ELSEVIER

Carbohydrate Research 337 (2002) 2301–2310

CARBOHYDRATE RESEARCH

www.elsevier.com/locate/carres

Crystal structure of penta-*O*-acetyl-β-D-galactopyranose with modeling of the conformation of the acetate groups

Devron P. Thibodeaux,[a] Glenn P. Johnson,[a] Edwin D. Stevens,[b,*] Alfred D. French[a,*]

[a]*Southern Regional Research Center, US Department of Agriculture, 1100 Robert E. Lee Boulevard, New Orleans, LA 70124, USA*
[b]*Department of Chemistry, University of New Orleans, New Orleans, LA 70148, USA*

Received 6 May 2002; accepted 1 June 2002

Dedicated to Professor Derek Horton on the occasion of his 70th birthday

Abstract

The crystal structure of penta-*O*-acetyl-β-D-galactopyranose was determined with Mo K_α radiation at 150 K to $R = 0.029$. The space group is $P2_12_12_1$, and the unit cell dimensions are, $a = 8.348$, $b = 9.021$ and $c = 25.418$ Å. The ring has the usual $^4C_1$ shape and O-6 is in the *tg* position as frequently observed for sugars having the axial *galacto* configuration at C-4. Conformations of the acetate groups were compared with those from the literature. Nearly eclipsed, '*Z*' conformations are found for the ester moieties, and the torsion angles for the sequence involving the ring hydrogen, carbon, alkoxy oxygen and carbonyl carbon for both **1** and related compounds are eclipsed-to-gauche. Orientations and conformations of the acetate substituents were modeled with both MM3 molecular mechanics and various levels of quantum mechanics theory. Higher levels of theory and more complete models provided better prediction of the experimental observations.

Keywords: Ab initio; Aldohexose; Boltzmann distribution; Carbohydrate; DFT; Isopropyl acetate; Triacetyl tetrahydropyran

1. Introduction

Increased use of agricultural products in new bio-based materials is likely to entail structural modifications to change their properties. Knowledge-based modification should be facilitated by a more complete understanding of structure–function relationships. One of the most common chemical modifications is acetylation. This reaction occurs naturally to some extent and is widely practiced in the laboratory and in industry. While there are numerous crystal structures of acetylated carbohydrates, no crystal structure for the title compound, penta-*O*-acetyl-β-D-galactopyranose (**1**), was available in the Cambridge Structural Database (CSD).[1] The following report remedies that situation. In addition, we have used **1** and related compounds to further explore the hypothesis[2–5] that crystal lattices induce more or less random deformations of intrinsically preferred molecular structures, similar to the distribution of geometries that would be found for molecules in an ideal gas.

2. Results and discussion

The details of the crystal structure determination are given in Table 1, and the ORTEP[6] depiction with atom numbering is shown in Fig. 1. Moderate thermal motion is greatest for the terminal ends of the acetate moieties, involving both the carbonyl oxygen atoms and the methyl groups, especially for O-7, C-8, O-9, C-12 and O-13. The crystal packing is illustrated in Fig. 2. The interacting extended chains from the C-5 atoms are visible in this *a*-axis projection. There are no short interatomic contacts in the crystal.

Bond lengths involving the carbon and oxygen atoms in Table 2 vary substantially. The equal C-1–O-1 and C-1–O-5 bond lengths (1.408 Å) at the anomeric center

* Corresponding authors. Tel.: +1-504-2806856 (E.D.S.); Tel.: +1-504-2864410 (A.D.F.)
E-mail addresses: estevens@uno.edu (E.D. Stevens), afrench@srrc.ars.usda.gov (A.D. French).

0008-6215/02/$ - see front matter
PII: S0008-6215(02)00266-5

are somewhat unusual for a β-configured equatorial anomeric bond. Such C-1–O-1 bonds in other crystals are shorter on average than C-1–O-5 bonds by 0.034 Å.[7] In keeping with the usual situation, both of those bonds are shorter than the C-5–O-5 bond (1.432 Å).

Table 1
Crystal data and structure refinement for penta-*O*-acetyl-β-D-galactopyranose (**1**)

Empirical formula	$C_{16}H_{22}O_{11}$
Formula weight	390.34
Temperature (K)	150(2)
Wavelength (Å)	0.71073
Crystal system, space group	Orthorhombic, $P2_12_12_1$
Unit cell dimensions	
a (Å)	8.3482(6)
b (Å)	9.0212(8)
c (Å)	25.418(2)
Volume (Å^3)	1914.2(3)
Z, calculated density (Mg/m^3)	4, 1.354
Absorption coefficient (mm^{-1})	0.116
F(000)	824
Crystal size (mm)	$0.15 \times 0.15 \times 0.15$
Theta range for data collection (°)	1.60–23.27
Limiting indices	$-9 \le h \le 8$, $-10 \le k \le 9$, $-26 \le l \le 28$
Reflections collected/unique	9055/2755 [$R_{int} = 0.0728$]
Completeness to theta = 23.27	100.0%
Absorption correction	Empirical
Max. and min. transmission	1.000000 and 0.880505
Refinement method	Full-matrix least-squares on F^2
Data/restraints/parameters	2755/0/331
Goodness-of-fit on F^2	0.561
Final *R* indices [$I > 2\sigma(I)$]	$R_1 = 0.0293$, $wR_2 = 0.0398$
R indices (all data)	$R_1 = 0.0869$, $wR_2 = 0.0456$
Absolute structure parameter	6.0(14)
Largest diff. peak and hole (e Å^{-3})	0.133 and −0.132

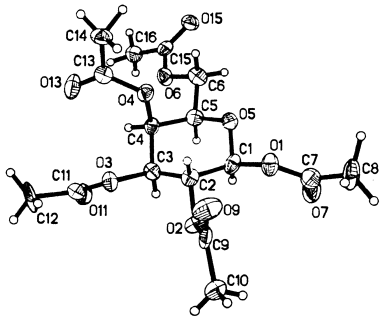


Fig. 1. ORTEP[6] for **1**, showing the numbering of the atoms. The thermal ellipsoids are shown at 50% probability.

