Investigation of the effects of enzymatic hydrolysis and hydroxylation on oriented semi-crystalline cellulose by *ex situ* polarisation dependent Raman spectroscopy

Monika Pranjic

M.Sc. by research

University of York

Chemistry

February 2021

Abstract

The aim of this research is to develop a robust and quantitative Raman-based enzymatic assay by investigating degraded, naturally oriented cellulose from celery (*Apium graveolens*) after enzymatic treatment with polarised confocal Raman spectroscopy. . The thesis comprises two main studies: a polarisation dependent study on oriented celery (*Apium graveolens*) cellulose and an enzymatic study. The aim of the work on naturally oriented cellulose was to capture the Raman signature and increase the understanding of the structural properties of typical, but highly oriented, natural semi-crystalline cellulose. Celery cellulose was used as a natural substrate for cellulase and LPMO activity and hence represents the spectral background of the Raman analytic assay. The Raman signature of celery cellulose at orientations of 0°- 90° in increments of 30° ± 1° was analysed using peak intensity ratio analysis (PIR), which showed characteristic changes of peak intensities with respect to the collection angle. These are consistent with the literature on Ramie fibers. The enzymatically treated cellulose study used cellulases Cel7A, Cel7B and LPMO *Sc*AA10 aka CelS2. The Raman spectra were analysed using PIR and principal component analysis (PCA). PCA was able to distinguish Cel7A and Cel7B treated samples from negative controls due to disruption of axial orientation and increased crystallinity, respectively. PIR results showed specific Raman signatures at 1095 cm^{-1} and 1120 cm^{-1} to be potential markers for cellulose modification at advanced stages of enzymatic degradation. PCA and PIR analyses on LPMO-treated cellulose showed that the samples were similar to its negative controls. This could either be due to the subtle surface changes not significantly affecting Raman spectra; or the differences in samples not reaching the Raman detection limit. The study represents a starting point for developing a more complex polarized Raman assay that requires multivariate calibration and classification models based on larger datasets.

List of Contents

List of Figures

List of Tables

Acknowledgements

First and foremost, I am grateful to my supervisors Prof. Paul Walton and Dr. Yvette Hancock for their valued assistance in the realization of this research. A great thank you also goes to Prof. Jane Thomas-Oates and my lab colleagues and friends Marcus Cameron, Alessandro Paradisi, Martin Steward, Peter Lindley, George Nunn, Daniel Diaz Romero and our visitors Joao Lourenco Franco Cairo, Cesar Terrasan and Thamy Correa and former lab colleagues Claire Fowler and Luisa Ciano. A great thank you goes to Saioa Urresti and Nicholas McGregor who equipped me with the enzymes for this work. A special thank you also goes to Prof. Roland Kröger and Wayne Twigger for their support regarding the Horiba Raman instrument and Chris Taylor who did the introduction to the MALDI-TOF instrument. An equally heartfelt thank you goes to the UOY Chemistry Graduation Office. I am thankful to my friends Ben and David Jones, James Stovold, Ben Coulson, Hazel Cheung, and Rachel Carr. A great thank you to everyone involved in the Oxytrain program; I really appreciate that I was given the opportunity to take part in it. Last but not least, I am thankful to my family for their continuous support in all situations. This project has received funding from the European Union’s Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No.722390.

Declaration

I declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References.

Work carried out by other persons;

- Protein expression and purification of LPMO *Sc*AA10 (CelS2) was carried out by Dr. Glyn Hemsworth.
- Cellulases Cel7A and Cel7B were provided by Dr. Mc Gregory.

1. Introduction

The worldwide increasing energy demand and concerns about the depletion of non-renewable fossil fuels (e.g. coal, petroleum and natural gas) are the driving forces for the industrial development of economically viable uses of renewable energy sources of low environmental impact.[1–3] Cellulosic biomass represents such an abundant carbon-neutral source which is used to produce second generation biofuels and various chemicals.[4,5] A continuous supply of plant-based biomass derives from forest residues like wood chip sources, forest debris, as well as agricultural waste matter from crop harvesting.[4] This makes cellulose an almost inexhaustible polymeric raw material[6]. The exploitation of these resources can have a significant impact on energy supply chains as they reduce the need for high-cost energy importation.[5] Also, the energy production of second-generation biofuels deriving from cellulosic biomass is more sustainable. It therefore represents a key material to secure a sustainable low carbon economy in the future.[7,8] The use of cellulose through conversion into mono- and oligosaccharides has since been realised.[9]

Cellulose is the most abundant biological material on earth with levels of 1.5 x 10^{12} tons annually and a half-life of several million years.[8,10,11] In industry, cellulose is mainly involved in paper and cardboard production.[9] Further applications involve the production of cellulose regenerate fibers and films, as well as the synthesis of cellulose esters and ethers.[9,11,12]

Cellulose is entirely composed of carbohydrate polymers (polysaccharides) that, once broken into their mono- or oligosaccharide components, can be used in 2nd generation bioethanol and biofuel production.[8] The strong glycosidic (COC) linkages in linear polymers of β-1,4-D-glucose of cellulose microfibrils can be cleaved enzymatically, which allows an environmentally friendly conversion of polysaccharide chains into high value components.[6,13] The molecular structure of cellulose I as a carbohydrate polymer with repeating β-D-gluco-pyranose units is shown in Figure 1.1. The units are covalently linked through acetal functions between the equatorial OH group of the C4 and C1 carbon atom (β-1,4-glucan). Every second glycosidic ring unit is rotated 180° in the plane to accommodate the preferred bond angles of the acetal oxygen bridges (Figure 1.1).[9]

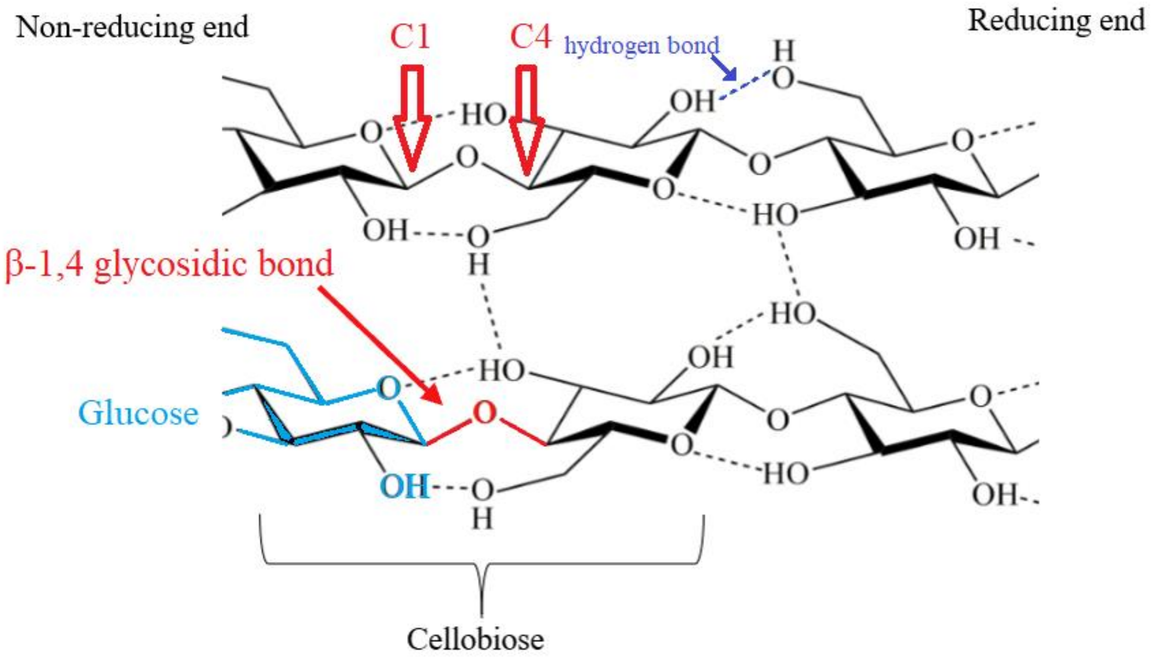


Figure 1. 1: Chemical structure of cellulose I. The non-reducing end is the monosaccharide residue in acetal form, while the reducing end of an oligosaccharide is the residue with hemiacetal functionality. The polymeric structure is formed by repeating units of cellobiose (β(1→4)-linked D-glucose).[14]

Cellulose has several crystalline allomorphs, which essentially differ in their ‘pattern’ of hydrogen bonds between cellulose chains: Natural cellulose I is produced by bacteria, algae and plants. Regenerated cellulose II is non-naturally produced by mercerisation of cellulose I fibers (alkali treatment, solubilisation and recrystallisation); and cellulose III can be obtained from cellulose II or cellulose I by treatment with liquid ammonia.[15–19] However, the work described within this thesis primarily focuses on the allomorphic forms of natural cellulose I. Cellulose Iα is dominant in organisms such as bacteria or algae. Cellulose Iβ is the main component in higher plants and therefore the most abundant form of cellulose found in nature.[20]

Natural cellulose is very resistant to degradation and industrially suitable biocatalysts for cellulose degradation are relatively slow and inefficient.[21] The recalcitrance of cellulose is linked to the highly crystalline structure which limits the accessibility of the glycosidic linkages .[22] The factors affecting the enzymatic degradation of cellulose are hence both substrate- and enzyme-related. [23–25] The slow degradation rate and the physicochemical recalcitrance of cellulosic biomass makes its biological processing costly and constrains its commercial usability.[26] Against this background, considerable effort has been expended in attempts to identify efficient biokatalysts.[27] Although some key factors in enzyme-substrate interactions have been elucidated, in-depth investigations of cellulose degradation processes are usually challenging because of the insolubility of cellulose substrates.[28]

Due to the challenging aspects of studying enzyme activity on insoluble substrates (e.g. heterogeneity of the system), a representative enzymatic assay is required, that displays cellulose activity on natural cellulose.[29–31] Experimental approaches to quantitatively analyse insoluble cellulose residues are often time-consuming, require specialized and in some cases expensive equipment, and are low in throughput.[9,32] Solid-state nuclear magnetic resonance spectroscopy (ssNMR) represents such an experimental approach, that can otherwise be used to obtain information about substrate compositions and supramolecular structures between biomass polymers.[33,34] Routinely used methods such as HPLC together with mass spectrometric analyses are restricted to the analysis of soluble mono- and oligosaccharides. Although insoluble crystalline cellulose products may be hydrolysed for complete analytical quantification, this prevents observing the natural supramolecular distribution of remaining cellulose chains.[35]

High-end imaging techniques such as scanning electron microscopy (SEM), atomic force microscopy (AFM) and ultra-violet (UV) microscopy revealed gross changes in the morphology of polysaccharide nanostructures after enzymatic treatment and are rich in physical analysis.[23,36,37] However, they ultimately lack chemical information.

Against this background, vibrational spectroscopy methods have evolved to a common analytical tool for a variety of applications in the field of cellulose research[38–44]: Raman scattering and Infrared absorption (IR) represent the main vibrational spectroscopic methods to analyse cellulose. Due to the continuous technological and scientific advances in the field of vibrational spectroscopy, native plant cell walls and biopolymer components, their distribution, linkages as well as interdependencies can be probed in a comparably fast and cheap manner.[32,40]

Raman spectroscopy is a vibrational spectroscopy method and as such related to near/ and mid-IR absorption spectroscopy. Similar to IR absorption techniques, Raman spectroscopy measures the vibrational frequencies of molecular bonds.[45] These frequencies depend on both the bond strength and mass of the bound atoms as well as intermolecular interactions. The ‘pattern’ of vibrational frequencies from a molecule is, therefore, highly characteristic of a given molecular species and, for solid samples, of the crystalline arrangement of those molecules.[45] While IR spectroscopy is essentially based on illuminating the sample with a broad range of wavelengths of IR light, and measuring which wavelengths are being absorbed energetically, a Raman spectrum is obtained by illuminating the sample with a single wavelength of a laser and analysing the change in wavelength, when the light is being deflected by a sample (inelastic scattering).[45]

The fundamental difference between Raman and IR spectroscopy is due to the following occurring physical-chemical event: Whereas IR absorption arises by a change in the dipole moment of a molecule (known as a separation of charge), Raman scattering is caused by a change in the polarizability of a molecule (an 'induced' dipole moment acquired in proportion to an electric field, the molecule has been subjected to).[46] The advantage of Raman spectroscopy compared to IR spectroscopy is therefore, that it can also be used in highly polar bond systems such as aqueous solutions which is essential for enzymatic systems.

In this study, polarization dependent, confocal Raman spectroscopy with univariate and multivariate analyses have been used to investigate degraded, naturally oriented cellulose from celery after enzymatic treatment (ex situ). Raman spectra may be readily recorded from solids, allowing to probe structural and chemical information such as crystallinity, polymorphism and hydrogen bonding of cellulose materials in a non-destructive, label-free and quantitative manner.[47–49] The study represents a starting point for developing a robust and quantitative polarized Raman assay.

The following sections briefly describe the hierarchical structure of cellulose; the Raman spectroscopic analysis of cellulose samples; as well as the mechanisms of the enzymes involved in the enzymatic cellulose degradation.

.

1.1 Natural cellulose substrates

The hierarchical structure of natural cellulose has been the subject of intense research, with frequent controversy over results and a consistent supply of new insight.[50] The progress was closely connected to the introduction and continued development of structure-analysis methods, such as NMR and X-ray crystallography. Detailed analysis and modelling of the various structural levels of cellulose was essential for the optimization of thermochemical reaction procedures and properties of cellulose-based products.[9] In general, there are two approaches to deconstruct biomass: the thermochemical and the biochemical approach.[51,52] The thermal techniques suffer from low selectivity and mixed heterogeneous products, as pyrolysis and gasification convert biomass into solid (char), biogas and liquid (biocrude/bio/oil) fractions.[10] Another drawback is that thermochemical biomass deconstruction includes high temperatures and environmentally unfriendly pre-treatment conditions with the use of acid, alkali and cellulosic solvents in the process of biofuel production.[53] The biochemical approach is currently costly but more environment friendly and relies on three main unit operations: the physical-chemical pre-treatment of the biomass, the enzymatic hydrolysis of polysaccharides to mono- and oligosaccharides and finally the fermentation of the sugars by microorganisms. The use of enzymes results in substantially more selective products and has a significantly lower environmental impact compared to thermochemical approaches.[54] With the introduction of solid-state NMR for cellulose analysis, it was established that all native celluloses are composites of two forms, Cellulose Iα and Iβ, which primarily differ in their pattern of hydrogen bonding. The microfibril crystal structures of celluloses and their Iα/Iβ ratio vary depending on their source of origin.[55,56] Cellulose Iα has a single-chain triclinic unit cell whereas cellulose Iβ (monoclinic unit cell) includes two parallel cellulose chains (Figure 1.2).[57] The individual cellulose chains are held together by van der Waals forces with intramolecular and intermolecular hydrogen bonds.[6]

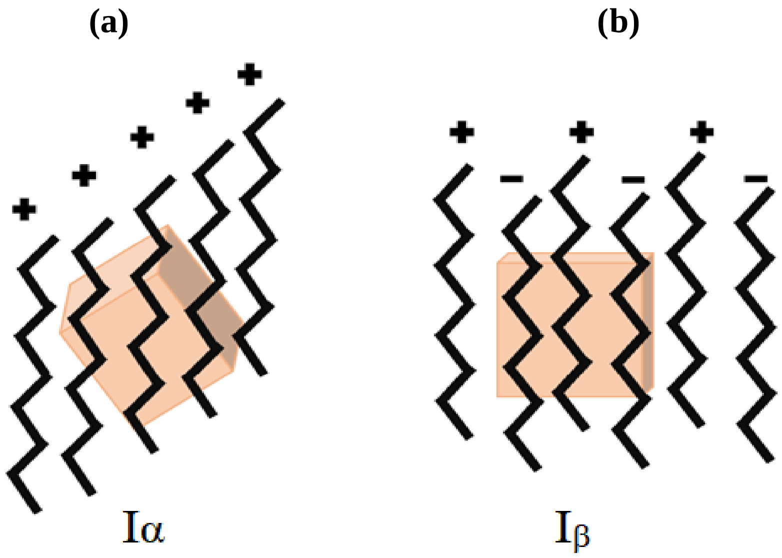


Figure 1. 2: (a) Schematic representation for cellulose type I structure Iα. (b) Diagram for cellulose Iβ with two parallel cellulose chains. The displacement of sheets is due to the hydrogen bonding. Black zig zag lines are cellulose chains.

Natural cellulose is packed in microfibrils containing a large number of crystalline (ordered) and paracrystalline/amorphous (disordered) crystal-surface chains. Bundles of microfibrils (sub-fibers) form the cellulosic fiber, as shown in Figure 1.3.

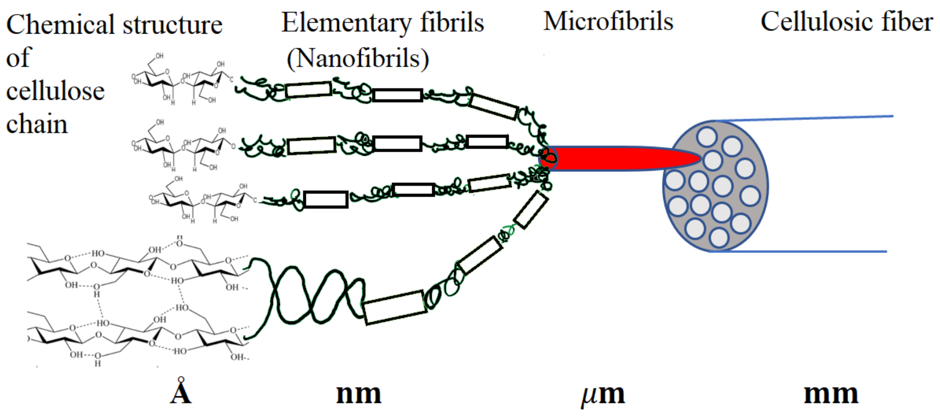


Figure 1. 3: : Hierarchical structure of cellulose showing all constituents at different scales.

1.1.1 Cellulosic substrate: Celery cellulose

This work focuses on the Raman analysis of enzymatic degradation on natural cellulose from Celery (*Apium graveolens*), which contains semi-crystalline cellulose typical for celluloses in higher plants. Celery cellulose is highly oriented in nature, which facilitates spectroscopic studies because of their uniform structure, as examined by Jarvis *et al.*[56,58–61] The cellulose fibers were extracted from the primary cell walls of celery collenchyma. Collenchyma forms the fundamental tissue in living (growing), elongated plant cells with irregularly thick cell walls. The cell wall consists of a complex matrix of celluloses, hemicelluloses and soluble proteins, which in celery collenchyma are mostly pectic polysaccharides (α(1→4)-D-galacturonosyl) and the hemicellulose xyloglucan (XG).[62] Cellulose and hemicellulose are complex polysaccharides of hydrophilic character.[63] Whilst cellulose consists of β (1→4) linked D glucose units which are arranged in a crystalline fashion forming several hydrogen bonds, hemicellulose consists of various monomeric units of arabinose, galactose, glucose, xylose and mannose in an amorphous structure.[64,65] Pectin is a structural acidic heteropolysaccharide that binds cells together.[62] The structures of pectin and XG are shown in Figure 1.4.

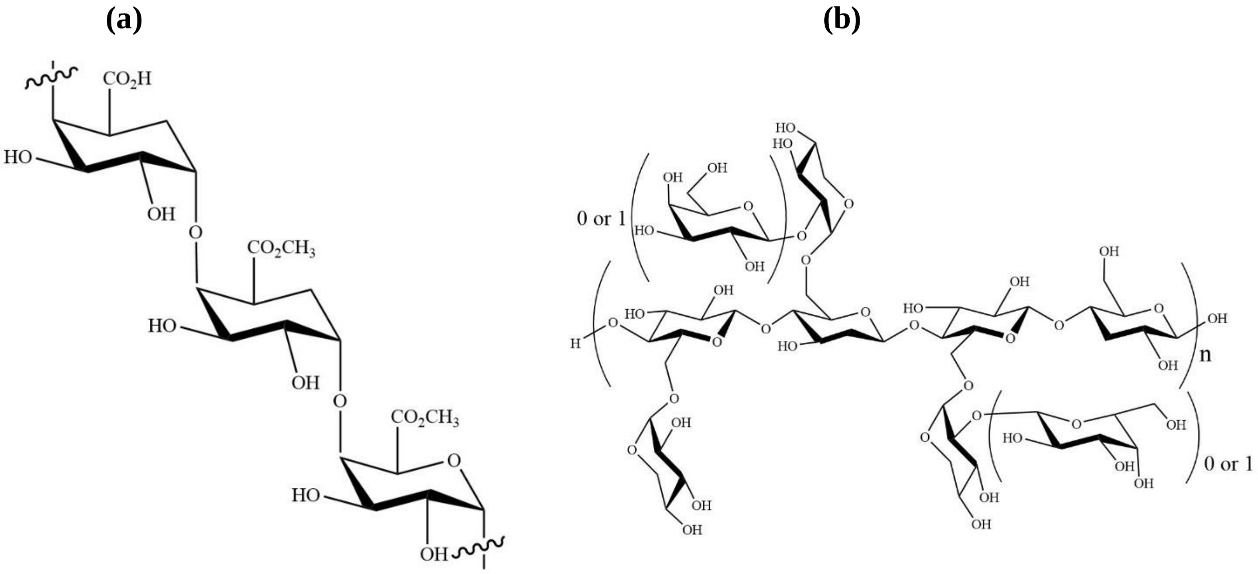


Figure 1. 4: : Chemical structure of (a) pectin and (b) xyloglucan, which are the main components of celery collenchyma next to cellulose.

The fiber wall morphology of the extracted cellulose was largely retained by the mechanical and chemical procedure described in literature.[58,60] The cellulose microfibrils in celery collenchyma are 2.9 to 3.0 nm in mean diameter with a particle depth distribution of 3.2 nm and a most probable structure of 24 chains arranged in eight hydrogen bonded sheets of three chains.[61] The approximate dimensions of the cellulose chains within the celery microfibril are shown in Figure 1.5.[66]

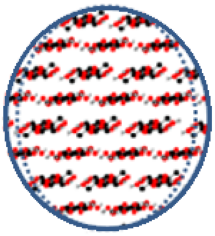


Figure 1. 5: Approximate dimensions of microfibrils from primary cell walls of celery as estimated by Jarvis *et al*. Their results were obtained using wide-angle X-ray scattering (WAXS) analysis.[66] The celery microfibril diameters are 3.2 x 3.0 nm, containing in average about 21 chains.

Figure 1.6 represents a model of a primary cell wall consisting of cellulose and hemicellulose and pectin between the plasma membrane and the middle lamella.[67]

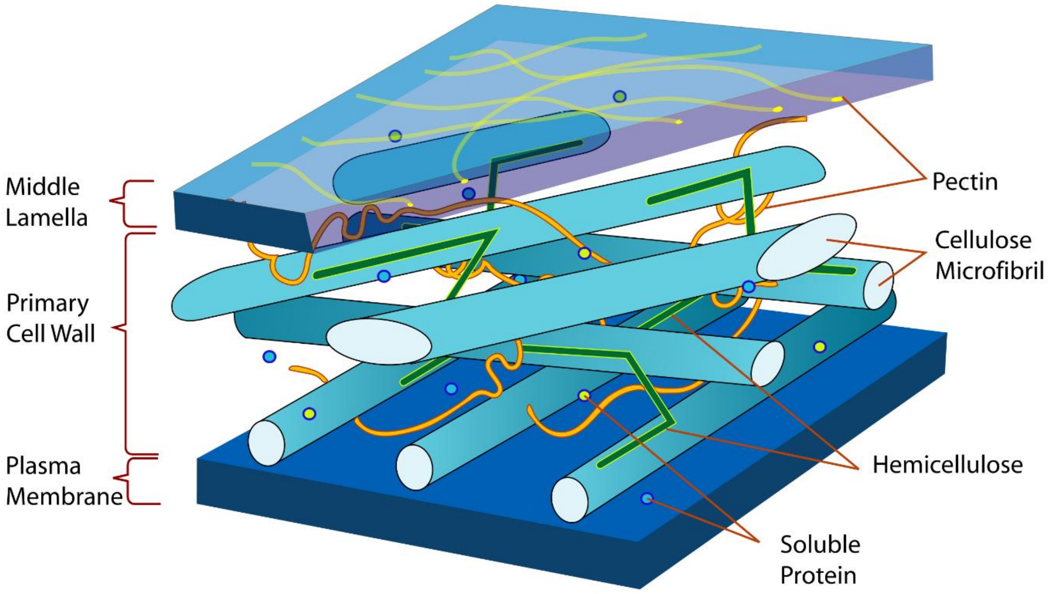


Figure 1. 6: : General representation of primary cell walls in plant. The different polysaccharide components form a complex matrix contributing to the primary cell wall enclosed by the middle lamella and plasma membrane. The primary cell wall consists of acidic pectin molecules, neutral pectin molecules, Ca^{2+}-bridges between pectin molecules, glycoprotein and cellulose microfibril. Image free for public use from www.wikimediaorg.

1.1.2 Cellulose crystallinity

A structural feature that affects the rate of enzymatic hydrolysis and hydroxylation of cellulose is the degree of crystallinity. For example, it has been shown that the paracrystalline region generally hydrolyse faster than the highly crystalline region.[68] Moreover, cellulase Cel7B which was used in this study, is known to only attack paracrystalline regions, whereas cellulase Cel7A can also attack crystalline regions. The crystallinity index (CI) is a quantitative indicator for the relative amount of structural order in a material given in percentage (%).[69] Multiple methods have been established to determine the crystallinity of cellulose. The results vary depending on the methods chosen and hence require cross-validation.[38,70,71] FTIR is often used for comparison purposes, whereas XRD and NMR analyses determine more accurately the crystallinity index.[69] One consideration is that XRD and NMR measure the average of the entire cellulose crystal, while carbohydrate active enzymes only interact with the substrate surface area that generally represent <10% of the whole substrate.[72] Moreover, other substrate characteristics such as fibrillation and macroscopic organization all influence the crystallinity measurement.[72]

1.2 Raman spectroscopy

Technological and scientific advances, like the development of laser sources since the 1960s, or multichannel Charge Coupled Device (CCD) and multiplex detectors, drastically increased the application of Raman spectroscopies in analytical research.[38] Raman spectroscopy was first applied in the early 1970s as an analytical tool to investigate cellulose materials.[43,73,74] Since then it has evolved to a common analytical technique in pulp and paper industry, in investigations on biomass and for the analysis of chain polarities, as well as inter- and intramolecular bonding.[73,75–80] Previous Raman studies also investigated changes in structure and properties of biomass during enzymatic treatment, determining that the crystallinity of cellulose increases.[41] Briefly, the basic experimental arrangement to perform confocal Raman spectroscopy includes the following steps: The excitation laser is focused on the sample via an optical fiber and the microscope objective. The backscattered light propagates into a fiber that acts as a pinhole. After passing the spectrometer, the signal is recorded by a CCD camera. A microscopic setup (white light source and a camera) is usually available for visual inspection of the sample (see Figure 1.7 a). A protractor was used for the angle-dependent alignment of cellulose samples, which were attached with minimal tension on a CaF_2 microscopy slide (see Figure 1.7 b).

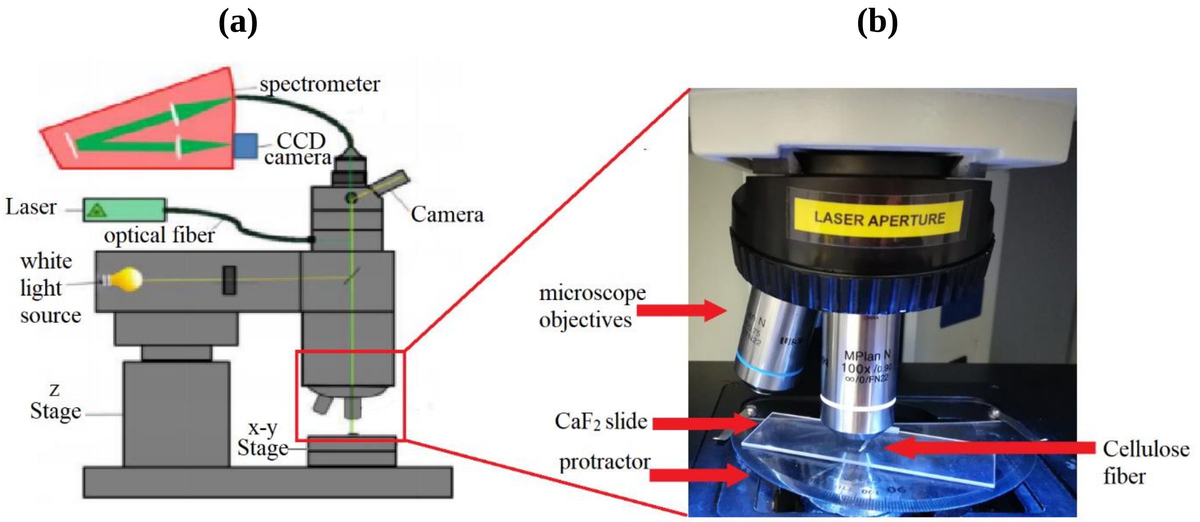


Figure 1. 7: : (a) Set up of a confocal Raman microscope. The system is equipped with a X- Y- stage and a Z-stage for mapping/scanning (Image free for public use from Gierlinger et al., 2013).[28] (b) Image of experimental setup: A protractor was used for cellulose orientation and analysed by confocal Raman spectroscopy.

For the understanding of physico-chemical information obtained by Raman spectroscopy, the following section presents the theoretical background on Raman spectroscopy in more detail.

1.2.1 The Raman effect

The Raman effect was first reported in 1928 by Sir C.V. Raman. It describes the change in the wavelength of light when a light beam is deflected by molecules (see Figure 1.8).[81] This inelastic scattering of photons probes the vibrational and the rotation-vibrational states of molecules based on the polarizability of the vibrating structure.[82]

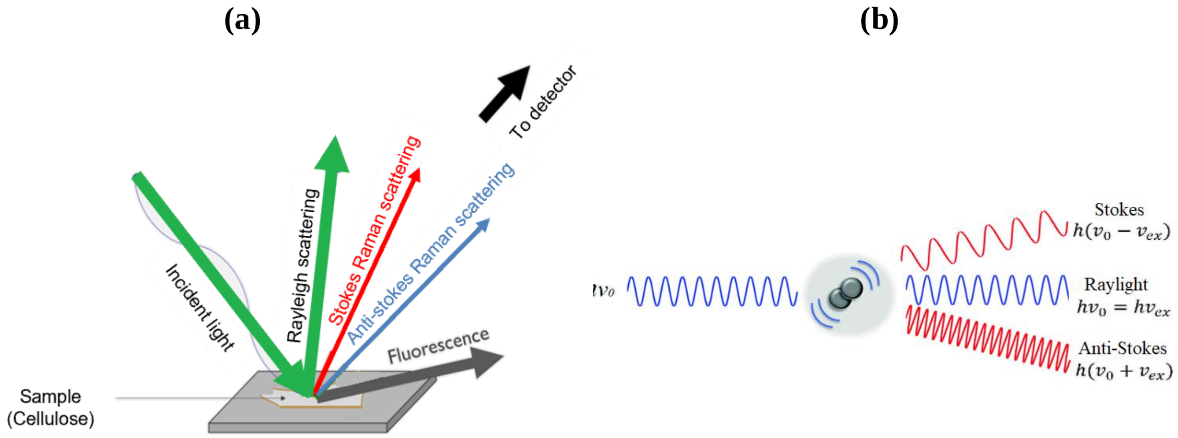


Figure 1. 8: : (a) Types of scattered light: Rayleigh scattering, Stokes Raman scattering, Anti-stokes Raman scattering, and (depending on chemical sample) fluorescence. (b) Types of scattered light shown as frequencies.

Molecules exist in various energy levels with the majority of molecules in their vibrational ground state.[46] The vibrational ground state v_0 is indicated with a thick line at the bottom of the Jablonski diagram in Figure 1.9. If a photon with a discrete energy of hv_0 collides with a molecule, one of three types of scattering may occur: elastic Rayleigh scattering where the energy of the incoming photon remains unchanged; Stokes scattering, where a photon with decreased energy $h(v_0 - v_{ex})$ is scattered; or anti-Stokes scattering, where a photon with increased energy $h(v_0 + v_{ex})$ may be scattered (see Figure 1.9). The majority of light is scattered at the same wavelength and frequency as the incident light (elastically). Only 1 in 10^6 – 10^8 photons are scattered inelastically, where the molecules fall back to a different vibrational energy state. Conventional Raman spectroscopy only measures the Stokes scattered photons since Stokes scattering results in more intense spectra than anti-Stokes scattering. The reason is that at room temperature the population of molecules in the vibrational ground state (v_0) is greater than in the excited state, as defined by the Maxwell Boltzmann distribution.[83] The electrons are excited to a short-lived, unobservable virtual energy state and return to their natural state following the scattering of the photon.[46] Infrared (IR) spectroscopy, on the other hand, is based on the absorption of a photon in the infrared region. If the molecular bond has a dipole moment and the photon is the same frequency as the frequency of a bond, absorption occurs.[46] The vibrational frequencies are observed as absorption peaks in the IR region of transmitted light.

