

This research was supported by 'Centre National de la Recherche Scientifique'. The authors wish to thank J. Mainville for reading the manuscript.

References

- ANDREAZZA, P., JOSSE, D., LEFAUCHEUX, F., ROBERT, M. C. & ZYSS, J. (1992). *Phys. Rev. B*, **45**, 7640–7649.
- BAERT, F., SCHWEISS, P., HEGER, G. & MORE, M. (1988). *J. Mol. Struct.* **178**, 29–48.
- BARZOUKAS, M., JOSSE, D., FREMEAUX, P. & ZYSS, J. (1987). *Non-linear Optical and Electroactive Polymers*, edited by P. N. PRASAD & D. R. ULRICH, pp. 69–103. New York, London: Plenum Press.
- BARZOUKAS, M., JOSSE, D., FREMEAUX, P., ZYSS, J., NICOD, J. F. & MORLEY, J. O. (1987). *J. Opt. Soc. Am. B*, **4**(6), 977–986.
- BECKER, P. J. & COPPENS, P. (1974). *Acta Cryst.* **A30**, 129–147.
- BLESSING, R. H. (1989). *J. Appl. Cryst.* **22**, 396–397.
- CLEMENTI, E. (1965). Tables of atomic functions, supplement to IBM. *J. Res. Rev.* **9**, 2.
- COPPENS, P. (1975). *Phys. Rev. Lett.* **34**, 98–100.
- COPPENS, P., GURU ROW, T. N., LEUNG, P., STEVENS, E. D., BECKER, P. J. & YANG, Y. W. (1979). *Acta Cryst.* **A35**, 63–72.
- COPPENS, P., LEISEROWITZ, L. & RABINOVICH, D. (1965). *Acta Cryst.* **18**, 1035–1038.
- DOCHERTY, V. J., PUGH, D. & MORLEY, J. O. (1985). *J. Chem. Soc. Faraday Trans. 2*, **81**, 1179–1192.
- DUNITZ, J. D., MAVERICK, E. F. & TRUEBLOOD, K. N. (1988). *Angew. Chem. Int. Ed. Engl.* **27**, 880–895.
- DUNITZ, J. D., SCHOMAKER, V. & TRUEBLOOD, K. N. (1988). *J. Phys. Chem.* **92**, 856–867.
- HANSEN, N. & COPPENS, P. (1978). *Acta Cryst.* **A34**, 909–921.
- HIRSHFELD, F. L. (1976). *Acta Cryst.* **A32**, 239–244.
- HIRSHFELD, F. L. (1977). *Theor. Chim. Acta*, **44**, 129–138.
- JOHNSON, C. K. (1965). *ORTEP*. Report ORNL-3794. Oak Ridge National Laboratory, Tennessee, USA.
- KANIS, D. R., RATNER, M. A. & MARKS, T. J. (1992). *J. Am. Chem. Soc.* **114**, 10338–10357.
- LALAMA, S. J. & GARITO, A. F. (1979). *Phys. Rev. A*, **20**, 1179–1194.
- LEDOUX, I. & ZYSS, J. (1982). *J. Chem. Phys.* **73**, 203–213.
- MEYERS, F., BREDAS, J. L. & ZYSS, J. (1992). *J. Am. Chem. Soc.* **114**, 2914–2921.
- OUДАР, J. L. (1977). *J. Chem. Phys.* **67**, 446–457.
- PERIGAUD, A., GONZALES, F. & CUNISSE, M. (1991). *Ann. Chim. Fr.* pp. 133–141.
- ROBINSON, F. N. H. (1967). *Bell. Syst. Tech. J.* pp. 913–956.
- ROSENFELD, R. E., TRUEBLOOD, K. N. & DUNITZ, J. D. (1978). *Acta Cryst.* **A34**, 828–829.
- SCHOMAKER, V. & TRUEBLOOD, K. N. (1968). *Acta Cryst.* **B24**, 63–76.
- SOUHASSOU, M., LECOMTE, C., BLESSING, R. H., AUBRY, A., ROHMER, M. M., WIEST, R., BENARD, M. & MARRAUD, M. (1991). *Acta Cryst.* **B47**, 253–266.
- STEWART, R. F. (1973). *J. Chem. Phys.* **58**, 1668–1676.
- STEWART, R. F., DAVIDSON, E. R. & SIMPSON, W. T. (1965). *J. Chem. Phys.* **43**, 175–187.
- TEHRUNE, R. W., MAKER, P. D. & SAVAGE, C. M. (1965). *Phys. Rev. Lett.* **14**, 681–684.
- TRUEBLOOD, K. N. (1990). *THMA11*. Univ. of California, Los Angeles, USA.
- ZYSS, J. (1979a). *J. Chem. Phys.* **70**, 3333–3340.
- ZYSS, J. (1979b). *J. Chem. Phys.* **70**, 3341–3349.
- ZYSS, J. (1979c). *J. Chem. Phys.* **71**, 909–916.
- ZYSS, J., CHAU VAN, T., DHENAUT, C. & LEDOUX, J. (1993). *Chem. Phys.* pp. 281–296.
- ZYSS, J. & LEDOUX, I. (1994). *Chem. Rev.* **94**, 77–105.
- ZYSS, J., NICOD, J. F. & COQUILLAY, M. (1984). *J. Chem. Phys.* **81**, 4160–4167.
- ZYSS, J. & OUDAR, J. L. (1982). *Phys. Rev. A*, **26**, 2028–2048.

Acta Cryst. (1995). **B51**, 209–220

An Extension of the *GROMOS* Force Field for Carbohydrates, Resulting in Improvement of the Crystal Structure Determination of α -D-Galactose

BY M. L. C. E. KOUWIJZER,* B. P. VAN EIJK, H. KOOLJMAN AND J. KROON

Department of Crystal and Structural Chemistry, Bijvoet Center for Biomolecular Research, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands

(Received 25 May 1994; accepted 3 October 1994)

Abstract

For carbohydrates the *GROMOS* force field has been extended to a more realistic all-atom model, with the use of parameters from the force field proposed by Ha, Giammona, Field & Brady [*Carbohydr. Res.* (1988), **180**, 207–221]. This extended and modified *GROMOS* force field has been used to simulate the crystal structures of seven monosaccharides. The results, compared with the experimental data, are satisfactory, and an overall improvement over those obtained with the comparable

Ha force field. The experimentally determined positions of the H atoms should be used with caution in this comparison, and the thermal parameters can only be used as indicators for the preservation of the symmetry during the simulation. The simulations gave rise to suspicion about the hydroxyl H-atom positions in two of the sugars. These two structures were redetermined by X-ray diffraction at low temperature. For β -D-glucose essentially the same structure was found as in the original publication. In the case of α -D-galactose, the new structure confirmed the results from the simulation, in contrast to earlier experimental determinations.

* Author to whom all correspondence should be addressed.

Introduction

When structural information about a molecular system is needed that is hard to obtain by experiments or quantum mechanical calculations, computer simulations can be used. Molecular dynamics simulations are widely applied for this purpose. The reliability of the information, however, depends strongly on the quality of the force field used. Force fields are often optimized for a special purpose (Grootenhuis & Haasnoot, 1993) or for a certain class of compounds, and therefore not always widely applicable. Our field of interest is mainly carbohydrate structures. In a previous study (Kouwijzer, van Eijck, Kroes & Kroon, 1993) we compared the *GROMOS* force field [with 'united atoms'; van Gunsteren (1987)] with an all-atom force field reported by Ha, Giammona, Field & Brady (1988), which was specifically optimized for α -D-glucose. Surprisingly, the crystal structure of this compound was reproduced slightly better by the *GROMOS* force field.

The *GROMOS* force field is designed for proteins, nucleotides or sugars in aqueous or apolar solvents. It uses the concept of united atoms: in the case of sugar molecules this means that the CH and CH₂ groups of atoms are replaced by one atom type. With the powerful computers now available, this time saving simplification is not so essential any more; a more realistic, full atom model can be used without paying a high price in computing time.

Our aim is to extend the *GROMOS* force field to an all-atom model, especially applicable to sugars. To do this, the missing parts of the *GROMOS* force field were taken from the Ha force field, and the new model was improved on the crystal structure of α -D-glucose (later, the simulation of β -D-galactose caused us to make another adjustment). Next, the new force field was tested on all seven monosaccharide crystal structures that have been determined experimentally by means of neutron or X-ray diffraction. For four of the sugars, simulations with the original *GROMOS* and the Ha force field were performed too, to investigate the quality with respect to one another.

The results of the simulations gave rise to suspicion about the positions of hydroxyl H atoms in β -D-glucose and α -D-galactose. Therefore, these compounds were crystallized and redetermined with X-ray diffraction at low temperature.

