



HAL
open science

Combining computational and experimental studies for a better understanding of cellulose and its analogs

Yoshiharu Nishiyama

► **To cite this version:**

Yoshiharu Nishiyama. Combining computational and experimental studies for a better understanding of cellulose and its analogs. David C. Baker. *Advances in Carbohydrate Chemistry and Biochemistry*, 80, Elsevier, pp.1-14, 2021, *Advances in Carbohydrate Chemistry and Biochemistry*, 10.1016/bs.accb.2021.10.002 . hal-04688311

HAL Id: hal-04688311

<https://hal.science/hal-04688311v1>

Submitted on 4 Sep 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Chapter 1. Combining computational and experimental studies for a better understanding of cellulose and its analogs

Yoshiharu Nishiyama

Université Grenoble Alpes, CNRS, CERMAV, Grenoble, France

published in *Advances in Carbohydrate Chemistry and Biochemistry vol 80*, 2021, David C. Baker Ed.
ISBN: 978-0-12-824628-3

1 Introduction

Cellulose and hemicelluloses are the most abundant components of biomass, and as such are the most important organic materials in terms of stock and production. Thus, their efficient transformation into useful materials is crucial for the development of sustainable materials. Since cellulose is resistant to very harsh treatments, it can be used, and has been used, without much understanding of processes at the molecular level. However, if we want to further reduce the energy input to attain equivalent or better functionality, or to extend further property range to replace materials from nonrenewable resources, a deep understanding of the molecular behavior of cellulose is needed. This includes interactions with other molecules as well as an understanding of its intrinsic properties as a function of its type of molecular assembly.

In Volume 67 of *Advances* published in 2012, [1] A. French discussed the advances and various options of computational chemistry approaches and the use of crystallographic data, mainly from single-crystal studies. The comparisons to the well-established crystal structures help in rationalizing the choice of assumptions used in modeling; furthermore, these comparisons are aids to further understand the stabilization factors of molecular associations.

In the meanwhile, a huge amount of reproducible experimental data is available, and a wide range of new experimental data is becoming more and more accessible. Those experimental data are very often analyzed only in a superficial way, but such data have the potential to further guide us in our understanding of carbohydrate structures and behavior, especially when combined with molecular modeling.

In a naïve view, having all physical processes understood, and with sufficient computational power at hand, we should be able to simulate molecular behaviors to reproduce reality that cannot be captured by experimental observations alone. However, even the very simple crystal structure of cellulose has not been predicted from scratch. To simulate crystals, the model must be guided by experimental constraints such as unit cell dimensions or starting coordinates. In a more complex system, there are often no experimental observations to contest the event in a simulation box. Yet who can trust a method that cannot predict a simpler system?

One can dream of establishing theoretical approaches that withstand experimental scrutiny, which can be then deployed for detailed insights that are not experimentally accessible. This is a goal that will not be realized anytime soon, as there are still always gaps between prediction and experimental observations whenever they are available, and the progress to merge the experimental and numerical world requires a significant amount of effort.

With modern computers and various force-fields, simulation of complex structure and behavior at the molecular level is becoming more and more accessible. To guarantee the validity of each theoretical approach to the real system, further efforts to bridge experimental data to the simulated systems are needed. In the present paper, I will draw my perspective of how computational efforts should be combined with well-designed experiments to be directly compared with molecular modeling.

2 From primitive models to atomistic details

2.1 Toward experimental structures with higher precision

Molecular models were used even before the appearance of computers to describe plausible structures from the very beginning of structural studies. The unit cell can be calculated from the positions of a small number of X-ray diffraction spots, and approximating cellulose as a rectangular slab, diverse allomorphs of cellulose can be described by their arrangement, qualitatively explaining the relative intensities appearing on X-ray fiber diffractions. A thorough historical account of X-ray structure analysis of cellulose has been recently given by Zugenmaier. [2] Not only pure cellulose, but cellulose complexed with a plethora of guest molecules has been studied in the first decades after the discovery of X-ray diffraction (1912), one of the famous examples being the temperature/concentration diagram of cellulose with sodium hydroxide. [3]

The description of molecules at low resolution, however, does not give direct insight on the molecular interactions. So more realistic molecular models are sought. Indeed once the rough structure of glucose residues is known, it seemed obvious that the hydroxy groups should be engaged in hydrogen bonding. But how? Fiber diffraction shrinks the three-dimensional diffraction information into two-dimensions, and the peak broadening due to the limited crystal size leads to overlapping of the peaks at higher resolution. Thus the detail of the

structure cannot be directly deduced from these data.

