CELLULOSE CRYSTAL STRUCTURE AND FORCE FIELDS

Malin Bergenstråhle,[1,2] James Matthews[3], Michael Crowley[3] and John Brady[2]
1 Wallenberg Wood Science Center, KTH, Stockholm, Sweden
2 Dep of Food Science, Cornell University, Ithaca, NY, USA
3 NREL, Golden CO, USA

ABSTRACT

Classical molecular mechanics force fields for carbohydrates are widely used for molecular dynamics simulations of crystalline cellulose, in particular, cellulose Iβ. To investigate the impact of choice of force field on crystalline cellulose structure and properties we have performed a comparative study of four different carbohydrate force fields. Molecular dynamics simulations applying the different force fields were performed on a solvated cellulose Iβ crystal. The crystal consisted of 36 cellulose chains, each of them 40 glucose units long, arranged in a crystal manner with a square cross section. These simulations show that the differences in force fields are of great importance for the resulting relaxed cellulose structure. The orientation of the hydroxymethyl groups is a key parameter and an indicator of different hydrogen bonding patterns that may be found in crystalline cellulose.

INTRODUCTION

Cellulose structure has been under investigation since the early days of polymer science. The fact that cellulose has many possibilities to form hydrogen bonds, both intra- and intermolecular with other chains or surrounding water render it possible to organize in several ways. Experimentally, this ambiguity is also supported, since a number of different crystal allomorphs have been determined for cellulose (cellulose I, II, III_I, III_{II}, IV_I, IV_{II}) by experimental techniques such as X-ray or neutron diffraction [1].
A key parameter to cellulose structure is the behavior of the rotation of the hydroxymethyl group. It has three low energy rotameric conformations, defined by two letters (*tg*, *gg* and *gt*) referring to the trans- and gauche states of the dihedral angles O5-C5-C6-O6 and C4-C5-C6-O6 respectively. The monomeric unit of cellulose, glucose, has in aqueous solution a preference for *gg* conformation of the hydroxymethyl group, followed by *gt*, whereas *tg* is least populated [2]. In contrast, native cellulose Iα and Iβ have all hydroxymethyls in *tg* conformation [3,4], making it unique among β(1-4) linked saccharide structures.
The two co-existing native crystalline polymorphs, cellulose Iα and Iβ, differ mainly in the packing arrangement of their hydrogen bonded sheets. Both native polymorphs are layered structures where sheets stabilized with hydrogen bonds are stacked on top of each other, with no hydrogen bonding interactions between the sheets. Cellulose Iβ is the most common form in higher plants, and it has two different hydrogen bonding patterns, pattern A and B [3]. Pattern A is dominating and is illustrated in Figure 1. It is characterized by two intramolecular hydrogen bonds, one from HO3 to the O5 oxygen on next glucose unit and another from HO6 to the O2 oxygen on next glucose. These two hydrogen bonds may only be present (given the glycosidic linkage conformation typically found in cellulose crystals) if the hydroxymethyl group is in *tg* conformation.

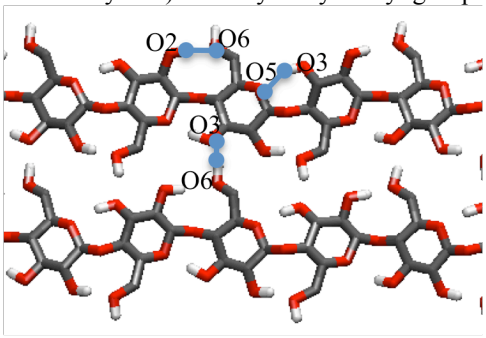


Figure 1: Intramolecular and intrasheet hydrogen bonds in cellulose Iβ hydrogen bonding pattern A

The increasing computational power and the development of force fields, has lead to better atomistic models and simulations of cellulose. Classical molecular mechanics force fields and molecular dynamics simulations have been widely used to study cellulose crystal structure, for instance in the studies reported in references [5-11]. These studies differ both in force fields and simulation protocols, but also in the modeled structure, such as length of cellulose chains, size of aggregate and solvated or vacuum simulation. The present study aims to make a fair comparison of four modern carbohydrate force fields, by simulating identical solvated cellulose crystal

structures under similar conditions.

