1 Introduction

Detailed crystalline structures of native cellulose have been extensively investigated in a lot of diffraction, spectroscopic, and computational studies. Those studies have indicated that hydrogen bonds between the OH groups of the glucosyl residues are considerably important for the construction of the cellulose crystal. The major network is constructed by the O6-H···O3' and O6-H···O2' intermolecular hydrogen bonds in both cellulose Iα and Iβ. In spite of the above structural information on native cellulose crystal, the strength of the hydrogen bonds has not been quantitatively discussed. It is clear, however, that quantitative information of the hydrogen bond energy plays very important roles in discussing various properties of native cellulose such as chemical stability, solubility, and mechanical strength.

In this study, we investigated the intermolecular O6-H···O3' and O6-H···O2' hydrogen bond energies of several cellulose models at the MP2 and DFT(B3LYP) levels of theory. This study, although several DFT studies have been made for structural analysis of native cellulose crystals, presents the first theoretical calculations to quantitatively evaluate the hydrogen bond energies. We employed one-sheet cellulose models shown in Fig. 2, where two or three chains of dimer to octamer are bound together via the intermolecular hydrogen bonds. For convenience, we name these models as 2×2mer, 3×6mer etc., where the first and the second numbers represent the chain dimension and the degree of polymerization, respectively. Although the cellulose Iα and Iβ allomorphs cannot be defined with these models due to the lack of the intersheet interaction, this also means that the present computational results correspond to the average of the hydrogen bond energies of the cellulose Iα and Iβ allomorphs – the present results will carry useful information for cellulose chemistry.
2 Computational Details

The DFT with the B3LYP functional was used for geometry optimization. The 6-31G(d) basis sets were used for C, O, and H atoms, where a diffuse function was added to all O atoms and a p-polarization function was added to H atoms of all OH groups. This basis set system is named BS-I, hereafter. The initial structure of the geometry optimization was selected according to the crystallographic native cellulose structures reported in the X-ray and neutron diffraction studies. For the 2×2mer model, we carried out frequency calculation and ascertained that the optimized geometry exhibited no imaginary frequency. Potential energy was evaluated by the MP2 and DFT/B3LYP methods with the DFT(B3LYP)-optimized geometries, where better basis set system (BS-II) was used. In BS-II, 6-311G(d) basis sets were used for all atoms, where a diffuse function was added to all O atoms and a p-polarization function was added to H atoms of all OH groups. The binding energy between the chains was evaluated as a difference in the potential energy between the cellulose model and the sum of the two or three single chain models. The basis set superposition error (BSSE) was estimated with the counterpoise method. The evaluated energy is the sum of the O6-H···O3' and O6-H···O2' intermolecular hydrogen bond energies. All calculations were carried out with Gaussian 03 program package.

3 Results and Discussion

Before investigating the hydrogen bond energy, we wish to discuss the geometry optimized at the DFT(B3LYP)/BS-I level of theory. Figure 3 compares the optimized geometry of the inner residues R3 and R4 of the 3×6mer model with the experimental geometries of cellulose Iα and Iβ. The DFT(B3LYP)-optimized bond lengths are in good agreement with the experimental ones. The dihedral angles $\psi$ and $\phi$ are also important geometrical parameters, as they determine the relative positions of the two glucosyl residues; see Fig. 1 for the definitions of $\psi$ and $\phi$. The optimized values, $\psi$={-92.9°, 96.7°}, are in the range of experimental values reported for cellulose Iα (ψ=-97.7° and ϕ=99.2°) and Iβ (ψ=-88.8° and ϕ=94.7° for center chain and ψ=-98.4° and ϕ=90.5° for corner chain). The calculated bond lengths and bond angles (in parenthesis) of the O6-H···O3' and the O6-H···O2' hydrogen bonds are 1.835 to 1.836 Å (164°) and 2.812 to 2.813 Å (124°), respectively, as shown in Table 1. These values are also in the range of the experimental values reported for cellulose Iα and Iβ. We obtained similar results also for the smallest 2×2mer model, as shown in Fig. 4(A). All these results indicate that the calculated geometry is the average of cellulose Iα and Iβ. This averaged geometry is reasonable because the one-sheet models employed do not have intersheet interaction, which is a main factor determining the difference in the geometry between cellulose Iα and Iβ. We therefore concluded that the DFT(B3LYP) geometry optimization well reproduced the structure of molecular sheets of native cellulose. Several theoretical studies have indicated that twist of the molecular sheet occurs during the geometry optimization and similar twist was also observed in this study, especially in the case of the models with larger chain length and dimension (see 3×6mer and 2×8mer in Fig. 2).

