^{13}C Chemical Shift Constrained Crystal Structure Refinement of Cellulose I_α and Its Verification by NMR Anisotropy Experiments

Raiker Witter,*,† Ulrich Sternberg,† Stephanie Hesse,‡,||,# Tetsuo Kondo,||,⊥ Frank-Th. Koch,‡ and Anne S. Ulrich†,§

Institute of Biological Interfaces, Forschungszentrum Karlsruhe, POB 3640, 76021 Karlsruhe, Germany; Institute of Optics and Quantum Electronics, Friedrich-Schiller-Universität Jena, Max-Wien-Platz 1, Jena 07743, Germany, and Forestry and Forest Products Research Institute, Matusnosato 1, Tsukuba, Ibaraki 305-8687, Japan; and Institute of Organic Chemistry, University of Karlsruhe, Fritz-Haber-Weg 6, 76131 Karlsruhe, Germany

Received November 15, 2005; Revised Manuscript Received June 27, 2006

ABSTRACT: The solid-state NMR assignments of the ^{13}C resonances of bacterial cellulose I_α were reinvestigated by INADEQUATE experiments on uniformly ^{13}C-enriched samples from *Acetobacter xylinum*. Additionally, we determined the principal chemical shift tensor components of each ^{13}C labeled site from a 2D iso-aniso RAI (recoupling of anisotropy information) spectrum acquired at magic angle spinning speed of 10 kHz. On the basis of these NMR data, the crystal structure of cellulose I_α was refined using the ^{13}C chemical shifts for target functions. Starting off with coordinates derived from neutron scattering, our molecular dynamics simulations yielded four ensembles of 200 structures, two ensembles for hydrogen bond scheme A and B and two ensembles for different chemical shift assignments I and II, giving 800 structures in total. These were subsequently geometry-optimized with the given isotropic chemical shift constraints applying crystallographic boundary conditions, to identify a structure for every ensemble that fit best to the experimental NMR data. The resulting four model structures were then assessed by simulating the chemical shift tensors (using the bond polarization theory) and comparing these values with the experimental chemical shift anisotropy information (obtained by RAI). The earlier neutron diffraction study had reported two possible occupation schemes for the hydrogen-bonded hydroxyl-groups (A, B) which connect the cellulose chains. From these two possibilities, our NMR results single out pattern A as the most probable structure. In this work, the first time crystallographic boundary conditions were applied for ^{13}C chemical shift structure refinement for molecular dynamics simulations and Newton−Raphson geometry optimization.

Introduction

Natural cellulose is a partially crystalline polymer of 1−4 linked β-D-glucose residues. Its structure and NMR investigations thereof are discussed by Sternberg et al.[1] VanderHart and Atalla[2] revealed the presence of two allomorphs, I_α and I_β, by CP-MAS ^{13}C NMR studies of highly crystalline native cellulose I. They published a first assignment for the resonances of C1, C4, and C6. The cluster of signals between 70 and 80 ppm was attributed to carbons C2, C3, and C5. These results were confirmed later using selectively ^{13}C-labeled cellulose[3,4] and by solid-state INADEQUATE NMR.[5] In both cases the C2, C3, and C5 chemical shifts were resolved and assigned. Recently, Kono et al.[6] assigned all ^{13}C signals to the respective carbon sites in the two different anhydroglucose rings of purified *Cladophora* cellulose (I_α) and tunicate cellulose (I_β). Additionally, Jaeger et al.[7] assigned all carbon sites in uniformly ^{13}C-enriched bacterial cellulose, for which slightly different isotropic chemical shift values were found. New results on cellulose based on correlation spectroscopy are given by Cadars et al.[8] and Sakellariou et al.[9]

* Corresponding author. E-mail: Raiker.Witter@ibg.fzk.de.
† Forschungszentrum Karlsruhe.
‡ Friedrich-Schiller-Universität Jena.
|| Forest Products Research Institute.
§ University of Karlsruhe.
⊥ Current address. Biomaterial Design Lab, Bioarchitecture Center(BAC) & Graduate school of Bioresource and Bioenvironmental Sciences, Kyushu University, 6-10-1, Hakozaki, Higashi-ku, Fukuoka 812-8581, Japan.
Current address. Centre of Excellence for Polysaccharide Research, Friedrich-Schiller-Universität Jena, Humboldtstrasse 10, Jena 07743, Germany.

In diffraction studies of cellulose fibers, the amorphous character of this microcrystalline material tends to produce poorly resolved diffraction patterns. Nevertheless, Reiling and Brickmann[10] constructed computer models of cellulose I_α from X-ray and electron diffraction studies,[11,12] and they performed force field refinements with periodic boundary conditions. First, precise atomic coordinates based on ^{13}C NMR chemical shift refinements were derived by Sternberg et al.[1] At the same time, the native cellulose structures were reinvestigated by Nishiyama et al.[13] with X-ray and neutron diffraction, which yielded information about hydrogen-bond networks. Hence, it is now of interest to compare the diffraction results with the newly refined NMR structures.

In this work, the recent crystal structure of Nishiyama et al.[13] is used as a starting model for the ^{13}C NMR structure refinement of cellulose I_α. The unit cell contains one chain consisting of two crystallographically different anhydroglucose units (see Figure 1). Therefore, 12 resonances should be observable by NMR. Their isotropic chemical shifts are used for direct structure refinement, based on a newly developed NMR force field which utilizes the chemical shift target functions introduced by Witter et al.[14,15,16]

Preparation of Bacterial Cellulose

Bacterial cellulose was produced by *Acetobacter xylinum* NQ-5 (strain ATCC 53582, from the collection of T. Kondo), and relevant details about their biosynthesis behavior in standard media can be found in Brown et al.[17] and Kondo et al.[18] This microbial strain is also quite robust for cultivation in ^{13}C-enriched media, as investigated by Hesse and Kondo.[19]

10.1021/ma052439n CCC: $33.50
Published on Web 08/09/2006

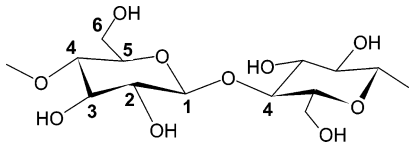

Figure 1. Two glucose residues of cellulose I.

