Lead Article

Acta Cryst. (1990). **B46**, 89–103

Crystallographic Studies of Carbohydrates*

BY G. A. JEFFREY

Department of Crystallography, University of Pittsburgh, Pittsburgh, PA 15260, *USA*

(*Received* 16 *August* 1989; *accepted* 6 *November* 1989)

Abstract

The monosaccharides which constitute the monomer units of many important industrial and biological macromolecules are well represented among the 2000 crystal structures of the carbohydrate class 45 of the Cambridge Structural Database. There are few examples of crystal structure analyses of the corresponding acids, but many of their calcium salts and calcium salt complexes. With the exception of the disaccharides and cyclodextrins, the oligosaccharides are not well represented, with less than ten trisaccharides, one tetrasaccharide and one hexasaccharide–iodide complex. Two important conformational factors are the *Hassel–Ottar effect* and the *anomeric effect*, both of which have been studied using crystallographic data. Hydrogen bonding is ubiquitous in carbohydrate crystals and generally involves all the hydroxyls as both donors and acceptors, and some of the ring and glycosidic oxygens as acceptors. These hydrogen bonds tend to form finite or infinite chains. In hydrates, these chains are linked through the water molecules to form networks. Cyclic hydrogen-bond systems are observed in the cyclodextrins. Long-chain alkylated carbohydrates provide a large class of thermotropic and lyotropic liquid crystals, and some non-ionic surfactants which have been shown to be useful for membrane-protein solubilization and crystallization.

Introduction

Crystals were very important in the formative years of carbohydrate chemistry (Fischer, 1891; Haworth, 1929; Hudson, 1941) when crystallization was a primary means of purification. Many carbohydrates were syrups or amorphous powders and, prior to chromatography and NMR spectroscopy, a crystalline product with a sharp melting point and a good microanalysis was a prerequisite to publication. The crystal morphology of a number of carbohydrates is described, with exquisite precision, in Groth's *Chemische Kristallographie*, Volume III (Groth, 1910). These include the trisaccharide raffinose hydrate and the tetrasaccharide stachyose hydrate.

Carbohydrates were among the first organic compounds to be studied by X-ray crystal structure analysis: pentaerythritol by Llewellyn, Cox & Goodwin (1937) and α-D-glucosamine hydrochloride and hydrobromide by Cox & Jeffrey (1939). This was followed by sucrose.NaBr.$2H_2O$ (Beevers & Cochran, 1947). Thereafter, only eight carbohydrate crystal structures were reported over the next 15 years. The first neutron diffraction analyses were of sucrose and α-D-glucose by Brown & Levy (1963, 1965).

With the advent of automatic diffractometers and the increasing success of direct methods for crystal structures with noncentrosymmetric space groups, the number of reported crystal structure analyses increased steadily to reach the present number of about 2000. This class of organic compounds shares with amino acids the distinction of having the greatest number of neutron diffraction analyses. The Cambridge Structural Database (class 45) contains a vast amount of structural data relating to the relationship between configuration, conformation and molecular packing in carbohydrate crystals, relatively little of which has been systematically utilized (*cf.* Allen, Kennard & Taylor, 1983).

In this article, the results of some of these analyses are discussed in the context of the principal types of molecule. The author apologizes, in advance, for the omission of some reader's '*pet structures*' at the expense of including his own.

A basic knowledge of the stereochemistry of carbohydrates is assumed (*cf.* Stoddard, 1971). Standard references for nomenclature are *The Carbohydrates: Chemistry, Biochemistry and Physiology* (Pigman, 1957), *The Carbohydrates: Chemistry and Biochemistry* (Pigman & Horton, 1970) and *Rodd's Chemistry of Carbon Compounds* (Vol. I, Part F, 1967, and Supplement Part FG, 1983).

* *Editorial note*: This invited paper is one of a series of comprehensive Lead Articles which the Editors invite from time to time on subjects considered to be timely for such treatment.

0108-7681/90/020089-15$03.00

© 1990 International Union of Crystallography

The aldoses and ketoses

These are the most important carbohydrates from the structural point of view because they are the monomers of so many important industrial and biological macromolecules. Apart from the celluloses, starches, chitins, carrageenans, blood sugars, cell-wall viral and bacterial polysaccharides, these monomers combine with other types of molecules to form glycolipids, glycoproteins and nucleic acids. These are the molecules of *glycobiology*. The crystal structure analyses of the monosaccharides provide basic structural data for modelling these more complex molecules for which direct structural information is less readily available.

As shown in Fig. 1,* crystal structures have been determined for most of the pentoses and hexoses in one or more of the isomeric forms or as a 1-*O*-methyl derivative. Notable exceptions are D-glyceraldehyde, the tetroses, erythrose and threose, and the aldohexoses, idose and gulose. Glyceraldehyde crystallizes as a cyclic dimer which is formed by a spontaneous aldol condensation reaction and is misnamed D,L-glyceraldehyde, the crystal structure of which is known (Senma, Taira, Osaki & Taga, 1973). The tetroses are only known as syrups. Idose isomerizes to the ketose, sorbose. Gulose has never been crystallized. It is isolated and purified as a crystalline $CaCl_2.H_2O$ complex for which there is no crystal structure analysis. The crystal structures of the pentaacetates of gulopyranose (PACDGP) and idopyranose (PAIDOP) have been determined.

The aldoses and ketoses exist in aqueous solution as a configurational mixture of four isomers, α- and β-pyranoses, α- and β-furanoses. The acyclic (Fischer) configuration is useful for classification purposes, but there is no evidence of any significant proportion in solution, with the possible exception of D-idose (*cf.* Angyal, 1968). In general, only one of these four isomers crystallizes and this is presumably the predominant species or the least soluble in the solvent used. This gives rise to the phenomenon of *mutarotation*, whereby a particular isomer in the crystal slowly changes to the equilibrium mixture of configurations after dissolution, with a corresponding change in optical rotation. Crystalline pyranoses predominate, with relatively few examples of crystalline furanoses. The furanose isomers of D-ribose and 2-deoxy-D-ribose, which are a feature of the nucleosides, nucleotides and nucleic acids, have never been isolated and crystallized. Configurational heterogeneity in solution tends to inhibit crystallization. For this reason 1-*O*-methyl derivatives, which cannot

* Literature references for the REFCODEs cited throughout this paper have been deposited with the British Library Document Supply Centre as Supplementary Publication No. SUP 52467 (22 pp.). Copies may be obtained through The Technical Editor, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England.

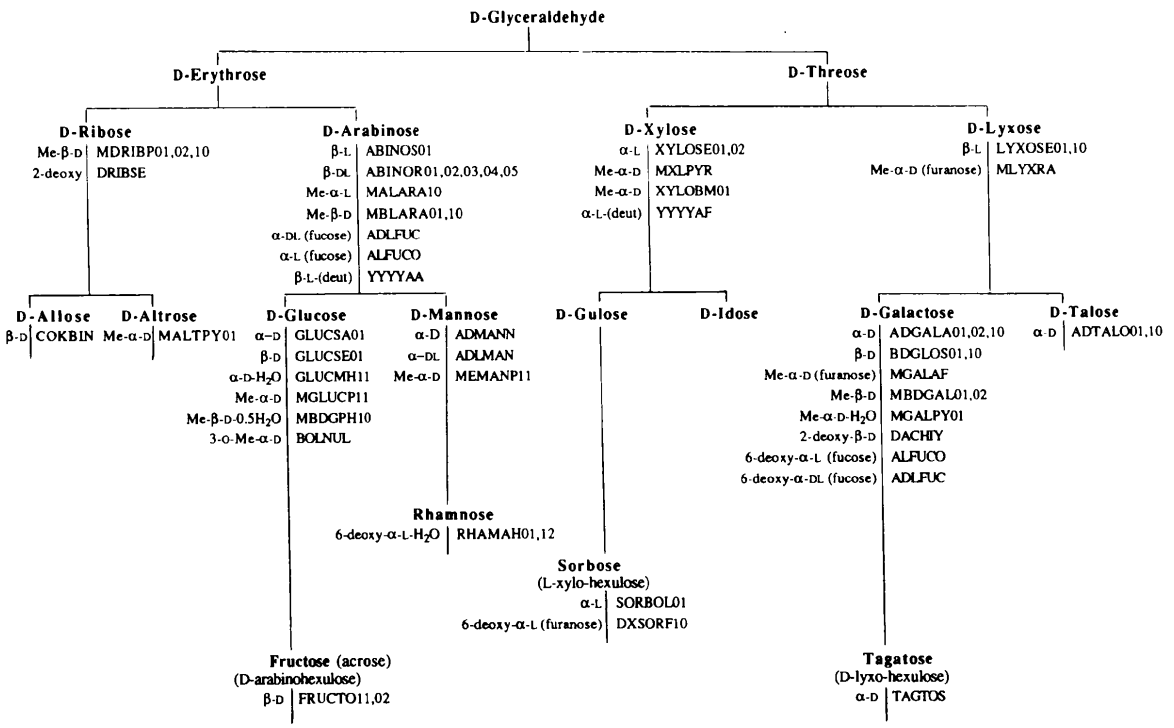


Fig. 1. The aldoses and ketoses. The pentoses ($C_5H_{10}O_5$) and hexoses ($C_6H_{12}O_6$) are pyranoses except where noted otherwise.

Table 1. *Examples of cocrystallization of α,β epimers*

	Percentage proportions α	β	REFCODE
6-Deoxy-α-L-sorbofuranose	95	5	DXSORF10
Sodium α-L-guluronate dihydrate (87 K)	90	10	DUDGUE
α-Lactose monohydrate	93	7	LACTOS10
	100	0	LACTOS03*
Laminarabiose hemihydrate (20% H_2O) (3-*O*-α-glucosyl-β-glucose; disorder between α anomer + $\frac{1}{2}H_2O$ and β anomer)	40	60	LAMBIO
α,β-Maltose	82	18	MALTOT
	84	16	
α,β-Melibiose monohydrate	72	28	MELIBM10
(6-*O*-α-D-galactopyranosyl-	85	15	MELIBM02
α,β-D-glucopyranose monohydrate)	80	20	MELIBM01
			MELIBM03
Mannobiose [*O*-β-D-mannopyranosyl-(1→4)-α,β-D-mannopyranose]	68	32	DIHTUJ
β-Galabiose [*O*-α-D-galactopyranosyl-(1→4)-β-D-galactopyranose]	56	44	CITSIH
Lactulose (4-*O*-β-D-galactopyranosyl-β-D-fructofuranose) (the remaining 15% is β-D-fructopyranose)	75	10	BOBKUY10

*Crystals grown from commercial samples of α-lactose monohydrate generally contain up to 10% of the β epimer. Very slow growth from ion-free solution gives crystals with no β epimer. The pure α crystals have a 0·9% lower unit-cell volume than the α,β mixture. An early analysis of *N*-acetyl-α-D-glucosamine reported ~20% β anomer, but the later work does not reproduce this (ACGLUA10; ACGLUA11).

epimerize, crystallize more readily. The furanosides for which crystal structures are available are methyl α-D-lyxo- (MLYXRA), ethyl 1-thio-α-D-gluco- (ETGLFR), 1-deoxy-1-nitro-β-D-ribo- (FACREG) and a $CaCl_2.2H_2O$ complex of β-D-mannofuranoside (MANCAC).

