Acta Cryst. (1978). B**34**, 1562–1569

The Crystal Structure of a β-(1 → 4) Linked Disaccharide, α-*N*,*N'*-Diacetylchitobiose Monohydrate

By Frode Mo

*Institutt for Røntgenteknikk, Universitetet i Trondheim-NTH, N-*7034 *Trondheim-NTH, Norway*

and Lyle H. Jensen

Department of Biological Structure, University of Washington, Seattle, Washington 98195, *USA*

(*Received* 12 *August* 1977; *accepted* 15 *November* 1977)

N,*N'*-Diacetylchitobiose, the repeating unit of chitin, contains two *N*-acetylglucosamine rings linked β-(1 → 4). The α anomer crystallizes with one molecule of water, $C_{16}H_{28}O_{11}N_2.H_2O$, in space group $P2_12_12_1$; $a = 11{\cdot}017$ (3), $b = 13{\cdot}066$ (4), $c = 13{\cdot}896$ (4) Å, $Z = 4$. A tangent-refinement procedure was used to solve the structure; refinement was by full-matrix least squares. The final R based on 2202 averaged F_o was 0·054; R_w was 0·041. Partial anomeric disorder with $\alpha:\beta \sim 90:10$ in the crystal was inferred from the analysis. There is a strong right-handed helical twist between the $^4C_1$ chair rings of the disaccharide, $\psi_H = +54°$, which prevents formation of the normal intramolecular hydrogen bond O(3')···O(5). The wide range of helical twist found so far in crystalline β-(1 → 4) disaccharide moieties, −12·5 to +80·5°, is evidence of a large flexibility in this glycosidic linkage. The fact that the intramolecular O(3')—H···O(5) bond does exist within a considerable range of ψ_H, emphasizes that it is only one of several factors determining the structure and conformation of the glycosidic bridge. The carbohydrate molecules are connected through a number of hydrogen bonds to form buckled ribbons along **c**. Other hydrogen bonds involving the water molecules link these ribbons together.

Introduction

N,*N'*-Diacetylchitobiose, or 2-acetamido-2-deoxy-4-*O*-(2-acetamido-2-deoxy-β-D-glucopyranosyl)-D-gluco-pyranose, or $(GlcNAc)_2$, contains two *N*-acetyl-D-glucosamine units linked β-(1 → 4). Its crystal structure is of interest for several reasons. It is the repeating unit of the natural fibrous polymer chitin and thus also related to other β-(1 → 4) linked polysaccharides, *e.g.* cellulose and mannan. Presumably. the geometry and properties of the glycosidic linkage are of particular significance for the polymer structures. Also of interest in this study were structure details of the *N*-acetyl group in substituted pyranosides and a description of the hydrogen-bonding system.

Like its monomer (GlcNAc or *N*-acetylglucosamine), $(GlcNAc)_2$ is an inhibitor of lysozyme (Rupley, 1964). Crystalline lysozyme–saccharide complexes have been studied by X-ray diffraction to a resolution of 2 Å. Fairly detailed data exist now both for tetragonal lysozyme–$(GlcNAc)_3$ (Blake, Johnson, Mair, North, Phillips & Sarma, 1967) and the triclinic lysozyme–GlcNAc and –$(GlcNAc)_2$ complexes (Kurachi, Sieker & Jensen, 1976). Thus, a further aim of the present investigation was to obtain information on possible conformational differences in the 'free' and complexed crystalline states, relating for instance to the flexibility of the glycosidic bridge. We report here the crystal structure of α-*N*,*N'*-diacetylchitobiose monohydrate. A preliminary account has been given previously (Mo & Jensen, 1975*a*).

Experimental

Single crystals of α-$(GlcNAc)_2$ monohydrate were grown by very slow evaporation of an aqueous 2-methyl-2,4-pentanediol solution. They are bisphenoidal; those with well developed faces exhibit point-group symmetry D_2. Varying amounts of crystals of a different habit were frequently obtained in the crystallization runs. They are rod-shaped, monoclinic and have been identified by structure analysis as a trihydrate of β-$(GlcNAc)_2$. Thus, crystals of both anomers of the disaccharide can develop from the same solution.

Crystals of the α form sufficiently large for X-ray work grow in about a week; however, all specimens examined by film were apparently affected to varying extents by lattice defects.

Crystal data

α-*N*,*N'*-Diacetylchitobiose monohydrate, $C_{16}H_{28}O_{11}N_2.H_2O$, FW 442·42, $a = 11{\cdot}017$ (3), $b = 13{\cdot}066$ (4), $c = 13{\cdot}896$ (4) Å at 19 (1)°C, $V = 2000{\cdot}3$ Å^3, λ(Mo $K\bar{\alpha}$) = 0·71069 Å, $D_x = 1{\cdot}469$ g cm^{-3} for

$Z = 4$, μ(Mo $K\bar{\alpha}$) = 1·36 cm^{-1}; space group $P2_12_12_1$; crystal size 0·60 × 0·35 × 0·25 mm.

Cell dimensions were determined from the setting angles of 16 reflexions. Repeated measurements during and after data collection showed the variation in each of these parameters to be within one standard deviation. The intensities of 5197 reflexions (excluding extinctions) were measured to a limit in sin θ/λ of 0·65 Å^{-1} with Nb-filtered Mo $K\alpha$ radiation on a computer-controlled diffractometer in the $\omega/2\theta$ scan mode. The scan rate in 2θ was 2° min^{-1}, basic scan width was 1·8°, and backgrounds were measured for 20 s at each end of the scan. Intensities below 2θ = 12° were remeasured semi-manually to minimize errors caused by the Nb K absorption edge. Three standard reflexions were monitored every 100 reflexions. The data were collected in two sets comprising the classes hkl and $\bar{h}\bar{k}l$. They were scaled according to the average decay curve of the three standards and corrected for coincidence loss but not for absorption, before conversion to F^2.

