ACTA CHEMICA SCANDINAVICA 14 (1960) 1357—1363

The Structure of 2-Deoxy-ribose

S. FURBERG

Institute of Chemistry, University of Oslo, Blindern-Norway

The structure of 2-deoxy-ribose in the crystalline state has been determined by X-ray crystallographic methods. The sugar is found to occur in the pyranose form and to belong to the β-series, the conformation being 1*a*3*e*4*a*. This is explained in terms of a general preference for axial substituents on the anomeric carbon atom. The implications of this view are discussed.

The sugar of deoxyribonucleic acid (DNA), 2-deoxy-ribose, has been rather extensively investigated by chemical methods, and a number of reviews on its chemistry have been published[1,2]. However, several features of the stereochemistry of the free sugar are still in some doubt. The ring form does not seem to be definitely established, although a pyranose structure is to be expected, especially in view of recent work on ribose[3]. The methyl glycosides are isolated and assigned to the α- and β-series[4], but the configuration at the anomeric carbon atom in the free sugar does not appear to be known. Finally, the conformation of the sugar has not been determined.

We have now established the structure of 2-deoxyribose in the crystalline state, using the methods of X-ray crystallography. Details of the structure analysis are given below. Projections in two directions of the electron density in the crystal have been calculated (Fig. 1), showing that the molecule has the structure depicted in Fig. 2. It is seen that the molecule occurs in the pyranose structure with the normal staggered "chair" form found in all sugars hitherto studied. The orientation of the anomeric hydroxyl group relative to the two others shows that the sugar belongs to the β-series. The hydroxyl groups on C_4 and C_1 are axial (*a*) whereas that attached to C_3 is equatorial (*e*). The structure may thus be described as 1*a*3*e*4*a*.

As pointed out by Hassel and Ottar[5], there are two possible conformations for each pyranose sugar, the one being derivable from the other by interconversion of the pyranose ring. In the case of β-2-deoxy-ribose the two conformations are 1*a*3*e*4*a* and 1*e*3*a*4*e*. The stabilities of pyranose conformations have been discussed by a number of investigators. Reeves[6] predicts the preferred conformation by means of three instability factors, whereas Barker and Shaw[3] derive the relative stabilities by adding the "overlaps"

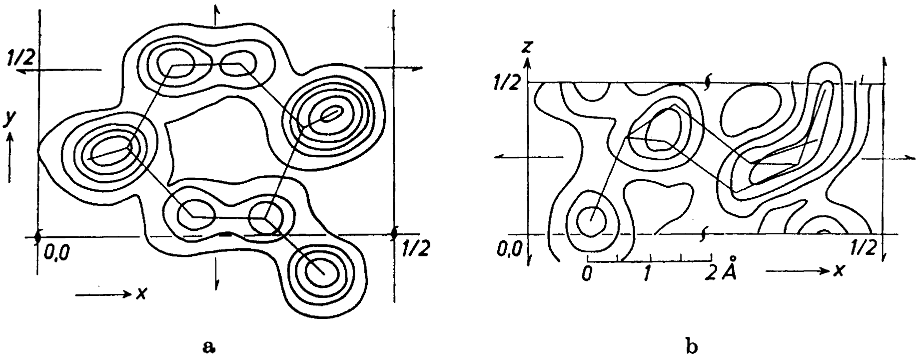

Fig. 1. Electron density projections on the (001) plane (a) and the (100) plane (b). Contours at 2,4,6,... e.Å^{-1}.

of non-bonded atoms. These methods have so far not been applied to deoxysugars but as both approaches consist in adding the unfavourable interactions in the molecule, they should be expected to be applicable also to these substances. In fact, the deoxysugars might provide an interesting test of the validity of the principles involved. In the case of 2-deoxy-ribose both methods predict that 1*e*3*a*4*e* is somewhat more stable than the observed form 1*a*3*e*4*a*, equatorial substituents being assumed to be more favourable than axial ones. The difference in stability is not great, and the solid state conformation may be determined by crystal forces. However, it is also possible that the molecular conformation 1*a*3*e*4*a* really is the more stable one, the reason, in case probably being that an equatorial hydroxyl group on the anomeric carbon atom C_1 may be an instability factor of importance in addition to those given by Reeves. This favours 1*a*-conformations, except possibly when there is an axial substituent in positions 3 or 5, in which case the effect is opposed by the repulsion between the two axial groups.

