Terahertz absorption by cellulose: Application to ancient paper artifacts

Article (Accepted Version)

Terahertz absorption by cellulose: Application to ancient paper artifacts

M. Peccianti[1], R. Fastampa[2,3], A. Mosca Conte[4], O. Pulci[4], C. Violante[4],
J. Łojewska[5], M. Clerici[6], R. Morandotti[7,8,9], and M. Missori[3]*
[1]*Emergent Photonics Lab (EPic) - Dept. of Physics and Astronomy, University of Sussex, BN1 9QH Brighton UK*
[2]*Dipartimento di Fisica, Università di Roma "Sapienza", Roma, Italy*
[3]*Istituto dei Sistemi Complessi, Consiglio Nazionale delle Ricerche, UOS Sapienza, Roma, Italy*
[4]*ETSF, Dipartimento di Fisica, Universitá di Roma Tor Vergata, Rome, Italy*
[5]*Faculty of Chemistry, Jagiellonian University, Ingardena 3, 30-060 Krakow, Poland*
[6]*School of Engineering, University of Glasgow, G12 8LT, Glasgow, UK*
[7]*INRS-EMT, University of Quebec, Varennes, Quebec J3X 1S2, Canada*
[8]*Institute of Fundamental and Frontier Sciences,*
University of Electronic Science and Technology of China, Chengdu 610054, China and
[9]*National Research University of Information Technologies, Mechanics and Optics, St. Petersburg, Russia*

Artifacts made of cellulose, such as ancient documents, pose a significant experimental challenge in the THz transmission spectra interpretation due to their small optical thickness. In this Letter we describe a method to recover the complex refractive index of cellulose fibers from the THz transmission data obtained on single freely standing paper sheets in the 0.2–3.5 THz range. By using our technique, we were able to eliminate Fabry-Perot effects and recover the absorption coefficient of the cellulose fibers. The obtained THz absorption spectra are explained in terms of absorption peaks of the cellulose crystalline phase superimposed to a background contribution due to a disordered hydrogen bonds network. The comparison between the experimental spectra with THz vibrational properties simulated by density functional theory calculations confirms this interpretation. In addition, evident changes in the THz absorption spectra are produced by natural and artificial aging on paper samples, whose final stage is characterized by a spectral profile with only two peaks at about 2.1 THz and 3.1 THz. These results could be used to provide a quantitative assessment of the state of preservation of cellulose artifacts.

PACS numbers: 87.15.-v, 78.40.Me, 81.70.-q, 63.50.-x

I. INTRODUCTION

Experimental methods probing the physical and chemical structure of ancient artifacts in a non-destructive way are of utmost importance for the understanding and the diagnosis of their unavoidable degradation processes. In particular, an accurate and quantitative appraisal of the artworks' states of preservation poses a significant challenge for ancient artifacts made of delicate organic matter. A particular example are those made of cellulose, the most abundant biopolymer on Earth [10]. For centuries, cellulose has been widely exploited in several important areas of human activities, ranging from building wood based constructions, through the realization of writing media in paper and finally as a valuable chemical agent to produce derivatized materials [11].

Cellulose is composed of an unbranched homopolysaccharide made up of β-D-glucopyranose units $(C_6H_{10}O_5)_n$ connected by β-(1,4)-glycosidic bonds. Such units are prone to form chains with n ranging from $\sim 10^2$ to $\sim 10^4$. The reason for the strong tendency of cellulose polymers to organize themselves into multiple parallel arrangements to finally create a crystalline structure is due to the presence of hydroxyl groups forming interchain hydrogen bonds (H-bonds). A hierarchical structure is then formed, and made up of elemental fibrils (diameter 2–4nm), fibrils (diameter $\sim$200nm) and finally fibers, whose diameter ranges from 1 to 10 μm [21]. Elemental fibrils are composed of highly ordered (crystalline) domains (of typical length $\sim$100nm) and of disordered (amorphous–like) regions, the latter providing paper material with its remarkable mechanical properties.

The degradation of cellulose artifacts depends on the environmental conditions [19]. It mostly proceeds in the amorphous regions, which are the most accessible to chemical agents [12], and it is mediated by two interconnected chemical processes: hydrolysis of the β-(1,4)-glycosidic bonds and oxidation of the β-D-glucopyranose units [14, 15, 39]. These are accompanied by the rearrangement of the H–bond network, which in turn gives rise to recrystallization. Such process swallows up the amorphous regions where both reactions proceed. In fact, at a macroscopic scale, degradation appears as a weakening of the mechanical properties of cellulose artifacts [25].