Weighted averages of F^2 and $\sigma(F^2)$ were calculated for pairs of equivalent reflexions; $\sigma(F_i^2) = \sigma(I_i).(\mathrm{Lp})^{-1}.(\mathrm{scale})$ where $\sigma^2(I_i) = \sigma^2_{i\,\mathrm{count}} + (SI_{i\,\mathrm{net}})^2$ and i = 1, 2. S was determined as 0·035 assuming that the differences, $\Delta_i = |F_i^2 - F^2_{\mathrm{ave}}|$ follow a normal distribution (Mo & Jensen, 1975*b*). Of 2600 reflexions, eight at $2\theta < 7\cdot5°$ were deleted because the Nb K edge was in the peak itself in this 2θ range. Another 390 reflexions with $F^2 \leq \sigma(F^2)$ were given zero weight.

Structure determination and refinement

The structure was solved by a modified version of the multisolution tangent-refinement program *TANNY* (Mo, 1973, 1977). The E map for the best phase model showed all 29 non-H atoms in the molecule, and in addition, one strong maximum later assigned to the O atom of a water molecule. All H atoms except one in the water were located in ΔF maps following anisotropic refinement of the heavier atoms.

The densest peak (0·85 e Å^{-3}) in the first ΔF map appeared at the position of the H atom attached equatorially to the anomeric C(1′) of the reducing ring. Least-squares refinement gave B values of this H atom in the range −1·5 to −2 Å^2 and C−H distances of 1·15 to 1·2 Å. Combined X-ray and neutron diffraction studies of various compounds have demonstrated that X-ray B values of H bonded to C are consistently low (*e.g.* Hanson, Sieker & Jensen, 1973) even when the bonded H scattering-factor curve of Stewart, Davidson & Simpson (1965) is used. Stewart (1976) has reassessed the average *error* in $B_{\text{X-ray}}$ for H bonded to C in sucrose at ∼−1·7 Å^2, in agreement with quantum-chemical calculations. In the present case, however, B of H(C1′) is definite nonpositive. This fact taken together with the long C−H bond suggests that the crystal contains a small amount of the β anomer in which the reducing O atom is equatorial. A ΔF map calculated subsequently without the contribution of H(C1′) had a diffuse but strong peak of maximum density ∼1 e Å^{-3} near the position of this atom. A β-O(1′) atom was placed 1·40 Å from C(1′) and population parameters for α and β-O(1′) were refined independently in alternating cycles together with the other variables. Although the H(C1′) and β-O(1′) positions remained separated by about 0·4 Å, the diffuseness of the density made a meaningful refinement of these atoms difficult. In the last least-squares cycles the coordinates of both H(C1′) and β-O(1′) were constrained; the final C−O distance was 1·27 Å. Refined values of the population parameters were 0·91 (1) for α-O(1′) and 0·11 (1) for β-O(1′), indicating that about 10% of the molecules are in the β form.

In a previous X-ray study of α-GlcNAc, Johnson (1966) found evidence indicating the presence of 20–25% of the β anomer in the crystal; however, no definite conclusion was reached as to possible α/β co-crystallization in a later study by Mo & Jensen (1975*b*) of this compound. We obtained stronger evidence for such co-crystallization in the present case. The poorer quality of the α-(GlcNAc)$_2$ crystals could, in fact, be caused largely by strain from the partial substitution of α-anomeric molecules in the lattice, *cf. The crystal structure*. Anomeric disorder was reported both for α-lactose monohydrate (Fries, Rao & Sundaralingam, 1971) and its hydrated complexes with $CaBr_2$ (Bugg, 1973) and $CaCl_2$ (Cook & Bugg, 1973) and also for α-melibiose monohydrate (Kanters, Roelofsen, Doesburg & Koops, 1976).

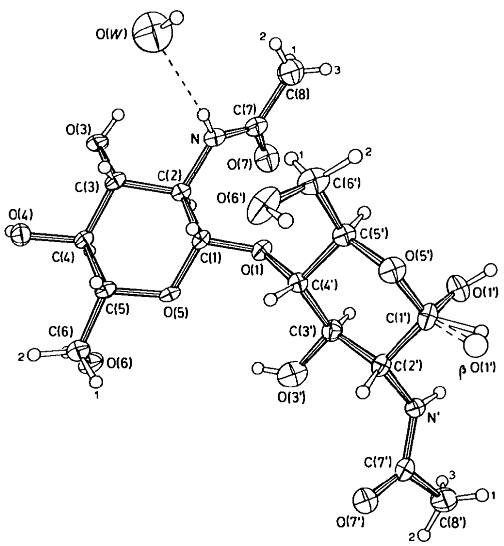

Fig. 1. Molecular conformation and atomic labelling. Sequential numbering of H atoms is indicated only where necessary. Thermal ellipsoids of the heavier atoms correspond to a 40% probability.

Table 1. *Final atomic parameters*

The positional parameters are × 10^4 for C, N, O and × 10^3 for H and β-O. The e.s.d.'s are in parentheses.

	x	y	z		x	y	z
C(1)	6182 (3)	3673 (2)	1982 (2)	C(1')	2780 (3)	2293 (3)	4319 (2)
C(2)	6195 (3)	4283 (3)	1055 (2)	C(2')	3759 (3)	2960 (3)	4777 (2)
C(3)	7420 (3)	4159 (2)	570 (2)	C(3')	4283 (3)	3693 (3)	4029 (2)
C(4)	8409 (3)	4490 (3)	1262 (2)	C(4')	4726 (3)	3091 (2)	3161 (2)
C(5)	8293 (3)	3906 (3)	2218 (2)	C(5')	3720 (3)	2418 (3)	2764 (2)
C(6)	9195 (3)	4258 (3)	2973 (2)	C(6')	4113 (4)	1716 (3)	1957 (3)
C(7)	4193 (3)	4525 (3)	357 (2)	C(7')	3843 (3)	3643 (3)	6439 (2)
C(8)	3196 (5)	4062 (6)	−243 (4)	C(8')	3150 (5)	4223 (5)	7181 (3)
O(1)	5069 (2)	3833 (2)	2436 (1)	O(1')	1818 (3)	2893 (2)	4075 (2)
O(3)	7505 (2)	4784 (2)	−273 (2)	O(3')	5250 (3)	4236 (2)	4473 (2)
O(4)	9589 (2)	4297 (2)	890 (2)	O(5')	3272 (2)	1765 (2)	3517 (2)
O(5)	7103 (2)	4029 (2)	2607 (1)	O(6')	5218 (3)	1189 (2)	2131 (2)
O(6)	9099 (3)	5325 (2)	3165 (2)	O(7')	4856 (2)	3295 (2)	6620 (2)
O(7)	4026 (2)	5339 (2)	781 (2)	N'	3264 (2)	3505 (2)	5597 (2)
N	5197 (2)	3982 (2)	424 (2)	O(W)	5895 (3)	2490 (3)	−1090 (3)

