JOURNAL
OF THE AMERICAN CHEMICAL SOCIETY

VOLUME 87, NUMBER 13 JULY 6, 1965

Physical and Inorganic Chemistry

The Crystal and Molecular Structure of the Cyclohexaamylose–Potassium Acetate Complex[1]

Albert Hybl,[2] Robert E. Rundle,[3] and Donald E. Williams

Contribution from the Institute for Atomic Research and Department of Chemistry, Iowa State University, Ames, Iowa. *Received August 13, 1964*

The crystal and molecular structure of the potassium acetate complex of cyclohexaamylose (CHA) has been solved using three-dimensional X-ray diffraction data. The α-D-glucose units are all in the pyranose staggered chair form with the C1 conformation (1a2e3e4e5e). CHA shows approximate sixfold molecular symmetry. The principal feature of the α-(1,4)-glucosidic linkage is a hydrogen bond (2.852 Å.) between atoms O_2 and O_3 of each contiguous pair of glucose residues. The linkage valence angle is 119.1°. The mean C–C and C–O distances are 1.528 and 1.426 Å. The bond angles for the glucose residues are within normal limits. An analysis of the anisotropic thermal motion shows that the CHA molecule is vibrating as a unit with superimposed glucose unit librations about the glucosidic linkages. A detailed structural model for V-amylose is proposed based on the CHA structure. The CHA molecules form a rigid framework permeated by continuous channels and cavities analogous to the zeolite molecular sieves. The K^+ ions are in distorted octahedral environments located in pockets on the outside of the CHA channels. The K^+ statistically occupy only about three of four symmetry equivalent sites in each unit cell. Two Ac^- anions are at pseudo-fluid sites inside the CHA channel framework. The crystals exhibit space group symmetry $P2_12_12$ with lattice parameters $a = 21.89$, $b = 16.54$, and $c = 8.30$ Å. Each unit cell contains 2 CHA, 3.08 KAc, and 19.4 H_2O. The final R value was 0.10 for observed data.

Introduction

The anomeric form of D-glucose is responsible for the remarkable differences in chemical and physical properties between cellulose and amylose. The flexibility of the amylose chain is the most important difference in the properties of cellulose and starch. This flexibility is not understood but is well illustrated by the many crystalline forms of starch, many with very different fiber repeat distances.

The conformation of the amylose chain depends on both the conformation of the glucose units and the geometry of the α-(1,4) linkages. From results on cuprammonium complex formation, Reeves[4,5] postulated that the amylose chain most likely contains combinations of either B1 and B3 boat forms or B1 and C1 forms. Bentley[6] claims that the reducing glucose in maltose is C1 while the nonreducing residue is a boat form intermediate between B1 and 3B. Recent n.m.r. studies of polymers of D-glucose by Rao and Foster[7] indicated that all of the glucopyranose residues have the C1 chair conformation.

(1) (a) Supported in part by a grant from the Corn Industries Research Foundation. Work was performed in part in the Ames Laboratory of the U. S. Atomic Energy Commission (Contribution No. 1529). Requests for reprints can be addressed to Document Library, Institute for Atomic Research, Iowa State University, Box 1129 ISU Station, Ames, Iowa 50012. Other inquiries should be addressed to the first-named author. (b) A more detailed form of this paper (or extended version, or material supplementary to this article) has been deposited as Document No. 8330 with the ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D. C. A copy may be secured by citing the document number and by remitting $2.50 for photoprints, or $1.75 for 35-mm. microfilm. Advance payment is required. Make checks or money orders payable to: Chief, Photoduplication Service, Library of Congress.

(2) Corn Industries Postdoctoral Fellow, 1961–1964.

(3) Deceased, Oct. 9, 1963.

(4) R. E. Reeves, *J. Am. Chem. Soc.*, 72, 1499 (1950).

(5) R. E. Reeves, *ibid.*, 76, 4595 (1954).

(6) R. Bentley, *ibid.*, 81, 1952 (1959).

(7) V. S. R. Rao and J. F. Foster, *J. Phys. Chem.*, 67, 951 (1963).

Rundle and co-workers[8-10] proposed a helical structure for the V form of amylose from X-ray studies of the butanol and iodine complexes. Although it was established that the helix contained six glucose residues per turn, no information concerning the conformation of the glucose units nor about the geometry of the linkages were obtained. Holló, Szejtle, and Toth[11] concluded from models that only the C1 and B1 glucose conformations were suited for formations of helices, whereas Freudenberg and Cramer[12] suggest the 3B conformation from similar studies.

Cyclohexaamylose (CHA), one of the Schardinger dextrins[13] obtained from starch by the action of *Bacillus macerans*, is composed of six glucopyranose rings linked together to form a macro-ring by α-(1,4)-glucosidic linkages. The obvious structural similarity and analogous complexing behavior between CHA and the amylose helix make CHA an ideal model compound for studying both the conformation of the glucose units and the geometry of the α-(1,4)-glucosidic linkage.

James, French, and Rundle[14] attempted an X-ray crystal structure analysis of the $CHA \cdot I_2$ complex. They located the iodine atoms but were unable to resolve the carbohydrate. The unit cell and space group data for several of the crystalline complexes of CHA have been determined by French.[15] The KAc complex appeared particularly favorable for a complete structure analysis.

