Pro gradu -tutkielma
Fysiikan suuntautumisvaihtoehto

Nanoscale Structural Study of Genetically Modified Arabidopsis Thaliana by X-Ray Radiation

Patrik Ahvenainen
30.01.2012

Ohjaaja:	Ritva Serimaa
Tarkastajat:	Ritva Serimaa
	Arto Annila

HELSINGIN YLIOPISTO
FYSIIKAN LAITOS

PL 64 (Gustaf Hällströmin katu 2)
00014 Helsingin yliopisto

HELSINGIN YLIOPISTO — HELSINGFORS UNIVERSITET — UNIVERSITY OF HELSINKI

Tiedekunta/Osasto — Fakultet/Sektion — Faculty	Laitos — Institution — Department
Matemaattis-luonnontieteellinen	Fysiikan laitos

Tekijä — Författare — Author
Patrik Ahvenainen

Työn nimi — Arbetets titel — Title

Nanoskaalan rakenteellinen röntgentutkimus geneettisesti muunnellusta lituruohosta

Oppiaine — Läroämne — Subject
Fysiikka

Työn laji — Arbetets art — Level	Aika — Datum — Month and year	Sivumäärä — Sidoantal — Number of pages
Pro gradu -tutkielma	Tammikuu 2012	44

Tiivistelmä — Referat — Abstract

Selluloosa on yleisin luonnossa esiintyvä materiaali. Tutkimalla geenimuunneltuja kasveja voidaan löytää uusia tapoja käyttää Maapallon rajallisia resursseja ihmiskunnan hyväksi.

Luonnonmateriaalin makroskooppisten ominaisuuksien kannalta on tärkeä tuntea sen rakenne nanometrin tarkkuudella. Materiaalia voidaan tutkia mittaamalla sen kiteisyysastetta ja näiden kiteiden kokoa. Selluloosakiteet ovat kiertyneet soluseinän ympärille pitkiksi kuitumaisiksi mikrofibrilleiksi, jotka muodostuvat kiteisen ja huonommin järjestäytyneen (amorfisen) selluloosan vuorottelevista alueista. Mikrofibrillikulma on se kulma, jonka nämä kuidut muodostavat solun akselin kanssa. Se korreloi materiaalin pitkittäissuuntaisen jäykkyyden kanssa. Laajakulmaröntgensirontamenetelmällä voidaan tutkia näitä nanotason ominaisuuksia näytettä vahingoittamatta.

Arabidopsis thaliana (lituruoho) on tärkeä kasvibiologian mallisyystemi. Se on laajalle levinnyt pieni kukkiva kasvi, jolla on sekä lyhyt elämänkierto että suhteellisen pieni ja kauttaaltaan tunnettu genomi. Röntgenmikrotomografialla otetut kuvat vahvistivat, että *Arabidopsiksen* solut ovat yleisesti ottaen pyöreitä eivätkä suorakulmaisia kuten puusolut. Kasvien solukkorakenne ei tuhoutunut röntgensirontamittauksissa.

Luonnollisen kosteuden omaavia näytteitä ei saatu mitattua riittävän tarkasti, joten kaikki näytteet mitattiin kuivina. Laajakulmaröntgensirontamenetelmällä mitattiin yhteensä 62 näytettä, joista 15 oli villin tyypin kasveja ja loput geneettisesti muunneltuja. Tutkimuksessa jatkokehitettiin tämän menetelmän data-analyysimetodeja heikosti sirottavalle lituruoholle sopiviksi.

Kaikkien mitattujen näytteiden kiteenleveydet olivat välillä 26–30 Å. Havainto sopii sen kanssa, että kiteenleveydet määräytyvät selluloosan biosynteesissä. Kiteisyysasteen vaihteluväli oli noin (20–30)% ja näytesarjojen keskiarvot olivat hiukan yli 25%. Keskimäärainen mikrofibrillikulma puolestaan vaihteli paljon näytteiden välillä aina kahdesta asteesta 21:een asteeseen ja sarjojen keskiarvotkin vaihtelivat enemmän kuin muiden ominaisuuksien. Tilastollisesti merkittäviä eroja näytesarjojen keskiarvoissa löytyi yhden sarjan kohdalla kiteisyysasteissa, yhden keskimääräisissä mikrofibrillikulmissa ja kahden kohdalla kiteenleveyksissä. Pienen näytekoon takia näita eroja voidaan pitää lähinnä jatkotutkimuksia ohjaavina eikä sitovina todisteina geenimuuntelun tuloksista.

Avainsanat — Nyckelord — Keywords
Laajakulmaröntgensironta (WAXS), röntgenmikrotomografia, Arabidopsis thaliana, selluloosa

Säilytyspaikka — Förvaringsställe — Where deposited
Kumpulan tiedekirjasto, Gustaf Hällströmin katu 2 (PL 64), 00014 Helsingin yliopisto

Muita tietoja — övriga uppgifter — Additional information

HELSINGIN YLIOPISTO — HELSINGFORS UNIVERSITET — UNIVERSITY OF HELSINKI

Tiedekunta/Osasto — Fakultet/Sektion — Faculty	Laitos — Institution — Department
Faculty of Science	Department of Physics

Tekijä — Författare — Author
Patrik Ahvenainen

Työn nimi — Arbetets titel — Title

Nanoscale Structural Study of Genetically Modified Arabidopsis Thaliana by X-Ray Radiation

Oppiaine — Läroämne — Subject		
Physics		
Työn laji — Arbetets art — Level	Aika — Datum — Month and year	Sivumäärä — Sidoantal — Number of pages
Pro Gradu thesis	January 2012	44

Tiivistelmä — Referat — Abstract

Cellulose is the most abundant natural material on Earth. New ways to utilize the limited resources on Earth for the benefit of the human kind can be found by studying genetically modified plants.

The nanoscale structure of organic matter is important to its macroscopic properties. Natural materials can be characterized by measuring the degree of crystallinity in the material and the average dimensions of the crystallites. The cellulose crystallites are helically wound as long fibrous microfibrils in which crystalline and less ordered (amorphous) regions alternate. The microfibril angle (MFA) is the angle of these microfibrils with respect to the cell axis and it is linked with the longitudinal stiffness of the material. A wide-angle X-ray scattering (WAXS) method can be used to study these nanoscale properties of the matter in a non-destructive manner.

Arabidopsis thaliana (AT) is an important model system for plant biology. It is a widely spread small flowering plant with a short life cycle and a relatively small genome which has been fully sequenced. X-ray microtomography confirmed that the AT cells are generally round rather than rectangular — unlike xylem cells. The cellular structure of the plants was not destroyed in the WAXS measurements.

The resolution obtained from the measurements of samples with natural humidity was not sufficient hence all the samples were measured dry. A total of 62 samples were measured with the WAXS set-up, 15 of them wild type plants and the rest genetically modified. In this study the WAXS data analysis methods were enchanced for the benefit of the weakly scattering *Arabidopsis* samples.

The mean crystallite width of all samples ranged from 26 to 30 Å. This is consistent with the crystallite width being determined during the biosynthesis of cellulose. The values for the degree of crystallinity ranged from approximately (20–30)% and all the average values were slightly above 25%. The mean microfibril angle varied greatly between the samples, all the way from 2 to 21 degrees. The mean values of different lines varied more for the MFA than for other properties. Statically significant differences between means of different lines were seen for one line in the degree of crystallinity, for one line in the MFA and for two lines in the crystallite width. Due to small sample sizes these differences should be considered mostly as indicative and not as conclusive evidence on the effects of the genetic modification.

Avainsanat — Nyckelord — Keywords
Wide-Angle X-ray Scattering (WAXS), X-ray Microtomography, Arabidopsis Thaliana, Cellulose

Säilytyspaikka — Förvaringsställe — Where deposited
Kumpula Campus Library, Gustaf Hällströmin katu 2 (PO Box 64), 00014 University of Helsinki

Muita tietoja — övriga uppgifter — Additional information

Contents

1 X-ray study of natural materials **3**
- 1.1 Cellulose in natural materials . . . 3
- 1.2 X-ray study of crystalline material . . . 3

2 Theory **5**
- 2.1 X-ray measurements . . . 5
  - 2.1.1 Crystallite size . . . 7
  - 2.1.2 Crystallinity . . . 8
  - 2.1.3 Microfibril angle . . . 9
- 2.2 Test statistics . . . 9

3 Material and methods **11**
- 3.1 Cellulose . . . 11
  - 3.1.1 Cellulose biosynthesis . . . 11
- 3.2 Comparison sample: Aspen . . . 12
  - 3.2.1 Wood structure and composition . . . 12
- 3.3 Model plant: Arabidopsis thaliana . . . 15
  - 3.3.1 General description. . . . 15
  - 3.3.2 Arabidopsis structure and composition . . . 15
  - 3.3.3 Nanoscale structure . . . 16

4 Experiments **17**
- 4.1 Sample preparation . . . 17
- 4.2 Experimental set-up . . . 18
  - 4.2.1 X-ray Microtomography . . . 18
  - 4.2.2 Wide Angle X-Ray Scattering . . . 19
- 4.3 Data analysis . . . 21
  - 4.3.1 Crystallite size . . . 23
  - 4.3.2 Crystallinity . . . 24
  - 4.3.3 Microfibril angle . . . 26
- 4.4 Results . . . 27
  - 4.4.1 X-ray Microtomography . . . 27
  - 4.4.2 Wide Angle X-Ray Scattering . . . 27

5 Discussion **33**
- 5.1 Air scattering . . . 33
- 5.2 Crystallite width approximation . . . 33
- 5.3 Moist samples . . . 34
- 5.4 Beam positioning . . . 35
- 5.5 Beam perpendicularity . . . 35
- 5.6 Scripting in data analysis automatization . . . 37
- 5.7 On Arabidopsis . . . 37

6 Conclusions **38**

Acknowledgements **39**

Bibliography **39**

Glossary **43**

Acronyms **44**

Appendices **45**

1 X-ray study of natural materials

Genetically modified plants may grow faster and bigger and be more resistant to pests and disease than their natural counterparts. However successful generic engineering calls for means and methods to quantify any induced effects. The X-ray method serves us to study plant structures quantitatively at nanometer resolution. When the entire genome of the plant has been sequenced and the plant has been thoroughly studied, it is easier to quantify the effects of the genetic modifications. For this reason we examined the nanostructure of *Arabidopsis thaliana* (*Arabidopsis*). Later these results may be used as a reference for other plants that have more direct agricultural or other socio-economic impact such as wood and grain. For instance X-ray scattering methods have been used to examine the paper-making properties of wood for industrial purposes.

1.1 Cellulose in natural materials

Cellulose is the most abundant biological material on Earth [1]. For this reason alone the study of this material has large economic and social impact. It is also a renewable energy source and a global carbon sink [2]. In wood cellulose is located mostly in the secondary cell wall and comprises 40% to 50% of dry wood mass [3, p. 54]. The cellulose molecules form microfibrils, long fiber-like strands in which crystalline (highly ordered) regions and amorphous (less ordered) regions alternate [3, p. 54]. The orientation and properties of the microfibrils and the crystalline properties of the material have vital significance to the structure of the living organism.

This study focuses on the flowering plant *Arabidopsis*. It is a model system for plant biology and its genome has been fully sequenced. In this study the *Arabidopsis* samples are compared to aspen samples to emphasize the similarity and differences in their structures as well as in the way X-ray methods can be used to examine them. The aspen samples represent common dry wood samples. The methods for their analysis are well-established at the Division of Materials Physics, Department of Physics, University of Helsinki [4–7]. The *Arabidopsis* samples in turn are smaller and weaker scatterers and therefore pose different challenges in their data analysis and measurements.

The *Arabidopsis* plant has a shorter life cycle than trees rendering it more suitable for scientific studies. Furthermore *Arabidopsis* is easy to grow in laboratory.

1.2 X-ray study of crystalline material

Electromagnetic radiation can resolve structural features down to its wavelength. The micrometer range can be accessed with visible light (some 400 to 700 nm in wavelength).

Detection of details of this resolution suffices for the function of human vision but the study of nanoscale structures necessitates smaller wavelengths of electromagnetic radiation. X-ray wavelengths are in the Ångström region (0.1 nm).

Crystalline material has by definition a periodically repeating structure. This lattice structure can be examined by X-ray scattering methods in a non-destructive way. The study of scattering to large angles is called Wide Angle X-ray Scattering (WAXS) in contrast to Small Angle X-ray Scattering (SAXS).

Macroscopic properties of natural materials correlate with many micrometer and nanometer scale structures. For example the X-ray scattering methods have been used to study the Microfibril Angle (MFA) in cellulosic fibers for over 75 years [8]. The MFA has been linked to the longitudinal stiffness of wood [9] as well as other properties that are important in paper-making [3, p. 15]. Thus the wood quality analysis by X-ray scattering has traditions in commercial applications.

The X-ray study of small plants has the advantage that only a little sample preparation — if any — is needed [a]. Thus the same sample can be analyzed later by any other, including invasive and destructive, method.

[a] The physical size of the sample related to its scattering properties is the determining factor on whether the sample needs to be cut thinner. A thick sample will absorb most of the radiation intensity and induce theoretical problems in the analysis. On the other hand a sample too thin will yield weak scattering intensities.

2 Theory

2.1 X-ray measurements

Interactions of X-rays with matter. High-energy photons interact mostly with the electrons in the matter. The most relevant interactions to X-ray measurements are elastic (Rayleigh) scattering, inelastic (Compton) scattering and photoelectric absorption. Other processes such as pair production happen only with high energies where the incident photon has more than twice the energy of the rest mass of the particle (and antiparticle). In photoelectric absorption the entire energy of the incident photon is transferred to an electron that escapes the target atom. In Compton scattering only a part of the energy is transferred, and the trajectory of the X-ray photon will change. Customarily it is said that the photon scatters from the atom. If the photon does not lose any energy in the impact, the interaction is called elastic scattering. Traditional X-ray diffraction methods are based on elastic scattering whereas absorption X-ray microtomography is based on the photoelectric absorption. Scattering interactions are also seen with microtomography but in standard absorption tomography they serve only to create background noise.

X-ray radiation can be produced in an X-ray tube in which electrons are emitted from the cathode material and accelerated with a high voltage towards the anode. When the electrons hit the anode material they will both emit the Bremsstrahlung radiation and raise the target electrons to higher energy levels (higher electron shells). When these electrons return to lower energy levels, they will emit a photon whose energy is equal to the energy difference between the two electron shells. This characteristic energy is seen as sharp peaks in the emission spectra that are called characteristic lines. If the acceleration voltage is U_0 then the maximum Bremsstrahlung radiation energy is $E_{BS} = eU_0$. Most of the energy generated is lost as heat in the X-ray tube that typically requires water-cooling. The characteristic lines depend on the anode material, which is typically tungsten, molybdenum or copper. These characteristic lines can be used to select a very narrow wavelength spectrum with a monochromator.

X-ray scattering. The lattice spacing d_{hkl} for a given X-Ray Diffraction (XRD) wavelength λ [b] determines the scattering angle 2θ as is given by Bragg's law [11] (Figure 1)

$$2d_{hkl} \sin\theta = \lambda. \tag{1}$$

[b] Copper Kα radiation of wavelength $\lambda = 1.541\,\text{Å}$ was used in the X-ray diffraction measurements. Kα energy is emitted by the electrons moving from the second lowest energy shell (L) to the lowest energy shell (K).

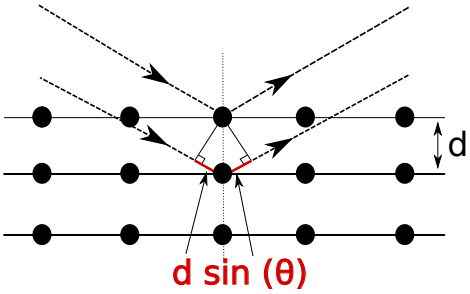


Figure 1. Bragg's law (Eq. (1)) originates from the elastic scattering of the incident photons (dotted lines) from the lattice plane. When $2d_{hkl}\sin\theta$ is equal to the wavelength λ, constructive interference will be observed.

The magnitude of the scattering vector q is here defined as

$$q = \frac{4\pi \sin\theta}{\lambda}, \tag{2}$$

where the Bragg angle θ is half the scattering angle (2θ) [10].

Bragg's law (Eq. (1)) and the definition of the scattering vector (Eq. (2)) for a reflection from the hkl-plane give the relation between the lattice spacing d_{hkl} and the magnitude of the scattering vector q_{hkl} by

$$q_{hkl} = \frac{4\pi \sin\theta}{\lambda} = \frac{4\pi \sin\theta}{2d_{hkl}\sin\theta} = \frac{2\pi}{d_{hkl}}. \tag{3}$$

Perpendicular transmission geometry. In the WAXS measurements, perpendicular transmission geometry was used (Figure 2). The advantage of this geometry over symmetrical transmission and symmetrical reflection geometries is that the whole Two-Dimensional (2D) diffraction pattern can be measured at the same time.