METHOD

Molecular dynamics simulations of cellulose fibrils were performed at 300 K with four different force fields: Carbohydrate Solution Force Field (CSFF) [12], GLYCAM06 [13], GROMOS 45a4 [14] and a newly developed CHARMM Carbohydrate Force Field (C35) [15,16,17]. The simulated structure was a solvated crystal consisting of 36 glucan chains, each with a length of 40 glucose units arranged in a cellulose Iβ manner, starting from the two proposed hydrogen bonding networks A and B [3], with a square cross-section. The SPC water model [18] was used with the GROMOS 45a4 force fiels and TIP3P water [19] with the other force fields. Simulations extended to 100 ns to allow the structure to relax. Simulations were conducted with the AMBER PMEMD program [20,21,22] (CSFF, GLYCAM06 and C35) and GROMACS 4.0.2 [23] (GROMOS 45a4).

RESULTS AND DISCUSSION

All four force fields propose different relaxed structures. Many of the structural properties took tens of nanoseconds to converge, and all simulations were simulated 100 ns. For all, the distribution of different conformations of the hydroxymethyl group was investigated. The average structures from the last few ns of each simulation are shown in Figure 2, with different colors indicating the conformations of the hydroxymethyl groups.

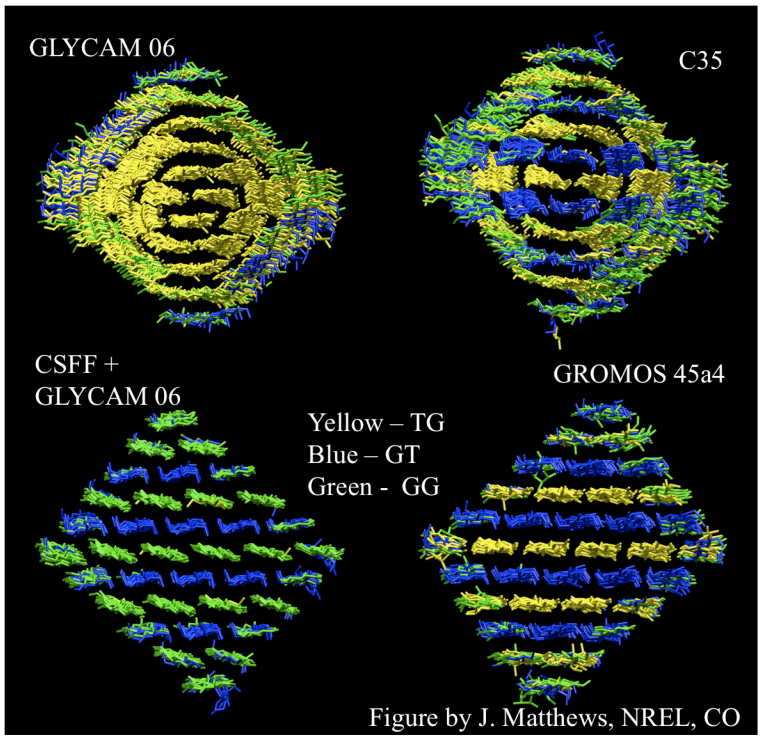


Figure 2: Average structures from the last few ns of each simulation. Colors indicate conformation of hydroxymethyl groups.

The GLYCAM06 force field is able to keep the crystal interior in a cellulose Iβ structure, but only when starting from network A. When starting from network B a structure similar to the CSFF structure evolves. The CSFF force field (and GLYCAM06 when starting from network B) produces a structure with alternating layers, where the chains in every second layer are tilted and hydrogen bonds are formed between the layers. Hydroxymethyl groups are found in *gt* and *gg* conformations. The C35 and GROMOS 45a4 force fields also have different hydrogen bonding patterns in every second layer. The hydroxymethyl groups are here altering between the *gt* and *tg* conformation in the crystal interior. There is no regular occurrence of hydrogen bonds between the layers. C35 and GLYCAM06 show a significant twist along the fiber axis, which is prevented by the interlayer hydrogen bonds in CSFF and existing but less significant in the GROMOS 45a4 structure.
The main conclusion we may draw from these results is that although all four force fields are developed for carbohydrates, their differences turn out to be of great importance when simulating cellulose fibril models at room temperature. The key point is the variety of hydrogen bonding patterns that may be formed within the

fibrils. Small differences in prerequisites for MD simulation of cellulose cause significant differences in final structure. It is of great importance to analyze these differences and try to point out what might cause them. It is also important to put a combined experimental and modeling effort into the understanding of cellulose structure. In addition to these simulations, high temperature (500 K) simulations were performed of the same structure with the same force fields. Interestingly, all force fields show common behavior at high temperature. This might be of importance when trying to understand how cellulose reacts to thermal treatment.