Computational methods

Extending the GROMOS parameter set

The potential energy V of a system of N atoms is calculated from the bond lengths b , bond angles θ , dihedrals φ and interatomic distances r , according to the *GROMOS* function

$$\begin{aligned}
 V = & \sum_{\text{bonds}} \frac{1}{2} K_b (b - b_0)^2 + \sum_{\text{angles}} \frac{1}{2} K_\theta (\theta - \theta_0)^2 \\
 & + \sum_{\text{dihedrals}} \sum_n K_\varphi [1 + \cos(n\varphi - \delta_n)] \\
 & + \sum_i^N \sum_{j>i}^{N'} [(q_i q_j / 4\pi\epsilon_0 r_{ij}) \\
 & + \{[C_{12}(i, j)] / r_{ij}^{12}\} - [C_6(i, j)] / r_{ij}^6].
 \end{aligned} \tag{1}$$

In this function, the parameters K_b , b_0 , K_θ and θ_0 denote the force constants for the bond lengths, the equilibrium bond lengths, the force constants for the bond angles, and equilibrium bond angles, respectively. For every dihedral there is at least one force constant K_φ , a multiplicity n and a phase shift δ . In the last term, only nonbonded interactions are taken into account (in other words, the prime in the second summation means that atoms j that are connected to i via one or two bonds are excluded). The electrostatic energy is calculated from the atomic partial charges q , and the van der Waals energy is approximated by the Lennard–Jones energy with the interaction parameters C_{12} and C_6 .

To extend the *GROMOS* parameter set, the parameters for the CH and CH₂ united atoms had to be replaced by parameters for separate C and H atoms. As a start, the Lennard–Jones parameters for both atom types were taken from the Ha force field (Ha, Giammona, Field & Brady, 1988), and so were the parameters for the bond lengths and angles that contain an aliphatic H atom. The remaining Lennard–Jones and dihedral parameters were taken from the original *GROMOS* force field.

For testing the new force field, molecular dynamics were used rather than energy minimization, to avoid optimizing to a local minimum in the structures and to study the time-dependent behavior of the systems.

A 25 ps simulation of the crystal structure of α -D-glucose (see Fig. 1; the simulation conditions are given in the next subsection) with this new force field showed already surprisingly good results. During this simulation, the deviations of the atomic coordinates with respect to the crystal structure determination were smaller than during the Ha simulation, but not as good as during the simulation with the original *GROMOS* force field. The quality of the temperature coefficients was about as good as in the Ha simulation, just slightly worse than the *GROMOS* simulation. The volume of the crystal cell, however, decreased about 5%, much more than in the simulations with the Ha force field (1.3%) or the original *GROMOS* force field (1.7%). To improve on this, several test simulations were carried out, varying the charges on the O atoms (compensated, for reasons of electroneutrality, with the C atoms), the Lennard–Jones parameters for hydroxyl H atoms and for the interaction between O atoms, and the force constants for the dihedral X–C–O–H. Eventually, improvements were made by using increased Lennard–Jones parameters for the interactions between O atoms. Later on, we added a

Table 1. Atomic partial charges (for atom names, see Fig. 1)

	Present work	Original GROMOS	Ha
C1	0.25	0.31	0.35
C2-C4	0.05	0.15	0.15
C5	0.06	0.20	0.10
C6	-0.05	0.15	0.05
O1-O4, O6	-0.55	-0.55	-0.65
O5	-0.36	-0.36	-0.40
H1-H62	0.10		0.10
HO1-HO6	0.40	0.40	0.40

Table 2. Force-field parameters for bond lengths [see (1)]

	K_b (kcal mol ⁻¹ Å ⁻²)	b_0 (Å)
O-H	750.0	1.000
C-H	674.6	1.099
C-C	600.0	1.520
C-O	600.0	1.430
C-OS	600.0	1.435

Table 3. Force-field parameters for bond angles [see (1)]

	K_θ (kcal mol ⁻¹ rad ⁻²)	θ_0 (°)
C-O-H	95.0	109.50
H-C-H	67.2	107.85
H-C-C	86.0	108.72
H-C-O	91.8	109.89
H-C-OS	90.4	107.24
C-C-C	60.0	109.50
C-C-O	68.0	109.50
C-C-OS	68.0	109.50
C-OS-C	80.0	109.50
OS-C-O	68.0	109.50

Table 4. Force-field parameters for dihedrals [see (1)]

For each bond one torsional term is used; the last four terms, however, are defined for each dihedral

	K_ϕ (kcal mol ⁻¹)	n	δ (°)
X-C-C-Y	1.40	3	0
X-C-O-Y	0.30	3	0
X-C-OS-Y	0.90	3	0
O-C-C-O	0.50	2	0
O-C-C-OS	0.50	2	0
C-C-C-O	0.10	2	0
C-C-C-OS	0.10	2	0

Table 5. Lennard-Jones parameters

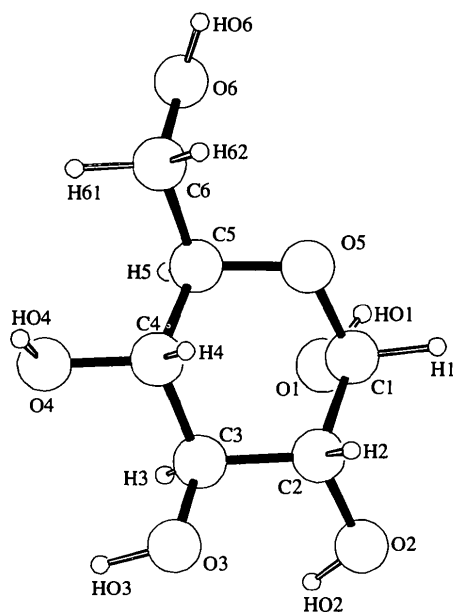
	C_6 (Å ⁶ kcal mol ⁻¹)	C_{12} (Å ¹² kcal mol ⁻¹)
C...C	393	4.28×10^5
C...O or OS	370	2.86×10^5
C...HO	41.4	6.40×10^3
C...H	49.1	2.99×10^4
O or OS...O	700	5.00×10^5
O or OS...HO	0	0
OS(-C-O-)HO	0	1.00×10^5
O or OS...H	44.6	1.86×10^4
OS...OS	700	1.80×10^5
HO...HO	0	0
HO...H	4.07	2.77×10^2
H...H	5.76	1.85×10^3

repulsion term between HO1 and O5, because of a collision between these atoms in the simulation of β -D-galactose.

All the resulting force-field parameters are given in Tables 1-5, where the ring O atom is referred to as OS and the hydroxyl H atoms as HO. Except for the repulsion mentioned in the previous paragraph, no special Lennard-Jones parameters for third neighbors are introduced; in the GROMOS force field this was necessary due to the use of united atoms.

Testing the new parameter set

To test this set of parameters, several molecular dynamics simulations were performed. All anhydrous monosaccharides with known crystal structure were studied: both the α - and β -form of D-glucose and D-galactose, β -D-allose, α -D-mannose and α -D-talose. The crystal structure of α -D-glucose has been determined by neutron diffraction (Brown & Levy, 1979), and the others by X-ray diffraction (Chu & Jeffrey, 1968; Ohanessian & Gillier-Pandraud, 1976; Sheldrick, 1976; Longchambon, Ohanessian, Avenel & Neuman, 1975; Kroon-Batenburg, van der Sluis & Kanters, 1984; Longchambon, Avenel & Neuman, 1976; Hansen & Hordvik, 1977; Ohanessian, Avenel & Kanters, 1977). They all crystallize in the orthorhombic space group $P2_12_12_1$, with four molecules per unit cell, except for α -D-mannose which has eight molecules per unit cell. The coordinates were retrieved from the Cambridge Structural Database (CSD; Allen, Davies, Galloy,

Fig. 1. Atomic numbering of α -D-glucose.

Johnson, Kennard, Macrae, Mitchell, Mitchell, Smith & Watson, 1991).

In two of the structure determinations one or more H atoms were removed by the CSD editors because of suspected coordinate errors. In β -D-galactose (Sheldrick, 1976), however, we included the original coordinates of HO3 (see Fig. 1), as they did not give rise to unusual bond lengths or angles. In α -D-mannose (Longchambon, Avenel & Neuman, 1976), we found that all the fractional y-coordinates of the H atoms of molecule (II) are wrong (not only some of them as the CSD editors expected); it seems that the minus signs from these coordinates have been omitted in their article. We included them and found no unusual bond lengths or angles anymore.

Simulation boxes were constructed for each of these compounds; they contained 9–24 unit cells to have sides of at least 20 Å. Inside the simulation boxes no symmetry was imposed, so each box contained 64–96 independent molecules. The cut-off radius was 9.5 Å. The systems were coupled to a pressure bath of 1 atm with a coupling constant of 3 ps and to a temperature bath with a coupling constant of 0.1 ps (Berendsen, Postma, van Gunsteren, DiNola & Haak, 1984). Bond lengths were constrained with the *SHAKE* algorithm (Ryckaert, Ciccotti & Berendsen, 1977). The time step for the simulations was 2 fs, and the energies and atomic coordinates were saved every 10 time steps. All simulations were started from the crystal structure and the systems were simulated for 60 ps. The first 10 ps were considered equilibration time and, therefore, not used in the analysis.

In our previous study (Kouwijzer, van Eijck, Kroes & Kroon, 1993), the temperature came out systematically too low. We have now traced down the origin of this effect to an incorrect implementation of the leap-frog scheme: between successive steps the velocity itself of each particle is averaged rather than its square (van Eijck, 1994). The temperatures are now better, the mean temperature during the simulations ranging from 297.8 to 298.7 K with r.m.s. deviations of 0.1 K.

