STRUCTURAL CHEMISTRY

ISSN: 2053-2296
journals.iucr.org/c

Glycosidic linkage, *N*-acetyl side-chain, and other structural properties of methyl 2-acetamido-2-deoxy-β-D-gluco-pyranosyl-(1$\rightarrow$4)-β-D-mannopyranoside monohydrate and related compounds

Wenhui Zhang, Reagan J. Meredith, Allen G. Oliver, Ian Carmichael and Anthony S. Serianni

Acta Cryst. (2020). C**76**, 287–297

Acta Cryst C STRUCTURAL CHEMISTRY
ISSN 2053-2296

Glycosidic linkage, *N*-acetyl side-chain, and other structural properties of methyl 2-acetamido-2-deoxy-β-D-glucopyranosyl-(1→4)-β-D-mannopyranoside monohydrate and related compounds

Wenhui Zhang,[a] Reagan J. Meredith,[a] Allen G. Oliver,[a] Ian Carmichael[b] and Anthony S. Serianni[a]*

[a]Department of Chemistry & Biochemistry, University of Notre Dame, Notre Dame, IN 46556-5670, USA, and [b]Radiation Laboratory, University of Notre Dame, Notre Dame, IN 46556-5670, USA. *Correspondence e-mail: aseriann@nd.edu

Received 16 October 2019
Accepted 3 February 2020

Edited by R. I. Cooper, University of Oxford, UK

Keywords: GlcNAc; glycopyranoside; mannopyranoside; DFT; disaccharide; crystal structure; glycans.

CCDC reference: 1981654

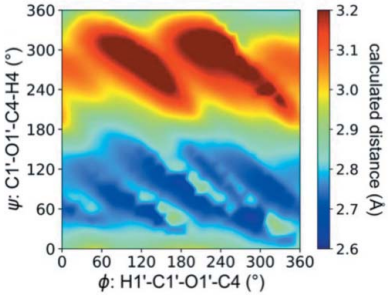
Supporting information: this article has supporting information at journals.iucr.org/c

The crystal structure of methyl 2-acetamido-2-deoxy-β-D-glycopyranosyl-(1→4)-β-D-mannopyranoside monohydrate, $C_{15}H_{27}NO_{11}\cdot H_2O$, was determined and its structural properties compared to those in a set of mono- and disaccharides bearing *N*-acetyl side-chains in βGlcNAc aldohexopyranosyl rings. Valence bond angles and torsion angles in these side chains are relatively uniform, but C—N (amide) and C—O (carbonyl) bond lengths depend on the state of hydrogen bonding to the carbonyl O atom and N—H hydrogen. Relative to *N*-acetyl side chains devoid of hydrogen bonding, those in which the carbonyl O atom serves as a hydrogen-bond acceptor display elongated C—O and shortened C—N bonds. This behavior is reproduced by density functional theory (DFT) calculations, indicating that the relative contributions of amide resonance forms to experimental C—N and C—O bond lengths depend on the solvation state, leading to expectations that activation barriers to amide *cis–trans* isomerization will depend on the polarity of the environment. DFT calculations also revealed useful predictive information on the dependencies of inter-residue hydrogen bonding and some bond angles in or proximal to β-(1→4) *O*-glycosidic linkages on linkage torsion angles ϕ and ψ. Hypersurfaces correlating ϕ and ψ with the linkage C—O—C bond angle and total energy are sufficiently similar to render the former a proxy of the latter.

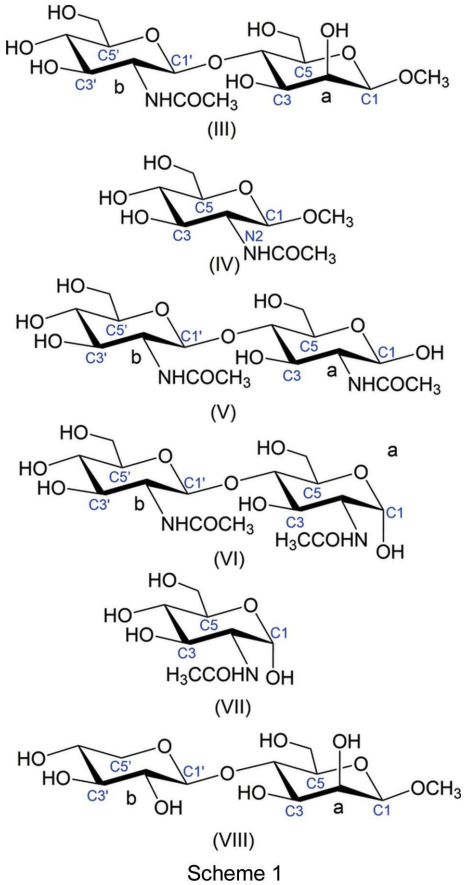
1. Introduction

N-Acetyl-D-glucosamine (GlcNAc) is a key constituent of many biologically important *N*- and *O*-glycans in eukaryotic cells (Rudd & Dwek, 1997). This 2-aminosugar contributes to the pentasaccharide core structure of *N*-glycans and to its elaboration to give complex and hybrid *N*-glycans (Fig. 1*a*) (Johansen *et al.*, 1961). Proteins that are *O*-glycosylated often contain GlcNAc in glycosidic linkage to other residues, such as found in the βGlcUA-(1→3)-βGlcNAc disaccharide repeating unit of hyaluronic acid (Kaye & Stacey, 1951). Post-translational mono-GlcNAcylation affects protein biological function, often occurring in crosstalk with *O*-phosphorylation (Yang *et al.*, 2006). The interplay between *O*-GlcNAcylation and *O*-phosphorylation is a key mechanism in the biological control of protein function (Hart *et al.*, 2011).

The co-translational introduction of *N*-glycans into proteins involves the initial insertion of a common precursor, $Glc_3Man_9GlcNAc_2$, *via* a dolichol bisphosphate donor. Subsequent restructuring of this 14-mer occurs partly in the endoplasmic reticulum (ER), but mostly in the Golgi apparatus. A series of glycosyltransferases and glycosidases, presumably distributed asymmetrically in the *cis*, *medial*, and *trans* Golgi network,



remodel the parent *N*-glycan into high-mannose, complex-type, and hybrid *N*-glycans. Of interest to the present work are the *N*-acetylglucosaminyltransferases (GNTs) required to assemble complex-type and hybrid chains, the most common of which are GNT-I, GNT-II, and GNT-III. The former two enzymes attach GlcNAc residues in a β-(1→2) linkage to αMan residues 4 and 4′ of the core pentasaccharide (Fig. 1*a*), respectively, thereby allowing further elongation of both arms with βGal and *N*-acetyl-α-neuraminic acid (Neu5Ac) to give mature complex-type *N*-glycans (Geisler & Jarvis, 2012). GNT-III catalyzes the addition of GlcNAc in a β-(1→4) linkage to βMan residue 3 [see (I) and (II) in Fig. 1] (Takahashi *et al.*, 2009; Nagae *et al.*, 2013). The introduction of β-(1→4)-linked GlcNAc apparently prevents further elongation of the *N*-glycan at GlcNAc 5 and 5′, resulting in truncated chains. GNT-III-mediated chain modification appears to confer new properties to the proteins that bears it, including effects on tumor metastasis and human longevity (Takahashi *et al.*, 2009; Ruhaak *et al.*, 2010; Andre *et al.*, 2007; Sultan *et al.*, 1997).



Scheme 1

The introduction of βGlcNAc at O4 of βMan 3 produces a 3,4,5-trisubstituted βMan residue, leading to increased steric interactions and the possibility of conformational distortion of one or more of the four *O*-glycosidic linkages involving βMan 3. These conformational effects have recently been investigated by Yamaguchi and co-workers (Hanashima *et al.*, 2018). The effects of structural context on linkage conformation are best evaluated using reference disaccharides in which context effects are presumed to be small or zero (Zhang *et al.*, 2017). Knowledge of the conformational properties of reference disaccharides permits a quantitative determination of whether context effects are significant, achieved by comparing the behavior of a reference disaccharide to that of the same disaccharide embedded in a larger oligosaccharide. While the natural inclination is to anticipate changes in linkage conformation as a function of context, other saccharide conformational elements may also be affected, such as the preferred geometries of hydroxy groups and other exocyclic appendages. An impediment to quantifying these effects has been the lack of a method that provides continuous models of conformational elements based exclusively on experimental data. This obstacle has been recently overcome with the development of *MA'AT* analysis, which uses circular statistics and redundant NMR *J*-couplings to derive single- and multi-state models of conformational behaviors in solution (Zhang *et al.*, 2017, 2019).

To characterize saccharide structures from a conformational standpoint, knowledge of their X-ray structures can be beneficial, if only to compare crystal conformations with those observed in solution to determine whether solvation (*e.g.* by explicit solvent water or by the crystalline lattice itself) affects conformational preference. In prior work, a group of biologically relevant disaccharides has been studied by crystallography, and the present work extends this effort to the βGlcNAc-(1→4)-βMan linkage shown in (II) (Fig. 1*b*). Methyl 2-acetamido-2-deoxy-β-D-glucopyranosyl-(1→4)-β-D-mannopyranoside, (III) [Scheme 1; the scheme shows the conventional atom numbering and identifies the **a** and **b** residues of (III), (V), (VII), and (VIII)], was prepared by chemical synthesis and crystallized, and its low-temperature crystal structure determined. The structural parameters observed in crystalline (III) are compared to those in methyl 2-acetamido-2-deoxy-β-D-glucopyranoside, (IV) (Hu *et al.*, 2011), 2-acetamido-2-deoxy-β-D-glucopyranosyl-(1→4)-2-acetamido-2-deoxy-β-D-glucopyranose, (V) (Mo, 1979), 2-acetamido-2-deoxy-β-D-glucopyranosyl-(1→4)-2-acetamido-2-deoxy-α-D-glucopyranose, (VI) (Mo & Jensen, 1978), 2-acetamido-2-deoxy-β-D-glucopyranose, (VII) (Mo & Jensen, 1975), and methyl β-D-xylopyranosyl-(1→4)-β-D-mannopyranoside, (VIII) (Zhang *et al.*, 2013).

2. Experimental

2.1. Chemical synthesis and crystallization of (III) (Fig. 2)

2.1.1. Methyl 3,6-di-*O*-benzyl-β-D-mannopyranoside (2). Methyl β-D-mannopyranoside (**1**) (5.90 g, 30.4 mmol) and di-*n*-butyltin oxide (17.0 g, 68.3 mmol) were added to anhydrous toluene (60 ml). After stirring at 373 K for 3 h, the reaction mixture was concentrated to 30 ml, and benzyl bromide (20 ml, 168 mmol) and tetrabutylammonium bromide (5.00 g, 15.5 mmol) were added. The resulting mixture was stirred at 373 K for an additional 20 h, and then concentrated *in vacuo*. The residue was dissolved in ethyl acetate, washed with distilled water, dried over anhydrous Na_2SO_4, and concen-

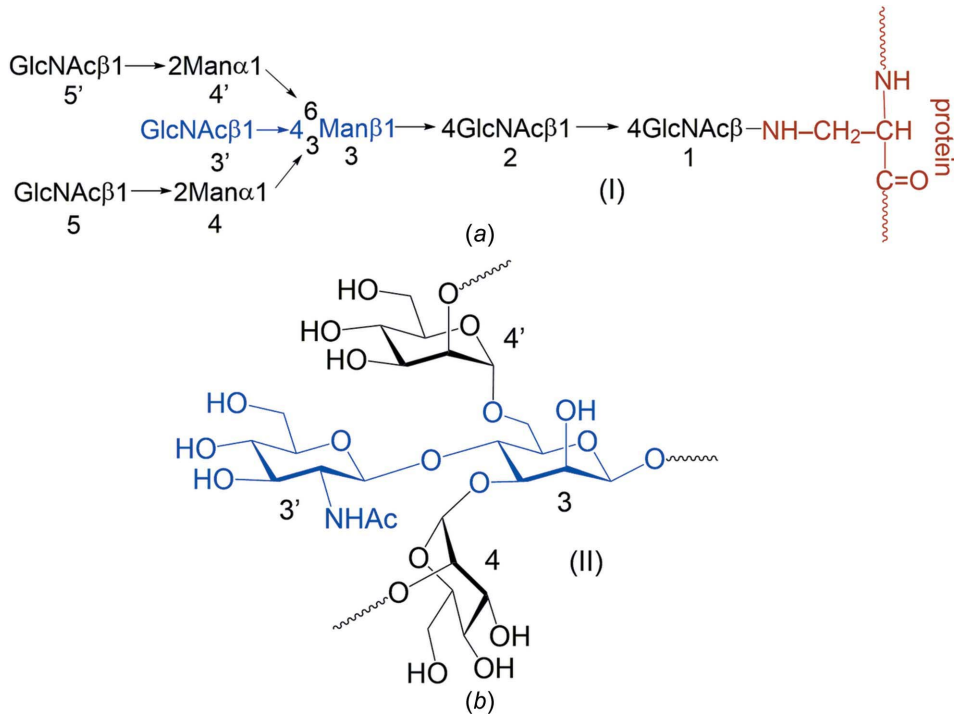


Figure 1
(*a*) The structure of a complex-type *N*-glycan, (I), containing a bisecting β-(1→4)-linked GlcNAc to βMan (disaccharide fragment highlighted in blue) of $Man_3GlcNAc_2$. (*b*) Tetrasaccharide subfragment of (I) containing residues 3, 3′, 3 and 4′, showing the molecular structure in the vicinity of the bisecting βGlcNAc.

trated to a syrup, which was purified by flash chromatography on silica gel to afford glycoside **2** (yield: 7.50 g, 20.1 mmol, 66%) (Qin & Grindley, 1994). In this and all of the following steps, flash column chromatography on silica gel (preparative scale) was performed on a Reveleris X2 flash chromatography system. A mixture of hexanes and ethyl acetate was used as the eluent.