The purpose of this paper is to report the results of a complete three-dimensional crystal structure analysis of the cyclohexaamylose–potassium acetate complex and to interpret these results in terms of their contribution to the better understanding of the chemistry and structure of starch.

Crystal Data

Crystals of cyclohexaamylose–potassium acetate were obtained by evaporation of an aqueous solution of CHA containing at least a twofold excess of KAc. When exposed to air the crystals effloresced. Subsequently all diffraction experiments were made with the apparatus enclosed in a polyethylene tent, with the relative humidity controlled to 60 ± 10%. The crystals are orthorhombic prisms bounded by {110} and {001} forms. The density of the crystals was measured to be 1.434 g./cc. by flotation in a mixture of bromobenzene and *o*-dichlorobenzene. The analytical composition of the unit cell was found to be $2(C_6H_{10}O_5)_6 \cdot 3.08(KAc) \cdot 19.4(H_2O)$. The KAc to carbohydrate ratio was obtained from a potassium analysis of thoroughly desiccated crystals. The amount of water was computed from the observed density of the crystals.

X-Ray diffraction photographs confirmed the orthorhombic space group $P2_12_12$ reported by French.[15] The lattice constants, measured by the back reflection Weissenberg method, are $a = 21.89$, $b = 16.54$, and $c = 8.30$ Å.

(8) R. E. Rundle and D. French, *J. Am. Chem. Soc.*, **65**, 558 (1943).
(9) R. E. Rundle and D. French, *ibid.*, **65**, 1707 (1943).
(10) R. E. Rundle, *ibid.*, **69**, 1769 (1947).
(11) J. Holló, J. Szejtle, and J. Toth, *Staerke*, **13**, 222 (1961).
(12) K. Freudenberg and F. Cramer, *Chem. Ber.*, **83**, 296 (1950).
(13) D. French, *Advan. Carbohydrate Chem.*, **12**, 189 (1957).
(14) W. T. James, D. French, and R. E. Rundle, *Acta Cryst.*, **12**, 385 (1959).
(15) D. French, Ph.D. Thesis, Iowa State University, 1942.

Collection and Treatment of Intensity Data

Two complete sets of X-ray diffraction intensity data were obtained using a General Electric single crystal orienter and scintillation counter with nickel-filtered copper X-radiation. Each set of data consisted of the intensities of 2559 reflections. The first set was obtained from two crystals using the moving crystal–moving counter method. The second set of data was obtained rapidly from one crystal by the stationary crystal–stationary counter method.

During the period that intensity measurements were being made, the reflections decreased steadily in intensity. An especially large decrease in the intensity of the (007) reflection was noted, and the crystal decomposition was frequently accompanied by cracking of the crystal parallel to the (007) planes.

Crystal no. 1 decomposed after 15 days of measurement, during which time 1734 intensities were obtained. The remainder of the moving crystal–moving counter data were obtained from crystal no. 2; at the end of 9 days of measurement crystal no. 2 was still intact, with the (007) reflection reduced to 80% of its initial value and the average reflection reduced to 85% of its original value.

Since it was desired to have a complete set of intensity data from one crystal, all 2559 reflections were remeasured rapidly by the stationary crystal–stationary counter method on crystal no. 3. At the end of the measurements the crystal appeared intact, although the intensity of the (007) reflection had declined to 55% of its initial value.

Several standard reflections, including the (007) reflection, were followed during the course of the measurements. A linear decline line was graphically estimated for these standard reflections. Subsequently all intensities were corrected back to zero time by use of this linear decline line for each crystal. Background counts were obtained from a plot of average background *vs.* 2θ.

The observed structure factors were obtained from the observed total and background counts by applying the Lorentz, polarization, and linear decline corrections. A counter nonlinearity correction was applied to the crude count totals. The method of estimating the errors in the structure factors has been discussed previously.[16] For these data, errors in addition to the usual statistical counting errors were estimated as 8% for both the total counts and the background counts. The weighting factors which were used for the least-squares refinement were $[s(F_o^2)]^{-2}$, where $s(F_o^2)$ is the estimated error in F_o^2. For purposes of calculating the discrepancy index, R, a reflection was designated as "unobserved" if $F_o^2 < 2.25s(F_o^2)$.

The data which were used for the structural analysis are primarily the stationary crystal data, using the moving crystal data for checking and correction where necessary. The intensity disagreement index, $\Sigma|I(\text{moving}) - I(\text{stationary})|/\Sigma I(\text{stationary})$, between the two sets of data was calculated to be 0.09. As a result of a comparison of the two sets of data, a total of 117 changes were made in the stationary crystal data. Of these, 71 could be ascribed to a specific cause, such as an incorrect angular setting; in these

(16) D. E. Williams and R. E. Rundle, *J. Am. Chem. Soc.*, **86**, 1660 (1964).

cases the properly scaled moving crystal intensity was inserted into the stationary data. In the remaining 46 cases the discrepancy could not be assigned to any specific cause, and the mean of the moving and stationary values was used, with an increased error estimate.

