Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

In silico studies of crystalline cellulose and its degradation by enzymes

Giovanni Bellesia,[a] Andrea Asztalos,[b] Tongye Shen,[c] Paul Langan,[d] Antonio Redondo[e] and S. Gnanakaran[f]*

[a]T-6 and CNLS, Theoretical Division, Los Alamos National Laboratory, MS B258, Los Alamos, NM 87545, USA, [b]Department of Physics, University of Notre Dame, Notre Dame, IN 46556, USA, [c]Center for Molecular Biophysics and Department of Biochemistry, Cellular and Molecular Biology, ORNL and University of Tennessee, Knoxville, TN 37996, USA, [d]Biosciences Division, Los Alamos National Laboratory, Los Alamos, NM 87545, USA, [e]Theoretical Division, Los Alamos National Laboratory, MS B210, Los Alamos, NM 87545, USA, and [f]T-6, Theoretical Division, Los Alamos National Laboratory, MS K710, Los Alamos, NM 87545, USA

Correspondence e-mail: gnana@lanl.gov

Received 5 May 2010
Accepted 23 July 2010

In this report, the current state of computational studies on crystalline cellulose is reviewed. The discussion is focused on fully atomistic molecular-dynamics simulations as well as on other computational approaches which are relevant in the context of enzymatic degradation of cellulose. Finally, possible directions and necessary improvements for future computational studies in this challenging research field are summarized.

1. Introduction

Cellulose, an assembly of polymers of glucose, is an important renewable energy resource from plant biomass. The efficient degradation of crystalline fibers of cellulose to glucose is a critical roadblock to lignocellulosic biofuel; it is caused by the unusually high thermal and mechanical stability of cellulose. The redundancy in hydrogen-bonding pattern and the intertwining of intramolecular and intermolecular hydrogen bonds ensure this high stability. Cellulose occurs in the woody cell wall as microfibrils of two distinct crystal phases, namely I_α and I_β (Atalla & VanderHart, 1984). Regeneration and mercerization yield cellulose II. Pretreatment of cellulose I and cellulose II with amines yields cellulose III_I and cellulose III_{II}, respectively (Clark & Parker, 1937). X-ray and neutron crystallography have provided molecular-level details of these different forms of crystalline cellulose (Langan *et al.*, 2001; Nishiyama *et al.*, 2002, 2003; Wada *et al.*, 2004, 2006). The availability of these cellulose structures has led to extensive theoretical work probing the structural and chemical properties of cellulose and its degradation.

The purpose of this short review is to assess the current state of *in silico* studies on cellulose crystals and to propose viable directions for future studies in this field. As of today, the majority of the computational studies on cellulose have been devoted to analysis of its different crystalline forms. Therefore, the focus of our discussion will be on fully atomistic molecular-dynamics (MD) simulations of various crystalline cellulose allomorphs, although MD simulations of soluble cellulose oligomers and coarse-grained computational models will also be considered. Our discussion will be limited to classical statistical mechanical and rule-based methods; hence, both quantum-mechanical and molecular-mechanics calculations (electronic structure calculations and potential energy optimizations) will not be considered. We also would like to point out that the choice of the referenced computational studies in this short review has been mostly motivated by the

Table 1
Basic methodological details of a set of representative molecular-dynamics simulation studies on crystalline cellulose.

The simulation lengths are given in ns and refer to the production runs.

