JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

COMMUNICATION

pubs.acs.org/JACS

Low-Temperature Neutron Diffraction Structures of *N*-Glycoprotein Linkage Models and Analogues: Structure Refinement and Trifurcated Hydrogen Bonds

Gianluca Cioci,[†] Amrita Srivastava,[‡] Duraikkannu Loganathan,[*,‡] Sax A. Mason,[*,§] Serge Pérez,[*,†] and Anne Imberty[*,∥]

[†]European Synchrotron Radiation Facility, BP220 Grenoble, France

[‡]Indian Institute of Technology Madras, Chennai 600036, India

[§]Institut Laue-Langevin, BP156 Grenoble, France

[∥]CERMAV-CNRS (affiliated with Université Joseph Fourier and ICMG), BP53 Grenoble, France

Supporting Information

ABSTRACT: The biological addition of oligosaccharide moieties to asparagine residues of *N*-glycoproteins influences the properties and bioactivities of these macromolecules. The low-temperature neutron crystal structures of three *N*-glycoprotein linkage models and analogues provide accurate characterization of the three-dimensional structure of the conserved GlcNAc−Asn linkage. These first crystal structures of N-acetylated carbohydrates obtained by neutron diffraction provide high-resolution geometrical parameters that can be used for force-field parametrization and subsequent molecular dynamics simulation of *N*-glycoproteins. The correct localization of hydrogen atoms demonstrates the occurrence of trifurcated hydrogen bonds and hydrophobic contacts.

Glycosylation is the most common post-translational modification of proteins, with more than 50% of proteins in eukaryotes being glycoproteins.[1] The glycan components modify the properties of the proteins (folding, solubility, resistance to proteolysis, masking of epitopes, etc.) and also serve as recognition determinants in many biological processes. N-Glycosylation is the most studied modification and consists of attachment of branched oligosaccharides through a 2-acetamido-2-deoxy-β-D-glucopyranosyl (GlcNAc) residue to the side chain of an asparagine (Asn) residue belonging to the consensus sequence Asn-Xaa-Ser/Thr.[2] Elucidation of the structure and conformation of glycoproteins, a challenging problem in glycobiology, is fundamental for obtaining a better understanding of their functions.

The conformation of the linkage between the protein and the glycan [e.g., between Asn and GlcNAc in compound **1** (Scheme 1)] determines the presentation of the sugar at the protein surface and therefore at the cell surface for membrane proteins. The X-ray structure of **1** has been solved.[3,4] A systematic X-ray crystallographic study of GlcNAcβAsn (**1**)[3,4] and GlcNAcβNAc (**3**), which are models of the *N*-glycoprotein linkage region, and their analogues has shown that the key torsion angle, Φ_N, is influenced to a greater extent by structural variation of the glycan part than of the aglycon group.[5] Analysis of the crystal structure also demonstrated the importance of bifurcated N−H···O and C−H···O hydrogen bonds that could stabilize the linkage conformation.[5,6]

Neutron diffraction has proved to be an essential tool for highly accurate determinations of molecular dimensions.[7] Full characterization of the intermolecular interactions and precise details of the geometry of the O−H···O hydrogen bonds in carbohydrate crystals has become possible through the use of single-crystal neutron diffraction.[8,9] Such data are also needed for parametrization of molecular mechanics force fields.

However, because of the reluctance of carbohydrates to crystallize in a form suitable for structural investigations and the even greater difficulty of obtaining large crystals, relatively few crystal structures have been determined by neutron diffraction experiments.[9,10] The structures have been obtained for furano- and pyranoforms of simple monosaccharides, a few disaccharides such as maltose and sucrose, and cyclodextrin. However, no neutron crystal structures are available for carbohydrates substituted with an *N*-acetamido group, such as GlcNAc, and therefore, no precise details of the geometry of this substituent or of the *N*-glycoprotein linkage. Furthermore, the N−H···O hydrogen bond and its bifurcated forms, which are driving forces in crystals of acetamido sugars,[6] have not been investigated with the same accuracy as the C−H···O linkage.