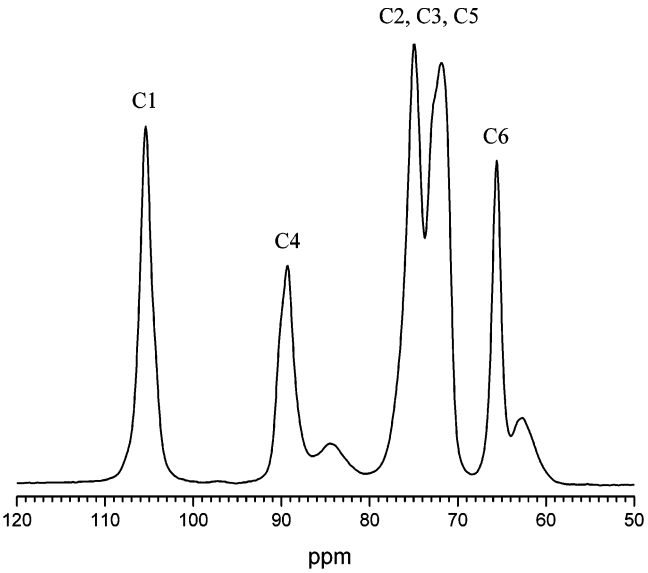

Figure 2. CP/MAS spectrum of uniformly ^{13}C-labeled cellulose from *A. xylinum*, showing the dominant I_α allomorph besides some amorphous contributions. Broadening due to homonuclear coupling can be clearly observed.

For cultivation, a total of 200 μL Schramm-Hestrin (SH)[20] medium, precultured for more than 10 days, was inoculated into 8.3 mL SH nutrient medium. Fully ^{13}C-enriched β-D-glucose-U-$^{13}C_6$ (^{13}C, 99%) was used as a carbon source for the final nutrient medium, while the preculture contained normal β-D-glucose. Bacteria were grown under static conditions, with the culture allowed to stand on a clean bench. After 14 days incubation at a temperature of 30 °C, the cellulose fleece was harvested and rinsed under running water for at least 24 h. It was then treated with 0.1 M NaOH solution at 80 °C for 4 h, and washed again with running water for at least 24 h. Finally, the cellulose fleece was not sterilized but instead covered with aluminum foil and dried in air at a temperature of 50 °C for about 24 h.

NMR Experiments

NMR experiments were carried out on a Bruker Avance 500 MHz wide-bore spectrometer at room temperature (298 K). The 1D CP/MAS ^{13}C NMR spectrum (Figure 2), and the 2D-refocused INADEQUATE[5] (Figure 3) and 2D-RAI[23] (Figure 4) data were obtained with a Bruker 4 mm H/F MAS probe. The MAS frequency was set to 10 kHz, except for the 2D-refocused INADEQUATE being run at 12.5 kHz. For cross-polarization, a ramped rf field with 5 ms proton irradiation was used. During data acquisition a standard two-pulse phase modulation ^{1}H-decoupling scheme[21] was applied with a decoupling power of 50 kHz. The recycle delay was set to 2 s for all experiments. For the 2D methods quadrature phase detection was carried out using the States method.[22] The 1D ^{13}C CP/MAS, 2D-refocused INADEQUATE, and 2D-RAI experiments were acquired with respective numbers of scans of 1024, 64, and 32. The indirect time dimension of the 2D experiments consisted of 128 data points with a dwell time of 80 μs for refocused INADEQUATE, and 64 points with a separation of 3× the rotor period (300 μs) for the RAI powder pattern recoupling. The total experiment times were 36 min, 4.8 h, and 1.2 h, respectively. The echo time of the refocused INADEQUATE was set to 3.04 ms. The 90 and 180° ^{13}C pulses were set to 2.4 μs and 5 μs.

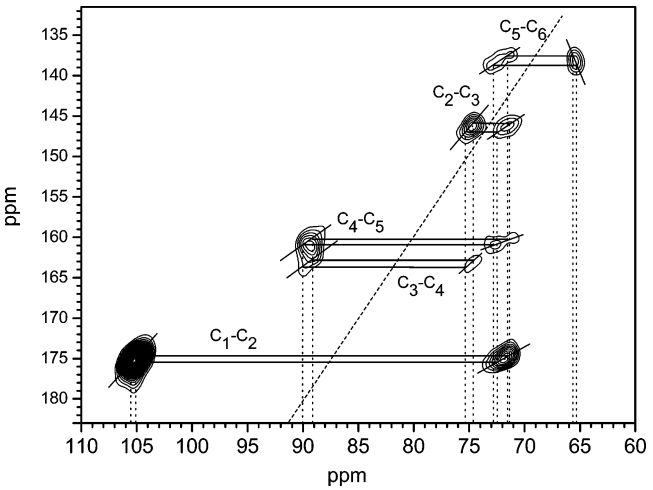

Figure 3. ^{13}C−^{13}C refocused INADEQUATE spectrum of cellulose I_α.

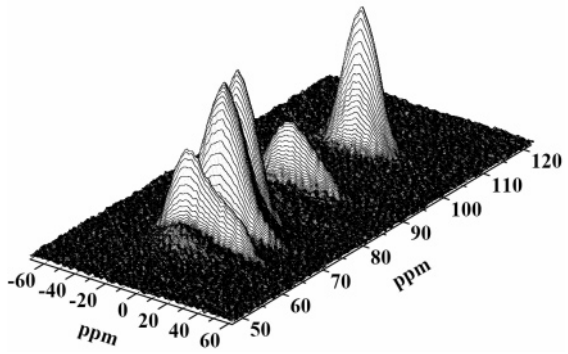

Figure 4. RAI spectrum of cellulose I_α, in which the individual CSA powder patterns are resolved along the isotropic dimension. The powder patterns were fitted to give an overall rmsd of 0.5 ppm for the tensor values.

Experimental NMR. Results

The ^{13}C NMR CP/MAS spectrum from the bacterially produced cellulose sample is shown in Figure 2. The sensitivity of the ^{13}C chemical shift to the chain conformations allows an unambiguous identification of the different allomorphs. This spectrum is dominated by the native allomorph of cellulose I_α, besides around 30% I_β and amorphous components.

The complete resonance assignment of the cellulose I_α signals is demonstrated in the refocused INADEQUATE spectrum (see Figure 3). The single quantum ^{13}C chemical shift is displayed in the horizontal dimension, and double quantum frequencies in the vertical dimension. The cross-peaks represent through-bond ^{13}C−^{13}C correlations between J-coupled carbons. From this, the following assignment of the two glucose units 1 and 2 was obtained: $C1_1$, 105.1 ppm; $C1_2$, 105.7 ppm; $C2_1$, 71.3 ppm; $C2_2$, 72.5 ppm; $C3_1$, 74.7 ppm; $C3_2$, 75.3 ppm; $C4_1$, 89.2 ppm; $C4_2$, 90.2 ppm; $C5_1$, 72.9 ppm; $C5_2$, 71.4 ppm; $C6_1$, 65.4 ppm; $C6_2$, 65.7 ppm.