2.1.2. Methyl 2-*O*-benzoyl-3,6-di-*O*-benzyl-β-D-manno-pyranoside (3). Compound **2** (7.50 g, 20.1 mmol) was dissolved in anhydrous toluene (60 ml) and anhydrous pyridine (8 ml) was added. Benzoyl chloride (2.50 ml, 21.5 mmol) was then added dropwise at 273 K and the reaction mixture was stirred at 273 K for 2 h. The mixture was then evaporated to dryness and purified by flash chromatography on silica gel, affording product **3** (yield: 8.40 g, 17.6 mmol, 87%).

2.1.3. 2-Deoxy-2-phthalimido-3,4,6-tri-*O*-acetyl-α-D-gluco-pyranosyl trichloroacetimidate (7). D-Glucosamine hydrochloride (**4**) (6.33 g, 29.2 mmol), Na_2CO_3 (3.10 g, 29.2 mmol), and phthalic anhydride (4.32 g, 29.2 mmol) were added to water (38 ml). The mixture was stirred at room temperature overnight and concentrated to dryness. The residue was dissolved in pyridine (100 ml), and Ac_2O (40 ml, 423 mmol) was added. After stirring at room temperature for 12 h, the mixture was concentrated at 313 K *in vacuo*. The residue was dissolved in ethyl acetate, washed with distilled water, dried over anhydrous Na_2SO_4, and concentrated to give compound

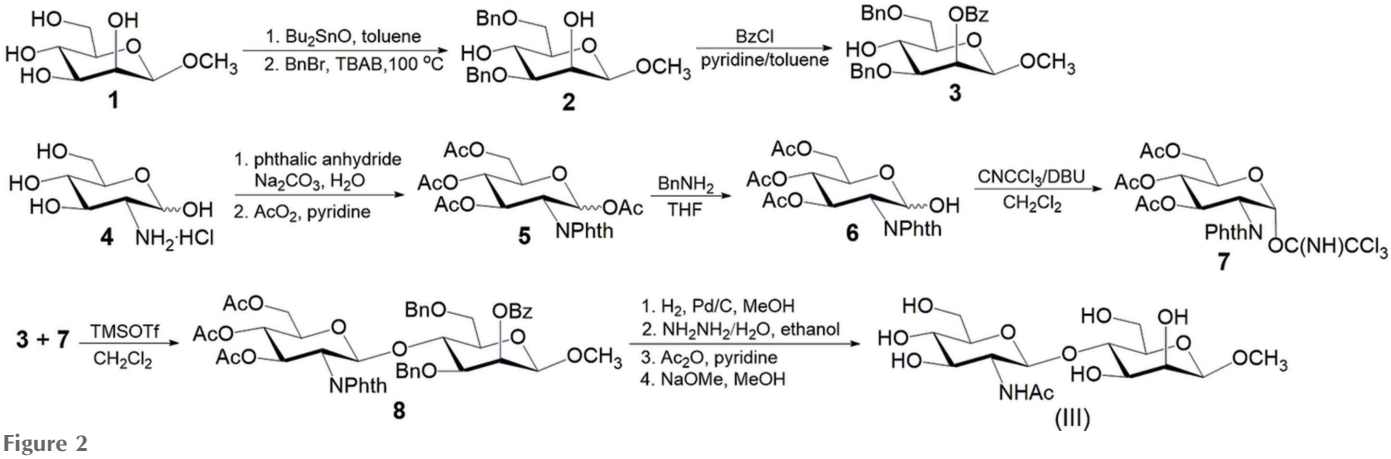


Figure 2
Synthetic route used to prepare disaccharide (III).

Table 1
N-Acetyl side-chain bond lengths calculated in different models of (IV) using density functional theory.

Model[a]	Number of hydrogen bonds[c]			C—N bond length (Å)			C—O bond length (Å)		
	Vacuum	MeOH[e]	IEFPCM[f]	Vacuum	MeOH	IEFPCM	Vacuum	MeOH	IEFPCM
1	0	0	0	1.373	1.359	1.359	1.218	1.231	1.232
2	3[f]	3[f]	3[f]	1.351[g]	1.341[g]	1.342	1.244[g]	1.250[g]	1.25
3	2	2	2	1.352	1.347	1.347	1.238	1.246	1.246
4	1	1	1	1.364	1.354	1.353[g]	1.228	1.237	1.238[g]
5	1	1	1	1.359	1.353	1.353	1.23	1.239	1.24
6	1[f]	1[f]	1[f]	1.368	1.356	1.356	1.22	1.234	1.234
7	1.5[f]	2[f]	2[f]	1.359	1.349	1.349	1.232	1.241	1.241
8	2[f]	2[f]	2[f]	1.359[g]	1.349	1.349	1.233[g]	1.243	1.243
9[b]	0	0	0	1.314	1.305	1.305	1.301	1.304	1.304

Notes: (*a*) *in silico* models of (IV) differ in the number of explicit 'solvent' water and/or methanol molecules included during geometry optimization. (*b*) Model was protonated at the carbonyl O atom. (*c*) After geometry optimization. (*d*) Calculations were conducted using self-consistent reaction field (SCRF) with implicit methanol solvent. (*e*) Calculations included effects of solvent water, which were treated using the SCRF and Integral Equation Formalism (polarizable continuum) model (IEFPCM); see *Calculation methods* (§2.2). (*f*) Includes hydrogen bond to acceptor water from the N—H hydrogen. (*g*) Geometry optimization conducted with C1—C2—N2—CO torsion angle in (IV) (see Scheme 1) constrained at 112°.

5. Compound **5** was dissolved in tetrahydrofuran (THF, 150 ml), the resulting solution was cooled in an ice bath, and benzylamine (3.82 ml, 35.0 mmol) was added. After stirring for 4 h at room temperature, the THF was removed *in vacuo*. The residue was dissolved in ethyl acetate, washed with 1 *N* aqueous HCl solution, saturated aqueous $NaHCO_3$ solution, and distilled water sequentially, and then dried over anhydrous Na_2SO_4. After concentration, crystallization from an ethyl acetate/hexane (~3:1 *v*/*v*) mixed solvent afforded pure compound **6** (yield: 8.00 g, 18.4 mmol, 63%). Compound **6** (3.00 g, 6.89 mmol) was dissolved in CH_2Cl_2 (30 ml), and trichloroacetonitrile (2.00 ml, 19.9 mmol) and several drops of 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU) were added. The reaction solution was stirred for 3 h at room temperature and concentrated *in vacuo*. Flash chromatography on silica gel gave trichloroacetimidate **7** (yield: 3.10 g, 5.35 mmol, 78%) (Liu & Wei, 2012; Schmidt & Michel, 1985).

2.1.4. Methyl 2-deoxy-2-phthalimido-3,4,6-tri-*O*-acetyl-*β*-D-glucopyranosyl-(1→4)-2-*O*-benzoyl-3,6-di-*O*-benzyl-*β*-D-mannopyranoside (8). A mixture of trichloroacetimidate **7** (750 mg, 1.30 mmol), methyl glycoside **3** (480 mg, 1.00 mmol), and molecular sieves (4 Å, 2.0 g) was dried under high vacuum, and then anhydrous CH_2Cl_2 (20 ml) was added. The solution was cooled to 273 K and treated with TMSOTf (20 ml, 0.11 mmol) under N_2. After 2 h, the reaction was quenched with the addition of a few drops of triethylamine and the molecular sieves were removed by filtration. The solution was concentrated to a syrup *in vacuo*, and the residue was purified by flash chromatography on silica gel to afford disaccharide **8** (yield: 790 mg, 0.88 mmol, 88%).

2.1.5. Methyl 2-acetamido-2-deoxy-*β*-D-glucopyranosyl-(1→4)-*β*-D-mannopyranoside, (III). Compound **8** (300 mg, 335 mmol) was dissolved in methanol (15 ml) and treated with Pd/C (10%, 100 mg) and H_2 overnight. The Pd/C catalyst was removed by filtration and the filtrate was concentrated to dryness *in vacuo*. The residue was dissolved in ethanol (20 ml) and hydrazine hydrate (3.00 ml) was added. After refluxing for 20 h, the reaction mixture was concentrated *in vacuo* to a syrup, which was dried under high vacuum. The dried residue was dissolved in pyridine (20 ml) and Ac_2O (4.00 ml, 42.3 mmol) was added. The mixture was stirred at room temperature overnight and concentrated *in vacuo*. After concentrating *in vacuo*, the residue was purified by flash chromatography on silica gel to give an acetylated disaccharide, which was treated with sodium methoxide in methanol (20 ml, pH > 10) overnight. Final product (III) was purified by crystallization from a mixture of water and methanol (yield: 90 mg, 230 mmol, 69%). Disaccharide (III) was recrystallized from water, characterized by ^{1}H and ^{13}C NMR spectroscopy (see Table S1 in the supporting information for ^{1}H and ^{13}C chemical shifts, and Figs. S1 and S2 for ^{1}H and ^{13}C{^{1}H} NMR spectra), and used in the X-ray structure analysis.

2.2. Calculation methods

Density functional theory (DFT) calculations were performed to investigate the effect of hydrogen bonding on the C—N and C—O bond lengths in the *N*-acetyl side chain of methyl 2-acetamido-2-deoxy-*β*-D-glucopyranoside, (IV). The calculations were performed in two steps to ensure accurate positions of explicit solvent water molecules with respect to (IV). The first step utilized molecular dynamics (MD) simulations to minimize the energy of the solvated molecule. The second step employed density functional theory (DFT) to fully optimize the monosaccharide with explicit solvent water molecules.

2.2.1. Solvation and initial energy minimization. The initial structure of (IV) was built using the *Carbohydrate Builder* module available at the *GLYCAM* web site (http://www.glycam.org) (Woods Group, 2019). The structure was solvated with TIP3P (Jorgensen *et al.*, 1983) water using a 10 Å buffer in a cubic box and the LEaP module in the *AMBER18* software package (Case *et al.*, 2018). The solvated structure then underwent energy minimization under constant volume (500 steps steepest descent, followed by 24500 steps of conjugate-gradient minimization). The *GLYCAM06* (Version j) force field (Kirschner *et al.*, 2008) was employed for energy minimization. The Cartesian coordinates of the minimized structure and solvent were extracted and examined using the molecular viewing software *Chem3D* (PerkinElmer Informatics). Only the three solvent water molecules that were

hydrogen bonded to the carbonyl O atom or amide H atom were retained in subsequent calculations.

2.2.2. Geometry optimizations. DFT calculations were conducted within *GAUSSIAN16* (Frisch *et al.*, 2016) using the B3LYP functional (Becke, 1993) and 6-311++G(d,p) basis set (McLean & Chandler, 1980; Frisch *et al.*, 1984) for geometry optimizations. Two sets of calculations were conducted, *i.e. in vacuo* and with the inclusion of implicit solvent modeling. Two implicit solvent models were employed, water and methanol, both using the self-consistent reaction field (SCRF) (Cances *et al.*, 1997) and the integral equation formalism (polarizable continuum) model (IEFPCM) (Cammi *et al.*, 2000). In all calculations, the C1—C2—N2—CO torsion angle in the *N*-acetyl side chain was initially set to 113° and allowed to optimize freely unless otherwise indicated. Geometry optimizations were conducted on nine models of (IV) that differed in the number of explicit solvent water molecules (Table 1) that were hydrogen bonded to the side chain. The C—N and C—O bond lengths were extracted from the optimized geometries using *Python* (Jones *et al.*, 2014).

2.3. Refinement

Crystal data, data collection, and structure refinement details are summarized in Table 2. The structure of (III) was solved using dual-space methods (Sheldrick, 2015*a*) and refined routinely. H atoms bonded to C atoms were included in idealized positions riding on the C atom to which they are bonded, with C—H = 1.00 (methine), 0.99 (methyl), or 0.98 Å (methylene), and U_{iso}(H) = 1.2U_{eq}(C) (methine and methylene) or 1.5U_{eq}(C) (methyl). Methyl H atoms were treated as a rotating model to minimize the electron-density contribution from these atoms. Hydroxy, water, and amide H atoms were initially located from a difference Fourier map. In the final model, hydroxy and water H atoms were refined as a rotating model to minimize their electron-density contribution, with U_{iso}(H) = 1.5U_{eq}(O) and O—H = 0.84 (hydroxy) or 0.85 Å (water). The amide H atom was refined freely. This model was adopted after some of the hydroxy H atoms refined unreasonably when treated as freely refined atoms.