ACKNOWLEDGMENTS

We thank the U.S. DOE Office of Science ASCR SciDAC program for funding. Author MB thanks Alf de Ruvo Memorial Foundation of SCA AB for financial support. Computer time was provided by the TACC Ranger cluster under the National Science Foundation Teragrid grant number MCB090159 and by the NREL Computational Sciences Center supported by the DOE Office of EERE under contract number DE-AC36-08GO28308

REFERENCES

1. Nishiyama, Y., J. Wood Sci., **2009**, 241–249
2. Mason, P.E., Neilson, G.W., Enderby, J.E., Saboungi, M.-L., Cuello, G., Brady, J.W. **2006,** J. Chem. Phys., 125, art. no. 224505
3. Nishiyama, Y., Langan, P. and Chanzy, H., *J. Am. Chem. Soc.* **2002**, 124, 9074 – 9082
4. Nishiyama, Y., Sugiyama, J., Chanzy, H. and Langan, P., *J. Am. Chem. Soc.* **2003**, 125, 14300 – 14306
5. French, A.D., Miller, D.P., Aabloo A., *Int. J. Biol. Macromol.* **1993**, 15, 30-36
6. Matthews, J.F., Skopec, C.E., Mason, P.E., Zuccato, P., Torget, R.W., Sugiyama, J., Himmel,M.E., Brady, J.W., *Carbohydr. Res.*, **2006**, 341, 138 - 152
7. Yui, T. and, Hayashi, S., *Biomacromolecules* **2007** *8*, 817-824
8. Heiner, A.P., Sugiyama, J., Teleman, O., *Carbohydr. Res.*, **1995,** 273, 207-223
9. Mazeau, K. and, Heux, L., *J. Phys. Chem. B* **2003,** *107*, 2394-2403
10. Bergenstråhle, M., Berglund, L.A. and Mazeau, K., *J. Phys. Chem. B* **2007** *111*, 9138-9145
11. Tanaka, F. Iwata, T., *Cellulose*, **2006**, 13, 509 - 517
12. Kuttel, M., Brady, J.W. and Naidoo, K.J. *J.Comput. Chem*. **2002**, 23, 1236 – 1243
13. Kirschner, K.N., Yongye, A.B., Tschampel, S.M., Daniels, C.R., Foley, B.L, Woods, R.J. *J. Comput. Chem*. **2008**, 29, 622-655
14. Lins, R.D. and Hünenberger, P.H., *J. Comput. Chem*. **2005**, 26, 1400 – 1412
15. Guvench, O., Greene, S.N., Kamath, G., Brady, J.W., Venable, R.M., Pastor, R.W., MacKerell, Jr., A.D., *J. Comput. Chem.*, **2008**, 29, 2543-2564
16. Hatcher, E. Guvench, O., MacKerell, Jr., A.D J. Chem. Theory Comput., **2009**, 5, 1315-1327
17. Guvench, O., Hatcher, E., MacKerell, Jr., A.D **2010**, submitted
18. Berendsen, H.J.C., Postma, J.P.M, van Gunsteren, W.F. and Hermans, J., **1981** in *Intermolecular Forces*, Pullman, B. Ed., Reidel:Dordrecht, The Netherlands
19. Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W., Klein, M.L., **1983**, *J. Chem. Phys.* 79, 926-935
20. Case, D.A, Darden, T.A., Cheatham, T.E, Simmerling, C.L., Wang, J.,Duke, R.E Luo, R., Walker, R.C., Zhang, W., Merz, K.M., Roberts, B., Wang, B., Hayik, S., Roitberg, A., Seabra, G., Kolossváry, I., Wong, K.F., Paesani, F., Vanicek, J., Wu, X., Brozell, S.R., Steinbrecher, T., Gohlke, H., Cai, Q., Ye, X., Wang, J., Hsieh, M.-J., Cui, G., Roe, D.R., Mathews, D.H., Seetin, M.G., Sagui, C., Babin, V., Luchko, T., Gusarov, S., Kovalenko, A. and Kollman, P.A., **2010**, AMBER 11, University of California, San Francisco
21. Grand, S.L., Goetz, A.W., Xu, D., Poole, D., and Walker, R.C., **2010** (in preparation)
22. Grand, S.L., Goetz, A.W., Xu, D, Poole, D. and Walker, R.C. **2010** (in preparation)
23. Hess, B., Kutzner, C., van der Spoel, D. and Lindahl, E. *J. Chem. Theory Comput.* **2008**, *4*, 435-447

Cellulose crystal structure and force fields

Presented by:
Malin Bergenstråhle
PhD
Wallenberg Wood Science Center, Stockholm, Sweden

