



HAL
open science

The Major Role of London Dispersion Interaction in the Assembly of Cellulose, Chitin, and Chitosan

Yiwei Li, Chunxia Yan, Yu Chen, Xuhui Han, Ziqiang Shao, Haisong Qi, Xiaodong Li, Yoshiharu Nishiyama, Tao Hu, Pan Chen

► **To cite this version:**

Yiwei Li, Chunxia Yan, Yu Chen, Xuhui Han, Ziqiang Shao, et al.. The Major Role of London Dispersion Interaction in the Assembly of Cellulose, Chitin, and Chitosan. *Cellulose*, 2023, 30, pp.8127-8138. 10.1007/s10570-023-05376-5 . hal-04177384

HAL Id: hal-04177384

<https://hal.science/hal-04177384>

Submitted on 4 Aug 2023

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.

1 The Major Role of London Dispersion Interaction in the Assembly 2 of Cellulose, Chitin, and Chitosan

3 Yiwei Li^{1,†}, Chunxia Yan^{1,†}, Yu Chen¹, Xuhui Han¹, Ziqiang Shao¹, Haisong Qi^{2,*}, Xiaodong Li¹,
4 Yoshiharu Nishiyama^{3,*}, Tao Hu^{4,*}, Pan Chen^{1,*}

5
6 ¹*School of Materials Science and Engineering, Beijing Institute of Technology, 100081, Beijing, China*

7 ²*State Key Laboratory of Pulp and Paper Engineering, South China University of Technology, Guangzhou*
8 *510640, China*

9 ³*Université Grenoble Alpes, CNRS, CERMAV, 38000 Grenoble, France*

10 ⁴*School of Materials Science and Engineering, State Key Laboratory of Advanced Special Steels, Shanghai*
11 *University, Shanghai 200444, China*

12 *Contributed equally.*

13 *Corresponding authors: qibs@suct.edu.cn; yoshi@cermav.cnrs.fr; taohu@shu.edu.cn; panchen@bit.edu.cn*

14

15 **Abstract:**

16 Noncovalent interactions are vitally important to understand the structural stability and
17 molecular assembly of cellulose and its analogue molecules. Using density functional theory in
18 conjunction with three popular generations of dispersion correction (D2, D3, D4), we
19 systematically estimate the strength of inter-chain interaction for several β -1,4-linked crystalline
20 polysaccharides (cellulose I α , I β , II, III₁, α -chitin, β -chitin, chitosan) and their building block
21 monomers (glucose, cellobiose). Switching on and off dispersion correction for both condensed
22 and isolated chains allow the extraction of the intra- and inter-chain London dispersion
23 interactions as well as the inter-chain electrostatic interaction. Regardless of the generations of
24 dispersion correction and allomorphs, the estimated inter-chain London dispersion interaction is
25 45~74 kJ/mol per pyranose ring comparable to the inter-chain electrostatic interaction (47~88
26 kJ/mol). The upper limit of the strength of inter- or intra-chain hydrogen bonds are estimated to
27 be 21~53 kJ/mol based on energy profiles of hydroxyl rotation. Our work quantitatively
28 highlights that it is the London dispersion interaction rather than the hydrogen bonding
29 interaction dominating in the tight assembly of polymer chains for β -1,4-linked crystalline
30 polysaccharides, regardless of the crystal allomorph and types as well as the generations of
31 dispersion correction of DFT. thus, London dispersion interaction should be preferentially
32 considered during their deconstruction, defibrillation, or dissolution processes.

33

34 **Introduction:**

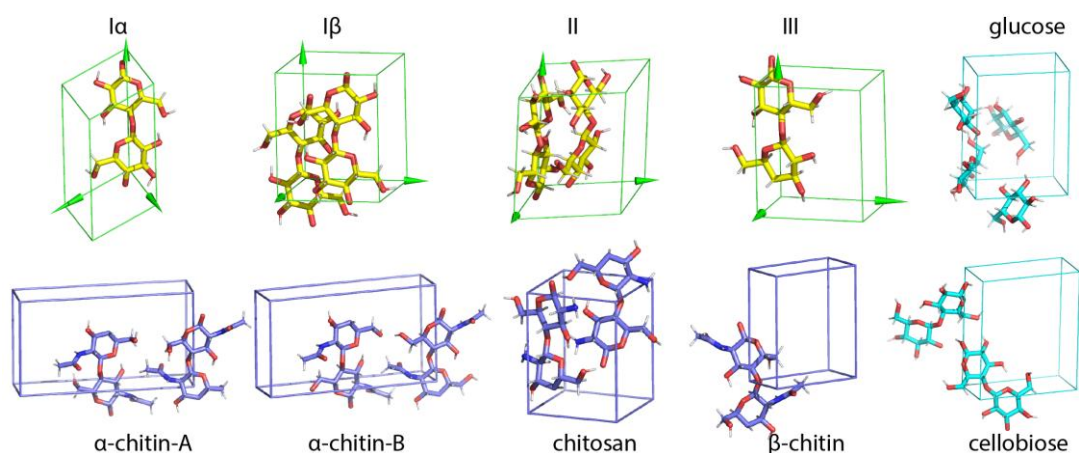
35 X-ray and neutron studies revealed abundant hydrogen bonds and regular network patterns
36 among the crystal structures of cellulose and chitin (Langan et al. [1999](#); Nishiyama et al. [2002](#);
37 Nishiyama et al. [2003](#); Wada et al. [2004](#); Deringer et al. [2016](#); Ogawa et al. [2019](#); Sikorski et al.
38 [2009](#); Nishiyama et al. [2011](#); Naito et al. [2016](#)). Some authors regarded hydrogen bonds as the
39 key factor governing the assembly of polymer chains, trying to explain many physical properties
40 of cellulose. Recent studies have gradually reshaped such a view. The impact of hydrogen bond
41 on the peeling-off of the cellotetrose (Bergensträhle et al. [2010](#)) and the proposal of the
42 hydrophobicity of cellulose (Lindman et al. [2010](#); Lindman et al. [2021](#); Medronho et al. [2014](#);
43 Glasser et al. [2012](#)) also argued that the contribution of hydrogen bonding to the insolubility of
44 cellulose was overemphasized.

