2009

Investigating the Structural Characteristics of Cellobiose-Water Solutions *via* Liquid-State Neutron Diffraction

Brad O'Dell

FORM C
COLLEGE SCHOLARS PROJECT APPROVAL

William Bradley O'Dell
Scholar

Fred Schell, David Baker
Mentor

Directly Observing Structural Characteristics of Cellobiose - Water Solutions via Liquid State Neutron Diffraction
Project Title

COMMITTEE MEMBERS
(Minimum 3 Required)

Name	Signature
DAVID C. BAKER	David C. Baker
FRED M. SCHELL	
T. FFRANCON WILLIAMS	T. Ffrancon Williams
SYLVIA E. MCLAIN	

PLEASE ATTACH A COPY OF THE SENIOR PROJECT TO THIS SHEET AND RETURN BOTH TO THE PROGRAM DIRECTOR. THIS PAGE SHOULD BE DATED AND COMPLETED ON THE DATE THAT YOUR DEFENSE IS HELD.

DATE COMPLETED 4/3/09

Investigating the Structural Characteristics of Cellobiose–Water Solutions *via* Liquid-State Neutron Diffraction

A Thesis

Presented for the

Bachelor of Arts

Degree

College Scholars Program

The University of Tennessee, Knoxville

Wm. Brad O'Dell

Fred M. Schell

David C. Baker

T. F. Williams

Sylvia E. McLain

April 2009

ABSTRACT:

Cellulosic recalcitrance to aqueous solvation presents considerable challenges to chemical processes based upon cellulose chemistry such as bioethanol production. Unlike higher cello-oligosaccharides, cellobiose with two glucose subunits exhibits considerable solubility in water. Neutron diffraction with isotopic substitution (NDIS) experiments and empirical pair structure refinement (EPSR) simulations have been performed on aqueous cellobiose solutions to determine the fundamental hydrogen-bonding characteristics of cello-saccharides in water. Results to be presented indicate significant changes to both the bulk water structure and intramolecular cellobiose hydrogen bonding.

Table of Contents

List of Figures

List of Tables

Abbreviations and Acronyms

ATLAS	Analysis of Time-of-Flight Diffraction Data from Liquid and Amorphous Samples
CN	coordination number
DP	degree of polymerization
DSC	differential scattering cross-section
EP	empirical potential
EPSR	Empirical Potential Structure Refinement
MD	molecular dynamics
NDIS	neutron diffraction with isotopic substitution
NMR	nuclear magnetic resonance spectroscopy
NOE	nuclear Overhauser effect
RDF	radial distribution function
SANDALS	Small Angle Neutron Diffractometer for Liquids and Amorphous Samples

1. Introduction

1.1 Cellulose: A Potentially Viable Bioethanol Feedstock

The term "cellulose" first appeared in the chemical literature when Payen[1] coined the term to describe a fraction of plant matter recalcitrant to digestion by nitric acid and subsequent treatment with sodium hydroxide. He established the chemical structure of this fraction to be identical to that of starch, $C_6H_{10}O_5$, and concluded that cellulose was an "isomeric" form. This discovery conceptually linked the previous knowledge that both starch[2] and wood[3, 4] produce D-glucose upon treatment with acid, uniting the chemistries of soluble and insoluble naturally occurring carbohydrates.

While the use of cellulosic biomass as an energy source is an ages-old concept in the form of burning plant matter to generate heat, the production of refined fuels from cellulose began en force in the latter half of the nineteenth century. By 1863, B. C. Tilgham had developed a patented process for cellulose extraction from plant matter and opened a mill in 1866 for this process. The main intention of this process was the production of bulk pulp; however, work by Mitscherlich in 1878 promoted use of the glucose hydrolyzed from cellulose in the Tilgham process as a fermentation feedstock for ethanol.[5] In the Twentieth Century, the growing dependence of transportation on refined petroleum and concerns about scarcity and energy security, arising primarily from World War II and the Cold War, prompted large commitments to the conversion of cellulose to ethanol as a primary process with notable successes occurring in the United States,[6] the then Soviet Union[7] and Germany.[8]

Processes for the conversion of cellulose to ethanol have a long and rich chemical history. However, the historically lower costs of refined petroleum fuels have led to their continued dominance of the global energy market. Though, as current markets see both

increasing demands for and decreasing supplies of petroleum, science and industry are revisiting cellulosic ethanol as a petrol-fuel alternative.

1.1.1 Lignocellulosic Biomass Conversion to Ethanol

Early processes for ethanol production from plant biomass, such as those mentioned previously, focused upon hydrolysis and fermentation of only cellulose purified from the plant feedstock. In contrast, current approaches are being designed to use hemicellulose and lignin, biological polymers occurring in coincidence with cellulose, along with the purified cellulose fraction.[5] Lignocellulosic biomass significantly increases the useable feedstock for ethanol and other chemical products in comparison with purified cellulose. However, processing of lignin, hemicellulose, and cellulose concurrently presents significant additional challenges, the most significant of which is the recovery of each material at its greatest feedstock potential without loss or inhibitive contamination from the other fractions.[9]

A brief overview of the lignocelluloic biomass to ethanol conversion process as detailed by Lee[9] and Kamm et al.[5] is presented here. Raw plant matter or other sources of cellulose, such as solid municipal waste, is initially processed mechanically for efficient storage, transport and chemical treatment. Next, the feedstock undergoes some form of "pre-treatment" intended to increase the surface area of the matter and to disrupt the natural aggregation of the principal biopolymers. Pre-treatment techniques vary from purely physical approaches such as ball milling and steam explosion to chemical methods including acid hydrolysis of the cellulose and hemicellulose components and solvation of the hydrophobic lignin component with organic solvents; combinations of physical and chemical pre-treatments are also common. Once the material has undergone pre-treatment and has been returned to biologically amenable conditions

(i.e., water solvent at neutral pH), the carbohydrate components are typically exposed to cellulolytic enzymes and microbial yeast for fermentation to ethanol while, the lignin is separated for use in either producing other chemical products or serving as a combustible fuel for powering the overall process.

The high yields of ethanol from enzymatic hydrolysis and microbial fermentation have led current research interests to focus on optimization of pre-treatment processes for increasing efficiency and economic viability of lignocellulosic ethanol.[9] This interest has intensified research in not only process engineering, but also in fundamental cellulose science where better understanding of the structure and structural interactions of cellulose in pure solid, biomass, and solution forms holds great for developing efficient and economical means of pre-treating plant biomass.

1.1.2 Cellulose Structure

Investigation of the structure of cellulose began during the infancy of modern structural analysis techniques. In 1926, just thirteen years after the first reports of crystal structure analysis by x-ray diffraction,[10] Sponsler and Dore reported a diffractometric analysis of crystalline cellulose[11] derived from ramie fibers. This work proved seminal to both the fields of carbohydrate chemistry and macromolecular structure analysis. With respect to carbohydrates, the structure determined by crystallography established the chair conformation of the glucopyranose ring and introduced the concept of the glycosidic linkage in structural terms. Furthermore, the report of a crystal structure in which a molecular chain extended beyond the unit cell, which was found to contain two glucopyranose rings, revolutionized ideas about the structure of both crystals and polymeric materials.[12]

Though having an early start in the science of materials structure analysis, cellulose has retained significant structural uncertainty over the last eighty years. In the original model structure, Sponsler and Dore included both 1→4 and 4→4 glycosidic linkages.[11] This inconsistency was removed in the model of Meyer and Mark[13] and the later revision by Meyer and Misch,[14] which contained only 1→4 linkages. A representative chemical model of this structure of cellulose is given as Figure 1.1. However, the Meyer models simultaneously raised new questions regarding alignment of adjacent cellulose chains in either parallel or anti-parallel directions and the extents of intra- and inter-chain hydrogen bonding in the solid state.

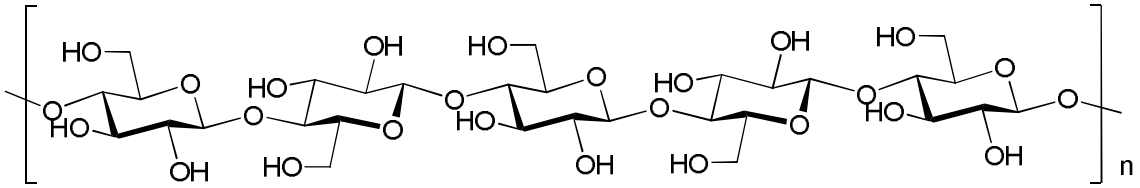

Figure 1.1: Representative cellulose chain fragment.

Attempts to answer the questions raised by both the Meyer model and other properties of solid cellulose, such as the nature of amorphous phases, have driven the last eight decades of cellulose structure research.[15, 16] The progressive discovery of consistent structural differences among celluloses in native form and after various chemical treatments led to the classification of six cellulose polymorphs. Native cellulose from plant cell walls is defined as cellulose I with cellulose II, III_I, III_{II}, IV_I and IV_{II} the results of chemical interconversion. In 1984, the known structural complexity was expanded by Atalla and VanderHart[17] who elucidated the structures of two naturally occurring polymorphic forms of native cellulose, cellulose I_α and cellulose I_β. Figure 1.2 shows the common polymorphic conversion processes.

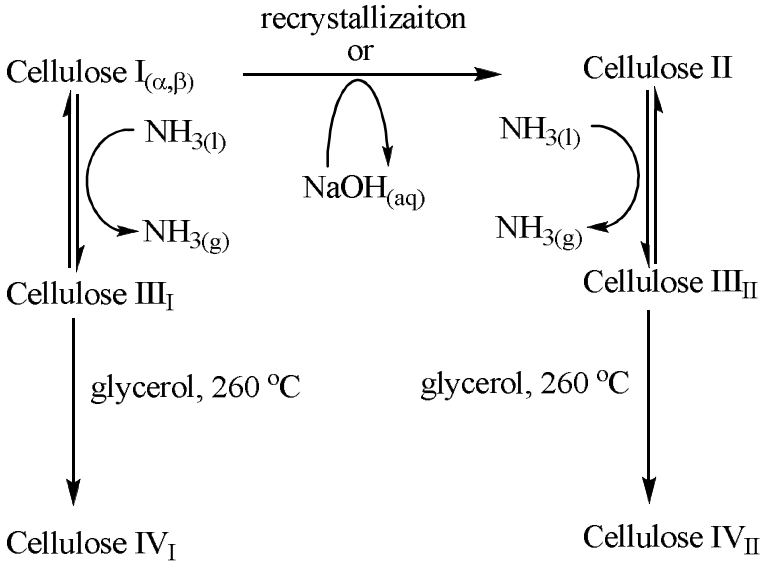

Figure 1.2: Polymorphs of cellulose. After O'Sullivan.[15]

1.1.3 Methods for Investigating Cellulosic Structure in Solution

While the structures of cellulosic materials have been investigated extensively in the solid state, relatively little is known about cellulose structure in solution. Studies on solution structure are largely hindered by the near-insolubility of cello-oligomers with degree of polymerization (DP) greater than seven in common solvent systems. As such, extensive spectroscopic characterization *via* NMR spectroscopy has only been completed with cello-oligomers in this DP range.[18] A small number of NMR studies of high-DP cellulose have been conducted in derivatizing[19] and ionic liquid[20-22] solvent systems. However structural information from these studies is limited, and any cellulose structure in such chemical environments cannot be considered representative of the structure in the conditions of biological systems or industrial processes of interest.