After the simulations, the mean unit-cell dimensions from the simulations can be compared with the experimental values. For a detailed analysis of the atomic coordinates, all the symmetry operations (translations and screw axes) were applied to give one molecule as an average in space and time. The mean fractional atomic coordinates can be compared with those known from the diffraction experiment, which are also averages in space (the crystal) and time (needed for the measurements). The r.m.s. deviations from the mean positions during the simulations can be compared with the temperature coefficients from the experiments. This is a very sensitive test: if the symmetry has been lost during the simulation, the r.m.s. deviations of the mean atomic positions will be much larger than those for a well behaved crystal. Alternatively, the comparison can be based on the bond angles and dihedrals and on hydrogen-

bonding schemes (which were calculated from the average atomic positions). The aliphatic H atoms are not taken into account in the comparison between simulation and experiment, in order not to obscure the comparison between the force fields and because of the uncertainty in the experimental information on these atoms.

For three of the monosaccharides, two independent experimental structure determinations are available. The differences in the reported structures of β -D-galactose and α -D-talose are small enough to restrict the comparison to one of them (Sheldrick, 1976; Hansen & Hordvik, 1977, respectively), but in the case of α -D-galactose (Ohanessian & Gillier-Pandraud, 1976; Sheldrick, 1976) the hydrogen-bonding schemes are quite different (Jeffrey & Shiono, 1977). In the present paper we report a redetermination of this structure, and comparisons between simulation and experiment will refer to these new data.

Four of the seven sugars were also simulated with the original *GROMOS* and the original Ha force field. These simulations are denoted with *G* and *H*, respectively. The simulation conditions were almost the same as mentioned before; the atomic partial charges in these simulations are also given in Table 1. The simulation of β -D-galactose with the original *GROMOS* force field stopped almost immediately after the equilibration (altogether after 10.26 ps), because of a collision between O5 and HO1. In the next section the results from the period 5–10 ps are reported.

Results

Unit cells, atomic positions and temperature coefficients

All unit cells decreased in size. The percentages for every compound, relative to each of the axes and to the experimental volume, are given in Table 6.

As mentioned before, the differences between the atomic coordinates from simulation and experiment were examined. To eliminate effects of the cell dimensions, the comparison was based on fractional coordinates and then converted to Å. The differences were averaged and these are given in Table 7, also for the simulations with the original *GROMOS* and Ha force fields for four of the structures. Because of the uncertainty in the positions of the H atoms in the X-ray diffraction, these values are given both for all the atoms and for the non-H atoms only.

To compare the r.m.s. deviations from the average coordinates in the simulations with the temperature coefficients from the diffraction experiments, first a scaling factor was calculated on the basis of isotropic coefficients of the C and O atoms only

$$S = \left(\sum_i U_{i,iso}^{md} \right) / \left(\sum_i U_{i,iso}^{exp} \right). \quad (2)$$

Table 6. Mean unit-cell parameters during simulations compared with the experimental values

For the simulations labeled *G* the original GROMOS force field was used, *H* refers to the original Ha force field.

	Δa (%)	Δb (%)	Δc (%)	ΔV (%)
α -D-Glucose	0.1	-3.0	1.5	-1.4
(<i>G</i>)	-0.9	0.3	-1.1	-1.7
(<i>H</i>)	-0.1	-4.8	3.8	-1.3
β -D-Glucose	-10.9	-3.3	10.5	-4.8
α -D-Galactose	-0.7	3.7	-4.9	-2.1
(<i>G</i>)	1.6	-1.7	-0.8	-0.9
(<i>H</i>)	-0.2	2.9	-2.9	-0.2
β -D-Galactose	-6.6	2.3	1.7	-2.8
(<i>G</i>)	3.2	-2.4	-1.6	-0.9
(<i>H</i>)	-12.7	5.1	1.4	-7.0
β -D-Allose	2.3	-1.9	-2.8	-2.5
(<i>G</i>)	0.3	0.6	-2.6	-1.8
(<i>H</i>)	1.6	-1.8	-0.7	-0.9
α -D-Mannose	1.5	-2.1	-3.5	-4.2
α -D-Talose	1.3	-3.8	-2.0	-4.5

Table 7. Mean differences in atomic coordinates (corrected for unit-cell dimensions) for all atoms, only the non-H atoms ($\Delta r = \langle |r_i^{\text{md}} - r_i^{\text{exp}}| \rangle$; between parentheses $|r_i^{\text{md}} - r_i^{\text{exp}}|$ is given, which should be zero when the differences are distributed randomly); and scaling and disagreement factors for the temperature coefficients for the C and O atoms [*S* and *R_T*, respectively, see (2) and (3)] and scaling factor for the H atoms, if the experimental value is known

	$\Delta r_{\text{C,O,H}}$ (Å)	$\Delta r_{\text{C,O}}$ (Å)	<i>S</i>	<i>R_T</i>	<i>S_H</i>
α -D-Glucose	0.23 (0.10)	0.22 (0.12)	1.25	0.11	1.08
(<i>G</i>)	0.11 (0.05)	0.11 (0.08)	0.94	0.08	0.97
(<i>H</i>)	0.34 (0.09)	0.33 (0.10)	1.28	0.17	1.31
β -D-Glucose	0.58 (0.34)	0.52 (0.35)	1.40	0.28	2.32
α -D-Galactose	0.22 (0.15)	0.18 (0.13)	1.24	0.17	1.58
(<i>G</i>)	0.17 (0.09)	0.13 (0.08)	1.58	0.27	1.80
(<i>H</i>)	0.23 (0.15)	0.21 (0.18)	1.74	0.20	5.16
β -D-Galactose	0.12 (0.04)	0.08 (0.04)	1.85	0.17	7.72
(<i>G</i>)	0.33 (0.28)	0.33 (0.31)	0.91	0.15	4.99
(<i>H</i>)	0.28 (0.14)	0.21 (0.13)	1.39	0.29	11.28
β -D-Allose	0.14 (0.08)	0.12 (0.08)	1.13	0.11	0.67
(<i>G</i>)	0.21 (0.16)	0.21 (0.17)	1.14	0.26	0.82
(<i>H</i>)	0.42 (0.37)	0.42 (0.40)	0.97	0.14	0.84
α -D-Mannose	0.45 (0.40)	0.44 (0.41)	1.03	0.22	
α -D-Talose	0.18 (0.09)	0.13 (0.10)	1.00	0.29	1.34

With this scaling factor, a disagreement factor for the three main anisotropic coefficients was calculated

$$R_T = \left(\frac{\sum_i \sum_{j=11,22,33} |U_{ij}^{\text{md}} - S U_{ij}^{\text{exp}}|}{\sum_i \sum_{j=11,22,33} |U_{ij}^{\text{md}}|} \right) \quad (3)$$

These factors are also given in Table 7. For most structures isotropic temperature factors for the H atoms were also reported, although the reliability of such data can be questionable. Isotropic scaling factors *S_H*

Table 8. Mean differences in bond angles and dihedrals ($\Delta\theta = \langle |\theta_i^{\text{md}} - \theta_i^{\text{exp}}| \rangle$ and $\Delta\varphi = \langle |\varphi_i^{\text{md}} - \varphi_i^{\text{exp}}| \rangle$), where those containing a H atom are taken apart; between parentheses $|\theta_i^{\text{md}} - \theta_i^{\text{exp}}|$ and $|\varphi_i^{\text{md}} - \varphi_i^{\text{exp}}|$ are given

	$(\Delta\theta)_{\text{non-H}}$ (°)	$(\Delta\theta)_{\text{H}}$ (°)	$(\Delta\varphi)_{\text{non-H}}$ (°)	$(\Delta\varphi)_{\text{H}}$ (°)
α -D-Glucose	1.7 (0.6)	1.5 (0.3)	2.6 (1.1)	5.3 (2.4)
(<i>G</i>)	2.4 (0.8)	1.4 (0.7)	2.4 (0.3)	4.2 (2.7)
(<i>H</i>)	1.4 (0.7)	1.9 (0.3)	2.0 (0.1)	9.8 (9.5)
β -D-Glucose	2.2 (0.9)	8.7 (1.4)	5.5 (0.1)	33.6 (20.9)
α -D-Galactose	2.0 (0.9)	0.8 (0.9)	2.4 (0.5)	12.1 (10.3)
(<i>G</i>)	2.5 (1.2)	0.9 (0.1)	4.2 (0.6)	10.5 (9.1)
(<i>H</i>)	2.1 (0.9)	10.1 (9.8)	2.5 (0.2)	14.8 (14.8)
β -D-Galactose	1.7 (1.3)	1.8 (0.1)	2.8 (0.1)	3.8 (1.0)
(<i>G</i>)	2.8 (1.2)	1.7 (1.2)	3.4 (1.1)	9.7 (4.6)
(<i>H</i>)	1.7 (1.4)	1.7 (0.5)	3.1 (0.2)	13.3 (0.6)
β -D-Allose	1.7 (0.7)	2.8 (2.8)	3.7 (1.2)	7.2 (1.5)
(<i>G</i>)	2.7 (0.8)	2.0 (1.7)	3.1 (0.6)	6.9 (6.9)
(<i>H</i>)	1.5 (0.7)	1.7 (1.1)	3.7 (1.2)	12.6 (8.8)
α -D-Mannose	2.0 (1.3)	3.6 (0.4)	5.0 (1.2)	13.0 (1.5)
α -D-Talose	1.7 (0.9)	2.4 (1.2)	3.6 (0.4)	4.1 (3.0)

[calculated as in (2), but now for H atoms only] are also given in Table 7.

