Determination of Cellulose Crystallinity by Terahertz-Time Domain Spectroscopy

ARTICLE *in* ANALYTICAL CHEMISTRY · MARCH 2014

Impact Factor: 5.64 · DOI: 10.1021/ac4035746 · Source: PubMed

2 AUTHORS:

Francisco Senna Vieira
NIST Boulder Laboratories
4 PUBLICATIONS **20** CITATIONS
SEE PROFILE

Celio Pasquini
University of Campinas
90 PUBLICATIONS **1,518** CITATIONS
SEE PROFILE

analytical chemistry

Article

pubs.acs.org/ac

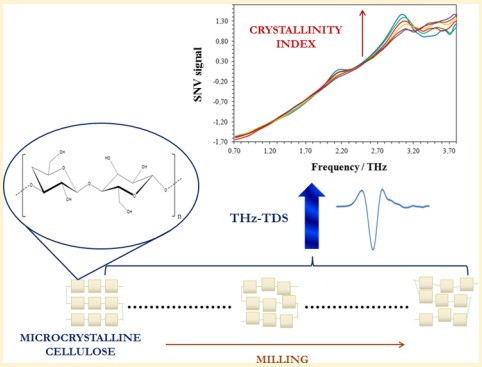
Determination of Cellulose Crystallinity by Terahertz-Time Domain Spectroscopy

Francisco Senna Vieira* and Celio Pasquini

Chemistry Institute, State University of Campinas−UNICAMP, Cidade Universitária, Campinas, São Paulo, 13083-970, Brazil

Supporting Information

ABSTRACT: Terahertz-time domain spectroscopy (THz-TDS) has the ability to probe the crystallinity of several materials, due to the interaction of THz radiation with optical phonons in crystal lattices. In this work, THz-TDS has been used to quantify the degree of crystallinity of microcrystalline cellulose (MCC) samples. The THz spectra of cellulose present absorption features which could be directly correlated with the crystallinity index (CI) obtained by means of the well-established powder X-ray diffraction (PXRD) technique. The effect of THz time-domain signal processing was investigated, and both univariate and multivariate, based on partial least-squares (PLS), regressions were carried out with the signal in the frequency domain to correlate the THz spectra with CI. Results show that the multivariate regression models based on spectral data, collected with the sample displaced from the focal plane of the THz optics to improve representativeness and measurement repeatability, present the best performance with external validation achieving an absolute root-mean-square error of prediction (RMSEP) of 4% for CI. This result compares well with the PXRD technique.



Microcrystalline cellulose (MCC) plays an important role as an excipient in pharmaceutical formulations.[1−6] Along with other parameters, the crystallinity index (CI) is essential to ensure that the tablets are stable after preparation. In 1999, Suzuki and Nakagami demonstrated that the compression energy is inversely proportional to the degree of crystallinity.[3] In the same paper, the authors emphasized the dramatic influence the crystallinity of cellulose has on the dissolution kinetics of acetaminophen tablets: for high crystalline cellulose, the solubilization rate decreases with crystallinity, whereas the inverse happens for samples of low crystallinity. Water swelling, among other properties, is also affected by the CI of cellulose.

The degree of crystallinity of cellulose is also important in other fields. For instance, it has been demonstrated that crystallinity plays a major role in the enzymatic hydrolysis of cellulose to produce glucose with implications in second generation bioethanol production.[7]

There are several techniques used for determining the crystallinity index (CI) of cellulose. The most employed technique is powder X-ray diffraction (PXRD). With this technique, at least three different methods to calculate CI are described.[5,8] The most common, and also the simplest one, is the peak subtraction method. Basically, this method consists of subtracting the intensity of a peak corresponding to one of the crystalline forms of cellulose, I_C, usually the [002] peak at 22°, from the amorphous halo intensity, I_{AM}, usually measured around 18°, and dividing the result by I_C. Other methods include deconvolution and amorphous halo fitting. Surprisingly, each of these methods yield significantly different results, even though they are derived from essentially the same data, the PXRD diffractogram.[8] Nevertheless, PXRD remains the most reliable technique for determining the crystallinity of cellulose, given the fact that it provides information directly related to crystal structure, and alternative methods, such as differential scanning calorimetry, would deteriorate cellulose due to the required elevated temperatures.

Solid-state ^{13}C nuclear magnetic resonance (NMR) is also used for evaluating the CI of cellulose.[1,8] With this technique, the CI is measured as the ratio between the areas of the C4 peaks corresponding to amorphous and crystalline chemical environments. Diffuse reflectance infrared Fourier transform spectroscopy (DRIFT) has also been used to determine the CI of cellulose.[9] In the deconvoluted DRIFT spectra of cellulose, a series of bands in the region between 1500 and 900 cm^{-1} dramatically change as the crystallinity of cellulose varies, due to differences in molecular environment. Near infrared Fourier transform Raman spectroscopy determines the CI of cellulose by calculating the ratio between two peaks and calibrating with a reference method.[10] All of the above techniques, however, are unable to measure a signal directly related to crystal structure but rather detect the influences of the chemical environments in NMR chemical shifts or infrared frequencies/bandwidths/intensities. Thus, in order to be truly reliable, the method should allow one to probe a signal directly, and preferentially exclusively, related to crystallinity.

