The occurrence of seven-fold helical molecular conformation in cellulose-phosphoric acid complex

Jia Hui Lim, Isabelle Morfin, Isabelle Jeacomine, Alfred French, Yoshiharu Nishiyama, Yu Ogawa

The occurrence of seven-fold helical molecular conformation in cellulose-phosphoric acid complex.

Jia Hui Lim,[1] Isabelle Morfin,[2] Isabelle Jeacomine,[1] Alfred D. French,[3] Yoshiharu Nishiyama,[1] Yu Ogawa[1, *]

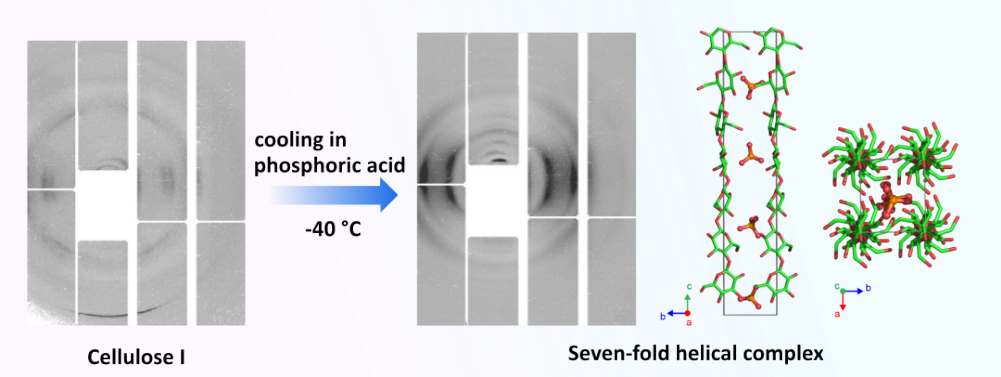
1 Univ. Grenoble Alpes, CNRS, CERMAV, 38000 Grenoble, France
2 Univ. Grenoble Alpes, CNRS, LiPhy, 38000 Grenoble, France
3 Southern Regional Research Center, U.S. Department of Agriculture, New Orleans, LA 70124, USA
* corresponding author: yu.ogawa@cermav.cnrs.fr

Abstract

Phosphoric acid is widely used for the swelling and hydrolysis of cellulose. The detailed description of molecular interactions between cellulose and phosphoric acid is essential for understanding and controlling these processes. Here, to obtain structural insights into the swelling behavior, we investigated the structural evolution of cellulose swollen in the concentrated phosphoric acid solution using X-ray fiber diffraction and solid-state NMR spectroscopy. We observed the formation of a crystalline complex of cellulose and phosphoric acid at -40 °C, where cellulose molecules adopt a seven-fold helical conformation. This structure is the second known cellulose-acid crystalline complex and the first cellulosic crystal consisting of seven-fold helical chains. Our observation highlights the conformational flexibility of cellulose molecules in the solvated states and the strong influence of cellulose-acid interactions on the packing and conformation of cellulose molecules.

Keywords: Cellulose; Phosphoric acid; Crystallosolvate; seven-fold helix

Graphical abstract



Introduction

Cellulose forms crystalline complexes with various polar solvent molecules. The crystal structure analysis of these solvated crystals, or crystallosolvates, has provided valuable insights into the molecular interactions between cellulose and solvent molecules and the conversion mechanisms between cellulose crystalline allomorphs. The crystalline complexes with alkalis and amines are among the most extensively studied cellulose crystallosolvates as intermediate structures in allomorphic conversions (Sobue et al. 1939; Lee et al. 1983; Okano and Sarko 1984, 1985; Porro et al. 2007; Wada et al. 2009). These molecules are the components of historic cellulose solvents such as aqueous alkali solutions and cupriethylenediamine hydroxide solutions. In recent years, new cellulose crystallosolvates were reported with industrially relevant solvent molecules such as N-methyl morpholine-N-oxide (NMMO) and ionic liquids (Song et al. 2018; Nishiyama et al. 2019; Endo et al. 2020).

Mineral acids, such as sulfuric and phosphoric acid, are essential reagents for processing cellulose materials. They are widely-used hydrolysis and swelling agents of cellulose. Concentrated mineral acid solutions are also direct solvents of cellulose. Despite their importance, the complexation of cellulose and acid molecules has not received attention compared to the solvent molecules mentioned above. This situation hinders us from understanding the molecular-level details of the interactions between cellulose and acids. We recently reported the cellulose-sulfuric acid complex as the first cellulose-acid crystallosolvate (Li et al. 2022). The complexation of cellulose and sulfuric acid occurs only at low temperatures, and cellulose molecules adopt a rare five-fold helical conformation (five glucose residues in two turns, or n=2.5) in the complex structure. Such a structure was not without precedent; earlier work with the gelling molecule xanthan had shown that its cellulose backbone could apparently form

both single- and double-helices with each strand having the n=2.5 conformation (French and Johnson, 2004; Moffat et al. 2016; Morris 2019).The complexation was likely driven by the deficiency of water around sulfuric acid molecules that recruit cellulose hydroxyl groups to form a stable hydrate-like structure (Li et al. 2022). This finding prompted us further to explore the complexation behavior of cellulose and mineral acids. In this study, we extended a similar approach to concentrated phosphoric acid.