Reference	Force field	Length	Solvent	'Infinite chain'	Allomorph
Heiner *et al.* (1995)	GROMOS87	1.0	No	Yes	I_α, I_β
Reiling & Brickman (1995)	???	0.75	Yes	Yes	I
Kroon-Batemburg *et al.* (1996)	GROMOS87	0.5	No	Yes	I_β, II
Hardy & Sarko (1996)	CHARMM20	0.1	No	Yes	I_α, I_β
Heiner & Teleman (1997)	GROMOS87	0.5	Yes	No	I_β
Heiner *et al.* (1998)	GROMOS87	1.0	Yes	No	I_β
Neyertz *et al.* (2000)	New†	0.35	No	Yes	I_α, I_β
Bierman *et al.* (2001)	GROMOS96	1.3	Yes	No	I_β
Ito *et al.* (2002)	GROMOS87	0.2	Yes	Yes	I_β
Mazeau & Heux (2003)	PCFF	0.4	No	Yes	I_α, I_β
Mazeau (2005)	PCFF	1.0	No	Yes	I_β
Yui *et al.* (2006)	GLYCAM04	2.0	Yes	No	I_β
Matthews *et al.* (2006)	Kuttel *et al.* (2002)	1.0	Yes	No	I_β
Bergenstrahle *et al.* (2007)	GROMOS96 45A4	10.0	No	Yes	I_β
Yui & Hayashi (2007)	GLYCAM04	1.0	No	No	I_α, III_I
Bergenstrahle *et al.* (2008)	GROMOS96 45A4	5.0	Yes	Yes	I_β
Mazeau & Rivet (2008)	PCFF	15.0	Yes‡	Yes	I_β
Bergenstrahle *et al.* (2009)	GROMOS96 45A4	5.0	Yes	Yes	I_β
Nishiyama *et al.* (2009)	GLYCAM04	10.0	No	No	I_β
Yui & Hayashi (2009)	GLYCAM04	2.5	Yes	No	III_I

† The authors report the details of a new all-atom force field for crystalline cellulose I_β. ‡ The solvent consisted of a small water droplet on the surface of the cellulose crystal.

ongoing research projects in our laboratory and is by no means complete or exhaustive.

2. All-atomistic MD simulation approaches

In Table 1, we show the main methodological details of a representative set of MD simulation studies on crystalline cellulose that have been published during the last 15 years. It is worth noting that the simulation time length has increased by two orders of magnitude, varying in the interval 0.1–15 ns. This is a remarkably slow growth, especially when compared with MD studies of other biomolecules such as proteins and nucleic acids.

Efforts are currently under way in different research groups to finalize MD simulations of cellulose crystalline systems on the hundreds-of-nanoseconds time scale (J. Matthews, personal communication; ongoing work in our laboratory). The extended conformational phase-space sampling obtained from these studies will firstly improve our knowledge of the relative stability of the different crystalline allomorphs and its dependency on the crystal thermal modes and hydration properties. Secondly, it will allow a more reliable validation of the empirical force fields used in cellulose simulations (GROMOS, CHARMM, GLYCAM *etc.*).

3. Comparison with X-ray and neutron diffraction data

MD simulations of the I_β, I_α and III_I allomorphs have mostly been focused on comparison with the experimental results obtained from X-ray and neutron diffraction studies (Kroon-Batemburg *et al.*, 1996; Mazeau & Heux, 2003; Mazeau, 2005; Yui *et al.*, 2006; Matthews *et al.*, 2006; Bergenstrahle *et al.*, 2007; Yui & Hayashi, 2007, 2009; Nishiyama *et al.*, 2009). Different MD simulation studies have reported different degrees of success in reproducing the shape and size of the experimental crystal unit cell as well as the rotational state population of the critical dihedral degrees of freedom. Bergenstrahle and coworkers, in their paper on thermal response in crystalline I_β cellulose (Bergenstrahle *et al.*, 2007), present an extended overview of both experimental and modeled unit cells. The majority of the studies (both experimental and computational) display similar values for both the angle γ (96–98°) and the unit-cell dimensions a, b and c. It is worth noticing that (i) the ratio a/b which, together with γ, defines the overall shape of the unit cell varies between 0.95 and 0.99 and (ii) the quantity $a\sin(\gamma)/b$, which defines the ratio between the intersheet distance and the interchain/intrasheet distance (the distance between cellulose chains within the same crystal sheet or layer), is always <1 (it varies between 0.94 and 0.98; the intersheet distance is smaller than the interchain/intrasheet distance).