Internal geometry and hydrogen bonds

To obtain an impression of the internal molecular geometry, the bond angles and dihedrals were calculated from the mean atomic coordinates.* Again, a comparison was made between simulations and experiments. The mean deviations are given in Table 8. To assess the hydrogen-bonding scheme from the mean atomic coordinates and the space-group symmetry, the program PLATON (Spek, 1990a) was used. We took as criteria to define a hydrogen bond that the distance between the two O atoms involved should be shorter than 3.54 Å, the distance between the H atom and the accepting O atom should be shorter than 2.60 Å, and the angle O—H··O should be larger than 110°.

Not only can the mean dihedral values from the simulations be compared with those from the experiments, but from the simulation also a distribution function can be made. This was done for the dihedrals which determine the conformation of the CH₂OH group, and for five dihedrals, each containing one of the five hydroxyl groups. From these functions it was seen that in α -D-glucose, β -D-galactose, β -D-allose and α -D-talose no dihedral transitions had occurred. When the hydrogen-bonding schemes are compared, the differences are marginal. Only in α -D-talose does one of the hydrogen bonds become bifurcated.

In β -D-glucose, most hydroxymethyl groups occupied the *gg* conformation (O6 *gauche* with respect to both O5

* Lists of structure factors, anisotropic thermal parameters, H-atom coordinates and complete geometry have been deposited with the IUCr (Reference: SE0154). Copies may be obtained through The Managing Editor, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England.

Table 9. *Hydrogen-bonding schemes in α -D-galactose from the two experimental structures and the simulations*

$r_{O...O}$ (Å) and $\theta_{O-H...O}$ (°) are given in parentheses. The abbreviation OGP stands for Ohanessian & Gillier-Pandraud (1976); BS for Sheldrick (1976); KEKK for our redetermination.

Acceptor according to	O1	O2	Donor O3	O4	O6
X-ray					
OGP	O6 (2.65, 114)	O3 (2.70, 168)	O4 (2.91, 147)	O3 (2.91, 106)	O2 (2.77, 161)
BS	O3 (3.11, 180)	O6 (2.75, 180)	O4 (2.92, 180)	O2 (2.89, 180)	O1 (2.66, 180)
KEKK	O6 (2.65, 166)	O3 (2.69, 166)	O4 (2.88, 168)	O2 (2.83, 167)	O2 (2.73, 151)
MD					
New force field	O6 (2.65, 169)	O3 (2.71, 171)	O4 (2.77, 160)	O2 (2.82, 160)	O2 (2.77, 154)
GROMOS force field	O6 (2.67, 161)	O3 (2.64, 175)	O4 (2.75, 164)	O2 (2.90, 160)	O2 (2.71, 152)
Ha force field	O6 (2.74, 160)	O3 (2.80, 156)	O4 (2.87, 152)	O2 (2.85, 161)	O2 (2.81, 146)

and C4, respectively), but 0.9% *tg* was found. This was due to one molecule that had a transition after 21 ps of simulation. In the distribution functions of the dihedrals with a hydroxyl group, no second maxima were found. Nevertheless, a difference was found in the hydrogen-bonding schemes, due to changes in the positions of the HO4 and HO3 atoms. In the experimental structure, a long O4—HO4...O2 ($r_{O...O} = 3.27$ Å, $\theta_{O-H...O} = 161^\circ$) bond is found. In the simulated structure, this bond is replaced by O4—HO4...O1 ($r_{O...O} = 3.03$ Å, $\theta_{O-H...O} = 106^\circ$, intramolecular) is also present. The in the hydrogen bond is used, it becomes bifurcated, since O4—HO4...O3 ($r_{O...O} = 2.88$ Å, $\theta_{O-H...O} = 106^\circ$) (intramolecular) is also present. The bond formed by O3—HO3 changes slightly; it becomes bifurcated.

Every distribution function of the dihedrals of α -D-galactose, except for O5—C1—O1—HO1, showed a second maximum. 1.6% *gt* was found instead of *tg*, because six molecules had changed their conformation temporarily (3–4 ps, once 12 ps). From the dihedrals containing a hydroxyl group, 1–2% was found in another conformation. When the hydrogen-bonding scheme is compared with those of the experimental structures (see Table 9), it is seen that the simulation shows a combination of both previously determined experimental structures.

In one of the two independent molecules of α -D-mannose (molecule II), no disorder in the dihedrals was found. In the other molecule, however, one quarter of the molecules (nine of the total of 72 simulated molecules) showed a transition in the C5—C6—O6—HO6 dihedral; at the same time the hydroxymethyl group occupied a position in between *gt* and *tg*. In one of the nine molecules involved in these changes, the dihedrals return to the old position after *ca* 25 ps. All the distribution functions of the other dihedrals did not show a second

maximum. The second 'conformation' of the aforementioned dihedrals appeared to be stabilized by a change in the hydrogen-bonding scheme; instead of the bond O6—HO6...O1 ($r_{O...O} = 2.73$ Å, $\theta_{O-H...O} = 162^\circ$), O6—HO6...O6 ($r_{O...O} = 2.78$ Å, $\theta_{O-H...O} = 157^\circ$) was formed.

In the simulations with the other force fields, generally somewhat more dihedral transitions were found, both with respect to the hydroxyl H atoms and to the hydroxymethyl group. In the simulations of α -D-galactose with the Ha force field, most distribution curves showed second and sometimes even third maxima.

Discussion

General remarks

On the computer we used, the calculations with the original GROMOS force field, where a hexapyranose molecule consists of 17 'atoms' instead of 24, are *ca* 1.8 times faster than the calculations with the all-atom models. With the computers available at the time when the GROMOS force field was developed, this decrease of computer time was much more important than it is now. In our opinion, it is only a small price to pay for the more realistic all-atom model.

For both the α - and β -form of D-galactose and for α -D-talose, two independently determined crystal structures have been published. The difference between the two can be seen as a measure for the quality of the structures, and should be taken into account when simulated structures are compared with the experimentally determined structures. Therefore, in Table 10 the information from Tables 6, 7 and 8 is given again for the comparison of the two independent experimental structures. From this table it is seen that the differences in the structure determinations of β -D-galactose and α -D-talose are very small, except for the temperature parameters. Obviously these parameters also contain considerable experimental and model errors. As noted before, in the determinations of α -D-galactose the positions of the hydroxyl H atoms are essentially different. The next subsection will discuss this structure in detail.

All unit cells decreased in size during the simulations, not only with the new force field, but also with the original GROMOS and Ha force fields. None of the three force fields, however, is generally better in this respect than the others. The systematic decrease is not alarming, but will be one of the observations on which we intend to optimize our parameter set.

The mean atomic positions are reasonable, compared with those obtained by experiments. Here too, none of the force fields is the best for all four structures, but the new force field performs very well in this comparison.

The scale parameters S and S_H from (2) and Table 7 would, in the ideal case, equal one, but it is difficult to

Table 10. As in Tables 6, 7 and 8, for comparison of two independent experimental structure determinations [for α -D-galactose the comparison is made between Ohanessian & Gillier-Pandraud (1976) and Sheldrick (1976)]

α -D-Galactose	β -D-Galactose	α -D-Talose	
Δa (%)	0.1	0.1	-0.2
Δb (%)	0.1	-0.0	-0.2
Δc (%)	-0.1	0.0	-0.3
ΔV (%)	0.0	0.1	-0.7
$\Delta r_{C,O,H}$ (Å)	0.41 (0.14)	0.06 (0.02)	0.04 (0.01)
$\Delta r_{C,O}$ (Å)	0.02 (0.01)	0.01 (0.00)	0.00 (0.00)
S	0.64	1.18	0.94
R_T	0.32	0.07	0.26
S_H	0.59		1.68
$(\Delta\theta)_{non-H}$ (°)	0.7 (0.1)	0.3 (0.2)	0.1 (0.0)
$(\Delta\theta)_H$ (°)	6.9 (4.8)	3.6 (2.2)	2.6 (0.1)
$(\Delta\varphi)_{non-H}$ (°)	0.8 (0.1)	0.4 (0.0)	0.2 (0.0)
$(\Delta\varphi)_H$ (°)	107.0 (71.0)	3.2 (2.6)	5.1 (5.1)

determine the temperature parameters accurately with X-ray crystallography. This is illustrated by Table 10, where purely experimental structures, all determined at room temperature, are compared. Likewise, the anisotropic disagreement factor R_T should be, ideally, very small. Nevertheless, when two experimental determinations of the same compound are compared, it can be up to 0.3. This points to almost no correlation: even when the temperature parameters of β -D-galactose (Sheldrick, 1976) are compared with the correlated ones of α -D-talose (Hansen & Horvik, 1977) according to (2), the value of R_T is 0.30. Therefore, only one reliable remark can be made here: the anisotropic temperature parameters are not useful indicators for the quality of a force field; only the comparison of isotropic temperature factors of the non-H atoms is useful to verify whether the space-group symmetry is retained during the simulation. Looking at the results presented in Table 7, it is seen that this is the case for all the simulations performed.