45 After a few years of debates, it seems that we have well established hitherto that other
46 noncovalent interactions, such as electrostatic interaction and London dispersion interactions,
47 are also responsible for the tight chain stacking (Jarvis [2023](#)). A more recent, detailed review
48 evaluated the “exaggerated” role of hydrogen bonds (Wohlert et al. [2023](#)) associated with the
49 properties of cellulose in the paste decades. Based on the linear tendencies of heat of evaporation
50 of analogue molecules with number of hydroxyl groups and molecular weights, the inter-chain
51 London dispersion interaction in crystalline cellulose I β is estimated to be 67 kJ/mol per glucose
52 while hydrogen bond energy is 24 kJ/mol (Nishiyama [2018](#)). Still, the first-principles-based
53 quantification of internal energy of the cellulose and chitin are rare (Deringer et al. [2016](#)). We
54 previously quantified the partition of the non-covalent interaction in chitin and chitosan based
55 on DFT-D2 calculations (Chen et al. [2021](#)) using energy decomposition analysis base on low-
56 dimension fragments (Deringer et al. [2016](#)). Here, we have extended this method for the
57 systematical analysis of cellulose, chitin, chitosan, and their allomorphs as well as their
58 monomers using three different generations of dispersion correction approach (D2 (Grimme
59 [2006](#)), D3 (Grimme et al. [2010](#)), D4 (Caldeweyher et al. [2017](#); Caldeweyher et al. [2019](#))) in which
60 polarizability was considered. In addition, the upper limit of hydrogen bond strength is estimated
61 based on the energy profile of hydroxyl rotation, providing a basic understanding of the
62 components of noncovalent of the crystalline polysaccharides.

63 **Computational Methods:**

64 **Model Construction:** Eight types of crystalline polysaccharides, named cellulose I α
65 (Nishiyama et al. [2003](#)), I β (Nishiyama et al. [2002](#)), II (Langan et al. [1999](#)), III_I (Wada et al. [2004](#)),
66 α -chitin-A (Sikorski et al. [2009](#), Deringer et al. [2016](#)), α -chitin-B (Sikorski et al. [2009](#), Deringer et
67 al. [2016](#)), β -chitin (Nishiyama et al. [2011](#)), chitosan (Naito et al. [2016](#))) whose atomic structures
68 are available. β -D-glucose and, β -D--cellobiose (Jeffrey [1968](#)) were also included. The unit cells

69 and the corresponding fragment of these crystals are all represented in Figure 1. The
 70 computational setup is identical to our former study (references), and a detailed description can
 71 be found in the following method sections.



72
 73 Figure 1. Unit cell representation of cellulose, chitin, chitosan, glucose, and cellobiose. (All
 74 crystal structures are deposited in glyco3D)

75 **Energy Minimization:** Periodic boundary condition (PBC) based DFT calculation were
 76 performed by using Quantum Espresso (QE) (Giannozzi et al. 2009; Giannozzi et al. 2017) and
 77 VSAP package. D2 (Grimme 2006) and D3 (Grimme et al. 2010) are implemented in QE and
 78 D4 (Caldeweyher et al. 2017; Caldeweyher et al. 2019) can be implemented to VASP (Hafner et
 79 al. 1997). The generalized gradient approximation (GGA) functional PBE (Perdew et al. 1996)
 80 was used for geometry optimization of crystals. The total energy and force convergence
 81 thresholds for ion minimization were set to 1.0e-6 Ry and 1.0e-5Ry/bohr, respectively. The
 82 kinetic energy cutoff value of the wave function was 160 Ry. The k-points were set to (2, 2, 2, 0,
 83 0, 0).

84 **Energy Decomposition Analysis:** The classical molecular mechanics (MM) represents the
 85 intermolecular noncovalent interaction as composed of Coulomb interactions between point
 86 charges of each paired atoms and Lennard-Jones potential interactions which includes the
 87 London dispersion and Pauli repulsion terms. In DFT calculation, the intermolecular energy is
 88 composed of terms such as electrostatics, exchange, induction and dispersion interaction etc. For
 89 both cases, we can simplify as dispersion energy and everything else as electrostatic interactions,
 90 the molecular interaction energy within a crystal can be written as equation 1.

$$91 \quad E_{int} = E_{elec} + E_{disp} \quad (1)$$

92 Further decomposition for intra-chain and inter-chain terms results in four energy terms,
 93 which is the interchain electrostatic energy (E_{inter_E}), the intrachain electrostatic energy (E_{intra_E}),

94 the interchain dispersion energy (E_{inter_D}), and the intrachain dispersion energy (E_{intra_D}) as
 95 represented by equation 2.

$$96 \quad E_{int} = E_{inter_E} + E_{intra_E} + E_{inter_D} + E_{intra_D} \quad (2)$$

97 We rely on the low dimensional fragments approach that is widely used in DFT-based
 98 materials simulations (Deringer et al. [2016](#)) to estimate the contribution of interchain interactions.
 99 In brief, DFT calculation was performed for the three-dimensional crystals (3D: raw unit cell
 100 applied with periodic boundary condition), relaxation of both atomic coordinates and crystal
 101 lattices, obtaining an optimized energy and structure (noted as E_{3D_disp} , corresponds to top left in
 102 Figure 2). One isolated structural fragment (1D) was constructed by computationally “cleaving”
 103 the lattice apart from the above fully relaxed 3D structure. This is achieved by leaving one chain
 104 within the supercell and enlarging the transverse lattices by a factor of two, which separates this
 105 chain from their periodic boundary images, as shown in (top right of) Figure 2. The total-energy
 106 computation was subsequently performed by freezing both the box size and coordinates, and the
 107 obtained energy is noted as E_{1D_disp} . This is slightly different from our previous work (Chen et al.
 108 [2022](#)), in which the enlarged box and atoms were both relaxed, and makes the E_{intra_E} truly
 109 constant, In the previous studies the E_{intra_E} was “assumed” to be unchanged which turned out
 110 to be not the case as can be seen in the comparison between Table 1 and Table S1. By switching
 111 off the dispersion correction and freezing the atoms in the 3D system when dispersion
 112 correction switched off and in 1D system when dispersion correction was turned on and off, two
 113 other energy terms can be generated, noted as E_{3D_nodisp} and E_{1D_nodisp} corresponding to the
 114 bottom left and right states in Figure 2, respectively. This is illustrated in equation 3, 4, and 5,
 115 respectively.