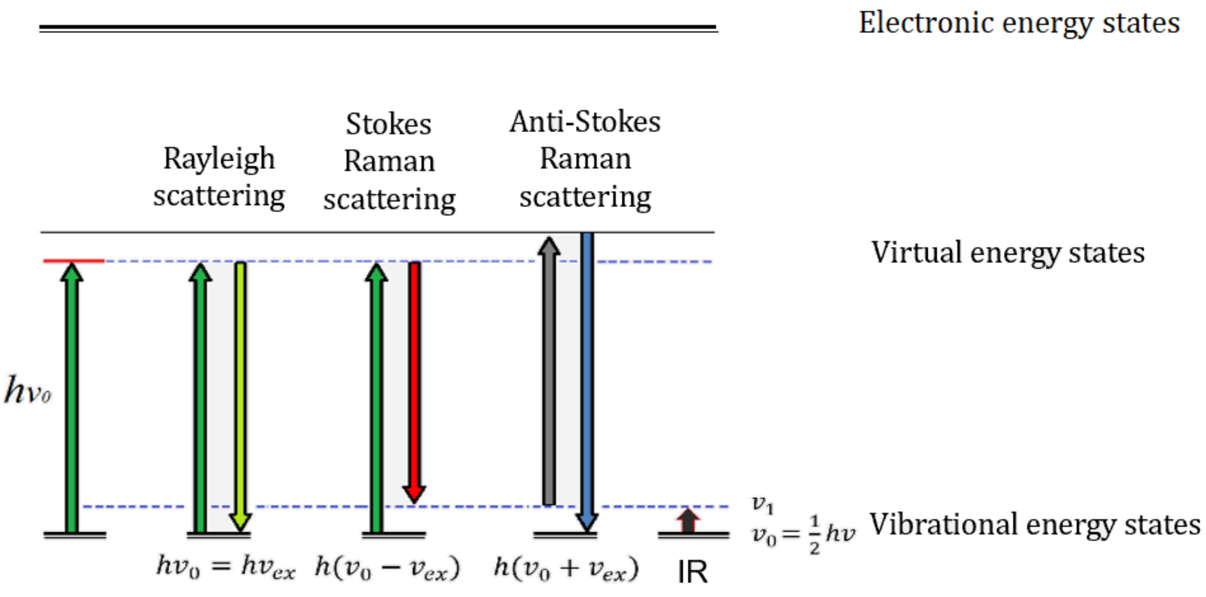


Figure 1. 9: : Jablonski diagram showing vibrational states for description of the Raman effect. The scattered photon h_{ν_0} deflects from a molecule either elastically (Rayleigh scattering) or inelastically (Raman scattering) The Raman effect is explained by the molecule's excitation to a virtual energy state, where the scattered photon has either lower (Stokes scattering) i.e. "red shift" or higher (Anti-Stokes scattering) energy i.e. "blue shift" in comparison to the incident photon (modified from Ferraro et al., 2003).[46]

1.2.2 Inelastic scattering of photons

Raman shifts involves a change in the energy of a photon from a polarized laser light source (UV or visible light).[84] The quantum approach describes the Raman effect as collision of photons with frequency v_0 and an equivalent discrete energy of hv_0, where h represents the Planck's constant

$h = 6.626 \times 10^{-34}$ J s. The frequency of its electromagnetic radiation is given by

Eq. 1

$$E = hv_0$$

where E is the energy of a photon given in joules and v_0 is the frequency given in hertz.[46]

The relation between the frequency of a photon and the wavelength is

Eq. 2

$$v = \frac{c}{\lambda}$$

where v is the frequency expressed in $\frac{1}{s}$, when c is the speed of light ($3 \times 10^{10} \frac{cm}{s}$) and λ is the wavelength given in cm. The most common spectroscopic parameter, the wavenumber, is given by

Eq. 3

$$\tilde{v} = \frac{v}{c} = \frac{1}{\lambda}$$

where the wavenumber units are given in cm^{-1}. By combining equations Eq. 1, Eq. 2 and Eq. 3, the mathematical expression of the energy of the incoming photon can be obtained as

Eq. 4

$$E = hv_0 = h\frac{c}{\lambda} = hc\tilde{v}_0.$$

Since the Raman effect can be related to the change in wavelength of the incident and scattered photon, the Raman shift Δv can equivalently be described as

Eq. 5

$$\Delta v(cm^{-1}) = \left(\frac{1}{\lambda_0} - \frac{1}{\lambda_i}\right) 10^7$$

where λ_0 and λ_i are the wavelengths of the incident and scattered photons in nm.[46]

1.2.3 Polarizability of molecules

According to the classical description, the atoms of any molecule are connected through chemical bonds which can vibrate. The vibrational frequency of the bond can be compared to a harmonic oscillator.[46] The quantum description of a vibrational energy level however is more precisely described as

Eq. 6

$$E_v = \left(\frac{h}{2\pi}\right)\sqrt{\frac{k}{m}}(v + \frac{1}{2})$$

where h is Planck's constant, m is the reduced mass and v is the quantum number of the vibrational energy state. The polarizability of a molecule is the ability of the incident electric field to generate an induced dipole moment in the molecule and can be written as

Eq. 7

$$\vec{P} = \vec{\vec{\alpha}}\vec{E}$$

where $\vec{P}$ is the induced dipole moment, and $\vec{\vec{\alpha}}$ is the polarizability tensor interacting with the electric field $\vec{E}$.[46] A vibrating molecule can change its polarizability which induces a dipole that allows the molecule to express a Raman scattered signal.[40,46,85] For a vibrational mode to be Raman active, the polarizability of the molecule $\vec{P}$ needs to change by the interaction of the polarizability tensor $\vec{\vec{\alpha}}$ and $\vec{E}$.

The polarizability α depends on the type of chemical bonding and on the electronic structure. In particular, the polarizability of anisotropic molecules can vary with position and interatomic distances, mainly depending on the symmetry of the molecule.

The electric field strength E of an electromagnetic wave (here: a laser light) varies according to

Eq. 8

$$E = E_0\cos{(2\pi v_0 t)}$$

where v_0 is the frequency of the electromagnetic radiation, and t is time.[46] From equations Eq. 7 and Eq. 8, the time dependent induced dipole moment $\vec{P}$ is

Eq. 9

$$\vec{P} = \vec{\vec{\alpha}}\overrightarrow{E_0}\cos(2\pi v_0 t)$$

With the assumption that a molecule is set in its equilibrium position, a perturbation in the electronic cloud will cause changes in the relative positions of the atoms, resulting in a change in the polarizability caused by the the electromagnetic wave.

If molecular vibrations are defined as the displacements of an atom as Q, then the displacement of each atom from the equilibrium position is given by

Eq. 10

$$Q = Q_0\cos(2\pi v_i t)$$

where Q_0 is the vibrational amplitude and v_i is the molecule's vibrating frequency. The polarizability α can then be written as a linear function of Q.[46] Thus

Eq. 11

$$\alpha = \alpha_0 + (\partial\alpha/\partial Q)_0 \times Q_0$$

For a specific vibrational mode to be Raman active, the change of the polarizability of the molecule and thus $(\partial\alpha/\partial Q)_0$ must be non-zero.[46] If the derivative tends to zero, then the vibrational mode does not alter the polarizability of the molecule significantly, and the intensity of the particular mode will be weak.

1.2.4 Raman intensity

For a Raman signal to be measurable, it must be greater than the dark current of the CCD detector. The Raman intensity is also dependent on the qualitative ability of a chemical bond as Raman scatterer. The Raman intensity is given by

$$I_R = I_0 \sigma_i D d_z \qquad \text{Eq. 12}$$

where I_R is the Raman intensity and I_0 is the laser intensity respectively in Watts.[46] The cross section σ_i is proportional to the probability of the laser photon being inelastically scattered and is strongly dependent on the change of polarizability. D is defined as the density of scatters and d_z is defined as the path length of the laser light within the investigated material. Other contributing factors involve the laser wavelength, the collection geometry and polarization. Higher wavelength lasers improve spectral resolution but at the expense of the peak intensity. The Raman efficiency is given by

$$Eff \propto \frac{1}{\lambda^4} \qquad \text{Eq. 13}$$

where Eff is the Raman efficiency inversely proportional to the laser wavelength, meaning that lower wavelengths provide higher sensitivity.[46]

2.7.3 Spatial resolution

The lateral spatial resolution (laser spot diameter) is

$$1.22\ \lambda\ /\ \mathrm{NA} = 0.7\ \mu\mathrm{m}.^{86} \qquad \text{Eq. 14}$$

The theoretical diffraction limit of the lateral (XY) spatial resolution of the optical microscope is determined by the laser wavelength and the NA of the objective, where

$$0.61\ \lambda\ /\ \mathrm{NA} = 0.3\ \mu\mathrm{m}. \qquad \text{Eq. 15}$$

The confocal depth (Z) spatial resolution however requires experimental determination by Raman mapping. The depth resolution is dependent on instrumental parameters as well as sample parameters, and sometimes interactions between the sample and the optics.[86] The instrumental depth resolution is sometimes approximated to be

$$4\ \lambda\ \mathrm{NA} = \sim 1.8\ \mu\mathrm{m}.^{87} \qquad \text{Eq. 16}$$

The optical processes during Raman microscopy however are more complex, wherefore the spatial resolution is often quoted to be in the order of 1 μm in diameter and 2 μm in depth.[87] Under the assumption of celery cellulose microfibrils being 3.2 nm in depth x 3.0 nm wide and comprising about 21 chains[66], each point measurement will sample approximately $\frac{2.0\ \mu\mathrm{m}}{3.2\ nm} \times \frac{1.0\ \mu\mathrm{m}}{3.0\ nm} = 625 \times 333 = 208333$ microfibrils or over 4.4 million chains.

1.2.5 Molecular vibrations

In a molecule of N atoms, each atom can move in all three dimensions (x, y and z), resulting in 3N degrees of freedom of motion. Along the three principal dimensions of rotation (x, y and z), three motions can be related to the translational mode of the entire molecule and three motions can be related to rotational modes of the entire molecule. As a result, the remaining degrees of freedom which relate to the vibrational motions in a molecule of N atoms is 3N – 6 for non-linear molecules. For linear molecules, there are 3N-5 vibrational degrees of freedom as rotation is not occurring along the molecular axis.[46] The frequency of molecular vibrations and the relative motions of a molecule are determined by its geometry.[88] Raman spectroscopy spectra will be specific for any molecule and its vibrational environment, providing a unique "fingerprint" of a sample on the molecular level. Figure 1.10 shows a schematic representation of the vibrational (a), rotational/translational (b) modes for a triatomic molecule.[46]

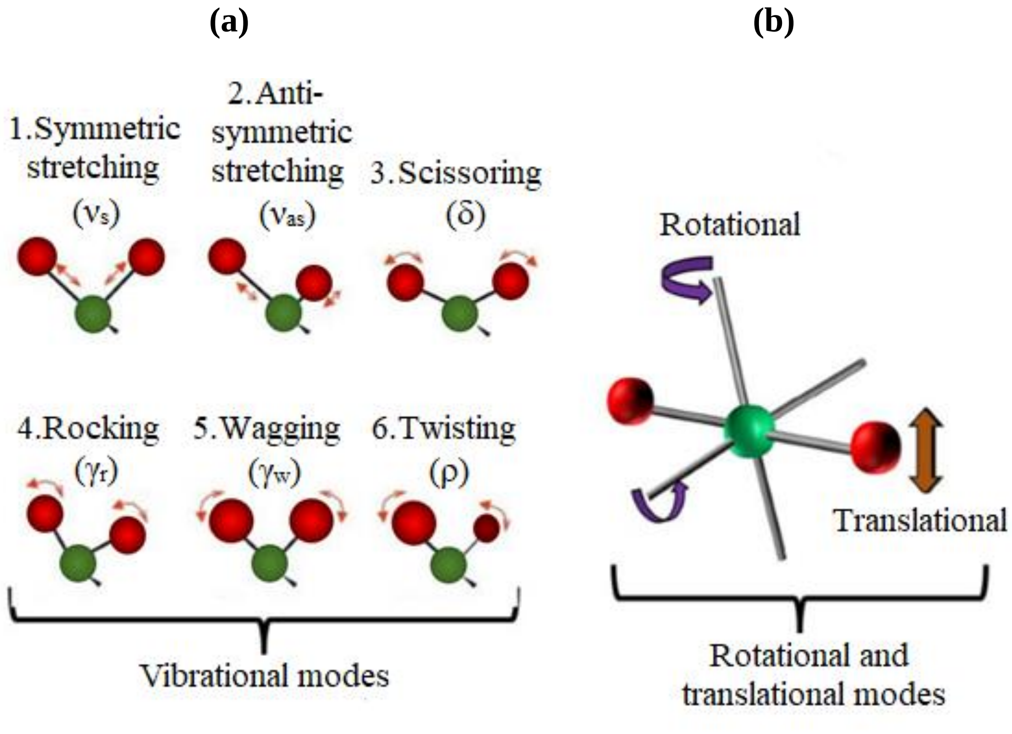


Figure 1. 10: The vibrational, rotational and translational modes in molecules. The arrows show the direction of motion of atoms in, above and below the plane, respectively. (a) Vibrational modes are shown in a 2D representation. Stretching modes can be subcategorized to symmetric and antisymmetric stretches; scissoring, rocking, wagging and twisting represent bending modes. (b) Rotational and translational modes are shown in a 3D representation (Modified from Ferraro et al., 2003).[46]

1.2.6 Polarization dependent Raman spectroscopy

While Raman spectroscopy provides information about the chemical composition of a sample under standard conditions, polarized Raman spectroscopy can provide further information, for example about the symmetry of vibrational modes and the orientation of the sample. Cellulose is a crystalline, polymeric sample, which means that it is anisotropic hence directional. This sample property is a necessary requirement to perform polarisation dependent Raman spectroscopy: The molecules have fixed positions and the molecular bonds can be aligned in a certain orientation in respect to the laser, resulting in very different spectra that can provide useful information about the crystal lattice.[89,90] Against the polarizability of molecules and the Raman intensity introduced in previous chapters, it can be summarized, that the direction of polarizability (the induced dipole moment) in relation to the polarised laser (and the electric field it produces) is proportional to the Raman intensity. As a result, molecular bonds that are oriented parallel to the polarization of the laser are more intense than bonds, that are oriented anti-parallel. The spectra obtained by most Raman spectrometers are represented on the x-axis as the Raman shift in wavenumber (cm^{-1}) versus intensity (counts or counts/second) on the y-axis.

1.3 Raman spectroscopic investigations on highly oriented cellulose

Although all cellulosic substrates have a similar chemical composition, often they have distinct properties such as size, morphology, crystal structure and crystallinity which result in a characteristic Raman spectrum (molecular fingerprint). Though Raman spectra of various cellulose substrates have been published, only one study has included the Raman analysis of celery cellulose fibers.[91] The first part of this work aims to investigate the full Raman spectrum of the natural cellulose substrate isolated from celery collenchyma which will be used for studies on enzymatic activity with univariate and multivariate analyses. The first step of this project was the analysis of the cellulosic substrate and the study of the intensity of the Raman spectrum as a function of polarization and orientation of the unmodified cellulose fibrils. Subsequently, the impact of microfibril orientation on the relative intensities of Raman bands and the directional character of most vibrational bonds contributing to a spectral feature will be investigated. Figure 1.11 shows the Raman signature of oriented Ramie fibers, which are almost perfectly parallel to the fiber axis and consist of almost pure cellulose. Ramie is a weakly branched, perennial, herbaceous plant. It is one of the oldest fiber crops and has been used for fabric production.[92]

Changing the fiber orientation of a highly oriented sample of crystalline cellulose leads to characteristic peak changes in the Raman intensity of almost all Raman bands, except for the two bands at ~1375 and 437 cm^{-1}, as they involves various bending modes of hydrogen, carbon and oxygen atoms.[32,44] The Raman bands at 1377 cm^{-1} was therefore used as reference band for peak intensity ratio (PIR) analyses.

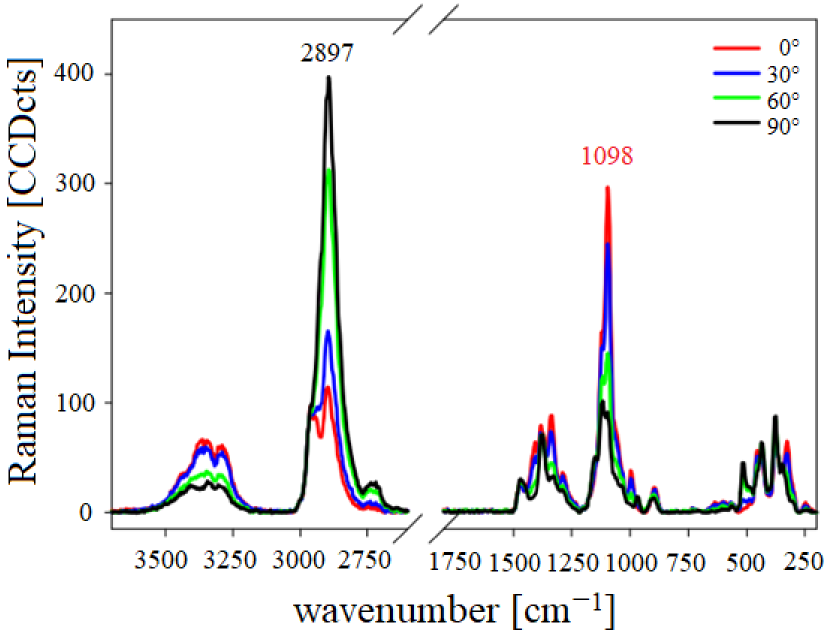


Figure 1. 11: : An example of Raman spectra of oriented Ramie fiber. The spectra show changes depending on the polarization direction of the incident laser from 0° (parallel to the electric vector of the laser) to 90° (perpendicular to the electric vector of the laser) in increments of 30. The spectra parameters are: 532 nm excitation, 0.25 s integration time and 10 accumulations, baseline corrected. Fiber parameters are >95% cellulose, high crystallinity and microfibrils aligned parallel to the fiber axis. Image free for unrestricted use (taken from Gierlinger et al., 2013).[32]

Whereas IR absorption arises by a change in the dipole moment, Raman scattering is caused by a change in the polarizability of a molecule by an incident electromagnetic wave. The Raman polarizability (peak intensity) depends on the angle of orientation between the molecular bond and the of the orientation of the polarized laser (electric vector). Specific chemical bonds show higher intensities when their arrangement is parallel to the electric vector of the incident laser light (0°). Because of the orientation-dependency, the cellulose fiber direction and the laser polarization have to be adjusted in a known and defined way in every Raman experiment.[32]

As shown in Figure 1.11, cellulose has broad Raman contribution located around 200 – 1500 cm^{-1} (fingerprint region) and 2500 – 3700 cm^{-1} (high wavenumber region). Normal coordinate analyses concluded that 60-80% of most individual Raman peaks in the cellulose spectrum contain combinatory vibrational contributions from several chemical groups.[93,94] Despite having the same chemical groups, every cellulose sample has a unique Raman signature due to the coherent vibrations of intra- and interspecific chains.[95] The following band assignment is typical for Raman spectra of celluloses: Previous normal coordinate calculations revealed the 250-950 cm^{-1} region to mainly contain vibrational contributions from so-called "heavy atoms" (involving C and O).[96] The heavy atom bending below 600 cm^{-1} involves skeleton CCC, COC, OCC and OCO bonds. The region above 800 cm^{-1} and below 950 cm^{-1} involves heavy atom angle bending, ring and CO stretches, as well as external bending modes of methylene groups. At 950 cm^{-1} to 1150 cm^{-1}, the most significant contributions of Raman bands involve ring stretching together with COC stretching motions. Angle bending of hydrogen atoms coupled to a heavy atom (methine bending, methylene bending and COH in-plane bending) are located between 1200-1450 cm^{-1}. The high wavenumber region above 2500 cm^{-1} involves characteristic group stretching modes such as CH stretches ~2850 cm^{-1} and OH stretches ~3500 cm^{-1}.[97] It is noticeable that the stretching modes are substantially more Raman active than the bending modes[98].

The peaks around 1100 cm^{-1} are widely considered to be unique Raman peaks for the glycosidic bond stretches in cellulose.[96] More precisely, the signature Raman bands at 1095 cm^{-1} (symmetric COC stretch) and 1120 cm^{-1} (asymmetric COC stretch) are commonly used as natural biomarkers in Raman imaging.[99] Other components (such as e.g. hemicelluloses) have very small contributions to these peaks.[100] It is therefore assumed that these Raman bands will provide a straightforward approach to contrast underlying mechanisms of enzymatic action on cellulose. The 1500 - 1800 cm^{-1} region does not contain any structural peaks in the Raman spectrum of native celluloses.[101] The oxidatively introduced carbonyl bonds (C=O) in LPMO treated cellulose are expected to appear at ~1700 cm^{-1}. This region does not contain any peaks in the native cellulose spectrum.

In this project, polarisation dependent confocal Raman spectroscopy has been used as the Raman spectroscopic method to investigate celery cellulose. Previous polarized Raman studies on celery fibers have been focused on the orientation of the fibrillar structure while applying micromechanical strain and these spectra have not been published in their entirety.[91] To our knowledge, the application of polarised Raman spectroscopy on highly oriented cellulose fibers for enzymatic degradation studies has not yet been exploited. Each average Raman spectrum comprises 50 – 150 individual Raman spectra. This can lead to difficulties in extracting hidden information, wherefore univariate peak intensity analysis (PIR) in combination with multivariate Principal component analysis (PCA) will be performed.[47]

1.4 Enzymatic cellulose deconstruction

Numerous enzymes are involved in the degradation of plant biomass. The following section is confined to enzymes involved in cellulose degradation and describes their chemical mode of action.

1.4.1 Cellulases

More than 130 glycoside hydrolases (GH) families have been classified that employ inverting or retaining hydrolysis mechanisms to cleave glycosidic linkages.[102] Enzymatic degradation of polysaccharides was traditionally thought to occur through the synergistic action of classical glycoside hydrolases. GHs can be divided into three main categories: endo-1,4-β-glucanases (EGs) that create new chain ends by cleaving internal glycosidic bonds within the amorphous region of cellulose; cellobiohydrolases (CBHs, also known as processive exocellulases[103]), which hydrolyses cellulose in a processive manner generating disaccharide units (cellobiose); and β-glucosidases which convert the disaccharide units into glucose.

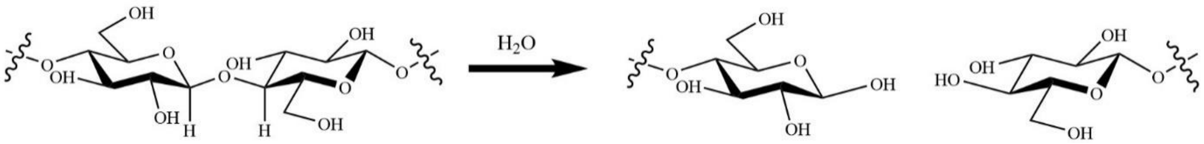


Figure 1. 12: General mechanism of cellulose degradation by cellulases.

Cellulolytic enzymes (GHs) hydrolyse glycosidic bonds via acid catalysis, which requires a proton donor and a nucleophile/base. Hydrolysis occurs via two main mechanisms, giving rise to either the retention or inversion of the anomeric configuration.

The products formed by insertion of water are essentially shown in the figure below.

The structural basis of this behaviour is a remarkable substrate binding region involving up to 10 pyranose sub/sites located in a groove or tunnel (1 and 2).[104,105] Cel7A from *Trichoderma reesei* is the most studied CBH.[106] In 1998, Divne and co-workers reported the first high-resolution crystal structure of Cel7A with bound substrate (see Figure 1.13). They noted that the intensity of enzyme substrate interactions increased towards the product site.[104] It has also been shown experimentally, that the product site also exhibits an anomeric selectivity, which influences the product profile.[107]

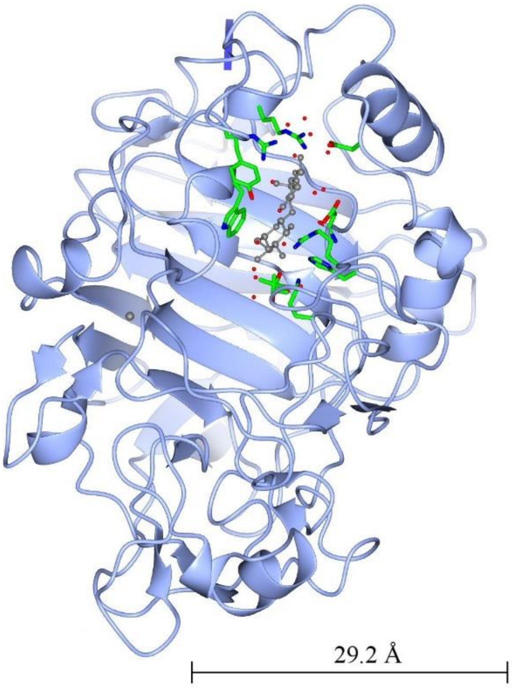


Figure 1. 13: Enzymatic structure of Cel7A (PDB 4V20). Image generated in CCP4MG.

The resulting idea is, that a characteristic product profile might lead to a characteristic pattern of cellulose degradation might lead to distinguishable features in the Raman spectra of degraded cellulose, that can be used as ‘markers’ for these kinds of enzymatic actions.

In this study, cellulases Cel7A and Cel7B have been used to study cellulose modifications on highly oriented cellulose. Three dimensional structures of cellulase Cel7A (CBH I) have shown that the active sites of CBHs are located inside the tunnel, and Cel7A acts in a progressive manner from the reducing end of cellulose.[108] The action of CBHs draws the cellulose chain away from its neighbouring chains, causing chain disruptions. Moreover, CBHs perform multiple hydrolysis reactions simultaneously without dissociating from the substrate.[109]

Cel7A (PDB 4V20) is a key industrial cellulase and a commonly used model system for fungal cellulases, disrupting cellulose chains in a processive manner.[108] Cel7A is a cellobiohydrolases.

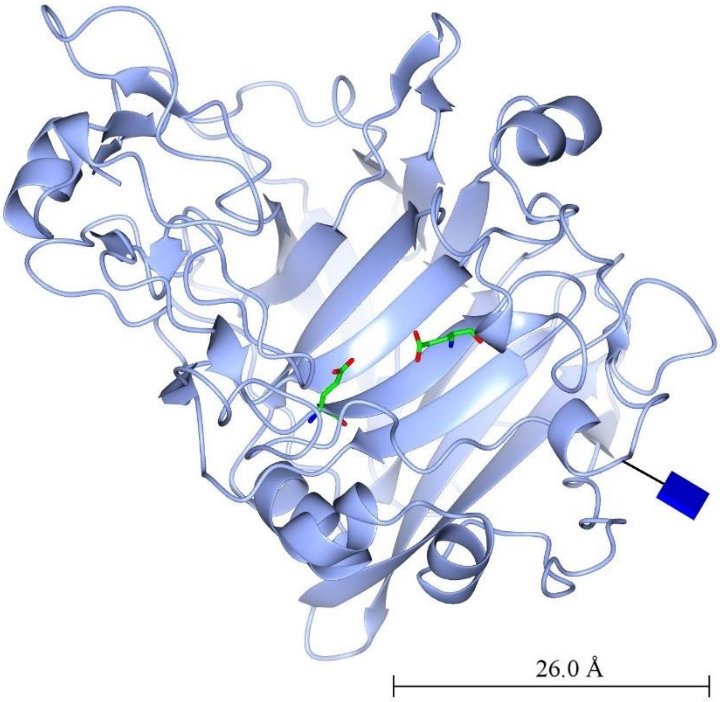


Figure 1. 14: Enzymatic structure of Cel7B (PDB 1A39). Image generated in CCP4MG.

Cel7B (PDB 1A39), which is an endo-1,4 glucanase, however is only active on paracrystalline (amorphous) cellulose regions.[110] Cel7B is a retaining beta-glycosidase hydrolase that serves a biological role as part of the secreted cellulolytic apparatus of the fungus.[111]

.

1.4.2 Lytic polysaccharide monooxygenases

A significant industrial progress has been made with the discovery of Lytic Polysaccharide Monooxygenases (LPMOs), which have since been introduced into commercial cellulolytic enzyme cocktails to ‘boost’ biological biomass degradation.[14,112] With the discovery of lytic polysaccharide monooxygenases (LPMOs) in 2010, a new class of oxidative enzymes were discovered that introduce chain breaks in glycosidic linkages in polysaccharides oxidatively and increase the rate of polysaccharide conversion in conjunction with conventional glycoside hydrolases such as classical cellulases.[14,11224] LPMOs are copper containing metalloenzymes which cleave polysaccharides or oligosaccharides at position C1 and/or C4 by activation of O_2.[113] Over 5700 known LPMO sequences are listed in the Carbohydrate Active enzyme (CAZy) database, grouped in seven different classes of “Auxiliary Activity” enzymes (AA) followed by an identification number, i.e. AA9-AA11, AA13-AA16.[114–118] LPMOs share a slightly distorted Immunoglobulin-like beta-sandwich core structure of two β- sheets comprising seven or eight β-strands. The helices and loops connecting the core β-strands are responsible for their structural diversity. The longer loops vary greatly among the different LPMOs and play a key role in substrate recognition.[119]

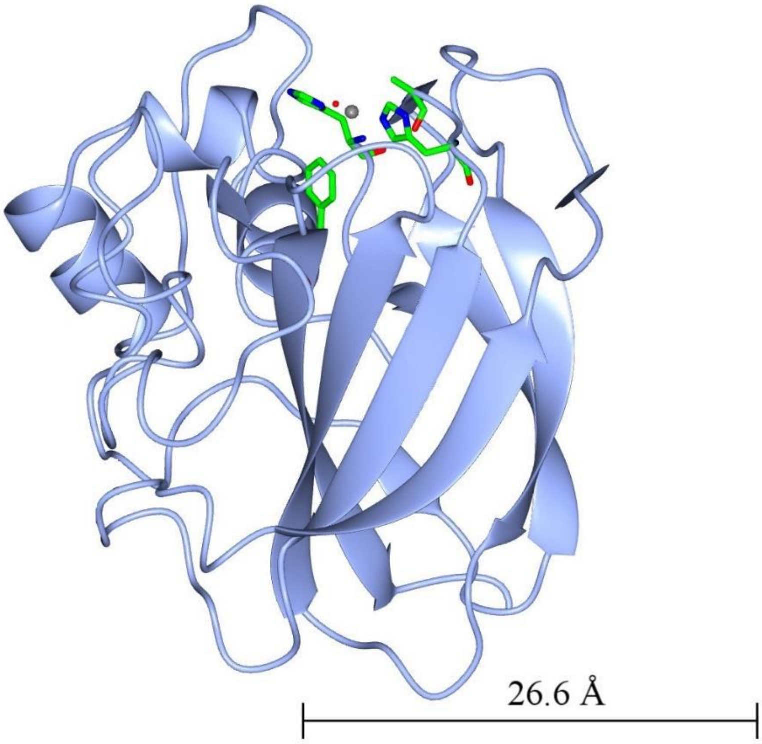


Figure 1. 15: Enzymatic structure of Cel7A (PDB 4OY7). Image generated in CCP4MG.

The active site of LPMOs contains a copper ion coordinated by an N-terminal histidine and an amine from a side chain of a second histidine. The three provided nitrogen ligands generate a T-shaped configuration to the copper ion known as the histidine brace.[113]

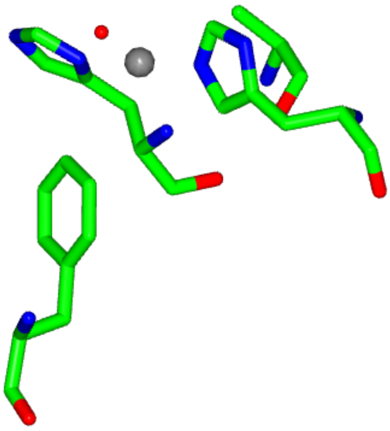


Figure 1. 16: Histidine Brace. Active site of ScAA10. Image generated in CCP4MG

The active site of LPMOs is exposed to the flat protein surface, where it can be in direct contact with the polysaccharide substrate. The catalytic cycle is believed to be initiated by the reduction of LPMO-Cu^{II} to LPMO-Cu^{I} by an external electron donor.[120] In this study Na-ascorbate was used, which is a common reducing agent used for in vitro experiments. The catalytic cycle leads to the oxidation of the C-H bonds (~100 kcal/mol).[120] In the study presented within this thesis, Na-ascorbate was used as external electron donor Cellulose active LPMOs include strict C1-oxidisers, strict C4-oxidisers and mixed C1/C4 oxidisers, although also C6 active LPMOs have been discussed.[121] LPMOs are copper containing enzymes that depolymerize polysaccharides through an oxidative mechanism involving regio- and stereoselective hydroxylation of cellulose at the C1 or C4 carbon to C=O. This subsequently cleaves the glycosidic bond and increases the amount of chain termini in the structure.

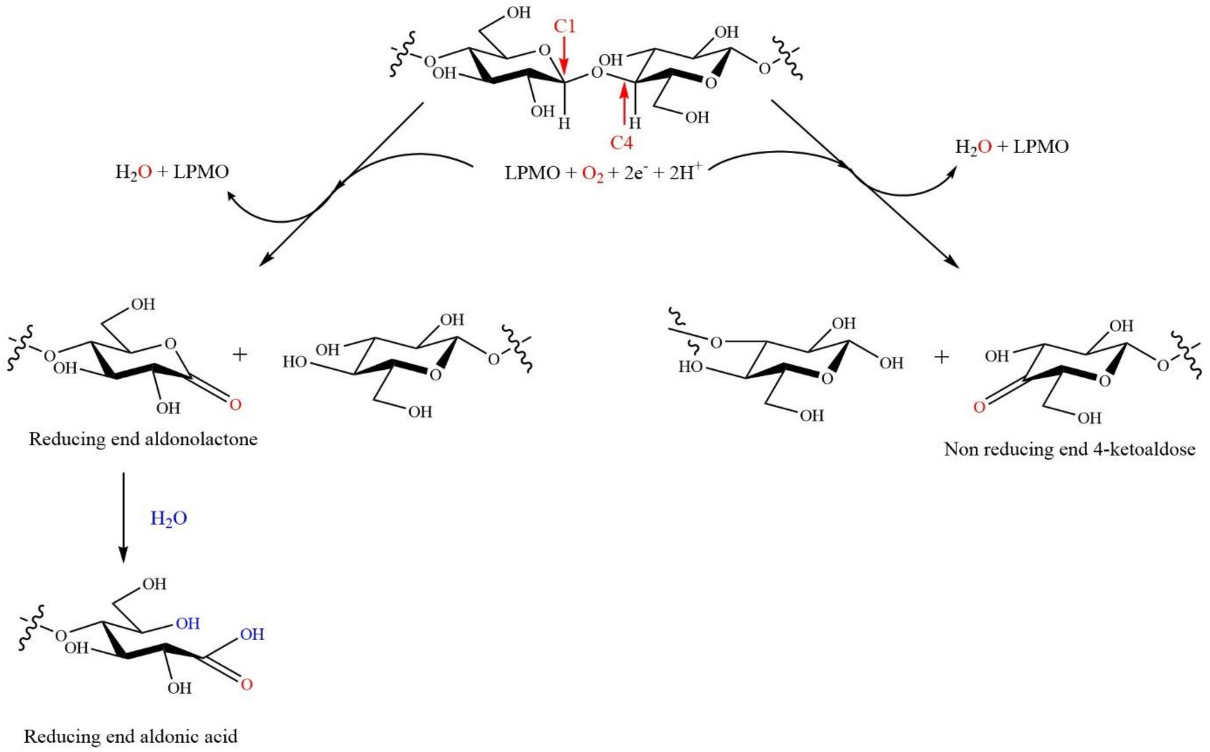


Figure 1. 17: General mechanism of cellulose degradation by LPMOs.