All angles and dihedrals without H atoms agree reasonably with the experimentally determined structures (see Table 8). The differences in the dihedrals with H atoms are acceptable, except for β -D-glucose (caused by the change in the hydrogen-bonding scheme of one H atom), which will be discussed later. A weak feature of the Ha force field is shown in this table: the dihedrals with H atoms do not agree too well with the experimental values. The new force field is an improvement in this respect, as it is also an improvement over the original GROMOS force field with respect to the bond angles without H atoms.

In the crystal structure determinations, no disorder was reported. In some of the simulations, however, dihedral transitions occurred which cause disorder in the structure. Nevertheless, the mean structures from most simulations where this was found deviate only a little more from the experimental structures than the others. Therefore, this phenomenon cannot be rejected as a failure of the force

field; when a crystal 1% disorder is present, this will not be found in the X-ray crystal structure determination. The largest deviations are found in the simulation of α -D-mannose, but these are also not alarming.

Whether the hydrogen-bonding schemes between experimental and simulated structures agree depends strongly on the agreement of the dihedrals with H atoms. Next to this, the criteria applied to define a hydrogen bond play a role. When extra bonds are found in the simulated structures, they often just meet the requirements. Only when *other* bonds are found than in the experimental structure, as in the case of α -D-galactose or β -D-glucose, is further investigation needed. Most of the crystal structure determinations used here have been published between 1975 and 1977. The methods available to determine the positions of the H atoms by X-ray diffraction are now more advanced. Therefore, we redetermined these two crystal structures, as discussed in the next subsections.

α -D-Galactose

The essential difference between the two independently determined crystal structures of α -D-galactose (see Tables 9 and 10) is in the positions of the hydroxyl H atoms. It is seen that these positions give rise to different hydrogen-bonding schemes, and that the molecular dynamics simulations do not agree completely with either one of them. All authors claim to have found the positions of the H atoms from a difference synthesis, but for the structure determined by Sheldrick (1976), this seems questionable when looking at the perfectly linear hydrogen bonds. The problem was discussed by Jeffrey & Shiono (1977), who compared these publications and suggested that they were both wrong, with Ohanessian & Gillier-Pandraud (1976) being the best. Hope (1978) has redetermined the structure at 83 K; unfortunately, only the cell parameters of that determination were published, not the coordinates.

To try and settle the matter, we prepared crystals according to the method described by Sheldrick (1961) and determined the structure again (see Appendix A). The structure appeared to be different from the two published structures (see Fig. 2). In Ohanessian & Gillier-Pandraud (1976), the position of only one hydroxyl group appeared to be wrong, and another was not very good. In Sheldrick's paper (1976), however, three of the five hydroxyl groups were situated in a wrong position, and the other two were situated just besides the favorable position.

The hydrogen-bonding scheme of the new structure determination (at liquid nitrogen temperature) is given in Table 9, and is most similar to the MD schemes. From these data it can be seen that the simulations significantly improved the partially incorrect experimental structure which we used as a starting point. In fairness, it should be added that Jeffrey & Shiono (1977) predicted the same

hydrogen-bonding scheme by careful analysis of the oxygen positions.

The coordinates from which the simulations of α -D-galactose started were from the structure determination of Ohanessian & Gillier-Pandraud (1976), which now appears to be the best of the two. In order to find out whether the force fields are able to correct a structure that is further from the real structure, new simulations were performed with the coordinates of Sheldrick (1976) as the origin. The same simulation conditions as before were applied. Unfortunately, the structure was not improved. With all three force fields many changes took place, but even during the last 10 ps the mean hydrogen-bonding scheme was the same as at the start. The average total energy during the simulation with the new force field was $150.92(4) \text{ kJ mol}^{-1}$ [$36.07(1) \text{ kcal mol}^{-1}$]; in the simulation started from the coordinates given by Ohanessian & Gillier-

Pandraud (1976), this was $136.9(1) \text{ kJ mol}^{-1}$ [$32.72(3) \text{ kcal mol}^{-1}$]. We continued this simulation by raising the temperature in 5 ps to 398 K, and after 40 ps at that temperature, lowering it again in 5 ps to 298 K. Then, 60 ps were simulated and the last 50 ps were analyzed. Here too, no significant improvement was made [the average total energy returned to the previous value: $150.8(2) \text{ kJ mol}^{-1}$ [$36.05(4) \text{ kcal mol}^{-1}$]]. The distribution function of the dihedral defining the conformation of the CH_2OH group shows that more molecules occupy another conformation than in the aforementioned simulations: almost 30%. Apparently, the energetical barriers to overcome the structural differences between the unfavorable structure determined by Sheldrick (1976) and our newly determined structure are too high for the short simulation times we used. In other words, the starting structure apparently occupies a rather deep local minimum in the potential energy surface, which it cannot leave during the simulations. Since the hydrogen bonds form a network, it is unlikely that a transition of a dihedral to form a better hydrogen bond will fit in this network and, therefore, the dihedral will easily return to the original position (unless many transitions occur simultaneously, which statistically will only happen once during a simulation time that is not possible to realize now). From our simulations it can be deduced that when most of the distribution curves of the dihedrals with a hydroxyl group show more than one maximum, this can be an indication that the simulation cannot leave the local minimum it is in. Since we did not see such behavior of the hydroxyl groups during other simulations presented in this paper, we have no suspicions about them.

β -D-Glucose

The experimental structure of β -D-glucose was determined by Chu & Jeffrey (1968). Neither the long intermolecular hydrogen bond predicted by their experiment, nor the bifurcated bond predicted by our simulation, is favorable. According to Chu & Jeffrey (1968), all H atoms 'were clearly revealed in the difference Fourier syntheses', but since the values of the C—O—H angles are between 102 and 128° , they may not be very reliable. The authors used stronger criteria to define a hydrogen bond than we do: they write that O4 does not act as a donor at all. At variance with this, Jeffrey (1990) wrote that hydroxyls in carbohydrate crystal structures function both as hydrogen-bond donors and acceptors, with very rare exceptions. Kanters, Kroon, Peerdeman & Vliegthart (1969) concluded from the IR spectrum that this hydroxyl group is one of these exceptions. In this spectrum a 'shoulder' is seen in the broad absorption band from the hydrogen-bonded hydroxyl groups, but it can be debated whether this is caused by a 'free' hydroxyl group. *Ab initio* calculations on $(\text{H}_2\text{O})_2$ showed that the shift in ν_{OH} can be small even though there is a

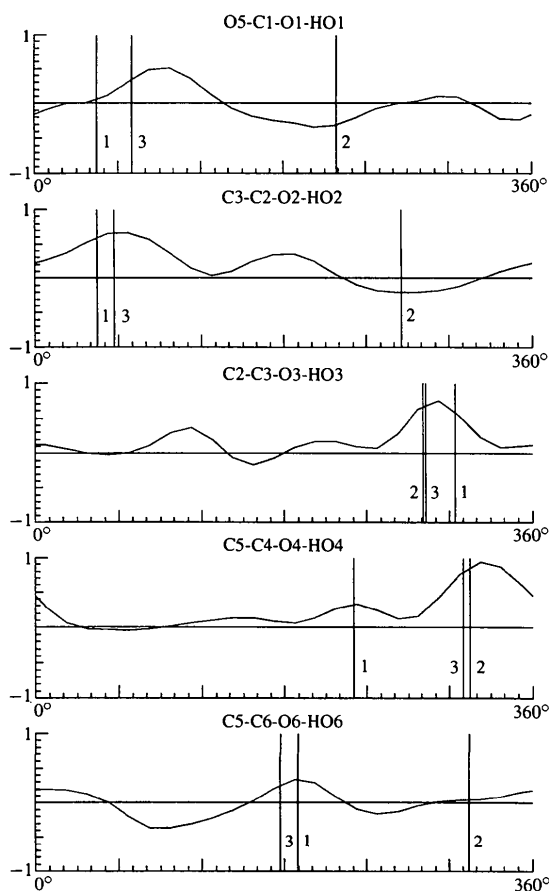


Fig. 2. Electron-density distributions (in $e \text{ \AA}^{-3}$) for the given dihedrals with an O—H distance of 0.82 \AA and a C—O—H angle of 109.5° , calculated from the low-temperature data of α -D-galactose. The values of the dihedrals according to the structure determination by Ohanessian & Gillier-Pandraud (1976), by Sheldrick (1976), and by ourselves at room temperature are marked 1, 2 and 3, respectively.

strong hydrogen bond formed (van de Rijdt-van Duijneveldt, van Duijneveldt, Kanters & Williams, 1984).

Here also, we prepared crystals (by recrystallization from ethanol and water) and redetermined the structure, as given in Appendix B. Apparently, the crystal structure determination by Chu & Jeffrey (1968) was correct. Nevertheless, the bifurcated hydrogen bond predicted by the simulation is reasonable. The mean value of the dihedral C5—C4—O4—HO4 is 159.5° , which is in the positive part of the electron density shown in Fig. 3. We suppose that this H atom is disordered and that both hydrogen bonds are formed in the crystal. Since it is not feasible to refine a disordered H atom from X-ray data, we cannot substantiate this experimentally.