THz time–domain spectroscopy (THz–TDS) is a suitable approach to study low–energy vibrational properties of biological materials in a non-destructive way. The THz photon energy range (about 1–40 meV) is particularly suitable to probe the H-bond between molecules that is of utmost importance in a biological system. This is the most widespread intra and intermolecular bonding responsible for the functioning and structure of biomolecules [22, 29]. Further, THz spectroscopy is an

* mauro.missori@isc.cnr.it

alternative method to X-ray diffraction to assess sample crystallinity [23]. Specifically, this is crucial for ancient paper artifacts where crystallinity strongly influences mechanical properties and durability [25].

A first attempt to characterize a historical paper using THz spectroscopy was made by using a partial least squares regression method applied to the transmission spectra, to estimate the chemical and mechanical properties of the samples [26]. Another study focused on the dependence of THz attenuation and refractive index spectra on paper density and water content. With the goal to reduce the inherent inaccuracy in the index reconstruction due to small optical delays in the time-domain THz signal (see Fig.1), the studied samples comprised several stacked paper sheets [6]. However, this approach prevents the characterization of unique ancient artifacts and may introduce unpredictable interference effects in the measured quantities. Finally, THz spectroscopy has also been used to evaluate the crystallinity of samples made of microcrystalline cellulose powders in the form of thick pellets as a function of the milling process time [28].

All these studies were considerably affected by inaccuracies connected with the extraction of the optical parameters from the THz signals obtained in both low refractive index, and thin paper samples. In this regime, substantial interference effects are superimposed to the transmission spectra thus hindering the reconstruction of the absorption and refractive index.

We present THz-TDS measurements performed on unaged, artificially aged and ancient paper samples. The recovered transmission spectra were analyzed to eliminate the interference Fabry-Perot (FP) effects and to obtain a precise determination of both cellulose THz absorption and refractive index. The obtained THz absorption spectra (normalized to account for the thickness and density of the samples) are explained in terms of absorption peaks of the cellulose crystalline phase superimposed to a background contribution due to a disordered hydrogen bonds network. The absorption peaks has been compared with predictions based on the vibrational energy levels obtained by density functional theory (DFT) *ab-initio* calculations, where cellulose has been modeled as a crystal.

II. EXPERIMENTAL DETAILS

Modern paper samples, obtained from the Netherlands Organization for Applied Scientific Research (TNO), are made of unbleached cotton linters, containing very low inorganic ingredients (ash content $<0.005\%$ in weight) and no additives or lignin [25]. They were artificially aged up to 47 days in closed vials (in air) at a relative humidity RH$=$ 59 % and a temperature T $=$ 90°C (see Supplemental Material for further details). Ancient samples were produced in various European countries, consisting of four specimens, made of linen and cotton cellulose fibers, bearing no print, named A1, B1 (both made in Perpignan, France, in 1413), A3 (made in Milan, Italy, in 1430), and N1 (made in Nuremberg, Germany, in the 16th century) presenting increasing levels of degradation [2, 14, 16, 19].

THz-TDS signals were acquired in transmission mode by using a Menlo Systems (Germany) TERA K15 THz-TDS equipped with photoconductive antennae. For all acquisition times, the scan range was set to 100ps and the data were collected with a time resolution of 30fs. In order to reduce statistical errors, different THz-TDS signals (typically 3 or 4) with and without samples were acquired. Since water vapor absorbs THz radiation, the sample compartment of the THz set-up was purged with N_2 until most water vapor absorption lines were indistinguishable from noise (nearly 60 min before the spectra acquisition). This procedure is also capable to remove free water from samples while it is expected that bound water remains in an amount of approximately 2% in mass [1, 3, 17]. THz-TDS signals were converted into their spectral representation by a fast Fourier transform (FFT). The usable range, where data are reliable, was found to be 0.2–3.5 THz. The frequency dependent dynamic range was about 80 dB at 0.35 THz and about 20 at 3.5 THz. Typical pulses obtained by the THz characterization set-up are shown in Fig.1 together with their FFT amplitude.