The synthesis, crystallization, and X-ray structures of compounds **2**, **3**, and **4** have been described previously.[5,11,12] For the present study, large crystals were grown as needles or plates with volumes between 0.5 and 20 mm^3. Monochromatic neutron data for all of the compounds were collected on diffractometer D19 at the Institut Laue-Langevin (ILL), which incorporates a very large curved 120 × 30° position-sensitive detector, using ω-step scans of 0.07° and a wavelength of 0.9468 Å. All of the crystals were cooled slowly to 20 K, and for compound **2** (GlcNAcβNHPr), two additional data sets were collected from another crystal at 120 and 295 K. The Bragg intensity data were processed with the ILL program RETREAT[13] and corrected for cryostat can attenuation and effective crystal absorption. The starting structural models were based on the previous X-ray structure determinations,[5,11,12] while all hydrogen atoms were located using neutron Fourier difference maps. The structures were refined by the full matrix least-squares method as implemented in SHELXL-97[14] using anisotropic displacement parameters for all atoms. Experimental and refinement statistics are given in Table S1 in the Supporting Information.

Received: April 21, 2011
Published: June 07, 2011

dx.doi.org/10.1021/ja203239j | *J. Am. Chem. Soc.* 2011, 133, 10042–10045

The unit cell parameters of compounds **2**, **3**, and **4** measured at 20 K and those measured at 120 and 295 K for compound **2** are given in Table 1, along with those obtained from the X-ray studies at 293 K.[5,11,12] The temperature decrease from room temperature to 20 K led to a decrease of 1% in the unit cell volume (range 1.01 to 1.03). The neutron structural determinations are more complete than the previous X-ray structure determinations in that all atoms were located very precisely and the H atom positions were correctly determined. Nevertheless, no general discussion of the molecular structure is warranted here, since for the non-H atoms the refinement essentially confirmed at a higher level of accuracy the features already presented in the previous reports of the X-ray structures.[5,11,12]

Since the present data are the first crystal structures of N-acetylated carbohydrates obtained by neutron diffraction, they provide high-resolution geometrical data that can be used for parametrization of GlcNAc. Structural parameters of the *N*-acetyl group at position C2 of the glucose rings (compounds **2** and **3**) are listed in Table 2. Bond lengths were obtained with an accuracy of 0.001 Å. The combined effect of diffraction by nuclei (rather than by electron density) and low temperature resulted in higher accuracy of the atom positions for the neutron structures than for the X-ray ones. The gain in estimated standard deviation (esd) for the geometrical parameters is small but significant (Table S2) and is even better for hydrogen bonds (Table S3).

The three compounds have an *N*-acetyl or *N*-propionyl group at position C1 (Figure 1) and are therefore suitable models for the *N*-glycoprotein linkage. The lengths of the C1−N1 bond and the openings of the O5−C1−N1 angles are reduced relative to C2−N2 and C1−C2 −N2 because of the exoanomeric effect. The preference of the Φ_N torsion angle (O5−C1−N1−C1′) for values close to −90° corresponds closely to previous observations based on crystal structures or ab initio calculations.[15,16]

The other valuable information that can be obtained from a high-resolution neutron crystal structure is the determination of H atom positions that are free of the asphericity effects observed in X-ray experiments.[7] First, it can be noted that the hydroxyl groups

Scheme 1. Model Compounds with *N*-Glycoprotein Linkages

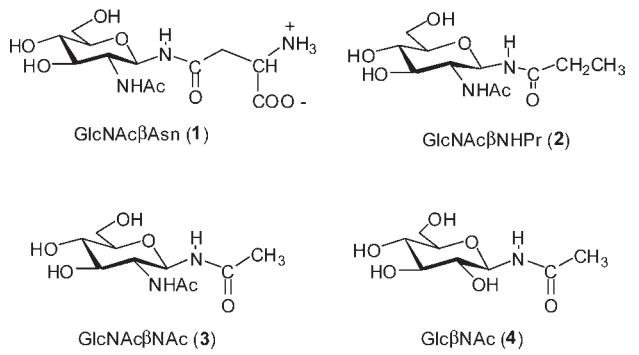

Table 2. Bond Lengths and Valence and Torsion Angles for the *N*-Acetyl Groups of the Compounds As Determined by Neutron Diffraction at 20 K