The resulting corrected stationary crystal data were used without further change in all subsequent work. These observed structure factors are listed in Table IV. A Wilson's plot for the data yielded an approximate scale factor and a temperature factor coefficient, B, of 4.14 Å.$^{-2}$. The scattering factor tables for H, C, O, and K used throughout this investigation are equivalent to those given in the "International Tables."[17]

Computational Procedures

(*i*) *Description of Programs Used in the Refinement.* Refinement of this structure was carried out using the Iowa State University IBM 7074 computer using programs written by A. H. The Fortran II structure-factor and least-squares program incorporated the triclinic structure factor (SF) expressions of the "International Tables,"[18] the slightly modified monoclinic expressions of Rollett and Davies,[19] and the orthorhombic space group expressions of Hybl and Marsh.[20] The least-squares (LS) normal equations are semi-diagonalized, a 3 × 3 (x, y, and z coordinates) and either a 2 × 2 (isotropic temperature and population parameters) or a 7 × 7 (anisotropic temperature and population parameters) matrix being collected for each atom. Adjustments to the over-all scale factor are obtained from a weighted average of the indicated changes for the individual atoms. The quantity minimized is $\Sigma w(F_o^2 - F_c^2)^2$. The program contains provisions for rigid-body (RB) refinement using the expressions developed in the next section. Up to 100 atoms can be refined either as independent atoms or as members of any one of up to nine RB groups.

A space group specific Fourier program was written in Fortran. The Beevers–Lipson summation technique was used for intervals of $^1/_{60}$ of the cell edge. A complete three-dimensional Fourier is computed in about 4 min.

(*ii*) *Theory of Rigid-Body Least Squares.* The rigid-body (RB) transformation can be written

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix}_i = \mathbf{O}^{-1}\left[\mathbf{R}\begin{pmatrix} U \\ V \\ W \end{pmatrix}_i + \begin{pmatrix} Tx \\ Ty \\ Tz \end{pmatrix}\right]$$

where $(x, y, z)_i$ are the unit cell fractional coordinates, $\mathbf{O}^{-1}$ is the inverse of the orthonormalization matrix, **R** is the best available RB rotation matrix, $(U, V, W)_i$ are the coordinates of the corresponding atoms of the RB unit relative to some molecular axes (say, for example, the molecule's inertia axes with the center of mass of the unit at the origin), and (Tx, Ty, Tz) are the best available RB translation components. The concept of RB–LS is to adjust only the rotation and translation components of the RB units in order to achieve an intermolecular refinement while holding rigid the "expected" intramolecular distances and angles.

The problem is to evaluate the components of an infinitesimal rotation matrix (**IR**) using the derivatives of the individual atoms. The usual form of the structure factor residual equation is

$$Vq = \sum_{i=1}^{N} (Px\ Py\ Pz)_i \begin{pmatrix} dx \\ dy \\ dz \end{pmatrix}_i - \Delta G$$

where $\Delta G = |F_o|^2 - |F_c|^2$ and $Px = (\partial|F_c|^2/\partial x)_i$, etc. The constrained set of coordinate differentials is related to the RB infinitesimals[21] by the relation

$$\begin{pmatrix} dx \\ dy \\ dz \end{pmatrix}_i = \mathbf{O}^{-1}\left[\mathbf{R}\begin{pmatrix} 0 & Rw & -Rv \\ -Rw & 0 & Ru \\ Rv & -Ru & 0 \end{pmatrix}\begin{pmatrix} U \\ V \\ W \end{pmatrix}_i + \begin{pmatrix} dTx \\ dTy \\ dTz \end{pmatrix}\right]$$

where Ru is the infinitesimal rotation around the u-axis, etc. The expression rearranges to

$$\begin{pmatrix} dx \\ dy \\ dz \end{pmatrix}_i = \mathbf{O}^{-1}\left[\mathbf{R}\begin{pmatrix} 0 & -W & V \\ W & 0 & -U \\ -V & U & 0 \end{pmatrix}_i\begin{pmatrix} Ru \\ Rv \\ Rw \end{pmatrix} + \begin{pmatrix} dTx \\ dTy \\ dTz \end{pmatrix}\right]$$

which is substituted into the SF residual equation. The modified residual equation is thus linear in the RB rotational and translational terms and can be used in the usual way to obtain the LS normal equations.

The post matrix multiplication of **R** by the **IR** matrix

$$\begin{pmatrix} 1 & Rw & -Rv \\ -Rw & 1 & Ru \\ Rv & -Ru & 1 \end{pmatrix}$$

will produce a new RB rotation matrix **R**; however, if used in this form, it will introduce small distortions into the transformed RB molecule owing to the slight nonorthogonality of the **IR** matrix. Whittaker[22] has shown how to orthonormalize this type of matrix. He defines $M^2 = 0.25\,(Ru^2 + Rv^2 + Rw^2)$ and $S^2 = 1 - M^2$. The components of the orthonormalized correction matrix (**C**) become $C11 = 0.25(Ru^2 - Rv^2 - Rw^2) + S^2$, $C12 = 0.5RuRv + RwS$, $C13 = 0.5RuRw - RvS$, $C21 = 0.5RuRv - RwS$, $C22 = 0.25(-Ru^2 + Rv^2 - Rw^2) + S^2$, $C23 = 0.5RvRw + RuS$, $C31 = 0.5RuRw + RvS$, $C32 = 0.5RvRw - RuS$, $C33 = 0.25(-Ru^2 - Rv^2 + Rw^2) + S^2$. The new undistorted **R** is obtained *via* **R**(new) = **R**(old)**C** and Tx(new) = Tx(old) + dTx, etc.