We then investigated the binding energy between the constituent chains in the 2×2mer model, as this is small enough to evaluate the hydrogen bond energy at the MP2/BS-II level. As presented in Table 2, the
Evaluation of OH–O Type Hydrogen Bond Energy in Native Cellulose by Quantum Chemical Calculations

Fig. 3  Geometry of the internal residues R3 and R4 of the 3×6mer model optimized at the DFT(B3LYP)/BS-I level of theory and experimental geometries of cellulose Iα and Iβ.  a) See Fig. 1 for the definition of ψ and φ. b) See Fig. 2 for the residues R3 and R4.

Table 1. Distances (Å) and angles (degree in parenthesis) of the intermolecular hydrogen bonds of the inner residues R3 and R4 of the 3×6mer model calculated at the DFT(B3LYP)/BS-I level and their experimental values of cellulose Iα and Iβ determined with neutron diffraction method.

|                      | Intermolecular hydrogen bond |          |          |
|----------------------|-----------------------------|----------|----------|
|                      | O6–H···O3'                  | O6–H···O2'|
| Calcd.               |                             |          |          |
| Residues A           | 1.836 (164.4)               | 2.812 (123.4) |
| Residues B           | 1.835 (164.3)               | 2.813 (123.7) |
| Expt.                 |                             |          |          |
| Cellulose Iα         |                             |          |          |
| Residue A            | 1.853 (154.0)               | 2.881 (134.8) |
| Residue B            | 2.176 (122.2)               | 2.791 (141.3) |
| Cellulose Iβ         |                             |          |          |
| Origin chain         | 1.779 (165.1)               |          |          |
| Center chain         | 2.040 (156.1)               | 2.544 (125.0) |

Table 2. Binding energies (BEs, kcal/mol) of the 2×2mer and the 2-deoxy-2×2mer models.

|                      | 2×2mer                  | 2-deoxy-2×2mer |
|----------------------|-------------------------|----------------|
|                      | MP2                     | DFT (B3LYP)    | MP2             | DFT (B3LYP)    |
| BSSE                 | -5.0                    | -1.6           | -4.1            | -1.5           |
| Zero-point energy    | -2.0                    | -2.0           | -1.8            | -1.8           |
| Corrected BE        | 16.7 (8.4)^a            | 11.5 (5.8)     | 10.1 (5.1)      |

a) Zero-point energy was calculated at the DFT(B3LYP)/BS-I method.
b) Binding energy per one hydroxymethyl group is shown in the parenthesis.
It is important to evaluate how much each of the O6-H···O3' and O6-H···O2' intermolecular hydrogen bonds contributes to the total binding energy. We investigated a 2-deoxygenated 2×2mer model shown in Fig. 4(B). This model contains the O6-H···O3' intermolecular hydrogen bond similar to that of the 2×2mer model but lacks the O6-H···O2' intermolecular hydrogen bond. As the O6-H···O3' intermolecular hydrogen bond energy corresponds to the binding energy of the 2-deoxygenated 2×2mer model, the O6-H···O2' intermolecular hydrogen bond energy can be estimated as the difference in the binding energy between the 2×2mer and 2-deoxygenated 2×2mer models. As shown in Table 2, the binding energy per one hydroxymethyl group of the 2-deoxygenated 2×2mer model was 5.8 kcal/mol at the MP2 level and 5.1 kcal/mol at the DFT(B3LYP) level. The difference in the binding energy per one hydroxymethyl group between the 2×2mer and the 2-deoxygenated 2×2mer models was 2.6 kcal/mol at the MP2 level and 2.5 kcal/mol with at the DFT(B3LYP) level. Thus, the intermolecular hydrogen bond energy was evaluated to be 5.8 kcal/mol for the O6-H···O3' bond and 2.6 kcal/mol for the O6-H···O2' bond. Note that the O6-H···O3' hydrogen bond is considerably stronger than the O6-H···O2' hydrogen bond. This result is consistent with the geometrical features that the O6-H···O3' distance (~1.9 Å) is much shorter than the O6-H···O2' distance (~3.1 Å); see Fig. 3 and Table 1. It should be mentioned that the molecular sheets in the 2-deoxygenated model are more intensively twisted than those of the 2×2mer one. This suggests that the above energy counted as hydrogen bond energy actually involves the energy originated from the difference in the twisting. The real native cellulose is much larger polymer with a larger chain dimension than the 2×2mer model. We need to examine whether or not the hydrogen bond energy is different between the 2×2mer model and the real native cellulose. We investigated the binding energy of the larger models with the DFT(B3LYP) method, as the DFT(B3LYP) method presents reliable interaction energy in this system; see above. As shown in Table 3, the value of the binding energy increased as the size of the model became larger; for instance, the