LPMO action results either in an aldonic acid at the reducing end of glucose molecules (C1 action), or a 4-keto-aldose at the non-reducing end (C4 action), depending on the activity of the enzyme.[113,122] It should be noted that LPMOs are not considered to decrystallise polymers to the same extent as cellulases.[122] In this study, LPMO *Sc*AA10 has been used to study cellulose modifications on highly oriented cellulose. The literature states *Sc*AA10 also known as CelS2 (PDB 4OY7) has C1 activity on glycosidic bonds of crystalline cellulose.[123]

1.5 Research Objectives

Drawing on previous work in the field of Raman studies on cellulose, this project seeks to determine the effect of enzyme action on highly oriented cellulose. A main goal of the Raman spectroscopic analysis is to identify Raman peaks that are significant for cellulose degradation and show how these spectral features can be correlated to the substrates' structural information. For this purpose, one LPMO *Sc*AA10 (also known as CelS2) and two cellulases Cel7A and Cel7B have been used. The enzymes were used to degrade enzymes and washed out to analyse the cellulose ex situ after enzymatic treatment. The original contribution of this study is the implementation of a new analytical tool to quantitatively assay enzymatic activity on semi-crystalline cellulose with polarization dependent Raman spectroscopy. The key hypothesis underpinning the purpose of this study is that the enzymatic modifications (C=O and OH bonds, as well as chain cleavages) of highly oriented cellulosic substrate is a newly introduced bond into the cellulose structure that can be detected and quantified by Raman spectroscopy with univariate and multivariate analysis. Moreover, changes in the crystalline surface structure should be detectable and quantifiable. In the work described within this thesis, two known cellulases (Cel7A, Cel7B) and one LPMO (*Sc*AA10, also known as CelS2) were used to investigate the effects of enzyme action on oriented cellulose with Raman spectroscopy.

All enzymes differ in their chemical mode of action. The hypothesis of this thesis is to identify 'patterns' within the Raman spectra of degraded cellulose that allow to classify and characterize these individual modes of action. To facilitate this task, two techniques of spectral analysis were used: peak intensity ratio (PIR) and principal component analysis (PCA). PIR analysis allows to compare relative peak intensities across samples, whereas PCA is used as a pattern recognition procedure.[124]

1.6 Outline of the thesis

The thesis comprises two main studies: a polarization dependent study on celery cellulose and an enzymatic study. In chapter 3, a comprehensive polarisation dependent Raman study on reference spectra of naturally highly oriented and semi-crystalline celery cellulose was performed. The purpose of this research work was to characterize the Raman signature of celery cellulose as it represents the sample background of the Raman-based assay. Chapter 4 investigates celery cellulose after enzymatic treatment with polarisation dependent Raman spectroscopy. Chapter 5 contains the discussion and conclusion.

2. Materials and Methods

2.1 Semi-crystalline cellulose fibers

The fibers used for this study were from celery, which was bought from the local supermarket (Waitrose). Well-oriented, semi-crystalline cellulose fibers from cell wall tissue (collenchyma) were isolated and prepared according to protocols from the literature.[56,60,125] Briefly, the fibers were pulled out by hand then placed into acetone for 5 min, rinsed with deionised water for 5 min and washed with 1 g/L Tween 20 in H_2O for 5 min. Tween 20 is a detergent (surfactant) that helps dissolve the plasma membrane and releases membrane proteins. The fibers were then treated with 1 mol/L HCl at 100 °C for 1 h followed by a mild alkali extraction in 0.1 mol/L NaOH for 30 min and lastly titrated to pH 6.5 in 20% (v:v) acetic acid. This procedure allows the removal of non-cellulosic polysaccharides.The fibers were then dried on a glass petri dish at room temperature.

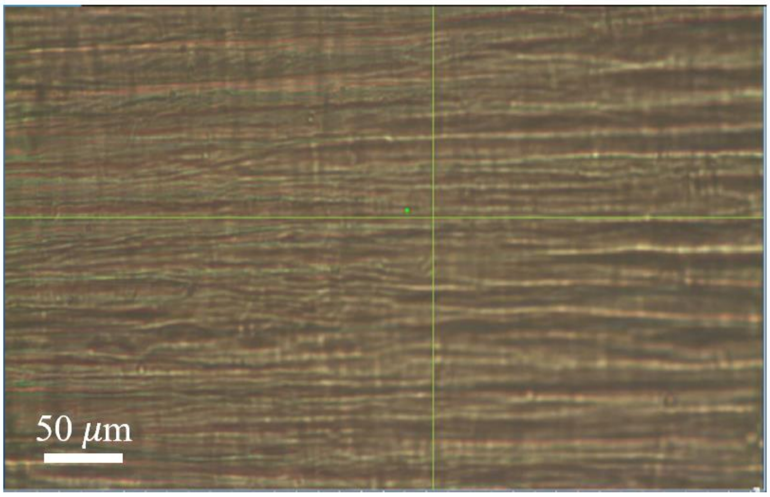


Figure 2. 1: x50 optical microscope image of crystalline cellulose fiber from celery, arranged horizontally.

2.2 Enzymes

*Sc*AA10 (also known as *CelS2*) was used for the enzymatic degradation of semi-crystalline cellulose by LPMOs. *Sc*AA10 (PDB 4OY7) was prepared by Dr. Hemsworth through heterologous expression in *E.coli* and provided by Dr. Urresti, University of York. Briefly, the protein was expressed in BL21 (DE) *E. coli*. Over night culture (ONC) was started with 15 mL LB medium at 37 °C and 180 rpm and transferred in 500 mL LB (inoculation with 5 mL ONC). The cell growth (37 °C and 180 rpm) was monitored by measuring the suspensions optical density at 600 nm (OD_{600}). At a concentration of 1 mM, (OD 0.6 – 0.8), Isopropyl β-D-1-thiogalactopyranoside (IPTG) was added which induces protein expression. The cell culture was left overnight at 16 °C and 180 rpm. The protein was then isolated from the periplasm and purified by cation exchange chromatography through a 5-ml HiTrap SP HP (GE Healthcare) column equilibrated in 50 mM sodium acetate, pH 5. The protein was then purified by size-exclusion chromatography on a HiLoad 26/60 Superdex 75 column (GE Healthcare) in 20 mM sodium acetate, pH 5.0, 250 mM NaCl. Peak fractions were pooled and concentrated by centrifugation on a Sartorius 10 kDa molecular weight cut off concentrator.[114]

CBH Cel7A (PDB 4V20) from *Asperillus fumigatus* and Cel7B (PDB 1A39) from *Humicola insolens* were expressed in *A. oryzae* strain JaL250 and purifeid as described previously.[111,126] Both cellulases were provided by Dr. McGregor, University of York. [126]

2.3 Enzymatic (Biochemical) Reactions

The experiments were conducted in triplicate at 38 °C. Cellulose fibers were cut in same-sized pieces of ~0.7 cm. An enzyme concentration of 1 μM and 20 mM MES buffer, pH 6.5, was used in a total reaction volume of 300 μL in 1.5 mL Eppendorf tubes. The concentration of reducing agent for LPMO treated replicates was 1 mM Na-ascorbate. The Eppendorf tubes, which were used as reaction vessels, were opened every 24 h to ensure aerobic conditions. At the same time, reducing agent was added to the respective samples to counteract depletion (every 24 h). Negative controls for S*c*AA10 contained either the LPMO enzyme without reducing agent, or the reducing agent without the enzyme. The LPMO enzyme is considered to be inactive without reducing agent. Negative controls for Cel7A and Cel7B only contained buffer. The samples were placed in a hybridization incubator to ensure thorough mixing. After

72 h, the remains of the cellulose fibers were withdrawn and washed thoroughly with deionized water. The washing step was repeated three times and the samples were dried in 1.5 mL Eppendorf tubes at room temperature before Raman analysis. The supernatant was subject to sugar analysis with MALDI-TOF Mass spectrometry as described in Vaaje-Kolstad *et al.*.[112]

2.4 MALDI-TOF spectrometry

The analysis of liquid fractions was performed by mass spectrometry for verification of enzyme activity. Briefly, 9-10 mg/ml 2,5-dihydroxybenzoic acid in 50% acetonitrile with 0.1% trifluoroacetic acid (TFA) was used as matrix. A 1:1 concentration of 1μL sample to matrix was pipetted and dried on a SCOUT-MTP 384 target plate (Bruker) and analysed on an Ultraflex III instrument (Bruker) by matrix-assisted laser desorption ionization-time of flight (MALDI-TOF), as described in literature.[112]

2.5 Acquisition of IR spectra (ATR-IR spectroscopy)

The IR beam enters the reflecting ATR crystal at an angle of typically 45° and is several times reflected at the crystal to sample interface. A fraction of light known as evanescent wave reaches into the sample. The evanescent wave is attenuated in the spectral regions where the sample absorbs energy, and the IR beam exits the ATR crystal where it is directed to the IR-detector.[127] The ATR-IR spectra of native cellulose were collected at the range 4000 – 250 cm^{-1}.

2.6 Raman experimental measurements

Dry cellulose microfibrils were rotated along the x-y plane to investigate the dependency between cellulose fibril orientation and direction of polarization. The incident electromagnetic wave transverses the z axis and is y- polarized. For Raman measurements, the fibers were taped to a CaF2 slide with minimum tension. Figure 2.1 shows the orientation of the celery fiber relative to the laser light.

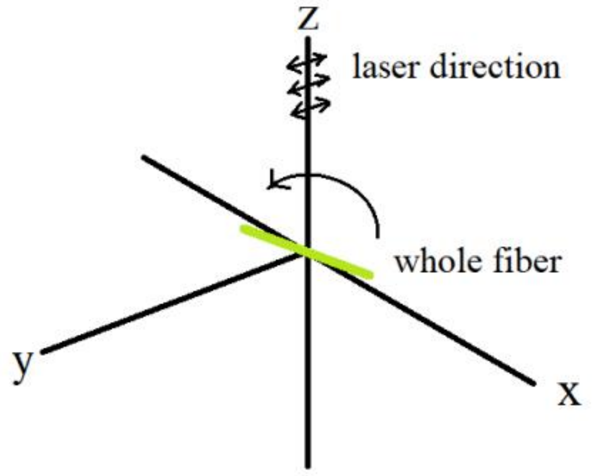


Figure 2. 2: Experimental setup for Raman measurements. Shown in green is the position of the celery fiber which was oriented along the x-y plane of the microscopy stage. The laser source is mounted on top of the stage, polarisation direction along the y-axis.

2.7 Raman Parameters

A HORIBA XploRA micro-Raman instrument equipped with a 100x objective and numerical aperture (NA) of 0.9 was used. The 100x objective lens was used to gain a smaller spot size and acquire Raman spectra of good quality. The cellulose spectra were collected with a laser wavelength of λ = 532 nm (green) with a 90 s integration time, two accumulations and three spectral windows. The spectral window is governed by the dispersion and the chip width of the CCD detector. Including the integration time and the windows, which requires the spectrometer to move to build up the spectrum, the total collection time for an individual Raman spectrum was 15 minutes. The laser filter was set to 25%, resulting in 1.9 mW power at the sample. All data collection was monitored real-time to prevent laser-induced damage to the samples. A greater slit width than 200 μm slit setting decreases the spectral resolution. The Raman light propagates through the entrance slit and grating mirror via a second mirror onto the Andor CCD detector.[128] Figure 2.2 shows a typical set up of a confocal Raman microscope. To achieve high spectral resolution, the highest density of 2400 lines/mm diffraction grating was used. The confocal pinhole enables the rejection of out-of-focal-plane signals and was set to 100 μm resolution depth. A greater resolution depth allows more than one focal plane of the sample to be measured.[129] Confocal settings were used meaning the spectra were only taken from single focal planes. These are optimized settings that allowed high spectral quality. The Raman spectra were measured at 200 cm^{-1} – 1800 cm^{-1} (fingerprint region) and 2500 cm^{-1} – 3700 cm^{-1} (high wavenumber region) at room temperature. As part of the Raman measurement acquisition, two types of corrections were performed: Intensity correction system (ICS) and dark correction. The ICS correction factor allows to yield the same results with different Raman acquisition settings. The dark current correction subtracts a pre-determined dark file from the spectrum to correct the detector response and pixel variation of a spectrum.

2.7.1 Laser wavelength

The Horiba XploRA Raman microscope comprises two different laser excitations: 532 nm (green) and 785 nm (near-IR). Near-IR lasers are most often used for lignocellulosic materials as they reduce fluorescence interference.[38,40] Moreover, longer laser wavelengths increase laser spot size.[86] Green and blue lasers represent the best compromise between Raman efficiency and analysed volume for crystalline cellulose materials.

2.7.2 Spectral resolution

A higher number of lines/mm in the spectral grating results in a greater spectral resolution. The grating disperses the scattered light onto the CCD detector (1650x200 pixels) by deflecting each wavelength at a different angle to generate the Raman spectrum. Besides the diffraction grating and the detector, the spectral resolution is also dependent on the laser wavelength and the spectrometer focal length. The focal length is the distance between the dispersing grating and the detector. The focal length is 200 mm for the HORIBA XploRA Raman microscope (low/medium resolution).[128] Against this background, the spectral resolution of the instrument is limited to 1 – 1.5 cm^{-1}.

2.8 Peak fitting and fitting uncertainties

The spectra were linear baseline corrected, total area normalized and 0.65 cubic spline smoothed in *Raman Tool Set* before curve fitting.[130] The fitting was performed in Igor Pro.[131] To ensure reproducibility, the peaks were fitted manually either individually or in small sections across the spectrum. Igor provides Gaussian, Lorentzian and Voigt fitting profiles. Gaussian fittings were applied for cellulose spectra as the samples within this study were dry cellulose fibers (solid residues). This facilitates statistical analyses as many physical and chemical processes are governed by Gaussian statistics.[132] The fitting uncertainties were measured by fitting the average wave ± the calculated standard error (SE). The maximum variance for each peak was represented as its fitting uncertainty.

2.9 Peak intensity ratio analysis

Peak positions were obtained from peak fittings in Igor Pro.[131] Peak-intensity ratios were calculated with respect to the peak at 1375 cm^{-1}, which comprises different hydrogen, oxygen and carbon containing bending modes and is considered to be independent of polarization angle.[133]

2.10 Statistical measurements

The number of Raman spectra needed to represent the average structure of a given cellulose sample reflects its optical heterogeneity. The optical heterogeneity (the scattered light intensity) is influenced by the chemical composition and structural properties (topology) of the sample.[48] The Raman signatures collected from one point in the sample to the next may vary. Raman spectra vary in peak intensity and spectral background, therefore statistical tests have been developed to determine the number of spectra required to fully characterize a cellulose sample. Initially, a pilot dataset of approximately 150 individual Raman spectra of dry crystalline cellulose before enzymatic treatment was acquired. From this dataset, a convergence test was performed. The convergence test compares the average waves, standard error (SE) of the mean and second order standard deviation (2SD) of a set of baseline corrected spectra. The individual Raman spectra were analysed by increasing the amount of randomly selected data in increments of 5 and comparing the observed differences. If these differences are sufficiently low (<1.7%), the average spectrum was defined as “converged”. The spectra are included in the appendix S1

2.11 Principal component analysis (PCA)

Principal component analysis (PCA) is an unsupervised multivariate analysis method used in data analysis to show variations in large datasets. More specifically, it highlights the components that differ in a Raman spectrum or simply suppresses noise by separating samples by distinguishable spectral features.[134] PCA converts correlations (or lack thereof) among datasets into a 2D or 3D cartesian graph. It reduces the dimensionality of a dataset by removing dependent datapoints, while retaining its variability.[135] The ‘relevant’ data is compressed into a PCA plot where clusters indicate highly correlated spectral features. The data reduction is achieved by transforming the spectra to a new set of uncorrelated variables. In Raman spectroscopy for example, the original spectrum axes report the intensity at wavenumbers. The PCA axes show the ‘intensity’ of contribution from each principal component (PC) to the spectrum. The spectral resolution of our system is approximately 1 cm^{-1} (see 2.7.4 Spectral resolution). The intensity of any given pixel will be correlated to the intensity of its neighbouring pixel, as the CCD pixel spacings are > 1 cm^{-1} across the spectrum. The number of peaks present in our spectra is ~ 35, the number of datapoints is significantly higher. PCA seeks to find the eigenvalues and eigenvectors of a covariance matrix, whereby it represents the data as a set of linear, orthonormal functions. PC1 captures most variance of the dataset, PC2 describes the most uncorrelated (orthogonal) variance to this, PC3 is again orthogonal to PC2and so on. The covariance matrix is described as

$$S = X'_0 * X_0/(n-1) \qquad \text{Eq. 17}$$

where X_0 are the mean-centered spectra in rows and n are the number of samples.[135] The diagonal of the covariance matrix gives the variance of the set of spectra at each wavenumber. In that way, a higher dimensional data matrix can be projected onto a low component subspace (maximal dimensionality reduction).[135,136]

Before PCA can be performed on a set of Raman spectra, the data must be pre-processed.[136] This is a multiple step procedure which includes

a) Cosmic ray/spike removal on individual spectra during monitored Raman spectra collection in LabSpec.[137]

b) Interpolation to align the spectral axis in IgorPro.[131,136]

c) Polynomial baseline subtraction in Raman Tool Set.[130,138]

d) Total area normalization in Raman Tool Set to remove artefacts from spectral acquisition. Source and environment fluctuations might add variability to Raman spectra that are not related to differences in the sample.[136]

e) 0.65 Cubic spline smoothing in Raman Tool Set.[130,138]

f) Outlier removal. A spectrum with a PC1 or PC2 score greater/lesser than 3 times the 2SD of PC1 or PC2 was removed (99% confidence level).[136]

PCA was then performed through R Studio.[139] The spectra were scaled so that the mean of each point was zero. The PCs were selected according to the Kaiser's criterion where only PCs covering >1% variance were taken into consideration.[140] The constituents representing the main variable features show the greatest distance from zero in their respective PC loading plots.

CHAPTER 3

Polarization dependent Raman measurements on celery cellulose

The purpose of the research work presented in this chapter was to characterize the Raman signature of the cellulose substrate obtained from celery collenchyma comprehensively with univariate and multivariate analyses, such as peak intensity ratio (PIR) and principal component analysis (PCA). Raman signatures of celery cellulose at different Raman orientations (0°- 90° in increments of 30±1°) are described on the micron scale. The first part of this chapter is devoted to the description of the vibrational spectra of celery cellulose from collenchyma cells with confocal Raman spectroscopy as the main spectroscopic method, complemented with ATR-IR spectroscopy. The second part presents the study of the angle-dependent transition of normalized bands in respect to the Raman collection angle. The third part contains the statistical analysis of Raman measurements. Statistical measures such as average spectra, standard error waves and 2nd order standard deviation of individual Raman spectra are listed in the appendix. Each Raman peak represents a combination of group frequencies, where the standard error of the mean will converge at different rates. The percentage standard error for the ratio of the majority of cellulose peaks was found to be converged for 50 Raman spectra. Therefore, it was concluded that it is necessary to collect a minimum of 50 random point spectra from a cellulose sample to ensure proper characterization. The convergence test was repeated for enzymatically treated cellulose on datasets comprising ~50 individual Raman spectra per sample.

PIR analysis was used to quantify the changes in peak intensity in respect to their collection angle (polarisation dependent Raman spectroscopy). PIR analysis undertaken by this study was able to identify the orientation of the majority of characteristic bonds. PCA was used to separate the samples by their cumulative spectral features and to provide a basis for deeper spectral analysis based on the PCA loadings. Although partially separated by PCA, the 30° and 60° spectra have shown to represented angle-dependent ‘transition states’, wherefore it was concluded that for the enzymatic study, only 0° and 90° spectra will be collected.

3.1 Raman and ATR-IR spectra of cellulose from celery collenchyma

Figure 3.1 shows a representative ATR-IR spectrum of highly oriented, semi-crystalline cellulose from celery (blue line) and a corresponding Raman spectrum where the sample orientation was perpendicular (90°) to the electric vector of the laser light (red line).

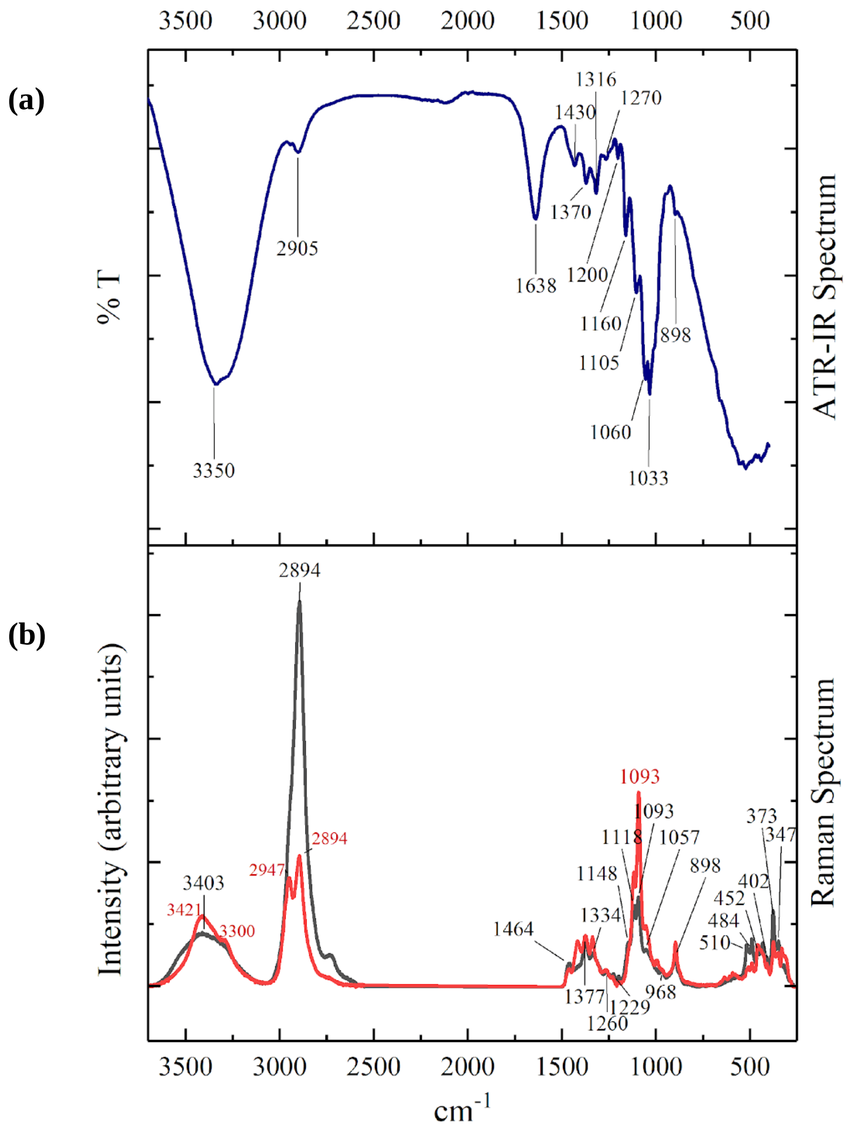


Figure 3. 1: (a) Infrared and (b) Raman spectrum of cellulose from celery collenchyma. Top trace IR spectrum in % transmission (%T): The lower the transmission value the greater the absorption (blue line). Lower trace Raman scattering in intensity (red line). For both spectra, the fingerprint region ranges from 250 - 1500cm^{-1} with annotation of the position of the main peaks. In the high wavenumber region, the annotation of the main peaks are the CH-stretching region (2880 - 2910 cm^{-1}) and the broad band of the OH-stretch (3000 - 3700 cm^{-1}).

IR and Raman spectroscopy are vibrational spectroscopy methods that can be used complementary to obtain the full molecular vibrational information of a sample, as the techniques are based on two discriminative principles. IR absorption is based on the intrinsic dipole moment of a molecular bond, whereas Raman spectroscopy is based on its polarizability. As a result, strong Raman scatterer modes tend to give weak bands in the IR spectra and vice versa. This can be seen at the OH band at ~3400 cm^{-1} which is significantly stronger in the IR spectrum than in the Raman spectrum, and the CH stretch at ~2900 cm^{-1} that results in a strong Raman signal, and a weak peak in the IR spectrum. Both spectra exhibit very characteristic peaks in the fingerprint region (200 – 1500 cm^{-1}). In correspondence to the specific sample environment, these bands exhibit distinct positions and widths. The cellulose microfibrils give a Raman signature comprising approximately 30 Raman bands per spectrum. Of these, the significant band contributions are summarized in Table 3.1.

Table 3. 1: Position of group assignments of the cellulose bands found in literature corresponding to celery fibers (Figure 3.1)[32,44,133,141,142]

Band position (cm^{-1})	Assignment according to the maximum contribution to bands as per region[32]
300 – 400	Ring torsions (out-of-plane deformations)
400 – 700	Heavy atom bending CO and ring modes
650 – 800	OH out-of-plane bending
800 – 950	Bending involving heavy atoms (CCC, COC, OCC, OCO), ring and CO stretches, external bending modes of methylene group
950 – 1150	Significant contributions from ring stretching, together with CO stretching motions
1200 – 1450	Considerable coupling between angle bending involving one bond to a hydrogen atom and the other to a heavy atom (methine bending, methylene bending and COH in plane bending)
2850 – 3000	CH and CH_2 stretching
3000 – 3500	OH and OH_2 stretching

3.2 Variation of Raman bands upon collection angle

The Raman intensity of the cellulose bands of highly oriented cellulose from celery collenchyma depend on the angle between the orientation of the cellulose microfibrils and the electric vector polarisation of the incident laser.[32] Figure 3.2 shows the microscopic images where celery was rotated from 0° to 90° in increments of 30. The alignment of collection angles were performed manually using a protractor.

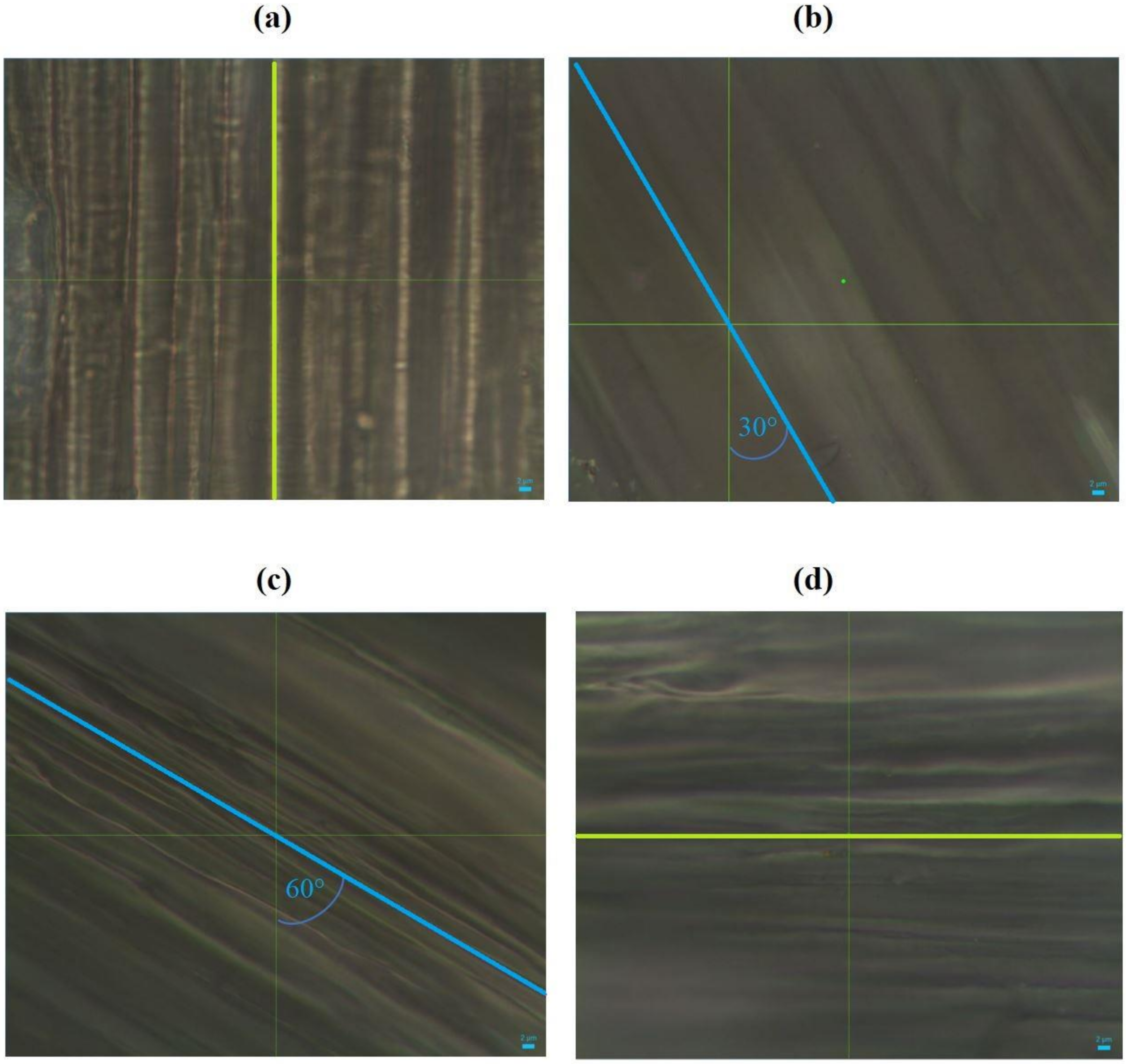


Figure 3. 2: x100 optical microscope image of celery cellulose fibers showing the collection angle at which the Raman spectra were collected. The polarisation of the laser is vertical wherefore: (a) 0° collection angle, (b) 30°, (c) 60° and (d) shows the fiber orientation for Raman spectra collected at 90° in respect to the laser polarisation.

Figure 3.3 shows the Raman spectra of cellulose in the Fingerprint region (200 - 1800 cm^{-1}) of crystalline cellulose from the parallel (0°) to the perpendicular polarized configuration (90°) in steps of 30. Figure 3.4 shows the respective spectra for the high wavenumber region. The spectra were collected at random points across the sample. The Raman spectrum from the undigested celery cellulose (control) mainly consists of cellulose, as well as hemicellulose residuals (peaks at around 830 cm^{-1}). According to literature, changing the sample orientation from parallel to the laser polarization (0°) to perpendicular (90°) leads to strong changes in the Raman intensity in almost all characteristic bands with the exception of the two bands at 1376 and 435 cm^{-1}.[32] The Raman band at 1376 cm^{-1} has therefore been used as reference point for baseline subtraction and PIR analysis. The Raman bands at 1465 cm^{-1} (HCH and HOC bending), 513 cm^{-1}, 490 cm^{-1} and 377 cm^{-1} (heavy atom stretching) show higher intensities at 90° than at 0°, thus indicating a vertical alignment of these groups (see Table 3.2). The first column of Table 3.2 shows the peak position of Raman bands as an average across the dataset. The second and third column show comparable peak positions found in literature and their annotations. Column 4 characterizes the maximum contribution of the vibrational modes in cellulose. A delocalization or a shift towards the lower wavenumber region when rotating the fiber axis from 0° to 90° was observed for the spectral features at 982 cm^{-1} and 810 cm^{-1}, generally assigned to heavy atom stretching (C-C and C-O stretch) and bending (C-C and C-O bend). The displacement was ~ 10 cm^{-1} as determined by peak fitting. In 1987, Wiley and Atalla presumed, that the bands around 900 cm^{-1} are delocalized based on geometrical cellodextrin calculations. The band at 900 cm^{-1} was hence assigned to the β-glyosidic linkage.[143] They further suggested through several empirical observations, that at ~ 900 cm^{-1}, relative intensity changes are proportional to the amount of disorder in cellulose, as the likely sites of disorder are the glyosidic linkages.[143] As expected and observed in this study, the glyosidic bonds of the glucopyranose ring units are oriented parallel along the fiber axis. To conclude, the bands that are most intense at 0° result from vibrations in which the maximum polarizability is parallel to the chain axis. For modes that are most intense at 90°, the maximum change in polarizability is perpendicular to the chain axis.[93]

Several aspects are noticeable in the spectra: Firstly, in general, symmetric stretches of bonds (COC bond at 1095 cm^{-1}) tend to yield stronger Raman bands when the orientation is parallel to the electric vector of the excitation laser. The direction of their maximum polarizability (polarizability tensor) is hence approximately parallel to the chain axis.[122] Totally symmetric vibrational modes are associated with the largest changes in polarizability with respect to the vibrational modulation of the bond.[98] Secondly, fine structures (shoulders etc. at e.g. 1145 cm^{-1}), referred to as Raman '*splits*', change in intensity relative to the electric vector of the excitation laser.[98] Both the 1095 and 1120 cm^{-1} peaks have traditionally been assigned to symmetric and antisymmetric COC stretches plus ring breathing, respectively.[99] The peak fitting results including statistical uncertainties are summarized in Table 3.2. The observed Raman bands are characteristic for natural cellulose Iβ and can be correlated to the corresponding vibrational modes of cellulose found in literature.[19,39,44,50,133,143–145] Overall, the results presented highlight two important considerations in the system: (a) the intensity/polarizability of vibrational modes in respect to the electric vector of the laser and (b) Raman peak splitting due to reorientations of the polarization axes.[10] This means that the vector of the polarizability matrices of the same chemical groups are oriented differently within the structure, resulting in two distinguishable peaks (a peak 'split'). The bands that are most intense at 0° result from vibrations in which the maximum polarizability is parallel to the chain axis. For modes that are most intense at 90°, the maximum change in polarizability is perpendicular to the chain axis[93].

