorption coefficient spectra. Generally speaking, the higher the frequency, the greater
210 the photon energy. Cellulose I—whether I_β or I_α—showed a common absorption peak at 3.04
211 THz. The THz absorption peak at 2.13 THz of cellulose I_β was located at a lower frequency
212 compared with that of cellulose I_α at 2.38 THz, which may be caused by the more stable

crystalline structure of cellulose I_β than that of cellulose I_α. The THz absorption peaks of cellulose II were distributed at a lower frequency, which may be related to the more stable and lower overall energy of cellulose II (Langan et al. 1999, 2001).

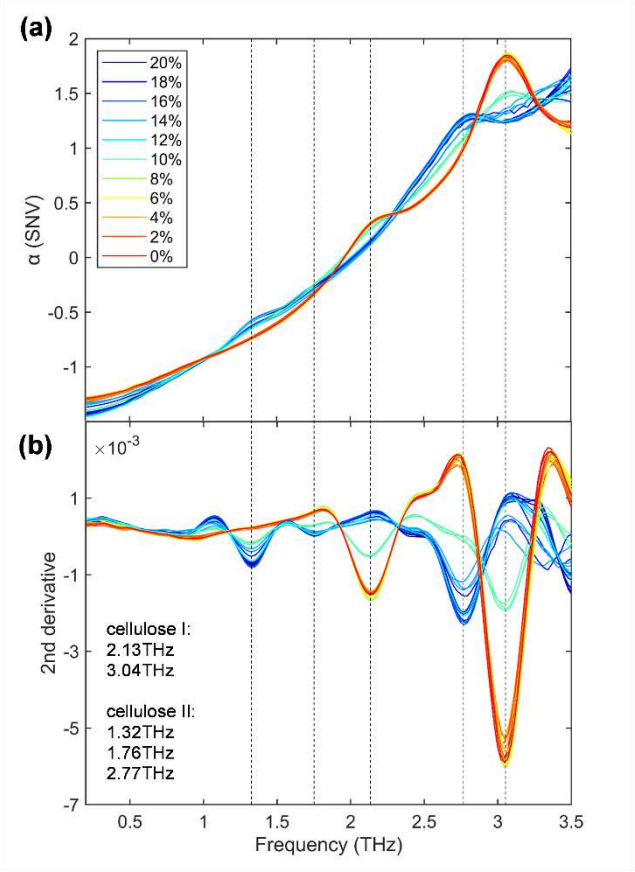
As we hypothesized in our previous study (Wang et al. 2020, 2021), THz absorption coefficient spectra are likely to detect various hydrogen bonds in cellulose, and a change in the hydrogen bonds in the cellulose crystalline structure will result in completely different absorption peaks, as shown in Fig. 2. The spectra of cellulose I had only two peaks (at 2.13 THz and 3.04 THz); for cellulose II, there were three absorption peaks (at 1.32 THz, 1.76 THz, and 2.77 THz). The number of absorption peaks might be related to the intersheet hydrogen bonds, which are the only bonds existing in cellulose II. However, the specific assignments of these peaks at the molecular level need further research.

Table 1 *2θ* and d-spacing values calculated from the XRD patterns and the peak positions of THz absorption coefficient spectra

	XRD pattern						THz absorption spectrum		
	2θ (°)			d-spacing (nm)			peak position (THz)		
	$(1\bar{1}0)$	(110)	(200)/(020)*	$(1\bar{1}0)$	(110)	(200)/(020)*	peak1	peak2	peak3
cellulose I_α (*Glaucocystis*)[a]	14.71	17.04	22.89	0.602	0.520	0.389	2.38	3.04	NaN
	14.58	16.93	22.78	0.607	0.524	0.390	2.38	3.04	NaN
	14.57	16.91	22.77	0.608	0.524	0.391	2.38	3.04	NaN
cellulose I_β (*Halocynthia*) [a]	14.89	16.7	22.98	0.595	0.531	0.387	2.13	3.04	NaN
	14.94	16.75	23.04	0.593	0.529	0.386	2.13	3.04	NaN
	15.03	16.84	23.1	0.589	0.527	0.385	2.13	3.04	NaN
cellulose I_β (without NaOH treated MCC)	14.50	17.00	22.50	0.611	0.522	0.395	2.13	3.04	NaN
	14.54	16.81	22.50	0.610	0.522	0.395	2.13	3.04	NaN
	14.54	16.99	22.50	0.611	0.522	0.395	2.13	3.04	NaN
cellulose II (20% NaOH treated MCC)	12.50	20.24	22.00	0.709	0.434	0.404	1.32	1.76	2.77
	12.50	20.45	21.09	0.708	0.440	0.404	1.32	1.76	2.77
	12.50	20.50	22.00	0.708	0.439	0.404	1.32	1.76	2.77

*The Miller indices for cellulose I and cellulose II were (200) and (020), respectively.