| Model      | Binding Energy (BE, kcal/mol) | Number of CH2OH groups | Binding Energy per CH2OH group (BE per CH2OH group) |
|------------|-------------------------------|-------------------------|---------------------------------------------------|
| 2×2mer     | 17.2                          | 2                       | 8.6                                               |
| 2×4mer     | 31.7                          | 4                       | 7.9                                               |
| 2×6mer     | 54.3                          | 6                       | 9.1                                               |
| 2×8mer     | 73.4                          | 8                       | 9.2                                               |
| 3×6mer     | 109.8                         | 12                      | 9.2                                               |

* Zero-point correction was not carried out due to the computational cost for the frequency calculation.

Table 3. DFT(B3LYP)/BS-1-calculated binding energy (BE, kcal/mol) and the number of the hydroxymethyl (CH2OH) groups involved in the intermolecular interaction in the employed models.

Fig. 4 Geometries of the 2×2mer and 2-deoxygenated 2×2mer models optimized at the DFT(B3LYP)/BS-1 level of theory. a) See Fig. 1 for the definition of Y', Y'', φ, and φ'.
binding energy was 17.2 kcal/mol for the 2×2mer model but increased to 109.8 kcal/mol when going to the 3×6mer model. This is because the number of the intermolecular hydrogen bonds increases with the increase in the size of the model. The binding energy per hydroxymethyl group was evaluated by dividing the total binding energy by the number of the hydroxymethyl groups involved in the intermolecular interaction; for instance, it was 8.6 kcal/mol for the 2×2mer model and 9.2 kcal/mol for the 3×6mer model, as shown in Table 3. Note that the DFT(B3LYP)/BS-I- calculated binding energy per one hydroxymethyl group was almost constant in the employed models, although 0.6 kcal/mol increase in the binding energy was observed upon going from the 2×2mer model to the 3×6mer one. This result indicates that the intermolecular hydrogen bond energy evaluated here does not significantly depend on the size of the model. Thus, the value of the intermolecular hydrogen bond energy of the real native cellulose can be discussed on the basis of those in the 2×2mer model, when the 0.6 kcal/mol difference is regarded as negligible. Note that the hydrogen bond energies of Tables 2 must be employed for discussion of cellulose as the zero-point energy is taken into account in Table 2.

4 Conclusions

We successfully evaluated the O6-H···O3’ and the O6-H···O2’ intermolecular hydrogen bond energies of cellulose model, employing various one-sheet models. The total intermolecular hydrogen bond energy was calculated to be 8.4 kcal/mol per hydroxymethyl group at the MP2 level and 7.6 kcal/mol at the DFT(B3LYP) level for the 2×2mer model, which can be applied to discussion of the intermolecular hydrogen bond energy of the real native cellulose. The energy contribution of the short O6-H···O3’ hydrogen bond was estimated to be 5.8 (5.1) kcal/mol, where the MP2- and DFT(B3LYP)-calculated values were presented without parenthesis and in parenthesis, respectively. This is about 70 % of the total hydrogen bond energy. The long O6-H···O2’ hydrogen bond provides the contribution of 2.6 kcal/mol at the MP2 level and 2.5 kcal/mol at the DFT(B3LYP) level, which is about one-half of the O6-H···O3’ hydrogen bond energy. These intermolecular hydrogen bond energies are considered to be average of those of cellulose Iα and Iβ.