WWSC
WALLENBERG WOOD SCIENCE CENTER

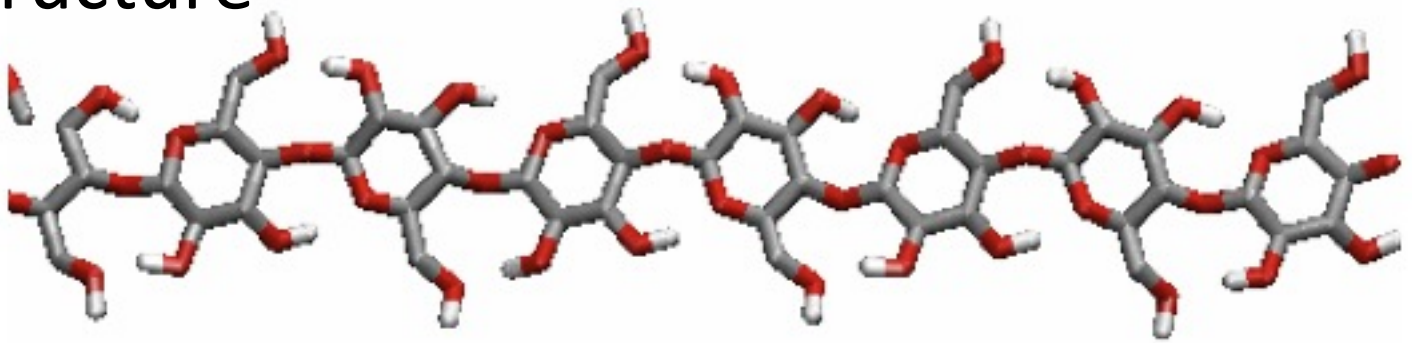
KTH VETENSKAP OCH KONST

CHALMERS

TAPPI
people resources solutions®

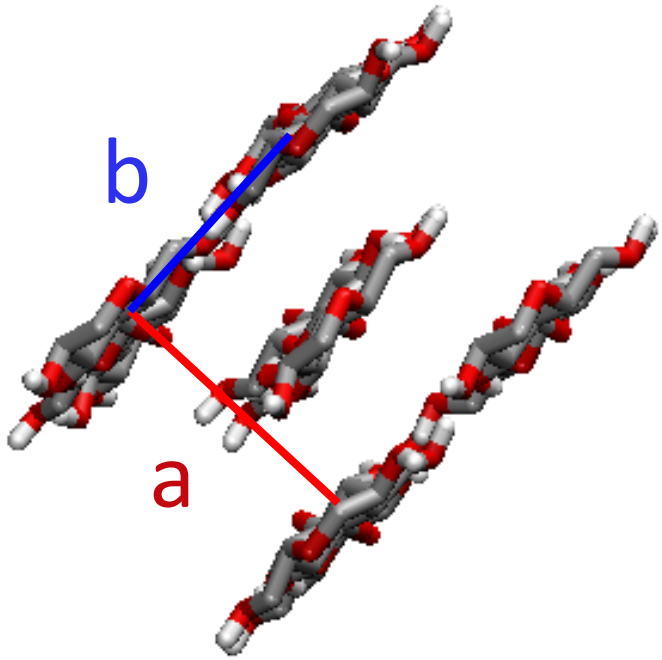
Cellulose structure

- Glucan chains aggregates into slender fibrils
- Fibrils have ordered structure, give strong X-ray diffraction patterns.
- Hydrogen bonds and vdW forces stabilize the structure



Cellulose Iβ

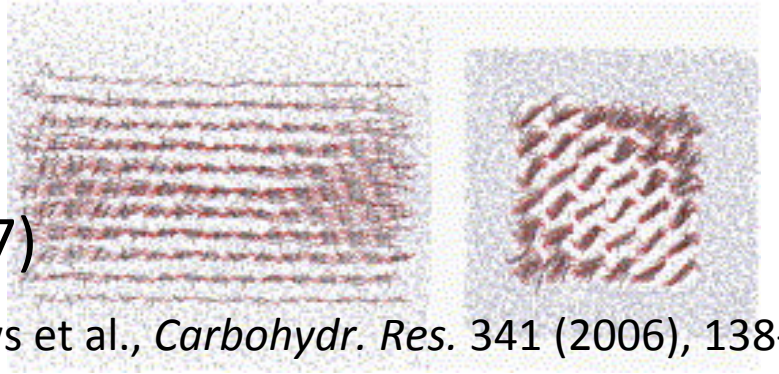
Parameter	Exp. value
a [nm]	0.78[1]
b [nm]	0.82[1]
c [nm]	1.03[1]
α nm	90.0[1]
β [nm]	90.0[1]
γ [nm]	96.5[1]
Density ρ [kg/m^3]	1636[1]
Torsion angle ω [°]	180[2]
Young's modulus in **c** direction [GPa]	134[3]