Greater stability of 1*a* substituents has been proposed as a possible explanation of the differences in reaction rates of α- and β-sugars. Thus the methyl β-pyranosides of glucose, mannose and galactose, in which the anomeric methoxyl group is equatorial, are hydrolysed more rapidly than the α-anomers. Edward [7] explained this in terms of electrostatic repulsive forces between the lone-pair electrons on the ring-oxygen atom and the polar group on C_1; this favours axial alkoxyl groups. Similarly, equatorial hydroxyl groups are oxidized more rapidly than axial ones [8]. However, for both types of reactions steric effects may provide an alternative explanation. Recently, the conformations of trans-di-halogeno-dioxanes have been studied in solid state and in solution. The results would seem to be in agreement with the preceding considerations, as the diaxial conformation is found in all cases [9].

The relation between this "1*e*-effect" and the powerful "Δ2-effect" of Reeves [6] should also be noted. The "Δ2-effect" corresponds to the conformation 1*e*2*a*, in other words to the combination of the two unfavourable factors 1*e* and 2*a*. Thus the "1*e*-effect" implies that the Δ2-condition must be unfavourable. Furthermore, it influences the relative importance of the two pyranose chair conformations also in other cases. The most notable deviations from previous stability considerations [3,6] would occur for sugars containing the configuration 1*a*2*a*3*e*⇌1*e*2*e*3*a*. Thus for α-lyxose, 1*a*2*a*3*e*4*e*⇌1*e*2*e*3*a*4*a*,

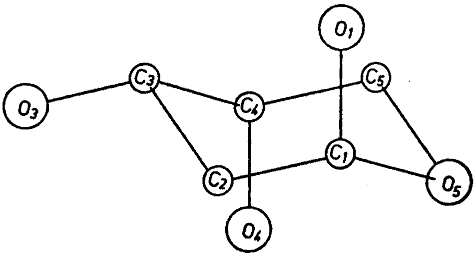

Fig. 2. The molecular structure of 2-deoxy-ribose.

the method of Reeves predicts conformational instability, which was also found in his studies of the cuprammin-complexes of methyl lyxosides. The approach of Barker and Shaw favours 1*e*2*e*3*a*4*a*, whereas the "1*e*-effect" gives preference to 1*a*2*a*3*e*4*e*. Proton magnetic resonance studies[11] have shown that α-lyxose tetra-acetate has the conformation 1*a*2*a*3*e*4*e* in solution.

If a 1*e* hydroxyl group represents an element of instability, the α-anomers of sugars as glucose, galactose and xylose should be more stable than the β-anomers, contrary to the view of Reeves[6]. As heat is evolved during the transformation from α- to β-glucose in aqueous solution, the latter anomer would appear to be the more stable one under these conditions[10]. The composition of the equilibrium mixture, in which the β-anomer predominates, points in the same direction[6]. However, the energy differences are so small that these observations do not represent conclusive evidence as regards the relative stability of the unhydrated molecules.