Diffraction studies of cellulose in solution are similarly limited. Small angle neutron and x-ray scattering experiments have revealed the bulk morphology of cellulose fibers with varying

degrees of hydration[23, 24], in aqueous suspension[25] and in gels formed from LiCl/*N*,*N*-dimethylacetamide solutions[26] though structural information for cellulose on the atomic length scale (1-10 Å) has yet to be attained. Recent solution state neutron diffraction studies of glucose[27-29] and trehalose[30] in aqueous solution have proven the viability of studying carbohydrates with the method. This prior work, along with the detailed structural information shown to be obtainable from other biological molecules by use of neutron diffraction with isotopic substitution coupled with computational simulation[31,32] inspired the investigations herein detailed.

1.2 Neutron Diffraction: A Powerful Tool for Investigating Hydrogenic Materials

Neutron diffraction is the principal structural technique employed in this investigation largely due to the importance of hydrogen bonding in the material of interest. The seminal work of Shull[33] established neutron diffraction as a useful probe for structural investigation of hydrogen-containing materials. This usefulness arises from the nature of the scattering interaction between an incident neutron and an objective atom. Unlike x-rays and electrons, which are scattered by the electrons surrounding an atomic nucleus, neutrons are scattered by nuclei themselves. This interaction, mediated by nuclear forces, does not produce a dependence of scattering intensity on atomic number and gives a neutron scattering intensity, or scattering length, for hydrogen that is comparable to other nuclei of interest in many systems[34] including organic and biological molecules. Both the nonsystematic variation of neutron scattering length with increasing atomic number and the relatively large neutron scattering length of hydrogen are demonstrated by Figure 1.3.

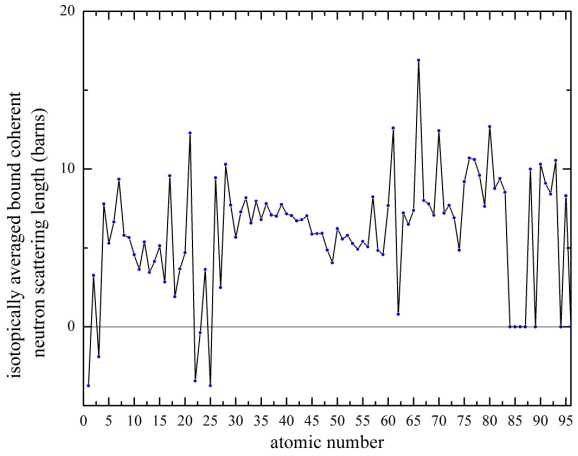

Figure 1.3: Neutron scattering length versus atomic number.

1.2.1 Fundamental Theory of Neutron Diffraction

Since the experimental details and results to be presented depend heavily upon neutorn diffraction techniques, a brief overview of relevant diffraction theory is presented. This discussion is drawn primarily from the theoretical presentations of Egelstaff,[35] Chieux[36] and Soper[37] with specific theoretical contributions referenced where appropriate.

Diffraction serves as a direct experimental probe of the structure of a sample material. The structure factor, *F(Q)*, for the case of neutron diffraction can be defined as

$$F(Q) = \sum_{ab}^{N} c_a c_b b_a b_b \left(2 - \delta_{ab}\right)\left[S_{a,b}(Q) - 1\right] \quad (1.1)$$

where $c_{a,b}$ and $b_{a.b}$ are the atomic concentration and neutron scattering length of atoms *a and b*, δ_{ab} is the Kronecker delta introduced to prevent double counting of atoms, and $S_{a,b}(Q)$ is the partial structure factor arising from the spatial distribution of atoms *b* about atoms *a* with *N*, the number of such partial structure factors, given by $N=M(M+1)/2$ where *M* is the number of distinct atom types in the system. The quantity *Q* is the magnitude of the momentum transfer vector of the scattered neutrons in reciprocal space and is related to the diffraction angle 2θ and incident neutron wavelength λ as shown in Eq (1.2). For cases of purely elastic scatter, **K' = K**

in magnitude; however, if energy transfer between the incident neutrons and the sample, or inelasticity, occurs, then **K'** varies in both magnitude and direction. Figure 1.4 also graphically defines *Q*.

$$Q = \frac{4\pi \sin\theta}{\lambda} \quad (1.2)$$

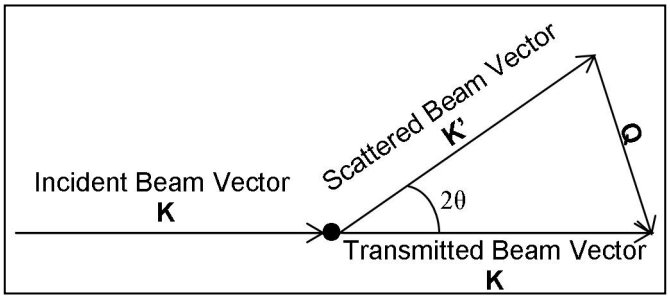

Figure 1.4: Graphical definition of *Q*.

The partial structure factor $S_{ab}(Q)$ is related to the radial distribution function (RDF) of atoms *b* about atoms *a*, $g_{a,b}(r)$ via Fourier transformation from reciprocal space to real space, namely,

$$S_{a,b}(Q) = 1 + \frac{4\pi\rho}{Q}\int_0^\infty \left(g_{a,b}(r) - 1\right)\sin(Qr)\,r\,dr \quad (1.3)$$

where *r* represents the distance separating atoms *b* and *a*. Since the RDF represents the probability of finding an atom *b* at distance *r* from atom *a*, local maxima in $g_{a,b}(r)$ represent average bond lengths between atoms in the system. Furthermore, since the total structure factor is measured on an absolute scale *via* normalization, integration of $g_{a,b}(r)$, such as

$$n_a^b(r) = 4\pi c_b \rho \int_{r_1}^{r_2} g_{a,b}(r)\, r^2 dr \quad (1.4)$$

yields coordination numbers, or the average number of atoms of type *b* found between distances r_1 *and* r_2 from atoms of type *a*.

While the goal of neutron diffraction experiments is clearly to obtain *F(Q)*, measuring the neutrons scattered by the sample in actuality yields the double differential cross section,

$$\frac{d^2\sigma}{d\Omega dE}(\lambda, 2\theta) = \left(\frac{\text{number of neutrons scattered into the solid angle } d\Omega \text{ with energy } dE \text{ at angle } 2\theta}{N\Phi(\lambda) d\Omega dE} \right) \quad (1.5)$$

where N is the number of atoms in the sample, $\Phi(\lambda)$ is the neutron flux, or number of neutrons passing through the sample per second, as a function of neutron wavelength λ. The solid angle $d\Omega$, in the above, represents an area of the detector illuminated by the scattered neutrons as demonstrated in Figure 1.5, while the energy dE represents a change in the incident neutron energy due to interaction with the sample.

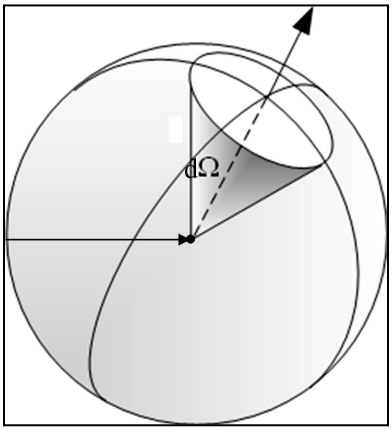

Figure 1.5: The solid angle *dΩ*.

For diffraction, only structural information is of interest; therefore, the dynamic information contained in the energy differences of the detected neutrons is effectively discarded by integrating Eq (1.5) over all energies, yielding

$$\frac{d\sigma}{d\Omega}(\lambda, 2\theta) = \left(\frac{\text{number of neutrons scattered into the solid angle } d\Omega \text{ at angle } 2\theta}{N\Phi(\lambda) d\Omega} \right) \quad (1.6).$$

The total DSC defined in Equation (1.6) contains two types of scattering by the sample, coherent scattering and incoherent scattering. Coherent scattering arises from interference of neutrons scattered by the spatial distribution of atoms in the sample while incoherently scattered neutrons do not interfere but are merely scattered from individual atoms. As such, the DSC is separated into its two components, as in Eq (1.7) with only the coherent scattering useful for structural analysis.

$$\frac{d\sigma}{d\Omega}_{total} = \frac{d\sigma}{d\Omega}_{coherent} + \frac{d\sigma}{d\Omega}_{incoherent} \quad (1.7)$$

At this point, the coherent DSC should be equated to the concentration, scattering lengths, and spatial arrangements of the atoms in the sample.

$$\frac{d\sigma}{d\Omega}(\lambda, 2\theta)_{coherent} = \sum_a c_a b_a^2 + \sum_{a,b} c_a c_b b_a b_b \left(2 - \delta_{ab}\right)\left(S_{a,b}(Q) - 1\right) \quad (1.8)$$

However, the integration over all neutron energy transfer performed in deriving the total DSC essentially assumes that all neutrons are scattered elastically, or without energy transfer to or from the sample. This assumption does not hold for most samples of interest, particularly liquids and samples containing light atoms, such as hydrogen, due to low energies of motion and invalidates Eq (1.8). Therefore, corrections must be introduced into both terms of the coherent DSC to account for inelasticity. Placzek[38] originally derived an approximate model of the dynamics of systems containing heavy atoms that can be subtracted from the coherent DSC to give only intensity from elastic scattering. A correction of the type $P(Q,\theta)$ is introduced to give

$$\frac{d\sigma}{d\Omega}(\lambda, 2\theta)_{coherent} = \left[\sum_a c_a b_a^2 + P(Q,\theta)\right] + \left[\sum_{a,b} c_a c_b b_a b_b \left(2 - \delta_{ab}\right)\left(S_{a,b}(Q) - 1\right) + P(Q,\theta)\right] \quad (1.9).$$

The first term results from scattering by atoms of a single type, or self scattering, while the second term results from scattering by atoms distributed about each other and is termed "distinct scattering". In fact, rearranging Eq (1.9) gives the familiar *F(Q)*

$$F(Q) = \sum_{a,b} c_a c_b b_a b_b \left(2 - \delta_{ab}\right)\left(S_{a,b}(Q) - 1\right) = \frac{d\sigma}{d\Omega}(\lambda, 2\theta)_{coherent} - \left[\sum_a c_a b_a^2 + P(Q,\theta)\right] - \left[P(Q,\theta)\right]$$

(1.10)

from which structural information about the sample can be derived by means of the partial structure factors and the resulting RDFs.

1.2.2 Neutron Diffraction with Isotopic Substitution (NDIS)

Eq (1.10) provides a direct theoretical approach to extracting structural information in the form of RDFs from a neutron diffraction measurement. In samples which contain a small number of distinct atom types, extraction of the partial RDFs can generally be attained by measuring chemically identical samples with different isotopic compositions. As discussed previously, neutrons scatter by interaction with the atomic nuclei of the objective sample. This allows neutrons to scatter differently from nuclei of different isotopes, and such contrast can be chemically incorporated into the sample of interest. Table 1.1 shows the bound coherent scattering lengths of different isotopes[34] commonly found in biological or organic molecules. It is of particular importance that the scattering lengths of hydrogen and deuterium vary significantly in magnitude and also vary in sign (a convention used to indicate a phase shift of 180 degrees in the scattering wave) since H/D substitution can often be introduced easily to organic and biological molecules.