Terahertz-time domain spectroscopy (THz-TDS) has the ability to probe the crystallinity of several materials, due to the

Received: November 5, 2013
Accepted: March 22, 2014
Published: March 23, 2014

   dx.doi.org/10.1021/ac4035746 | *Anal. Chem.* 2014, 86, 3780−3786

interaction of THz radiation with optical phonons in crystal lattices.[11−13] Crystalline materials present well-defined absorption peaks, whereas amorphous media absorb THz radiation in a featureless way. Given this characteristic, THz spectroscopy has been used to differentiate amorphous from crystalline states, as well as distinguish between several crystalline phases, of a large variety of compounds.[12,14−17] For instance, the crystallinity degree of lactose has been investigated in a real time crystallization process by acquiring THz spectra while hydrating the amorphous saccharide.[18] Walther et al. investigated the spectra of fructose, sucrose, and glucose, in both crystalline and amorphous states, at temperatures ranging from 10 to 300 K. They concluded that every observed feature in the spectra were due to long-range vibrations, since the loss of symmetry in the amorphous state of the sugars resulted in featureless spectra.[19] Those and many other examples reiterate the idea that most THz resonances in solids are essentially due to interactions between optical phonons and the incident radiation.

In that sense, THz-TDS arises as an alternative to determine directly such a property. As previously mentioned, THz radiation is responsible for long-range periodic vibrations in crystals. Thus, it is expected that absorption bands directly related to the degree of crystallinity of the substrate are observed. In this paper, the feasibility of using THz-TDS as a tool for determining the CI of microcrystalline cellulose was explored. Different treatments of the time-domain THz signal were evaluated, and both univariate and multivariate regression methods were tested employing PXRD peak subtraction as the reference method. It must be emphasized, however, that every crystallinity measurement in cellulose samples is unable to quantify this property in an absolute way. Given the fact that cellulose is a natural polymer, it is not possible to fully control its crystallinity, as it would with a synthesizable organic compound. Therefore, even the CI calculated from the reference method does not represent the actual crystalline/amorphous proportion within the probed sample.

EXPERIMENTAL PROCEDURES

Ball-Milling of the Cellulose. In order to prepare samples with different levels of crystallinity, the MCC samples were first dried at 40 °C for 7 days, and then submitted to ball-milling in a planetary ball mill (Bench-Top Planetary Automatic Ball Mill-MSK-SFM-1, MTI Corporation) at 400 rpm, using Al_2O_3 jars and ZrO_2 spheres. MCC samples used were Comprecel M101, Avicel 101, and Avicel 102. Several ball-millings sequences were carried out over a period of 7 months, always employing 6 g of dried MCC. Three different milling procedures were performed. In a first screening process, cellulose was ball milled for 4 h, with 20 min intervals of milling, followed by 10 min of pause to avoid overheating of cellulose. During this process, samples were collected every hour. The remaining two procedures were performed up to a maximum time of 100 min, being this time range found adequate for quantitative analysis. A detailed description of these procedures is shown in Table 1. A total set of 19 cellulose samples with distinct CI values was employed for calibration (univariate and PLS) and method validation. Both calibration and validation sets were comprised of samples prepared on different days.

Recrystallization of the Cellulose. Cellulose samples milled for 4 h were recrystallized by adding water, a method which has been previously reported in the literature with various procedures.[20,21] In this work, the procedure found to better restore MCC to its original crystallinity degree was simply to mix the amorphous cellulose with water overnight and then redry them at 40 °C. First, three samples of the 4 h milled cellulose were mixed with different amounts of water (30, 40, and 50% m/m) and put to rest for 12 h. After that period, this sample set was dried at 40 °C for an additional 24 h.

Table 1. Description of the Milling Processes and Sample Collection

batch	milling interval/min	pause/min	cellulose samples collection time/min	total milling time/min
1	20	10	60	240
2	10	10	20	100
3[a]	15	10	25	95

[a]The last milling step was reduced to 10 min.

Terahertz Measurements. Terahertz spectra were acquired by employing a high-speed asynchronous optical sampling (ASOPS) THz spectrometer (HASSP-THz, Laser Quantum). In this system, a 10 W 532 nm laser is used to pump two Ti:sapphire oscillators, both yielding 1.0 W in mode-lock operation, with a 1 GHz repetition rate. The asynchronous optical sampling is achieved by maintaining a 2 kHz difference between the repetition rates of the oscillators. A small fraction of the laser beam is deflected to two photodiodes, one for each Ti:sapphire oscillator, producing a signal which provides the repetition rate feedback required for maintaining the established offset. The repetition rate control is made by applying high voltage to piezoelectric crystals installed in two mirrors in one of the Ti:sapphire oscillators (slave oscillator), always taking the repetition rate of the second oscillator as reference (master oscillator). A GaAs interdigitated semiconductor antenna is used for generating the THz pulse, and a GaP crystal is employed for electro-optical detection. The THz measurements are started by an optical trigger system, which consists in deflecting part of the beams generated by both the oscillators into a two photon detector. When the phase difference between the lasers is zero, the trigger is activated. With this setup, a pulse containing frequencies from 0.3 to 6.0 THz is generated. A simplified scheme of the THz-HASSP setup is presented in Figure 1, and a more detailed description of the ASOPS principle and the THz-HASSP spectrometer can be found elsewhere.[22,23]

Cellulose samples were measured as Teflon pellets (thickness = 0.9 mm, diameter = 12.3 mm), yielding a usable spectral window ranging from 0.3 to 4.5 THz. Prior to spectra acquisition, the sample compartment of the spectrometer was purged with N_2 until water vapor absorption lines were indistinguishable from noise (nearly 20 min). THz spectra were recorded by averaging 500 scans, yielding approximately 4 min per sample (the samples are exchanged without requiring a new N_2 purge). Cellulose concentration was set to 15.0% (m/m), and duplicate pellets for each sample were prepared. Calibration and validation sets were constructed from samples processed and measured over a period of 7 months.