The distribution of C–O lengths in **1** was checked with a B3LYP/6-31G* quantum mechanics (QM) geometry optimization of the crystal structure with all torsion angles constrained to their values in the crystal. Those QM values were 1.398, 1.403 and 1.430 Å. While systematically shorter, the calculated values support the unusually similar experimental lengths (with $\sigma = 0.004$ Å) for C-1–O-1 and C-1–O-5.

Two other geometry optimizations of **1** were carried out. The first constrained all of the torsions except those involving hydrogen atoms. All constraints were removed in the last energy minimization, which yielded an energy 2.02 kcal/mol lower. This suggests that there is relatively little torsional deformation in the crystal.

Seven 2,3,4,6 tetra-*O*-acetyl-β-D-galactopyranose fragments were found in the CSD, differing chemically only in their substitution at C-1. These fragments are identified by the refcodes of their parent crystal structures, i.e., DUDQUO,[8] GEYMIG,[9] OACGAP,[10] TEPSAI (two rings),[11] WIPFAC[12] and ZUTPIN,[13] defined in Table 3. The average lengths of the bonds in their acetate groups at the individual positions are also listed in Table 2. Given the extensive variation in these other determinations, the bond lengths in **1** are within the expected ranges. There was another structure, KACDAT (2,3,4,6-tetra-*O*-acetyl-1-bromo-D-galactopyranosyl cyanide).[14] However, despite an *R* value of 0.047, several of its bond lengths were implausibly short. For example, C-6–O-6 was 1.387 Å, the corresponding carbonyl bond was 1.104 Å, and the carbonyl carbon-to-methyl distance at the C-4 position was 1.402 Å. These short bond lengths in KACDAT probably result from the dominance of the X-ray scattering by electron-rich bromine.

Table 3 also gives the conversion of the refcodes for six other aldohexapyranose penta-acetates, PACALP,[15] PACDGP,[16] PAIDOP,[17] PATPYS,[18] ZZZQPG01[19] AND ZZZRCA01,[20] along with their crystal densities. The crystal density (Table 1) of **1**, 1.354 Mg/m^3, is similar to densities for these related molecules. All structures in Table 3 were determined at room temperature.

Cremer–Pople[21] ring puckering parameters are shown in Table 4 for **1**, and the other structures listed in Table 3. With a θ value of 2.45°, the ring shape of **1** is close to a perfect chair ($\theta = 0.0°$). This is also shown by the torsion angles of the ring (Table 4), with nearly equal (but with opposite sign) values for the three pairs of torsion angles (those centered on C-5–O-5 and O-5–C-1, C-4–C-5 and C-1–C-2, and C-2–C-3 and C-3–C-4). The puckering amplitude (0.601 Å) of **1** indicates somewhat more puckering than most pyranosyl rings. Crystalline β-D-galactose[22] (BDGLOS10) has θ and *Q* values of 4.83° and 0.598 Å, respectively. Seventeen other α- and β-galactose residues with limited or no substitution had θ values ranging from 1.04 to 12.5°

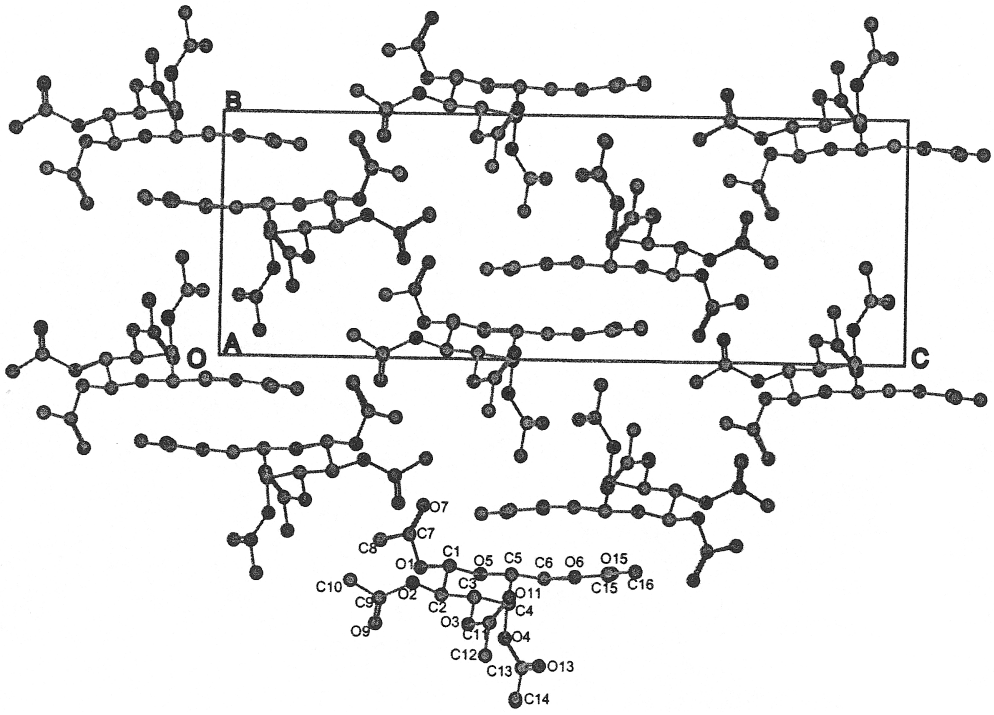


Fig. 2. Packing of **1** and its unit cell in projection along the *a*-axis. The carbon and oxygen atoms are labeled and hydrogen atoms are omitted.

and Q values ranging from 0.523 to 0.608 Å. Earlier modeling studies of native galactopyranose indicated substantial preference for the $^4C_1$ shape.[23]

All structures in Table 4 happen to have $^4C_1$ conformations, even when there are four axial substituents. It appears from the penta-acetate structures that high numbers of acetate substituents in axial configurations tend to flatten the ring, with reduced puckering amplitude (Q). All except the extra-flat PAIDOP are within usual ranges.

The primary alcohol group in **1** is in the *tg* position (O-5–C-5–C-6–O-6 = 178.4°), one of two favored orientations for sugars with the *galacto* configuration at C-4. In BDGLOS10, it has the *gt* orientation (57.8°).[22] Tvaroška and co-workers.[24] recently studied different orientations of primary alcohol groups by NMR and QM calculations. Their work on molecules with the *galacto* configuration showed that, when hydrogen bonding was not permitted, the *gt* and *tg* conformations were similar in energy, but the *gg* conformation had higher energy. When the configuration at *C*-4 is *gluco* (with *O*-4 equatorial), gt and gg are the preferred forms. The variations in the substituent orientations for **1** and the seven other β-galactosyl residues are indicated in Fig. 3, and values for their primary alcohol groups are given in Table 3.