Table 1.1: Bound coherent scattering lengths of nuclei common to organic and biological molecules.

Element	Isotope	Scattering Length (femtometers)
Hydrogen	^{1}H	-3.7406
	^{2}H	6.671
Carbon	^{12}C	6.6511
	^{13}C	6.19
Nitrogen	^{14}N	9.37
	^{15}N	6.44
Oxygen	^{16}O	5.803
	^{17}O	5.78
	^{18}O	5.84

Once a series of isotopically labeled samples equivalent to the number of unique atom pairs have been measured and the individual data sets corrected, differences between the total structure factors can be taken to isolate contributions from individual atom pairs. The small number of required isotopic contrasts in systems such as H_2O [39] and HF [40] have allowed isolation of all individual partial structure factors, and associated RDFs. However, systems of greater chemical complexity generally require a number of isotopic contrasts that either cannot be attained chemically or that could not be measured practically during a typical grant of instrument usage time at a spallation or reactor source. Therefore, methods alternative to first-order difference NDIS are commonly employed for samples of present interest.

1.2.3 Empirical Potential Structure Refinement

Computational modeling of liquid-state neutron diffraction samples has become a useful resource in overcoming limitations of the first-order difference method discussed above[4-7].

atomic arrangement of the model. EPSR is a computational method developed for deriving structural models of liquids and amorphous glasses that are consistent with experimental diffraction data. Other structural refinement techniques common in interpretation of diffraction patterns, such as the Reverse Monte Carlo[41] and Rietveld[42] methods, attempt to minimize the difference between the diffraction pattern or structure factor measured experimentally and that derived from simulation. In contrast, EPSR implements an "empirical" energy potential to account for the structural differences between the computational model and the measured sample. Refinement proceeds by minimizing the empirical potential (EP) by the Monte Carlo approach. Atoms in the model are moved at random; all moves that reduce the EP are accepted while those that increase the EP are accepted or rejected based on Boltzmann statistics. As such, EPSR represents a Monte Carlo simulation constrained to reproduce molecular structure by experimental diffraction.[43]

An EPSR simulation begins with the creation of a model distribution of atoms or modeling box. This box is composed of molecules, sets of atoms "bonded" by rigid harmonic forces, mixed to give atomic compositions and number density equal to the measured samples. Forces between atoms that are not bonded are governed by a potential function with Lennard–Jones (12-6) and Coulombic terms, shown as Eq. 1.11.

$$U_{ab}\left(r_{ij}\right) = 4\varepsilon_{ab}\left(\left(\frac{\sigma_{ab}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ab}}{r_{ij}}\right)^{6}\right) + \frac{q_a q_b}{4\varepsilon_0 r_{ij}} \quad (1.11)$$

In the above, ε_{ab} represents the potential well depth, and σ_{ab} represents the distance at which the Lennard–Jones terms contribute zero to the interaction energy. The parameters needed for Eq. 1.11, the reference potential for the simulation, are taken from force field

parameterizations of the system of interest. Figure 1.6 provides a conceptual representation of the EPSR refinement process.

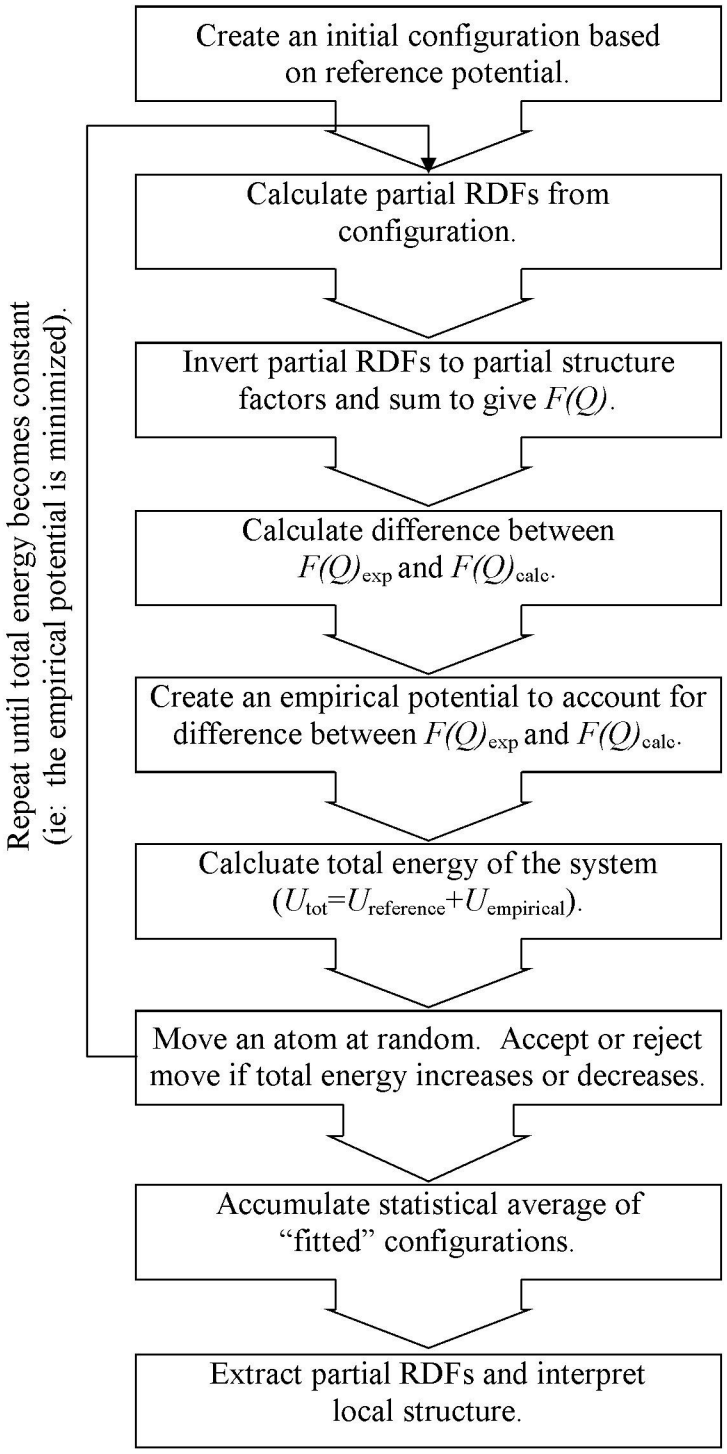

Figure 1.6: Schematic representation of the EPSR method.

1.2.4 Instrumentation for Measurement of the Differential Scattering Cross Section

All diffraction measurements presented here were made using the Small Angle Neutron Diffractometer for Amorphous and Liquid Samples (SANDALS) at the ISIS Pulsed Neutron Facility, Chilton, Didcot, UK. SANDALS is a time-of-flight diffractometer using a high-flux incident white beam of neutrons ranging from 0.5 Å to 3.5 Å in wavelength. The detectors on SANDALS are arranged in 18 groups covering 3.5° to 37.5° 2θ (0.1 $Å^{-1}$ to 50 $Å^{-1}$ in Q space) to take advantage of the long-wavelength neutrons for measurements at low Q and also to minimize effects of atomic recoil in the sample due to lower energy transfer from the incident neutrons in the forward scattering direction. SANDALS is also equipped with a neutron transmission monitor that allows for *in situ* comparison of the total scatter from the sample with that predicted by isotopic composition.

2. Statement of the Problem

The production of ethanol from cellulose and cellulosic biomass, as stated previously, requires extensive chemical pre-treatments to increase the surface area or to solubilize the cellulose component. These processes typically contribute significant cost to ethanol production. In fact, economic studies of cellulosic ethanol production have cited reduced pre-treatment costs as a necessity for the economic success of cellulosic ethanol.[44]

Water-based pre-treatments could potentially reduce the costs of cellulosic ethanol if the intrinsic insolubility of bulk cellulose in pure water could be overcome. The extensive hydrogen-bonding network of crystalline cellulose is commonly attributed to the biopolymer's insolubility in water and most other common solvents. Therefore, studying the hydrogen-bonding interactions of a soluble cello-oligosaccharide–water system will enhance fundamental understanding of the changes from bulk solute and solvent necessary for dissolution in water and inform further studies with higher oligosaccharides.

Liquid-state neutron diffraction with isotopic substitution is an ideal technique for structural investigation of hydrogen-bonded liquids and solute–solvent systems due to the high relative sensitivity of neutrons to hydrogen and to the easily discernible isotopic contrast between hydrogen and deuterium. NDIS in conjunction with Empirical Pair Structure Refinement can be used to derive pair radial distribution functions and quantify changes in hydrogen bonding in solute and solvent and new interactions between the two.

The objectives of this study should therefore be clear. A model cellulose compound is studied in aqueous solution by NDIS and EPSR. Particular attention is paid to hydrogen bonding as revealed by radial distribution functions and coordination numbers derived from the EPSR model. Since the EPSR models are relied upon heavily in interpreting molecular structure,

comparisons of the models with prior spectroscopic and computational studies will be made for further physical validation of the model. These correlations will also serve as qualitative evaluation of the performance of the EPSR technique itself with the forty-five atom solute under consideration, which is the largest to-date to be considered in substantial detail.

3. Structural Aspects of Cellobiose–Water Solutions

3.1 Cellobiose as a Cellulose Model

Cellobiose, β-D-glucopyranosyl-(1→4)-β-D-glucopyranose, shown below as Figure 3.1 is the repeating disaccharide subunit of cellulose. Representing "monomeric" cellulose, the structure of cellobiose has been studied extensively both to gain information about the disaccharide itself and for extrapolation to cellulose from the chemically simpler system. The early recognition of cellobiose as the repeat unit in the crystal structure of cellulose[13] gave great significance to determining the crystalline structure of cellobiose, which was achieved by Jacobson et al.[45] in 1961 with minor revisions by Chu and Jeffrey[46] in 1968.

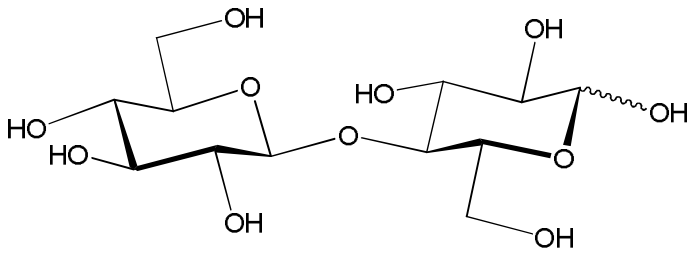

Figure 3.1: Cellobiose (β-D-glucopyranosyl-(1→4)-β-D-glucopyranose)

The considerable solubility of cellobiose in polar solvents, particularly water, has made it an ideal molecule for spectroscopic studies of conformation, primarily by NMR methods. From the time of the first NMR spectrometers capable of resolving carbohydrate resonances to the present day, various determinations of coupling constants[47-49] and spin relaxation behavior[50] have been made to characterize the relative solution-state conformations of the two pyranose rings about the glycosidic linkage. Furthermore, more recent spectroscopic studies have sought to quantify the populations of intramolecular cellobiose hydrogen bonds in solution.[51-53]

Cellobiose has also often served as a cellulose model in computational studies. Again, from the first stereochemical approaches[54] to modern quantum mechanical methods,[55] questions regarding the glycosidic conformation of the molecule have been prime. Molecular dynamics

(MD) simulations have allowed for computational studies of cellobiose both *in vacuo*[56] and in aqueous solution.[57-59] These studies have addressed intramolecular hydrogen bonding in cellobiose as a solute and intermolecular hydrogen bonding between the disaccharide and solution, in addition to the question of glycosidic conformation.