The emergence of computation power in the 1970s allowed relatively large searches of potential structures to find the one that agrees the best with the observed diffraction data, and realistic molecular models have been proposed. However, this trial-and-error approach always leaves some doubts. Even when reasonable agreement with a set of intensity data is obtained, we cannot be free of doubt as we might have missed a structure model, which in fact agrees better with data and thus represents the “real” structure.

A breakthrough appeared when fiber diffraction data from larger crystals were collected for major cellulose and chitin allomorphs, in combination with data reduction tools. Indeed, diffraction data from highly oriented and highly crystalline β -chitin already had been published in 1969 by Blackwell, [4] but the experiment was not set up for collecting data to high resolution. Since the beginning of the present century, sample preparation and programs to extract high-resolution data and wide-area X-ray detectors became available leading to a series of high-resolution fiber diffraction studies. [5–9]

When individual diffraction intensities could be evaluated to higher resolution, some structural details can be directly obtained from the diffraction data by simple Fourier synthesis. This gives much more confidence to the basic structure compared to the previous trial-and-error approaches. Furthermore, neutron diffraction that became available to the fiber community [10] in the 1990s, allowed direct localization of hydrogen atoms implicated in hydrogen bonding.

The reliability of experimental structure from fiber diffraction primarily depends on the size of the crystal and unit cell. If the unit cell is small and the size of the individual crystals are large, the structure can be directly determined by X-ray/neutron fiber diffraction with very high confidence, such as β -chitin that has only one monomer as the asymmetric unit. [8,11]8,11 In this case even the water molecules in the di-hydrate [12] could be directly located, and the whole hydrogen-bond network could be determined.

3 The limit of experimental determination

3.1 Limitation in resolution

This is, however, not the end of the story. Not all allomorphs give high-resolution fiber diffraction data, and not all structures are static or simple. If the structure contains inherent disorder, it is almost impossible to locate the various positions from the experimental data alone.

Furthermore, even for neutron diffraction, the effect of small difference of hydrogen position is subtle on the overall diffraction intensity distribution, and thus tautomers with only small differences in location cannot be clearly distinguished. For example, a revisit of the structures of cellulose II and III_I by using molecular modeling and density functional theory all agreed against the experimentally determined structures, suggesting an alternative hydrogen-bond

system. [13]

3.2 Inherent structural disorder and dynamics

Cellulose and chitin are among the most simple carbohydrate polymers, and yet they present difficulties when it comes to the noncrystalline regions, which include crystal surface and disorganized or vitrified regions. As homopolymers, they tend to crystallize, but the long-chain topology impedes reorganization and leaves a large fraction of the samples in the noncrystalline state. This is inherent with polymers that tend to be immobilized in a metastable state due to their topological constraints, and it is not a special property of cellulose and its analogs. The noncrystalline fraction of chain segments can take various conformations and intermolecular interactions. The complexity of the description of the ensemble of noncrystalline cellulose is incomparable to that of crystalline cellulose.

Many small molecules that strongly interact with cellulose or its analogs are important components of their solvents. These molecules often form complex crystals with cellulose with different levels of order. Typical historical examples are aqueous alkali components such as sodium hydroxide and alkyl amines, including ammonia and ethylenediamine, and water.

As an extension to the series of fiber diffraction structural refinements, the crystal structures of many amine complexes have been revisited. For the amine complexes, in particular, the crystal entity of the starting structure is not strongly modified, and the amine acts to intercalate the structure, thus leading to high-resolution data if the starting sample has a large crystal size and a high degree of orientation. [14–18]

In many cases, the amine acts as strong hydrogen-bond acceptor, but not as donor, and presents local disorder. In these cases, assistance with molecular modeling turned out to be useful. In the *alpha*-chitin–hydrazine complex, the nitrogen atoms were directly located in the Fourier maps from fiber diffraction.

