THE JOURNAL OF PHYSICAL CHEMISTRY A

ARTICLE

pubs.acs.org/JPCA

# Ab Initio Study of Structure and Interconversion of Native Cellulose Phases

Tomáš Bučko,[*,†,‡,#] Daniel Tunega,[§,‡] János G. Ángyán,[*,∥] and Jürgen Hafner[⊥]

[†]Department of Physical and Theoretical Chemistry, Faculty of Natural Sciences, Comenius University, Mlynská Dolina, SK-84215 Bratislava, Slovakia

[‡]Slovak Academy of Sciences, Institute of Inorganic Chemistry, Dubravska cesta 9, SK-84236 Bratislava, Slovakia

[§]Institute for Soil Research, University of Natural Resources and Life Sciences, Peter-Jordan-Strasse 82, A-1190 Vienna, Austria

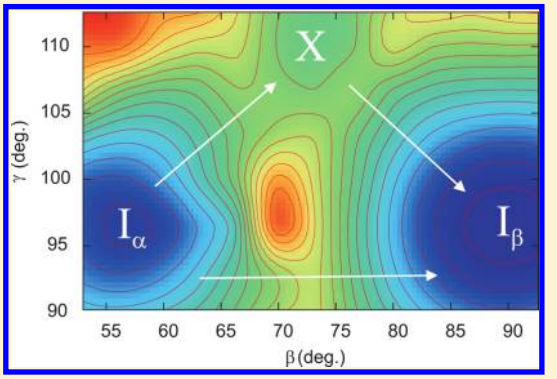
[∥]Équipe de Modélisation Quantique et Cristallographique, CRM2, Institut Jean Barriol, UMR 7036, CNRS—Nancy-Université, F-54506 Vandoeuvre-lès-Nancy, France

[⊥]Universität Wien, Wien 1090, Austria

[#]Fakultät für Physik and Center for Computational Materials Science, Universität Wien, Sensengasse 8/12, Wien 1090, Austria

Supporting Information

**ABSTRACT:** Dispersion-interaction corrected DFT simulations are performed to study the structure of two allomorphs of native cellulose I. Good agreement between theoretical and experimental data is achieved. Two H-bond patterns, previously identified experimentally, are examined for both allomorphs. The transition mechanism for the conversion between the phase $I_\alpha$ and $I_\beta$ is studied by means of constrained relaxations. New metastable intermediate phase occurring on the $I_\alpha \rightarrow I_\beta$ route is identified, and the corresponding structural data are reported.




## 1. INTRODUCTION

Cellulose, an unbranched homopolysaccharide, is one of the most abundant renewable biopolymers on the Earth. The cellulose structure consists of linear chains of D-glucopyranose units linked via $\beta(1\rightarrow4)$-glycosidic bonds (Figure 1). Four different basic crystalline forms of the cellulose have been identified so far.[1] In the native cellulose I, occurring, e.g., in woody cell walls, the neighboring chains are connected by interchain hydrogen bonds giving rise to a layered structure, in which van der Waals forces are the dominant stacking interactions. Two known allomorphs of the natural cellulose, $I_\alpha$ and $I_\beta$, coexist in natural materials, and their relative ratio depends on the source of cellulose.[1]

The crystal structures of cellulose $I_\alpha$ and $I_\beta$ have been studied by means of synchrotron X-ray and neutron fiber diffraction by Nishiyama et al. and described in a series of papers.[2−4] Two different intrasheet H-bond patterns have been observed[2,3] (Figure 2). The question has been raised in the literature, whether these two mutually exclusive patterns occur as static configurations in different regions of the crystal or whether they are dynamically interchangeable. Nishiyama et al.[4] performed neutron diffraction measurements of $I_\beta$ at two different temperatures ($T$ = 295 K and $T$ = 15 K). It was shown that the relative occupancies of the two H-bond schemes (∼80% A, ∼20% B) are basically independent of the temperature. Furthermore, Hartree−Fock and Perdew−Burke−Ernzerhof (PBE)[5] DFT calculations of one-layer models of cellulose, reported in the same work, have shown that the bonding scheme A is energetically more favorable than the alternative pattern, the energy differences being 31 kJ/mol (HF) and 44 kJ/mol (PBE), respectively (note that 1 mole is related to the number of disaccharide units). Owing to this large energy difference, it is unlikely that the two H-bond schemes can be dynamically interconverted in typical experimental conditions. Moreover, the lattice parameter $b$ calculated for the scheme B is significantly larger (by ∼0.3 Å) than the experimental value, while the parameter $b$ computed for the H-bonding scheme A is in reasonable agreement with the experiment. The energetic instability of the microcrystal with the H-bond scheme B has also been predicted by the theoretical study of Mazeau.[6] All these results indicate that structural domains with the H-bond scheme B may exist only as static configurations at defects, interfaces, or surfaces. In this work, among other features, the effect of the H-bond pattern on the structure and on the energetics of phases $I_\alpha$ and $I_\beta$ will be examined by means of DFT simulations.

The cellulose $I_\alpha$ can be transformed into phase $I_\beta$ by a hydrothermal treatment at $T$ = 260 °C in an alkaline solution.[7,8] The

**Received:** June 21, 2011
**Revised:** July 22, 2011
**Published:** July 29, 2011

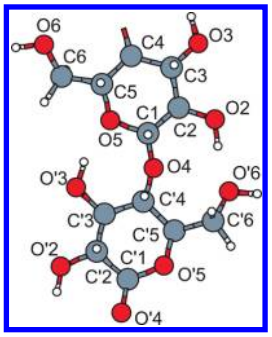



**Figure 1.** Two D-glucopyranose units linked via $\beta(1-4)$-glycosidic bonds. The atomic labeling has been adapted from ref 35. The prime has been added to distinguish atoms in the two D-glucopyranosyl units.

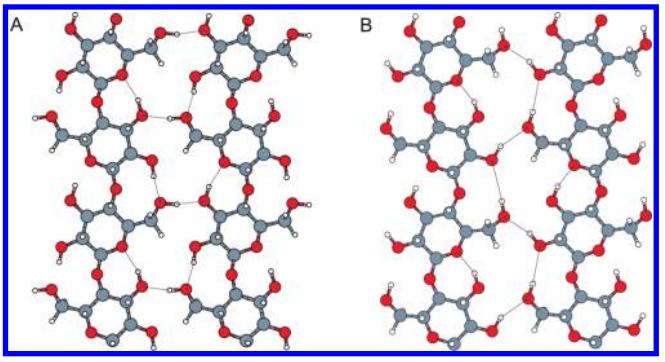



**Figure 2.** Two different H-bond networks observed experimentally in cellulose $I_\beta$ (refs 2 and 3).

transition occurs also in an inert atmosphere of helium gas at a temperature above 260 °C.[9] As this process is irreversible, cellulose $I_\alpha$ is considered as a metastable phase. It follows from the comparison of cellulose $I_\alpha$ and $I_\beta$ structures that the phase transformation can be realized either via rotation of cellulose chains around their axis or via a relative slippage of chains past one another.[3] On the basis of force-field molecular dynamics simulations, Hardy and Sarko[10] proposed a break–slip model, in which the transition is initiated by a heating-induced H-bond breakage, freeing thus the chains in the $I_\alpha$ phase for rotation and sliding into the more stable $I_\beta$ allomorph. Note, however, that the full transformation path was not reported in ref 10. It has been shown in an experimental work by Wada et al.[11] that the thermally induced lattice expansion indeed triggers the transformation from $I_\alpha$ to $I_\beta$. However, the observed lattice expansion is insufficient to allow a 180° rotation of the molecular chains. Hence a relative slippage of the cellulose sheets remains the only realistic scenario for the transformation. Despite its relative simplicity, this transition model has not yet been examined in a theoretical study.