In addition to the extensive studies of cellobiose in the literature, the water solubility of cellobiose is sufficient for preparation of solutions approaching one molar which is important in establishing statistical significance in the measured scatter in comparison with the naturally high scatter of water alone.

3.2 Experimental Neutron Diffraction

A series of cellobiose–water samples with increasing isotopic substitution were measured at molar ratios of 1:63 cellobiose:water which corresponds to ~0.88 M solutions. D-(+)-cellobiose (β-D-glucopyranosyl-(1→4)-β-D-glucopyranose, 99.0%) was purchased from Sigma-Aldrich Chemie GmbH, and D_2O (99.9%) was purchased from Cambridge Isotopes. All solutions were prepared by mass and those containing D_2O were prepared using OD-exchanged cellobiose previously lyophilized from D_2O. Table 1 summarizes the samples measured.

Table 3.1: **Cellobiose–water solutions measured by neutron diffraction**.

Sample
Cellobiose OD in D_2O
Cellobiose OD in 87.5% D_2O 12.5% H_2O
Cellobiose OD in 75% D_2O 25% H_2O
Cellobiose OD in 62.5% D_2O 37.5% H_2O
Cellobiose OD in HDO
Cellobiose OD in 25% D_2O 75% H_2O
Cellobiose OD in H_2O

Sample solutions were transferred to Ti–Zr alloy sample cans with a flat-plate geometry and sample thickness of 1.0 mm. Diffraction data were collected on the SANDALS (small angle neutron diffractometer for amorphous and liquid samples) at the ISIS pulsed neutron facility of

Rutherford Appleton Laboratory, Chilton, UK. Samples were maintained at 298 K during measurement, and diffraction patterns for each sample were collected for a total of 1500 μA hours. Diffraction data were also collected using each of the empty Ti–Zr cells under identical conditions to allow for background subtraction. As the SANDALS diffractometer is equipped with a transmission monitor that measures the total scattering cross-section of the measured sample, the transmission from each sample was compared against theoretical transmission values with each sample's transmission in agreement with the theoretical value within 10%.

Corrections for absorption, multiple, container and self scattering and inelasticity effects were performed and the resulting data was then converted to *F(Q)* using Gudrun, a program derived from the ATLAS suite[37] and available at the ISIS facility.

3.3 EPSR simulation

An EPSR model was constructed using 8 α-cellobiose molecules, 12 β-cellobiose molecules and 1263 water molecules with initial cellobiose bond lengths and angles taken from the crystal structure of β-methyl-cellobiose methanol solvate[49] the α:β cellobiose ratio taken from NMR measurements. The atomic density of the box was set to 0.108 atoms $Å^{-1}$ as the density of 1:63 cellobiose:H_2O was determined as 1.08 g mL^{-1}. Table 2 shows the starting reference potentials and partial charges[60-62] used in the EPSR refinement. Two models were constructed based on modifications of the AMBER[63] and CHARMM[64] force fields and compared as tests for bias from the starting potentials in the final structures obtained.

Table 3.2. EPSR reference potentials. Atom labels match those shown in Figure 3.2.

	AMBER			Brady-CHARMM		
Atom	ε/kJ mol^{-1}	σ/Å	q_e	ε/kJ mol^{-1}	σ/Å	q_e
O_w	0.65000	3.1660	-0.84760	0.65000	3.1660	-0.84760
H_w	0.00000	0.0000	0.42380	0.00000	0.0000	0.42380
O1(') OL	0.87864	1.6612	-0.39120	0.41840	1.6500	-0.40000
C	0.35982	1.9080	0.20797	0.13389	2.0000	0.15500
O(')	0.88031	1.7210	-0.63819	0.80375	1.7650	-0.06980
H(')	0.00000	0.0000	0.42380	0.00000	0.0000	0.42380
M	0.065689	1.3870	0.029486	0.18828	1.3400	0.09000

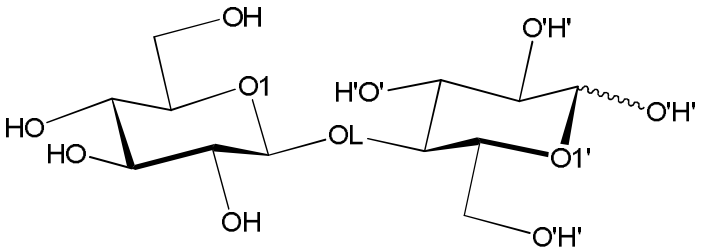

Figure 3.2: Atomic labels for EPSR modeling.

3.4 Experimental NMR Spectroscopy

Determinations of anomeric ratios of cellobiose solutions were conducted using a Bruker AMX spectrometer equipped with a 5-mm broadband inverse probe and variable-temperature unit and operating at resonance frequencies of 400 MHz for ^{1}H and 100 MHz for ^{13}C. All spectra were processed offline using MestReC 4.9 (MestReLab Research SL, Santiago de Compostela, Spain).

^{13}C NMR spectroscopy was used to determine anomeric ratios of α cellobiose to β cellobiose at the 1:63 cellobiose:water concentration studied by neutron diffraction. Samples of deuterium-exchanged cellobiose in D_2O and fully *proteo* cellobiose in H_2O, with 10% D_2O added for field locking the instrument, were prepared by mass and transferred to standard 5-mm sample tubes. Inverse gated ^{13}C spectra acquired using 3.5 a second relaxation delay and 8 K transients were obtained to allow for integration of carbon resonances and direct determination of the anomeric ratios. In each case, the signal-to-noise ratio of the resulting spectra exceeded

30:1, and sample temperatures were maintained at 25 ^{o}C ± 0.2 ^{o}C during the course of acquisition.

3.5 Evaluating the Cellobiose–Water EPSR Model by Correlation with Spectroscopic and Computational Studies

Figures 3.3 and 3.4 below show the experimental structure factors (circles) and those derived from the EPSR models (lines) based on the AMBER and Brady–CHARMM force field parameters.

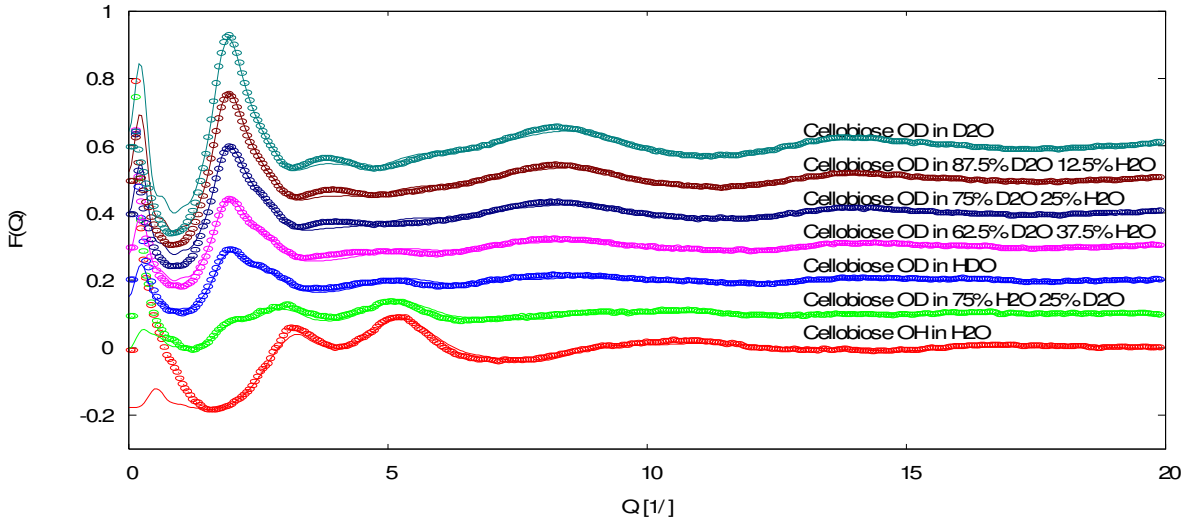

Figure 3.3: Experimental and AMBER-based model structure factors.

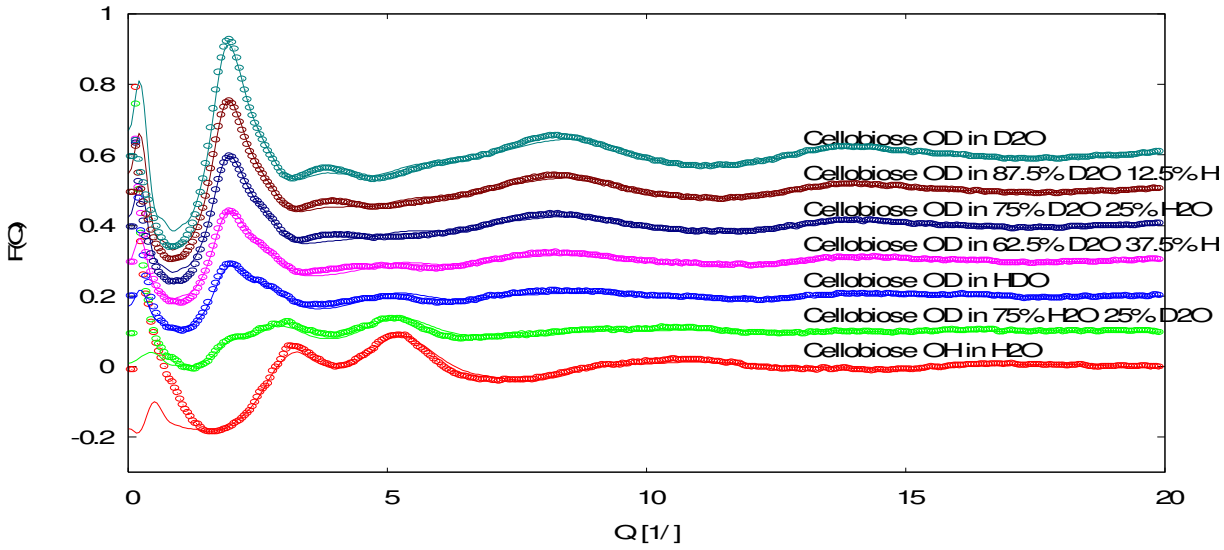

Figure 3.4: Experimental and Brady–CHARMM-based structure factors.

As a Reverse Monte Carlo method, EPSR has the potential to generate different structural models that "fit" the experimental data equally well.[43] Therefore, particular effort has been made to correlate properties of the EPSR models to results from measurements of the cellobiose–water system by NMR spectroscopy and from molecular dynamics simulations of cellobiose that explicitly include water as solvent.