To extract the full information about the chemical shift tensor we used the recoupling sequence RAI (recoupling of anisotropy information).[23,24] This experiment correlates the highly resolved isotropic chemical shift of each ^{13}C with the quasi-static powder pattern of the same site. The advantage of the RAI pulse scheme is that it can be applied at high MAS spinning speeds that average out the homonuclear dipolar interactions between carbons in a uniformly labeled sample. For cellulose, 10 kHz MAS turned out to be sufficient. The heteronuclear ^{1}H−^{13}C dipole coupling can be suppressed by proton decoupling, for

Table 1. Experimental ^{13}C Isotropic Chemical Shifts (δ_{iso}) with the Principal Axis Values (δ_{11}, δ_{22}, δ_{33}) of Cellulose I$_\alpha$, as Obtained from the RAI Spectrum of Figure 4

site	δ_{iso} (ppm)	δ_{11} (ppm)	δ_{22} (ppm)	δ_{33} (ppm)
		Glucose Unit 1		
C1	105.1	90.3	105.0	120.0
C2	71.3	53.2	72.2	88.5
C3	74.7	57.5	75.3	91.3
C4	89.2	65.3	93.1	109.2
C5	72.9	53.1	74.5	91.1
C6	65.4	35.4	73.2	87.6
		Glucose Unit 2		
C1	105.7	90.9	105.6	120.6
C2	72.5	53.6	74.2	89.7
C3	75.3	58.6	75.8	91.4
C4	90.2	65.9	94.1	110.6
C5	71.4	53.3	72.3	88.6
C6	65.7	35.7	73.5	87.9

which the ^{1}H-power was increased during the ^{13}C-pulses to satisfy the decoupling condition $\omega^{13}_C \approx 2\omega^1_H$, see Ishii et al.[25] Figure 4 shows that the isotropic ^{13}C chemical shifts are recorded with high resolution in the direct dimension, while the CSA powder patterns are recoupled in the indirect dimension of this 2D iso-aniso RAI spectrum. We are aware that the powder patterns are to some extent artificially broadened by the increasing number of recoupling pulses with increasing increments of t_1. However, this did not have any significant impact on the powder pattern fitting, and the mean deviation of the chemical shift tensor values is only 0.5 ppm. The tensor values were extracted from the 2D RAI spectrum using the software DMfit[26] which applies the Häberlen-Mehring-Spiess convention: $|\delta_{33} - \delta_{iso}| \geq |\delta_{11} - \delta_{iso}| \geq |\delta_{22} - \delta_{iso}|$.[27] The chemical shift anisotropy information is listed in Table 1.

Theory

In this work, the first time crystallographic boundary conditions were applied for ^{13}C chemical shift structure refinement for molecular dynamics simulations and geometry optimization. These calculations are based on the COSMOS (COmputer Simulations of MOlecular Structures) force field[28,29] with coordinate dependent charges. It utilizes the bond polarization theory,[30,31] a semiempirical approach, to calculate chemical shift tensors. This NMR property can be considered as a sum of one-electron operators, hence the chemical shift tensor is expressed within the BPT approach as

$$\delta_{\alpha\beta} = \langle\Psi_0|\hat{\delta}_{\alpha\beta}|\Psi_0\rangle = \sum_i^{i\in A} \sum_{\alpha'\beta'} D^i_{\alpha\alpha'} D^i_{\beta\beta'} (n_i \delta_i^{\alpha'\beta'} + n_i^2 A_i^{\alpha'\beta'} [\langle\chi^i_A|\hat{V}|\chi^i_A\rangle - \langle\chi^i_B|\hat{V}|\chi^i_B\rangle]) \quad (1)$$

The matrix elements $D_{\alpha\alpha'}$ describe the coordinate transformation from the bond orbital frame to the reference frame. The first sum runs over all bond contributions of atom A. The bond polarization matrix elements are given (in atomic units) by

$$\langle\chi^i_\lambda|\hat{V}|\chi^i_\lambda\rangle = \sum_x^{\text{charges}} \sum_k h_k^2 \int \phi^\lambda_k(r) \frac{Q_x}{|R_x - r|} \phi^\lambda_k(r)\, \mathrm{d}r^3 \quad (2)$$

with the charges Q_x at position R_x, the Slater type orbitals[32] $\phi^\lambda_k(r)$, and the bond orbital coefficients h_k. The first sum runs over all atomic charges of the molecular system. The bond tensor increments $\delta_i^{\alpha\beta}$ and polarization tensor parameters $A_i^{\alpha\beta}$ are obtained by calibration procedures.[30] To obtain these parameters, a collection of crystal structures and single-crystal chemical shift measurements[33,34] has to be used as an input for establishing a set of linear equations. To some extent, *ab initio* results can also be included in this parametrization procedure. Once the parameters $\delta_i^{\alpha\beta}$ and $A_i^{\alpha\beta}$ have been determined, only the matrix elements $\langle c^i_\lambda|\hat{V}|\chi^i_\lambda\rangle$ and the occupation numbers n_i[35] need to be calculated. Introducing point charges in the expression for the potential $\hat{V}$ leads to compact analytic expressions for the integrals, hence calculations within the BPT approach are highly efficient. In eq 1 there are two sums, the first runs over all bond contributions of the atom under consideration, and the second runs over all polarizing charges of the electrostatic potential $\hat{V}$. Hence the computational cost for a chemical shift calculation is proportional to the number of atoms N, once the charges are known.

As can be seen from eq 2, accurate atomic charges are also a prerequisite for BPT chemical shift calculations. The chemical shifts in this theory are proportional to bond polarization integrals that account for the change of the chemical shift caused by surrounding charge distributions. However, since semiempirical polarization parameters are introduced in the chemical shift calculations, the absolute values of the charges are not of concern. The polarization parameters $A_i^{\alpha\beta}$, on the other hand, will depend on the type of *ab initio* calculation and on the procedure for the population analysis. The atomic charges can be calculated within the BPT approach[36] in a manner analogous to eq 2:

$$Q_A = \sum_i^{i\in A} (n_i q_i + n_i^2 A_i^q [\langle\chi^i_A|\hat{V}|\chi^i_A\rangle - \langle\chi^i_B|\hat{V}|\chi^i_B\rangle]) \quad (3)$$

Overlap contributions are omitted when calculating the bond polarization integrals. By investigating the charge equations, it is obvious that the charges on atom A, Q_A, have to be estimated from all other charges Q_x. By taking the factors Q_x out of the integrals, we end up with a system of linear equations for Q_x and Q_A, with the sum over $n_i\, q_i$ as inhomogeneities:

$$Q_A = \sum_i^{i\in A} \left(n_i q_i + n_i^2 A_i^q \sum_x Q_x \left[\left\langle \chi^i_A \left| \frac{1}{|R_x - r|} \right| \chi^i_A \right\rangle - \left\langle \chi^i_B \left| \frac{1}{|R_x - r|} \right| \chi^i_B \right\rangle \right] \right) \quad (4)$$