However, the hydrogen positions involved in the hydrogen bonds could not be determined experimentally. The DFT method was used to check different scenarios of hydrogen bonds given the chitin structure and the nitrogen positions of the hydrazine. The O-3 is donating hydrogen bonds to either (a) a nitrogen atom of hydrazine or (b) to the ring oxygen O-5 of the neighboring residue. The energy of the DFT-optimized structure is significantly lower for scenario (a), but further molecular dynamics simulations suggest that hydrogen bonds can be quickly exchanging in the structure. Even the two nitrogen atoms of the hydrazine can be flipping relative to each other in the short simulation time. When fiber diffraction data is not conclusive in determining the structure on its own, a combination with DFT is also proving to be powerful. [19]

More recently, N-methylmorpholine N-oxide (NMMO), known as a solvent for the Lyocell process and some types of emerging cellulosedissolving ionic liquids, is also found to form cellulose-complex crystals. [20–22] However, those complex crystals have a large unit cell, and although sharp diffraction peaks are observed, even the unit cells are not yet determined for most of them. Further

analyses are expected to follow in the coming years with the wider availability of microfocus X-ray beams that allows recording of oriented patterns, thus enabling the determination of the unit cell. [23]

4 Modeling crystals and prediction of experimental response

Since neither an experimental nor a purely theoretical approach would give a satisfactory description of the complex system, we need to combine the two approaches. Experimental data contribute to validate some aspects of the numerical approaches, while simulation helps interpret the experimental data.

4.1 Force-field modeling and first principles calculations

To get better insight on the noncrystalline regions, we need a reliable model that can at least reproduce crystalline structures. The modeling approach, in general, consists of finding a rule based on existing observations and its extrapolation into the unknown.

A classic force-field based modeling consists of sets of more or less simple rules acting between a combination of chemically bonded and nonbonded atoms, mimicking a real system. This set of rules is called a force field. The rules are defined in a way so as to reproduce at best a set of observations such as the density of a liquid made up of simple molecules, the heat of evaporation, or its spin–spin coupling constants directly linked to the conformation. Those experimental values are training sets and lead to good reproduction of related properties. However, properties based on different physical processes are not necessarily well reproduced.

A typical example can be seen in the Lennard–Jones function parameters describing the nonbonded interaction in a series of force-field versions of GROMOS. [24] In the earlier version, crystal lattice constants were used as training sets, and these reproduced the unit cells of cellulose allomorphs fairly well. In a more recent version, the training sets were liquid/gas phase-change data, and the simulations of cellulose allomorphs gave systematically larger unit cells. This is in part because the real interactions energy profile does not follow the simple Lennard–Jones function, and the highly compact state and dilated state cannot be described by the same parameters. [24]

Torsional parameters are another delicate set of parameters to consider, again since the energy profile of torsion is dependent on many factors and does not necessarily have a simple shape. The ad hoc parameters are defined to favor certain conformations to reproduce quantum calculations or observations, but their reliability in the condensed phase is not certain. In GROMOS, these torsion parameters had to be modified to reproduce experimentally observed primary alcohol conformations of cellulose. [25] Thus extrapolation to an unknown system is still problematic.

Compared to force-field methods, quantum chemical calculations are expected to be more universal. Indeed, density functional theory (DFT) with dispersion correction proved to be powerful in reproducing the experimental unit cell of cellulose allomorphs when sufficiently accurate crystallographic structural data are given as the starting coordinates. [26] In this case the dispersion interaction due to the synchronized polarization is taken into account empirically, in a similar way as the attraction term of the Lennard–Jones potential. This additional consideration with empirical correction improves the agreement of simulated unit cells of molecular crystals without compromising the relatively universal nature of quantum chemical calculations. Still, DFT calculations come with a cost of much heavier numerical calculation since the electron density distribution has to be recalculated at each change in atomic position.

4.2 Simulation of wave scattering

Scattering of a wave such as that of an X-ray, neutron or electron reflects directly the spatial distribution of scattering length density. The physical origin and form of scattering length depends on the type of wave, but can be calculated in a straightforward way. The simplest is by using tabulated atomic scattering factors that assume a spherical distribution around the center of the atom, which is an approximation good enough for currently available diffraction data on cellulose and its analogs.

Depending on the sample symmetry or texture, the scattering data can have one- to three-dimensional character, with increased information with higher dimensions. The simplest case is that of a spherically averaged sample, in which case the average scattering can be calculated by the Debye formula, from the pair distribution function of an explicit model. [27]

In many cases, there is a cylindrical symmetry around the chain direction. The scattering for such symmetry can be calculated from a two-dimensional pair–distance distribution. [28] If fibrils with fiber symmetry are further organized with more complex texture, this can be taken into account with a relatively simple transformation.