The α and β epimers can cocrystallize, as shown by the examples in Table 1. The α and β ratio depends on the temperature and solvent of crystallization and is not necessarily reproducible between investigators. Epimer cocrystallization is more common in disaccharides than in monosaccharides, which is understandable. Accommodating a misfit at one C—OH is a lesser defect in a structure with eight hydroxyls per molecule than in one with only four.

An extreme example of cocrystallization from a configurational mixture in solution is provided by the crystal structure of the disaccharide lactulose, 4-*O*-β-D-galactopyranosyl-D-fructose (BOBKUY10). The crystals contained a mixture of isomers in which the fructose moiety was β-D-fructofuranose, α-D-fructofuranose and β-D-fructopyranose in the ratio 0·745:0·100:0·155, see Fig. 2. The phenomenon was identified by ^{13}C cross-polarization magic-angle spinning (CP MAS) NMR spectroscopy. The deconvolution of the electron distribution in the X-ray analysis for the fructose residue would not have been possible without the prior NMR spectroscopy. In all these molecules, the galacto components are superimposable and the perturbation of the intermolecular hydrogen bonding is relatively minor (Wood, Jeffrey, Pfeffer & Hicks, 1983). An example of cocrystallization of a preparative mixture of isomers was observed in the crystal structure of α/β-methyl 2-chloro-2-deoxy-D-galactopyranoside (COGALP) (Hoge & Trotter, 1969).

Solid-state ^{13}C NMR spectroscopy has been applied to a number of carbohydrates (Pfeffer, 1984) using the CP MAS technique with powder samples. More interesting to the crystallographer is single-crystal ^{13}C NMR spectroscopy which can provide the complete chemical shift tensor for each carbon atom. These are plotted using the *ORTEP* plot (Johnson, 1965). This method requires large crystals and a specialized single-crystal NMR spectrometer. The only carbohydrate to which it has been applied, see Fig. 3, is methyl α-D-glucopyranoside (Sastry, Takegoshi & McDowell, 1987). The combination of X-ray or neutron crystallography with NMR spectroscopy is a promising future area of crystallographic research (*cf.* Etter, 1989; Etter, Hoge & Vojta, 1988; Jeffrey & Yeon, 1986; Harris, 1988).

In the hexopyranoses, the ring conformation is invariably $^4C_1$-D (or $^5C_2$-D), irrespective of configuration, because this places the primary alcohol C(6)H_2OH group in the favored equatorial orientation relative to the ring. The possible exception is α-D-idopyranose, where an equatorial primary alcohol group with a $^4C_1$ ring conformation implies four axial hydroxyls. This conformation could be stabilized in the crystal by O(1)···O(3) and O(2)···O(4) intramolecular hydrogen bonds. In the pentaacetate of α-D-idopyranose (PAIDOP), the ring

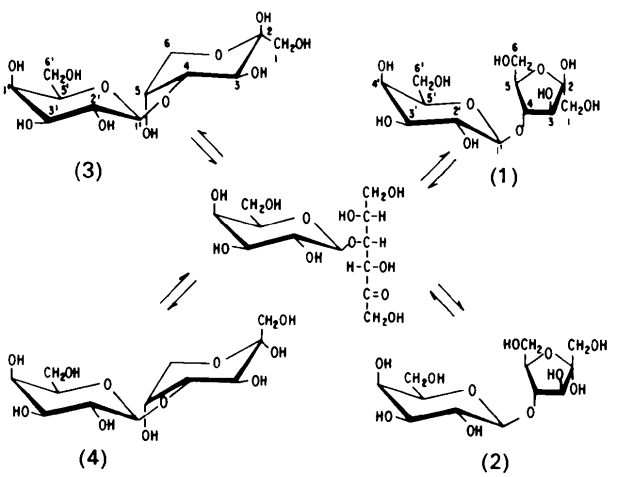


Fig. 2. The configurational equilibrium of 4-*O*-β-D-galactopyranosyl-D-fructose (lactulose); the isomers (1), (2) and (3) occur in the crystal structure in the ratio 0·745:0·100:0·155.

conformation is $^4C_1$-D. In the pentapyranoses, where this group is absent, both $^4C_1$-D and $^1C_4$-D conformations are observed. The conformer observed in the crystal, and predominating in solution, is that with the least number of axially oriented hydroxyl groups. When this criterion is ambiguous, intramolecular hydrogen bonding may determine which conformer occurs in the crystal. This is the case with methyl 5-thio- and methyl 1,5-dithio-α-D-ribopyranosides, where the $^1C_4$ ring conformation is stabilized by an O(2)H···O(4) intramolecular hydrogen bond (Girling & Jeffrey, 1973*a–c*).

The pyranose ring is relatively rigid and its conformation varies only slightly with configuration and molecular packing in the crystal structures of the pyranoses and methyl pyranosides. In terms of the Cremer & Pople (1975) puckering parameters, the range of values is $Q = 0{\cdot}55$–$0{\cdot}58$ Å with θ within 5° of 0 or 180°. Because of ring closure there is a range of C—C bond lengths between 1·50 and 1·56 Å. These ring conformations were adequately modelled using a molecular-mechanics program (*MM*1) which had been parameterized for alkanes and alcohols, with some modifications for the anomeric effect (*MM*1-*CARB*) (Jeffrey & Taylor, 1980; Tvaroska & Perez, 1986).

The '*standard*' dimensions for pyranoside residues, given in Table 2, for use in the molecular modelling of polysaccharides have been derived from surveys of accurate single-crystal analyses of relevant small molecules (Arnott & Scott, 1972; Jeffrey & Taylor, 1980).

The furanose ring is far from rigid and in the absence of constraining substituents can undergo *pseudo-rotation*. The factors governing the conformation of the ribofuranoside and 2-deoxyribofuranoside components of the crystal structures of the nucleosides and nucleotides have been analyzed in a very influential paper by Sundaralingam (1969).

In the hexapyranoses, the orientation of the primary alcohol group C(6)OH is of interest. In a 4-deoxyhexopyranose, *i.e.* with no hydroxyl on C(4), the three staggered orientations are equally likely. These are defined as *gg*, *gt* and *tg*, where the first symbol refers to the torsion angle O(5)—C(5)—C(6)—O(6) and the second to C(4)—C(5)—C(6)—O(6) [$g \simeq \pm 60°$, $t \simeq 180°$; an alternative nomenclature refers only to O(5)—C(5)—C(6)—O(6) as $+g = gt$, $-g = gg$, $t = tg$].

With an equatorial hydroxyl on C(4), *i.e.* the $^4C_1$-D-*gluco* configuration, the *tg* orientation is *forbidden*. With an axial hydroxyl on C(4), *i.e.* the $^4C_1$-D-*galacto* configuration, the *gg* orientation is *forbidden*. A survey of crystal structures by Perez, St Pierre & Marchessault (1978) showed that the permitted orientations were almost equally populated. The *forbidden* orientation was only observed in special cases where it is stabilized by intramolecular hydrogen bonding. This conformational factor was first recognized by Hassel & Ottar (1947). It is more generally referred to as the *peri* interaction or the 1,3-*syndiaxial* interaction. It is important in both cyclic and acyclic carbohydrates when there are C—OH bonds on C(n) and C($n+2$) in parallel alignment. It is believed to be associated with the unfavorable alignment of the parallel C—OH bond dipoles, but a convincing theoretical analysis is not yet available. It is not necessarily reproduced by molecular mechanics (Jeffrey & Taylor, 1980).

Twofold disorder can occur between the two permitted orientations as observed in the crystal structures of *meso*-erythritol (MERYOL03) and α-L-

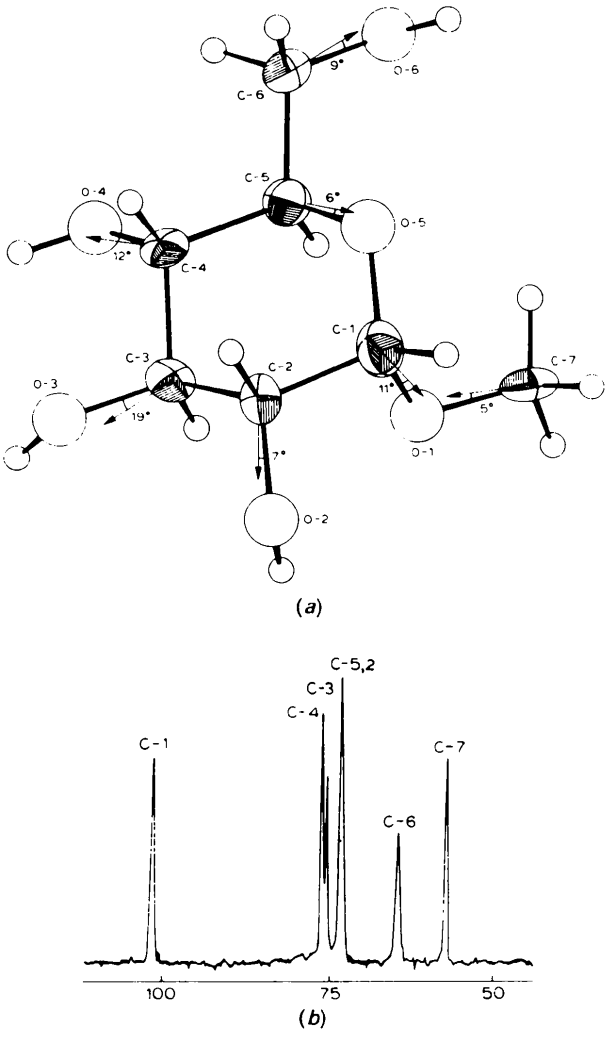


Fig. 3. (*a*) *ORTEP* representation of the ^{13}C chemical shielding tensors from a single crystal of methyl α-D-glucopyranoside. The arrows indicate the direction of greatest shielding. (*b*) ^{13}C CP MAS spectrum from powdered methyl α-D-glucopyranoside (p.p.m. from Me_4Si) (from Sastry, Takegoshi & McDowell, 1987).

Table 2. *Standard molecular dimensions for pyranosides (from Jeffrey & Taylor, 1980)*

Values in parentheses are from Arnott & Scott (1972).