X-ray absorption. When X-ray photons propagate through a material, part of their energy will be absorbed in the material. The intensity loss (dI) in an infinitesimal distance dz is proportional to the intensity I and the distance through

$$dI = -I\mu(z)dz \tag{4}$$

where the proportionality coefficient $\mu(z)$ is called the linear attenuation coefficient. By integrating Eq. (4) we get

$$\int dI/I = \int \mu(z)dz$$

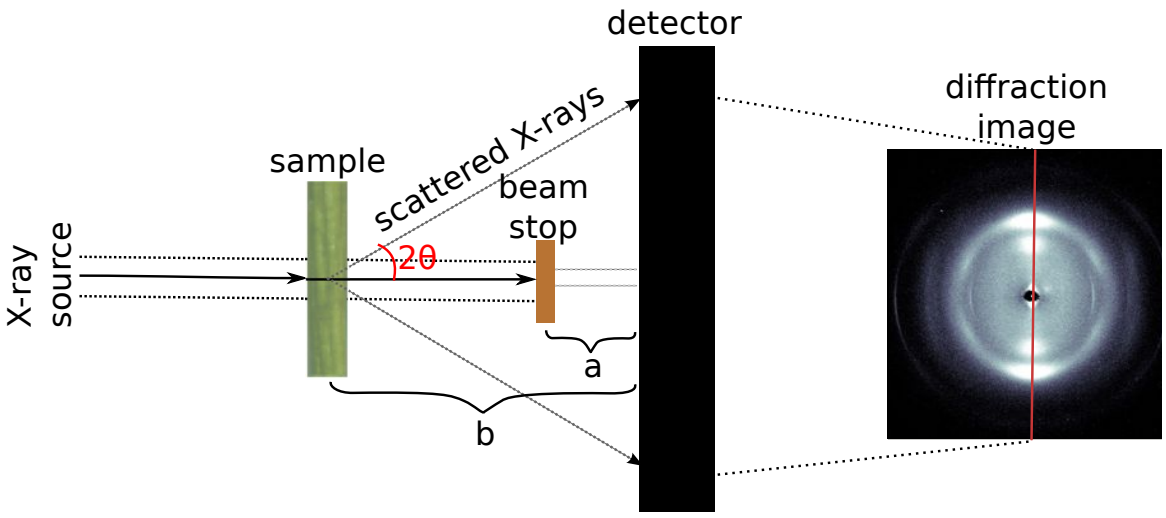


Figure 2. A schematic cross-sectional view of the perpendicular transmission geometry, viewed from above. The red line shows where the cross-sectional intensity is represented in the two-dimensional diffraction pattern. The beam stop to detector distance (a) was 35 mm and the sample to detector distance (b) was 95 mm. The figure shows a collimated beam, but the monochromatic X-ray beam can also be focused on the detector.

and when we assume a constant μ and set the initial intensity at zero distance to I_0 ($I(z = 0) = I_0$) we get the well known Beer-Lambert law for absorption:

$$I(z) = I_0 e^{-\mu z}. \tag{5}$$

2.1.1 Crystallite size

An infinitely large crystal would yield a delta peak for each lattice reflection in the 2D diffraction image. The finite size of the crystal causes widening of the reflection peaks.

The average crystallite size L is thus related to the width of the diffraction peak. This relation has been traditionally formulated as the Scherrer equation [12, 13]

$$L = \frac{K\lambda}{B\cos\theta}, \tag{6}$$

where K is a shape factor constant (approx. 0.9 [c]), λ is the X-ray wavelength, B is the Full Width at Half Maximum (FWHM) of the diffraction peak and θ is the Bragg angle (half of the scattering angle 2θ). The size L is given in the direction perpendicular to the scattering planes [10].

[c] The shape factor K of the cellulose 200-reflection is 0.9 [4]. Other shape factors are around unity. The shape factor value is based on the assumption of a Gaussian reflection peak.

Actual instruments cannot measure delta peaks but they always have some finite resolution, which is not caused by the sample but stem from hardware limitations. This effect is called the instrumental broadening effect. It will be taken into account by calculating the FWHM as $B = \sqrt{B_{meas}^2 - B_{inst}^2}$, where B_{meas} is the measured peak FWHM and B_{inst} is the instrumental broadening factor [d]. In other words, the FWHM B in Eq. (6) represents real peak broadening caused by the finite crystallite size and not by the instrumentation artifacts. Here we have to assume that both the 200-peak and the instrumental broadening peak are approximately Gaussian [14, 15].

The strongest reflection for cellulose is 200 [7]. It will be used to determine the crystallite width. Moreover the 200-reflection does not over-lap significantly with other reflections [7]. The reason for this will be explained in subsection 3.1, on p. 11.

2.1.2 Crystallinity

The degree of crystallinity of the sample C can be estimated from the scattering intensity from the entire sample and from the fitted amorphous part of the sample [15] as

$$C = 1 - \frac{\int I_{amorf}}{\int I_{sample}}, \tag{7}$$

where I_{sample} is the total intensity from the sample integrated in 2θ-scale and I_{amorf} is the amorphous background intensity. They are both integrated over the same 2θ range and using the same azimuthal sector, typically 180°. Similar equation has been used to approximate crystallinity in secondary xylem of Norway spruce [15] [e]. If the intensity values were equidistant, a simple sum would be sufficient, but as they are not equidistant in the 2θ scale numerical integration is needed.

Note that scattering from absorbed water in the sample (intensity I_{water}) and from air (intensity I_{air}) are not of interest. Hence these backgrounds should be subtracted from the measured intensity I_{meas} components. The scattering intensity of the sample is thus

$$I_{sample} = I_{meas} - I_{water} - I_{air}. \tag{8}$$

[d] A value of 0.3° (in 2θ scale) was used for the instrumental broadening factor. It can be estimated by measuring the FWHM of a diffraction peak originating from a very large crystal that should have a very narrow peak.

[e] The equation in question was a sum over intensity values rather than an integral. When the intensity values in the 2θ range are equidistant then the sum over the intensity values is the same as an integral over the values.

The numerical integration of Eq. (7) can be done by using the trapezoidal rule,

$$\int_a^b f(x) \approx (b-a)\frac{f(a)+f(b)}{2} \tag{9}$$

where we use a linear approximation in each interval $[a, b]$.

2.1.3 Microfibril angle

The 200-reflection was used to determine the MFA for the same reasons as it was used to determine the crystallite width. In reciprocal space the reflections are seen as rings that intersect the Ewald sphere. For a cell with rectangular cross section (e.g. aspen) and non-zero MFA the rings from all four cell facets are resolved in the reciprocal space. Consequently eight spots are seen on the Ewald sphere. These are reproduced in the real space 2D detector as intensity peaks at the same scattering angle 2θ but with different azimuthal angles ϕ. This can be used to determine the MFA provided that the cells have rectangular cross section. [7] If the cross section is more round than rectangular, other analysis methods have to be employed.

2.2 Test statistics

We are interested in the expectation value $\langle X \rangle$ of the parameter X based on $\langle x \rangle$ measured from our sample of size n. The null hypothesis is that $H_0 : \langle X \rangle = \langle x \rangle$. Assuming that X is normally distributed with variance σ, the transformation

$$Z = \frac{\langle x \rangle - \langle X \rangle}{\sigma/\sqrt{n}} \tag{10}$$

yields a (0,1) normally distributed variable [16] ($Z \sim N(0, 1)$, mean 0, variance 1). Values of cumulative distributive function for given value of Z define a level of significance to reject the null hypothesis ($\langle X \rangle \neq \langle x \rangle$). This is a two-sided test. We can similarly conduct a one-sided test to conclude whether the expectation value from our sample is larger than from the 0-population ($H_0 : \langle X \rangle \leq \langle x \rangle$) or smaller ($H_0 : \langle X \rangle \geq \langle x \rangle$).

In case the variance of the 0-population is unknown — like in most real measurements — we can estimate the 0-population variance with the standard deviation of the sample s [16]. This test statistic is now given by [16, p. 65]

$$t = \frac{\overline{x} - \langle X \rangle}{s/\sqrt{n}} \tag{11}$$

where $\overline{x}$ is the sample [f] mean, s is the sample standard deviation and n is the sample size.

[f] As in sample from the entire population of interest.

With a large sample size n this is also a normal distribution [g] assuming the null hypothesis. For small values of n one must use Student's t-distribution with number of degrees of freedom $n-1$ [16, p. 65] [17, p. 170]. The null hypothesis here is that the population mean is equal to the given value $\langle X \rangle$ in the case of a two-sided t-test. Once this t-value is determined from Eq. (11), a p-value is obtained from the appropriate Student's t-distribution. The p-value represents the probability that under the null hypothesis a value as or more extreme than the test statistic can be observed and it is used to report the evidence against the null hypothesis. If this p-value is below the chosen threshold the null hypothesis is rejected. Typical significance values are 0.1, 0.05 and 0.01 which can be considered the threshold values for weak, strong and very strong evidence against the null hypothesis [17, p. 157].

The two-sample t-test can be used to test whether two populations have the same mean. In this case the t-value is given by [16, p. 66]

$$t = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{\hat{s}_1^2/n_1 + \hat{s}_2^2/n_2}} \tag{12}$$

where $\overline{x_i}$ is the sample mean, n_i is the sample size and $\hat{s}_i$ is the unbiased estimator of the variance of the sample. In practice $\hat{s}_i$ is replaced by the standard deviation of the sample s_i and the degrees of freedom are estimated so that the distribution of the test statistic can be approximated by the ordinary Student's t distribution where the number of degrees of freedom is no longer $n_1 + n_2 - 2$ [16, p. 66]. The effective degrees of freedom $\mathcal{F}$ can be approximated by for example the Welch-Satterthwaite equation [18] [16, p. 66] [h]

$$\mathcal{F} = \frac{\left(s_1^2/n_1 + s_2^2/n_2\right)^2}{\frac{\left(s_1^2/n_1\right)^2}{(n_1-1)} + \frac{\left(s_2^2/n_2\right)^2}{(n_2-1)}}. \tag{13}$$

If we have two samples with the same sample size ($n_1 = n_2 = n$) Eq. (13) reduces to

$$\mathcal{F} = \frac{\left(s_1^2 + s_2^2\right)^2 (n-1)}{s_1^4 + s_1^4} \tag{14}$$

and if we further have the case of equal variances $s_1 = s_2 = s$ we see that our approximation is accurate:

$$\mathcal{F} = \frac{\left(s^2 + s^2\right)^2 (n-1)}{s^4 + s^4} = \frac{\left(4s^4\right)(n-1)}{2s^4} = 2(n-1). \tag{15}$$

[g] Formally we can state that $\lim_{n\to\infty} t = Z \sim N(0,1)$

[h] In [16, p. 66] the equations 30 and 31 are the same as Eq. (13), just expressed in different format and without a reference to the original paper published in 1947 [18].

3 Material and methods

3.1 Cellulose

It bears repeating that cellulose is the most abundant natural material on Earth [1]. In nature crystalline cellulose is found in the cellulose I allomorph in two phases, Iα and Iβ [7, 19]. A big concentration of the Iβ allomorph is found in wood and cotton whereas Iα is most abundant in some algae and bacteria [19]. Other allomorphs of cellulose are man-made or rare compared to cellulose I and therefore are not discussed in this work.

The cellulose Iβ cell is monoclinic and belongs to space group $P2_1$ (Figure 3b) [20]. The unit cell parameters are $a = 7.78\,\text{Å}$, $b = 8.20\,\text{Å}$, $c = 10.38\,\text{Å}$ and the unit cell angles are $\alpha = \beta = 90.0°$ and $\gamma = 96.5°$ [20]. The cellulose Iα cell on the other hand is triclinic and belongs to space group $P1$ (Figure 3a) [21]. The respective unit cell parameters are $a = 6.72\,\text{Å}$, $b = 5.96\,\text{Å}$, $c = 10.40\,\text{Å}$ and the unit cell angles are $\alpha = 118.1°$, $\beta = 114.8°$ and $\gamma = 80.4°$ [21]. These two crystal phases (Iα and Iβ) can be found along the same microfibril [21, 22]. The relative amounts of these two phases vary by the cellulose origin. Some cellulose I has been found both in nearly pure [i] Iα and Iβ consistencies [21]. The (theoretical) diffraction patterns of the two phases are very similar (Figure 4). Cellulose Iβ is thermodynamically more stable than Iα [22] which can be converted to Iβ by annealing [19].

An infinitely large lattice would produce distinct delta peaks for each reflection that could be easily separated (Figure 4). However, the finite size of the crystallites causes broadening of the diffraction peaks as given by the Scherrer equation (Eq. (6)). At the crystallite size of 27 Å (FWHM of 3.0) the 200-reflection peak is virtually the only easily recognizable peak. The 110- and $1\bar{1}0$-reflection peaks overlap and the 004-reflection peak is overlapped by other peaks (Figure 5). The 200-reflection and 004-reflection planes for cellulose Iβ are shown in Figure 6.

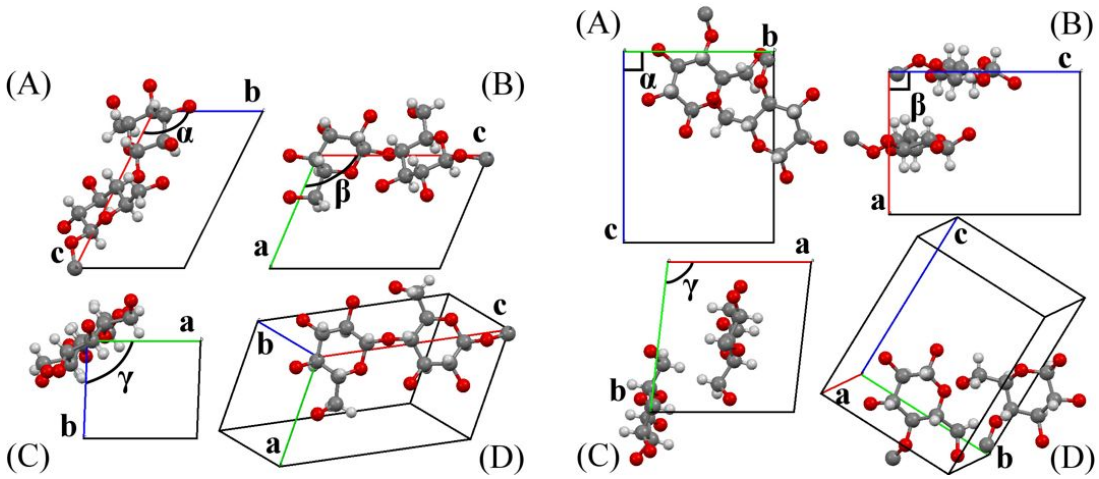
It has been suggested that the microfibrils of angiosperms contain only a single cellulose crystallite. Cellulose crystallite sizes have been reported between $(2.1 \times 2.4)\,\text{nm}^2$ and $(3.2 \times 3.1)\,\text{nm}^2$ in cross-sectional dimensions. [24]

3.1.1 Cellulose biosynthesis

The elementary cellulose microfibrils will be formed when the cellulose synthase catalytic subunit (cesA) chains emerge from the rosette terminal complexes [j] at the plasma membrane of the cell. The rosette has a six-fold symmetry and a 25 to 30 nm diameter. The subunits are believed to contain six synthetic units that each polymerize a glucan chain. These chains

[i] Nearly pure refers here to approximately 90%.

[j] Rosette is a structure observed only in vascular plants such as *Arabidopsis*. However, analogous structures referred as terminal complexes (TC) have been observed in all cellulose-synthesizing organisms.



(a) Cellulose Iα (tricilinc)

(b) Cellulose Iβ (monoclinic)

Figure 3. Cellulose I phases. The individual views shown are (A) along the a-axis, (B) along the b-axis, (C) along the c-axis and (D) in an arbitrary 3D angle. The lattice is defined by the length of the unit cell vectors $\vec{a}$, $\vec{b}$ and $\vec{c}$ and the angles α, β and γ between them [23]. Colors indicate atom type; red: oxygen (O), dark grey: carbon (C) and light grey: hydrogen (H). Chemical formula for Cellulose I is $C_{12}H_{20}O_{10}$ [20, 21]. Figure drawn with Mercury from the data provided in [20, 21].

form the elementary microfibrils. This process determines the cross-sectional dimensions of the microfibrils and explains why the cellulose crystallite width is very plant-specific. [19, 25]

3.2 Comparison sample: Aspen

The thickness of the sample determines the intensity of the scattering. As the thickness of the sample increases the likelihood of multiple scattering increases. Multiple scattering violates the assumptions of X-ray scattering theory whereas the thickness of the sample affects the measured intensity and thus it is an important practical consideration.