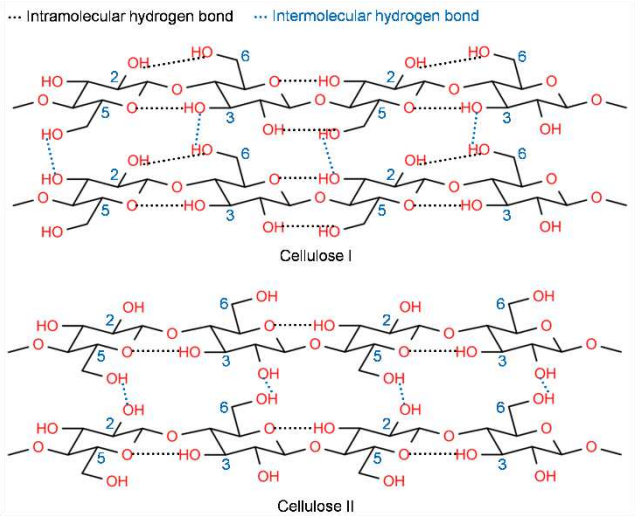
[a]The values were adopted from Wang et al. (2020).



228

229 Fig. 2 THz spectra of cellulose treated with NaOH of different concentrations. (a) SNV and smoothed

230 THz absorption coefficient spectra, (b) second derivative of the THz absorption coefficient spectra.



231

232 Fig. 3 Schematic representation of the intra- and intermolecular hydrogen bonds in cellulose I (top) and

233 cellulose II (bottom).

Crystallinity and relative content of cellulose (determined by XRD and THz)

The crystallinity index (CrI) calculated from the XRD patterns was determined by the following equation:

$$CrI = \frac{S_{Cr}}{S_{Cr} + S_{Am}} \tag{7}$$

where S_{Cr} and S_{Am} are the sums of the integrated intensity of six crystalline peaks and the amorphous intensity curves, respectively. The detailed process of the peak deconvolution with the pseudo-Voigt profile is described above.

CrI was divided into CrI1 (for cellulose I) and CrI2 (for cellulose II), which were calculated as follows:

$$CrI1 = \frac{S_{Cr1}}{S_{Cr1} + S_{Cr2}} \cdot CrI \tag{8}$$

$$CrI2 = \frac{S_{Cr2}}{S_{Cr1} + S_{Cr2}} \cdot CrI \tag{9}$$

where S_{Cr1} and S_{Cr2} are the sums of the integrated intensity of the crystalline peaks of cellulose I and cellulose II, respectively, and CrI is the crystallinity index calculated by Eq. (7).

CrI1 and CrI2 can be used to evaluate the relative content of cellulose I and cellulose II in the samples to grasp the progress of the NaOH treatment. The correlations of the concentrations of NaOH with CrI1 and CrI2 are shown in Figs. 5 (a) and (b), respectively. CrI1 and CrI2 showed completely opposite trends with the change in the NaOH concentration, where CrI1 decreased with the increase in the NaOH concentration, indicating a decrease in the cellulose I in the samples. In contrast, CrI2 increased with the NaOH concentration, which showed a transformation from cellulose I to cellulose II. Both Figs. 5 (a) and (b) showed inflection points at the 10% NaOH concentration, which indicated that the crystalline lattice transformed from cellulose I to cellulose II, as shown by the XRD patterns in Fig. 1 (c).

As shown in Fig. 2, the THz absorption coefficient spectra of the 10% NaOH–treated sample also showed mixed characteristics of cellulose I and cellulose II. Given the simplicity of the THz absorption coefficient and the above identification of the absorption peaks of cellulose I and cellulose II, the THz absorption coefficient spectra of the cellulose I without NaOH treatment and the cellulose II with 20% NaOH treatment can be written as two seventh-order Fourier series

(denoted as $f(0)$ and $f(20)$). Furthermore, the expression of the THz absorption coefficient spectra of the samples treated with NaOH of other concentrations can be written as $f = r1 \cdot f(0) + r2 \cdot f(20)$, where *r1* and *r2* are coefficients determined by the curve-fitting process, and *r1* + *r2* = 1, since the samples only have two types of cellulose crystalline. As shown in Fig. 4, the fitted results reproduced the measured THz absorption coefficient spectra well. The details of the fitting of $f(0)$ and $f(20)$ are provided in Supplementary Information.