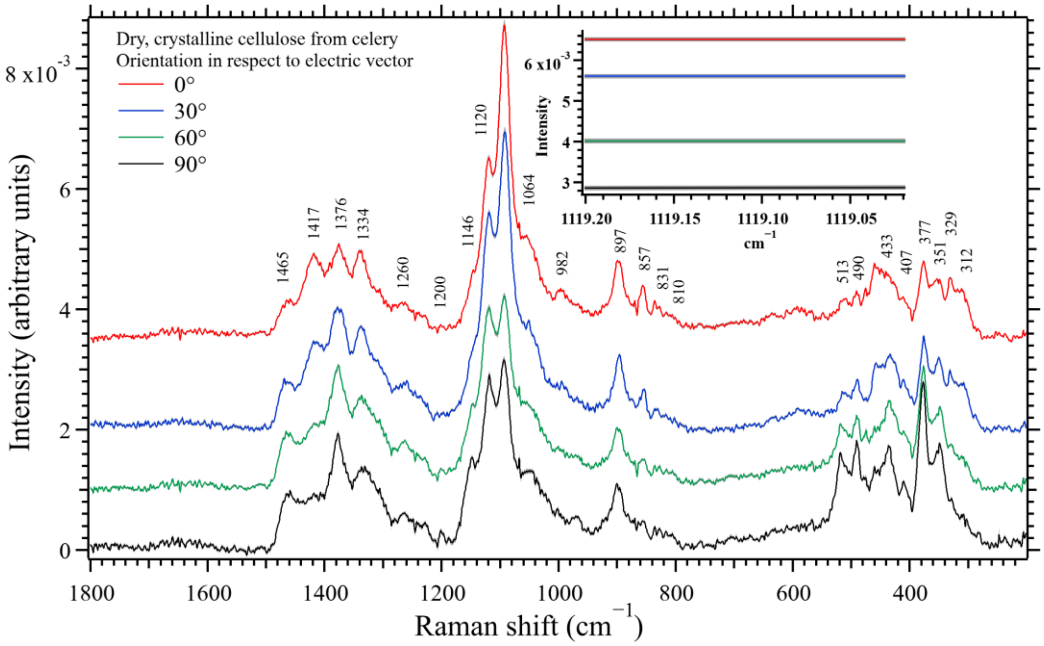


Figure 3. 3: Average Raman spectra in the fingerprint region (1800 cm^{-1}-200 cm^{-1}) of dry, semi-crystalline cellulose fibers showing peak assignments (Table 1). The fibers were orientated in respect to the electric vector of the laser from 0° to 90° in steps of 30. The spectra are vertically shifted using an arbitrary offset to aid

visualization. The graph shows from top to bottom the average spectrum at 0° (red) [n=141], 30° (blue) [n=78], 60° (green) [n=126] and 90° (black) [n=191].

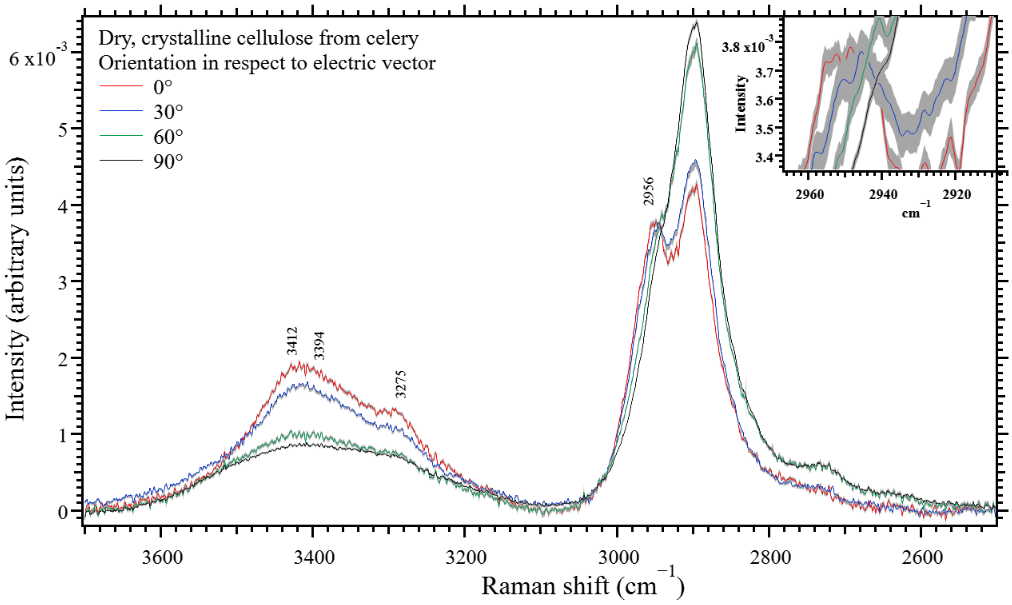


Figure 3. 4: Average Raman spectra in the high wavenumber region (3700 cm^{-1}-2500 cm^{-1}) of dry, semi-crystalline cellulose fibers showing peak assignments (Table 3.3). The fibers were orientated in respect to the electric vector of the laser from 0° to 90° in steps of 30. The graph shows the average spectrum at 0° (red) [n=141], 30° (blue) [n=78], 60° (green) [n=126] and 90° (black) [n=191]. The low standard error envelope was added to each spectrum in grey (see sub-window).

A convergence test (Table 3.2) and PCA analysis (Figure 3.5) were performed on these data sets to verify the statistical measurements and to complement the PIR analysis.

Table 3. 2: : Raman bands for untreated cellulose fibers from celery collenchyma. Showing peak positions obtained in this work, peak locations from literature[19,23,39,44,50,125,133,144,145,149], band assignments and maximum vibrational contributions to bands, respectively. Included in this table are the peak intensity ratio calculations from this work in respect to the band at 1376 cm^{-1}. Symbo ‖ shows peaks, that were most intense at 0°, ⊥ at 90°, - for peak splitting and ⁄ were peaks most intense at 60°.

Peak location (cm -1)	Literature comparison	Vibrational spectrum Cellulose Band assignment	Maximum contribution to bands as per region (cm -1) according to literature	Peak ratios			
				0°	30°	60°	90°
329 ± 2 ‖			300 – 400 Predominantly ring torsions (out-of-plane deformations)	0.35 ± 0.01	0.37 ± 0.01	0.30 ± 0.02	0,20 ± 0.01
351 ± 1 ⊥	337-520	γ (CCC), γ (COC), γ (OCC), γ (OCO) skeletal bending γ (CCH), γ (COH) methane bending movement (CC) and (CO) within the glucopyranose ring units		0.50 ± 0.06	0.22 ± 0.01	0.22 ± 0.01	0.81± 0.04
377 ± 1 ⊥				0.59 ± 0.07	0.52 ± 0.01	0.87 ± 0.04	1.22 ± 0.05
407 ± 1 ⊥			400 – 700 Heavy atom bending CO and ring modes 650 – 800 OH out-of-plane bending	0.25 ± 0.04	0.64 ± 0.02	1.31 ± 0.05	0.32 ± 0.1
433 ± 3 ⊥				0.31 ± 0.02	0.12 ± 0.00	0.30 ± 0.10	0.85 ± 0.1
458 ± 3 ‖				0.65 ± 0.03	0.45 ± 0.01	0.90 ± 0.10	0.39 ± 0.06
476 ± 0 ⊥				-	0.32 ± 0.01	0.42 ± 0.07	0.36 ± 0.03
490 ± 1 ⊥				0.25 ± 0.01	-	0.38 ± 0.04	0.62 ± 0.05
513 ± 2 ⊥				0.23 ± 0.01	0.24 ± 0.01	0.67 ± 0.05	0.55 ± 0.03
810 ± 10 ⁄			Below 950 Angle bending involving heavy atoms (CCC, COC, OCC, OCO), ring and CO stretches, external bending modes of methylene group	0.10 ± 0.01	0.17 ± 0.01	0.59 ± 0.03	0.10 ± 0.01
831 ± 4 ‖				0.20 ± 0.00	0.06 ± 0.01	0.11 ± 0.01	0.12 ± 0.01
857 ± 2 ‖	860	ν (CC)		0.38 ± 0.01	0.10 ± 0.01	0.13 ± 0.01	0.21 ± 0.01
897 ± 1 ‖	896	Glucose ring deformation		0.62 ± 0.01	0.23 ± 0.01	0.23 ± 0.01	0.53 ± 0.02
	898	β-glycosidic linkage			0.48 ± 0.02	0.56 ± 0.01	
982 ± 10 ‖	970-1050	CC and CO vibration, γ (HCH)	950 – 1150 Significant contributions from ring stretching, together with CO stretching motions	0.25 ± 0.00	0.17 ± 0.00	0.20 ± 0.03	0.18 ± 0.03
1064 ± 2 ⁄	1050-1145	ν (CO) in hydrogen bonding systems		1.08 ± 0.01	0.85 ± 0.02	1.10 ± 0.03	1.02 ± 0.03
1093 ± 0 ‖	1090	Skeletal deformation		3.02 ± 0.06	2.11 ± 0.05	1.75 ± 0.04	1.63 ± 0.05
1119 ± 1 -	1120	νs (COC) νas (COC)		1.81 ± 0.03	1.62 ± 0.06	2.10 ± 0.05	1.95 ± 0.06
1146 ± 3 ⊥				0.58 ± 0.00	0.57 ± 0.02	0.80 ± 0.02	0.75 ± 0.02
1199 ± 0 ⊥	1200		1200 – 1450 Considerable coupling between angle bending involving one bond to a hydrogen atom and the other to a heavy atom (methine bending, methylene bending and COH in plane bending)	-	-	0.09 ± 0.02	0.15 ±0.01
1260 ± 6 /	1262	ρ (HCH), γr (HOC), γr (COH)		0.31 ± 0.01	0.31 ± 0.01	0.43 ± 0.02	0.35 ± 0.01
1312 ± 1 ‖	1315	γ (CCH), γ (COH) and γ (CH$_2$) bending		0.38 ± 0.01	0.49 ± 0.02	-	-
1335 ± 3 ⊥	1337-1378	γw (HCH) γr (CCH), γr (COH),		0.88 ± 0.01	0.75 ± 0.02	1.19 ± 0.03	1.07 ± 0.03
1376 ± 2 N	1376	γ (HCC), γ (HCO) and γ (HOC)		1.00 ± 0.01	1.00 ± 0.03	1.00 ± 0.03	1.00 ± 0.04
	1377	γw (HCH) and γw (COH)			0.74 ± 0.02	0.87± 0.02	
1465 ± 1 ⊥	1462	δ (HCH)		0.36 ± 0.00	0.39 ± 0.02	0.64 ± 0.01	0.58 ± 0.02
2909 ± 24 ⊥	2890 - 2905	ν (CH), ν (CH2)	2850 – 3000 CH stretch	N/A			
2956 ± 23 -	2944 - 2968	ν (CH), ν (CH2)					
3394 ± 1 ‖	3100 - 3600	ν (OH)	3000 – 3500 OH stretch				
3275 ± 10 ‖	3270	ν (Interchain 2OH…6OH…6OH…3O)					
3412 ± 5 ‖	3410	ν (Interachain 6OH…3OH…5O)					

ν_s: symmetric stretching, ν_{as}: asymmetric stretching; δ: scissoring; γw: wagging; γr: rocking, ρ: twisting

3.2.1 PCA analysis of angle dependent Raman spectra of celery cellulose

A convergence test (see Appendix S1) and PCA analysis were performed on these data sets to verify the statistical measurements and to complement the PIR analysis. PCA is a pattern recognition procedure, allowing to reduce a large dataset to its main differentiating components. Using PCA this study was able to show highly correlated spectral features which vary in respect to the electric vector of the laser light. The collection angle effects the whole spectrum. 0° (number of spectra n=141) and 90° (n=191) spectra show the greatest spectral changes, whereas Raman spectra at 30° (n=78) and 60° (n=126) represent ‘transition spectra’ between these two variances. Although separated by PCA, the separation of 0° and 30°, as well as 60° and 90° is less clear.

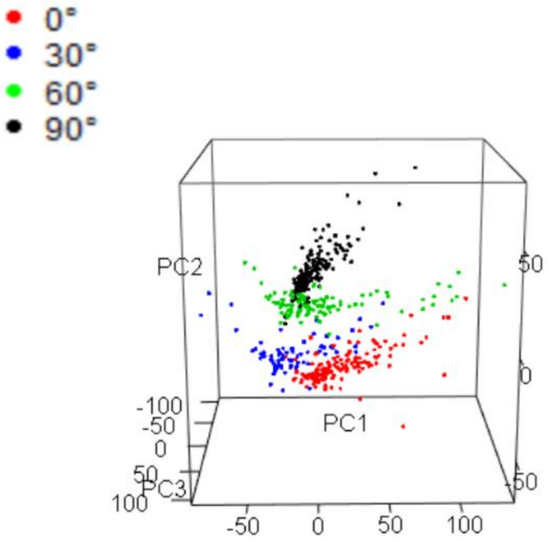


Figure 3. 5: PCA plot of celery cellulose Raman spectra in the fingerprint region (200 – 1800 cm-1). The ‘relevant’ data is comprised into a PCA plot where clusters indicate highly correlated spectral features. The analysis shows angle dependent transition values of peak intensities. 0° (n=141) and 90° (n=191) spectra do differ the most, 30° (n=78) and 60° (n=126) show similarities/ transitions.

PIR analysis was used to quantify the changes in peak intensity in respect to their collection angle (polarisation dependent Raman spectroscopy). PIR analysis undertaken by this study was able to identify the orientation of the majority of characteristic bonds. PCA was used to separate the samples by their cumulative spectral features and to provide a basis for deeper spectral analysis based on the PCA loadings. Although partially separated by PCA, the 30° and 60° spectra have shown to represented angle-dependent ‘transition states’, wherefore it was concluded that for the enzymatic study, only 0° and 90° spectra will be collected.

4. Raman spectroscopy investigation of the effects of enzyme action on oriented cellulose

Abstract:

The next step in this project was the *ex-situ* Raman analysis of enzymatically degraded crystalline cellulose fibers from celery collenchyma. The study in this chapter is motivated by providing a reliable quantitative method for the characterization of enzymatic degradation of cellulosic fibers. For this purpose, natural celery cellulose was degraded by one of three enzymes: CelS2 (here *Sc*AA10), Cel7A or Cel7B. Cel7A is a cellobiohydrolase (CBH), whereas Cel7B is an endoglucanase (EG). The enzymes therefore differ substantially in their mode of action. This chapter discusses the effects of enzymatic degradation on oriented cellulose fibers. Polarized Raman spectra were collected at 0° and 90° (approximately 50 individual Raman spectra per sample) and enzymatic activity was verified with MALDI-TOF analysis. Structural changes on semi-crystalline celery cellulose after Cel7A and Cel7B cellulase treatment were determined with Raman spectroscopy after 72 h. However, structural changes in the Raman spectra of LPMO treated samples have not been observed. The polarized Raman spectra show distinct differences after Cel7A compared to Cel7B degradation. It was therefore concluded that it might be possible to elucidate and differentiate enzymatic mechanisms on crystalline and amorphous cellulose regions at advanced stages of enzyme degradation.

4.1 Raman spectral analysis reveals changes in cellulose-specific bands

Figure 4.1 and Figure 4.2 show the average Raman spectra from Cel7A (a), Cel7B (c) and *Sc*AA10 (e) degraded cellulose, as well as negative controls. The spectra in Figure 4.1 were collected at random points across the sample at a fiber axis orientation of 0° to the electric vector of the laser, whereas the spectra in Figure 4.2 were collected at 90°. Peak fittings of the spectra have been performed to determine the peak location and their intensity. After Cel7A digestion, the resulting Raman spectra show predominant changes in the 900-1150 cm^{-1} region. These bands indicate the presence of pectinic sugars, probably residuals from the natural fibers. The high wavenumber region of Cel7A degraded cellulose shows an increase peaks at both 2850 and 2950 cm^{-1} (CH and CH_2 stretch) compared to the negative control. The average spectrum of Cel7B degraded cellulose are more distinct compared to the negative control, indicating an increase in crystallinity. This is indicated by a decreased band width of the Raman peaks. The Raman spectrum of LPMO *Sc*AA10 degraded cellulose did not show any clear results besides the extinguished peaks around 850 cm^{-1} after the addition of reducing agent (ascorbic acid) which are normally attributed to hemicellulose peaks. These might indicate that the reducing agent itself also interacts with the cellulose sample. This might indicate that the majority of the Raman peak signals do not arise from LPMO degraded cellulose.

The 1095 and 1123 cm^{-1} peaks have traditionally been assigned to symmetric and antisymmetric COC stretch plus ring breathing, respectively.[99]Both of these peaks showed changes after enzymatic digestion by the cellulases but showed no change after LPMO treatment. A previous study on plant polymers found that the Raman bands at 1219 - 1471 cm^{-1} are also sensitive to xylan thus might have little diagnostic value for cellulose analyses.[96] Also the peaks around 1219 cm^{-1} are possibly closely related to COH and CH groups of both celluloses and xylan which also contribute to the 1313 cm^{-1} peak.[96] The 1471 cm^{-1} is associated with cellulose and xylan, possibly related to the OH and CH_2 vibrations[96] The high wavenumber region ~2800 cm^{-1} is associated with symmetric and asymmetric CH stretches and the ~3000-3700 cm^{-1} region is known as OH stretching region. The band assignments of significant bands are summarized in Table 4.1

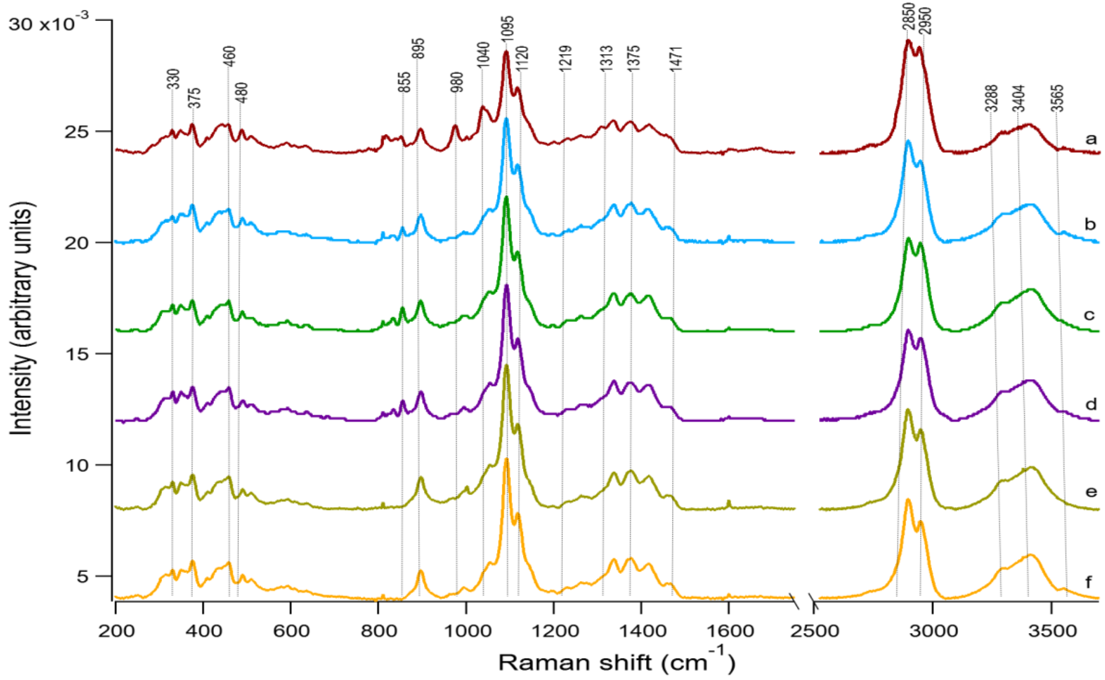


Figure 4.1: Comparison of normalized Raman spectra collected at 0° in respect to electric vector, n= number of spectra. (a) hydrolysed cellulose (Cel7A treatment) n=22, (b) negative control (buffer) n= 45, (c) hydrolysed cellulose (Cel7B treatment) n=59, (d) negative control (buffer + enzyme *Sc*AA10) n=48, (e) oxidized cellulose (*Sc*AA10 treatment) n=63, (f) negative control (buffer + Na-ascorbate) n=54.

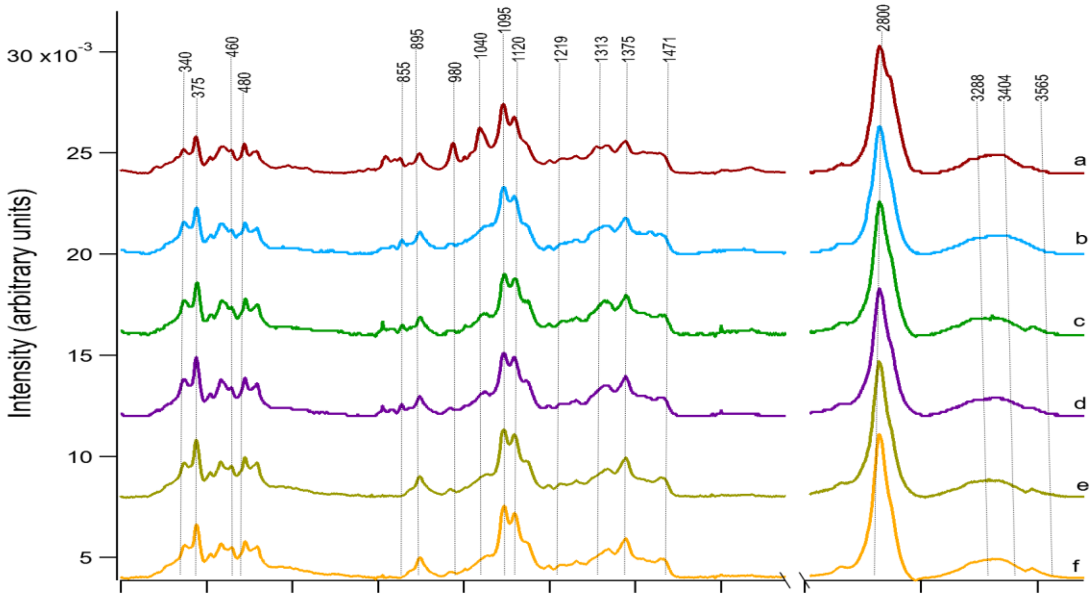


Figure 4.2: Comparison of normalized Raman spectra collected at 90° in respect to electric vector, n= number of spectra. (a) hydrolysed cellulose (Cel7A treatment) n=50, (b) negative control (buffer) n= 42, (c) hydrolysed cellulose (Cel7B treatment) n=54, (d) negative control (buffer + enzyme *Sc*AA10) n = 65,(e) oxidized cellulose (*Sc*AA10 treatment) n=50, (f) negative control (buffer + Na-ascorbate) n=50.

The changes observed in the Raman spectrum indicates that the molecular composition of the cellulose fiber changes after cellulase treatment.

Table 4. 1: Main peak changes of enzymatically treated cellulose and corresponding peak assignments from literature.[44,133,141,142]

Literature (cm^{-1})	**Peakfitting (cm^{-1})**	**Band assignment**	Ref.
1465	1465 ± 1	δ(HCH)	44,141
1378	1376 ± 2	δ(HCC), δ(HCO), δ(HOC)	44,141
1262 1256 1121	1260 ± 6	$\rho(\mathrm{HCH})$, γ_r(HCH), γ_r(HOC), γ_r(COH) [origin of assignment: ‘amorphous region’ of cellulose] δ(CH), δ(COH) [origin: hemicelluloses]	133
1094	1092 ± 1	ν(COC) glyosidic, assymetric	44,141
971	992 ± 22	$\rho(\mathrm{CH_2})$	44,141
854	857 ± 2	(COC) skeletal mode of α-anomers [origin of assignment: pectin]	142
817	810 ± 10	ν(COH) ring [origin of assignment: pectin]	142
380	377 ± 1	δ(CCC) ring	44,141
1465	1465 ± 1	δ(HCH)	44,141

The statistical convergence test (average of individual Raman spectra, their standard error and 2nd order standard deviation) can be found in the appendix (see Appendix S2).

4.1.1 Raman analysis reveals changes in cellulose-specific signature peaks of hydroxylated cellulose - univariate and multivariate analyses

Figure 4.3 shows the comparison of the intensity of cellulose-specific Raman bands at 1095 and 1120 cm^{-1} from celery collenchyma with and without enzyme treatment. The spectra were total area normalized for comparison and the standard errors (red) represent the corresponding fitting uncertainties. The two cellulose signature peaks, 1095 and 1123 cm^{-1}, show similar shape for all the samples.

The amplitude of the peak intensity at 1095 cm^{-1} decreases by 16% after 72h of Cel7A treatment when the crystalline sample orientation was 0° along the fiber axis in respect to electric vector of the laser; and 35% for spectra collected at 90°. The Raman bands at 1120 cm^{-1} after Cel7A treatment show decreased peak intensities of 20% at 0°, and 62% for spectra collected at 90°.

According to literature, the activity of endoglucanase Cel7B (*Humicola insolens*) is restricted to amorphous cellulose regions. An increase of crystallinity in the Raman spectrum was observed through the intensified peaks at 90°, showing an increase of 17% at 1095 cm^{-1} and 10% at 1120 cm^{-1}. At 0° collected spectra show a decreased peak intensity by 12% at 1095 cm^{-1} and 11% at 1120 cm^{-1}.

The LPMO *Sc*AA10 treated cellulose samples, however, show no significant changes compared to their corresponding negative controls, indicating that the substrate changes are subtle compared to cellulase action.

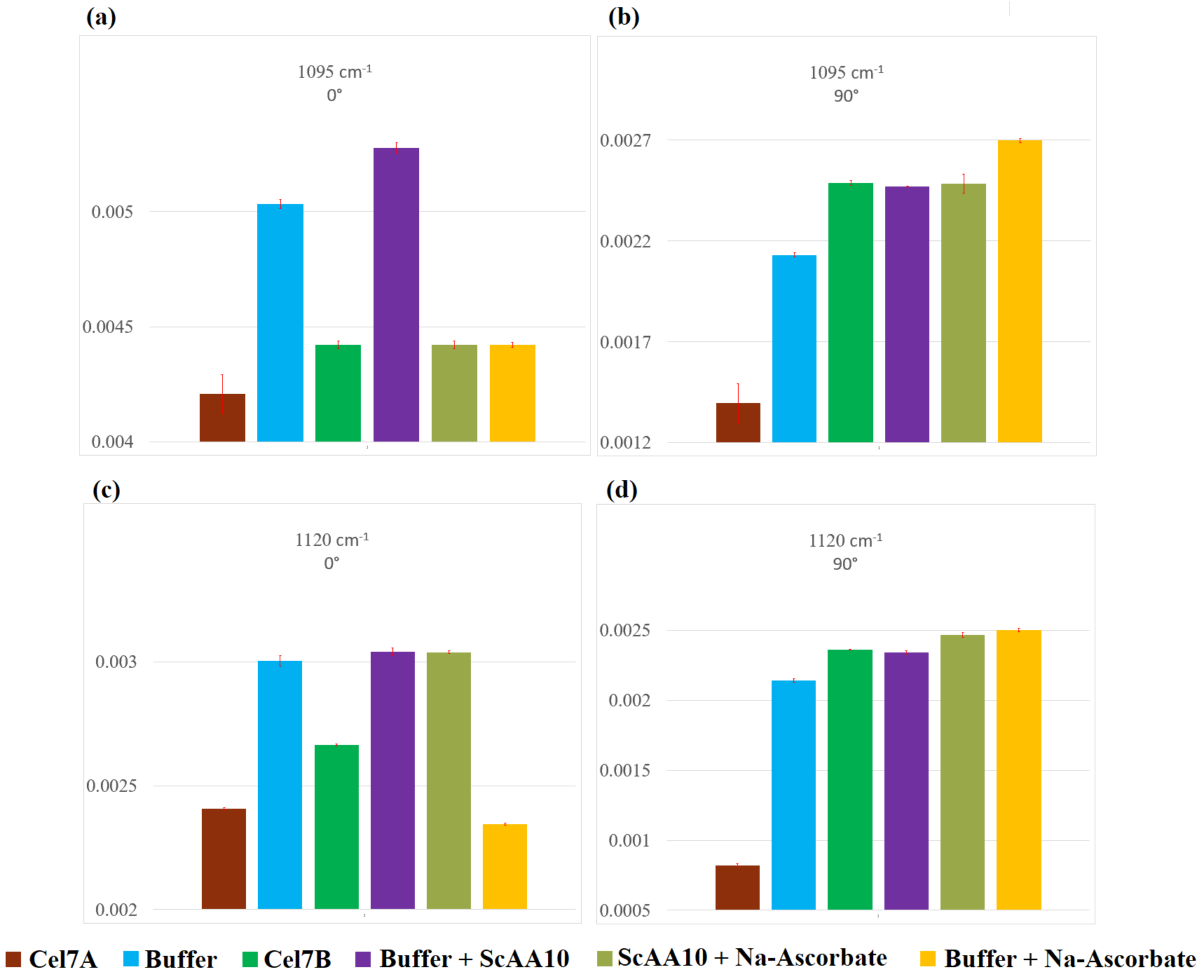


Figure 4.3: Comparison of cellulose's signature Raman peaks at 1095 cm^{-1} and 1120 cm^{-1} in enzymatically degraded celery cellulose. All spectra are normalized. Red bars indicate the standard error. Samples include Cel7A degraded cellulose, Negative control cellulose – Buffer, Cel7B degraded cellulose, Negative control cellulose – Buffer with *Sc*AA10, *Sc*AA10 degraded cellulose, Negative control – Buffer with reducing agent Na-ascorbate. **(a): Peak ratio 1095 cm^{-1} at 0° , (b) Peak ratio 1095 cm^{-1} at 90°, (c) Peak ratio 1120 cm^{-1} at 0°, (d) Peak ratio 1120 cm^{-1} at 90°**. The peak heights in semi-crystalline cellulose for cellulose specific peaks shows significant variance of cellulose peak intensities after Cel7A and Cel7B treatment compared to their negative control, whereas changes in *Sc*AA10 degraded cellulose remain within the standard error of its negative controls.

4.2 Optical microscopy on enzymatically degraded cellulose

The Horiba XploRA Raman instrument used for Raman spectroscopic analysis is also a confocal microscope. Figure 4.4 (a) shows disruption of the otherwise highly oriented fiber orientation after Cel7A treatment, showing that the enzyme has been acted on the substrate. This is in alignment with the current literature, stating that Cel7A works in a disruptive manner, snapping individual cellulose chains out of the crystalline structure.[126] A change in the morphology of the substrate after Cel7B treatment (Figure 4.4 (c)) is not clearly visible. A comprehensive study will be needed to further characterize the substrate changes and allow to learn about the mechanism of the enzyme on the substrate, e.g., advanced microscopic techniques such as EM would allow better resolutions.

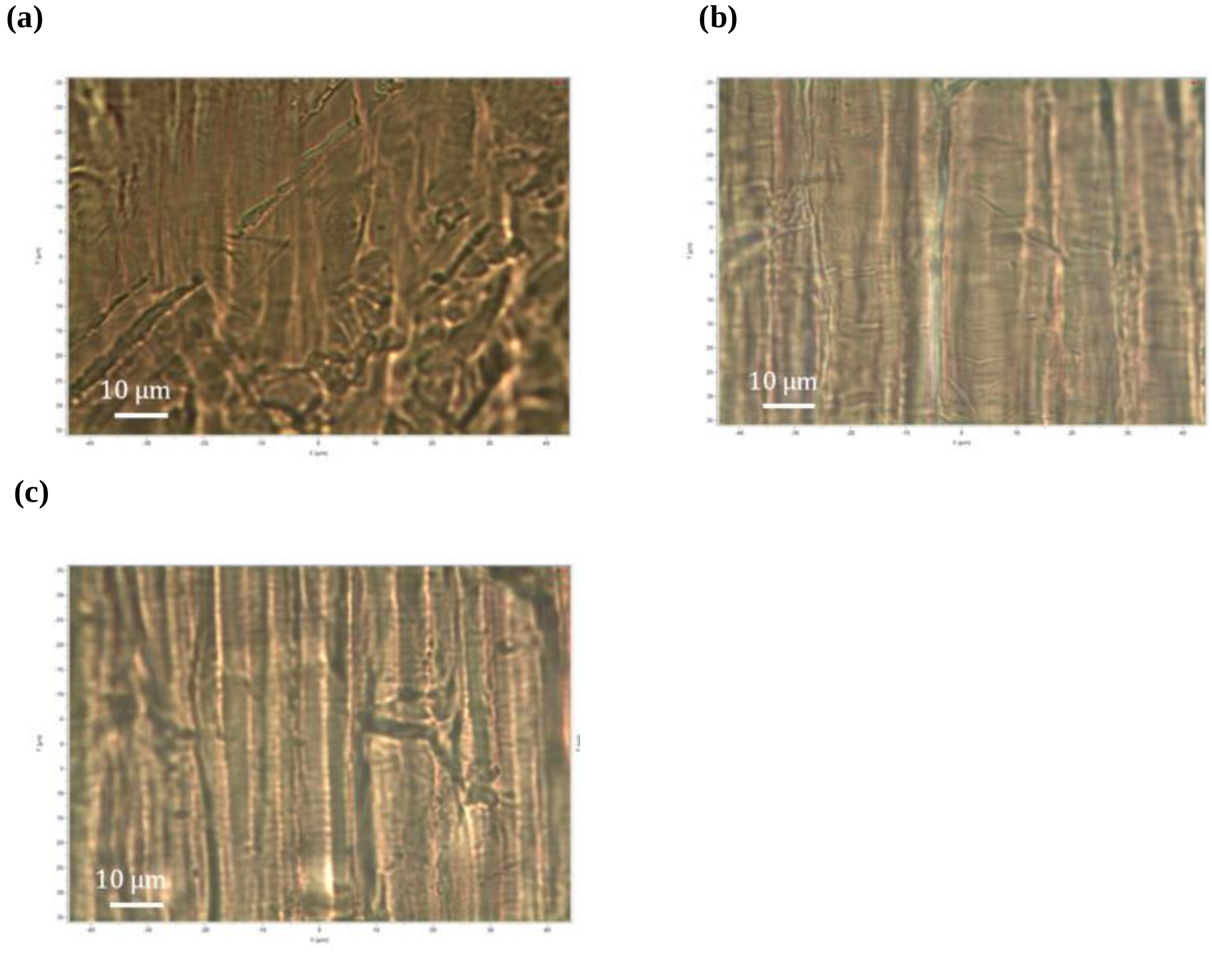


Figure 4.4: Optical microscopy image of cellulose fibers during Raman analysis, x100 optical lens. The fibers are vertically oriented to the microscopy stage, the electric vector of the laser is vertical. (a) Cel7A degraded cellulose, (b) negative control, (c) Cel7B degraded cellulose

4.3 PCA indicates highly correlated Raman bands across the whole spectrum of Cel7A degraded cellulose

Figure 4.5 compares the PCA scores of Cel7A treated cellulose (black) and the negative control data (red) from 0° (a) and 90° (b) experiments, respectively. The Cel7A treated cellulose sample is clearly distinguishable from the negative control. Interestingly, the cluster that contains a significant spectral component are observed to spread along PC2. Along PC1, the spectra remain clustered. Examination of the loading plots revealed that the separation is due to the bands at 800, 1000 and 1200 cm^{-1} attributed to cellulose and hemicellulose specific peaks. The corresponding Loading Plots PC1, PC2 and PC3 (as well as PC4) show high correlations across almost all Raman bands. The average Raman spectra are shown in Figure 4.6 and 4.7.