$$116 \quad E_{intraD} = (E_{1D_disp} - E_{1D_nodisp})/N \quad (3)$$

$$117 \quad E_{inter_D} = (E_{3D_disp} - E_{3D_nodisp})/N - E_{intra_D} \quad (4)$$

$$118 \quad E_{inter_E} = E_{3D_nodisp}/N \quad (5)$$

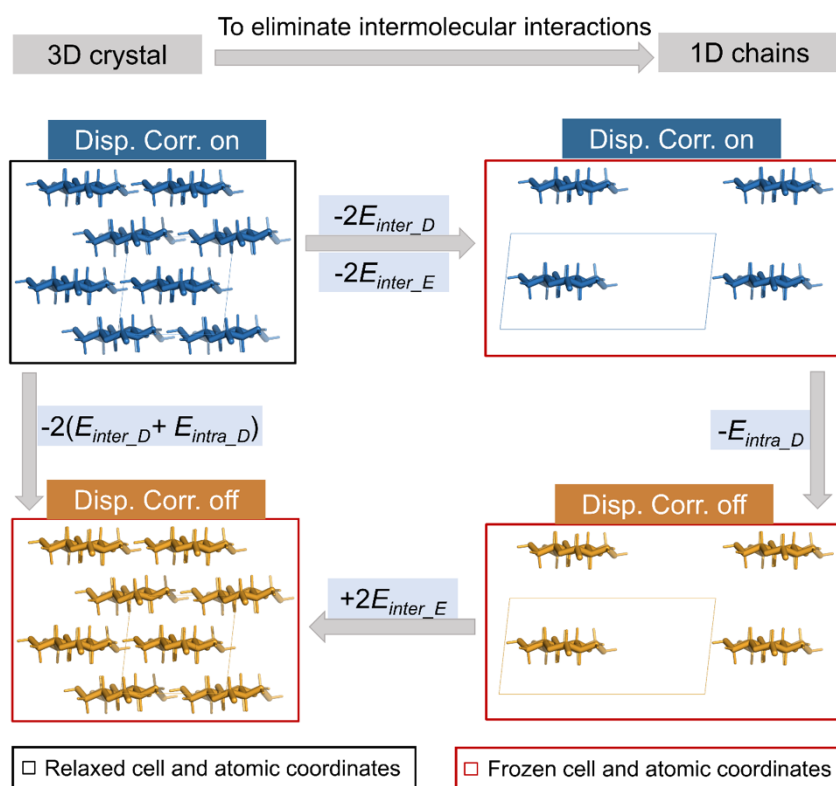
119 Where N stands for the number of residue per unit cell.

120 The cohesion energy of crystals per residue equals to the sum of E_{inter_D} and E_{inter_E}
 121 (equation 6).

$$122 \quad E_{cobe_E} = E_{inter_D} + E_{inter_E} \quad (6)$$

123 The graphical illustration of such energy decomposition with cellulose I β as a trial is shown in
 124 Figure 2. On the graph, E_{1D_disp} is total energy of one isolated structural fragment with
 125 dispersion correction switched on. In contrast, E_{1D_nodisp} is the one without dispersion correction.

126 For E_{3D_disp} and E_{3D_nodisp} , the 3D subscript means a standard unit cell with periodic
 127 boundary condition applied and thus the crystal is three dimensional infinite. Disp. and Nodisp
 128 indicated whether or not the dispersion correction is applied.

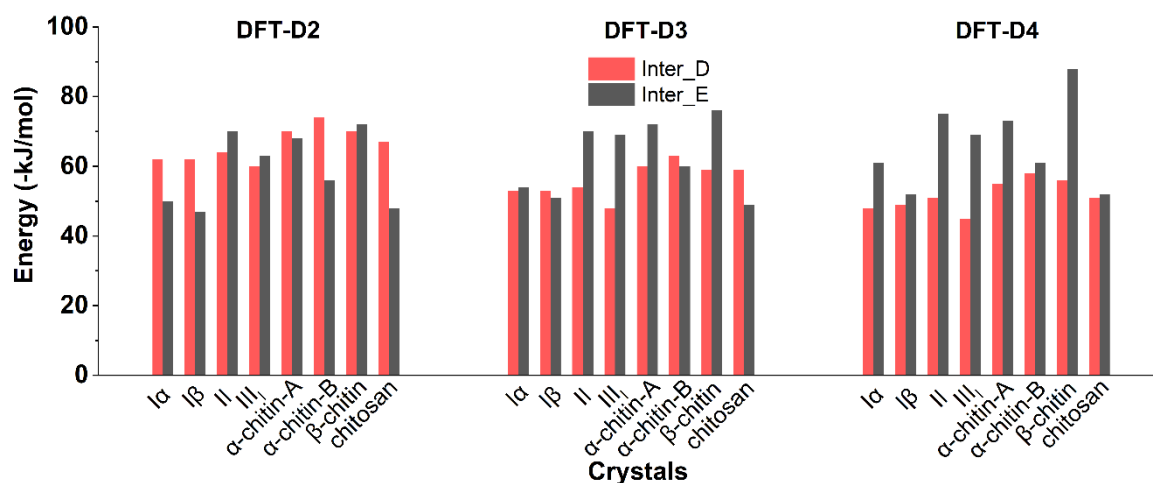


129
 130 Figure 2. Illustration of the condensed 3D fragment and isolated 1D fragment with dispersion
 131 correction switched on/off used for the energy composition calculation.

132
 133 **Table 1. Decomposed energy (E_{inter_E} , E_{inter_D} , E_{intra_D} , E_{Cohe_E}) of crystals.**

Energy (-kJ/mol)		Glu	CB	I α	I β	II	III ₁	α -chitin-A	α -chitin-B	β -chitin	Chitosan
D2	Intra_D	59	68	70	79	77	76	105	104	105	83
	Inter_D	98	79	62	62	64	60	70	74	70	67
	Inter_E	145	107	50	47	70	63	68	56	72	48
	Cohe_E	243	186	111	109	134	123	139	130	141	116
D3	Intra_D	38	44	52	52	51	51	70	69	70	54
	Inter_D	83	66	53	53	54	48	60	63	59	59
	Inter_E	148	111	54	51	70	69	72	60	76	49
	Cohe_E	231	177	107	105	124	117	132	124	136	108
D4	Intra_D	62	69	78	79	76	77	104	104	104	82
	Inter_D	76	61	48	49	51	45	55	58	56	51
	Inter_E	151	117	61	52	75	69	73	61	88	52
	Cohe_E	227	178	109	100	126	114	128	119	144	103

134



135
 136 Figure 3. Histogram of intermolecular London dispersion and electrostatic energies (the two
 137 components of E_{CohE}) based on DFT-D2 (left), D3(middle), and D4(right).