In this work we report a detailed study on the mechanism of interconversion between native cellulose phases $I_\alpha$ and $I_\beta$ using periodic DFT calculations with dispersion corrections. The paper is organized as follows: in section 2, the methodology and simulation details used in this study are introduced, the structural optimizations of phases $I_\alpha$ and $I_\beta$ are described in section 3, the mechanism of the phase transition between $I_\alpha$ and $I_\beta$ is explored in section 4, and the most important results are summarized in section 5.

## 2. METHODOLOGY

**2.1. Total Energy Calculations.** Periodic DFT calculations have been performed using the VASP code.[12−15] The Kohn–Sham equations have been solved variationally in a plane-wave basis set using the projector augmented wave (PAW) method of Blöchl,[16] as adapted by Kresse and Joubert.[17] The exchange-correlation energy was described by the PBE generalized gradient approximation.[5] The plane-wave cutoff was set to 800 eV, a k-point mesh with $2 \times 2 \times 2$ points has been used to sample the Brillouin zone. The convergence criterion for the electronic self-consistency cycle, measured by the change in the total energy between successive iterations, was set to $10^{-7}$ eV/cell. In order to account for London dispersion interactions, the PBE-D2 correction scheme proposed by Grimme[18] has been used. The standard set of parameters proposed in ref 18 has been used, with a cutoff radius for the pair-potential describing the dispersion forces set to 40 Å. The implementation of the D2 correction in VASP has been detailed in ref 19.

**Table 1. Experimental[4] and Theoretical Lattice Parameters for Cellulose $I_\beta$[a]**

| method | $V$ (Å$^3$) | $a$ (Å) | $b$ (Å) | $c$ (Å) | α (deg) | $\beta$ (deg) | $\gamma$ (deg) |
|---|---|---|---|---|---|---|---|
| exp (15 K) | 644 | 7.64(1) | 8.18(1) | 10.37(1) | 90.0 | 90.0 | 96.6(1) |
| exp (295 K) | 656 | 7.76(1) | 8.20(1) | 10.37(1) | 90.0 | 90.0 | 96.5(1) |
| PBE (A) | 716 | 8.37 | 8.23 | 10.45 | 90.0 | 90.0 | 96.0 |
| PBE-D2 (A) | 636 | 7.57 | 8.14 | 10.39 | 90.0 | 90.0 | 96.5 |
| PBE-D2 (B) | 655 | 7.51 | 8.55 | 10.30 | 90.0 | 90.1 | 98.2 |

[a] Theoretical results are reported for two alternative H-bonding schemes A and B; see Figure 2. The values in parentheses are standard deviations for experimentally measured lattice parameters. The $R$ factors for the refined experimental data are 0.241 (15 K) and 0.244 (295 K).

**2.2. Structural Optimization.** Structural optimizations have been performed with the optimization engine GADGET[20] based on delocalized internal coordinates.[21] GADGET allows one to perform atomic and lattice optimizations with arbitrary geometric constraints, which can be used to study relaxed cross sections of the full multidimensional potential energy surface of a periodic solid. The structures were considered to be relaxed when all components of the gradient vector were smaller than $3 \times 10^{-4}$ atomic units. In the optimized structures, all stress–tensor components were smaller than 0.05 kbar. Unless otherwise stated, all degrees of freedom, i.e., atomic positions and lattice parameters have been relaxed. The use of geometry constraints will be always explicitly indicated.

## 3. STRUCTURAL CHARACTERIZATION

**3.1. Cellulose $I_\beta$.** The crystal structure of the phase $I_\beta$ has been studied by means of X-ray and neutron fiber diffraction by Nishiyama at al.[4] The space group of cellulose $I_\beta$ is $P112_1$, the crystallographic cell contains two formula units ($Z = 2$). The experimental lattice parameters determined in low-temperature measurements ($T = 15$ K) are $a = 7.64$ Å, $b = 8.18$ Å, $c = 10.37$ Å, and $\gamma = 96.5°$. Note that when the measurements were performed at room temperature ($T = 295$ K) the value of $a$ increased to 7.76 Å[4] while the values of other lattice parameters changed less significantly (Table 1). This thermal sensitivity of the lattice parameter $a$ should be kept in mind when comparing zero-temperature calculations with experiment. In the cellulose $I_\beta$ unit cell, the D-glucose chains are parallel to the $c$ direction and the neighboring chains form H-bonded sheets parallel to the $(c,b)$ plane (Figure 3). The cellulose sheets are stacked in the direction parallel to the vector **a**; the average spacing between the two cellulose sheets is $d = 1/2a$. The experimental values for this parameter are 3.82 and 3.88 Å for $T = 15$ and 273 K, respectively.

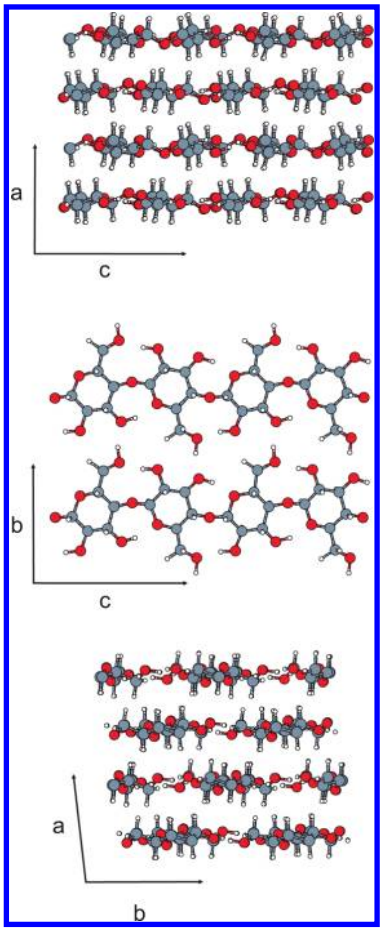



**Figure 3.** Monoclinic cell geometry of cellulose $I_\beta$ with the H-bond network A.

The results presented in this section were obtained using a simulation cell corresponding to one crystallographic cell of cellulose $I_\beta$. It is well-known that standard DFT functionals fail to describe van der Waals forces, which play a crucial role in the stabilization of the three-dimensional lattice structure of cellulose. For instance, the structure of cellulose $I_\beta$ optimized using the standard PBE functional differs significantly from the experimental one (see Table 1): the lattice parameter $a$ is overestimated by 0.72 Å and the intersheet spacing is too large by 0.36 Å compared to experiment.[4] These results demonstrate that the standard PBE functional is not suitable for an accurate description of the cellulose structure. In order to fix this problem, a semiempirical correction for the London dispersion interactions (PBE-D2) proposed by Grimme[18] has been used in this work. We have shown in our recent study[19] that this method improves significantly the prediction of the structure and the elastic properties of a wide range of materials including cellulose $I_\beta$. Very recently, the PBE-D2 method has been used by Li et al.[22] to study the crystal structure of cellulose I polymorphs.