	GlcNAc*β*Pr (**2**)	GlcNAc*β*NAc (**3**)	Glc*β*NAc (**4**)
Bond Lengths (Å)			
NAc on C1			
C1−N1	1.425(1)	1.433(2)	1.4290(7)
N1−C1′	1.359(1)	1.359(2)	1.3506(7)
C1′−O1′	1.231(1)	1.229(2)	1.2382(9)
C1′ −C2′	1.519(1)	1.508(2)	1.5021(8)
NAc on C2			
C2 −N2	1.451(1)	1.448(2)	
N2− C1″	1.344(1)	1.346(2)	
C1″−O1″	1.235(2)	1.236(2)	
C1″−C2″	1.504(2)	1.498(2)	
Valence Angles (deg)			
NAc on C1			
O5−C1−N1	107.52(8)	106.9(2)	108.31(5)
C1−N1−C1′	120.55(8)	120.4(1)	121.80(4)
N1−C1′−O1′	123.0(1)	122.9(1)	121.61(6)
N1−C1′ −C2′	114.91(9)	115.0(1)	116.19(5)
NAc on C2			
C1−C2 −N2	110.99(8)	111.2(1)	
C2−N2− C1″	122.55(8)	122.8(1)	
N2 −C1″−O1″	122.3(1)	122.1(1)	
N2 −C1″−C2″	116.44(9)	115.9(1)	
Torsion Angles (deg)			
NAc on C1			
O5−C1−N1−C1′ (Φ_N)	−88.4(1)	−87.8(2)	−93.90(6)
C1−N1−C1′−C2′	169.33(8)	173.0(1)	−178.83(5)
H1−C1−N1−HN	−165.3(3)	−165.0(4)	−155.3(1)
NAc on C2			
C1−C2−N2−C1″	128.72(9)	127.4(1)	
C2−N2−C1″−C2″	−178.73(8)	−177.8(1)	
H2−C2−N2−HN	−172.3(3)	−173.1(4)	

Table 1. Unit Cell Parameters Determined for the Crystal Structures of Compounds 2, 3, and 4

compound	conditions	space group	*a* (Å)	*b* (Å)	*c* (Å)	*β* (deg)	*V* (Å^3)
GlcNAc*β*Pr (**2**), H_2O	neutron (*T* = 20 K)	$P2_1$	4.9321(1)	7.8162(1)	18.4639(2)	91.798(1)	711.44(2)
GlcNAc*β*Pr (**2**), H_2O	neutron (*T* = 120 K)	$P2_1$	4.9349(1)	7.8394(2)	18.6902(11)	91.602(2)	715.62(3)
GlcNAc*β*Pr (**2**), H_2O	neutron (*T* = 295 K)	$P2_1$	4.9382(3)	7.9039(4)	18.6902(11)	91.480(3)	729.25(7)
GlcNAc*β*Pr (**2**), H_2O[a]	X-ray (*T* = 293 K)	$P2_1$	4.946(2)	7.908(3)	18.702(6)	91.55(2)	731.2(5)
GlcNAc*β*NAc (**3**), H_2O	neutron (*T* = 20 K)	$P2_1$	4.9430(1)	7.8063(1)	17.2384(3)	91.907(2)	664.80(2)
GlcNAc*β*NAc (**3**), H_2O[b]	X-ray (*T* = 293 K)	$P2_1$	4.9409(7)	7.879(4)	17.674(4)	91.41(1)	687.8(4)
Glc*β*NAc (**4**)	neutron (*T* = 20 K)	$P2_12_12_1$	7.8012(1)	9.339(1)	14.0228(1)	90	1021.6(2)
Glc*β*NAc (**4**)[c]	X-ray (*T* = 293 K)	$P2_12_12_1$	7.8642(13)	9.423(4)	14.008(4)	90	1038(6)