138

139 Results and Discussion:

140 Estimation of London Dispersion Interaction

141 Figure 3 and Table 1 show the decomposed energy for all crystals as a function of
 142 generation types of dispersion correction. The $(E_{intra_D}, E_{inter_D}, E_{inter_E}, E_{CohE_E})$ of Iβ is
 143 estimated as (79, 62, 47, 109) kJ/mol per glucose based on DFT-D2, respectively. This
 144 quantification of E_{inter_D} is close to the empirical estimation of 67 kJ/mol (Nishiyama [2018](#)).
 145 Comparing the four allomorphs of cellulose within D2 framework, the E_{inter_D} of II is the largest
 146 (64 kJ/mol per glucose), while that of III₁ is the smallest (60 kJ/mol per glucose), and Iα/Iβ is in
 147 the middle (62 kJ/mol per glucose). If we simply normalize the energy by the volume of their
 148 unit cell and considering the relative volume of Iβ being as 1, we obtained xxx for cellulose and
 149 xxx for chitin, respectively. Overall, the estimated E_{inter_D} for all polysaccharide crystals ranges
 150 from 60 to 74 kJ/mol per glucose, the E_{intra_D} vary from 76 to 105 kJ/mol per glucose, and the
 151 E_{inter_E} vary from 47 to 70 kJ/mol based on DFT-D2 calculation. In fact, one can find (Table 1)
 152 in crystalline monomer and dimer that the E_{intra_D} (59, 68 kJ/mol) and E_{inter_D} (98, 79 kJ/mol)
 153 are (although different but) already comparable to that of polymer crystals. The slightly higher
 154 value of E_{inter_D} and smaller value of E_{intra_D} in small molecular crystal than polymer crystal is
 155 simply ascribed to the increased molecular weight of repeat unit (162 for anhydrous glucose, 171
 156 for cellobiose, and 180 for glucose). Different hydroxy groups per glucose also leads to different
 157 number of inter-molecular hydrogen bonds per residue, which is 5 for glucose, 4 for cellobiose
 158 and 2 for II, III₁, chitin, and 1 for Iα/Iβ and chitosan. Because the component nature of

159 hydrogen bonds is mostly electrostatic interaction, the E_{inter_E} in small molecular crystals is both
160 much larger than its E_{inter_D} (Figure S1) and E_{inter_E} of polymer crystals, respectively. For a quick
161 and rough calculation, we simply divide the E_{inter_E} by the number of hydrogen bonds per
162 residue, the strength of the single hydrogen bond in mono and dimer crystals can be estimated to
163 be under 30 kJ/mol in glucose and 27 kJ/mol in cellobiose. Similarly, the E_{inter_E} in cellulose II,
164 III_I, and chitin are larger than I α , I β , and chitosan, simply due to one more inter-chain hydrogen
165 bond per glucose residue. More measurement of hydrogen bonds interaction will be discussed in
166 the last section.

167 Based on the DFT-D2 calculation, for the β -1,4-linked crystalline polysaccharides, the
168 London dispersion interaction represents by 48~58% of the total cohesion energy of the
169 polymer crystal regardless of polymer categories and types of correction used as can be
170 overviewed in Figure 3 (left) and Table 3.

171

172 **The Impact of Three Generations of Dispersion Correction**

173 When dispersion correction was modulated from D2 to the other two generations (D3 and
174 D4), the four energy terms (E_{intra_D} , E_{inter_D} , E_{inter_E} , E_{Cobe_E}) of cellulose I β varied to (52, 53, 51,
175 105 kJ/mol) for D3 and (79, 49, 52, 100 kJ/mol) for D4, showing the monoclinic decreasing of
176 E_{inter_D} and the increase of E_{inter_E} and resulting the E_{inter_E} slightly over E_{inter_D} . Such tend of
177 reverse between E_{inter_D} and E_{inter_E} is also applicable for other crystals as presented in Figure Sxx.
178 In D2, the dispersion coefficients for each atom species were constant no matter their chemical
179 contexts. In D3 and D4, the local electron polarizability effect was accounted for and the
180 dispersion coefficients are automatically adjusted according to their local chemical environment,
181 resulting in more expensive calculations and different energy values. Such update leads to slightly
182 difference in predicted unit cell parameters, especially the slightly expansion of unit cell
183 parameter a , which reflected the relatively less tight packing in D3 and D4, and therefore
184 reduced E_{inter_D} and increased E_{inter_E} . Still, the E_{inter_D} takes 41~55% of the total intermolecular
185 interactions for D3, and 35~48% for D4, as shown in Table S7.

186 No matter which type of dispersion correction, the E_{inter_D} and E_{intra_D} of chitin is always
187 higher than cellulose. This is simply ascribed to the larger molecular weight of the repeat unit
188 (which is 162 Da for cellulose, and 203 Da for chitin).