Without spherical or cylindrical symmetry, the intensity distribution can only be treated as three-dimensional and consequently becomes much more challenging

For small molecules where single crystals are available, nonspherical electron density distribution is experimentally accessible, and can be compared to quantum mechanical calculations. [29]

4.3 Molecular vibration

Molecular vibrations are directly reflecting the bond rigidity and can be obtained from the normal mode analysis of the energy landscape close to the energy minima. The normal mode analysis of molecular systems are routine and are implemented in many software suites. However, up to now, no theoretical calculation is able to quantitatively reproduce the vibrational spectra of

cellulose or carbohydrates in general.

Vibrational spectroscopy includes infrared and Raman spectroscopy, sum frequency generation spectroscopy, and neutron inelastic scattering. The spectral band positions of vibrational spectra, usually expressed as cm^{-1} , correspond to the frequencies of the normal modes of vibration and are common to all types of vibrational spectroscopy and depend just on the structure. However, the number of bands corresponds to the number of freedoms, and thus three times the number of constituent atoms of the system, leading to a huge number of modes, making it impossible to find correspondence with the experimental spectra. The density of normal modes on the frequency scale is in fact the phonon density of states, which is directly accessible by neutron inelastic scattering [30]; however, such experimental data are still scarce.

For other types of vibrational spectroscopy, the spectral intensity of each band depends on the interaction mechanism with the probe-transition dipole moment for infrared absorption cross-section or change in polarizability for Raman scattering cross-section. Those also can be calculated by quantum mechanical calculations. Several program packages integrate such aspects that were recently applied to amino acids. [31]

Infrared spectroscopy and Raman spectroscopy are widely available in many laboratories, and routinely reported spectra are reproducible including those at low temperature [32] or with isotopic replacement. The intensity calculation would help the validation of calculation and band assignments, but such studies are still in their infancy, and the simulated infrared spectra of the OH stretching region of cellulose [33] is still quite different from that from experimental data. In most studies in the last decade, only the frequencies of the vibrational modes are compared to those from experimental studies.

Sum frequency generation (SFG) spectroscopy, first developed as a surface-sensitive technique, is emerging to characterize crystalline cellulose [34–36] and its organization at higher levels in the submicrometer length scale. [37–39] It is based on a two-photon process combining typically infrared absorption data and Raman shifts, and is sensitive to both the transition dipole moment and polarizability tensors. Theoretical calculations of spectral intensities of each of the vibrational modes are crucial to exploit such data for further molecular understanding.

4.4 Nuclear magnetic resonance

Chemical shifts in nuclear magnetic resonances (NMR) spectroscopy reflects the magnetic shielding by the electron cloud around the nuclei, and is again directly linked to the electron distribution that can be calculated by density functional theory. Typical agreement between calculated chemical shifts and experimental values for ^{13}C are about 2ppm when the hydrogen positions are optimized from the experimental structure. This is of the order of the influence of conformation of primary hydroxyl groups, but much larger than experimental precision. [40]

4.5 Elastic tensors

The elastic modulus of cellulose crystals as an indicator of ultimate material properties has been of interest and has been studied both theoretically and experimentally, but both suffer with inherent problems of unverifiable assumptions.

For example, the classic way to measure lattice elasticity is to apply a load on a bundle of fibers while measuring the lattice strain from the displacement of X-ray diffraction peak positions. In general, it assumes that all load is transferred to the crystalline domain as pure tensile stress, and only the crystalline domain is contributing to the Bragg spots. Given the very thin microfibril structure of native cellulose, where considerable amount of surface chains are aligned to the crystalline core, those assumptions are certainly oversimplifying. The most frequently used sample is a bundle of ramie fibers in which the typical crystal size is about 6 nm, where about 40% of the chains are on the surface. Despite the high reproducibility of the experiment, the interpretation relies much on the nontrivial mechanical models.

Since DFT with dispersion corrections can reproduce the cellulose crystal structures with relatively high fidelity, one could expect the elastic properties at the equilibrium position to be relatively well reproduced. With such modeling, not only the Young modulus along the chain, but also the whole elastic tensor can be estimated, [41, 42] and both have been calculated mainly on cellulose I.