3.2.1 Wood structure and composition

Wood structure. The center of the wood is called a pith [k] and concentric year rings are formed around it by seasonal variations of earlywood and latewood. The inner part of the tree changes to a completely dead heartwood over time. The outer part which is responsible for the ascent of sap is called sapwood. [3, 4]

[k] Formed during the first year of growth [3, p. 2].

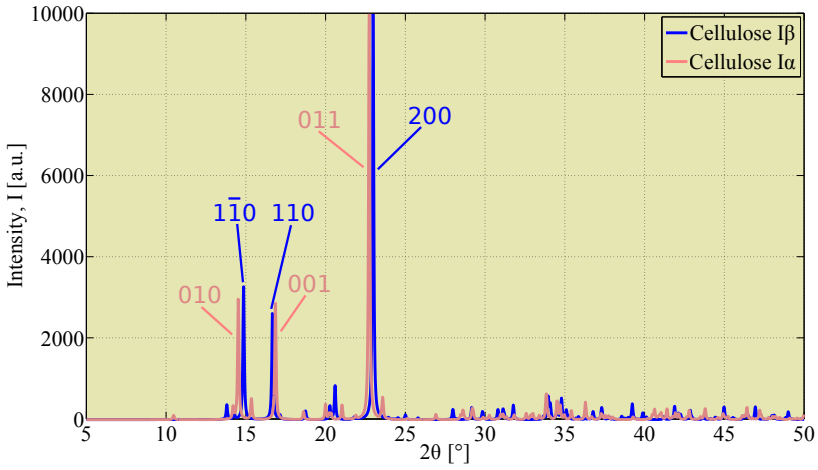


Figure 4. Cellulose Iα and Iβ reflections based on a theoretical model for a very large crystallite size. Strongest reflections are marked on the figure using the *hkl*-notation. Peaks are calculated with Mercury from the data provided in [20, 21].

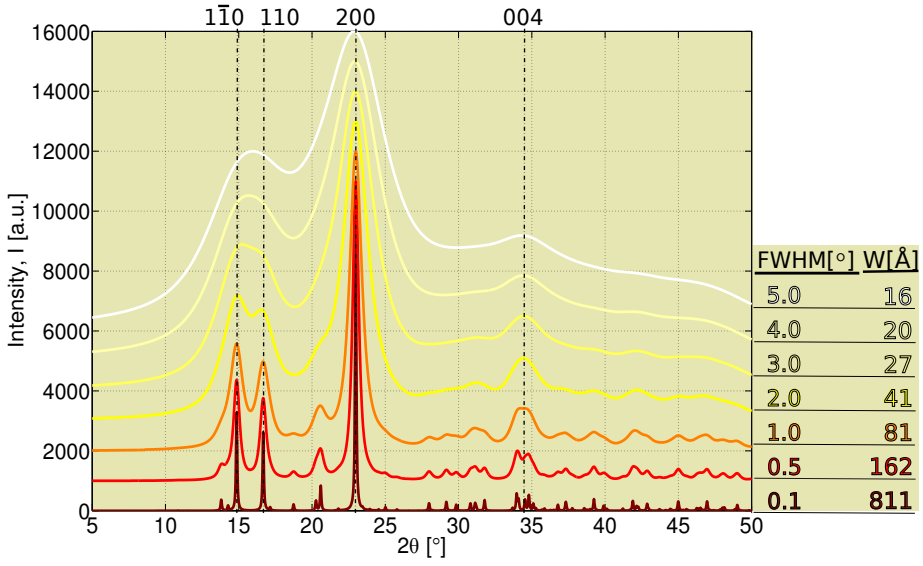


Figure 5. The broadening of the cellulose Iβ diffraction peaks as a function of the full width at half maximum (FWHM) and the corresponding crystallite width (W) as calculated with Eq. (6) using the 200-reflection peak position ($2\theta = 22.99°$). Diffraction intensities calculated with Mercury from the data provided in [20].

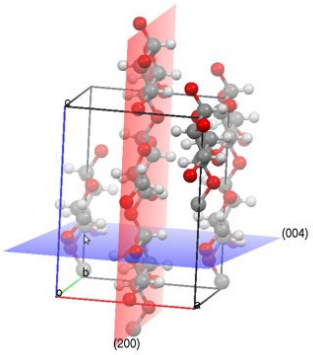


Figure 6. For cellulose Iβ the most often used reflections are the reflections from the 200 plane (perpendicular to unit cell vector $\vec{a}$, red) and the reflections from the 004 plane (perpendicular to unit cell vector $\vec{c}$, blue). Figure drawn with Mercury from the data provided in [20].

Cell wall structure. The wood cell (xylem) wall layers are from outside to inside: Middle Lamella (ML), primary wall, outer layer of the secondary wall (S_1), middle layer of the secondary wall (S_2), inner layer of the secondary wall (S_3) and warty layer [3, p. 13]. The outermost layers, ML and primary wall, are difficult to separate and therefore are usually considered as Compound Middle Lamella (CML) [5].

The outer S_1 layer and the inner S_3 layer have transversely oriented microfibrils, whereas the middle, thicker secondary (S_2) has axially oriented microfibrils. This forms a S-Z-S helical organization to the secondary wall [9]. The 3 nm diameter microfibrils are organized in $15-60$ nm diameter bundles called macrofibrils [26]. In coniferous wood the thickness of the S_1 and the S_3 layer is of the order of $0.1\,\mu$m whereas the S_2 layer thickness is $1-5\,\mu$m [5]. The S_2 layer contains up to 90% of the wood fiber mass [8].

The characteristics, such as thickness and the MFA of the S_2 layer have a decisive influence on the fiber stiffness as well as on other paper-making properties [3, p. 15] [27]. In particular, the longitudinal stiffness (longitudinal modulus of elasticity) of wood depends on the MFA of the secondary wall S_2 [9].

Composition. The main components of xylem are cellulose, hemicelluloses, and lignin. The cellulose can be considered as a skeleton surrounded by a hemicellulose-matrix and a lignin encrusting. [3, p. 12] The cellulose is formed both in crystalline and amorphous regions whereas the matrix of hemicelluloses and lignin is amorphous [4]. The cellulose in the primary cell wall has mainly a low crystallinity and no defined orientation [6] and therefore has negligible effect on average MFA.

Figure 7. A photograph of a sample from the wild type *Arabidopsis thaliana* main stem. Sample was dried in room temperature. Sample length is approximately 19.0 mm.

3.3 Model plant: Arabidopsis thaliana

3.3.1 General description.

The flowering plant *Arabidopsis* (Figure 7) — commonly known as wall cress or mouse-ear cress — is an important model system for plant biology. Its main advantages for genome analysis are a short life cycle [1], small size, large number of offspring and a (relatively) small nuclear genome [28, 29]. Most of the five-chromosome 125-megabase genome was sequenced in year 2000 and when published, the 25,498 gene genome was the largest sequenced gene set [29]. Thousands of *Arabidopsis* lines have been grown with mutations affecting almost all aspects of plant growth [28].

Arabidopsis is distributed widely in nature throughout Europe, Asia and North America. The mostly studied Wild Types (WTs) are Columbia and Landsberg, which are the accepted ecotype standards for genetic and molecular studies. [28] The Columbia ecotype originates from the USA and the Landsberg from Germany [30].

3.3.2 Arabidopsis structure and composition

Cell wall structure. The *Arabidopsis* hypocotyl cell wall has been studied using field-emission Scanning Electron Microscopy (SEM). The thickest cell wall of the freeze-fractured hypocotyls at final hypocotyl length was the outer epidermal wall ($\sim 1\,\mu$m) [31]. The study also suggested that the physical lower limit for plant cell wall could be from 50 to 60 nm [31].

Under appropriate growth conditions (short day) the *Arabidopsis* hypocotyl has been shown to undergo secondary growth and thus the plant can be considered as a model system for secondary xylem development. The main advantage of *Arabidopsis* over trees is that it has much shorter generation time and faster growth. [32] Secondary cell wall thickness of wild type plants have been cited as $1\,\mu$m with significantly lower values found on mutant lines [26].

[1] The life cycle is approximately six weeks from germination to mature seed [28].

Composition. The cellulose content (mass percentage of dry weight) of Columbia WT *Arabidopsis* has been measured to be $(26.8 \pm 0.6)\%$ [33] and around 30% for Landsberg erecta WT [34,35]. The two cellulose I phases, Iα and Iβ, have been suggested to be found in similar proportions in Columbia WT leaves [36]. The lignin content of Columbia WT was measured to be $(22.1 \pm 0.7)\%$ [33].

3.3.3 Nanoscale structure

Crystallite size. The crystallite size of Columbia WT *Arabidopsis* leaves has been estimated to be around 30 Å in cross-sectional direction based on interior and surface distribution of cellulose chains and data from solid state ^{13}C Nuclear Magnetic Resonance (NMR) analysis. [36]

Crystallinity. There are significant differences in the cell wall crystallinity of WT *Arabidopsis* plants depending on the growth temperature [m] but the degree of crystallinity have been found to be around 20% to 30% using an X-ray diffractometer [37].

Microfibril angle. The cellulose MFA of WT *Arabidopsis* root and pith walls have been studied using field emission SEM imaging [38,39] and the MFA of the root walls by polarized-light microscopy [39]. The WT *Arabidopsis* root and pith wall microfibrils were oriented transversely with the deviation of the MFA being less than 30° from the transverse direction [38].

X-ray studies. *Arabidopsis* plants have been studied in the past using X-ray scattering techniques to study the cellulose microfibril angles [33, 37], the degree of crystallinity in the plant wall [37] and the mean crystallite width [37, 40].

[m] Namely $(26.6 \pm 2.5)\%$ for plants grown at 21 ℃ and $(17.8 \pm 1.8)\%$ for plants grown at 29 ℃. [37].

4 Experiments

4.1 Sample preparation

Arabidopsis thaliana preparation. The *Arabidopsis* samples were cut from the main stem of the plant. The samples were then frozen and stored in a freezer. The samples were delivered by Prashant Pawar from professor Ewa J. Mellerowicz's group at the Department of Forest Genetics and Plant Physiology, Swedish University of Agricultural Sciences (SLU) where they were grown in a green house under controlled conditions.

The moist samples were taken out to dry in room temperature at least 12 hours before measurements after being cut into approximately 10 to 15 mm long pieces (Figure 7). The variation of the sample moisture during the two-hour measurement was assumed to insignificant and all samples were assumed to have reached an equilibrium condition based on the drying experiment (Figure 9). Small variations in the sample moisture due to variations in sample size, shape and ambient conditions were taken into account in the crystallinity fit by adding a water scattering background to the fitting model.

Drying experiment. The drying of the stem samples was studied by measuring the mass of the sample as a function of time with a milligram scale. The sample was in room temperature — similar to measurement conditions. During the first hour of drying the mass loss was linear and approximately 25% of the initial mass was lost and a further 60% during the next four hours after which there was no significant mass loss (Figure 9). Ambient temperature was (300 ± 1) K and air humidity (40 ± 5)%. The measured sample (Figure 7) length was 19.0 mm and its starting (moist) mass was 28.8 mg. The sample was kept frozen at 255 K before measurements.

The measurement was done using a 100 μg resolution A&D ER-180A (A&D Engineering, Inc., California, USA) electronic analytical balance, Figure 8. Photographs of the scale reading were taken at 120 second intervals [n]. The photographs were cropped in batch mode and the readings were read from the images with an optical character recognition software package written for Matlab [o]. The time stamps of the images were read from the EXIF data with Matlab function `imfinfo`. A Canon EOS 550D digital single-lens reflex camera was used along with a Hähnel Giga T Pro II wireless timer remote in interval shooting mode to

[n] Photographs were taken at sets of 60 and with 120 second intervals and the sets themselves were repeated with a 60 second interval.

[o] Original code `Optical Character Recognition (OCR)` was retrieved from Matlab File Exchange (`Mathworks.com`) and was written by Diego Orlando. The code separates the image into components using Matlab `bwlabeln` and compares each component against a template and chooses the character whose template has the highest correlation value with that component. The templates were redone especially for this experiment and the original code was modified in other ways as well for the drying experiment.

Figure 8. The A&D ER-180A (A&D Engineering, Inc., California, USA) electronic analytical balance was used to measure the mass loss of the *Arabidopsis* sample.

record the images.

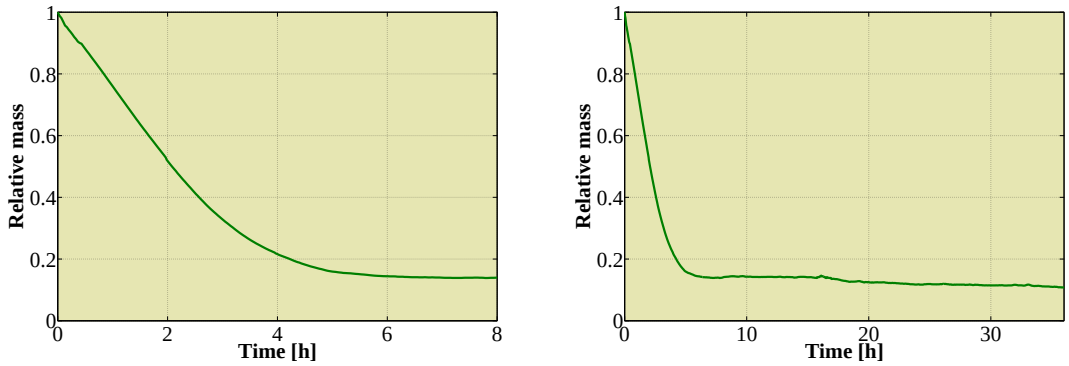
4.2 Experimental set-up

4.2.1 X-ray Microtomography

X-ray microtomography measurements were conducted at the Laboratory of Microtomography, University of Helsinki using the custom-made Nanotom 180NF device supplied by Phoenix|Xray Systems + Services GmbH (Wunstorf, Germany) [41–44]. Cone beam geometry was used with the high-power nanofocus transmission type X-ray tube with an end-window type tungsten anode. A total of 1440 X-ray transmission images (or projections) were taken at 0.25° intervals covering one entire circle of revolution around the *Arabidopsis* stem. The sample was attached with beeswax vertically on top of the high-precision computer-controlled rotation/translation stage. In order to reduce random noise each projection was averaged over 10 individual images, each of 1.0 s exposure time. The X-ray intensity was measured using a metal-oxide-semiconductor (CMOS) flat-panel detector of 2304 × 2284 pixels (50 μm pixel size; Hamamatsu Photonics, Japan). X-ray tube voltage was set to 60 kV and the current to 200 μA for the WT sample and 50 kV/250 μA for the genetically modified sample [p]. Total measurement time was approximately two hours for each sample.

The 3D reconstruction was done at full resolution with the *Datos|x rec* software supplied by device manufacturer Phoenix|Xray Systems + Services GmbH. Visualization of the

[p] Chosen samples were WT replicate #9 from construct #1 and replicate #3 of line 1a2 from construct #2.



(a) (b)

Figure 9. The *Arabidopsis* sample lost over 80% of its initial mass as water within the first five hours of drying (a). The curves were smoothed (averaged) over 11 data points with Matlab `smooth` and data points were taken at 2 min intervals. The depicted 36-hour interval (b) consists of 1089 data points.

reconstructed volumes was done with the *VGStudio MAX 1.2.1* (Volume Graphics GmbH, Germany) software. Random noise was removed from the 3D reconstructions using a 3 x 3 x 3 voxel kernel Gaussian convolution filter.

4.2.2 Wide Angle X-Ray Scattering

The measurement set-up used in the WAXS measurements (Figure 10) has been used previously in the University of Helsinki to study among other things wood and high-density polyethene [45]. A copper anode X-ray generator and a monochromator were used to produce Cu Kα radiation (wavelength $\lambda = 1.541$ Å). The size of the beam was adjusted with slits to be approximately $0.5 \times 0.5\,\mathrm{mm}^2$. The mar345 image plate detector (Marresearch GmbH, Norderstedt, Germany) was used with the maximum scanning diameter (345 mm) and 150 μm pixel size. A copper beam stop was used to cut away most of the primary beam intensity.

The wide angle X-ray scattering experiments were performed using perpendicular transmission geometry (Figure 2). The measurement time for each sample was determined to be two hours. An air background of two hours was measured as well as an equally long dark background during which there was no X-ray beam. The sample to detector distance was 95 mm and the 2θ scale was calibrated using a silver behenate (AgBeh) and a lanthanum hexaboride (LaB_6) calibration sample.

After each measurement, the diffraction pattern (Figure 11) was read from the area detector and imported to Matlab for data analysis.

The data used in this study combines measurements from three measurement set-ups.