The full resolution of the equipment, which corresponds to a fast Fourier transform (FFT) of a time window ranging from 30 to 300 ps, is 3 GHz. The pellets were measured both at the THz focus, which comprises an area of 0.5 mm, and out of focus, so that the THz pulse could probe a larger area of the pellet, thus minimizing heterogeneity effects. In the latter, the sample was displaced toward the focusing mirror until a circular

dx.doi.org/10.1021/ac4035746 | Anal. Chem. 2014, 86, 3780−3786

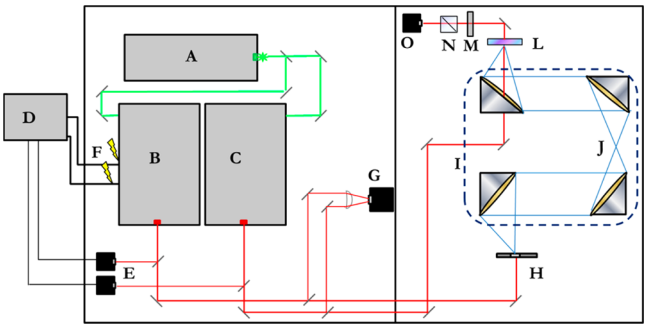


Figure 1. Simplified scheme of the THz-TD HASSP spectrometer. (A) pump laser (532 nm); (B) Ti:Sapphire oscillator (800 nm, slave, pump); (C) Ti:Sapphire oscillator (800 nm, master, probe); (D) repetition rate control unit; (E) photodiode detector for repetition rate feedback; (F) piezos high voltage input; (G) optical trigger two photon detector; (H) GaAs photoconductive antenna; (I) Au toroidal mirrors; (J) THz focus; (L) GaP electro-optical crystal; (M) half wave plate; (N) polarizing beam splitter; (O) photodiode detector.

area of approximately 5 mm diameter was probed by the THz radiation (~1.7 cm away from the focus).

PXRD Measurements. A XRD6000 Shimadzu X-ray diffractometer was employed for acquiring the PXRD patterns. Diffractograms were recorded from 5° to 50°, with a scan speed equal to 2° min^{-1}. CI values were obtained by the peak subtraction method, using the intensity of the peak at 22.64° (I_C) as the crystalline signal, and the 18.22° signal (I_{AM}) as the amorphous halo. CI was calculated according to eq 1:

$$\mathrm{CI} = \frac{I_C - I_{AM}}{I_C} \times 100 \quad (1)$$

Although this peak is attributed only to the cellulose I form, it is clear that this polymorph is predominant in the used MCC samples, given the lower intensities of the remaining peaks corresponding to other crystalline states.

Particle Size Distribution Measurements. Particle size distributions were determined for both milled and recrystallized cellulose samples, as well as the original MCC, by using a particle size analyzer (Mastersizer 2000, Malvern Instruments Corporate). Cellulose samples were added to the flow cell of the instrument until a pre-established threshold for the light scattering was reached. Measurements were made under 1750 rpm and 30% ultrasound intensity. A few drops of diluted commercial surfactant were also added in each sample prior to measurement, in order to prevent particle aggregation and coagulation.

SIGNAL PREPROCESSING

Several preprocessing treatments were carried out, in order to evaluate the best way to extract useful information from the acquired THz spectra. To remove the etalon effect found in the original spectra,[24] two approaches were followed: (1) truncating the FFT window to 46 ps, corresponding to a nominal resolution of 62 GHz, which severely reduces the spectral nominal resolution; (2) truncating the FFT window to 200 ps, corresponding to a nominal resolution of 6 GHz, followed by a FFT low pass filter to reduce noise. The lower limit for every FFT window was set to 30 ps, since no signal is present before this elapsed time. Henceforth, these approaches will be referred to as FFT-46 and FFT-200-filter, respectively. The number of frequency values describing the absorption spectra of cellulose (spectral variables) resulting from each process is 44 and 460, respectively, for FFT-46 and FFT-200-filter.

Both spectra in the frequency domain were further treated with standard normal variate algorithm (SNV) to correct baseline fluctuations.[25] Attempts were made to correlate the processed THz spectra with CI determined by using PXRD data both with univariate regression and partial least-squares regression (PLS). FFT transforms of the original spectra were carried out using the Matlab 6.5 software; SNV and PLS were performed using the Unscrambler 10.2 and the FFT filter using Origin 8.1.

PRXD patterns presented a considerable noise, as expected for measurements of highly amorphous materials. Thus, a Savitsky-Golay smoothing algorithm (31 points) was performed prior to CI calculation.

RESULTS AND DISCUSSION

Comparison between Terahertz Spectra and PXRD Patterns. Figure 2A shows THz spectra of cellulose samples with different milling times. The original MCC sample presents two characteristic peaks at 2.15 and 3.03 THz. However, even in this high crystalline cellulose, a featureless absorption which increases with frequency can be observed. This is an evidence of an elevated amount of amorphous material. This behavior is expected, since the CI of MCC samples usually ranges from 60 to 90%.[8] Also, it must be reminded that this CI does not necessarily correspond to the actual proportion of crystalline cellulose within the sample, but rather to a relative measurement of the amount of ordered polymer. Clearly, even the most crystalline cellulose, the original MCC, exhibits a PXRD pattern quite different than one obtained from a purely crystalline material. In that sense, THz spectra are coherent with the PXRD patterns.