Interestingly, MD simulation studies that obtain a value for γ that differs from that found in the majority of the studies ($\gamma \simeq 90°$) also give both a/b and $a\sin(\gamma)/b$ values that are larger than 1 (a larger intersheet distance than interchain/intrasheet distance; Matthews *et al.*, 2006; Bergenstrahle *et al.*, 2007). MD simulation data also show overall good agreement with experiments in the analysis of the rotational state population for the hydroxymethyl group (essential for defining the intrasheet and intersheet hydrogen-bond networks), *i.e.* a dominance of the *tg* conformation for cellulose I_β and of the *gt* conformation for cellulose III_I.

4. Thermodynamic analysis and structural observables

Aside from structural analysis of the crystal unit-cell dimensions and the internal degrees of freedom, MD simulation studies have considered other relevant structural/thermodynamic features of crystalline cellulose. MD data analyses and calculations have shed new light on (i) the relative stability of the different hydrogen bonds (intersheet, intrasheet/interchain, intrasheet/intrachain) within the crystal and their thermal behavior (Heiner *et al.*, 1995; Neyertz *et al.*, 2000; Ito *et al.*, 2002; Mazeau & Heux, 2003; Mazeau, 2005; Yui & Hayashi, 2007, 2009; Nishiyama *et al.*, 2009), (ii) the hydration properties of crystal cellulose surfaces (Heiner *et al.*, 1998; Heiner & Teleman, 1997; Biermann *et al.*, 2001), (iii) the crystal thermal response and its thermal expansion coefficients (Bergenstrahle *et al.*, 2007), (iv) the crystal bulk mechanical properties (Mazeau & Heux, 2003; Bergenstrahle *et al.*, 2007)

and (v) some dynamical aspects of the cellulose–water interface (NMR spin-lattice relaxation times and the calculation of force pulling of a single cellulose chain at the crystalline–liquid interface; Bergenstrahle *et al.*, 2008, 2009).

In particular, the studies on crystalline cellulose hydration (Heiner & Teleman, 1997) show that only the surface layer of crystalline cellulose is affected by the surrounding water solvent and that O atoms O2, O3 and O6 are extensively involved in a stable and ubiquitous cellulose–water hydrogen-bond network. A simulation study by Bergenstrahle *et al.* (2009) confirms that the cellulose–water hydrogen-bond network contributes to lowering the number of hydrogen bonds between neighboring cellulose chains on the surface and therefore facilitates the desorption process on the crystalline cellulose surface.

5. The supramolecular twist

Simulations of crystalline cellulose have been performed using both 'infinite' crystals (in which the cellulose chains are covalently bonded with their nearest images along their main backbone axis) and finite-length microcrystals (see Table 1). Interestingly, solvated cellulose microcrystals typically show a persistent right-handed supramolecular twist originating from breaking of the twofold screw-axis (2_1) symmetry of the cellulose chains (Matthews *et al.*, 2006; Yui *et al.*, 2006; Yui & Hayashi, 2007). Conversely, the twofold screw-axis (2_1) symmetry is conserved when 'infinite' crystals are considered. It is possible that the strain imposed by the covalent bond between neighboring images may prevent the twisting of the helix in the 'infinite' crystal. Preliminary results from MD simulations that considered an identical primary system with both finite and 'infinite' chain lengths with the GLYCAM force field show twisting only in the finite case.

The supramolecular twisting could be an inherent property only of cellulose fibers with a large aspect ratio and not of the bacterial cellulose from which crystal structures have been obtained. However, simulations of diffraction patterns from cellulose fibers of varying sizes and twists indicate that a small amount of twisting is not inconsistent with the crystallographic data, even from large highly crystalline bacterial fibers (A. French & Y. Nishiyama, private communication). Alternatively, it is also possible that nonpolarizable force fields may be inadequate to describe a strongly hydrogen-bonded system such as a cellulose crystal. Therefore, whether the supramolecular twist is an artifact of the empirical force fields used in MD simulations or a consequence of the chiral nature of the cellulose chains (French & Johnson, 2009; Selinger *et al.*, 2001) is still a matter of debate.