An obvious place to begin comparison is the glycosidic configuration or relative rotation of the two pyranose rings of cellobiose about the β-(1→4) glycosidic linkage. Rotation about this bond is typically considered the greatest form of conformational flexibility present in cellobiose and other glucose oligosaccharides as the pyranose ring pucker can generally be assumed as static in the $^4C_1$ chair conformation.[65] Glycosidic linkage conformations of oligosaccharides are commonly described in terms of the angles φ and ψ, which are analogous to the angles of the same name defined by Ramachandran[66] for peptide chains. Figure 3.5 defines both the angles φ and ψ and the $^4C_1$ ring pucker for cellobiose.

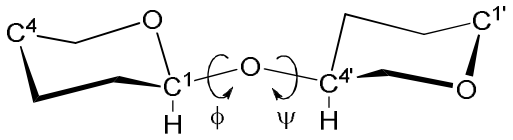

Figure 3.5: $^4C_1$ chair conformation and φ ψ angles for cellobiose.

NMR measurements of the vicinal scalar couplings between hydrogen and carbon atoms across the glycosidic linkage of an oligosaccharide can be related to the φ,ψ torsion angles by use of a Karplus-type relationship.[67] Two such studies of cellobiose are reported in the literature with reported φ,ψ vales of 32.2° -21.1° [49] and 45.4° -37.7°. [48] These values vary substantially from the angles observed in the crystal structure (24.33° -47.4°) from which the solute molecules input into the EPSR model were built. Migration of the φ,ψ angles from the starting crystal structure values to the time-averaged conformation from the NMR studies indicates the model is reproducing cellobiose conformation similarly to the solution, instead of solid state. Such migration of the average glycosidic torsion angles is observed in both models evaluated with the AMBER model giving angles of (24.15 -23) and the Brady–CHARMM model giving angles of (26.34 -33.16). Figure 3.6 summarizes this conformational agreement by showing the φ,ψ angles from the crystal structure, NMR measurements, and EPSR models in the form of a Ramachandran plot. Also plotted are the φ,ψ pairs for the cellobiose molecules in one configurational snapshot of each model to demonstrate the region of conformational space explored by the simulations.

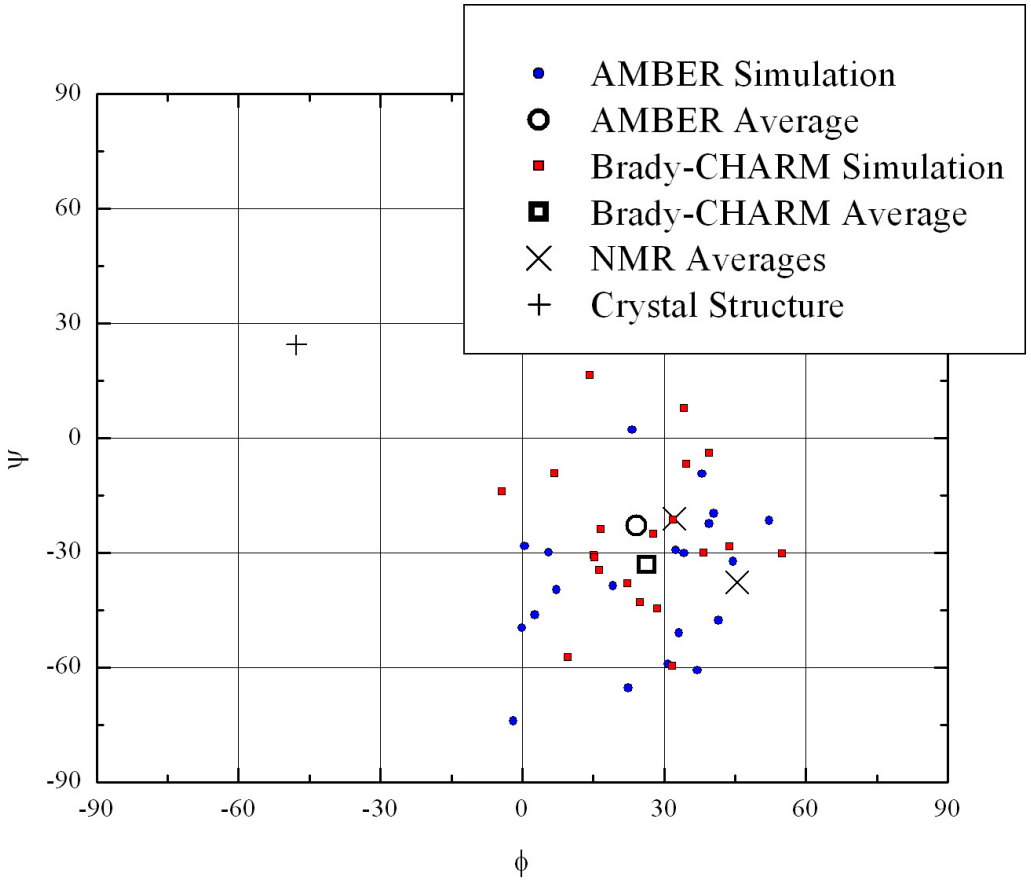

Figure 3.6: Ramachandran plot of *φ*, *ψ* angles for cellobiose–water model.

In addition to conformation about the glycosidic linkage, investigation of the hydration of the linkage oxygen has also been of interest in MD studies. A study comparing hydration of malto- and cello-oligosaccharides of varying DP[58] showed that disaccharides of both linkage types exhibit similar hydration with respect to the hydroxyl groups; however cellobiose showed a greater number of "doubly hydrogen bonded" waters that bridged the two pyranose rings. As such, the distribution of water oxygens about the linkage oxygen should be considerable. Figure 3.7 shows the RDFs for the OL-Ow pair from both the AMBER and Brady–CHARMM models that demonstrate a clear distribution of water oxygens about the linkage with coordination numbers of 1.11 and 1.02 for the two models, respectively.

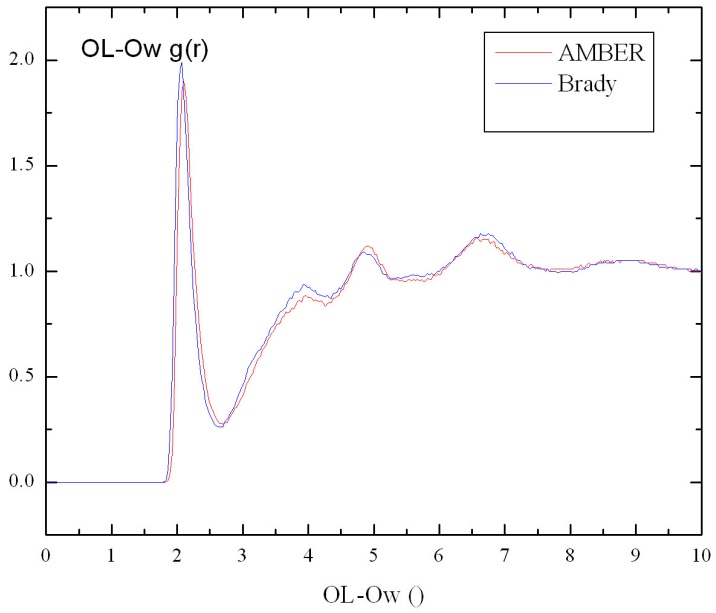

Figure 3.7: RDF for the OL-Ow atom pair.

While the glycosidic linkage is a dominant feature in examination of the conformation and hydration of oligosaccharides, the rotation of the exocyclic hydroxymethyl groups (carbons 6 and 6') introduces additional conformational variation that propagates into different solvation patterns[68] around the molecule. From both experiment[69, 70] and simulation[71-74] three predominant hydroxymethyl conformations have been determined with C4–C5–C6–O6 torsion angles, ω, of 60°, -60° and 180°. These conformations are typically referred to as gauche–gauche (*gg*), trans–gauche (*tg*) and gauche–trans (*gt*), respectively, which refers to the gauche or trans orientation of O5 with respect to O6 and C4.[75] Figure 3.8 demonstrates these hydroxymethyl conformations in the form of Newman projections.

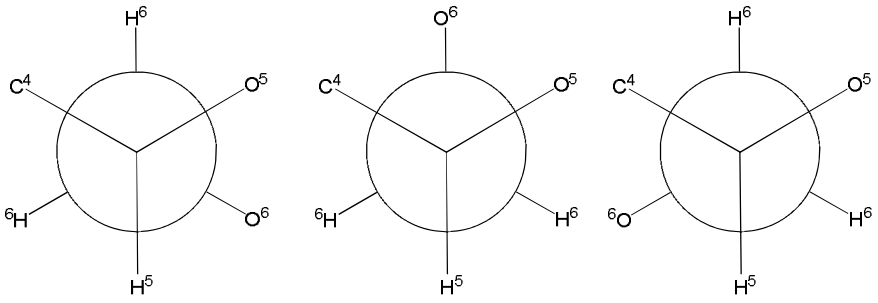

Figure 3.8: Newman projections of the *gt*, *gg* and *gt* hydroxymethyl conformations.

The conformational probabilities of the exocyclic hydroxymethyl groups from the AMBER-derived EPSR model are shown below in Figure 3.9; the Brady–CHARMM model showed similar conformational behavior. It can be seen clearly that cellobiose in the EPSR models is adopting a *tggt* hydroxymethyl conformation. This conformation has been demonstrated as energetically stable in MD simulations of cellobiose.[76] However, NMR measurements of the geminal coupling constants of the hydroxymethyl protons of both cellobiose[77] and α-methyl cellobioside[70] give values between -12.3 Hz and -12.6 Hz which, when deconvolved via the method of Stenutz et al.,[78] correspond to conformer populations of 40–50%:40–50%:5% *gt*:*gg*:*tg*. These findings correspond with earlier studies of hydroxymethyl conformation in glucose,[69] which show essentially no population of the *tg* conformation, along with a recent combined neutron diffraction and molecular dynamics study of glucose in aqueous solution that indicates *gg* and *gt* as the important conformations.[29] In light of the experimental results, it would seem plausible that the model assumed the local minimum energy associated with the *tg* conformation for the non-reducing ring and was not able to rotate through the *gg* and *gg* conformations. The possible effects of this non-physical conformation on the hydration of cellobiose as predicted by the EPSR models will be addressed in the following section.