Table 2
Experimental details.

Crystal data	
Chemical formula	$C_{15}H_{27}NO_{11}\cdot H_2O$
M_r	415.39
Crystal system, space group	Orthorhombic, $P2_12_12_1$
Temperature (K)	120
a, b, c (Å)	11.4170 (6), 12.6782 (6), 13.2102 (7)
V (Å^3)	1912.14 (17)
Z	4
Radiation type	Cu $K\alpha$
μ (mm^{-1})	1.09
Crystal size (mm)	0.13 × 0.11 × 0.10
Data collection	
Diffractometer	Bruker PHOTON-II
Absorption correction	Numerical (*SADABS*; Krause *et al.*, 2015)
T_{min}, T_{max}	0.863, 1.000
No. of measured, independent and observed [$I > 2\sigma(I)$] reflections	31774, 3635, 3510
R_{int}	0.047
(sin $\theta/\lambda)_{max}$ (Å^{-1})	0.614
Refinement	
$R[F^2 > 2\sigma(F^2)]$, $wR(F^2)$, S	0.044, 0.117, 1.05
No. of reflections	3635
No. of parameters	266
H-atom treatment	H atoms treated by a mixture of independent and constrained refinement
$\Delta\rho_{max}$, $\Delta\rho_{min}$ (e Å^{-3})	0.54, −0.38
Absolute structure	Flack x determined using 1469 quotients $[(I^+) - (I^-)]/[(I^+) + (I^-)]$ (Parsons *et al.*, 2013)
Absolute structure parameter	−0.07 (7)

Computer programs: *APEX3* (Bruker, 2015), *SAINT* (Bruker, 2015), *SHELXT2018* (Sheldrick, 2015*a*), *SHELXL2018* (Sheldrick, 2015*b*), *Mercury* (Macrae *et al.*, 2020), *CIFTAB* (Sheldrick, 2015*b*), *PLATON* (Spek, 2009) and *publCIF* (Westrip, 2010).

3. Results and discussion

3.1. General structural features and crystal packing description of (III)

The crystal structure of (III) (Fig. 3) reveals the βGlcNAc and βMan aldohexopyranosyl rings in $^4C_1$ conformations, and an internal *O*-glycosidic linkage characterized by *phi* (ϕ; H1′—C1′—O1′—C4) and *psi* (ψ; C1′—O1′—C4—H4) torsion angles of 42.9 and −15.5°, respectively. The exocyclic hydroxymethyl (CH_2OH) groups appended to the βGlcNAc and βMan rings adopt *gg* conformations characterized by O5′/O5—C5′/C5—C6′/C6—O6′/O6 torsion angles of −56.3 and −70.1°, respectively. An inter-residue hydrogen bond between O3H (donor) and O5′ is observed, characterized by an O3···O5′ internuclear distance of 2.71 Å. The *N*-acetyl side chain in the βGlcNAc residue adopts a *trans* configuration (C2′—N2′—C_{car}—CH_3 torsion angle of 177.1°), and the conformation about the C2′—N2′ bond positions the C_{car}=O_{car} and C2′—H2′ bonds (pseudo-torsion O7′—C7′···C2′—H2′; see Fig. 3) in a near-eclipsed orientation (see Table 2 for refinement details).

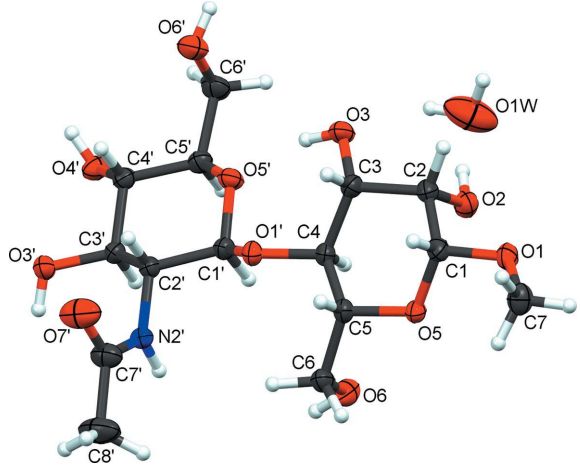


Figure 3
The labeling scheme for (III). Displacement ellipsoids are depicted at the 50% probability level. H atoms are shown as spheres of arbitrary radii.

All of the hydroxy groups in (III) are involved in hydrogen bonding, as are both water H atoms. Surprisingly, the second water H atom that was located from a difference Fourier map is not oriented towards nearby acceptor atoms, but rather into a void space within the lattice. As described above, O3 forms an intramolecular hydrogen bond to O5′; all other hydrogen-bond interactions are intermolecular. Glycosidic linkage atom O1′ and the methoxy atom O1 do not participate in hydrogen bonding. Atom O1 is spatially close to a neighboring hydroxy O6 atom, however, the contact is long and would be bifurcated with the O6···O2 hydrogen bond (see Table 3 for details). Ring atom O5 serves as an acceptor in a hydrogen bond from N2′ (N2′···O5^{i}; for symmetry codes, see Table 3; the relationship is by a screw axis along *b*). This hydrogen bond is bifurcated with O2^{i}. Hydroxy atom O4 only donates a hydrogen bond to O3^{ii} (a second screw axis parallel to the *b* axis) and O6′ only donates to O3′^{iii} (a third, unique, screw axis along *b*). All other hydroxy groups are both donors and acceptors. Amide atom O7′ is an acceptor in a hydrogen bond from the water of crystallization (O1*W*···O7′^{iv}; related by a screw axis along the crystallographic *c* axis). The water molecule is also a donor in a hydrogen bond to O6′^{v}. In turn, the water molecule serves as an acceptor in a hydrogen bond from O2. Hydroxy atom O6 has a strong hydrogen bond to O2^{i}, that is also slightly bifurcated and much weaker to O1^{i}. Atom O6 is the acceptor in a hydrogen bond from O3′ (O3′···O6^{i}). The overall motif is a three-dimensional hydrogen-bonded network (Fig. 4). Graph-set analysis of the strutcure reveals that the hydrogen bonds involving the water molecule are both *D* (discrete) contacts, the O3···O5′ interaction is an *S* (self-interacting) contact and all other local contacts are *C* (chain) hydrogen bonds (Etter *et al.*, 1990).

The following discussion treats structurally related βGlcNAc residues at the nonreducing ends of (III), (V), and (VI), and the βGlcNAc rings (IV_A) and (IV_B). GlcNAc residues at the reducing ends of (V) and (VII), and the αGlcNAc ring in (VII) are treated when anomeric configuration and related structural factors assist the discussion. Structural properties of the βMan ring of (III) (residue **a**) (see Scheme 1 for the identification of residues **a** and **b**) are compared to those of the βMan ring of (VIII) (residue **a**). Structural parameters in (III)–(VIII), which are summarized in Table S2 in the supporting information, informed the following discussion.

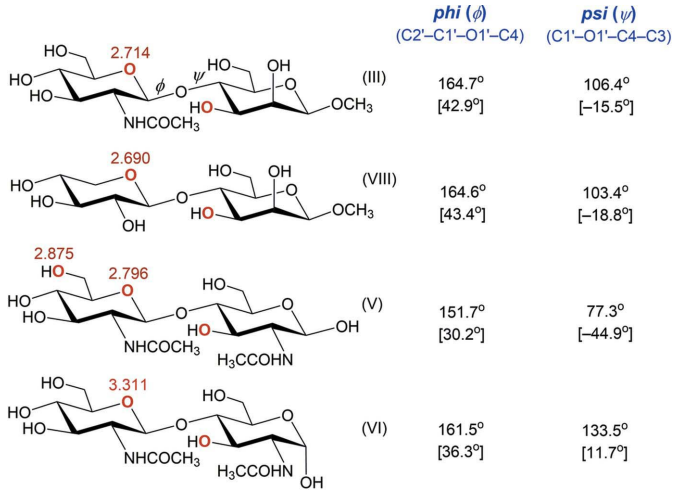


Figure 5
Phi (ϕ) and *psi* (ψ) glycosidic torsion angles in (III), (V), (VI), and (VIII) in their crystal structures, defined as C2′—C1′—O1′—C4 and C1′—O1′—C4—C3, respectively. Torsion angles for the conventional definitions of ϕ (H1′—C1′—O1′—C4) and ψ (C1′—O1′—C4—H4) are shown in brackets. Values shown in red are the O3···O5′ and O3···O6′ internuclear distances (in Å; atoms O3, O5′ and O6′ are shown in bold red) observed in the crystal structures. The atom numbering in these disaccharides is identical to the numbering shown in Scheme 1.

3.2. *O*-Glycosidic linkage conformations in (III), (V), (VI), and (VIII), and inter-residue hydrogen bonding

The conformation of the internal *O*-glycosidic linkage in (III) resembles that in (VIII), as reflected by the virtually

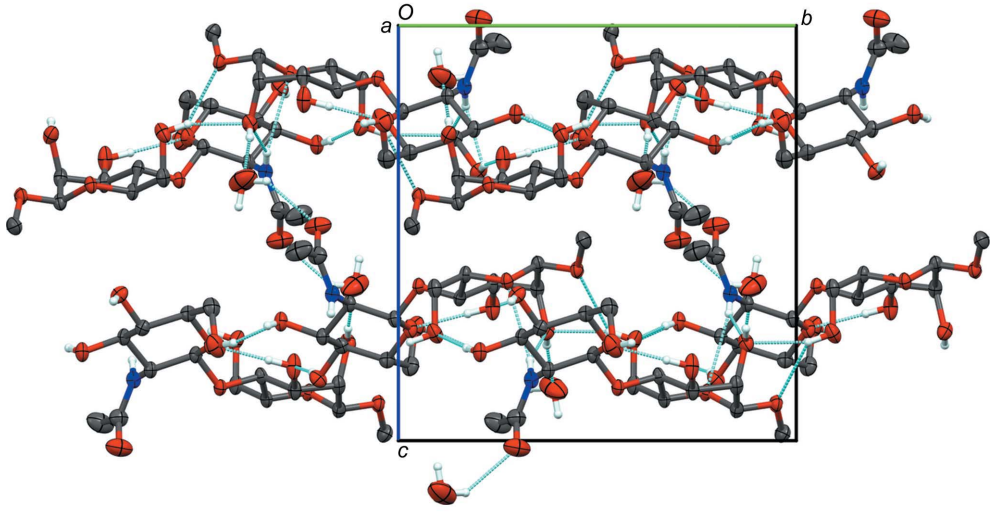


Figure 4
Packing diagram of (III), viewed along the *a* axis. C—H hydrogens have been omitted for clarity. Hydrogen bonds are depicted as blue dashed lines.

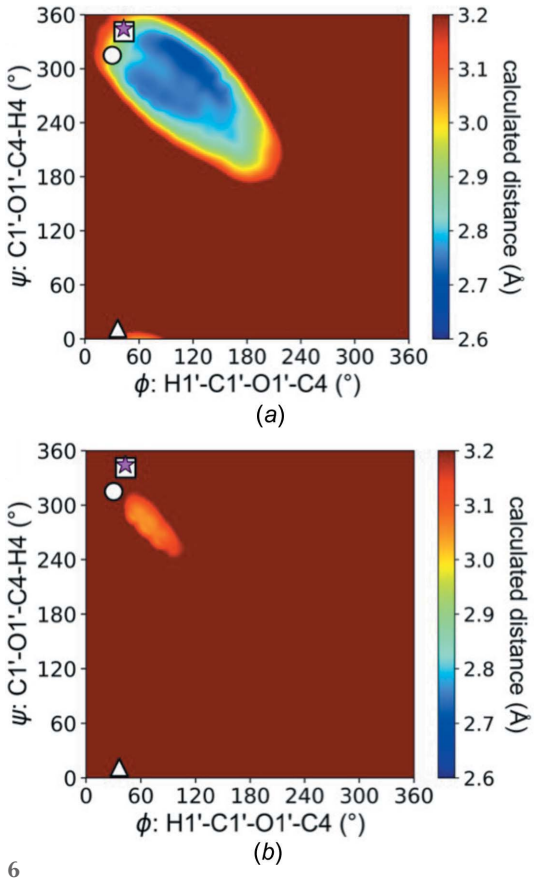


Figure 6
Internuclear distances for (*a*) O3···O5′ and (*b*) O3···O6′ as a function of ϕ and ψ determined by DFT calculations using (IX) as a model. For O3···O5′, ϕ and ψ values of ∼45–180 and 240–340°, respectively, yield distances compatible with inter-residue hydrogen bonding. For O3···O6′, ϕ and ψ values that permit hydrogen bonding are more limited (40–120 and 260–320°, respectively) in structures having O5′—C5′—C6′—O6′ torsion angles in residue **b** near +60° (*gt* rotamer). *Phi*/*psi* values observed in the crystal structures of (III) (star), (V) (circle), (VI) (triangle), and (VIII) (square) are superimposed on each map.

identical C2′—C1′—O1′—C4 torsion angles that report on *phi* (ϕ), and the similar C1′—O1′—C4—C3 torsion angles (Δ = 3°) that report on *psi* (ψ) (Fig. 5). The internuclear distances between O3 and O5′ in (III) and (VIII) are also similar (∼2.70 Å; Δ = 0.024 Å) and consistent with inter-residue hydrogen bonding. Rotameric properties of the C3—O3 bonds in residues **a** of (III) and (VIII) are also similar. For example, the C2—C3—O3—H torsion angles are 156 and 168°, respectively, which places O3H in close proximity to O5′. The O3—H···O5′ pseudo-bond angles in residues **a** of (III) and (VIII) are 164 and 157°, respectively, which approach the ∼180° value considered optimal for hydrogen bonding (Kroon & Kanters, 1974).