To establish the basis for ^{13}C chemical shift calculations, we calibrated the parameters q_i and A_i^q against the calculated charges of a set of 175 molecules consisting of H, C, N, O, F, Si, P, S, Cl, and Zn atoms.[16] These calculations were performed using the 6-31G(d, p) basis set, and the atomic charges were obtained by NBO (natural bond orbital) population analysis of the ab initio charge distributions. The BPT and *ab initio* charges of small molecules correlated very well with $R = 0.996$. For details of the parametrization procedure and formalism, see Witter et al.[16]

The computational time for setting up all charge equations is proportional to N^2. Calculation of the charges means solving this set of linear equations, for which the number of floating point operations is proportional to N^3. In the COSMOS force field the atomic charges are recalculated in every step of the MD simulation or geometry optimization, and the BPT charges are used in the calculation of the Coulomb energy part of the force field. These Coulomb interaction energies have to be scaled in order to reach reasonable molecular interaction energies. A scaling factor of 0.87 was determined by comparing the interaction energies of a series of small molecules.[40] It is

thus possible to include all mutual polarizations in the Coulomb energy part of the force field. This turned out to be essential not only for highly charged systems but also for hydrogen bonds, since this bond type has a highly electrostatic nature. The BPT parametrization for the charge and chemical shift calculations is given in Sternberg et al.[1]

To apply the COSMOS force field for NMR-based structure calculations, a pseudo-force term has to be introduced, which depends on the difference between the observed and calculated chemical shifts:

$$F_j = k^{\mathrm{CS}}(\delta^{\mathrm{theo}} - \delta^{\mathrm{exp}})\frac{\partial\delta^{\mathrm{theo}}}{\partial x_j} \quad (5)$$

Since the chemical shifts can be expressed in terms of the atomic polarization energies,[16,37] the individual force constants for different atom types can be derived from the polarization parameters of the chemical shift and charges:

$$k^{\mathrm{CS}} = \sum_A \sum_i^{i\in A} \frac{2A_i^q}{n_i^2 (A^i)^2} \quad (6)$$

To derive forces from the pseudo-energies in addition to the chemical shifts, their derivatives with respect to the coordinates have to be calculated (see eq 5). This procedure is much more demanding than the chemical shift calculation alone, since derivatives of the bond polarization integrals with respect to the coordinates have to be calculated. The computational cost depends, as a first approximation, on the charge calculation which is proportional to the cube of the number of atoms N^3. Calculations on systems of about 10^4 atoms are feasible within a day on current standard GHz machines with an average performance of about 500 Mflops/s.

To perform realistic crystal simulations, the force field has to maintain strict lattice periodicity throughout the calculations:

$$\vec{F}(\vec{r}) = \vec{F}(\vec{r} + i\vec{a} + j\vec{b} + k\vec{c}), \{i,j,k\} = 0, \pm 1, \pm 2 \quad (7)$$

For every part of a molecule that is not within the unit cell, a code is stored to update the positions of the atoms, forces and charges from the central unit cell (analogous to eq 7). All other sites in the first or if necessary in a second surrounding of the central unit cell are generated temporarily by translations from the central unit cell. In the case of cellulose I_α, two translations were necessary, since the electrostatic cutoff (20 Å) was larger than the cell dimension. (The influence of the charge distribution on the lattice chemical shift will not give any significant contributions at distances above 20 Å.) The translation procedure to generate neighbor unit cells was not only performed for the intermolecular electrostatic and van der Waals energy part, but also within the calculations of the integrals needed for the BPT charge and chemical shift calculations. In this way, we included the polarizing influence of the crystal lattice strictly periodically in the interaction energies and chemical shifts.

Structure Refinement Results

Model Building and Evaluation. The starting point for our MD simulations and geometry optimization of cellulose I_α was the recently published crystal structure of Nishiyama et al.[13] The authors had performed X-ray and neutron diffraction experiments on deuterated samples. The cell parameters differ slightly from former electron diffraction results,[12] and for the first time it was possible to refine the proton (=deuteron) positions from diffraction data. In this seminal paper, two

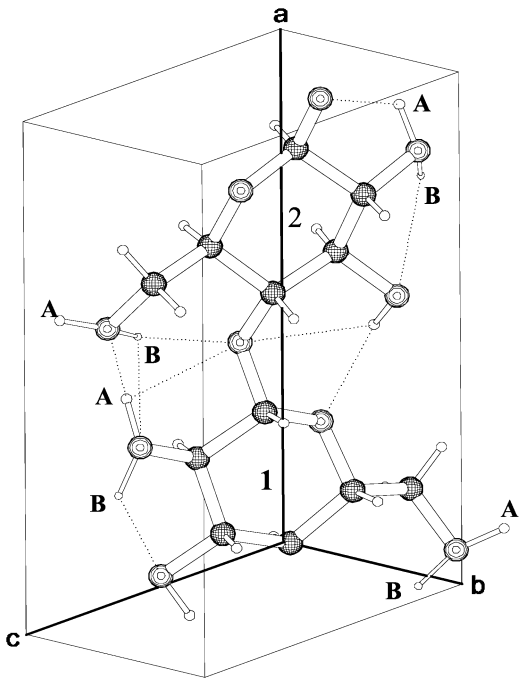

Figure 5. Unit cell of cellulose I_α according to the neutron diffraction study of Nishiyama et al.[13] (space group P1, with cell parameters a = 6.717, b = 5.962, c = 10.4, α = 118.08°, β = 114.8°, and γ = 80.37°). The two nonequivalent glucose rings are denoted with 1 and 2 (according to our subsequent MD-aided assignment, see below), and the two alternative hydrogen-bond schemes with A and B. The crystallographic axis a points upward along the cellulose chain direction (the unit cell is viewed from slightly below the origin of *a*/*b*/*c*).

possible hydrogen-bond networks (occupation schemes denoted with A and B in Figure 5) of the hydroxyl groups were postulated according to two alternative occupation schemes. These two networks should coexist within the cellulose crystallites, and the authors derived occupation numbers for the alternative positions. From these data we generated two coordinate input files for our COSMOS molecular dynamics simulations, representing the two alternative hydrogen-bond schemes. To simulate the situation of an infinite polymer chain, we added two further glucose rings to each end of the unit cell which by itself contains two glucose moieties.