Phosphoric acid is widely used as a swelling agent of cellulose. It dissolves cellulose at room temperature, and subsequent regeneration of cellulose yields amorphous cellulose. Such amorphous cellulose, so-called phosphoric acid swollen cellulose or PASC, is the most widely used form of amorphous cellulose and a standard substance for assays of cellulolytic enzyme activity (Zhang et al. 2006; Hall et al. 2010). While sulfuric acid is more common in literature, phosphoric acid is also used to hydrolyze cellulose to prepare cellulose oligomers and nanocrystals (Isogai and Usuda 1991; Camarero Espinosa et al. 2013; Billès et al. 2016). Furthermore, like sulfuric acid, concentrated phosphoric acid solution can dissolve cellulose at high concentrations. Cellulose molecules form a liquid crystalline phase in the resulting highly concentrated dope (Boerstoel et al. 2001). The anisotropic molecular organization in the dope can be exploited to obtain highly oriented cellulose II fibers. These fibers, called Fiber B, exhibit improved mechanical properties compared to the other commercially regenerated cellulose fibers (Northolt et al. 2001).

Given its importance in the processing of cellulose, it is crucial to better understand the molecular interactions between cellulose and phosphoric acid. Here, we investigated the structure of cellulose swollen in a concentrated phosphoric acid solution at low temperatures. Using X-ray fiber diffraction and ^{13}C solid-state nuclear magnetic resonance (NMR) spectroscopy, we observed the formation of a crystallosolvate of cellulose and phosphoric acid at -40 °C. Interestingly, the structure has an exceptionally long fiber repeat of 37 Å, making it the first example of seven-fold helical molecular conformation in cellulosic crystals. This extended form, seven-fold helix is unlike the collapsed seven-fold V_7 amylose structure (Nishiyama et al. 2010).

Material and Methods

X-ray scattering measurement

X-ray scattering measurements were carried out at the D2AM beamline at the European Synchrotron Radiation Facility (ESRF) using an X-ray beam of 16 KeV and hybrid pixel detectors (D5 and WOS). Bundles of well-aligned flax fibers were packed in glass capillaries and immersed in 83 wt% H_3PO_4 aqueous solution at room temperature. The specimens were then mounted on a low-temperature specimen holder and placed under a vacuum. The diffraction patterns were acquired at temperatures ranging from room temperature to -40°C with an exposure time of 3 seconds. The background scattering from the solvent in the patterns was subtracted numerically using in-house programs.

Solid-state ^{13}C NMR spectroscopy

Solid-state ^{13}C CP/MAS NMR analysis was performed on swollen and dry regenerated flax fibers using a Bruker Avance III 400 spectrometer (100 MHz, ^{13}C). For the swollen samples, spectra were obtained at room and low temperatures (23, -20, and -40°C). Approximately 50 mg of flax fibers were swollen in 83 wt% H_3PO_4 solution and allowing them to swell for 10 minutes. The fibers were then mixed with ca. 10 mg of glass wool and packed into the zirconia rotor. The specimen temperature was controlled using a Bruker BCU-II cooling unit. The spinning rate was set at 9 kHz with 2 ms contact time.

Model construction and analysis

Models of cellulose molecules in seven-fold conformations were built using a glucose residue of the corner chain of cellulose I_β. Each residue was rotated about the chain axis by 51° (7_1), 103° (7_2), 154° (7_3), 206° (7_4), 257° (7_5), and 309° (7_6) and shifted along the chain axis by 5.29 Å compared to the adjacent residue. The models were visualized and analyzed using Pymol software (Schrödinger and DeLano 2020). The glycosidic torsion angles of the generated helices were compared with the crystal structure conformations of cellulose allomorphs and di- and oligosaccharide crystals from the Protein Data Base with cellobiose linkage as well as the conformational energy map of cellobiose based on SMD continuum solvation and B3LYP/6-31+G(d) energies (French et al., 2012, 2021).

Results and Discussion

Crystal structure of cellulose-phosphoric acid complex

We performed in situ low-temperature X-ray scattering measurements to monitor the changes in the crystal structure of flax cellulose immersed in concentrated H_3PO_4 aqueous solution. We chose a concentration of 83 wt% since more concentrated solutions could dissolve cellulose fiber at room or low temperature. Furthermore, this concentration is commonly used in the preparation of amorphous cellulose.

Fig. 1A shows the X-ray fiber diffraction diagram of flax fiber immersed in 83 wt% H_3PO_4 at room temperature (20 °C) for 15 minutes. The room-temperature diagram shows characteristic reflections of cellulose I_β, i.e., equatorial 1 -1 0 and 1 1 0, and meridional 0 0 4. Thus, the room-temperature immersion in 83 wt% H_3PO_4 did not affect the crystal structure of flax cellulose for at least 15 min.

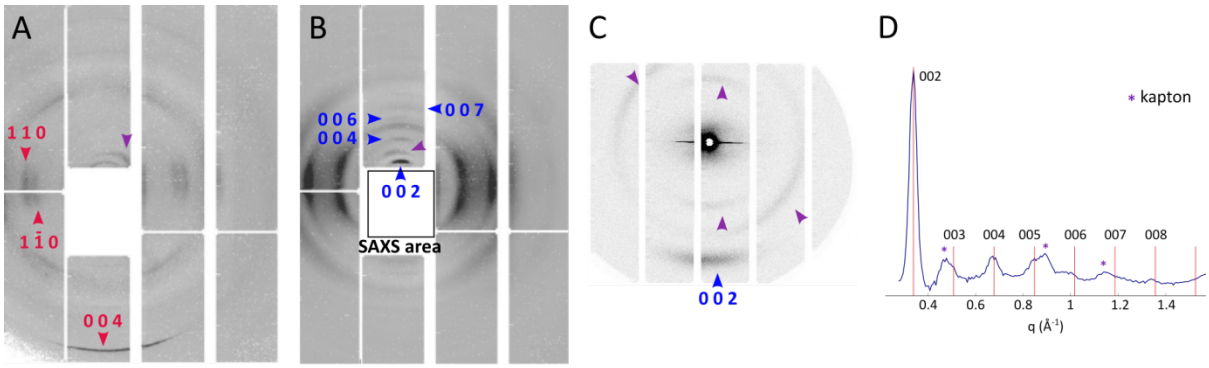


Fig. 1 X-ray fiber diffraction patterns of flax cellulose immersed in 83 wt% phosphoric acid solution at (A) room temperature and (B, C) -40 °C. Panel (C) corresponds to the small angle area indicated in panel (B). Reflections indicated by purple arrowheads correspond to the Kapton window used in the low-temperature vacuum cell. (D) Meridional profile of the diffraction diagram in (B). Reflection indices are shown above peaks.