[1] Y. Nishiyama et al, J.A.C.S 2002 [2] K. Bock et al J. Carboh. Ch. 1994 [3] I. Sakurada et al, J. Polym. Sci. 19621

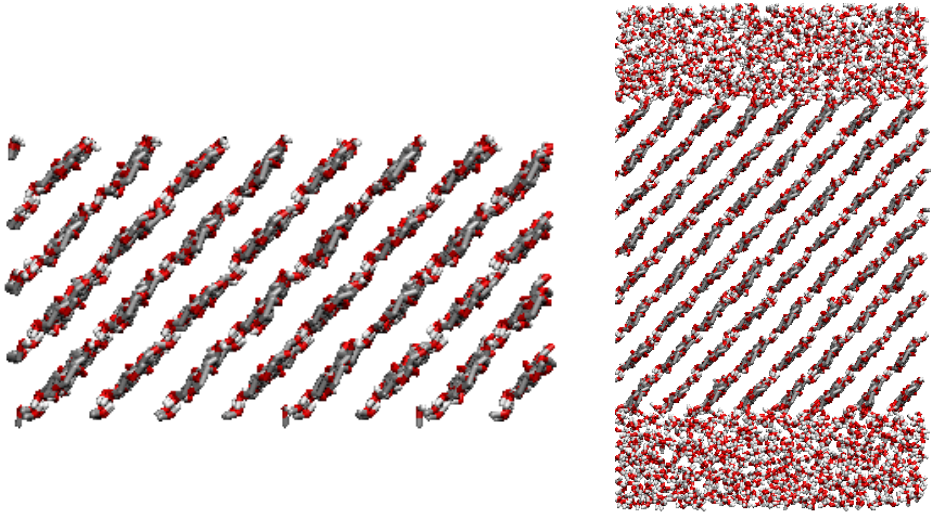
Computer simulations of cellulose

- Minicrystal
  (French et al. 1993,
  Matthews et al. 2006, Yui et al. 2007)



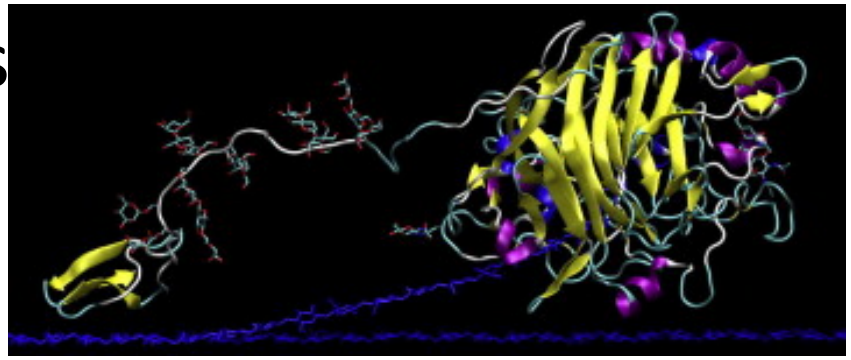
Matthews et al., *Carbohydr. Res.* 341 (2006), 138-152

- Bulk crystal
  (Heiner et al. 1995,
  Mazeau and Heux 2003,
  Bergenstråhle et al. 2007)
- Surface dynamics
  (Bergenstråhle et al. 2008)



