Polymer Vol. 37 No. 10, pp. 1833–1839, 1996

0032-3861/96/$15.00 + 0.00

Molecular dynamics simulations and diffraction-based analysis of the native cellulose fibre: structural modelling of the I-α and I-β phases and their interconversion

Barry J. Hardy*
Physical and Theoretical Chemistry Laboratory, Oxford University, South Parks Road, Oxford OX1 3QZ, UK

and Anatole Sarko
Department of Chemistry, College of Environmental Science and Forestry, SUNY, Syracuse, NY 13210, USA
(Revised 17 August 1995)

In this report we describe the building of diffraction-based models of the two phases of cellulose I, which subsequently are subjected to molecular dynamics simulation. The models showed an interesting variety of behaviour, including glycosidic and exocyclic torsional motion and isomerization, hydrogen-bond breakage and formation, individual and collective chain motion, and sheet deformation in the non-hydrogen-bonding direction. The I-α phase exhibited a greater dynamic range of behaviour than the I-β phase, including considerable movement of glycosidic torsions away from initial diffraction-based positions and considerable relative motion of the chains. Based on motions observed in the simulations, we suggest a *break–slip* model for the I-$\alpha \rightarrow$I-β phase transition, which proposes that the transition is initiated by heating-induced hydroxymethyl and hydroxyl side-group torsional rotations accompanied by hydrogen-bond breakage. Chains of the I-α phase are hence freed for rotation and sliding into the more stable I-β morphology. This model was tested with molecular mechanics refinement of likely intermediate structures. The results suggested that a facile transformation path is available via such a mechanism. Copyright © 1996 Published by Elsevier Science Ltd.

(Keywords: cellulose; molecular dynamics; X-ray diffraction)

INTRODUCTION

Based on ^{13}C nuclear magnetic resonance (n.m.r.) measurements, VanderHart and Atalla[1] proposed the existence of two different phases in native cellulose (cellulose I). More recently, the electron diffraction experiments of Sugiyama, Vuong and Chanzy demonstrated[2] the coexistence of the two phases (I-α and I-β) as regions of differing crystalline structure that occur contiguously along the microfibril. The I-α phase is predominant in bacterial celluloses whereas the I-β structure is the major component in higher plants such as ramie and cotton. The I-α phase is metastable and converts to I-β upon annealing, by heating with alkali to 260°C, or heating to 280°C in an inert atmosphere of He[3,4]. The latter condition indicates that the conversion proceeds in the solid state through a physical phase transformation. Although static and dynamic modelling studies[5–12] have recently been reported on cellulose, solid-state molecular simulations of cellulose have apparently not been attempted, primarily due to