When cooled down to -40 °C (Fig. 1B and 1C), the diagram showed diffraction features different from cellulose I_β, indicating the crystalline complex formation of cellulose and phosphoric acid. The diffraction pattern of the cellulose-phosphoric acid complex is poorly resolved, with only a few strong reflections in the observed wide-angle area. On the meridian of the fiber diagram, a new set of reflection spots appeared at d = 18.5, 9.3, 6.2, 5.3 Å (Fig. 1D). No other reflection appears on the meridian in the smaller angle region (Fig. 1C). This meridional pattern indicates that

149 the fiber repeat of the complex crystal is 37 Å, increased from that of cellulose I_β (10
150 Å). The aforementioned meridional reflections were thus indexed, from the utmost to
151 the outermost, as 0 0 2, 0 0 4, 0 0 6, and 0 0 7, respectively (Fig. 1D). The long 37-Å
152 fiber repeat suggests cellulose molecules likely adopt a seven-fold helical
153 conformation in the complex. The rise per residue, h, is thus 5.28 Å for this helix, the
154 longest among the known cellulosic crystals (cf. 5.19 Å for cellulose I_β).

155 We refined the unit cell based on 11 independent reflections indicated in Fig.
156 2. The refined unit cell is orthorhombic with the cell parameters of a = 6.19 Å, b = 6.95
157 Å, and c = 37.05 Å. In Fig. 2, the reflection positions predicted from the unit cell are
158 superimposed on the diffraction pattern remapped in the polar coordinate. They are in
159 good agreement with the observed reflection positions. The unit cell volume is 1594
160 $Å^3$. The volumes of each component molecule are 164 $Å^3$ for glucose residue based
161 on cellulose I_β crystal structure (Nishiyama et al. 2002), 75 $Å^3$ for phosphoric acid
162 based on its hemihydrate crystal structure (Smith et al. 1955), and 30 $Å^3$ for water.
163 Thus, the unit cell contains only one cellulose chain composed of seven glucose
164 residues (1148 $Å^3$). If the molar ratio between phosphoric acid and water remains
165 unchanged at 1:1 in the 83 wt% solution and the complex, the unit cell may
166 accommodate four phosphoric acid and four water molecules (420 $Å^3$). This
167 stoichiometric ratio makes the sum of the component volume, 1568 $Å^3$, nearly equal
168 to the unit cell volume. The strong 0 0 2 and 0 0 4 intensities, corresponding to
169 periodicities of 18.5 and 9.3 Å along the fiber axis, support the even number of
170 phosphoric acid molecules in the unit cell.

171

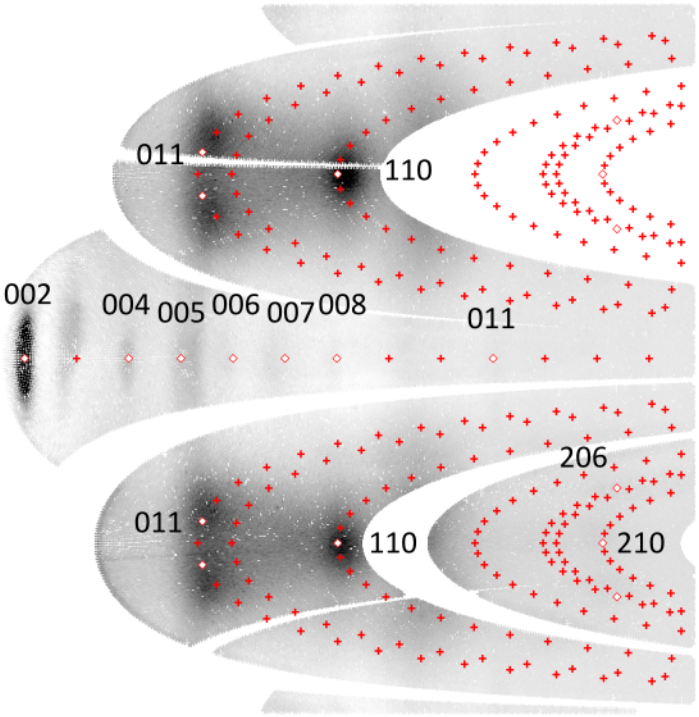


Fig. 2 X-ray diffraction pattern of flax cellulose immersed in 83 wt% H_3PO_4 in polar coordinates. Reflection positions predicted based on the refined orthorhombic unit cell parameters are indicated with red crosses. 11 reflections used for the unit cell refinement are indicated with white circles.

While the calculated diffraction positions fit well with the experimental ones, the real unit cell is bigger than the one-chain orthorhombic unit cell. No two-fold rotation axis can be conceived along either unit-cell axis due to the low symmetry of cellulose molecule in the seven-fold structure. This unit cell thus does not satisfy the symmetry conditions of the orthorhombic systems. The crystallographically correct unit cell should have each axis at least twice as long as the current proposed values. Nevertheless, we believe that the molecular packing and structural details in the complex can be discussed based on this small one-chain cell since it is likely an asymmetric unit of the real unit cell.