Experimentally, the elastic tensor is not easily accessible, but a number of directly related relations can be obtained, including the compressibility, measured under hydrostatic pressure, and the strain tensor, measured under uniaxial stretching. In this case anisotropic Poisson's ratios, ν_{31} and ν_{32} , which correspond to the contractions perpendicular to the chain direction with respect to the chain elongation when tensile load is applied along the chain, can be also measured. The ν_{31} and ν_{32} correspond to the contraction parallel to and perpendicular to the pyranose plane, respectively. DFT simulations consistently expect a crystal modulus of cellulose I along the chain direction to be of the order of 200 GPa. The compressibility along the three major directions of the crystals can be estimated by either simulating hydrostatic pressure or from the elastic tensor derived from a combination of deformations. There are some variations among different calculation approaches, as the elastic tensor does not seem to be strictly symmetric as continuum mechanics is assuming. It was in the range of 50–60 TPa^{-1} perpendicular to the pyranose plane, 5 TPa^{-1} along the pyranose plane and perpendicular to the chain direction and 1.64–1.72 TPa^{-1} along the chain direction. The compressibility values of the corresponding experimental values of 49.6 ± 0.4 , 6.5 ± 1 and $1.71 \pm 0.02 \text{TPa}^{-1}$ [43] thus agreed within 20% of the DFT values. However, the calculated Poisson's ratio ν_{32} was much smaller, by a factor of 2–4. In other words DFT calculations predicted much smaller coupling between the axial deformation and perpendicular deformation. Investigation of such discrepancies and the experimental observation should give more clues to the understanding of crystalline cellulose, and the improvement of modeling approaches

5 Understanding the underlying interactions and thermodynamics

Theoretical and experimental structural details allow us to investigate the underlying interactions stabilizing a system. The efforts of directly finding bonds in an electron density [44] is one direction to recognize chemical bonds, including hydrogen bonds. This approach does not, however, give quantitative energy values.

The importance of the dispersion interaction in the molecular cohesion of polysaccharides can readily be seen in the huge impact of the dispersion correction to the calculated unit cell. The role of hydrogen bonds in determining the properties of cellulose is often overemphasized, [45] but the hydrogen-bond geometry does not necessarily explain the energy differences (stabilities) among the cellulose allomorphs. [46] The contribution of hydrogen bonds to the intermolecular cohesion can be estimated from the geometry combined with a general extrapolation of thermodynamic data of small molecules, which leads to about 30% of the total cohesion being attributed to the hydrogen bonds in native cellulose. [47]

The effect of finite temperature sometimes has been taken into account for the DFT analysis of crystal properties. [42, 48] In systems involving solvents, entropy plays an important role for accumulation of cosolvents [49] or adsorption of polysaccharides on solid surfaces. [50] Approximations to estimate such effects seem to give reasonable results in the above examples, but the experimental check remains rather qualitative. More accurate thermodynamic experimental data such as accurate heat capacity and the adsorption isotherm would become necessary

6 Conclusions

Over the past decades, the quantity and quality of data on cellulosic systems have significantly increased. Those data are often analyzed with help of computational approaches to extract molecular insights, while data are also serving to check the gap between the theoretical approaches and molecular reality. Here I have overviewed the simulation efforts that can be somehow challenged by experimental data. The progress toward understanding needs further effort in combining experimental and theoretical work. The key for this is experimental designs to probe theoretical calculations, and development of tools to translate simulation into experimentally observable equivalents.

References

- [1] Alfred D. French. Combining Computational Chemistry and Crystallography for a Better Understanding of the Structure of Cellulose. *Adv. Carbohydr. Chem. Biochem.*, 67:19–93, jan 2012.