[a] Cambridge Structural Database (CSD) code AVUTUG.[5] [b] CSD code CAKFAV.[11] [c] CSD code RESJEE.[12]

dx.doi.org/10.1021/ja203239j |*J. Am. Chem. Soc.* 2011, 133, 10042–10045

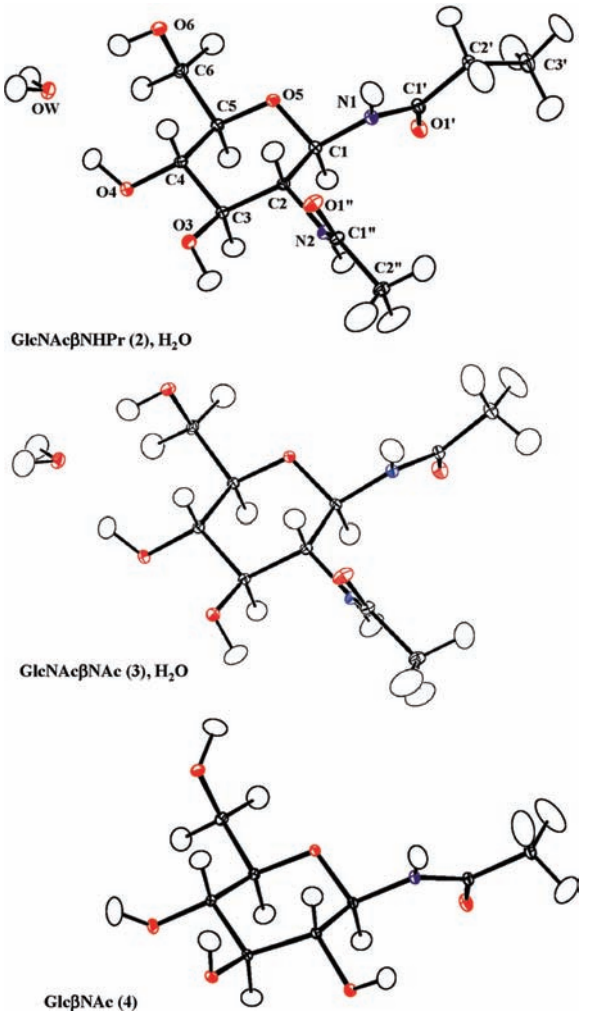

Figure 1. ORTEP representations (50% probability level) of the neutron crystal structures of compounds **2**, **3**, and **4**. Labeling of non-hydrogen atoms is given for compound **2**. PLATON software[19] was used.

on the pyranosyl rings are not pointed toward adjacent hydroxyl groups. The occurrence of such a crown of hydrogen bonds around monosaccharides has been the subject of controversy,[17,18] and the present study confirms that intermolecular contacts are favored over specific intramolecular orientations in the solid state.

A precise analysis of all of the short contacts involving hydrogen atoms was performed for **2** and **3**, which were previously described as having an unusual piling type of packing with stabilizing bifurcated hydrogen bonds.[6] Short distances are listed in Table 3, while full details of the heavy-atom distances and valence angles are available in Table S3. Compounds **2** and **3** exhibit the same packing arrangement, with columns of closely stacked carbohydrate rings assembled around water molecules. These water molecules establish a perfect helicoidal network, as each donates two hydrogens to O3 and O4 of two adjacent molecules along the *b* axis and accepts hydrogens from O4 and O6 of a third one (Figure 2A). The localizations of the hydrogen atoms could not be correctly deduced in the X-ray structure of **2** and were only partially determined for compound **3**. In both low-temperature neutron crystal structures, the four O···H distances involving each water molecule lie between 1.72 and 1.80 Å, demonstrating the special role of these water molecules in the architecture.