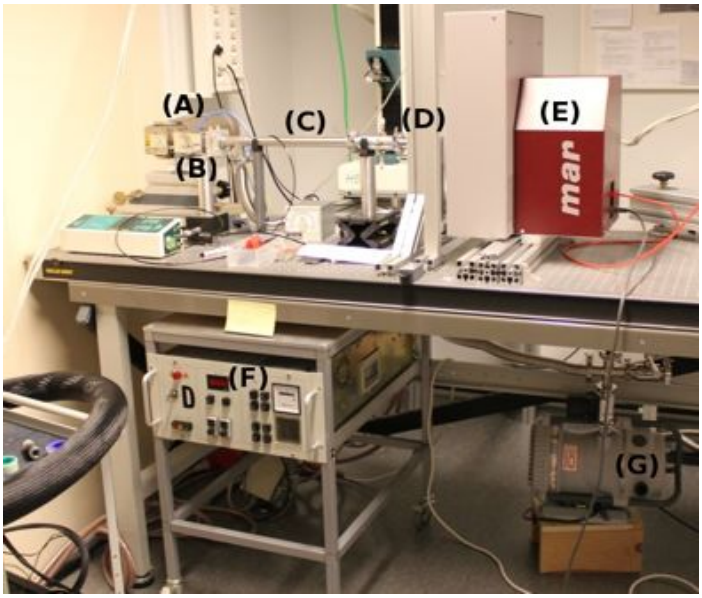


Figure 10. The measurement set-up used in the WAXS measurements. **(A)** PANalytical (PANalytical B.V., The Netherlands) PW2213/20 copper anode XRD glass point focus X-ray tube, **(B)** collimating Montel multilayer monochromator (Incoatec), **(C)** Vacuum tube, **(D)** sample holder (not visible), **(E)** mar345 (Marresearch GmbH, Norderstedt, Germany) image plate detector, **(F)** Siemens Kristalloflex K710H high-voltage generator (used voltage was 36.0 kV and current 25.0 mA) and **(G)** Edwards (Edwards Limited, United Kingdom) XDS-10 oil-free dry scroll vacuum pump.

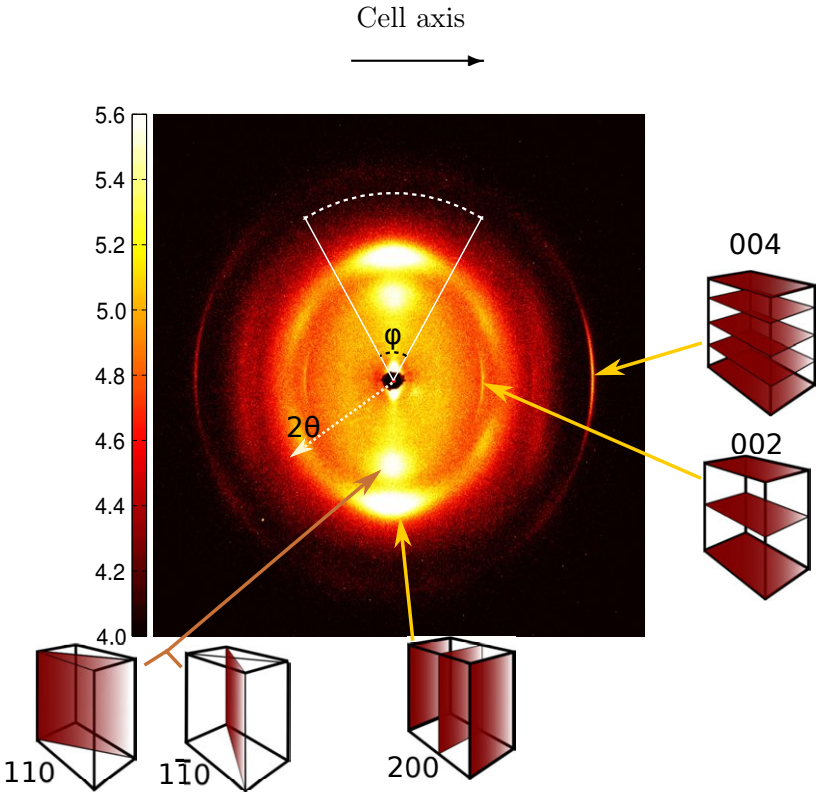


Figure 11. A diffraction pattern from an aspen sample. The strongest diffraction peak is caused by the reflection from the 200-plane (200), the combined reflection peak of the reflections from the 110 and 1$\bar{1}$0 planes (110, 1$\bar{1}$0) is the second strongest and the 004-reflection (004) is also clearly visible as well as the 002-reflection (002). The azimuthal angles φ were defined in the way that the depicted arc is from 60° to 120°, increasing counter clock-wise. The 2θ angle grows radially from the center of the image where $2\theta = 0°$. Colors indicate the logarithm of an arbitrary intensity scale with white spots showing the highest intensity.

They are all nearly identical, but after any equipment is moved or any beam parameter adjusted, the next measurement is considered to be of a new set-up. In total there were four set-ups in this project but as the X-ray intensity was significantly lower in one set-up those results were not used. The air scattering correction and 2θ scale calibration and related measurements are conducted always anew for each set-up.

4.3 Data analysis

The diffraction pattern analysis and the consequent data analysis were performed in the MATLAB scientific computing software (The MathWorks Inc., Natick, MA, USA). The center of the image was masked from the analysis in the region of $2\theta < 4°$ around the beam stop shadow. The 2D detector pixel coordinates were transformed to polar coordinates (2θ / q, ϕ)

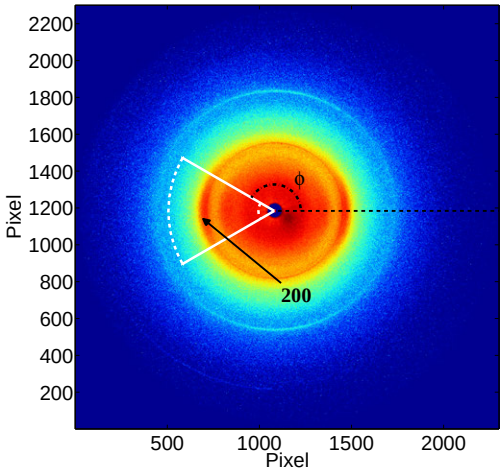


Figure 12. The sector used to calculate the 200-reflection intensity. The azimuthal angle (ϕ) of the chosen sector (marked in white) extends from 150° to 210° with the black dotted line indicating $\phi = 0°$ and it increases counter-clockwise. The 60° sector was chosen to cover the entire 200 reflection peak (200) for each sample. The 2θ angle grows radially from the center of the image where $2\theta = 0°$ (see Figure 11). The diffraction profile is from the *Arabidopsis* sample ΔFC2 6c5 (1). The reflection intensities are notably lower than those from the aspen sample, Figure 11, even though the measurement time was twice as long.

where ϕ is the azimuthal angle (Figure 12) and the radial component was twice the Bragg angle (2θ-angle) or the wavelength invariant q-scale.

A two-hour air background was measured alongside a two-hour dark current (i.e. no X-ray beam, shutter closed). Air scattering is obtained by subtracting the dark current from the air background. This is then subtracted from the measured scattering pattern, Figure 13. It is important that air background during the measurement is close to the subtracted air background since the scattering from the *Arabidopsis* sample is so weak that the air scattering dominates at small 2θ angles ($22\theta < 20°$). Absorption correction is done after the air background subtraction and it is followed by a geometrical correction.

The data was combined from three set-ups, whose analysis was done separately. Afterwards these data were combined by first saving the data from each set-up to a Matlab variable and then by reading in all these values and calculating averages. This can be a time-consuming process and changing the value of one variable can change many averages and a single modification can lead to many time-consuming corrections. For this reason a Matlab script was used to a) combine the data and calculate the averages, b) write the results in LaTeX format and to c) create figures from the averages. If this is done it is easy to update the results and the barrier for correcting or fine-tuning your results is lowered.

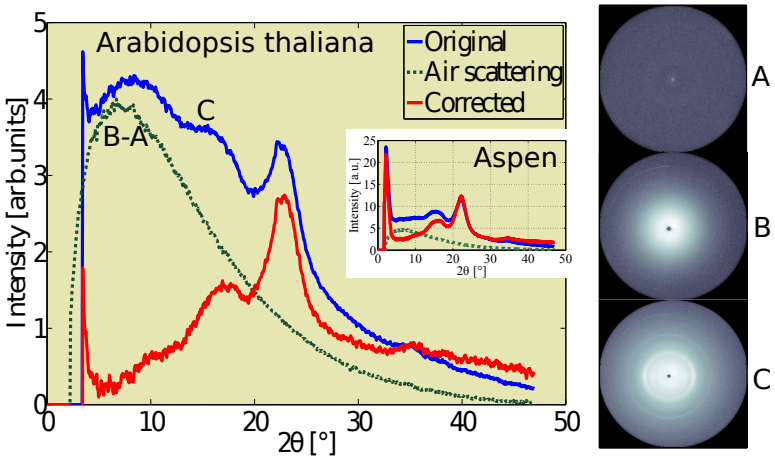


Figure 13. Air scattering correction for the 200-reflection section show in Figure 11. An air background (dark green dotted line) of the identical measurement time as for the sample measurements was measured and subtracted from the original data (solid blue line on top) and the resulting corrected intensity (solid red line) was used in the analysis. On the right are typical dark current (A), air background (B) and measurement (C) scattering patterns for *Arabidopsis*. The intensity curves are obtained from the diffraction images by integrating over an azimuthal angle (such as that shown in Figure 12) and are shown here as a function of the 2θ angle. The letters indicate that the blue line is obtained by integrating the diffraction pattern C and the air scattering by integrating the difference of air background (B) and dark current (A). The small inset shows the difference to a typical wood sample (aspen) where the scattering is stronger and air background is less significant.

4.3.1 Crystallite size

The Crystallite size was calculated using an implementation of the Scherrer equation, Eq. (6). The half-value breadths and peak positions were calculated by fitting two straight lines in the 200-diffraction peak intensity figure (Figure 14). The intensity was integrated over an azimuthal angle of 60° around the 200-reflection (Figure 12) after the air diffraction background was subtracted. The amorphous sulphate lignin background was subtracted before crystallite size analysis.

A Gaussian fit was done for the 200-reflection to find the peak value. This fitting was performed to a pre-determined 2θ range around the 200 peak. The region was selected to be 20..24.5° for all samples. The top of the 200 peak may not be very Gaussian in shape; in some cases it was quite flat. For this reason the established method from wood analysis, in which a small 2θ region was selected from the top, did not produce reliable and repeatable results.

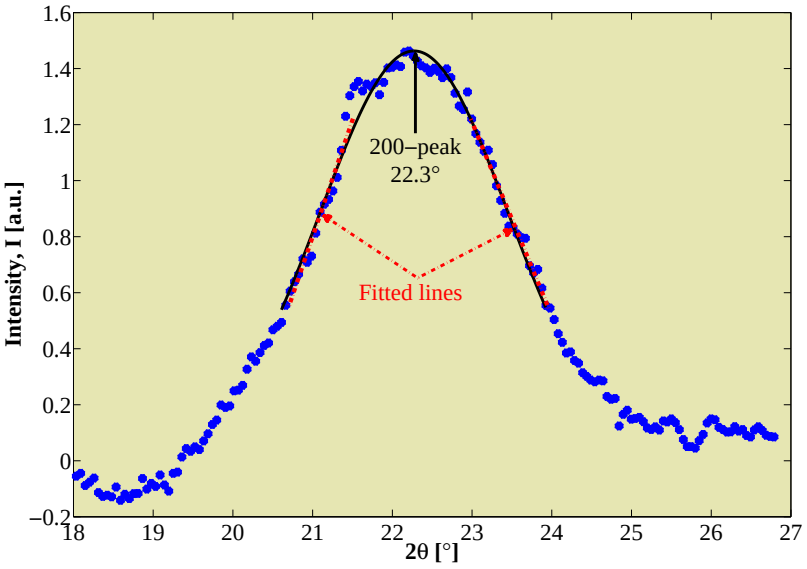


Figure 14. The crystallite size is determined by fitting two lines (red) to the 200-reflection peak in the intensity figure. The half-value breadth is calculated as the distance of the two fitted lines at half of the maximum intensity of the peak.

4.3.2 Crystallinity

Amorphous background and oriented crystalline cellulose reflections were fitted to the measured sample intensity. Sulfate lignin was chosen as the amorphous background model. Known cellulose reflection peaks were used to fit 22 Gaussian reflection peaks. The cellulose reflection peak positions were calculated for cellulose Iβ using the freeware program PowderCell from the data provided in [20]. The peak widths were adjusted slightly for a better fit to the non-perfect crystalline material of the sample. The same peak parameters were used for all samples. An exemplary fit for the crystallinity is shown in Figure 15.

The air scattering was approximated to be equal in magnitude to the air background measured separately (for each set-up), see section 4.3, p. 22 [q]. As there is no accurate *a priori* information about the exact degree of moisture in each sample (it is known to be very low compared to living plants), a water background was fitted to the measured intensity curve and subtracted from it before the crystallinity analysis (as per Eqs. (7,8)). The crystallinity was then calculated using Eq. (7) and the 2θ angle was limited to $13° < 2\theta < 40°$.

Originally the fitting was done using only one background (amorphous) and the crystalline

[q] The air scattering for two samples was different since the image plate scanner had been opened for inspection. Air scattering correction was fitted for these two samples to compensate for the change in the measured air scattering intensity similarly to the water background (see Figure 29 in Appendix C on p. 51).

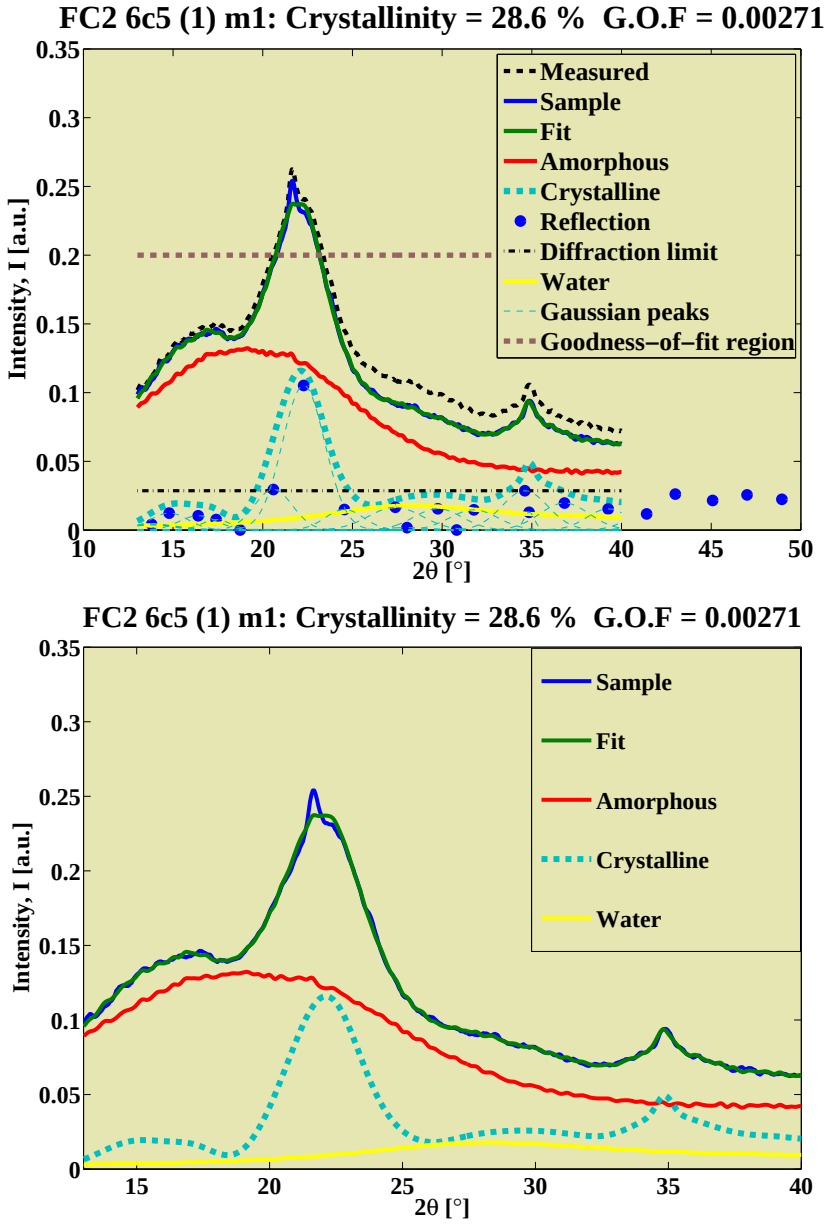


Figure 15. The crystallinity of the sample was calculated by fitting (solid green line) an amorphous background (red solid line) and a crystalline reflection curve (teal dotted line) modified from known perfect crystal reflection values. The positions of the used crystalline reflections are shown as blue circles on the upper figure along with the Gaussian fits used for crystallinity (thin teal dotted lines, on the upper figure). Water background (solid yellow line) is subtracted from the measured intensity (black dotted line on the upper figure) to produce sample intensity (solid blue line). The sample represents a typical value obtained for crystallinity.

reflections. Additionally a non-negative least square fit was used to obtain the coefficients for the fitted parameters. This fit was done using the intrinsic Matlab `lsqnonneg` function that does not allow one to set upper boundaries or give initial guesses for the coefficients.