The lattice parameters computed using the PBE-D2 method are compared with experimental data in Table 1. For the structure with the H-bond pattern A (Figure 2A), the predicted lattice parameters are in a very good agreement with low temperature experimental data. The computed intersheet spacing $d$ is 3.78 Å and the estimated error is well below 1% for all lattice parameters. The structure with the alternative H-bond pattern B (Figure 2B) is energetically less favorable than the pattern A, the total energy difference $\Delta E_{A,B} = E_A - E_B$ being as large as ~48 kJ/mol. Moreover, the cell parameters corresponding to the H-bond pattern B differ significantly from the experiment: the parameter $b$ is overestimated by ~0.3 Å and the value of $c$ is underestimated by ~0.1 Å. Similar results have been reported in the PBE study of one cellulose layer performed by Nishiyama et al.[4] ($b$ = 8.23 Å, $c$ = 10.46 Å (scheme A), $b$ = 8.45 Å, $c$ = 10.32 Å (scheme B), $\Delta E_{A,B}$ = 44 kJ/mol). An even larger energy difference between A and B ($\Delta E_{A,B} \approx 100$ kJ/mol) has been predicted by Mazeau,[6] who used classical force-field and molecular dynamics to study the structure of microcrystals of cellulose $I_\beta$. As the probability for swapping between H-bond schemes A and B is approximately proportional to $\exp\{-\Delta E_{A,B}/k_B T\}$, the dynamical interconversion between the two H-bonding patterns is expected to be a quite rare event. As suggested by Nishiyama et al.,[4] the H-bond disorder should be essentially static and the occurrence of the pattern B is probably limited to domains where structural distortions are more likely, such as defects, surfaces, and interfaces in microfibers.

**Table 2. Selected Structural Intrachain Parameters (Dihedral Angles Φ, Ψ, χ, χ′, and Bond Angle τ) for Cellulose $I_\beta$[a]**

| | exp | | PBE-D2 | |
|---|---|---|---|---|
| | chain 1 | chain 2 | chain 1 | chain 2 |
| Φ (O5−C1−O4−C4) | −99 | −89 | −93 | −94 |
| Ψ (C1−O4−C4−C5) | −142 | −147 | −143 | −145 |
| τ (C1−O4−C4) | 115 | 116 | 118 | 118 |
| χ (O6−C6−C5−O5) | 170 | 158 | 168 | 165 |
| χ′ (O6−C6−C5−C4) | −70 | −83 | −74 | −77 |

[a] The atomic labeling is as in Figure 1.

On the basis of $^{13}$C NMR spectroscopy, Kono et al.[23,24] have shown that two equally populated nonequivalent glucopyranose residues exist in the cellulose $I_\beta$. As the unit cell of $I_\beta$ contains two different D-glucose chains with slightly different internal structures,[2] all glucopyranose units within one chain should be equivalent. This fact is correctly reproduced by our simulations. The selected internal parameters, namely, the glucosidic torsion angles Φ and Ψ, the bond angle τ, and the torsion angles representing conformation of the hydroxymethyl group χ and χ′ measured for the optimized structure with the H-bond scheme A are compared with the experimental data in Table 2. According to our simulations, the structures of the two cellulose chains in the unit cell are slightly different but the glucopyranosyl rings within one chain are equivalent. Although the computed internal parameters fall into the range of experimentally measured values, the theoretically predicted structural difference between the two chains is less significant than in the experiment. Similar conclusions have been reported in the theoretical study of Li et al.,[22] obtained by the same PBE-D2 approach implemented in the simulation package Quantum Espresso.[25] The lengths of the intra- and intersheet H-bonds for the two nonequivalent cellulose sheets are shown in panels a and b of Figure 4. As the lengths of equivalent H-bonds are very similar (the largest difference being <0.1 Å), the strength of the H-bonds in the two nonequivalent sheets should be about the same.

**3.2. Cellulose $I_\alpha$.** Cellulose $I_\alpha$ crystallizes in space group $P1$; the crystallographic unit cell is triclinic and it contains one disaccharide unit ($Z$ = 1). The experimentally measured cell parameters[3] are $a$ = 6.72 Å, $b$ = 5.96 Å, $c$ = 10.40 Å, α = 118.1°, $\beta$ = 114.8°, and $\gamma$ = 80.4°. In order to investigate the $I_\alpha \rightarrow I_\beta$ transformation, the unit cell

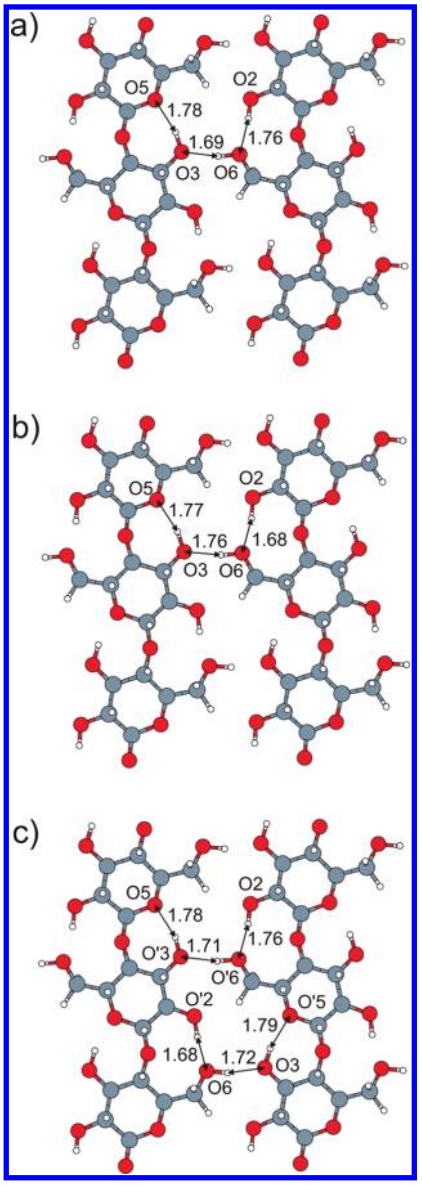



**Figure 4.** Lengths of hydrogen bonds in the two inequivalent sheets of cellulose $I_\beta$ (a, b) and in the sheet of cellulose $I_\alpha$ (c).

of the $I_\alpha$ phase had to be transformed in the setting of $I_\beta$, by defining a new set of lattice vectors related to the original lattice vectors via following transformations: $\mathbf{a}' = \mathbf{a} + \mathbf{b}$, $\mathbf{b}' = \mathbf{a} - \mathbf{b}$, and $\mathbf{c}' = \mathbf{c}$. The relation between the crystallographic and the transformed unit cell (space group *C*1) is illustrated by Figure 5. The cell parameters for the C-centered experimental unit cell are $a$ = 9.70 Å, $b$ = 8.20 Å, $c$ = 10.40 Å, $\alpha = 90°$, $\beta = 54.6°$, and $\gamma = 96.9°$. The volume of the transformed unit cell is doubled compared to the original *P*1 cell, and the transformed unit cell contains two disaccharide units ($Z' = 2$). The average intersheet spacing in $I_\alpha$ can be defined as $d = (1/2a)(1 - \cos^2(\beta) - \cos^2(\gamma))^{1/2}$. The experimental value of $d$ is 3.91 Å, which is only ~0.03 Å longer than that for the cellulose $I_\beta$ (3.88 Å) measured at a comparable temperature.