Crystal structure prediction

An exciting subject is the *ab initio* prediction of crystal structures. We have recently started such work on six

hexapyranoses (van Eijck, Mooij & Kroon, 1995). Here we can investigate the behavior of the force fields upon energy minimization. We found that for each molecule the standard *GROMOS* force field was able to generate some thousand possible crystal structures within an energy window of $41.86 \text{ kJ mol}^{-1}$ (10 kcal mol^{-1}). The experimental structure was always among them, but the deviations in dihedral angles with a hydroxyl group varied considerably: they were within 20° for the standard *GROMOS* force field, 41° for the Ha force field, and 24° for our new force field. The mean differences $(\Delta\phi)_H$, as defined in Table 8, are 7.3, 13.3 and 6.3° , respectively. In this work we can also obtain some information about the reliability of the calculated energies. The average energy differences between the experimental structure and the calculated structure with the lowest energy for the three force fields were $12.560(3.0)$, $12.98(3.1)$ and 8.79 kJ mol^{-1} ($2.1 \text{ kcal mol}^{-1}$), respectively. Assuming that the experimental structure is one of those with the lowest energy, we conclude that the new force field is at this moment the best one available for the prediction of conformations and hydrogen-bond structures of monosaccharides, but there is still room for improvement.

Concluding remarks

Our aim, an improved all-atom force field for carbohydrates, was fulfilled satisfactorily. The extension of the original *GROMOS* force field to a more realistic all-atom force field, with some parameters from the Ha force field and adaptation of some nonbonded parameters, performs very well in the cases investigated here. With respect to the mean atomic positions during the simulations, the new force field means a significant improvement over the Ha force field. A weak feature of the Ha force field appears to be the description of the dihedrals with H atoms; the new force field is an improvement on that too. Bond angles without H atoms are described better by the all-atom force fields than by the *GROMOS* force field.

Not all data from diffraction experiments can be used to optimize force field parameters without distinction. The unit-cell dimensions and positions of the non-H atoms are very reliable, and so should be used, along with the internal coordinates retrieved from them. The positions of the H atoms are determined reliably with neutron diffraction, but with X-ray diffraction, one should be careful. They might be incorrect; only when they appear to be correct, can they be used for the optimization of a parameter set also. Finally, the temperature parameters should preferably not be used to the optimization, since they do not only contain the temperature movement of the atoms, but also experimental errors. They are useful, however, to evaluate the simulation. The temperature factors should be calculated from the simulation, and when these factors are

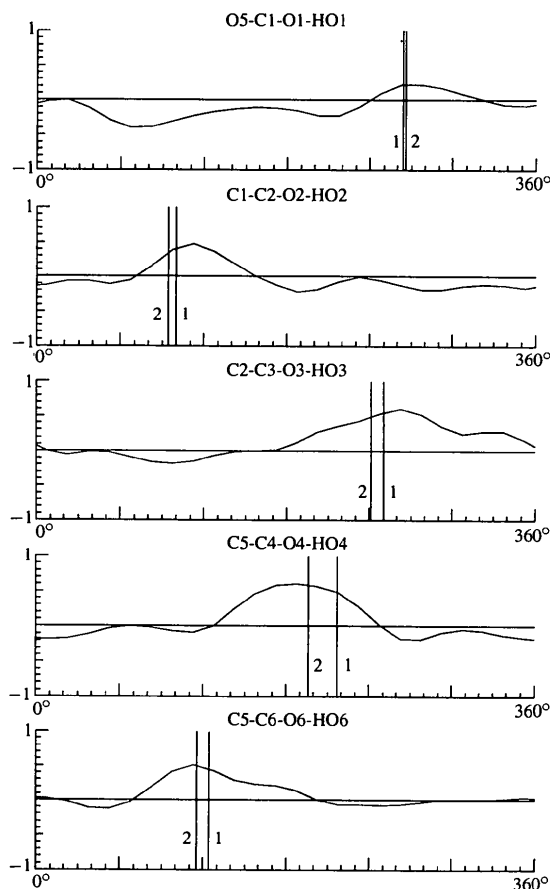


Fig. 3. Electron-density distributions (in $e \text{ \AA}^{-3}$) for the given dihedrals with an O—H distance of 0.82 \AA and a C—O—H angle of 109.5° , calculated from the low-temperature data of β -D-glucose. The values of the dihedrals according to the structure determination by Chu & Jeffrey (1968) and by ourselves at room temperature are marked 1 and 2, respectively.

Table 11. *Crystal data, data collection and refinement parameters for α -D-galactose*

Crystal data	
Chemical formula	C ₆ H ₁₂ O ₆
Molecular weight	180.16
Crystal system	Orthorhombic
Space group	<i>P</i> 2 ₁ 2 ₁ 2 ₁
<i>a</i> (Å)	5.8999 (8)
<i>b</i> (Å)	7.8433 (12)
<i>c</i> (Å)	15.685 (4)
<i>V</i> (Å ³)	725.8 (2)
<i>Z</i>	4
<i>D_x</i> (Mg m ⁻³)	1.649
Radiation type	Cu K α (Ni-filtered)
Wavelength (Å)	1.54184
No. of reflections for cell parameters	25
θ range for cell parameters (°)	17.15–35.10
μ (mm ⁻¹)	1.25
Temperature (K)	95 (2)
Crystal color	Colorless
Crystal form	Plate shaped
Crystal size (mm)	0.3 × 0.3 × 0.1
Data collection	
Diffractometer	Enraf–Nonius CAD-4
Data collection method	$\omega/2\theta$ scans
Absorption correction	None
No. of measured reflections	1539
No. of independent reflections	1341
<i>R</i> _{int}	0.0486
θ_{\max} (°)	75.03
Range of <i>h</i> , <i>k</i> , <i>l</i>	–7 → <i>h</i> → 0 –9 → <i>k</i> → 9 –19 → <i>l</i> → 0
No. of standard reflections	3
Frequency of standard reflections (min)	60
Intensity decay (%)	5
Refinement	
Refinement on	<i>F</i> ²
<i>R</i> { <i>F</i> ² > 2 σ (<i>F</i> ²)}	0.0590
<i>wR</i> (<i>F</i> ²)	0.1634
<i>S</i>	1.121
No. of parameters used	145
Weighting scheme	$w = 1/[\sigma^2(F_o^2) + (0.1317P)^2 + 0.20P]$, where $P = [\max(F_o^2, 0) + 2F_c^2]/3$
(Δ/σ) _{max}	–0.007
$\Delta\rho_{\max}$ (e Å ⁻³)	0.60
$\Delta\rho_{\min}$ (e Å ⁻³)	–0.24
Source of atomic scattering factors	<i>International Tables for Crystallography</i> (1992, Vol. C, Tables 4.2.6.8 and 6.1.1.4)
Computer programs	
Data collection	Locally modified CAD-4 software (Enraf–Nonius, 1989)
Cell refinement	SET4 (de Boer & Duisenberg, 1984)
Data reduction	HELENA (Spek, 1990b)
Structure solution	SHELXS86 (Sheldrick, 1990)
Structure refinement	SHELXL93 (Sheldrick, 1993)
Preparation of material for publication	PLATON (Spek, 1990a)

Table 12. *Fractional atomic coordinates and equivalent isotropic thermal parameters (Å²) for α -D-galactose*

$$U_{eq} = (1/3)\sum_i \sum_j U_{ij} a_i^* a_j^*$$

	<i>x</i>	<i>y</i>	<i>z</i>	<i>U</i> _{eq}
C1	0.5172 (5)	0.2544 (3)	0.1541 (2)	0.0151 (6)
H1	0.344 (7)	0.303 (6)	0.174 (2)	0.018
C2	0.6632 (4)	0.1811 (3)	0.2256 (2)	0.0132 (5)
H2	0.585 (6)	0.078 (5)	0.251 (2)	0.016
C3	0.8894 (4)	0.1255 (3)	0.1887 (2)	0.0134 (6)
H3	0.944 (7)	0.234 (6)	0.163 (2)	0.016
C4	0.8496 (5)	–0.0084 (3)	0.1185 (2)	0.0139 (6)
H4	0.993 (7)	–0.040 (5)	0.090 (2)	0.017
C5	0.6889 (5)	0.0689 (4)	0.0523 (2)	0.0141 (6)
H5	0.767 (7)	0.164 (5)	0.034 (3)	0.017
C6	0.6130 (5)	–0.0549 (4)	–0.0166 (2)	0.0166 (6)
H61	0.537 (7)	–0.145 (6)	0.007 (3)	0.020
H62	0.502 (7)	0.009 (5)	–0.059 (3)	0.020
O1	0.6248 (4)	0.3977 (3)	0.1222 (1)	0.0202 (5)
H10	0.508 (8)	0.464 (6)	0.092 (3)	0.030
O2	0.6885 (4)	0.3017 (3)	0.2931 (1)	0.0135 (4)
H20	0.745 (8)	0.374 (6)	0.277 (3)	0.020
O3	1.0325 (4)	0.0564 (3)	0.2536 (1)	0.0162 (5)
H30	1.077 (8)	0.154 (6)	0.282 (3)	0.024
O4	0.7678 (3)	–0.1653 (3)	0.1531 (11)	0.0158 (5)
H40	0.614 (8)	–0.163 (6)	0.165 (3)	0.024
O5	0.4816 (4)	0.1246 (2)	0.0914 (1)	0.0153 (5)
O6	0.8095 (4)	–0.1071 (3)	–0.0624 (1)	0.0198 (5)
H60	0.761 (8)	–0.180 (6)	–0.105 (3)	0.030

the number of atoms in incorrect positions is low, and they do not cooperate in a hydrogen-bond network, the simulation can overcome the energy barriers and predict the correct structure. We see this feature of the molecular dynamics simulations as a sign of promising quality already.