The orientations of acetate groups relative to the ring have been of interest for more than 30 years.[13,25–27] The acetate groups themselves are mostly planar because of the sp^2 hybridization of the carbonyl carbon. As substituents of primary alcohols, they tend to be in gauche orientations with respect to both of the hydrogen atoms on C-6 and trans to C-5. When acetates are substituents of secondary alcohols on rings, their carbonyl carbon atoms tend to eclipse the methine hydrogen atoms on the sugar rings. However, the original modeling work was unable to confirm that finding.[27] The carbonyl oxygen atoms eclipse the ring carbon atoms, making the so- called '*Z*' conformation, with a number of explanations given by Eliel and Wilen.[28] Of particular importance is the partial double-bond character of the acyloxy bond. This is indicated by their short lengths (ca. 1.345 Å, Table 2).

Initially, we attempted to rationalize the disposition and conformation of the acetate groups that substitute the secondary hydroxyl groups with a computer model of isopropyl acetate (**2**, Fig. 4). The propyl moiety represents three atoms of the sugar ring in this model and avoids issues relating to axial and equatorial substitution. It contains all of the atoms used to define the torsional energies for movement of the acetate groups. From Fig. 1, it appears that there would be minimal contact among neighboring acetate groups, so this reduced model seemed appropriate. The torsion angle definitions for **2** are shown in Fig. 4.

An overview of the energy surface of **2** was obtained with an empirical force field (MM3)[29–31] calculation, shown in Fig. 5(a). Full rotation about the alkoxy bond, defined by the H–C–O–C torsion angle ϕ, is considerably easier than rotation about the acyloxy

bond, defined by the C–O–C=O torsion ψ, with barriers of only 4.5 kcal/mol for ϕ and 16 kcal/mol for ψ. The latter barrier, for rotation about a partial double bond, was compatible with the quantum mechanics (QM) study of less-hindered methyl acetate (13.2 kcal at the MP3/6-31G** level) by Wiberg and Laidig.[32] Cox, Armishaw and Wardell used another empirical function to compute the energy for rotation about the acyloxy bond in the C-3 position in ZUTPIN.[13]

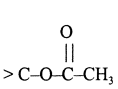
The experimental torsion angles from the four acetates on the secondary and anomeric hydroxyl groups of **1** are plotted on Fig. 5(a), which seems to provide a good accounting for the observations. Because of the rotational symmetry of this surface, there are two identical global minima. The barrier between them of about 1.0 kcal/mol corresponds to a structure with eclipsing of the H–C and O–C (alkoxy) bonds and of the O–C (acyloxy) and C=O bonds.

The 167 experimental acetate group structures on secondary alcohol positions from the literature as well as **1** (see Section 4) are all found in the region of the global minima (Fig. 5(b)). This confirms the general utility of the reduced model and the approximate validity of the MM3 calculation. Consistent with the greater torsional rigidity of the acyloxy bond, the range of observed crystal structures in the C–O–C=O direction is considerably less than the range of observed H–C–O–C values.

Table 2
Heavy-atom bond lengths for **1** and related structures (standard deviations in parentheses)

O(5)–C(1)	1.408(5)
C(1)–C(2)	1.530(6)
C(2)–C(3)	1.485(5)
C(3)–C(4)	1.526(5)
C(4)–C(5)	1.514(5)
O(5)–C(5)	1.432(4)
C(5)–C(6)	1.490(6)

Acetate groups



Atom #	C–O	O–C	C=O	C–CH_3
1	1.408(4)	1.385(5)	1.212(5)	1.473(6)
2	1.430(4)	1.348(4)	1.201(4)	1.485(6)
3	1.440(4)	1.371(4)	1.195(4)	1.513(6)
4	1.454(4)	1.357(4)	1.196(4)	1.505(6)
6	1.436(4)	1.346(4)	1.209(4)	1.466(6)
Average values from seven 2,3,4,6-tetra-*O*-acetyl-D-galactose residues*				
2	1.444(16)	1.355(14)	1.195(13)	1.493(19)
3	1.440(13)	1.346(21)	1.186(21)	1.479(14)
4	1.444(9)	1.346(8)	1.192(8)	1.488(14)
6	1.441(11)	1.335(27)	1.190(9)	1.483(15)
Average	1.442(12)	1.345(19)	1.191(13)	1.486(16)

* DUDQUO,[8] GEYMIG,[9] OACGAP,[10] TEPSAI (two rings),[11] WIPFAC[12] and ZUTPIN.[13]

Despite finding all of the structures in the region of lowest energy, however, the distribution of observed structures on the calculated surface was not fully satisfying. A major issue was that the barrier at $\phi = 0°$, $\psi = 0°$ seemed too high to account for such a large fraction of the observed conformations in that region. Another problem was that the observed structures did not extensively populate the regions with the ϕ values larger than $\pm 40°$, the location of the global minima. A third problem is that the relative MM3 energy for the 180°, 0° conformations was only 1.34 kcal/mol, suggesting that some structures could be found in this secondary region. After all, three observed ϕ,ψ points in Fig. 5(b) have energies corresponding to more than 1.4 kcal/mol. On this surface the average energy for the crystal structures was 0.73 kcal/mol.

To see if the fit between the observed structures and an energy surface could be improved, we calculated the energies for variation of ϕ for **2** with ψ held to 0°, using HF, B3LYP and MP2 QM theory, all with the 6-31G* basis set. The QM results were very similar to each other and had the same general features as the MM3 results. However, the eclipsed 0°, 0° energy was lower by 0.3 kcal at the MP2 and HF levels than the MM3 result, and the B3LYP level was 0.6 kcal/mol lower. Also, all QM values for the $\phi = 180°$, $\psi = 0°$ conformation were higher than the MM3 value by about 1 kcal/mol. That higher value would reduce the chances of finding experimental structures in that region. Another difference was that the B3LYP level gave an energy about 3 kcal/mol higher than MM3 at the $\phi = 120°$, $\psi = 0°$ barrier conformation, and the other two levels were 0.75 kcal/mol higher still. Because the largest problem with the fit of experimental data with the MM3 surface was the high barrier at $\phi = 0°$, $\psi = 0°$, we chose B3LYP theory to construct a limited-range energy surface.