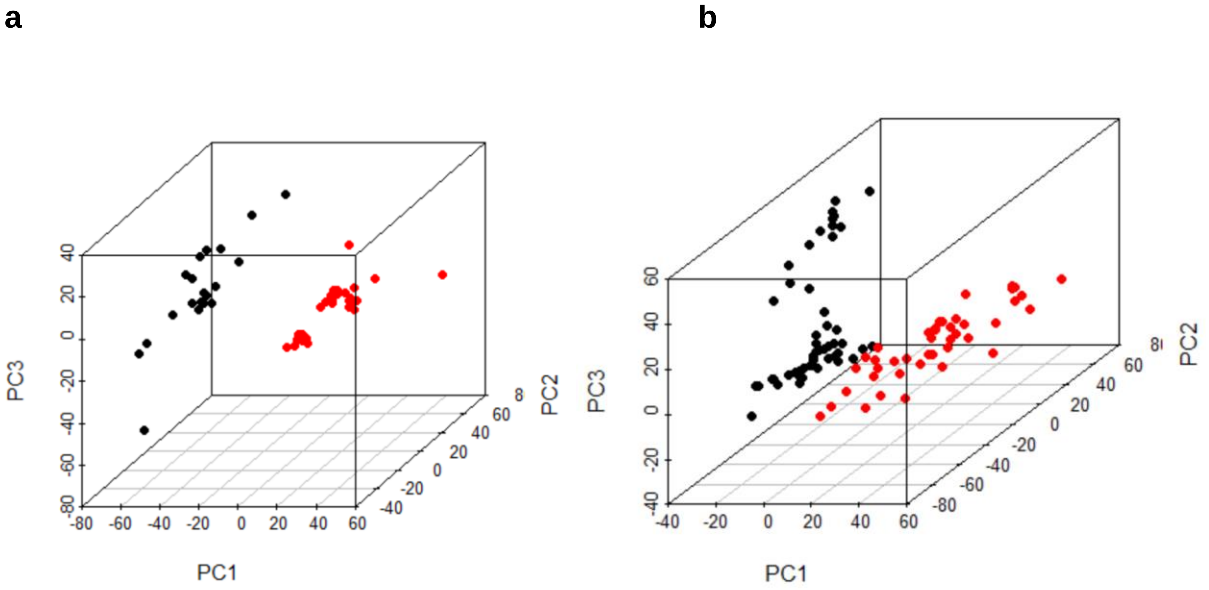


Figure 4.5: PCA analysis results of Cel7A treated cellulose sample at 0° (a) and 90° (b). Cel7A degraded cellulose is shown in black, negative control is shown in red. (a) PC1 covers 45% variance, PC2 = 13% variance, PC3 = 9% variance. (b) PC1 = 29% variance, PC2 = 19%, PC3 = 15%.

4.4 Cel7B degradation increased macromolecular order and crystallinity of cellulose samples as detected with Raman spectroscopy

Figure 4.8 provides the PCA scores of Cel7B treated cellulose (blue) to its negative control (red) at 0° (a) and 90° (b), respectively. The corresponding Loading Plots PC1, PC2 and PC3 (as well as PC4) show that the biggest variance is caused by spectral background, meaning that differences were not caused by individual peaks but by subtle changes in peak widths. This might indicate that the amount of amorphous cellulose regions drastically decreased and, consequently the crystallinity of the sample increased. The resulting crystal structure is therefore more ordered. Raman spectra of crystals or anisotropic samples compared to isotropic samples (e.g., liquids) have sharper peaks and narrower peak widths. Due to the physical coherences of these property, solid samples are usually fitted mathematically with a Gaussian fit whereas peaks of liquids and gasses follow a Voight-Lorentzian distribution.[132,146] However, further work will be needed to understand if this could be used to establish a correlation between the visual, physical and chemical signs of enzymatic cellulose degradation of amorphous regions. The average Raman spectra are shown in Figure 4.9 and 4.10.

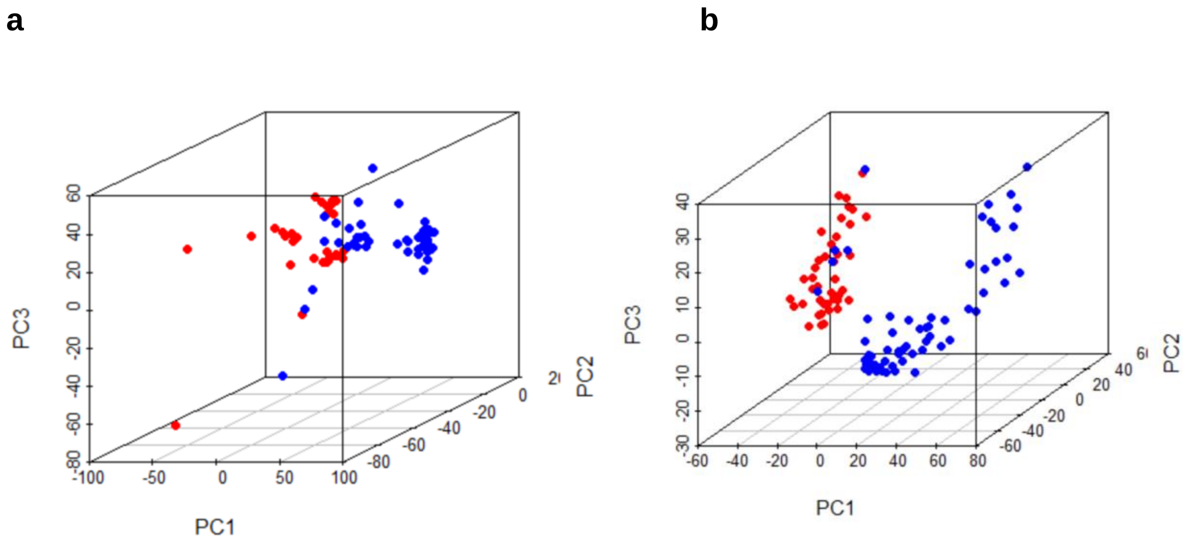


Figure 4.6: PCA analysis results of Cel7B treated cellulose sample at 0° (a) and 90° (b). Cel7B degraded cellulose is shown in blue, negative control is shown in red. (a) PC1 covers 42% variance, PC2 = 12% variance, PC3 = 8% variance. (b) PC1 = 30% variance, PC2 = 16%, PC3 = 8%.

4.5 *Sc*AA10 degraded cellulose fibers

The univariate and multivariate analyses on *Sc*AA10 treated cellulose samples do not show clear results when the vibrational Raman spectra compared to the negative control. Here, the spectra of *Sc*AA10 degraded cellulose of two different replicas are shown, at 0° (Figure 4.11) and 90° (Figure 4.12), respectively. Two negative controls for *Sc*AA10 treated cellulose samples were included. The first negative control included the fiber in addition of LPMO enzyme, but without the reducing agent Na-ascorbate. It is assumed that the enzyme is cannot function in the absence of reducing agent. The second negative control contained only the Na-ascorbate with buffer. The peaks around 850 cm^{-1} were identified as xylan specific peaks, which explains their depletion after treatment with Na-ascorbate.[96] The sub windows in Figure 16 and 17 show the peak at 897 cm-1 assigned to the β-glyosidic linkage which might indicate distortion in the crystalline cellulose conformation. The relative peak intensity for the replicate 2 at 90° shows no statistical variation and differences in replicate 1 are due to background. The analysis of the liquid fraction after the enzymatic degradation of cellulose fibers with MALDI-TOF mass spectrometry revealed C1 and C4 oxidized mono- and oligosaccharides in the liquid fraction of the replicate sample for which Raman spectra were collected at 90°, however this could not be verified for the 0° sample (Figure 4.11).

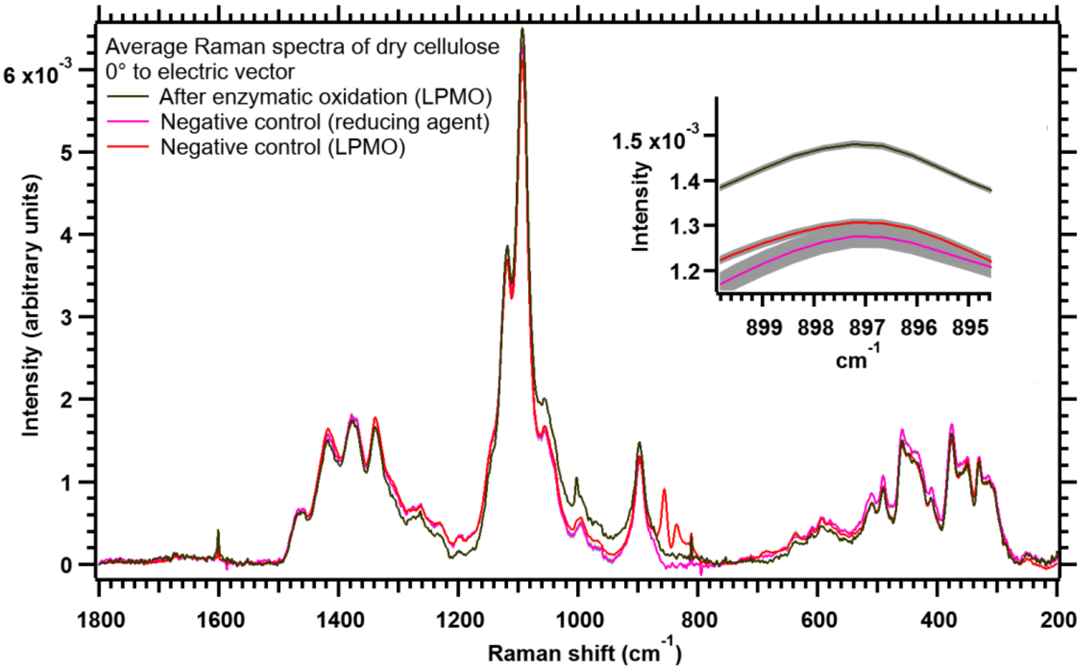


Figure 4.7: Replicate 1. Raman spectrum of fingerprint region of isolated and dried cellulose from celery collenchyma, washed and dried after treatment with *Sc*AA10 for 72 h. Raman spectra were collected at 0° in respect to electric vector. The graph includes the average spectra of *Sc*AA10 treated cellulose fiber (black) comprising 63 individual spectra, negative control with reducing agent Na-ascorbate (magenta) comprising 54 individual spectra, negative control with enzyme and without reducing agent (red) comprising 48 spectra.

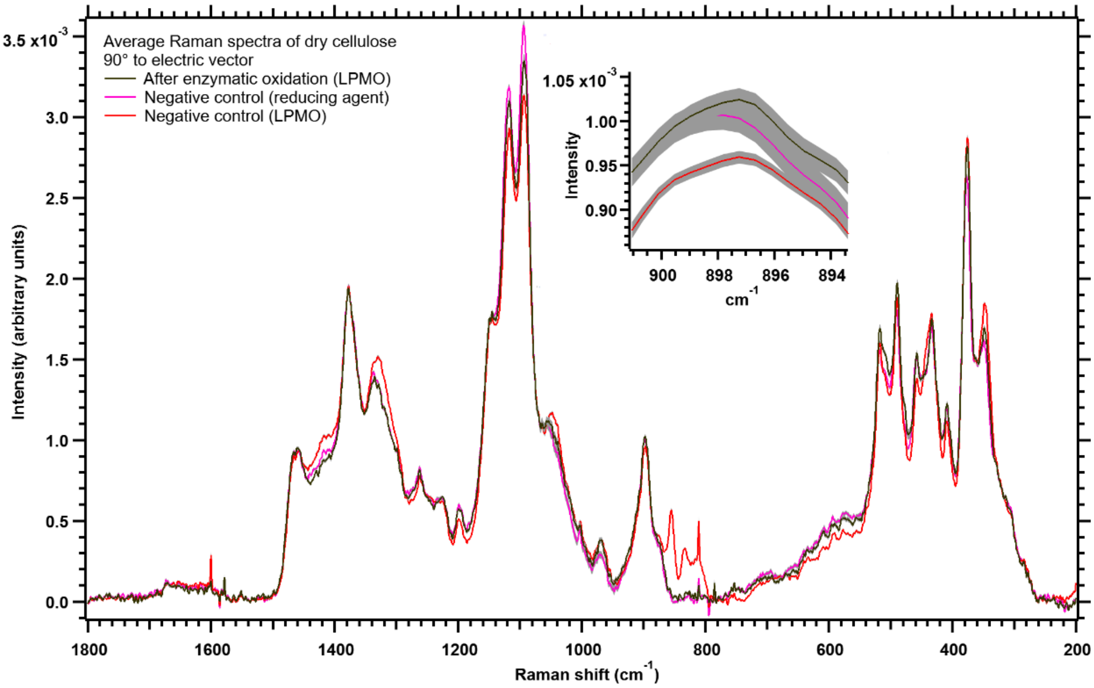


Figure 4.8: Replicate 2. Raman spectrum of the fingerprint region of isolated and dried cellulose from celery collenchyma, washed and dried after treatment with *Sc*AA10 for 72 h. Raman spectra were collected at 90° in respect to electric vector. The graph includes the average spectra of *Sc*AA10 treated cellulose fiber (black) comprising 51 individual spectra, negative control with Na-ascorbate (magenta) comprising 50 individual spectra, negative control with enzyme and without reducing agent (red) comprising 65 individual Raman spectra.

4.6 MALDI-TOF analysis

The conducted MALDI-TOF experiment was a qualitative analysis of the soluble phase of enzyme experiments. The aim was to identify oligosaccharides that were produced by the enzymes and hence verify enzymatic activity. MALDI-TOF mass spectrometry is particularly useful in investigating cellulose-degrading enzyme activity because it allows the detection of the different oligosaccharides that are released after the oxidation of the solid cellulose substrate. At pH 6.5 and in presence of a 100-fold excess of Na-ascorbate the enzyme produced soluble fractions in only one of the replicates, producing aldonic acids and 4-keto-aldoses with different degrees of polymerization. In the single measurement that showed activity, the products seen is comparable with the activity reported by Forsberg *et al.*.[147] This earlier study stated that it is difficult to show soluble *Sc*AA10-generated products when using substrates with high a degree of polymerization (DP) and the amounts of soluble products can be minor.[147]

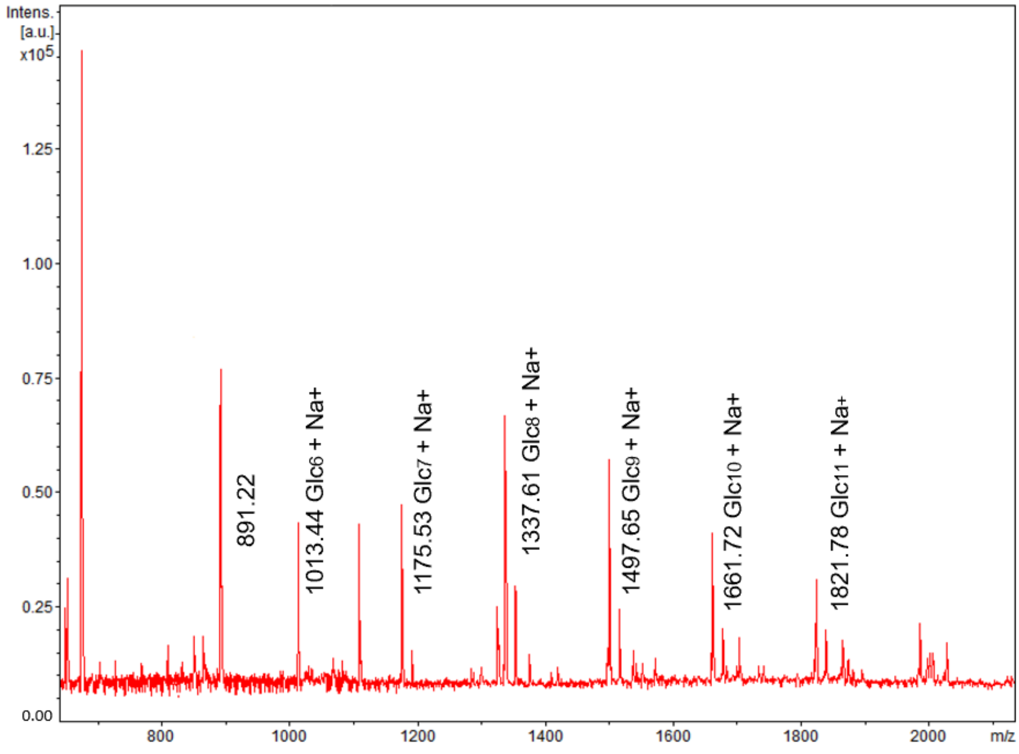


Figure 4.9: MALDI-TOF analysis of *Sc*AA10 on celery cellulose shown in red. Cellulose + ascorbate (1mM), pH 6.5: Panel show soluble native (Glc_{5-} - Glc_{10}) and oxidized cello-oligosaccharides. Only the main peaks in each oligosaccharide cluster are labelled. The oxidized oligosaccharide is observed as sodium adducts, as is commonly seen for carbohydrates containing carboxylic groups[148]. Glcal = aldonic acid; Glc-2 = oxidation from R-OH to R=O (measured molecular weight). Observed masses in the Glc clusters are Glc_6 + Na (1013.44); Glc_7/ Glc_{7-}/ Glc_7al (1173.53/1175.53/1191.53); Glc_{8-}/ Glc_8al (1337.61/1353.61); Glc_{9-}/ Glc_9al (1497.65/1513.65); Glc_{10-}/ Glc_{10}al (1661.72/1677.72); Glc_{11-}/ Glc_{11}al (1837.78).

Figure 4.10 shows an overview of the qualitative MALDI-TOF analyses. Cel7B and Cel7A show several peaks that can be related to sugar residues, whereas *Sc*AA10 show some cellulose activity, although these could not be verified by Raman analyses.

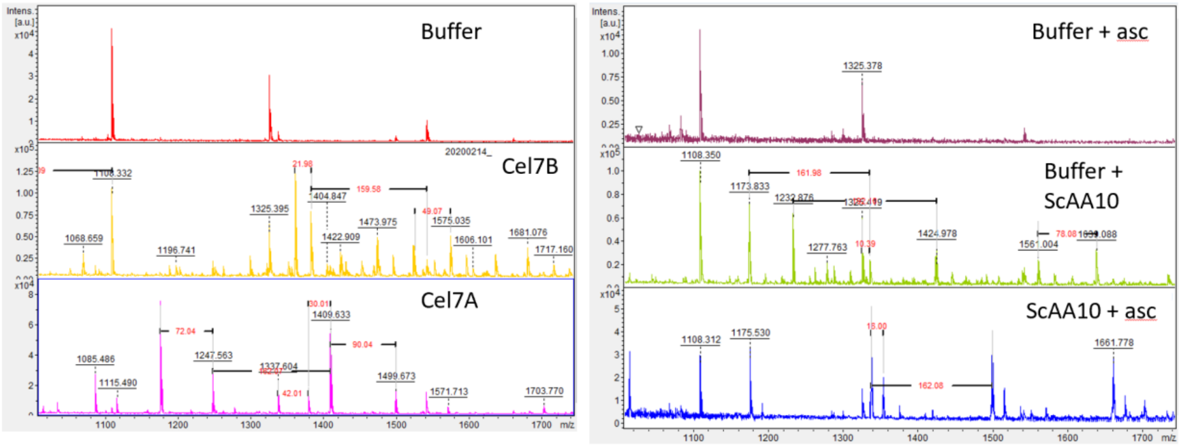


Figure 4.10: Overview of MALDI-TOF results. All enzymes show a pattern of saccharide fractions. The negative controls Buffer and Buffer + asc are comparably clean (peaks resulting from buffer). Buffer + ScAA10 (although used as negative control ‘leaking’ action) and ScAA10 + asc show a characteristic pattern of oligosaccharides, too; so do Cel7A and Cel7B.

4.7 PCA of LPMO samples

Based on the results of chapter 4.5 and Chapter 4.6, PCA was performed to identify any spectral information, that might be hidden. The PCA analysis results of LPMO *Sc*AA10 treated cellulose are shown in Figure 4.14. PCs according to the Kaisers’ criterion (variance >1%) were taken into consideration to try to detect the carbonyl stretch introduced by the oxidative mode of enzymatic LPMO action at around 1770 cm^{-1}.[101] Figure 23 show the PCA plots of the 1100 cm^{-1} – 1800 cm^{-1} region of Raman spectra collected at 90° in respect to the electric vector of the excitation laser. PC2, PC3, PC4 and PC6 were representatively chosen based on the results of the loading plots, as these were showing some variation in the region of interest. However, the analysis of the current samples do not give clear results. Corresponding Raman spectra are shown in Figure 4.12, respectively.

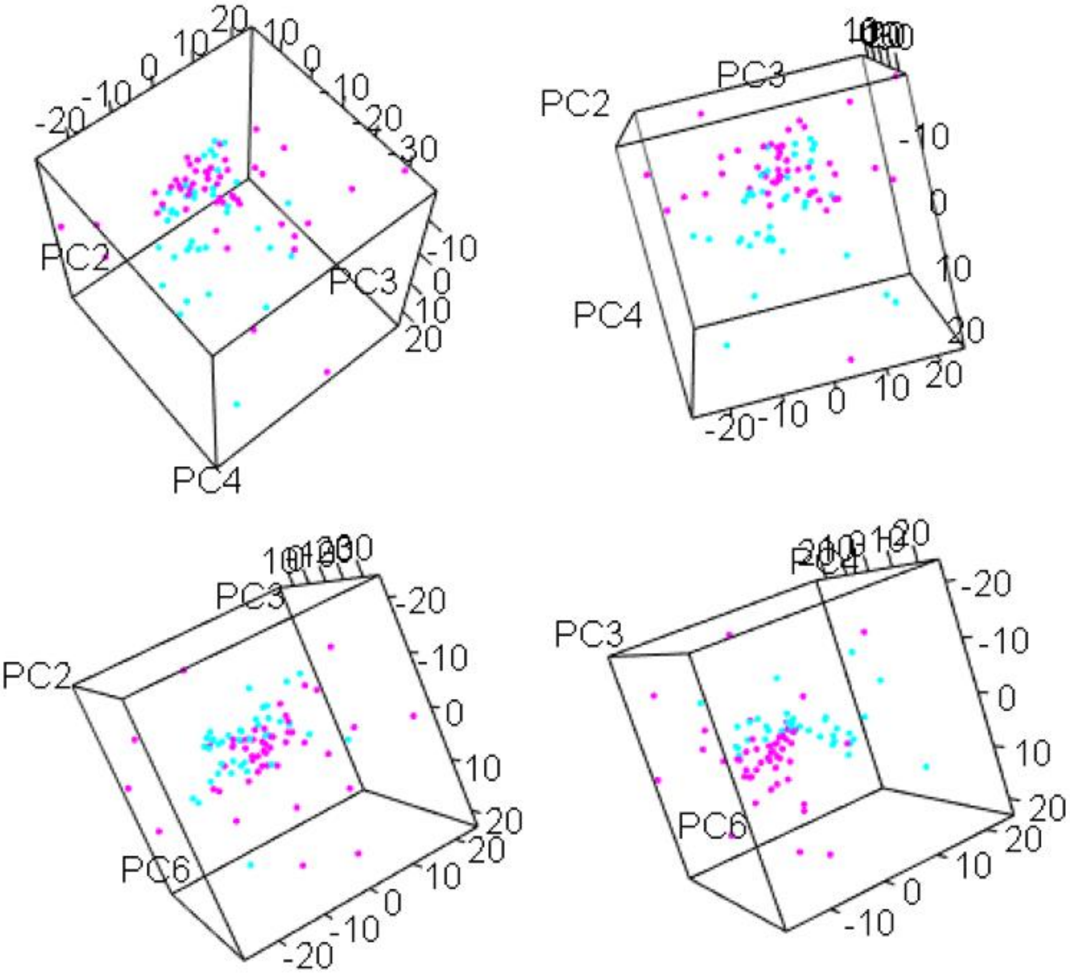


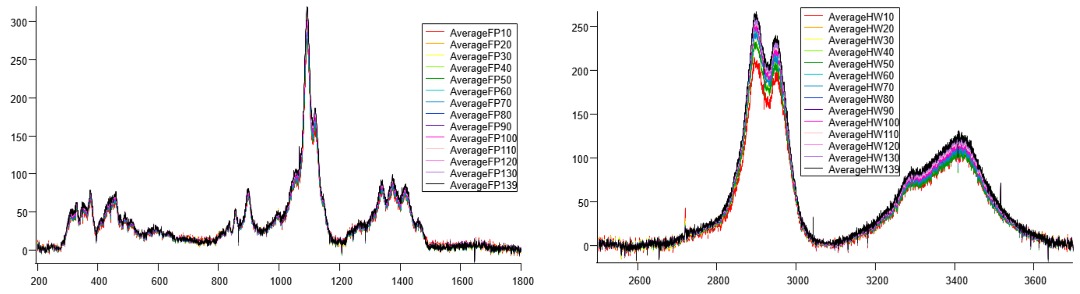
Figure 4.11: PCA for LPMO *Sc*AA10 treated celery cellulose (2nd replicate), *Sc*AA10 treated cellulose shown in (cyan) [n=51] with negative control (magenta) [n=50]. Top plots show PC2, PC3 and PC4 (A) whereas the plots on the bottom show PC2, PC3 and PC6 (B). The PCs represent PC2= 11% variance, PC3= 8%, PC4= 5% and PC6 = 3.

5. Discussion and Conclusions

Raman spectroscopy has been utilised to characterize highly oriented cellulose after LPMO and cellulase treatment, as it was hypothesized to provide a sensitive, non-destructive and label-free method that can be used quantitatively to determine cellulose-specific changes. The spectra were collected at random points across the sample. LPMO *Sc*AA10 treated cellulose did not show any significant differences to the negative controls. The current setup might therefore not be sensitive enough to detect the subtle surface changes after LPMO treatment or the weak Raman intensity of carbonyl (C=O) bonds, because some enzymatic activity could be qualitatively detected by MALDI-TOF mass spectrometry. However, more LPMOs should be tested to evaluate the current findings. *Sc*AA10 has C1 activity, future research could include LPMOs with C4-oxidation activity such as LsAA9A to see if changes on the cellulose surface appeared. Moreover, it might be that ascorbate is interacting with the cellulose samples, as peaks in the Raman spectra around 850cm^{-1} diminished for Na-ascorbate treated samples. It would have been beneficial if another reductant could have been used and investigated, such as hydrogen peroxide. The Raman analytical approach supports detection of enzyme activity and characterization of cellulases. In this project, the activity of well-known cellulases Cel7A and Cel7B were characterised based on morphological changes in the supramolecular structure of highly oriented, semi-crystalline celery cellulose. The enzymatic activity on the investigated materials was verified by MALDI-TOF analysis of the soluble fraction. The Raman data of cellulase-treated cellulose collected in this study suggest, that the cellulose specific signature peaks at 1095 and 1120 cm^{-1} may be correlated with the degree of enzymatic degradation at advanced stages of degradation using univariate and multivariate analysis methods. These peak positions are commonly used for Raman imaging techniques. It was found that Cel7A disrupted glycosidic bonds had a significant macromolecular impact on the structure of celery microfibrils (up to 18 ± 2% decrease of the symmetric COC stretch). Investigations on the full spectral range of Cel7A degraded cellulose, in consideration with microscopic observations, lead to the conclusion, that enzymatically degraded celery fibers increase the degree of disorder in the sample and reveal residues of incorporated hemicelluloses that have not been fully removed. The optical image of Cel7A degraded cellulose confirmed changes in the morphology of the fiber structure, and hence confirm enzyme activity. Cel7B action is restricted to amorphous (disordered) cellulose regions. An increase of crystallinity was observed through a decrease of peak widths across the spectrum and increased peak intensity of cellulose specific bands of 13.5 ± 3.5% at 90° and decreased cellulose specific peak intensities of 11.5 ± 0.5 % at 0° (the direction of disorder is lateral). Principal component analysis (PCA) provided an opportunity to distinguish the datasets into subgroups (cluster) and it is hoped that in the future it can be used to detect specific levels of cellulose degradation. However, other parameters such as the peak width need to be taken into consideration and an applicable quantification method needs to be developed. Future work involves the cross-validation of PCA results, and the development of a valid statistical model based on larger datasets which results in a straightforward Raman-based assay for highly efficient carbohydrate active enzymes on the basis of PCA/LDA (linear discriminant analysis). This requires multivariate calibration and classification models based on larger datasets. Advanced microscopic techniques, such as EM would allow higher resolution of microscopic images and could potentially visualize and complement current findings. Chemical specific information about crystallinity changes of the samples could be complemented with WAXS and solid-state NMR. Further experiments, such as heavy atom labelling of cellulose with D_2O by cellulases, although affecting enzyme kinetics, might lead to a new OD-Raman band in the cellulose spectrum which would represent an affirmative marker to quantify enzyme activity.

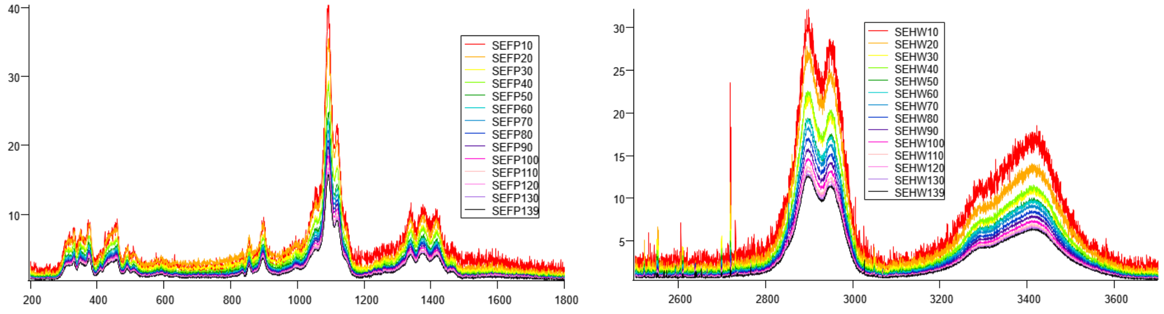
Appendix S1: Convergence Test on natural celery cellulose

0°



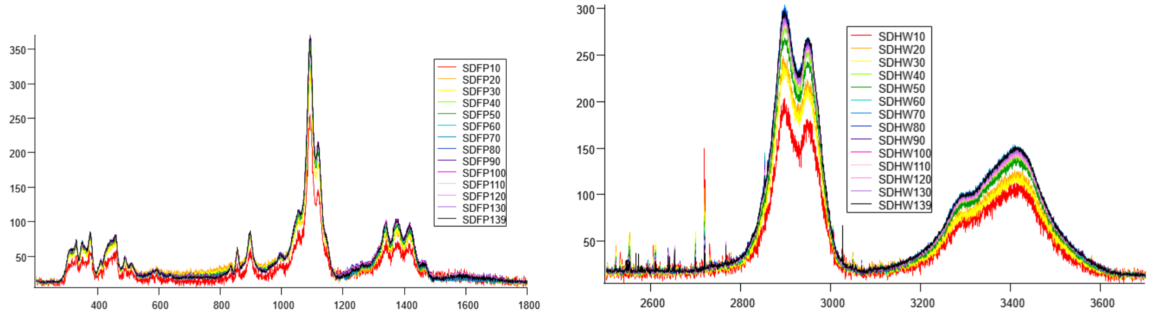
(a) (b)

Figure S1.1: Averaged Raman spectra of celery cellulose at 0° in the (a) fingerprint region and (b) high wavenumber region (n=139).



(a) (b)

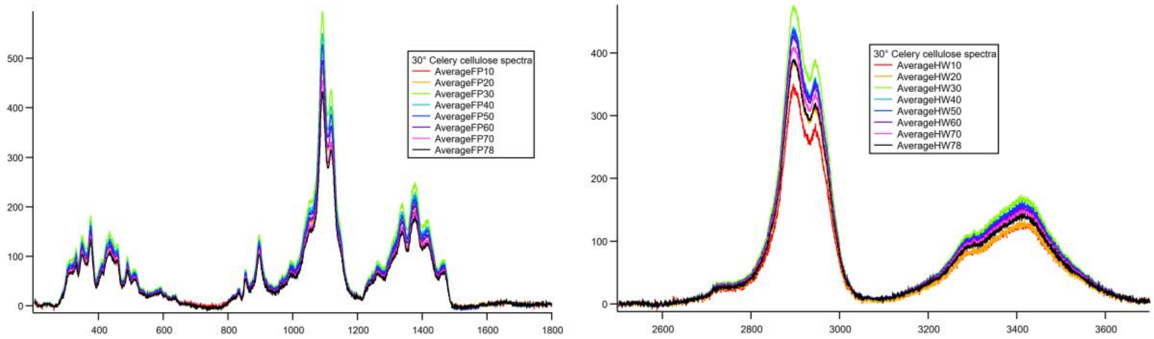
Figure S1.2: Convergence of the standard error of the mean as a function of the number of spectra in the average at 0° in the (a) fingerprint region and (b) high wavenumber region (n=139).