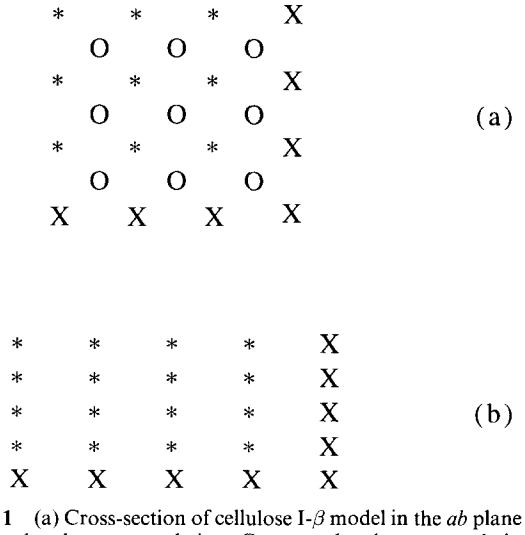
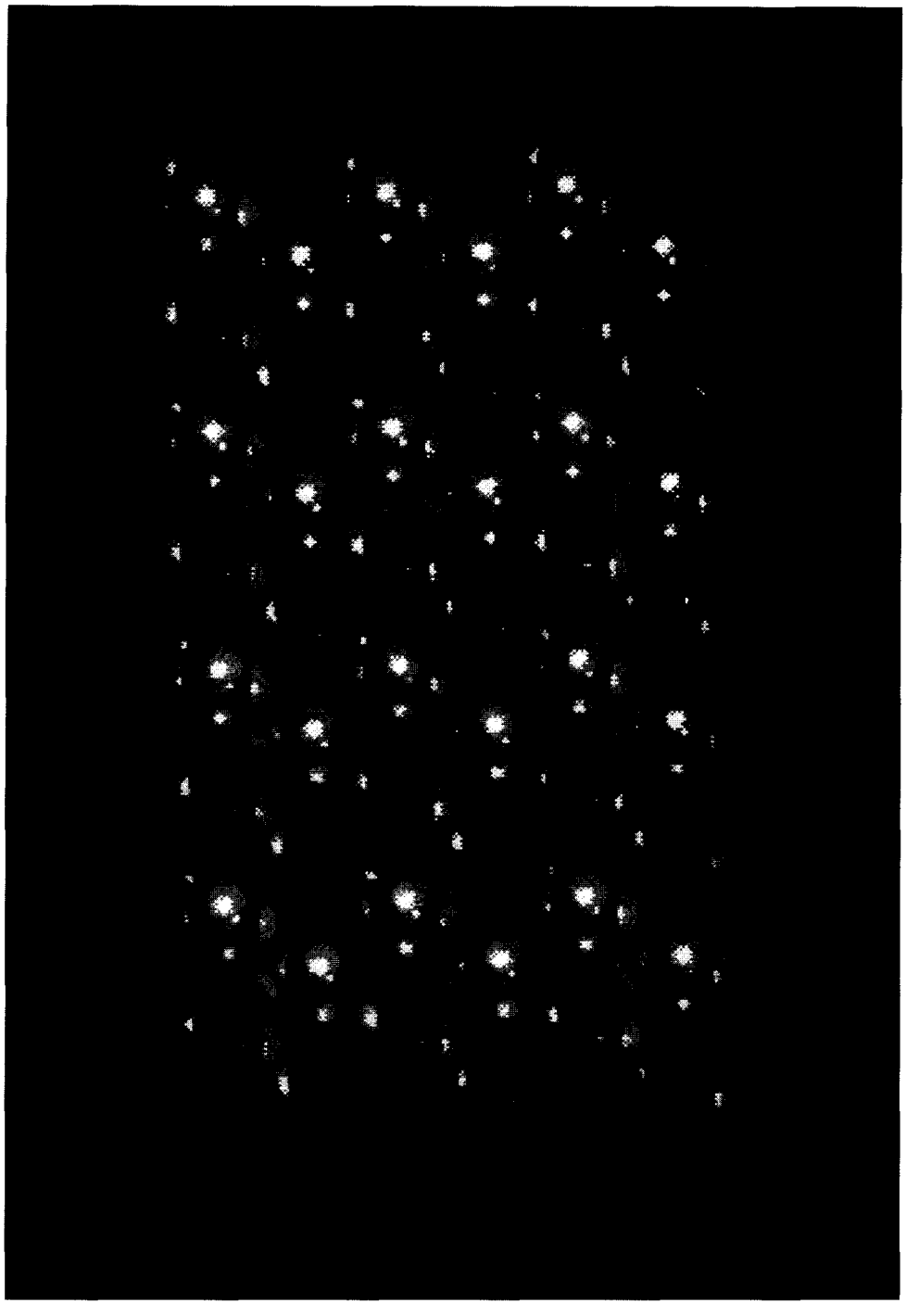


Figure 1 (a) Cross-section of cellulose I-β model in the *ab* plane where $*$ = translated corner chain, ○ = translated centre chain, and $\times$ = imaged chain of $*$ on opposite side of section. (b) Cross-section of cellulose I-α model in the *ab* plane where $*$ = translated corner chain and $\times$ = imaged chain of $*$ on opposite side of section