Table 3. Short Contacts (Hydrogen-Bonding and van der Waals) Observed in the Neutron and X-Ray Crystal Structures of 2 and 3[a]

	GlcNAcβPr (**2**), H_2O			GlcNAcβNAc (**3**), H_2O	
	neutron	neutron	X-ray[b]	neutron	X-ray[c]
T (K)	20	295	293	20	293
Piling Interactions					
H-bonding					
N1−H···O1′[d]	1.974(3)	2.032(8)	2.21	2.011(5)	2.30(2)
C2′−HA···O1′[d]				2.571(6)	2.66(4)
N2−H···O1″[e]	1.775(3)	1.792(7)	1.93	1.784(4)	1.96(3)
C1−H···O1″[e]	2.418(3)	2.460(7)	2.53	2.402(4)	2.56(2)
C3−H···O1″[e]	2.647(3)	2.696(9)	2.73(6)	2.688(4)	2.80(2)
C6−HA···O1″[f]	2.676(4)	2.80(1)	3.00	2.729(5)	2.94(3)
van der Waals					
C1−H···H−C2[e]	2.186(4)	2.206(8)	2.49	2.196(5)	2.54(3)
C4−H···H−C5[d]	2.276(4)	2.334(9)	2.49(8)	2.294(5)	2.62(4)
Other Interactions					
H-bonding					
C3′−HA···O1″[g]	2.663(4)	2.78(1)	2.87		
O3−H···O6[h]	1.753(3)	1.817(7)	2.12(9)	1.745(4)	2.03(4)
OW−H···O3[i]	1770(4)	1.788(9)	n.a.[l]	1.751(5)	1.90(4)
OW−H···O4[j]	1.735(4)	1.747(9)	1.91	1.724(5)	1.91(3)
O4−H···OW	1.760(3)	1.793(8)	n.a.[l]	1.751(5)	1.94(4)
O6−H···OW	1.798(3)	1.846(8)	2.38(9)	1.799(4)	1.94(3)
C2″−HA···O5[h]	2.534(5)	2.67(1)	2.64(8)	2.350(6)	2.54(4)
van der Waals					
C2′−HB···HC−C3′[k]	2.390(7)	2.54(3)	2.62		

[a] Distances involving hydrogen atoms are given in Å; esd values for hydrogen positions in the X-ray CSD files are given when available. [b] CSD code AVUTUG.[5] [c] CSD code CAKFAV.[11] [d] $x-1, y, z$. [e] $x+1, y, z$. [f] $x+1, y-1, z$. [g] $1-x, -{}^1/_2+y, -z$. [h] $x, y+1, z$. [i] $1-x, -{}^1/_2+y, 1-z$. [j] $2-x, -{}^1/_2+y, 1-z$. [k] $2-x, -{}^1/_2+y, -z$. [l] Water molecules are not properly oriented to participate in the hydrogen-bonding network.

Compounds **2** and **3** crystallize in the $P2_1$ space group with stacking of the carbohydrate rings along the *a* axis (Figure 2B and 2C). This packing interaction has been described as being driven by two strong N−H···O═C hydrogen bonds involving the two *N*-acetyl (or *N*-propionyl) groups. The carbonyl group of the *N*-acetyl moiety at C2 also appears to be involved in additional weak hydrogen bonds by accepting hydrogen from carbons C1 and C3 of the adjacent ring. At 20 K, the H···O distances are 2.42 and 2.65 Å for **2** and 2.40 and 2.69 Å for **3**. Such distances are shorter than the sum of van der Waals contacts and represent weak hydrogen bonds in the classification of Desiraju and Steiner.[20] The CO group of each *N*-acetyl at C2 establishes a trifurcated hydrogen bond with the neighboring molecules, thereby stabilizing the stacking of rings. These trifurcated H-bonds consist of one conventional and two nonconventional H-bonds, and the C−H···O contacts having the longest distances are probably very weak. It should be noted that the assignment of C3−H···O═C hydrogen bonds was not obvious at room temperature; indeed, the added effects of errors in C−H distances and lattice expansion at higher temperature resulted in apparent distances on the order of 2.70−2.80 Å.