The *O*-glycosidic linkage conformations in (V) and (VI) differ significantly from one another and from the internal linkages in (III) and (VIII) (Fig. 5). In (VI), ϕ is similar to values observed in (III) and (VIII) (Δ = ∼3°), but ψ increases by ∼29°. Consequently, the internuclear distance between O3 and O5′ in (VI) increases to 3.311 Å, a value too large to support inter-residue hydrogen bonding. The C2—C3—O3—H torsion angle in residue **a** of (VI) (143°), and the O3′—H···O5 pseudo-bond angle (143°) are both smaller than the corresponding values in (III) and (VIII), suggesting a weaker interaction, if any, between O3H and O5′.

O-Glycosidic linkage conformation in (V) deviates significantly from those found in (III), (VI), and (VIII) (Fig. 5). Both ϕ and ψ are affected, the latter more than the former [Δ = ∼12° for ϕ and Δ = ∼27° for ψ, relative to (III) and (VIII)]. While the absolute difference for ψ is similar for (V) and (VI) relative to (III)/(VIII), these angles change in opposite directions. Interestingly, the ϕ and ψ torsion angles in (V) yield an O3···O5′ internuclear distance of 2.796 Å, which would allow inter-residue hydrogen bonding. The C2—C3—O3—H torsion angle (−149°) and O3—H···O5′ pseudo-bond angle (133°) in (V) differ considerably from those in (III), (VI), and (VIII), however, which may weaken the interaction. In addition, the exocyclic hydroxymethyl conformation in residue **b** of (V) (*gt* rotamer; *gt* = *gauche–trans*) orients O3 relatively close to O6′ (2.875 Å), resulting in additional inter-residue hydrogen bonding. However, the O3—H···O6′ pseudo-bond angle in (VI) is small (139°), which may render the interaction weak.

Theoretical calculations of the dependence of the O3···O5′ and O3···O6′ internuclear distances on ϕ and ψ in β-(1→4)-linked disaccharides were conducted using density functional theory (DFT) and methyl β-lactoside (methyl β-D-galacto-

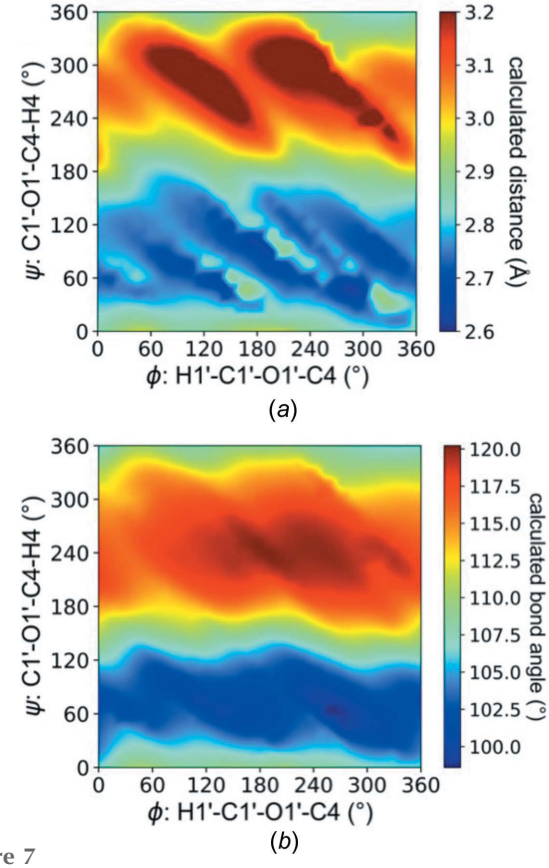


Figure 7
(*a*) Internuclear O3···O1′ distances and (*b*) O1′—C4—C3 bond angles in (IX) as a function of ϕ and ψ determined by DFT. The C1′—O1′—C4 and C4—C3—O3 bond angles showed no correlation with the O3···O1′ internuclear distance (data not shown; see text).

Table 3
Hydrogen-bond geometry (Å, °).

D—H···A	D—H	H···A	D···A	D—H···A
O2—H2*O*···O1*W*	0.84	1.89	2.702 (4)	163
O3—H3*O*···O5′	0.84	1.89	2.714 (3)	165
O6—H6*O*···O1^{i}	0.84	2.57	3.143 (3)	126
O6—H6*O*···O2^{i}	0.84	2.01	2.814 (3)	160
O3′—H3′*O*···O6^{i}	0.84	1.89	2.729 (3)	175
O4′—H4′*O*···O3^{ii}	0.84	1.93	2.767 (3)	171
O6′—H6′*O*···O3′^{iii}	0.84	1.95	2.755 (3)	160
N2′—H2′*N*···O2^{i}	0.88 (4)	2.33 (5)	3.133 (4)	154 (4)
N2′—H2′*N*···O5^{i}	0.88 (4)	2.62 (4)	3.312 (4)	137 (3)
O1*W*—H1*WA*···O7′^{iv}	0.85	2.17	2.833 (5)	135
O1*W*—H1*WB*···O6′^{iii}	0.85	1.95	2.795 (4)	178

Symmetry codes: (i) $-x+1, y+\frac{1}{2}, -z+\frac{1}{2}$; (ii) $-x+2, y+\frac{1}{2}, -z+\frac{1}{2}$; (iii) $-x+2, y-\frac{1}{2}, -z+\frac{1}{2}$; (iv) $-x+\frac{3}{2}, -y+1, z-\frac{1}{2}$.

pyranosyl-(1→4)-β-D-glucopyranoside), (IX), as a model (Fig. 6). These calculations were originally conducted for studies of the linkage torsion angles in (IX) (Zhang *et al.*, 2017). The *phi* (ϕ) and *psi* (ψ) torsion angles in (IX) were rotated in 15° increments, yielding 576 geometry-optimized conformers. Internuclear distances and bond angles were extracted from these conformers to generate the contour plots in Figs. 4 and 6 using *Python*. A small region of ϕ/ψ space correlates with internuclear distances compatible with hydrogen bonding, especially for O3···O6′. In the latter case, the exocyclic hydroxymethyl group must adopt the *gt* conformation in residue **b** to allow hydrogen bonding. Superimposition of crystallographic distances in (III), (V), (VI), and (VIII) on these theoretical maps explains why (V) accommodates hydrogen bonding between O3 and O5′/O6′, and (VI) does not. The DFT calculations also reveal that, while the O3—O1′ internuclear distance in β-(1→4)-linked disaccharides is expected to be relatively constant and independent of linkage conformation, a discernable dependence is nevertheless observed (Fig. 7). This behavior appears to correlate with the O1′—C4—C3 bond angle. In contrast, the C4—C3—O3 bond angle is essentially unaffected by ϕ and ψ, whereas the C1′—O1′—C4 bond angle is affected but does not correlate with the O1′···O3 internuclear distance (data not shown). However, the C1′—O1′—C4 bond angle appears to be a proxy for the total energy of the molecule, with smaller angles associated with lower-energy structures (see Fig. S3 in the supporting information).

3.3. *N*-Acetyl side-chain structural properties in (III)–(VII)

Average structural parameters in the *N*-acetyl side chains of (III)–(VII) are shown in Fig. 8 (throughout the following discussion and as shown in Fig. 8, *N*-acetyl side-chain atoms are denoted N2, C_{car}, O_{car}, and CH_3 to allow comparisons between different structures). The C2—N2 and C_{car}—O_{car} bond lengths are relatively constant, giving standard deviations of ±0.005 Å. The amide bond, N2—C_{car}, exhibits the largest variability, giving a standard deviation of 0.012 Å. Residue **b** in (VI) contains the shortest N2—C_{car} bond (1.317 Å), whereas residue **b** in (III) contains the longest bond (1.348 Å), for an overall difference of 0.031 Å. Shorter amide bonds suggest greater double-bond character and larger activation barriers for *cis–trans* isomerization (Hu *et al.*, 2010). Prior studies of noncarbohydrate systems have shown that activation barriers to amide bond rotation depend on solvation state, with nonpolar solvents lowering the activation barriers relative to polar and/or hydrogen-bonding solvents (Wiberg *et al.*, 1995; Pluth *et al.*, 2008). Nonpolar solvents are believed to stabilize the relatively nonpolar transition state for amide bond rotation while destabilizing the more polar ground state (Wiberg *et al.*, 1995; Pluth *et al.*, 2008). An inspection of the crystal structures of (III) and (VI) reveals different hydrogen-bonding interactions involving O_{car} and the N—H hydrogen (only internuclear distances < 3.0 Å between the heavy atoms involved in hydrogen bonds were considered in this treatment, since those with greater distances are comparatively weak and are not likely to be significant contributors to the observed N2—C_{car} and C_{car}—O_{car} bond-length effects). For (III), the N—H hydrogen is involved in weak bifurcated hydrogen bonds to O2^{i} and O5^{i}, and the carbonyl O atom serves as a monoacceptor from a water molecule (2.840 Å), although not in an optimal hydrogen-bonding geometry [C_{car}—O_{car}···H bond angle of 122/107° (bifurcated); N2—C_{car}—O_{car}···H pseudo-torsion angle of 142/115° (bifurcated)] (Wiberg *et al.*, 1995; Pluth *et al.*, 2008). In (VI), the N—H hydrogen serves as a donor to a water molecule, and the carbonyl O atom serves as a monoacceptor in a near-optimal hydrogen-bonding geometry (C_{car}—O_{car}···H bond angle of 119°; N2—C_{car}—O_{car}···H pseudo-torsion angle of −176°), assuming that the two lone-pair orbitals on the carbonyl O atom are directional and lie in the —NH—(CO)—CH_3 plane (Wiberg *et al.*, 1995; Pluth *et al.*, 2008). If the hydrogen-bonding interactions in (VI) are, on average,

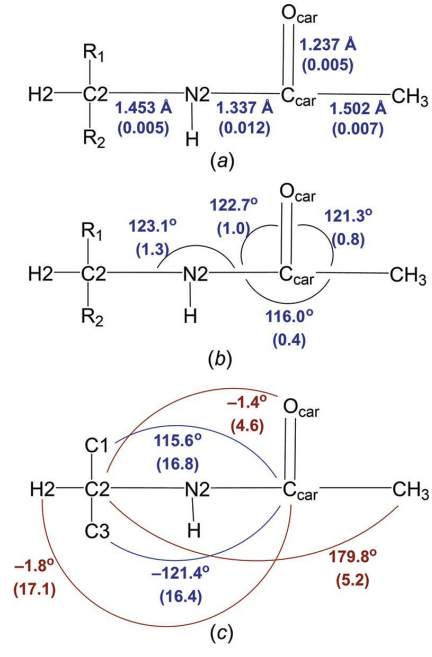


Figure 8
Averaged structural parameters for the *N*-acetyl side chains in (III)–(VII) obtained from the data in Table S1 (see supporting information), showing (*a*) bond lengths, (*b*) valence bond angles, and (*c*) torsion angles. Mean values are shown, with s.u. values in parentheses.

assumed to be stronger than those in (III), the effect enhances the double-bond character of the amide bond. This finding may be explained by postulating that the relative contributions of resonance forms of the amide bond (Fig. 9) differ in nonpolar and polar solvents, with the former favoring forms containing an amide C—N single bond, and the latter favoring forms containing an amide C═N double bond. When O_{car} is involved as a hydrogen-bond acceptor, resonance forms containing the amide C═N double bond are more favored, leading to a shorter bond and higher activation barrier to rotation.

