doi.org/10.26434/chemrxiv.11560173.v1

Bottom-up Chemoenzymatic Synthesis Towards Novel Fluorinated Cellulose-like Materials

Peterson de Andrade, Juan Munoz, Giulia Pergolizzi, Valeria Gabrielli, Sergey Nepogodiev, Dinu Iuga, Laszlo Fábián, Rinat Nigmatullin, Marcus A. Johns, Robert Harniman, Stephen J. Eichhorn, Jesús Angulo, Yaroslav Khimyak, Robert Field

Submitted date: 09/01/2020 • Posted date: 13/01/2020

Citation information: de Andrade, Peterson; Munoz, Juan; Pergolizzi, Giulia; Gabrielli, Valeria; Nepogodiev, Sergey; Iuga, Dinu; et al. (2020): Bottom-up Chemoenzymatic Synthesis Towards Novel Fluorinated Cellulose-like Materials. ChemRxiv. Preprint. https://doi.org/10.26434/chemrxiv.11560173.v1

Understanding the fine details of self-assembly of building blocks into complex hierarchical structures represents a major challenge en route to the design and preparation of soft matter materials with specific properties. Enzymatically-synthesised cellodextrins are known to have limited water solubility beyond DP9, a point at which they self-assemble into particles resembling the anti-parallel cellulose II crystalline packing. We have prepared and characterized a series of site-selectively fluorinated cellodextrins of different degrees of fluorination and substitution patterns by chemoenzymatic synthesis. The structural characterization of these materials at different length scales, combining advanced NMR and microscopy methods, showed that multiply 6-fluorinated cellodextrin chains assembled into particles presenting morphological and crystallinity features that are unprecedented for cellulose-like materials. In contrast, the introduction of a single fluorine atom per cellodextrin chain had a minor impact on materials structure. Our work emphasizes the strength of combining chemoenzymatic synthesis, fluorinated building blocks and advanced NMR and microscopy methods for the thorough characterization of hierarchical structures, leading to the controlled design of new biomaterials with specific properties.

Bottom-up chemoenzymatic synthesis towards novel fluorinated cellulose-like materials

Peterson de Andrade[1‡], Juan C. Muñoz-García[2‡], Giulia Pergolizzi[1‡], Valeria Gabrielli[2], Sergey A. Nepogodiev[1], Dinu Iuga[3], László Fábián,[2] Rinat Nigmatullin[4], Marcus A. Johns[4], Robert Harniman[5], Stephen J. Eichhorn[4], Jesús Angulo[2], Yaroslav Z. Khimyak[*2], Robert A. Field[*1,6]

[1]*Department of Biological Chemistry, John Innes Centre, Norwich NR4 7UH, UK*
[2]*School of Pharmacy, University of East Anglia, Norwich Research Park, Norwich NR4 7TJ, UK*
[3]*Department of Physics, University of Warwick, Coventry CV4 7AL, UK*
[4]*Bristol Composites Institute, CAME School of Engineering, University of Bristol, Bristol BS8 1TR, UK*
[5]*School of Chemistry, University of Bristol, Bristol BS8 1TS, UK*
[6]*Present address: Department of Chemistry and Manchester Institute of Biotechnology, University of Manchester, Manchester M1 7DN, UK*

‡ Equal contribution

[*]Corresponding authors e-mail address: rob.field@jic.ac.uk; robert.field@manchester.ac.uk; y.khimyak@uea.ac.uk

Abstract

Understanding the fine details of self-assembly of building blocks into complex hierarchical structures represents a major challenge *en route* to the design and preparation of soft matter materials with specific properties. Enzymatically-synthesised cellodextrins are known to have limited water solubility beyond DP9, a point at which they self-assemble into particles resembling the anti-parallel cellulose II crystalline packing. We have prepared and characterized a series of site-selectively fluorinated cellodextrins of different degrees of fluorination and substitution patterns by chemoenzymatic synthesis. The structural characterization of these materials at different length scales, combining advanced NMR and microscopy methods, showed that multiply 6-fluorinated cellodextrin chains assembled into particles presenting morphological and crystallinity features that are unprecedented for cellulose-like materials. In contrast, the introduction of a single fluorine atom per cellodextrin chain had a minor impact on materials structure. Our work emphasizes the strength of combining chemoenzymatic synthesis, fluorinated building blocks and advanced NMR and microscopy methods for the thorough characterization of hierarchical structures, leading to the controlled design of new biomaterials with specific properties.

Introduction

Cellulose is an abundant natural biopolymer used extensively in industry as raw material for the production of paper, textile, food thickeners, explosives, etc.[1,2] The contemporary use of cellulose is based on nanosized cellulose particles (nanocellulose), which is a promising class of renewable material due to its intrinsic characteristics and potential for broad advanced industrial applications.[3–5] Hence, the design of nanocellulose-based soft-matter materials relies on the formation of assembly-driven structures, and may prompt to changes in mechanical properties or bring additional functionality.[6–10]

Production of cellulose nanocrystals and nanofibrillated cellulose, the main classes of nanocellulose, are both top-down methodologies, based on the isolation of nanocellulose from cellulosic biomass, which requires high energy consumption.[5] Moreover, functionalization of nanocellulose to meet requirements for specific applications often requires harsh chemical conditions (*i.e.* strong acids and bases). As an alternative, significant progress has been made in the synthesis of nanocellulosic materials based on a bottom-up approach *via* enzymatic polymerization.[11–13] This approach is highly flexible and is projected to enable the generation of tailor-made functionalized nanocellulosic materials, by opening up the possibility to direct the nanoarchitecture and properties of cellulose oligomers.

Developing strategies to functionalize biocompatible materials giving rise to tuneable properties is becoming increasingly important to their integration into biomedical products.[14,15,16] In this scenario, the site-selective introduction of probes into cellulose to report on local structure and to modulate material properties is of great importance. Fluorine is well-known for its unique physicochemical properties,[17] which find widespread use in drugs[18] and agrochemicals[19] due to its remarkable impact on conformation, permeability, potency, metabolism, etc.[20,21] In addition, the absence of fluorine in biological systems makes the introduction of ^{19}F nuclei a powerful reporter of local structure and environment in complex systems.[22–29] Nonetheless, its use remains underexplored with respect to biomaterials.

Synthetic strategies for the regio- and stereoselective (deoxy)fluorination of monosaccharides have been successful.[30–33] In contrast, the corresponding top-down derivatization of cellulose[8,34] is complicated by solubility challenges, resulting in incomplete control of the sites and extent of fluorination.[35] The bottom-up chemical synthesis of structurally defined cellodextrins, including 3-fluorinated compounds,[32] has been achieved recently by automated glycan assembly.[36] Enzymatic synthesis presents an alternative option to achieve the bottom-up preparation of functionalized oligo- and poly-saccharides[37] in a regio- and stereo-controlled manner.[38,39] For instance, glycoside phosphorylases (GPs)[12,40–43] have recently received attention for the synthesis of amylose and cellulose-like materials. In particular, cellodextrin phosphorylase (CDP, EC 2.4.1.49) has emerged as a powerful tool for the synthesis of differently functionalized cellulose oligomers giving rise to a variety of nanostructures (sheets,[11,44,45] rods,[46] or ribbons[47]) depending on the nature of the substrate.[11,48–52]

Herein, we have prepared enzymatically produced cellodextrins (EpCs) with different fluorination patterns. Monofluorinated EpCs (2F-, 3F- and 6F-EpC) were obtained by CDP-mediated oligomerisation of α-D-glucose1-phosphate (Glc-1P) as donor and deoxy-fluoro-cellobioses as acceptor substrates. Multiply 6-fluorinated EpC (multi-6F-EpC) was prepared from 6-deoxy-6-fluoro-α-D-glucose1-phosphate (6F-Glc-1P) and cellobiose as donor and acceptor substrates, respectively.

The thorough structural characterization of these biomaterials has been performed by a combination electron microscopy, X-ray diffraction, Raman spectroscopy and advanced Nuclear Magnetic Resonance experiments. The presence of a single fluorine atom per cellodextrin chain did not exert a substantial impact on the morphology and crystalline structure of the material. In contrast, the presence of multiple 6-deoxy-6-fluoro-glucose units yielded an unprecedented crystalline allomorph never reported before for a cellulose-like material, thus constituting a new long-range ordering motif for it. Our findings demonstrate how enzymatic catalysis from functionalized monosaccharide building blocks enables the generation of site-specific fluorinated cellulose oligomers with unique structural features, which may open up way for new a class of cellulose-like materials, and highlight the value of advanced NMR methods in detailing the molecular level characterization.

Results and discussion

Enzymatic synthesis of fluorinated cellodextrins

*Synthesis of 2-, 3- and 6-monofluorinated cellodextrins (2F-EpC, **4**; 3F-EpC, **5**; and 6F-EpC, **6**)*

CDP uses glucose as an acceptor substrate only poorly compared to cellobiose and longer cello-oligosaccharides,[11,52] to produce cellodextrins containing deoxy-fluoro-glucose at the reducing terminus (compounds **4**-**6**). We therefore initially used cellobiose phosphorylase (CBP) (PRO-GH94-004) to synthesize monofluorinated cellobiose analogues (**1**-**3**) (Figure 1A) for use as acceptors for CDP (Figure 1B). CBP was incubated at 37 °C with 1:1 ratio of Glc-1P and deoxy-fluoro-glucose (2F-, 3F- or 6F-Glc) for 16 h, at which point TLC showed *ca.* 80% conversion into the disaccharides **1** and **3**, and *ca.* 60% into **2**. The different conversion efficiencies may be rationalized based on a study of *Cellulomoinas uda* cellobiose phosphorylase,[53] in which k_{cat}/K_m values for 2F-Glc (2.4%), 3F-Glc (0.013%) and 6F-Glc (31%) acceptors are substantially lower than that of the parent Glc substrate, but all three compounds are indeed productive substrates. CBP was removed from reaction mixture by affinity chromatography (His6-tag nickel column purification) and the desired products were purified by gel filtration chromatography. The purification successfully removed residual deoxy-fluoro-glucose acceptors, but small amounts of cellobiose required removal by HPLC to obtain compounds **1**-**3** in high purity for characterization (†ESI Figures S1-S3). However, the monofluorinated cellodextrins **4**-**6** could be obtained in one-pot reactions from the respective sugar-1P and fluorinated glucoses.

Once Glc-1P consumption was almost complete in the CBP reactions, more Glc-1P (4 eq.) was added together with CDP and the reactions were incubated at 37 °C shaking for 16 h. A white precipitate was formed and isolated by centrifugation, followed by resuspension and washing with MQ water to remove enzyme, salts and any soluble sugar. Further 4 eq. of Glc-1P were added to the supernatant, and CDP reaction was further incubated to produce more fluorinated EpCs. In this manner, monofluorinated cellodextrins were obtained with reasonable overall yield based on consumed fluoro-glucose [47% (Glc-β-1,4)$_n$-2F-Glc (**4**), 30% (Glc-β-1,4)$_n$-3F-Glc (**5**), 32% (Glc-β-1,4)$_n$-6F-Glc (**6**)]. MALDI-TOF mass spectrometry analysis showed these materials to have an average DP *ca*. 9, while the unsubstituted cellodextrin (EpC, **8**) produced under the same reaction conditions averaged *ca.* DP 8 (†ESI Figure S7). Traces of longer monofluorinated cellodextrins were also observed . Solution-state ^{19}F NMR analysis in 1 M NaOD (†ESI Figure S4) showed two singlets for each material,

reflecting reducing terminal anomers, with peaks at -195.21 and -195.26 ppm (2F-EpC, **4**), -190.86 and -197.19 ppm (3F-EpC, **5**) and -232.55 and -234.05 ppm (6F-EpC, **6**).

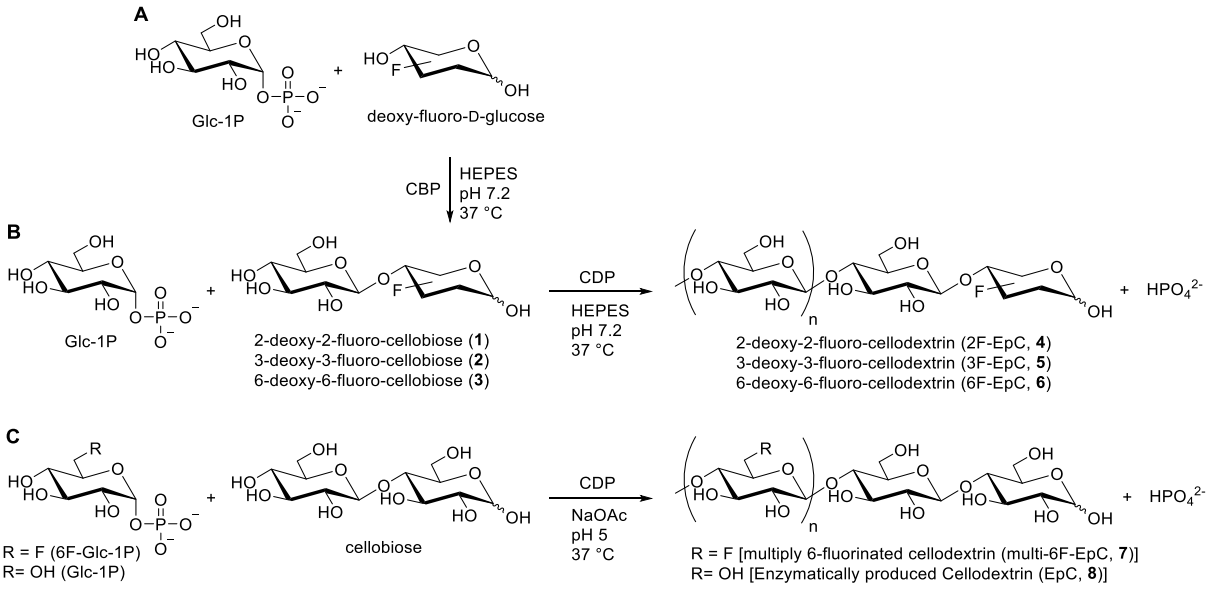


Figure 1. **Enzymatic synthesis of fluorinated cellodextrins**. (**A**) Cellobiose phosphorylase (CBP) catalysed reaction of α-D-glucose 1-phosphate (Glc-1P) and deoxy-fluoro-D-glucose (2F-Glc, 3F-Glc or 6F-Glc), followed by (**B**) cellodextrin phosphorylase (CDP) catalyzed oligomerization with Glc-1P and monofluorinated cellobioses, to afford enzymatically-produced fluorinated cellodextrins (2F-EpC, **4**; 3F-EpC, **5;** and 6F-EpC, **6**). (**C**) CDP-catalysed reaction of 6-deoxy-6-fluoro-α-D-glucose 1-phosphate (6F-Glc-1P) or Glc-1P and cellobiose as acceptor, to produce multiply 6-fluorinated cellodextrin (multi-6F-EpC, **7**) or the parent Enzymatically produced Cellodextrin (EpC, **8**), respectively.

Synthesis of multiply 6-fluorinated cellodextrin (multi-6F-EpC, 7)

We also investigated CDP-mediated oligomerization using the chemically modified glucosyl donor 6F-Glc-1P (Figures S5-S6)[33] and cellobiose as acceptor (Figure 1C) to achieve higher structural impact by placing multiple fluorine atoms along the cellodextrin (multiply 6-fluorinated cellodextrin, multi-6F-EpC, **7**). The initial tests using 6F-Glc as an acceptor to obtain a fully 6F-substitued cellodextrin proved very slow and inefficient. Alternatively, CBP was tested with 6F-Glc as an acceptor to generate a difluorinated cellobiose, which could be a better substrate for CDP. However, only trace amounts of the product were detected; prompting us to choose the natural acceptor cellobiose. CDP was incubated at 37 °C with a 6:1 ratio of 6F-Glc-1P and cellobiose for 72 h. The resulting white precipitate was isolated by centrifugation, followed by re-suspension and washing with MilliQ water to give **7** with 64% yield. ^{19}F solution-state NMR analysis of **7** dissolved in 1 M NaOD (†ESI Figure S4) showed one major singlet at -233.25 ppm, which may correspond to fluorine from the 6F-Glc internal repeating units, and three smaller singlets at -233.29, -233.31 and -233.35 ppm from 6F-Glc close to the reducing terminal and the non-reducing terminal unit. Analysis by MALDI-TOF mass spectrometry revealed that multi-6F-EpC **7** had a higher average DP (*ca.* 10) than the parent EpC (*ca.* DP 8) and that longer chains, up to DP 15, could also be observed in the multiple 6-fluorinated material (†ESI Figure S7).

Morphological characterization

Electron Microscopy (EM) and Atomic Force Microscopy (AFM)

Transmission electron microscopy (TEM) was initially used to observe the morphological differences between EpC and fluorinated EpCs, which were prepared for analysis only by dilution of concentrated suspensions obtained after purification of precipitates formed during enzymatic synthesis. As expected, the TEM images of the monofluorinated 2F-EpC (**4**), 3F-EpC (**5**) and 6F-EpC (**6**) (†ESI Figure S8) show a very similar morphology to EpC (**8**, Figure 2A, a). This crystalline sheet-like morphology is well-known for enzymatically synthesised cello-oligosaccharides,[11,49] including derivatised cellulose, such as acrylated cellulose[45] and cellulose conjugated with oligo(ethylene glycol).[47] On the other hand, multi-6F-EpC (**7**) particles formed predominantly into significantly shorter platelets (< 100 nm length) (Figure 2B, a). These differences were further confirmed by AFM imaging using samples prepared by depositing diluted sample suspensions on freshly cleaved mica (Figure 2A, b and c). Although a few long platelets are present in multi-6F-EpC (**7**), their fraction is smaller than in EpC (**8**). As reported in the literature,[11,49] the thickness of EpC (**8**) platelets was found to be ~5 nm. Similar thickness was observed for long platelets of multi-6F-EpC (**7**, Figure 2B, c).

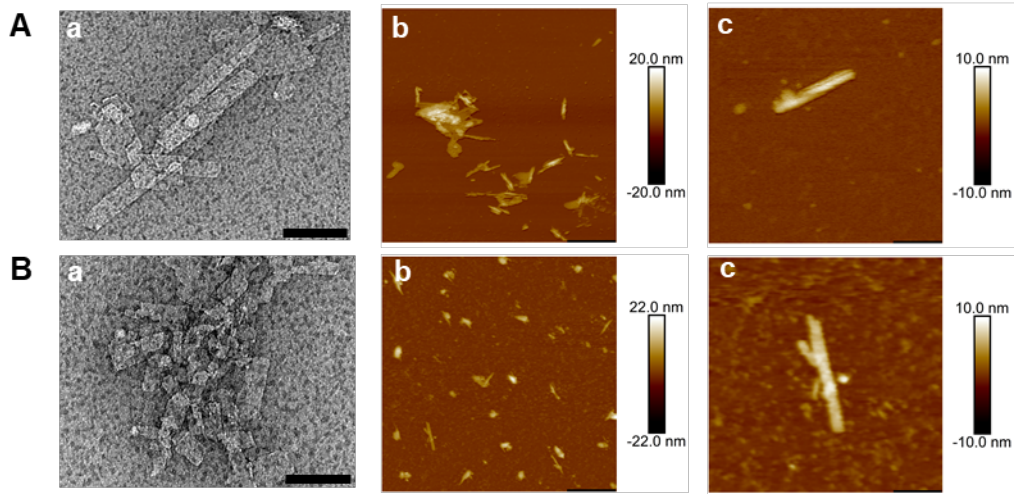


Figure 2. TEM (**a**) and AFM (**b** -2 x 2 μm and **c** - 0.5 x 0.5 μm) images of enzymatically produced cellodextrin EpC (**8**, **A**) and multi-6F-EpC (**7**, **B**). Scale bars correspond to 100 nm. Scale bars on the right of **b** and **c** are height measurements.