6. Cellulose oligomers in aqueous solution

Molecular-dynamics simulations of cellulose oligomers in water solution date back to the work of Hardy & Sarko (1993*a*,*b*). Although limited to a timescale of a few hundred picoseconds, those computational studies present a thorough statistical mechanical analysis of the relevant conformational degrees of freedom, the intramolecular hydrogen bonds and the solvation properties of cellobiose, cellotetraose and cellooctaose oligomers. Recently, cellulose oligomers (cellobiose, cellotetraose and cellohexaose) have been studied using explicit solvent replica-exchange molecular dynamics (REMD; Shen *et al.*, 2009). Both the simulation timelength (192–672 ns) and the use of REMD allowed an extensive sampling of the oligomers' conformational phase space.

In contrast to what is typically observed in non-interacting polymers, cellulose oligomers have been shown to become more rigid as the degree of polymerization increases. This trend is mainly a consequence of the increased number of intrachain hydrogen bonds in the longer chains. In particular, hydrogen bonds between adjacent cellulose units tend to restrain the flexibility of the glycosidic linkage and therefore to increase the oligomer persistence length. This study also served as a measure of the quality of the GLYCAM force field when applied to cellulose oligomers in an aqueous environment.

7. Coarse-grained and statistical mechanical models

To our knowledge, the only coarse-grained molecular model for cellulose has been proposed recently by Bu and coauthors in their study of the interactions of the carbohydrate-binding module (CBM) from *Trichoderma reesei* (represented at atomistic resolution) with crystalline cellulose I_β (Bu *et al.*, 2009). The coarse-grained model, based on a three-beads representation of the glucose unit and fitted against fully atomistic MD data, was used to build a hydrophobic surface of cellulose I_β.

Recently, a statistical mechanical two-dimensional lattice model has been used to analyze the thermodynamics of both intrachain and interchain/intrasheet hydrogen bonds in single layers of crystalline cellulose I_β over a large temperature interval (Shen & Gnanakaran, 2009). It was found that multiple alternative hydrogen-bond patterns can exist within the crystalline layer and that such 'plasticity' of the hydrogen-bond network greatly contributes to the stability of the layers over a wide range of temperatures.

8. Interactions of enzymes with crystalline cellulose

The interactions between enzymes and crystalline cellulose I_β have been the subject of a series of extensive molecular-dynamics studies focusing on the dynamics of cellobiohydrolase I (CBHI) from *T. reesei* on a crystalline I_β cellulose surface (Nimlos *et al.*, 2007; Zhong *et al.*, 2008, 2009; Bu *et al.*, 2009; Beckham *et al.*, 2010). Aside from confirming the main role of hydrophobic forces in the enzyme–cellulose interaction process, these studies revealed important features such as an induced-fit conformational change of both the CBM alone and the complete CBHI molecule upon binding (enabling the critical hydrophobic interaction between tyrosine and the cellulose surface). It was also observed that the CBM tends to diffuse away from hydrolyzed glycosidic bonds and that the potential energy surface for the CBM–cellulose interaction

displays stable energy minima corresponding to a cellobiose unit along a chain on the hydrophobic face of crystalline cellulose. A docking study of the CBMI with crystalline cellulose I_α has recently been published by Yui *et al.* (2010). This study confirms the preferential binding of the CBMI to the hydrophobic (110) surface. The two deepest minima in the binding potential energy surface were found to have a separation along the main axis of the cellulose fiber consistent with the twofold helical symmetry of the cellulose chain (cellobiose unit) and to occur when the CBMI was placed in an antiparallel orientation with respect to the main axis of the cellulose fiber.

Owing to length and time-scale limitations, it is not possible to simulate and analyze the entire crystalline cellulose-degradation process using fully atomistic computational models. A multi-resolution computational approach was recently employed in the study of the dynamics of individual enzymes on a crystalline cellulose surface. In this study, a coarse-grained model for the crystalline cellulose surface was coupled with a fully atomistic representation of the enzyme (Bu *et al.*, 2009).