To solve the question whether the two hydrogen-bond schemes can coexist or whether they interconvert spontaneously, we performed MD simulations for 100 ps, starting with either occupation schemes A and B. The coordinates and atomic charges were recalculated every 0.5 fs. This *NTV* (conversion of the particle number *N*, temperature *T*, and volume *V*) simulation was run at 293 K to create structures near all minima that could be populated at room temperature. A total of 200 coordinate snapshots were stored, and the dihedral angle trajectories of the hydroxyl protons were sampled, too. An interconversion between the two hydrogen-bond networks would either show up as 180° flips in the dihedral angles, or alternatively the protons could jump from one minimum of the hydrogen-bond potential to the other minimum. The latter effect cannot be simulated using regular force fields, since it involves a breaking of the O−H bonds. Both mechanisms must proceed in a highly cooperative manner, hence a statistical distribution over both occupation schemes at the same time can be excluded for energetic reasons (see Table 2). Our MD simulations showed that both hydrogen-bonded structures represent well-defined minima, and an interconversion would have to surmount a very high activation energy barrier. Starting with either occupation scheme A or B, any 180° flips of the hydroxyl groups were not

Table 2. Energy Contributions and Chemical Shift Differences of the Original and the Chemical Shift Refined Cellulose I_α Structures (Lowest Energy Search from 2 × 2 × 200 Structures)

structure and method	hydrogen-bond scheme	NMR assignment of glucose units	total energy (kJ/mol)	van der Waals energy (kJ/mol)	electrostatic energy (kJ/mol)	pseudoenergy (kJ/mol)	RMSD of CS values (ppm)
neutron diffraction	A	I	161.8	40.4	−622.6	2970.4	5.4
		II				2991.0	5.4
CS optimized	A	I	−1111.6	67.0	−1369.1	71.0	0.83
		II	−1249.7	93.8	−1557.6	323.0	1.77
neutron diffraction	B	I	876.4	71.1	65.2	3677.3	6.0
		II				3705.1	6.0
CS optimized	B	I	−1061.7	139.6	−1389.7	389.0	1.93
		II	−1294.6	77.9	−1624.8	702.8	2.61

observed. The dihedral angles nevertheless fluctuated within a wide angular range of about 50°, in a similar manner for both occupation schemes. Figure 6 illustrates the trajectories of the hydroxyl proton on C6 (which forms hydrogen-bonded bridges connecting the glucose chains laterally to form sheets), as well as on C1 and C3 (which are engaged in intrachain hydrogen bonds that stabilize the chain conformation). These observed fluctuations agree with the results by Heiner et al.,[38] although the electrostatic stabilization of our structures is much higher. Note that in our calculations all polarization effects are included in the Coulomb part of the energy, which is not usually taken into account in most other force fields.

Chemical Shift Constrained Geometry Optimizations. The 200 coordinate snapshots that had been sampled during the MD simulations were used for geometry optimizations with ^{13}C isotropic chemical shifts as target functions. First, all structures were purely geometry optimized, and in a second run the chemical shift pseudo-forces were switched on. All these calculations were performed for the central unit cell flanked by two shells of surrounding unit cells by triclinic translations in every step of the optimization.

The assignment of the NMR shifts to the different carbon positions of the two distinct glucose units 1 and 2 had been obtained from the refocused INADEQUATE experiment (see Figure 3). However, this assignment left open the question as to which set of signals belongs to which glucose ring. Therefore, both possible assignments (designated by I and II) were used in the structure refinements. This had to be done for both hydrogen-bond occupation schemes A and B, hence 2 × 2 × 200; i.e., calculations on 800 structures had to be ultimately performed. The geometry optimization for a large subset of structures can be performed with the back-end version of COSMOS-NMR. This program optimizes all structures in a directory with predefined parameters from an interactive run of COSMOS. We performed a search for the lowest energy structure of each of the 2 × 2 = 4 subsets of coordinates, see Table 2.

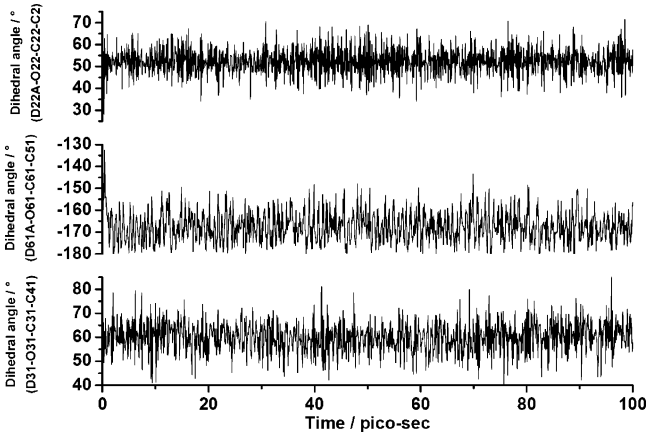

Figure 6. Dihedral angle trajectories for three representative hydroxyl groups (on C2, C3, and C6) forming hydrogen bonds in cellulose I_α. The trajectories are illustrated for occupation scheme A.

We calculated the force field energy contributions and the ^{13}C chemical shifts for the original unit cell (neutron diffraction data, see Table 2) taking into account the two shells of neighboring cells. Comparing the results of the unrefined structures with hydrogen-bond schemes A and B, a negative electrostatic energy is obtained only in the case of A, though the total force field energy is positive in both cases (which is a relative measure of the stability of the lattice).

Subsequently, the neutron diffraction structures A and B were geometry optimized by applying chemical shift pseudo-forces. These pseudo-forces were scaled up until the pseudo-energies reached the level of the electrostatic energy (scaling factor 10, and 3 ppm width of the potential, see Witter et al.[14]). Even though we used ^{13}C chemical shifts as target parameters, this does not mean that the pseudo-forces act only on the carbons. All atoms that contribute to the polarization of a carbon bond will experience the pseudo-forces and are therefore influenced by the geometry optimization.

As outlined above, one question that had to be solved is the assignment of the chemical shifts to the two discrete glucose units 1 and 2, since there is no evidence from the NMR experiments to support either choice. We thus compared the results of the geometry optimization procedure for both assignments to derive criteria for the preference of one or the other. The energies resulting from this chemical shift refinement procedure are given in Table 2. The first remarkable observation is that deep minima for the electrostatic and total energy are obtained for both hydrogen-bond schemes A and B, which makes a spontaneous interconversion of the two forms very unlikely at room temperature. Taking only the total energy into account, occupation scheme B would seem to be the more favorable one. In most cases, however, chemical shift pseudo-forces tend to drive the structures energetically uphill, hence we took the sum of the total and pseudo-energy as the main criterion for the most preferable structure. In this case, we have to select hydrogen-bond scheme A and assignment I ("A−I" as the most favorable structure). This structure is moreover in the 6th position (of all 800 optimized MD conformations) with regard to the total energy, and it has the lowest rms deviation between calculated and experimental chemical shifts (see Table 2).