The calculations have been performed using a simulation cell corresponding to a single transformed cell described above. The cell geometry of the relaxed structure is shown in Figure 6, and the corresponding optimized values of the lattice parameters for the H-bond schemes A and B are compared with experimental

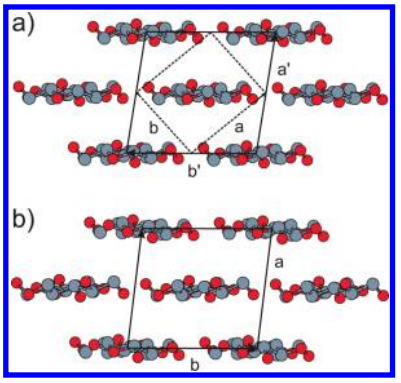



**Figure 5.** The crystallographic (dashed line) and the simulation unit cell (solid line) used in this study for cellulose $I_\alpha$ (a). For sake of comparison, the simulation unit cell (identical with the crystallographic cell) for the phase $I_\beta$ is shown in panel b. The cells are viewed along the [001] direction.

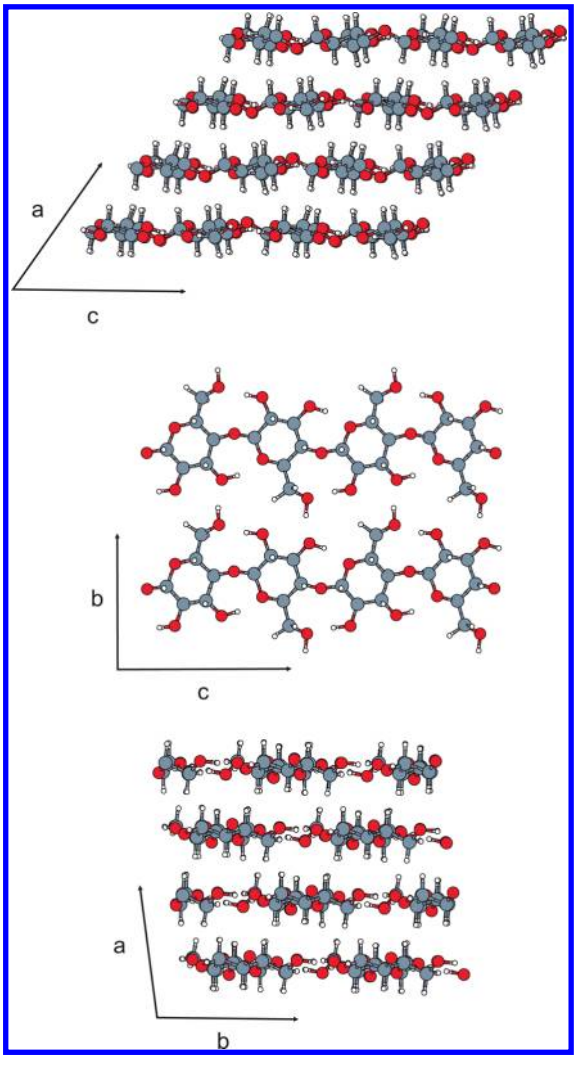



**Figure 6.** Triclinic cell geometry of cellulose $I_\alpha$ with the H-bond network A.

data in Table 3. As in the case of allomorph $I_\beta$, the H-bond pattern B is energetically strongly disfavored (the energy difference being ~48 kJ/mol) and the cell geometry is significantly

**Table 3. Lattice Parameters of Cellulose $I_\alpha$[a]**

| method | $V$ (Å$^3$) | $a$ (Å) | $b$ (Å) | $c$ (Å) | α(°) | $\beta$(°) | $\gamma$(°) |
|---|---|---|---|---|---|---|---|
| exp (293 K) | 666.7 | 9.70(1) | 8.20(1) | 10.40(1) | 90.0(1) | 54.6(1) | 96.9(1) |
| PBE-D2 (A) | 641.7 | 9.38 | 8.13 | 10.39 | 90.0 | 55.0 | 96.7 |
| PBE-D2 (B) | 658.1 | 9.30 | 8.53 | 10.30 | 89.5 | 55.2 | 98.8 |

[a] The experimental cell reported in ref 3 is modified to be comparable with the unit cell of phase $I_\beta$. Theoretical results are reported for two alternative H-bonding schemes A and B; see Figure 2. The values in parentheses are standard deviations for experimentally measured lattice parameters. The $R$ factor for the refined experimental structure is 0.178.

**Table 4. Selected Structural Intrachain Parameters (Dihedral Angles Φ, Ψ, χ, χ′, and Bond Angle τ) for Cellulose $I_\alpha$[a]**

| | exp. | | PBE-D2 | |
|---|---|---|---|---|
| | residue 1 | residue 2 | residue 1 | residue 2 |
| Φ (O5−C1−O4-C4) | −98 | −99 | −94 | −94 |
| Ψ (C1−O4−C4-C5) | −138 | −140 | −143 | −145 |
| τ (C1−O4−C4) | 116 | 116 | 118 | 117 |
| χ (O6−C6−C5-O5) | 167 | 166 | 164 | 166 |
| χ′ (O6−C6−C5-C4) | −75 | −74 | −79 | −76 |

[a] The atomic labeling is as in Figure 1.

distorted compared to the experimental structure. For the H-bonding pattern A, DFT-D2 predicts a reasonable structure. The largest error with respect to the experimental data is found for the lattice parameter $a$ (dominated by van der Waals interactions) that is underestimated by ∼3%. It should be noted, however, that the measurement reported in ref 3 has been made at $T$ = 293 K. No low temperature data, which would be more appropriate for comparison with our simulations, are available for the phase $I_\alpha$. It is therefore likely that a significant part of the difference in the computed lattice parameter $a$ is due to thermal effects. A crude estimate of the low temperature value of parameter $a$ can be obtained by assuming that the thermal lattice expansion of the cellulose $I_\alpha$ is similar to that for $I_\beta$ (see Table 1). Using a simple extrapolation scheme, we estimate that the low temperature value of $a$ is ∼9.55 Å, which is only 0.17 Å larger than the computed value. The computed intersheet spacing $d$ = 3.78 Å is identical with the value obtained for the allomorph $I_\beta$.

Using $^{13}$C spectroscopy, Kono et al.[23,24] have shown that, in analogy to the $I_\beta$ phase, two equally populated nonequivalent glucopyranosyl residues exist in cellulose $I_\alpha$ too. As the crystallographic cell for this phase contains only a fragment of one chain, it can be concluded that the nonequivalent residues must be located on the same chain. It is evident from the values of internal parameters compiled in Table 4 that our simulations reproduce this behavior correctly, although the theoretical values differ slightly from experimental results. Almost identical values for the parameters Φ, Ψ, and χ have been reported in the PBE-D2 study of Li et al.[22]

Figure 4c shows the calculated H-bond distances linking the chains in one sheet of the phase $I_\alpha$ (note that all sheets in $I_\alpha$ are symmetry equivalent). As expected, the lengths of intrasheet H-bonds involving atoms from two different glucopyranosyl units (O2H···O′6 and O′2H···O6) are not exactly the same; their difference is 0.08 Å. On the basis of their lengths, the strength of H-bonds in the optimized structure can be estimated to be about the same as those in the two different sheets in $I_\beta$.