The correct localization of hydrogen atoms in neutron crystal structures also helps in demonstrating the importance of hydrophobic

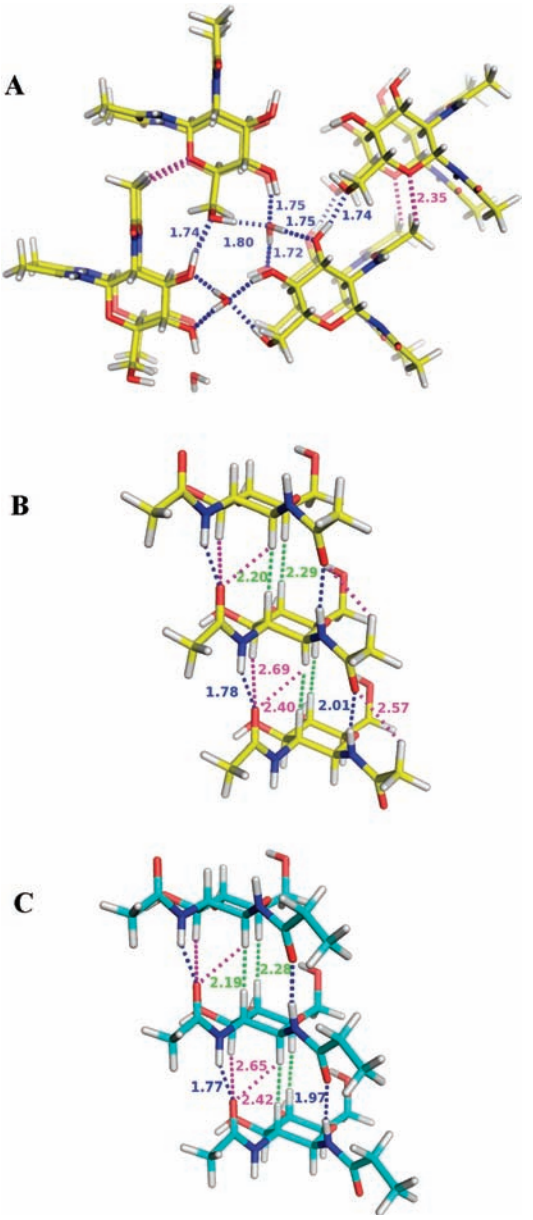

Figure 2. Representations of selected hydrogen bonds and short contacts in the neutron crystal structures of compound **3** (yellow) and **2** (cyan): (A) view of **3** along the *b* axis; (B) view of **3** along the *a* axis; (C) view of **2** along the *a* axis. Dotted lines show classical OH···O or NH···O hydrogen bonds (blue), weak CH···O hydrogen bonds (pink) and van der Waals contacts (green).

contacts (e.g., the two short contacts involving the pairs C1−H/C2−H and C4−H/C5−H). Very short H···H distances between 2.20 and 2.30 Å were observed in the low-temperature neutron diffraction structure and are listed as van der Waals contacts in Table 3. With a distance larger than 2.5 Å, they could not be assigned on the basis of the X-ray structures measured at room temperature.

Besides yielding new information relevant to the field of structural glycobiology, the present work illustrates how the accurate X—H distances obtained using neutrons may then be used, for example, in generating more precise and useful experimental structures. X-ray diffraction data, in contrast to neutron density data, cannot provide anisotropic displacement parameters for H atoms, a major problem in charge-density analysis of molecular crystals.[21] Furthermore, neutron diffraction bond lengths and statistical mean values are classically used in parametrization of molecular mechanics force fields. High-quality data are also of prime interest for testing computational methods, and the neutron structure of glucose has been utilized for this.[22] The present structures will therefore be of interest for testing and improving force fields in order to develop better tools for the modeling of *N*-glycoproteins.