Long-range structural characterization by powder X-ray diffraction (PXRD)

The PXRD patterns of the monofluorinated 2F-EpC (**4**), 3F-EpC (**5**) and 6F-EpC (**6**) are virtually indistinguishable from the diffraction pattern of EpC (**8**) (Figure 3). This result indicates that the monofunctionalized cellodextrin-like molecules arrange as cellulose type II allomorph, with three intense and sharp peaks at $2\theta = 12°$, 20° and 23° (d-spacing of 0.74, 0.44 and 0.39 nm, respectively) representing $(1\bar{1}0)$, (110) and (020) planes.[54,55] On the other hand, the experimental PXRD pattern reported for multi-6F-EpC (**7**) does not correspond to any allomorph previously described for cellulose (Figure 3). The pattern shows two well defined peaks at $2\theta = 15°$ and 23° (d-spacing of 0.59 and 0.39 nm, respectively), as well as four different broad components at $2\theta = 21°$, 25°, 30° and 36° (d-spacing of 0.42, 0.36, 0.30 and 0.25 nm, respectively).

In order to verify possible similarities with previously reported cellulose structural organisations, we predicted and compared the PXRD spectra of multi-6F-EpC (**7**) to each known allomorph (†ESI Figure S9 and Table S1). Remarkably, the observed peak positions of multi-6F-EpC (**7**) are unique when compared to the diffraction patterns of the known allomorphs (Figure 3, †ESI Figure S9 and Table S1), hence demonstrating the formation of a new crystalline structure for this new cellulose-like material.

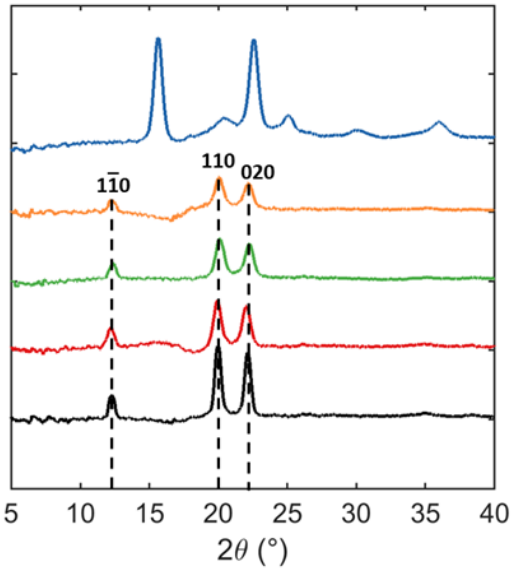


Figure 3. Powder X-ray diffraction patterns of EpC (**8**, black), 2F-EpC (**4**, red), 3F-EpC (**5**, green), 6F-EpC (**6**, orange) and multi-6F-EpC (**7**, blue).

Molecular level characterization

Raman spectroscopy

Figure 4 shows typical Raman spectra of 2F-EpC (**4**), 3F-EpC (**5**), 6F-EpC (**6**), multi-6F-EpC (**7**) and EpC (**8**). The bands located at *ca.* 1462 (HOC and HCH stretching), 1265 (HCC and HCO stretching) and 576 cm^{-1} (heavy atom stretching) and the dominance of the band located at *ca*. 354 cm^{-1} over the band located at *ca*. 379 cm^{-1} (both heavy atom stretching) confirmed that EpC (**8**) arranges into a cellulose type II structure.[56] **4**, **5** and **6** are very similar to **8**, as expected for a single fluorine atom (at the reducing end) per oligosaccharide chain. The weak band located at 487 cm^{-1} for the monofluorinated EpCs is probably an amalgamation of the 480 and 496 cm^{-1} bands as a result of the single fluorine present in each chain.

In contrast, the multi-6F-EpC (**7**) spectrum is significantly different owing to the presence of multiple fluorine atoms. Multi-6-fluorination results in new Raman bands located at 480, 496 and 924 cm^{-1}. The first two bands correspond to the presence of CH_xF, whilst the third relates to CH_2F, confirming modification at the C6 position.[57] Multi-6-fluorination also results in the shift of multiple bands, including 1462 to 1451 cm^{-1} and 1265 to 1268 cm^{-1}, as well as the loss of others, such as the band located at 576 cm^{-1} (Table S2). Most significantly, the band associated with the glycosidic linkage located at 1097 cm^{-1} is shifted to 1088 cm^{-1}. This provides some evidence that the crystal structure of the multi-6F-EpC material is neither cellulose type I nor type II.

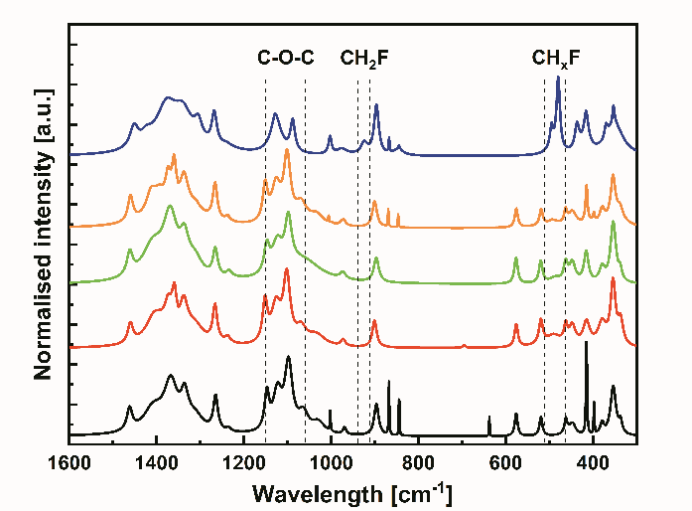


Figure 4. Normalised, deconvoluted Raman spectra for EpC (black), 2F-EpC (red), 3F-EpC (green), 6F-EpC (orange) and multi-6F-EpC (blue). Dashed lines correspond to boundaries of bands associated with C-O-C stretching (C-O-C) and presence of fluorinated carbon groups (CH_2F and CH_xF).

Solid-state Nuclear Magnetic Resonance (SSNMR)

Direct polarization ^{19}F NMR experiments (without ^{1}H decoupling) were carried out at 60 kHz MAS rate for the 2F-EpC (**4**), 3F-EpC (**5**), 6F-EpC (**6**) and multi-6F-EpC (**7**) (Figure 5a). A single very broad and asymmetric peak was observed for the monofluorinated materials, centred at -190 (2F-EpC, **4**), -197 (3F-EpC, **5**) and -232 ppm (6F-EpC, **6**), respectively, in good agreement with the solution ^{19}F NMR data (†ESI Figure S4). 3F-EpC (**5**) and 6F-EpC (**6**) showed broad peaks, with line widths at half height of 11.9 and 9.4 kHz, respectively (Figure 5c), while multi-6F-EpC (**7**) showed a sharper (3.8 kHz width at half height) Lorentzian-shaped peak (Figure 5c). ^{1}H-decoupled ^{19}F ($^{19}F\{^{1}H\}$) NMR spectra of 3F-EpC recorded at slower MAS rate showed an even broader ^{19}F peak (†ESI Figure S10), indicating that (i) fast MAS is more efficient at decoupling that radiofrequency decoupling (fast MAS decouples both ^{19}F-^{19}F homonuclear dipolar coupling as well as heteronuclear ^{1}H-^{19}F coupling), and (ii) the large ^{19}F line widths of 2F-EpC (**4**), 3F-EpC (**5**), 6F-EpC (**6**) and multi-6F-EpC (**7**) in the fast MAS spectra (Figure 5a) are mostly due to the large heterogeneity of ^{19}F chemical environments. This can be easily understood considering that these materials assemble into particles with a specific crystalline packing (cellulose type II[54] for 2F-EpC (**4**), 3F-EpC (**5**) and 6F-EpC (**6**), and a new organisation for muti-6F-EpC (**7**); Figures 5d and 6a), and the ^{19}F nucleus is extremely sensitivity to chemical environment. Upon assembly of nanocellulose the ^{19}F atoms of each cellulose chain can occupy any position within the nanofibril (surface, core, far from or nearby other fluorinated residues, etc), hence presenting non-equivalent environments within the packing of EpC (Figure 5c). Assuming that ^{19}F-^{1}H dipolar interactions are practically eliminated at fast MAS, the peak broadening reflects a multitude of orientations sampled by the C-F bonds.

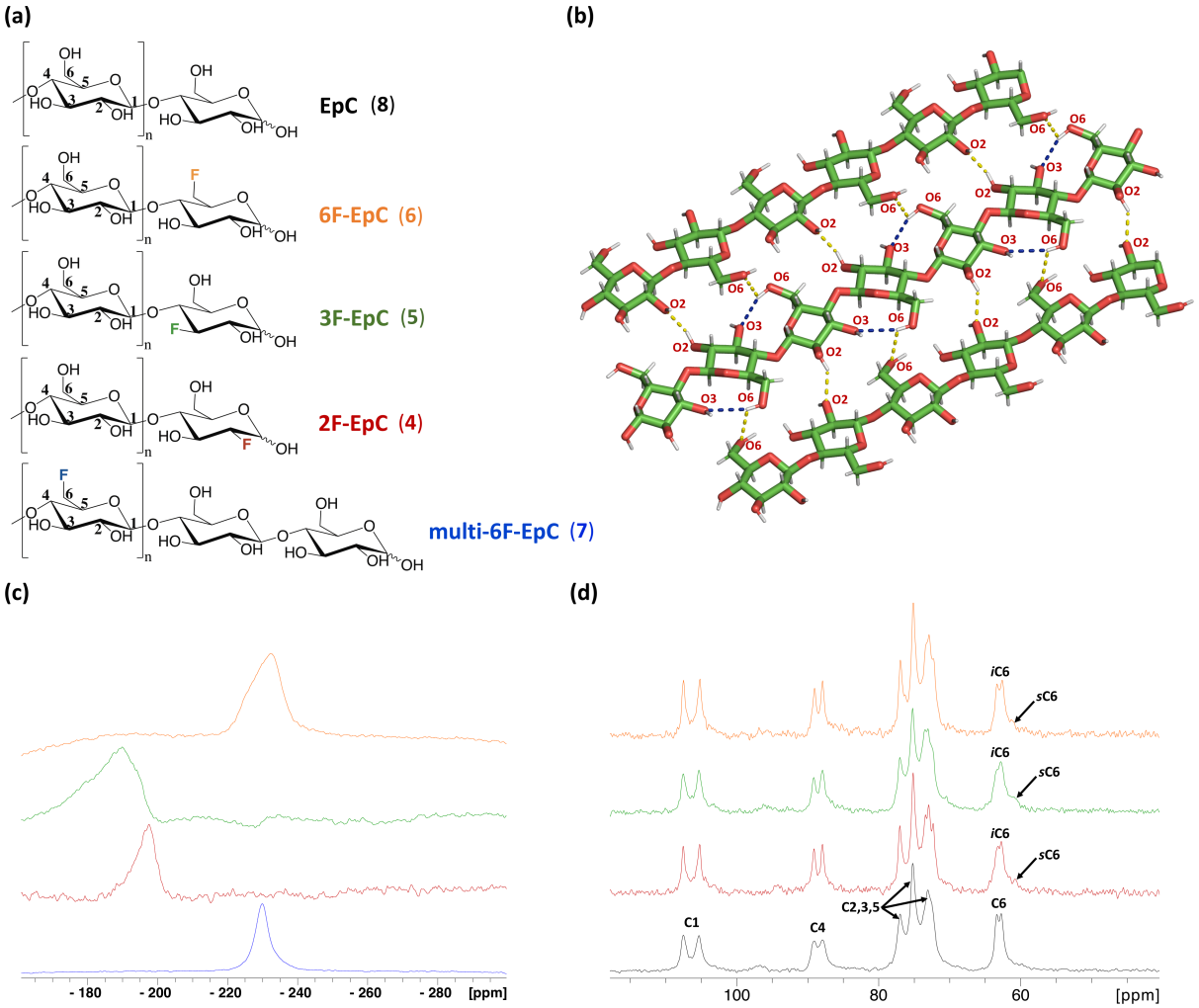


Figure 5. (a) Chemical structures of non-modified EpC (**8**), monofluorinated 2- (**4**), 3- (**5**) and 6F-EpC (**6**), and multi-6F-EpC (**7**). **(b)** 3D model of the crystalline packing of cellulose II allomorph based on the origin-center-origin (o-c-o) chains.[55] The O3-O6 intra-chain (blue dashes) and O2-O2 and O6-O6 inter-sheet (yellow dashes) hydrogen bonds are shown. It should be noted that the substitution of all -OH groups at C6 with fluorine atoms precludes the formation of O6-O6 inter-sheet hydrogen bonds during self-assembly. Note: the intra-chain hydrogen bonds are only shown for the center chain for simplicity.[55] **(c)** Direct detection ^{19}F MAS NMR spectra of multi-6F-EpC (**7**, blue) and 2F- (**4**, red), 3F- (**5**, green) and 6F-EpC (**6**, orange) powders, acquired at 60 kHz MAS rate and 800 MHz ^{19}F frequency (20 T magnetic field). **(d)** ^{1}H-^{13}C CP/MAS NMR spectra of EpC powder acquired at 10 kHz MAS rate (black), and 2F- (**4**, red), 3F- (**5**, green) and 6F-EpC (**6**, orange) 10 wt% dispersions acquired at 6 kHz MAS, and 100 MHz ^{13}C frequency.

To characterize the structural organization of 2F-EpC (**4**), 3F-EpC (**5**), 6F-EpC (**6**) and multi-6F-EpC (**7**) materials at the molecular level, ^{1}H-^{13}C CP/MAS experiments were carried out. Each type of cellulose allomorph presents a characteristic ^{13}C NMR fingerprint.[58,59] Monofluorinated EpCs (**4**, **5** and **6**) 10 wt% dispersions showed the characteristic cellulose II ^{1}H-^{13}C CP fingerprint, typical of non-modified EpC (**8**) (Figure 5d). The only noticeable difference was the presence of a broad peak at *ca.* 61 ppm, which is characteristic of a surface/disordered population of C6 (*s*C6, Figure 5d).[59,60] Hence, the peak at 63 ppm represents the interior/ordered domains of C6 (*i*C6, Figure 5d). The *s*C6 broad peak is typically observed in bacterial cellulose, which consists of cellulose particles containing both I_α and I_β crystalline domains and disordered regions.[60] Indeed, surface/disordered and interior/ordered domains are typically found in nanocrystalline cellulose (CNC)[59], bacterial cellulose (BC)[60] and plant cell

walls (PCWs)[61]. We note that the presence of a fluorine atom substituting the 3-hydroxyl group of glucose might affect the formation of the characteristic O3-O6 intra-chain hydrogen bond between adjacent glucose residues in cellulose II allomorph (Figure 5b). On the other hand, the formation of the O2-O2 and O6-O6 inter-sheet hydrogen bonds of cellulose II would be affected in 2F- and 6F-EpC, respectively (Figure 5b). Spectral deconvolution of the *s*C6 and *i*C6 peaks of ^{1}H-^{13}C CP/MAS NMR spectra enabled us to estimate the relative surface area (RSA) of the EpC particles constituting each material (†ESI Figure S11, Table S3). RSAs of about 16-23 % were obtained for the three monofluorinated EpCs, with 3F-EpC showing the highest value (23 %; †ESI Table S3). Similar values of surface area have been determined before for bacterial cellulose.[60]

The ^{1}H-^{13}C CP/MAS NMR spectrum of multi-6F-EpC (**7**) showed a pattern of peaks that does not correspond to either cellulose I, II or III allomorphs (†ESI Table S4, S5). This is evidenced by the appearance of C1 (105.9 ppm) as a singlet peak in **7**, which is a singlet in cellulose type I_{α} and III_{I}, and a doublet in cellulose type I_{β}, II and and III_{II} (†ESI Table S4).[62,63] Notably, multi-6F-EpC (**7**) presents some spectral features in common with cellulose III_{I}, i.e. C1 and non-fluorinated C6 peaks are singlets and appear at very similar chemical shifts (104.8 and 62.1 ppm, respectively; †ESI, Table S5). However, the PXRD pattern of multi-6F-EpC (**7**, Figure 3) does not correspond to either cellulose III_{I}[64] or III_{II}[63] allomorphs, or any other cellulose allomorphs reported so far (†ESI Table S1). Hence, **7** assembles into a crystalline organization which is unprecedented for a cellulose-type material. The formation of this novel structural motif is also supported by the new features observed in the Raman spectra, which do not correspond to either cellulose I or II (Figure 3, †ESI Table S4).