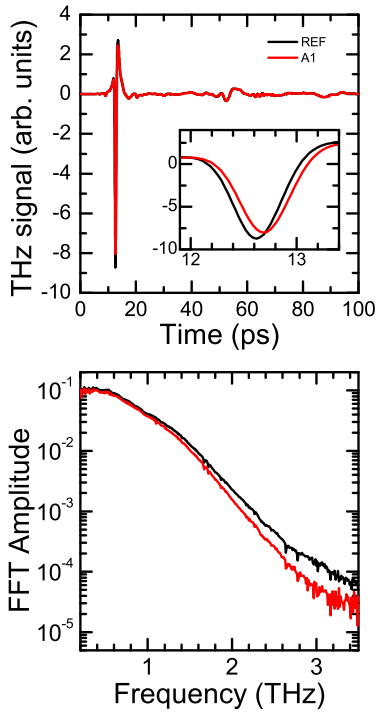

Figure 1. Traces of the reference and A1 sample pulses in the transmission THz-TDS experiment (upper panel), and the power spectral density of the reference and sample pulses (lower panel).

A paper sample consists of a random assembly of cellulose fibers and voids and their relative amount is sample dependent [1, 4]. In order to assess the volume fraction v of cellulose fibers in the samples and recover their absolute THz absorption coefficient, the density (ρ_{paper}) of each paper sample was measured. Then $v = V_c/V_{paper} = \rho_{paper}/\rho_c$, where V_c and V_{paper} are the volume occupied by cellulose fibers and the volume of the paper sheets used in this study, respectively, while $\rho_c = 1.5 g/cm^3$ is the average density of cellulose fibers [18] (see Supplemental Materials for details [7]).

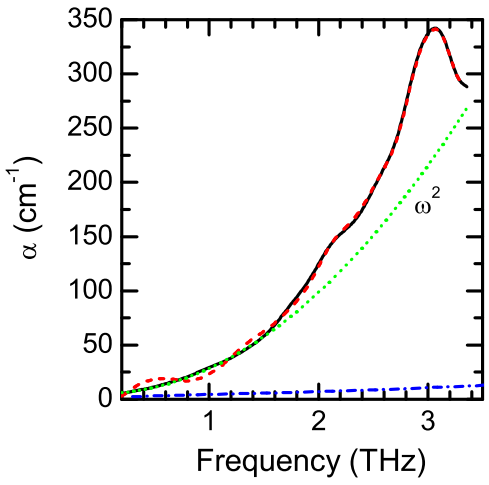

Figure 2. Absorption coefficient curves of cellulose fibers in sample A1 without (dashed red) and with (black) removal of the FP oscillations. A quadratic fit of the initial black curve (green dot line) and an estimation of the bound water (2% of sample's mass) spectral contribution [30] (blue dashed line) are also shown.

The frequency dependent behavior of the complex refraction index $\hat{n}(\omega) = n(\omega) + i\alpha(\omega)\, c/2\omega$ (c is the speed of light and $\omega/2\pi$ is the frequency) can be extracted from the THz-TDS transmission mode signals by using the experimental complex transfer function $\hat{T}_{exp}(\omega_j)$ (j is a data array index):

$$\hat{T}_{exp}(\omega_j) = \frac{\hat{E}_{sample}(\omega_j)}{\hat{E}_{ref}(\omega_j)} \qquad (1)$$

where $\hat{E}_{sample}(\omega)$ and $\hat{E}_{ref}(\omega)$ are the spectral amplitude and phase with and without the sample in the THz line, respectively [8]. The extraction of the optical parameters can be obtained by equating the theoretical and experimental transfer functions: $\hat{T}(\hat{n}, \omega_j) = \hat{T}_{exp}(\omega_j)$ in which the analitycal function $\hat{T}(\hat{n}, \omega)$ is obtained as a solution of the inverse scattering problem for the electromagnetic waves. Due to the coherence of our source, when THz radiation propagates within a sample, multiple internal reflections give rise to the well–known FP effect [8, 34, 35]. Consequentely, the resulting expression can not be analytically inverted in order to express the optical parameters in term of the experimental quantities.

To recover the spectral behavior of these parameters in paper sheets, we have developed a numerical procedure optimized for thin absorbent samples of low refractive index (Fig. 2). In particular, it has been necessary to minimize the instrumental errors involved in determining the subtle time difference between the reference pulses and those propagating in the sample ($\sim$ 100fs). See Supplemental Material and bibliography therein for further details [7].

The spectra of the optical parameters of each sample have been obtained by averaging over the different acquisitions performed.