A more reliable fit is obtained by giving reasonable starting values for the coefficients and upper boundaries for the parameters. Namely the cellulose reflection peaks should follow the model to some extent; if they did not, the fitting was not reliable in any case. Upper boundaries were set so that the 200-reflection peak can have stronger intensity than any other peak and peaks for reflections $1\bar{1}0$, 110, 102 and 004 can have stronger intensities than any other peak excluding 200. Lower boundaries were set to zero (i.e. non-negative fit) although negative values could make sense if negative air scattering correction [r] would be done.

The starting values for the new fit were taken from the old fit (now called initial fit) with just the strongest reflections (200, $1\bar{1}0$, 110, 102 and 004; other peaks labeled as weak reflections) and a single amorphous background. The starting values for the coefficients of the strong reflections were taken as the coefficients from the initial fit, same as for the amorphous background. Other backgrounds were assumed to be weaker than the amorphous background and their starting value was half the coefficient for the amorphous background. The starting value for the weak reflections was the minimum value of the strong reflections. Fitting was done using a linear least squares fit (Matlab `lsqlin`). The degree of crystallinity as given by Eq. (7) was calculated using a numerical integration method. The chosen method was the trapezoidal formula, Eq. (9).

4.3.3 Microfibril angle

For the microfibril angle azimuthal integral was taken over a half-circle of the image, 180 degrees in azimuthal angles. One value was integrated for each azimuthal angle in the 180-degree region (180 integrals). The 2θ region over which each value was calculated was taken to be approximately 21° to 26° (Figure 16). The asymmetry was chosen so that the 102 peak would not overlap significantly over the 200 peak.

[r] In air scattering the X-ray photons scatter from air before and after the sample. This scattering pattern is always measured since it needs to be subtracted from the measurement scattering pattern to obtain scattering from the sample. The air scattering background measurement was done separately and it was assumed that the scattering from air was identical in the background measurement and in the sample measurement. This assumption is usually a reasonable approximation even though the air background may change over time. Here the air scattering correction means that the air background would have changed from the time that the air background was measured to the time the sample measurement was conducted. The correction would be a fit that would correct for this change in the air background. A positive air scattering correction is in fact in done in Appendix C on p. 51.

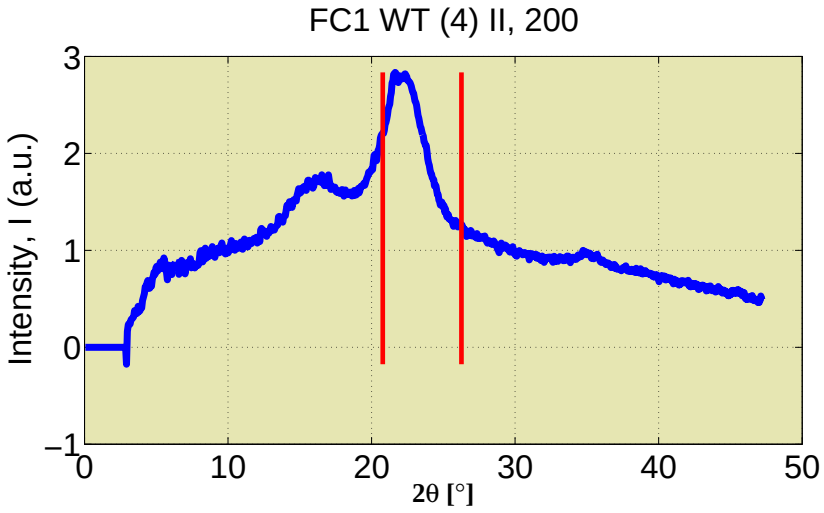


Figure 16. The red lines in the intensity profile indicate the 2θ region used in integrating in the azimuthal direction (each azimuthal integral is taken over the indicated 2θ region.)

4.4 Results

4.4.1 X-ray Microtomography

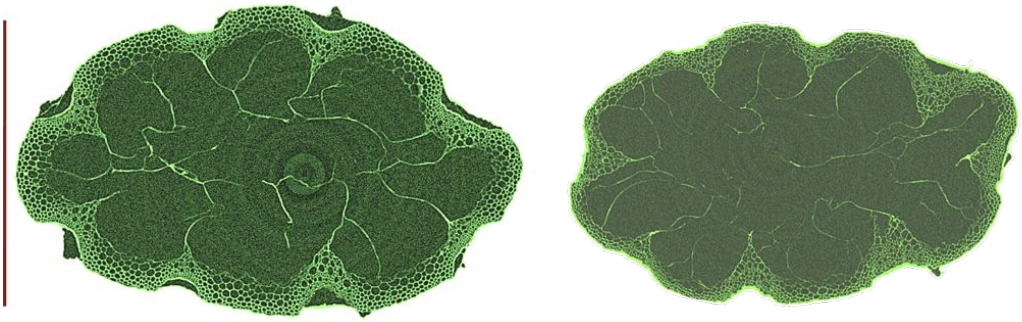
Tomography measurements were conducted on the WAXS samples [s] to study the overall structure of the stem. The samples were dried and handled as discussed in the sample preparation subsection (Section 4.1 on p. 17). A cross-section of the stem is shown in the slice reconstructions of a WT plant and genetically modified plant (FC2-1a2 (3)) in Figures 17a and 17b. Reconstructions from a part of the stem of the genetically modified plant and the WT are shown in Figures 18 and 19. Note that the coloring is for visual effect only and that the stems were squashed slightly when they were positioned on the sample holder during the WAXS measurements. Also for the 3D reconstructions all voxels with low intensity were removed for visualization purposes.

We can see from Figures 17a and 17b that the cells are generally more round than rectangular and thus it is not valid to use analysis methods that assume rectangular cells. The microfibril angle determination can be done by studying reflections from different cell walls [7] — a method that requires rectangular cells in order for it to be applied.

4.4.2 Wide Angle X-Ray Scattering

All the WT *Arabidopsis* samples from construct #2 (FC2) have very similar 200-reflection intensities, Figure 20. A good fit (Goodness Of Fit, $GOF < 0.01$) was obtained by fitting an

[s] Samples were first measured with WAXS.



(a) (b)

Figure 17. A slice taken from a tomographic reconstruction volume of a section of the stem from a a) WT plant and b) genetically modified plant (FC2-1a2 (3)). Larger images in Appendix A, p. 46. Scale bar is 1.0 mm.

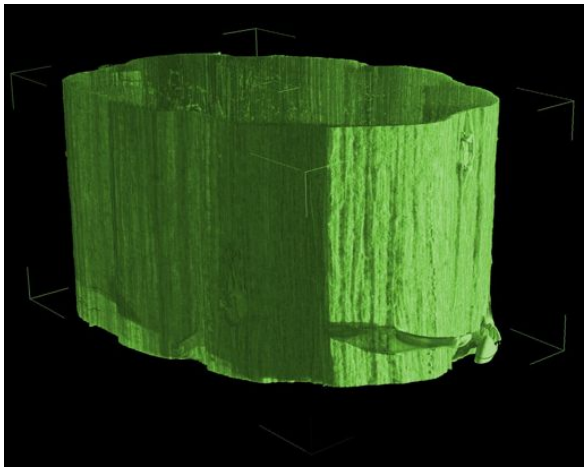


Figure 18. A tomographic reconstruction volume of a section of the stem from a genetically modified plant (FC2-1a2 (3)). The bounding box of the volume selection is marked on the image. Artificial lighting effects are introduced to the volume.

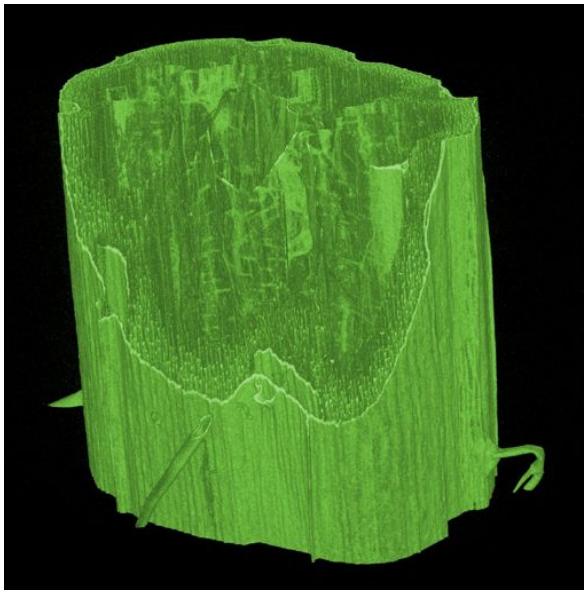


Figure 19. A tomographic reconstruction volume of a section of the stem from a WT plant; the plane is cut from the volume to reveal 3D structure. Artificial lighting effects are introduced to the volume.

amorphous lignin background and cellulose Iβ reflection peaks to the intensity curve of the 200-reflection (see Figure 15).

When the diffraction pattern of the *Arabidopsis* sample was compared to a wood sample (aspen) it was clearly seen that the diffraction was much weaker (Figure 21). With weaker diffraction from the sample other factors such as the air scattering reduction and amorphous background fitting are more important.

The individual results for the crystallinity C, crystallite width W and mean microfibril angle MFA for each sample are listed in Appendix B, Table 1 (FC1, WT), Table 2 (FC1, genetically modified), Table 3 (FC2, WT) and Table 4 (FC2, genetically modified).

The average crystallite width W for each line was consistently between 27 Å and 28 Å as can be seen in Figure 22. While there is very strong evidence ($p < 0.01$) that the means of lines 1a2 and 4a9 of construct #2 differ from the respective WT mean, the absolute difference in the means is smaller than the estimated resolution of the analysis method. Due to the small sample size the measurements do not conclusively show a decrease in the crystallite width for the two genetically modified lines. However, they do suggest that further studies with larger sample sizes should be conducted if possible.

The cellulose crystallite width is consistent with results determined from a NMR study [36] [t].

[t] The study cited cross-sectional dimensions of 30 Å for Columbia WT *Arabidopsis* leaves.

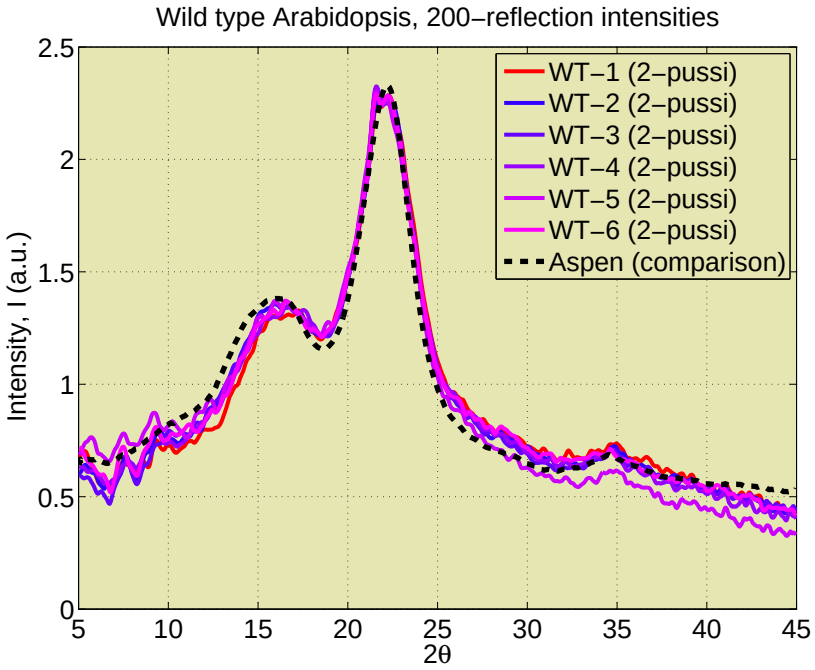


Figure 20. Comparison wild type *Arabidopsis* sample 200-reflection intensities from six samples (solid lines). An aspen intensity profile is shown here for reference (dotted line). The intensities have been fitted for zero-level and scaling in the region of $15° < 2\theta < 25°$. The intensity profiles shown here are smoothed over 13 data points using Matlab command `smooth`.

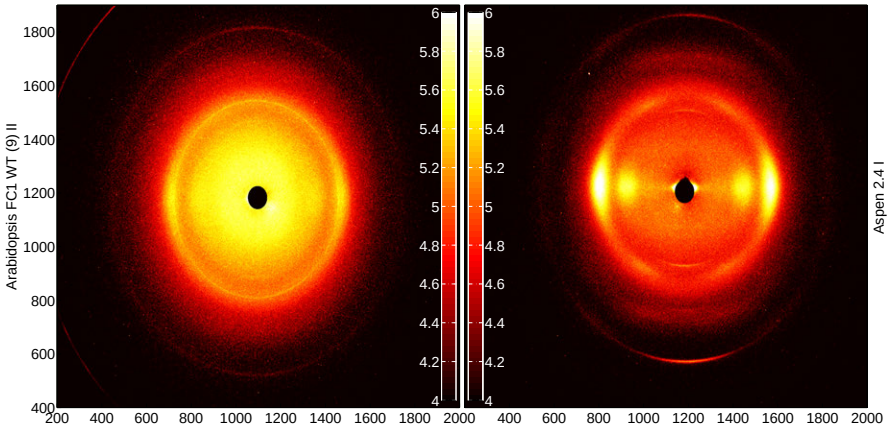


Figure 21. Comparison of the WT *Arabidopsis* (left) diffraction pattern with the diffraction pattern from an aspen sample (right). The peak-intensity is much higher in the aspen measurement even though the measurement time was only one hour, half of the measurement time of the *Arabidopsis* sample.

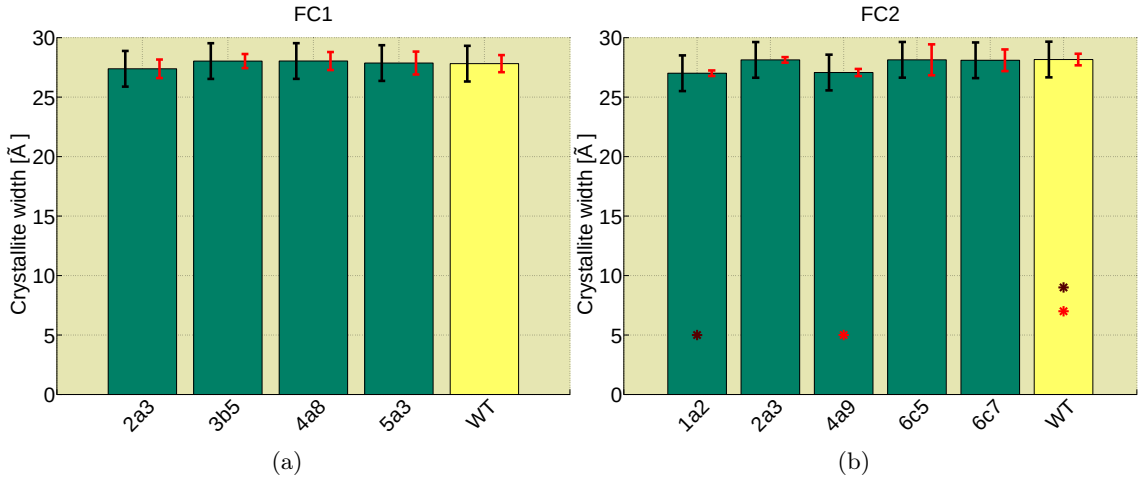


Figure 22. (a) Crystallite width comparison for construct #1 (FC1), line means. Sample count $n = 9$ except for 2a3 and 5a3 for which $n = 7$. (b) Crystallite width comparison for construct #2 (FC2), line means. Sample count $n = 3$ except for WT for which $n = 6$. WT = Wild type. Black error bar represents methodology inaccuracy (± 1.5 Å) and red error bars the standard deviation of each line. Asterisks indicate statically significant differences in means between two lines with $p = 0.05$ using a one-sided two-sample t-test (Matlab function `ttest2`, see section 2.2).