Bond lengths (Å)		Valence angles (°)		Torsion angles (°)	
C—C(ring)	1·526 (1·523)	C—C—C(ring)	110·4 (110·5)	(Subject to ring-closure requirements)	
C—C(exo)	1·516 (1·514)	C—C—C(exo)	112·5 (112·7)		
C—O(exo)	1·420 (1·426)	C—C—O(ring)	110·0 (110·0)		
		C—C—O(exo)	109·7 (109·6)		
Axial glycosidic bond; α-D-$^4C_1$					
C(5)—O(5)	1·434 (1·436)	C(5)—O(5)—C(1)	114 (114·0)	C—C—C—C(ring)	53 (53)
C(1)—O(5)	1·419 (1·419)	O(5)—C(1)—O(1)	112·1 (111·6)	C—C—C—O(ring)	56 (55)
C(1)—O(1)	1·398 (1·415)			C—C—O—C(ring)	60 (61)
Equatorial glycosidic bond; β-D-$^4C_1$					
C(5)—O(5)	1·426 (1·436)	C(5)—O(5)—C(1)	112 (112·0)	C—C—C—C(ring)	53 (58)
C(1)—O(5)	1·428 (1·429)	O(5)—C(1)—O(1)	108 (107·3)	C—C—C—O(ring)	57 (55)
C(1)—O(1)	1·385 (1·389)			C—C—O—C(ring)	64 (61)

sorbopyranose (SORBOL01). The *forbidden* orientation is never observed in the crystalline cyclodextrins, but is a controversial feature of one of the models for the structure of cellulose II (Blackwell, Gardner, Kolpak, Minke & Claffey, 1980).

The heptoses are relatively unimportant and few occur naturally, alone or in combination. Some crystal structures have been reported, *i.e.* D-*altro*-3-heptulopyranose, coreose (COROSE10); D-*glycero*-β-D-heptopyranose, β-D-*gluco*-heptose (BDGHEP); α-D-*manno*-2-heptulopyranose (ADMHEP); D-*manno*-3-heptulopyranose monohydrate (MANHEP). None of these molecules have seven-membered rings.

Crystal structures have been reported for the biologically important *N*-acetyl-2-amino-2-deoxypyranoses of α-glucose (ACGLUA11), α-galactose (AGALAM01,10) and β-mannose monohydrate (NACMAN10). *N*-Acetyl-α-galactosamine has the primary alcohol group in the unfavorable *gg* conformation, stabilized by an O(4)H···O(6) intramolecular hydrogen bond.

Other biologically interesting crystal structure analyses are methoxyneuraminic acid trihydrate (MNURAC), β-D-*N*-acetylneuramic acid dihydrate (sialic acid) (SIALAC), its methyl ester monohydrate (ANEUME), and *N*-acetylmuramic acid, NAM (AMURAC). In all four molecules the pyranose ring has the expected $^4C_1$ conformation, with all the bulky substituents equatorial. The plane of the *N*-acetyl groups is approximately normal to the pyranose ring. The molecules have extensive hydrogen bonding in the crystal, and in NAM there is an NH···O=C intramolecular hydrogen bond which influences the orientation of the *N*-acetyl group.

The alditols

The alditols, or sugar alcohols, have a geneological family related to that of the aldoses. They differ from the aldoses in having members of the series which have *meso* configurations. As shown in Fig. 4, the crystal structures of all the pentitols and hexitols

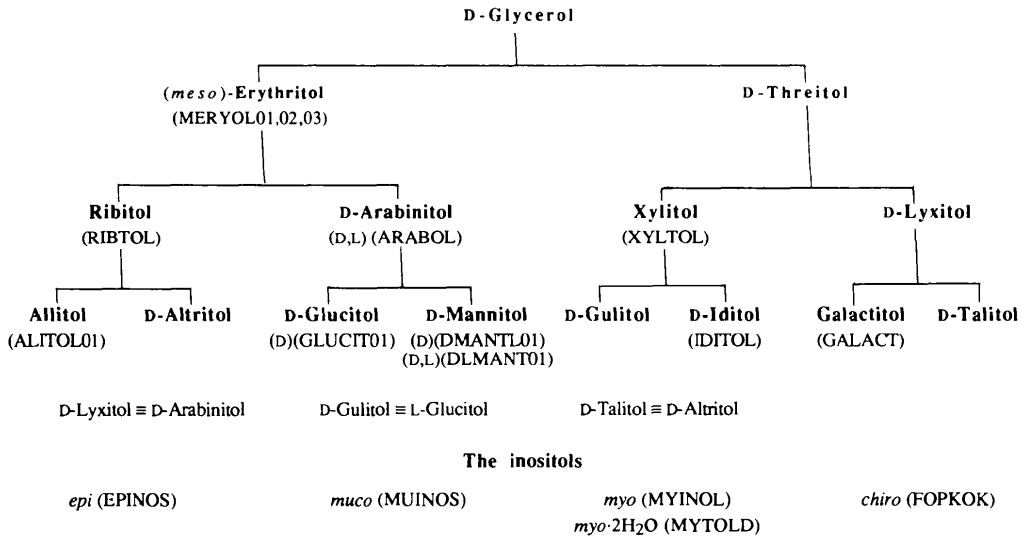


Fig. 4. The alditols.

have been determined, except for D-talitol (D-altritol) which does not occur naturally. These crystal structures provide examples of two interesting comparisons between solid-state and solution conformations.

(1) *Meso* configurations can have enantiomorphic conformations, both in solution and in the crystalline state. This occurs with ribitol and xylitol. In ribitol, the carbon chain is bent so that two conformational enantiomers occur in the same crystal in the centrosymmetric space group $P2_1/c$, see Fig. 5. The carbon chain is also bent in xylitol and there are two conformational enantiomers, but they appear in different crystals in space group $P2_12_12_1$. As the crystals dissolve, the conformers equilibrate to an optically inactive solution. With a large xylitol single crystal, this dissolution–equilibration process might be slow enough to provide an example of *conformational mutarotation*.

(2) The alditols provide a clear example of a single conformation in the crystal going to multiple conformations in solution. The optically active alditols have very low molecular optical rotations, never exceeding $[\alpha_D]$ water of 3·5° (Pigman, 1957). This requires that there is a population of rotamers in solution. On crystallization, only one is observed, even in the case of D-mannitol where there is more than one crystal structure (DMANTL; DMANTL01). No *conformational polymorphism* is observed with these compounds.

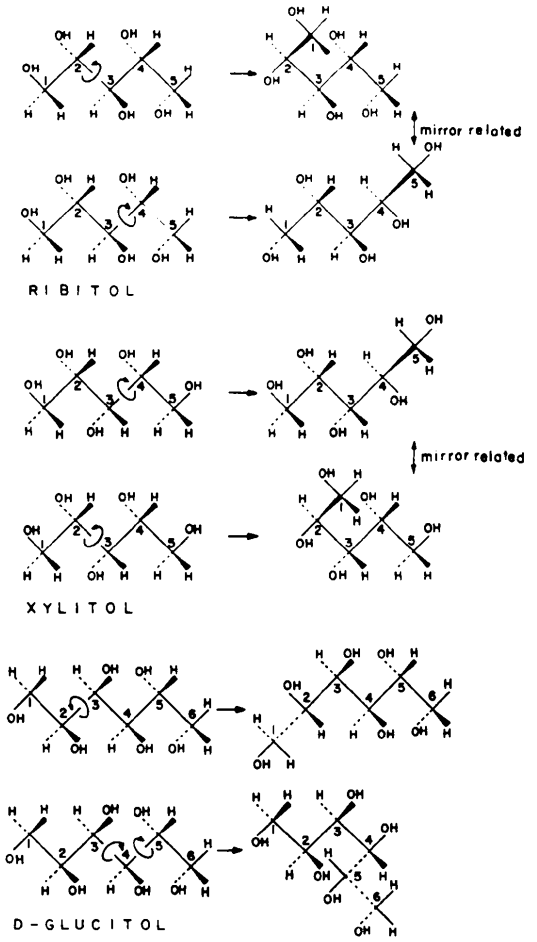


Fig. 5. The conformational changes in ribitol, xylitol and D-glucitol necessary to avoid the unfavorable 1,3-syndiaxial, *peri*, interaction between C(n)—OH and C(n+2)—OH.

Table 3. *Conformation of the pentitols and hexitols in the crystalline state* (*Jeffrey & Kim*, 1970)

Alditol	Configuration	Carbon-chain conformation	Crystal structure
D,L-Arabinitol	LDD	Straight	ARABOL
Ribitol	DDD	Bent	RIBTOL
Xylitol	DLD	Bent	XYLTOL
Allitol	DDDD	Bent	ALITOL01
Galactitol	DLLD	Straight	GALACT
D-Glucitol	DLDD	Bent	GLUCIT01
D-Iditol	LDLD	Bent	IDITOL
D-Mannitol	LLDD	Straight	DMANTL01
D-Talitol	LLLD	Bent	Predicted

The conformation in the crystal is controlled by the Ottar–Hassel effect referred to previously. When the C(n)—OH and C(n + 2)—OH bonds are parallel in the straight-chain conformation, one or more of the C—C—C—C torsion angles adopts the *gauche* conformation to avoid this interaction, as shown in Fig. 5.

The solid-state conformation of the alditols can therefore be predicted from the configuration at each carbon atom along the chain. When the configuration is the same at alternate carbon atoms, the carbon chain is bent, as shown in Table 3.

There may be two ways of avoiding this unfavorable '*peri*' interaction, as in the case of D-glucitol and D-iditol, which gives rise to two conformers which are not enantiomers, see Fig. 5 (Jeffrey & Kim, 1970). In D-glucitol (GLUCIT) the bent chain conformer has a *gauche* C(2)—C(3)—C(4)—C(5) torsion angle, whereas in the D-glucitol–pyridine complex (SORBPY20), the conformation is that in which both C(2)—C(3)—C(4)—C(5) and C(3)—C(4)—C(5)—C(6) are *gauche*.

In the heptitols *meso*-L-*glycero*-L-gulohepitiol (MGGHEP) and *glycero*-L-alloheptitol (FAJLOR), the same rule applies. The former has configuration LLDLL with a *gauche* torsion angle at C(3)—C(4)—C(5)—C(6). The latter is DLLLL with a *gauche* torsion angle at C(4)—C(5)—C(6)—C(7). A similar conformation factor applies to the acetylated alditols in non-polar solutions as studied by NMR spectroscopy (Horton & Wander, 1969).

The rotomeric population and hence the low optical activity of the D-alditols in aqueous solution is an indication that this conformational effect is attenuated by solvation, probably through the for-

mation of C(n)OH···(H_2O)···C(n + 2)OH hydrogen bonds.