In all cases, the significant drop in total energy upon applying chemical shift pseudo-forces is a clear indication that the calculated chemical shifts are of high quality. After geometry optimization we reach minima with small pseudo-energies and with chemical shift values that lie within the error-range of our experimental NMR data. In the case of the A−I structure, the pseudo-energy of 71 kJ/mol is only about 5% of the electrostatic energy. Our results are therefore dominated by the fundamental electrostatic interactions within the cellulose lattice, and the chemical shift differences as such do not translate into large

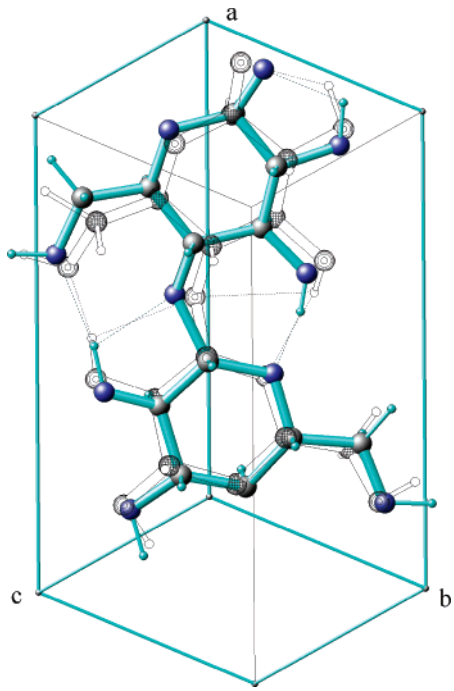

Figure 7. Least-squares superposition of our chemical shift optimized structure A−I (occupation scheme A, chemical shift assignment I, see Table 2) with the original neutron diffraction structure of Nishiyama et al.[13] (shown as the transparent model). The rms difference for all atoms is 0.57 Å, and 0.37 Å for heavy atoms alone. The unit cell is viewed from slightly below the origin of *a*/*b*/*c*).

Table 3. Experimental ^{13}C Chemical Shifts and Principal CS Tensor Components (See Table 1), Compared to the Calculated Values for the Chemical Shift Optimized Structure A−I

	experimental CSA parameters				calculated CSA parameters			
site	δ_{iso} (ppm)	δ_{11} (ppm)	δ_{22} (ppm)	δ_{33} (ppm)	δ_{iso} (ppm)	δ_{11} (ppm)	δ_{22} (ppm)	δ_{33} (ppm)
Glucose Unit 1								
$C1_1$	105.1	90.3	105.0	120.0	103.4	74.2	105.7	130.5
$C2_1$	71.3	53.2	72.2	88.5	72.1	48.8	76.4	91.0
$C3_1$	74.7	57.5	75.3	91.3	74.6	51.3	79.3	93.0
$C4_1$	89.2	65.3	93.1	109.2	87.4	59.3	97.8	105.3
$C5_1$	72.9	53.1	74.5	91.1	73.4	47.6	80.4	92.3
$C6_1$	65.4	35.4	73.2	87.6	65.3	34.2	74.1	87.5
Glucose Unit 2								
$C1_2$	105.7	90.9	105.6	120.6	105.3	77.7	108.6	129.7
$C2_2$	72.5	53.6	74.2	89.7	73.1	47.7	78.0	93.6
$C3_2$	75.3	58.6	75.8	91.4	75.3	53.0	80.4	92.6
$C4_2$	90.2	65.9	94.1	110.6	89.6	57.2	101.8	109.9
$C5_2$	71.4	53.3	72.3	88.6	71.8	46.5	75.2	93.5
$C6_2$	65.7	35.7	73.5	87.9	65.7	35.0	74.5	87.5

structural differences. To test the reliability of the structures, a least-squares superposition of our chemical shift optimized structure A−I and the original neutron diffraction coordinates was performed. The rms difference of the two cellulose chain fragments is 0.57 Å for all atoms, and this difference drops to only 0.37 Å if only heavy atoms are superimposed (see Figure 7).

Most atom positions of our chemical shift refined structure are within the intrinsic error-range of the fiber diffraction analysis, but the NMR structure is clearly different in several aspects; see Figure 7. These differences may be a consequence of the refinement procedure of the diffraction data, or they might be attributed to the different sources and/or treatments of the bacterial cellulose samples. As result of our MD simulations and energy calculations we can clearly reject a statistical distribution or a rapid interconversion of the hydroxyl protons between the two postulated hydrogen-bond schemes. A conversion from structure A to B can only be possible if a cooperative switch of all intra- and intermolecular hydrogen bonds happens at the same time. Such conversion is highly unlikely at 293 K. Therefore, a coexistence of the two occupation schemes within the sample remains the only description. In the NMR spectra of bacterial cellulose, however, such distinct domains should give rise to spectral splittings, which were not observed in high-resolution MAS spectra of unlabeled bacterial cellulose, which show less broadening than labeled cellulose (see Sternberg et al.[1]). Therefore, the coexistence of domains with different hydrogen-bond networks that was suggested from neutron diffraction, must be attributed to the different source of this material (*Glaucocystis nostochinearum*) or to the special treatment that had been used to exchange the hydrogen-bonded protons with deuterons.

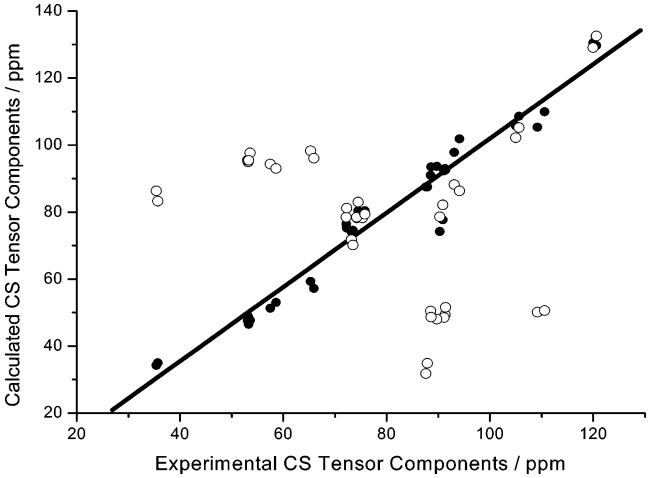

Figure 8. Calculated principal ^{13}C chemical shift tensor components (δ_{11}, δ_{22}, δ_{33}), plotted against the experimental values. The values after optimization (best A−I) with isotropic chemical shift-pseudo forces are displayed as filled circles (●), while the result evaluated from the original (nonoptimized) diffraction structure are open circles (○).