The combination of fast MAS ^{1}H-^{13}C CP, low MAS ^{1}H,^{19}F-decoupled ^{19}F-^{13}C CP, water-polarization transfer (WPT) solid-state NMR and ^{13}C, COSY and HSQC solution NMR experiments enabled the assignment of the ^{13}C spectrum of multi-6F-EpC (**7**, Figure 6) to be made. ^{1}H,^{19}F-decoupled ^{19}F-^{13}C CP enabled the assignment of C6, C4 and C5 peaks of the fluorinated residues (Figure 6a). The highest intensity peak was assigned to C6 (83.8 ppm), as it is the carbon atom closest to 6F (1.3 Å). The peak at 73.1 ppm corresponds to C4 and C5 sites, based on their proximity to fluorine (Figure 6a, †ESI Figure S12), while C2 and C3 are too far away to cross-polarize from fluorine effectively. ^{13}C DEPT135, COSY and HSQC solution NMR experiments confirmed this assignment (†ESI Figure S13), with the methylene carbons of the fluorinated (C6) and non-fluorinated (C6*) glucose units appearing in antiphase with respect to the CH carbons (Figure 6a). Importantly, the ^{13}C peaks at 81.9 and 73.1 ppm observed in the CP spectrum did not appear on the ^{13}C DEPT135 nor ^{1}H-^{13}C HSQC solution NMR experiments carried out for a diluted dispersion of multi-6F-EpC (**7**). Hence, these peaks most likely correspond to the immobile interior carbons (*i*C6 and *i*C2,3,4,5, respectively) that are too broad to be detectable by solution NMR. The solution-NMR-observed C6 and C2,3,4,5 peaks were therefore assigned to surface/disordered domains (*s*C6 and *s*C2,3,4,5, respectively). The assignment of *s*C6 and *i*C6 was further validated by water polarization transfer CP (WPT-CP) NMR experiments (Figure 6b).[65] The peak intensity in WPT-CP experiments depends on the distance and relative mobility of bound water at the particle surface and the number of interacting water molecules at a particular site. Hence, peaks corresponding to surface domains will show faster WPT growth at short mixing times than interior domains, as we have recently observed for BC.[60] At sufficiently long mixing times, WPT become homogeneous for both surface and interior domains due to the efficient spin diffusion. Figure 6b shows the WPT factors for a 25 wt% dispersion of multi-6F-EpC (**7**) at 16 ms mixing time (under our experimental conditions, homogenization of surface-interior water polarization transfer is achieved around 200 ms). A much higher WPT factor was observed for the *s*C6

(83.8 ppm) compared to the *i*C6 peak (81.8 ppm), confirming the assignment of *s*C6 and *i*C6 peaks to surface and interior domains, respectively (Figure 6b). Also, *s*C2,3,4,5 showed higher WPT compared to *i*C2,3,4,5 (Figure 6b), in agreement with solution NMR data where the *s*C2,3,4,5 and *i*C2,3,4,5 peaks are visible and invisible, respectively (Figure 6a). Spectral deconvolution of *s*C6 and *i*C6 peaks indicated that multi-6F-EpC (**7**) presents a RSA of ~54 % (†ESI Figure S11, Table S3), very similar to the RSA values that we reported for nanocrystalline cellulose.[66]

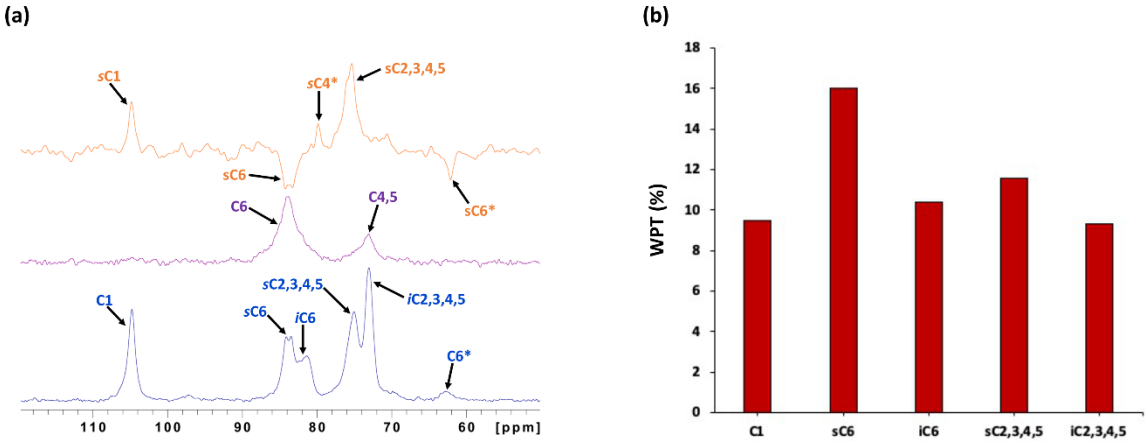


Figure 6. (a) ^{1}H-^{13}C CP (blue) and ^{1}H,^{19}F - decoupled ^{19}F-^{13}C CP NMR (purple) spectra of multi-6F-EpC (**7**) powder acquired at 60 and 15 kHz MAS spinning, respectively, and 212.5 MHz ^{13}C frequency. The ^{13}C DEPT135 spectrum of a 1 wt% dispersion of multi-6F-EpC (**7**) in D_2O is shown for comparison (orange). * Low intensity peaks corresponding to the non-fluorinated glucose units of **7** at the reducing terminal of each cellodextrin chain. **(b)** Bar graph showing the water polarization transfer (WPT) factors, in percentage, determined for each carbon peak of multi-6F-EpC 25 wt% hydrogel using a mixing time of 16 ms. The higher WPT factor observed for *s*C6 and *s*C2,3,4,5 compared to *i*C6 and *i*C2,3,4,5 demonstrates the increased solvation of the former, hence being assigned to surface domains.

Conclusions

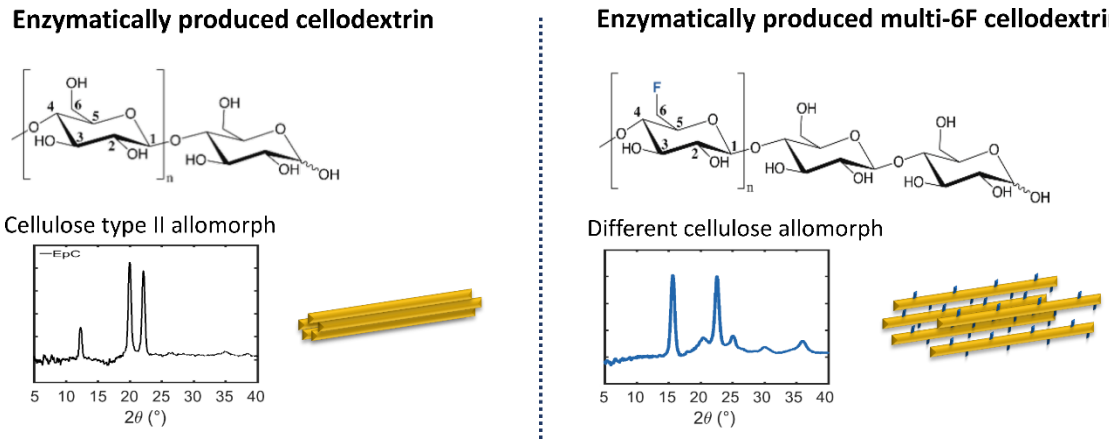
We have developed a bottom-up approach for the chemoenzymatic synthesis and characterization of novel mono and multiply 6-fluorinated cellodextrins from functionalized monosaccharide building blocks. Enzymatic catalysis has been employed for the rapid production of tailor-made fluorinated materials based on substrates with specific fluorination pattern. CBP and CDP proved to be robust enzymes for accepting fluorinated substrates (monofluorinated cellobioses and 6F-Glc-1P, respectively) in order to produce novel nanocellulosic materials. Comparison between the enzymatically produced cellodextrin (EpC) and its monofluorinated analogues (2F-, 3F- and 6F-EpC) clearly shows that the presence of a single fluorine atom in the reducing end glucose unit does not affect to a large extent the crystalline structure and morphology of these materials. In contrast, the presence of multiple 6-fluorinated glucose units along the oligosaccharide chain (multi-6F-EpC) yielded a novel cellodextrin material of different morphology and crystallinity. Importantly, multi-6F-EpC shows a unique long-range ordering motif unprecedented for cellulose-based biomaterials, indicated by diffraction and spectroscopic methods. Advanced solid-state NMR methods enabled the detailed characterization of these novel materials, deciphering the water-exposed and interior chemical environments for different carbon sites. Our findings highlight considerable potential of chemoenzymatic synthesis for generating novel glycomaterials of controlled morphology and molecular structure via easy-to-prepare building blocks of specific

substitution pattern. Our work opens promising avenues for the substitution of chemical methods for material preparation by environmentally friendly procedures.

Acknowledgements

We thank the GelEnz consortium, which is funded by EPSRC (Grant Research Number: IUK 59000 442149). Work at the John Innes Centre is supported by the UK BBSRC Institute Strategic Program on Molecules from Nature - Products and Pathways [BBS/E/J/000PR9790] and the John Innes Foundation; the BBSRC. BBSRC, EPSRC, and Innovate UK: IBCatalyst (Grant BB/M02903411). The UK 850 MHz solid-state NMR Facility used in this research was funded by EPSRC and BBSRC (contract reference PR140003), as well as the University of Warwick including via part funding through Birmingham Science City Advanced Materials Projects 1 and 2 supported by Advantage West Midlands (AWM) and the European Regional Development Fund (ERDF). The Engineering and Physical Sciences Research Council (EPSRC) is acknowledged for provision of financial support (EP/N033337/1) for J.C.M.-G., J.A. and Y.Z.K. We are also grateful for UEA Faculty of Science NMR facility. V.G. would like to acknowledge the support of BBSRC Norwich Research Park Bioscience Doctoral Training Grant (BB/M011216/1).

Graphical abstract



References

(1) Habibi, Y. Key Advances in the Chemical Modification of Nanocelluloses. *Chem. Soc. Rev.* **2014**, *43* (5), 1519–1542. https://doi.org/10.1039/c3cs60204d.

(2) French, A. D.; Kim, H. J. Cotton Fiber Structure. In *Cotton Fiber: Physics, Chemistry and Biology*; Fang, D. D., Ed.; Springer International Publishing: Cham, 2018; pp 13–39. https://doi.org/10.1007/978-3-030-00871-0_2.

(3) Eichhorn, S. J.; Dufresne, A.; Aranguren, M.; Marcovich, N. E.; Capadona, J. R.; Rowan, S. J.; Weder, C.; Thielemans, W.; Roman, M.; Renneckar, S.; et al. *Review: Current International Research into Cellulose Nanofibres and Nanocomposites*; 2010; Vol. 45. https://doi.org/10.1007/s10853-009-3874-0.

(4) Abitbol, T.; Rivkin, A.; Cao, Y.; Nevo, Y.; Abraham, E.; Ben-Shalom, T.; Lapidot, S.; Shoseyov, O. Nanocellulose, a Tiny Fiber with Huge Applications. *Curr. Opin. Biotechnol.* **2016**, *39* (I), 76–88. https://doi.org/10.1016/j.copbio.2016.01.002.

(5) Dufresne, A. Nanocellulose Processing Properties and Potential Applications. *Curr. For.*

Reports **2019**. https://doi.org/10.1007/s40725-019-00088-1.
(6) Klemm, D.; Cranston, E. D.; Fischer, D.; Gama, M.; Kedzior, S. A.; Kralisch, D.; Kramer, F.; Kondo, T.; Lindström, T.; Nietzsche, S.; et al. Nanocellulose as a Natural Source for Groundbreaking Applications in Materials Science: Today's State. *Mater. Today* **2018**, *21* (7), 720–748. https://doi.org/10.1016/j.mattod.2018.02.001.
(7) Eichhorn, S. J. Cellulose Nanowhiskers: Promising Materials for Advanced Applications. *Soft Matter* **2011**, *7* (2), 303–315. https://doi.org/10.1039/c0sm00142b.
(8) Heinze, T.; El Seoud, O. A.; Koschella, A. *Principles of Cellulose Derivatization*; 2018. https://doi.org/10.1007/978-3-319-73168-1_4.
(9) Genix, A. C.; Oberdisse, J. Nanoparticle Self-Assembly: From Interactions in Suspension to Polymer Nanocomposites. *Soft Matter* **2018**, *14* (25), 5161–5179. https://doi.org/10.1039/c8sm00430g.
(10) Dannert, C.; Stokke, B. T.; Dias, R. S. Nanoparticle-Hydrogel Composites: From Molecular Interactions to Macroscopic Behavior. *Polymers (Basel).* **2019**, *11* (2). https://doi.org/10.3390/polym11020275.
(11) Hiraishi, M.; Igarashi, K.; Kimura, S.; Wada, M.; Kitaoka, M.; Samejima, M. Synthesis of Highly Ordered Cellulose II in Vitro Using Cellodextrin Phosphorylase. *Carbohydr. Res.* **2009**, *344* (18), 2468–2473. https://doi.org/10.1016/j.carres.2009.10.002.
(12) O'Neill, E. C.; Field, R. A. Enzymatic Synthesis Using Glycoside Phosphorylases. *Carbohydr. Res.* **2015**, *403*, 23–37. https://doi.org/10.1016/j.carres.2014.06.010.
(13) Hata, Y.; Fukaya, Y.; Sawada, T.; Nishiura, M.; Serizawa, T. Biocatalytic Oligomerization-Induced Self-Assembly of Crystalline Cellulose Oligomers into Nanoribbon Networks Assisted by Organic Solvents. *Beilstein J. Nanotechnol.* **2019**, *10*, 1778–1788. https://doi.org/10.3762/bjnano.10.173.
(14) Wijekoon, A.; Fountas-Davis, N.; Leipzig, N. D. Fluorinated Methacrylamide Chitosan Hydrogel Systems as Adaptable Oxygen Carriers for Wound Healing. *Acta Biomater.* **2013**, *9* (3), 5653–5664. https://doi.org/10.1016/j.actbio.2012.10.034.
(15) Akula, S.; Brosch, I. K.; Leipzig, N. D. Fluorinated Methacrylamide Chitosan Hydrogels Enhance Cellular Wound Healing Processes. *Ann. Biomed. Eng.* **2017**, *45* (11), 2693–2702. https://doi.org/10.1007/s10439-017-1893-6.
(16) Patil, P. S.; Fathollahipour, S.; Inmann, A.; Pant, A.; Amini, R.; Shriver, L. P.; Leipzig, N. D. Fluorinated Methacrylamide Chitosan Hydrogel Dressings Improve Regenerated Wound Tissue Quality in Diabetic Wound Healing. *Adv. Wound Care* **2019**, *8* (8), 374–385. https://doi.org/10.1089/wound.2018.0887.
(17) O'Hagan, D. Understanding Organofluorine Chemistry. An Introduction to the C-F Bond. *Chem. Soc. Rev.* **2008**, *37* (2), 308–319. https://doi.org/10.1039/b711844a.
(18) Purser, S.; Moore, P. R.; Swallow, S.; Gouverneur, V. Fluorine in Medicinal Chemistry. *Chem. Soc. Rev.* **2008**, *37* (2), 320–330. https://doi.org/10.1039/b610213c.
(19) Theodoridis, G. *Chapter 4 Fluorine-Containing Agrochemicals: An Overview of Recent Developments*; 2006; Vol. 2. https://doi.org/10.1016/S1872-0358(06)02004-5.
(20) Meanwell, N. A. Fluorine and Fluorinated Motifs in the Design and Application of Bioisosteres for Drug Design. *J. Med. Chem.* **2018**, *61* (14), 5822–5880. https://doi.org/10.1021/acs.jmedchem.7b01788.
(21) Gillis, E. P.; Eastman, K. J.; Hill, M. D.; Donnelly, D. J.; Meanwell, N. A. Applications of Fluorine in Medicinal Chemistry. *J. Med. Chem.* **2015**, *58* (21), 8315–8359. https://doi.org/10.1021/acs.jmedchem.5b00258.
(22) Daniel Martínez, J.; Valverde, P.; Delgado, S.; Romanò, C.; Linclau, B.; Reichardt, N. C.; Oscarson, S.; Ardá, A.; Jiménez-Barbero, J.; Javier Cañada, F. Unraveling Sugar Binding Modes to DC-SIGN by Employing Fluorinated Carbohydrates. *Molecules* **2019**, *24* (12). https://doi.org/10.3390/molecules24122337.
(23) Diercks, T.; Infantino, A. S.; Unione, L.; Jiménez-Barbero, J.; Oscarson, S.; Gabius, H. J. Fluorinated Carbohydrates as Lectin Ligands: Synthesis of OH/F-Substituted N-Glycan Core Trimannoside and Epitope Mapping by 2D STD-TOCSYreF NMR Spectroscopy. *Chem. - A Eur. J.* **2018**, *24* (59), 15761–15765. https://doi.org/10.1002/chem.201803217.
(24) Van Straaten, K. E.; Kuttiyatveetil, J. R. A.; Sevrain, C. M.; Villaume, S. A.; Jiménez-

Barbero, J.; Linclau, B.; Vincent, S. P.; Sanders, D. A. R. Structural Basis of Ligand Binding to UDP-Galactopyranose Mutase from Mycobacterium Tuberculosis Using Substrate and Tetrafluorinated Substrate Analogues. *J. Am. Chem. Soc.* **2015**, *137* (3), 1230–1244. https://doi.org/10.1021/ja511204p.

(25) Nartowski, K. P.; Malhotra, D.; Hawarden, L. E.; Sibik, J.; Iuga, D.; Zeitler, J. A.; Fábián, L.; Khimyak, Y. Z. 19 F NMR Spectroscopy as a Highly Sensitive Method for the Direct Monitoring of Confined Crystallization within Nanoporous Materials . *Angew. Chemie* **2016**, *128* (31), 9050–9054. https://doi.org/10.1002/ange.201602936.

(26) Preslar, A. T.; Lilley, L. M.; Sato, K.; Zhang, S.; Chia, Z. K.; Stupp, S. I.; Meade, T. J. Calcium-Induced Morphological Transitions in Peptide Amphiphiles Detected by 19F-Magnetic Resonance Imaging. *ACS Appl. Mater. Interfaces* **2017**, *9* (46), 39890–39894. https://doi.org/10.1021/acsami.7b07828.

(27) Pinney, J. R.; Melkus, G.; Cerchiari, A.; Hawkins, J.; Desai, T. A. Novel Functionalization of Discrete Polymeric Biomaterial Microstructures for Applications in Imaging and Three-Dimensional Manipulation. *ACS Appl. Mater. Interfaces* **2014**, *6* (16), 14477–14485. https://doi.org/10.1021/am503778t.

(28) Ghuman, H.; Hitchens, T. K.; Modo, M. A Systematic Optimization of 19F MR Image Acquisition to Detect Macrophage Invasion into an ECM Hydrogel Implanted in the Stroke-Damaged Brain. *Neuroimage* **2019**, *202* (April), 116090. https://doi.org/10.1016/j.neuroimage.2019.116090.

(29) Li, C.; Lutz, E. A.; Slade, K. M.; Ruf, R. A. S.; Wang, G.-F.; Pielak, G. J. 19F NMR Studies of α-Synuclein Conformation and Fibrillation. *Biochemistry* **2009**, *48* (36), 8578–8584. https://doi.org/10.1021/bi900872p.

(30) Uhrig, M. L.; Lantaño, B.; Postigo, A. Synthetic Strategies for Fluorination of Carbohydrates. *Org. Biomol. Chem.* **2019**, *17* (21), 5173–5189. https://doi.org/10.1039/c9ob00405j.

(31) Denavit, V.; St-Gelais, J.; Tremblay, T.; Giguère, D. Exploring the Chemistry of Non-Sticky Sugars: Synthesis of Polyfluorinated Carbohydrate Analogues of d-Allopyranose. *Chem. - A Eur. J.* **2019**, *25* (39), 9272–9279. https://doi.org/10.1002/chem.201901346.

(32) Yu, Y.; Tyrikos-Ergas, T.; Zhu, Y.; Fittolani, G.; Bordoni, V.; Singhal, A.; Fair, R. J.; Grafmüller, A.; Seeberger, P. H.; Delbianco, M. Systematic Hydrogen-Bond Manipulations To Establish Polysaccharide Structure–Property Correlations. *Angew. Chemie - Int. Ed.* **2019**, *58* (37), 13127–13132. https://doi.org/10.1002/anie.201906577.

(33) Zhu, J. S.; McCormick, N. E.; Timmons, S. C.; Jakeman, D. L. Synthesis of α-Deoxymono and Difluorohexopyranosyl 1-Phosphates and Kinetic Evaluation with Thymidylyl- and Guanidylyltransferases. *J. Org. Chem.* **2016**, *81* (19), 8816–8825. https://doi.org/10.1021/acs.joc.6b01485.