- [2] Peter Zugenmaier. Order in cellulose: Historical review of crystal structure research on cellulose. *Carbohydr. Polym.*, 254(August 2020):117417, feb 2021.
- [3] Hiroshi Sobue, Heinz Kiessig, and Kurt Hess. Das System Cellulose–Natriumhydroxyd–Wasser in Abhängigkeit von der Temperatur. *Zeitschrift für Phys. Chemie*, 43B(1):309–328, may 1939.
- [4] J. Blackwell. Structure of β -chitin or parallel chain systems of poly- β -(1 \rightarrow 4)-N-acetyl-D-glucosamine. *Biopolymers*, 7(3):281–298, mar 1969.
- [5] Yoshiharu Nishiyama, Paul Langan, and Henri Chanzy. Crystal Structure and Hydrogen-Bonding System in Cellulose I β from Synchrotron X-ray and Neutron Fiber Diffraction. *J. Am. Chem. Soc.*, 124(31):9074–9082, aug 2002.
- [6] Yoshiharu Nishiyama, Junji Sugiyama, Henri Chanzy, and Paul Langan. Crystal Structure and Hydrogen Bonding System in Cellulose I α from Synchrotron X-ray and Neutron Fiber Diffraction. *J. Am. Chem. Soc.*, 125(47):14300–14306, nov 2003.
- [7] Masahisa Wada, Henri Chanzy, Yoshiharu Nishiyama, and Paul Langan. Cellulose III I Crystal Structure and Hydrogen Bonding by Synchrotron X-ray and Neutron Fiber Diffraction. *Macromolecules*, 37(23):8548–8555, nov 2004.
- [8] Yoshiharu Nishiyama, Yasutomo Noishiki, and Masahisa Wada. X-ray Structure of Anhydrous β -Chitin at 1 Å Resolution. *Macromolecules*, 44(4):950–957, feb 2011.
- [9] Philip-Kunio Naito, Yu Ogawa, Daisuke Sawada, Yoshiharu Nishiyama, Tadahisa Iwata, and Masahisa Wada. X-ray crystal structure of anhydrous chitosan at atomic resolution. *Biopolymers*, 105(7):361–368, jul 2016.
- [10] P. Langan, R. C. Denny, A. Mahendrasingam, S. A. Mason, and A. Jaber. Collecting and Processing Neutron Fibre Diffraction Data from a Single-Crystal Diffractometer. *J. Appl. Crystallogr.*, 29 PART 4:383–389, 1996.
- [11] Daisuke Sawada, Yoshiharu Nishiyama, Paul Langan, V. Trevor Forsyth, Satoshi Kimura, and Masahisa Wada. Direct Determination of the Hydrogen Bonding Arrangement in Anhydrous β -Chitin by Neutron Fiber Diffraction. *Biomacromolecules*, 13(1):288–291, jan 2012.
- [12] Daisuke Sawada, Yoshiharu Nishiyama, Paul Langan, V. Trevor Forsyth, Satoshi Kimura, and Masahisa Wada. Water in crystalline fibers of dihydrate β -chitin results in unexpected absence of intramolecular hydrogen bonding. *PLoS One*, 7(6):4–11, 2012.

- [13] Pan Chen, Yu Ogawa, Yoshiharu Nishiyama, Malin Bergenstråhle-Wohlert, and Karim Mazeau. Alternative hydrogen bond models of cellulose II and III_I based on molecular force-fields and density functional theory. *Cellulose*, 22(3):1485–1493, jun 2015.
- [14] Yasutomo Noishiki, Yoshiharu Nishiyama, Masahisa Wada, and Shigenori Kuga. Complexation of α -chitin with aliphatic amines. *Biomacromolecules*, 6(4):2362–2364, 2005.
- [15] Y Noishiki, Y Nishiyama, M Wada, S Okada, and S Kuga. Inclusion complex of beta-chitin and aliphatic amines. *Biomacromolecules*, 4(4):944–949–944–949, 2003.
- [16] Daisuke Sawada, Yoshiharu Nishiyama, Loukas Petridis, R. Parthasarathi, S. Gnanakaran, V. Trevor Forsyth, Masahisa Wada, and Paul Langan. Structure and dynamics of a complex of cellulose with EDA: Insights into the action of amines on cellulose. *Cellulose*, 20(4):1563–1571, 2013.
- [17] Daisuke Sawada, Yu Ogawa, Satoshi Kimura, Yoshiharu Nishiyama, Paul Langan, and Masahisa Wada. Solid-solvent molecular interactions observed in crystal structures of β -chitin complexes. *Cellulose*, 21(2):1007–1014, 2014.
- [18] Daisuke Sawada, Yu Ogawa, Yoshiharu Nishiyama, Eiji Togawa, Satoshi Kimura, and Paul Langan. Molecular Interactions in an α -Chitin/Hydrazine Complex: Dynamic Hydrogen Bonds and Improvement of Polymeric Crystallinity. *Cryst. Growth Des.*, 16(6):3345–3352, 2016.
- [19] Yu Ogawa, Philip-Kunio Naito, and Yoshiharu Nishiyama. Hydrogen-bonding network in anhydrous chitosan from neutron crystallography and periodic density functional theory calculations. *Carbohydr. Polym.*, 207(November 2018):211–217, mar 2019.
- [20] Yoshiharu Nishiyama, Shirin Asaadi, Patrik Ahvenainen, and Herbert Sixta. Water-induced crystallization and nano-scale spinodal decomposition of cellulose in NMMO and ionic liquid dope. *Cellulose*, 26(1):281–289, 2019.
- [21] Guangjie Song, Jian Yu, Meichun Ding, and Jun Zhang. A Novel Cellulose/Ionic Liquid Complex Crystal. *Cryst. Growth Des.*, 18(8):4260–4264, aug 2018.
- [22] Takatsugu Endo, Shota Yoshida, and Yoshifumi Kimura. Self-Assembly and Complexation of Cellulose/Ionic Liquid at High Cellulose Concentration: Anion Dependence. *Cryst. Growth Des.*, 20(10):6267–6271, oct 2020.
- [23] Guangjie Song, Jun Zhang, and Yoshiharu Nishiyama. Twisted pseudo-tetragonal orthorhombic lamellar crystal in cellulose/ionic liquid spherulite. *Cellulose*, 27(10):5449–5455, 2020.