* To whom correspondence should be addressed



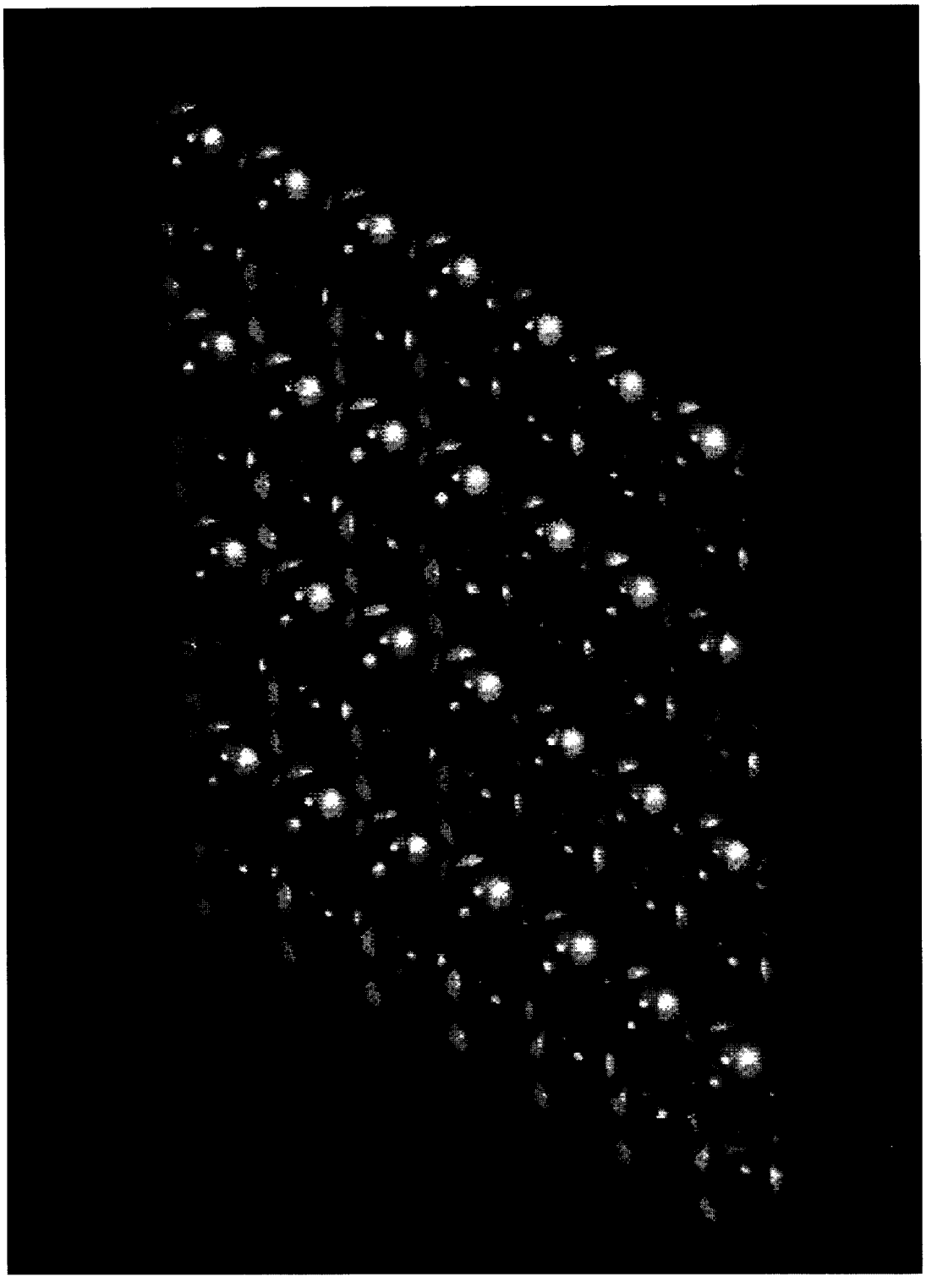
(a)

Figure 2 (a) Cellulose I-β model viewed perpendicular to the fibre axis

computational limitations. We therefore initiated simulation studies, using high-performance parallel computing machinery, to construct detailed atomistic models of the different polymorphic phases of cellulose, to examine their equilibrium structure and dynamics and to investigate the transformation pathways between the phases. In this report (preliminary results have already been presented elsewhere[13]) we describe the building of diffraction-based models of the two phases of cellulose I, which subsequently are subjected to molecular dynamics (MD) simulation and molecular mechanics (MM) modelling.