(34) López Durán, V.; Larsson, P. A.; Wågberg, L. Chemical Modification of Cellulose-Rich Fibres to Clarify the Influence of the Chemical Structure on the Physical and Mechanical Properties of Cellulose Fibres and Thereof Made Sheets. *Carbohydr. Polym.* **2018**, *182* (August 2017), 1–7. https://doi.org/10.1016/j.carbpol.2017.11.006.

(35) Kasuya, N.; Iiyama, K.; Meshitsuka, G.; Okano, T. A Structural Study of Fluorinated Cellulose: Crystallization of Fluorinated Cellulose by Conversion Treatments Used for Cellulose. *Carbohydr. Polym.* **1997**, *34* (4), 229–234. https://doi.org/10.1016/s0144-8617(97)00068-4.

(36) Yu, Y.; Gim, S.; Kim, D.; Arnon, Z. A.; Gazit, E.; Seeberger, P. H.; Delbianco, M. Oligosaccharides Self-Assemble and Show Intrinsic Optical Properties. *J. Am. Chem. Soc.* **2019**, *141* (12), 4833–4838. https://doi.org/10.1021/jacs.8b11882.

(37) Okamoto, E.; Kiyosada, T.; Shoda, S. I.; Kobayashi, S. Synthesis of Alternatingly 6-O-Methylated Cellulose via Enzymatic Polymerization of a Substituted Cellobiosyl Fluoride Monomer Catalyzed by Cellulase. *Cellulose* **1997**, *4* (2), 161–172. https://doi.org/10.1023/A:1018479621509.

(38) Kadokawa, J. I. Precision Polysaccharide Synthesis Catalyzed by Enzymes. *Chem. Rev.* **2011**, *111* (7), 4308–4345. https://doi.org/10.1021/cr100285v.

(39) Kobayashi, S. New Developments of Polysaccharide Synthesis via Enzymatic Polymerization. *Proc. Jpn. Acad. Ser. B. Phys. Biol. Sci.* **2007**, *83* (8), 215–247.

https://doi.org/10.2183/pjab/83.215.
(40) Pergolizzi, G.; Kuhaudomlarp, S.; Kalita, E.; Field, R. A. Glycan Phosphorylases in Multi-Enzyme Synthetic Processes. *Protein Pept. Lett.* **2017**, *24* (8). https://doi.org/10.2174/0929866524666170811125109.
(41) Kadokawa, J. I. α-Glucan Phosphorylase-Catalyzed Enzymatic Reactions Using Analog Substrates to Synthesize Non-Natural Oligo-and Polysaccharides. *Catalysts* **2018**, *8* (10). https://doi.org/10.3390/catal8100473.
(42) Kadokawa, J. I. Enzymatic Preparation of Functional Polysaccharide Hydrogels by Phosphorylase Catalysis. *Pure Appl. Chem.* **2018**, *90* (6), 1045–1054. https://doi.org/10.1515/pac-2017-0802.
(43) Loos, K.; Kadokawa, J. Synthesis of Polysaccharides II: Phosphorylase as Catalyst. In *Enzymatic Polymerization towards Green Polymer Chemistry*; Kobayashi, S., Uyama, H., Kadokawa, J., Eds.; Springer Singapore: Singapore, 2019; pp 47–87. https://doi.org/10.1007/978-981-13-3813-7_3.
(44) Yataka, Y.; Sawada, T.; Serizawa, T. Enzymatic Synthesis and Post-Functionalization of Two-Dimensional Crystalline Cellulose Oligomers with Surface-Reactive Groups. *Chem. Commun.* **2015**, *51* (63), 12525–12528. https://doi.org/10.1039/c5cc04378f.
(45) Wang, J.; Niu, J.; Sawada, T.; Shao, Z.; Serizawa, T. A Bottom-Up Synthesis of Vinyl-Cellulose Nanosheets and Their Nanocomposite Hydrogels with Enhanced Strength. *Biomacromolecules* **2017**, *18* (12), 4196–4205. https://doi.org/10.1021/acs.biomac.7b01224.
(46) Yataka, Y.; Sawada, T.; Serizawa, T. Multidimensional Self-Assembled Structures of Alkylated Cellulose Oligomers Synthesized via in Vitro Enzymatic Reactions. *Langmuir* **2016**, *32* (39), 10120–10125. https://doi.org/10.1021/acs.langmuir.6b02679.
(47) Nohara, T.; Sawada, T.; Tanaka, H.; Serizawa, T. Enzymatic Synthesis of Oligo(Ethylene Glycol)-Bearing Cellulose Oligomers for in Situ Formation of Hydrogels with Crystalline Nanoribbon Network Structures. *Langmuir* **2016**, *32* (47), 12520–12526. https://doi.org/10.1021/acs.langmuir.6b01635.
(48) Samain, E.; Lancelon-Pin, C.; Férigo, F.; Moreau, V.; Chanzy, H.; Heyraud, A.; Driguez, H. Phosphorolytic Synthesis of Cellodextrins. *Carbohydr. Res.* **1995**, *271* (2), 217–226. https://doi.org/10.1016/0008-6215(95)00022-L.
(49) Serizawa, T.; Kato, M.; Okura, H.; Sawada, T.; Wada, M. Hydrolytic Activities of Artificial Nanocellulose Synthesized via Phosphorylase-Catalyzed Enzymatic Reactions. *Polym. J.* **2016**, *48* (4), 539–544. https://doi.org/10.1038/pj.2015.125.
(50) Serizawa, T.; Fukaya, Y.; Sawada, T. Self-Assembly of Cellulose Oligomers into Nanoribbon Network Structures Based on Kinetic Control of Enzymatic Oligomerization. *Langmuir* **2017**, *33* (46), 13415–13422. https://doi.org/10.1021/acs.langmuir.7b03653.
(51) Brinkmann, A.; Chen, M.; Couillard, M.; Jakubek, Z. J.; Leng, T.; Johnston, L. J. Correlating Cellulose Nanocrystal Particle Size and Surface Area. *Langmuir* **2016**, *32* (24), 6105–6114. https://doi.org/10.1021/acs.langmuir.6b01376.
(52) O'Neill, E. C.; Pergolizzi, G.; Stevenson, C. E. M.; Lawson, D. M.; Nepogodiev, S. A.; Field, R. A. Cellodextrin Phosphorylase from Ruminiclostridium Thermocellum: X-Ray Crystal Structure and Substrate Specificity Analysis. *Carbohydr. Res.* **2017**, *451*, 118–132. https://doi.org/10.1016/j.carres.2017.07.005.
(53) Nidetzky, B.; Eis, C.; Albert, M. Role of Non-Covalent Enzyme-Substrate Interactions in the Reaction Catalysed by Cellobiose Phosphorylase from Cellulomonas Uda. *Biochem. J.* **2000**, *351* (3), 649–659. https://doi.org/10.1042/0264-6021:3510649.
(54) Langan, P.; Nishiyama, Y.; Chanzy, H. X-Ray Structure of Mercerized Cellulose II at 1 Å Resolution. *Biomacromolecules* **2001**, *2* (2), 410–416. https://doi.org/10.1021/bm005612q.
(55) Langan, P.; Nishiyama, Y.; Chanzy, H. A Revised Structure and Hydrogen-Bonding System in Cellulose II from a Neutron Fiber Diffraction Analysis. *J. Am. Chem. Soc.* **1999**, *121* (43), 9940–9946. https://doi.org/DOI 10.1021/ja9916254.
(56) Schenzel, K.; Almlöf, H.; Germgård, U. Quantitative Analysis of the Transformation Process of Cellulose I → Cellulose II Using NIR FT Raman Spectroscopy and Chemometric Methods. *Cellulose* **2009**, *16* (3), 407–415. https://doi.org/10.1007/s10570-009-9286-0.
(57) Tammer, M. G. Sokrates: Infrared and Raman Characteristic Group Frequencies: Tables and

Charts. *Colloid Polym. Sci.* **2004**, *283* (2), 235. https://doi.org/10.1007/s00396-004-1164-6.
(58) Mittal, A.; Katahira, R.; Himmel, M. E.; Johnson, D. K. Effects of Alkaline or Liquid-Ammonia Treatment on Crystalline Cellulose: Changes in Crystalline Structure and Effects on Enzymatic Digestibility. *Biotechnol. Biofuels* **2011**, *4*, 1–16. https://doi.org/10.1186/1754-6834-4-41.
(59) Larsson, P. T.; Wickholm, K.; Iversen, T. A CP/MAS 13C NMR Investigation of Molecular Ordering in Celluloses. *Carbohydr. Res.* **1997**, *302* (1–2), 19–25. https://doi.org/10.1016/S0008-6215(97)00130-4.
(60) Muñoz-García, J. C.; Corbin, K. R.; Hussain, H.; Gabrielli, V.; Koev, T.; Iuga, D.; Round, A. N.; Mikkelsen, D.; Gunning, P. A.; Warren, F. J.; et al. High Molecular Weight Mixed-Linkage Glucan as a Mechanical and Hydration Modulator of Bacterial Cellulose: Characterization by Advanced NMR Spectroscopy. *Biomacromolecules* **2019**. https://doi.org/10.1021/acs.biomac.9b01070.
(61) Wang, T.; Hong, M. Solid-State NMR Investigations of Cellulose Structure and Interactions with Matrix Polysaccharides in Plant Primary Cell Walls. *J. Exp. Bot.* **2015**, *67* (2), 503–514. https://doi.org/10.1093/jxb/erv416.
(62) Wada, M.; Heux, L.; Isogai, A.; Nishiyama, Y.; Chanzy, H.; Sugiyama, J. Improved Structural Data of Cellulose IIII Prepared in Supercritical Ammonia. *Macromolecules* **2001**, *34* (5), 1237–1243. https://doi.org/10.1021/ma001406z.
(63) Wada, M.; Heux, L.; Nishiyama, Y.; Langan, P. X-Ray Crystallographic, Scanning Microprobe X-Ray Diffraction, and Cross-Polarized/Magic Angle Spinning 13C NMR Studies of the Structure of Cellulose III(II). *Biomacromolecules* **2009**, *10* (2), 302–309. https://doi.org/10.1021/bm8010227.
(64) Wada, M.; Chanzy, H.; Nishiyama, Y.; Langan, P. Cellulose IIII Crystal Structure and Hydrogen Bonding by Synchrotron X-Ray and Neutron Fiber Diffraction. *Macromolecules* **2004**, *37* (23), 8548–8555. https://doi.org/10.1021/ma0485585.
(65) White, P. B.; Wang, T.; Park, Y. B.; Cosgrove, D. J.; Hong, M. Water-Polysaccharide Interactions in the Primary Cell Wall of Arabidopsis Thaliana from Polarization Transfer Solid-State NMR. *J. Am. Chem. Soc.* **2014**, *136* (29), 10399–10409. https://doi.org/10.1021/ja504108h.
(66) Nigmatullin, R.; Harniman, R.; Gabrielli, V.; Muñoz-García, J. C.; Khimyak, Y. Z.; Angulo, J.; Eichhorn, S. J. Mechanically Robust Gels Formed from Hydrophobized Cellulose Nanocrystals. *ACS Appl. Mater. Interfaces* **2018**, *10* (23). https://doi.org/10.1021/acsami.8b05067.

Supporting information

Bottom-up chemoenzymatic synthesis towards novel fluorinated cellulose-like materials

Peterson de Andrade[1‡], Juan C. Muñoz-García[2‡], Giulia Pergolizzi[1‡], Valeria Gabrielli[2], Sergey A. Nepogodiev[1], Dinu Iuga[3], László Fábián,[2] Rinat Nigmatullin[4], Marcus A. Johns[4], Robert Harniman[5], Stephen J. Eichhorn[4], Jesús Angulo[2], Yaroslav Z. Khimyak[*2], Robert A. Field[*1,6]

[1]*Department of Biological Chemistry, John Innes Centre, Norwich NR4 7UH, UK*
[2]*School of Pharmacy, University of East Anglia, Norwich Research Park, Norwich NR4 7TJ, UK*
[3]*Department of Physics, University of Warwick, Coventry CV4 7AL, UK*
[4]*Bristol Composites Institute, CAME School of Engineering, University of Bristol, Bristol BS8 1TR, UK*
[5]*School of Chemistry, University of Bristol, Bristol BS8 1TS, UK*
[6]*Present address: Department of Chemistry and Manchester Institute of Biotechnology, University of Manchester, Manchester M1 7DN, UK*

‡ Equal contribution

[*]Corresponding author e-mail address: rob.field@jic.ac.uk; robert.field@manchester.ac.uk; y.khimyak@uea.ac.uk

Table of Contents

1. Materials and methods

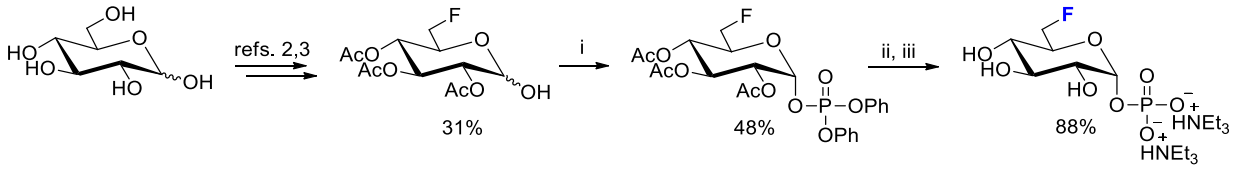
1.1. General materials and methods

Chemicals were commercially obtained as reagent grade and used without any purification. Deoxy-fluoro-D-glucoses (2F-, 3F- and 6F-Glc) and α-D-glucose 1-phosphate disodium salt hydrate (Glc-1P) were purchased from Toronto Research Chemicals (Canada) and Sigma-Aldrich (UK), respectively. Cellobiose phosphorylase (CBP) (PRO-GH94-004) was kindly provided by Prozomix Limited (UK) and Milli-Q (MQ) H_2O was used to prepare all buffers. Thin-layer chromatography (TLC) was performed on pre-coated silica gel 60 F_{254} plates (Merck) and compounds were visualized by UV irradiation (λ 254 nm) and/or by spraying TLC with staining solution (2% orcinol w/v in $EtOH/H_2O/H_2SO_4$ 15:1:2 v/v/v) followed by heating. Biotage SP4 flash chromatography system was used for purification of protected monosaccharides using normal phase (pre-packed SNAP cartridges) and the monofluorinated cellobiose analogues were purified by HPLC (Thermo Scientific Dionex Ultimate 3000) on a Luna OH column (5 µm HILIC 200 Å, 250 × 10 mm, Phenomenex) using 5 mM ammonium formate buffer (5%) and acetonitrile (95%) at 5 mL/min in isocratic elution over 25 min. Detection was performed by charged aerosol detector (CAD) with power function 1.00, data collection rate 10 Hz and nebulizer temperature 25 °C. Products were lyophilized using a Labconco FreeZone Benchtop freeze dryer. ^{1}H, ^{13}C, ^{31}P and ^{19}F NMR spectra were recorded on a Bruker Avance III 400 MHz and/or Bruker Avance Neo 600 MHz spectrometers at 298 K. Chemical shifts recorded in D_2O are reported with respect to the solvent residual peak at 4.79 ppm in ^{1}H NMR. High resolution mass spectra were acquired in a Synapt G2-Si mass spectrometer (Waters, UK) using electrospray ionisation (positive or negative mode). Optical rotations were measured at 20 °C using a Perkin-Elmer Model 341 polarimeter.

1.2. Expression and purification of cellodextrin phosphorylase (CDP)

A recombinant plasmid (pET15b) containing the CDP gene from *Ruminiclostridium thermocellum* (YM4 strain) was transformed into *E. coli* BL21 (DE3) cells and grown as described previously.[1] Briefly, 1 L of LB medium containing the transformant and carbenicillin (100 µg/mL) was incubated at 37 °C with shaking (200 rpm) until OD_{600} around 0.6. Heterologous protein expression was induced by adding isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 1 mM and incubating for 4 hours at 30 °C with shaking (180 rpm). The cells were harvested by centrifugation (4,000 × g, 20 min), re-suspended in lysis buffer (50 mM HEPES, pH 7.5, 100 mM NaCl, EDTA-free protease inhibitor cocktail tablet, 0.02 mg/mL DNaseI), lysed by cell disruption (30 Kpsi, constant flow) and the supernatant containing the recombinant proteins was separated from cell debris by centrifugation (20,000 × g, 30 min). Proteins were purified at 4 °C using an ÄKTA pure FPLC system (GE Healthcare). The supernatant was loaded to a 5 mL HisTrap™ HP column (GE healthcare) pre-equilibrated with buffer A (50 mM Tris-HCl, pH 8, 50 mM glycine, 5% glycerol, 500 mM NaCl, 20 mM imidazole). The column was washed with buffer A to remove unbound proteins followed by elution of bound proteins with buffer B (50 mM Tris-HCl, pH 8, 50 mM glycine, 5% glycerol, 500 mM NaCl, 500 mM imidazole). Further purification was carried out by gel filtration chromatography (Superdex S200 16/600 column, GE Healthcare) with 20 mM HEPES, pH 7.5, 150 mM NaCl, 1 mL/min. Fractions containing CDP were pooled and concentrated using Amicon Ultra-15 Centrifugal Filter (30,000 MW cut off) and the enzyme concentration (5.7 mg/mL) was determined by NanoDrop™ spectrophotometer (Thermo Fisher Scientific, UK). The His-tag CDP was stored in aliquots at -80 °C until required.