Kinetic models have been used to analyze the physico-chemical properties of the whole hydrolysis process. In these models, the enzyme and substrate concentrations evolve deterministically according to a set of coupled ordinary differential equations (reaction-rate equations). Although the deterministic approach gives realistic results for large, well mixed and thermally equilibrated systems, it cannot capture important spatial details related to the structure of the crystalline cellulose substrate and to the specificity of the binding sites. An in-depth analysis of these models is beyond the scope of this review article; therefore, we refer the interested reader to the review articles by Zhang & Lynd (2004) and by Bansal *et al.* (2009) and to the recent studies by Zhou and coworkers (Zhou, Hao *et al.*, 2009; Zhou, Schüttler *et al.*, 2009) and by Ting *et al.* (2009).

9. Future directions

Future *in silico* studies on crystalline cellulose need to consider longer timescale MD simulations. All-atom MD simulation timescales of the order of hundreds of nanoseconds to submicroseconds are possible (considering the current advances in parallel high-performance computing) for some of the systems mentioned in this review. These long-time simulations will improve the conformational phase-space sampling and will be essential to obtain more consistent data on both the crystal thermal modes and the mechanical properties of crystalline cellulose fibrils. In addition, they will serve as a reliable tool for testing the quality of the different empirical force fields used in cellulose simulations. When an extensive thermodynamics analysis of the system of interest is needed, standard long-time MD simulations should be coupled with enhanced sampling algorithms such as REMD. Advanced free-energy calculation algorithms (Darve *et al.*, 2008) will also be essential for analyzing the relative free-energy change involved in cellulose crystal shape transitions (for example, from cellulose I_β to cellulose III_I) and in the binding process between cellulases and the cellulose fibril surface.

More efforts need to be directed towards the development of coarse-grained molecular models for cellulose crystals. These models will complement fully atomistic MD simulations and will allow the exploration of both time and size scales that are currently not accessible using atomistic calculations. An important potential application of these coarse-grained models is the study of large-scale plant and wood biomass systems (made up of cellulose fibrils, hemicellulose and lignin) over extended simulation times. The main challenges in the development of such models will be associated with the current limited knowledge of both the plant cell wall at the molecular and supramolecular levels and the nature of the interactions (covalent and intermolecular) between cellulose fibrils, hemicellulose and lignin.

Additional comparative studies on different cellulose allomorphs are also needed. These studies need to be carried out under identical conditions with the same force field so that the structural and chemical properties can be probed on the same platform. Future comparative all-atom simulation studies are expected to provide an understanding of the molecular forces that lead to cellulose adopting different crystal forms. Recent experiments show that when cellulose I_β is converted into cellulose III_I (*via* NH_3 pretreatment; Wada *et al.*, 2004) the enzymatic hydrolysis rate increases 2–5 times (Chundawat *et al.*, manuscript in preparation). Currently, we are carrying out all-atom MD comparative studies of cellulose I_β and cellulose III_I. The goal of these studies is to gain a better understanding of the relative differences between the two allomorphs in terms of structural, thermal and solvation properties and of their impact on the enzyme binding and digestibility of cellulose. An important aspect of these comparative simulation studies is the choice of the shape of the cellulose microfibrils. Indeed, the selection of the shape (hexagonal, diagonal or square; Matthews *et al.*, 2006) may have a significant effect on determining the microfibril (meta)stability. For example, in a hypothetical comparative study of the structural and solvation properties of two cellulose allomorphs (*i.e.* cellulose I_β and cellulose III_I) a rational choice would be to study two microfibrils (for cellulose I_β and cellulose III_I, respectively) with a similar form factor (hexagonal, diagonal or square), a similar solvent-accessible surface area and a similar ratio between hydrophobic and hydrophilic surfaces.