By iterative application of the RB–LS process one obtains a refined rotation matrix and translation vector for each RB unit. The translation components parallel to the molecular axes system are given by

$$\begin{pmatrix} Tu \\ Tv \\ Tw \end{pmatrix} = \mathbf{R}^{-1}\begin{pmatrix} Tx \\ Ty \\ Tz \end{pmatrix}$$

Determination of the Structure of Cyclohexaamylose

(*i*) *Analysis of the Projection Data.* The (hk0) Patterson projection shown in Figure 1 exhibits a re-

(17) J. A. Ibers in "International Tables for X-ray Crystallography," Vol. III, K. Lonsdale, Ed., Kynoch Press, Birmingham, England, 1962, p. 202.
(18) N. F. M. Henry and K. Lonsdale, Ed., "International Tables for X-ray Crystallography," Vol. I, Kynoch Press, Birmingham, England, 1952, p. 374.
(19) J. S. Rollett and D. R. Davies, *Acta Cryst.*, 8, 125 (1955).
(20) A. Hybl and R. E. Marsh, *ibid.*, 14, 1046 (1961).
(21) H. Goldstein, "Classical Mechanics," Addison-Wesley, Reading, Mass., 1959, p. 124.
(22) E. T. Whittaker, "Analytical Dynamics," 4th Ed., Dover Publications, New York, N. Y., 1944, p. 7 ff.

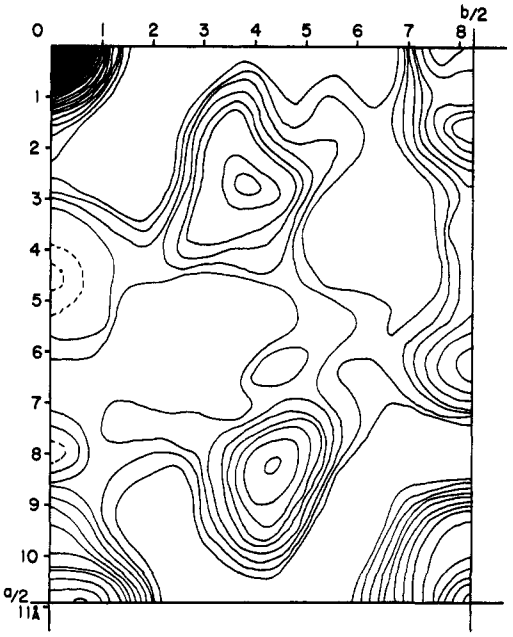

Figure 1. The (*hk*0) Patterson map of cyclohexaamylose.

markable resemblance to the Patterson pattern of a benzene ring. By equating a glucose residue to each of the carbon atoms in the benzene ring, the orientation of the macro-ring structure of CHA can be easily recognized; two *para*-glucose units must lie on the *x*-axis with the *y*-axis separating the remaining glucose units. The *z*-axis is normal to the plane of the macro-ring and coincident with the pseudo-hexagonal symmetry axis of CHA.

A detailed trial model was next formulated. (The naming of the atoms is explained in Figure 4.) The pyranose structure of the glucose residues was assumed to occur in the normal staggered chair form as has been found in all sugars hitherto studied. The atomic coordinates of α-D-glucose reported by McDonald and Beevers[23] were transformed for use as a "reference" residue. It was further assumed that the six glucosidic linkage oxygen atoms were coplanar and arranged in a regular hexagonal constellation with the distance along any side equal to the O_1-to-O_4 separation in the reference glucose residue. The final assumption involved the amount of rotation about an axis through the O_1 and O_4 atoms. This was assumed to be determined by a 2.80-Å. O_2–O_3 interglucose hydrogen bond. Combining this set of assumptions with the results of the Patterson map gave a trial structure and, subsequently, an electron density projection onto (001).

The (001) Fourier confirmed the general character of the trial model. Each glucose residue was resolved into two peaks with atoms C_1, C_2, O_2, and O_5 clustered into the smaller peak and atoms C_3, C_4, C_5, C_6, O_3, and O_6 clustered into the larger peak. A 10° counterclockwise rotation about the *z*-axis was indicated. The improved resolution of a subsequent electron density map revealed some new features interpreted as water peaks appearing just outside the cyclodextrin ring near the C_1 clusters.

(*ii*) *Analysis in Three Dimensions.* In many instances heavy atoms incorporated into the crystal assist and expedite the structural analysis. In this instance the potassium atoms provided no help in solving the structure. The first set of gravimetric results erroneously indicated a ratio of one K^+ to one CHA molecule and resulted in a fruitless search for K^+ at the special positions along the twofold axis.

We concluded that most of the atoms had to lie close to planes separated by *c*/7 Å. with only a few intercalated atoms from the fact that the (007) reflection had near maximal intensity. A three-dimensional unsharpened Patterson map clearly confirmed the sixfold molecular symmetry of CHA; the benzene-like pattern was evident on almost every layer except for one large nonbenzene peak with maximum on the 17–60th layer along *c*. We interpreted this peak as arising from the overlapping of several interglucose vectors between molecules related by a 2_1 axis with the *z*-value of the layer corresponding to the distance of separation between two pseudo-mirror planes. We noted that atoms O_1, C_1, and C_4 roughly define a pseudo-mirror plane; that is, atoms C_2, C_3, and O_3 mirror into atoms O_5, C_5, and C_6 while atoms O_1, C_1, and C_4 mirror into themselves and atoms O_2 and O_6 have no correspondents.