1.3. Chemical synthesis of 6-deoxy-6-fluoro-α-D-glucose 1-phosphate (6F-Glc-1P)

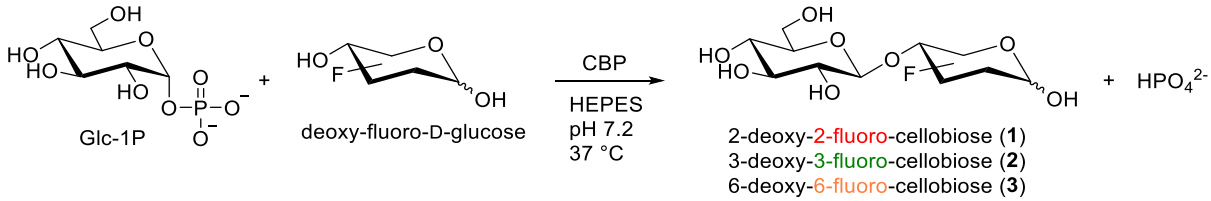


Reagent and conditions: i) *n*-BuLi, $(PhO)_2POCl$, THF, -78 °C to room temperature, overnight; ii) PtO_2, $H_{2(g)}$, EtOH, room temperature, 24 h; iii) Et_3N:H_2O:MeOH (1:3:7, v/v/v), room temperature, 48 h.

n-BuLi (2 mL of 1.6 M in hexane; 3.20 mmol; 1.2 eq.) was added dropwise to a solution of 6-deoxy-6-fluoro-2,3,4-tri-*O*-acetyl-D-glucopyranose (822 mg; 2.67 mmol; 1 eq.) (synthesized from D-glucose in five steps)[2,3] in anhydrous THF (40 mL) at -78 °C under N_2 atmosphere and stirred for 15 min. Diphenyl chlorophosphate (664 µL; 3.20 mmol; 1.2 eq.) was added dropwise and the reaction mixture was allowed to warm to room temperature overnight. The reaction was quenched with NH_4Cl saturated solution (30 mL) and partitioned with EtOAc (3 x 30 mL). The organic phase was washed with NaCl saturated solution (2 × 30 mL), dried over $MgSO_4$, filtered, concentrated under vacuum and purified by flash chromatography [cartridge SNAP 25g; solvent: Hexane/EtOAc; gradient: 0-20%, 20-20% and 20-30% (v/v); flow: 25 mL/min] to afford the product 6-deoxy-6-fluoro-2,3,4-tri-*O*-acetyl-α-D-glucopyranosyl 1-diphenylphosphate[4] in 48% yield (690 mg; 1.28 mmol). The deprotection steps of the synthesised phosphate (600 mg; 1.11 mmol) were performed with PtO_2 (50 mg; 0.22 mmol; 0.2 eq.) in absolute ethanol (25 mL) at room temperature under H_2 atmosphere for 48 h. After catalyst removal by filtration, the crude was concentrated under vacuum, dissolved in MeOH (20 mL) followed by addition of Et_3N (10 mL) and concentrated again. Lastly, the residue was dissolved in Et_3N:H_2O:MeOH (1:3:7, v/v/v) (45 mL) and stirred for 48 h at room temperature. After concentration under vacuum, the crude was dissolved in water and freeze dried to afford 6-deoxy-6-fluoro-α-D-glucose 1-phosphate triethylammonium salt[4] in 88% yield (457 mg; 0.98 mmol). ^{1}H NMR (400 MHz, D_2O) δ 5.49 (1 H, dd, $J_{1,P}$ 7.0 Hz, $J_{1,2}$ 3.5 Hz, H1), 4.83-4.58 (2 H, m, H6a, H6b), 3.97 (1 H, dd, $J_{5,F}$ 30.3 Hz, $J_{4,5}$ 10.6 Hz, H5), 3.78 (1 H, t, $J_{2,3}$ = $J_{3,4}$ 9.5 Hz, H3), 3.60-3.51 (2 H, m, H4, H2), 3.20 (10 H, q, *J* 7.4 Hz, 2x HN($C\underline{H}_2CH_3)_3$), 1.28 (16 H, t, *J* 7.3 Hz, 2x HN($CH_2C\underline{H}_3)_3$). ^{13}C NMR (101 MHz, D_2O): δ 94.61 (C1), 82.80 (C6), 81.13 (C6), 72.56 (C3), 71.48 (C2), 71.36 (C5), 71.17 (C5), 68.27 (C4), 46.62 (HN($\underline{C}H_2CH_3)_3$), 8.19 (HN($CH_2\underline{C}H_3)_3$). ^{19}F NMR (^{1}H-decoupled, 376 MHz, D_2O): δ -236.30. ^{31}P NMR (162 MHz, D_2O): δ -1.07. HRMS (ESI): *m/z* calculated for $C_6H_{11}FO_8P$ [M-H]$^-$: 261.0181; found 261.0173.

1.4. Enzymatic synthesis of fluorinated cellodextrins

1.4.1. <u>Synthesis of 2-, 3- and 6-monofluorinated cellobioses **1-3**</u>



To a solution of deoxy-fluoro-glucose (F-Glc) (18.2 mg, 100 mM stock solution in Milli-Q (MQ) water, 1 mL, 1 eq., 13 mM final concentration) and Glc-1P (disodium salt hydrate, 30.4 mg, 100 mM stock solution in MQ water, 1 mL, 1 eq., 13 mM final concentration) in HEPES buffer (80 mM stock solution in MQ water, 5 mL, pH 7.2), was added cellobiose phosphorylase (CBP, 1 mg/mL stock solution, 1 mL, 1 mg) and the reaction was incubated at 37 °C shaking

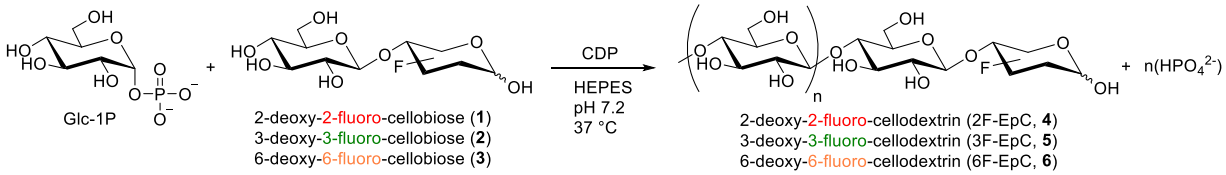
for 16 h. The reaction mixture was centrifuged, and the supernatant was passed through HisTrap HP (1 mL) and High Q (5 mL) Bio-Rad columns to remove CBP and remaining Glc-1P, respectively. The pre-purified mixture was freeze dried, dissolved in 2 mL MQ water and purified by gel filtration chromatography (Toyopearl TSK HW40S, 1.6 × 90 cm, MQ water, 0.5 mL/min) followed by HPLC (Luna OH 5 µm HILIC 200 Å, 250 × 10 mm, Phenomenex, 5 mL/min in isocratic elution over 25 min, 5% 5 mM ammonium formate buffer and 95% acetonitrile) to afford the title compounds **1**-**3**:

2-Deoxy-2-fluoro-cellobiose (**1**).
Yield: 26% (9 mg; 0.029 mmol); Rf = 0.40 (isopropanol:NH_4OH:H_2O, 6:3:1); $[\alpha]_D^{20}$ +32.5 (c 1, CH_3OH). ^{1}H NMR (600 MHz, D_2O) δ 5.47 (0.45 H, d, $J_{1,2}$ 3.9 Hz, H1α), 4.95 (0.5 H, dd, $J_{1,2}$ 7.9 Hz, $J_{1,F}$ 2.4 Hz, H1β), 4.54 (1 H, two d, $J_{1',2'}$ 8.0 Hz, H1'), 4.48 (0.45 H, ddd, $J_{2,F}$ 49.5 Hz, $J_{2,3}$ 9.5 Hz, $J_{1,2}$ 3.9 Hz, H2α), 4.22-4.07 (1 H, m, H2β, H3α), 4.02-3.87 (3.5 H, m, H5α, H3β, H6$_a$β, H6$_a$α, H6'$_a$, H6$_b$α), 3.83 (0.5 H, dd, $J_{6a,6b}$ 12.4 Hz, $J_{5,6}$ 5.0 Hz, H6$_b$β), 3.78-3.72 (2 H, m, H4α, H4β, H6'$_b$), 3.67 (0.5 H, ddd, $J_{4,5}$ 9.9 Hz, $J_{5,6a}$ 5.0 Hz, $J_{5,6b}$ 2.2 Hz, H5β), 3.56-3.50 (2 H, m, H3', H5'), 3.47-3.42 (1 H, m, H4'), 3.37-3.32 (1 H, m, H2'). ^{13}C NMR (151 MHz, D_2O): δ 102.48 (C1'), 93.44 (C1$_\beta$), 93.21 (C2$_\beta$), 91.99 (C2$_\beta$), 90.55 (C2$_\alpha$), 89.44 (C1$_\alpha$), 89.32 (C2$_\alpha$), 78.02 (C4$_\alpha$), 77.97 (C4$_\beta$), 75.97 (C5'), 75.46 (C3'), 74.89 (C5$_\beta$), 73.11 (C2'), 72.76 (C3$_\beta$), 69.94 (C5$_\alpha$), 69.92 (C3$_\alpha$), 69.48 (C4'), 60.64 (C6'), 59.83 (C6$_\beta$), 59.68 (C6$_\alpha$). ^{19}F NMR (^{1}H-decoupled, 376 MHz, D_2O) δ -199.08, -199.27. HRMS (ESI): *m/z* calculated for $C_{12}H_{21}FO_{10}Na^+$ $[M+Na]^+$: 367.1011; found: 367.1013.

3-Deoxy-3-fluoro-cellobiose (**2**).
Yield: 12% (4 mg; 0.012 mmol); Rf = 0.29 (isopropanol:NH_4OH:H_2O, 6:3:1); $[\alpha]_D^{20}$ +28.7 (c 1, CH_3OH). ^{1}H NMR (600 MHz, D_2O) δ 5.28 (0.5 H, t, $J_{1,2}$ = $J_{1,F}$ 3.8 Hz, H1α), 4.77-4.66 (1 H, m, H3α, H1β), 4.64-4.51 (1.5 H, m, H3β, H1'), 4.04-3.96 (2 H, m, H5α, H4α, H4β, H6$_a$β), 3.94-3.90 (2 H, m, H6$_a$α, H6'$_a$, H6$_b$α), 3.88-3.82 (1 H, m, H2α, H6$_b$β), 3.72 (1 H, dd, $J_{6'a,6'b}$ 12.5 Hz, $J_{5,6'b}$ 5.9 Hz, H6'$_b$), 3.63-3.60 (0.5 H, m, H5β), 3.57 (0.5 H, ddd, $J_{2,F}$ 14.2, $J_{2,3}$ 9.1, $J_{1,2}$ 8.0 Hz, H2β), 3.53-3.49 (1 H, m, H3'), 3.48-3.44 (1 H, m, H5'), 3.42-3.37 (1 H, m, H4'), 3.33-3.28 (1 H, m, H2').^{13}C NMR (151 MHz, D_2O): δ 102.39 (C1'), 95.29 (C3$_\beta$), 95.05 (C1$_\beta$), 94.08 (C3$_\beta$), 93.73 (C3$_\alpha$), 92.53 (C3$_\alpha$), 91.99 (C1$_\alpha$), 76.04 (C5'), 75.69 (C5$_\alpha$), 75.56 (C3'), 75.49 (C4$_\alpha$), 73.84 (C5$_\beta$), 73.27 (C2'), 72.76 (C2$_\beta$), 70.07 (C2$_\alpha$), 69.86 (C4$_\beta$), 69.61 (C4'), 60.67 (C6'), 59.74 (C6$_\beta$), 59.61 (C6$_\alpha$). ^{19}F NMR (^{1}H-decoupled, 376 MHz, D_2O) δ -192.35, -197.01 HRMS (ESI): *m/z* calculated for $C_{12}H_{21}FO_{10}Na^+$ $[M+Na]^+$: 367.1011; found: 367.1014.

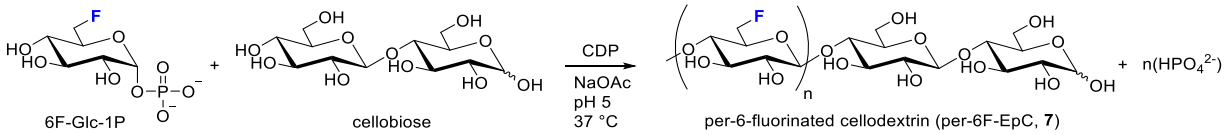
6-Deoxy-6-fluoro-cellobiose (**3**).
Yield: 23% (8 mg; 0.023 mmol); Rf = 0.34 (isopropanol:NH_4OH:H_2O, 6:3:1); $[\alpha]_D^{20}$ +17.0 (c 1, CH_3OH). ^{1}H NMR (600 MHz, D_2O) δ 5.27 (0.55 H, d, $J_{1,2}$ 3.8 Hz, H1α), 4.96-4.85 (1 H, m, H6$_{a,b}$α), 4.76-4.65 (1.5 H, m, H1β, H6$_{a,b}$β), 4.52 (1 H, d, $J_{1',2'}$ 7.9 Hz, H1'), 4.15-4.06 (0.55 H, m, H5α), 3.94 (1 H, dd, $J_{6'a,6'b}$ 12.3 Hz, $J_{5,6'a}$ 2.2 Hz, H6'$_a$), 3.87 (0.55 H, t, $J_{2,3}$ = $J_{3,4}$ 9.3 Hz, H3α), 3.82-3.72 (2.5 H, m, H5β, H4α, H4β, H6'$_b$), 3.69-3.65 (0.45 H, m, H3β), 3.62 (0.55 H, dd, $J_{2,3}$ 9.8 Hz, $J_{1,2}$ 3.8 Hz, H2α), 3.56-3.49 (2 H, m, H3', H5'), 3.47-3.43 (1 H, m, H4'), 3.37-3.31 (1.5 H, m, H2', H2β). ^{13}C NMR (151 MHz, D_2O): δ 102.62 (C1'), 95.90 (C1$_\beta$), 91.96 (C1$_\alpha$), 82.39 (C6$_\alpha$), 82.12 (C6$_\beta$), 81.28 (C6$_\alpha$), 81.01 (C6$_\beta$), 77.80 (C4$_\alpha$), 77.56 (C4$_\beta$), 75.97 (C5'), 75.48 (C3'), 74.10 (C3$_\beta$), 73.76 (C2$_\beta$), 73.33 (C5$_\beta$), 73.15 (C2'), 71.24 (C3$_\alpha$), 71.13 (C2$_\alpha$), 69.42 (C4'), 69.02 (C5$_\alpha$), 60.56 (C6'). ^{19}F NMR (^{1}H-decoupled, 376 MHz, D_2O) δ -233.68, -234.24. HRMS (ESI): *m/z* calculated for $C_{12}H_{21}FO_{10}Na^+$ $[M+Na]^+$: 367.1011; found: 367.1014.

1.4.2. Synthesis of 2-, 3- and 6-monofluorinated cellodextrins (2F-EpC, 4; 3F-EpC, 5 and 6F-EpC, **6**)



The monofluorinated cellodextrins were synthesised in a one-pot reaction. After CBP removal, more Glc-1P (110 mg, 4 eq.) was added together with cellodextrin phosphorylase (CDP, 0.9 mg/mL stock solution, 0.7 mL, 0.63 mg) and the reaction was incubated at 50 °C while shaking for 16 h. A white precipitate was formed and isolated by centrifugation, followed by re-suspension and washing with MQ water (3×). To the supernatant, more Glc-1P was added (90 mg) and the reaction incubated at 50 °C shaking for 12 h. A white precipitate was formed again and isolated by centrifugation, followed by re-suspension and washing with MQ water (3×). The precipitates were combined to give reasonable final yields [47% 2F-EpC (**4**), 30% 3F-EpC (**5**) and 32% 6F-EpC (**6**)]. The white solid was analysed by MALDI-ToF (DP8 on average) and solution-state ^{19}F NMR (^{1}H-decoupled, 376 MHz, 1 M NaOD) δ -195.21 and -195.26 ppm (**4**); -190.86 and -197.19 ppm (**5**); -232.55 and -234.05 ppm (**6**).

1.4.3. Synthesis of multiply 6-fluorinated cellodextrin (multi-6F-EpC, **7**)



Cellodextrin phosphorylase (CDP, 5.7 mg/mL stock solution, 300 µL, 1.7 mg) was added to a solution of cellobiose (5.5 mg, 16.1 µmol, 1 eq.) and 6-deoxy-6-fluoro-α-D-glucose 1-phosphate (6F-Glc-1P) (triethylammonium salt, 45 mg, 96.6 µmol, 6 eq.) in 500 µL NaOAc buffer (200 mM, pH 5.0) and the reaction was incubated at 37 °C for 24 h with shaking (300 rpm). A white precipitate started forming and more CDP was added (1-fold) and the reaction continued in the same conditions for 48 h (72 h in total). The white precipitate was isolated by centrifugation, re-suspended and washed with MQ water (4×) and freeze dried to afford the multiple 6-fluorinated cellodextrin **7** in 64% yield (17 mg; 10.3 µmol). The white solid was analysed by MALDI-ToF (DP10 on average) and solution-state ^{19}F NMR (^{1}H-decoupled, 565 MHz, 1 M NaOD) δ -233.25, -233.29, -233.31, -233.35 ppm.

1.5. Matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS)

Samples suspended in MQ water were mixed with equal volume of 2,5-dihydroxybenzoic acid (DHB) matrix (10 mg/mL in 30% acetonitrile in MQ water), spotted on a target plate (Bruker MTP 384 Polished Steel TF Target) and analysed on an Autoflex™ Speed MALDI-TOF/TOF mass spectrometer (Bruker Daltonics™ GmbH**,** Coventry, UK). The instrument was controlled by a flexControl™ (version 3.4, Bruker) method optimised for peptide detection and calibrated using peptide standards (Bruker). All spectra were processed with flexAnalysis™ (3.4, Bruker).

1.6. Electron microscopy (EM)

Transmission Electron Microscopy (TEM) images were viewed on a Thermo Fisher Talos F200C transmission electron microscope at 200kV (Thermo Fisher UK Ltd, Cambridge, UK) using a Gatan OneView 4k X4K digital camera (Gatan, Abingdon, UK) to record DM4 files. 400 mesh EM Resolution Formvar/Carbon coated copper grids were glow discharged, the samples were suspended in 200 µL MQ water and 5 µL drop was pipetted onto the grid for 1 minute. The excess was then blotted off with filter paper and 5 µL of 2% uranyl acetate was pipetted onto the grid for 30 seconds then blotted off with filter paper and allowed to dry.

1.7. Atomic force microscopy (AFM)

Samples of fluorinated cellodextrins were deposited onto freshly cleaved mica substrates by drop-casting of diluted suspensions (0.02 mg/ml). Cast drops were allowed to dry before SPM investigation in an ambient environment. SPM measurements were conducted using a Multi-Mode VIII microscope with Nanoscope V controller operating under non-resonant PeakForce feedback control (Bruker, CA, USA). SCANASYST-FLUID+ cantilevers were employed with nominal spring constants of 0.7 N/m (Bruker, CA, USA). Real-time analysis of the force-separation curves were collected as part of the Peakforce control mechanism allowing the physical tip-sample interaction forces and sample topography to be mapped concurrently.

1.8. Powder X-ray diffraction (PXRD)

X-ray diffraction data were collected using a single crystal diffractometer (Rigaku Synergy S, Cu X-ray tube, 50kV-1mA) with Cu Kα radiation (λ = 0.154 nm). Samples of enzymatically produced cellodextrin (EpC), 2F-EpC, 3F-EpC, 6F-EpC and multi-6F-EpC were placed in a 96-well plate and analysed using an XtalCheck screening plate mounted on the diffractometer. The collected diffraction images were integrated between diffraction angles (2Θ) 5 and 40°. All the samples were analysed in powder form and after gentle grinding with mortar and pestle.

The simulated powder patterns were generated using Mercury[5] and the published crystal structures of cellulose I_α,[6] I_β,[7] II,[8] III_I[9] and III_{II}.[10]

1.9. Raman spectroscopy

Raman images were obtained using a Raman spectrometer (Renishaw, UK). Raman spectra were acquired using a 785 nm wavelength laser (NIR) and 41 mW laser power at the sample for excitation of Raman scattering. The sample was focused with a 50× objective lens (numerical aperture: 0.7, vertical resolution: 1.6 µm) with a lateral resolution of 684 nm. Due to the narrow spectrum acquisition range, each complete spectrum is a composite of three individual spectra centred at 590, 1090 and 1490 cm^{-1} respectively. Each individual spectrum was obtained from 10 exposures of 20 seconds each. Overlap of the individual spectra enabled their normalisation to produce the complete spectrum. Three complete spectra were acquired for each sample. Raman bands, background fluorescence and photon spikes were fitted for each complete spectrum using Lorentzian curve functions in Fityk.[11] This enabled the generation of clean spectra for each sample and the identification of the wavelength and intensity for each band.