The total energy computed for the $I_\alpha$ phase is slightly higher but comparable to the energy of the allomorph $I_\beta$, the difference being only ∼2 kJ/mol. A practically identical energy difference between $I_\alpha$ and $I_\beta$ has been reported in the PBE-D2 study of Li et al.[22] and a similar result has been obtained also in the force-field study of Neyertz et al.[26] (∼8 kJ/mol). These results are consistent with the observation that both allomorphs occur in the nature under similar conditions, but the phase $I_\beta$ is thermodynamically slightly more stable.[11]

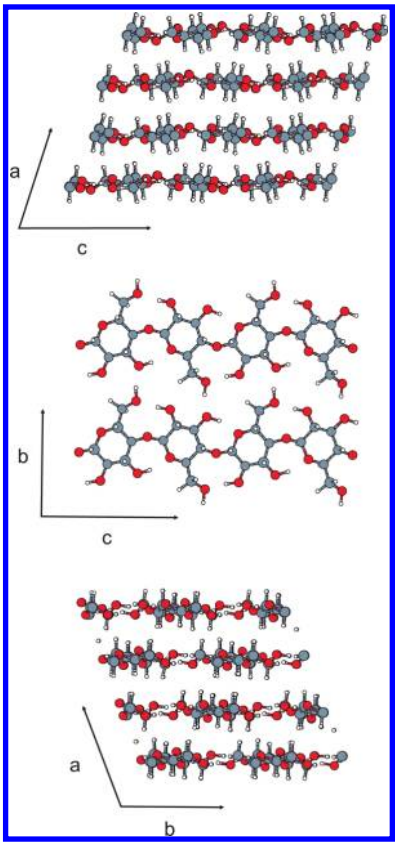



**Figure 7.** Triclinic cell geometry of metastable phase X occurring on the transition path between cellulose $I_\alpha$ and $I_\beta$.

**Table 5. Lattice Parameters and Relative Total Energies Computed for Three Phases Present on the Transformation Path $I_\alpha \rightarrow I_\beta$[a]**

| phase | $\Delta E_{tot}$ (kJ/mol) | $V$ (Å$^3$) | $a$ (Å) | $b$ (Å) | $c$ (Å) | α (deg) | $\beta$ (deg) | $\gamma$ (deg) | $d$ (Å) |
|---|---|---|---|---|---|---|---|---|---|
| α | 2 | 641.7 | 9.376 | 8.128 | 10.389 | 90.0 | 55.0 | 96.7 | 3.78 |
| X | 13 | 663.5 | 8.813 | 8.109 | 10.371 | 89.9 | 73.2 | 109.8 | 3.95 |
| $\beta$ | 0 | 635.8 | 7.566 | 8.136 | 10.394 | 90.0 | 90.0 | 96.5 | 3.78 |

[a] The newly identified metastable phase is designated X. The relative total energies are computed with respect to the most stable phase $I_\beta$.

## 4. PHASE TRANSITION

It has been recognized by Nishiyama et al.[3] that the most likely route for the $I_\alpha \rightarrow I_\beta$ transformation is a relative slippage of the cellulose sheets along each other. It follows by comparison of structures of allomorphs $I_\alpha$ and $I_\beta$ that this slippage can be associated with a change of the lattice angle $\beta$, while the parameters $b$, $c$, α, and $\gamma$ vary only negligibly. Also the intrasheet geometric parameters, such as the H-bond distances, are expected to change only moderately and no rupture of covalent or

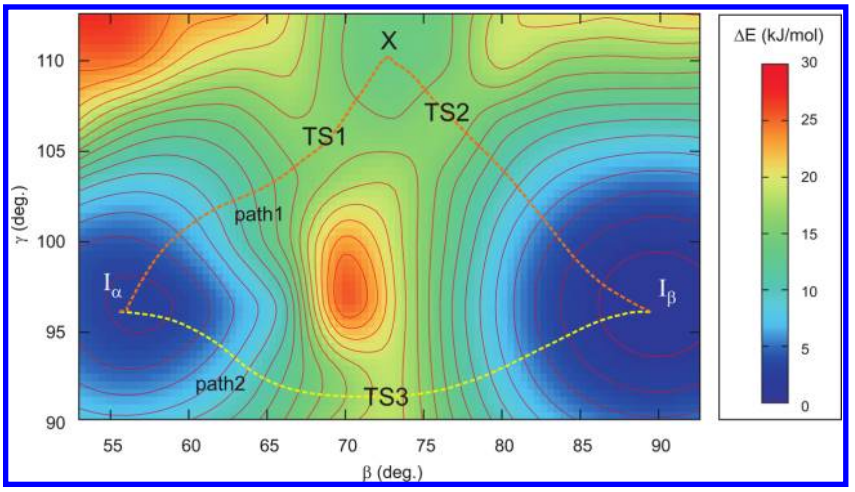



**Figure 8.** Potential energy surface for cellulose I as a function of two independent variables—lattice parameters $\beta$ and $\gamma$.

hydrogen bonds is needed for this transformation. In the rest of this section we describe a series of DFT simulations that we have performed in order to elucidate this transition mechanism.

In a first attempt to determine the energy profile for the transformation, a set of constrained relaxations has been made, where the parameter $\beta$ was kept constant for each intermediate structure. The initial configurations have been obtained by a linear interpolation between the relaxed structures of the $I_\alpha$ and $I_\beta$ phases, respectively. The resulting potential energy profile indicated the existence of a metastable intermediate phase, hereafter designated by X (see Figure 7). The structure of the newly identified phase has been refined by performing a full optimization of all atomic positions and lattice parameters. The calculated lattice geometry and energetics are compared with those for both allomorphs $I_\alpha$ and $I_\beta$ in Table 5. The .cif file containing the structure of the relaxed phase X is provided as Supporting Information. The phase X lies about midway between the two isomorphs with a value of 73° for the $\beta$ angle. Interestingly, the lattice angle $\gamma$ deviates significantly from a value of ~97° (characteristic for both $I_\alpha$ and $I_\beta$) up to 110°. The cell volume of the phase X is increased to 664 Å$^3$ and the intersheet spacing $d$ becomes 3.95 Å, i.e., ~4% larger than that for the $I_\alpha$ phase. The energy difference between phases X and $I_\beta$ is 13 kJ/mol. About half of this value (6 kJ/mol) can be attributed to the weakening of van der Waals interactions caused by increased intersheet spacing. The existence of a metastable intermediate phase on the transformation route between $I_\alpha$ and $I_\beta$ has been observed experimentally by Wada et al.[11] Unfortunately, the only relevant experimental information about this phase is that the intersheet spacing is probably increased by about 6% as compared to the $I_\alpha$ phase. This feature seems to be in a reasonable agreement with our simulations.