	x	y	z	B (Å²)		x	y	z	B (Å²)
H(C1)	635 (2)	290 (2)	188 (2)	0·6 (0·5)	H(C2')	439 (3)	244 (2)	504 (2)	2·7 (0·7)
H(C2)	602 (2)	502 (2)	120 (2)	1·1 (0·5)	H(C3')	371 (3)	419 (2)	385 (2)	2·7 (0·7)
H(C3)	756 (2)	344 (2)	44 (2)	0·3 (0·5)	H(C4')	541 (2)	268 (2)	330 (2)	0·8 (0·5)
H(C4)	838 (2)	527 (2)	135 (2)	1·0 (0·5)	H(C5')	307 (3)	290 (3)	250 (2)	3·6 (0·8)
H(C5)	840 (2)	314 (2)	215 (2)	2·0 (0·6)	H(C6'1)	426 (4)	215 (3)	147 (3)	4·1 (1·0)
H(C61)	903 (3)	383 (2)	355 (2)	3·4 (0·8)	H(C6'2)	337 (3)	121 (2)	188 (2)	2·6 (0·7)
H(C62)	1004 (3)	410 (2)	281 (2)	2·4 (0·7)	H(C8'1)	249 (4)	382 (3)	726 (3)	6·5 (1·3)
H(C81)	297 (5)	464 (4)	−60 (4)	8·0 (1·7)	H(C8'2)	373 (4)	447 (3)	763 (3)	6·5 (1·2)
H(C82)	342 (4)	357 (3)	−65 (3)	4·8 (1·1)	H(C8'3)	289 (4)	483 (3)	699 (3)	6·2 (1·5)
H(C83)	250 (6)	393 (5)	−8 (5)	11·8 (2·3)	H(O1')	120 (4)	258 (3)	397 (3)	4·4 (1·1)
H(O3)	689 (4)	465 (4)	−59 (4)	8·4 (1·5)	H(O3')	586 (4)	439 (3)	419 (3)	4·9 (1·0)
H(O4)	981 (4)	496 (4)	66 (4)	8·6 (1·5)	H(O6')	510 (5)	43 (5)	252 (5)	14·8 (2·5)
H(O6)	954 (4)	560 (4)	283 (3)	9·7 (1·5)	H(N')	259 (3)	384 (3)	553 (2)	2·8 (0·8)
H(N)	535 (3)	340 (2)	14 (2)	2·2 (0·7)	H(OW)	547 (4)	183 (4)	−99 (3)	6·5 (1·2)
H(C1')	207	187	478	0·5 (0·5)	β-O(1')	233	168	493	10·7 (2·0)

All attempts to locate the missing H atom in the water molecule were unsuccessful insofar as both positional and thermal parameters of this atom refined toward unacceptable values. Refinement of the other H atom, H(*W*1), was normal. A possible explanation of this situation is that the water molecule is rotationally disordered about the O(*W*)—H(*W*1) bond. The mode of disorder must allow a hydrogen bond to be formed between N attached to the nonreducing ring and O(*W*).

At the end of the refinement, based on 2202 *F* values, *R* was 0·054 and R_w* = 0·041, weights $w = 1/\sigma^2(F)$. Final average and maximum parameter shifts of C, O and N atoms were 0·02 and 0·10 of the e.s.d. respectively; corresponding figures for H: 0·04 and 0·20. The largest maxima and minima, of magnitude 0·2–0·25 e Å^{-3}, in the final ΔF map are near the O(3') position; two maxima ~0·2 e Å^{-3} are near H(C3'). The

* $R = \Sigma||F_o| - K|F_c||/\Sigma|F_o|$; $R_w = [\Sigma w(|F_o| - K|F_c|)^2/\Sigma wF_o^2]^{1/2}$.

Table 2. *Some endo- and exocyclic torsion angles*

The full atomic sequence is given only for exocyclic angles. Sign convention of torsion angles is that of Klyne & Prelog (1960).

Nonreducing ring		Reducing ring	
Endocyclic			
C(1)C(2)	60·0°	C(1')C(2')	56·4°
C(2)C(3)	−56·7	C(2')C(3')	−54·4
C(3)C(4)	54·3	C(3')C(4')	54·5
C(4)C(5)	−54·8	C(4')C(5')	−55·7
C(5)O(5)	59·1	C(5')O(5')	60·3
O(5)C(1)	−62·1	O(5')C(1')	−60·9
Exocyclic			
O(1)C(1)C(2)N	−58·4	O(1')C(1')C(2')N'	56·5
NC(2)C(3)O(3)	60·7	N'C(2')C(3')O(3')	62·0
O(3)C(3)C(4)O(4)	−65·0	O(3')C(3')C(4')O(1)	−68·1
O(4)C(4)C(5)C(6)	62·0	O(1)C(4')C(5')C(6')	68·2
C(5)O(5)C(1)O(1)	−179·9	C(5')O(5')C(1')O(1')	60·4
O(5)C(5)C(6)O(6)	−65·6	O(5')C(5')C(6')O(6')	−74·5
C(4)C(5)C(6)O(6)	56·7	C(4')C(5')C(6')O(6')	46·0

remaining four or five maxima in this range correspond to build-up of deformation density. There are several maxima and minima of smaller magnitude in the general area. Positional parameters of the atoms are given in Table 1.* Atomic form factors for C, O and N were those of Doyle & Turner (1968); for H the values of Stewart *et al.* (1965) were used.