1.10. Characterization of cellodextrins by solution and solid-state NMR

^{1}H-^{13}C cross-polarization solid-state NMR experiments of the single 2-, 3- and 6- monofluorinated EpC 10 wt% hydrogels were performed at 5 °C using a Bruker Avance III spectrometer equipped with a 4 mm triple resonance probe operating at frequencies of 400.2 MHz (^{1}H) and 100.6 MHz (^{13}C). Each gel was packed into a kel-f insert, sealed using a plug and a screw, and spun at 6 kHz. For the powder samples, experiments were carried out at the UK 850 MHz solid state NMR facility (Warwick). All the experiments were run at frequencies of 850.2 MHz (^{1}H), 799.8 MHz (^{19}F) and 213.8 MHz (^{13}C). ^{1}H-decoupled ^{1}H-^{19}F CP (^{1}H-^{19}F{1H}), ^{1}H-^{13}C CP with ^{1}H and ^{19}F decoupling (^{1}H-^{13}C{^{1}H,^{19}F}), and ^{19}F-^{13}C CP with ^{1}H and ^{19}F decoupling (^{19}F-^{13}C{^{1}H,^{19}F}) experiments were run using a 2.5 mm HFX H13894 probe, at 15 kHz MAS spinning and 12 ± 3 °C. A field strength of 67 or 83 kHz was employed for ^{1}H decoupling. The CP contact time was optimised to 1 ms and a relaxation delay of 3 s was employed. A 90° pulse of 4 and 3 µs was used for ^{1}H and ^{19}F, respectively. ^{1}H-undecoupled ^{19}F spectra with ^{19}F background suppression (^{19}F-bs), and ^{1}H-decoupled and ^{19}F-undecoupled ^{1}H-^{13}C CP (^{1}H-^{13}C{1H}) experiments were run using a 1.3 mm X/Y/H-F H13863 probe, at 60 kHz MAS spinning and 27 ± 3 °C. The CP contact time was optimised to 3 ms and a relaxation delay of 3 s was employed. A 90° pulse of 1.5 µs was employed for ^{1}H and ^{19}F. All spectra were referenced with respect to TMS.

EpC solution state NMR characterization was carried out in a Bruker Avance I spectrometer equipped with a 5 mm probe operating at frequencies of 499.69 MHz (^{1}H). Around 600 µL of dispersion (2 w/V%) in 99.9% D_2O (Sigma-Aldrich®) was pipetted into a 5 mm NMR tube at room temperature. Both phase-sensitive ^{1}H-^{13}C HSQC experiment with ^{1}H-^{13}C correlation via double inept transfer and ^{1}H-^{1}H COSY experiment with multiple quantum filter and gradient ratio for artifact suppression were acquired with 256 increments in the F1 dimension, 8 number of scans and a relaxation delay of 2s. 8k number of scans were acquired for DEPT135 experiment with decoupling and a pulse length of 11.25 µs.

Solution state NMR characterization of a 0.5 wt% dispersion of multi-6F-EpC was carried out in an Avance Bruker 800.23 MHz spectrometer equipped with a 5 mm inverse triple-resonance probe. Phase-sensitive ^{1}H-^{13}C HSQC experiment with ^{1}H-^{13}C correlation via double inept transfer was acquired at 293K with 126 increments in the F1 dimension, 8 number of scans and a relaxation delay of 3s. ^{1}H-^{1}H COSY experiment with multiple quantum filter and gradient ratio for artifact suppression was acquired at 278K with 256 increments in the F1 dimension, 12 number of scans and a relaxation delay of 2 seconds. Finally, DEPT135 experiment was acquired in a 600.18 MHz Avance Bruker spectrometer equipped with a cryoprobe with a total of 30k scans and a relaxation delay of 1s.

2. Results

2.1. Enzymatic synthesis of fluorinated cellobioses and cellodextrins

2.1.1. Nuclear magnetic resonance (NMR) characterization

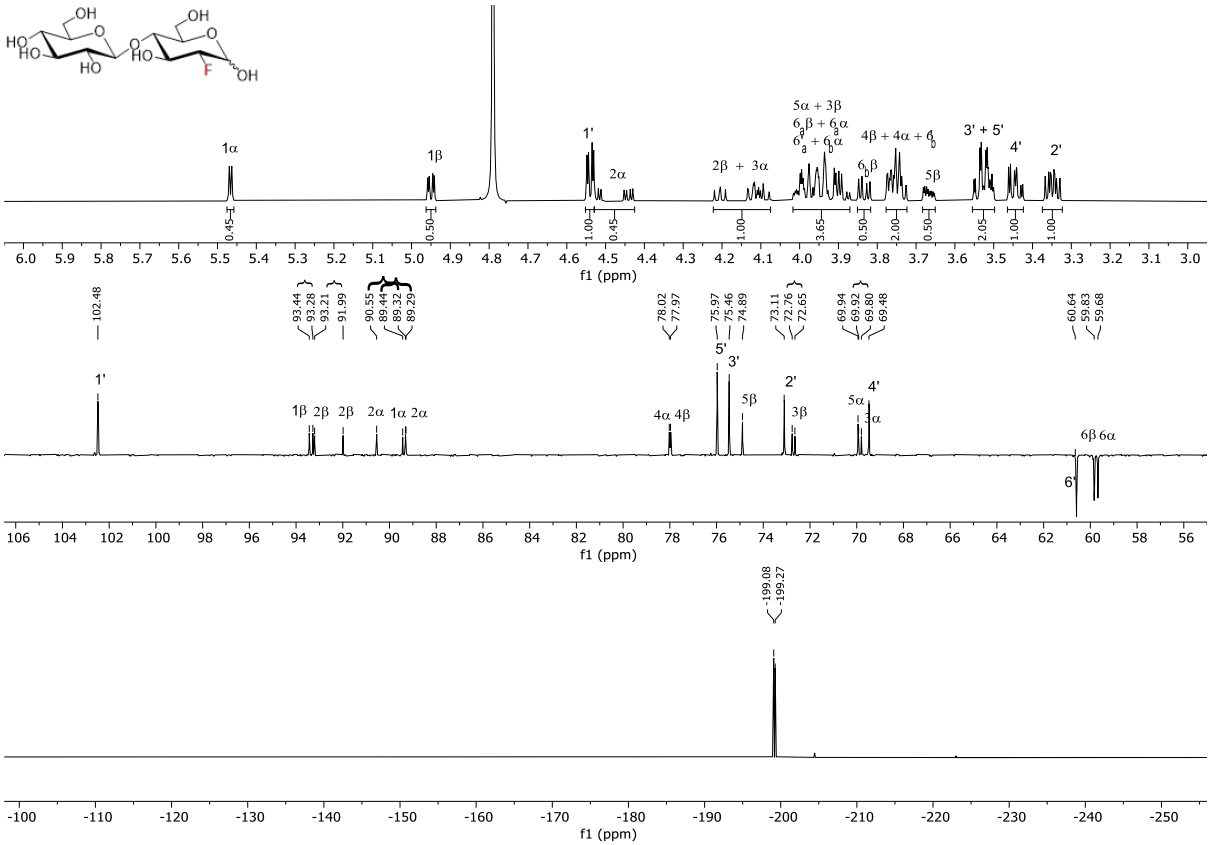


Figure S1. ^{1}H, ^{13}C and $^{19}F\{^{1}H\}$ NMR of 2-deoxy-2-fluoro-cellobiose (**1**).

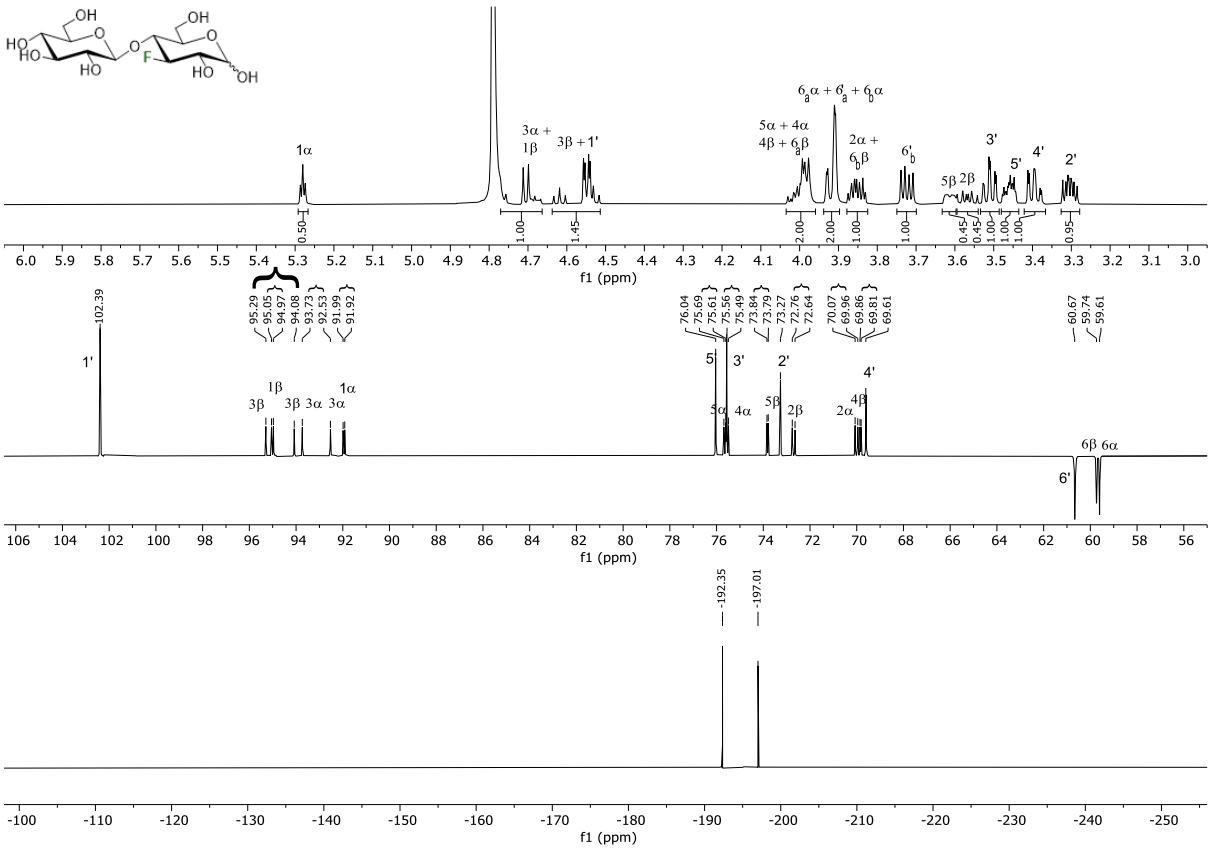


Figure S2. ^{1}H, ^{13}C and $^{19}F\{^{1}H\}$ NMR of 3-deoxy-3-fluoro-cellobiose (**2**).

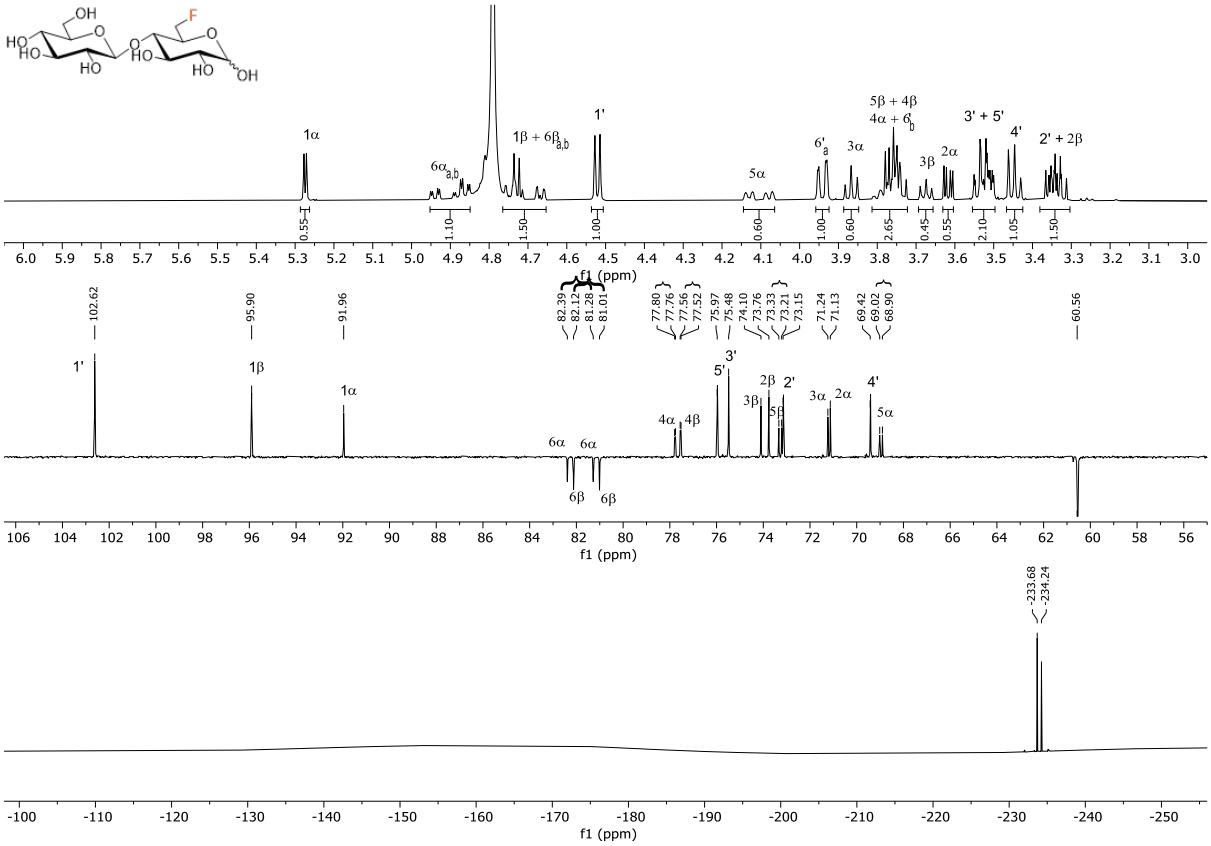


Figure S3. ^{1}H, ^{13}C and $^{19}F\{^{1}H\}$ NMR of 6-deoxy-6-fluoro-cellobiose (**3**).

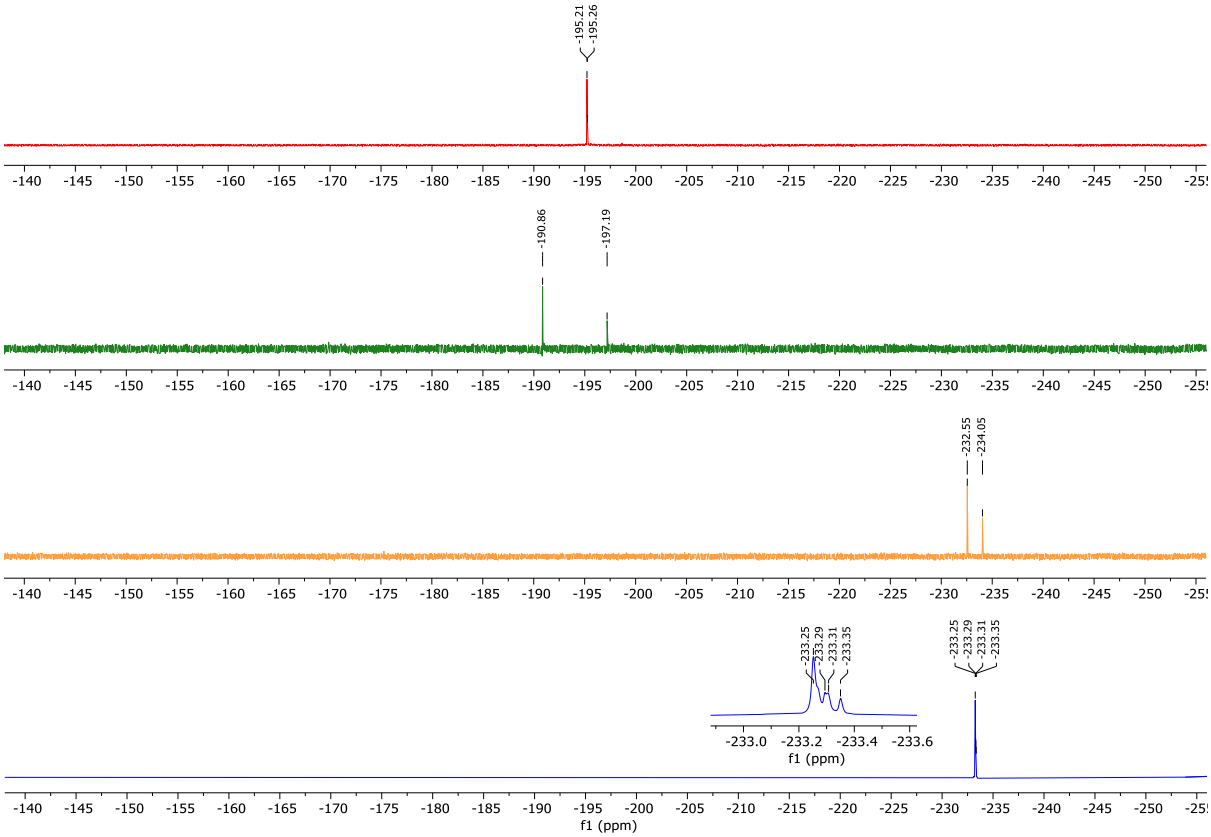


Figure S4. $^{19}F\{^{1}H\}$ NMR spectra (1 M NaOD) of 2-, 3- and 6-monofluorinated cellodextrins (2F-EpC, **4**; 3F-EpC, **5** and 6F-EpC, **6**) and multiply 6-fluorinated cellodextrin (multi-6F-EpC, **7**).