- [24] Pan Chen, Yoshiharu Nishiyama, and Karim Mazeau. Atomic partial charges and one Lennard-Jones parameter crucial to model cellulose allomorphs. *Cellulose*, 21(4):2207–2217, 2014.
- [25] Pan Chen, Yoshiharu Nishiyama, and Karim Mazeau. Torsional Entropy at the Origin of the Reversible Temperature-Induced Phase Transition of Cellulose. *Macromolecules*, 45(1):362–368, jan 2012.
- [26] Tomáš Bučko, Jürgen Hafner, Sébastien Lebègue, and János G. Ángyán. Improved Description of the Structure of Molecular and Layered Crystals: Ab Initio DFT Calculations with van der Waals Corrections. *J. Phys. Chem. A*, 114(43):11814–11824, nov 2010.
- [27] Yan Zhang, Hideyo Inouye, Michael Crowley, Leiming Yu, David Kaeli, and Lee Makowski. Diffraction pattern simulation of cellulose fibrils using distributed and quantized pair distances. *J. Appl. Crystallogr.*, 49(6):2244–2248, dec 2016.
- [28] Yoshiharu Nishiyama, Glenn P Johnson, and Alfred D French. Diffraction from nonperiodic models of cellulose crystals. *Cellulose*, 19(2):319–336, apr 2012.
- [29] Da’san M.M. Jaradat, Stefan Mebs, Lilianna Checińska, and Peter Luger. Experimental charge density of sucrose at 20 K: bond topological, atomic, and intermolecular quantitative properties. *Carbohydr. Res.*, 342(11):1480–1489, 2007.
- [30] Martin Müller, Christoph Czihak, Helmut Schober, Yoshiharu Nishiyama, and Gero Vogl. All disordered regions of native cellulose show common low-frequency dynamics. *Macromolecules*, 33(5):1834–1840, 2000.
- [31] Abdelali Boukaoud, Younes Chiba, and Djamel Sebbar. A periodic DFT study of IR spectra of amino acids: An approach toward a better understanding of the N-H and O-H stretching regions. *Vib. Spectrosc.*, 116(February):103280, sep 2021.
- [32] Clemens M. Altaner, Yoshiki Horikawa, Junji Sugiyama, and Mike C. Jarvis. Cellulose I β investigated by IR-spectroscopy at low temperatures. *Cellulose*, 21(5):3171–3179, oct 2014.
- [33] Heath D. Watts, Mohamed Naseer Ali Mohamed, and James D. Kubicki. A DFT study of vibrational frequencies and ^{13}C NMR chemical shifts of model cellulosic fragments as a function of size. *Cellulose*, 21(1):53–70, 2014.
- [34] Christopher M. Lee, James D. Kubicki, Bingxin Fan, Linghao Zhong, Michael C. Jarvis, and Seong H. Kim. Hydrogen-Bonding Network and OH Stretch Vibration of Cellulose: Comparison of Computational Modeling with Polarized IR and SFG Spectra. *J. Phys. Chem. B*, 119(49):15138–15149, 2015.