Results and discussion

The molecular conformation

The overall molecular conformation of α-$(GlcNAc)_2$ is the $^4C_1$ chair (Fig. 1). Atoms in the nonreducing ring are shown unprimed, those in the reducing ring are primed. Endocyclic torsion angles (Table 2) vary from 54·3 to 62·1 and 54·4 to 60·9° in the unprimed and primed pyranose rings respectively. Both ranges are narrower than the corresponding mean ranges calculated by Hirotsu & Shimada (1974) for four β-(1 → 4) linked disaccharides, β-cellobiose (Chu & Jeffrey, 1968), methyl β-cellobioside (Ham & Williams, 1970), α-lactose (Fries *et al.*, 1971) and β-lactose (Hirotsu & Shimada, 1974): 52·5–64·2 and 49·0–64·0°. Both hydroxymethyl groups in α-$(GlcNAc)_2$ are in the *gauche–gauche* or *g–g* conformation, torsion angles of the primed unit, τ[O(5′)C(5′)C(6′)O(6′)] = −74·5 and τ[C(4′)C(5′)C(6′)O(6′)] = 46·0° are well outside the ranges given by Fries *et al.* (1971) for nine carbohydrates with hydroxymethyl groups in this conformation.

The orientation of the *N*-acetyl groups is defined by torsion angles about the bonds C(2)–N (ζ_N) and about the vectors C(2)–C(7) (χ) in Table 3. In both rings, the C(7)–C(8) bond is oriented *g–g* relative to C(1)–C(2) and C(2)–C(3). Thus, O(6) and O(7) are on the same side of the parent ring plane when the hydroxymethyl group has the *g–g* conformation. A similar orientation is adopted by the *N*-acetyl group in α-GlcNAc (Mo & Jensen, 1975*b*), χ_1 = −67·3, χ_2 = 77·8°, and is also seen in molecular plots of the complexes of triclinic lysozyme with GlcNAc and $(GlcNAc)_2$ (Kurachi *et al.*, 1976) and of tetragonal lysozyme with GlcNAc and $(GlcNAc)_3$ (Imoto, Johnson, North, Phillips & Rupley, 1972).

Parameters in Table 3 describing nonplanar distortions of the *N*-acetyl groups follow the convention of Winkler & Dunitz (1971). In this system, τ is a measure of the twist about the N–C(7) bond, χ_C and χ_N give out-of-plane bending at C(7) and N respectively. Deviations from planarity are small in the unprimed unit; the largest contribution to nonplanarity in the primed *N*-acetyl group comes from χ_N : 8·3°. A slightly pyramidal conformation at the N atom of an amide group may be introduced at very modest energy cost (Winkler & Dunitz, 1971) or may even correspond to an energy minimum (Ramachandran, Lakshminarayanan & Kolaskar, 1973).

Perhaps the most noteworthy structural feature of α-$(GlcNAc)_2$ is the conformation at the glycosidic bridge. Table 4 gives torsion (Sundaralingam, 1968) and pseudotorsion angles (Rohrer, 1972), the latter defined here as the rotation about the vector C(1)–C(4′), in six determined structures. Although not exhaustive, this table adequately shows the observed range of twist between β-(1 → 4) linked pyranose rings. The relative twist can be given alternatively as the average of ψ_1[O(5)C(1)C(4′)C(3′)] and ψ_2[C(2)C(1)C(4′)C(5′)]. This helicity parameter, termed ψ_H, is +54° for α-$(GlcNAc)_2$. Only the xylobiose unit of an aldotriuronic acid (Moran & Richards, 1973), hereinafter ALDXX,

* Lists of structure factors and anisotropic thermal parameters have been deposited with the British Library Lending Division as Supplementary Publication No. SUP 33224 (16 pp.). Copies may be obtained through The Executive Secretary, International Union of Crystallography, 13 White Friars, Chester CH1 1NZ, England.

Table 3. *Conformational parameters of the N-acetyl groups*

(*a*) Orientational parameters follow the sign convention of Klyne & Prelog (1960).

	Nonreducing ring		Reducing ring	
χ_1	C(1)C(2)C(7)C(8)	−81·1°	C(1′)C(2′)C(7′)C(8′)	−51·8°
χ_2	C(3)C(2)C(7)C(8)	65·7	C(3′)C(2′)C(7′)C(8′)	92·8
ζ_N	C(1)C(2)N C(7)	100·5	C(1′)C(2′)N′ C(7′)	138·7
ζ'_N	C(3)C(2)N C(7)	−137·0	C(3′)C(2′)N′ C(7′)	−98·9

(*b*) Parameters describing nonplanarity are defined according to Winkler & Dunitz (1971).

	Nonreducing ring		Reducing ring	
ω_1	C(8)C(7)N C(2)	−173·9°	C(8′)C(7′)N′ C(2′)	−179·7°
ω_2	O(7)C(7)N H(N)	179·2	O(7′)C(7′)N′ H(N′)	−173·8
ω_3	O(7)C(7)N C(2)	3·0	O(7′)C(7′)N′ C(2′)	−2·1
ω_4	C(8)C(7)N H(N)	2·3	C(8′)C(7′)N′ H(N′)	8·6
$\tau = \frac{1}{2}(\omega_1 + \omega_2)$	mod 2π	−177·3		−176·8
$\chi_C = \omega_1 - \omega_3 + \pi$	mod 2π	3·1		2·4
$\chi_N = -\omega_1 + \omega_4 + \pi$	mod 2π	−3·8		8·3

Table 4. *Conformational parameters of the β-(1 → 4) bridge in some disaccharide units*

Torsion angles φ_1, φ_1', φ_2 and φ_2' are according to Sundaralingam (1968). Pseudotorsion angles ψ_1, ψ_1', ψ_2 and ψ_2' define rotations about the vector C(1) → C(4'). All parameters in the table have been calculated from published coordinates.

		(I)	(II)	(III)	(IV)	(V)	(VI)
φ_1	O(5)C(1)O(1)C(4')	−91·1	−94·2	−64	−76·3	−79·5	−81·9
φ_1'	C(2)C(1)O(1)C(4')	152·0	146·1	166	166·5	161·5	159·8
φ_2	C(1)O(1)C(4')C(3')	80·3	96·0	87	106·4	133·5	161·5
φ_2'	C(1)O(1)C(4')C(5')	−160·7	−142·8	−151	−132·3	−106·8	−79·6
ψ_1	O(5)C(1)C(4')C(3')	−10·7	−0·2	26	24·5	49·1	81·0
ψ_1'	O(5)C(1)C(4')C(5')	125·9	144·6	162	171·1	−165·9	−142·9
ψ_2	C(2)C(1)C(4')C(5')	−13·9	2·6	18	40·2	58·8	80·0
ψ_2'	C(2)C(1)C(4')C(3')	−150·4	−142·1	−118	−106·5	−86·3	−56·1
$\psi_H = \frac{1}{2}(\psi_1 + \psi_2)$		−12·5	1	22	32·5	54	80·5