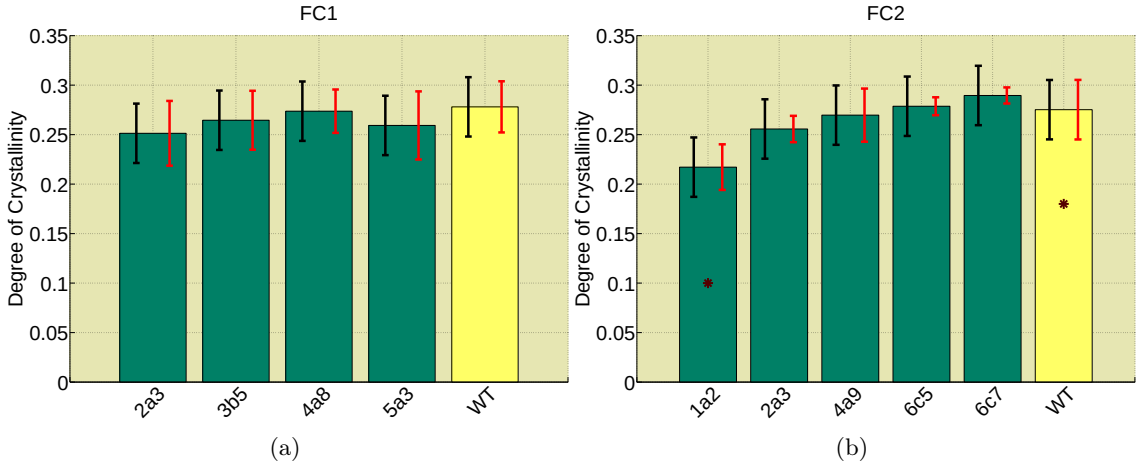
Similar values [u] have also been quoted for Norway spruce using similar X-ray measurement techniques than used in this study [6, 46]. A crystallite width of 35 Å was reported for the *acw1* mutation of *Arabidopsis* in a reflection mode X-ray study [40]. This value is slightly higher than any individual result obtained for the crystallite width.

The degree of crystallinity obtained was similar to that reported [v] for the cell wall of *Arabidopsis* [37]. The t-test suggests strong evidence ($p < 0.05$) against the hypothesis of equal degree of crystallinity means for the 1a2 line and the WT for construct #2 (Figure 23). However, the t-test is susceptible to outliers in the data, especially with small sample sizes. For this reason the measurements can be considered to be only indicative of a difference in the means and not conclusive. Again, further measurements with larger sample sizes should be carried out.

Large variation was seen in all lines in the mean MFA value (Figure 24). For construct #2 the mean values of the genetically modified lines did not have statically significant differences with respect to the WT line (the null hypothesis of the other mean being larger could not be rejected with $p = 0.05$). However with construct #1 there was strong evidence that line 3b5 has a larger mean than the WT (one-sided two-sample t-test with p=0.05).

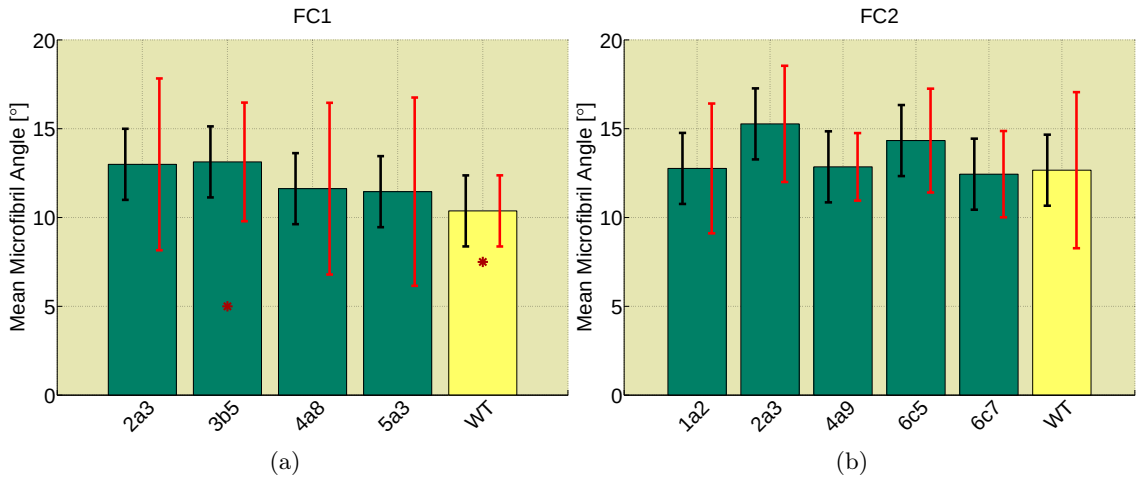
[u] Mean width varied between 28.9 Å and 30.9 Å as a function of the year ring of the samples [46].

[v] Assuming growth temperature of approximately 21 °C. Degree of crystallinity was reported as (23.6±2.5)% for WT *Arabidopsis* plants grown at 21 °C (and (17.8 ± 1.8)% for those grown at 29 °C). [37]



(a) (b)

Figure 23. (a) Degree of crystallinity comparison for construct FC1, line means. Sample count $n = 9$ except for 2a3 and 5a3 for which $n = 7$. (b) Degree of crystallinity comparison for construct ΔFC2, line means. Sample count $n = 3$ except for WT for which $n = 6$. WT = Wild type. Black error bar represents methodology inaccuracy (± 0.03) and red error bars the standard deviation of each line. Asterisks indicate statically significant differences in means between two lines with $p = 0.05$ using a one-sided two-sample t-test (Matlab function `ttest2`, see section 2.2).



(a) (b)

Figure 24. (a) Mean Microfibril Angle, $\langle$`MFA`$\rangle$ comparison for construct FC1, line means. Sample count $n = 9$ except for 2a3 and 5a3 for which $n = 7$. (b) Mean Microfibril Angle, $\langle$`MFA`$\rangle$ comparison for construct ΔFC2, line means. Sample count $n = 3$ except for WT for which $n = 6$. WT = Wild type. Black error bar represents methodology inaccuracy ($\pm 2°$) and red error bars the standard deviation of each line. Asterisks indicate statically significant differences in means between two lines with $p = 0.05$ using a one-sided two-sample t-test (Matlab function `ttest2`, see section 2.2).

5 Discussion

5.1 Air scattering

The scattering from air was noted to dominate at small 2θ-angles ($2\theta < 20°$). Therefore the air scattering must be carefully taken into account when doing measurements of such weakly scattering samples as the *Arabidopsis* main stem (see Figure 21 on p. 30). The air scattering reduction mostly influenced the estimates of crystallinity but did not have a big impact on the crystallite width. During this study the WAXS set-up was rebuilt three times and it was noted that using the air scattering calculated from the other set-up resulted in an unreliable estimation for the degree of crystallinity in the sample.

5.2 Crystallite width approximation

The shape of the 200-reflection peak was assumed to be Gaussian in the shape factor of Eq. (6). This introduces some inaccuracy to the crystallite width calculations since the measured peak was quite clearly non-symmetrical for some samples. However it can be considered to be a reasonable approximation with the given precision of the method.

It should also be noted that it is not clear which is the best peak position to use in the Scherrer equation (Eq. (6)); should one use the theoretical value based on lattice constants, literature value for cellulose Iβ 200 peak or the peak position from the measurement. The 200 peak is well isolated (Figure 4) except for the 102-reflection, which affects the shape of the 200 peak so that the peak may not be exactly where the 200-reflection is based on lattice constants. Due to the inaccuracies in the measurement and 2θ scale calibration there is some inaccuracy in the absolute (and relative) values of the 2θ scale. Also, using cellulose Iβ peak value is problematic since there is some crystalline cellulose Iα in *Arabidopsis*. Here it was decided to use the peak value from the measurements and the peak FWHM from the measurements as these are both affected by all the aforementioned considerations.

We should look at how significant the change is in the Scherrer equation (Eq. (6)) with respect to what value we use. From the Nishiyama 2002 cellulose Iβ lattice constants [20] ($a = 7.78\,\text{Å}$) we get for the 200 peak position (from Eq. (1))

$$2\theta = 2\arcsin\left(\frac{\lambda}{2(a/2)}\right) \approx 22.8°. \tag{16}$$

From the fitted Gaussian peaks from each measurement the values for the peak position ranged from 22.0° to 22.4°. Since the crystallite width approximation is proportional to the inverse of the cosine of half the scattering angle ($W \propto 1/cos(2\theta/2)$) it does not matter much which value is used since $\cos\alpha \approx 1$ with small α. If we take 22° and 24° to be the lowest

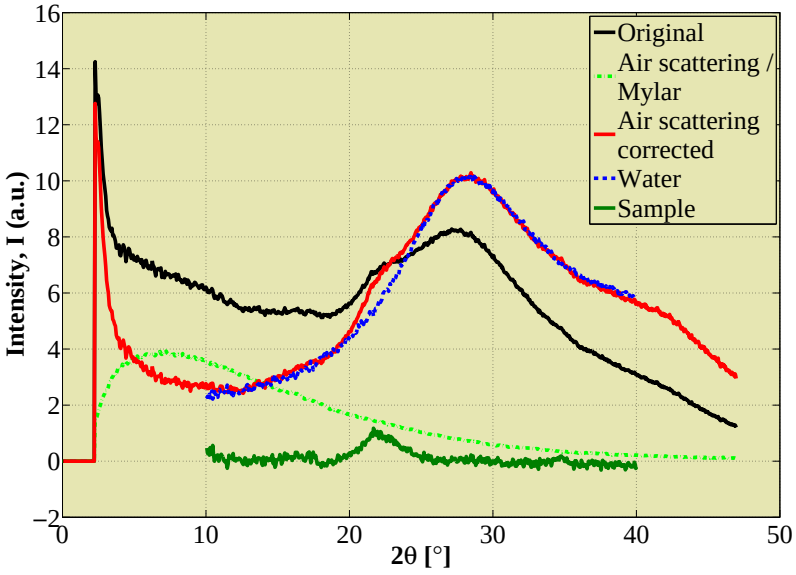


Figure 25. Intensity of the 200-reflection from a moist sample: uncorrected original measurement (black); used air and mylar scattering (light green); after air scattering, absorption and geometrical correction (red); fitted amorphous water background (blue) and scattering from the sample (dark green). As can be seen, the scattering from the sample is very small compared to air and water scattering corrections.

and highest reasonable values for the diffraction angle of the 200-reflection (with Cu $K\alpha$ radiation wave length) the ratio of W_{22° and W_{24° is

$$\frac{W_{22^\circ}}{W_{24^\circ}} \approx 1.0036.$$

This means that the choice of the value for the 200 peak position is not significant with respect to the crystallite width value and its error.

5.3 Moist samples

We considered also measuring samples with natural humidity enclosed in a mylar film to maintain humidity. Our measurement showed that the scattering from water is too strong compared to diffraction from the cellulose. After air scattering, absorption and geometrical corrections (as performed for all samples) the obtained intensity was dominated by water scattering (Figure 25). A water background was fitted to the corrected intensity in the least square sense and the fitted curve was subtracted from the corrected intensity. The 2θ-scale was limited to $10^\circ < 2\theta < 40^\circ$ due to the limited 2θ-range of the water background

data. The resulting intensity curve could then be considered to originate from scattering from the sample. While the curve had peaks around the theoretical cellulose Iβ 200- and 004-reflection peak positions, the statistics of the resulting intensity was deemed too weak for any further analysis. This implies that it is not reasonable to measure living (humid) parts from *Arabidopsis* stem, i.e. *in vivo* measurements — at least not with a measurement set-up similar to the one used in this study.

5.4 Beam positioning

For set-up IV a system of step motors was constructed that allowed sample movement in the plane parallel to the X-ray beam. While the step motors were not used to automate the measurements (which would have been possible) they were used to locate the beam with respect to the sample holder. The metal ring that has a hole in it for the sample was aligned so that the beam went through the center of the hole. This was achieved by the following procedure:

1. moving the sample holder with the step motors until half of the beam intensity was lost in the negative y-direction and marking the position down,
2. moving the sample holder with the step motors to the positive y-direction until half of the beam intensity was again lost and marking this position down,
3. moving the sample holder to the middle of the two previous positions and
4. repeating points 1.–3. in the x-direction.

Note that this would not have been possible to do with the WAXS set-up alone since the detector must be read constantly and the WAXS detector does not give out live intensity values. This was done using a SAXS detector [w]. This procedure is recommended if the beam location with respect to the sample holder must be determined very precisely (sub-millimeter precision, small or thin samples). While the small number of samples measured with this set-up (IV) cannot provide good statistics, the more precise knowledge of the X-ray beam location seemed to improve the quality of the measurements and made positioning the sample much more simple.

5.5 Beam perpendicularity

It was noticed that the X-ray beam was not perpendicular with respect to the detector. Usually it is assumed that the distance from the primary beam along the detector surface

[w] Bruker AXS HI-STAR area detector

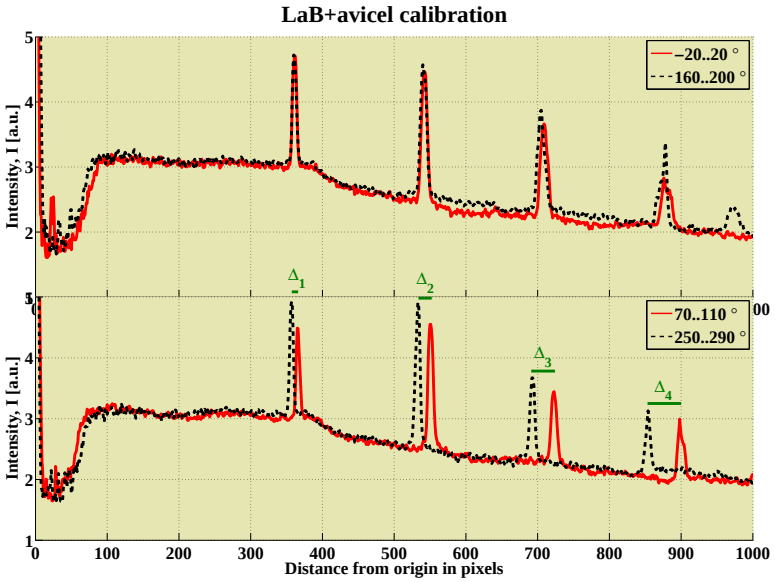


Figure 26. With a lanthanum hexaboride calibration sample we can clearly see four sharp peaks (upper image). Selecting a different 40-degree sector in the vertical direction (upper image) did not change the positions of the peaks. However, in the horizontal direction (lower image) there is a significant difference in the positions of the peaks ($\Delta_i, i = 1..4$).

can be directly mapped to a 2θ scale regardless of the azimuthal angle. The usual assumpion is that the mapping from one azimuthal angle (or sector) is valid for any other azimuthal angle as well. This assumption is true only if the incoming beam is perpendicular to the detector surface. We found that there is a difference in azimuthal mapping (Figure 26) in the horizontal direction of the detector. In the vertical direction a similar difference was not seen. We assume this to mean that the X-ray beam is not perpendicular to the detector in the horizontal direction. This was further confirmed by the fact that turning the detector in the horizontal direction resulted in a circular diffraction pattern using a similar crystalline calibration sample in other measurements with the same device.

Moreover when the calibration is done for a 180- or 360-degree sector, it is not valid in all azimuthal angles if the beam is not perpendicular to the detector. It is valid in the vertical direction in our case, however, and since our 200 peak was located in this direction the calibration was valid. Had our sample been positioned vertically, our 200 peak would have been in the horizontal direction. In that case a single 2θ calibration would not be sufficient and one would need to calculate multiple 2θ scales. With multiple 2θ scales one would use for each point the one closest to it in the azimuthal angle (or some interpolation of the (two) closest). We did not find that necessary in our case since our most important data is in

the vertical direction. If we had measured the crystallite length, we would have used the 004-reflection, which is in the horizontal direction and multiple 2θ scales would have been needed.

5.6 Scripting in data analysis automatization

On a technical note, it was discovered that using small high-level language [x] scripts both reduces the time needed for simple tasks and is less error-prone. One problem is of course that the number of repetitions needed should be known beforehand as there is no reason to write code for tasks that are only carried out once. If the script is not needed again, writing it takes extra time although it can still be used as a means of documentation and it can still serve the function to increase the repeatability of the data analysis work flow. However, if the data analysis always requires input from the user [y] this will eventually consume more time if the number of samples is large or the samples need to be analysed many times (as is usually the case if analysis method improvement is considered). Basic data analysis should not become unnecessarily strenuous as the user should mainly focus on the special circumstances of the data or conclusions from the results. Further on, humans are not good at repetitive tasks as it is difficult to maintain attention over extended periods of time.