To confirm the seven-fold helical molecular conformation, we conducted the solid-state NMR spectroscopy of the cellulose-phosphoric acid complex. We measured CP/MAS ^{13}C NMR spectra of cellulose swollen in 83 wt% H_3PO_4 solution at three temperature conditions, i.e., room temperature (23 °C), -20 °C, and -40 °C.

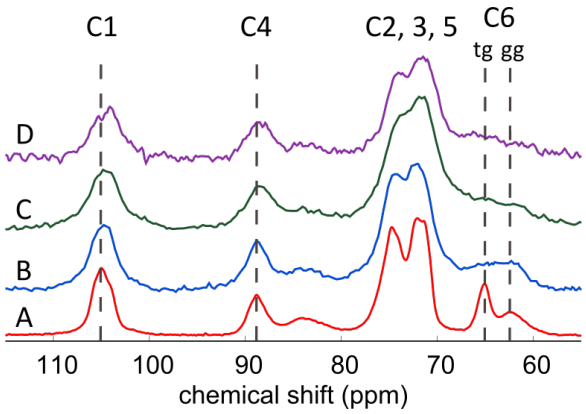


Fig. 3 CP/MAS ^{13}C NMR spectra of flax cellulose under different conditions. (a) intact flax cellulose and (b-d) flax cellulose swollen in 83 wt% H_3PO_4 and measured at (b) room temperature, (c) -20 °C, and (d) -40 °C.

As shown in Fig. 3, the spectra of swollen samples (Fig. 3b-d) are different from that of intact flax cellulose (Fig. 3a), especially in the C6 region (60-68 ppm) (Horii et al. 1983). The sharp signal at 66 ppm of cellulose I_β, corresponding to the *tg* conformer, disappeared in all the swollen samples. The swelling in phosphoric acid thus resulted in the loss of the original cellulose I_β crystal structure regardless of the temperature. The room-temperature swollen sample had a broad C6 signal at 62 ppm corresponding to the *gg* conformer, while even more broader peaks were observed at -20 °C and -40 °C (Fig. 3c and 3d).

The loss of cellulose I_β crystal structure at room temperature contradicts the X-ray diffraction result at room temperature (Fig. 1A). This discrepancy likely arose from the duration of swelling. We carried out the X-ray scattering measurement at room temperature after 15 min of swelling. In comparison, the sample preparation time and measurements of solid-state NMR experiments were ca. 30 min and 5 hours, respectively. Furthermore, the swelling of flax fiber was laterally constrained by the glass capillary in the X-ray scattering experiments, which may have further slowed down the swelling process at room temperature.

Compared to the drastically altered C6 signals, the changes in the other carbon signals are less significant. The C1 and C4 upfield shifted by 1 and 0.5 ppm, respectively, at -40 °C compared to those of cellulose I_β. These upfield shifts suggest

the conformational change of the cellulose molecular chain around the glycosidic linkage at the low temperature. The observation is consistent with the occurrence of the seven-fold helical chain conformation in the complex predicted based on the X-ray diffraction results. The change in the C1 and C4 chemical shifts is less pronounced at the higher temperature.

The chemical shifts of C1 and C4 are known susceptible to the conformational changes of glycosidic linkage of β-1,4 (e.g. cellulose and xylan) and α-1,4 linked glycans (e.g. amylose) (Veregin et al. 1987; Porro et al. 2007; Falcoz-Vigne et al. 2017). The magnitude of the observed C1 upfield shift is similar to those of the solvation-induced conformation changes in cellulose and xylan crystals. For instance, the C1 signal shifts up-field by ca. 1 ppm in Na-cellulose II with three-fold helical cellulose chains compared to cellulose I with two-fold helices. On the other hand, the same conformational change gives a larger up-field shift of the C4 signal, by 9 ppm, in the transition from cellulose I to Na-cellulose II (Porro et al. 2007).

The small C4 upfield shift likely indicates the similarity of the molecular conformation in the complex to the two-fold one in cellulose I. The seven-fold symmetry allows six possible helices: 7_1, 7_2, 7_3, 7_4, 7_5, and 7_6 helices. The nominal rotation angles between adjacent glucose residues are 51° (7_1), 103° (7_2), 154° (7_3), 206° (7_4), 257° (7_5), 309° (7_6), respectively. The first three are right-handed, whereas the last three are considered left-handed because a left-handed helical thread that connects equivalent points on helices with a rotation angle greater than 180° has a shorter path by proceeding in a left-handed manner. The 7_3 and 7_4 helices thus have high conformational similarities to the two-fold helices of native cellulose, which have a 180° rotational angle between residues.