Despite the reduced barrier at $\phi = 0°$, $\psi = 0°$, the B3LYP/6-31G* map for **2**, shown in Fig. 5(c), had a similar problematic distribution of the conformations. Most of the structures were still between the minima, rather than distributed around them. The average corresponding energy was reduced to 0.34 kcal/mol. We were able to calculate a surface with B3LYP/6-311 + G* theory (not shown) and optimized geometries. The average energy was reduced further to 0.29 kcal/mol and the distribution problem remained.

At this point, it seemed that more time-consuming levels of QM were unlikely to improve the prediction of the observed structures. Instead, a chemically more complete model of the structure was needed. Our improved model includes a tetrahydropyran ring, substituted with three adjacent equatorial acetyl groups (**3**,

Table 3
Other crystals examined in this work

REFCODE	Ref.	Compound	*R* factor	Density (Mg/m^3)	O-6 Conf. O-5–C-5–C-6–O-6 (°)
Tetra-acetylated galactose residues					
DUDQUO	8	5-Phenyl-1-(2,3,4,6-tetra-*O*-acetyl-β-D-galacto-pyranosyl)-4-trifluoromethyl-1,2,3-triazole	0.046	1.407	77.6 *gt*
GEYMIG	9	3,5-Di-*O*-acetyl-6-deoxy-1,2-*O*-isopropylidene-6-thio-α-D-glucofuranose-6-*S*-(2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranosyl)	0.051	1.283	70.2 *gt*
OACGAP	10	1,2,3,6-Tetra-*O*-acetyl-4-*O*-(2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranosyl)-α-D-galactopyranose	0.052	1.309	175.8 *tg*
TEPSAI (2)*	11	2-(Indol-3-yl)ethyl 2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranoside	0.062	1.319	73.9, 75.6 *gt*
WIPFAC	12	1-Isothiocyanato 2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranoside	0.051	1.330	178.9 *tg*
ZUTPIN	13	Phenyl 2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranoside	0.048	1.291	177.93 *tg*
Aldohexopyranose penta-acetates					
PACALP	15	Penta-*O*-acetyl-α-D-altropyranose	0.043	1.352	−71.5 *gg* eq *O*-4
PACDGP	16	Penta-*O*-acetyl-α-D-gulopyranose	0.048	1.316	−60.6 *gg* ax *O*-4
PAIDOP	17	Penta-*O*-acetyl-α-D-idopyranose	0.046	1.346	−167.1 *tg* ax *O*-4
PATPYS	18	Penta-*O*-acetyl-α-D-talopyranose	0.041	1.380	159.9 *tg* ax *O*-4
ZZZQPG01	19	Penta-*O*-acetyl-β-D-glucopyranose	0.073	1.299	79.6 *gt* eq *O*-4
ZZZRCA01	20	Penta-*O*-acetyl-α-D-glucopyranose	0.046	1.314	64.7 *gt* eq *O*-4

Other carbohydrate acetates with $R<0.05$ that furnished acetate group conformations in Fig. 5*

BARQAM, BIKWOH-10 (2), BOSWOV, BUFTIF (2), CAPZAU10, CAVBEG10 (2), CEQFAF, COFLUE, CUJJIA, DIDYUK, DIGTAO, FANKAG, FIWLIG, GUNSEN, HIPMAU, JERJIZ, KACJED, KOSLUZ, LETTAF, MEAGPY, MTAGLV, NOYVOM (2), NUZRUV (2), PULCHB11, QAMVEF (2), RIKHIC, RIWZIG (2), RUGZIC01, SAPMEB, TERWIW, VAFPOH, VIMZUM10, WEHTIM (2), WODTOY, XAZLAL (2), YAFMAT, YODROY (2), ZETNOB, ZIZHUL, ZIZJAT, ZIZJEX, ZIZJIB

* REFCODES followed by (2) indicate that there were two carbohydrate rings that furnished acetate groups.

Fig. 4). The resulting MM3 energy surface for the central acetate group is shown in Fig. 5(e and f). The differences between them and Fig. 5(a and b) are subtle, but provide a better fit to the observations. In particular, the central barrier is substantially reduced, to just over 0.2 kcal/mol, and the minima are located closer to the $\phi = 0°$, $\psi = 0°$ conformation.

The potential energy in MM3 calculations is easily partitioned. The torsional energies dominate the energy surfaces, with ranges of more than 19 kcal/mol for both **2** and **3**. The high values arise from rotations about the partial double bonds. The major reason for the lowered central barrier was relative dipole–dipole stabilization at the $\phi = 0°$, $\psi = 0°$ conformation for **3**, compared to that for **2**. On Fig. 5(e and f), the barrier at $\phi = 120°$, $\psi = 0°$ to complete rotation is increased by about 1 kcal/mol, and the most strained conformations of **2** are even more strained in **3**. The average energy on the MM3 surface for **3** was reduced to 0.25 kcal/mol, but the distribution was still not satisfying.

A restricted-range, B3LYP/6-31G* energy surface for **3** is shown in Fig. 5(g). This surface is more consistent with the observed crystal structures. The $\phi = 0°$, $\psi = 0°$ conformation is now the global minimum, and there are secondary minima of 0.15 kcal at $\phi = \pm 28°$, $\psi = 0°$. With that improvement, the distribution becomes fairly even over the lowest energy region. However, the average energy for the observed structures is about 0.3 kcal/mol, lower than the value expected for random distortions in an ideal gas at room temperature (0.592 kcal/mol), according to the Boltzmann equation.