A digital low-pass filter (0.14 THz resolution) was then applied at the averaged spectra to eliminating the noise present especially at frequencies higher than 2.5 THz. The resulting curves for the absorption coefficients, normalized for sample thickness and density, are shown in Fig. 3 for all measured samples in the range (0.2–3.5 THz) where data can be considered reliable. Due to small differences in the overall intensity of the spectra, the optical parameters have an uncertainty of about $\pm$10% in their intensity. This value can be attributed to a non-uniformity in terms of samples thickness and density as well as to small variations of the free water content in cellulose (due humidity fluctuation in the sample compartment of the THz set-up during measurements). This uncertainty also includes absorption by the small amount of lignin (around 2% in mass) that could be expected in the linen fibers of ancient samples. Indeed, it has been demonstrated that the THz spectra of lignin do not show peaks in the 0.2-4 THz range but a continuous increasing absorption [26].

The THz vibrational properties have been simulated by DFT calculations to interpret the experimental spectra. Unaged cellulose has been modeled as a crystal in its most common crystalline form (I_β phase) at 0 K. Van der Waals (VdW) interactions have been included by using a VdW-DF2 [13] nonlocal density functional to achieve a better description of the H-bonds [32, 33, 36–38, 40–43]. Computational details are reported in the Supplemental Material [7].

III. RESULTS AND DISCUSSION

The THz experimental absorption spectra of cellulose fibers in ancient and modern paper samples, free of *FP* oscillations, are shown in Fig. 3 (a typical refractive index spectrum is shown in the Supplemental Material [7]). Overall the absorption increases as a function of frequency: from 0.2 to about 1.6 THz a monotonic increase is evident. Superimposed to this trend two clear spectroscopic features can be observed for all samples: a shoulder at about 2.1 THz, and a peak at about 3 THz.

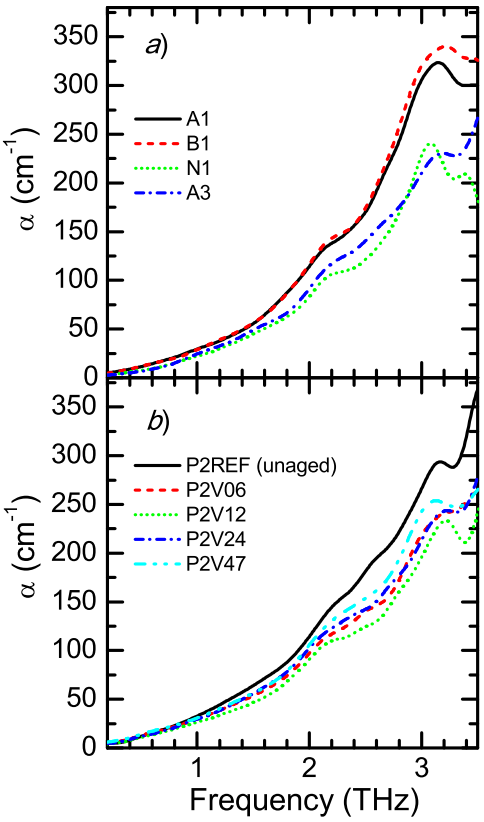

Figure 3. Absorption coefficients *vs* frequency of cellulose fibers in the ancient samples (panel *a)*) and modern samples artificially aged 0 (unaged), 6, 12, 24 and 47 days (panel *b)*) .

For some samples, minor features appear at intermediate frequencies, more evident for the artificially aged paper. A significant variation of the absorption coefficient occurs for all samples, going from about 10 cm^{-1} at ~ 0.2 THz to more than 200 cm^{-1} at 3 THz.

The interpretation of the THz spectra must be based on the knowledge of the microscopical structure of cellulose. A monotonic increase of the absorption is typically observed in amorphous solids [24] and biological macromolecular complexes formed by H-bond networks [20]. These absorption spectra can be generally approximated by a function $C\cdot\omega^{\beta}$ where C is a numerical coefficient which is expected to be proportional to the density of H–bonds, and β is an exponent which is approximately 2 in glassy materials [24, 27]. A similar behavior to that described for biosystems rich in H–bonds is then expected also for cellulose where disordered H-bond networks are present. In all paper samples, we found that the absorption coefficient curve from 0.2 to 1.6 THz can be well approximated by $A + C\cdot\omega^2$ where A is a constant that for modern samples is located in the range 9 – 11 cm^{-1} and for ancient samples is located in the range 4 – 9 cm^{-1}. The presence of the non–vanishing absorption contribution A is potentially connected with the existence of about a 2% in mass of bound water in the sample (see Fig.2). The numerical coefficient C is 26.0 cm^{-1}/THz^2 for the unaged sample P2REF while it has lower values for artificially and naturally aged samples up to 21.5 cm^{-1}/THz^2 for P2V47 and 18.6 cm^{-1}/THz^2 for N1 (due to small differences in the overall intensity of the spectra the uncertainty on the C values is $\pm 10\%$). We interpret these data as a reduction of the H-bond density in the amorphous cellulose network with aging.