■ ASSOCIATED CONTENT

Supporting Information. Data collection and structure refinement details and tables of structural parameters. This material is available free of charge via the Internet at http://pubs.acs.org. Crystal structures have been deposited at the Cambridge Crystallographic Data Centre with deposition numbers CCDC 820947−820951.

■ AUTHOR INFORMATION

Corresponding Author

*loganath@iitm.ac.in; mason@ill.eu; serge.perez@esrf.eu; anne.imberty@cermav.cnrs.fr

■ ACKNOWLEDGMENT

We are grateful to the Indo-French Centre for Promotion of Advanced Research (IFCPAR) for financial support.

■ REFERENCES

(1) Dwek, R. A. *Chem. Rev.* **1996**, *96*, 683.
(2) Spiro, R. G. *Glycobiology* **2002**, *12*, 43R.
(3) Delbaere, L. T. *Biochem. J.* **1974**, *143*, 197.
(4) Ohanessian, J.; Avenel, D.; Neuman, A.; Gillier-Pandraud, H. *Carbohydr. Res.* **1980**, *80*, 1.
(5) Lakshmanan, T.; Sriram, D.; Priya, K.; Loganathan, D. *Biochem. Biophys. Res. Commun.* **2003**, *312*, 405.
(6) Loganathan, D.; Aich, U. *Glycobiology* **2006**, *16*, 343.
(7) Allen, F. H.; Bruno, I. J. *Acta Crystallogr.* **2010**, *B66*, 380.
(8) Jeffrey, G. A.; Takagi, S. *Acc. Chem. Res.* **1978**, *11*, 264.
(9) Steiner, T.; Saenger, W. *J. Am. Chem. Soc.* **1992**, *114*, 10146.
(10) Pérez, S.; Gautier, C.; Imberty, A. In *Oligosaccharides in Chemistry and Biology: A Comprehensive Handbook*; Ernst, B., Hart, G., Sinay, P., Eds.; Wiley-VCH: Weinheim, Germany, 2000; pp 969 ff.
(11) Sriram, D.; Lakshmanan, T.; Loganathan, D.; Srinivasan, S. *Carbohydr. Res.* **1998**, *309*, 227.
(12) Sriram, D.; Sreenivasan, H.; Srinivasan, S.; Priya, K.; Vishnu Thirtha, M.; Loganathan, D. *Acta Crystallogr.* **1997**, *C53*, 1075.
(13) Wilkinson, C.; Khamis, H. W.; Stansfield, R. D.; McIntyre, G. J. *J. Appl. Crystallogr.* **1988**, 21.
(14) Sheldrick, G. M. *SHELXL-97: Program for Crystal Structure Refinement*; University of Göttingen: Göttingen, Germany, 1997.
(15) Mohamed Naseer Ali, M.; Aich, U.; Pérez, S.; Imberty, A.; Loganathan, D. *Carbohydr. Res.* **2009**, *344*, 355.
(16) Mohamed Naseer Ali, M.; Aich, U.; Varghese, B.; Pérez, S.; Imberty, A.; Loganathan, D. *J. Am. Chem. Soc.* **2008**, *130*, 8317.
(17) French, A. D.; Csonka, G. I. *Cellulose* [Online early access]. DOI: 10.1007/s10570-011-9539-6. Published Online: April 11, 2011.
(18) Klein, R. A. *J. Am. Chem. Soc.* **2002**, *124*, 13931.
(19) Spek, A. L. *Acta Crystallogr.* **2009**, *D65*, 148.
(20) Desiraju, G. R.; Steiner, T. *The Weak Hydrogen Bond in Structural Chemistry and Biology*; Oxford University Press: New York, 1999.
(21) Whitten, A. E.; Spackman, M. A. *Acta Crystallogr.* **2006**, *B62*, 875.
(22) Kouwijzer, M. L. C. E.; Eijck, B. P. V.; Kroes, S. J.; Kroon, J. *J. Comput. Chem.* **1993**, *14*, 1281.