Final calculations for **2** and **3** were carried out with the B3LYP/6-31G* geometries at the B3LYP/6-311+G* level. They are shown in Fig. 5(d and h). The surface for **2** (Fig. 5(d)) is very similar to the surface for **2** with full minimization at the B3LYP/6-311+G* level (not shown), confirming the utility of the single point calculations and the geometry of the 6-31G* minimizations. The surface for **3** (Fig. 5(h)) has a deeper minimum at the $\phi = 0°$, $\psi = 0°$ conformation, compared to

Table 4
Critical torsion angles for **1** and Cremer–Pople puckering parameters for **1** and related structures

O-5–C-1–C-2–C-3	56.8	C-5–O-5–C-1–O-1	−179.4	
C-1–C-2–C-3–C-4	−53.3	O-5–C-1–O-1–C7	−95.1	
C-2–C-3–C-4–C-5	54.6	C-1–C-2–O-2–C9	103.2	
C-3–C-4–C-5–O-5	−56.9	C-2–C-3–O-3–C-11	−159.2	
C-4–C-5–O-5–C-1	61.5	C-3–C-4–O-4–C-13	−101.8	
C-5–O-5–C-1–C-2	−61.4	C-4–C-5–C-6–O-6	−58.2	
		C-5–C-6–O-6–C-15	−168.5	
Sugar	Q (Å)	θ (°)	ϕ* (°)	N axial groups
1 (GAL_AC)	0.601	2.46	107	1
DUDQUO	0.578	4.12	291	1
GEYMIG	0.561	10.52	321	1
OACGAP	0.580	1.88	81	1
TEPSAI(1)	0.584	5.67	25	1
TEPSAI(2)	0.576	5.23	49	1
WIPFAC	0.574	3.05	14	1
ZUTPIN	0.547	7.37	314	1
ZZZOPG01	0.598	4.87	314	0
ZZZRCA01	0.573	2.08	340	1
PACALP	0.538	8.52	294	3
PACDGP	0.538	6.95	305	3
PATPYS	0.527	2.95	216	3
PAIDOP	0.488	8.87	316	4

* Values of ϕ are relatively unimportant for these chair structures.

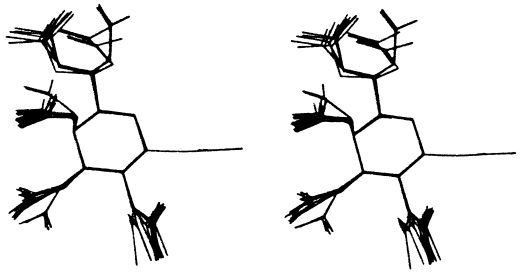


Fig. 3. Superimposition of **1** on DUDQUO, GEYMIG, OACGAP, TEPSAI (two rings), WIPFAC and ZUTPIN, shown as a stereo pair (see Table 3 for refcode definitions).

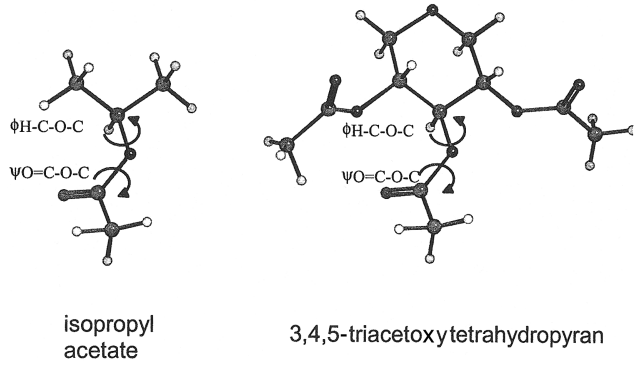


Fig. 4. The model molecules (isopropyl acetate, **2**) and 3,4,5-triacetoxytetrahydropyran (**3**) used to calculate the energy surfaces shown in Fig. 5.

Fig. 5(g). This deeper minimum results in an increase in the average energy, to 0.48 kcal/mol, and the distribution shown in Fig. 6. As proposed earlier,[2–5] the number of structures decreases exponentially as the energy increases. The curve-fit equation is Relative Frequency $= e^{-\Delta E/0.488}$, with a squared correlation coefficient (R^2) of 0.994 and a standard deviation in the fit parameter, 0.488, of 0.003. This is equivalent to the energy distribution in an ideal gas at 245 K and, in comparison to a room-temperature distribution, corresponds to an average underestimation of relative energy in our calculations of 0.108 kcal/mol. To clarify, this temperature of 245 K has nothing to do with the temperature of the determination, as the conformation of the molecule did not change when the crystal was cooled during the experiment. According to the Boltzmann equation, high-energy structures are more probable at higher temperatures. The temperature of 245 K indicates that these details of the structures are less distorted by crystal packing than would be expected for isolated molecules at room temperature. The 'temperature' in the present work is somewhat less than the values that we found earlier.[4,5]

3. Conclusions

The crystal structure of penta-*O*-acetyl-β-D-galactopyranose was, except for the knowledge of the