5.7 On Arabidopsis

The cellulose in the cell wall of the *Arabidopsis* stem has a similar crystallite width to wood (around 28 Å for Arabidopsis and 30 Å for e.g. Norway Spruce [7]). This similarity can well originate from the similarity in the cellulose biosynthesis as the number of glucan chains in the microfibril depends on the terminal complex structure [19]. A genetic modification on *Arabidopsis* can influence the cellulose biosynthesis and cellulose crystallinity [2]. Since the diffraction pattern is very weak compared with the aspen samples measured with the same set-up, one could try to increase the intensity by creating a powder-like sample that could be studied for the degree of crystallinity. Naturally this sample cannot be studied for MFA since the orientation information is lost. However, it would be possible to cut the stem in smaller pieces and measure through all the pieces while maintaining the cell axis orientation. In practise this could be difficult since a small variation in the angles of the pieces would result in a change in the perceived mean MFA. Further on, this should be done without damaging the cell level structure of the plant.

[x] Namely Matlab and python.

[y] Naturally the user still needs to verify the validity of all steps in the data analysis and some action is always required from the user in special circumstances where the algorithm is not working properly.

6 Conclusions

Our results show that it is possible to study *Arabidopsis* plants at the nanometer scale using a single stem with WAXS. The resulting diffraction pattern is not strong but with careful analysis differences between samples can be discerned when it comes to parameters such as the average crystallite width, relative degree of crystallinity in the sample and the mean MFA.

The crystallite width of the *Arabidopsis* stem samples can be measured reliably and is (28 ± 2) Å [z]. This is consistent with crystallite width being determined by the cellulose biosynthesis and suggests that the biosynthesis was not affected destructively. Large variation was found between the MFA of different samples.

X-ray microtomography results suggest that the X-ray measurements did not damage the cell level structure of the samples. They also show that since the cells are not rectangular one cannot employ analysis methods that rely on cell rectangularity such as those that are used to study the MFA in wood.

Measurements are suggestive of a reduced mean crystallite width and mean degree of crystallinity of line 1a2 of construct #2 with respect to the WT. Similarly they suggest a reduction of mean width of line 4a9 of the same construct against the respective WT. For construct #1 the only statically significant difference in means was the increase in the mean MFA for line 3b5 against the WT. Further measurements should be carried out along these lines to verify or rule out these differences.

[z] All individual crystallite width values were within 26.4 and 29.7 Å verifying that there are no outliers in the data.

Acknowledgements

I would like to thank professor Ewa J. Mellerowicz and her group at the Department of Forest Genetics and Plant Physiology, Swedish University of Agricultural Sciences (SLU) for providing the *Arabidopsis thaliana* samples. In particular Mr. Prashant Pawar is credited for sample preparation.

References

[1] C. Payne, M. Himmel, M. Crowley, and G. Beckham, "Decrystallization of oligosaccharides from the cellulose Iβ surface with molecular simulation," *The Journal of Physical Chemistry Letters*, vol. 2, pp. 1546–1550, 2011.

[2] T. Arioli, L. Peng, A. Betzner, J. Burn, W. Wittke, W. Herth, C. Camilleri, H. Höfte, J. Plazinski, R. Birch, *et al.*, "Molecular analysis of cellulose biosynthesis in arabidopsis," *Science*, vol. 279, no. 5351, p. 717, 1998.

[3] E. Sjöström, *Wood Chemistry*. Academic Press, 2nd ed., 1993.

[4] S. Andersson, *A study of the nanostructure of the cell wall of the tracheids of conifer xylem by X-ray scattering*. PhD thesis, University of Helsinki, 2006.

[5] M.-P. Sarén, *Characterisation of properties of coniferous wood tracheids by x-ray diffraction, laser scattering and microscopy*. PhD thesis, University of Helsinki, 2006.

[6] M. Peura, *Studies on the cell wall structure and on the mechanical properties of Norway spruce*. PhD thesis, University of Helsinki, 2007.

[7] K. Pirkkalainen, *Nanoscale studies of materials using X-ray synchrotron radiation and mesoscopic modelling*. PhD thesis, University of Helsinki, 2011.

[8] J. Long, A. Conn, W. Batchelor, and R. Evans, "Comparison of methods to measure fibril angle in wood fibres," *Appita journal*, vol. 53, no. 3, pp. 206–209, 2000.

[9] L. Donaldson, "Microfibril angle: measurement, variation and relationships: a review," *IAWA Journal*, vol. 29, no. 04, pp. 387–396, 2008.

[10] U. Vainio, *Characterisation of cellulose- and lignin-based materials using X-ray scattering methods*. PhD thesis, University of Helsinki, 2007.

[11] W. Bragg, "The diffraction of short electromagnetic waves by a crystal," *Proceedings of the Cambridge Philosophical Society*, vol. XVII, pp. 43–57, 1913.

[12] A. Patterson, "The Scherrer formula for X-ray particle size determination," *Physical Review*, vol. 56, no. 10, p. 978, 1939.

[13] Z. Wang, B. Hsiao, B. Fu, L. Liu, F. Yeh, B. Sauer, H. Chang, and J. Schultz, "Correct determination of crystal lamellar thickness in semicrystalline poly (ethylene terephthalate) by small-angle X-ray scattering," *Polymer*, vol. 41, no. 5, pp. 1791–1797, 2000.

[14] S. Andersson, H. Wikberg, E. Pesonen, S. Maunu, and R. Serimaa, "Studies of crystallinity of scots pine and norway spruce cellulose," *Trees-Structure and Function*, vol. 18, no. 3, pp. 346–353, 2004.

[15] K. Pirkkalainen, M. Peura, K. Leppänen, A. Salmi, A. Meriläinen, P. Saranpää, and R. Serimaa, "Simultaneous X-ray diffraction and X-ray fluorescence microanalysis on secondary xylem of norway spruce," *Wood science and technology*, 2011. Submitted for review.

[16] H. Kerttunen, *Datan käsittely*. CSC - Tieteellinen laskenta Oy, 2nd ed., 2001.

[17] L. Wasserman, *All of Statistics: a Concise Course in Statistical Inference*. Springer Texts in Statistics, Springer, 2004.

[18] B. L. Welch, "The generalization of 'student's' problem when several different population variances are involved," *Biometrika*, vol. 34, pp. 28–35, 1947.

[19] I. M. Saxena and R. M. Brown, "Cellulose biosynthesis: Current views and evolving concepts," *Annals of Botany*, vol. 96, no. 1, pp. 9–21, July 2005.

[20] Y. Nishiyama, P. Langan, and H. Chanzy, "Crystal structure and hydrogen-bonding system in cellulose Iβ from synchrotron X-ray and neutron fiber diffraction," *Journal of the American Chemical Society*, vol. 124, no. 31, pp. 9074–9082, 2002.

[21] Y. Nishiyama, J. Sugiyama, H. Chanzy, and P. Langan, "Crystal structure and hydrogen bonding system in cellulose Iα from synchrotron X-ray and neutron fiber diffraction," *Journal of the American Chemical Society*, vol. 125, no. 47, pp. 14300–14306, 2003.

[22] R. M. Brown Jr., "Cellulose structure and biosynthesis," *Pure Appl. Chem*, vol. 71, no. 5, pp. 767–775, 1999.

[23] J. R. Hook and H. Hall, *Solid State Physics*. John Wiley and sons, second ed., 1991.

[24] L. Davies and P. Harris, "Atomic force microscopy of microfibrils in primary cell walls," *Planta*, vol. 217, no. 2, pp. 283–289, 2003.

[25] M. Doblin, I. Kurek, D. Jacob-Wilk, and D. Delmer, "Cellulose biosynthesis in plants: from genes to rosettes," *Plant and Cell Physiology*, vol. 43, no. 12, p. 1407, 2002.

[26] J. Takahashi, U. Rudsander, M. Hedenström, A. Banasiak, J. Harholt, N. Amelot, P. Immerzeel, P. Ryden, S. Endo, F. Ibatullin, *et al.*, "KORRIGAN1 and its aspen homolog PttCel9A1 decrease cellulose crystallinity in Arabidopsis stems," *Plant and cell physiology*, vol. 50, no. 6, p. 1099, 2009.

[27] I. Cave, "Theory of X-ray measurement of microfibril angle in wood," *Wood science and technology*, vol. 31, no. 4, pp. 225–234, 1997.

[28] D. Meinke, J. Cherry, C. Dean, S. Rounsley, and M. Koornneef, "Arabidopsis thaliana: a model plant for genome analysis," *Science*, vol. 282, no. 5389, p. 662, 1998.

[29] Arabidopsis Gene Initiative, "Analysis of the genome sequence of the flowering plant Arabidopsis thaliana," *Nature*, vol. 408, no. 6814, p. 796, 2000.

[30] The Arabidopsis Information Resource, "Ecotype search." http://www.arabidopsis.org/servlets/Search?action=new_search&type=ecotype, Aug. 2011.

[31] P. Derbyshire, K. Findlay, M. McCann, and K. Roberts, "Cell elongation in Arabidopsis hypocotyls involves dynamic changes in cell wall thickness," *Journal of experimental botany*, vol. 58, no. 8, p. 2079, 2007.

[32] N. Chaffey, E. Cholewa, S. Regan, and B. Sundberg, "Secondary xylem development in Arabidopsis: a model for wood formation," *Physiologia Plantarum*, vol. 114, no. 4, pp. 594–600, 2002.

[33] C. MacMillan, S. Mansfield, Z. Stachurski, R. Evans, and S. Southerton, "Fasciclin-like arabinogalactan proteins: specialization for stem biomechanics and cell wall architecture in Arabidopsis and Eucalyptus," *The Plant Journal*, vol. 62, no. 4, pp. 689–703, 2010.

[34] S. Turner and C. Somerville, "Collapsed xylem phenotype of Arabidopsis identifies mutants deficient in cellulose deposition in the secondary cell wall," *The Plant Cell Online*, vol. 9, no. 5, p. 689, 1997.

[35] P. Szyjanowicz, I. McKinnon, N. Taylor, J. Gardiner, M. Jarvis, and S. Turner, "The irregular xylem 2 mutant is an allele of korrigan that affects the secondary cell wall of Arabidopsis thaliana," *The Plant Journal*, vol. 37, no. 5, pp. 730–740, 2004.

[36] R. Newman, L. Davies, and P. Harris, "Solid-state ^{13}C nuclear magnetic resonance characterization of cellulose in the cell walls of Arabidopsis thaliana leaves," *Plant physiology*, vol. 111, no. 2, p. 475, 1996.

[37] M. Fujita, R. Himmelspach, C. H. Hocart, R. E. Williamson, S. D. Mansfield, and G. O. Wasteneys, "Cortical microtubules optimize cell-wall crystallinity to drive unidirectional growth in Arabidopsis," *Plant J*, vol. 66, pp. 915–28, June 2011.

[38] D. Burk and Z. Ye, "Alteration of oriented deposition of cellulose microfibrils by mutation of a katanin-like microtubule-severing protein," *The Plant Cell Online*, vol. 14, no. 9, p. 2145, 2002.

[39] T. Baskin, G. Beemster, J. Judy-March, and F. Marga, "Disorganization of cortical microtubules stimulates tangential expansion and reduces the uniformity of cellulose microfibril alignment among cells in the root of Arabidopsis," *Plant physiology*, vol. 135, no. 4, p. 2279, 2004.

[40] S. Sato, T. Kato, K. Kakegawa, T. Ishii, Y. Liu, T. Awano, K. Takabe, Y. Nishiyama, S. Kuga, S. Sato, *et al.*, "Role of the putative membrane-bound endo-1, 4-β-glucanase KORRIGAN in cell elongation and cellulose synthesis in Arabidopsis thaliana," *Plant and Cell Physiology*, vol. 42, no. 3, p. 251, 2001.

[41] P. A. Penttilä, A. Várnai, K. Leppänen, M. Peura, A. Kallonen, P. Jääskeläinen, J. Lucenius, J. Ruokolainen, M. Siika-aho, L. Viikari, and R. Serimaa, "Changes in submicrometer structure of enzymatically hydrolyzed microcrystalline cellulose," *Biomacromolecules*, vol. 11, no. 4, pp. 1111–1117, 2010. PMID: 20329744.

[42] P. Penttilä, J.-P. Suuronen, S. Kirjoranta, M. Peura, K. Jouppila, M. Tenkanen, and R. Serimaa, "X-ray characterization of starch-based solid foams," *Journal of Materials Science*, vol. 46, pp. 3470–3479, 2011. 10.1007/s10853-011-5252-y.

[43] M. Tahvanainen, T. Rotko, E. Mäkilä, H. Santos, D. Neves, T. Laaksonen, A. Kallonen, K. Hämäläinen, M. Peura, R. Serimaa, *et al.*, "Tablet preformulations of indomethacin-loaded mesoporous silicon microparticles," *International Journal of Pharmaceutics*, vol. 422, pp. 125–131, 2011.

[44] J. Virtasalo, E. Bonsdorff, M. Moros, K. Kabel, A. Kotilainen, D. Ryabchuk, A. Kallonen, and K. Hämäläinen, "Ichnological trends along an open-water transect across a large marginal-marine epicontinental basin, the modern baltic sea," *Sedimentary Geology*, vol. 241, pp. 40–51, 2011.

[45] A. Salmi, A. Merilainen, E. Hæggström, M. Torkkeli, and R. Serimaa, "Combining X-rays and ultrasound to determine micro-elasticity," in *Ultrasonics Symposium, 2007. IEEE*, pp. 54–57, IEEE, 2007.

[46] M. Peura, M. Müller, U. Vainio, M. Sarén, P. Saranpää, and R. Serimaa, "X-ray microdiffraction reveals the orientation of cellulose microfibrils and the size of cellulose crystallites in single Norway spruce tracheids," *Trees-Structure and Function*, vol. 22, no. 1, pp. 49–61, 2008.

[47] Encyclopædia Britannica, "angiosperm." Encyclopædia Britannica Online Academic Edition, http://www.britannica.com/EBchecked/topic/24667/angiosperm, July 2011.

Glossary

2θ

The 2θ-scale is the scale for the diffraction as seen on the detector. It is calculated from the center of the diffraction patter in perpendicular transmission mode (Figure 11). The 2θ angle is called the scattering angle and is twice the Bragg angle, Eq. (1), and is related to the q-scale by Eq. (2). 5, 8, 9, 19, 21–24, 26, 33, 34, 36, 37, 44

angiosperm

A flowering plant that comprises of roots, a stem and leaves and belongs to the largest group of plants (angiosperms) [47]. 11

Arabidopsis thaliana

a small flowering plant that is an important model system for identifying genes and their functions. Commonly known as wall cress or mouse-ear cress. 3, 11, 15–18, 22, 27, 29, 31, 33, 35, 37, 38, 44

hypocotyl

"The stem is an aerial axis of the plant that bears leaves and flowers and conducts water and minerals from the roots and food from the site of synthesis to areas where it is to be used. The main stem of a plant is continuous with the root system through a transition region called the *hypocotyl*. In the developing embryo, the hypocotyl is the embryonic axis that bears the seedling leaves (cotyledons)" [47]. 15

Microfibril Angle

The orientation of the cellulose microfibrils in the secondary cell wall with respect to the longitudinal cell axis. The microfibrils are organized in helical windings around the cell. The variation in the average microfibril angle of wood is assumed to be mostly affected by the S_2 layer [6]. 4, 9, 14, 16, 29, 31, 37, 38, 44

S_1

outer layer of the secondary wall in wood cells. 14

S_2

middle layer of the secondary wall in wood cells. Contains most of the mass of the wood fibre and therefore the longitudinal stiffness of wood is very dependent on the microfibrillar angle of this layer. 14, 43

S_3

inner layer of the secondary wall in wood cells. 14

Small Angle X-ray Scattering

an X-ray measuring method that focuses on scattering to small 2θ angles ($2\theta < 5°$). 44, *See* XRD & WAXS

Wide Angle X-ray Scattering

an X-ray measuring method that focuses on scattering to large 2θ angles ($2\theta > 5°$). 6, 19, 35, 44, *See* XRD & SAXS

X-Ray Diffraction

a scattering process from crystalline samples. The term WAXS is used to cover also the scattering from partly amorphous samples, such as the samples measured in this paper. For large 2θ angles ($2\theta > 5°$) the term Wide Angle X-ray Diffraction (WAXD) is sometimes used. 44