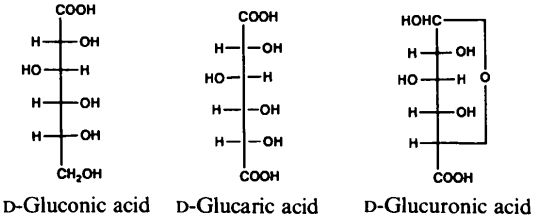
Another example where solvation affects conformation is observed in *trans-O-β*-D-glucopyranosyl methylacetoacetate which has a large positive optical rotation in water and a large negative rotation in 1,4-dioxane. A crystal structure analysis (GLPMAC10) provided a stereochemical basis for interpreting this observation in terms of a conformational difference with solvent (Ruble & Jeffrey, 1974).

The cyclitols

There are nine inositols, $C_6H_6(OH)_6$, the configurations of which differ by the axial or equatorial disposition of the hydroxyl groups around the chair-shaped cyclohexane ring. Crystal structure analyses have been reported for *chiro-* (FOPKOK), *epi-* (EPINOS), *muco-* (MUINOS), *myo-* (MYINOL) and *myo*-inositol dihydrate (MYTOLD). The crystal structures of three complexes are also reported: *epi*-inositol.$SrCl_2.5H_2O$ (EPINSR), *myo*-inositol.-$CaBr_2.5H_2O$ (MYINCA10) and *myo*-inositol.$MgCl_2.$-$4H_2O$ (INSMGC). These structures have no conformational surprises. The cyclohexane ring is a chair and the conformation is that with the minimum number of axial hydroxyl groups. In *epi*-inositol (EPINOS), where parallel C—OH bonds on C(n) and C(n + 2) are unavoidable, the ring is distorted by a repulsive interaction between the oxygen atoms. The alternative formation of an intramolecular hydrogen bond is not observed in the crystal structure (Jeffrey & Kim, 1971). In the salt complexes, the cation and anion coordinations are similar to those observed in the salts and complexes of the carbohydrate acids discussed below.

The carbohydrate acids

The aldonic (I), aldaric (II) and alduronic (III) acids are related to the alditols and aldoses, for example:



D-Gluconic acid (I) D-Glucaric acid (II) D-Glucuronic acid (III)

Like the alditols, some aldaric acids have *meso* configurations and are optically inactive. The aldonic and aldaric acids lose water to form cyclic five-membered 1,4-lactones and cyclic six-membered 1,5-lactones. The crystal structure of D-gluconic acid monohydrate (CAKZAP) shows that the molecule has the bent-chain conformation with the bend at C(2)—C(3). Bent-chain conformations are also observed in the crystal structures of various D-gluconate salts (NH_4^+, $Mn^{2+}.2H_2O$, Na, K, $K.2H_2O$, Pb^{2+}) (Lis, 1983). The alduronic acids form 1,4-lactones which are bicyclic with fused five-membered rings. Crystal structures have been reported for the following lactones: D-glucono-1,5- (GLULAC10), D-arabino-1,4- (ARALAC), D-ribono-1,4- (BAGZAK), D-galactono-1,4- (GALLAD), D-gulono-1,4- (GULONO10), D-glucono-1,4- (DGLACM10) and *β*-D-glucurono-1,4-lactone (GLULAD).

The only aldaric acid crystal structure is that of galactaric acid (mucic acid, BIVTUV). This compound is insoluble in cold water. The hydrogen-bond structure in the crystal is exceptionally strong since it includes a cyclic carboxylic acid dimer and a homodromic four-membered ring, as shown in Fig. 6. This is an example where the hydrogen-bond energy exceeds the solvation energy plus entropy at room temperature. D-Glucaric acid is deliquescent, but its crystal structure has not been determined.

The alduronic acids and the alduronate anions are pyranoses and can exist as *α* or *β* epimers. The ring conformation is $^4C_1$-D with the equatorial orientation of the carboxylate group. This is the conformation in the crystal structure of *α*-D-glucuronamide (AGLCAM). *N,N*-Diethyl-D-gluconamide (DAYWAB) and *N*-cyclohexyl-D-gluconamide (CIBWIT) have the acyclic straight-chain conformation, while in *N*-isopropyl-D-gluconamide (DAYVUU) the acyclic chain conformation is bent. The D-gluconate ion in the crystal structure of potassium D-gluconate has a straight-chain conformation in one form and a bent-chain conformation in a second form. These structures were studied by neutron diffraction (Panagiotopoulos, Jeffrey, LaPlaca & Hamilton, 1974) and provide an interesting example of alternate modes of hydrogen bonding, which is discussed below.

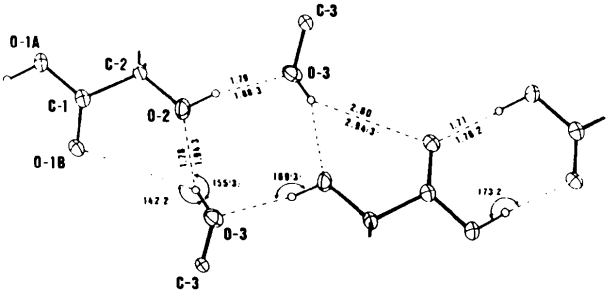


Fig. 6. Strong hydrogen bonding in the crystalline structure of mucic acid, an aldaric acid which is insoluble in cold water. Distances in Å, angles in °.

Alkali-earth salts and complexes

The mono- and disaccharides from alkali-earth salt complexes and the carbohydrate acids form salts which are relatively easy to prepare and crystallize. The calcium compounds, in particular, have received much attention owing to the importance of calcium in human metabolic processes. The crystal structures of more than twenty carbohydrate salts and complexes have been reported including four disaccharide complexes. The calcium coordination and hydration have been analyzed in detail by Einspahr & Bugg (1977, 1980) and Cook & Bugg (1977).

It would be of interest to know how the cation coordination modifies the cooperativity of the carbohydrate hydrogen bonding (*cf.* Jeffrey & Mitra, 1983) and whether this has any relationship to the effects of salts on the hydrogen bonding in water.

The anhydro sugars

These compounds are formed by the elimination of water between hydroxyls of the pyranoses and furanoses thereby forming fused bicyclic or tricyclic ring structures. They are stable molecules which generally crystallize well and there are about 40 crystal structure analyses of this type reported. Like the cyclic alkanes, these are molecules with rigid conformations which can be predicted very successfully by standard methods of molecular mechanics (*cf.* Engler, Andose & von Schleyer, 1973; Jeffrey & Park, 1979; Ceccarelli, Ruble & Jeffrey, 1980). The molecules have a higher ratio of ether oxygens to hydroxyls than the pyranoses and pyranosides, which influences the intermolecular hydrogen bonding as illustrated by the crystal structure of 1,6-anhydro-β-D-galactopyranose (Ceccarelli *et al.*, 1980).

The ascorbic acids and ascorbates

The crystal structure of L-ascorbic acid has been studied by X-ray and neutron diffraction (LASCAC01,10; YYYYAD), and that of D-isoascorbic acid by X-rays (IASCOR10). The iso acid differs only in the configuration of C(5) in the side chain. Crystal structures are also reported for sodium, calcium dihydrate and thallium ascorbates (NAASCB; CAASCO01,02; TLASCB) and sodium monohydrate isoascorbate (SIASCB).

The interest in these structures has been in identifying the protolytic hydroxyl, O(3)H, and the distortions of the ideally planar enofuranose ring arising from the repulsion of the three substituent oxygen atoms (Kanters, Roelofsen & Alblas, 1977). The molecule of isoascorbic acid and one of the two molecules per cell of ascorbic acid are close to planar with endocyclic torsion angles less than 2·1°. The other ascorbic acid molecule has a twist distortion similar to that observed in the isoascorbate ion, but less than that in the ascorbate ion. These structures are extensively hydrogen bonded and the anionic O(3) of the isoascorbate ion accepts four hydrogen bonds from two hydroxyls and two waters. These structures merit closer study using modern molecular-modelling methods.

The oligosaccharides

The oligosaccharides are natural products in which two or more monosaccharides are linked by condensation of a reducing group [C(1)OH in aldoses, C(2)OH in ketoses] with another hydroxyl. This linkage may be $1\rightarrow n$, where n is 1 to 6, except 5. If two reducing groups are involved, the disaccharide is non-reducing as in sucrose and $\alpha\alpha'$-trehalose. There are 32 crystal structures of unsubstituted disaccharides, as shown in Table 4, of which five involve hexitols. There are six crystal structures of the fully acetylated disaccharides, which include octaacetylcellobiose (ACELLO), octaacetyl-β-maltose (ZZZTUC10) and octaacetylsucrose (ZZZSTI01). There are eight trisaccharide crystal structures and only one tetrasaccharide.

One linear oligoamylose has been crystallized and analyzed as the polyiodide complex, that of *p*-nitrophenyl-α-D-maltohexopyranoside. The maltohexose molecules form a double helical structure which encloses the polyiodide chain in a way which is analogous to the cyclodextrin inclusion compounds (Hinricks, Büttner, Steifa, Betzel, Zabel, Pfannemüller & Saenger, 1987).

The torsion-angle homologies between similar linkages among the crystal structures of the unsubstituted oligosaccharides are shown in Table 5. For a particular type of linkage, the torsion angles generally fall between limits of $\pm 30°$, irrespective of the monomers involved. These limits are defined by the inter-residue van der Waals interactions and lie within the permitted regions of simple hard-sphere Ramanchran potential-energy maps. The finer details are determined by intramolecular inter-residue hydrogen bonds and the intermolecular hydrogen bonds and van der Waals forces in the crystal. Because of the rotational freedom of the hydroxyl groups, the accurate prediction of oligosaccharide conformations is a rigorous test of empirical force-field parameters in molecular modelling.