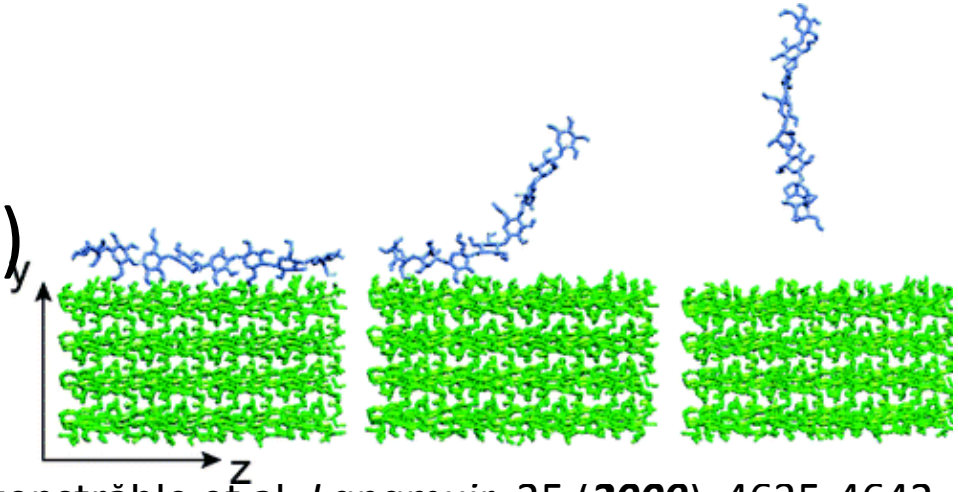
Computer simulations of cellulose

- Enzyme-cellulose interactions (Zhong et al. 2009)



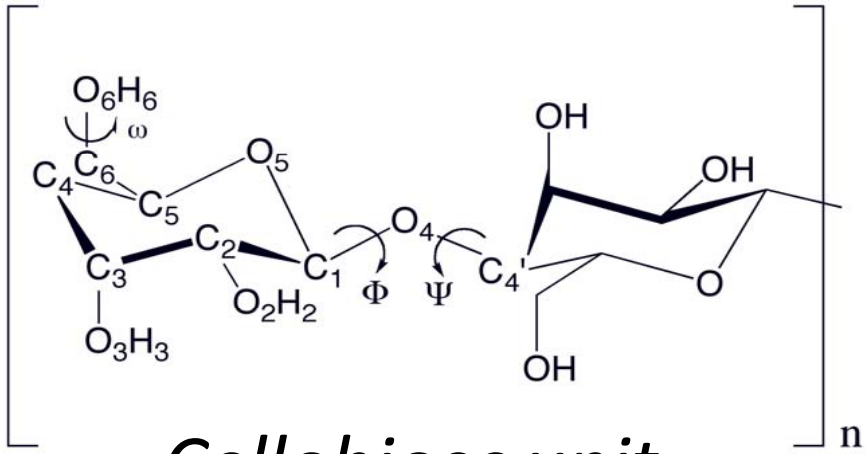
Zhong et al. *Carbohydr. Res.*, 344 (**2009**) 1984-1992

- AFM pulling (Bergenstråhle et al. 2009)



Bergenstråhle et al. *Langmuir*, 25 (***2009***) 4635-4642

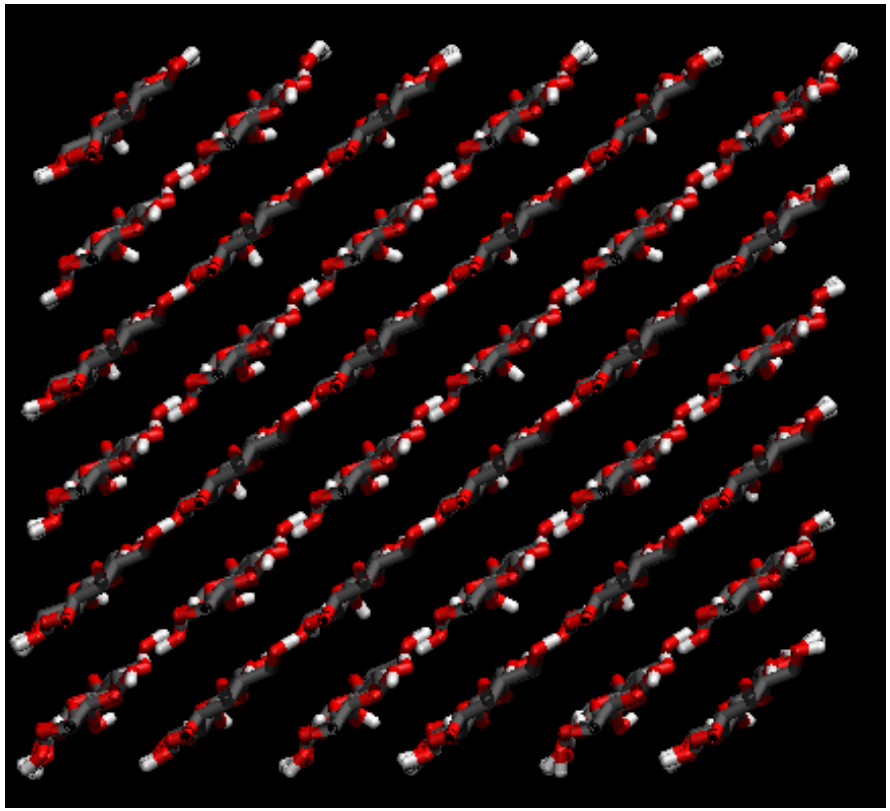
Differences in cellulose models



Cellobiose unit

- Chain length
- Aggregate size
- Solvated or vacuum
- Force Fields
- Simulation protocol