The existence of a metastable phase on the transformation path between $I_\alpha$ and $I_\beta$ that differs from the structures of both stable phases in the value of the lattice angles $\beta$ and $\gamma$ indicates that the transformation path is probably not a simple function of the $\beta$ angle only. As such a transformation path links three different potential energy minima ($I_\alpha$, X, and $I_\beta$), it must necessarily proceed via two different barriers. In order to learn more about the transformation mechanism, we performed exploratory Parrinello−Rahman molecular dynamics simulations[27] at low temperature ($T$ = 10 K) in combination with the slow-growth method,[28,29] in which the parameter

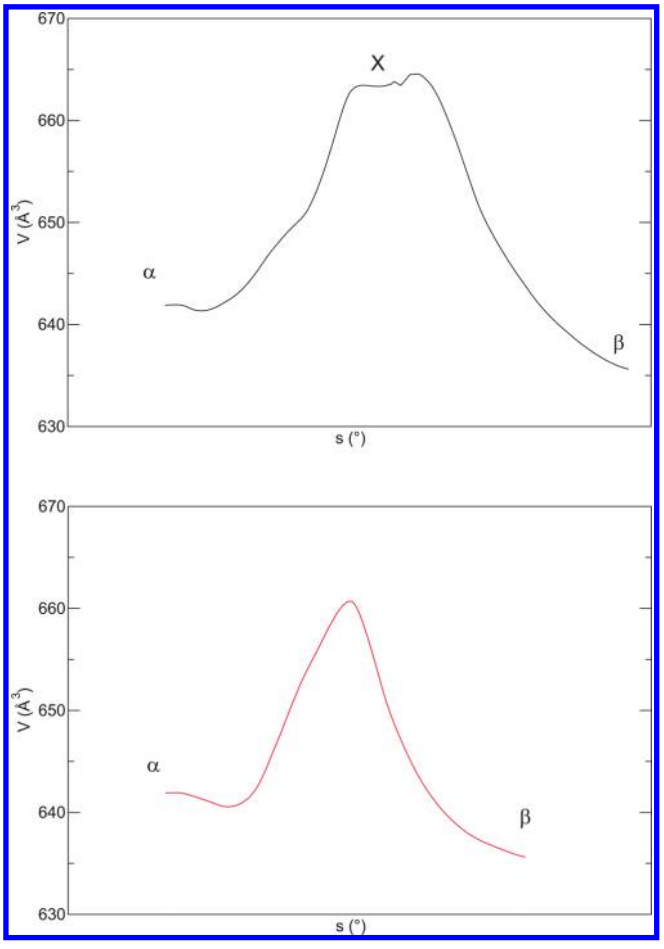



**Figure 9.** Cell volume as a function of the path length ($s$) for the two transformation mechanisms identified for the $I_\alpha$ to $I_\beta$ transition (see Figure 8): path 1 (above), path 2 (below).

$\beta$ was linearly changed from the value characteristic for phase $I_\alpha$ to that for the phase $I_\beta$. Note, that this type of simulation technique has been used in molecular chemistry to study free-energy profiles of chemical reactions.[28,29] Our exploratory MD simulations revealed the existence of an alternative transformation mechanism

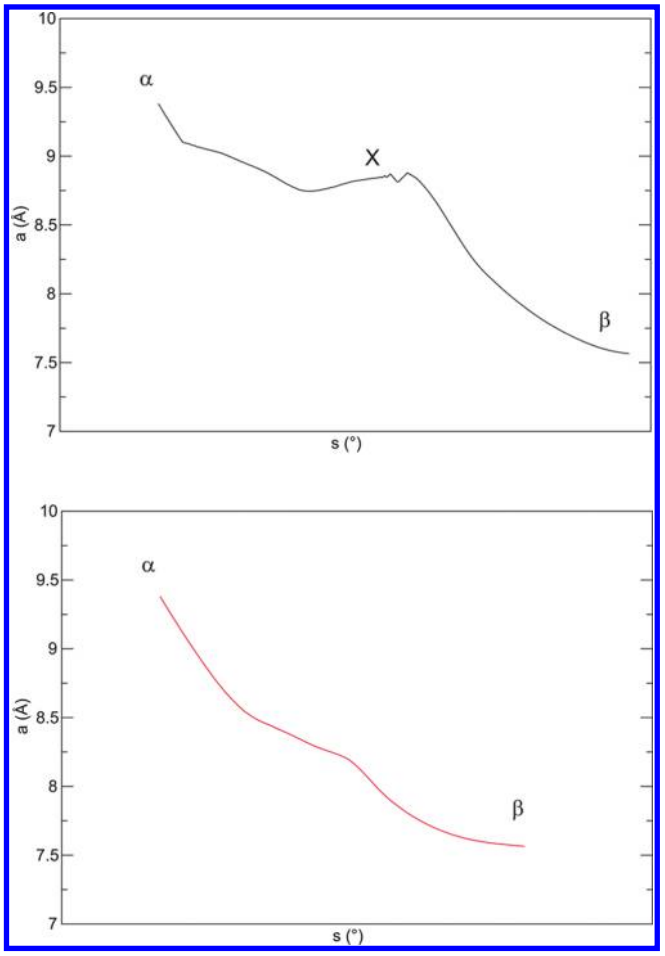



**Figure 10.** The lattice parameter $a$ as a function of the transition path length ($s$) for the two mechanisms identified for the $I_\alpha \rightarrow I_\beta$ transformation (see Figure 8): path 1 (above), path 2 (below).

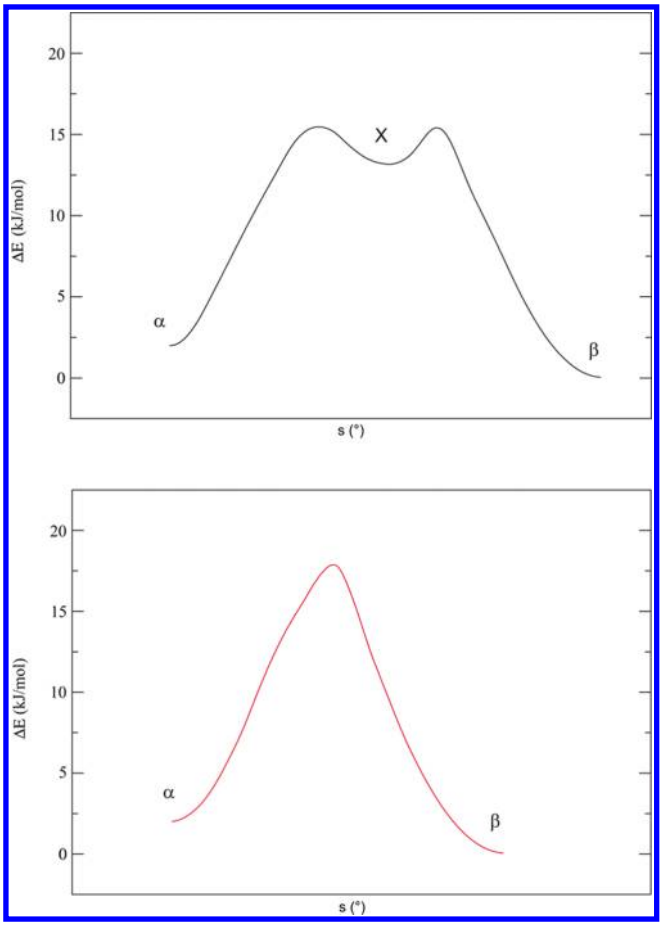



**Figure 11.** Potential energy profiles for the two mechanisms identified for the $I_\alpha \rightarrow I_\beta$ transformation (see Figure 8): path 1 (above), path 2 (below). The relative energies $\Delta E$ are computed with respect to the energy of relaxed cellulose $I_\beta$. The symbol $s$ stands for the transition path length.