The $1\alpha\rightarrow 4$ and $1\beta\rightarrow 4$ linkages are important because they occur in the amyloses and celluloses, respectively. These linkages are favorable for inter-residue hydrogen bonding and, consequently, the torsion angle ranges are somewhat narrower. In contrast, the $1\rightarrow 6$ linkages involve three linkage

Table 4. *Crystal structures of oligosaccharides*

(*a*) Disaccharides (excluding salts and complexes)		
2-*O*-β-Glucopyranosyl α-glucopyranose	Sophorose.H_2O	(SOPROS)
3-*O*-β-Glucopyranosyl α-glucopyranose	Laminarabiose.-0·19H_2O	(LAMBIO)
4-*O*-α-Glucopyranosyl α-glucopyranose	α-Maltose	(MALTOT)
4-*O*-β-Galactopyranosyl α-glucopyranose	α-Lactose.H_2O	(LACTOS01)
6-*O*-α-Galactopyranosyl α-glucopyranose	Melibiose.H_2O	(MELIBM01)
1-*O*-α-Glucopyranosyl α-glucopyranoside	α,α′-Trehalose	(DEKYEX)
1-*O*-α-Glucopyranosyl α-glucopyranoside	α,α′-Trehalose.-2H_2O	(TREHAL01)
3-*O*-α-Glucopyranosyl methyl 1-*O*-α-glucopyranoside	Methyl nigeroside	(MOGLPR)
4-*S*-α-Glucopyranosyl methyl 1-*O*-α-glucopyranoside		
4-*S*-α-Glucopyranosyl methyl 1-*S*-α-glucopyranoside		(CIBYAN)
4-*O*-α-Glucopyranosyl β-glucopyranose	β-Maltose.H_2O	(MALTOS10)
4-*O*-β-Glucopyranosyl β-glucopyranose	Cellobiose	(CELLOB)
6-*O*-β-Glucopyranosyl β-glucopyranose	Gentiobiose	(GENTBS)
4-*O*-β-Galactopyranosyl β-glucopyranose	β-Lactose	(BLACTO)
6-*O*-α-Galactopyranosyl β-glucopyranose	Melibiose.H_2O	(MELIBM01)
4-*O*-α-Glucopyranosyl methyl 1-*O*-β-glucopyranoside	β-Methylmaltoside	(MMALTS)
4-*O*-β-Glucopyranosyl methyl 1-*O*-β-glucopyranoside	β-Methylcellobioside	(MCELOB)
6-*O*-Glucopyranosyl methyl 1-*O*-β-glucopyranoside		(DINSOI)
3-*O*-α-Glucopyranosyl β-D-fructopyranose	Turanose	(TURANS)
5-*O*-α-Glucopyranosyl β-D-fructopyranose	Leucrose.H_2O	(LEUCRO)
6-*O*-α-Glucopyranosyl α-D-fructofuranose	Isomaltose.H_2O	(IMATUL)
4-*O*-β-Galactopyranosyl α-D-fructofuranose	Lactulose	(BOBKUY10)
1-*O*-α-Glucopyranosyl β-D-fructofuranoside	Sucrose	(SUCROS)
4-*O*-β-Galactopyranosyl α-D-mannopyranose		(DICMEH)
4-*O*-β-Mannopyranosyl α-D-mannopyranose		(DIHTUJ)
2-*O*-α-Mannopyranosyl methyl 1-*O*-mannopyranoside		(FABTOW)
4-*O*-α-Galactopyranosyl β-D-galactopyranose	Galobiose	(CITSIH)
2-Amino-2-deoxy-3-*O*-β-glucopyranosyl β-D-galactopyranose	Chondrosamine.H_2O	(CHONDM)
4-*O*-α-Glucopyranosyl D-glucitol	Maltitol	(BIZHIB)
4-*O*-β-Glucopyranosyl D-glucitol		(BDGPGL)
6-*O*-α-Glucopyranosyl D-glucitol	Isomaltitol	(BAVCAC)
1-*O*-α-Glucopyranosyl D-mannitol dihydrate		(BAGZEO)
4-*O*-β-Galactopyranosyl l-rhamnitol		(GAPRAM10)
(*b*) Trisaccharides		
O-α-D-Glucopyranosyl-(1→6)-*O*-α-D-glucopyranosyl-(1→4)-α-D-glucopyranose	Panose	(Imberty & Perez, 1988)
O-α-D-Galactopyranosyl-(1→6)-*O*-β-D-fructofuranosyl-(2→1)-α-D-glucopyranoside dihydrate	Planteose.2H_2O	(PLANTE10)
O-α-D-Glucopyranosyl-(1→3)-*O*-β-D-fructofuranosyl-(2→1)-α-D-glucopyranoside hydrate	Melezitose.H_2O	(MELEZT01)
O-α-D-Galactopyranosyl-(1→6)-α-*O*-glucopyranosyl-(1→2)-β-D-fructofuranoside pentahydrate	Raffinose.5H_2O	(RAFINO)
O-α-D-Glucopyranosyl-(1→2)-*O*-β-D-fructofuranosyl-(1→2)-β-D-fructofuranoside	1-Kestose	(KESTOS)
O-α-D-Glucopyranosyl-(1→2)-*O*-β-D-fructofuranosyl-(6→2)-β-D-fructofuranoside monohydrate	6-Kestose.H_2O	(CELGIJ)
O-β-D-Mannopyranoside-(1→4)-*O*-β-mannopyranosyl-(1→4)-*O*-α-D-mannopyranose trihydrate	Mannotriose.3H_2O	(COFMEP10)
O-α-D-Mannopyranoside-(1→3)-β-D-mannopyranoside-(1→4)-2-acetamido-2-deoxy-α/β-D-glucopyranose		(MPYAGL)
(*c*) Tetrasaccharides		
O-α-D-Galactopyranosyl-(1→6)-*O*-α-D-galactopyranosyl-(1→6)-*O*-α-D-glucopyranosyl-(1→2)-α-D-fructofuranoside hydrate	Stachyose.H_2O	(STACHY)

bonds, so that the residues are further separated and inter-residue hydrogen bonding is less common.

The most common inter-residue hydrogen bond is from a hydroxyl on one residue to the ring oxygen, O(5), on the adjacent residue. This occurs from O(3′)H to O(5) in cellobiose (CELLOB02) (Fig. 7), methyl cellobioside (MCELOB), β-galabiose (CITSIH), α-lactose monohydrate (LACTOS10), β-lactose (BLACTO), 4-*O*-β-D-mannopyranosyl-α-D-mannopyranose (DIHTUJ), lactose.$CaCl_2$.7H_2O (LACTCC10), and twice in mannotriose trihydrate (COFMEP10). This bonding requires that the pyranose ring oxygens are on opposite sides of the long molecular axis and leads to a linear polymer. Another conformation of cellobiose in which the ring oxygens are on the same side of the long molecular axis could be stabilized by intramolecular hydrogen bonds between O(2)H···O(3′) and O(6′)H···O(5). This conformation has not been observed. It would lead to β-glycosidic helical or cyclic polymers analogous to the α-glycosidic amyloses and cyclodextrins.

The O(3′)H···O(2)H or O(2′)H···O(3)H bonds that occur in α-maltose (MALTOT), β-maltose monohydrate (MALTOS11), methyl β-maltoside hydrate (MMALTS) and phenyl α-maltoside (PHMALT) require that the pyranose rings are oriented so that their ring oxygens are on the same side of the long molecular axis. This leads to a cyclic polymer in the cyclodextrins and helical structures in the amyloses.

In the crystal structure of sucrose, there are two inter-residue bonds, O(1′)H···O(2)H and O(6′)H···O(5), both from the fructofuranose to the glucopyranose residue. This neutron diffraction analysis provides a rare example of a four-center hydrogen bond with H···O distances of 2·20, 2·49 and 2·52 Å with O—H···O angles of greater than 90°.

Table 5. *Homologies in oligosaccharide linkage bond-torsion angles* (°)

	glu-py—fru-fu								acetylated	glu-py—glu-py
1α–2'	SUCROS	RAFINO	PLANTE,10	CELTIJ	MELEZT	CELGIJ	KESTOS	STACHY	ZZZSTI01	SOPROS
O5-C1-O1-C2'	+107.8	+81.7	+108.7	+89.6	+99.8	+89.6	+84.7	+109.6	+93.2	-78.9
C1-O1-C2'-O2'	-44.8	+11.4	-26.2	-54.5	-30.6	-54.5	-65.8	-47.3	-21.8	-139.8

	glu-py—glu-py							gal-py—gal-py	glu-py—glu-py acetyl
1α–4'	MALTOT	MALTOS	MMALTS,10	PANOSE	IPMALT	PHMALT		CITSIH	ZZZTUC01
O5-C1-O1-C4'	+116.1	+121.7	+109.3	+92.9	+72.5	+108.5	+110.5	+98.1	+84.2
C1-O1-C4'-C5'	-118.0	-109.7	-108.8	-131.3	-155.0	-139.4	-139.5	+157.7	-154.8

	glu-py—glu-py		gal-py—glu-py			man-py—man-py				gal-py—fru-py	gal-py—fru-fu	gal-py—gal-py acetyl
1β–4'	CELLOB	MCELOB	LACTOS01	BLACTO	LACTCC10	COFMEP10		MPYAGL	DIHTUJ	DICMEH	BOBKUY10	OACGAP
O5-C1-O1-C4'	-77.8	-91.1	-92.6	-70.9	-76.9	-71.9	-93.9	-75.7	-96.0	-83.2	-82.1	-81.6
C1-O1-C4'-C5'	-127.3	-160.7	-143.0	-131.5	-136.7	-131.6	-150.4	-130.7	-148.2	-109.9	-164.4	+123.6

	glu-py—fru-fu		glu-py—glu-py	man-py—man-py
1α–3	TURANS	MELEZT	MOGLPR	MPYAGL
O5-C1-O1-C3'	+99.2	+78.4	+99.9	+60.5
C1-O1-C3'-C4'	+111.2	+143.1	+106.2	+97.0

	gal-py—glu-py			glu-py—fru-fu		glu-py—glu-py
1α–6	RAFINO	MELIBM	STACHY	IMATUL	PLANTE10	PANOSE
O5-C1-O1-C6'	+71.1	+76.5	+64.8	+76.9	+58.5	+71.4
C1-O1-C6'-C5'	-169.6	-173.9	-174.9	+143.6	+172.5	+71.4
O1-C6'-C5'-O5'	-64.7	-63.4	-62.1	-64.6	+63.5	+75.7

	glu-py—glu-py	xyl-py—glu-py
1β–6	GENTBS	MCRZMA
O5-C1-O1-C6'	-58.3	-77.5
C1-O1-C6'-C5'	-156.3	-104.1
O1-C6'-C6'-O5'	-61.6	+66.4

The hydrogen bonding in this crystal structure is also unusual in that there are six other hydrogen bonds with H···O greater than 2·2 Å and only one less than 1·8 Å.

There can also be *indirect inter-residue hydrogen bonding* which includes a water molecule. This is observed in $\alpha\alpha'$-trehalose dihydrate (TREHAL01), melibiose hydrate (MELIBM10), methyl β-maltoside hydrate (MMALTS) (Fig. 7) and sophorose hydrate (SOPROS). *Indirect intra-residue hydrogen bonding*, where a water molecule bonds to hydroxyls on adjacent carbon atoms, *i.e.* C(2) C(3) or the ring oxygen O(5) and the primary alcohol C(6)H_2OH, has been reported in the heavily hydrated structure of the *p*-nitrophenyl α-D-maltopyranoside triiodide complex. This conclusion, however, is based on O···O(*W*) separations, since the hydrogen atoms were not located in this low-resolution X-ray analysis (Hinricks *et al.*, 1987). Similar chelation of water molecules is invoked in studies of carbohydrate solvation (Lemieux & Pavia, 1969), and may be the explanation for the conformational equilibrium of alditols in solution discussed above.