We thank Dr A. Sicherer-Roetman for her assistance with the preparation of the β -D-glucose crystals, and Drs A. J. M. Duisenberg and A. M. M. Schreurs for the data collection.

Appendix A. The crystal structure of α -D-galactose

Details of cell data, data collection and refinement for α -D-galactose are summarized in Table 11.

All H atoms were located from a difference electron-density synthesis. The positions were refined; the thermal parameters were taken to be 1.2 or 1.5 times the isotropic thermal parameter of the carrier C or O atom, respectively. The fractional atomic coordinates and isotropic thermal parameters are given in Table 12; bond lengths and angles are listed in Table 13.

The main difference between the two published crystal structure determinations (Ohanessian & Gillier-Pandraud, 1976; Sheldrick, 1976) was found in the positions of the hydroxyl H atoms. Therefore, the electron density on a circle was calculated (before the allocation of the H atoms) at a fixed O—H distance of 0.82 Å and a fixed C—O—H angle of 109.5° (Sheldrick, 1993) by varying the dihedral C—C—O—H. In Fig. 2 these electron densities are given as a function of a

considerably higher than those found experimentally, it means that the symmetry has been lost during the simulation.

Furthermore, the molecular dynamics simulations appear to be able to correct the positions of H atoms from an incorrect crystal structure determination, as is seen from the simulations of α -D-galactose. As long as

Table 13. Geometric parameters (Å, °) of α -D-galactose

O1—C1	1.384 (3)	O1—H10	0.98 (5)
O2—C2	1.427 (4)	O2—H20	0.70 (5)
O3—C3	1.430 (4)	O3—H30	0.92 (5)
O4—C4	1.429 (3)	O4—H40	0.93 (5)
O5—C1	1.431 (3)	O6—H60	0.93 (5)
O5—C5	1.436 (3)	C1—H1	1.13 (4)
O6—C6	1.424 (4)	C2—H2	1.01 (4)
C1—C2	1.527 (4)	C3—H3	1.00 (4)
C2—C3	1.519 (4)	C4—H4	0.99 (4)
C3—C4	1.540 (4)	C5—H5	0.92 (4)
C4—C5	1.531 (4)	C6—H61	0.92 (5)
C5—C6	1.520 (4)	C6—H62	1.06 (4)
C1—O5—C5	112.6 (2)	C6—O6—H60	107 (3)
O1—C1—O5	113.4 (2)	O1—C1—H1	104 (2)
O1—C1—C2	108.2 (2)	O5—C1—H1	107 (2)
O5—C1—C2	108.7 (2)	C2—C1—H1	116 (2)
O2—C2—C1	110.7 (2)	O2—C2—H2	107 (2)
O2—C2—C3	112.4 (2)	C1—C2—H2	109 (2)
C1—C2—C3	108.9 (2)	C3—C2—H2	109 (2)
O3—C3—C2	110.9 (2)	O3—C3—H3	115 (2)
O3—C3—C4	109.9 (2)	C2—C3—H3	101 (2)
C2—C3—C4	109.5 (2)	C4—C3—H3	110 (2)
O4—C4—C3	111.6 (2)	O4—C4—H4	104 (2)
O4—C4—C5	112.9 (2)	C3—C4—H4	111 (2)
C3—C4—C5	108.0 (2)	C5—C4—H4	109 (2)
O5—C5—C4	111.0 (2)	O5—C5—H5	108 (3)
O5—C5—C6	104.3 (2)	C4—C5—H5	103 (3)
C6—C5—C4	114.3 (2)	C6—C5—H5	116 (3)
O6—C6—C5	107.6 (2)	O6—C6—H61	112 (3)
C1—O1—H10	106 (3)	O6—C6—H62	109 (2)
C2—O2—H20	109 (4)	C5—C6—H61	110 (3)
C3—O3—H30	101 (3)	C5—C6—H62	109 (2)
C4—O4—H40	113 (3)	H61—C6—H62	108 (4)

dihedral. In each plot the values of the dihedral in the unpublished structure determinations are given.

Next to the low-temperature determination, data were collected at room temperature to compare the molecular dynamics simulation with this structure determination. The procedure followed differed from the low-temperature determination in the use of empirical absorption correction and in the placing of the H atoms on calculated (from the electron-density distributions as in Fig. 2) positions (final $wR_2 = 0.1128$, $R_1 = 0.0445$). The values of the dihedrals in this structure determination are also given in Fig. 2; they differ at most 10° from the values obtained at low temperature.

Appendix B. The crystal structure of β -D-glucose

Details of cell data, data collection and refinement for α -D-glucose are summarized in Table 14.

The aliphatic H atoms were located from a difference electron-density synthesis. The hydroxyl H atoms were placed on the position (with an O—H distance 0.82 Å and a C—O—H angle 109.5°) with the highest electron density. These electron densities, as a function of a dihedral, are given in Fig. 3. The positions were refined (the hydroxyl H atoms independent of the carrier O atom); the thermal parameters were taken to be 1.2 or 1.5 times the isotropic thermal parameter of the carrier C or O atom, respectively. The fractional atomic coordinates

Table 14. Crystal data, data collection and refinement parameters for α -D-glucose

Crystal data	
Chemical formula	C ₆ H ₁₂ O ₆
Molecular weight	180.16
Crystal system	Orthorhombic
Space group	<i>P</i> 2 ₁ 2 ₁
<i>a</i> (Å)	6.5949 (11)
<i>b</i> (Å)	9.014 (2)
<i>c</i> (Å)	12.720 (2)
<i>V</i> (Å ³)	756.2 (2)
<i>Z</i>	4
<i>D_x</i> (Mg m ⁻³)	1.582
Radiation type	Mo <i>K</i> α (Zr-filtered)
Wavelength (Å)	0.71073
No. of reflections for cell parameters	25
θ range for cell parameters (°)	13.91–19.74
μ (mm ⁻¹)	0.14
Temperature (K)	95 (2)
Crystal color	Colorless
Crystal form	Block shaped
Crystal size (mm)	0.5 × 0.4 × 0.3
Data collection	
Diffractometer	Enraf–Nonius CAD-4
Data collection method	$\omega/2\theta$ scans
	$\Delta\omega = 1.20 + 0.35\tan\theta$
Absorption/extinction correction	Empirical (<i>DIFABS</i> ; Walker & Stuart, 1983)
<i>T_{min}</i>	0.791
<i>T_{max}</i>	1.484
No. of measured reflections	1119
No. of independent reflections	1084
<i>R_{int}</i>	0.0769
θ_{max} (°)	27.49
Range of <i>h, k, l</i>	−8 → <i>h</i> → 0 −10 → <i>k</i> → 11 −16 → <i>l</i> → 0
No. of standard reflections	3
Frequency of standard reflections (min)	60
Intensity decay (%)	2.4
Refinement	
Refinement on	<i>F</i> ²
<i>R</i> [<i>F</i> ² > 2σ(<i>F</i> ²)]	0.0476
<i>wR</i> (<i>F</i> ²)	0.1324
<i>S</i>	1.078
No. of parameters used	145
Weighting scheme	$w = 1/[\sigma^2(F_o^2) + (0.0922P)^2 + 0.30P]$, where $P = [\max(F_o^2, 0) + 2F_c^2]/3$
($\Delta\sigma$) _{max}	−0.001
$\Delta\rho_{max}$ (e Å ⁻³)	0.36
$\Delta\rho_{min}$ (e Å ⁻³)	−0.33
Source of atomic scattering factors	<i>International Tables for Crystallography</i> (1992, Vol. C, Tables 4.2.6.8 and 6.1.1.4)
Computer programs	
Data collection	Locally modified CAD-4 software (Enraf–Nonius, 1989)
Cell refinement	SET4 (de Boer & Duisenberg, 1984)
Data reduction	HELENA (Spek, 1990b)
Structure solution	SHELXS86 (Sheldrick, 1990)
Structure refinement	SHELXL93 (Sheldrick, 1993)
Preparation of material for publication	PLATON (Spek, 1990a)

and isotropic thermal parameters are given in Table 15; bond lengths and angles are listed in Table 16.