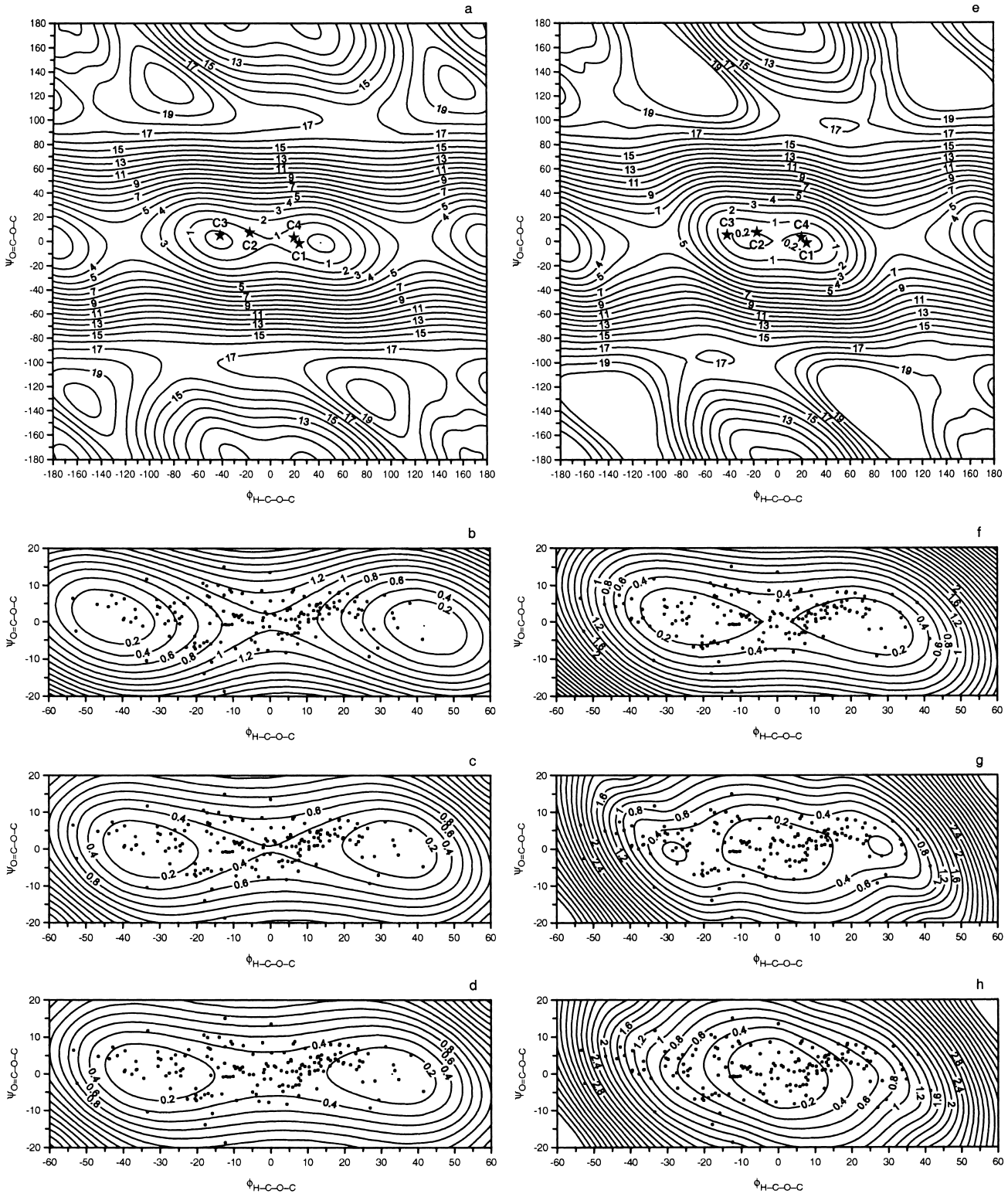


Fig. 5. Energy surfaces for isopropyl acetate (**2**) (a,b,c,d) and 3,4,5-triacetoxytetrahydropyran (**3**) (e,f,g,h) models. (a) All possible conformations of the alkoxy (H–C–O–C) and acyloxy (C–O–C=O) bonds of **2**, calculated with MM3(96), $\epsilon = 1.5$. The surface has two-fold rotational symmetry. The orientations and conformations of the acetate groups on C-1, C-2, C-3 and C-4 of **1** are shown as stars (★). (b) The central region from Fig. 5(a), with three acetate groups from **1** and from 164 other examples from the literature as dots (●). (c) Surface for **2**, similar to (b), but calculated with B3LYP/6-31G* QM and full optimization. (d) Surface for **2** calculated with single point energies at the B3LYP/6-311+G* level for B3LYP/6-31G* geometries. Even at this level for **2**, most of the crystal structures are between the two lowest energy regions. (e) Energy surface for **3**, calculated with MM3, as in (a). (f) The region of the crystal structures calculated for **3** with MM3. (g) The B3LYP/6-31G* surface for **3**. (h) The single point energy surface for **3**, calculated as in (d).

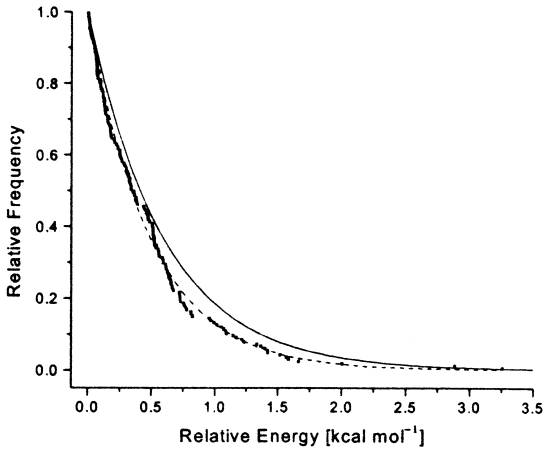


Fig. 6. Cumulative frequency plot for the conformational energies[4,5] of the 167 observed structures (°) shown in Fig. 5(h). The x-axis values are the corresponding energies of the observed structures on Fig. 5(h). The y-axis values are $y = 1-(p-1)/167$, where $p =$ the position in a sorted list of the 167 energy values, from lowest to highest. The fitted exponential decay line (dashed) corresponds to random distortions at about 245 K. The solid line corresponds to a room-temperature Boltzmann distribution. The latter line indicates that 18% of the structures have an energy greater than 1 kcal/mol, while the dashed line indicates that 14% of the structures have energies greater than 1 kcal/mol.

primary alcohol orientation, predictable. The ring had the usual shape, and the acetate groups were in their usual orientations as found in seven analogous 2,3,4,6 tetra-*O*-acetyl-β-D-galactose fragments. Both the *gt* and *tg* conformations of the primary alcohol group are almost equally probable. In this case the *tg* orientation was found, unlike the *gt* orientation for the native molecule in its crystal.

From a theoretical standpoint, acetates are interesting because they offer an opportunity to study the shape properties of carbohydrates without one of the usual obstacles, namely the severe multiple minimum problem that results from the many hydroxyl group orientations that can occur. Acetate groups are confined to a small area of conformation space. In addition, hydroxyl orientations in crystals are often different from those preferred in isolated models, not the case for the acetates.

Once a sufficiently high level of theory was employed for a reasonably complete molecular model, the energy surface provided a plausible explanation for the observed orientations and conformations of the acetate groups. Because the 3,4,5-triacetoxytetrahydropyran model was considerably more predictive than the isopropyl acetate model, we conclude that the orientations of the individual acetate groups are affected by their neighboring substituents. Partitioning of the MM3 energies suggested that this effect results from electrostatic interactions.