(a) (b)

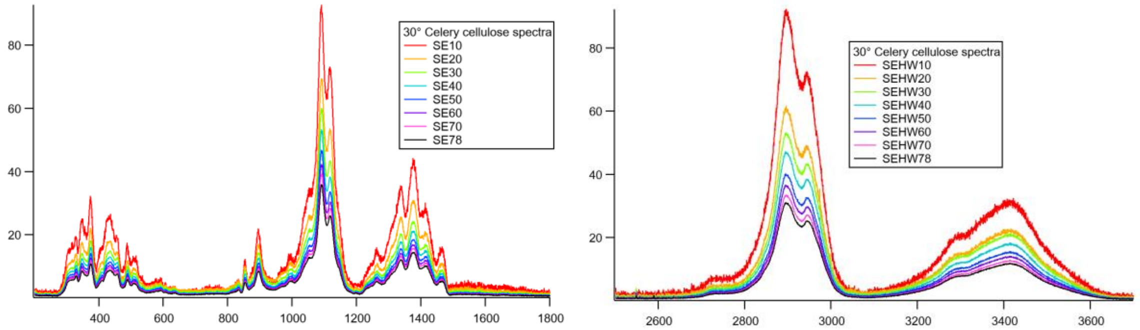
Figure S1.3: Convergence of the 2nd order standard deviation as a function of the number of spectra in the average at 0° in the (a) fingerprint region and (b) high wavenumber region (n=139).

30°



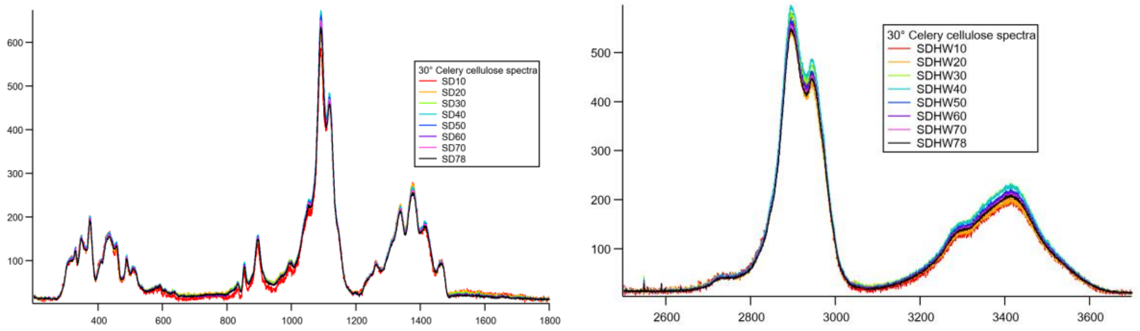
(a) (b)

Figure S1.4: Averaged Raman spectra of celery cellulose at 30° in the (a) fingerprint region and (b) high wavenumber region (n=78).



(a) (b)

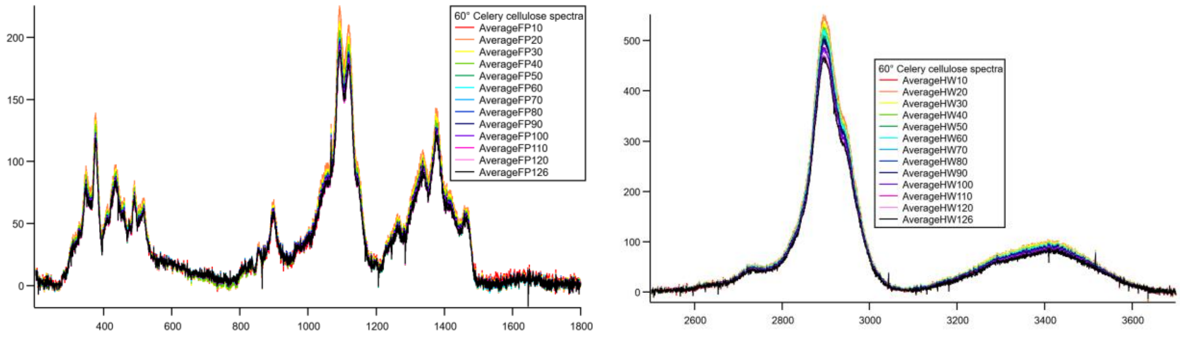
Figure S1.5: Convergence of the standard error of the mean as a function of the number of spectra in the average at 30° in the (a) fingerprint region and (b) high wavenumber region (n=78).



(a) (b)

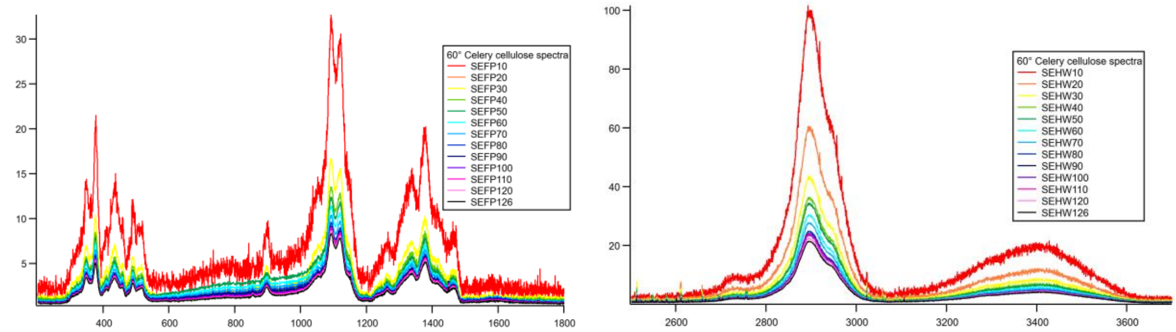
Figure S1.6: Convergence of the 2nd order standard deviation as a function of the number of spectra in the average at 30° in the (a) fingerprint region and (b) high wavenumber region (n=78).

60°



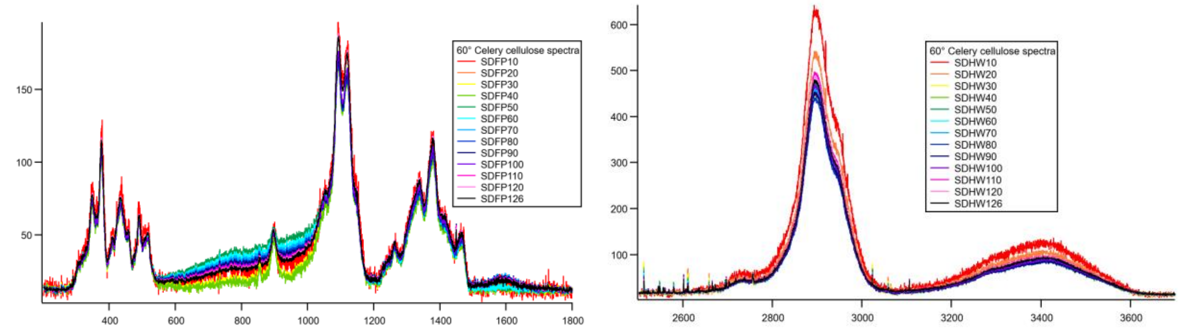
(a) (b)

Figure S1.7: Averaged Raman spectra of celery cellulose at 60° in the (a) fingerprint region and (b) high wavenumber region (n=126).



(a) (b)

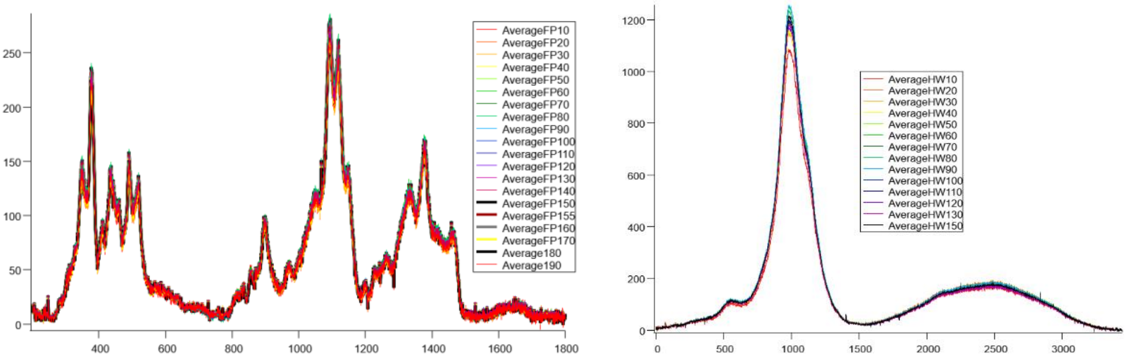
Figure S1.8: Convergence of the standard error of the mean as a function of the number of spectra in the average at 60° in the (a) fingerprint region and (b) high wavenumber region (n=126).



(a) (b)

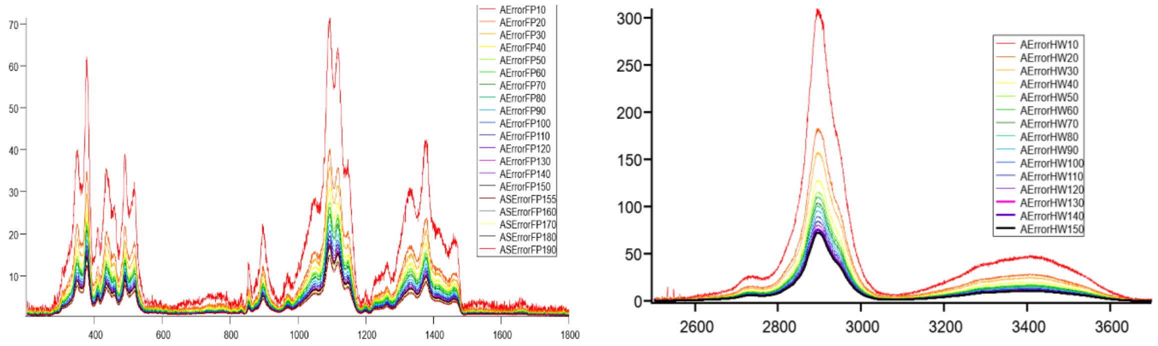
Figure S1.9: Convergence of the 2^{nd} order standard deviation as a function of the number of spectra in the average at 60° in the (a) fingerprint region and (b) high wavenumber region (n=126).

90°



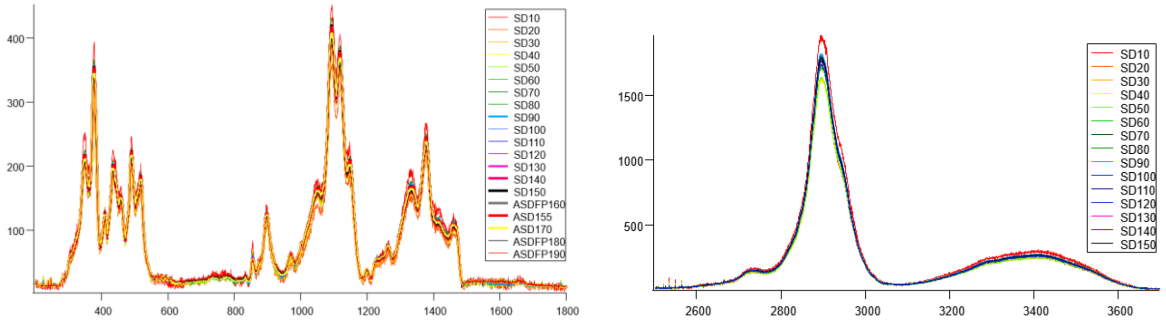
(a) (b)

Figure S1.10: Averaged Raman spectra of celery cellulose at 90° in the (a) fingerprint region (n=190) and (b) high wavenumber region (n=150).



(a) (b)

Figure S1.11: Convergence of the standard error of the mean as a function of the number of spectra in the average at 90° in the (a) fingerprint region (n=190) and (b) high wavenumber region (n=150).



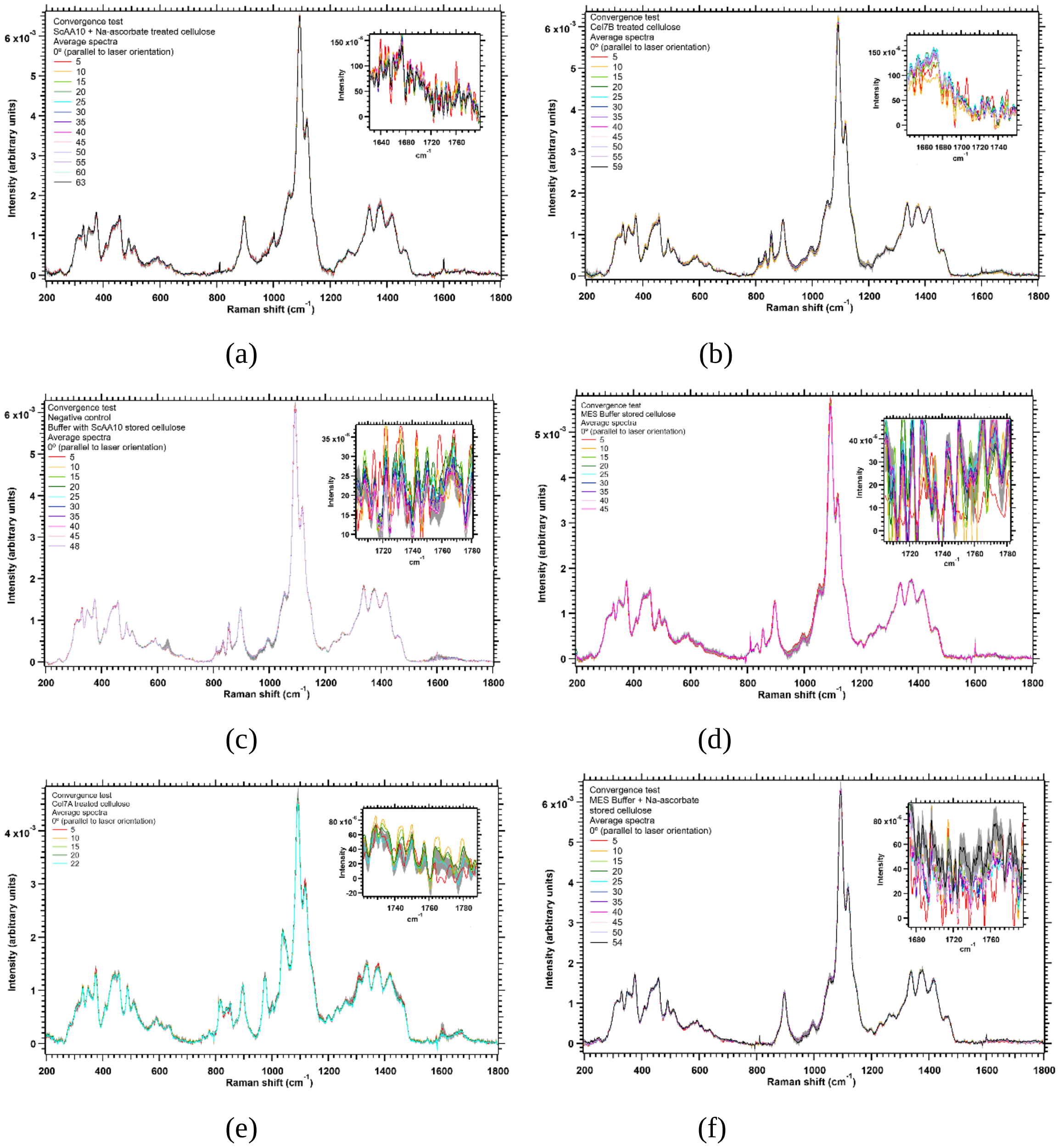
(a) (b)

Figure S1.12: Convergence of the 2^{nd} order standard deviation as a function of the number of spectra in the average at 90° in the (a) fingerprint region (n=190) and (b) high wavenumber region (n=150).

Appendix S2: Convergence test on enzymatically treated cellulose

0°, Fingerprint region

Average Raman spectra



(a)

(b)

(c)

(d)

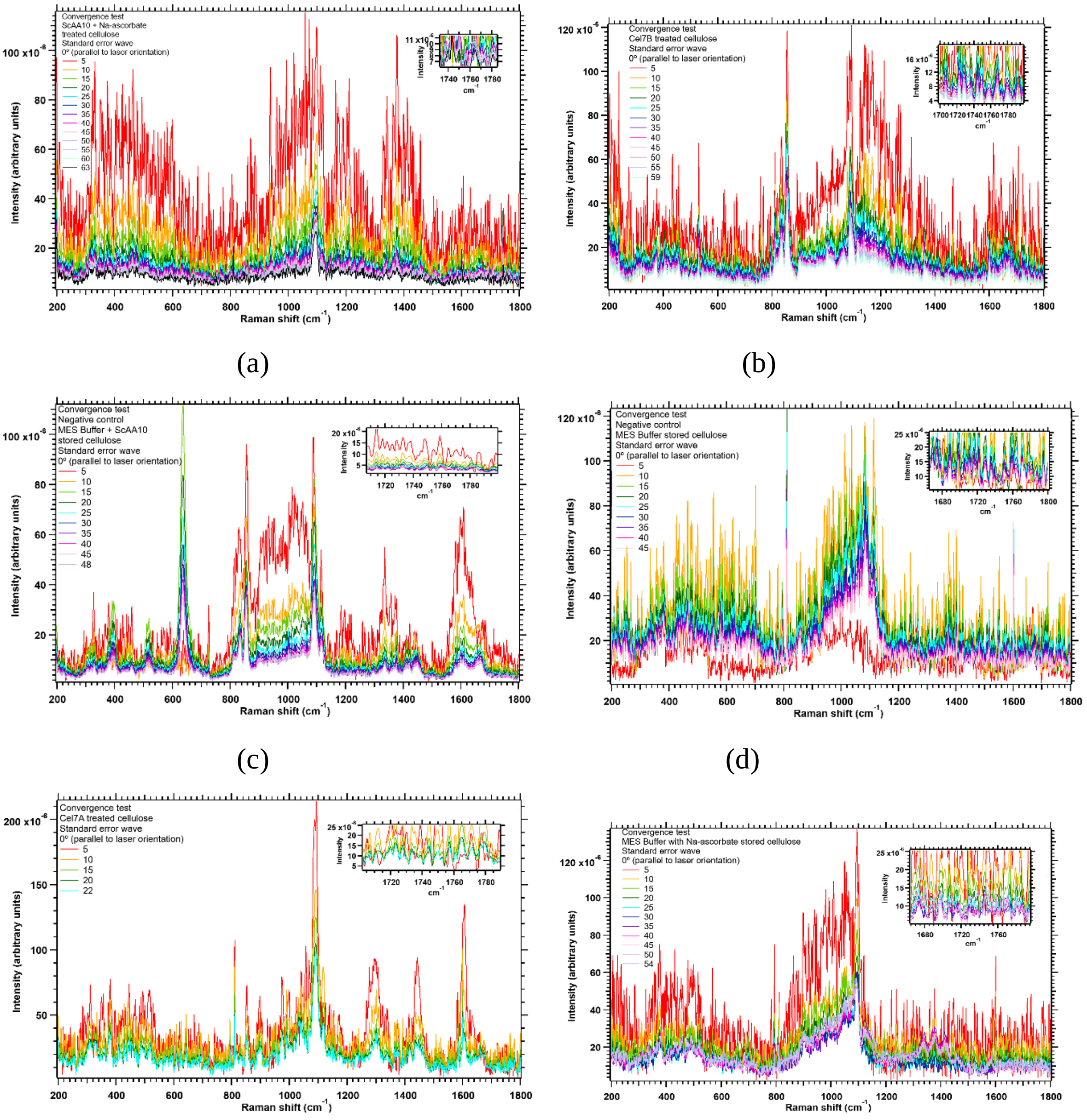
(e)

(f)

Figure S2.1: Convergence test of averaged Raman spectra taken in the fingerprint region (200 – 1800) for dry cellulose after enzymatic treatment. Collection angle of Raman spectra was 0°. Standard error of average Raman spectrum comprising all individual Raman spectra is shown in grey. (a) oxidized cellulose (*Sc*AA10 treatment) n=63, (b) hydrolysed cellulose (Cel7B treatment) n=59, (c) Negative control (buffer + enzyme *Sc*AA10) n = 48, (d) Negative control (buffer) n= 45, (e) hydrolysed cellulose (Cel7A treatment) n=22, (f) Negative control (buffer + Na-ascorbate) n=54. The averaged spectra have been polynomial baseline subtracted, total area normalized, and cubic spline smoothed.

0°, Fingerprint region

Standard error wave



(a) (b)

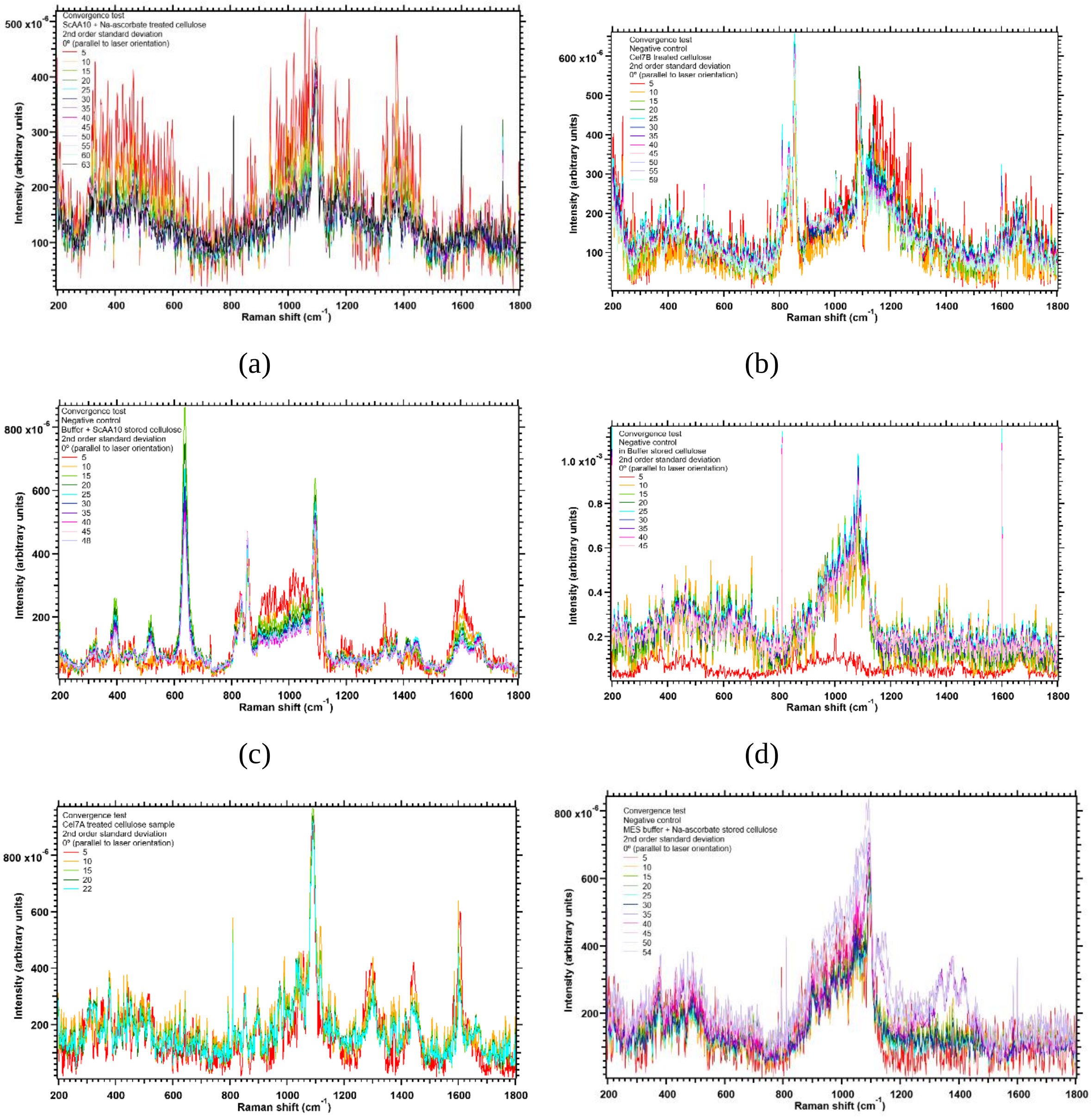
(c) (d)

(e) (f)

Figure S2.2: Convergence test of the standard error of the mean as a function of the number of spectra in the average, as indicated by the legend for the fingerprint region (200 – 1800) for dry cellulose after enzymatic treatment. Collection angle of Raman spectra was 0°. The standard error waves are of the average Raman spectra shown in Figure S1. (a) oxidized cellulose (*Sc*AA10 treatment) n=63, (b) hydrolysed cellulose (Cel7B treatment) n=59, (c) Negative control (buffer + enzyme *Sc*AA10) n = 48, (d) Negative control (buffer) n= 45, (e) hydrolysed cellulose (Cel7A treatment) n=22, (f) Negative control (buffer + Na-ascorbate) n=54. As more spectra are added to the average, the SE can be seen to be decreasing. The spectra are described as converged if the difference between standard error waves is within ~ 1.5% of the intensity of the spectra.

0°, Fingerprint region

2nd order standard deviation



(a) (b)

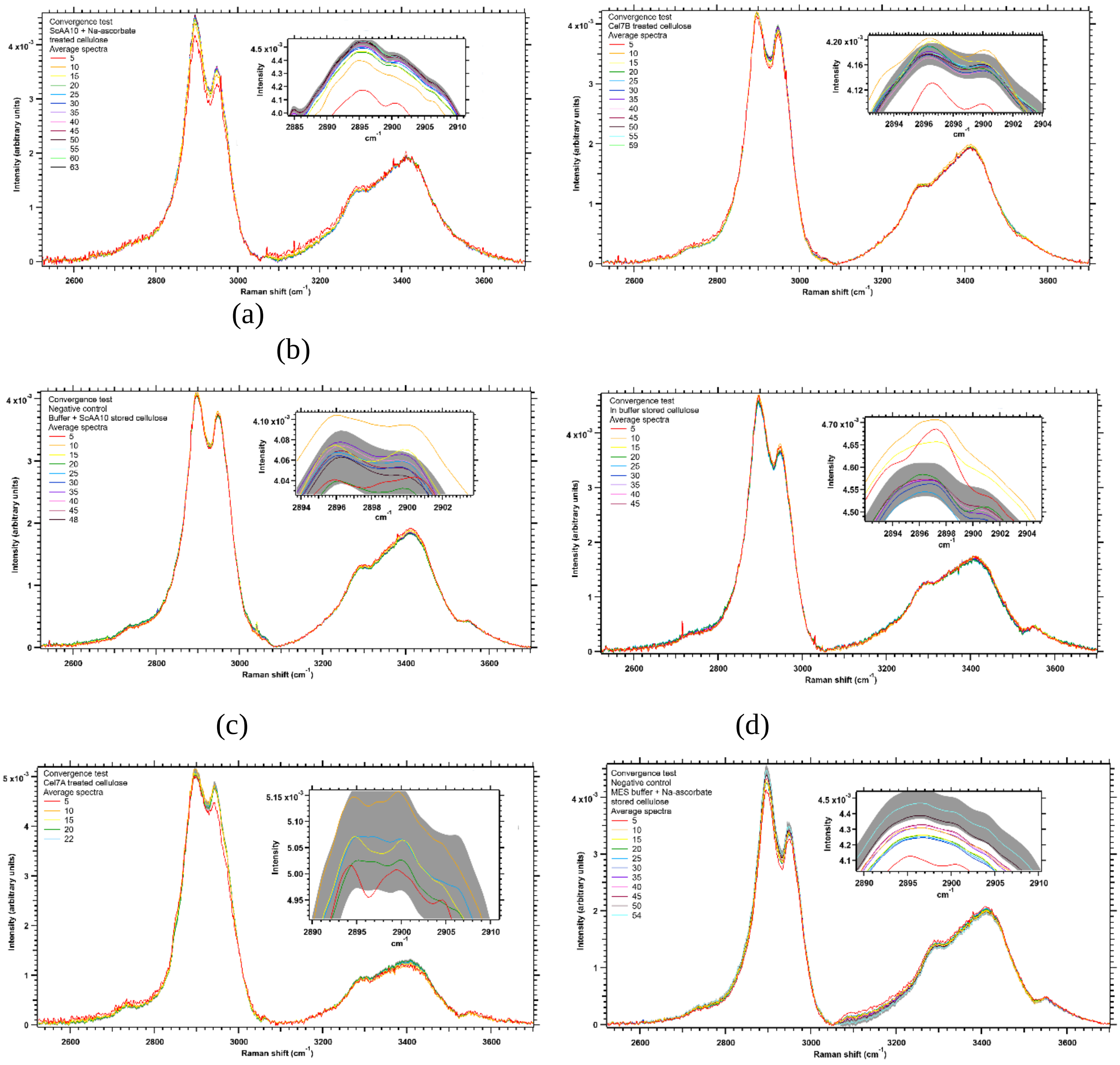
(c) (d)

(e) (f)

Figure S2.3: Convergence test of the 2nd order standard deviation of the mean as a function of the number of spectra in the average, as indicated by the legend for the fingerprint region (200 – 1800) for dry cellulose after enzymatic treatment. Collection angle of Raman spectra was 0°. The 2nd order standard deviation waves are of the average Raman spectra shown in Figure S1. (a) oxidized cellulose (*Sc*AA10 treatment) n=63, (b) hydrolysed cellulose (Cel7B treatment) n=59, (c) Negative control (buffer + enzyme *Sc*AA10) n = 48, (d) Negative control (buffer) n= 45, (e) hydrolysed cellulose (Cel7A treatment) n=22, (f) Negative control (buffer + Na-ascorbate) n=54.

0°, High wavenumber region

Average Raman spectra



(a)

(b)

(c)

(d)

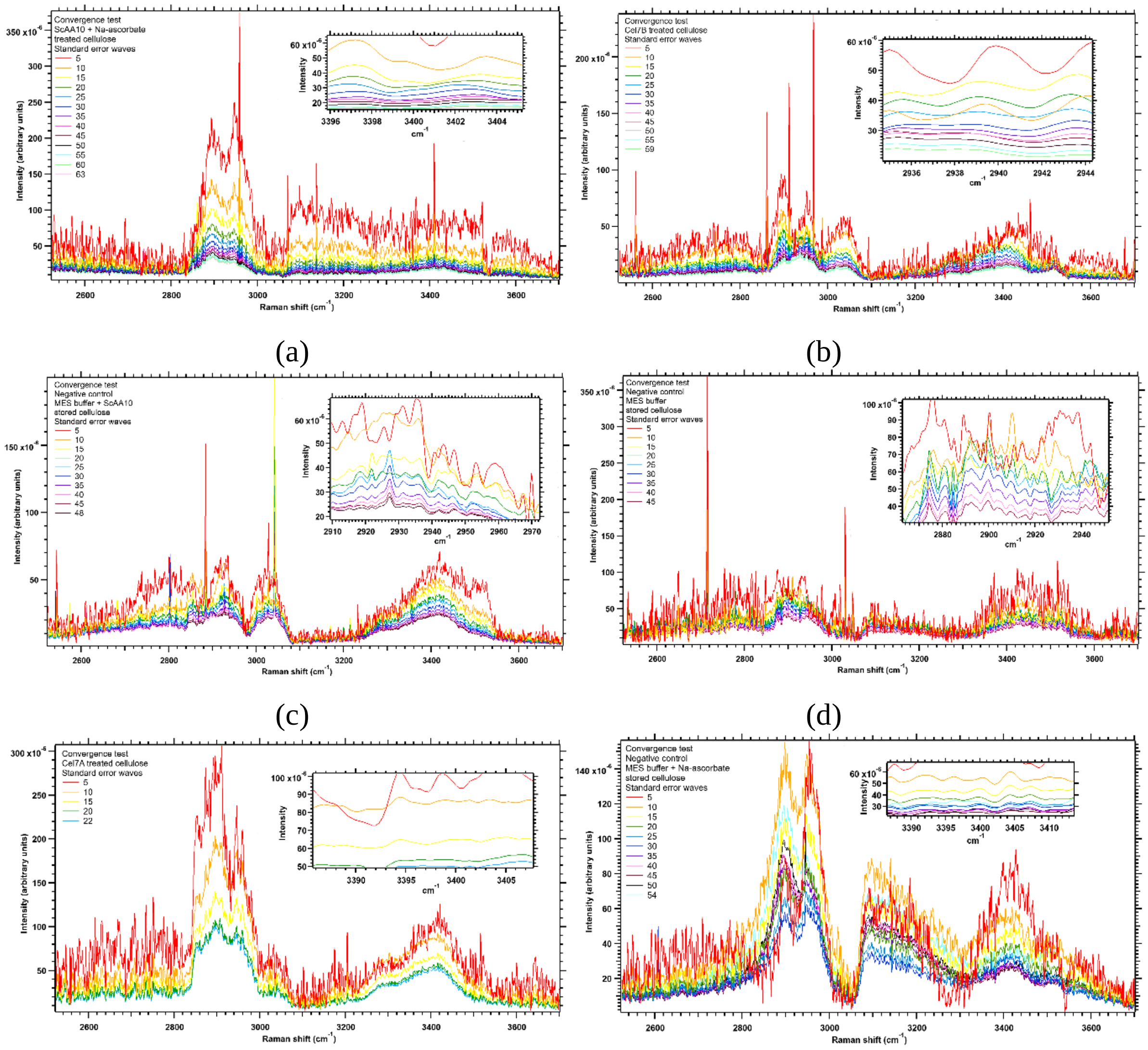
(e)

(f)

Figure S2.4: Convergence test of averaged Raman spectra taken in the high wavenumber region (2500 – 3700) for dry cellulose after enzymatic treatment. Collection angle of Raman spectra was 0°. Standard error of average Raman spectrum comprising all individual Raman spectra is shown in grey. (a) oxidized cellulose (*Sc*AA10 treatment) n=63, (b) hydrolysed cellulose (Cel7B treatment) n=59, (c) Negative control (buffer + enzyme *Sc*AA10) n = 48, (d) Negative control (buffer) n= 45, (e) hydrolysed cellulose (Cel7A treatment) n=22, (f) Negative control (buffer + Na-ascorbate) n=54. The averaged spectra have been polynomial baseline subtracted, total area normalized, and cubic spline smoothed.