As the crystallinity of cellulose decreases due to the milling process, both peaks become less intense, and the amorphous absorption seems to predominate. For milling times longer than 120 min, however, THz spectra become rather insensitive to changes in the degree of crystallinity of cellulose, unlike the PXRD patterns, in which a slight crystalline aspect is still distinguishable from the amorphous halo up to a milling time of 180 min. This lack of sensitivity of the THz technique may be due to the fact that experiments were carried out at room temperature (295 K), a fact that has been observed for several crystalline systems studied by THz spectroscopy.[26−28] At this temperature, the population of the excited state corresponding to the transition with frequency equal to 3 THz is already elevated, reducing the probability of transition and, consequently, the peak intensity. Thus, in order to be able to quantify very low crystallinity degrees in cellulose samples, the measurements should be made, in the future, at lower temperatures.

Even though room temperature spectra do not present sufficient resolution and peak intensity to compare with theoretical data, the fact that both peaks fade in the amorphous state leads to the conclusion that such features are related to long-range intermolecular vibrations. Thus, in principle, both peaks can be used to quantify the degree of crystallinity in cellulose samples. Although relevant to the present study and other studies carried out in crystalline materials investigated by THz spectroscopy, it is not the purpose of this paper to attribute those resonances to their respective vibrational modes/crystalline phases. Instead, in the present work, the

dx.doi.org/10.1021/ac4035746 | Anal. Chem. 2014, 86, 3780−3786

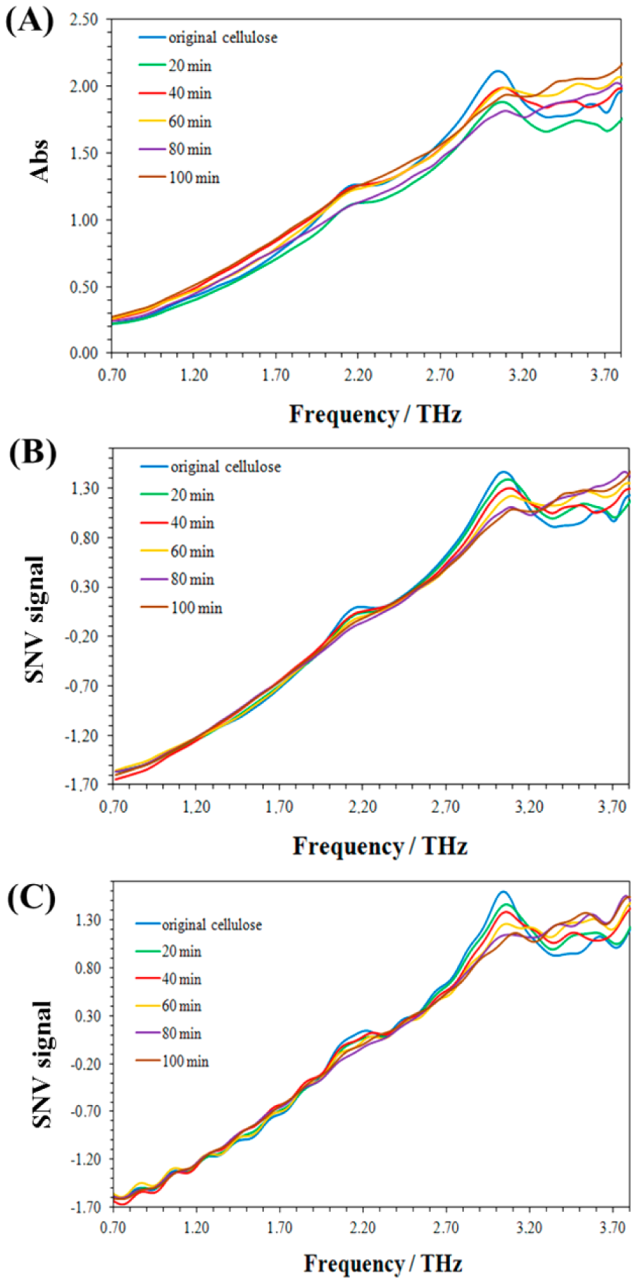


Figure 2. Spectra of milled cellulose obtained by measuring the pellets displaced from the THz focus: (A) FFT-46 original spectra, (B) FFT-46 spectra treated with SNV, and (C) FFT-200-filter spectra treated with SNV.

spectral features are empirically demonstrated to be related to crystallinity, by comparison with PXRD data, and therefore used to quantify this property.

PXRD patterns exhibited a similar behavior, exactly as expected (Figure 3). The original MCC pattern presents peaks typical of a substrate which is mainly composed of cellulose type I, with a slight contribution of cellulose II.[29] As the cellulose is progressively milled, all peaks tend to diminish and the PXRD pattern acquires the aspect of an amorphous material. Such observations are coherent with the behavior of the THz spectra.