The fitting results of the THz absorption coefficient spectra were evaluated with the coefficient of determination R^2:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(X_i - Y_i)^2}{\sum_{i=1}^{n}(X_i - \bar{X})^2} \quad (10)$$

where X_i, Yi, and $\bar{X}$ are the intensity of the THz absorption coefficient spectra after SNV and smoothing, the fitted THz absorption coefficient spectra, and the average of the THz absorption coefficient spectra, respectively. Table 2 summarizes the calculated R^2; the values were good enough, indicating credible fitting of the THz absorption coefficient spectra.

The correlations of the NaOH concentrations with the coefficients *r1* and *r2* are shown in Figs. 5 (c) and (d). These correlations were highly similar with those of CrI1 and CrI2, which were calculated from the XRD patterns. This indicated that the coefficients obtained from the fitting of the THz absorption coefficient can be used to evaluate the relative content of cellulose I and cellulose II. Furthermore, for the samples which only have two types of cellulose crystalline components, the coefficients *r1* and *r2* may give a more accurate relative content of each crystalline component, since the integrated intensity of the THz mass absorption coefficient peak was very possible to directly reflect the amount of the cellulose crystalline without the interference of the amorphous region on the THz absorption coefficient spectra (Wang et al. 2021).

Table 2 Calculated R^2 values of the fitted THz absorption coefficient spectra

	NaOH concentration										
	0%	2%	4%	6%	8%	10%	12%	14%	16%	18%	20%
R^2	1.0000	0.9994	0.9998	0.9978	0.9999	0.9994	0.9998	0.9996	0.9999	0.9998	1.0000
	1.0000	0.9999	0.9991	0.9974	0.9997	0.9984	0.9999	0.9999	1.0000	0.9999	0.9999

1.0000	0.9990	0.9994	0.9982	0.9998	0.9988	0.9997	0.9997	1.0000	0.9998	1.0000

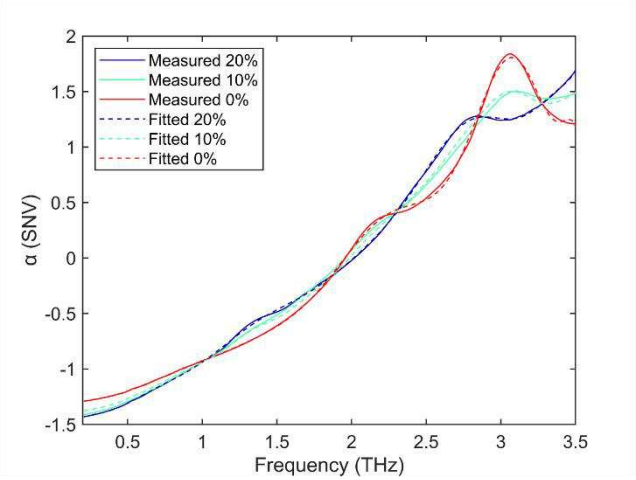


Fig. 4 THz absorption coefficient spectra of cellulose treated with 20% (blue), 10% (green), and 0% (red) NaOH that are measured and curve fitted using a Fourier series.

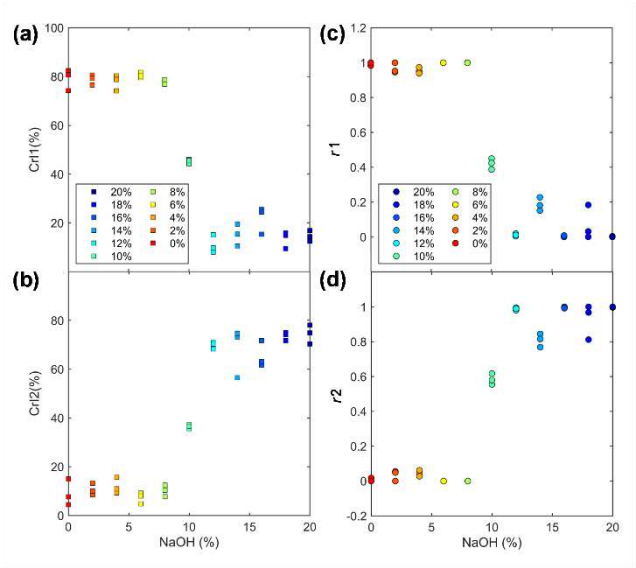


Fig. 5 Correlation of the concentrations of NaOH with (a) CrI1 and (b) CrI2 calculated from the XRD patterns, and coefficients (c) *r*1 and (d) *r*2 fitted from the THz absorption coefficient spectra.