(mechanism 2, vide infra) with a single potential energy barrier, in which the $\gamma$ angle decreases slightly in the transition region. All these results clearly show that the phase transformation between $I_\alpha$ and $I_\beta$ is not a simple process that could be described by the variation of the parameter $\beta$ only.

In order to explore systematically the potential energy surface relevant for the phase transformation, a series of constrained relaxations has been performed, in which the lattice angles $\beta$ and $\gamma$ were fixed, whereas all the remaining atomic and lattice degrees of freedom were free to relax. The energies were computed for 104 points organized in a regular 13 × 8 grid and characterized by different values of $\beta$ and $\gamma$. The region of configuration space sampled in such a way is defined by the intervals $53 \leq \beta \leq 93$, and $90 \leq \gamma \leq 113$. The contour plot of the resulting potential energy surface is shown in Figure 8, and an animation showing a 3D view of the surface is provided as Supporting Information. The optimal transformation paths on the two-dimensional potential energy surface have been identified using the nudged elastic band method.[30] The energies and the gradients needed for these calculations have been computed numerically using the two-dimensional bicubic Lagrange interpolation scheme.[31] In the first transformation path (mechanism 1), the transition structure connecting $I_\alpha$ to X is situated at $\beta \approx 69°$ and $\gamma \approx 106°$ (TS1), while the transition $X \rightarrow I_\beta$ proceeds via a saddle point at $\beta \approx 77°$ and $\gamma \approx 108°$ (TS2). The alternative transformation route (mechanism 2) is identical with the mechanism revealed by our exploratory MD run and can be described as a direct $I_\alpha \rightarrow I_\beta$ transition with only one barrier near the point $\beta \approx 73°$ and $\gamma \approx 93°$ (TS3).

As we have already pointed out, the volume and the intersheet spacing in phase X are significantly increased compared to the phases $I_\alpha$ and $I_\beta$. A similar increase of the volume is observed in the transition region of mechanism 2, the maximal value of volume being 661 Å$^3$ (see Figure 9). Concomitantly, the intersheet spacing is increased to ~3.92 Å. As shown in Figure 10, according to the mechanism 2 the parameter $a$ decreases monotonically along the transformation from $I_\alpha$ to $I_\beta$. In mechanism 1, the parameter $a$ is a slightly more complicated function of the path length with a local maximum occurring in the region characteristic for the metastable phase X. The remaining three lattice parameters $b$, $c$, and $\alpha$ change only little in the course of both transformation paths.

The potential energy profiles for the two transformation paths are shown in Figure 11. The barriers for mechanism 1 are ~13 kJ/mol ($I_\alpha \rightarrow$ X) and ~2 kJ/mol ($X \rightarrow I_\beta$). The barrier computed for mechanism 2 is slightly higher (~15 kJ/mol); therefore mechanism 1 seems to be more favorable. As the difference in barriers is only small, in the order of 2 kJ/mol, we cannot exclude that both transformation paths can occur in reality.

## 5. CONCLUSIONS

Periodic dispersion−interaction corrected DFT simulations of the structure and of the mutual interconversion of two stable allomorphs of native cellulose I have been performed. In general, direct comparison of experimental and theoretical structural data is problematic because the experimentally measured cellulose microcrystals are typically small and contain only a few hundred chains.[4] In such microcrystals, a significant percentage of chains is located on the external surface, and their structure can be expected to deviate from the bulk structure. As shown in the theoretical study of Matthews et al.,[32] deformations of the crystal structure, such as twisting along the cellulose chain, can be expected for finite-sized crystals of cellulose. Despite these complications, in this work we have succeeded to obtain a very good agreement between the experimental and theoretical structures. Not only the computed lattice parameters are close to the experimentally measured values but our simulations were able to reproduce even subtle structural details in the intrachain geometry such as the nonequivalence of glucopyranosyl rings in cellulose $I_\alpha$. Two alternative and mutually exclusive interchain H-bond patterns have been examined. It has been shown, in agreement with earlier computational studies,[6,4] that pattern A is energetically more favorable than the alternative pattern B, the difference in energy being as large as ∼48 kJ/mol for both allomorphs of cellulose I. These results are consistent with the experimental observation that the chains linked via pattern A are more populated than those with the H-bond scheme B.

The transition mechanism for the $I_\alpha \rightarrow I_\beta$ transformation proposed in the experimental work of Nishiyama et al.[3] has been studied. In this mechanism, the transformation is realized via relative slippage of the cellulose chains along each other without any disruption of the hydrogen bonds. With a constrained relaxation technique, two alternative mechanisms have been identified. Mechanism 1 involves the formation of an intermediate metastable phase. Two barriers are present on the transition path with heights of ∼13 kJ/mol for $I_\alpha \rightarrow$ X, and ∼2 kJ/mol for X $\rightarrow I_\alpha$. The corresponding transition structures can be defined by points with $\beta = 69°$, $\gamma = 106°$, and $\beta = 77°$, $\gamma = 108°$, respectively. The structure of the intermediate phase X has been characterized too. It has been found that the intersheet spacing is increased by ∼4% as compared to the stable phase $I_\alpha$. The possible existence of such a high-temperature metastable structure has been reported in the experimental work of Wada et al.[11] The experimentally observed increase of the intersheet spacing in the intermediate structure is ∼6%, which is in good agreement with our simulation results.

The alternative transition route (mechanism 2) is a direct $I_\alpha \rightarrow I_\beta$ transformation with a single energy barrier occurring near the point $\beta = 73°$, $\gamma = 93°$. As in the mechanism 1, the cell volume and the intersheet spacing increase in the transition region. Mechanism 2 is less favorable from the energetical point of view (the computed barrier is 15 kJ/mol); however, the difference in barriers (∼2 kJ/mol) is comparable with the value of thermal energy $k_BT$ for the observed transition temperature (>260 °C). For that reason we expect that both mechanisms can occur in the nature.

In summary, we have reported a first ab initio study of the $I_\alpha \rightarrow I_\beta$ transformation in cellulose. Due to computational limitations, we considered only the simplest, nevertheless realistic, mechanism for the transition proposed in the literature.[3,11] We have demonstrated that this mechanism requires an increase of the intersheet spacing at a certain point of the transition. The anisotropic thermal expansion of cellulose I causing an increase of the intersheet spacing is a well-known experimental fact.[33] Hence it can be expected that the elevated temperature needed for the transition to occur leads to a decrease of the energy barriers for transformation. An alternative transition mechanism discussed in the literature[3,10,11] would require a rotation of whole cellulose chains around their axis. Obviously, in order to realize such a process, simultaneous rupture of all the H-bonds connecting the rotating chain with the neighboring chains would be needed. Although the elevated temperature at which transition occurs may partly weaken the interchain H-bonds, the observed thermal expansion of the crystal lattice is not large enough for the 180° rotation of molecular chains[11] and hence this mechanism is unlikely to occur. Finally, we cannot exclude that the temperature may play an even more radical role in the phase transformation than expected on the basis of the transition model analyzed in this study. For instance, Matthews et al.[34] have suggested in their recent force-field molecular dynamics study that the two-dimensional hydrogen bond network characteristic for the low-temperature crystal form of cellulose $I_\beta$ can transform into a three-dimensional hydrogen bond network at the transition temperature. Although this result might have been affected by the finite size of the structural model and the quality of the force field used in the simulation, it demonstrates that more complex routes for the $I_\alpha \rightarrow I_\beta$ can exist. The mechanism and energetics of possible more complex transition routes should be examined systematically in the future studies.