Finally, for the study of the entire biomass-degradation process, we believe that both stochastic and mechanistic kinetic models should be considered as essential modeling tools for complementing both all-atom MD and coarse-grained simulations. Future studies need to capture some of the factors that affect the heterogeneous cellulose catalysis process that are poorly understood. It is well known that the overall efficiency of this heterogeneous catalysis process depends on factors such as adsorption, desorption and diffusion rates on the insoluble cellulose substrate and on processivity. Currently, we are constructing a coarse-grained stochastic dynamical model for simulating the overall hydrolysis of crystalline cellulose. In this model, the catalysis process

is broken down into distinct parts related to different kinetic events performed by individual particles (enzymes). These events are essentially chemical reactions that take place on the surface of cellulose catalyzed by enzymes (adsorption, breakage of hydrogen bonds, cleavage of glycosidic bonds, desorption) and constitute the main elements of this model. Coordination of these events may happen based on Gillespie's algorithm (Gillespie, 1976) by constructing a numerical realization of overall hydrolysis in time or by following and updating the state (based on some predefined rules) of each individual particle in the system.

We acknowledge funding from the LANL LDRD program. The authors would like to thank an anonymous reviewer of this manuscript for providing insightful comments and suggestions. GB would also like to thank S. Chundawat for useful discussions.