Since cellulose is made of amorphous and crystalline regions, the peaks due to the crystalline phase are expected to be superimposed to the disordered H-bond spectral behavior. To separate the crystalline contribution from the disordered one, we have performed a subtraction of the $A + C\cdot\omega^2$ background term, extending the background fit deduced from the initial behavior (0.2 – 1.6 THz) up to about 3.3 THz, where the curves remain reliable. Results of such a subtraction are shown in Fig. 4 for all the samples. Furthermore, examples of the measurement uncertainties in the absorption curves are shown in the Supplemental Material [7].

The two groups of spectra exhibit different characteristic features. It is evident that ancient samples spectra show profiles with only two main peaks at about 2.1 and 3.1 THz. Instead, modern samples have more complex peak profiles that evolve as a function of the aging time from an evident many–peak profile, at about 2.1, 2.5 and 3.1 THz for the pristine specimen (P2REF), to a simplified two-peak profile, at about 2.1 and 3.1 THz, for the most artificially aged sample (P2V47). This confirms that the aging and degradation of paper is a complex process in which chemical and structural phenomena compete. As stated in previous works the aging process starts with a first chemical modification of the amorphous regions and of the external surfaces of the crystalline regions, later followed by material consumption, production of degradation byproducts [5], and resulting mass loss [16]. In parallel, a reorganization of the supramolecular structure of cellulose takes place with a consequent increasing of the sample crystallinity.

Based on all behaviors that can be observed in Fig.4, it can be inferred that the THz two-peak profile spectrum is representative of an extended degradation process, for both the ancient and the modern samples. The presence of a clear two-peak profile in the spectral range (1–3.3 THz) can, therefore, be regarded as the indicator of an advanced aging stage. Indeed, the spectroscopic features at about 2.1 and 3.1 THz were identified as due to the crystalline phase absorption as found in microcrystalline cellulose powders [28].

Our DFT simulations confirm that the spectral profile is composed of several peaks associated with long-range cellulose crystal phononic modes. The calculated peaks at about 2.2, 2.7, 3.0 and 3.3 THz appear to be in qualitative agreement with the experimental ones.

Indeed, it is reasonable that the numerical analysis could not be in perfect quantitative agreement with

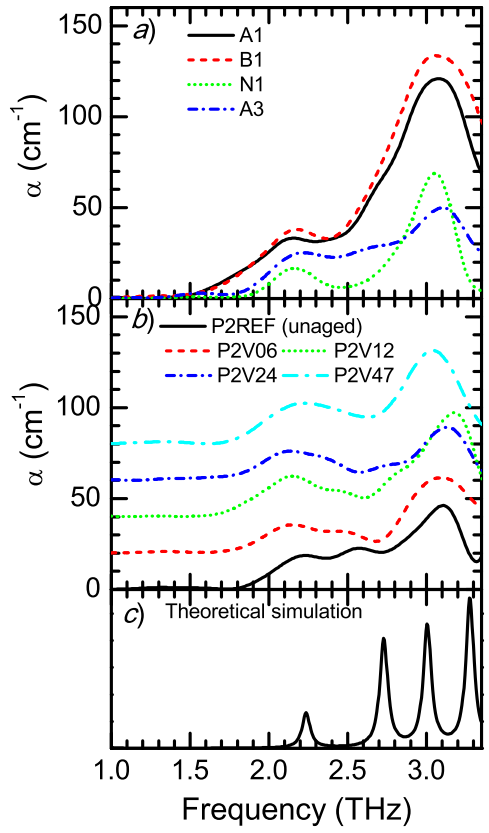

Figure 4. Panels *a)* e *b)* depict the absorption coefficients of cellulose fibers obtained by subtraction of the backgrounds from the curves reported in Fig.3 (see text for details) in the range 1–3.3 THz. For modern samples, each curve is shifted of 20 cm^{-1} along the vertical scale for clarity. In panel *c)* the theoretical THz absorption spectrum of a perfect crystal of unoxidized cellulose, with a Lorentzian broadening of 0.03 THz, is reported.

the experiments, since the DFT simulation is based on the assumption of a perfect crystal lattice. In the real supramolecular structure of cellulose many interchain hydrogen bonds are formed via water molecules. The collective dynamics of hydrated cellulose cannot be simulated by using the available DFT methods because the molecular arrangements of water molecules in the cellulose structure is not known. Therefore, the THz vibrational frequencies can be slightly different from those of a crystal made only be cellulose polymers.