This work was financially supported by a Grant-in-Aid for Young Scientists (B) (No. 17K18008), Grant-in-Aid for JSPS Fellows (No. 20.1219), Grant-in-Aids for Specially Promoted Research No. 22000009, Grand-Challenge project from the Ministry of Education, Science, Sports, and Culture. SGI Altix4700 workstations of Institute for Molecular Science (Okazaki, Japan) and PC cluster computers of our laboratory were used.

References

1) A review of crystalline structure of cellulose: P. Zugtenmaier, “Conformation and packing of various crystalline cellulose fibers”, Progress in Polymer Science, Vol.26, pp. 1341-1417 (2001).
2) K. H. Gardivert and J. Blackwell, “The structure of native cellulose”, Biopolymers, Vol.13, pp. 1975-2001 (1974).
3) A. Sacko and R. Muggli, “Packing analysis of carbohydrates and polysaccharides. III. valonia cellulose and cellulose II”, Macromolecules, Vol.4, pp. 486-494 (1974).
4) (a) J. Sugiyama, T. Okano, H. Yamamoto and F. Horii, “Transformation of Valonia cellulose crystals by an alkaline hydrothermal treatment”, Macromolecules, Vol.23, pp. 3196-3198 (1990). (b) J. Sugiyama, R. Vuong and H. Chanzy, “Electron diffraction study on the two crystalline phases occurring in native cellulose from an algal cell wall”, Macromolecules, Vol.24, pp. 4168-4175 (1991).
5) (a) Y. Nishiyama, J. Sugiyama, H. Chanzy and P. Langan, “Crystal structure and hydrogen bonding system in cellulose Iα from synchrotron X-ray and neutron fiber diffraction” Journal of American Chemical Society, Vol.125, pp.14300-14306 (2003). (b) Y. Nishiyama, P. Langan and H. Chanzy, “Crystal structure and hydrogen-bonding system in cellulose Iβ from synchrotron X-ray and neutron fiber diffraction” Journal of American Chemical Society, Vol.124, pp. 9074-9082 (2002).
6) T. Imai, J. Sugiyama, T. Itoh and F. Horii, “Almost pure Iα Cellulose in the cell wall of glutucoestis” Journal of Structural Biology, Vol.127, pp. 248-257 (1999).
7) Y. Nishiyama, G. P. Johnson, A. D. French, V. T. Forsyth and P. Langan, “Neutron crystallography, molecular dynamics, and quantum mechanics studies of the nature of hydrogen bonding in cellulose Iβ”, Biomacromolecules, Vol.9, pp. 3133-3140 (2008).
8) R. H. Atalla and D. L. VanderHart “Native cellulose: a composite of two distinct crystalline forms” Science, Vol.223, pp. 283-285 (1984).
9) P. S. Belton, S. F. Tanner, N. Cartier and H. Chazy, “High-resolution solid-state carbon-13 nuclear magnetic resonance spectroscopy of tunicin, an animal cellulose” Macromolecules, Vol.22, pp. 1615-1617 (1989).
10) (a) F. Horii, A. Hirai and R. Kitamaru, “Solid-state high-resolution 13C NMR studies of regenerated cellulose samples with different crystallinities” Polymer Bulletin, Vol.8, pp. 167-170 (1982). (b) F. Horii, A. Hirai and R. Kitamaru, “Solid-state 13C-NMR study of conformations of oligosaccharides and cellulose” Polymer Bulletin, Vol.10, pp. 357-361 (1983).
11) U. Sternberg, F. T. Koch, W. Prieß and R. Witter, “Crystal structure refinements of cellulose polymorphs using solid state $^{13}$C chemical shifts” Cellulose, Vol.10, pp. 189-199 (2003).
12) H. Kono and Y. Numata, “Structural investigation of cellulose Iα and Iβ by 2D RFDR NMR spectroscopy: determination of sequence of magnetically inequivalent D-glucose units along cellulose chain” Cellulose, Vol.13, pp. 317-326 (2006).
13) P. T. Larsson, V. L. Westemarck Finkenstadt and R. P. Millane, “Crystal structure of valonia cellulose II”, Macromolecules, Vol.31, pp. 7776-7783 (1998).
14) D. L. VanderHart and R. H. Atalla, “Studies of microstructure in native celluloses using solid-state $^{13}$C NMR” Macromolecules, Vol.17, pp. 1465-1472 (1984).