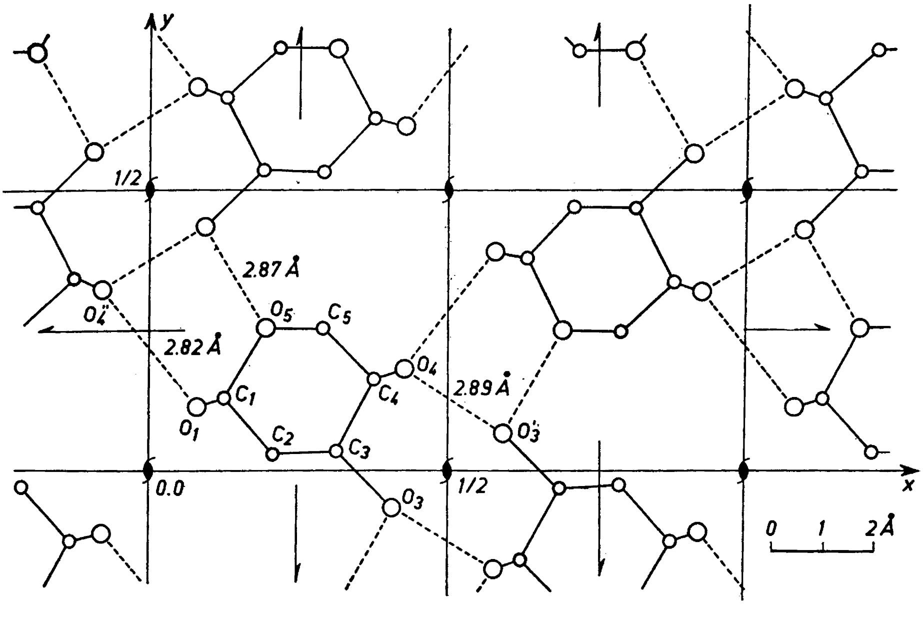

Fig. 3. The *c*-projection of the structure. Broken lines indicate hydrogen bonds.

Table 1. Atomic coordinates as fractions of the corresponding cell edge.

	x	*y*	*z*		*x*	*y*	*z*
C_1	0.130	0.131	0.312	H_1	0.05	0.13	0.45
C_2	0.210	0.030	0.418	H_2	0.23	0.04	0.62
C_3	0.316	0.029	0.239	H_2'	0.15	0.94	0.43
C_4	0.376	0.158	0.233	H_3	0.30	0.00	0.03
C_5	0.286	0.255	0.135	H_4	0.43	0.16	0.05
O_1	0.080	0.155	0.051	H_5	0.34	0.35	0.19
O_3	0.404	0.060	0.324	H_5'	0.27	0.23	0.94
O_4	0.419	0.185	0.496	H_{O_1}	0.02	0.18	0.01
O_5	0.188	0.252	0.311	H_{O_3}	0.13	0.38	0.46
				H_{O_4}	0.48	0.14	0.50

In Table 1 the atomic coordinates deduced from the final difference maps are given. The mean deviation in the *x* coordinates from the two projections is 0.015 Å and largest discrepancy is 0.04 Å. These figures may be taken as an indication of the accuracy of the coordinates. The corresponding bond lengths and bond angles are given in Table 2.

Table 2. Bond lengths (in Å units) and bond angles.

Bond lengths		Bond angles	
C_1-C_2	1.50	$O_5C_1C_2$	111°
C_2-C_3	1.49	$O_1C_1C_2$	118°
C_3-C_4	1.53	$O_1C_1O_5$	107°
C_4-C_5	1.53	$C_1C_2C_3$	107°
C_5-O_5	1.41	$C_2C_3C_4$	111°
O_5-C_1	1.45	$C_3C_4C_5$	108°
C_1-O_1	1.40	$C_3C_4O_4$	109°
C_3-O_3	1.44	$C_5C_4O_4$	112°
C_4-O_4	1.40	$C_4C_5O_5$	109°
		$C_5O_5C_1$	112.5°
		$C_2C_3O_3$	113.5°
		$C_4C_3O_3$	107°

The bond lengths and bond angles derived in the present investigation are given in Table 2. The bond lengths do not deviate significantly from the normal values of 1.53 Å and 1.42 Å for C-C and C-O distances, respectively. The bond angle at the ring oxygen atom is found to be 112.5°, somewhat greater than the other ring angles, which lie in the range 107° to 111°. The mean value of the external angles is 111°, but some of them show possibly significant variations from the tetrahedral angle, as is found also in other sugar structures[12].

Only three intermolecular distances are shorter than 3.0 Å. As there are three hydroxyl hydrogens in the molecule, these interactions are likely to be hydrogen bonds. It will be seen from Fig. 3 that one of the bonds connect O_3 and the ring oxygen atom O_5. Thus the hydrogen at O_3 is forming this bond. The hydrogen at O_4 must then take part in the hydrogen bond O_4——O_3' and that at O_1 in O_1—O_4''. The lengths of these bonds are 2.87 Å, 2.89 Å and 2.82 Å, respectively.