References: (I) Methyl β-cellobioside (Ham & Williams, 1970). (II) α-Lactose (Fries, Rao & Sundaralingam, 1971). (III) $(GlcNAc)_2$ in triclinic lysozyme complex (Kurachi, Sieker & Jensen, 1976). (IV) β-Cellobiose (Chu & Jeffrey, 1968). (V) α-$(GlcNAc)_2$ (this paper). (VI) Xylobiose unit in an aldotriuronic acid (Moran & Richards, 1973).

has a larger right-handed helical twist of +80·5°.* Methyl β-cellobioside with a left-handed twist of −12·5° is at the other extreme of the angle range. For $(GlcNAc)_2$ in the triclinic lysozyme complex we calculated an intermediate twist, $\psi_H \sim +22°$, which is about 10° less than that of cellobiose. $(GlcNAc)_3$ in the tetragonal lysozyme complex appears more distorted. No angle values have been calculated, but comparison of models with an *ORTEP* drawing (Johnson, 1965) of the *B* and *C* site rings suggests that the intra-ring twist is of similar magnitude or possibly somewhat larger than in cellobiose.

The wide range of ψ_H values in Table 4 is evidence of a remarkably high degree of flexibility in the β-(1 → 4) glycosidic bridge. Omitting ALDXX, which contains no hydroxymethyl group near the bridge, this range is about 67°. With increasing right-handed twist, O(3') is moved away from O(5) until no intramolecular H bond can be formed for large ψ_H values. In β-lactose ($\psi_H = +39°$) the O(3')···O(5) distance is in the normal range for an O—H···O bond. Corresponding distances in α-$(GlcNAc)_2$ ($\psi_H = +54°$) and the xylobiose unit of ALDXX ($\psi_H = +80·5°$) are 3·311 and 3·964 Å, respectively, which are too great to allow formation of a normal intramolecular hydrogen bond. It is interesting to note, however, that the O(3')—H···O(5) bond persists in all studied β-(1 → 4) linked disaccharides in the ψ_H range −12·5 to +39°. Clearly, this hydrogen bond constrains only mildly the conformational freedom about the glycosidic bridge.

Bond lengths and angles

Bond lengths and angles with their e.s.d.'s are listed in Tables 5 and 6. Mean values and ranges of endocyclic C—C bonds are 1·522 and 0·022 Å for the unprimed ring, 1·523 and 0·008 Å for the primed ring. Unprimed and primed rings have means of 1·425 and 1·421 Å, with ranges 0·009 and 0·001 Å, respectively, in exocyclic C—O bonds excluding those involving O(1) and O(1'). With one exception, relative bond lengths of the two five-atom sequences C—O_{ring}—C_{ano}—O—*R* are in good agreement with theoretical models of the anomeric effect in the β-pyranosidic and α-pyranose moieties (Jeffrey, Pople & Radom, 1974). The extreme shortening of the α-anomeric C(1')—O(1') bond to 1·361 Å may, in part, reflect difficulties in handling disorder in this part of the molecule by the least-squares method. Shortening of this bond was also observed in α-melibiose monohydrate (1·359 Å) for which the α:β disorder ratio was 80:20 (Kanters *et al.*, 1976) and in α-lactose-$CaBr_2.7H_2O$ (1·365 Å)

* A left-handed helical conformation was assigned to this disaccharide unit in the original study.

Table 5. *Bond lengths* (Å) *with standard deviations*

Nonreducing ring		Reducing ring	
C(1)–C(2)	1·515 (4)	C(1')–C(2')	1·526 (5)
C(2)–C(3)	1·518 (4)	C(2')–C(3')	1·527 (5)
C(3)–C(4)	1·517 (4)	C(3')–C(4')	1·519 (4)
C(4)–C(5)	1·537 (4)	C(4')–C(5')	1·519 (5)
C(5)–O(5)	1·427 (4)	C(5')–O(5')	1·438 (4)
C(1)–O(5)	1·414 (4)	C(1')–O(5')	1·419 (4)
C(1)–O(1)	1·395 (4)	C(1')–O(1')	1·361 (5)
		C(1')–β-O(1')	1·27
C(3)–O(3)	1·431 (4)	C(3')–O(3')	1·421 (4)
C(4)–O(4)	1·422 (4)	C(4')–O(1)	1·449 (4)
C(5)–C(6)	1·517 (5)	C(5')–C(6')	1·511 (5)
C(6)–O(6)	1·423 (5)	C(6')–O(6')	1·420 (5)
C(2)–N	1·460 (4)	C(2')–N'	1·449 (4)
N–C(7)	1·317 (5)	N'–C(7')	1·346 (4)
C(7)–O(7)	1·231 (5)	C(7')–O(7')	1·230 (4)
C(7)–C(8)	1·506 (7)	C(7')–C(8')	1·490 (6)

with $\alpha:\beta \sim 88:12$ (Bugg, 1973). However, the C(1′)–O(1′) bond was normal (1·391 Å) in the parent α-lactose–$CaCl_2.7H_2O$ complex with an $\alpha:\beta$ ratio of ~95:5 (Cook & Bugg, 1973).