One of our early three-dimensional Fourier maps revealed an unusually large peak near G_2 which led us to question the accuracy of the original potassium analysis. A subsequent series of accurate analytical determinations for potassium indicated that there were 3.08 K^+ per unit cell. The Fourier maps showed that they must be distributed statistically into the four symmetry equivalent positions. The combined interpretation of the available evidence yielded a set of atomic parameters which was refined to the final results.

Refinement of the Structure

(*i*) *Refinement of Rigid-Body Parameters.* Rigid-body least squares was used successfully during the (*hk*0) refinement of the angular orientation of the trial model about the *z*-axis. Several additional calculations were made to determine the convergence range of the method. The hexagonal nature of the trial model limited the possible range to 60°. The convergence range was not symmetric about the minimum; it extended to 25° on one side but to only 11° on the other, giving a total convergence range of 36°.

(*ii*) *Refinement of Individual Atom Parameters.* The starting set of atomic parameters was refined through use of 20 cycles of SFLS. The first five cycles were run using only the 627 most intense reflections with unit weighing. Further refinement used all of the three-dimensional data and counter statistical weights. The identification and locations of additional water peaks W_K, W_5, W_6, W_7, and Ac^- were obtained from Fourier maps computed at various stages during the refinement.

Several of these interstitial atomic positions were refined with variable population parameters. The potassium position is a case of joint statistical occupancy (discussed more fully later). The scattering power of this position was represented by *f*-curve for K^+ with a variable population parameter.

Anisotropic temperature parameter refinement was started with the 15th cycle. Positions for the 21 hydro-

(23) T. R. R. McDonald and C. A. Beevers, *Acta Cryst.*, 5, 654 (1952).

TABLE I

THE FINAL ATOMIC PARAMETERS AND THEIR STANDARD DEVIATIONS

EXCEPT FOR THE ISOTROPIC TEMPERATURE FACTORS, ALL PARAMETER VALUES HAVE BEEN MULTIPLIED BY 10**4. THE ANISOTROPIC TEMPERATURE FACTORS ARE IN THE FORM T = EXP(-B11*H**2-B22*K**2-B33*L**2-B12*H*K-B13*H*L-B23*K*L).

ATOM	CX	CY	CZ	B11	B22	B33	B12	B13	B23
G1C1	1522	2467	3514	13	36	88	-8	2	4
S	4	6	11	2	4	13	5	9	14
G1C2	1243	2892	4975	11	30	104	-3	-4	-8
S	4	6	12	2	4	13	4	8	12
G1C3	573	2630	5192	11	35	68	-2	7	11
S	4	6	11	2	4	12	5	8	12
G1C4	221	2819	3653	10	31	86	4	-2	-8
S	3	6	11	1	4	12	4	8	13
G1C5	536	2410	2198	10	44	79	1	15	-8
S	4	7	11	2	5	13	5	8	14
G1C6	243	2660	635	13	70	74	16	10	15
S	4	8	12	2	6	12	6	8	17
G1O2	1603	2700	6379	15	39	113	-1	-19	-24
S	3	4	9	1	3	10	3	7	10
G1O3	300	3070	6494	14	52	109	12	5	-49
S	3	5	9	1	3	11	4	7	11
G1O4	-360	2466	3854	10	27	91	0	8	10
S	2	4	7	1	2	9	3	6	9
G1O5	1166	2669	2134	12	40	83	-2	10	32
S	3	4	8	1	3	9	3	6	9
G1O6	519	2265	-707	28	110	90	29	12	1
S	4	7	10	2	7	11	7	9	16
G2C1	2445	-494	3535	11	34	96	1	-4	26
S	4	6	12	2	4	14	4	9	13
G2C2	2595	29	4996	13	35	82	3	-17	18
S	4	6	11	2	4	11	5	8	14
G2C3	2117	672	5179	12	33	71	-5	-2	-1
S	4	6	11	2	4	13	4	8	12
G2C4	2042	1156	3621	8	33	71	2	6	-3
S	3	6	11	1	4	12	4	7	13
G2C5	1938	582	2185	13	38	52	15	12	13
S	4	6	10	2	4	12	5	8	12
G2C6	2010	1071	617	30	49	60	25	-3	13
S	5	7	12	3	5	13	7	11	15
G2O2	2642	-472	6401	20	38	96	-2	-27	24
S	3	4	8	1	3	10	4	8	10
G2O3	2276	1232	6418	20	41	87	2	-16	-12
S	3	4	8	1	3	10	4	7	10
G2O4	1497	1636	3826	9	25	96	2	9	11
	2	4	7	2	5	9	3	5	8
G2O5	2434	0	2141	16	36	102	9	15	17
S	3	5	8	1	3	9	4	6	11
G2O6	1934	575	-721	40	60	124	13	-6	-11
S	5	6	10	3	4	13	6	11	14
G3C1	886	-2944	3564	11	32	150	9	10	6
S	4	6	14	2	4	17	4	10	16
G3C2	1300	-2873	5005	13	35	123	6	-1	31
S	4	6	13	2	4	15	5	9	14
G3C3	1505	-1987	5190	11	35	68	1	1	13
S	4	6	10	2	4	12	4	8	12
G3C4	1803	-1706	3649	8	37	59	3	-3	8
S	3	6	10	1	4	11	4	8	12
G3C5	1374	-1858	2201	16	37	61	-4	4	-19
S	4	6	11	2	4	12	5	8	13
G3C6	1702	-1677	609	23	56	91	-10	23	-25
S	5	8	13	2	6	14	6	10	16
G3O2	1005	-3179	6389	17	51	124	-1	-8	63
S	3	5	9	1	4	11	4	7	12
G3O3	1946	-1930	6464	14	50	85	1	-7	14
S	3	5	8	1	3	10	4	6	10
G3O4	1872	-852	3840	9	29	79	1	7	0
S	2	4	7	1	2	8	3	5	8
G3O5	1202	-2675	2142	15	38	67	-3	11	-13
S	3	4	7	1	3	9	3	6	9
G3O6	1323	-1791	-698	33	94	103	-34	3	-19
S	4	7	11	2	6	13	6	9	16
G1W	3291	-1601	583	44	67	159	47	37	4
S	5	6	12	3	5	15	7	12	15
K+	3167	-122	-684	29	58	171	-4	1	-24
S	2	2	4	1	2	5	2	4	5
G3W	3677	1244	563	47	49	152	-24	-28	-6
S	5	6	11	3	4	15	6	11	14
WK	4274	-54	-2674	154	75	925	48	309	131
S	13	12	33	14	8	85	19	63	58
W5	4315	-161	3857	168	100	750	-9	317	166
S	14	13	34	16	12	75	23	61	56
AC-O1	590	565	-738	53	138	364	77	1	-110
S	14	22	40	8	22	71	24	43	69
AC-O2	286	-916	-916	31	149	268	14	12	42
S	10	21	32	5	21	49	18	28	57
AC-C	211	-83	-767	13.0	ISOTROPIC TEMPERATURE FACTOR				
S	23	44	61	1.6					
AC-ME	646	-267	-918	18.0	ISOTROPIC TEMPERATURE FACTOR				
S	32	45	95	2.3					
W6	0	0	2653	23	77	699	15	0	0
S			72	8	22	142	24		
W7	320	-130	6183	15.6	ISOTROPIC TEMPERATURE FACTOR				
S	16	25	47	1.1					