Acronyms

2D Two-Dimensional. 6, 7, 21

Arabidopsis *Arabidopsis thaliana*. 3, 11, 15–18, 22, 27, 29, 31, 33, 35, 37, 38, 43

cesA cellulose synthase catalytic subunit. 11

CML Compound Middle Lamella. 14

FWHM Full Width at Half Maximum. 7, 8, 11, 33

GOF Goodness Of Fit. 27, 51

MFA Microfibril Angle. 4, 9, 14, 16, 29, 31, 37, 38, 43

ML Middle Lamella. 14

NMR Nuclear Magnetic Resonance. 16, 29

SAXS Small Angle X-ray Scattering. 4, 35, 44

SEM Scanning Electron Microscopy. 15, 16

WAXS Wide Angle X-ray Scattering. 4, 6, 19, 27, 33, 35, 38, 44

WT Wild Type. 15, 16, 18, 27, 29, 31, 38

XRD X-Ray Diffraction. 5, 44

Appendices

List of Appendices

A X-ray Microtomography Reconstruction images **46**

B Individual Wide Angle X-ray Scattering Results **48**

C Air Scattering Correction for Crystallinity Fit **51**

A X-ray Microtomography Reconstruction images

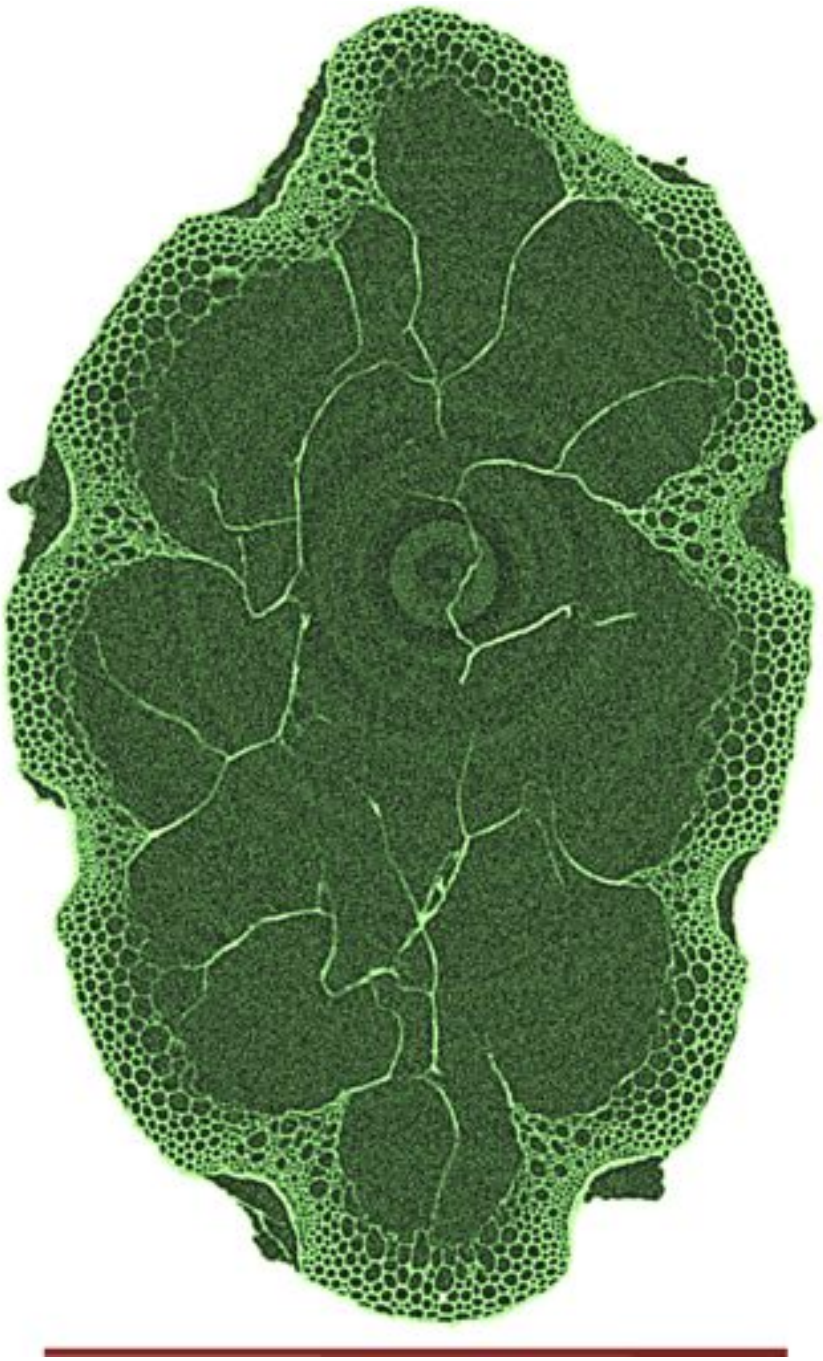


Figure 27. A slice taken from a tomographic reconstruction volume of a section of the stem from a WT plant (construct #1, replicate #9). Scale bar is 1.0 mm.

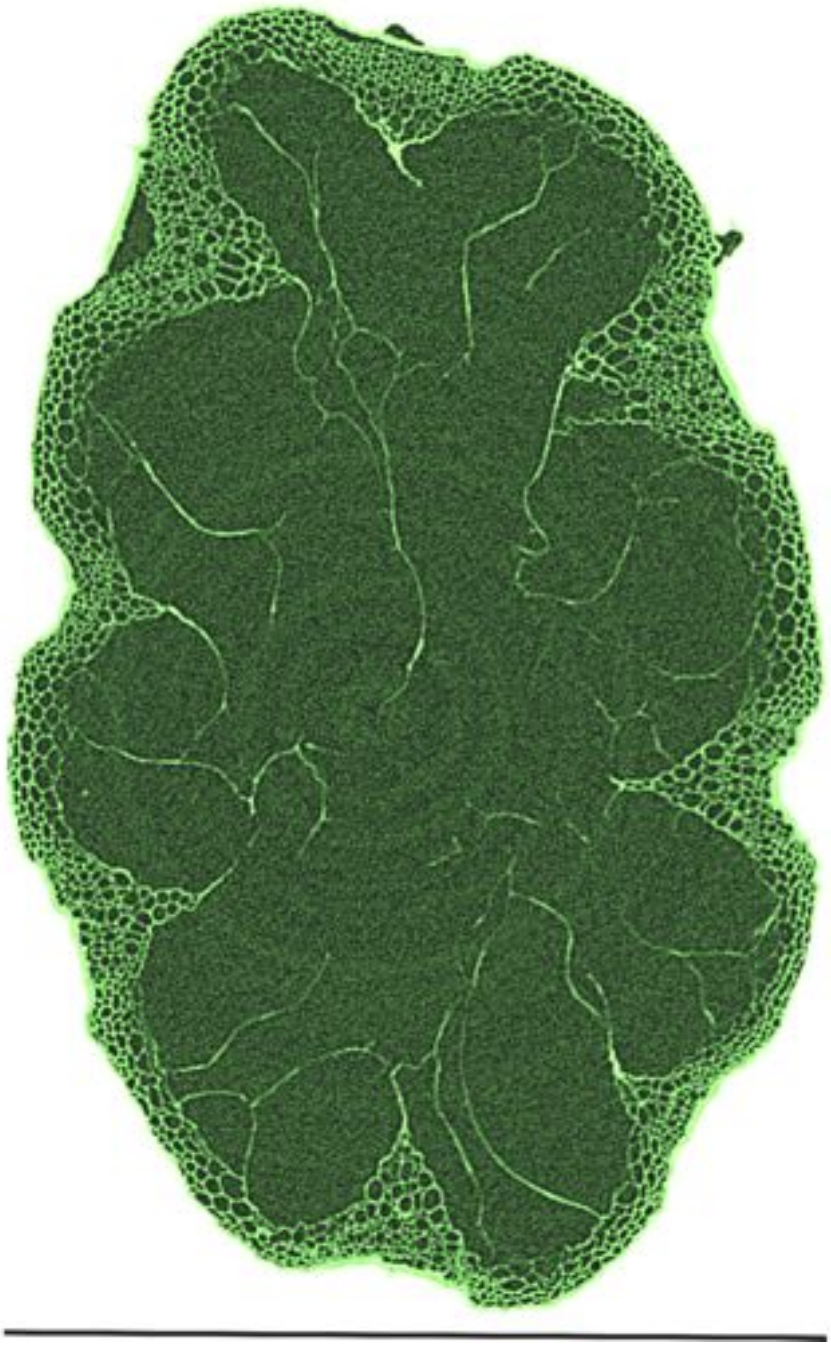


Figure 28. A slice taken from a tomographic reconstruction volume of a section of the stem from genetically modified plant (construct #2, line 1a2, replicate #3). Scale bar is 1.0 mm.

B Individual Wide Angle X-ray Scattering Results

Table 1. The results from the crystallite width, degree of crystallinity and microfibril mean angle calculations for individual samples from *Arabidopsis* main stem, construct FC1, wild type plants.

Sample name	**Crystallite width,** W **[±1.5 Å]**	**Crystallinity,** C **[±0.03]**	**Microfibril angle (mean),** $\langle \mathtt{MFA} \rangle$**[±2 °]**
FC1 WT (1)	27.7	0.29	10
FC1 WT (2)	26.5	0.29	12
FC1 WT (3)	27.8	0.24	8
FC1 WT (4)	28.1	0.24	9
FC1 WT (5)	28.4	0.27	10
FC1 WT (6)	27.1	0.29	10
FC1 WT (7)	29.0	0.29	12
FC1 WT (8)	28.1	0.29	3
FC1 WT (9)	27.8	0.32	8
mean ± std	**27.8±0.7**	**0.28±0.03**	**9±3**

Table 2. The results from the crystallite width, degree of crystallinity and mean microfibril angle calculations for individual samples from *Arabidopsis* main stem, construct FC1, genetically modified plants.

Sample name	**Crystallite width,** W **[±1.5 Å]**	**Crystallinity,** C **[±0.03]**	**Microfibril angle (mean),** $\langle \mathtt{MFA} \rangle$**[±2 °]**
FC1 2a3 (1)	27.0	0.26	10
FC1 2a3 (2)	26.4	0.29	9
FC1 2a3 (3)	28.0	0.26	10
FC1 2a3 (4)	28.1	0.26	20
FC1 2a3 (6)	26.5	0.27	10
FC1 2a3 (7)	28.3	0.19	16
FC1 2a3 (8)	27.4	0.24	8
mean ± std	**27.4±0.8**	**0.25±0.03**	**12±4**
FC1 3b5 (1)	28.4	0.31	17
FC1 3b5 (2)	28.9	0.27	13
FC1 3b5 (3)	27.4	0.28	10
FC1 3b5 (4)	28.4	0.25	20
FC1 3b5 (5)	28.2	0.25	19
FC1 3b5 (6)	27.0	0.24	10
FC1 3b5 (7)	28.1	0.24	12
FC1 3b5 (8)	28.2	0.24	9
FC1 3b5 (9)	27.7	0.31	16
mean ± std	**28.0±0.6**	**0.26±0.03**	**14±4**
FC1 4a8 (1)	27.4	0.28	11
FC1 4a8 (2)	27.4	0.27	14
FC1 4a8 (3)	28.0	0.25	21
FC1 4a8 (4)	27.9	0.26	3
FC1 4a8 (5)	28.7	0.26	10
FC1 4a8 (6)	29.0	0.30	10
FC1 4a8 (7)	27.3	0.29	9
FC1 4a8 (8)	29.2	0.25	10
FC1 4a8 (9)	27.4	0.31	15
mean ± std	**28.0±0.8**	**0.27±0.02**	**11±5**
FC1 5a3 (1)	27.6	0.29	12
FC1 5a3 (3)	27.0	0.28	2
FC1 5a3 (4)	28.0	0.23	5
FC1 5a3 (5)	28.4	0.25	21
FC1 5a3 (6)	29.7	0.20	5
FC1 5a3 (7)	27.4	0.28	10
FC1 5a3 (8)	26.9	0.29	20
mean ± std	**27.9±1.0**	**0.26±0.03**	**11±7**
mean ± std	**27.9±0.8**	**0.26±0.03**	**12±5**

Table 3. The results from the crystallite width, degree of crystallinity and mean microfibril angle calculations for individual samples from *Arabidopsis* main stem, construct FC2, wild type plants.

Sample name	**Crystallite width,** W **[±1.5 Å]**	**Crystallinity,** C **[±0.03]**	**Microfibril angle (mean),** $\langle \mathtt{MFA} \rangle$**[±2 °]**
FC2 WT (1)	27.7	0.29	9
FC2 WT (2)	28.1	0.30	16
FC2 WT (3)	28.4	0.31	9
FC2 WT (4)	28.6	0.26	18
FC2 WT (5)	27.5	0.23	11
FC2 WT (6)	28.7	0.26	10
mean ± std	**28.2±0.5**	**0.28±0.03**	**12±4**

Table 4. The results from the crystallite width, degree of crystallinity and mean microfibril angle calculations for individual samples from *Arabidopsis* main stem, construct FC2, genetically modified plants.

Sample name	**Crystallite width,** W **[±1.5 Å]**	**Crystallinity,** C **[±0.03]**	**Microfibril angle (mean),** $\langle \mathtt{MFA} \rangle$**[±2 °]**
FC2 1a2 (1)	26.9	0.22	11
FC2 1a2 (2)	26.9	0.19	4
FC2 1a2 (3)	27.3	0.24	16
mean ± std	**27.0±0.2**	**0.22±0.02**	**11±6**
FC2 2a3 (1)	28.3	0.27	18
FC2 2a3 (2)	27.9	0.26	17
FC2 2a3 (3)	28.3	0.24	9
mean ± std	**28.1±0.2**	**0.26±0.01**	**14±5**
FC2 4a9 (1)	26.8	0.25	10
FC2 4a9 (2)	26.9	0.30	17
FC2 4a9 (3)	27.4	0.25	11
mean ± std	**27.1±0.3**	**0.27±0.03**	**13±4**
FC2 6c5 (1)	29.3	0.29	15
FC2 6c5 (2)	26.8	0.28	14
FC2 6c5 (3)	28.3	0.27	12
mean ± std	**28.1±1.3**	**0.28±0.01**	**14±1**
FC2 6c7 (1)	27.6	0.29	15
FC2 6c7 (2)	27.6	0.30	10
FC2 6c7 (3)	29.1	0.28	12
mean ± std	**28.1±0.9**	**0.29±0.01**	**12±2**
mean ± std	**27.7±0.6**	**0.26±0.02**	**13±4**

C Air Scattering Correction for Crystallinity Fit

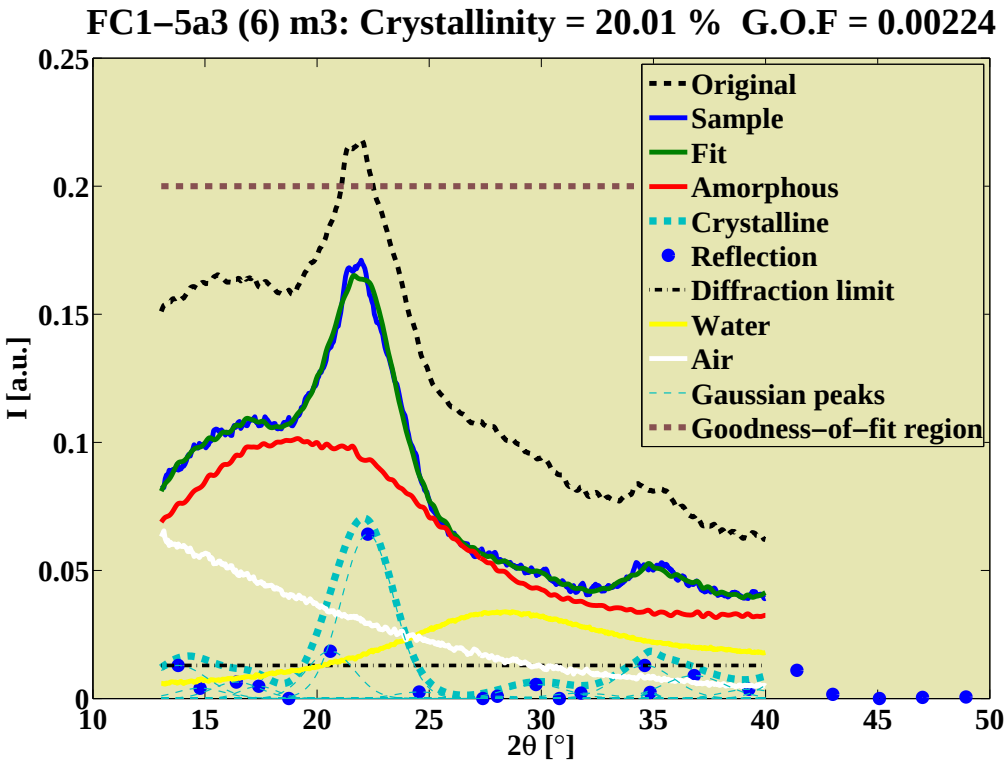


Figure 29. For samples FC1-5a3 (5) and FC1-5a3 (6) the air scattering reduction was not good. Therefore air scattering intensity (white solid line) was added as a fitting parameter which provided sensible degree of crystallinity and a good value for the Goodness Of Fit.