Table 5
Atomic coordinates ($\times 10^4$) and equivalent isotropic displacement parameters ($Å^2 \times 10^3$) for penta-*O*-acetyl-β-D-galactopyranose. U(eq) is defined as one third of the trace of the orthogonalized U_{ij} tensor

	x	y	z	U(eq)
C(1)	2592(5)	8526(5)	8408(2)	29(1)
O(1)	1505(3)	8569(3)	7985(1)	33(1)
C(2)	3886(5)	9696(5)	8310(2)	24(1)
O(2)	4792(3)	9220(3)	7864(1)	27(1)
C(3)	4936(5)	9780(5)	8781(2)	25(1)
O(3)	6108(3)	10932(3)	8706(1)	28(1)
C(4)	3960(5)	10029(4)	9281(2)	24(1)
O(4)	3124(3)	11436(3)	9233(1)	28(1)
C(5)	2692(5)	8836(5)	9324(2)	27(1)
O(5)	1698(3)	8834(3)	8865(1)	27(1)
C(6)	1644(6)	8978(6)	9796(2)	33(1)
O(6)	2715(3)	8930(3)	10238(1)	30(1)
C(7)	841(5)	7240(6)	7821(2)	43(1)
O(7)	1211(4)	6078(3)	8027(1)	52(1)
C(8)	−364(7)	7481(8)	7407(2)	52(2)
C(9)	4523(5)	9884(5)	7397(2)	30(1)
O(9)	3670(4)	10947(3)	7347(1)	46(1)
C(10)	5394(8)	9144(7)	6960(2)	39(1)
C(11)	7447(5)	10815(5)	9015(2)	36(1)
O(11)	7688(3)	9785(3)	9300(1)	39(1)
C(12)	8513(7)	12159(7)	8950(2)	54(2)
C(13)	3742(5)	12618(5)	9493(2)	33(1)
O(13)	4961(3)	12582(3)	9740(1)	45(1)
C(14)	2699(6)	13963(6)	9418(2)	33(1)
C(15)	2076(5)	8772(4)	10721(2)	25(1)
O(15)	643(3)	8702(3)	10785(1)	32(1)
C(16)	3295(6)	8653(7)	11134(2)	36(1)

Among the factors not included in the modeling component of the present work are separate treatment of axial and equatorial configuration and the location relative to the ring oxygen, etc. Those factors, which surely skew the distribution in different ways, seem to be accommodated fairly well by our triacetyl model. However, further studies, to include the anomeric acetate substituent, are warranted.

4. Experimental

Synthesis and crystallization of **1** were reported by Bates and associates.[33] A single crystal suitable for data collection was mounted on a Bruker SMART single-crystal X-ray diffractometer with a 1K CCD detector. The crystal was cooled to 150(2) K in a stream of cold nitrogen gas, and X-ray diffraction data were collected using Mo K_α radiation. Unit cell dimensions were obtained by least-squares refinement of positions of 1408 observations in the range $4.7° < 2\theta < 40.0°$. The

Table 6
Hydrogen coordinates ($\times 10^4$) and isotropic displacement parameters ($Å^2 \times 10^3$) for penta-*O*-acetyl-β-D-galactopyranose (**1**)

	x	*y*	*z*	*U*(eq)
H(5)	332(3)	792(3)	9342(11)	7(9)
H(6B)	107(3)	979(3)	9813(11)	5(10)
H(6A)	85(4)	821(3)	9812(12)	33(13)
H(1)	309(4)	779(4)	8456(14)	35
H(4)	477(3)	992(3)	9594(10)	5(9)
H(3)	553(3)	886(3)	8787(10)	4(9)
H(2)	343(4)	1058(4)	8225(11)	24(12)
H(14C)	159(4)	1378(4)	9605(13)	63(14)
H(14B)	244(4)	1411(4)	9001(13)	55(13)
H(14A)	275(6)	1451(5)	9731(18)	120(20)
H(8C)	−30(6)	664(5)	7145(18)	120(20)
H(8B)	−149(7)	788(7)	7620(20)	210(40)
H(8A)	−8(5)	813(5)	7188(16)	70(20)
H(10C)	644(5)	900(5)	6978(16)	80(20)
H(10B)	513(5)	963(5)	6585(15)	87(18)
H(10A)	497(5)	819(5)	6921(15)	90(20)
H(12C)	958(4)	1192(5)	8954(15)	67(18)
H(12B)	816(4)	1257(4)	8608(12)	57(16)
H(12A)	861(5)	1253(4)	9369(14)	97(17)
H(16C)	417(5)	923(4)	11091(15)	50(16)
H(16B)	291(4)	876(4)	11497(13)	63(15)
H(16A)	385(5)	792(4)	11134(17)	70(20)

structure was solved by direct methods and optimized by full-matrix least-squares refinement of F^2 including all observations $> 2\sigma(I)$. All hydrogen atoms were located in a difference Fourier synthesis and included in the least-squares refinement with isotropic thermal parameters. The thermal parameter of H-1 refined to a negative value, so it was constrained in the refinement to a value of $1.2\times$ the equivalent isotropic temperature factor of atom C-1. Other details of the data collection and refinement are included in Table 1. The Flack absolute structure parameter,[34] 6.0(1.4) indicates that the reported D-configuration is incorrect. However, we feel that this is due to the lack of heavier atoms in the structure and the resulting low level of anomalous scattering rather than the chemically improbable L-structure. Final atomic coordinates are given in Tables 5 and 6.

Empirical force field calculations were carried out with MM3(96)[30] and a dielectric constant of 1.5. Levels of theory for QM calculations on carbohydrates have been discussed.[35] The QM calculations were done with Jaguar[36] using its pseudospectral method for increased computing speed. Except for the final B3LYP/6-311+G*//B3LYP/6-31G* calculations shown in Fig. 5(d and h), all reported values were based on full energy minimization with constraints on only the torsion angles used to define the energy surface. The increment of variation for the torsion angles was 10°. Symmetry was exploited for the QM calculations.

Structures were selected from the CSD for comparison with **1** and with the energy surfaces based on the following criteria: an aldohexapyranose ring with substituent oxygen atoms and a carbon at C-6, and the presence of the isopropyl acetate fragment. Also, crystallographic *R* values were 0.05 or less. Acetate groups at anomeric centers were not included, except in Fig. 5(a and e). Instead of using the experimental coordinates of the hydrogen atoms to define the H–C–O–C torsion angles, the hydrogen atoms were repositioned with Chem-X[†] for comparison with the energy surfaces. As is normal in X-ray diffraction studies, including the present work, hydrogen atoms are not always well located.