DFT calculations on nine *in silico* models of methyl 2-acetamido-2-deoxy-β-D-glucopyranoside, (IV), reveal the sensitivities of the N2—C_{car} and C_{car}—O_{car} bond lengths to the hydrogen-bonding state of the amide side chain (Table 1, and Figs. S4 and S5 in the supporting information). The results support the contention that carbonyl hydrogen bonding reduces the N2—C_{car} bond lengths and may slightly elongate the C_{car}—O_{car} bond lengths (models 2–6), whereas N—H hydrogen bonding exerts smaller effects on these bond lengths (model 6 relative to model 4). This conclusion is consistent with the experimental data (Fig. 8) that indicate a greater change in the N2—C_{car} bond length than in the C_{car}—O_{car} bond length in response to different hydrogen-bonding patterns. When hydrogen bonding to O_{car} is replaced by formal protonation (model 9), the bond-length effects are amplified, with the N2—C_{car} bond shortened by ∼0.06 Å and the C_{car}—O_{car} bond lengthened by an equivalent amount relative to bond lengths in the vacuum model (model 1). These results may have implications for *N*-acetyl side-chain behavior in more complex oligosaccharides, where states of solvation and/or hydration could affect the activation barriers to *cis–trans* isomerization and indirectly influence the binding of oligosaccharides to biological receptors and/or *O*-glycosidic linkage conformation when the side chain is in close proximity to the linkage (Wiberg *et al.*, 1994; Kirby, 1983*a*). Since *N*-glycosidic linkages in the *N*-glycans of glycopeptides and glycoproteins are structurally related to *N*-acetyl side chains (Fig. 10), the effects of solvent on their behaviors should mimic those for *N*-acetyl side chains.

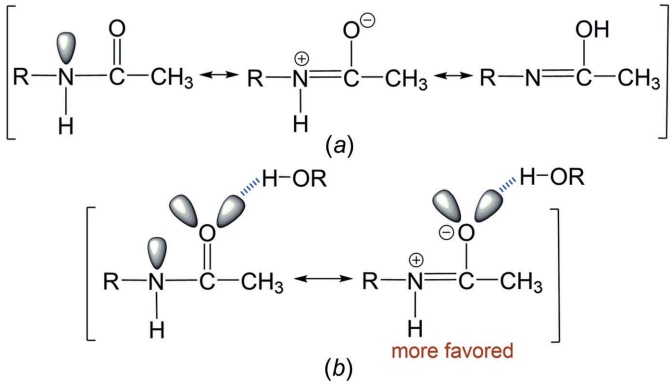


Figure 9
Resonance models of an *N*-acetyl side chain (*a*) *in vacuo* and in nonpolar solvents, and (*b*) in polar and/or hydrogen-bonding solvents. In part (*b*), the hydrogen bond to the carbonyl O atom shifts the relative contributions of the resonance forms towards those bearing a C═N double bond. Participation of the N—H hydrogen as a hydrogen-bond donor appears to exert weaker but reinforcing effects on the C—N and C—O bond lengths (see Table 1).

Valence bond angles in *N*-acetyl side chains are relatively constant, with standard deviations of 0.4–1.3° (Fig. 8). Bond angles involving C_{car} are asymmetric, with the N2—C_{car}—O_{car} and O_{car}—C_{car}—CH_3 angles very similar and ∼6° larger than the N2—C_{car}—CH_3 angle. This behavior is similar to that observed in *O*-acetyl side chains (Wiberg *et al.*, 1994; Kirby, 1983*b*).

Torsion angles that report on the rotameric properties of the C2—N2 bond show significant variability, with standard deviations of ∼17°. In general, geometries about the C2—N2 bond that orient H2 and O_{car} in eclipsed arrangements, or nearly so, are favored, similar to the behavior of *O*-acetyl side chains (Turney *et al.*, 2017, 2019). By comparison, the rotameric behavior of the amide bond, N2—C_{car}, is highly constrained, favoring a geometry that orients C2 *anti* to CH_3, although some mobility is observed, as reflected in the standard deviation of ∼5°.

3.4. Aldohexopyranosyl ring and exocyclic hydroxymethyl conformations in (III)–(VIII)

The aldohexopyranosyl rings in (III)–(VIII) adopt $^4C_1$ conformations in the crystals, although small deviations from idealized chairs are indicated by the Cremer–Pople puckering parameters (Table 4). Values of θ vary over the range 0–11°, with the ideal $^4C_1$ conformer observed only in the βGlcNAc residue in (III) (θ = 0°) and the greatest distortion from $^4C_1$ observed in (IV_B) (θ = 11°). Within (III), (V), (VI), and (VIII), the most distorted ring geometries occur in (V), with both βGlcNAc residues giving θ values of 4.8 and 8.7°. The direction of chair distortion favors θ values ranging from approximately 270 to 90° (*i.e.* one continuous half of the pseudorotational itinerary at θ = 90°), which encompasses the $^{C1}S_{C5}$, $B_{C2,C5}$, $^{O5}S_{C2}$, $^{C3,O5}B$, $^{C3}S_{C1}$, $B_{C1,C4}$, and $^{C5}S_{C1}$ forms (Table 4).

Other indicators of conformational variability of the aldohexopyranosyl rings in (III)–(VII) are the endocyclic C1—C2—C3—C4 and C1—O5—C5—C4 torsion angles, with average values of −52.69±3.31 and 63.26±3.54°, respectively, indicating roughly equivalent deviations along both pathways.

Exocyclic hydroxymethyl groups in (III)–(VIII) assume either *gg* (H5 *anti* to O6; *gg* = *gauche–gauche*) or *gt* (C4 *anti* to O6) conformations. Average values of the O5—C5—C6—O6 torsion angles for residues adopting *gg* and *gt* conformations are −64.64±6.80 and 66.08±7.4°, respectively. For disaccharides (III), (V), (VI), and (VIII), the hydroxymethyl confor-

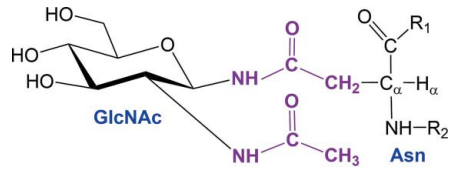


Figure 10
The structure of βGlcNAc appended to Asn in the *N*-linked glycoprotein, highlighting the similarities in the structure of the linkage and that of the adjacent *N*-acetyl side chain.

Table 4
Cremer–Pople puckering parameters for the aldohexopyranosyl rings in (III)–(VIII).

Compound[a]	θ (°)	ϕ (°)	Q (Å)	q_2 (Å)	q_3 (Å)	Conformer[b]
(III), residue **a**	4.3 (3)	40 (4)	0.598 (3)	0.046 (3)	0.597 (3)	$^{C3}S_{C1}$
(III), residue **b**	0.0 (3)	185 (7)	0.575 (3)	0.024 (3)	0.576 (3)	—
(IV_A)	11.4 (3)	302.0 (12)	0.595 (3)	0.118 (3)	0.583 (3)	$B_{C2,C5}$
(IV_B)	7.1 (2)	0.0 (2)	0.585 (3)	0.078 (2)	0.580 (3)	$^{C3,O5}B$
(V), residue **a**	4.8 (3)	19 (4)	0.568 (3)	0.044 (3)	0.566 (3)	$^{C3}S_{C1}$
(V), residue **b**	8.7 (3)	338 (2)	0.577 (3)	0.089 (9)	0.570 (3)	$^{O5}S_{C2}$
(VI), residue **a**	0.9 (3)	55 (3)	0.572 (3)	0.006 (3)	0.572 (3)	$B_{C1,C4}$
(VI), residue **b**	2.3 (3)	97 (6)	0.580 (3)	0.029 (3)	0.580 (3)	$^{C5}S_{C1}$
(VII)	3.8 (2)	274 (3)	0.582 (2)	0.031 (2)	0.581 (2)	$^{C1}S_{C5}$
(VIII), residue **a**	4.52	336	0.5752	0.0468	0.5734	$^{O5}S_{C2}$

Notes: (*a*) see Scheme 1 for the definitions of residues **a** and **b** in (III), (V), (VI), and (VIII). (*b*) *B* = boat and *S* = twist-boat or skew. Parameters shown for (IV)–(VIII) were extracted from crystal structures reported in Hu *et al.* (2011), Mo (1979), Mo & Jensen (1978), Mo & Jensen (1975), and Zhang *et al.* (2013), respectively.

mation in residue **b** influences the inter-residue hydrogen bonding involving O3H as a donor. In (V), residue **b** adopts a *gt* conformation, which allows O6′ (in addition to O5′) to participate as a hydrogen-bond acceptor with O3H. This interaction is not possible when the hydroxymethyl group adopts a *gg* (or *tg*) conformation.

3.5. Other structural considerations

Exocyclic O5—C1—O1 bond angles in all of residues (III)–(VIII) fall into two groups, those associated with rings having the α-anomeric configuration [112.18±0.40°; residue **a** in (VI) and (VII)] and those associated with rings having the β-anomeric configuration (107.41±0.61°; remaining eight residues). This dependence has been described previously; α-anomers are subject to the *endo*-anomeric effect and β-anomers are not (Juaristi & Cuevas, 1995; Kirby, 1983*b*), rendering sp^2-like character to C1 in the former, resulting in an increased O5—C1—O1 bond angle. Similar arguments pertain to the endocyclic C1—O5—C5 bond angles, which also fall into two groups. Rings having the α-anomeric configuration give larger values [114.72±0.33°; residue **a** in (VI) and (VII)] than rings having the β-anomeric configuration [112.09±0.55°; remaining residues except residue **b** in (III)]. The exception is residue **b** of (III), which contains a C1—O5—C5 bond angle of 115.47°, significantly larger than expected.

For the βGlcNAc rings of residue **b** in disaccharides (III), (V), and (VI), and in monosaccharides (IV_A) and (IV_B), the following average endocyclic bond lengths were calculated with their corresponding s.u. values [indicated atoms would be primed in (III), (V), and (V), but not in (IV_A) and (IV_B); for clarity, unprimed atoms are used here]: C1—C2 = 1.525±0.007 Å, C2—C3 = 1.529±0.007 Å, C3—C4 = 1.520±0.006 Å, C4—C5 = 1.525±0.011 Å, C5—O5 = 1.436±0.006 Å, and O5—C1 = 1.423±0.007 Å. A larger s.u. value is observed for the C4—C5 bond, whose length is likely affected, among other factors, by the rotameric states of the C4—C5, C5—O5, and C4—O4 bonds (Carmichael *et al.*, 1993). The former two are dictated by ring conformation and do not vary appreciably (by ∼10°; the C3—C4—C5—O5 torsion angle ranges from −55 to −64° and the C4—C5—O5—C1 torsion angle ranges from 59 to 69°). Greater variability in the C3—C4—O4—H torsion angles is observed (57 to 133°), although these changes may not be sufficient to explain the observed behavior. Other factors may pertain, such as different hydrogen bonding to O4/O5 and/or differences in the conformation of the exocyclic hydroxymethyl group, although the latter is not expected to be significant based on DFT calculations (data not shown).

4. Conclusions

Structural parameters in the *N*-acetyl side chains in saccharides are relatively uniform with respect to bond and torsion angles, but bond lengths, especially for the amide N2—C_{car} and carbonyl C_{car}—O_{car} bonds, are sensitive to the state of hydrogen bonding. DFT calculations support the empirical observations and generally show O_{car} serving as a hydrogen-bond acceptor in which both lone-pair orbitals are directional, although this behavior is not uniformly observed in both the crystal structures and the DFT calculations (the donor H atom may orient between the lone pairs, implying a kidney-shaped electron distribution around the O atom). As discussed previously (Wiberg *et al.*, 1995; Pluth *et al.*, 2008), amide solvation can influence the activation barriers to *cis–trans* isomerization, with increasingly higher barriers correlating with increasing solvent polarity and/or increased formal hydrogen bonding to O_{car}. The *N*-acetyl side chains in complex saccharides could exhibit different kinetics properties depending on their states of solvation, a prediction that can be tested experimentally using saturation–transfer and related NMR methods (Hu *et al.*, 2010). The biological implications of this behavior could be significant in terms of ligand binding to receptors.

Given the important role that the βGlcNAc-(1→4)-βGlcNAc linkage plays in glycobiology (Fig. 1), modern low-temperature crystal structures of methyl 2-acetamido-2-deoxy-β-D-glucopyranosyl-(1→4)-2-acetamido-2-deoxy-α- and β-D-glucopyranoside should be determined, since current databases contain crystal structures of only reducing disaccharides (VI) and (V), and their reported structures contain disorder caused by the presence of minor anomers and/or other factors.

DFT calculations on model disaccharide (IX) confirm that a relatively limited region of ϕ/ψ space yields O3···O5′ internuclear distances compatible with inter-residue hydrogen bonding in β-(1→4) *O*-glycosidic linkages, and an even more restricted space is compatible with O3···O6′ hydrogen bonding. The calculated hypersurface shows that relatively large but correlated changes in ϕ and ψ are capable of supporting hydrogen bonding between O3H and O5′. Calculated hypersurfaces correlating ϕ and ψ with the O3···O1′ internuclear distance and the O1′—C4—C3 bond angle are similar and indicate that the former distance is not fixed but responds to changes in the latter bond angle brought about by altered linkage conformation. In β-(1→4) linkages, transglycoside C1′—O1′—C4 bond angles appear to serve as

proxies of total energy; hypersurfaces correlating ϕ and ψ and either total energy or C1′—O1′—C4 bond angle contain virtually identical local and global minima. Whether a similar behavior is associated with two-bond *O*-glycosidic linkages in general remains to be determined.

Funding information

Funding for this research was provided by: National Science Foundation (grant Nos. CHE 1402744 and CHE 1707660 to AS); Department of Energy Office of Science, Office of Basic Energy Sciences (award No. DE-FC02-04ER15533 to NDRL 5249).