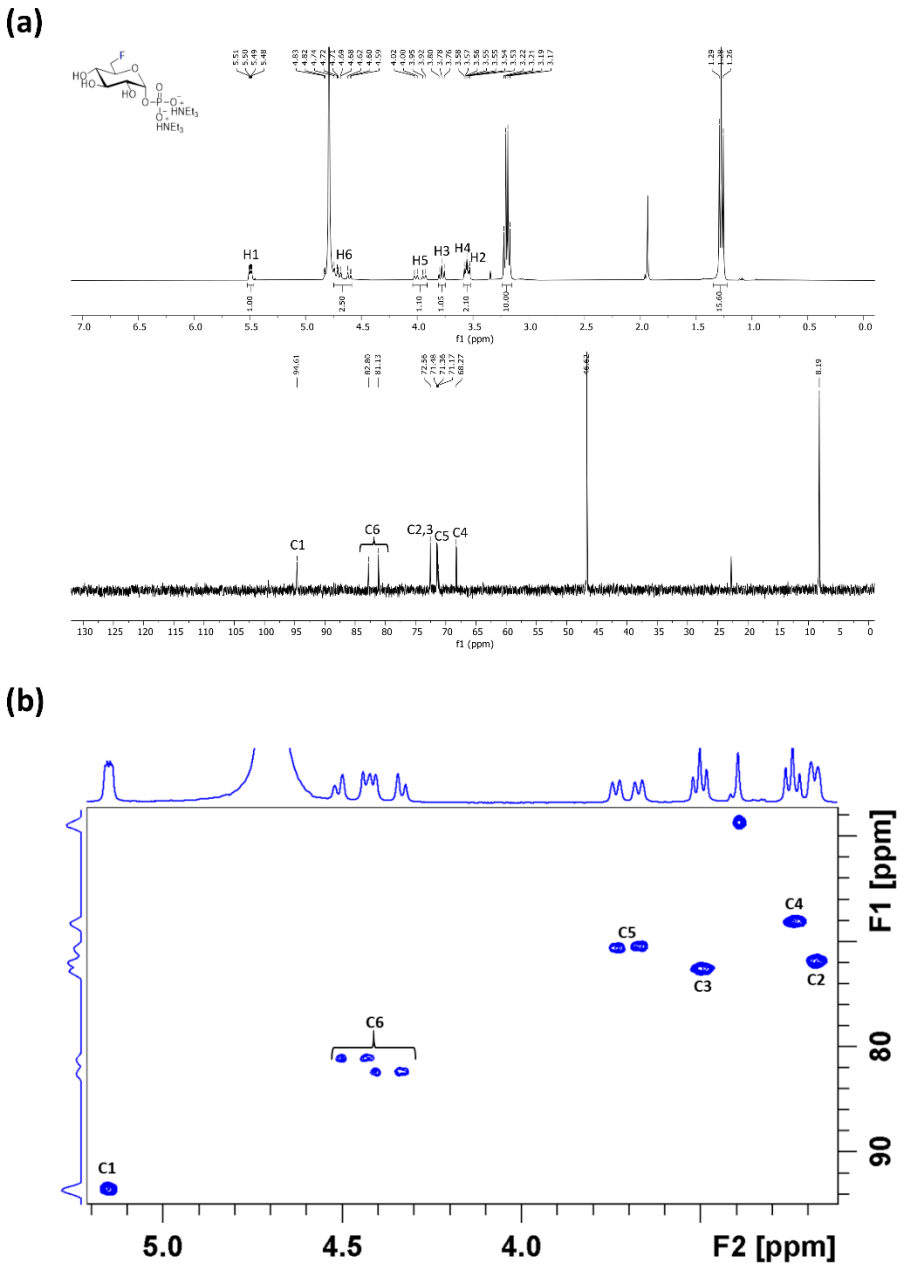


Figure S5. (a) ^{1}H (top) and ^{13}C (bottom) in D_2O, and **(b)** ^{1}H-^{13}C HSQC NMR spectra of 6-deoxy-6-fluoro-α-D-glucose 1-phosphate (triethylammonium salt) 2mM in Tris-d_{11} pH 7.4 (25 mM, NaCl 100 mM). Peak assignment is shown.

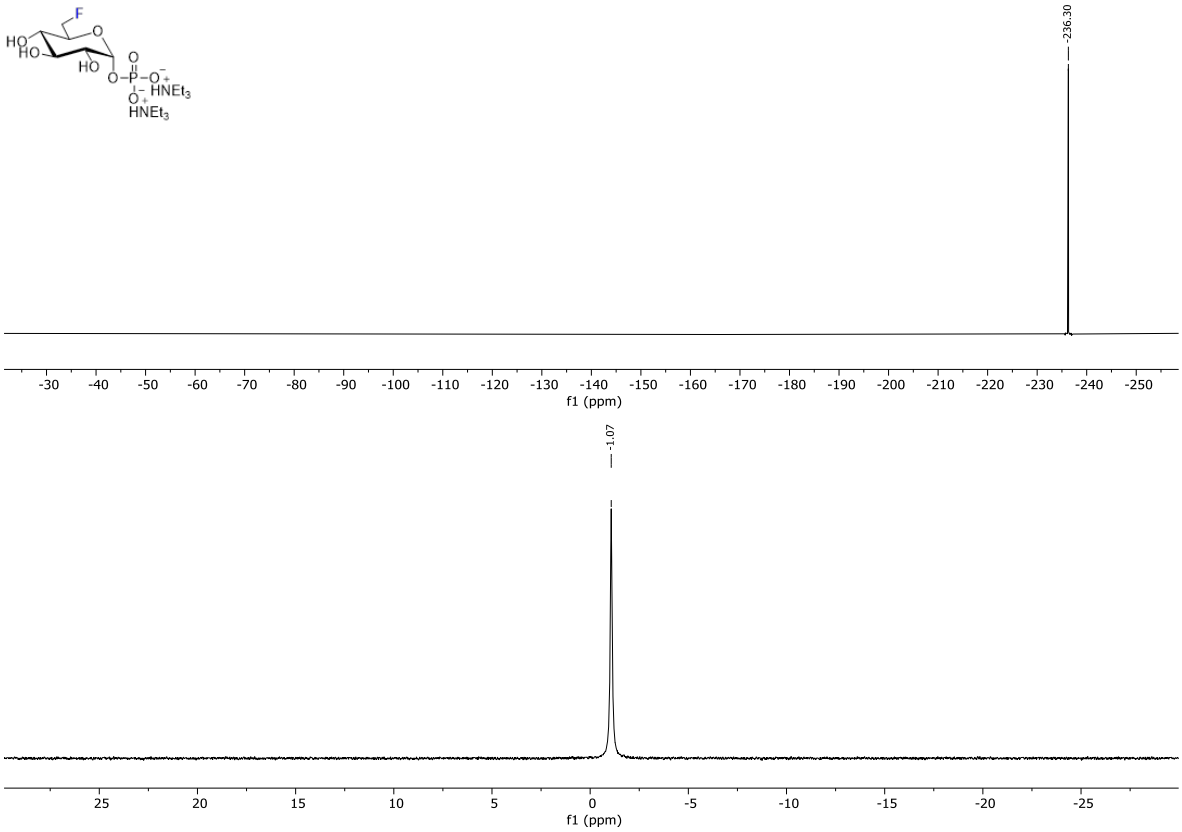


Figure S6. $^{19}F\{^{1}H\}$ (top) and ^{31}P (bottom) NMR of 6-deoxy-6-fluoro-α-D-glucose 1-phosphate (triethylammonium salt).

2.1.2. Matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS) characterization

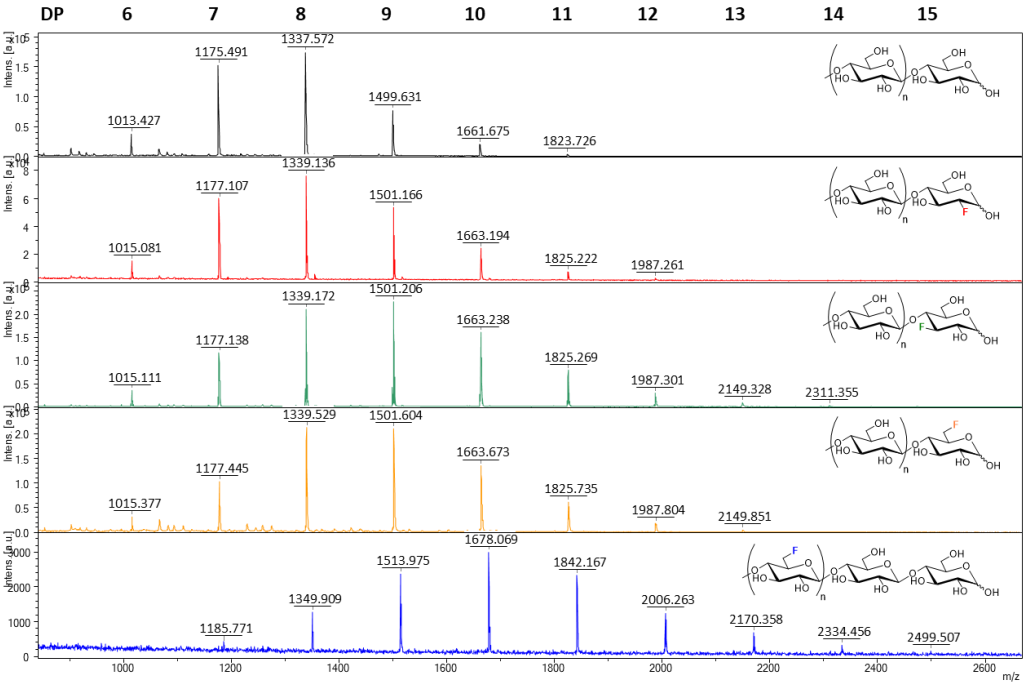


Figure S7. MALDI-TOF spectra comparison between enzymatically produced cellodextrin (EpC, **8**, black line), 2-, 3- and 6-monofluorinated cellodextrins (2F-EpC, **4**, red line; 3F-EpC, **5**, green line; and 6F-EpC, **6**, orange line) and multiply 6-fluorinated cellodextrin (multi-6F-EpC, **7**, blue line).

2.2. Morphological characterization

2.2.1. Transmission electron microscopy (TEM) characterization

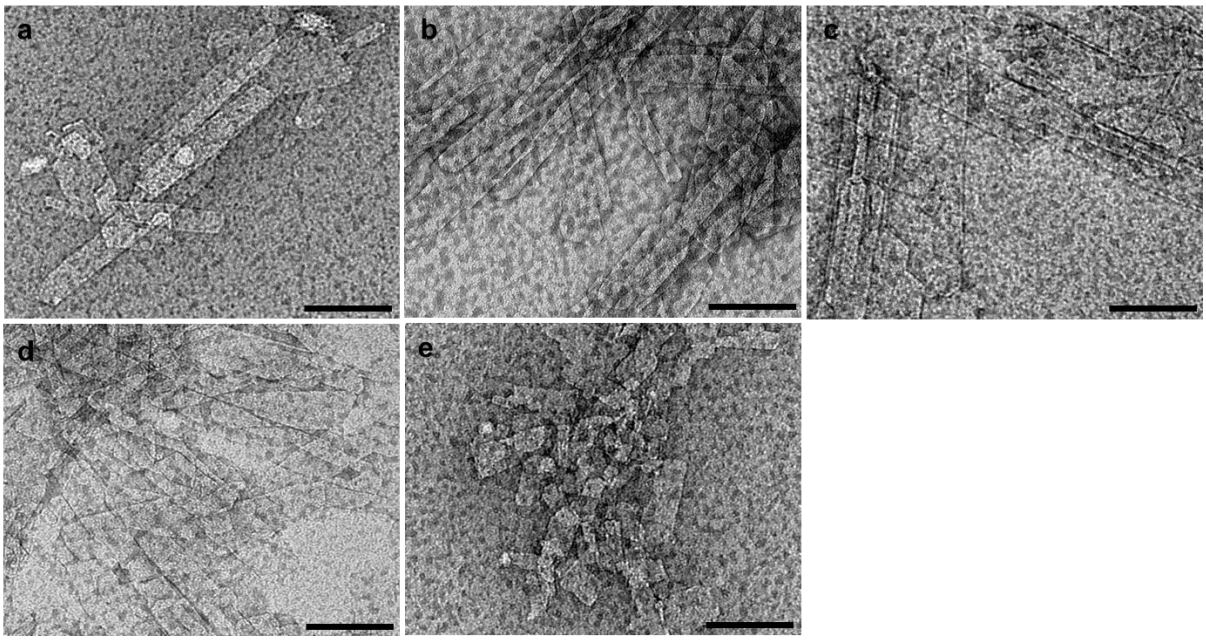


Figure S8. TEM images of enzymatically produced cellodextrin EpC (**a**) and enzymatically produced fluorinated cellodextrins 2F-EpC (**b**), 3F-EpC (**c**), 6F-EpC (**d**) multi-6F-EpC (**e**) negatively stained with 2% uranyl acetate. Scale bars correspond to 100 nm.

2.3. Long-range structural characterization

2.3.1. Powder X-ray diffraction (PXRD) patterns

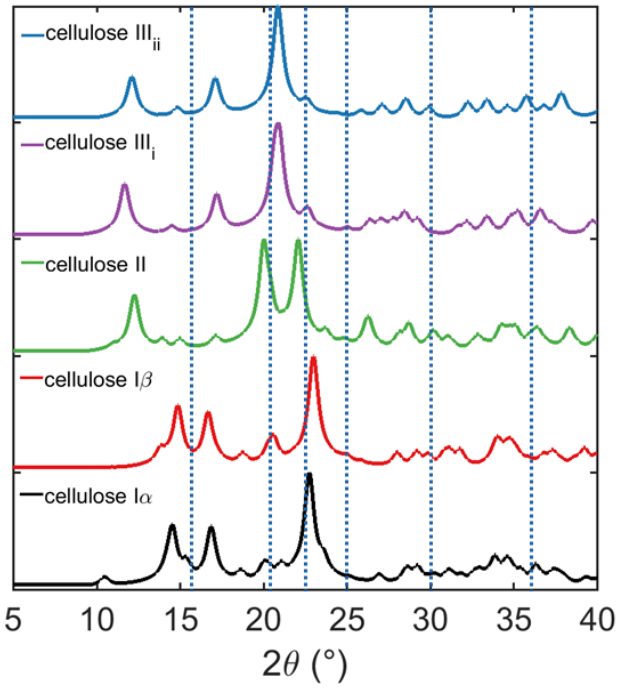


Figure S9. Predicted diffraction patterns for cellulose types I_α (black), I_β (red), II (green), III_I (purple) and III_{II} (blue). The blue dashed lines represent the experimental diffraction angles obtained for multi-6F-EpC (the diffraction pattern is reported in the main text, Figure 4b).

Table S1. Comparison of the predicted *2Θ*(°) and *d-spacing* values for cellulose II, III_I and III_{II} with the experimental values obtained for multi-6F-EpC (**7**). Peak indices refer only to cellulose II PXRD pattern (Figure 3).

	PEAK 1 (1$\bar{1}$0)		**PEAK 2 (110)**		**PEAK 3 (020)**	
	2Θ(°)	*d-spacing (nm)*	*2Θ*(°)	*d-spacing (nm)*	*2Θ*(°)	*d-spacing (nm)*
cellulose II	12.24	0.722	20.02	0.443	22.08	0.402
cellulose III_I	11.66	0.758	17.18	0.516	20.88	0.425
cellulose III_{II}	12.1	0.731	17.1	0.518	20.86	0.425
multi-6F-EpC	15.62	0.567	22.56	0.394		

2.4. Molecular characterization: local structure

2.4.1. Raman spectroscopy

Table S2. Raman spectroscopy band positions for EpC, 2F-EpC, 3F-EpC, 6F-EpC, and multi-6F-EpC samples.

Band no.	Sample band centre [cm^{-1}]				
	EpC	2F-EpC	3F-EpC	6F-EpC	multi-6F-EpC
1	337	337	338	340	-
2	354	354	354	354	351
3	379	379	379	379	370
4	398	-	-	399	-
5	415	415	416	415	416
6	-	-	-	-	437
7	447	448	448	447	-
8	462	463	463	463	-
9	-	-	-	-	480
10		491	487	493	495
11	519	519	520	519	-
12	576	576	576	576	-
13	637	-	-	-	-
14	-	695	-	-	-
15	844	-	-	844	845
16	867	-	-	867	867
17	896	897	897	896	896
18	-	-	-	-	924
19	969	969	973	969	975
20	1002	-	-	999	1003
21	1029	1030	-	1030	-
22	1064	1065	1064	1065	-
23	1097	1098	1098	1097	1088
24	1123	1122	1124	1122	1127
25	1147	1148	1147	1147	-
26	1235	1235	1232	1234	1239
27	1265	1265	1265	1265	1268
28	1314	1312	1312	1308	1304

29	1336	1337	1336	1336	-
30	-	-	-	1344	1340
31	1368	1361	1369	1362	-
32	-	1374	-	1375	1379
33	-	-	-	1396	-
34	1407	1404	1409	1414	1424
35	1462	1462	1461	1461	1451

2.4.2. Solid-state and solution-state NMR spectroscopy

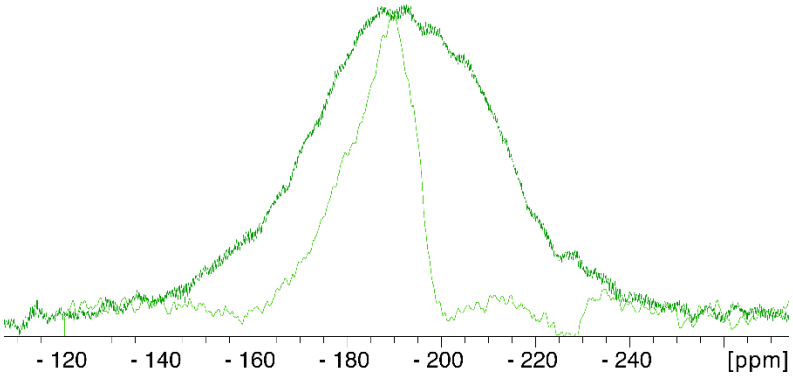


Figure S10. ^{19}F NMR spectrum of 3F-EpC powder acquired with (light green) and without (dark green) ^{1}H decoupling at 15 kHz and 60 kHz MAS rate, respectively, and 800 MHz ^{19}F frequency

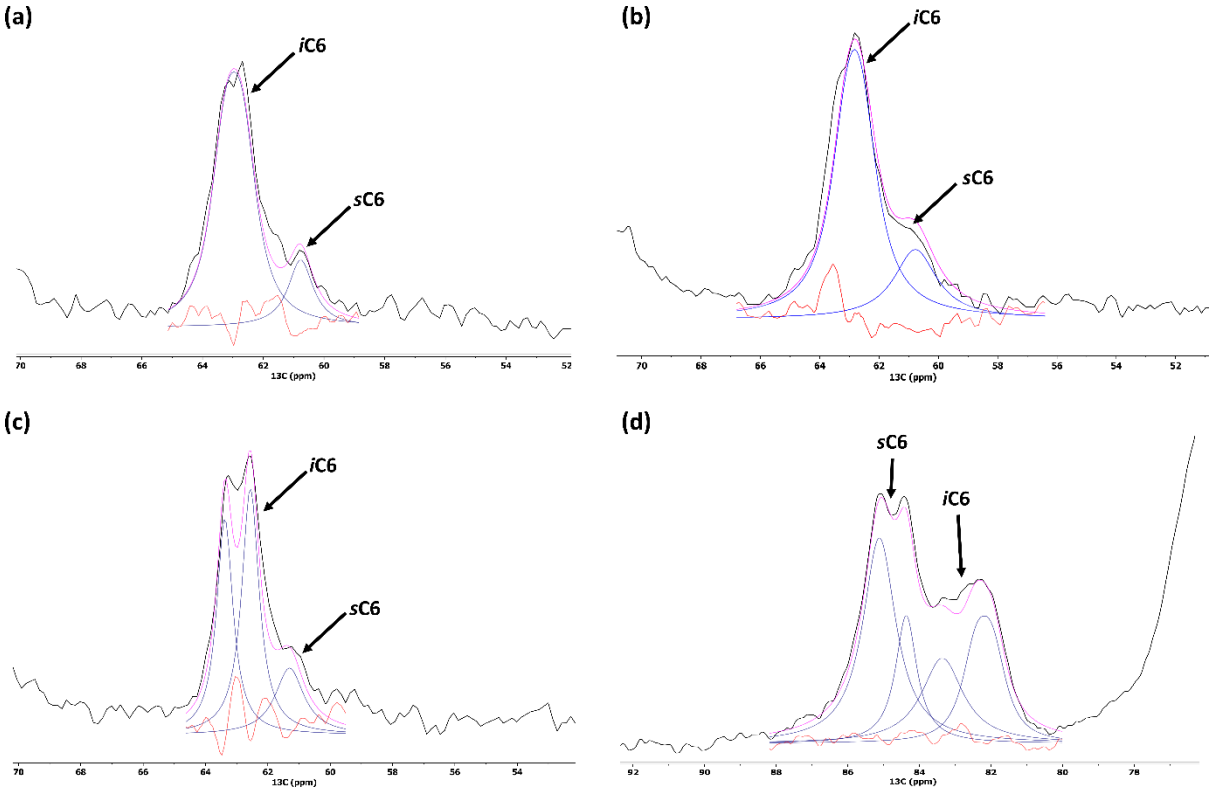


Figure S11. Spectral deconvolution of the *s*C6 and *i*C6 peaks of ^{1}H-^{13}C CP spectra of 2F-EpC (a), 3F-EpC (b), 6F-EpC (c) and multi-6F-EpC (d). The experimental and deconvoluted spectra are shown with black and blue lines, respectively. The sum and residuals spectra are shown in pink and red, respectively.