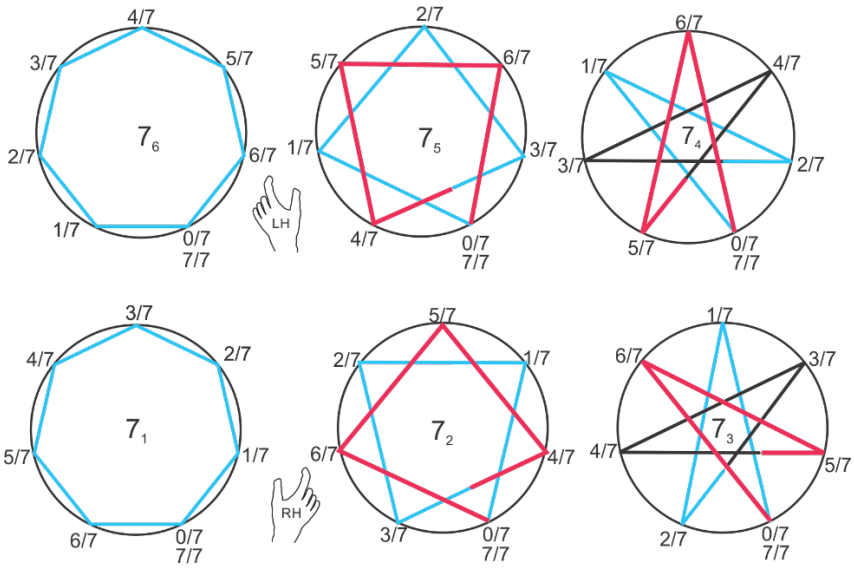


Fig. 4. The various 7-fold helices viewed down the helix axis. Blue colors indicate that the lines that connect equivalent points are in the first helix turn, the red lines (for the 2- and 3-turn repeats) are in the last turn (closest to the viewer) and the black lines (for the 3-turn repeat structures) are in the second turn.

Fig. 4 illustrates the possible seven-fold helices with circles representing the projected helical thread that connects equivalent points (French and Johnson, 2009). Straight lines connect arbitrarily assigned equivalent points. It is initially helpful to think of the equivalent points as the linkage oxygen atoms. In that case, the lines can represent the O1—O4 virtual bonds of the glucose residue. The numbers around the circles indicate the fractional height along the z-axis for the structure repeat, regardless of the number of turns. For the right-handed helix in the lower left corner, 7_1, with seven residues in one turn, the height of the equivalent points starts at 0 and increases toward the viewer after a counter-clockwise rotation of about 51.4° to the second point at 1/7 of the helix repeat. A second 51° rotation gives a height of 2/7 of the repeat and so-forth until one complete turn brings the equivalent point back to the 7/7 point where the helix generation process begins for a second turn.

For the 7_6 helix above it, also with seven residues in one turn, clockwise rotations of 51.4° each produce the analogous left-handed helix. If a rotation of 309° had been applied in a counter-clockwise sense from the 0/7 point, it would have brought the equivalent point to the 6/7 point on the right-handed lower drawing. But

that is in the exact same place as the 1/7 point on the left-handed helix and the path is much shorter. This is why this left-handed helix is said to have 7_6 helix symmetry.

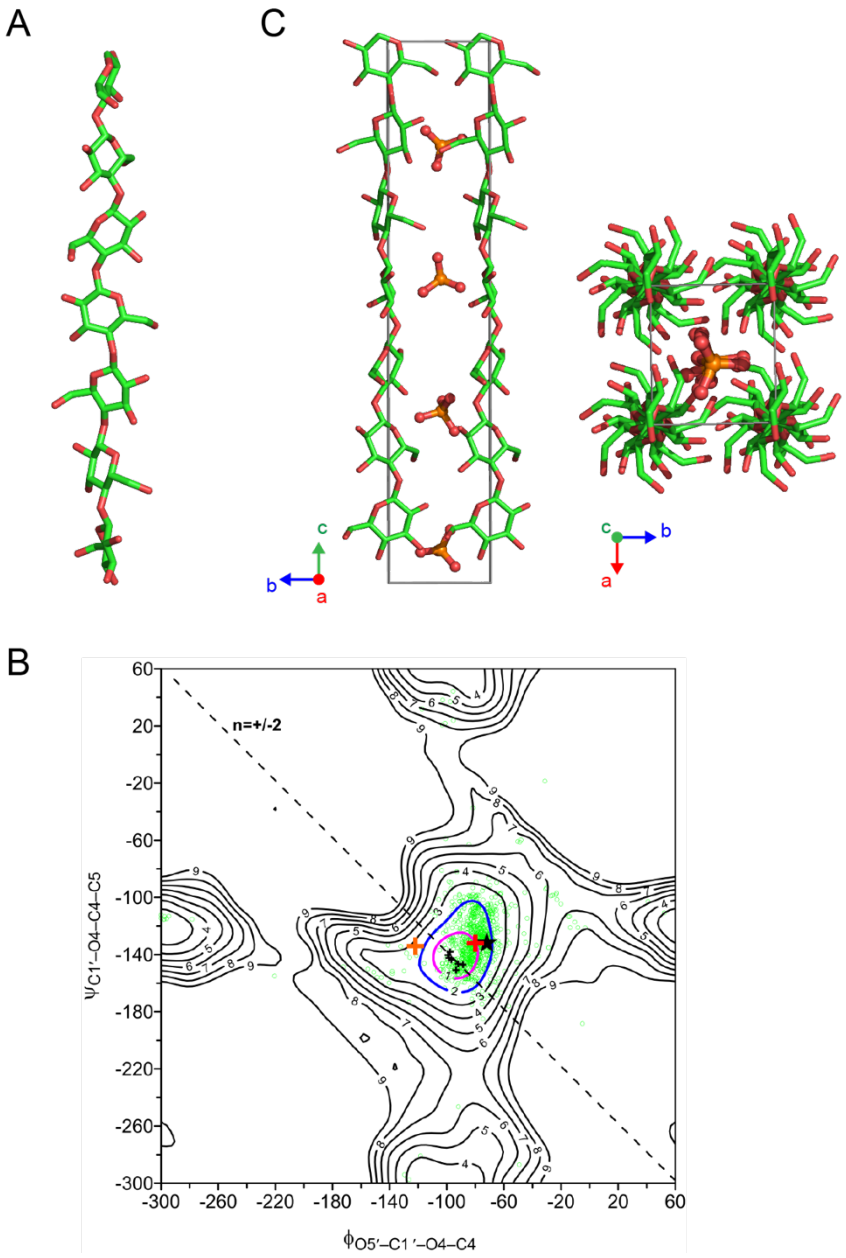
Looking at the 7_5 helix in the upper middle, the equivalent points are spaced at 102.9° clockwise increments. Thus 3.5 residues will be needed to complete the first turn, the path indicated by the blue line that connects equivalent points. The path of these lines connecting equivalent points for the second turn of the helix is indicated by the red lines, taking again 3.5 residues to complete the second turn. Here, the 1/7 point indicates that the equivalent point has moved towards the viewer by one seventh of the two-turn structure repeat distance.