TABLE II

ASSUMED H ATOM PARAMETERS

COORDINATES ARE TIMES 10**4.

ATOM	CX	CY	CZ	B
G1H1	1995	2660	3343	3.5
G1H2	1257	3543	4776	3.7
G1H3	551	1982	5440	3.7
G1H4	185	3471	3468	3.5
G1H5	516	1755	2327	3.7
G1H6A	291	3313	494	4.5
G1H6B	-240	2502	672	4.5
G2H1	2791	-965	3397	3.5
G2H2	3035	322	4791	3.7
G2H3	1683	385	5480	3.7
G2H4	2439	1540	3409	3.5
G2H5	1493	282	2248	3.7
G2H6A	2468	1335	579	4.5
G2H6B	1672	1554	590	4.5
G3H1	749	-3575	3415	3.5
G3H2	1706	-3241	4779	3.7
G3H3	1111	-1604	5460	3.7
G3H4	2244	-2002	3475	3.5
G3H5	967	-1482	2307	3.7
G3H6A	2095	-2080	501	4.5
G3H6B	1858	-1052	627	4.5

TABLE III

VARIABLE POPULATION PARAMETERS AND MULTIPLICITY FACTORS

ATOM	P1	S	REMARKS
K+	0.885	.007	JOINT OCCUPANCY (SEE TEXT)
WK	1.0	.032	
W5	1.0	.034	
W6	.411	.032	SPECIAL POSITION
W7	.5	.027	OISORDERED AT SPECIAL POSITION
AC-	.500	-	OISORDERED AT SPECIAL POSITION

TABLE IV

OBSERVED AND CALCULATED STRUCTRURE FACTORS

(THIS TABLE HAS BEEN PLACED ON FILE WITH THE AMERICAN DOCUMENTATION INSTITUTE, LIBRARY OF CONGRESS, WASHINGTON 25, D. C.)

TABLE V

INTRAMOLECULAR DISTANCES

ALL INTRAMOLECULAR DISTANCES WITHIN CYCLOHEXAAMYLOSE AND WITHIN THE ACETATE ION WHICH ARE LESS THAN 3.0 ANGSTOMS ARE LISTED. THE MEAN ESTIMATED STANDARD DEVIATION FOR C-C DISTANCES IS S = 0.013 A, C-O S = 0.011 A, O-O S = 0.010 A FOR DISTANCES WITHIN THE ACETATE ION S = 0.3 A.