Recrystallization Assays and Particle Size Evaluation. Considering that an absolute attribution of the absorption features was not carried out in this work, one should be aware

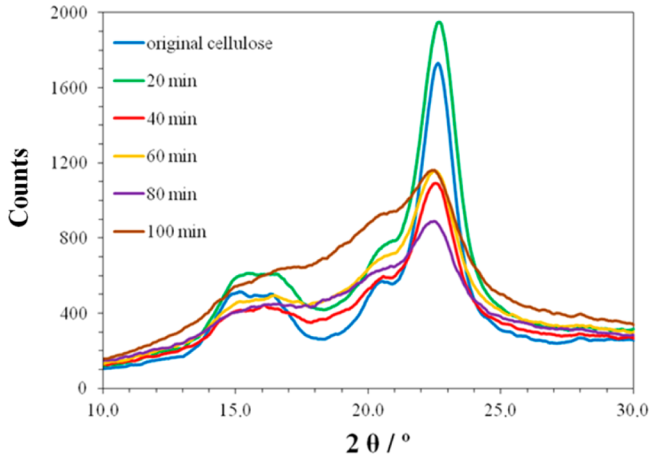


Figure 3. PXRD patterns of MCC samples with different milling times.

of possible noncausal correlation between peak intensities and crystallinity degree. Since samples of different crystallinity degree were obtained by milling the originally high crystalline cellulose, it is pertinent to verify whether or not the particle size could be spuriously correlated with the changes observed in the spectral absorbance features observed in the THz region. Therefore, samples with different crystallinity degrees were assayed for their particle size distribution and revealed the absence of correlation between particle average size and size distribution and the crystallinity degree determined by PXRD. Table 2 shows the average values for the particle size for some selected samples of milled cellulose.

Table 2. Comparison between Particle Size and Degree of Crystallinity Obtained[a]

milling time/min	CI (PXRD)	CI (THz)	average particle size/μm
0	84.72	86.91	115
20	78.39	76.44	70
40	65.78	67.03	80
100	39.66	41.33	76

[a]THz crystallinities found by univariate regression of absorbances at 3.03 THz are also presented. R^2 obtained for the regression of particle size on CI (PXRD) = 0.2559.

In order to reinforce the hypothesis that the features observed in the THz spectra of cellulose are related exclusively to its crystallinity and not influenced by uncontrolled physical parameters, such as particle size distribution, recrystallized cellulose samples were also analyzed. THz spectra of these samples, along with their respective particle size distributions, are presented as Supporting Information. From those data, it becomes clear that THz spectra, in the same way that PXRD patterns, are quite insensitive to different particle size distributions. Even when samples present drastic heterogeneity in that parameter, such as bimodal distributions, the spectra remain virtually the same.

Quantitatively, the observations do not change: the difference obtained for either the 2.15 or the 3.03 THz peak intensities between the recrystallized samples is within the standard deviation obtained between replicates of the original MCC sample. Such results are also coherent with PXRD data: CI calculated for the recrystallized samples and for three replicates of the original MCC are found to be equal to 80.9 ± 1.8 and 82.2 ± 1.2, respectively. Thus, in spite of the high

dx.doi.org/10.1021/ac4035746 | Anal. Chem. 2014, 86, 3780−3786

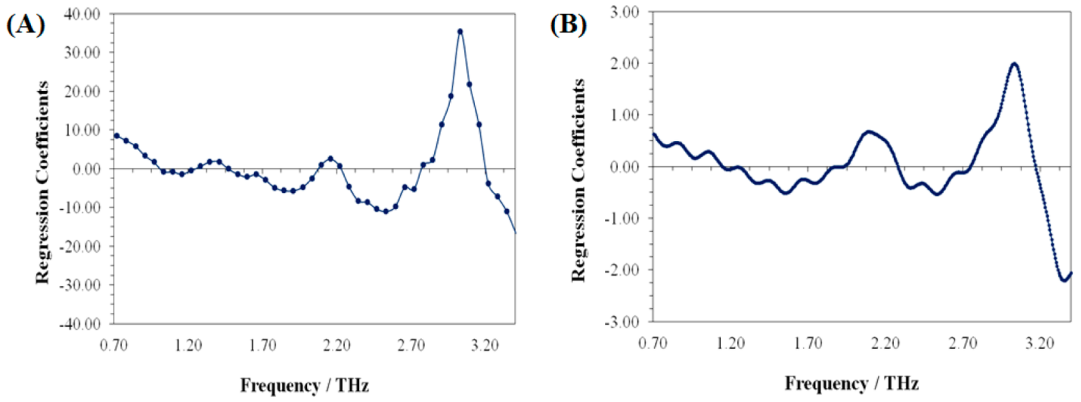


Figure 4. Regression coefficients for PLS models calculated with spectral data set obtained by (A) FFT-46 and (B) FFT-200-filter. Spectral range: 0.7−3.4 THz.

background interference found in every THz spectra, it is clear that relevant information related to the crystallinity of cellulose can be easily obtained from them.

Univariate Regression. In spite of the clear effect of crystallinity on the THz spectra, the CI cannot be directly extracted from the original data. This is due to the high variation in the spectra baselines, caused mainly by pellet heterogeneity and possibly light scattering. In an attempt to correct this effect, spectra were submitted to a SNV preprocessing. The intensities of the 3.03 THz peak, after this treatment, are directly related either to the milling time or the CI calculated from PXRD data. Correlation plots between these quantities with the FFT-46 and the FFT-200-filter THz spectra, obtained by employing 15 samples with the CI values in the range 25−85%, yielded determination coefficients equal to 0.9250 and 0.9158, respectively. No significant difference was observed between the two spectral treatments, mainly due to the large width of the analyzed band.