As evidenced in a previous study, the hydrolysis induced by the artificial aging of P2 samples results in an increasing depolymerization of cellulose chains [15]. The degree of polymerisation decreases from about 1700 to about 700 passing from sample P2REF to the most artificially aged sample P2V47. In addition, the concentration variation (final-initial) for the oxidized groups increases: for –CHO groups of (0.045 ± 0.001)mmol/(g of cellulose) and for –COOH groups of (0.072 ± 0.003)mmol(g of cellulose). This suggests that some THz vibrational modes can be modified as degradation–induced structural modifications take place. Since degradation primarily happens in the amorphous regions [25], the observed THz spectral changes could be associated to the modification of vibrational modes in the recrystallized amorphous regions but also to the progressive degradation of the external surfaces of the crystalline domains, with a consequent enhancements of finite size effects in the phononic modes.

In the present study we have successfully obtained a precise determination of the THz absorption spectral profile of cellulose fibers in single freely standing paper sheets in the 0.2–3.5 THz range. This was achieved by numerically extracting the complex refractive index from our data in which the *FP* oscillations where accurately removed. By using this approach, the absorption coefficient of cellulose fibers in ancient and modern samples artificially aged was obtained. We found that the spectral behaviour can be explained as the superposition of a disordered *H*-bond network background with a many-peak profile, due to both amorphous and crystalline regions, respectively. The complex evolution of the spectra as a function of natural and artificial aging has been explained with a reduction of the *H*-bond density in the amorphous cellulose network and a parallel increase of the sample crystallinity. This in turn results, for advanced aging, in a simplified two-peak profile at about 2.1 and 3.1 THz.

These results demonstrate a significant potential in the analysis of the structural properties of cellulose and paper artifacts. While small cut samples were used for experimental convenience, THz-TDS can be performed even on whole sheets of ancient paper or even on the numerous pages bounded in a book [31]. To this aim, suitable sample holders are needed to keep safely the artifacts in the correct positions during the THz transmission measurements are needed. Nitrogen-purge was used to reduce water-vapor absorption lines, thus improving the signal to noise ratio and extending the measurable THz spectral range. However, it is also possible to only purge or put in vacuum the THz beam path up to the sample surface and leave the artifact in air. It is worth noting that this technique is non-destructive and able to provide a quantitative assessment of the state of preservation of cellulose artifacts, being therefore suitable for use in ancient and precious pieces of art.

ACKNOWLEDGMENTS

This work was funded by the Ministero dell'Istruzione, dell'Università e della Ricerca (MIUR) (Progetto Premiale "THEIA"), and by the NSERC and MESI in Canada. We thank the Istituto Centrale per il Restauro e la Conservazione del Patrimonio Archivistico e Librario (Roma, Italy) for its support. M.P. acknowledges the support of the Marie Curie Action MC-CIG, REA grant agreement 630833. We aknowledge CINECA for cpu

hours provided by projects: IsB05_LEONAR_E, IsB03_PAPER, IsB05_LEONARDO. M.C. acknowledges support from EPSRC (Grants No. EP/P009697/1 and EP/P51133X/1). All relevant data present in this publication can be accessed at: http://dx.doi.org/10.5525/gla.researchdata.405

[1] I. Brueckle. Structure and properties of dry and wet paper, In G. Banik and I. Brueckle, editors, *Paper and Water: A guide for conservators*, chapter 4, pages 81–119, Routledge, London and New York (2011).

[2] C. Corsaro, D. Mallamace, J. Łojewska, F. Mallamace, L. Pietronero, and M. Missori, Molecular degradation of ancient documents revealed by 1H HR-MAS NMR spectroscopy, Sci. Rep. **3**, 2896 (2013).

[3] C. Corsaro, D. Mallamace, S. Vasi, L. Pietronero, F. Mallamace, and M. Missori, The role of water in the degradation process of paper using 1H HR-MAS NMR spectroscopy, Phys. Chem. Chem. Phys. **18**, 33335-33343 (2016).

[4] M. De Spirito, M. Missori, M. Papi, G. Maulucci, J. Teixeira, C. Castellano, and G. Arcovito, Modifications in solvent clusters embedded along the fibers of a cellulose polymer network cause paper degradation, Phys. Rev. E **77**, 041801 (2008).

[5] A.-L. Dupont, A. Seemann, and B. Lavédrine, Capillary electrophoresis with electrospray ionisation-mass spectrometry for the characterisation of degradation products in aged papers, Talanta **89**, 301–309 (2012).