References

André, S., Kozár, T., Schuberth, R., Unverzagt, C., Kojima, S. & Gabius, H.-J. (2007). *Biochemistry*, **46**, 6984–6995.
Becke, A. D. (1993). *J. Chem. Phys.* **98**, 5648–5652.
Bruker (2015). *APEX3* and *SAINT*. Bruker–Nonius AXS Inc., Madison, Wisconsin, USA.
Cammi, R., Mennucci, B. & Tomasi, J. (2000). *J. Phys. Chem. A*, **104**, 5631–5637.
Cancès, E., Mennucci, B. & Tomasi, J. (1997). *J. Chem. Phys.* **107**, 3032–3041.
Carmichael, I., Chipman, D. M., Podlasek, C. A. & Serianni, A. S. (1993). *J. Am. Chem. Soc.* **115**, 10863–10870.
Case, D. A., *et al.* (2018). *AMBER 2018*. University of California, San Francisco, USA.
Etter, M. C., MacDonald, J. C. & Bernstein, J. (1990). *Acta Cryst.* B**46**, 256–262.
Frisch, M. J., Pople, J. A. & Binkley, J. S. (1984). *J. Chem. Phys.* **80**, 3265–3269.
Frisch, M. J., *et al.* (2016). *GAUSSIAN16*. Revision B.01. Gaussian Inc., Wallingford, CT, USA. https://gaussian.com/gaussian16/.
Geisler, C. & Jarvis, D. L. (2012). *J. Biol. Chem.* **287**, 7084–7097.
Hanashima, S., Suga, A. & Yamaguchi, Y. (2018). *Carbohydr. Res.* **456**, 53–60.
Hart, G. W., Slawson, C., Ramirez-Correa, G. & Lagerlof, O. (2011). *Annu. Rev. Biochem.* **80**, 825–858.
Hu, X., Zhang, W., Carmichael, I. & Serianni, A. S. (2010). *J. Am. Chem. Soc.* **132**, 4641–4652.
Hu, X., Zhang, W., Oliver, A. G. & Serianni, A. S. (2011). *Acta Cryst.* C**67**, o146–o150.
Johansen, P. G., Marshall, R. D. & Neuberger, A. (1961). *Biochem. J.* **78**, 518–527.
Jones, E., Oliphant, T. & Peterson, P. (2014). *SciPy: Open Source Scientific Tools for Python*. https://www.scipy.org/.
Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. (1983). *J. Chem. Phys.* **79**, 926–935.
Juaristi, E. & Cuevas, G. (1995). *The Anomeric Effect*, pp. 3–13. Boca Raton: CRC Press.
Kaye, M. A. G. & Stacey, M. (1951). *Biochem. J.* **48**, 249–255.
Kirby, A. J. (1983*a*). *The Anomeric Effect and Related Stereoelectronic Effects at Oxygen*, pp. 42–44. Berlin: Springer-Verlag.
Kirby, A. J. (1983*b*). *The Anomeric Effect and Related Stereoelectronic Effects at Oxygen*, pp. 52–61. Berlin: Springer-Verlag.
Kirschner, K. N., Yongye, A. B., Tschampel, S. M., González-Outeiriño, J., Daniels, C. R., Foley, B. L. & Woods, R. J. (2008). *J. Comput. Chem.* **29**, 622–655.
Krause, L., Herbst-Irmer, R., Sheldrick, G. M. & Stalke, D. (2015). *J. Appl. Cryst.* **48**, 3–10.
Kroon, J. & Kanters, J. A. (1974). *Nature*, **248**, 667–669.
Liu, R. & Wei, A. (2012). *J. Carbohydr. Chem.* **31**, 384–419.
Macrae, C. F., Sovago, I., Cottrell, S. J., Galek, P. T. A., McCabe, P., Pidcock, E., Platings, M., Shields, G. P., Stevens, J. S., Towler, M. & Wood, P. A. (2020). *J. Appl. Cryst.* **53**, 226–235.
McLean, A. D. & Chandler, G. S. (1980). *J. Chem. Phys.* **72**, 5639–5648.
Mo, F. (1979). *Acta Chem. Scand.* **33a**, 207–218.
Mo, F. & Jensen, L. H. (1975). *Acta Cryst.* B**31**, 2867–2873.
Mo, F. & Jensen, L. H. (1978). *Acta Cryst.* B**34**, 1562–1569.
Nagae, M., Yamanaka, K., Hanashima, S., Ikeda, A., Morita-Matsumoto, K., Satoh, T., Matsumoto, N., Yamamoto, K. & Yamaguchi, Y. (2013). *J. Biol. Chem.* **288**, 33598–33610.
Parsons, S., Flack, H. D. & Wagner, T. (2013). *Acta Cryst.* B**69**, 249–259.
Pluth, M. D., Bergman, R. G. & Raymond, K. N. (2008). *J. Org. Chem.* **73**, 7132–7136.
Qin, H. & Grindley, T. B. (1994). *J. Carbohydr. Chem.* **13**, 475–490.
Rudd, P. M. & Dwek, R. A. (1997). *Crit. Rev. Biochem. Mol. Biol.* **32**, 1–100.
Ruhaak, L. R., Uh, H.-W., Beekman, M., Koeleman, C. A. M., Hokke, C. H., Westendorp, R. G. J., Wuhrer, M., Houwing-Duistermaat, J. J., Slagboom, P. E. & Deelder, A. M. (2010). *PLoS One*, **5**, e12566.
Schmidt, R. R. & Michel, J. (1985). *J. Carbohydr. Chem.* **4**, 141–169.
Sheldrick, G. M. (2015*a*). *Acta Cryst.* A**71**, 3–8.
Sheldrick, G. M. (2015*b*). *Acta Cryst.* C**71**, 3–8.
Spek, A. L. (2009). *Acta Cryst.* D**65**, 148–155.
Sultan, A. S., Miyoshi, E., Ihara, Y., Nishikawa, A., Tsukada, Y. & Taniguchi, N. (1997). *J. Biol. Chem.* **272**, 2866–2872.
Takahashi, M., Kuroki, Y., Ohtsubo, K. & Taniguchi, N. (2009). *Carbohydr. Res.* **344**, 1387–1390.
Turney, T., Pan, Q., Sernau, L., Carmichael, I., Zhang, W., Wang, X., Woods, R. J. & Serianni, A. S. (2017). *J. Phys. Chem. B*, **121**, 66–77.
Turney, T., Zhang, W., Oliver, A. G. & Serianni, A. S. (2019). *Acta Cryst.* C**75**, 1166–1174.
Westrip, S. P. (2010). *J. Appl. Cryst.* **43**, 920–925.
Wiberg, K. B., Marquez, M. & Castejon, H. (1994). *J. Org. Chem.* **59**, 6817–6822.
Wiberg, K. B., Rablen, P. R., Rush, D. J. & Keith, T. A. (1995). *J. Am. Chem. Soc.* **117**, 4261–4270.
Woods Group (2019). *GLYCAM Web*. Complex Carbohydrate Research Center (CCRC), University of Georgia, Athens, GA, USA. http://www.glycam.org.
Yang, W. H., Kim, J. E., Nam, H. W., Ju, J. W., Kim, H. S., Kim, Y. S. & Cho, J. W. (2006). *Nat. Cell Biol.* **8**, 1074–1083.
Zhang, W., Meredith, R., Pan, Q., Wang, X., Woods, R. J., Carmichael, I. & Serianni, A. S. (2019). *Biochemistry*, **58**, 546–560.
Zhang, W., Oliver, A. G., Vu, H. M., Duman, J. G. & Serianni, A. S. (2013). *Acta Cryst.* C**69**, 1047–1050.
Zhang, W., Turney, T., Meredith, R., Pan, Q., Sernau, L., Wang, X., Hu, X., Woods, R. J., Carmichael, I. & Serianni, A. S. (2017). *J. Phys. Chem. B*, **121**, 3042–3058.

supporting information

Acta Cryst. (2020). C**76**, 287-297 [https://doi.org/10.1107/S2053229620001515]

Glycosidic linkage, *N*-acetyl side-chain, and other structural properties of methyl 2-acetamido-2-deoxy-*β*-D-glucopyranosyl-(1→4)-*β*-D-mannopyranoside monohydrate and related compounds

Wenhui Zhang, Reagan J. Meredith, Allen G. Oliver, Ian Carmichael and Anthony S. Serianni

Computing details

Data collection: *APEX3* (Bruker, 2015); cell refinement: *SAINT* (Bruker, 2015); data reduction: *SAINT* (Bruker, 2015); program(s) used to solve structure: SHELXT2018 (Sheldrick, 2015a); program(s) used to refine structure: *SHELXL2018* (Sheldrick, 2015b); molecular graphics: *Mercury* (Macrae *et al.*, 2020); software used to prepare material for publication: CIFTAB (Sheldrick, 2015b), *PLATON* (Spek, 2009) and *publCIF* (Westrip, 2010).

(I)

Crystal data

$C_{15}H_{27}NO_{11}·H_2O$
M_r = 415.39
Orthorhombic, $P2_12_12_1$
a = 11.4170 (6) Å
b = 12.6782 (6) Å
c = 13.2102 (7) Å
V = 1912.14 (17) Å^3
Z = 4
$F(000)$ = 888
D_x = 1.443 Mg m^{-3}
Cu $K\alpha$ radiation, λ = 1.54178 Å
Cell parameters from 9648 reflections
θ = 3.4–71.1°
μ = 1.09 mm^{-1}
T = 120 K
Block, colorless
0.13 × 0.11 × 0.10 mm

Data collection

Bruker PHOTON-II diffractometer
Radiation source: Incoatec micro-focus
Detector resolution: 7.41 pixels mm^{-1}
combination of ω and φ–scans
Absorption correction: numerical (SADABS; Krause *et al.*, 2015)
T_{min} = 0.863, T_{max} = 1.000
31774 measured reflections
3635 independent reflections
3510 reflections with $I > 2\sigma(I)$
R_{int} = 0.047
θ_{max} = 71.2°, θ_{min} = 4.8°
h = −14→13
k = −15→15
l = −13→15

Refinement

Refinement on F^2
Least-squares matrix: full
$R[F^2 > 2\sigma(F^2)]$ = 0.044
$wR(F^2)$ = 0.117
S = 1.05
3635 reflections
266 parameters
0 restraints
Primary atom site location: dual
Secondary atom site location: difference Fourier map
Hydrogen site location: mixed
H atoms treated by a mixture of independent and constrained refinement
$w = 1/[\sigma^2(F_o^2) + (0.0702P)^2 + 1.0254P]$ where $P = (F_o^2 + 2F_c^2)/3$

$(\Delta/\sigma)_{max}$ = 0.055
$\Delta\rho_{max}$ = 0.63 e Å^{-3}
$\Delta\rho_{min}$ = −0.36 e Å^{-3}

Absolute structure: Flack x determined using 1469 quotients [(I+)-(I-)]/[(I+)+(I-)] (Parsons *et al.*, 2013)
Absolute structure parameter: −0.07 (7)

Special details

Geometry. All esds (except the esd in the dihedral angle between two l.s. planes) are estimated using the full covariance matrix. The cell esds are taken into account individually in the estimation of esds in distances, angles and torsion angles; correlations between esds in cell parameters are only used when they are defined by crystal symmetry. An approximate (isotropic) treatment of cell esds is used for estimating esds involving l.s. planes.