METHODS

The I-β phase initial conditions were based on the Woodcock–Sarko X-ray fibre diffraction structure of ramie cellulose[14]. The structure consists of a two-chain monoclinic unit cell with unit-cell parameters: $a = 7.79$ Å, $b = 8.25$ Å, $c = 10.34$ Å, $\alpha = 90°$, $\beta = 90°$ and $\gamma = 96.8°$. Chains are located at the corner and the centre of the cell so as to form parallel sheets. The sheet passing through the centre of the unit cell is staggered by $-c/4$ relative to the corner chain sheet. The sheets contain interchain hydrogen bonds in the plane of the sheets but not between them. A super unit cell of 18

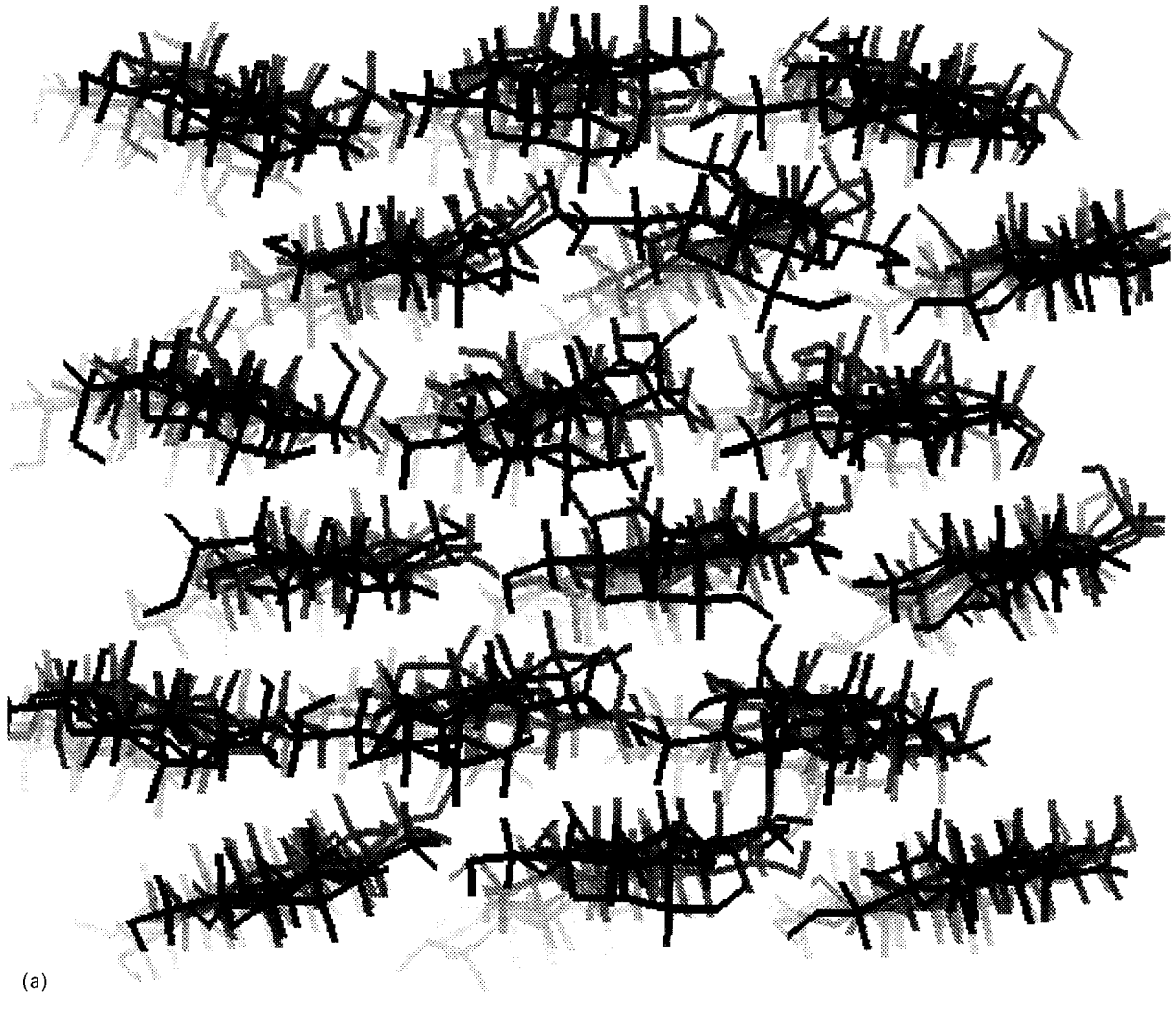


(b)