0°, High wavenumber region

Standard error wave

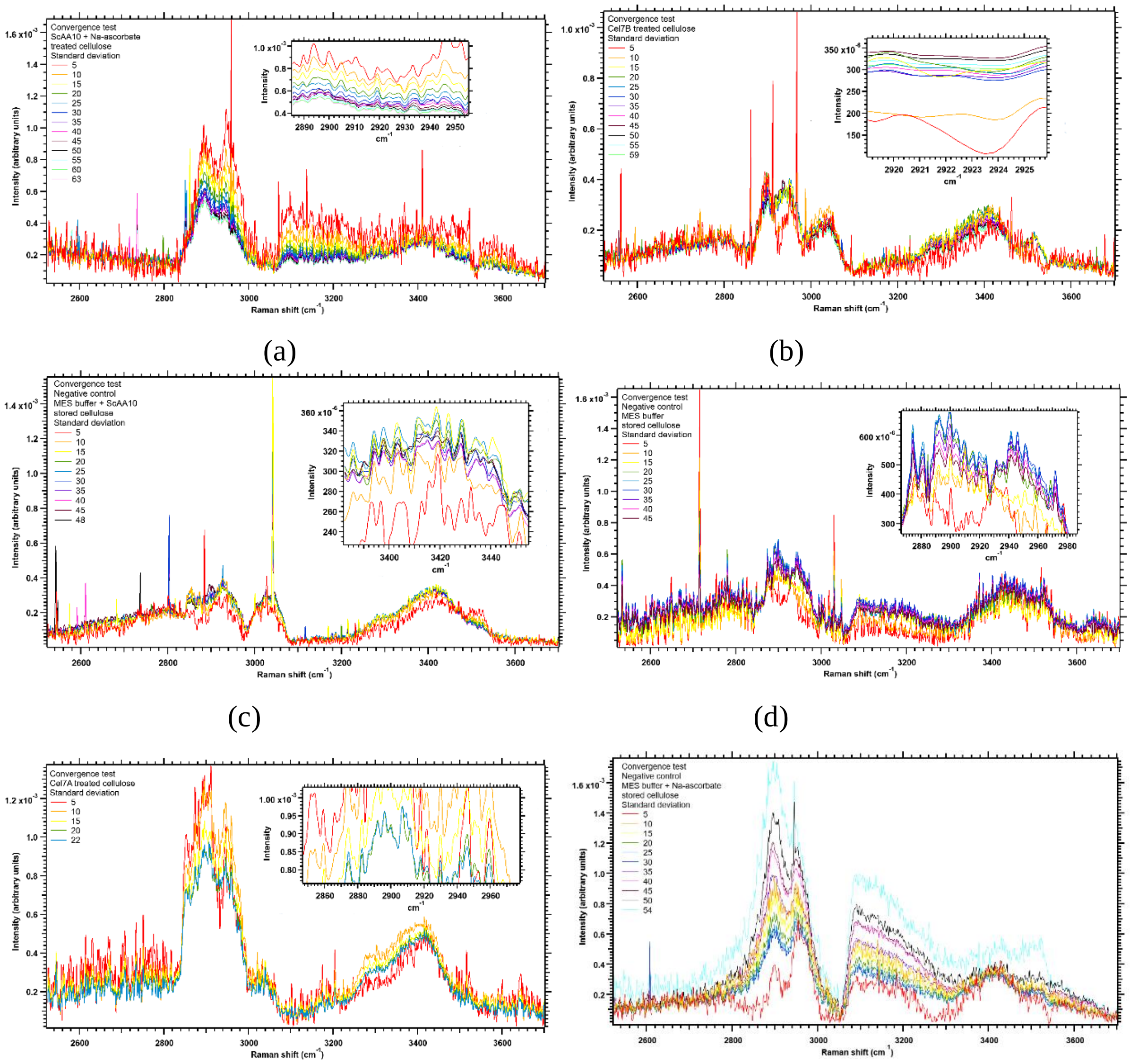


(a) (b) (c) (d) (e) (f)

Figure S2.5: Convergence test of the standard error of the mean as a function of the number of spectra in the average, as indicated by the legend for the high wavenumber region (2500 – 3700) for dry cellulose after enzymatic treatment. Collection angle of Raman spectra was 0°. The standard error waves are of the average Raman spectra shown in Figure S4. (a) oxidized cellulose (*Sc*AA10 treatment) n=63, (b) hydrolysed cellulose (Cel7B treatment) n=59, (c) Negative control (buffer + enzyme *Sc*AA10) n = 48, (d) Negative control (buffer) n= 45, (e) hydrolysed cellulose (Cel7A treatment) n=22, (f) Negative control (buffer + Na-ascorbate) n=54. As more spectra are added to the average, the SE can be seen to be decreasing. The spectra are described as converged if the difference between standard error waves is within ~ 1.5% of the intensity of the spectra.

0°, High wavenumber region

2nd order standard deviation



(a) (b)

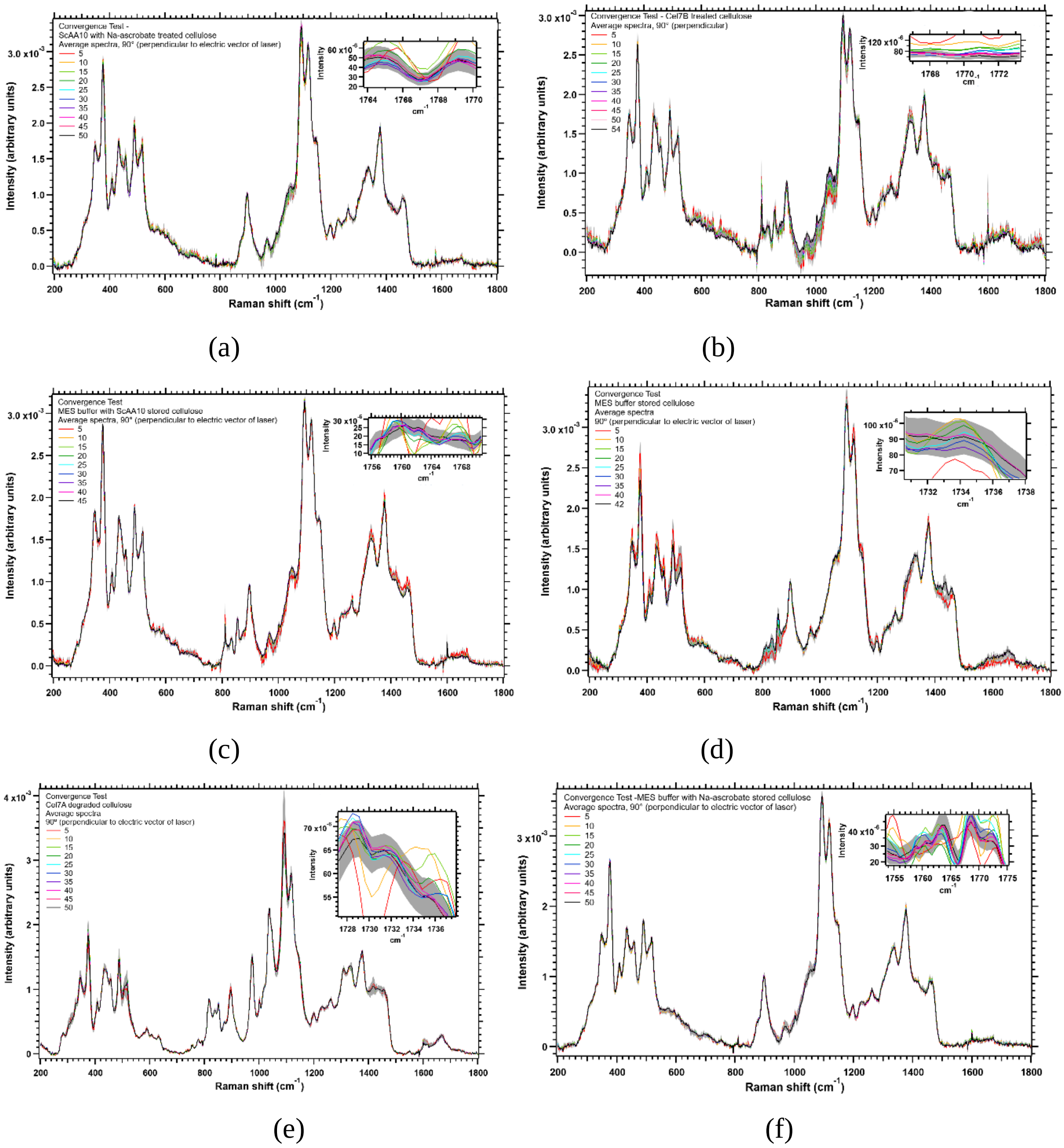
(c) (d)

(e) (f)

Figure S2.6: Convergence test of the 2nd order standard deviation of the mean as a function of the number of spectra in the average, as indicated by the legend for the high wavenumber region (2500 – 3700) for dry cellulose after enzymatic treatment. Collection angle of Raman spectra was 0°. The 2nd order standard deviation waves are of the average Raman spectra shown in Figure S4. (a) oxidized cellulose (*Sc*AA10 treatment) n=63, (b) hydrolysed cellulose (Cel7B treatment) n=59, (c) Negative control (buffer + enzyme *Sc*AA10) n = 48, (d) Negative control (buffer) n= 45, (e) hydrolysed cellulose (Cel7A treatment) n=22, (f) Negative control (buffer + Na-ascorbate) n=54.

90°, Fingerprint region

Average Raman spectra



(a)

(b)

(c)

(d)

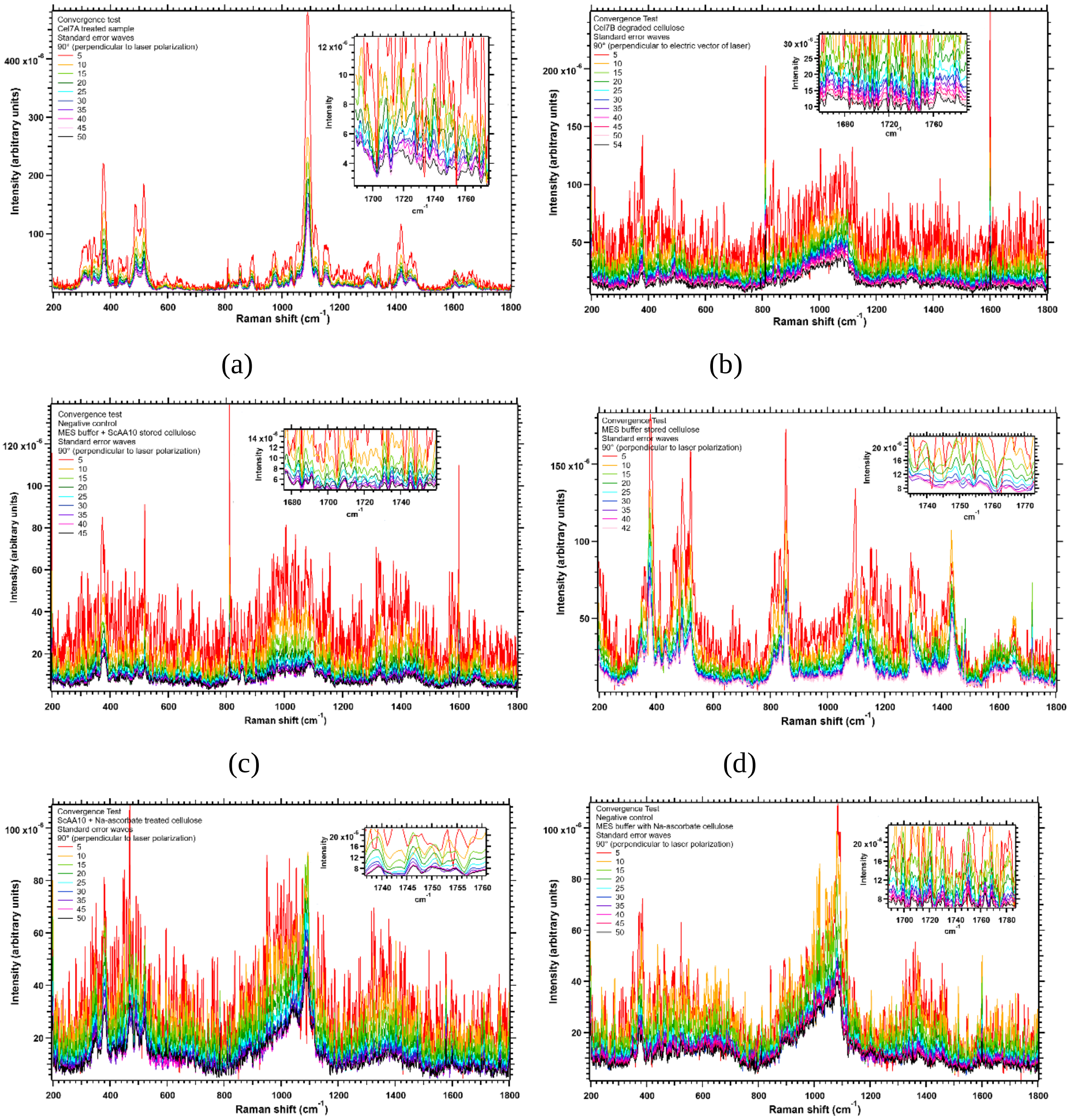
(e)

(f)

Figure S2.7: Convergence test of averaged Raman spectra taken in the fingerprint region (200 – 1800) for dry cellulose after enzymatic treatment. Collection angle of Raman spectra was 90°. Standard error of average Raman spectrum comprising all individual Raman spectra is shown in grey. (a) oxidized cellulose (*Sc*AA10 treatment) n=50, (b) hydrolysed cellulose (Cel7B treatment) n=54, (c) Negative control (buffer + enzyme *Sc*AA10) n = 65, (d) Negative control (buffer) n= 42, (e) hydrolysed cellulose (Cel7A treatment) n=50, (f) Negative control (buffer + Na-ascorbate) n=50. The averaged spectra have been polynomial baseline subtracted, total area normalized, and cubic spline smoothed.

90°, Fingerprint region

Standard error wave



(a) (b)

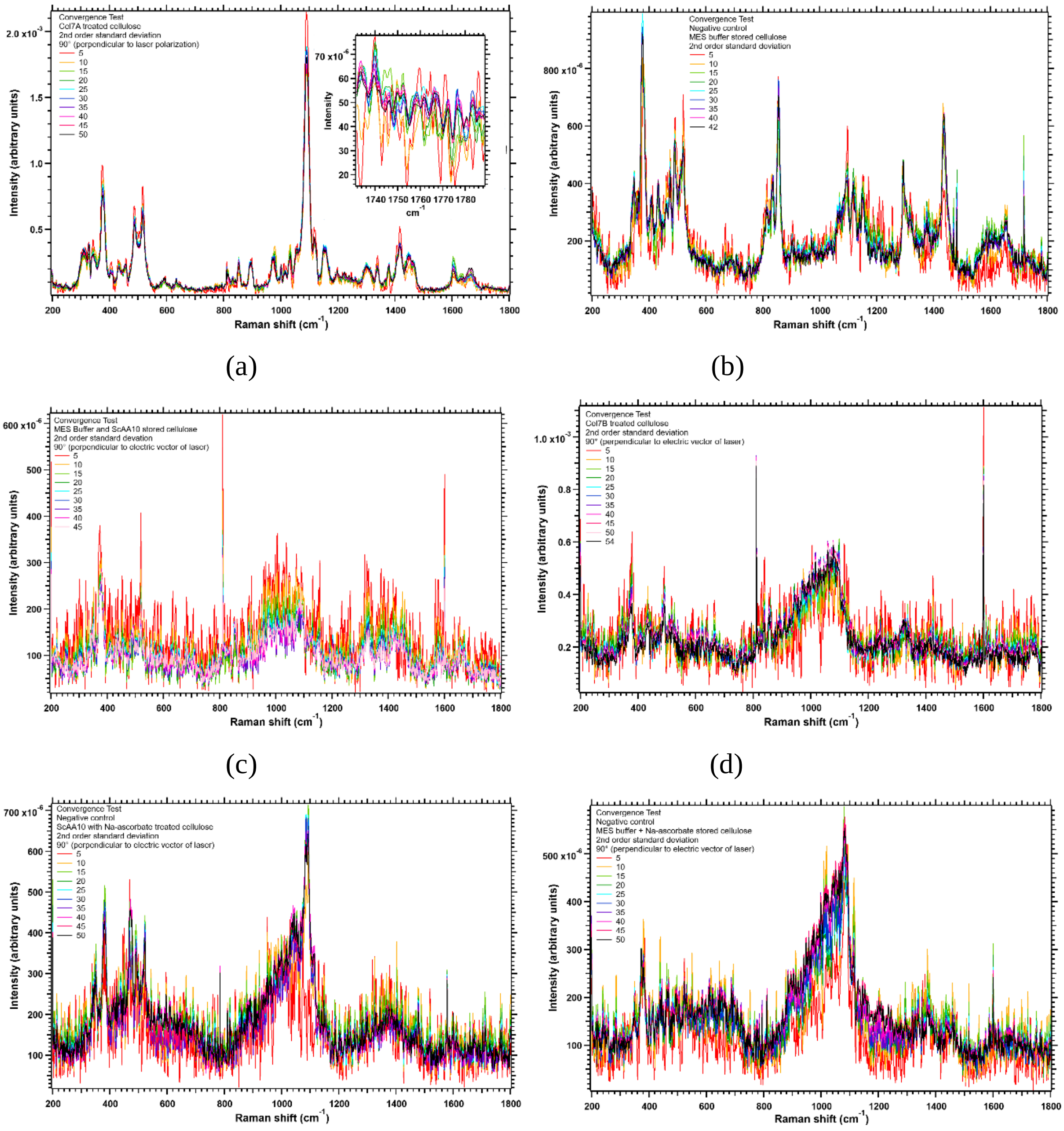
(c) (d)

(e) (f)

Figure S2.8: Convergence test of the standard error of the mean as a function of the number of spectra in the average, as indicated by the legend for the fingerprint region (200 – 1800) for dry cellulose after enzymatic treatment. Collection angle of Raman spectra was 90°. The standard error waves are of the average Raman spectra shown in Figure S7. (a) oxidized cellulose (*Sc*AA10 treatment) n=50, (b) hydrolysed cellulose (Cel7B treatment) n=54, (c) Negative control (buffer + enzyme *Sc*AA10) n = 65, (d) Negative control (buffer) n= 42, (e) hydrolysed cellulose (Cel7A treatment) n=50, (f) Negative control (buffer + Na-ascorbate) n=50. As more spectra are added to the average, the SE can be seen to be decreasing. The spectra are described as converged if the difference between standard error waves is within ~ 1.5% of the intensity of the spectra.

90°, Fingerprint region

2nd order standard deviation



(a)

(b)

(c)

(d)

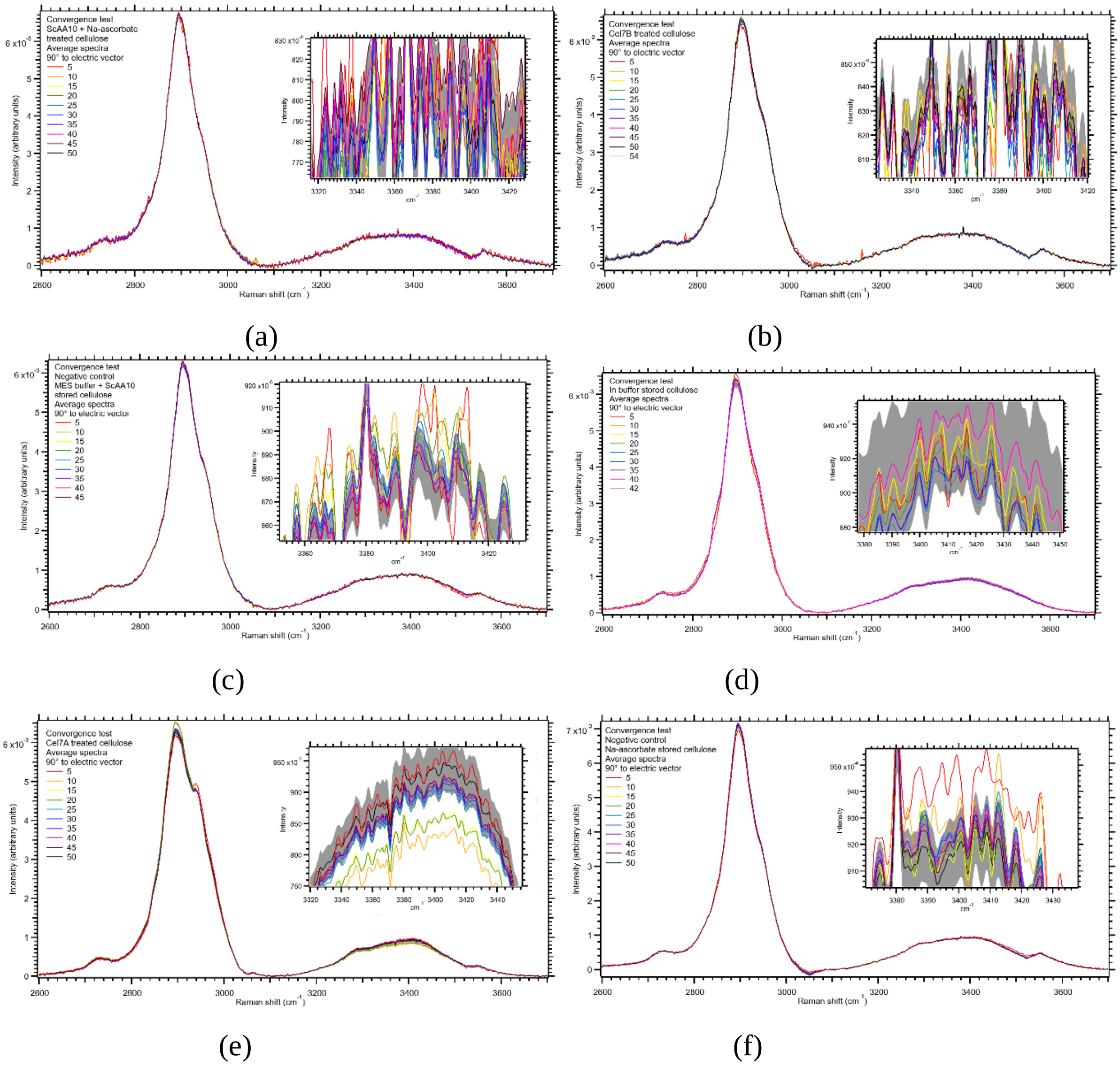
(e)

(f)

Figure S2.9: Convergence test of the 2nd order standard deviation of the mean as a function of the number of spectra in the average, as indicated by the legend for the fingerprint region (2500 – 3700) for dry cellulose after enzymatic treatment. Collection angle of Raman spectra was 90°. The 2nd order standard deviation waves are of the average Raman spectra shown in Figure S7. (a) oxidized cellulose (*Sc*AA10 treatment) n=50, (b) hydrolysed cellulose (Cel7B treatment) n=54, (c) Negative control (buffer + enzyme *Sc*AA10) n = 65, (d) Negative control (buffer) n= 42, (e) hydrolysed cellulose (Cel7A treatment) n=50, (f) Negative control (buffer + Na-ascorbate) n=50.

90°, High wavenumber region

Average Raman spectra

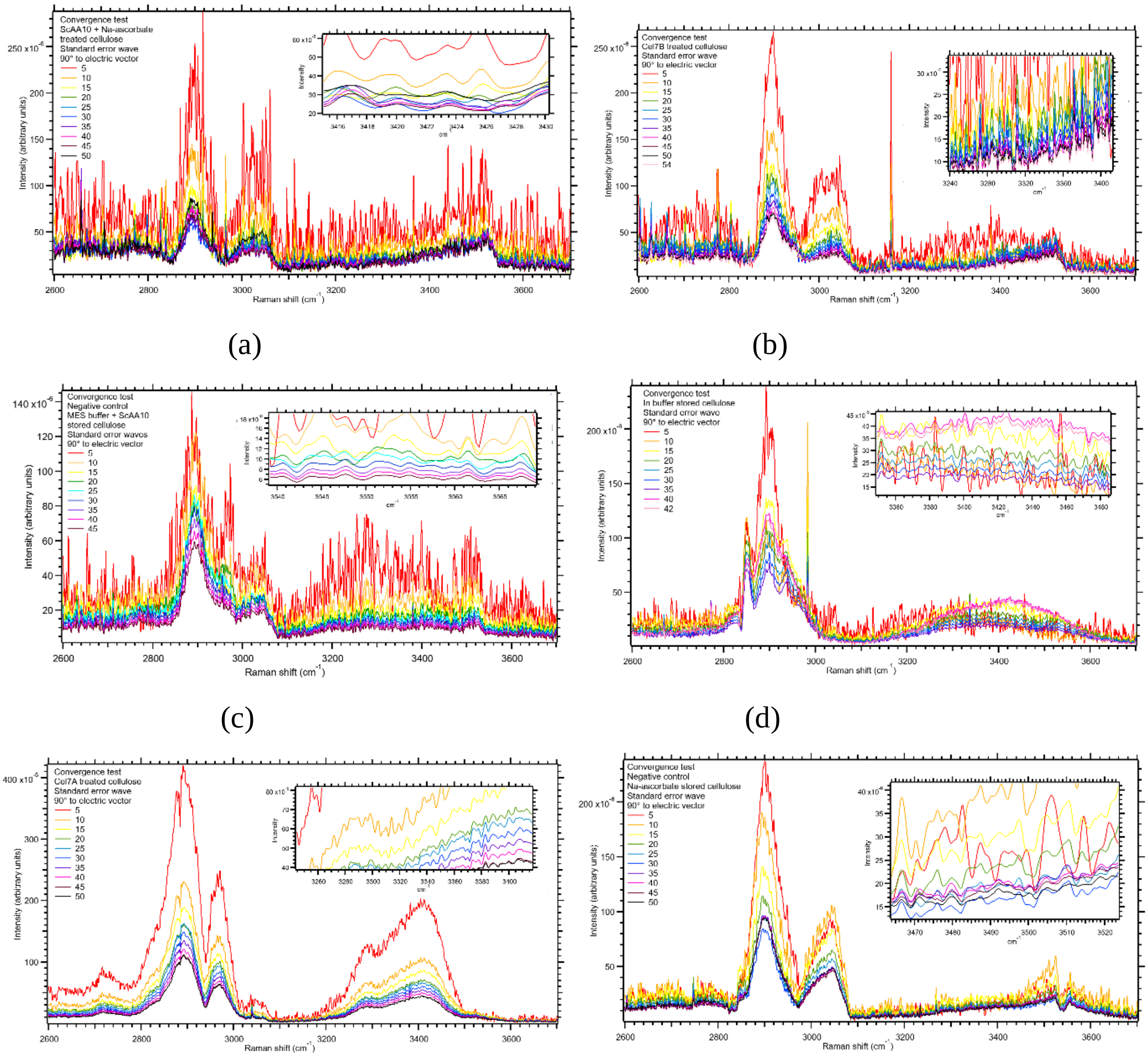


(a) (b) (c) (d) (e) (f)

Figure S2.10: Convergence test of averaged Raman spectra taken in the high wavenumber region (2500 – 3700) for dry cellulose after enzymatic treatment. Collection angle of Raman spectra was 90°. Standard error of average Raman spectrum comprising all individual Raman spectra is shown in grey. (a) oxidized cellulose (*Sc*AA10 treatment) n=50, (b) hydrolysed cellulose (Cel7B treatment) n=54, (c) Negative control (buffer + enzyme *Sc*AA10) n = 65, (d) Negative control (buffer) n= 42, (e) hydrolysed cellulose (Cel7A treatment) n=50, (f) Negative control (buffer + Na-ascorbate) n=50. The averaged spectra have been polynomial baseline subtracted, total area normalized, and cubic spline smoothed.

90°, High wavenumber region

Standard error wave



(a) (b)

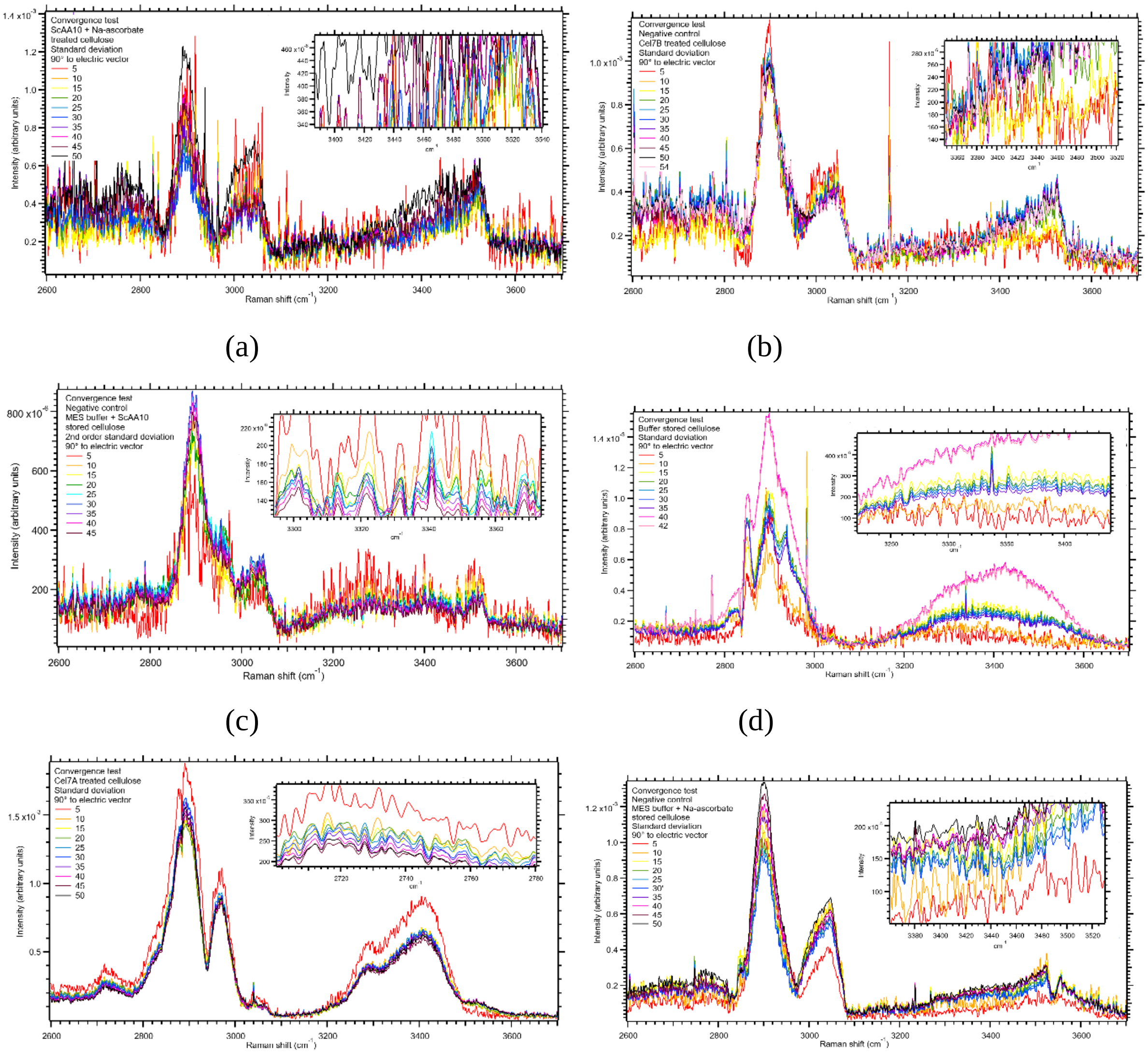
(c) (d)

(e) (f)

Figure S2.11: Convergence test of the standard error of the mean as a function of the number of spectra in the average, as indicated by the legend for the high wavenumber region (2500 – 3700) for dry cellulose after enzymatic treatment. Collection angle of Raman spectra was 90°. The standard error waves are of the average Raman spectra shown in Figure S10. (a) oxidized cellulose (*Sc*AA10 treatment) n=50, (b) hydrolysed cellulose (Cel7B treatment) n=54, (c) Negative control (buffer + enzyme *Sc*AA10) n = 65, (d) Negative control (buffer) n= 42, (e) hydrolysed cellulose (Cel7A treatment) n=50, (f) Negative control (buffer + Na-ascorbate) n=50. As more spectra are added to the average, the SE can be seen to be decreasing. The spectra are described as converged if the difference between standard error waves is within ~ 1.5% of the intensity of the spectra.

90°, High wavenumber region

2nd order standard deviation



(a)

(b)

(c)

(d)

(e)

(f)

Figure S2.12: Convergence test of the 2nd order standard deviation of the mean as a function of the number of spectra in the average, as indicated by the legend for the high wavenumber region (2500 – 3700) for dry cellulose after enzymatic treatment. Collection angle of Raman spectra was 90°. The 2nd order standard deviation waves are of the average Raman spectra shown in Figure S10. (a) oxidized cellulose (*Sc*AA10 treatment) n=50, (b) hydrolysed cellulose (Cel7B treatment) n=54, (c) Negative control (buffer + enzyme *Sc*AA10) n = 65, (d) Negative control (buffer) n= 42, (e) hydrolysed cellulose (Cel7A treatment) n=50, (f) Negative control (buffer + Na-ascorbate) n=50.