References

Atalla, R. H. & VanderHart, D. L. (1984). *Science*, **223**, 283–285.
Bansal, P., Hall, M., Realff, M. J., Lee, J. H. & Bommarius, S. (2009). *Biotechnol. Adv.* **27**, 833–848.
Beckham, G. T., Matthews, J. F., Bomble, Y. J., Bu, L., Adney, W. S., Himmel, M. E., Nimlos, M. R. & Crowley, M. F. (2010). *J. Phys. Chem. B*, **114**, 1447–1453.
Bergenstrahle, M., Berglund, L. A. & Mazeau, K. (2007). *J. Phys. Chem. B*, **111**, 9138–9145.
Bergenstrahle, M., Thormann, E., Nordgren, N. & Berglund, L. A. (2009). *Langmuir*, **25**, 4635–4642.
Bergenstrahle, M., Wohlert, J., Larsson, P. T., Mazeau, K. & Berglund, L. A. (2008). *J. Phys. Chem. B*, **112**, 2590–2595.
Biermann, O., Hädicke, E., Koltzenburg, S. & Müller-Plathe, F. (2001). *Angew. Chem. Int. Ed.* **40**, 3822–3825.
Bu, L., Beckham, G. T., Crowley, M. F., Chang, C. H., Matthews, J. F., Bomble, Y. J., Adney, W. S., Himmel, M. E. & Nimlos, M. R. (2009). *J. Phys. Chem. B*, **113**, 10994–11002.
Clark, G. L. & Parker, E. A. (1937). *J. Phys. Chem.* **41**, 777–786.
Darve, E., Rodriguez-Gomez, D. & Pohorille, A. (2008). *J. Chem. Phys.* **128**, 144120.
French, A. D. & Johnson, G. P. (2009). *Cellulose*, **16**, 959–973.
Gillespie, D. T. (1976). *J. Comput. Phys.* **22**, 403–434.
Hardy, B. J. & Sarko, A. (1993*a*). *J. Comput. Chem.* **14**, 831–847.
Hardy, B. J. & Sarko, A. (1993*b*). *J. Comput. Chem.* **14**, 848–857.
Hardy, B. J. & Sarko, A. (1996). *Polymer*, **37**, 1833–1839.
Heiner, A. P., Kuutti, L. & Teleman, O. (1998). *Carbohydr. Res.* **306**, 205–220.
Heiner, A. P., Sugiyama, J. & Teleman, O. (1995). *Carbohydr. Res.* **273**, 207–223.
Heiner, A. P. & Teleman, O. (1997). *Langmuir*, **13**, 511–518.
Ito, T., Hirata, Y., Sawa, F. & Shirakawa, N. (2002). *Jpn. J. Appl. Phys.* **41**, 5809–5814.
Kroon-Batemburg, L. M. J., Bouma, B. & Kroon, J. (1996). *Macromolecules*, **29**, 5695–5699.
Kuttel, M., Brady, J. W. & Naidoo, K. J. (2002). *J. Comput. Chem.* **23**, 1236–1243.
Langan, P., Nishiyama, Y. & Chanzy, H. (2001). *Biomacromolecules*, **2**, 410–416.
Matthews, J. F., Skopec, C. E., Mason, P. E., Zuccato, P., Torget, R. W., Sugiyama, J., Himmel, M. E. & Brady, J. W. (2006). *Carbohydrate Res.* **341**, 138–152.
Mazeau, K. (2005). *Cellulose*, **12**, 339–349.
Mazeau, K. & Heux, I. (2003). *J. Phys. Chem. B*, **107**, 2394–2403.
Mazeau, K. & Rivet, A. (2008). *Biomacromolecules*, **9**, 1352–1354.
Neyertz, S., Pizzi, A., Merlin, A., Maigret, B., Brown, D. & Deglise, X. (2000). *J. Appl. Polymer Sci.* **78**, 1939–1946.
Nimlos, M. R., Matthews, J. F., Crowley, M. F., Walker, R. C., Chukkapalli, G., Brady, J. W., Adney, W. S., Cleary, J. M., Zhong, L. & Himmel, M. E. (2007). *Protein Eng. Des. Sel.* **20**, 179–187.
Nishiyama, Y., Johnson, G. P., French, A. D., Forsyth, T. & Langan, P. (2009). *Biomacromolecules*, **9**, 3133–3140.
Nishiyama, Y., Langan, P. & Chanzy, H. (2002). *J. Am. Chem. Soc.* **124**, 9074–9082.
Nishiyama, Y., Sugiyama, J., Chanzy, H. & Langan, P. (2003). *J. Am. Chem. Soc.* **125**, 14300–14306.
Reiling, S. & Brickman, J. (1995). *Macromol. Theory Simul.* **4**, 725–743.
Selinger, J. V., Spector, M. S. & Schnur, J. M. (2001). *J. Phys. Chem. B*, **105**, 7157–7169.
Shen, T. & Gnanakaran, S. (2009). *Biophys. J.* **96**, 3032–3040.
Shen, T., Langan, P., French, A. D., Johnson, G. P. & Gnanakaran, S. (2009). *J. Am. Chem. Soc.* **131**, 14786–14794.
Ting, C. L., Makarov, D. E. & Wang, Z.-G. (2009). *J. Phys. Chem. B*, **113**, 4970–4977.
Wada, M., Chanzy, H., Nishiyama, Y. & Langan, P. (2004). *Macromolecules*, **37**, 8548–8555.
Wada, M., Nishiyama, Y. & Langan, P. (2006). *Macromolecules*, **39**, 2947–2952.
Yui, T. & Hayashi, S. (2007). *Biomacromolecules*, **8**, 817–824.
Yui, T. & Hayashi, S. (2009). *Cellulose*, **16**, 151–165.
Yui, T., Nishimura, S., Akiba, S. & Hayashi, S. (2006). *Carbohydr. Res.* **341**, 2521–2530.
Yui, T., Shiiba, H., Tsutsumi, Y., Hayashi, S., Miyata, T. & Hirata, F. (2010). *J. Phys. Chem. B*, **114**, 49–58.
Zhang, Y.-H. P. & Lynd, L. R. (2004). *Biotechnol. Bioeng.* **88**, 797–824.
Zhong, L., Matthews, J. F., Crowley, M. F., Rignall, T., Talon, C., Cleary, J. M., Walker, R. C., Chukkapalli, G., McCabe, C., Nimlos, M. R., Brooks, C. L. II, Himmel, M. E. & Brady, J. W. (2008). *Cellulose*, **15**, 261–273.
Zhong, L., Matthews, J. F., Hansen, P. I., Crowley, M. F., Cleary, J. M., Walker, R. C., Nimlos, M. R., Brooks, C. L. III, Adney, W. S., Himmel, M. E. & Brady, J. W. (2009). *Carbohydr. Res.* **344**, 1984–1992.
Zhou, W., Hao, Z., Xu, Y. & Schüttler, H.-B. (2009). *Biotechnol. Bioeng.* **104**, 275–289.
Zhou, W., Schüttler, H.-B., Hao, Z. & Xu, Y. (2009). *Biotechnol. Bioeng.* **104**, 261–274.