Table 6. *Valency angles (°) with standard deviations*

Nonreducing ring		Reducing ring	
O(5)C(1)C(2)	110·0 (2)	O(5′)C(1′)C(2′)	109·6 (3)
C(1)C(2)C(3)	109·3 (2)	C(1′)C(2′)C(3′)	110·0 (3)
C(2)C(3)C(4)	109·0 (2)	C(2′)C(3′)C(4′)	109·8 (3)
C(3)C(4)C(5)	110·3 (3)	C(3′)C(4′)C(5′)	110·7 (3)
C(4)C(5)O(5)	110·4 (2)	C(4′)C(5′)O(5′)	109·2 (2)
C(5)O(5)C(1)	112·9 (2)	C(5′)O(5′)C(1′)	114·5 (3)
O(5)C(1)O(1)	107·7 (2)	O(5′)C(1′)O(1′)	112·4 (3)
C(2)C(1)O(1)	108·3 (2)	C(2′)C(1′)O(1′)	109·0 (3)
C(1)C(2)N	111·2 (3)	C(1′)C(2′)N′	110·0 (3)
C(3)C(2)N	111·9 (2)	C(3′)C(2′)N′	111·7 (3)
C(2)C(3)O(3)	111·2 (3)	C(2′)C(3′)O(3′)	107·5 (3)
C(4)C(3)O(3)	108·0 (3)	C(4′)C(3′)O(3′)	111·2 (3)
C(3)C(4)O(4)	112·0 (2)	C(3′)C(4′)O(1)	106·8 (3)
C(5)C(4)O(4)	107·6 (2)	C(5′)C(4′)O(1)	109·0 (2)
C(4)C(5)C(6)	113·1 (3)	C(4′)C(5′)C(6′)	114·4 (3)
O(5)C(5)C(6)	107·8 (2)	O(5′)C(5′)C(6′)	106·2 (3)
C(5)C(6)O(6)	112·2 (3)	C(5′)C(6′)O(6′)	114·4 (3)
C(2)N C(7)	122·0 (3)	C(2′)N′ C(7′)	124·8 (3)
N C(7)O(7)	123·8 (3)	N′ C(7′)O(7′)	123·9 (3)
N C(7)C(8)	115·9 (4)	N′ C(7′)C(8′)	115·3 (3)
O(7)C(7)C(8)	120·2 (4)	O(7′)C(7′)C(8′)	120·8 (3)
Bridge		O(5′)C(1′)β-O(1′)	111
C(1)O(1)C(4′)	116·3 (2)	O(1′)C(1′)β-O(1′)	103
		C(2′)C(1′)β-O(1′)	111

Angles involving H

Type	Number	Range	Mean	σ_{ave}
X*CH	35	103·1–113·5	108·9	1·7
CCH (methyl)	6	99·1–128·1	110·9	3·2
X*C(1′)H(C1′)	3	83·6–121·7	107·5	–
COH	6	100·9–123·1	110·9	3·3
CNH	4	111·1–126·8	118·2	2·1
HCH	8	94·2–131·3	108	4

*X = C,O,N.

Comparison of the pyranosidic C–C and C–O bond lengths in seven β-(1 → 4) linked disaccharides or disaccharide residues* shows that the observed spread in each bond is of the order 0·015 to 0·035 Å. The C(1)–O(1) and O(1)–C(4′) bonds in the bridge have relatively narrow ranges, 0·015 and 0·016 Å, respectively. These bonds are 1·395 and 1·449 Å in α-$(GlcNAc)_2$.

The mean value of endocyclic C–C–X angles (X = C, O, N) of α-$(GlcNAc)_2$ is 109·8° with a spread of 1·5° in both rings. The ring C–O–C angles are 112·9 and 114·5° for the unprimed (β) and primed (α) units respectively. A similar dependency on anomeric configuration was found in α-lactose and methyl β-maltopyranoside (Chu & Jeffrey, 1967; *cf.* also Arnott & Scott, 1972).

The average spread in endo- and exocyclic C–C–X angles not involving O(1) or O(1′) is 3·1 and 4·1°, respectively, for the seven β-(1 → 4) linked disaccharide units. Individual ranges are from 1·4 to 6·3°. The angles O(5)–C(1)–O(1) and C(2)–C(1)–O(1) at the glycosidic bridge have rather small spreads, 1·3 and 2·7°, respectively, about their average values 107·5 and 108·8°. By contrast, both angles C(3′)–C(4′)–O(1) and C(5′)–C(4′)–O(1) vary over wide ranges of about 8°. The different properties of exocyclic angles at C(1) and C(4′) are accompanied by similar differences in twist of the two rings relative to the glycosidic bridge. Excluding $(GlcNAc)_2$ from the protein study, the range of φ_1 and φ_1' (relative twist of unprimed ring) in Table 4 is only approximately 20°, compared with a range of 80° for φ_2 and φ_2' (relative twist of primed ring). Differences in strain associated with these conformational dissimilarities thus seem to be absorbed largely in the exocyclic angles at C(4′). The bridge angle itself is remarkably constant. In six of the seven structures it ranges from 115·8 (methyl β-cellobioside) to 117·1° (α-lactose). The mean is 116·4° [*cf.* 116·3° in α$(GlcNAc)_2$]. The xylobiose

* These include the structures in Table 4 except (III), but in addition β-lactose (Hirotsu & Shimada, 1974) and β-$(GlcNAc)_2$ (Mo, to be published).

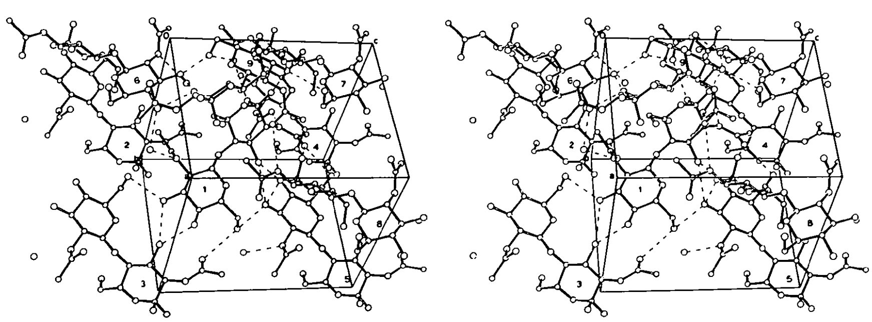

Fig. 2. Stereoscopic packing diagram with hydrogen bonds shown as broken lines. Molecules are numbered according to the symmetry code in Table 7.