Fractional atomic coordinates and isotropic or equivalent isotropic displacement parameters (Å^2)

	x	*y*	*z*	U_{iso}*/U_{eq}
O1	0.48814 (19)	0.05124 (16)	0.41109 (17)	0.0261 (5)
O2	0.6316 (2)	0.12664 (18)	0.26391 (18)	0.0297 (5)
H2O	0.694514	0.123890	0.231040	0.044*
O3	0.8165 (2)	0.26309 (18)	0.3214 (2)	0.0375 (6)
H3O	0.831442	0.325608	0.304740	0.056*
O5	0.46341 (18)	0.22663 (15)	0.38386 (16)	0.0229 (4)
O6	0.4021 (2)	0.40801 (18)	0.25833 (19)	0.0333 (5)
H6O	0.396247	0.470235	0.237091	0.050*
C1	0.5412 (3)	0.1493 (2)	0.4249 (2)	0.0260 (6)
H1	0.555290	0.162869	0.498453	0.031*
C2	0.6563 (3)	0.1508 (2)	0.3666 (3)	0.0278 (7)
H2	0.710694	0.096636	0.395223	0.033*
C3	0.7102 (3)	0.2603 (2)	0.3782 (3)	0.0278 (7)
H3	0.729379	0.271926	0.451277	0.033*
C4	0.6235 (3)	0.3454 (2)	0.3445 (2)	0.0225 (6)
H4	0.610864	0.341536	0.269695	0.027*
C5	0.5071 (3)	0.3316 (2)	0.4004 (2)	0.0230 (6)
H5	0.520768	0.341720	0.474487	0.028*
C6	0.4142 (3)	0.4088 (2)	0.3659 (3)	0.0278 (7)
H6A	0.338301	0.390145	0.397368	0.033*
H6B	0.435587	0.480728	0.388579	0.033*
C7	0.3951 (3)	0.0344 (3)	0.4831 (3)	0.0332 (7)
H7	0.355933	−0.032676	0.468376	0.050*
H7B	0.338307	0.092122	0.478290	0.050*
H7C	0.427790	0.032205	0.551693	0.050*
O1′	0.66607 (18)	0.44811 (15)	0.37228 (17)	0.0230 (4)
O3′	0.79463 (19)	0.79521 (16)	0.28686 (18)	0.0271 (5)
H3′O	0.732411	0.828864	0.276453	0.041*
O4′	0.95356 (19)	0.70168 (18)	0.14871 (17)	0.0288 (5)
H4′O	1.022889	0.716772	0.164297	0.043*
O5′	0.82289 (18)	0.47020 (17)	0.26980 (18)	0.0263 (5)
O6′	1.0620 (2)	0.4591 (2)	0.2716 (2)	0.0432 (6)
H6′O	1.116571	0.415298	0.264381	0.065*
O7′	0.6476 (2)	0.7022 (3)	0.5169 (2)	0.0500 (7)
N2′	0.6031 (2)	0.66577 (19)	0.3531 (2)	0.0230 (5)

H2′N	0.549 (4)	0.669 (3)	0.306 (3)	0.032 (10)*
C1′	0.7086 (3)	0.5098 (2)	0.2929 (2)	0.0228 (6)
H1′	0.656049	0.504968	0.232526	0.027*
C2′	0.7183 (3)	0.6233 (2)	0.3298 (2)	0.0217 (6)
H2′	0.765990	0.623696	0.393246	0.026*
C3′	0.7818 (2)	0.6900 (2)	0.2504 (2)	0.0224 (6)
H3′	0.734637	0.691021	0.186682	0.027*
C4′	0.9013 (3)	0.6431 (2)	0.2284 (2)	0.0229 (6)
H4′	0.951576	0.646349	0.290341	0.027*
C5′	0.8866 (3)	0.5284 (2)	0.1942 (2)	0.0248 (6)
H5′	0.841401	0.527089	0.129370	0.030*
C6′	1.0007 (3)	0.4698 (3)	0.1793 (3)	0.0329 (7)
H6′A	0.984427	0.399040	0.151026	0.039*
H6′B	1.050084	0.508570	0.130177	0.039*
C7′	0.5770 (3)	0.7035 (3)	0.4459 (3)	0.0339 (7)
C8′	0.4566 (4)	0.7499 (4)	0.4575 (4)	0.0536 (11)
H8′A	0.437020	0.754848	0.529577	0.080*
H8′B	0.399391	0.704760	0.423240	0.080*
H8′C	0.455062	0.820527	0.427306	0.080*
O1W	0.8050 (4)	0.1101 (3)	0.1256 (3)	0.0732 (12)
H1WA	0.841110	0.168685	0.124123	0.110*
H1WB	0.845280	0.065235	0.158386	0.110*

Atomic displacement parameters (Å²)

	U^{11}	U^{22}	U^{33}	U^{12}	U^{13}	U^{23}
O1	0.0289 (11)	0.0187 (9)	0.0309 (12)	−0.0042 (9)	0.0014 (9)	0.0015 (8)
O2	0.0268 (11)	0.0266 (10)	0.0355 (13)	−0.0040 (9)	0.0075 (9)	−0.0052 (9)
O3	0.0197 (11)	0.0219 (11)	0.0709 (18)	0.0011 (9)	0.0055 (11)	0.0015 (11)
O5	0.0207 (10)	0.0172 (10)	0.0308 (12)	−0.0024 (8)	−0.0036 (8)	0.0024 (8)
O6	0.0299 (12)	0.0275 (11)	0.0424 (14)	−0.0046 (9)	−0.0078 (10)	0.0114 (10)
C1	0.0264 (15)	0.0196 (14)	0.0319 (17)	0.0001 (12)	−0.0060 (12)	0.0020 (12)
C2	0.0234 (15)	0.0193 (14)	0.0406 (18)	0.0001 (11)	−0.0057 (13)	−0.0016 (12)
C3	0.0197 (14)	0.0212 (15)	0.0426 (19)	−0.0021 (12)	−0.0023 (13)	0.0003 (13)
C4	0.0206 (14)	0.0178 (13)	0.0292 (16)	−0.0017 (11)	−0.0026 (11)	−0.0031 (11)
C5	0.0225 (14)	0.0168 (12)	0.0297 (16)	−0.0021 (11)	−0.0007 (11)	−0.0009 (11)
C6	0.0221 (14)	0.0212 (14)	0.0401 (19)	0.0002 (12)	0.0027 (13)	0.0023 (12)
C7	0.0382 (18)	0.0280 (15)	0.0332 (18)	−0.0065 (14)	0.0029 (14)	0.0040 (13)
O1′	0.0237 (10)	0.0179 (10)	0.0274 (11)	−0.0035 (8)	−0.0007 (8)	0.0002 (8)
O3′	0.0234 (10)	0.0185 (10)	0.0395 (13)	−0.0008 (8)	0.0027 (9)	−0.0001 (9)
O4′	0.0211 (10)	0.0329 (12)	0.0324 (12)	−0.0040 (9)	0.0027 (9)	0.0068 (9)
O5′	0.0202 (10)	0.0233 (10)	0.0355 (12)	0.0035 (8)	0.0056 (9)	0.0026 (9)
O6′	0.0300 (13)	0.0417 (15)	0.0580 (17)	0.0118 (11)	−0.0065 (11)	−0.0122 (12)
O7′	0.0394 (15)	0.077 (2)	0.0337 (15)	0.0129 (15)	−0.0020 (11)	−0.0078 (14)
N2′	0.0173 (11)	0.0213 (11)	0.0304 (14)	0.0006 (10)	−0.0007 (10)	−0.0010 (10)
C1′	0.0198 (13)	0.0193 (13)	0.0293 (16)	−0.0002 (11)	0.0016 (11)	0.0009 (11)
C2′	0.0187 (13)	0.0181 (13)	0.0284 (15)	0.0000 (11)	−0.0005 (11)	0.0000 (11)
C3′	0.0185 (13)	0.0192 (13)	0.0294 (16)	−0.0008 (10)	−0.0003 (11)	0.0001 (11)

C4′	0.0192 (13)	0.0229 (13)	0.0267 (14)	−0.0009 (11)	−0.0013 (11)	0.0020 (11)
C5′	0.0202 (14)	0.0257 (14)	0.0285 (16)	−0.0001 (12)	0.0033 (11)	0.0003 (12)
C6′	0.0258 (16)	0.0314 (16)	0.041 (2)	0.0035 (14)	0.0078 (14)	−0.0024 (14)
C7′	0.0290 (17)	0.0392 (18)	0.0336 (18)	0.0070 (14)	0.0015 (13)	−0.0009 (14)
C8′	0.036 (2)	0.078 (3)	0.047 (2)	0.022 (2)	−0.0018 (17)	−0.017 (2)
O1W	0.074 (2)	0.083 (3)	0.063 (2)	0.035 (2)	0.0301 (19)	0.0254 (18)

Geometric parameters (Å, º)

O1—C1	1.395 (4)	O4′—C4′	1.420 (4)
O1—C7	1.442 (4)	O4′—H4′O	0.8400
O2—C2	1.419 (4)	O5′—C1′	1.431 (4)
O2—H2O	0.8400	O5′—C5′	1.439 (4)
O3—C3	1.427 (4)	O6′—C6′	1.412 (5)
O3—H3O	0.8400	O6′—H6′O	0.8400
O5—C1	1.430 (4)	O7′—C7′	1.236 (5)
O5—C5	1.438 (3)	N2′—C7′	1.349 (4)
O6—C6	1.428 (4)	N2′—C2′	1.454 (4)
O6—H6O	0.8400	N2′—H2′N	0.88 (4)
C1—C2	1.523 (5)	C1′—C2′	1.523 (4)
C1—H1	1.0000	C1′—H1′	1.0000
C2—C3	1.527 (4)	C2′—C3′	1.530 (4)
C2—H2	1.0000	C2′—H2′	1.0000
C3—C4	1.531 (4)	C3′—C4′	1.516 (4)
C3—H3	1.0000	C3′—H3′	1.0000
C4—O1′	1.437 (3)	C4′—C5′	1.532 (4)
C4—C5	1.530 (4)	C4′—H4′	1.0000
C4—H4	1.0000	C5′—C6′	1.513 (4)
C5—C6	1.513 (4)	C5′—H5′	1.0000
C5—H5	1.0000	C6′—H6′A	0.9900
C6—H6A	0.9900	C6′—H6′B	0.9900
C6—H6B	0.9900	C7′—C8′	1.503 (5)
C7—H7	0.9800	C8′—H8′A	0.9800
C7—H7B	0.9800	C8′—H8′B	0.9800
C7—H7C	0.9800	C8′—H8′C	0.9800
O1′—C1′	1.395 (4)	O1W—H1WA	0.8499
O3′—C3′	1.425 (3)	O1W—H1WB	0.8500
O3′—H3′O	0.8400		

C1—O1—C7	111.5 (2)	C6′—O6′—H6′O	109.5
C2—O2—H2O	109.5	C7′—N2′—C2′	121.6 (3)
C3—O3—H3O	109.5	C7′—N2′—H2′N	118 (3)
C1—O5—C5	111.2 (2)	C2′—N2′—H2′N	120 (3)
C6—O6—H6O	109.5	O1′—C1′—O5′	106.3 (2)
O1—C1—O5	107.0 (2)	O1′—C1′—C2′	108.3 (2)
O1—C1—C2	108.7 (2)	O5′—C1′—C2′	109.5 (2)
O5—C1—C2	109.6 (2)	O1′—C1′—H1′	110.9
O1—C1—H1	110.5	O5′—C1′—H1′	110.9

O5—C1—H1	110.5	C2′—C1′—H1′	110.9
C2—C1—H1	110.5	N2′—C2′—C1′	110.7 (2)
O2—C2—C1	108.0 (2)	N2′—C2′—C3′	111.7 (2)
O2—C2—C3	111.9 (3)	C1′—C2′—C3′	109.7 (2)
C1—C2—C3	108.0 (2)	N2′—C2′—H2′	108.2
O2—C2—H2	109.6	C1′—C2′—H2′	108.2
C1—C2—H2	109.6	C3′—C2′—H2′	108.2
C3—C2—H2	109.6	O3′—C3′—C4′	109.9 (2)
O3—C3—C2	108.2 (3)	O3′—C3′—C2′	109.6 (2)
O3—C3—C4	112.3 (3)	C4′—C3′—C2′	109.9 (2)
C2—C3—C4	110.5 (2)	O3′—C3′—H3′	109.2
O3—C3—H3	108.6	C4′—C3′—H3′	109.2
C2—C3—H3	108.6	C2′—C3′—H3′	109.2
C4—C3—H3	108.6	O4′—C4′—C3′	108.3 (2)
O1′—C4—C5	105.9 (2)	O4′—C4′—C5′	108.9 (2)
O1′—C4—C3	110.2 (2)	C3′—C4′—C5′	109.3 (2)
C5—C4—C3	109.9 (2)	O4′—C4′—H4′	110.1
O1′—C4—H4	110.2	C3′—C4′—H4′	110.1
C5—C4—H4	110.2	C5′—C4′—H4′	110.1
C3—C4—H4	110.2	O5′—C5′—C6′	105.9 (2)
O5—C5—C6	108.1 (2)	O5′—C5′—C4′	109.7 (2)
O5—C5—C4	109.5 (2)	C6′—C5′—C4′	114.2 (3)
C6—C5—C4	112.9 (2)	O5′—C5′—H5′	109.0
O5—C5—H5	108.7	C6′—C5′—H5′	109.0
C6—C5—H5	108.7	C4′—C5′—H5′	109.0
C4—C5—H5	108.7	O6′—C6′—C5′	111.2 (3)
O6—C6—C5	111.2 (3)	O6′—C6′—H6′A	109.4
O6—C6—H6A	109.4	C5′—C6′—H6′A	109.4
C5—C6—H6A	109.4	O6′—C6′—H6′B	109.4
O6—C6—H6B	109.4	C5′—C6′—H6′B	109.4
C5—C6—H6B	109.4	H6′A—C6′—H6′B	108.0
H6A—C6—H6B	108.0	O7′—C7′—N2′	122.7 (3)
O1—C7—H7	109.5	O7′—C7′—C8′	121.6 (3)
O1—C7—H7B	109.5	N2′—C7′—C8′	115.7 (3)
H7—C7—H7B	109.5	C7′—C8′—H8′A	109.5
O1—C7—H7C	109.5	C7′—C8′—H8′B	109.5
H7—C7—H7C	109.5	H8′A—C8′—H8′B	109.5
H7B—C7—H7C	109.5	C7′—C8′—H8′C	109.5
C1′—O1′—C4	115.7 (2)	H8′A—C8′—H8′C	109.5
C3′—O3′—H3′O	109.5	H8′B—C8′—H8′C	109.5
C4′—O4′—H4′O	109.5	H1WA—O1W—H1WB	109.5
C1′—O5′—C5′	115.4 (2)		