Force field comparison



- Solvated finite crystal
- DP 40
- 36 chains in cellulose Iβ
- 100 nanoseconds at 300 K
- 4 different Force Fields

Force fields

- GLYCAM 06[1]
- CSFF [2]
- GROMOS 45a4[3]
- CHARMM 35[4-6]

1. Kirschner, K.N., Yongye, A.B., Tschampel, S.M., Daniels, C.R., Foley, B.L, Woods, R.J. *J. Comput. Chem*. **2008**, 29, 622-655
2. Kuttel, M., Brady, J.W. and Naidoo, K.J. *J.Comput. Chem*. **2002**, 23, 1236 – 1243
3. Lins, R.D. and Hünenberger, P.H., *J. Comput. Chem*. **2005**, 26, 1400 – 1412
4. Guvench, O., Greene, S.N., Kamath, G., Brady, J.W., Venable, R.M., Pastor, R.W., MacKerell, Jr., A.D., *J. Comput. Chem.*, **2008**, 29, 2543-2564
5. Hatcher, E. Guvench, O., MacKerell, Jr., A.D J. Chem. Theory Comput., **2009**, 5, 1315-1327
6. Guvench, O., Hatcher, E., MacKerell, Jr., A.D **2010**, submitted

Structural features of cellulose

Conformation of hydroxymethyl group

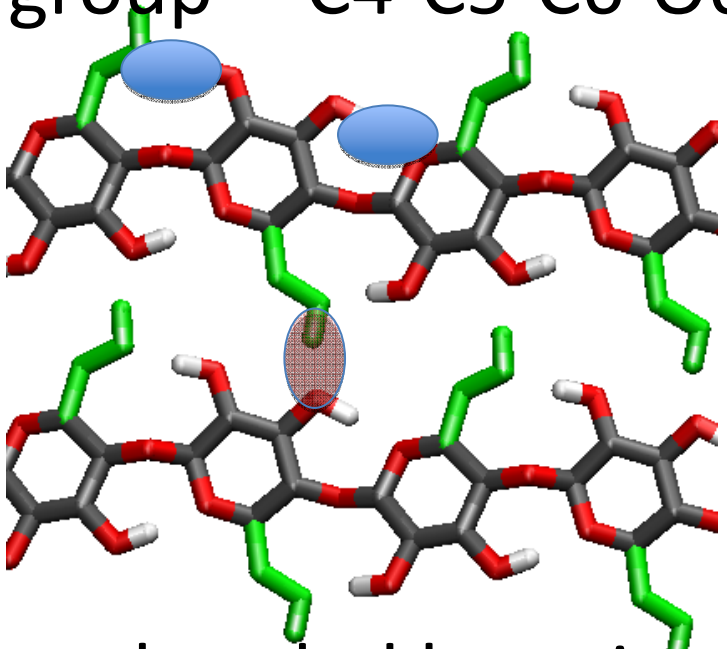
O5-C5-C6-O6

C4-C5-C6-O6

TG (180°)

GG (300 °)

GT (60°)