The upper right diagram depicts the 7_4 helix. To get to the 1/7 point, a clockwise rotation of 154.3° is needed, indicated by the blue line starting from the 0/7 point. Each turn will have (7÷3=) 2.3333 of the connecting lines. The height of the first turn is reached after two lines connecting equivalent points plus one-third of the next line. The second 360° turn of the helix includes the remaining black two-thirds of the O2—O3 virtual bond, all of the O3—O4 virtual bond, and two-thirds of the O4—O5 virtual bond, indicated by the black lines. That is followed by the red set of 2.3333 lines that bring the third sequence to the structure repeating point of the three-turn helix (7/7).

A few more observations complete these explanations. If a structure has only two residues per turn, the distance between equivalent points is the same regardless of whether a clockwise or counter-clockwise rotation is used to reach the second point 180° away. A slight twist in either direction will create a helix that has a shorter distance to the next equivalent point in either the clockwise or the counter-clockwise direction. If the distance is shorter, then there will be more residues per turn. That is why there are no helices with descriptions of fewer than two residues per turn. Also, it is important to remember that the helical description applies to any given point, such as the C6 atom, and its equivalent points in other members of the helix. For this reason, a given rotation would apply to all the atoms, and the chirality of the glucose residue itself would be maintained, even though the sense, or chirality of the helix changes from right- to left-handed. These distinct helical conformations, including the seven-fold helix in the present structure and the five-fold helix in the sulfuric acid complex are among the different reasons why it is not useful to describe cellulose as having a cellobiose repeating unit (French 2017).

We built a model of the 7_3 and 7_4 helical cellulose chains using the corner chain glucose residue of cellulose I_β (Fig. 5a) (Nishiyama et al. 2002). The two glycosidic

torsion angles, φ (O5-C1-O1-C4) and ψ (C1-O1-C4-C5), of the 7_4 helix are -80° and -132°, respectively, notably close to φ = -89° (center), -99° (corner) and ψ = -147°, -142° in cellulose I_β. The 7_3 helix has a larger discrepancy in these torsion angles from the two-fold helix, with φ = -122° and ψ = -134°. Thus, the 7_4 helix only slightly twists away from the two-fold helix. This conformation is energetically stable due to its similarity to the 2_1 helical conformation, as shown in the conformational energy map based on the B3LYP density functional theory calculations of cellobiose (Fig. 4b) (French et al. 2012, 2021). Its φ/ψ point (red cross in Fig. 5b) is within the 1 kcal/mol contour relative to the global minimum. This contour circumscribes the known neat (un-complexed) cellulose allomorph conformations (black crosses). On the other hand, the 7_3 helix corresponds to an energy of about 2.7 kcal/mol. Unlike the chosen 7_4 model, which lies roughly in the center of the observed cellodextrin conformations from the Protein Data Bank (green squares), the 7_3 helix model is in a very sparsely populated area (French et al. 2012). Compared to the five-fold helix found in the sulfuric acid complex, the 7_4 helix is more energetically stable. The most plausible helical conformation in the sulfuric acid complex is a 5_3 helix with φ = -72° and ψ = -132°, giving an energy of ca. 1.8 kcal/mol (black star in Fig 4b).



317

318 **Fig. 5** (a) Cellulose chain in 7_4 helical conformation. (b) Conformational energy map
319 of cellobiose based on SMD continuum solvation and B3LYP/6-31+G(d) energies. Red
320 and orange crosses correspond to the current structures with 7_4 and 7_3 helical
321 conformations. The black star corresponds to the 5_3 helical conformation in the sulfuric
322 acid complex. The black cross and green squares are for conformations of cellulose
323 allomorphs and for di- and oligosaccharide crystals from the Protein Data Base with
324 cellobiose linkage, respectively. The dotted diagonal line represents two-fold screw
325 conformations. (c) Molecular packing model of the cellulose-phosphoric acid complex.

326

Based on the above-mentioned consideration of the conformational energy and the solid-state NMR results, the 7_4 helix is the most likely conformation formed in the phosphoric acid complex. We thus built a molecular packing model of the cellulose-phosphoric acid complex composed of 7_4 cellulose helices, as shown in Fig. 5c. In this model, cellulose chains are at the corners of the unit cell, leaving room for phosphoric acid molecules at the center of the cell. Four phosphoric acid molecules stack along the fiber axis in this central space. Considering the drastic change in the C6 signal in the solid-state NMR spectra, the hydroxymethyl group should directly interact with the phosphoric acid molecule. On the other hand, the complex may retain the intramolecular O3-H…O5 hydrogen bond, the common feature in the neat cellulose allomorphs, considering its conformational similarity with the 2_1 helix. The distance between O3 and O5 in the 7_4 helix is 2.84 Å, only 0.1 Å longer than that in cellulose I_β (Nishiyama et al., 2002).