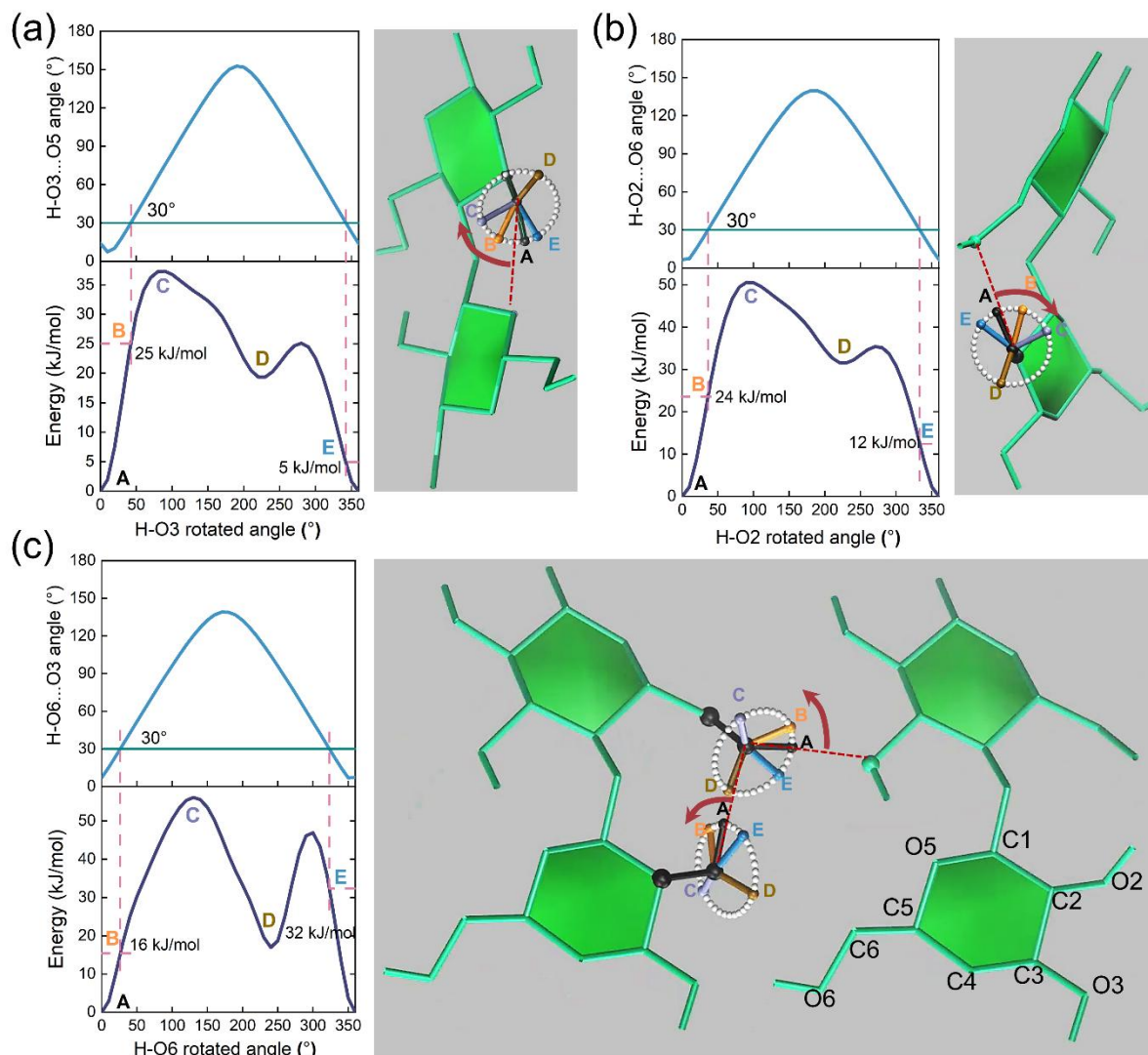
189

190 **Estimations of hydrogen bond strength**

191 Although more than one hundred years have passed since the first proposal for hydrogen
192 bond (Huggins¹⁹⁷¹; Derewenda et al. ²⁰²¹), the estimation of its range of the strength is still
193 under development (Emamian et al. ²⁰¹⁹). To computationally estimate one hydrogen bond
194 interaction between small molecules such as water, one can simply separate the hydrogen
195 bonding paired molecules and estimated the energy difference as hydrogen bond strength since
196 the London dispersion interaction is small and thus can be neglected. However, for hydrogen
197 bonds in cellulose or chitin, one cannot simply do so because of other electrostatic interactions
198 and the increased London dispersion due to increased molecular weight. In addition, dividing the
199 E_{inter_E} by the number of interchain hydrogen bonds would overestimate the hydrogen bond
200 contribution, because other multipolar electrostatic interactions also contribute to E_{inter_E} . In the
201 textbook (Mark ²⁰²³), an hydrogen bonds are often judged by an arbitrary geometric factors:
202 donor (H)-acceptor length < 0.27 nm and H-donor-acceptor angle $< 30^\circ$. When such criteria are
203 not fulfilled, they fall into categories of Coulomb interaction. Based on this, we developed an
204 approach by extracting one chain or one sheet out of the 3D crystal (as shown in Figure 4) and
205 rotating the hydroxyl group around the C-O bond. Single point energy calculations were run at
206 each point by freezing all the atoms. Only a proton is moving so the London dispersion
207 interaction can be regarded as nearly constant and the energy difference can be regarded as an
208 indicator of hydrogen bond and partial contribution from electrostatic repulsion interaction. This
209 method is similar to the study by Estácio et al (Estácio et al. ²⁰⁰⁴) which shows that this
210 hydrogen bond energy value is overestimated and can be considered as an upper limit of
211 hydrogen bond strength. Details for each hydrogen bond energies in β -chitin are also provided in
212 Figure S2.

213 Taking the intrachain hydrogen bond of $I\beta$ as an example, the HO3-O3 was rotated around
214 the O3-C3 bond with a stepwise increment of 10° starting from the initial energy minimum
215 (labeled as A) and single point energy was calculated at each frame, as shown in Figure 4. The
216 variation of total energy is purely ascribed to the movement of hydrogen (HO3) atom. The
217 difference between the optimized energy and when H-O...H angle becomes 30° is considered as
218 the strength of intra-chain O3H...O5 hydrogen bond (labeled as B in the Figure 4a). The
219 estimation of another intra-chain (O2-HO2...HO6) hydrogen bond was done in a similar way by
220 applying rotation for O2-HO2 around the C2-O2 bond as shown in Figure 4b. The estimation
221 of inter-chain hydrogen bond (O6-HO6...O3) requires the simultaneous rotations of C2-O2-
222 HO2 and C6-O6-HO6 angles, because the sole rotation of C6-O6-HO6 will induce short
223 HO6...HO2 contact ($< 1 \text{ \AA}$) that is unreasonable. To avoid this short proton-proton contact,
224 starting from the energy minima, the HO2 hydroxyl and HO6 hydroxyls are rotated in an

225 opposite direction as indicated in the red arrow in Figure 4c. The rotation of HO2 in Figure 4c
 226 and 4b follows the same direction. The total energy variation of simultaneous rotation of HO2
 227 and HO6 was subtracted by the energy profiles of the sole HO2 rotation in Figure 4b, and the
 228 result is shown in Figure 4c. The estimation of other crystals follows the receipt of $I\beta$.



229
 230 Figure 4. Total energy variation of $I\beta$ chain as the function of H-O rotated angle around bond
 231 C-O. The arrow indicates the rotation trajectory of hydroxyl groups. A, B, C, D, E in energy
 232 profiles corresponds to the same label in molecular snapshots which indicates the selected
 233 frames during the rotation trajectory of hydroxyls. The red arrows indicate the rotational
 234 direction of hydroxyls from 0 to 360 degree.

235 By setting the universal hydrogen bond criteria ($H\dots\text{acceptor} < 0.27$ nm and $H-D-A < 30^\circ$),
 236 the three major hydrogen bond ($E_{HO3\dots O5}$, $E_{HO2\dots O6}$, $E_{HO6\dots O23}$) of $I\beta$ can be qualitatively
 237 estimated to have the upper limit of 25 kJ/mol, 24 kJ/mol, and 31 kJ/mol, with DFT-D2 as
 238 labeled in the dashed line in Figure 4a, 4b, and 4c. The slightly higher value of $E_{HO6\dots O3}$ than the

239 other two can be ascribed to the additional contribution of electrostatic attraction from the
 240 HO6...O2 pair, since the rupture of HO6...O3 also alters the HO6...O2 distance during the
 241 rotation of hydroxyl group (Figure 4 & S3). Summary of all hydrogen bond strength is provided
 242 in Table 2, showing the range from 14 to 33 kJ/mol, which is similar to that in the estimation of
 243 alcohol hydrogen bond (24 kJ/mol) of small analogs (Nishiyama [2018](#)). This indicates that the
 244 hydrogen bond strength in cellulose, chitin, and chitosan is not particularly strong, but is similar
 245 to their smaller analogs (such as glucose, Table S2). **For the chitin and chitosan, the strength of**
 246 **NH...OC hydrogen bond is 26 or 34 kJ/mol, within the same magnitude of the strength of**
 247 **OH...OH hydrogen bond of cellulose.** All the hydrogen bond strength estimation depends very
 248 little on the generation types of London dispersion correction, as shown in Table S3, S4 (13~34
 249 kJ/mol for D3) and S5 (14~32 kJ/mol for D4), for all crystals.