## ■ ASSOCIATED CONTENT

**Supporting Information.** Structural data (.cif) for the newly identified phase of cellulose (X) and animation (.mpg) showing the potential energy surface for cellulose I as a function of lattice parameters $\beta$ and $\gamma$. This material is available free of charge via the Internet at http://pubs.acs.org.

## ■ AUTHOR INFORMATION

### Corresponding Author

*E-mail: tomas.bucko@univie.ac.at; Janos.Angyan@crm2.uhp-nancy.fr.

## ■ ACKNOWLEDGMENT

The authors are grateful to Professor M. Nespolo (Nancy) for helpful discussions. This work has been supported by the VASP project, by the project GE 1676/1-1 of the priority program SPP 1315 (German Grant Agency), by the project VEGA − 1/0520/10 (Slovak Grant Agency), and by the ANR (Agence National de Recherche) via Contract Number ANR-07-BLAN-0272 (Wademecom). T.B. is grateful to University of Nancy for invited professorship.

## ■ REFERENCES

(1) Pérez, S.; Mackie, B. *Structure and morphology of cellulose.* http://glyco3d.cermav.cnrs.fr/glyco3d/lessons/cellulose/index.html, 2001.

(2) Nishiyama, Y.; Langan, P.; Chanzy, H. *J. Am. Chem. Soc.* **2002**, *124*, 9074–9082.

(3) Nishiyama, Y.; Sugiyama, J.; Chanzy, H.; Langan, P. *J. Am. Chem. Soc.* **2003**, *125*, 14300–14306.

(4) Nishiyama, Y.; Johnson, G. P.; French, A. D.; Forsyth, V. T.; Langan, P. *Biomacromolecules* **2008**, *9*, 3133–3140.
(5) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.
(6) Mazeau, K. *Cellulose* **2005**, *12*, 339.
(7) Horii, F.; Yamamoto, H.; Kitamaru, R.; Tanahashi, M.; Higuchi, T. *Macromolecules* **1987**, *20*, 2946–2949.
(8) Yamamoto, H.; Horii, F. *Macromolecules* **1993**, *26*, 1313–1317.
(9) Debzi, E.; Chanzy, H.; Sugiyama, J.; Tekely, P.; Excoffier, G. *Macromolecules* **1991**, *24*, 6816–6822.
(10) Hardy, B.; Sarko, A. *Polymer* **1996**, *37*, 1833–1839.
(11) Wada, M.; Kondo, T.; Okano, T. *Polym. J.* **2003**, *35*, 155–159.
(12) Kresse, G.; Hafner, J. *Phys. Rev. B* **1993**, *48*, 13115–13118.
(13) Kresse, G.; Hafner, J. *Phys. Rev. B* **1994**, *49*, 14251–14269.
(14) Kresse, G.; Furthmüller, J.; Hafner, J. *Phys. Rev. B* **1994**, *50*, 13181–13185.
(15) Kresse, G.; Furthmüller, J. *Comput. Mater. Sci.* **1996**, *6*, 15–50.
(16) Blöchl, P. E. *Phys. Rev. B* **1994**, *50*, 17953–17979.
(17) Kresse, G.; Joubert, D. *Phys. Rev. B* **1999**, *59*, 1758–1775.
(18) Grimme, S. *J. Comput. Chem.* **2006**, *27*, 1787–1799.
(19) Bucko, T.; Hafner, J.; Lebègue, S.; Ángyán, J. G. *J. Phys. Chem. A* **2010**, *114*, 11814–11824.
(20) Bucko, T.; Hafner, J.; Ángyán, J. G. *J. Chem. Phys.* **2005**, *122*, 124508.
(21) Baker, J.; Kessi, A.; Delley, B. *J. Chem. Phys.* **1996**, *105*, 192–212.
(22) Li, Y.; Lin, M.; Davenport, J. W. *J. Phys. Chem. C* **2011**, *115*, 11533–11539.
(23) Kono, H.; Yunoki, S.; Shikano, T.; Fujiwara, M.; Erata, T.; Takai, M. *J. Am. Chem. Soc.* **2002**, *124*, 7506–7511.
(24) Kono, H.; Erata, T.; Takai, M. *Macromolecules* **2003**, *36*, 3589–3592.
(25) Giannozzi, P.; Baroni, S.; Bonini, N.; Calandra, M.; Car, R.; Cavazzoni, C.; Ceresoli, D.; Chiarotti, G. L.; Cococcioni, M.; Dabo, I.; Dal Corso, A.; de Gironcoli, S.; Fabris, S.; Fratesi, G.; Gebauer, R.; et al. *J. Phys.: Condens. Matter* **2009**, 21.
(26) Neyertz, S.; Pizzi, A.; Merlin, A.; Maigret, B.; Brown, D.; Deglise, X. *J. Appl. Polym. Sci.* **2000**, *78*, 1939–1946.
(27) Parrinello, M.; Rahman, A. *Phys. Rev. Lett.* **1980**, *45*, 1196–1199.
(28) Woo, T.; Margl, P.; Blöchl, P.; Ziegler, T. *J. Phys. Chem. B* **1997**, *101*, 7877–7880.
(29) Hu, H.; Yun, R.; Hermans, J. *Mol. Simul.* **2002**, *28*, 67–80.
(30) Jonsson, H.; Mills, G.; Jacobsen, K. W. Nudged Elastic Band Method for Finding Minimum Energy Paths of Transitions. In *Classical and Quantum Dynamics in Condensed Phase Simulations*; Berne, B. J., Ciccotti, G., Coker, D. F., Eds.; World Scientific: Singapore, 1998; p 385.
(31) Press, W. H.; Flannery, B. P.; Teukolsky, S. A.; Vetterling, W. T. *Numerical Recipes in FORTRAN 77: The Art of Scientific Computing*, 2nd ed.; Cambridge University Press: Cambridge and New York, 1992; Vol. 1, p 99.
(32) Matthews, J.; Skopec, C.; Mason, P.; Zuccato, P.; Torget, R.; Sugiyama, J.; Himmel, M.; Brady, J. *Carbohydr. Res.* **2006**, *341*, 138.
(33) Wada, M. *J. Polym. Sci., Part B: Polym. Phys.* **2002**, *40*, 1095–1102.
(34) Matthews, J. F.; Bergenstrahle, M.; Beckham, G. T.; Himmel, M. E.; Nimlos, M. R.; Brady, J. W.; Crowley, M. F. *J. Phys. Chem. B* **2011**, *115*, 2155–2166.
(35) Wada, M.; Nishiyama, Y.; Chanzy, H.; Forsyth, T.; Langan, P. *Powder Diffr.* **2008**, *23*, 92–95.