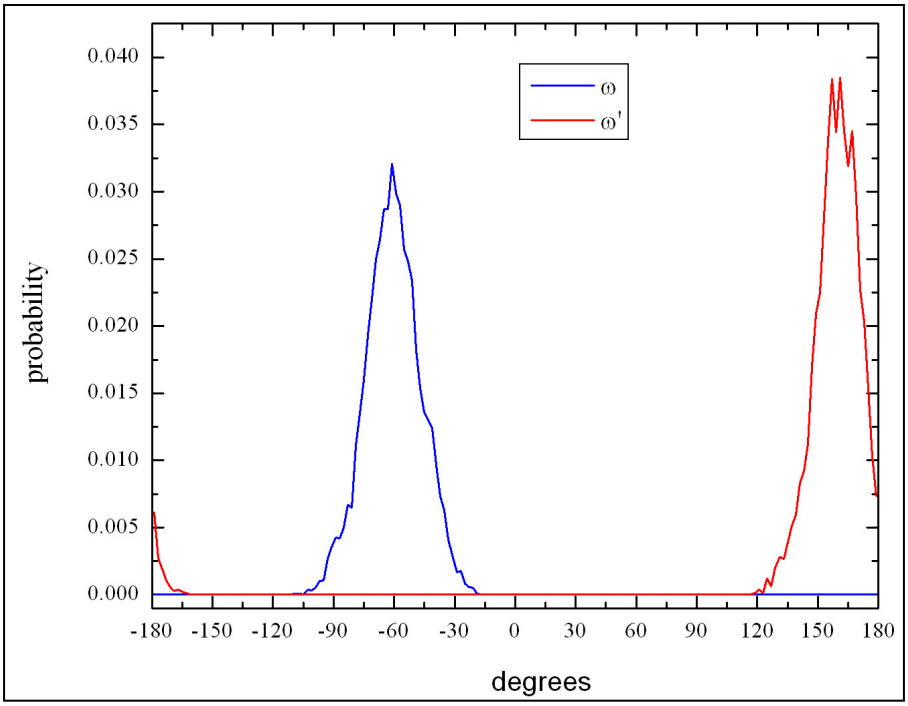

Figure 3.9: Average exocyclic hydroxymethyl conformation probabilities from the AMBER-based EPSR model

3.6 Intra- and Intermolecular Hydrogen-Bonding Interactions in the Cellobiose–Water System

Analysis of the cellobiose–water models derived from EPSR simulation, should begin with examination of the water structure as water constituted ~80 at. % of the system. Figures 3.10–3.12 show the RDFs derived from the EPSR models for the Hw-Hw, Ow-Hw, and Ow-Ow atom pairs plotted against the RDFs of neat water,[79] while Table 3.3 summarizes the coordination numbers (CNs) derived from the same RDFs by integration.

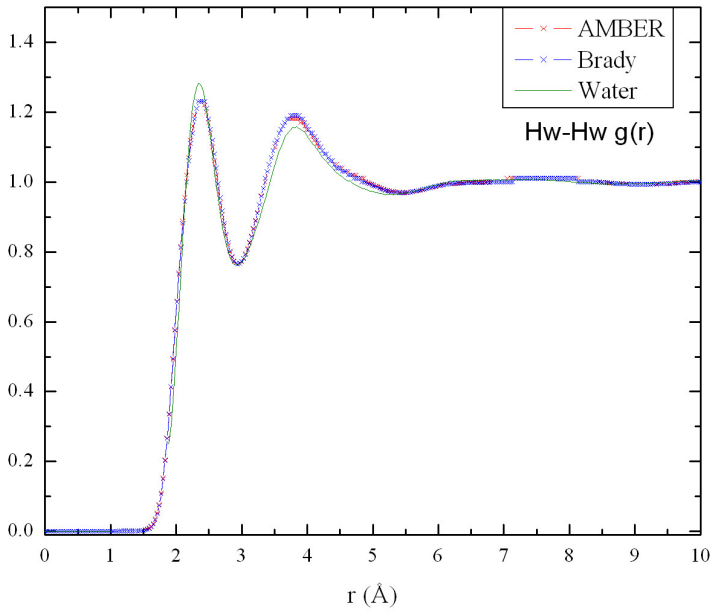

Figure 3.10: Hw-Hw RDFs for AMBER- and Brady–CHARMM-based EPSR models and pure water.

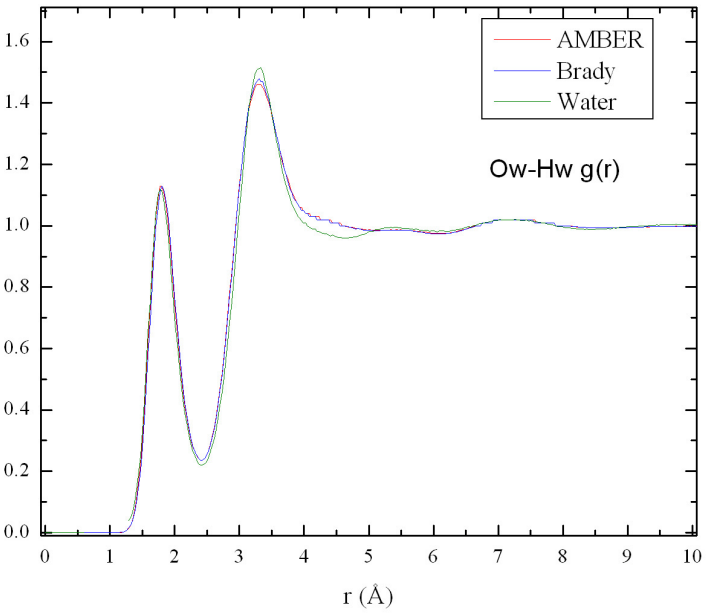

Figure 3.11: Ow-Hw RDFs for AMBER- and Brady–CHARMM-based EPSR models and pure water.

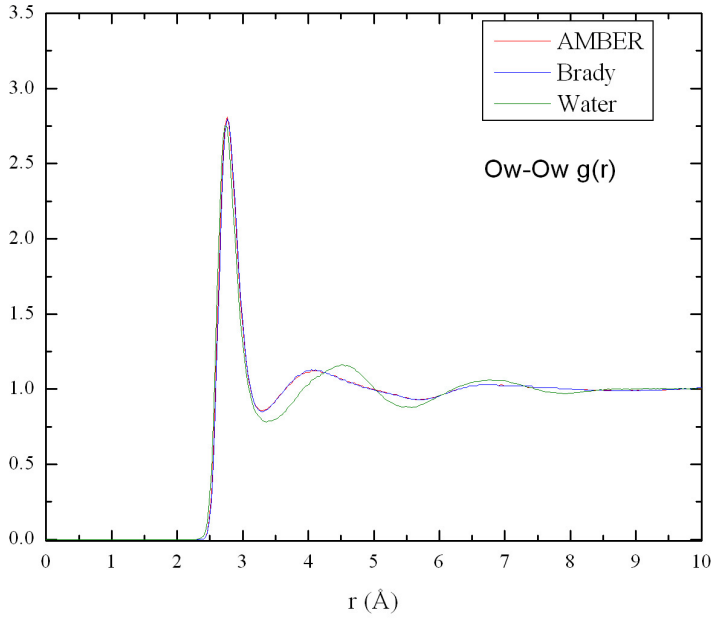

Figure 3.12: Ow-Ow RDFs for AMBER- and Brady–CHARMM-based EPSR models and pure water.

Table 3.3: Bulk water coordination numbers.

	Hw-Hw	Ow-Hw	Ow-Ow
Pure water	~4-5	~1.8	~4.5-5
Cellobiose–Water	4.40	1.57	3.73

Two characteristics of the RDFs and derived CNs suggest changes to the bulk water structure upon addition of cellobiose as solute. First, the Ow-Hw and Ow-Ow CNs clearly decrease, indicating a reduction in the number of water–water hydrogen bonds. This reduction has been observed in studies of other hydrogen-bond donating/accepting bioorganic solutes with similar interpretation of the effect.[80] Additionally, in Figure 3.12, a shift to lower radius of the maximum associated with the water second coordination sphere is observed. A systematic study by Mason and Brady[81] of the O-O RDFs generated for water by various computational water

models, including SPC/E, has suggested that this RDF is particularly sensitive to only minor changes in the tetrahedral coordination geometry of water. However, a similar shift in the water coordination sphere was observed in an NDIS study of trehalose,[30] an α-(1→1)-linked glucose disaccharide, and was interpreted as a significant condensation of the second coordination sphere comparable to the effects of extreme pressures.

The suggestion that cellobiose promotes substantial changes to the hydrogen-bonding structure of water is in stark contrast to the results of earlier NDIS studies of glucose in aqueous solution.[28, 82] In glucose solutions with concentration as high as five molal, corresponding to roughly 50 wt. % glucose, no perturbation to the bulk water structure is noted.[28] In fact, it has been suggested by Sidu et al. that the distribution of the five hydroxyl groups about the pyranose ring of glucose allows water to form "a continuous ring of water [molecular] density around the plane of the glucose ring" allowing the bulk water structure to incorporate the solute with little change in overall structure.[82] In this theory lies a potential explanation for the greater destructuring effects of cellobiose noted in this study; the glycosidic linkage of cellobiose introduces anisotropy into the distribution of hydroxyl groups about the two glucose rings potentially preventing the formation of a "ring" of coordinating waters and requiring greater change in the bulk water structure for solute accommodation.

The intramolecular hydrogen bonding of cellobiose is also considered. In the solid state, cellobiose exhibits a clear hydrogen bond between hydroxyl OH3' and ring oxygen O5.[46] Numerous spectroscopic and MD studies have attempted to quantify the population of this bond in solution. Leeflang et al. measured the exchange rates of hydroxyl protons by NOE NMR measurements to extract hydroxyl chemical shift information and concluded that OH3' did not exhibit the downfield shift characteristic of participation in hydrogen bonding.[51] Alternatively,

Poppe et al. directly observed the chemical shifts of hydroxyl protons for a number of glucose disaccharides in supercooled water and concluded that OH3' did exhibit a downfield shift suggesting a population of the OH3'–O5 hydrogen bond of roughly 50%.[52] Molecular dynamics results generally concur on roughly 50% to 60% population.[57-59] In the course of EPSR refinement of the NDIS structure factors collected in this study, it became of interest to investigate the presence of this bond. Therefore, the EPSR models were adjusted to distinguish between oxygen and exchangeable hydrogen atoms of the reducing and nonnreducing rings, which, as is obvious from the potential parameters used for each, would typically be considered equivalent atom types in an NDIS/EPSR study. Figure 3.13 below shows the RDFs derived from the O1-H' pairs of the EPSR models, representing all hydroxyl protons of the reducing ring in distribution about the oxygen (O5) of the non-reducing ring. Clearly, the lack of any significant local maxima below 10 Å provides no evidence for the existence of the OH3'–O5 hydrogen bond in this system, though it should be noted that the averaging across all hydroxyl groups of the reducing ring could greatly reduce the prevalence of distributions arising exclusively from one hydroxyl group.

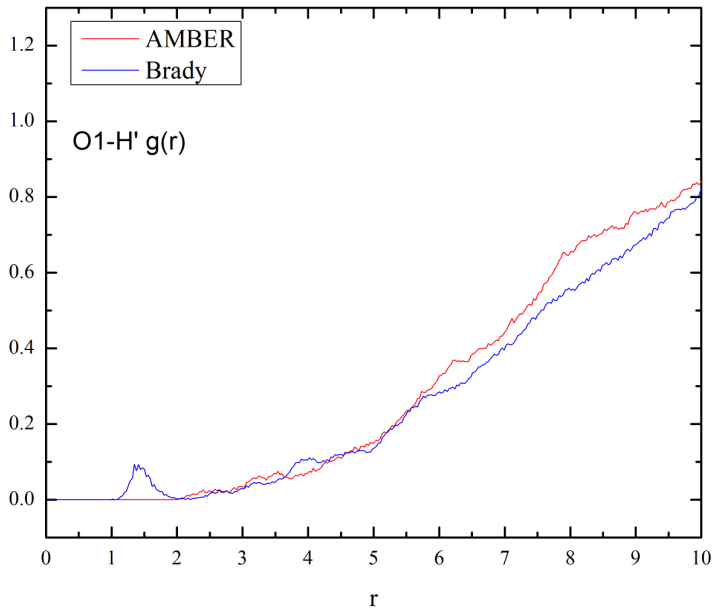

Figure 3.13: O1-H' RDFs from AMBER- and Brady-CHARMM-based EPSR models.