Table 3. Observed (F_o) and calculated (F_c) structure factors.

hkl	F_o	F_c
200	30	−34
400	51	−47
600	27	−24
800	15	−15
10 0 0	17	+17
12 0 0	< 3	− 1
110	13	−13
210	11	− 8
310	6	+ 6
410	6	+ 7
510	14	+14
610	2	0
710	3	− 3
810	3	− 4
910	< 3	+ 1
10 1 0	< 3	+ 4
11 1 0	4	− 6
12 1 0	3	− 3
020	16	−15
120	48	−51
220	< 2	+ 3
320	44	−41
420	< 2	− 2
520	5	− 4
620	< 2	− 2
720	5	+ 5
820	< 3	+ 3
920	< 3	+ 3
10 2 0	< 3	+ 1
11 2 0	< 3	+ 1
130	< 2	0
230	37	−38
330	4	+ 4
430	39	−36
530	< 2	− 2
630	5	+ 5
730	4	− 4
830	16	+19
930	= 3	+ 4
10 3 0	3	− 3
11 3 0	= 3	− 2
12 3 0	= 3	+ 3
040	11	+11
140	7	+ 7
240	56	−52
340	49	+46
440	29	+30
540	2	+ 3
640	13	+17
740	13	−16
840	13	−14
940	< 3	− 2
10 4 0	5	+ 7
11 4 0	< 3	− 1
12 4 0	< 2	− 2
150	28	−28
250	19	−19
350	3	+ 4
450	4	− 3
550	14	+15
650	< 3	+ 2
750	6	+ 5
850	6	+ 6
950	< 3	+ 3
10 5 0	< 3	+ 2
11 5 0	4	− 3
12 5 0	3	− 3
060	7	− 6
160	< 3	− 3
260	9	+10
360	21	+22
460	4	− 6
560	< 3	− 2
660	< 3	− 1
760	8	−10
860	3	+ 5
960	< 3	+ 3
10 6 0	< 3	− 3
11 6 0	< 2	0
170	14	+11
270	23	+22
370	< 3	0
470	4	+ 6
570	9	− 9
670	3	− 3
770	3	− 4
870	< 3	+ 2
970	< 3	+ 1
10 7 0	6	− 7
11 7 0	< 3	+ 2
080	8	+ 8
180	15	−12
280	13	−12
380	6	+ 5
480	5	+ 4
580	5	+ 4
680	7	+ 7
780	< 3	− 2
880	< 3	− 2
980	< 2	0
10 8 0	< 2	0
11 8 0	3	− 3

Table 3. Continued.

hkl	F_o	F_c
190	14	+14
290	5	− 6
390	4	− 4
490	<3	− 2
590	10	−10
690	<3	− 1
790	<3	− 3
890	<2	+ 1
0 10 0	10	−10
1 10 0	4	− 6
2 10 0	7	+ 7
3 10 0	<3	− 1
4 10 0	<3	+ 1
5 10 0	<3	+ 1
6 10 0	3	− 3
7 10 0	<2	− 1
1 11 0	<2	− 2
2 11 0	<2	− 1
3 11 0	<2	0
4 11 0	<2	0
0 12 0	<2	0
1 12 0	6	− 5
2 12 0	2	− 2
3 12 0	<2	− 2
101	19	−22
201	14	−16
301	13	+12
401	8	+11
501	<6	+ 3
601	10	+ 9
701	<6	− 1
801	<6	− 8
901	<7	+ 5
002	14	−12
102	<5	+ 2
202	28	−27
302	32	−30
402	13	+11
502	34	−36
602	26	+25
702	<7	0
802	<7	− 5
902	<7	− 3
103	10	+10
203	14	+16
303	8	− 7
403	<6	− 3
503	<7	− 7
603	<7	0
004	9	+ 9
104	<7	+ 8
204	<7	− 6
304	9	+ 9
404	9	−10
504	<7	+ 1
604	<7	+ 1

THE STRUCTURE ANALYSIS

A commercial sample of D-2-deoxyribose from Mann Research Laboratories, New York, was used for the investigation. Single crystals big enough for X-ray work were grown by precipitating the compound from a solution in acetone, the solution being placed in a desiccator containing excess of ethyl acetate. The crystals obtained were elongated along the *c* axis, with (100) as the dominating face.