References

1. Lin R, Cheng J, Ding L, et al. Bioresource Technology Enhanced dark hydrogen fermentation by addition of ferric oxide nanoparticles using Enterobacter aerogenes. *Bioresour Technol*. 2016;207:213-219. doi:10.1016/j.biortech.2016.02.009

2. Tan H, Corbin KR, Fincher GB. Emerging Technologies for the Production of Renewable Liquid Transport Fuels from Biomass Sources Enriched in Plant Cell Walls. 2016;7(December):1-18. doi:10.3389/fpls.2016.01854

3. Arnold M, Tainter JA, Strumsky D. Productivity of innovation in biofuel technologies. *Energy Policy*. 2019;124(March 2018):54-62. doi:10.1016/j.enpol.2018.09.005

4. Daioglou, V., Stehfest, E., Wicke, B., Faaij, A. and van Vuuren, D.P. (2016), Projections of the availability and cost of residues from agriculture and forestry. GCB Bioenergy, 8: 456-470. https://doi.org/10.1111/gcbb.12285

5. Junginger M, Faaij A, Broek R Van Den, Koopmans A, Hulscher W. Fuel supply strategies for large-scale bio-energy projects in developing countries . Electricity generation from agricultural and forest residues in Northeastern Thailand. 2001;21:259-275. https://doi.org/10.1016/S0961-9534(01)00034-4.

6. Nanomaterials G, Klemm D, Klemm D, et al. Reviews Nanocelluloses : A New Family of Nature-Based Materials Angewandte. 2011:5438-5466. doi:10.1002/anie.201001273

7. Arias-chalico T, Ghilardi A, Masera O. Energy for Sustainable Development Spatial and temporal projection of fuelwood and charcoal consumption in Mexico. *Energy Sustain Dev*. 2014;19:39-46. doi:10.1016/j.esd.2013.11.007

8. Himmel ME, Ding SY, Johnson DK, Adney WS, Nimlos MR, Brady JW, Foust TD. Biomass recalcitrance: engineering plants and enzymes for biofuels production. Science. 2007 Feb 9;315(5813):804-7. doi: 10.1126/science.1137016

9. Klemm D, Heublein B, Fink H, Bohn A. Polymer Science Cellulose : Fascinating Biopolymer and Sustainable Raw Material Angewandte. 2005:3358-3393. doi:10.1002/anie.200460587

10. Dumitriu S, ed. *Polysaccharides Structural Diversity and Functional Versatility*. 2nd Edition. New York; 2005. *J. Am. Chem. Soc.* 2005, 127, 28, 10119 https://doi.org/10.1021/ja0410486

11. Laxmeshwar SS, Kumar DJM, Viveka S, Nagaraja GK. Preparation and Properties of Biodegradable Film Composites Using Modified Cellulose Fibre-Reinforced with PVA. 2012;2012. doi:10.5402/2012/154314

12. Heinze T, Liebert T. Unconventional methods in cellulose functionalization. 2001;26. https://doi.org/10.1016/S0079-6700(01)00022-3

13. Smith MR, Khera E, Wen F. Engineering novel and improved biocatalysts by cell surface display. *Ind Eng Chem Res*. 2015;54(16):4021-4032. doi:10.1021/ie504071f

14. Hemsworth GR, Johnston EM, Davies GJ, Walton PH. Lytic Polysaccharide Monooxygenases in Biomass Conversion. *Trends Biotechnol*. 2015;33(12):747-761. doi:10.1016/j.tibtech.2015.09.006

15. Langan P, Nishiyama Y, Chanzy H. A Revised Structure and Hydrogen-Bonding System in Cellulose II from a Neutron Fiber Diffraction Analysis. 1999;(11):9940-9946. doi:10.1021/ja9916254

16. Nishiyama Y, Langan P, Chanzy H. Crystal Structure and Hydrogen-Bonding System in Cellulose I from Synchrotron X-ray and Neutron Fiber Diffraction. 2002:9074-9082. doi:10.1021/ja0257319

17. Nishiyama Y, Sugiyama J, Chanzy H, Langan P. Crystal Structure and Hydrogen Bonding System in Cellulose I r from Synchrotron X-ray and Neutron Fiber Diffraction. 2003:14300-14306. doi:10.1021/ja037055w

18. Wada M, Chanzy H, Nishiyama Y, Langan P. Cellulose III I Crystal Structure and Hydrogen Bonding by Synchrotron X-ray and Neutron Fiber Diffraction. 2004:8548-8555. doi:10.1021/ma0485585

19. Ralph SA. Cellulose I crystallinity determination using FT – Raman spectroscopy : univariate and multivariate methods. 2010:721-733. doi:10.1007/s10570-010-9420-z

20. Asma A, Anders N, Spiess AC, Baldrian P, Benallaoua S. Biomass and Bioenergy Insights from enzymatic degradation of cellulose and hemicellulose to fermentable sugars – a review. *Biomass and Bioenergy*. 2020;134(February):105481. doi:10.1016/j.biombioe.2020.105481

21. Horn SJ, Vaaje-Kolstad G, Westereng B, Eijsink VG. Novel enzymes for the degradation of cellulose. *Biotechnol Biofuels*. 2012;5(1):45. Published 2012 Jul 2. doi:10.1186/1754-6834-5-45

22. Beeson WT, Vu V V., Span EA, Phillips CM, Marletta MA. Cellulose Degradation by Polysaccharide Monooxygenases. *Annu Rev Biochem*. 2015. doi:10.1146/annurev-biochem-060614-034439

23. Eibinger M, Sattelkow J, Ganner T, Plank H, Nidetzky B. Single-molecule study of oxidative enzymatic deconstruction of cellulose. *Nat Commun*. 2017;8(1):4-10. doi:10.1038/s41467-017-01028-y

24. Eibinger M, Ganner T, Bubner P, et al. Cellulose Surface Degradation by a Lytic Polysaccharide Monooxygenase and Its Effect on Cellulase Hydrolytic. 2014;289(52):35929-35938. doi:10.1074/jbc.M114.602227

25. Chabbert B, Habrant A, Herbaut M, et al. Action of lytic polysaccharide monooxygenase on plant tissue is governed by cellular type. 2017;(July):1-9. doi:10.1038/s41598-017-17938-2

26. Ezeilo UR, Zakaria II, Huyop F, Abdul R. Enzymatic breakdown of lignocellulosic biomass : the role of glycosyl hydrolases and lytic polysaccharide monooxygenases. 2017;2818(May). doi:10.1080/13102818.2017.1330124

27. Contreras F, Pramanik S, Rozhkova AM, Zorov IN, Korotkova O, Sinitsyn AP, Schwaneberg U, Davari MD. Engineering Robust Cellulases for Tailored Lignocellulosic Degradation Cocktails. Int J Mol Sci. 2020 Feb 26;21(5):1589. doi: 10.3390/ijms21051589

28. Loose JSM, Arntzen MØ, Bissaro B, Ludwig R, Eijsink VGH, Vaaje-Kolstad G. Multi-point precision binding of substrate protects LPMOs from self-destructive off-pathway processes. *Manuscr With Minor Revis Submitt*. 2018. doi:10.1021/acs.biochem.8b00484

29. Sun P, Frommhagen M, Haar MK, et al. Mass spectrometric fragmentation patterns discriminate C1- and C4-oxidised cello-oligosaccharides from their non-oxidised and reduced forms. *Carbohydr Polym*. 2020;234(November 2019):115917. doi:10.1016/j.carbpol.2020.115917

30. Ara A. Analytical Methods Combination of MALDI-TOF MS and UHPLC-ESI-MS for the characterization of lytic polysaccharide monooxygenase activity †. 2020:149-161. doi:10.1039/c9ay01774g

31. Keller MB, Felby C, Labate CA, et al. ORIGINAL RESEARCH PAPER A simple enzymatic assay for the quantification of C1-specific cellulose oxidation by lytic polysaccharide monooxygenases. *Biotechnol Lett*. 2020;42(1):93-102. doi:10.1007/s10529-019-02760-9

32. Gierlinger N, Keplinger T, Harrington M, Schwanninger M. Raman imaging of lignocellulosic feedstock. 2013;(November 2019). doi:10.5772/50878

33. Kim SH, Lee CM, Kafle K. Characterization of crystalline cellulose in biomass : Basic principles , applications , and limitations of XRD , NMR , IR , Raman , and SFG. 2013;30(12):2127-2141. doi:10.1007/s11814-013-0162-0

34. Bootten TJ, Harris PJ, Melton LD, Newman RH. Solid-state 13 C-NMR spectroscopy shows that the xyloglucans in the primary cell walls of mung bean (Vigna radiata L .) occur in different domains : a new model for xyloglucan ± cellulose interactions in the cell wall. 2004;55(397):571-583. doi:10.1093/jxb/erh065

35. Frommhagen M, Westphal AH, Hilgers R, et al. Quantification of the catalytic performance of C1-cellulose-specific lytic polysaccharide monooxygenases. 2018:1281-1295. doi:10.1007/s00253-017-8541-9

36. Chalak A, Villares A, Moreau C, et al. Biotechnology for Biofuels Influence of the carbohydrate - binding module on the activity of a fungal AA9 lytic polysaccharide monooxygenase on cellulosic substrates. *Biotechnol Biofuels*. 2019:1-10. doi:10.1186/s13068-019-1548-y

37. Nomura T, Minami E, Kawamoto H. RSC Advances ultraviolet microscopy †. 2020:7460-7467. doi:10.1039/c9ra09435k

38. Agarwal UP. Analysis of Cellulose and Lignocellulose Materials by Raman Spectroscopy: A Review of the Current Status. Molecules. 2019 Apr 27;24(9):1659. doi: 10.3390/molecules24091659

39. Modified C, Using C, Spectroscopy R, Background C. Characterizing Modified Celluloses Using Raman Spectroscopy. 2019:1-5. https://www.horiba.com/fileadmin/uploads/Scientific/Documents/Raman/Specy_Workbench-Characterizing_Modified_Celluloses_Using_Raman.pdf

40. Makarem M, Lee CM, Kafle K, et al. *Probing Cellulose Structures with Vibrational Spectroscopy*. Vol 26. Springer Netherlands; 2019. doi:10.1007/s10570-018-2199-z

41. Cao Y, Lu Y, Huang Y. NIR FT-Raman study of biomass (Triticum aestivum) treated with cellulase. 2004;693:87-93. doi:10.1016/j.molstruc.2004.02.017

42. Ewanick SM, Thompson WJ, Marquardt BJ, Bura R. Real-time understanding of lignocellulosic bioethanol fermentation by Raman spectroscopy. Biotechnol Biofuels. 2013 Feb 20;6(1):28. doi: 10.1186/1754-6834-6-28.

43. Blackwell, J. Spectroscopy R. Infrared and Raman Spectroscopy of Cellulose. 1977:206-218.

44. Konnerth J, Eder M, Fratzl P, Gierlinger N, Luss S, Ko C. Cellulose microfibril orientation of Picea abies and its variability at the micron-level determined by Raman imaging. 2010;61(2):587-595. doi:10.1093/jxb/erp325

45. Simon Webster KJB. Raman Spectroscopy for Pharma - Principles and Applications. *Pharm Technol*. 17(6). https://www.pharmtech.com/view/raman-spectroscopy-pharma-part-1-principles-and-applications.

46. Ferraro JR, Nakamoto K, Brown CW. *Introductory Raman Spectroscopy*.; 2003. doi:10.1002/jrs.1407

47. Zhang X, Chen S, Xu F. Combining Raman Imaging and Multivariate Analysis to Visualize Lignin , Cellulose , and Hemicellulose in the Plant Cell Wall. 2017;(June):1-7. doi:10.3791/55910

48. Butler HJ, Ashton L, Bird B, et al. Using Raman spectroscopy to characterize biological materials. *Nat Protoc*. 2016;11(4):664-687. doi:10.1038/nprot.2016.036

49. Reports A, Raman P. Polarized measurements in Raman microscopy. 2007:326-350. doi:10.1039/b605698a

50. Atalla R. The Structures Of Cellulose. 2014;(March). doi:10.1557/PROC-197-89

51. Tanger P, Field JL, Jahn CE, Defoort MW, Leach JE, Allison GG. Biomass for thermochemical conversion : targets and challenges. 2013;4(July):1-20. doi:10.3389/fpls.2013.00218

52. Brethauer S, Studer MH. Biochemical Conversion Processes of Lignocellulosic Biomass to Fuels and Chemicals – A Review. 2015;69(10):572-581. doi:10.2533/chimia.2015.572

53. Karimi K, Shafiei M, Kumar R. Progress in Physical and Chemical Pretreatment of Lignocellulosic Biomass. In: Gupta VK, Tuohy MG, eds. *Biofuel Technologies: Recent Developments*. Berlin, Heidelberg: Springer Berlin Heidelberg; 2013:53-96. doi:10.1007/978-3-642-34519-7_3

54. Ragauskas AJ. The Path Forward for Biofuels and Biomaterials The Path Forward for Biofuels and Biomaterials. 2006;(February). doi:10.1126/science.1114736

55. Atalla, R.H., & Isogai, A. (2005). Recent Development in Spectroscopic and Chemical Characterization of Cellulose. doi:10.1201/9781420030822.ch5

56. Kennedy CJ, Cameron GJ, Şturcová A, et al. Microfibril diameter in celery collenchyma cellulose: X-ray scattering and NMR evidence. *Cellulose*. 2007;14(3):235-246. doi:10.1007/s10570-007-9116-1

57. Moon RJ, Martini A, Nairn J, Youngblood J, Martini A, Nairn J. *Chem Soc Rev*.; 2011. doi:10.1039/c0cs00108b

58. Kennedy CJ, Cameron ÆGJ, Altaner C, Wess ÆTJ, Jarvis MC. Microfibril diameter in celery collenchyma cellulose : X-ray scattering and NMR evidence Microfibril diameter in celery collenchyma cellulose : X-ray scattering and NMR evidence. 2007;(May). doi:10.1007/s10570-007-9116-1

59. Fenwick KM, Jarvis MC, Apperley DC. Estimation of Polymer Rigidity in Cell Walls of Growing and Nongrowing Celery Collenchyma by Solid-State Nuclear Magnetic Resonance in Vivo. Plant Physiol. 1997 Oct;115(2):587-592. doi: 10.1104/pp.115.2.587

60. Šturcova A, His I, Apperley DC, Sugiyama J, Jarvis MC. Structural details of crystalline cellulose from higher plants. *Biomacromolecules*. 2004;5(4):1333-1339. doi:10.1021/bm034517p

61. Thomas LH, Forsyth VT, Sturcova A, et al. Structure of Cellulose Microfibrils in Primary Cell Walls from Collenchyma. *Plant Physiol*. 2013;161(1):465-476. doi:10.1104/pp.112.206359

62. Chen D, Harris PJ, Sims IM, Zujovic Z, Melton LD. Polysaccharide compositions of collenchyma cell walls from celery (Apium graveolens L.) petioles. *BMC Plant Biol*. 2017;17(1):1-13. doi:10.1186/s12870-017-1046-y

63. Bauer S, Arpa-sancet MP, Finlay JA, Callow ME, Callow JA, Rosenhahn A. Adhesion of Marine Fouling Organisms on Hydrophilic and Amphiphilic Polysaccharides. 2013. doi:10.1021/la3038022

64. Rubin EM. Genomics of cellulosic biofuels. 2008;454(August):841-845. doi:10.1038/nature07190

65. Scheller HV, Ulvskov P. Hemicelluloses. 1871. doi:10.1146/annurev-arplant-042809-112315

66. Jarvis MC. Philosophical Transactions of the Royal Society B Structure of native cellulose microfibrils , the starting point for nanocellulose manufacture. 2018;376(August 2017). doi:10.1098/rsta.2017.0045

67. Other D, Carpita NC. Update on Mechanisms of Plant Cell Wall Biosynthesis : How Plants Make Cellulose and. 2011;155(January):171-184. doi:10.1104/pp.110.163360

68. Beckham GT, Matthews JF, Peters B, Bomble YJ, Himmel ME, Crowley MF. Molecular-Level Origins of Biomass Recalcitrance : Decrystallization Free Energies for Four Common Cellulose Polymorphs. 2011:4118-4127. doi:10.1021/jp1106394

69. Park S, Baker JO, Himmel ME, Parilla PA, Johnson DK. Cellulose crystallinity index: measurement techniques and their impact on interpreting cellulase performance. *Biotechnol Biofuels*. 2010;3:10. Published 2010 May 24. doi:10.1186/1754-6834-3-10

70. Ahvenainen, P., Kontro, I. & Svedström, K. Comparison of sample crystallinity determination methods by X-ray diffraction for challenging cellulose I materials. *Cellulose* **23,** 1073–1086 (2016). https://doi.org/10.1007/s10570-016-0881-6

71. Rongpipi S, Ye D, Gomez ED, Gomez EW, Bartley LE. Progress and Opportunities in the Characterization of Cellulose – An Important Regulator of Cell Wall Growth and Mechanics. 2019;9(March):1-28. doi:10.3389/fpls.2018.01894

72. Aïssa, K. (2020). The potential of using carbohydrate-binding modules (CBMs) for the characterization and modification of cellulose surfaces (T). University of British Columbia. Retrieved from https://open.library.ubc.ca/collections/ubctheses/24/items/1.0388734

73. Atalla RH, C SN. Laser-induced fluorescence in cellulose. *J Chem Soc*. 1972;19:1049-1050. doi: https://doi.org/10.1039/C39720001049

74. Atalla RH, Dimick BE. Preliminary communication l&man-spectral evidence for differences between the conformations cellulose I. 1975;39:7-9. https://doi.org/10.1016/S0008-6215(00)82656-7

75. Chiriu D, Carlo P, Giancarlo R, Marcello C, Giorgia S, Carbonaro CM. Ageing of ancient paper : A kinetic model of cellulose degradation from Raman spectra. 2018;(February):1802-1811. doi:10.1002/jrs.5462

76. Proniewicz LM, Paluszkiewicz C, Weseøucha-birczyn A, Konieczna A. FT-IR and FT-Raman study of hydrothermally degradated cellulose. 2001;596:163-169. doi: 10.1016/S0022-2860(01)00706-2

77. Agarwal U, Crystallinity UC. An Overview of Raman Spectroscopy as Applied to Lignocellulosic Materials Chapter 9 An Overview of Raman Spectroscopy as Applied to Lignocellulosic Materials. 2014;(January 1999).

78. You A, Be MAY, In I. Infrared and Raman Spectra of the Cellulose from the Cell Wall of Valonia ventricosa. 2008;4375(December 2003).

79. Regeneration I, Lattice N. Cellulose: Its Regeneration in the Native Lattice. 185:522-524.

80. Atalla RH, Whitmore RE. Raman Spectral Evidence for Molecular Orientation in Native Cellulosic Fibers. 1980:1717-1719. doi:10.1021/ma60078a066

81. © 1928 Nature Publishing Group. 1928. arxiv.iacs.res.in:8080/jspui/bitstream/10821/877/1/nature 1928.pdf

82. Fenn MB, Xanthopoulos P, Pyrgiotakis G, Grobmyer SR, Pardalos PM, Hench LL. Raman Spectroscopy for Clinical Oncology. 2011;2011. doi:10.1155/2011/213783

83. Hugh D. Young, Roger A. Freedman, Lewis Ford - University Physics with Modern Physics (12th Edition)-Addison Wesley (2007).pdf.

84. Gierlinger N, Luss S, König C, Konnerth J, Eder M, Fratzl P. Cellulose microfibril orientation of Picea abies and its variability at the micron-level determined by Raman imaging. *J Exp Bot*. 2010;61(2):587-595. doi:10.1093/jxb/erp325

85. Cael JJ, Gardner KH, Koenig JL, Blackwell J. Infrared and Raman spectroscopy of carbohydrates . Paper V . Normal coordinate analysis of cellulose I Infrared and Raman spectroscopy of carbohydrates . Paper V Normal coordinate analysis of cellulose I. 2008;1145(September). https://doi.org/10.1063/1.430558

86. Depth Resolution of the Raman Microscope: Optical Limitations and Sample Characteristics. https://www.spectroscopyonline.com/view/depth-resolution-raman-microscope-optical-limitations-and-sample-characteristics.

87. Spatial resolutions. https://www.horiba.com/uk/scientific/products/raman-spectroscopy/raman-academy/raman-faqs/what-is-the-spatial-resolution-of-a-raman-microscope/.

88. Richardd L. McCreery, *Raman Spectroscopy for Chemical Analysis*. © 2000 John Wiley & Sons, Inc. doi:10.1002/0471721646

89. Crut, Aurélien & Maioli, Paolo & Fatti, Natalia & Vallée, Fabrice. (2014). Optical absorption and scattering spectroscopies of single nano-objects. Chemical Society reviews. 43. doi: 10.1039/c3cs60367a.

90. Stenqvist, Björn & Bialik, Erik & Lund, Mikael. (2015). Cellulose-Water Interactions: Effect of electronic polarizability. Nordic Pulp & Paper Research Journal. 30. 026-031. doi: 10.3183/NPPRJ-2015-30-01-p026-031.

91. Eichhorn BBSJ. Elastic coils : deformation micromechanics of coir and celery fibres. 2010:1-11. doi:10.1007/s10570-009-9373-2

92. Mitra S, Fibres A, Chakrabarti K, et al. Ramie : The Strongest Bast Fibre of Nature Ramie : The Strongest Bast Fibre of Nature Sabyasachi Mitra. 2014;(May). doi:10.13140/2.1.3519.5842

93. Schroeder, Leland & Gentile, Victor & Atalla, Rajai. (1986). Nondegradative Preparation of Amorphous Cellulose. Journal of Wood Chemistry and Technology. 6. doi: 10.1080/02773818608085213.

94. Edwards HG, Farwell DW, Webster D. FT Raman microscopy of untreated natural plant fibres. Spectrochim Acta A Mol Biomol Spectrosc. 1997 Nov;53A(13):2383-92. doi: 10.1016/s1386-1425(97)00178-9

95. Schenzel, K., Fischer, S. NIR FT Raman Spectroscopy–a Rapid Analytical Tool for Detecting the Transformation of Cellulose Polymorphs. *Cellulose* **8,** 49–57 (2001). https://doi.org/10.1023/A:1016616920539

96. Zeng Y, Yarbrough JM, Mittal A, Tucker MP, Vinzant TB. In situ label - free imaging of hemicellulose in plant cell walls using stimulated Raman scattering microscopy Biotechnology for Biofuels In situ label - free imaging of hemicellulose in plant cell walls using stimulated Raman scattering microscopy. *Biotechnol Biofuels*. 2016. doi:10.1186/s13068-016-0669-9

97. Wiley, J. H. (1986). Raman Spectra of Celluloses. The University of Puget Sound, Tacoma, WA. Retrieved from https://smartech.gatech.edu/bitstream/handle/1853/5748/wiley_jh.pdf

98. Tuschel D. Practical Group Theory and Raman Spectroscopy , Part II : Application of Polarization. 2014;(March). https://www.spectroscopyonline.com/view/practical-group-theory-and-raman-spectroscopy-part-ii-application-polarization

99. Raman Analysis and Imaging of Cellulose Materials: Wood cell wall and nanocellulose-PP composite. https://www.tappi.org/content/events/08nano/papers/08nan05.pdf.

100. Agarwal UP. Raman imaging to investigate ultrastructure and composition of plant cell walls: Distribution of lignin and cellulose in black spruce wood (Picea mariana). *Planta*. 2006;224(5):1141-1153. doi:10.1007/s00425-006-0295-z

101. Tammer, M. G. Sokrates: Infrared and Raman characteristic group frequencies: tables and charts. *Colloid Polym Sci* 283, 235 (2004). https://doi.org/10.1007/s00396-004-1164-6

102. Arfi Y, Shamshoum M, Rogachev I, Peleg Y, Bayer EA. Integration of bacterial lytic polysaccharide monooxygenases into designer cellulosomes promotes enhanced cellulose degradation. *Proc Natl Acad Sci*. 2014;111(25):9109-9114. doi:10.1073/pnas.1404148111

103. Chundawat SPS, Nemmaru B, Hackl M, et al. Molecular origins of reduced activity and binding commitment of processive cellulases and associated carbohydrate-binding proteins to cellulose III. *J Biol Chem*. 2021;296:100431. doi:10.1016/j.jbc.2021.100431

104. Divne, C., Ståhlberg, J., Teeri, T., & Jones, A. (1998). High-resolution crystal structures reveal how a cellulose chain is bound in the 50 Å long tunnel of cellobiohydrolase I from Trichoderma reesei. *Journal of Molecular Biology*, *275*(2), 309-325. https://doi.org/10.1006/jmbi.1997.1437

105. Breyer WA, Matthews BW. REVIEW A structural basis for processivity. 2001:1699-1711. doi:10.1101/ps.10301.Hyaluron

106. Olsen, Johan & Kari, Jeppe & Windahl, Michael & Borch, Kim & Westh, Peter. (2019). Molecular recognition in the product site of cellobiohydrolase Cel7A regulates processive step length. Biochemical Journal. 477. doi: 10.1042/BCJ20190770

107. Westh, P., Kari, J., Olsen, J., Borch, K., Jensen, K., and Krogh, K. B. R. M. (May 1, 2014) PCT International Patent Appl. WO/2014/064115. C12N 9/00 Ed. https://doi.org/10.1074/jbc.M114.604264

108. Mudinoor AR, Goodwin PM, Rao RU, et al. Biotechnology for Biofuels Interfacial molecular interactions of cellobiohydrolase Cel7A and its variants on cellulose. *Biotechnol Biofuels*. 2020:1-16. doi:10.1186/s13068-020-1649-7

109. Wilson DB, Irwin DC. Genetics and Properties of Cellulases. 2014;(May). doi:10.1007/3-540-49194-5

110. Karlsson J, Momcilovic D, Wittgren B, Tjerneld F, Brinkmalm G. Enzymatic Degradation of Carboxymethyl Cellulose Hydrolyzed by the Humicola insolens and Cel7B , Cel12A and Cel45Acore from Trichoderma reesei. 2002;63:32-40. doi:10.1002/bip.1060

111. Ducros M, Tarling CA, Zechel DL, et al. Anatomy of Glycosynthesis : Structure and Kinetics of the Humicola insolens Cel7B E197A and E197S Glycosynthase Mutants. 2003;10:619-628. doi:10.1016/S

112. Vaaje-Kolstad G, Westereng B, Horn SJ, et al. An Oxidative Enzyme Boosting the Enzymatic Conversion of Recalcitrant Polysaccharides. *Science (80-)*. 2010;330(6001):219-222. doi:10.1126/science.1192231

113. Quinlan RJ, Sweeney MD, Lo Leggio L, et al. Insights into the oxidative degradation of cellulose by a copper metalloenzyme that exploits biomass components. *Proc Natl Acad Sci*. 2011;108(37):15079-15084. doi:10.1073/pnas.1105776108

114. Hemsworth GR, Henrissat B, Davies GJ, Walton PH. Discovery and characterization of a new family of lytic polysaccharide monooxygenases. *Nat Chem Biol*. 2014;10(2):122-126. doi:10.1038/nchembio.1417

115. Vu V V, Beeson WT, Phillips CM, Cate JHD, Marletta MA. Determinants of regioselective hydroxylation in the fungal polysaccharide monooxygenases. *J Am Chem Soc*. 2014;136(2):562-565. doi:10.1021/ja409384b

116. Lo Leggio L, Simmons TJ, Poulsen JCN, et al. Structure and boosting activity of a starch-degrading lytic polysaccharide monooxygenase. *Nat Commun*. 2015;6. doi:10.1038/ncomms6961

117. Couturier M, Ladevèze S, Sulzenbacher G, et al. Lytic xylan oxidases from wood-decay fungi unlock biomass degradation. *Nat Chem Biol*. 2018;14(3):306-310. doi:10.1038/nchembio.2558

118. Sabbadin F, Hemsworth GR, Ciano L, et al. An ancient family of lytic polysaccharide monooxygenases with roles in arthropod development and biomass digestion. *Nat Commun*. 2018;9(1). doi:10.1038/s41467-018-03142-x

119. Vaaje-Kolstad G, Forsberg Z, Loose JS, Bissaro B, Eijsink VG. Structural diversity of lytic polysaccharide monooxygenases. *Curr Opin Struct Biol*. 2017;44:67-76. doi:10.1016/j.sbi.2016.12.012

120. Dell WBO, Agarwal PK, Meilleur F. Oxygen Activation at the Active Site of a Fungal Lytic Polysaccharide Monooxygenase. *Angew Chemie (International ed)*. 2016:1-5. doi:10.1002/anie.201610502

121. Chiu E, Hijnen M, Bunker RD, et al. Structural basis for the enhancement of virulence by viral spindles and their in vivo crystallization. *Proc Natl Acad Sci*. 2015;112(13):201418798. doi:10.1073/pnas.1418798112

122. Vermaas J V, Crowley MF, Beckham GT, Payne CM. Effects of Lytic Polysaccharide Monooxygenase Oxidation on Cellulose Structure and Binding of Oxidized Cellulose Oligomers to Cellulases. 2015. doi:10.1021/acs.jpcb.5b00778

123. Forsberg Z, Mackenzie AK, Sørlie M, et al. Structural and functional characterization of a conserved pair of bacterial cellulose-oxidizing lytic polysaccharide monooxygenases. *Proc Natl Acad Sci U S A*. 2014;111(23):8446-8451. doi:10.1073/pnas.1402771111

124. Kokot S, Czarnik-matusewicz B, Ozaki Y. Two-Dimensional Correlation Spectroscopy and Principal Component Analysis Studies of Temperature-Dependent IR Spectra of Cotton – Cellulose. 2002;67:456-469. doi:10.1002/bip.10163

125. Thomas LH, Forsyth VT, Adriana Š, et al. Structure of Cellulose Microfibrils in Primary Cell Walls. 2013;161(January):465-476. doi:10.1104/pp.112.206359

126. Biology S, Moroz O V, Shaghasi T, Harris P V, Wilson KS, Davies GJ. research communications The three-dimensional structure of the cellobiohydrolase Cel7A from Aspergillus fumigatus ˚ resolution research communications. 2015;(December 2014):114-120. doi:10.1107/S2053230X14027307

127. Lekshmi GS, Tamilselvi R, Geethalakshmi R, et al. Journal of Colloid and Interface Science Multifunctional oil-produced reduced graphene oxide – Silver oxide composites with photocatalytic , antioxidant , and antibacterial activities. *J Colloid Interface Sci*. 2022;608:294-305. doi:10.1016/j.jcis.2021.08.048

128. Spectral resolution. Retrieved from https://www.horiba.com/uk/scientific/products/raman-spectroscopy/raman-academy/raman-faqs/what-factors-affect-spectral-resolution-in-a-raman-spectrometer/.

129. Confocal Laser Scanning Microscope. Retrieved from https://mlm503.weebly.com/index.html.

130. Raman Tool Set. Retrieved from https://sourceforge.net/projects/ramantoolset/.

131. Igor Pro 7. Retrieved from https://www.wavemetrics.com/news/igor-pro-7-released.

132. Yuan X, Mayanovic RA. An Empirical Study on Raman Peak Fitting and Its Application to Raman Quantitative Research. 2017;71(10):2325-2338. doi:10.1177/0003702817721527

133. Gierlinger N, Keplinger T, Harrington M. Imaging of plant cell walls by confocal Raman microscopy. *Nat Protoc*. 2012;7(9):1694-1708. doi:10.1038/nprot.2012.092

134. Bai Y., Liu Q. Denoising Raman spectra by Wiener estimation with a numerical calibration dataset, Biomed. Opt. Express 11, 200-214 (2020). https://doi.org/10.1364/BOE.11.000200

135. Jollife I. T., Principal component analysis. 1986. New York: Springer-Verlag, 1986. Retrieved from http://cda.psych.uiuc.edu/statistical_learning_course/Jolliffe%20I.%20Principal%20Component%20Analysis%20(2ed.,%20Springer,%202002)(518s)_MVsa_.pdf

136. Gautam R, Vanga S, Ariese F, Umapathy S. Review of multidimensional data processing approaches for Raman and infrared spectroscopy. *EPJ Tech Instrum*. 2015. doi:10.1140/epjti/s40485-015-0018-6

137. Suite S, Powerful S. LabSpec 6. Retrieved from https://www.horiba.com/en_en/products/detail/action/show/Product/labspec-6-spectroscopy-suite-software-1843/.

138. Online VA, He S, Zhang W, et al. Analytical Methods Baseline correction for Raman spectra using an improved asymmetric least squares method †. 2014:4402-4407. doi:10.1039/c4ay00068d

139. R Studio Desktop. Retrieved from https://www.rstudio.com/products/rstudio/download/.

140. Kaiser HF. The Application of Electronic Computers to Factor Analysis. *Educational and Psychological Measurement*. 1960;20(1):141-151. doi:10.1177/001316446002000116

141. Gierlinger N, Schwanninger M. Chemical Imaging of Poplar Wood Cell Walls by Confocal Raman Microscopy. 2006;140(April):1246-1254. doi:10.1104/pp.105.066993.1246

142. Szymańska-chargot M, Cybulska J, Zdunek A. Sensing the Structural Differences in Cellulose from Apple and Bacterial Cell Wall Materials by Raman and FT-IR Spectroscopy. 2011:5543-5560. doi:10.3390/s110605543

143. Wiley H, Atua RH. Band assignments. *The Structures of Cellulose.* 1987(June): 151-168. doi:10.1021/bk-1987-0340.ch008

144. Hayakawa D, Nishiyama Y, Mazeau K, Ueda K. Evaluation of hydrogen bond networks in cellulose I b and II crystals using density functional theory and Car e Parrinello molecular dynamics. *Carbohydr Res*. 2017;449:103-113. doi:10.1016/j.carres.2017.07.001

145. Gierlinger N. New insights into plant cell walls by vibrational microspectroscopy. *Appl Spectrosc Rev*. 2017;4928:1-35. doi:10.1080/05704928.2017.1363052

146. Uhlin, K.I., Atalla, R.H. & Thompson, N.S. Influence of hemicelluloses on the aggregation patterns of bacterial cellulose. Cellulose 2, 129–144 (1995). https://doi.org/10.1007/BF00816385

147. Forsberg Z, Vaaje-kolstad G, Westereng B, et al. Cleavage of cellulose by a CBM33 protein. 2011;20:1479-1483. doi:10.1002/pro.689

148. Coenen GJ, Bakx EJ, Verhoef RP, Schols HA, Voragen AGJ. Identification of the connecting linkage between homo- or xylogalacturonan and rhamnogalacturonan type I. 2007;70:224-235. doi:10.1016/j.carbpol.2007.04.007

149. Adriana S, His I, Wess TJ, Cameron G, Jarvis MC. Polarized Vibrational Spectroscopy of Fiber Polymers : Hydrogen Bonding in Cellulose II. 2003:1589-1595. doi:10.1021/bm034295v