In considering the intermolecular hydrogen bonding between cellobiose and water, it became apparent in the H-Ow and O-Hw RDFs that both EPSR models demonstrated greater cellobiose–water coordination for the hydroxyl groups of the reducing ring in comparison to the non-reducing ring. Figure 3.14 below provides an example of this "hydration preference" from the AMBER based EPSR model.

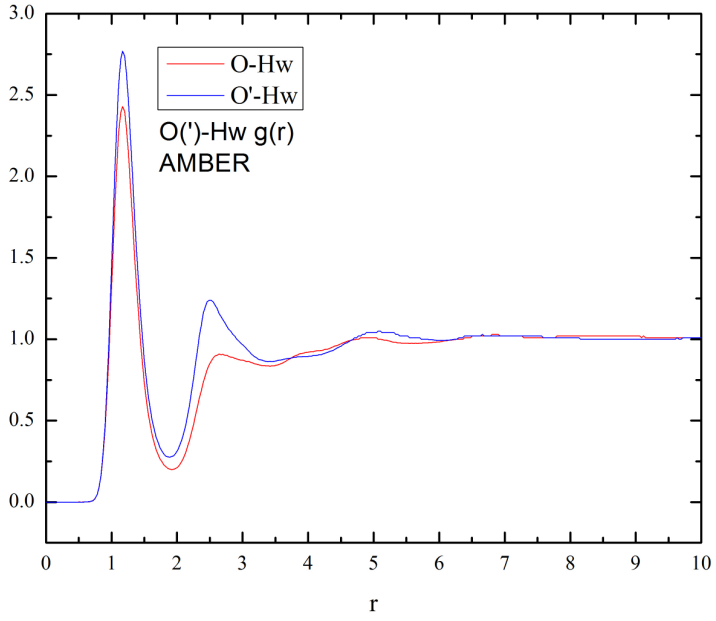

Figure 3.14: O(')-Hw "Hydration Preference" example from AMBER-based EPSR model.

Evaluation of this effect led to concern that the non-physical exocyclic hydroxymethyl conformation, discussed previously, could be distorting the average hydration of the molecule. To investigate the effects of hydroxymethyl conformation on solvation, the hydrogen bond donor and acceptor density[83] was mapped onto the Connelly water-accessible surfaces of a series of cellobiose molecules with differing hydroxymethyl conformations and this surface area for each ring at half-maximum density was calculated using the Sybyl8.0 suite of programs (Tripos Associates, St Louis, MO). Table 3.4 summarizes the results of the density map predictions that clearly demonstrate a strong dependence of the density of hydrogen bonding sites available for interaction with water on the hydroxymethyl conformation. In light of these results, the cellobiose–water coordination numbers derived from the EPSR models are presented, in Table 3.5, with the disclaimer that further refinement of the structural models is necessary to ensure a physically consistent representation of the hydration behavior.

Table 3.4: Hydrogen bond donor acceptor densities for different hydroxymethyl conformations of cellobiose.

	tggt	*gggt*	*gggg*	*gtgg*	*gtgt*
maximum density (donor or acceptor/ Å$^{2)}$	0.2486	0.1638	0.1976	0.1996	0.1984
reducing ring density surface area at half maximum (Å^2)	26.41	63.69	42.07	49.76	45.12
nonreducing ring density surface area at half maximum (Å^2)	13.62	91.28	38.51	53.83	53.16
"hydration preference"	reducing	non	equal	equal	non

Table 3.5: Cellobiose–Water Coordination Numbers.

	Coordination Number	
sites	AMBER	Brady-CHARM
H-Ow	0.19	0.07
H'-Ow	0.73	0.75
O-Hw	1.30	0.82
O'-Hw	1.47	1.37
OL-Hw	0.79	0.82
O1-Hw	0.91	0.42
O1'-Hw	1.02	1.05

3.7 Conclusions

Neutron diffraction with isotopic substitution experiments coupled with Empirical Potential Structure Refinement have revealed structural information on the atomic length scale from a model cellobiose–water system. In the cellobiose–water solutions measured, changes to both the bulk water hydrogen-bonding network and the solid-state hydrogen bonding of cellobiose were observed. In addition, conformational flexibility, particularly the exocyclic groups, has been shown to affect overall solvation of the cello-saccharide solute. These results demonstrate the fundamental importance of hydrogen bonding in cellulose solubility and have provided foundational data and insight for future structural studies of higher cello-oligomers in aqueous solution.

4. References

1. Payen, A., Memoire sur la composition du tissu proper des plantes et du ligneux. *Comptes Rendus* **1838,** *7*, 1052-1056.
2. Kirchhoff, G., *Technologicheski Zhurnal* **1812,** *9*, 3.
3. Braconnot, H., Sur la conversion du corps ligneux en gomme, en sucre et un acide d'une nature particuliere, par le moyen de l'acide sulfurique. *Annales de Chimie* **1819,** *12*, 172-195.
4. Melsens, G. F., *Dinglers Polytechnic Journal* **1856,** *138*, 426.
5. Kamm, B.; Kamm, M.; Gruber, P. R.; Kromus, S., Biorefinery Systems--An Overview. In *Biorefineries--Industrial Processes and Products*, Kamm, B.; Gruber, P. R.; Kamm, M., Eds. Wiley-VCH: Weinheim, 2006; Vol. 1, pp 3-33.
6. Harris, E. E.; Beglinger, E., Madison Wood Sugar Process. *Industrial Engineering and Chemistry* **1946,** *38* (9), 890-895.
7. Prescott, S. C.; Dunn, C. G., *Industrial Microbiology*. third ed.; McGraw-Hill: New York, 1959.
8. Kosaric, N.; Vardar-Sukan, F., Potential Sources of Energy and Chemical Products. In *The Biotechnology of Ethanol*, Rohr, M., Ed. Wiley-VCH: Weinheim, 2001.
9. Lee, S., Ethanol from Lignocellulosics. In *Handbook of Alternative Fuel Technologies*, Lee, S.; Speight, J. G.; Loyalka, S. K., Eds. CRC Press: Boca Raton, 2007; pp 343-376.

10. Bragg, W. L., The Structure of some crystals as indicated by their diffraction of X-rays. *Proccedings of the Royal Society of London A* **1913,** *89*, 248-77.

11. Sponsler, O. L.; Dore, W. H., The structure of ramie cellulose as dreived from x-ray data. *Fourth Colloid Symposium Monograph* **1926,** *41*, 174-202.

12. Young, R. A.; Rowell, R. M., *Cellulose: Structure, Modification, and Hydrolysis*. Wiley: New York, 1986.

13. Meyer, K. H.; Mark, H., Uber den Bau des krystallisierten Anteils der Cellulose. *Berichte der Deutschen Chemischen Gesellschaft* **1928,** *61B*, 593-614.

14. Meyer, K. H.; Misch, H., Positions des atomes dans la nouveau modele spacial de la cellulose. *Helvetica Chimica Acta* **1936,** *20*, 232-244.

15. O'Sullivan, A. C., Cellulose: the structure slowly unravels. *Cellulose* **1997,** *4*, 173-203.

16. Zugenmaier, P., *Crystalline Cellulose and Derivatives*. Springer: Heidelberg, 2008.

17. Atalla, R. H.; VanderHart, D. L., Native Cellulose: A Composite of Two Distinct Crystalline Forms. *Science* **1984,** *223* (4633), 283-285.

18. Flugge, L. A.; Blank, J. T.; Petillo, P. A., Isolation, Modification, and NMR Assignments of a Series of Cellulose Oligomers. *Journal of the American Chemical Society* **1999,** *121*, 7228-7238.

19. Nehls, I.; Wagenknecht, W.; Philipp, B.; Stscherbina, D., Characterization of cellulose and cellulose derivatives in solution by high resolution ^{13}C NMR spectrometry. *Progress in Polymer Science* **1994,** *19*, 29-78.

20. Remsing, R. C.; Swatloski, R. P.; Rogers, R. D.; Moyna, G., Mechanism of cellulose dissolution in the ionic liquid 1-n-butyl-3-methylimidazolium chloride: a ^{13}C and $^{35}Cl/^{37}Cl$ NMR relaxation study on model systems. *Chemical Communications* **2006,** *12*, 1271-1273.

21. Brendler, E.; Fischer, S.; Leipner, H., ^{7}Li NMR as probe for solvent-cellulose interactions in cellulose dissolution. *Cellulose* **2001,** *8* (4), 283-288.

22. El-Kafrawy, A., Investigation of the Cellulose/LiCl/Dimethylacetamide and Cellulose/LiCl/*N*-Methyl-2-Pyrrolidinone Solutions by ^{13}C NMR Spectroscopy. *Journal of Applied Polymer Science* **1982,** *27*, 2435-2443.

23. Missori, M.; Mondelli, C.; De Spirito, M.; Castellano, C.; Biccheri, M.; Schweins, R.; Arcovito, G.; Papi, M.; Castellano, A. C., Modifications of Mesoscopic Structure of Cellulose in Paper Degradation. *Physical Review Letters* **2006,** *97* (23), 2380011-2380014.

24. De Spirito, M.; Mauro, M.; Massimiliano, P.; Giueseppe, M.; Jose, T.; Carlo, C.; Giuseppe, E., Modifications in solvent clusters embedded along the fibers of a cellulose

polymer network cause paper degradation. *Physical Review E, Statistical, Nonlinear, and Soft Matter Physics* **2008,** *77* (4), 41801.

25. Sugiyama, M.; Iigima, H.; Hara, K.; Nakamura, A.; Hiramatsu, N.; Maeda, Y., Structural investigation on aqueous suspsension of microcrystalline cellulose. *Transactions of the Materials Research Society of Japan* **2000,** *25* (3), 743-746.

26. Ishii, D.; Tatsumi, D.; Matsumoto, T.; Murata, K.; Hayashi, H.; Yoshitani, H., Investigation of the Structure of Cellulose in LiCl/DMAc Solution and Its Gelation Behavior by Small-Angle X-Ray Scattering Measurements. *Macromolecular Bioscience* **2006,** *6*, 293-300.

27. Mason, P.; Neilson, G.; Enderby, J.; Saboungi, M.; Brady, J., Structure of aqueous glucose solutions as determined by neutron diffraction with isotopic substitution experiments and molecular dynamics calculations. *Journal of Physical Chemistry B* **2005,** *109*, 13104-111.

28. Mason, P. E.; Neilson, G. W.; Barnes, A. C.; Enderby, J. E.; Brady, J. W.; Saboungi, M. L., Neutron diffraction studies on aqueous solutions of glucose. *Journal of Chemical Physics* **2003,** *119* (6), 3347-3353.

29. Mason, P. E.; Neilson, G. W.; Enderby, J. E.; Saboungi, M. L. C., Gabriel; Brady, J. W., Neutron diffraction and simulation studies on the exocyclic hydroxymethyl conformation of glucose. *The Journal of Chemical Physics* **2006,** *125*, 224505;1-9.

30. Pagnotta, S. E.; Ricci, M. A.; Bruni, F.; Mclain, S. E.; Magazu, A., Water structure around trehalose. *Chemical Physics* **2008,** *345*, 159-163.

31. McLain, S. E.; Soper, A. K.; Terry, A. E.; Watts, A., Structure and Hydration of L-Proline in Aqueous Solutions. *Journal of Physical Chemistry B* **2007,** *111*, 4568-4580.