Attempts were made to calibrate the 2.15 THz peak with CI, but the THz intensities yielded poor correlations with cellulose CI. This is due to the very low intensity of such band, which results in a correspondingly low sensitivity. Even though the photometric signal-to-noise ratio of the THz spectrophotometer at 3.03 THz is slightly lower than at 2.15 THz ($SNR_{3.03}$ = 0.84 $SNR_{2.15}$), the 2.15 THz absorption band of cellulose is nearly 4 times less intense than the 3.03 THz. Thus, such low sensitivity is expected. It must be pointed out, however, that the method presented herein could eventually be optimized for working with the 2.15 THz peak by using more concentrated pellets. This alternative approach could be useful when only narrow band THz spectrometers are available.

Since the THz focus spot of the spectrometer is only 0.5 mm wide, it is reasonable to consider that pellet homogeneity plays an important role in the quality of the spectra and on its representativeness of the sample composition. In fact, by performing several scans in the same pellet, but in different spots, different spectra were obtained. This observation lead to a simple procedure to minimize the effect of pellet heterogeneity: measure the same samples again, but displaced from the focus toward the focusing mirror of the THz spectrometer, in a way that a larger circular area (5 mm diameter) of the pellet is probed by the THz pulse. Resulting spectra from this approach are shown in Figure 2B,C, for both pretreatments employed. The correlations between this new set of spectra and the CI calculated from PXRD are considerably more significant, confirming the hypothesis that the pellet homogeneity must be taken into account whenever the quantitative determination of crystallinity is aimed. Once again, both FFT-46 and FFT-200-filter spectra of 15 samples were tested, and the same observations previously reported apply to this case (determination coefficients obtained were equal to 0.9651 and 0.9628, for the truncated and FFT filtered spectra, respectively). Only one disadvantage could be noticed regarding the out-of-focus measurements: the etalon effect became stronger. Since this effect is related to the angle in which the radiation reaches the sample, the incidence of a conical beam probably generates a more complex etalon containing different frequency contributions. Thus, the low pass FFT filter employed was unable to fully remove this source of noise (Figure 2C). This higher etalon effect, however, did not disturb the 3.03 THz band and did not deteriorate the quantitative results. They were actually improved, as reported above for the univariate regression.

Partial Least Squares (PLS) Regression. Even though crystallinity can be directly related to the 3.03 THz absorption peak, partial least-squares (PLS) regression models were also evaluated, in an attempt to improve the obtained correlations and better predict the CI of unknown samples. This approach is justified by the fact that not only the 3.03 THz peak contribute to CI determination but also the spectral region beyond this high frequency peak. Since amorphous cellulose absorbs THz radiation continuously at higher frequencies, the remaining part of the spectrum can also be used to extract information correlated with the degree of crystallinity of cellulose. Thus, prior to using a specific PLS model, different spectral ranges were tested, with both FFT-46 and FFT-200-filter approaches. The low frequency limit was set to 0.7 THz for all measurements, since no spectral features were observed below this point, and the upper edge was tested from 3.3 to 4.0 THz. PLS models were calculated using 15 samples by the NIPALS algorithm, and an internal full cross validation was employed. For every model, the first factor explained over 90% of the variance, showing a high correlation among the spectral variables.

The best results were obtained with the upper limit established at 3.4 THz, for both FFT-46 and FFT-200-filter spectra. Models built with higher frequencies yielded worse and unstable results, particularly for the FFT-46 spectra, which contain less variables and therefore are more sensitive to slight variations in the spectral range. This behavior is justified by the fact that additional variables present beyond 3.4 THz are not sufficiently relevant to compensate for the reduced signal-to-

dx.doi.org/10.1021/ac4035746 | Anal. Chem. 2014, 86, 3780−3786

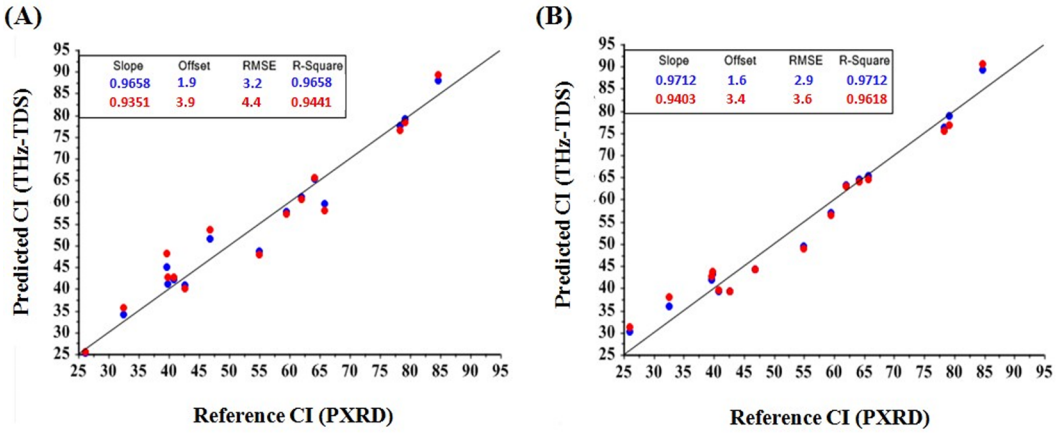


Figure 5. Plots of the CI values obtained by PXRD versus the CI values predicted by THZ-TD PLS models calculated in the spectral range from 0.7 to 3.4 THz: (A) FFT-46 and (B) FFT-200-filter spectra. Blue dots: calibration results. Red dots: full cross validation results.

noise ratio observed in that region. Thus, only models built up to 3.4 THz will be further discussed.