250 **Table 2. The type and maximum strength of intra- and inter-chain hydrogen bond (HB)**
 251 **in crystals.**

	Energy (-kJ/mol)	I α	I β	II	III ₁	α -chitin-A	α -chitin-B	β -chitin	Chitosan
Intra-	O3H...O5	25	26	21	21	27	32	27	27
HB	O2H...O6	23	24	--	--	--	--	--	--
	O6H...O3	33	31	--	--	--	--	--	--
	O6H...O2	--	--	16	21	--	--	--	--
Inter-	O2H...O6	--	--	19	23	--	--	--	--
HB	O2H...O2	--	--	14	--	--	--	--	--
	O6H...O6	--	--	30	--	--	20	--	--
	O6H...OC	--	--	--	--	13	--	17	--
	NH...OC	--	--	--	--	33	30	33	--
	O6H...N	--	--	--	--	--	--	--	27

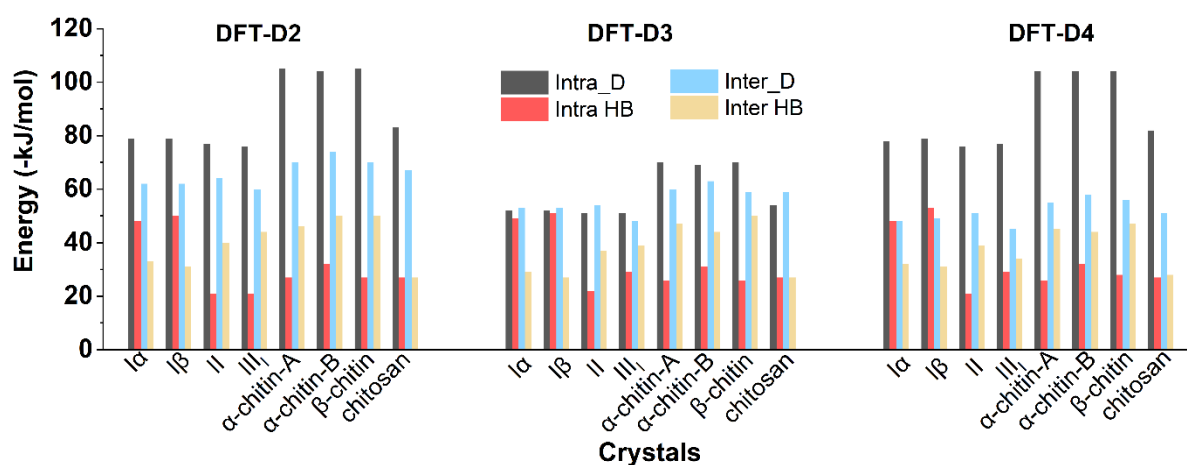
252 Each anhydrous glucose residue in I β contains three free hydroxyl groups that form one
 253 interchain hydrogen bond and two intra-chain hydrogen bond on average, the strength of which
 254 is <31 and <50 kJ/mol per glucose, proving that both the inter- and intra-chain hydrogen **bond**
 255 **strengths are less stronger than the corresponding London dispersion interactions (62 and 79**
 256 **kJ/mol per glucose)** respectively. This picture is regardless of types of dispersion correction (52,
 257 51, 53, 27 of E_{intra_D} , E_{intra_HB} , E_{inter_D} , E_{inter_HB} for D3 and 79, 53, 49, 31 for D4). A similar
 258 expression also applies to cellulose I α that exhibits similar structural features as I β , as shown in
 259 Table 3.

260 **Table 3. Comparison between London dispersion interaction and hydrogen bond.**

		I α	I β	II	III ₁	α -chitin-A	α -chitin-B	β -chitin	Chitosan
D2	Intra_D	79	79	77	76	105	104	105	83
	Intra HB	48	50	21	21	27	32	27	27

	Inter_D	62	62	64	60	70	74	70	67
	Inter HB	33	31	40	44	46	50	50	27
D3	Intra_D	52	52	51	51	70	69	70	54
	Intra HB	49	51	22	29	26	31	26	27
	Inter_D	53	53	54	48	60	63	59	59
	Inter HB	29	27	37	39	47	44	50	27
D4	Intra_D	78	79	76	77	104	104	104	82
	Intra HB	48	53	21	29	26	32	28	27
	Inter_D	48	49	51	45	55	58	56	51
	Inter HB	32	31	39	34	45	44	47	28

261



262

263 Figure 5. Histogram of (inter-chain and intra-chain) London dispersion interaction and hydrogen
 264 bond strength in crystals.

265 For other cellulose allomorphs (II and III₁) and chitin, as well as chitosan, the intra-chain
 266 hydrogen bond (O3H...O5) is retained but the planar hydrogen-bond network disappears due to
 267 the conformational variation of the exocyclic hydroxymethyl group, which is *tg* in native
 268 cellulose, *gg/gt* in α -chitin, and *gt* in the rest. The number of inter-chain hydrogen bonds increases
 269 from 1 to 2 per residue and the intra-chain one decreases from 2 to 1, thus accompanying with
 270 the increased total strength of inter-chain hydrogen bonds. However, the strength of inter-chain
 271 hydrogen bonds is always below 50 kJ/mol from D2, 50 kJ/mol from D3, and 47 kJ/mol from
 272 D4, showing that none of them exceeds their corresponding inter-chain London dispersion
 273 interaction (see Table 3). Regarding the intra-chain interactions, hydrogen bonds contribute far
 274 less than dispersion interaction (<32 kJ/mol versus >50 kJ/mol, respectively) as also shown in
 275 Table 3. The same picture between intra- and inter-chain dispersion interactions and hydrogen
 276 bonds also holds for other polysaccharide analogues. As shown in the Table S6, the component
 277 of hydrogen bonding interaction in the total interchain interaction in crystals varies between 23%

278 and 40% while that of London dispersion energy fluctuates between 35% and 58% (Table S7),
279 and the electrostatic interactions overall occupied from 42% to 61% (Table S8).