The complexation with phosphoric acid caused unexpected changes in the molecular conformation of cellulose. This complex is the first crystal structure consisting of β-1,4 glycan chains in a seven-fold helical conformation. Together with the five-fold helical conformation observed in the cellulose-sulfuric acid complex, the seven-fold helix in this complex indicates the strong influence of the cellulose-acid interactions on the packing and conformation of cellulose molecules. The long pitch per glucose residue is another characteristic of this complex structure. In most cellulose crystallosolvates, this value decreases compared to cellulose I, presumably due to the solvation-induced structural relaxation. The stretched molecular conformation may result from the spatial coordination of cellulose and phosphoric acid molecules. A further crystallographic analysis based on an improved diffraction data set is necessary to elucidate such finer structural details.

As aforementioned, the diffraction pattern of the phosphoric acid complex was poorly resolved, and the radial width of reflections remained approximately the same as the original cellulose I_β. These observations are in striking contrast with the cellulose-sulfuric acid complex, whose radial width of reflections were significantly narrower than the original cellulose I_β. We attributed this apparent increase of crystallite size in the sulfuric acid complex to the increase in the interchain distance of cellulose due to the incorporation of sulfuric acid molecules in the crystalline lattice (Li et al. 2022). Although its unit cell is significantly smaller than the sulfuric acid complex,

the interchain distance also increased in the phosphoric acid complex compared to cellulose I_β. This lack of increase in the crystallite size implies the partial amorphization or disorganization of the cellulose crystal domain during the swelling process in phosphoric acid. While the preparation conditions of the complexes are nearly identical for these two acid species, the swelling and complexation behaviors are distinct between sulfuric and phosphoric acids.

Conclusion

We reported the formation of the cellulose-phosphoric acid crystalline complex based on X-ray diffraction measurement and solid-state NMR spectroscopy at -40 °C. This complex is the second known cellulose-acid complex and the first crystal structure consisting of cellulose molecules in a seven-fold helical conformation. The occurrence of the seven-fold conformation in this complex structure highlights the conformational flexibility of cellulose in its solvated states. The range of structures includes the Na-cellulose II structure with n=-3, the sulfuric acid complex with a 5_3 helical conformation (n=-2.5), the present phosphoric acid structure with a 7_4 helical conformation (n=-2.333), and cellulose itself with n=2. We are currently conducting a further investigation of the complexation behaviors of cellulose and mineral acids and a more detailed crystallographic analysis of these complexes. Such efforts will improve our understanding of the molecular interactions between cellulose and acids, which is crucial for the optimal processing of cellulosic biomass.

Competing interests

The authors have no relevant financial or non-financial interests to disclose.

Consent for publication

All authors have provided their consent for publication.

Availability of data

All data generated or analyzed during this study are included in this article.

Authors' contributions

I.M. and Y.O. conceptualized the study. J.H.L., I.M., I.J., Y.N., and Y.O. performed the experiments. J.H.L., Y.N., A.D.F., and Y.O. analyzed the data. J.H.L., A.D.F., and Y.O. wrote the manuscript. All authors read and approved the final manuscript.

Acknowledgment and funding

We acknowledge Agence Nationale de la Recherche (ANR grant number: ANR-21-CE29-0016-1 and ANR-11-EQPX-0010 under the "Investissements d'Avenir" program) and Glyco@Alps (ANR-15-IDEX-02) for the financial support and the NanoBio-ICMG platform (FR 2607) for granting access to the NMR facility. ESRF is acknowledged for the provision of beamtime (experiment number A02-1-902, D2AM beamline). We acknowledge Benjamin Richter and Yves Watier at the Sample Environment Service of ESRF, who designed the temperature-controlled specimen holder.

Ethics approval and consent to participate

Not applicable

References

Billès E, Onwukamike KN, Coma V, et al (2016) Cellulose oligomers production and separation for the synthesis of new fully bio-based amphiphilic compounds. Carbohydr Polym 154:121–128. https://doi.org/10.1016/j.carbpol.2016.07.107

Boerstoel H, Maatman H, Westerink JB, Koenders BM (2001) Liquid crystalline solutions of cellulose in phosphoric acid. Polymer 42:7371–7379. https://doi.org/10.1016/S0032-3861(01)00210-5

Camarero Espinosa S, Kuhnt T, Foster EJ, Weder C (2013) Isolation of Thermally Stable Cellulose Nanocrystals by Phosphoric Acid Hydrolysis. Biomacromolecules 14:1223–1230. https://doi.org/10.1021/bm400219u

Endo T, Yoshida S, Kimura Y (2020) Self-Assembly and Complexation of Cellulose/Ionic Liquid at High Cellulose Concentration: Anion Dependence. Cryst Growth Des 20:6267–6271. https://doi.org/10.1021/acs.cgd.0c00346

Falcoz-Vigne L, Ogawa Y, Molina-Boisseau S, et al (2017) Quantification of a tightly adsorbed monolayer of xylan on cellulose surface. Cellulose 24:3725–3739. https://doi.org/10.1007/s10570-017-1401-z

French AD, Johnson GP (2004) What crystals of small analogs are trying to tell us about cellulose structure. Cellulose 11:5-22. https://doi.org/10.1023/B:CELL.0000014765.94239.fe

French AD, Johnson GP (2009) Cellulose and the twofold screw axis: modeling and experimental arguments. Cellulose 16:959-973. https:// DOI 10.1007/s10570-009-9347-4

French AD (2017) Glucose, not cellobiose, is the repeating unit of cellulose and why that is