The crystal structure of laminarabiose quarterhydrate (LAMBIO) is interesting because it is a cocrystal of laminarabiose and laminarabiose hemihydrate. The crystal structure of turanose (TURANS01) contains a rare example of a *non-hydrogen bonding* hydroxyl. The closest approach to O(2)H of the fructopyranose residue is an O(3') of the same molecule with an H···O distance of 2·64 Å and an O—H···O angle of 97°. This non-bonding hydroxyl is identified with an absorption peak at 3600 cm^{-1} in the solid-state infrared spectrum (Kanters, Gaykema & Roelofsen, 1978).

The trisaccharides (Table 5) have curled conformations in their crystal structures. The higher homologues of the 1–4 linked glucoses which might be models for the celluloses and amyloses have not been crystallized for X-ray structure analysis. The tetrasaccharide stachyose hydrate (STACHY,10) has the glucopyranosyl–fructofuranoside component of sucrose and the galactopyranosyl–glucopyranose component of raffinose and melibiose with very similar linkage torsion angles, see Table 5.

The cyclodextrins (cyclooligoamyloses) provide excellent crystals and many crystal structure analyses have been reported. These compounds are cyclic 1–4 linked α-D-glucopyranosides with 6(α), 7(β) and 8(γ) monosaccharide residues. The molecules are shaped like truncated cones with the interior voids occupied by solvent or guest molecules. The molecular conformations are relatively rigid with the $^4C_1$ pyranose rings, and little variation in the torsion angles of the linkage bonds. The primary alcohol groups show the same two preferred orientations as are observed in the monopyranoses and pyranosides. There are three packing modes:

Type *A*: herringbone packing, when the voids are filled with water, methanol or ethanol.

Type *B*: lateral packing forming layers, when the voids are filled with larger guest species.

Type *C*: cylindrical packing, with guests which protrude beyond the voids or with ionic guests.

More than 80 crystal structures of the cyclodextrins and their complexes have been reported, including some neutron diffraction analyses. This field is well reviewed elsewhere (Saenger, 1984).

The anomeric effect

The anomeric effect is an important configurational and conformational influence in carbohydrate chemistry (Kirby, 1983). It was first observed by Edwards (1955) and developed primarily by Lemieux & Chu (1958) and Lemieux (1959). It is a particular application of the more general *gauche effect* (Wolfe, 1972) to carbohydrates. The anomeric carbon atom C(1) is unique in the pyranoses [as is C(2) in the furanoses] in that it is bonded to two more electronegative atoms, *i.e.*, the ring oxygen O(5) and the glycosidic oxygen O(1) or a halide atom in a glycosyl fluoride or chloride. This leads to a *configurational* preference for an axial electronegative substituent at the anomeric carbon atom, to bond-length differences and to preferred *gauche conformations* about the glycosidic bond (the *exo-anomeric effect*; Lemieux, Koto & Vorsin, 1979).

The first crystallographic observation of the anomeric effect was in the difference in C—O bond lengths in the crystal structure of *trans*-2,5-dichloro-1,4-dioxane (Altona, Knobler & Romers, 1963). A few years later, there were sufficient accurate crystal structure analyses of pyranoses and pyranosides that a systematic shortening of the C(1)—O(1) bonds was noted (Berman, Chu & Jeffrey, 1967). An investigation of this effect provided one of the earlier examples of collaboration between crystallography and quantum mechanics. Using successively, methanediol (Jeffrey, Pople & Radom, 1972), methoxymethanol (Jeffrey, Pople & Radom, 1974) and dimethoxymethanol (Jeffrey, Pople, Binkley & Vishveshwara, 1978), it was demonstrated that the bond shortening and preferred *gauche* conformation of the glycosidic bonds in pyranoses and methyl pyranosides was a consequence of an electronic distribution in the hemiacetal and acetal moiety of these molecules. These theoretical calculations reproduced the C—O bond-length differences and the preferred *gauche* conformations. The results were consistent with additional π-bonding in the C—O bonds from the C-2*p* orbital in the energetically favorable conformation. Attempts to observe this bonding effect in an experimental charge-density study of D,L-arabinose were unconvincing (Longchambon, Gillier-Pandraud, Weist, Rees, Mitschler, Feld, Lehman & Becker, 1985). A theoretical charge-density study of methanediol and fluorodimethyl ether also failed to show significant features which were consistent with this electronic interpretation (Pichon-Pesme & Hansen, 1989). A reinterpretation may be necessary.

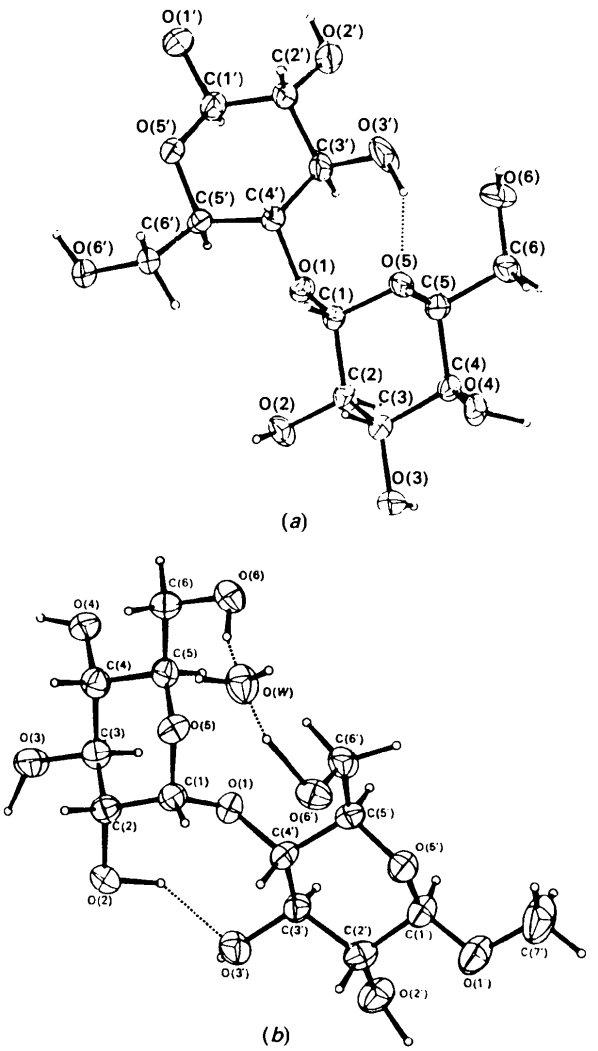


Fig. 7. (*a*) O(3′)H···O(5) inter-residue hydrogen bond in cellobiose (CELLOB). (*b*) O(2)H···O(3′) and O(6)H···O(*W*)···HO(6′) hydrogen bonds in methyl β-maltopyranoside hydrate.

Hydrogen bonding in carbohydrate crystal structures

The carbohydrates provide some of the simplest of the biological molecules for the analysis of hydrogen bonding, in the sense that they can contain only two functional groups; the hydroxyls, which can be donors or acceptors, and the ring and glycosidic oxygens, which are only acceptors. On the other hand, the prediction of hydrogen bonding by molecular modelling is more difficult than for nucleotides and peptides because of the orientational freedom of the C—OH group. For $>$NH···O=C$<$ and $>$NH···N$\lesssim$ bonds, the positions of the hydrogen atoms are determined by the non-hydrogen

Table 6. *Crystal structures of carbohydrate mesogens*

All crystal structures have head-to-head bilayers except MEGA-9 and MEGA-11. *CC*: crystal-to-crystal transitions. *MP*: melting point (crystal→liquid crystal). *CP*: clearing point (liquid crystal→liquid).

Compound	REFCODE or reference	*CC* (K)	*MP* (K)	*CP* (K)
Octyl 1-*O*-α-glucopyranoside	van Koningsveld, Jansen & Straathof (1988)			
Octyl 1-*O*-α-glucopyranoside.$\frac{1}{2}H_2O$	Jeffrey, Yeon & Abola (1987)	325	328–345	389
Octyl 1-*O*-α-glucopyranoside.H_2O	Jeffrey, Yeon & Abola (1987)			
Decyl 1-*O*-α-glucopyranoside	DECGPY10		346–349	403
Heptyl 1-*S*-α-glucopyranoside	van Doren, van der Giest & van Bolhuis (1989)		370	411
Heptyl 1-*S*-β-mannopyranoside	BIFPIP		333	424
Octyl 1-*S*-β-xylopyranoside	DARSUK07	345	373	387
N-(*n*-Octyl)gluconamide	DOJWII,* FAKFUS	353, 372	431	432
1-Deoxy(*N*-methyloctamido)-D-glucitol (MEGA-8)	Jeffrey & Maluszynska (1989)	335	345	371
1-Deoxy(*N*-methylnonamido)-D-glucitol† (MEGA-9)	FALKAE		363	
1-Deoxy(*N*-methylundecanamido)-D-glucitol (MEGA-11)	Jeffrey & Maluszynska (1989)		365	386

* Misidentified as *N*-octyl-D-gluconate.
† Misnamed nonanoyl-*N*-methylglucamide.

molecular structure. For C—OH···O bonds, a variable torsion-angle parameter is introduced for every hydrogen bond considered.

Since there are a relatively larger number of neutron diffraction and high-precision X-ray analyses, an analysis of the rules governing hydrogen-bonding in the monosaccharides and disaccharides was possible (Jeffrey & Mitra, 1983). These rules depend upon two concepts:

(1) maximize the hydrogen-bond interactions by including all hydroxyls and as many ring and glycosidic oxygens as possible, using both two- and three-center bonds (Jeffrey, 1978, 1987);

(2) maximize cooperativity (non-additivity) by forming as many finite and infinite chains of hydrogen bonds as possible. It is interesting to note that the application of these rules did not provide a means of predicting the type of hydrogen-bond structure for a particular carbohydrate molecule. Four different patterns were identified and the crystal structures were divided approximately equally between them.

With very rare exceptions, hydroxyls function as both hydrogen-bond donors and acceptors. The anomeric hydroxyl tends to be a strong donor and a weak acceptor compared with the other hydroxyls. It is not always possible, for stereochemical and packing reasons, for ring and glycosidic oxygens to be acceptors. Because there are more acceptors than donors, three-centered (bifurcated-donor) bonds constitute about 20% of the total number (Ceccarelli, Jeffrey & Taylor, 1981). Bifurcated-acceptor bonds postulated for the water dimer (Pimentel & McClelland, 1960) are never observed, except in combination with three-center bonds.