Figure 2 (b) Cellulose I-α model viewed perpendicular to the fibre axis

chains was constructed through the translation of these chains so as to form a lattice that has the cross-section in the *ab* plane shown in *Figure 1a*. The super unit cell of 18 chains was subjected to regular periodic boundary conditions in the *ab* plane so as to reproduce image chains (X) as far out as required for calculation of the non-bonded pair interactions. Along the *c* axis the two-ring cellobiose repeat unit was translated to form chains of eight glucose rings. The eight-ring chains were subjected to periodic boundary conditions along the *c* axis direction with an accompanying joining of the chain ends across the boundary, i.e. the eighth ring at the top of the chain in the central super unit cell was glycosidically bonded to an image of the first ring of the chain located in the adjacent image unit cell above. The final model can be viewed perpendicular to the fibre axis in *Figure 2a*.

The I-α phase initial conditions were based on the



(a)

Figure 3 View along the c axis of (a) cellulose I-β after 100 ps of dynamics and (b) cellulose I-α after 50 ps of dynamics

electron diffraction structure of *Microdictyon tenuius*[2]. The structure consists of a one-chain triclinic unit cell with unit-cell parameters; $a = 6.74$ Å, $b = 5.93$ Å, $c = 10.36$ Å, $\alpha = 117°$, $\beta = 113°$ and $\gamma = 81°$. Chains are located at the corner of the cell so as to form a series of parallel sheets, which are staggered in increasing fashion by $-c/4$, $-c/2$, $-3c/4$, etc., in the direction of the (1 1 0) plane. A super unit cell of 16 chains was constructed through the translation of the corner chain to form a lattice, which has the cross-section arrangement in the *ab* plane shown in *Figure 1b* and the structure viewed perpendicular to the fibre axis shown in *Figure 2b*.

The above I-α and I-β models were used as initial conditions for MD simulations carried out under *NVE* (i.e. constant number of atoms, volume and energy) conditions for the super unit cell. The component unit cells themselves were not explicitly constrained, and adjacent chains in the central super cell fluctuated independently. The following procedures were followed using the CHARMM[15] molecular modelling program and the PARM 20 parameter set:

(1) The initial diffraction-based coordinates were subjected to 250 steps of a steepest-descent minimization, followed by 250 steps of conjugate-gradients minimization, after which the r.m.s. gradient was 0.12 kcal mol^{-1} Å^{-1} for the I-β phase and 0.24 kcal mol^{-1} Å^{-1} for the I-α phase.

Table 1 Energy summary for the I-α →I-β transition

Model	Δz^a (Å)	Relative energy (kcal/unit cell)
I-α	−5.3	25.2
O6 → *tg*	−5.3	3.5
$\alpha \rightarrow 90°$	−5.3	1.9
translate	−4.3	3.0
translate	−3.3	2.9
translate	−2.3	2.4
translate	−1.3	1.7
I-β	0.0	0.0

[a] Δz refers to chain stagger between successive corner chains in a two-chain equivalent I-α unit cell. In this cell the centre chain sheet stagger is about −2.7 Å relative to the corner sheet that passes through the unit-cell origin

(2) MD simulations were started from minimized coordinates. Initial velocities were chosen from a Maxwell–Boltzmann distribution at −80°C. The system was heated incrementally (10°C per 250 fs) until the desired simulation temperature (20°C) was reached, at which point the system was allowed to equilibrate for 20 ps, during which velocity rescaling was employed when the

Figure 3 continued

system deviated by more than 3°C from the target temperature. Production trajectories of 100 ps were then carried out. The Verlet algorithm[16] was used to integrate Newton's equations of motion for the system at constant energy using a time step of 1 fs. A shifted potential with a 7.5 Å cutoff was used for non-bonded interactions.

The MM refinement of cellulose I-α and I-β phases and the intermediate structures likely to be encountered in their interconversion were conducted using the program PS79[17]. The energies were calculated using a Lennard-Jones 6–12 function set with parameters developed at SUNY-ESF for the analysis of polysaccharide crystal structures. No partial charges were used, but a $1/r^3$ hydrogen-bond term, adjusted to give a $-4\,\text{kcal mol}^{-1}$ energy well at an O–O distance of 2.8 Å, was incorporated.