- [35] Christopher M. Lee, Naseer M. A. Mohamed, Heath D. Watts, James D. Kubicki, and Seong H. Kim. Sum-Frequency-Generation Vibration Spectroscopy and Density Functional Theory Calculations with Dispersion Corrections (DFT-D2) for Cellulose I α and I β . *J. Phys. Chem. B*, 117(22):6681–6692, jun 2013.
- [36] Mohamadamin Makarem, Christopher M. Lee, Kabindra Kafle, Shixin Huang, Inseok Chae, Hui Yang, James D. Kubicki, and Seong H. Kim. *Probing cellulose structures with vibrational spectroscopy*, volume 26. Springer Netherlands, 2019.
- [37] Mohamadamin Makarem, Yoshiharu Nishiyama, Xiaoran Xin, Daniel M. Durachko, Ying Gu, Daniel J. Cosgrove, and Seong H. Kim. Distinguishing Mesoscale Polar Order (Unidirectional vs Bidirectional) of Cellulose Microfibrils in Plant Cell Walls Using Sum Frequency Generation Spectroscopy. *J. Phys. Chem. B*, 124(37):8071–8081, 2020.
- [38] Shixin Huang, Mohamadamin Makarem, Sarah N. Kiemle, Hossein Hamedi, Moujhuri Sau, Daniel J. Cosgrove, and Seong H. Kim. Inhomogeneity of Cellulose Microfibril Assembly in Plant Cell Walls Revealed with Sum Frequency Generation Microscopy. *J. Phys. Chem. B*, 122(19):5006–5019, 2018.
- [39] Shixin Huang, Sarah N. Kiemle, Mohamadamin Makarem, and Seong H. Kim. Correlation between crystalline cellulose structure and cellulose synthase complex shape: a spectroscopic study with unicellular freshwater alga *Micrasterias*. *Cellulose*, 27(1):57–69, 2020.
- [40] Oleg E. Shklyaev, James D. Kubicki, Heath D. Watts, and Vincent H. Crespi. Constraints on I β cellulose twist from DFT calculations of ¹³C NMR chemical shifts. *Cellulose*, 21(6):3979–3991, 2014.
- [41] Fernando L. Dri, Louis G. Hector, Robert J. Moon, and Pablo D. Zavattieri. Anisotropy of the elastic properties of crystalline cellulose I β from first principles density functional theory with Van der Waals interactions. *Cellulose*, 20(6):2703–2718, 2013.
- [42] Fernando L. Dri, Shun Li Shang, Louis G. Hector, Paul Saxe, Zi Kui Liu, Robert J. Moon, and Pablo D. Zavattieri. Anisotropy and temperature dependence of structural, thermodynamic, and elastic properties of crystalline cellulose I β : A first-principles investigation. *Model. Simul. Mater. Sci. Eng.*, 22(8), 2014.
- [43] Guangjie Song, Christine Lancelon-Pin, Pan Chen, Jian Yu, Jun Zhang, Lei Su, Masahisa Wada, Tsunehisa Kimura, and Yoshiharu Nishiyama. Time-Dependent Elastic Tensor of Cellulose Nanocrystal Probed by Hydrostatic Pressure and Uniaxial Stretching. *J. Phys. Chem. Lett.*, 12(15):3779–3785, apr 2021.

- [44] Alfred D. French, Monica Concha, Michael K. Dowd, and Edwin D. Stevens. Electron (charge) density studies of cellulose models. *Cellulose*, 21(2):1051–1063, apr 2014.
- [45] Wolfgang G. Glasser, Rajai H. Atalla, John Blackwell, Malcolm M. Brown, Walther Burchard, Alfred D. French, Dieter O. Klemm, and Yoshiharu Nishiyama. About the structure of cellulose: Debating the Lindman hypothesis. *Cellulose*, 19(3):589–598, 2012.
- [46] Yoshiharu Nishiyama, Paul Langan, Masahisa Wada, and V. Trevor Forsyth. Looking at hydrogen bonds in cellulose. *Acta Crystallogr. Sect. D Biol. Crystallogr.*, 66(11):1172–1177, 2010.
- [47] Yoshiharu Nishiyama. Molecular interactions in nanocellulose assembly. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.*, 376(2112):20170047, feb 2018.
- [48] Gregory J. O. Beran, Joshua D. Hartman, and Yonaton N. Heit. Predicting Molecular Crystal Properties from First Principles: Finite-Temperature Thermochemistry to NMR Crystallography. *Acc. Chem. Res.*, 49(11):2501–2508, nov 2016.
- [49] Pan Chen, Yoshiharu Nishiyama, Jakob Wohlert, Ang Lu, Karim Mazeau, and Ahmed E. Ismail. Translational Entropy and Dispersion Energy Jointly Drive the Adsorption of Urea to Cellulose. *J. Phys. Chem. B*, 121(10):2244–2251, mar 2017.
- [50] Saina Kishani, Tobias Bensefelt, Lars Wågberg, and Jakob Wohlert. Entropy drives the adsorption of xyloglucan to cellulose surfaces – A molecular dynamics study. *J. Colloid Interface Sci.*, 588:485–493, 2021.