Table 7. *The geometry of the hydrogen-bonding system*

Symmetry code for subscripts

(1)	x,	y,	z	(4)	$\frac{1}{2}-x$,	$1-y$,	$\frac{1}{2}+z$	(7)	$-\frac{1}{2}+x$,	$\frac{1}{2}-y$,	$1-z$
(2)	$\frac{1}{2}-x$,	$1-y$,	$-\frac{1}{2}+z$	(5)	$\frac{3}{2}-x$,	$1-y$,	$\frac{1}{2}+z$	(8)	$\frac{1}{2}+x$,	$\frac{1}{2}-y$,	$1-z$
(3)	$\frac{3}{2}-x$,	$1-y$,	$-\frac{1}{2}+z$	(6)	$-\frac{1}{2}+x$,	$\frac{1}{2}-y$,	$-z$	(9)	$1-x$,	$-\frac{1}{2}+y$,	$\frac{1}{2}-z$
								(10)	$-1+x$,	y,	z

D—H···A	D···A	H···A	$(\mathrm{H}\cdots A)^*_{corr}$	$(\angle D\text{—H}\cdots A)^*_{corr}$
O(3)—H(O3)···$O(6)_3$	2·803 (4) Å	2·04 (5) Å	1·91 Å	150°
O(4)—H(O4)···$O(3')_3$	2·754 (4)	1·96 (5)	1·94	139
O(6)—H(O6)···$O(7')_3$	3·031 (4)	2·32 (5)	2·12	155
N—H(N)···$O(W)_1$	2·970 (5)	2·17 (3)	2·05	150
O(1′)—H(O1′)···$O(7')_7$	2·831 (4)	2·04 (4)	1·88	163
O(3′)—H(O3′)···$O(3)_5$	2·808 (4)	2·22 (4)	2·12	126
N′—H(N′)···$O(7)_4$	2·951 (4)	2·10 (3)	1·96	164
O(W)—H(OW)···$O(4)_6$	2·757 (5)	1·77 (5)	1·78	174
β-O(1′)—H(β-O1′)···$O(3)_9$	2·52			
Intermolecular O···O contact:				
O(6′)···$O(1)_9$	3·152 (3)	2·09	2·23	156

* Distances and angles involving H have been recalculated (corr.) assuming lengths of 0·98 Å for the O—H (Brown & Levy, 1973) and 1·015 Å for the N(sp^2)—H bonds (Lehmann, Verbist, Hamilton & Koetzle, 1973).

bridge angle in ALDXX, 113·8°, falls outside the range and corresponds to a normal C—O—C angle in methyl β-pyranosides (Moran & Richards, 1973).

The bridge angle in β-(1 → 4) linked disaccharides is apparently quite insensitive to changes in intermolecular bonding, and also to substantial changes in helical twist. The structure analyses of α-(GlcNAc)$_2$ and ALDXX show that contraction of this angle may take place at extreme right-handed twists but does not depend primarily on the existence of an intramolecular hydrogen bond O(3′)···O(5). It is suggested instead that a major steric factor is repulsion between the H atoms bonded to C(1) and C(4′). The average uncorrected X-ray H(C1)···H(C4′) distance of disaccharides in the ψ_H range −12·5 to +54° is 2·22 Å, the value of α-(GlcNAc)$_2$ is 2·24 Å. For $\psi_H = +80{\cdot}5°$, as in ALDXX, this distance has increased to 2·50 Å.

Bond lengths and angles in the N-acetyl groups correspond reasonably well with average values for the peptide unit compiled by Ramachandran, Kolaskar, Ramakrishnan & Sasisekharan (1974). Valency angles involving H in the unprimed group are more distorted relative to averages (~7°). As noted before, the amide C′—N bond length varies considerably in different structures (Mo & Jensen, 1975*b*). In the present case, C(7)—N and C(7′)—N′ differ by 6–7 standard deviations with the shorter bond in the most planar, unprimed group. A correlation between this bond length and nonplanar distortions of the amide group has been proposed (Ealick & van der Helm, 1975, 1977). Both N atoms of α-(GlcNAc)$_2$ are donors in hydrogen bonds.

Except for C(1′)—H(C1′) at 1·15 Å, the pyranosidic C—H bonds range from 0·90 to 1·06; the mean is 0·99 Å. The range of methyl C—H bonds is 0·82–0·95, mean 0·90 Å, and O—H bonds vary from 0·76 to 0·99 Å, except for O(6′)—H(O6′) at 1·14 Å. The mean value is 0·86 Å. A summary of valency angles involving H is given in Table 6.

The crystal structure

Fig. 2 shows the molecular packing and part of the hydrogen-bond network. Both rings are involved in about the same number of nonpolar interactions. Contacts are distributed fairly well over the exposed parts of the unprimed unit, while more than half of the contacts to the primed ring involve C(7′), O(6′), O(7′) and, in particular, methyl C(8′)H$_3$. The shortest distances are: C(8′)···O(7)$_4$* 3·140 (6), C(7′)···H(O4)$_5$ 2·59 (5), C(8′)···H(O6)$_5$ 2·71 (5), O(5′)···H(C2)$_9$ 2·44 (3) and H(C8′2)···H(O6)$_5$ 1·94 (6) Å.

The geometry of the hydrogen-bonding system is summarized in Table 7. Distances and angles involving H have been corrected (Mo & Sivertsen, 1971), assuming that the systematic errors in the X-ray coordinates of these atoms are along the parent covalent bonds. All O and N atoms take part in hydrogen bonding except O(5) and O(5′) in the rings, O(1) in the glycosidic bridge and O(6′). One infers from their geometry that some of the bonds are relatively weak, in particular O(3′)···O(3)$_5$ and O(6)···O(7′)$_3$. The short intermolecular distance between β-O(1′) and O(3) indicates a hydrogen bond, probably with O(3) as acceptor, thus compensating for the absent O(1′)···

* The subscript 4 denotes molecule at $\frac{1}{2}-x$, $1-y$, $\frac{1}{2}+z$. The symmetry code is explained in Table 7.

O(7′) bond in the β anomer. A similar situation was observed in the α/β disordered structures of α-lactose monohydrate (Fries *et al.*, 1971) and α-melibiose monohydrate (Kanters *et al.*, 1976). Normalizing the length of C(1′)–β-O(1′) would make the short β-O(1′)$\cdots$O(3) contact even shorter, and it seems therefore that the accommodation of β anomers in the lattice may involve some strain. In general, the relative effect on the lattice energy from co-crystallized second anomer should be smaller with increasing number of sugar residues in the molecule. *A priori*, therefore, one would expect anomeric disorder to become more common in crystals of higher oligomers.