For modelling purposes the cellulose I-α unit cell was converted into its two-chain equivalent cell similar in *ab* projection to cellulose I-β. The parameters of this cell were: $a = 9.64$ Å, $b = 8.25$ Å, $c = 10.36$ Å, $\alpha = 90.4°$, $\beta = 123.5°$ and $\gamma = 97.4°$. The conversion of this unit cell to that of I-β was modelled using variable parameters in refinements of the intermediate structures, which included only the chain positions, the ϕ and ψ torsion angles, and the rotations of the hydroxymethyl groups. Even though the lattice was allowed to vary, the distance between the chains within the sheet (i.e. the *b* axis of the unit cells) was held constant at 8.3 Å. All refinements were conducted at 270°C.

RESULTS AND DISCUSSION

The models of cellulose I, studied here by MD simulation showed an interesting variety of behaviour, including

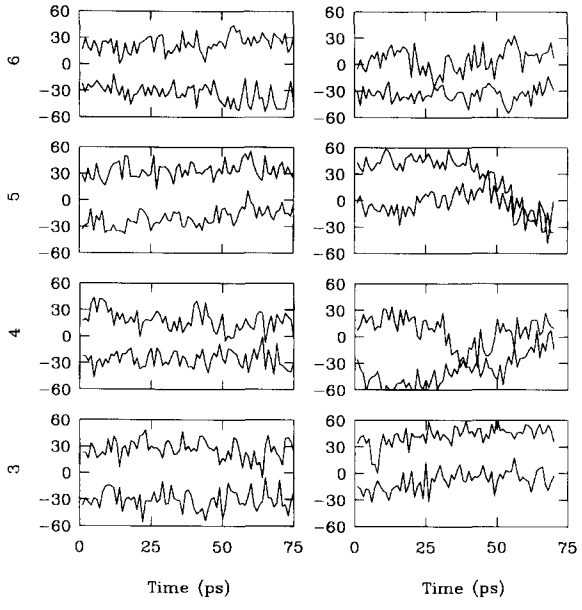


Figure 4 Glycosidic torsion-angle time series for four consecutive linkages (3–6) along a cellulose I-β chain (left) and cellulose I-α chain (right). The ϕ trajectories are predominantly positive and above the more negative ψ trajectories (ϕ = torsion angle H1–C1–O4–C4, ψ = torsion angle C1–O4–C4–H4)

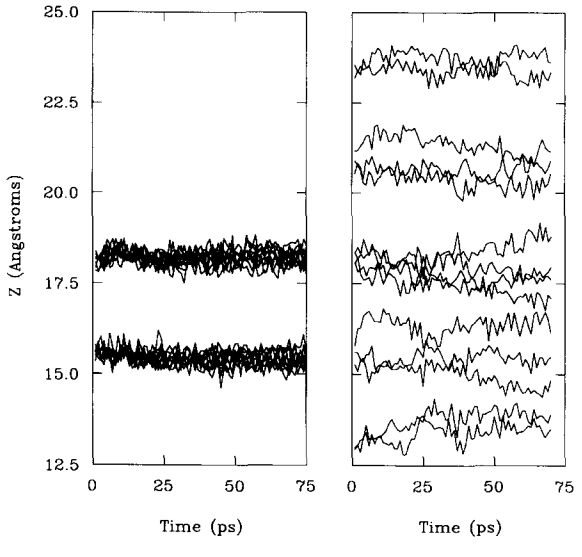


Figure 5 Time series for the chain centre displacement (Z) along the c axis for the 18 central super unit cell chains of cellulose I-β (left) and the equivalent 16 chains of cellulose I-α (right). The chain centre was defined as the glycosidic oxygen between rings 4 and 5

glycosidic and exocyclic torsional motion and isomerization, hydrogen-bond breakage and formation, individual and collective chain motion, and sheet deformation in the non-hydrogen-bonding direction. The consequences of the motions can be seen in the views of both phases during the simulations shown in *Figure 3*. Limited disorder along the chains is clearly seen, arising principally from glycosidic torsion fluctuations. Chains in both phases also exhibit rocking rotations about the c axis, which turn the chains away from ideal lattice positions where all chains are lined up in parallel sheets.