In Figure 4, the regression coefficients for both FFT-46 and FFT-200-filter treatments are shown. It is clear that, regardless of the employed pretreatment, the same qualitative behavior is observed; the 3.03 THz band is the spectral region that most contributes to the calibration. It can also be noticed that the multivariate regression coefficients values for the FFT-200-filter spectra present a periodic wave-like profile throughout the used THz frequency range. This behavior suggests that the coefficients of the regression were adjusted to compensate for the etalon effect superimposed to the main spectral features employed for CI prediction.

Figure 5 shows plots of the CI reference values obtained by PXRD versus the predicted values by the PLS models based on THz-TD, with their respective values of RMSECV and determination coefficients (R^2) calculated for the spectral range between 0.7 and 3.4 THz for the best models. It can be noticed that the determination coefficients are actually quite similar to those obtained in the univariate analysis. Also, the fact that a single factor in the majority of the models explains most of the variance corroborates the hypothesis that all information related to crystallinity is being extracted from the spectra in a straight way. Thus, it is possible that multivariate models are unnecessary for determining CI in cellulose with THz spectroscopy.

External Validation. A third set comprised of 4 MCC samples was used in order to validate both univariate and PLS calibrations. CI values for this sample set covered nearly the whole evaluated range (34, 50, 55, and 76%). RMSEP values for each calibration are listed in Table 3. Once again, it can be noticed no significant difference when PLS models are employed to predict the CI of cellulose. When compared to the aggregated variance of the PXRD method, it becomes clear that no significant difference in precision can be observed between the methods (Table 4).

Table 3. External Validation Errors of Prediction (RMSEP) Obtained for the Main PLS and Univariate Models Tested

model	RMSEP	no. factors	R^2
PLS/FFT-46	3.8	3	0.9441
PLS/FFT-200-filter	4.1	1	0.9618
univariate/FFT-46[a]	3.9		0.9651
univariate/FFT-200-filter[a]	4.0		0.9628

[a]Results obtained employing only the values of the peak absorbance at 3.03 THz.

Table 4. Precision Estimates for the Evaluated Models[a]

model	aggregated standard deviation (%)	aggregated variance $(\%)^2$	*F*-test result
univariate FFT-46	3.4	11.6	1.22
univariate FFT-200-filter	3.4	11.9	1.18
PLS FFT-46	5.0	25.0	1.78
PLS FFT-200-filter	4.0	16.0	1.90

[a]Degrees of freedom = 4; aggregated variance of the PXRD method = 14.07; $F_{\text{critical } 4/4}\ (0.05) = 6.4$.

It must be reminded that calibration and validation sets were constructed from samples of different origins, which were milled and measured over a period of 7 months. The fact that such time lapse and sample variability did not impart a high prediction error in the analysis is an evidence of the robustness of the method. Additionally, the fact that no gain was observed for the PLS calibration further confirms the straightforward relationship between crystallinity and the investigated spectral region.

CONCLUSION

In this paper, it has been demonstrated the potential of the THz-TDS as a new technique for quantifying the CI of cellulose employing a fast procedure carried out at ambient temperature. Results obtained with both univariate and PLS calibrations were comparable with PXRD and presented some remarkable advantages, such as dismissing the use of dangerous X-ray radiation and shortening the analysis time, 4 min per sample, instead of 22 min. Probably, the results can be improved by obtaining the cellulose spectra at lower temperatures, and some additional experiments will be carried out in the near future to demonstrate this hypothesis.

A careful evaluation of the sample measurement procedure is also relevant to the quality of the models. The best results were obtained by probing the sample pellet during spectrum acquisition over a broader area, improving the representativeness of its THz spectrum. Data pretreatments are fundamental to minimize variability sources not correlated with CI, and the best results were obtained by employing SNV followed by

dx.doi.org/10.1021/ac4035746 | Anal. Chem. 2014, 86, 3780−3786

either univariate calibration, or PLS, a multivariate regression technique. Univariate results did not present a considerable difference in performance when compared to PLS models, indicating that, for CI determination, multivariate models are possibly unnecessary.

Another important feature of the observed THz spectra of cellulose that needs to be addressed is the 2.15 THz peak. Even though in this work such band provided poor results, it is quite possible that the same method could eventually be optimized for using that peak. Further assays must be carried out to discover the proper conditions for a quantitative analysis, but the qualitative observations made in the current work indicate that it may eventually be possible.

THz spectra are, in a way, analogous to XRD patterns. Even though fundamentally different phenomena are involved, both techniques are able to directly measure signals associated with the crystalline structure of the substrate. Therefore, in a near future, it is expected that THz spectroscopy may enable the determination of crystallinity degree in a way that precludes comparison with a reference method.

■ ASSOCIATED CONTENT

Supporting Information

Additional information as noted in text. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: franciscosennavieira@gmail.com. Phone: +55 (19) 3521-3037.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors are grateful to FAPESP (Proc. No. 2011/13777-8); Mr. Mário Shissun Toma and Renata Magueta, members of the staff of Chemistry Institute of UNICAMP, respectively, for manufacturing several accessories for the THz spectrometer and for obtaining the PXRD diffractograms; and Prof. Celso Aparecido Bertran for kindly providing the ball mill and particle size distribution instruments. F.S.V. is grateful to FAPESP for the Ph.D. fellowship (Proc. No. 2013/00502-6). This is a contribution of the INCTAA (FAPESP, Proc. No. 2008/57808-1 and CNPq Proc. No. 573894/2008-6).