Hydrates are more common in disaccharides than in monosaccharides. This is understandable because monosaccharides tend to be spheroidal molecules which can pack very efficiently. The dumbbell-shaped disaccharides pack less efficiently and leave voids for water molecules. In the hydrate structures, the water molecules always donate two bonds, but may accept one or two. When they accept one, the coordination ranges from planar to pyramidal with no bimodal distribution. Water molecules which accept one bond form branch points from one to two finite or infinite chains. Those that accept two bonds form crossing points which link the chains into nets.

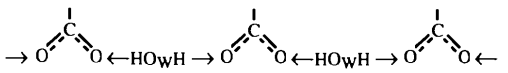
Cyclic systems of hydrogen bonds are common in the cyclodextrin hydrates, where they are classified as *homodromic, antidromic* and *heterodromic,* of which the homodromic is the more stable (Saenger, 1979; Chacko & Saenger, 1981). This concept is an extension of the sequential, double-donor and double-acceptor classification for water trimers (Hankins, Moskovitz & Stillinger, 1970). Cyclic systems are rare in the mono- and disaccharide crystal structures.

Proton disorder is observed in the cyclodextrin hydrates, but not in the oligosaccharide hydrates. It seems likely that this disorder is *via* a 'flip-flop' conformational mechanism involving reorientation of the hydroxyl groups, rather than the configurational bond-breaking mechanism of an enol–keto transformation (Saenger, 1982).

When the covalent O—H bond lengths are normalized to internuclear distances (*e.g.* 0·97 Å) (Jeffrey & Lewis, 1978; Allen, 1986), the O—H···O hydrogen bonds observed by X-ray diffraction in carbohydrates have a distribution of H···O bond lengths ranging from 1·72 to 2·09 Å, with a maximum at 1·83 Å. The O—H···O bond angles have a distribution with a maximum at ~160°. As the O—H···O angles get closer to 90°, the H···O distances become longer owing to non-bonded repulsions between neighbor atoms (Savage & Finney, 1986). Three-center bonds can be symmetrial with

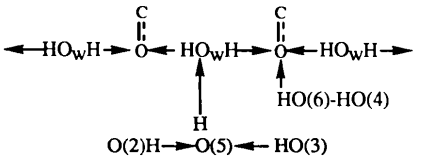
H···O ~2·0 Å and O—H···O ~120°, to unsymmetrical with H···O distances of 1·8 and 2·5 Å and angles of ~170 and ~110°.

The molecular packing in crystals tends to compress hydrogen bonds to bond lengths which are shorter than equilibrium gas-phase values (Lesyng, Jeffrey & Maluszynska, 1988). This effect is unpredictable, however. In the crystal structure of D-glucitol, the hydrogen bonds form two separate infinite chains (Park, Jeffrey & Hamilton, 1971). In one chain, the bonds are shorter than average [H···O = 1·691 (8), 1·723 (8), 1·729 (8) Å], while in the other chain they are longer than average [H···O = 1·914 (9), 2·018 (9), 2·218 (8) Å]. The two forms of potassium D-gluconate monohydrate studied by neutron diffraction (Panagiotopoulos, Jeffrey, LaPlaca & Hamilton, 1974) illustrate two types of hydrogen-bond cooperativity. In the *A* form, the gluconate ion has the straight-chain conformation stabilized by an intramolecular O(4)H···O(2) hydrogen bond. The hydrogen bonding is cooperative through the π-bond system of the carboxylate groups,*



In this *A* form, the cations are coordinated entirely by the gluconate hydroxyls and there is no cation–anion contact.

In the *B* form, the gluconate ion has the bent-chain conformation and the hydrogen bonding is characteristic of a non-polar carbohydrate with infinite and finite chains,



Only one of the carboxylate oxygens is involved in hydrogen bonding, the other is in the coordination shell of the cation. In this structure the cooperativity is through the infinite chains of σ bonds.

Carbohydrate liquid crystals and surfactants

It was noticed by Fischer & Helferich (1911) that long-chain *n*-alkyl glucosides had double melting points, but despite the pioneering work of Lehman (1908), the connection with liquid-crystal behavior was not made until many years later (Noller & Rockwell, 1938). More recent interest suggests that all *n*-alkyl substituted mono- or disaccharides with chain lengths of more than six carbon atoms form thermotropic liquid crystals. This appears to be irrespective of whether the carbohydrate moiety is pyranose, furanose or acyclic and irrespective of the nature of the linkage to the alkyl chain (Jeffrey, 1986).

Lyotropic liquid-crystal behavior is less common, since it requires a level of solubility in water greater than 50%. For example, the *n*-alkyl β-glucopyranosides are very soluble and display both thermotropic and lyotropic properties, whereas the α anomers are less soluble and display only thermotropic properties. Octyl β-glucopyranoside exhibits the classical lyotropic phase behavior associated with the soaps (Brown & Wolken, 1979), displaying successively laminar, cubic, hexagonal and micellar structures with progressive interpenetration of the crystals by water at room temperature (Chung & Jeffrey, 1989).

Octyl β-glucopyranoside is a *'famous'* non-ionic surfactant used in the solubilization and crystallization of membrane proteins (Baron & Thompson, 1975). This application prompted the synthesis and use of a number of *n*-alkylamido derivatives of D-glucitol and *n*-alkyl gluconamides (Goodby, Marcus, Chin, Finn & Pfannemüller, 1988). Crystal structures have been determined for a number of these compounds, see Table 6.

There are two types of molecular packing in these crystal structures. One is a bilayer structure in which the 'core' of the bilayer is a double row of hydrogen-bonded carbohydrate moieties. The alkyl chains extend laterally from this core and intercalate with those on the adjacent layers. These are referred to in Table 6 as *head-to-head bilayer structures*. The other mode of molecular packing is a monolayer with a single row of hydrogen-bonded carbohydrate moieties. The alkyl chains also extend laterally and intercalate with those on adjacent layers. These are referred to as *head-to-tail monolayer structures*. All these compounds exhibit crystal-to-crystal phase transitions prior to the transitions to the liquid crystal (melting points). These melting points tend to be independent of chain length, whereas the transitions to an isotropic liquid (clearing points) are dependent on alkyl chain length. This suggests that, contrary to expectation, the transition from the crystal lattice to the liquid-crystal molecular clusters involves melting of the hydrogen-bonded carbohydrate segments rather than the van der Waals-bonded alkyl chains.

Concluding remarks

There has been a noticeable decrease in the number of carbohydrate crystal structure analyses reported in recent years. Interest is focusing on the more complex molecules of *glycobiology*; the blood sugars, glycolipids, and the oligosaccharide components of bacterial and cell-wall polysaccharides. Crystallogra-

* A similar phenomenon in the enol form of β-diketone moieties is referred to as *resonance-assisted* hydrogen bonding (Gilli, Bellucci, Ferretti & Bertolasi, 1989).

phers have not yet developed methods for crystallizing these compounds and the principal experimental structural tools are the two-dimensional NOE NMR methods. An increasing number of crystal structures are reported for glycoproteins and for protein–carbohydrate complexes, but hitherto the resolution is rarely sufficient to provide conformational information.

The preparation of this article and the author's research reported therein has been supported by research grants from the Institute of General Medical Sciences, National Institutes of Health.

References

ALLEN, F. H. (1986). *Acta Cryst.* **B42**, 515–522.

ALLEN, F. H., KENNARD, O. & TAYLOR, R. (1983). *Acc. Chem. Res.* pp. 146–153.

ALTONA, C., KNOBLER, C. & ROMERS, C. (1963). *Acta Cryst.* **16**, 1217–1225.

ANGYAL, S. J. (1968). *Aust. J. Chem.* **21**, 2737–2746.

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

BARON, C. & THOMPSON, T. E. (1975). *Biochem. Biophys. Acta*, **382**, 276–285.

BEEVERS, C. A. & COCHRAN, W. (1947). *Proc. R. Soc. London Ser. A*, **190**, 257–272.

BERMAN, H., CHU, S. S. C. & JEFFREY, G. A. (1967). *Nature (London)*, **157**, 1576–1577.

BLACKWELL, J., GARDNER, K. H., KOLPAK, F. J., MINKE, R. & CLAFFEY, W. B. (1980). *Fibre Diffraction Methods*, ACS Symposium Series 141, edited by A. D. FRENCH & K. C. H. GARDNER, ch. 19. Washington, DC: American Chemical Society.

BROWN, G. H. & WOLKEN, J. J. (1979). *Liquid Crystals and Biological Structures*. New York: Academic Press.

BROWN, G. M. & LEVY, H. A. (1963). *Science*, **141**, 921–924.

BROWN, G. M. & LEVY, H. A. (1965). *Science*, **147**, 1038–1041.

CECCARELLI, C., JEFFREY, G. A. & TAYLOR, R. (1981). *J. Mol. Struct.* **70**, 255–271.

CECCARELLI, C., RUBLE, J. R. & JEFFREY, G. A. (1980). *Acta Cryst.* B**36**, 861–865.

CHACKO, K. K. & SAENGER, W. (1981). *J. Am. Chem. Soc.* **103**, 1708–1715.

CHUNG, Y. J. & JEFFREY, G. A. (1989). *Biochem. Biophys. Acta*. In the press.

COOK, W. J. & BUGG, C. E. (1977). *Metal–Ligand Interactions in Organic Chemistry and Biochemistry*, Vol. II, edited by B. PULLMAN & W. GOLDBLUM, pp. 231–256. Dordrecht: D. Reidel.

COX, E. G. & JEFFREY, G. A. (1939). *Nature (London)*, **143**, 894–895.

CREMER, D. & POPLE, J. A. (1975). *J. Am. Chem. Soc.* **97**, 1354–1358.

DOREN, H. VAN, VAN DER GIEST, R. & VAN BOLHUIS, F. (1989). *Carbohydr. Res.* In the press.

EDWARDS, J. T. (1955). *Chem. Ind. (London)*, pp. 1102–1104.

EINSPAHR, H. & BUGG, C. E. (1977). *Calcium Binding Proteins and Calcium Function*, edited by R. H. WASSERMAN, R. A. CORRANDO, E. CARAFOLI, R. H. KRETSINGER, D. H. MACLERNN & F. E. SIEGEL, pp. 13–20. New York: North-Holland.

EINSPAHR, H. & BUGG, C. E. (1980). *Acta Cryst.* B**36**, 264–271.

ENGLER, E. M., ANDOSE, J. D. & VON SCHLEYER, P. (1973). *J. Am. Chem. Soc.* **95**, 8005–8025.

ETTER, M. C. (1989). Editor. *NMR and X-ray Crystallography*, Trans. Am. Crystallogr. Assoc., Vol. 24. New York: American Crystallographic Association.

ETTER, M. C., HOGE, R. C. & VOJTA, G. M. (1988). *Crystallogr. Rev.* **1**, 281–338.