Conclusions

In this study, cellulose I was treated with NaOH of different concentrations (from 0% to 20%) to change the crystalline structure. XRD patterns and THz absorption coefficient spectra of samples were obtained. The XRD patterns and the THz absorption coefficient spectra can be

295 divided into three types; when the concentration of NaOH during processing was lower than 10%,
296 equal to 10%, and higher than 10%, the XRD patterns and the THz absorption coefficient spectra
297 showed a typical profile of cellulose I, mixed characteristics of cellulose I and cellulose II, and a
298 typical profile of cellulose II, respectively. For the THz absorption coefficient spectra of cellulose
299 I, only two absorption peaks (at 2.13THz and 3.04 THz) could be observed. By contrast, three
300 absorption peaks (at 1.32 THz, 1.76 THz, and 2.77 THz) could be observed for cellulose II. The
301 XRD patterns were deconvoluted by using a pseudo-Voigt profile, and the calculated crystalline
302 index can be further divided into CrI1 and CrI2 to indicate the relative content of cellulose I and
303 cellulose II in the samples, respectively. CrI1 decreased as the concentration of NaOH increased,
304 whereas CrI2 increased with the NaOH concentration. The THz absorption coefficient spectra of
305 cellulose I without NaOH treatment and cellulose II treated with 20% NaOH were fitted by the
306 seventh-order Fourier series ($f(0)$ and $f(20)$). After this process, all the THz absorption
307 coefficient spectra could be fitted as a mixture of these two formulae by multiplying two
308 coefficients ($r1$ and $r2$), which meant that $f = r1 \cdot f(0) + r2 \cdot f(20)$. The correlations
309 between the concentrations of NaOH and $r1$ and $r2$ showed similar changing trends of CrI1
310 and CrI2. Remarkably, the transformation of the crystalline lattice from cellulose I to cellulose II
311 after the NaOH treatment can be traced by THz absorption coefficient spectra the coefficients $r1$
312 and $r2$ are able to describe the relative content of cellulose I and cellulose II in a sample. The
313 THz signal can be measured at room temperature and does not require sample pretreatment, the
314 measurement and analysis processes are rapid and simple compared with the XRD patterns.
315 Combined with the results of previous research, THz we believe THz-TDS has the potential to
316 become a vital tool in the research and understanding of cellulose crystallography.

317 **Acknowledgments** This work was supported by JSPS KAKENHI Grant Number 21H02255 and
318 16H02559. The XRD measurement was supported by Radioisotope Research Center, Nagoya
319 University. And the first author also thanks for the Mitsutani scholarship for living support.

Reference

Born B, Weingärtner H, Bründermann E, Havenith M (2009) Solvation dynamics of model peptides probed by terahertz spectroscopy. observation of the onset of collective network motions. J Am Chem Soc 131:3752–3755.

Brown RM (2004) Cellulose Structure and Biosynthesis: What is in Store for the 21st Century? J Polym Sci Part A Polym Chem 42:487–495.

Debzi EM, Chanzy H, Sugiyama J, et al (1991) The $I_\alpha \rightarrow I_\beta$ Transformation of Highly Crystalline Cellulose by Annealing in Various Mediums. Macromolecules 24:6816–6822.

Dinand E, Vignon M, Chanzy H, Heux L (2002) Mercerization of primary wall cellulose and its implication for the conversion of cellulose I → cellulose II. Cellulose 9:7–18.

French AD (2014) Idealized powder diffraction patterns for cellulose polymorphs. Cellulose 21:885–896.

Halonen H, Larsson PT, Iversen T (2013) Mercerized cellulose biocomposites: A study of influence of mercerization on cellulose supramolecular structure, water retention value and tensile properties. Cellulose 20:57–65.

Kafle K, Greeson K, Lee C, Kim SH (2014) Cellulose polymorphs and physical properties of cotton fabrics processed with commercial textile mills for mercerization and liquid ammonia treatments. Text Res J 84:1692–1699.

Kolpak FJ, Blackwell J (1976) Determination of the Structure of Cellulose II. Macromolecules 9:273–278.

Kroon-Batenburg LMJ, Kroon J (1997) The crystal and molecular structures of cellulose I and II. Glycoconj J 14:677–690.

Langan P, Nishiyama Y, Chanzy H (2001) X-ray structure of mercerized cellulose II at 1 Å resolution. Biomacromolecules 2:410–416.

Langan P, Nishiyama Y, Chanzy H (1999) A revised structure and hydrogen-bonding system in cellulose II from a neutron fiber diffraction analysis. J Am Chem Soc 121:9940–9946.