X-Ray single crystals diagrams of the Weissenberg type were taken around the *c* and *b* axis, using copper radiation. They showed that the crystals are orthorhombic, with unit cell dimensions $a = 11.37$ Å, $b = 10.64$ Å and $c = 4.86$ Å (all ± 0.5 %). The space group is $P2_12_12_1$ as reflexions *h*00, 0*k*0 and 00*l* are absent for odd values of the indices. Four molecules in the unit cell correspond to a calculated density of 1.51 g/cm^3.

The intensities of the reflexions were estimated visually by comparison with a standard scale. Difficulties were encountered, as the crystals which were big enough to give a satisfactory diagram proved to be aggregates of two or more crystals. The true single crystals found were so small that only the strongest reflexions were recorded. The intensity values derived are therefore probably somewhat less accurate than those usually obtained by this procedure. 69 of the 146 *hk*0 reflexions and 22 of the 63 *h*0*l* reflexions obtainable with Cu*K*α radiation were recorded with measurable intensity.

To determine the *c* projection of the structure, the approximate position of the molecule was first deduced by means of Bragg—Lipson structure factor graphs. These graphs also indicated the signs of four low-order strong reflexions (240, 400, 120 and 320). Using these signs as a basis, the "multiplication rule" was applied to derive the signs of 16 reflexions expressed by the sign of the 340 reflexion. The two corresponding electron density maps both clearly showed a six-membered ring system, but only one of them indicated a 1,3,4-substituted sugar. The very first set of structure factors based on this structure gave a remarkable agreement with the observed values, the reliability factor $R = \Sigma||F_{\text{obs}}| - |F_{\text{calc}}||/\Sigma|F_{\text{obs}}|$ having the value 0.28. The structure was further refined by a number of difference synthesis. The process was complicated by the fact that in the projection there are hydrogen atoms close to nearly all the heavier atoms. It was found essential for the refinement to include their contribution to the structure factors. The positions of most of them were indicated by the difference maps, but some had to be placed by stereochemical considerations. The final value obtained for R was 0.10. A common isotropic temperature factor $\exp(-4.0 \sin^2\Theta/\lambda^2)$ was used for all the atoms. As the experimental data are rather unsatisfactory, further refinements were not carried out. In Table 3 observed and calculated structure factors are compared. It may be remarked that the 16 signs originally deduced all were correct.

The relative z coordinates of the atoms were calculated from the known c projection and from standard values of the bond lengths. The molecule was then placed so as to give reasonable lengths to the intermolecular hydrogen bonds indicated by the c projection. This postulated structure was refined by difference syntheses, the final value of R being 0.10 also in this case. However, the small number of reflexions and the overlap of atoms make the coordinates less reliable than those derived from the c projection. The electron density projection is shown in Fig.1 b.

The investigation was supported by a grant from *Norsk Hydros Fond.*

REFERENCES

1. Overend, W. G. *Advances in Carbohydrate Chem.* **8** (1953) 45.
2. Overend, W. G. and Stacey M. *The Nucleic Acids*, Academic Press Inc., New York, 1955, Vol. I, p. 48.
3. Barker, G. R. and Shaw, D. F. *J. Chem. Soc.* **1959** 584.
4. Deriaz, R. E., Overend, W. G., Stacey, M. and Wiggins, L. F. *J. Chem. Soc.* **1949** 2836.
5. Hassel, O. and Ottar, B. *Acta Chem. Scand.* **1** (1947) 929.
6. Reeves, R. E. *J. Am. Chem. Soc.* **72** (1950) 1499.
7. Edward, J. T. *Chem. & Ind. (London)* **1955** 1102.
8. Bentley, R. *J. Am. Chem. Soc.* **79** (1957) 1720.
9. Altona, C., Romers, C. and Havinga, E. *Tetrahedron Letters* **1** (1959) 16.
10. Riiber, C. N. *Ber.* **57** (1924) 1599.
11. Lemieux, R. U., Kullnig, R. K., Bernstein, H.J. and Schneider, W.G. *J. Am. Chem. Soc.* **80** (1958) 6098.
12. McGeachin, H. McD., and Beevers, C. A. *Acta Cryst.* **10** (1957) 227.

Received February 10, 1960