ATOM	ATOM	DIST	ATOM	ATOM	DIST	ATOM	ATOM	DIST
G1C1	G1C2	1.530		G6O5	2.353	G3C1	G3C2	1.504
	G1C3	2.517					G3C3	2.482
	G1C4	2.909	G1O5	G1O6	2.831		G3C4	2.867
	G1C5	2.422		G2O4	2.327		G3C5	2.376
	G1O2	2.416					G3O2	2.391
	G1O3	3.779	G2C1	G2C2	1.526		G3O5	1.439
	G1O5	1.426		G2C3	2.470		G4C4	2.435
	G2C4	2.450		G2C4	2.869		G4O4	1.417
	G2O4	1.398		G2C5	2.378			
				G2O2	2.418	G3C2	G3C3	1.541
G1C2	G1C3	1.540		G2O5	1.417		G3C4	2.491
	G1C4	2.493		G3C4	2.451		G3C5	2.875
	G1C5	2.888		G3O4	1.410		G3O2	1.412
	G1O2	1.443					G3O3	2.429
	G1O3	2.436	G2C2	G2C3	1.499		G3O5	2.409
	G1O5	2.393		G2C4	2.498		G4O4	2.366
	G2O4	2.352		G2C5	2.888			
				G2O2	1.435	G3C3	G3C4	1.509
G1C3	G1C4	1.524		G2O3	2.417		G3C5	2.507
	G1C5	2.513		G2O5	2.397		G3O2	2.465
	G1O2	2.464		G3O4	2.355		G3O3	1.435
	G1O3	1.433						
	G1O4	2.340	G2C3	G2C4	1.530	G3C3	G3O4	2.329
	G1O5	2.852		G2C5	2.520		G3O5	2.852
	G2O4	2.843		G2O2	2.435		G4O4	2.852
				G2O3	1.426			
G1C4	G1C5	1.546		G2O4	2.377	G3C4	G3C5	1.546
	G1C6	2.519		G2O5	2.842		G3C6	2.533
	G1O3	2.400		G3O4	2.807		G3O3	2.387
	G1O4	1.411					G3O4	1.429
	G1O5	2.434	G2C4	G2C5	1.540		G3O5	2.420
	G6C1	2.435		G2C6	2.498			
				G2O3	2.381	G3C5	G3C6	1.533
G1C5	G1C6	1.505		G2O4	1.443		G3O4	2.410
	G1O4	2.397		G2O5	2.429		G3O5	1.403
	G1O5	1.445					G3O6	2.411
	G1O6	2.423	G2C5	G2C6	1.541		G4O4	2.795
	G2O4	2.809		G2O4	2.414			
				G2O5	1.451	G3C6	G3O5	2.352
G1C6	G1O4	2.998		G2O6	2.413		G3O6	1.379
	G1O5	2.371		G3O4	2.745			
	G1O6	1.425				G3O2	G3O3	2.918
			G2C6	G2O5	2.366		G4O3	2.863
G1O2	G1O3	2.919		G2O6	1.391		G4O4	2.795
	G2O3	2.839						
	G2O4	2.764	G2O2	G2O3	2.931	G3O3	G3O4	2.819
				G3O3	2.853		G3O6	2.731
G1O3	G1O4	2.808		G3O4	2.784			
	G6O2	2.863				G3O5	G3O6	2.786
			G2O3	G2O4	2.824		G4O4	2.353
G1O4	G6C1	1.417		G2O6	2.717			
	G6C2	2.367						
	G6C3	2.852	G2O5	G2O6	2.782	AC-O1	AC-C	1.36
	G6C5	2.795		G3O4	2.342	AC-O2	AC-C	1.39
	G6O2	2.795				AC-C	AC-ME	1.97

TABLE VI

VALENCE ANGLES FOR THE GLUCOSE RESIDUES

THE MEAN ESTIMATED STANDARD DEVIATION FOR THE VALENCE ANGLES IS S = 0.8 DEGREES. * DENOTES AN ATOM IN AN ADJACENT GLUCOSE RESIDUE.

ANGLE	G1	G2	G3	ANGLE	G1	G2	G3
O1C1C2	106.8	106.6	108.1	C3C4O4	105.7	106.1	104.8
O1C1O5	111.0	111.9	110.9	C3C4C5	109.9	110.4	110.3
C2C1O5	108.0	109.0	109.8	O4C4C5	108.2	108.1	108.4
C1C2O2	108.7	109.5	110.1	C4C5O5	108.9	108.6	110.2
C1C2C3	110.1	109.5	109.1	C4C5C6	111.3	108.4	110.7
O2C2C3	111.4	112.2	113.1	O5C5C6	107.0	104.5	106.4
C2C3O3	110.1	111.4	109.4	C5O5C1	115.0	112.0	113.4
C2C3C4	108.9	111.1	109.5	C5C6O6	111.6	110.6	111.7
O3C3C4	108.5	107.2	108.3	C1O1C4*	119.1	119.4	118.8

TABLE VII

CONFORMATION ANGLES FOR THE GLUCOSE RESIDUES

FOR THE ATOMS C4C5C6O6 THE CONFORMATION ANGLE OF THE DIRECTED BOND C5 TO C6 IS DEFINED AS THE ANGLE, MEASURED COUNTER-CLOCKWISE, THAT THE PROJECTION OF THE BOND C5 TO C4 MAKES RELATIVE TO THE PROJECTION OF THE BOND C6 TO O6 WHEN ONE LOOKS IN THE DIRECTION OF THE BOND C5 TO C6. * DENOTES AN ATOM IN AN ADJACENT GLUCOSE RESIDUE.