C7—O1—C1—O5	−76.2 (3)	C4—O1′—C1′—C2′	164.8 (2)
C7—O1—C1—C2	165.5 (3)	C5′—O5′—C1′—O1′	−175.5 (2)
C5—O5—C1—O1	176.3 (2)	C5′—O5′—C1′—C2′	−58.7 (3)
C5—O5—C1—C2	−66.1 (3)	C7′—N2′—C2′—C1′	122.2 (3)
O1—C1—C2—O2	56.1 (3)	C7′—N2′—C2′—C3′	−115.3 (3)

O5—C1—C2—O2	−60.5 (3)	O1′—C1′—C2′—N2′	−64.9 (3)
O1—C1—C2—C3	177.3 (3)	O5′—C1′—C2′—N2′	179.6 (2)
O5—C1—C2—C3	60.7 (3)	O1′—C1′—C2′—C3′	171.5 (2)
O2—C2—C3—O3	−59.7 (3)	O5′—C1′—C2′—C3′	55.9 (3)
C1—C2—C3—O3	−178.4 (3)	N2′—C2′—C3′—O3′	59.1 (3)
O2—C2—C3—C4	63.7 (3)	C1′—C2′—C3′—O3′	−177.8 (2)
C1—C2—C3—C4	−55.0 (3)	N2′—C2′—C3′—C4′	180.0 (2)
O3—C3—C4—O1′	−69.3 (3)	C1′—C2′—C3′—C4′	−57.0 (3)
C2—C3—C4—O1′	169.7 (3)	O3′—C3′—C4′—O4′	−64.2 (3)
O3—C3—C4—C5	174.2 (3)	C2′—C3′—C4′—O4′	175.2 (2)
C2—C3—C4—C5	53.3 (3)	O3′—C3′—C4′—C5′	177.3 (2)
C1—O5—C5—C6	−173.6 (3)	C2′—C3′—C4′—C5′	56.7 (3)
C1—O5—C5—C4	63.0 (3)	C1′—O5′—C5′—C6′	−177.5 (3)
O1′—C4—C5—O5	−174.8 (2)	C1′—O5′—C5′—C4′	58.8 (3)
C3—C4—C5—O5	−55.8 (3)	O4′—C4′—C5′—O5′	−174.1 (2)
O1′—C4—C5—C6	64.7 (3)	C3′—C4′—C5′—O5′	−55.9 (3)
C3—C4—C5—C6	−176.3 (3)	O4′—C4′—C5′—C6′	67.2 (3)
O5—C5—C6—O6	−70.0 (3)	C3′—C4′—C5′—C6′	−174.6 (3)
C4—C5—C6—O6	51.4 (3)	O5′—C5′—C6′—O6′	−56.3 (3)
C5—C4—O1′—C1′	−134.7 (3)	C4′—C5′—C6′—O6′	64.5 (4)
C3—C4—O1′—C1′	106.4 (3)	C2′—N2′—C7′—O7′	−1.6 (5)
C4—O1′—C1′—O5′	−77.6 (3)	C2′—N2′—C7′—C8′	177.0 (3)

Hydrogen-bond geometry (Å, º)

D—H···*A*	*D*—H	H···*A*	*D*···*A*	*D*—H···*A*
O2—H2*O*···O1*W*	0.84	1.89	2.702 (4)	163
O3—H3*O*···O5′	0.84	1.89	2.714 (3)	165
O6—H6*O*···O1^{i}	0.84	2.57	3.143 (3)	126
O6—H6*O*···O2^{i}	0.84	2.01	2.814 (3)	160
O3′—H3′*O*···O6^{i}	0.84	1.89	2.729 (3)	175
O4′—H4′*O*···O3^{ii}	0.84	1.93	2.767 (3)	171
O6′—H6′*O*···O3′^{iii}	0.84	1.95	2.755 (3)	160
N2′—H2′*N*···O2^{i}	0.88 (4)	2.33 (5)	3.133 (4)	154 (4)
N2′—H2′*N*···O5^{i}	0.88 (4)	2.62 (4)	3.312 (4)	137 (3)
O1*W*—H1*WA*···O7′^{iv}	0.85	2.17	2.833 (5)	135
O1*W*—H1*WB*···O6′^{iii}	0.85	1.95	2.795 (4)	178

Symmetry codes: (i) −*x*+1, *y*+1/2, −*z*+1/2; (ii) −*x*+2, *y*+1/2, −*z*+1/2; (iii) −*x*+2, *y*−1/2, −*z*+1/2; (iv) −*x*+3/2, −*y*+1, *z*−1/2.

Selected structural parameters[a] *in (III)–(VIII).*

Parameter	*β*ManOCH$_3$ (III) (residue **a**)[b]	*β*ManOCH$_3$ (VIII)	*β*GlcNAcOR (III) (residue **b**)	*β*GlcNAcOH (V) (residue **a**)	*β*GlcNAcOR (V) (residue **b**)	*β*GlcNAcOH (VI) (residue **a**)	*β*GlcNAcOR (VI) (residue **b**)	*β*GlcNAcOCH$_3$ (IV$_A$)[c]	*β*GlcNAcOCH$_3$ (IV$_B$)	*β*GlcNAcOH (VII)

Bond lengths (Å)										
C1—C2	1.523	1.524	1.522	1.522	1.522	1.526	1.515	1.533	1.531	1.534
C2—C3	1.527	1.532	1.532	1.521	1.531	1.527	1.517	1.530	1.534	1.530
C3—C4	1.531	1.526	1.515	1.531	1.516	1.520	1.516	1.527	1.525	1.521
C4—C5	1.529	1.535	1.532	1.536	1.507	1.519	1.537	1.526	1.523	1.528
C5—C6	1.512	1.516	1.512	1.501	1.499	1.512	1.516	1.519	1.514	1.514
C1—O1	1.395	1.390	1.393	1.389	1.389	1.361	1.395	1.389	1.387	1.390
C1—O5	1.429	1.422	1.430	1.427	1.429	1.418	1.414	1.418	1.423	1.433
C2—N2			1.453	1.450	1.446	1.450	1.460	1.456	1.455	1.457
C2—O2	1.419	1.424								
C3—O3	1.427	1.430	1.424	1.430	1.424	1.421	1.431	1.430	1.424	1.430
C4—O4/O1	′1.438	1.439	1.420	1.448	1.425	1.448	1.422	1.424	1.425	1.435
C5—O5	1.439	1.427	1.440	1.429	1.436	1.438	1.427	1.443	1.435	1.448
C6—O6	1.430	1.432	1.412	1.413	1.415	1.419	1.423	1.430	1.430	1.416
O1—CH_3	1.443	1.435						1.445	1.439	
N2—C_{car}			1.348	1.332	1.321	1.345	1.317	1.342	1.343	1.346
C_{car}—O_{car}			1.237	1.243	1.246	1.231	1.230	1.237	1.235	1.235
C_{car}—CH_3			1.504	1.497	1.495	1.490	1.506	1.508	1.510	1.508
O3···O5	2.714	2.690		2.796		3.311				
O3···O1	2.984	2.989		3.099		2.886				
O3···O6	3.803			2.875		4.828				
Bond angles (°)										
C5—O5—C1	111.24	112.05	115.47	112.59	112.13	114.49	112.92	111.9	111.8	114.95
O5—C1—O1	106.96	108.28	106.41	107.71	106.87	112.46	107.68	107.64	107.76	111.89
C1—O1—CH_3	111.42	113.54						112.39	112.43	
C1′—O1′—C4			115.67		117.07		116.34			
C2—N2—C_{car}			121.69	122.90	124.91	124.83	123.89	122.6	121.8	122.17
N2—C_{car}—C_{Me}			115.79	116.33	115.91	115.29	115.80	116.1	116.4	116.01
N2—C_{car}—O_{car}			122.73	121.49	121.26	123.94	123.89	122.6	122.9	123.06
O_{car}—C_{car}—C_{Me}			121.47	122.17	122.73	120.73	120.24	121.3	120.7	120.93
O3—H···O5	164.39	157.07		132.75		142.95				
Torsion angles (°)										
C1—C2—C3—C4	-55.06	-51.08	-56.83	-52.05	-48.16	-54.44	-56.66	-47.9	-50.1	-54.60

C1—O5—C5—C4	62.96	63.92	58.79	61.82	67.41	60.36	59.12	69.3	66.5	62.45
C3—C4—C5—O5	-55.87	-57.57	-55.94	-54.82	-58.61	-55.73	-54.68	-63.90	-56.56	-58.35
O5—C5—C6—O6	-70.12	76.26	-56.29	-60.61	58.55	-74.59	-65.53	64.3	65.2	-60.71
C1—C2—N2—C_{car}			122.01	100.50	113.72	138.69	100.49	108.2	100.0	140.89
C3—C2—N2—C_{car}			-115.23	-135.18	-122.48	-98.90	-136.95	-128.1	-137.2	-96.77
C2—N2—C_{car}—C_{Me}			177.07	-173.70	178.39	-179.61	-173.90	179.1	-179.1	169.86
H2—C2—N2—C_{car}			3.65	-12.50	-6.73	24.94	-16.95	-9.80	-19.20	22.25
C2—N2—C_{car}—O_{car}			-1.58	5.16	-5.28	-2.07	2.94	-1.2	0.8	-9.65
C2—C3—O3—H	156.03	168.11	-81.76	-148.65	-152.13	143.25	-51.20	-157.87	-112.65	-118.33
C4—C3—O3—H	33.78	44.57		91.57		23.13				
H3—C3—O3—H	-86.36	-75.20		-31.91		-97.92				
C3—C4—O4—H			133.26		125.98		99.93	70.45	56.71	43.27
C5—C6—O6—H	-140.49	101.86	166.57	80.71	-178.65	88.34	-91.61	-80.28	-80.95	-159.60
C2—C1—O1—CH_3/H (φ_a)[b]	165.59	173.61		161.75		-163.57		169.48	175.49	-170.76
O5—C1—O1—CH_3/H (φ_b)	-76.16	-66.00		-80.33		74.69		-71.05	-66.63	68.64
H1—C1—O1—CH_3/H (φ_c)	44.23	53.87		40.24		-42.50		49.18	54.04	-52.07
C2′—C1′—O1′—C4 (φ'_a)			164.72		151.65		161.49			
O5′—C1′—O1′—C4 (φ'_b)			-77.63		-90.28		-79.57			

H1′—C1′—O1′—C4 (φ'_c)	42.92	30.20	36.30
C1′—O1′—C4—C3 (ψ'_a)	106.41	77.27	133.47
C1′—O1′—C4—C5 (ψ'_b)	-134.78	-162.31	-106.89
C1′—O1′—C4—H4 (ψ'_c)	-15.50	-44.86	11.69

Notes: (*a*) to simplify structural comparisons between like residues in (III)–(VII), atom representations in this table differ, in some cases, from those found in their X-ray structures. These changes involve primed atoms that normally distinguish atoms in the two residues of (III), (V) and (VI), atom numberings, and/or the use of `car′ subscripts to denote carbonyl C and O atoms in the *N*-acetyl side chains of (III)–(VII). (*b*) *O*-Glycosidic torsion angles φ, φ' and ψ', which specify the rotational properties of the C1—O1, C1′—O1′ and O1′—C4 bonds, respectively, in (III), (V) and (VI), can be defined by three different vicinal pathways, which are distinguished by the a–c subscripts. (*b*) See Scheme 1 for the definitions of residues **a** and **b** in (III), (V) and (VI). (*c*) (IVA) and (IVB) refer to the two molecules of (IV) observed in the reported crystal structure (Hu *et al.*, 2011). (*d*) Data for (IV)–(VII) were taken from Hu *et al.* (2011), Mo (1979), Mo & Jensen (1978) and Mo & Jensen (1975), respectively. **[Two notes (*b*)?]**

^{1}H and ^{13}C chemical shifts for disaccharide (III)

				^{1}H chemical shift (ppm)[a]					
Residue	H1	H2	H3	H4	H5	H6	H6′	OC**H$_3$**	–NHCOC**H$_3$**
(residue **a**)[b]	4.555	4.007	3.729	*β*Man 3.672	3.417	3.833	3.650	3.514	
(residue **b**)	4.524	3.728	3.544	*β*GlcNAc 3.441	3.498	3.920	3.729		2.051
				^{13}C chemical shift (ppm)[a]					
	C1	C2	C3	C4	C5	C6	O**C**H$_3$	–NH**C**OCH$_3$	–NHCO**C**H$_3$
(residue **a**)	103.57	72.36	74.41	*β*Man 80.18	77.43	63.17	59.50		
(residue **b**)	104.27	58.24	76.11	*β*GlcNAc 72.45	78.57	63.28		177.27	24.79

Notes: (*a*) in ^2H_2O at 22 °C; in ppm relative to external DSS. H6′ is defined as the more shielded C6 hydrogen; (*b*) see Scheme 1 for definitions of the **a** and **b** residues in (III).