FISCHER, E. (1891). *Ber. Dtsch. Chem. Ges.* **24**, 1836–1845, 2683–2687.

FISCHER, E. & HELFERICH, B. (1911). *Justus Liebigs Ann. Chem.* **383**, 68–71.

GILLI, G., BELLUCCI, F., FERRETTI, V. & BERTOLASI, V. (1989). *J. Am. Chem. Soc.* **111**, 1023–1028.

GIRLING, R. L. & JEFFREY, G. A. (1973*a*). *Carbohydr. Res.* **27**, 257–260.

GIRLING, R. L. & JEFFREY, G. A. (1973*b*). *Acta Cryst.* B**29**, 1006–1011.

GIRLING, R. L. & JEFFREY, G. A. (1973*c*). *Acta Cryst.* B**29**, 1102–1111.

GOODBY, J. W., MARCUS, M. A., CHIN, E., FINN, P. L. & PFANNEMÜLLER, B. (1988). *Liq. Cryst.* **3**, 1569–1581.

GROTH, P. (1910). *Chemische Kristallographie*, Vol. III, pp. 380–381, 430–452. Leipzig: Wilhelm Engelmann.

HANKINS, D., MOSKOVITZ, J. W. & STILLINGER, F. (1970). *J. Chem. Phys.* **53**, 4544–4554.

HARRIS, R. (1988). *J. Chem. Soc. Faraday Trans.* 1, **84**, 3649–3655.

HASSEL, O. & OTTAR, B. (1947). *Acta Chem. Scand.* **1**, 929–942.

HAWORTH, W. N. (1929). *The Constitution of Sugars*. London: Arnold.

HINRICKS, W., BÜTTNER, G., STEIFA, M., BETZEL, CH., ZABEL, V., PFANNEMÜLLER, B. & SAENGER, W. (1987). *Science*, **238**, 205–208.

HOGE, R. & TROTTER, J. (1969). *J. Chem. Soc. A*, pp. 267–271.

HORTON, D. & WANDER, J. D. (1969). *Carbohydr. Res.* **10**, 279–288.

HUDSON, C. S. (1941). *J. Chem. Educ.* **18**, 353–357.

IMBERTY, A. & PEREZ, S. (1988). *Carbohydr. Res.* **181**, 41–55.

JEFFREY, G. A. (1978). *Hydrogen Bonding in Carbohydrates*, Proc. Symp. Org. Cryst. Chem., Institute of Chemistry, Adam Mickiewicz Univ., Poznan, Poland.

JEFFREY, G. A. (1986). *Acc. Chem. Res.* **19**, 168–173.

JEFFREY, G. A. (1987). *Patterson and Pattersons*, pp. 193–221. Oxford: Oxford Univ. Press/International Union of Crystallography.

JEFFREY, G. A. & KIM, H. S. (1970). *Carbohydr. Res.* **14**, 207–216.

JEFFREY, G. A. & KIM, H. S. (1971). *Acta Cryst.* B**27**, 1812–1817.

JEFFREY, G. A. & LEWIS, L. (1978). *Carbohydr. Res.* **60**, 179–182.

JEFFREY, G. A. & MALUSZYNSKA, H. (1989). *Acta Cryst.* B**45**, 447–452.

JEFFREY, G. A. & MITRA, J. (1983). *Acta Cryst.* B**39**, 469–480.

JEFFREY, G. A. & PARK, Y. J. (1979). *Carbohydr. Res.* **74**, 1–5.

JEFFREY, G. A., POPLE, J. A., BINKLEY, J. S. & VISHVESHWARA, S. (1978). *J. Am. Chem. Soc.* **100**, 373–379.

JEFFREY, G. A., POPLE, J. A. & RADOM, L. (1972). *Carbohydr. Res.* **25**, 117–131.

JEFFREY, G. A., POPLE, J. A. & RADOM, L. (1974). *Carbohydr. Res.* **38**, 81–95.

JEFFREY, G. A. & TAYLOR, R. (1980). *J. Comput. Chem.* **1**, 99–109.

JEFFREY, G. A. & YEON, Y. (1986). *Acta Cryst.* B**42**, 410–413.

JEFFREY, G. A., YEON, Y. & ABOLA, J. E. (1987). *Carbohydr. Res.* **169**, 1–11.

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

KANTERS, J. A., GAYKEMA, W. P. J. & ROELOFSEN, G. (1978). *Acta Cryst.* B**34**, 1873–1881.

KANTERS, J. A., ROELOFSEN, G. & ALBLAS, B. P. (1977). *Acta Cryst.* B**33**, 1906–1912.

KIRBY, A. J. (1983). In *Relativity and Structure Concepts in Organic Chemistry*, Vol. 15. Berlin, New York: Springer-Verlag.

KONINGSVELD, H. VAN, JANSEN, J. C. & STRAATHOF, A. J. J. (1988). *Acta Cryst.* C**44**, 1054–1057.

LEHMAN, O. (1908). *Flüssige Kristalle und die Theorien des Lebens Barth*. Liepzig.

LEMIEUX, R. U. (1959). *Abstr. Am. Chem. Soc. Meet.* **135**, 5E.

LEMIEUX, R. U. & CHU, P. (1958). *Abstr. Am. Chem. Soc. Meet.* **133**, 31N.
LEMIEUX, R. U., KOTO, S. & VORSIN, D. (1979). *Anomeric Effect. Origin and Consequences*, ACS Symposium Series 87, edited by W. A. SZARAK & D. HORTON. Washington, DC: American Chemical Society.
LEMIEUX, R. U. & PAVIA, A. A. (1969). *Can. J. Chem.* **47**, 4441–4446.
LESYNG, B., JEFFREY, G. A. & MALUSZYNSKA, H. (1988). *Acta Cryst.* B**44**, 193–198.
LIS, T. (1983). *Carbohydr. Res.* **122**, 23–29.
LLEWELLYN, F. J., COX, E. G. & GOODWIN, T. H. (1937). *J. Chem. Soc.* pp. 883–894.
LONGCHAMBON, F., GILLIER-PANDRAUD, R., WEIST, R., REES, B., MITSCHLER, A., FELD, R., LEHMAN, M. S. & BECKER, P. (1985). *Acta Cryst.* B**41**, 47–56.
NOLLER, C. R. & ROCKWELL, W. C. (1938). *J. Am. Chem. Soc.* **60**, 2076–2077.
PANAGIOTOPOULOS, N., JEFFREY, G. A., LAPLACA, S. & HAMILTON, W. C. (1974). *Acta Cryst.* B**30**, 1421–1430.
PARK, Y. J., JEFFREY, G. A. & HAMILTON, W. C. (1971). *Acta Cryst.* B**27**, 2393–2401.
PEREZ, S., PIERRE, ST. & MARCHESSAULT, R. (1978). *Can. J. Chem.* **56**, 2866–2871.
PFEFFER, P. E. (1984). *J. Carbohydr. Chem.* **3**, 613–639.
PICHON-PESME, V. & HANSEN, N. K. (1989). *J. Mol. Struct. (Theochem.)*, **183**, 151–160.
PIGMAN, W. (1957). Editor. *The Carbohydrates: Chemistry, Biochemistry and Physiology*. New York: Academic Press.
PIGMAN, W. & HORTON, D. (1970). Editors. *The Carbohydrates: Chemistry and Biochemistry*. New York: Academic Press.
PIMENTEL, G. C. & MCCLELLAND, A. L. (1960). *The Hydrogen Bond*. San Francisco: W. H. Freeman.
Rodd's Chemistry of Carbon Compounds (1967). 2nd ed., edited by S. COFFEY. Supplement FG (1983), edited by M. F. ANSELL. Amsterdam: Elsevier.
RUBLE, J. & JEFFREY, G. A. (1974). *Carbohydr. Res.* **38**, 61–69.
SAENGER, W. (1979). *Nature (London)*, **279**, 343–344.
SAENGER, W. (1982). *Nature (London)*, **296**, 581–583.
SAENGER, W. (1984). *Inclusion Compounds*, edited by J. L. ATWOOD, J. E. D. DAVIES & D. D. MACNICOL, pp. 231–259. New York: Academic Press.
SASTRY, D. L., TAKEGOSHI, K. & MCDOWELL, C. A. (1987). *Carbohydr. Res.* **165**, 161–171.
SAVAGE, H. & FINNEY, J. (1986). *Nature (London)*, **322**, 717–720.
SENMA, M., TAIRA, Z., OSAKI, K. & TAGA, T. (1973). *J. Chem. Soc. Chem. Commun.* p. 880.
STODDARD, J. F. (1971). *Stereochemistry of Carbohydrates*. New York: Wiley.
SUNDARALINGAM, M. (1969). *Biopolymers*, **7**, 821–830.
TVAROSKA, I. & PEREZ, S. (1986). *Carbohydr. Res.* **149**, 389–410.
WOLFE, S. (1972). *Acc. Chem. Res.* **5**, 102–111.
WOOD, R. A., JEFFREY, G. A., PFEFFER, P. E. & HICKS, K. B. (1983). *J. Am. Chem. Soc.* **105**, 2128–2133.

Acta Cryst. (1990). B**46**, 103–111

Electron Microscopy Study of the Structure of Metastable Oxides Formed in the Initial Stage of Copper Oxidation. V. $Cu_4O_{0.75}$*

BY R. GUAN

Laboratory of Atomic Imaging of Solids, Institute of Metal Research, Academia Sinica, 110015 *Shenyang, People's Republic of China*

H. HASHIMOTO

Faculty of Engineering, Okayama University of Science, Okayama 700, *Japan*

AND K. H. KUO

Beijing Laboratory of Electron Microscopy, Academia Sinica, PO Box 2724, *Beijing, People's Republic of China*

(*Received* 22 *June* 1988; *accepted* 18 *September* 1989)

Abstract

Besides Cu_4O, Cu_4O-S_1 and Cu_4O-S_2 reported previously, three allotropes of a new metastable oxide of copper, α-$Cu_4O_{0.75}$, β-$Cu_4O_{0.75}$ and γ-$Cu_4O_{0.75}$, have been observed by high-resolution electron microscopy in an early stage of the oxidation of copper. The atomic positions of Cu and O have been determined by comparing the observed structure images with simulated ones calculated using the theories of electron diffraction and image formation. These new suboxides have oxygen-deficient Cu_4O structures and the O atoms occupy not only the central positions in the tetrahedra consisting of four Cu atoms, but also the central positions in the octahedra consisting of six Cu atoms. The differences between α-, β- and γ-$Cu_4O_{0.75}$ are due to the distribution of O atoms.

* Project partly supported by the National Natural Science Foundation of the People's Republic of China, under grant 888600.

0108-7681/90/020103-09$03.00 © 1990 International Union of Crystallography