280 When hydroxyl groups is rotated around their corresponding CO bond (Figure 4), the
281 energy barrier during the C-O-H angle rotation may reach 50~60 kJ/mol per residue. One may
282 simply think this should be considered as the hydrogen bond strength. Such thought is improper
283 because the energy variation induced by the rotation of hydroxyl group includes both hydrogen
284 bond and other repulsions or attractions between hydrogen and nearby atoms. This repulsion or
285 attraction occurs in polymers but not for small molecules (such as water) in isolated state due to
286 steric effect of adjacent atoms in polymer chain. One obvious evidence is the bimodal shape of
287 the energy profile, where an energy minimum occurs between 200° and 250°, and is due to the
288 electrostatic attraction between mobile hydrogen and its adjacent oxygen. Their geometry
289 parameters at this minimum are far beyond the standard hydrogen bond criteria. At this low
290 energy minimum, the nearby electrostatic repulsion is the most minor and the energy variation is
291 always less than 32 kJ/mol in comparison to the initial state (20 kJ/mol in Figure 4a, 32 kJ/mol
292 in Figure 4b, 18 kJ/mol in Figure 4c), which is much smaller than the energy barrier. Therefore,
293 our estimated energy difference is already the upper limit of hydrogen bond strength.

294 The fact that hydrogen bond strength estimated here is slightly higher than those reported
295 for small molecules is that one hydroxyl in polysaccharide crystal structures acts as both donor
296 and acceptor because of the hydrogen bonding network, resulting in constructing correlation
297 among each hydrogen bond. Influencing one may also partially interrupt others. The
298 cooperativity in hydrogen bond network is also reported to enhance the strength of a single
299 hydrogen bond, but the extent is limited (Qian [2008](#); Masella et al. [2000](#)). The hydrogen bond
300 strength is also reported to be context-dependent in protein, which is stronger in the inner
301 hydrophobic core than surface but never found to dominate the structural stability.(Deechongkit
302 et al. [2004](#)) A similar principle can be parallel to crystalline polysaccharides according to our
303 estimation and London dispersion interaction.

304

305 **Conclusion:**

306 To summarize, using DFT calculation and modulation of dispersion correction and the
307 single energy calculation with rotating hydroxyl groups, we have systematically quantified the
308 London dispersion interaction and strength of hydrogen bond of cellulose, chitin, and chitosan
309 and their monomer and dimers. We are able to confirm the fact that inter-chain London
310 dispersion interaction exceeds the strength of inter-chain hydrogen bond within the lattice
311 energy of cellulose I β as Nishiyama reported. In addition, the intra-chain London dispersion

312 interaction was proved to also be stronger than the intra-chain hydrogen bonds for I β crystals.
313 Moreover, these findings not only adapt to cellulose I β , but also to other cellulose allomorphs
314 (I α , II, III_I) and other β -(1,4)-crystalline polysaccharides (chitin and chitosan). Although the
315 alteration of different generations of dispersion interactions alters slightly the absolute value of
316 intermolecular dispersion and electrostatic as well as hydrogen bonding energies due to slightly
317 different unit cell parameters comparing to experimental observation, the picture that London
318 dispersion interaction excesses hydrogen bonding interaction always stands. Our quantification
319 provides direct evidence that refutes the hydrogen bonding dominated dissolution mechanism of
320 cellulose and chitin. One may argue that although one hydrogen bond is not strong enough, the
321 activation energy of peeling-off abundant hydrogen bonds along a polymer chain would be
322 strong. This is true, but the peeling-off energy required to resist London dispersion interaction in
323 crystal would also be strong, and still larger than the total energy of many hydrogen bonds.

324

325 **Ethics approval and consent to participate**

326 All authors state that they adhere to the Ethical Responsibilities of Authors. In addition, this
327 article does not contain any studies with human participants or animal performed by any of the
328 authors.

329 **Competing Interests**

330 Not applicable

331 **Availability of data and materials**

332 Not applicable

333 **Competing interests**

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

335 **Funding**

336 This research was supported by the Beijing Natural Science Foundation (2232064).

337 **Author Contributions**

338 YL & CY: Investigation, data analysis, original draft. Y C & XH: Investigation. ZS & XL:
339 Resources and review. HQ & Y. Nishiyama & TH: Supervision and review. PC: Project
340 administrator, investigation, supervision, original draft and review. All authors read and approved
341 the final manuscript.

342 **Acknowledgements**

343 PC thanks the Beijing Natural Science Foundation (2232064).

344

345 References

- 346 Bergensträhle M, Wohler J, Himmel M E, Brady J W (2010) Simulation studies of the
347 insolubility of cellulose. *Carbohydr. Res* 345(14): 2060–2066.
348 <https://doi.org/10.1016/j.carres.2010.06.017> Caldeweyher E, Bannwarth C, Grimme S
349 (2017) Extension of the D3 dispersion coefficient model. *J Chem Phys* 147 (3): 034112.
350 <https://doi.org/10.1063/1.4993215>
- 351 Caldeweyher E, Ehlert S, Hansen A, Neugebauer H, Spicher S, Bannwarth C, Grimme S (2019)
352 A generally applicable atomic-charge dependent London dispersion correction. *J Chem*
353 *Phys*: 150 (15), 154122. <https://doi.org/10.1063/1.5090222>
- 354 Chen P, Nishiyama Y, Putaux J-L, Mazeau K (2014) Diversity of potential hydrogen bonds in
355 cellulose I revealed by molecular dynamics simulation. *Cellulose* 21 (2): 897-908.
356 <https://doi.org/10.1007/s10570-013-0053-x>
- 357 Chen P, Nishiyama Y, Wohler J (2021) Quantifying the influence of dispersion interactions on
358 the elastic properties of crystalline cellulose. *Cellulose* 28 (17): 10777-10786.
359 <https://doi.org/10.1007/s10570-021-04210-0>
- 360 Chen P, Ogawa Y, Nishiyama Y, Bergensträhle-Wohler M, Mazeau K (2015) Alternative
361 hydrogen bond models of cellulose II and III₁ based on molecular force-fields and density
362 functional theory. *Cellulose* 22 (3), 1485-1493. <https://doi.org/10.1007/s10570-015-0589-z>
- 363 Chen P, Zhao C J, Wang H Y, et al. (2022) Quantifying the Contribution of the Dispersion
364 Interaction and Hydrogen Bonding to the Anisotropic Elastic Properties of Chitin and
365 Chitosan[J]. *Biomacromolecules* 23(4): 1633–1642.
366 <https://doi.org/10.1021/acs.biomac.1c01488>
- 367 Deechongkit S, Nguyen H, Powers E, et al. (2004) Context-dependent contributions of
368 backbone hydrogen bonding to β -sheet folding energetics. *Nature* 430: 101–105.
369 <https://doi.org/10.1038/nature02611>
- 370 Derewenda Z S (2021) On the centennials of the discoveries of the hydrogen bond and the
371 structure of the water molecule: the short life and work of Eustace Jean Cuy (1897–1925).
372 *Acta Crystallogr., Sect. A*. <https://doi.org/10.1107/S205327332100>
- 373 Deringer V L, Englert U, Dronskowski R (2016) Nature, strength, and cooperativity of the
374 hydrogen-bonding network in α -chitin. *Biomacromolecules* 17 (3): 996-1003.
375 <https://doi.org/10.1021/acs.biomac.5b01653>
- 376 Emamian S, Lu T, Kruse H, Emamian H (2019) Exploring Nature and Predicting Strength of
377 Hydrogen Bonds: A Correlation Analysis Between Atoms-in-Molecules Descriptors,
378 Binding Energies, and Energy Components of Symmetry-Adapted Perturbation Theory. *J.*
379 *Comput. Chem* 40: 2868-2881. <https://doi.org/10.1002/jcc.26068>
- 380 Estácio S G, Cabral do Couto P, Costa Cabral B J, et al. (2004). Energetics of Intramolecular
381 Hydrogen Bonding in Di-substituted Benzenes by the ortho-para Method. *J. Phys. Chem. A*
382 108(49), 10834-10843. <https://doi.org/10.1021/jp0473422>
- 383 Glasser W G, Atalla R H, Blackwell J, Malcolm B R, et al. (2012) About the structure of cellulose:
384 debating the Lindman hypothesis. *Cellulose* 19 (3): 589-598.
385 <https://doi.org/10.1007/s10570-012-9691-7>
- 386 Grimme S (2006) Semiempirical GGA-type density functional constructed with a long-range
387 dispersion correction. *J Comput Chem* 27 (15), 1787-99.
388 <https://doi.org/10.1002/jcc.20495>
- 389 Grimme S, Antony J, Ehrlich S, Krieg H (2010) A consistent and accurate ab initio
390 parametrization of density functional dispersion correction (DFT-D) for the 94 elements
391 H-Pu. *J Chem Phys* 132 (15): 154104. <https://doi.org/10.1063/1.3382344>
- 392 Huggins M L (1971) 50 Years of Hydrogen Bond Theory. *Angew. Chem., Int. Ed.*
393 *Engl.*10(3):147-152. <https://doi.org/10.1002/anie.197101471>