Hydrogen bonded layer in cellulose Iβ

Average structures after 100 ns

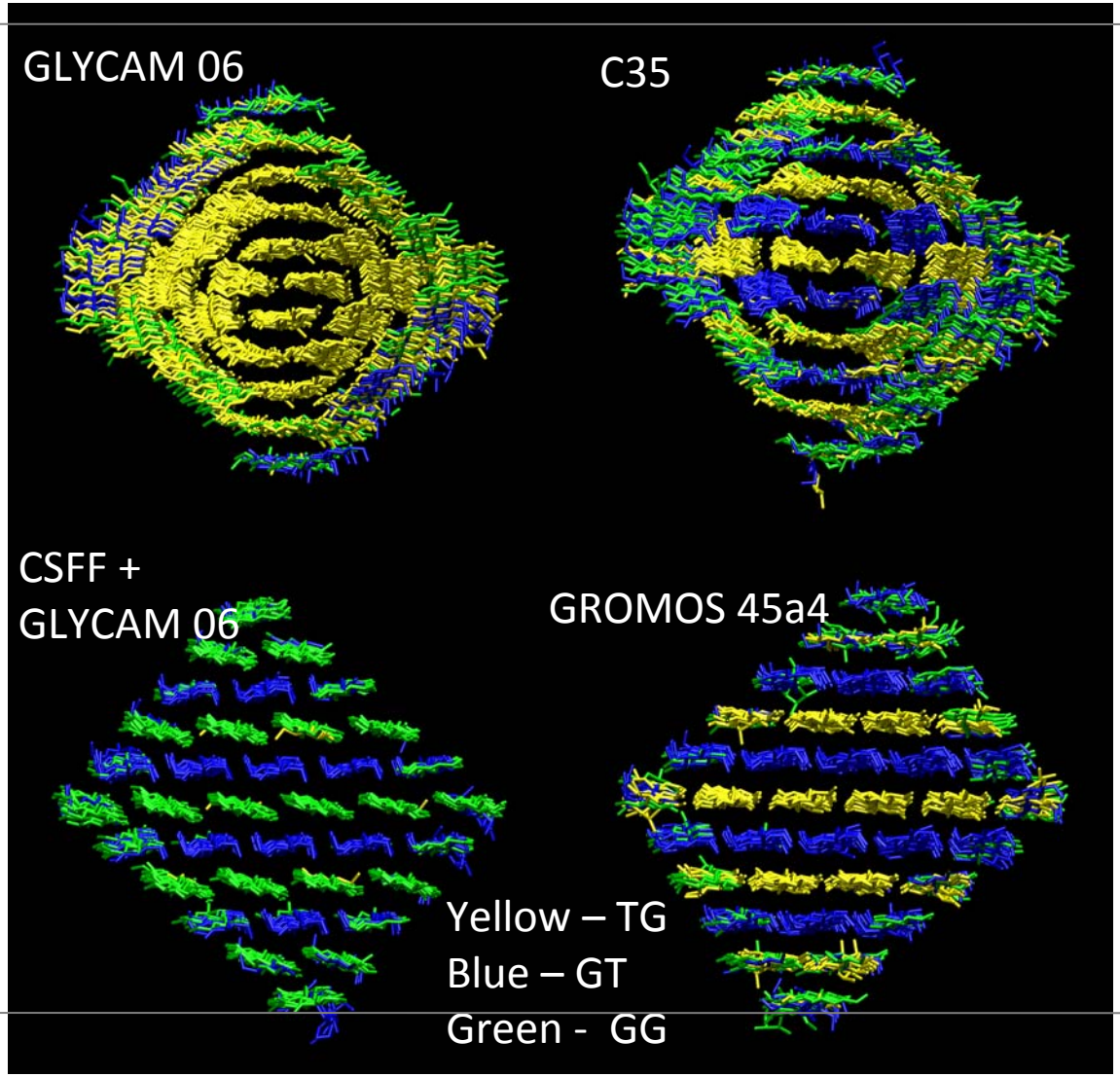


Figure by J. Matthews, NREL, CO

Conclusions

- All FFs give different structures.
- Structure is extremely sensitive to small variations in parameters.
- Analysis of FF differences required
- Comparison with experimental data important

Acknowledgments

Coauthors
James Matthews (NREL, Golden, CO)
Michael Crowley (NREL, Golden, CO)
John Brady (Cornell University, Ithaca, NY)

Financial Support
U.S. Department of Energy, Office of Science
Alf de Ruvo Memorial Foundation, SCA AB

Thank you

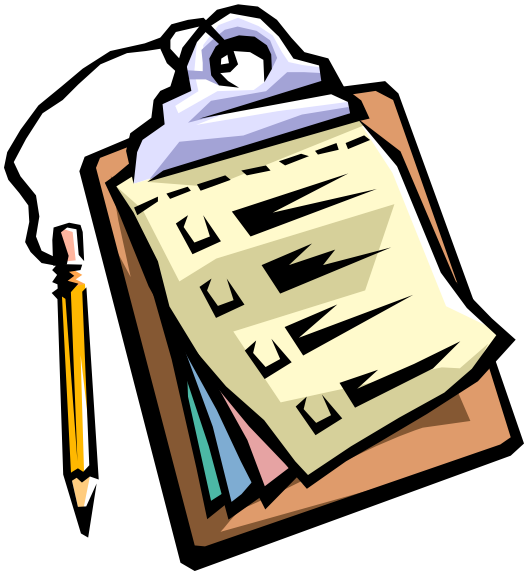
PRESENTED BY

Malin Bergenstråhle

PhD

Wallenberg Wood Science Center

malbe@kth.se



Please remember to turn in your evaluation sheet...

WWSC
WALLENBERG WOOD SCIENCE CENTER

KTH VETENSKAP OCH KONST
CHALMERS