5. Supplementary material

Complete crystallographic data for the structural analysis have been deposited with the Cambridge Crystallographic Data Centre, CCDC No. 192406. Copies of this information may be obtained free of charge from the Director, Cambridge Crystallographic Data Centre, 12 Union Road, Cambridge, CB2 1EZ, UK. (fax: +44-1223-336033, e-mail: deposit@ccdc.cam.ac.uk or www: http://www.ccdc.cam.ac.uk).

Acknowledgements

The crystal was a gift from Professor John Robyt, Department of Biochemistry, Biophysics and Molecular Biology, Iowa State University, Ames, IA. Dr. John Vercellotti of V-Labs consulted on nomenclature. A draft of the manuscript was read by Professors Gabor Csonka of Budapest University of Technology and Economics and Paul Baures of Kansas State University, Dr. Jenn-Huei Lii of the University of Georgia and Dr. William E. Franklin, retired.

References

1. Allen, F. H.; Kennard, O. *Chem. Des. Automation News* **1993**, *8*, 31–37.
2. Bürgi, H.-B.; Dunitz, J. D. *Acta Crystallogr., Sect. B* **1988**, *44*, 445–448.
3. Bye, E.; Schweizer, B.; Dunitz, J. D. *J. Am. Chem. Soc.* **1982**, *104*, 5893–5898.
4. French, A. D.; Kelterer, A.-M.; Cramer, C. J.; Johnson, G. P.; Dowd, M. K. *Carbohydr. Res.* **2000**, *326*, 305–322.
5. French, A. D.; Kelterer, A.-M.; Johnson, G. P.; Dowd, M. K.; Cramer, C. J. *J. Mol. Graphics Modelling* **2000**, *18*, 95–107.

[†] Chem-X is no longer available.

6. Michael N. Burnett and Carroll K. Johnson, Oak Ridge National Laboratory.
7. Tvaroška, I.; Bleha, T. *Adv. Carbohydr. Chem. Biochem.* **1989**, *47*, 45–123.
8. Hager, C.; Miethchen, R.; Reinke, H. *J. Prakt. Chem.–Chem.-Zeitung* **2000**, *342*, 414–420.
9. Carpenter, G. B.; Kristen, H. *Acta Crystallogr., Sect. C: Cr. Str. Commun.* **1988**, *44*, 1051–1054.
10. Foces-Foces, C.; Cano, F. H.; Garcia-Blanco, S. *Acta Crystallogr., Sect. B* **1981**, *37*, 1270–1275.
11. Tomic, S.; Kojic-Prodic, B.; Magnus, V.; Lacan, G.; Duddeck, H.; Hiegemann, M. *Carbohydr. Res.* **1995**, *279*, 1–19.
12. Selkti, M.; Kassab, R.; Lopez, H. P.; Villain, F.; de Rango, C. *J. Carbohydr. Chem.* **1999**, *18*, 1019–1032.
13. Cox, P. J.; Armishaw, O. A.; Wardell, J. L. *J. Chem. Res. (S)* **1996**, 140–141. Alternatively; *J. Chem. Res. (M)* **1996**, 815–826.
14. Parkanyi, L.; Kalman, A.; Somsak, L.; Farkas, I. *Carbohydr. Res.* **1987**, *168*, 1–5.
15. Ollis, J.; James, V. J.; Stevens, J. D. *Cryst. Struct. Commun.* **1975**, *4*, 215–220.
16. James, V. J.; Stevens, J. D. *Cryst. Struct. Commun.* **1977**, *6*, 119–122.
17. Luger, P.; Paulsen, H. *Carbohydr. Res.* **1976**, *51*, 169–178.
18. Kopf, J.; Koll, P. *Acta Crystallogr., Sect. B* **1981**, *37*, 1127–1129.
19. Jones, P. G.; Sheldrick, G. M.; Kirby, A. J.; Glenn, R. *Z. Kristallog. Kristallg. Kristallp. Kristallch.* **1982**, *161*, 245–251.
20. Jones, P. G.; Sheldrick, G. M.; Kirby, A. J.; Glenn, R. *Z. Kristallog. Kristallg. Kristallp. Kristallch.* **1982**, *161*, 237–243.
21. Cremer, D.; Pople, J. A. *J. Am. Chem. Soc.* **1975**, *97*, 1354–1358.
22. Sheldrick, B. *Acta Crystallogr., Sect. B* **1976**, *32*, 1016.
23. Dowd, M. K.; French, A. D.; Reilly, P. J. *Carbohydr. Res.* **1994**, *264*, 1–19.
24. Tvaroška, I.; Taravel, F. R.; Utille, J. P.; Carver, J. P. *Carbohydr. Res.* **2002**, *337*, 353–367.
25. Gabbay, S. M.; Sundararajan, P. R.; Marchessault, R. H. *Biopolymers* **1972**, *11*, 79–94.
26. Leung, F.; Marchessault, R. H. *Can. J. Chem.* **1973**, *51*, 1215–1222.
27. Perez, S.; St.-Pierre, J.; Marchessault, R. H. *Can. J. Chem.* **1978**, *56*, 2866–2871.
28. Eliel, E. L.; Wilen, S. H. *Stereochemistry of Organic Compounds*; John Wiley and Sons: New York, 1994; pp 618–619.
29. Allinger, N. L.; Rahman, M.; Lii, J. H. *J. Am. Chem. Soc.* **1990**, *112*, 8293–8307.
30. MM3 is available to academic users from the Quantum Chemistry Program Exchange, Creative arts, Indiana University, Indiana University, Bloomington, IN. Commercial users may obtain MM3 from Tripos Associates, St. Louis, MO.
31. Allinger, N. L.; Zhu, Z.-q.; Chen, K. *J. Am. Chem. Soc.* **1992**, *114*, 6120–6133.
32. Wiberg, K. B.; Laidig, K. E. *J. Am. Chem. Soc.* **1987**, *109*, 5935–5943.
33. F. J. Bates and Associates, *Polarimetry, Saccharimetry and the Sugars*; US Government Printing Office: Washington, DC 1942; p. 488.
34. Flack, H. D. *Acta Crystallogr., Sect. A* **1983**, *39*, 881–987.
35. Csonka, G. I. *J. Mol. Struct. (THEOCHEM)* **2002**, *584*, 1–4.
36. Jaguar 4.0, Schrödinger, Inc., Portland, OR, USA.