The crystal structure was also determined at room temperature, using the same method for the determination of the hydroxyl H atoms (final $wR_2 = 0.1417$,

Table 15. Fractional atomic coordinates and equivalent isotropic thermal parameters (\AA^2) for β -D-glucose
$$U_{eq} = (1/3)\sum_i \sum_j U_{ij} a_i^* a_j^* a_i a_j$$

	x	y	z	U_{eq}
O1	-0.0867 (3)	-0.0669 (3)	0.2738 (2)	0.0236 (5)
H10	-0.166 (6)	-0.001 (5)	0.244 (3)	0.035
O2	-0.0666 (3)	-0.1059 (2)	0.4943 (2)	0.0189 (5)
H20	-0.093 (7)	-0.172 (5)	0.483 (3)	0.028
O3	0.3522 (3)	-0.1143 (2)	0.5644 (2)	0.0183 (5)
H30	0.332 (7)	-0.095 (5)	0.603 (3)	0.028
O4	0.5889 (3)	0.1307 (2)	0.4793 (2)	0.0207 (5)
H40	0.642 (6)	0.077 (5)	0.510 (3)	0.031
O5	0.2092 (3)	0.0612 (2)	0.2769 (1)	0.0154 (4)
O6	0.6081 (3)	0.0941 (3)	0.2052 (2)	0.0223 (5)
H60	0.588 (7)	0.086 (5)	0.148 (3)	0.033
C1	0.0367 (4)	0.0194 (3)	0.3384 (2)	0.0156 (5)
H1	-0.040 (6)	0.116 (4)	0.369 (3)	0.019
C2	0.1053 (4)	-0.0766 (3)	0.4291 (2)	0.0121 (5)
H2	0.159 (5)	-0.160 (4)	0.404 (3)	0.015
C3	0.2717 (4)	-0.0061 (3)	0.4940 (2)	0.0124 (5)
H3	0.205 (5)	0.085 (4)	0.534 (3)	0.015
C4	0.4413 (4)	0.0492 (3)	0.4218 (2)	0.0132 (5)
H4	0.504 (5)	-0.036 (4)	0.403 (2)	0.016
C5	0.3518 (4)	0.1477 (3)	0.3367 (2)	0.0145 (5)
H5	0.281 (6)	0.224 (4)	0.376 (3)	0.017
C6	0.5038 (5)	0.2096 (3)	0.2592 (2)	0.0200 (6)
H61	0.422 (6)	0.269 (4)	0.213 (3)	0.024
H62	0.610 (6)	0.266 (4)	0.296 (3)	0.024

Table 16. Geometric parameters (\AA , $^\circ$) of β -D-glucose

O1—C1	1.394 (4)	O1—H10	0.88 (4)
O2—C2	1.429 (3)	O2—H20	0.64 (4)
O3—C3	1.427 (3)	O3—H30	0.54 (4)
O4—C4	1.422 (3)	O4—H40	0.71 (4)
O5—C1	1.431 (3)	O6—H60	0.74 (4)
O5—C5	1.439 (3)	C1—H1	1.08 (4)
O6—C6	1.424 (4)	C2—H2	0.89 (4)
C1—C2	1.511 (4)	C3—H3	1.06 (4)
C2—C3	1.513 (4)	C4—H4	0.90 (3)
C3—C4	1.531 (4)	C5—H5	0.97 (4)
C4—C5	1.519 (4)	C6—H61	0.96 (4)
C5—C6	1.513 (4)	C6—H62	0.98 (4)
C1—O5—C5	112 (2)	C6—O6—H60	117 (4)
O1—C1—O5	106.8 (2)	O1—C1—H1	113 (2)
O1—C1—C2	107.8 (2)	O5—C1—H1	111 (2)
O5—C1—C2	109.3 (2)	C2—C1—H1	109 (2)
O2—C2—C1	108.1 (2)	O2—C2—H2	112 (2)
O2—C2—C3	109.7 (2)	C1—C2—H2	109 (2)
C1—C2—C3	113.1 (2)	C3—C2—H2	105 (2)
O3—C3—C2	109.0 (2)	O3—C3—H3	112 (2)
O3—C3—C4	109.1 (2)	C2—C3—H3	107 (2)
C2—C3—C4	109.8 (2)	C4—C3—H3	110 (2)
O4—C4—C3	111.1 (2)	O4—C4—H4	105 (2)
O4—C4—C5	109.3 (2)	C3—C4—H4	102 (2)
C3—C4—C5	109.5 (2)	C5—C4—H4	119 (2)
O5—C5—C4	108.3 (2)	O5—C5—H5	110 (2)
O5—C5—C6	106.8 (2)	C4—C5—H5	104 (2)
C4—C5—C6	115.0 (2)	C6—C5—H5	113 (2)
O6—C6—C5	111.4 (2)	O6—C6—H61	113 (2)
C1—O1—H10	103 (3)	O6—C6—H62	105 (2)
C2—O2—H20	105 (4)	C5—C6—H61	103 (2)
C3—O3—H30	105 (5)	C5—C6—H62	111 (2)
C4—O4—H40	105 (3)	H61—C6—H62	114 (3)

$R_1 = 0.0507$). The differences between the values of the dihedrals in both structure determinations were at most 9° . In Fig. 3 the values we found at room temperature and the values from the structure from Chu & Jeffrey (1968) are given for comparison.

References

- ALLEN, F. H., DAVIES, J. E., GALLOY, J. J., JOHNSON, O., KENNARD, O., MACRAE, C. F., MITCHELL, E. M., MITCHELL, G. F., SMITH, J. M. & WATSON, D. G. (1991). *J. Chem. Inf. Comp. Sci.* **31**, 187–204.
- BERENDSEN, H. J. C., POSTMA, J. P. M., VAN GUNSTEREN, W. F., DiNOLA, A. & HAAK, J. R. (1984). *J. Chem. Phys.* **81**, 3684–3690.
- BOER, J. L. DE & DUISENBERG, A. J. M. (1984). *Acta Cryst.* **A40**, C410.
- BROWN, G. M. & LEVY, H. A. (1979). *Acta Cryst.* **B35**, 656–659.
- CHU, S. S. C. & JEFFREY, G. A. (1968). *Acta Cryst.* **B24**, 830–838.
- EIJCK, B. P. VAN (1994). *Mol. Simul.* **13**, 221–230.
- EIJCK, B. P. VAN, MOOLU, W. T. M. & KROON, J. (1995). *Acta Cryst.* **B51**, 99–103.
- Enraf-Nonius (1989). *CAD-4 Software*. Version 5.0. Enraf-Nonius, Delft, The Netherlands.
- GROOTENHUIS, P. D. J. & HAASNOOT, C. A. G. (1993). *Mol. Simul.* **10**, 75–95.
- GUNSTEREN, W. F. VAN (1987). *GROMOS. Groningen Molecular Simulation Program Package*. Univ. of Groningen, The Netherlands.
- HA, S. N., GIAMMONA, A., FIELD, M. & BRADY, J. W. (1988). *Carbohydr. Res.* **180**, 207–221.
- HANSEN, L. K. & HORDVIK, A. (1977). *Acta Chem. Scand. A*, **31**, 187–191.
- HOPE, H. (1978). *Am. Cryst. Assoc. Ser. 2*, **6**, 22–23.
- JEFFREY, G. A. (1990). *Acta Cryst.* **B46**, 89–103.
- JEFFREY, G. A. & SHIONO, R. (1977). *Acta Cryst.* **B33**, 2700–2701.
- KANTERS, J. A., KROON, J., PEERDEMAN, A. F. & VLIEGENTHART, J. A. (1969). *Nature*, **222**, 370–371.
- KOUWIJZER, M. L. C. E., VAN EIJCK, B. P., KROES, S. J. & KROON, J. (1993). *J. Comp. Chem.* **14**, 1281–1289.
- KROON-BATENBURG, L. M. J., VAN DER SLUIS, P. & KANTERS, J. A. (1984). *Acta Cryst.* **C40**, 1863–1865.
- LONGCHAMON, F., AVENEL, D. & NEUMAN, A. (1976). *Acta Cryst.* **B32**, 1822–1826.
- LONGCHAMON, F., OHANESSIAN, J., AVENEL, D. & NEUMAN, A. (1975). *Acta Cryst.* **B31**, 2623–2627.
- OHANESSIAN, J. & GILLIER-PANDRAUD, H. (1976). *Acta Cryst.* **B32**, 2810–2813.
- OHANESSIAN, J., AVENEL, D. & KANTERS, J. A. (1977). *Acta Cryst.* **B33**, 1063–1066.
- RIJDT-VAN DULNEVELDT, J. G. C. M. VAN DE, VAN DULNEVELDT, F. B., KANTERS, J. A. & WILLIAMS, D. R. (1984). *J. Mol. Struct.* **109**, 351–366.
- RYCKAERT, J. P., CICCOTTI, G. & BERENDSEN, H. J. C. (1977). *J. Comp. Phys.* **23**, 327–341.
- SHELDRIK, B. (1961). *J. Chem. Soc.* pp. 3157–3158.
- SHELDRIK, B. (1976). *Acta Cryst.* **B32**, 1016–1020.
- SHELDRIK, G. M. (1990). *Acta Cryst.* **A46**, 467–473.
- SHELDRIK, G. M. (1993). *SHELXL93. Program for the Refinement of Crystal Structures*. Univ. of Göttingen, Germany.
- SPEK, A. L. (1990a). *Acta Cryst.* **A46**, C-34.
- SPEK, A. L. (1990b). *HELENA. Program for Data Reduction*. Laboratorium voor Kristal- en Structuurchemie, Univ. of Utrecht, The Netherlands.
- WALKER, N. & STUART, D. (1983). *Acta Cryst.* **A39**, 158–166.