430 important. Cellulose 24:4605–4609 https://doi.org/10.1007/s10570-017-1450-3
431 French AD, Johnson GP, Cramer CJ, Csonka GI (2012) Conformational analysis of cellobiose
432 by electronic structure theories. Carbohydr Res 350:68–76.
433 https://doi.org/10.1016/j.carres.2011.12.023
434 French AD, Montgomery DW, Prevost NT, Edwards JV, Woods RW (2021) Comparison of
435 cellooligosaccharide conformations in complexes with proteins with energy maps for
436 cellobiose. Carbohydr Polym 264:118004
437 Hall M, Bansal P, Lee JH, et al (2010) Cellulose crystallinity – a key predictor of the enzymatic
438 hydrolysis rate. FEBS J 277:1571–1582. https://doi.org/10.1111/j.1742-
439 4658.2010.07585.x
440 Horii F, Hirai A, Kitamaru R (1983) Solid-state 13C-NMR study of conformations of
441 oligosaccharides and cellulose. Polym Bull 10:357–361.
442 https://doi.org/10.1007/BF00281948
443 Isogai A, Usuda M (1991) Preparation of low-molecular weight celluloses using phosphoric
444 acid. J Jpn Wood Res Soc Jpn
445 Lee DM, Blackwell J, Litt MH (1983) Structure of a cellulose II-hydrazine complex.
446 Biopolymers 22:1383–1399. https://doi.org/10.1002/bip.360220510
447 Li W, Ogawa Y, Perez J, et al (2022) Fivefold Helical Cellulose Trapped in a Sulfuric Acid
448 Framework. Cryst Growth Des 22:20–25. https://doi.org/10.1021/acs.cgd.1c00664
449 Moffat J, Morris VJ, Al-Assaf S, Gunning AP (2016) Visualisation of xanthan conformation by
450 atomic force microscopy. Carbohydr Polym 148:380–389.
451 https://doi.org/10.1016/j.carbpol.2016.04.078
452 Morris ER (2019) Ordered conformation of xanthan in solutions and “weak gels”:
453 Single helix, double helix – or both? Food Hydrocoll 86:18–25.
454 https://doi.org/10.1016/j.foodhyd.2017.11.036
455 Nishiyama Y, Mazeau K, Morin M, Cardoso MB, Chanzy H, Putaux J-L (2010) Molecular and
456 crystal structure of 7-fold V-amylose complexed with 2-propanol. Macromolecules,
457 43(20):8628-8636
458 Nishiyama Y, Asaadi S, Ahvenainen P, Sixta H (2019) Water-induced crystallization and nano-
459 scale spinodal decomposition of cellulose in NMMO and ionic liquid dope. Cellulose
460 26:281–289. https://doi.org/10.1007/s10570-018-2148-x
461 Nishiyama Y, Langan P, Chanzy H (2002) Crystal Structure and Hydrogen-Bonding System in
462 Cellulose Iβ from Synchrotron X-ray and Neutron Fiber Diffraction. J Am Chem Soc
463 124:9074–9082. https://doi.org/10.1021/ja0257319
464 Northolt MG, Boerstoel H, Maatman H, et al (2001) The structure and properties of cellulose
465 fibres spun from an anisotropic phosphoric acid solution. Polymer 42:8249–8264.
466 https://doi.org/10.1016/S0032-3861(01)00211-7
467 Okano T, Sarko A (1984) Mercerization of cellulose. I. X-ray diffraction evidence for
468 intermediate structures. J Appl Polym Sci 29:4175–4182.
469 https://doi.org/10.1002/app.1984.070291247
470 Okano T, Sarko A (1985) Mercerization of cellulose. II. Alkali–cellulose intermediates and a
471 possible mercerization mechanism. J Appl Polym Sci 30:325–332.
472 https://doi.org/10.1002/app.1985.070300128
473 Porro F, Bédué O, Chanzy H, Heux L (2007) Solid-State 13C NMR Study of Na−Cellulose
474 Complexes. Biomacromolecules 8:2586–2593. https://doi.org/10.1021/bm0702657
475 Schrödinger L, DeLano W (2020) PyMOL. http://www.pymol.org/pymol
476 Smith JP, Brown WE, Lehr JR (1955) Structure of Crystalline Phosphoric Acid. J Am Chem Soc

77:2728–2730. https://doi.org/10.1021/ja01615a013
Sobue H, Kiessig H, Hess K (1939) Das System Cellulose–Natriumhydroxyd–Wasser in Abhängigkeit von der Temperatur. Z Für Phys Chem 43B:309–328. https://doi.org/10.1515/zpch-1939-4324
Song G, Yu J, Ding M, Zhang J (2018) A Novel Cellulose/Ionic Liquid Complex Crystal. Cryst Growth Des 18:4260–4264. https://doi.org/10.1021/acs.cgd.8b00754
Veregin RP, Fyfe CA, Marchessault RH, Taylor MG (1987) Correlation of 13C chemical shifts with torsional angles from high-resolution, 13C-C.P.-M.A.S. N.M.R. studies of crystalline cyclomalto-oligosaccharide complexes, and their relation to the structures of the starch polymorphs. Carbohydr Res 160:41–56. https://doi.org/10.1016/0008-6215(87)80302-6
Wada M, Heux L, Nishiyama Y, Langan P (2009) The structure of the complex of cellulose I with ethylenediamine by X-ray crystallography and cross-polarization/magic angle spinning 13C nuclear magnetic resonance. Cellulose 16:943–957. https://doi.org/10.1007/s10570-009-9338-5
Zhang Y-HP, Cui J, Lynd LR, Kuang LR (2006) A Transition from Cellulose Swelling to Cellulose Dissolution by o-Phosphoric Acid: Evidence from Enzymatic Hydrolysis and Supramolecular Structure. Biomacromolecules 7:644–648. https://doi.org/10.1021/bm050799c