15) A. A. Backer, W. Helbert, J. Sugiyama and M. J. Miles, “High-resolution atomic force microscopy of native valonia cellulose I microcrystals”, Structural Biology, Vol.119, pp. 129-138 (1997).
16) M. Koyama, W. Helbert, T. Imai, J. Sugiyama and B. Henrissat, “Parallel-up structure evidences the molecular directionality during biosynthesis of bacterial cellulose” Proceedings of the natural academy of sciences of the united states of America, Vol.94, pp. 9091-9095 (1997).
17) (a) A. Aabloo, A. D. French, R. H. Mikelsaar and A. J. Pertsin, “Studies of crystalline native celluloses using potential energy calculations” Cellulose, Vol.1, pp. 161-168 (1994). (b) A. Aabloo and A. D. French, “Preliminary potential energy calculations of cellulose Iα crystal structure” Macromolecular Theory and Simulations, Vol.3, pp. 185-191 (1994).
18) A. D. French and D. P. Miller, A. Aabloo, “Miniature crystal models of cellulose polymorphs and other carbohydrates” International Journal of Biological Macromolecules, Vol.15, pp. 30-36 (1993).
19) A. P. Heiner, J. Sugiyama and O. Teleman, “Crystalline cellulose Iα and Iβ studied by molecular dynamics simulation”, Carbohydrate Research, Vol.273, pp. 207-223 (1995).
20) L. M. J. Kroon-Batenburg and J. Kroon, “The crystal and molecular structures of cellulose I and II”, Glycoconjugate Journal, Vol.14, pp. 677-690 (1997).
21) R. J. Viëtor, K. Mazeau, M. Lakin and S. Pérez, “A priori crystal structure prediction of native celluloses.” Biopolymers, Vol.54, pp. 342-354 (2000).
22) (a) K. Mazeau and L. Heux, “Molecular dynamics simulations of bulk native crystalline and amorphous structures of cellulose” Journal of Physical Chemistry B, Vol.107, pp. 2394-2403 (2003). (b) K. Mazeau, “Structural micro-heterogeneities of crystalline Iβ-cellulose”, Cellulose, Vol.12, pp. 339-349 (2005). (c) K. Mazeau and A. Rivert, “Wetting the (110) and (100) surfaces of Iβ cellulose studied by molecular dynamics” Biomacromolecules, Vol.9, pp. 1352-1354 (2008).
23) M. Bergenstrëhl, L. A. Berglund and K. Mazeau, “Thermal response in crystalline Iβ cellulose: a molecular dynamics study” Journal of Physical Chemistry B, Vol.111, 9138-9145 (2007).
24) I. Simon, H. Scheraga and R. S. John Manley “Structure of cellulose. 1. Low-energy conformations of single chains” Macromolecules, Vol. 21, pp. 983-990 (1988).
25) I. Simon, L. Glasser and H. Scheraga, “Structure of cellulose. 1. Low-energy conformations of single chains” Macromolecules, Vol.21, pp. 999-998 (1988).
26) T. Uto and T. Yui “DFT optimization of isolated molecular chain sheet models constituting native cellulose crystal structures” ACS Omega, Vol.3, pp. 8050-8058 (2018).
27) T. Uto, S. Mawatari and T. Yui, “Theoretical study of the structural stability of molecular chain sheet models of cellulose crystal allomorpha”, Journal of Physical Chemistry B, Vol.118, pp. 9313-9321 (2014).
28) A. D. Becke, “Density-functional exchange-energy approximation with correct asymptotic behavior” Physical Review A, Vol.38, pp. 3098-3100 (1988).
29) (a) S. F. Boys and F. Bernardi, “The calculation of small molecular interactions by the differences of separate total energies. Some procedures with reduced errors” Molecular Physics, Vol.19, 553-566 (1970). (b) S. Simon, M. Duran and J. J. Dannenberg, “How does basis set superposition error change the potential surfaces for hydrogen - bonded dimers?” Journal of Chemical Physics Vol.105, 11024-11031 (1996).
30) Gaussian 03, Revision C.02, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople, Gaussian, Inc., Wallingford CT, 2004.