Molecules are laced together in infinite ribbons in the **c** direction by a number of hydrogen bonds (molecules 3, 1 and 5 constitute part of one such ribbon):

$$
\begin{array}{c}
\hspace{22em} O(7') \\
\hspace{22em} \uparrow \\
O(4)_5 \rightarrow O(3')_1 \rightarrow O(3)_5 \rightarrow O(6)_1 \cdots O(4)_1 \rightarrow O(3')_3 \rightarrow O(3)_1 \rightarrow O(6)_3 \\
\downarrow \\
O(7')_3
\end{array}
$$

Symmetry-related ribbons, *e.g.* the 3, 1, 5 and 6, 9, 7 ribbons, are connected through a right-handed helical system of hydrogen bonds, the sense donor–acceptor advancing along −**a**: $N_1 \rightarrow O(W)_1 \rightarrow O(4)_6 \cdots N_6 \rightarrow O(W)_6 \rightarrow O(4)_{10}$. Hydrogen bonds $O(1')_8 \rightarrow O(7')_1$ and $O(1')_1 \rightarrow O(7')_7$ form a second link between these ribbons. Hydrogen bonds involving β-O(1′) would serve a similar function. Parallel ribbons are also linked directly *via* hydrogen bonds of the type N′ → O(7).

The primary hydroxyl O(6′) points towards a void in the crystal structure. Its interaction with O(1) of a neighbour molecule is probably very weak (Table 7).

The XRAY 72 system (Stewart, Kruger, Ammon, Dickinson & Hall, 1972) was used for structure refinement and analyses of molecular geometry. Molecular plots were made by *ORTEP* (Johnson, 1965). Thanks are due to Dr J. A. Rupley, University of Arizona, Tucson, Arizona, for supplying samples of the disaccharide, and to Dr L. C. Sieker for helpful comments on crystallization. Grants from the National Institutes of Health (USPHS Grant GM-10828) and Norges Tekniske Høgskoles Fond are gratefully acknowledged.

References

Arnott, S. & Scott, W. E. (1972). *J. Chem. Soc. Perkin Trans.* 2, pp. 324–335.

Blake, C. C. F., Johnson, L. N., Mair, G. A., North, A. C. T., Phillips, D. C. & Sarma, V. R. (1967). *Proc. R. Soc. London Ser. B*, **167**, 378–388.

Brown, G. M. & Levy, H. A. (1973). *Acta Cryst.* B**29**, 790–797.

Bugg, C E. (1973). *J. Am. Chem. Soc.* **95**, 908–913.

Chu, S. S. C. & Jeffrey, G. A. (1967). *Acta Cryst.* **23**, 1038–1049.

Chu, S. S. C. & Jeffrey, G. A. (1968). *Acta Cryst.* B**24**, 830–838.

Cook, W. J. & Bugg, C. E. (1973). *Acta Cryst.* B**29**, 907–909.

Doyle, P. A. & Turner, P. S. (1968). *Acta Cryst.* A**24**, 390–397.

Ealick, S. E. & van der Helm, D. (1975). *Acta Cryst.* B**31**, 2676–2680.

Ealick, S. E. & van der Helm, D. (1977). *Acta Cryst.* B**33**, 76–80.

Fries, D. C., Rao, S. T. & Sundaralingam, M. (1971). *Acta Cryst.* B**27**, 994–1005.

Ham, J. T. & Williams, D. G. (1970). *Acta Cryst.* B**26**, 1373–1383.

Hanson, J. C., Sieker, L. C. & Jensen, L. H. (1973). *Acta Cryst.* B**29**, 797–808.

Hirotsu, K. & Shimada, A. (1974). *Bull. Chem. Soc. Jpn*, **47**, 1872–1879.

Imoto, T., Johnson, L. N., North, A. C. T., Phillips, D. C. & Rupley, J. A. (1972). *The Enzymes*, Vol. VII, 3rd ed., edited by P. D. Boyer, pp. 665–868. New York: Academic Press.

Jeffrey, G. A., Pople, J. A. & Radom, L. (1974). *Carbohydr. Res.* **38**, 81–95.

Johnson, C. K. (1965). *ORTEP*. Report ORNL-3794. Oak Ridge National Laboratory, Tennessee.

Johnson, L. N. (1966). *Acta Cryst.* **21**, 885–891.

Kanters, J. A., Roelofsen, G., Doesburg, H. M. & Koops, T. (1976). *Acta Cryst.* B**32**, 2830–2837.

Klyne, W. & Prelog, V. (1960). *Experientia*, **16**, 521–523.

Kurachi, K., Sieker, L. C. & Jensen, L. H. (1976). *J. Mol. Biol.* **101**, 11–24.

Lehmann, M. S., Verbist, J. J., Hamilton, W. C. & Koetzle, T. F. (1973). *J. Chem. Soc. Perkin Trans.* 2, pp. 133–137.

Mo, F. (1973). *Acta Cryst.* B**29**, 1796–1807.

Mo, F. (1977). *Acta Cryst.* B**33**, 641–649.

Mo, F. & Jensen, L. H. (1975*a*). *Acta Cryst.* A**31**, S110.

Mo, F. & Jensen, L. H. (1975*b*). *Acta Cryst.* B**31**, 2867–2873.

Mo, F. & Sivertsen, B. K. (1971). *Acta Cryst.* B**27**, 115–128.

Moran, R. A. & Richards, G. F. (1973). *Acta Cryst.* B**29**, 2770–2783.

Ramachandran, G. N., Kolaskar, A. S., Ramakrishnan, C. & Sasisekharan, V. (1974). *Biochim. Biophys. Acta*, **359**, 298–302.

Ramachandran, G. N., Lakshminarayanan, A. V. & Kolaskar, A. S. (1973). *Biochim. Biophys. Acta*, **303**, 8–13.

Rohrer, D. C. (1972). *Acta Cryst.* B**28**, 425–433.

Rupley, J. A. (1964). *Biochim. Biophys. Acta*, **83**, 245–255.

Stewart, J. M., Kruger, G. J., Ammon, H. L., Dickinson, C. & Hall, S. R. (1972). XRAY 72 system. Tech. Rep. TR-192. Computer Science Center, Univ. of Maryland, College Park, Maryland.

Stewart, R. F. (1976). *Acta Cryst.* A**32**, 182–185.

Stewart, R. F., Davidson, E. R. & Simpson, W. T. (1965). *J. Chem. Phys.* **42**, 3175–3187.

Sundaralingam, M. (1968). *Biopolymers*, **6**, 189–213.

Winkler, F. K. & Dunitz, J. D. (1971). *J. Mol. Biol.* **59**, 169–182.