Langan P, Sukumar N, Nishiyama Y, Chanzy H (2005) Synchrotron X-ray structures of cellulose Iβ and regenerated cellulose II at ambient temperature and 100 K. Cellulose 12:551–562.

348 Markelz AG (2008) Terahertz dielectric sensitivity to biomolecular structure and function. IEEE J Sel
349 Top Quantum Electron 14:180–190.
350 Nishiyama Y, Langan P, Chanzy H (2002) Crystal structure and hydrogen-bonding system in cellulose
351 I_β from synchrotron X-ray and neutron fiber diffraction. J Am Chem Soc 124:9074–9082.
352 O'sullvian AC (1997) Cellulose: the structure slowly unravels. Cellulose 4:173–207
353 Oh SY, Dong IY, Shin Y, et al (2005) Crystalline structure analysis of cellulose treated with sodium
354 hydroxide and carbon dioxide by means of X-ray diffraction and FTIR spectroscopy. Carbohydr Res
355 340:2376–2391.
356 Peccianti M, Fastampa R, Mosca Conte A, et al (2017) Terahertz Absorption by Cellulose: Application
357 to Ancient Paper Artifacts. Phys Rev Appl 7:1–7.
358 Pickwell E, Wallace VP (2006) Biomedical applications of terahertz technology. J Phys D Appl Phys
359 39:.
360 Plusquellic DF, Siegrist K, Heilweil EJ, Esenturk O (2007) Applications of terahertz spectroscopy in
361 biosystems. ChemPhysChem 8:2412–2431.
362 Reid M, Fedosejevs R (2006) Terahertz birefringence and attenuation properties of wood and paper.
363 Appl Opt 45:2766–2772.
364 Schenzel K, Almlöf H, Germgård U (2009) Quantitative analysis of the transformation process of
365 cellulose I → cellulose II using NIR FT Raman spectroscopy and chemometric methods. Cellulose
366 16:407–415.
367 Schenzel K, Fischer S (2001) NIR FT Raman spectroscopy - A rapid analytical tool for detecting the
368 transformation of cellulose polymorphs. Cellulose 8:49–57.
369 Strachan CJ, Taday PF, Newnham DA, et al (2005) Using terahertz pulsed spectroscopy to quantify
370 pharmaceutical polymorphism and crystallinity. J Pharm Sci 94:837–846.
371 Trafela T, Mizuno M, Fukunaga K, Strlič M (2013) Quantitative characterisation of historic paper
372 using THz spectroscopy and multivariate data analysis. Appl Phys A Mater Sci Process 111:83–90.
373 Updegraff DM (1969) Semimicro determination of cellulose inbiological materials. Anal Biochem
374 32:420–424.

375 Vieira FS, Pasquini C (2014) Determination of cellulose crystallinity by terahertz-time domain
376 spectroscopy. Anal Chem 86:3780–3786.
377 Wada M, Okano T, Sugiyama J (1997) Synchrotron-radiated X-ray and neutron diffraction study of
378 native cellulose. Cellulose 4:221–232.
379 Wang H, Horikawa Y, Tsuchikawa S, Inagaki T (2020) Terahertz time-domain spectroscopy as a novel
380 tool for crystallographic analysis in cellulose. Cellulose 27:9767–9777.
381 Wang H, Tsuchikawa S, Inagaki T (2021) Terahertz time-domain spectroscopy as a novel tool for
382 crystallographic analysis in cellulose: the potentiality of being a new standard for evaluating
383 crystallinity. Cellulose 28:5293–5304.
384 Xie L, Yao Y, Ying Y (2014) The application of terahertz spectroscopy to protein detection: A review.
385 Appl Spectrosc Rev 49:448–461.
386 Yada H, Nagai M, Tanaka K (2009) The intermolecular stretching vibration mode in water isotopes
387 investigated with broadband terahertz time-domain spectroscopy. Chem Phys Lett 473:279–283.
388 Zeitler JA, Kogermann K, Rantanen J, et al (2007) Drug hydrate systems and dehydration processes
389 studied by terahertz pulsed spectroscopy. Int J Pharm 334:78–84.
390 Zsigmondy R, Scherrer P (1912) Bestimmung der inneren Struktur und der Größe von Kolloidteilchen
391 mittels Röntgenstrahlen. Kolloidchem Ein Lehrb 277:387–409.
392 Zugenmaier P (2001) Conformation and packing of various crystalline cellulose fibers. Prog Polym Sci
393 26:1341–1417.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SPinfo.docx