[6] T. Hattori, H. Kumon, and H. Tamazumi, Terahertz spectroscopic characterization of paper, In *35th International Conference on Infrared, Millimeter, and Terahertz Waves (IRMMW-THz 2010)* (2010).

[7] See Supplemental Material at [URL will be inserted by publisher] for details on: THz TDS set-up, samples thickness and density measurements, DFT simulations, extraction algorithm, measurement uncertainties in the absorption curves.

[8] P. U. Jepsen, D. G. Cooke, and M. Koch, Terahertz spectroscopy and imaging - modern techniques and applications, Laser Photonics Rev. **5** (1), 124–166 (2011).

[37] A. K. Kelkkanen, B. I. Lundqvist, and J. K. Norskov, Density functional for Van der Waals forces accounts for hydrogen bond in benchmark set of water hexamers, Chem. Phys. **131**, 046102 (2009).

[10] D. Klemm, B. Philipp, T. Heinze, U. Heinze, and W. Wagenknecht, *Comprehensive Cellulose Chemistry Volume I*, WILEY-VCH Verlag GmbH, Weinheim (1998).

[11] D. Klemm, B. Heublein, H. P. Fink, and A. Bohn, Cellulose: fascinating biopolymer and sustainable raw material, Angew. Chem. Int. Ed. **44**, 3358–3393 (2005).

[12] H. A. Krassig. *Cellulose: Structure, Accessibility, and Reactivity*. Gordon and Breach Science Publishers, Singapore (1993).

[13] K. Lee, E. D. Murray, L. Kong, B. I. Lundqvist, and D. C. Langreth, Higher-accuracy Van der Waals density functional, Phys. Rev. B **82**, 081101(R) (2010).

[14] P. Łojewski, P. Miśkowiec, M. Missori, A. Lubańska, L. Proniewicz, and J. Łojewska, Ftir and UV/Vis as methods for evaluation of oxidative degradation of model paper: DFT approach for carbonyl vibrations, Carbohydr. Polym. **82**, 370 (2010).

[15] T. Łojewski, K. Zieba, A. Knapik, J. Bagniuk, A. Lubańska, and J. Łojewska, Evaluating paper degradation progress. cross-linking between chromatographic, spectroscopic and chemical results, Appl. Phys. A **100**, 809–821 (2010).

[16] T. Łojewski, K. Ziebla, A. Knapik, J. Bagniuk, A. Lubańska, and J. Łojewska, Furfural as a marker of cellulose degradation. a quantitative approach, Appl. Phys. A **100**, 809–821 (2010).

[17] D. Mallamace, S. Vasi, M. Missori, and C. Corsaro, New insight into hydration and aging mechanisms of paper by the line shape analysis of proton NMR spectra, Il Nuovo Cimento C **39**, 309–319 (2016).

[18] M. Missori, O. Pulci, L. Teodonio, C. Violante, I. Kupchak, J. Bagniuk, J. Łojewska, and A. Mosca Conte, Optical response of strongly absorbing inhomogeneous materials: application to paper degradation, Phys. Rev. B **89**, 054201 (2014).

[19] A. F. Mosca Conte, O. Pulci, M. Misiti, J. Łojewska, L. Teodonio, C. Violante, and M. Missori, Visual degradation in Leonardo da Vinci's iconic self-portrait: a nanoscale study, Appl. Phys. Lett. **104**, 224101 (2014).

[20] M. Naftaly and R. E. Miles, Terahertz time-domain spectroscopy for material characterization, Proceedings of the IEEE **95** (5), 1658–1665 (2007).

[21] A. C. O'Sullivan, Cellulose: the structure slowly unravels, Cellulose **4**, 173–203 (1997).

[22] Joo-Hiuk Son, *Terahertz Biomedical Science and Technology*, CRC Press (2014).

[23] C. J. Strachan, T. Rades, K. C. Newnham, D. A. Gordon, Michael Pepper, and Philip F. Taday, Using terahertz pulsed spectroscopy to study crystallinity of pharmaceutical materials, Chemical Physics Letters **390**, 20–24 (2004).

[24] U. Strom and P. C. Taylor, Temperature and frequency dependences of the far-infrared and microwave optical absorption in amorphous materials, Phys. Rev. B **16** (12), 5512–5522 (1977).

[25] L. Teodonio, M. Missori, D. Pawcenis, J. Łojewska, and F. Valle, Nanoscale analysis of degradation processes of cellulose fibers, Micron **91**,75 (2016).