394 Jarvis M C (2023) Hydrogen bonding and other non-covalent interactions at the surfaces of
395 cellulose microfibrils. *Cellulose* 30: 667-687. <https://doi.org/10.1007/s10570-022-04954-3>
396 Jeffrey G (1968) The refinement of the crystal structures of β -D-glucose and cellobiose. *Acta*
397 *Crystallogr., Sect. B: Structural Crystallography and Crystal Chemistry* 1968, 24 (6), 830-838.
398 <https://doi.org/10.1107/S0567740868003250>
399 Langan P, Nishiyama Y, Chanzy H (1999) A revised structure and hydrogen-bonding system in
400 cellulose II from a neutron fiber diffraction analysis. *J. Am. Chem. Soc* 121 (43): 9940-9946.
401 <https://doi.org/10.1021/ja9916254>
402 Lindman B, Karlström G, Stigsson L (2010) On the mechanism of dissolution of cellulose.
403 *Journal of molecular liquids* 156 (1): 76-81.
404 <https://doi.org/10.1016/J.MOLLIQ.2010.04.016>
405 Lindman B, Medronho B, Alves L Norgren M, Nordenskiöld L (2021) Hydrophobic interactions
406 control the self-assembly of DNA and cellulose. *Q. Rev. Biophys* 54: 33541444.
407 <https://doi.org/10.1017/S0033583521000019>
408 Mark A, Andrey A, Cathrine B, et al. (2023) GROMACS 2023 Manual (Version 2023). Zenodo.
409 <https://doi.org/10.5281/zenodo.7588711>
410 Masella M, Flament J P (2000) Influence of cooperativity on hydrogen bond networks. *Mol*
411 *Simul* 24(1-3): 131-156. <https://doi.org/10.1080/08927020008024192>
412 Medronho B, Lindman B (2014) Competing forces during cellulose dissolution: from solvents to
413 mechanisms. *Curr Opin Colloid Interface Sci* 19 (1): 32-40.
414 <https://doi.org/10.1016/j.cocis.2013.12.001>
415 Naito P K, Ogawa Y, Sawada D, Nishiyama Y, Iwata T, Wada M (2016) X- ray crystal structure
416 of anhydrous chitosan at atomic resolution. *Biopolymers* 105 (7), 361-368.
417 <https://doi.org/10.1002/bip.22818>
418 Nishiyama Y (2018) Molecular interactions in nanocellulose assembly. *Philos. Trans. Royal Soc.*
419 *A* 376 (2112): 20170047. <https://doi.org/10.1098/rsta.2017.0047>
420 Nishiyama Y, Langan P, Chanzy H (2002) Crystal structure and hydrogen-bonding system in
421 cellulose I β from synchrotron X-ray and neutron fiber diffraction. *J. Am. Chem. Soc* 124
422 (31): 9074-9082. <https://doi.org/10.1021/ja0257319>
423 Nishiyama Y, Noishiki Y, Wada M (2011) X-ray structure of anhydrous β -chitin at 1 Å resolution.
424 *Macromolecules* 44 (4): 950-957. <https://doi.org/10.1021/ma102240r>
425 Nishiyama Y, Sugiyama J, Chanzy H, Langan P (2003) Crystal structure and hydrogen bonding
426 system in cellulose I α from synchrotron X-ray and neutron fiber diffraction. *J. Am. Chem.*
427 *Soc* 125 (47): 14300-14306. <https://doi.org/10.1021/ja0257319>
428 Ogawa Y, Naito P -K, Nishiyama Y (2019) Hydrogen-bonding network in anhydrous chitosan
429 from neutron crystallography and periodic density functional theory calculations.
430 *Carbohydr. Polym* 207: 211-217. <https://doi.org/10.1016/j.carbpol.2018.11.042>
431 Qian X (2008) The effect of cooperativity on hydrogen bonding interactions in native cellulose
432 I β from ab initio molecular dynamics simulations, *Mol Simul*, 34 (2): 183-191.
433 <https://doi.org/10.1080/08927020801961476>
434 Sikorski P, Hori R, Wada M (2009) Revisit of alpha-chitin crystal structure using high resolution
435 X-ray diffraction data. *Biomacromolecules* 10 (5): 1100-1105.
436 <https://doi.org/10.1021/bm801251e>
437 Wada M, Chanzy H, Nishiyama Y (2004) Cellulose III $_1$ Crystal Structure and Hydrogen
438 Bonding by Synchrotron X-ray and Neutron Fiber Diffraction. *Macromolecules* 37 (23):
439 8548-8555. <https://doi.org/10.1021/ma0485585>
440 Wohler M, Benselfelt T, Wägberg L, Furó I, Berglund L A, Wohler J (2023) Cellulose and the
441 role of hydrogen bonds: not in charge of everything. *Cellulose* 29 (1): 1-23.
442 <https://doi.org/10.1007/s10570-021-04325-4>