Table 4. Orientations of the δ_{33} Tensor Components

site	deviation of the δ_{33} tensor component from the C−O bond direction (deg)
$C1_1$	26.8
$C2_1$	7.0
$C3_1$	6.3
$C4_1$	4.4
$C5_1$	6.0
$C6_1$	7.9
$C1_2$	18.5
$C2_2$	6.9
$C3_2$	4.2
$C4_2$	2.8
$C5_2$	8.2
$C6_2$	6.2

Comparing our new coordinates to the earlier NMR refined cellulose structure (Sternberg et al.[1]), the superposition of all atoms gives rms differences of 0.47 and 0.40 Å for the heavy atoms, thus preserving all general features of the former structural model. The slight differences can be attributed to the smaller unit cell of our new A−I model and some differences in the NMR data. The newly refined cellulose I_α structure with hydrogen-bond scheme A therefore satisfies the features of both the diffraction experiments as well as the earlier NMR data.

^{13}Chemical Shift Tensors. Another strong indication for the validity of our new structure is the fact that the calculated chemical shift tensors fit much better to the experimental data, once we had performed the isotropic chemical shift refinement. The full ^{13}C chemical shift tensor information was calculated

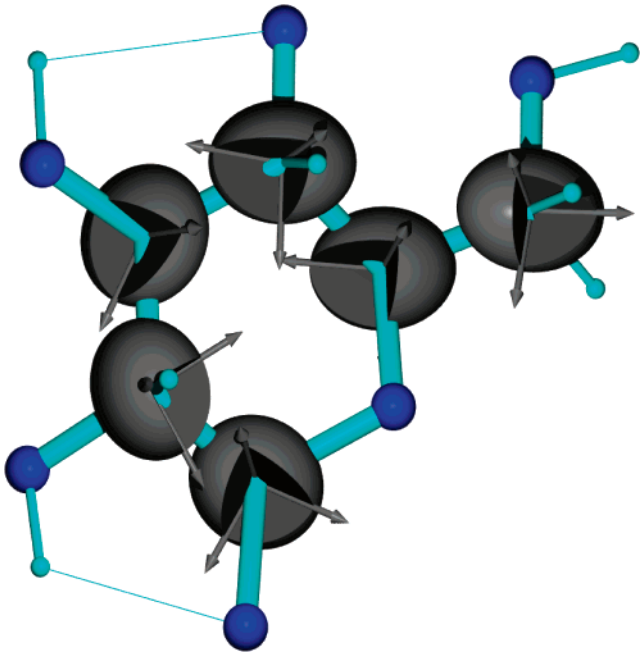

Figure 9. Orientations of the ^{13}C chemical shift tensors in a glucose ring calculated from our refined cellulose structure. The δ_{33} principal components of the tensors are aligned nearly parallel with the C−O bond directions. In the case of carbon C1, which has two oxygen atoms as neighbors, the δ_{22} component lies along the bisector of the O−C−O angle.

by BPT and is compared to the experimental data in Table 3. With the exception of δ_{11} for carbons $C1_1$ and $C1_2$, the differences between the calculated and experimental components are 5 ppm. As can be seen from Figure 8, the overall correlation is rather good (filled circles). Notably, the chemical shift components derived from the original diffraction coordinates gave no correlation at all (open circles). This must be regarded as strong evidence for the validity of the new NMR-refined structure. The remaining deviation between theoretical and experimental values of the $C1_1$ and $C1_2$ carbon sites may arise from the anomeric effect of the two adjacent ether oxygen atoms, which contain lone electron pairs that are not separately considered in the bond polarization treatment. Experimentally, some deviation may come from the very small isotropic chemical shift difference between $C1_1$ and $C1_2$, which makes a separation of the CSA patterns in the RAI experiment complicated.

From the chemical shift anisotropy calculations it is also possible to obtain the orientation of the CS tensors, which is not straightforward experimentally, hence the characteristic rules for carbohydrates can be tested. As discussed by Koch et al.,[39] for carbon atoms carrying a hydroxyl group the δ_{33} tensor component should be pointing toward the neighboring oxygen atom.

(see Table 4). As seen in Figure 9, the δ_{33} directions deviate by only a few degrees from the C−O bond direction. In the case of C1, which is bound to two oxygen atoms, the δ_{33} component lies in the C1−O1−C5 plane, and the δ_{22} direction is aligned with the bisector of the angle formed by the three atoms.

General Features of the Cellulose I_α Structure. In all structure calculations, the only negative energy contribution comes from electrostatic interactions including polarization. This means that the cellulose lattice is held together by Coulomb forces. As the COSMOS-NMR force field does not contain any explicit hydrogen-bond terms, these Coulomb interactions must

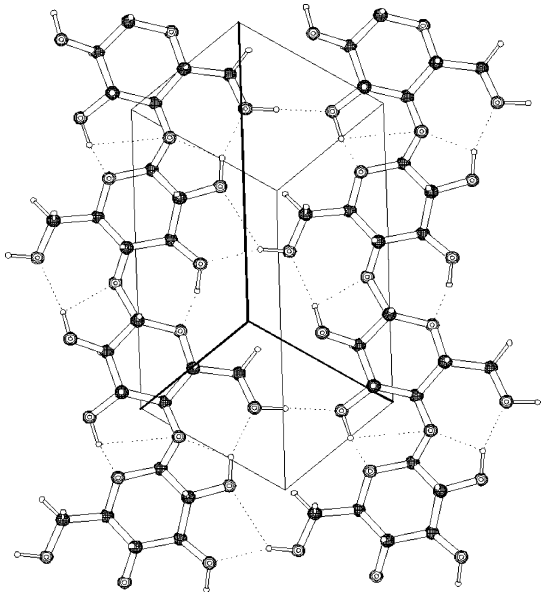

Figure 10. Two glucose chains in the cellulose I_α structure obtained by our geometry optimization with chemical shift constraints. The chains consist of 1−4 linked β-D-glucose residues, which form sheets that are connected by hydrogen bonds of the C6 hydroxyl groups.

be attributed to the hydrogen bonds as well as interchain and intersheet interactions. Figure 10 illustrates how two cellulose chains are connected by hydrogen bonds to form sheets. These sheets do not form hydrogen bonds to adjacent sheets but interact with them electrostatically. The hydrogen-bond network with bifurcated intrachain hydrogen bonds differs in some aspects from the original neutron diffraction result. van der Waals interactions do not have any stabilizing effect on the lattice in our calculations. It should be mentioned that earlier molecular mechanics simulations of cellulose I_α and I_β using the GROMOS force field had failed to produce negative total energies (Heiner et al.[37]). Because of the neglect of polarization, the stabilizing effect of the electrostatics seems to have been underestimated in their calculations.