Example glycosidic torsion trajectories are shown in *Figure 4*. The torsional positions for the I-β phase remain

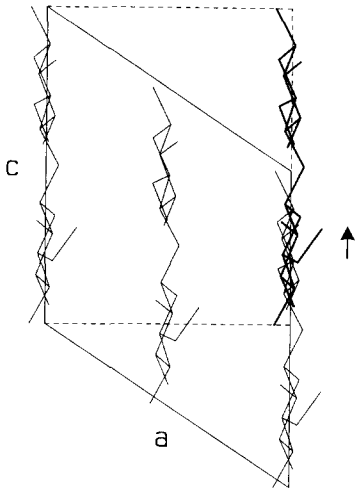


Figure 6 Model of the transformation of cellulose I-α to I-β, viewed in the b direction of both (two-chain) unit cells. The initial positions of the chains are shown as a thin line, corresponding to the two-chain unit cell of I-α. The sheet on the right is gradually pulled upwards until the I-β structure has been reached (bold line). The projection of the I-β unit cell is shown as a dashed line. (Hydrogen atoms are not shown)

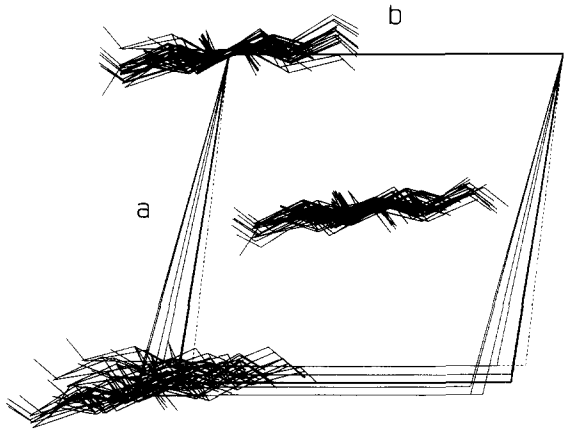


Figure 7 Superposition of the structures shown in *Table 1*, viewed along the c axis. The two-chain unit cell of cellulose I-α is drawn as a bold line and the unit cell of cellulose I-β is drawn as a dashed line. The origins of all superimposed unit cells are coinciding at the top left corner. (Hydrogen atoms are not shown)

in relatively constant average positions about which fluctuations of order $\pm 10°$ occur. In contrast, several of the I-α torsions show considerable drift away from initial conditions, e.g. from ca. $+30°$ to $-30°$. In both phases hydroxymethyl group transitions are observed in which hydrogen bonds are broken and re-formed. The positions of chain centres along the direction of the c axis are shown for both phases in *Figure 5*. The positions of the I-β chains remain in a stable up-and-down pattern (nine chains are displaced by $-c/4$ from nine others). In contrast the I-α chains show considerable motion and slip relative to each other along the c axis.

The question immediately arises regarding the accuracy of the parametrization of these models. We, in fact, suspect that the strength of hydrogen-bonding and hydrophobic interactions may be underestimated in this parameter set. The effective temperature of the simulations may thus be considerably higher than room

temperature. Therefore, we may be observing events more representative of those induced by heating as the much higher phase transition temperature of 260–280°C is approached. Additional parametrization studies are required to examine this issue further. Another difficulty we observed with the simulations is that both structures continued to heat during the production phases, suggesting that a considerably longer equilibration period may be required than used here.

Transformation of I-α to I-β

As shown above, the I-α phase exhibited a greater dynamic range of behaviour than the I-β phase, including considerable movement of glycosidic torsions away from initial diffraction-based positions and considerable relative motion of the chains. Based on motions observed in the simulations, we suggest a *break–slip* model for the phase transition, which proposes that the transition is initiated by heating-induced hydroxymethyl and hydroxyl side-group torsional rotations accompanied by hydrogen-bond breakage. Chains of the I-α phase are hence freed for rotation and sliding into the more stable I-β morphology.