Table S3. Relative surface area (RSA) values derived from deconvolution of the sC6 and iC6 peaks of ^{1}H-^{13}C CP spectra of the three monofluorinated EpCs and multi-6-EpC powders. RSA were calculated as the ratio of the *s*C6 peak area against the total C6 area (*s*C6 + *i*C6).

	Relative surface area (%)
2F-EpC	16
3F-EpC	23
6F-EpC	20
multi-6F-EpC	54

The ^{1}H-^{13}C CP/MAS spectrum of multi-6F-EpC (**7**) was compared to all previously reported cellulose allomorphs to identify differences and similarities on ^{13}C chemical shifts (Tables S4, S5).

Table S4. ^{13}C chemical shifts reported in the literature for all the knowns cellulose allomorphs. * Experimental values obtained in-house for enzymatically produced cellulose (EpC).

Cellulose allomorph	**Chemical shift (ppm)**		
	C1	**C4**	**C6**
cellulose Iα (Glaucocystis)[12]	105	89.7, 88.8	65.3
cellulose Iβ (tunicin)[12]	105.7, 103.9	88.7, 88.0	65.5, 64.9
cellulose II[10]	107.0, 104.7	88.6, 87.4	62.9, 62.2
*cellulose II (EpC)**	107.0, 104.8	88.6, 87.5	62.9, 62.2
cellulose III_I[12]	104.8	87.8	62.3
Cell III_{II}[10]	106.5, 105.7	88.5, 87.3	62.5, 62.1

Table S5. ^{13}C chemical shifts of multi-6F-EpC (**7**) acquired in solid-state (top) and solution state (bottom) NMR.

Chemical shifts (ppm)					
Solid-state NMR					
C1	***s*C6**	***i*C6c**	***s*C2,3,4,5**	***i*C2,3,4,5**	**C6***
104.8	84.2, 83.5	81.9	75.2	73.1	62.6
Solution state NMR					
***s*C1**	***s*C6**	***s*C4***	***s*C2,3,4,5**	***s*C6***	
104.8	84.3, 83.3	79.9	76.0, 75.3	62.1	

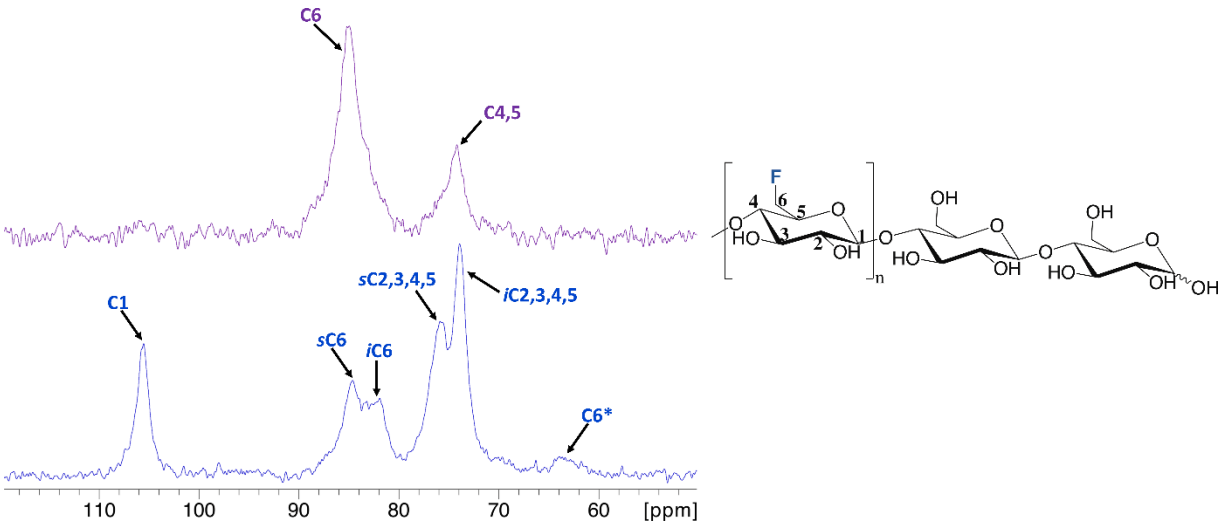


Figure S12. Comparison of the ^{1}H-^{13}C CP (blue) and $^{19}F\{^{1}H\}$-^{13}C CP (purple) NMR spectra of multi-6F-EpC powder acquired at room temperature, 15 kHz MAS spinning and a ^{13}C frequency of 212.5 MHz.

The high-resolution solution NMR spectra (COSY and HSQC, Figure S13a) obtained for a diluted dispersion of EpC (**8**) enabled the detailed assignment of the two anomeric spin systems and the internal and non-reducing terminal protons. On the other hand, the higher presence of chemical environments in multi-6F-EpC (**7**) complicated the full spectral assignment (COSY and HSQC, Figure S13b), but we could still assign the β-spin system (belonging to the cellobiose unit) and the 6-fluorinated non-reducing terminal (*nr*).

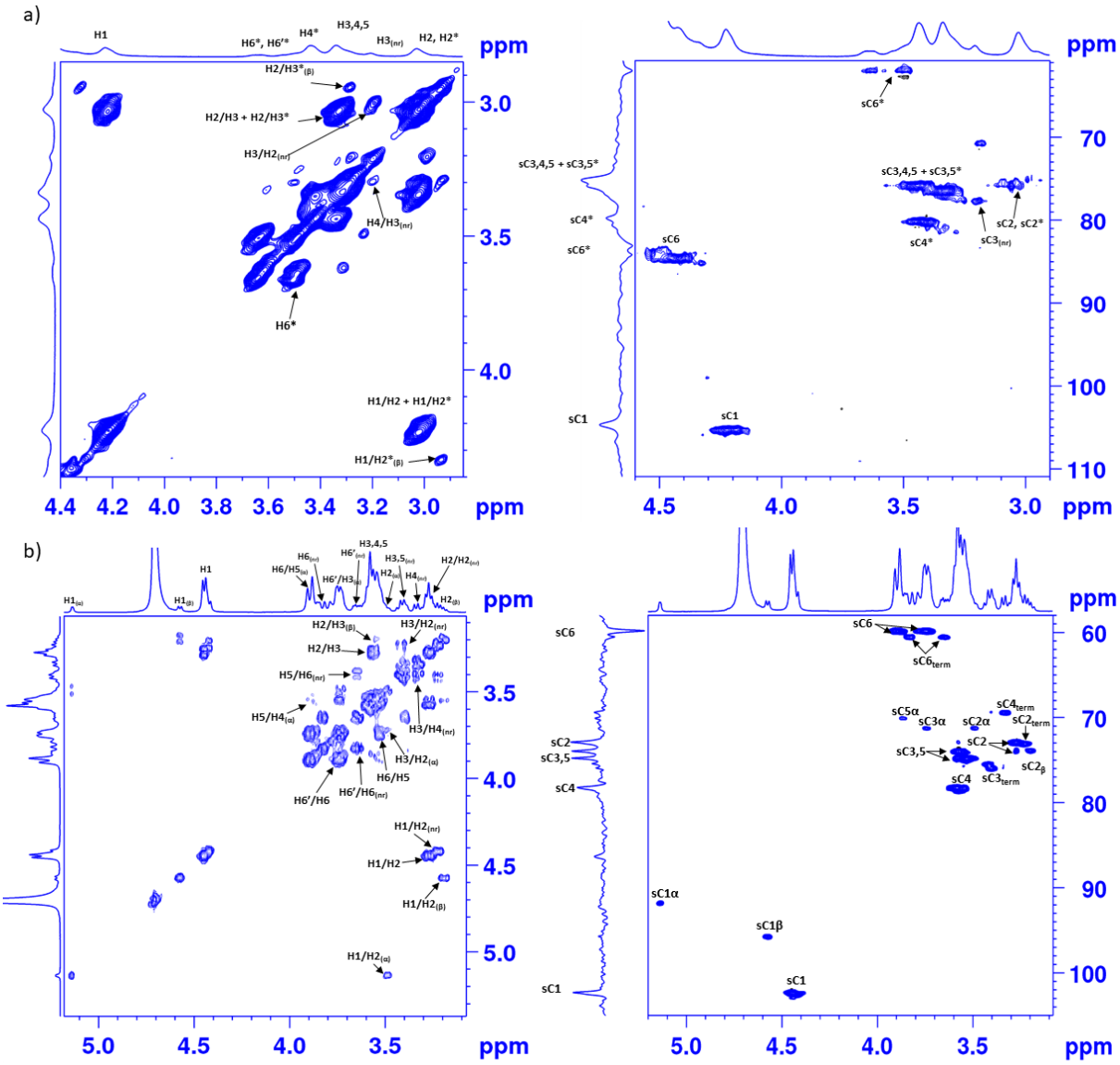


Figure S13. a) COSY (right) and HSQC (left) of a 0.5 wt% dispersion of multi-6F-EpC (**7**). b) COSY (right) and HSQC (left) of a 2 wt% dispersion of EpC (**8**). In the cross-peak nomenclature, α- and β- refer to the corresponding anomeric proton spin system on the reducing end, whereas *nr* indicates the non-reducing terminal.

To characterize the internal dynamics of these materials, ^{1}H-^{13}C CP/MAS NMR experiments at varying contact times were carried out for EpC and multi-6F-EpC (Figure S14). Very similar build-up curves were obtained for all carbon peaks in both powdered samples, which suggest that, at the molecular level, the CP-observable domains (i.e. rigid) of EpC and multi-6F-EpC nanofibrils present very similar dynamics.

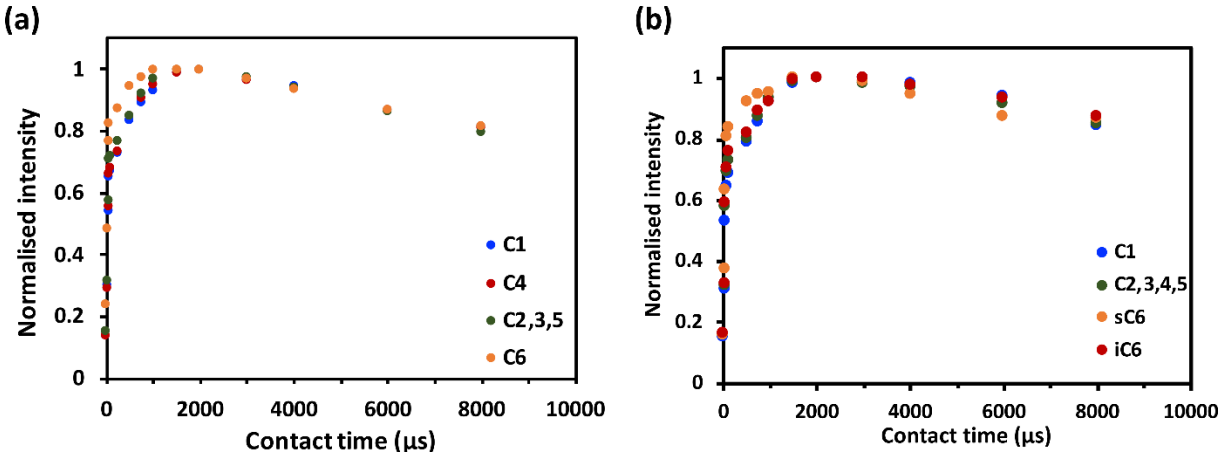


Figure S14. ^{1}H-^{13}C CP build-up curve of EpC (a) and multi-6F-EpC (b) powders acquired at 25 °C, 12 kHz MAS rate and a ^{13}C frequency of 100.6 MHz.

References

(1) O'Neill, E. C.; Pergolizzi, G.; Stevenson, C. E. M.; Lawson, D. M.; Nepogodiev, S. A.; Field, R. A. Cellodextrin Phosphorylase from Ruminiclostridium Thermocellum: X-Ray Crystal Structure and Substrate Specificity Analysis. *Carbohydr. Res.* **2017**, *451*, 118–132. https://doi.org/10.1016/j.carres.2017.07.005.

(2) Arai, M. A.; Yamaguchi, Y.; Ishibashi, M. Total Synthesis of Agalloside{,} Isolated from Aquilaria Agallocha{,} by the 5-O-Glycosylation of Flavan. *Org. Biomol. Chem.* **2017**, *15* (23), 5025–5032. https://doi.org/10.1039/C7OB01004D.

(3) Withers, S. G.; Percival, M. D.; Street, I. P. The Synthesis and Hydrolysis of a Series of Deoxy- and Deoxyfluoro-α-d- "Glucopyranosyl" Phosphates. *Carbohydr. Res.* **1989**, *187* (1), 43–66. https://doi.org/https://doi.org/10.1016/0008-6215(89)80055-2.

(4) Zhu, J. S.; McCormick, N. E.; Timmons, S. C.; Jakeman, D. L. Synthesis of α-Deoxymono and Difluorohexopyranosyl 1-Phosphates and Kinetic Evaluation with Thymidylyl- and Guanidylyltransferases. *J. Org. Chem.* **2016**, *81* (19), 8816–8825. https://doi.org/10.1021/acs.joc.6b01485.

(5) Macrae, C. F.; Bruno, I. J.; Chisholm, J. A.; Edgington, P. R.; McCabe, P.; Pidcock, E.; Rodriguez-Monge, L.; Taylor, R.; Van De Streek, J.; Wood, P. A. Mercury CSD 2.0 - New Features for the Visualization and Investigation of Crystal Structures. *J. Appl. Crystallogr.* **2008**, *41* (2), 466–470. https://doi.org/10.1107/S0021889807067908.

(6) Nishiyama, Y.; Sugiyama, J.; Chanzy, H.; Langan, P. Crystal Structure and Hydrogen Bonding System in Cellulose I(Alpha) from Synchrotron X-Ray and Neutron Fiber Diffraction. *J Am Chem Soc* **2003**, *125* (47), 14300–14306. https://doi.org/10.1021/ja037055w.

(7) Nishiyama, Y.; Langan, P.; Chanzy, H. Crystal Structure and Hydrogen-Bonding System in Cellulose Ibeta from Synchrotron X-Ray and Neutron Fiber Diffraction. *J Am Chem Soc* **2002**, *124* (31), 9074–9082.

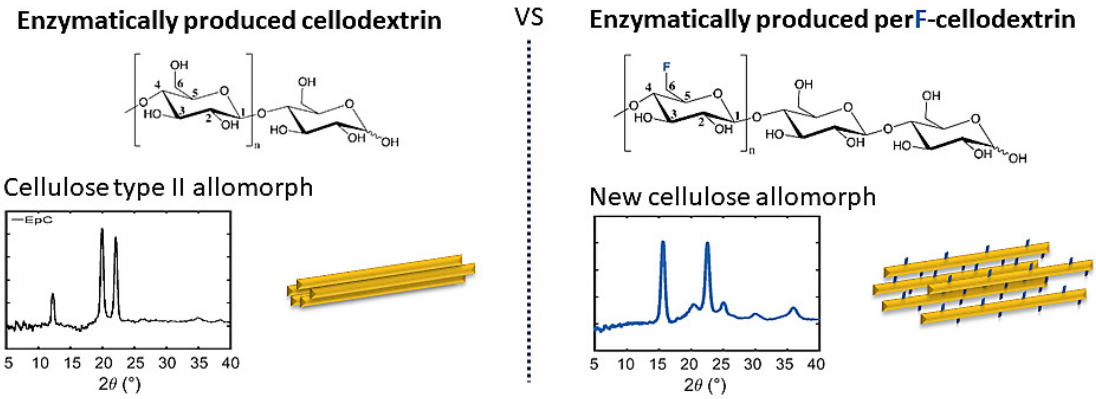
(8) Langan, P.; Nishiyama, Y.; Chanzy, H. X-Ray Structure of Mercerized Cellulose II at 1 Å Resolution. *Biomacromolecules* **2001**, *2* (2), 410–416. https://doi.org/10.1021/bm005612q.

(9) Wada, M.; Chanzy, H.; Nishiyama, Y.; Langan, P. Cellulose IIII Crystal Structure and Hydrogen Bonding by Synchrotron X-Ray and Neutron Fiber Diffraction. *Macromolecules* **2004**, *37* (23), 8548–8555. https://doi.org/10.1021/ma0485585.

(10) Wada, M.; Heux, L.; Nishiyama, Y.; Langan, P. X-Ray Crystallographic, Scanning Microprobe X-Ray Diffraction, and Cross-Polarized/Magic Angle Spinning 13C NMR Studies of the Structure of Cellulose III(II). *Biomacromolecules* **2009**, *10* (2), 302–309. https://doi.org/10.1021/bm8010227.

(11) Wojdyr, M. {\it Fityk}: A General-Purpose Peak Fitting Program. *J. Appl. Crystallogr.* **2010**, *43* (5 Part 1), 1126–1128. https://doi.org/10.1107/S0021889810030499.

(12) Wada, M.; Heux, L.; Isogai, A.; Nishiyama, Y.; Chanzy, H.; Sugiyama, J. Improved Structural Data of Cellulose IIII Prepared in Supercritical Ammonia. *Macromolecules* **2001**, *34* (5), 1237–1243. https://doi.org/10.1021/ma001406z.



ChemRxiv_Khimyak_Field_graphic abstract.tif (129.25 KiB) view on ChemRxiv • download file