We finally note that the intrachain hydrogen bonds are responsible for the special geometry of the glycosidic linkage, which leads to the characteristic chemical shift of $C4_1$ and $C4_2$ near 90 ppm that is only observed in crystalline cellulose polymorphs. In the amorphous state or in cellulose solutions this value moves upfield by about 10 ppm (Sternberg et al.[1]).

Acknowledgment. We thank the FZK and the DFG-CFN for the NMR infrastructure.

Supporting Information Available: Cellulose I_α structure and NMR data, in pdb formats. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Sternberg, U.; Koch, F.-Th.; Priess, W.; Witter, R. *Cellulose* **2003**, *10*, 189.
(2) VanderHart, D. L.; Atalla, R. H. *Macromolecules* **1984**, *17*, 1465.
(3) Erata, T.; Shikano, T.; Yunoki, S.; Takai, M. *Cellulose Commun.* **1997**, *4*, 128.
(4) Kono, H.; Yunoki, S.; Shikano, T.; Fujiwara, M.; Erata, T.; Takai, M. *J. Am. Chem. Soc.* **2002**, *124*, 7506.
(5) Lesage, A.; Bardet, M.; Emsley, L. *J. Am. Chem. Soc.* **1999**, *121*, 10987.
(6) Kono, H.; Erata, T.; Takai, M. *Macromolecules* **2003**, *36*, 5131.
(7) Jäger, C.; Pauli, J.; Schmauder, H.-P. Poster: "Observation of the Glycosidic Bond and Complete Ring Assignment of the ^{13}C−Signals of Bacterial Cellulose. Presented at Magnetic Resonance Spectroscopy Division of the German Chemical Society, Leipzig, Germany, 2003.

(8) Cadars, S.; Lesage, A.,; Emsley, L. *J. Am. Chem. Soc.* **2005**, *127*, 4466.
(9) Sakellariou, D.; Brown, S. P.; Lesage, A.; Hediger, S.; Bardet, M.; Meriles, C. A.; Pines, A.; Emsley, L. *J. Am. Chem. Soc.* **2003**, *125*, 4376.
(10) Reiling, S.; Brickmann, J. *Macromol. Theory Simul.* **1995**, *4*, 725.
(11) Sugiyama, J.; Persson, J.; Chanzy, H. *Macromolecules* **1991**, *24*, 2461.
(12) Sugiyama, J.; Voung, R.; Chanzy, H. *Macromolecules* **1991**, *24*, 4168.
(13) Nishiyama, Y.; Sugiyama, J.; Chanzy, H.; Langan, P. *J. Am. Chem. Soc.* **2003**, *125*, 143.
(14) Witter, R.; Priess, W.; Sternberg, U. *J. Comput. Chem* **2002**, *23*, 298.
(15) Witter, R.; Seyfart, L.; Greiner, G.; Reissmann, S.; Weston, J.; Sternberg, U. *J. Biomol. NMR* **2002**, *24*, 277.
(16) Witter, R. Structure Elucidation with the COSMOS-NMR Force Field. Dissertation, FSU Jena, 2003; www.dissertation.de, ISBN:38982565588.
(17) Brown, R. M. Jr.; Willison, J. H. M.; Richardson, C. L. *Proc. Natl. Acad. Sci. U.S.A.* **1976**, *73*, 4565.
(18) Kondo, T.; Nojiri, M.; Hishikawa, Y.; Togawa, E.; Romanovicz, D.; Brown, R. M. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99/22*, 14008.
(19) Hesse, St.; Kondo, T. *Carbohydr. Polym.* **2005**, 60/4, 457.
(20) Schramm, M.; Hestrin, S. *J. Gen. Microbiol.* **1954**, *11*, 123.
(21) Bennett, A. E.; Rienstra, C. M.; Auger, M.; Lakshmi, K. V.; Griffin, R. G. *J. Chem. Phys.* **1995**, *103*, 6951.
(22) States, D. J.; Haberkorn, R. A.; Ruben, D. J. *J. Magn. Reson.* **1982**, *48*, 286.
(23) Witter, R.; Hesse, St.; Sternberg, U. *J. Magn. Reson.* **2003**, *161*, 35.
(24) Witter, R.; Sternberg, U.; Ulrich, A. S. *J. Am. Chem. Soc.* **2006**, *128*, 2236.
(25) Ishii, Y.; Ashida, J.; Terao, T. *Chem. Phys. Lett.* **1995**, *246*, 439.
(26) Massiot, D.; Fayon, F.; Capron, M.; King, I.; Calve, S. Le; Alonso, B.; Durand, J. O.; Bujoli, B.; Gan, Z.; Hoatson, G. *Magn. Reson. Chem.* **2002**, *40*, 70.
(27) Haeberlen, U. In *Advances in Magnetic Resonance, Suppl. 1*; Waugh, J. S., Ed.; Academic Press: New York, 1993. Mehring, M. *Principles of High-Resolution NMR in Solids*, 2nd. ed.; Springer-Verlag: Berlin, 1983.
(28) Möllhoff, M.; Sternberg, U. *J. Mol. Mod.* **2001**, *7*, 90.
(29) Sternberg, U.; Koch, F.-T.; Losso, P. COSMOS, www.cosmos-software.de. The COSMOS-NMR force field (back-end version) can be obtained from the authors.
(30) Sternberg, U. *J. Mol. Phys.* **1988**, *63*, 249.
(31) Sternberg, U.; Priess, W. *J. Magn. Reson.* **1997**, *125*, 8.
(32) Slater, J. C. *Phys. Rev.* **1930**, *36*, 57.
(33) Veeman, W. S. *Prog. NMR Spectrosc.* **1984**, *20*, 193.
(34) Sherwood, M. H.; Alderman, D. W.; Grant, M. G. *J. Magn. Reson.* **1989**, *84*, 466.
(35) O'Keefe, M.; Brese, N. E. *J. Am. Chem. Soc.* **1991**, *113*, 3226.
(36) Koch, F.-T.; Möllhoff, M.; Sternberg, U. *J. Comput. Chem.* **1994**, *15*, 524.
(37) Sternberg, U.; Witter, R.; Ulrich, A. S. *Annu. Rep. NMR Spectrosc.* **2004**, *52*, 53.
(38) Heiner, A. P.; Sugiyama, J.; Teleman, O. *Carbohydr. Res.* **1995**, *273*, 207.
(39) Koch, F.-Th.; Priess, W.; Witter, R.; Sternberg, U. *Macromol. Chem. Phys.* **2000**, *201*, 1930.
(40) Witter, R.; Möllhoff, M.; Sternberg, U. Manuscript in preparation.

MA052439N