[26] T. Trafela, M. Mizuno, K. Fukunaga, and M. Strlič, Quantitative characterisation of historic paper using THz spectroscopy and multivariate data analysis, Appl. Phys. A **111**, 83–90 (2013).

[27] S. Tsuzuki, N. Kuzuu, H. Horikoshi, K. Saito, K. Yamamoto, and M. Tani, Influence of OH-group concentration on optical properties of silica glass in terahertz frequency region, Applied Physics Express **8**, 072402 (2015).

[28] F. S. Vieira and C. Pasquini, Determination of cellulose crystallinity by terahertz-time domain spectroscopy, Anal. Chem. **86**, 3780–3786 (2014).

[29] M. Walther, B. M. Fischer, and P. U. Jepsen, Noncovalent intermolecular forces in polycrystalline and amorphous saccharides in the far infrared, Chem. Phys. **288**, 261–268 (2003).

[30] J. Xu, Kevin W. Plaxco, and S. James Allen, Absorption spectra of liquid water and aqueous buffers between 0.3 and 3.7THz, The Journal of Chemical Physics **124**, 036101 (2006).

[31] A. Redo-Sanchez, B. Heshmat, A. Aghasi, S. Naqvi, M. Zhang, Justin Romberg, and Ramesh Raskar, Terahertz time-gated spectral imaging for content extraction through layered structures, Nature Communications **7**, 12665 (2016).

[32] A. Calzolari and M. Buongiorno Nardelli, Dielectric properties and Raman spectra of ZnO from a first principles finite-differences/finite-fields approach, *Sci. Rep.* **3**, 2999 (2013).

[33] M. Dion, H. Rydberg, E. Schroder, D. Langreth, and I. B. Lundqvist, Van der Waals density functional for general geometries, *Phys. Rev. Lett.* **92**, 246401 (2004).

[34] L. Duvillaret, F. Garet, and J. L. Coutaz, Highly precise determination of optical constants and sample thickness in terahertz time-domain spectroscopy *Appl. Opt.* **38**(2), 409–415 (1999).

[35] L. Duvillaret, F. Garet, and J.-L. Coutaz, A reliable method for extraction of material parameters in terahertz time-domain spectroscopy, *IEEE Journal of Selected Topics in Quantum Electronics* **2**(3), 739 (1996).

[36] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G.L. Chiarotti, M. Cococcioni, I. Dabo, A. Dal Corso, S. de Gironcoli, S. Fabris, G. Fratesi, R. Gebauer, U. Gerstmann, C. Gougoussis, A. Kokalj, M. Lazzeri, L. Martin-Samos, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Sclauzero, P. Ari Seitsonen, A. Smogunov, P. Umari, and R. M. Wentzcovitch, Quantum espresso: a modular and open-source software project for quantum simulations of materials, *Journal of Physics: Condensed Matter* **21**(39), 395502 (2009).

[37] A. K. Kelkkanen, B. I. Lundqvist, and J. K. Norskov, Density functional for van der waals forces accounts for hydrogen bond in benchmark set of water hexamers, *Chem. Phys.* **131**, 046102 (2009).

[38] A. Mosca Conte, O. Pulci, R. Del Sole, A. Knapik, J. Bagniuk, J. Łojewska, L. Teodonio, and M. Missori, Visual degradation in Leonardo da Vinci's iconic self-portrait: a nanoscale study, *e-J. Surf. Sci. Nanotech.* **10**, 569–574 (2012).

[39] A. Mosca Conte, O. Pulci, A. Knapik, J. Bagniuk, R. Del Sole, J. Łojewska, and M. Missori, Role of cellulose oxidation in the yellowing of ancient paper. *Phys. Rev. Lett.* **108**, 158301 (2012).

[40] J. Perdew and Y. Wang, Accurate and simple density functional for the electronic exchange energy: Generalized gradient approximation, *Phys. Rev. B* **33**, 8800(R) (1986).

[41] R. Sabatini, T. Gorni, and S. de Gironcoli, Nonlocal van der waals density functional made simple and efficient, *Phys. Rev. B* **87**, 041108(R) (2013).

[42] D. Vanderbilt, Soft self-consistent pseudopotentials in a generalized eigenvalue formalism, *Phys. Rev. B* **41**, 7892(R) (1990).

[43] C. Violante, L. Teodonio, A. Mosca Conte, O. Pulci, I. Kupchak, and M. Missori, An ab-initio approach to cultural heritage: The case of ancient paper degradation, *Physica Status Solidi B* **252**, 112–117 (2015).