■ REFERENCES

(1) Ek, R.; Wormald, P.; Ostelius, J.; Iversen, T.; Nyström, C. *Int. J. Pharm.* **1995**, *125*, 257−264.
(2) Buckton, G.; Yonemochi, E. *Eur. J. Pharm. Sci.* **2000**, *10*, 77−80.
(3) Suzuki, T.; Nakagami, H. *Eur. J. Pharm. Biopharm.* **1999**, *47*, 225−230.
(4) Edge, S.; Steele, D. F.; Chen, A.; Tobyn, M. J.; Staniforth, J. N. *Int. J. Pharm.* **2000**, *200*, 67−72.
(5) Rowe, R. C.; McKillop, A. G.; Bray, D. *Int. J. Pharm.* **1994**, *101*, 169−172.
(6) Kothari, S. H.; Kumar, V.; Banker, G. S. *Int. J. Pharm.* **2002**, *232*, 69−80.
(7) Fan, L. T.; Lee, Y.-H.; Beardmore, D. H. *Biotechnol. Bioeng.* **1980**, *22*, 177−199.
(8) Park, S.; Baker, J. O.; Himmel, M. E.; Parilla, P. A.; Johnson, D. K. *Biotechnol. Biofuels* **2010**, *3*, 10.
(9) Hulleman, S. H. D.; van Hazendonk, J. M.; van Dam, J. E. G. *Carbohydr. Res.* **1994**, *261*, 163−172.
(10) Agarwal, U. P.; Reiner, R. R.; Ralph, S. A. *J. Agric. Food Chem.* **2013**, *61*, 103−113.
(11) Mantsch, H. H.; Naumann, D. *J. Mol. Struct.* **2010**, *964*, 1−4.
(12) McIntosh, A. I.; Yang, B.; Goldup, S. M.; Watkinson, M.; Donnan, R. S. *Chem. Soc. Rev.* **2012**, *41*, 2072−2082.
(13) Baxter, J. B.; Guglietta, G. W. *Anal. Chem.* **2011**, *83*, 4342−4368.
(14) Strachan, C. J.; Taday, P. F.; Newnham, D. A.; Gordon, K. C.; Zeitler, J. A.; Pepper, M.; Rades, T. *J. Pharm. Sci.* **2005**, *94*, 837−846.
(15) Aaltonen, J.; Gordon, K. C.; Strachan, C. J.; Rades, T. *Int. J. Pharm.* **2008**, *364*, 159−169.
(16) Otsuka, M.; Nishizawa, J.; Shibata, J.; Ito, M. *J. Pharm. Sci.* **2010**, *99*, 4048−4053.
(17) Strachan, C. J.; Rades, T.; Newnham, D. A.; Gordon, K. C.; Pepper, M.; Taday, P. F. *Chem. Phys. Lett.* **2004**, *390*, 20−24.
(18) McIntosh, A. I.; Yang, B.; Goldup, S. M.; Watkinson, M.; Donnan, R. S. *Chem. Phys. Lett.* **2013**, *558*, 104−108.
(19) Walther, M.; Fischer, B. M.; Uhd Jepsen, P. *Chem. Phys.* **2003**, *288*, 261−268.
(20) Ago, M.; Endo, T.; Hirotsu, T. *Cellulose* **2004**, *11*, 163−167.
(21) Kocherbitov, V.; Ulvenlund, S.; Kober, M.; Jarring, K.; Arnebrant, T. *J. Phys. Chem. B* **2008**, *112*, 3728−3734.
(22) Bartels, A.; Thoma, A.; Janke, C.; Dekorsy, T.; Dreyhaupt, A.; Winnerl, S.; Helm, M. *Opt. Express* **2006**, *14*, 430−437.
(23) Bartels, A.; Cerna, R.; Kistner, C.; Thoma, A.; Hudert, F.; Janke, C.; Dekorsy, T. *Rev. Sci. Instrum.* **2007**, *78*, 035107.
(24) Namkung, H.; Kim, J.; Chung, H.; Arnold, M. A. *Anal. Chem.* **2013**, *85*, 3674−3681.
(25) Barnes, R. J.; Dhanoa, M. S.; Lister, S. J. *Appl. Spectrosc.* **1989**, *43*, 772−777.
(26) Smirnova, I. N.; Sapozhnikov, D. A.; Kargovsky, A. V.; Volodin, V. A.; Cherkasova, O. P.; Bocquet, R.; Shkurinov, A. P. *Vib. Spectrosc.* **2012**, *62*, 238−247.
(27) Siegrist, K.; Bucher, C. R.; Mandelbaum, I.; Walker, A. R. H.; Balu, R.; Gregurick, S. K.; Plusquellic, D. F. *J. Am. Chem. Soc.* **2006**, *128*, 5764−5775.
(28) Harsha, S. S.; Grischkowsky, D. *J. Phys. Chem. A* **2010**, *114*, 3489−3494.
(29) Ishikawa, A.; Okano, T.; Sugiyama, J. *Polymer* **1997**, *38*, 463−468.