In order to probe the merit of the break–slip model further, we undertook the MM calculation diagrammed in *Figure 6*, in which we gradually pulled one $-c/4$ staggered sheet up relative to its neighbouring sheet to a final position of $+c/4$ stagger. In this sequence, the hydroxymethyl groups of the chains in the two-chain equivalent of the cellulose I-α diffraction-based unit cell were first rotated into *tg* positions and the unit-cell angle α was forced from 90.4° to 90° to place all chains within a given sheet at the same height. The sheet was then translated in increments of 1 Å until it reached the I-β position, corresponding to a total displacement of a half-unit cell-length of 5.3 Å. At each step along this pathway the structures were refined by energy minimization; the results are displayed in *Table 1*. The initial relaxation of the I-α structure results in a large drop in energy as the MM functions locate a local minimum. Translation of the structure then shows the crossing of a relatively modest barrier after a translation of 1–2 Å, following which the energy profile is all downhill to the I-β structure. As shown in *Figure 7* the chains undergo considerable rotations and translations during this transformation, and there is a slight loosening of the structure. The final I-β structure then packs more tightly again.

New MD simulations recently reported by Heiner *et al.*[18] show several interesting detailed features and results that are consistent with the results observed herein. In this study the monoclinic I-β phase was found to have an energy 2.1 kcal mol^{-1} lower than the I-α phase. The difference in energy was primarily due to electrostatic interactions, which contrasts with the work of French *et al.*[10], who primarily attributed the difference to van der Waals energy. During MD the chains in the 'odd' (2 0 0) sheets of the I-β structure were found to rotate to form an angle of 11° with the (2 0 0) 'even' sheets; this can be compared to the smaller angle of 3° in the diffraction-based structure[14]. The structure of the I-β 'odd' sheets was found to resemble most closely that of the I-α structure. These observations tend to support the break–slip model, as the I-α chains can rotate and translate during the transition event. Uncertainty remains, however, in the accuracy of the details of all these simulations, and comparisons with experiment remain challenging[18]. Considerable work remains to improve the parameters used in the simulations. As a starting point in this regard, small-molecule[19] *ab initio*, n.m.r. and crystallography studies have been initiated to yield improved parametrization for new cellulose simulations. This work will enable further validation of simulation against experiment and enhance the reliability of further, more extensive simulations of the cellulose fibre and its environment.

REFERENCES

1 VanderHart, D. L. and Atalla, R. H. *Macromolecules* 1984, **17,** 1465
2 Sugiyama, J., Vuong, R. and Chanzy, H. *Macromolecules* 1991, **24,** 4168
3 Sugiyama, J., Persson, J. and Chanzy, H. *Macromolecules* 1991, **24,** 2461
4 Debzi, E. M., Chanzy, H., Sugiyama, J., Tekely, P. and Excoffier, G. *Macromolecules* 1991, **24,** 6816
5 Hardy, B. J., PhD Dissertation, Syracuse University, 1990
6 Hardy, B. J. and Sarko, A. *J. Comput. Chem.* 1993, **14,** 831
7 Hardy, B. J. and Sarko, A. *J. Comput. Chem.* 1993, **14,** 848
8 Simon, I., Scheraga, H. A. and Manley, R. St John *Macromolecules* 1988, **21,** 983
9 Aabloo, A. and French, A. D. *Macromol. Theory. Simul.* 1994, **3,** 185
10 French, A. D., Miller, D. P. and Aabloo, A. *Int. J. Biol. Macromol.* 1993, **15,** 30
11 French, A. D., Tran, V. H. and Perez, S. *ACS Symp. Ser.* 1990, **430,** 191
12 Henrissat, B., Perez, S., Tvaroska, I. and Winter, W. T. *ACS Symp. Ser.* 1987, **340,** 38
13 Cellulose Division, American Chemical Society, Fall '92 Washington DC Meeting, Abstracts, 1992; Polymer Division, American Chemical Society, Anaheim Meeting, 1995; *ACS Polym. Prepr.* 1995, **36,** 640
14 Woodcock, C. and Sarko, A. *Macromolecules* 1980, **13,** 1183
15 Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S. and Karplus, M. *J. Comput. Chem.* 1993, **4,** 187
16 Verlet, L. *Phys. Rev.* 1967, **159,** 98
17 Zugenmaier, P. and Sarko, A. *Biopolymers* 1976, **15,** 2121
18 Heiner, A. P., Sugiyama J. and Teleman, O. *Carbohydr. Res.* in press
19 Gutierrez, A., Hardy, B. J., Lesiak, K., Seidl, E. and Widmalm, G. submitted for publication