32. Mclain, S. E.; Soper, A. K.; Watts, A., Water structure around dipeptides in aqueous solutions. *European Journal of Biophysicis* **2008,** *37*, 647-55.

33. Shull, C. G.; Wollan, E. O., X-Ray, Electron, Neutron Diffraction. *Science* **1948,** *108*, 69-75.

34. Sears, V., Neutron scattering lengths and cross sections. *Neutron News* **1992,** *3* (3), 29-37.

35. Egelstaff, P. A., *Thermal Neutron Scattering*. Academic Press: London, 1965.

36. Chieux, P., Introductory Theory. In *Neutron Diffraction*, Dachs, H., Ed. Springer-Verlag: Berlin, 1978; Vol. 6.

37. Soper, A. K.; Howells, W. S.; Hannon, A. C. *ATLAS—analysis of time-of-flight diffraction data from liquid and amorphous samples*; ISIS Pulsed Neutron Source, Rutherford Appleton Laboratory: 1989.

38. Placzek, G., Scattering of neutrons by systems of heavy nuclei. *Physical Review* **1952,** *86*, 377-388.

39. Soper, A. K., Orientational correlation function for molecular liquids: The case of liquid water. *Journal of Chemical Physics* **1994,** *101* (8), 6888-6901.

40. Mclain, S. E.; Benmore, C. J.; Siewenie, J. E.; Urquidi, J.; Turner, J. F. C., On the structure of liquid hydrogen fluoride. *Angewandte Chemie-International Edition* **2004,** *43* (15), 1952-1955.

41. McGreevy, R. L.; Pusztai, L., Reverse Monte Carlo simulation: a new technique for the determination of disordered structures. *Molecular Simulation* **1988,** *1* (6), 359-367.

42. Rietveld, H. M., Line Profiles of neutron powder-diffraction peaks for structure refinement. *Acta Crystallographica* **1967,** *22*, 151.

43. Soper, A. K., Partial structure factors from disordered materials diffraction data: an approach using empirical potential structure refinement. *Physical Reviews B* **2005,** *72* (10), 104204-16.

44. Kaylen, M.; Van Dyne, D. L.; Choi, Y.-S.; Blase, M., Economic feasibility of producing ethanol from lignocellulosic feedstocks. *Bioresource Technology* **2000,** *72*, 19-32.

45. Jacobson, R. A.; Wunderlich, J. A.; Lipscomb, W. N., The crystal and molecular structure of cellobiose. *Acta Crystallographica* **1961,** *14*, 598-607.

46. Chu, S. S. C.; Jeffrey, G. A., Refinement of crystal structures of β-D-glucose and cellobiose. *Acta Crystallographica, Section B.* **1968,** *24*, 830-838.

47. Parfondry, A.; Cyr, N.; Perlin, A. S., ^{13}C-^{1}H Inter-Residue Coupling in Disaccharides, and the Orientations of Glycosidic Bonds. *Carbohydrate Research* **1977,** *59*, 299-309.

48. Sugiyama, H.; Hisamichi, K.; Usui, T.; Sakai, K.; Ishiyama, J. I., A study of the conformation of β-1,4-linked glucose oligomers, cellobiose to cellohexaose, in solution. *THEOCHEM* **2000,** *556*, 173-77.

49. Cheetham, N. W. H.; Dasgupta, P.; Ball, G. E., NMR and modeling studies of disaccharide conformation. *Carbohydrate Research* **2003,** *338*, 955-962.

50. Lepri, A.; Marchettini, N.; Pogliani, L.; Rossi, C.; Ulgiati, S., NMR Structural Investigation of Cellobiose and Glucose. *Magnetic Resonance in Chemistry* **1987,** *25*, 521-23.

51. Leeflang, B. R.; Vliegenthart, J. F. G.; Kroon-Batenburg, L. M. J.; van Eijck, B. P.; Kroon, J., A ^{1}H NMR and MD study of intramolecular hydrogen bonds in methyl beta-cellobioside. *Carbohydrate Research* **1992,** *230*, 41-61.

52. Poppe, L.; van Halbeek, H., NMR spectroscopy of hydroxyl protons in supercooled carbohydrates. *Nature Structural Biology* **1994,** *1* (4), 215-.

53. Bernet, B.; Vasella, A., Intra- and Intermolecular H-Bonds of Alcohols in DMSO ^{1}H NMR Analysis of Inter-Residue H-Bonds in Selected Oligosaccharides: Cellobiose, Lactose, *N*,*N'*-Diacetylchitobiose, Maltose, Sucrose, Agarose, and Hyaluronates. *Helvetica Chimica Acta* **2000,** *83*, 2055-2072.

54. Rees, D. A.; Skerrett, R. J., Conformational Analysis of Cellobiose, Cellulose, and Xylan. *Carbohydrate Research* **1968,** *7*, 334-348.

55. French, A. D.; Johnson, G. P., Quantum mechanics studies of cellobiose conformations. *Canadian Journal of Chemistry* **2006,** *84*, 603-612.

56. Hardy, B. J.; Sarko, A., Conformational Analysis and Molecular Dynamics Simulation of Cellobose and Large Cello-oligomers. *Journal of Computational Chemistry* **1993,** *14* (7), 831-847.

57. Hardy, B. J.; Sarko, A., Molecular Dynamics Simulation of Cellobiose in Water. *Journal of Computational Chemistry* **1992,** *14* (7), 848-857.

58. Umenura, M.; Yuguchi, Y.; Hirotsu, T., Hydration at glycosidic linkages of malto- and cello-oligosaccharides in aqueous solution from molecular dynamics simulation: Effect of conformational flexibility. *THEOCHEM* **2005,** *730*, 1-8.

59. Pereira, C. S.; Kony, D.; Baron, R.; Muller, M.; van Gunsteren, W. F.; Hunenberger, P. H., Conformational and dynamical properties of disaccharides in water: a molecular dynamics study. *Biophysical Journal* **2006,** *90*, 4337-44.

60. Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P., The missing term in effective pair potentials. *Journal of Physical Chemistry* **1987,** *91* (24), 6269-71.

61. Glennon, T. M.; Merz, K. M., Jr., A carbohydrate force field for AMBER and its application to the study of saccharide to surface adsorption. *THEOCHEM* **1997,** *395-6*, 157-71.

62. Guvench, O.; Greene, S. N.; Kamath, G.; Brady, J. W.; Venable, R. M.; Pastor, R. W.; Mackerell, A. D., Additive Empirical Force Field for Hexopyranose Monosaccharides. *Journal of Computational Chemistry* **2008,** *29* (15), 2543-2564.

63. Weiner, S. J., An all atom force field for simulations of proteins and nucleic acids. *Journal of Computational Chemistry* **1986,** *7*, 230-252.

64. Mackerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E., III; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M., All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *Journal of Physical Chemistry B* **1998,** *102* (18), 3586-3616.

65. Shallenberger, R. S., *Advanced Sugar Chemistry: Principles of Sugar Stereochemistry*. AVI Publishing Co., Inc.: Westport, CT, 1982.

66. Ramachandran, G. N.; Ramakrishnan, C.; Venkatachalam, C. M., Determination of the possible conformations of the residues linked in a polypeptide chain. *Biopolymers* **1965,** *3* (5), 591-2.

67. Tvaroska, I.; Hricovini, M.; Petrakova, E., An attempt to derive a new Karplus-type equation of vicinal proton-carbon coupling constants for C-O-C-H segments of bonded atoms. *Carbohydrate Research* **1989,** *189*, 359-362.

68. Mason, P. E.; Neilson, G. W.; Enderby, J. E.; Saboungi, M. L.; Brady, J. W., Neutron Diffraction and Computer Simulation Studies of D-Xylose. *Journal of the American Chemical Society* **2005,** *127*, 10991-10998.

69. Nishida, Y.; Ohrui, H.; Meguro, H., H-1 NMR Studies of (6R)- and (6S)-Deuterated D-Hexoses: Assignemnet of the Preferred Rotamers about C5-C6 Bond of D-Glucose and D -Galactose Derivatives in Solutions. *Tetrahedron Letters* **1984,** *25* (15), 1575-1578.

70. Olsson, U.; Serianni, A. S.; Stenutz, R., Conformational Analysis of Beta-Glycosidic Linkages in ^{13}C Labeled Glucobiosides Using Inter-residue Scalar Coupling Constants. *Journal of Physical Chemistry B* **2008,** *112*, 4447-4453.

71. Brady, J. W., Molecular dynamics simulations of a-D-glucose. *Journal of the American Chemical Society* **1986,** *108* (26), 8153-8160.

72. Brady, J. W., Molecular dynamics simulations of β-D-glucopyranose. *Carbohydrate Research* **1987,** *165*, 306-312.

73. Tvaroska, I.; Kozar, T., *Theoretica Chimica Acta* **1986,** *70*, 99-114.

74. Kroon-Batenburg, L. M. J.; Kroon, J., Sovlent Effect on the Conformation of the Hydroxymethyl Group Established by Molecular Dynamics Simulations of Methyl β-D-Glucoside in Water. *Biopolymers* **1990,** *29*, 1243-1248.

75. Marchessault, R. H.; Perez, S., Conformations of the hydroxymethyl group in crystalline aldohexopyranoses. *Biopolymers* **1979,** *18* (9), 2369-74.

76. Hardy, B. J.; Sarko, A., Molecular Dynamics Simulation of Cellobiose in Water. *Journal of Computational Chemistry* **1993,** *14* (7), 848-857.

77. Roslund, M. U.; Tahtinen, P.; Miemitz, M.; Sjoholm, R., Complete Assignments of the ^{1}H and ^{13}C chemical shifts and $J_{H,H}$ coupling constnats in NMR spectra of D-glucopyranose and all D-glucopyranosyl-D-glucopyranosides. *Carbohydrate Research* **2008,** *343*, 101-12.

78. Stenutz, R.; Carmichael, I.; Widmalm, G.; Serianni, A. S., Hydroxymethyl Group Conformation in Saccharides: Structural Dependencies of $^{2}J_{H,H}$, $^{3}J_{H,H}$ and $^{1}J_{C,H}$ Spin-Spin Coupling Constants. *Journal of Organic Chemistry* **2002,** *67* (3), 949-958.

79. Soper, A. K.; Rossky, P. J., Special Issue on Liquid Water and Aqueous Solutions. *Chemical Physics* **2000,** *258*, 121-37.

80. Mclain, S. E.; Soper, A. K.; Watts, A., Structural Studies on the Hydration of L-Glutamic Acid in Solution. *Journal of Physical Chemistry B* **2006,** *110*, 21251-21258.

81. Mason, P.; Brady, J. W., "Tetrahedrality" and the Relationship between Collective Structure and Distribution Functions in Liquid Water. *Journal of Physical Chemistry B* **2007,** *111*, 5669-5679.

82. Sidhu, K. S.; Goodfellow, J. M.; Turner, J. Z., Effect of molecular shape and electrostatic interactions on the water layer around polar and apolar groups in solution. *Journal of Chemical Physics* **1999,** *110* (16), 7943-7950.

83. Keil, M. "Moldellierung and Vorhersage von Strukturen biomolekularer Assoziate auf der Basis von statistischen Datenbankanalyses". University of Darmdstadt, Darmstadt, Germany, 2002.