ANGLE	G1	G2	G3	ANGLE	G1	G2	G3
O5C1C2C3	58.4	59.9	58.6	O1C1C2O2	61.2	62.2	62.1
C1C2C3C4	-57.3	-53.8	-55.9	O2C2C3O3	63.3	64.9	62.6
C2C3C4C5	55.6	51.9	53.6	O3C3C4O4	-68.1	-69.3	-71.0
C3C4C5O5	-55.8	-54.4	-55.2	O4C4C5C6	71.4	77.0	73.2
C4C5O5C1	60.8	63.0	59.8				
C5O5C1C2	-61.9	-65.8	-62.6	O2C2C1O5	-179.4	-176.8	-176.7
O3C3C4C5	175.3	173.9	172.8	O2C2C3*O3*	-9.0	-7.6	-4.4
C3C4C5O6	-173.6	-167.4	-172.7				
C4C5C6O6	-178.1	178.7	-178.6				

gen atoms attached to the carbon atoms of CHA were computed assuming tetrahedral bonding and a C–H distance of 1.09 Å. Their structure factor contributions were included in the last two SFLS cycles. Refinement was terminated when the coordinate shifts were less than $^1/_4$ to $^1/_2$ the standard deviation of the parameters. The final discrepancy value was 0.10 for the observed reflections and 0.12 for all data.

The final nonhydrogen coordinates and temperature parameters are listed in Table I. The assumed hydrogen atom coordinates are given in Table II. The population parameters which were refined or having values other than unity are in Table III. The structure-factor list is in Table IV (on file with the American Documentation Institute, Library of Congress, Washington 25, D. C.).

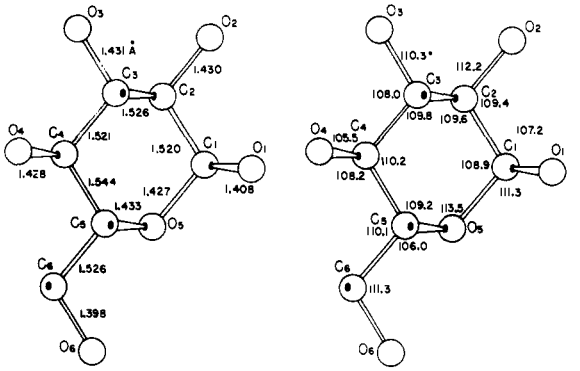

Figure 2. The conformation of the glucose residue. The mean values of the bond distances and valence angles for the glucose residue are illustrated.

Discussion of Results

(*i*) *Accuracy of Refinement.* The parameters for CHA, G_1W, K^+, and G_3W, which form the rigid framework of the crystal, have been determined to about 0.01-Å. accuracy. (The naming of the atoms is explained in Figure 4.) This crystal framework is permeated by continuous channels and cavities analogous to the zeolite structures in which SiO_4 and AlO_4 tetrahedra form molecular sieves. The water molecules, K^+, and Ac^- are poorly ordered and loosely accommodated into the relatively large cavities and channels, which precludes an accurate determination of their position. The framework atoms are all very well resolved in electron density sections as indicated by peak shapes and numbers of contours. The peaks for atoms W_K and W_5 are more diffuse. Only approximate positions and orientations are known for the intrachannel Ac^- and water molecules which are very poorly resolved in the electron density maps. Some additional acetate is inferred to be present only by the requirement of charge balance, indicating that it may have appreciable mobility within its intracrystalline cage.

(*ii*) *The Bond Distances and Angles.* The intramolecular distances and valence angles are listed in Tables V and VI. These values have not been corrected for the effect of thermal librations.

The mean bond distance and angle values for the glucose residue are shown in Figure 2. The average C–C and C–O distances are 1.528 ± 0.013 and 1.426 ± 0.011 Å. The average C_1–O bond length is 1.418 Å. The average over the other C–O bonds (except C_6–O_6) is 1.431 Å. However, this slight difference is near the limit of accuracy for this determination. The lengths of the C_6–O_6 bonds are most affected by thermal motion which accounts for their shortening.

The large shortening of the corresponding C_1–O bonds reported for cellobiose by Jacobson, Wunderlich, and Lipscomb[24] or for β-arabinose by Hordvik[25] was not observed. Brown and Levy,[26] however, did find a slight but significant shortening of the anomeric C_1–O_1 bond in α-D-glucose.

The carbon valence angles inside CHA range from 106 to 112°. The mean oxygen valence angles in the ether linkages are: C_1–O_1–C_4 (glucosidic linkage), 119.1°; C_1–O_5–C_5, 113.5°.

(*iii*) *The Conformation of the Glucose Residue.* The α-D-glucose units are all in the pyranose staggered chair form with the C1 conformation (1a2e3e4e5e) as shown in Figure 4. Atoms O_2, C_2, C_1, O_5 and O_3, C_3, C_4, C_5, C_6, O_6 form two approximately parallel planes. The deviations from these planes average less than 0.02 and 0.06 Å., respectively. The tendency for minimum overlaps of nonbonded atoms is clearly evident in the figure.

The six conformation angles of the ring alternate in sign and fall within the range of magnitudes 52 to 66°. Table VII lists the observed conformation angles in each glucose residue. The range is wider than found in cellobiose. The corresponding values in the three glucose units follow the same pattern. The divergence of the conformation angles is probably due to the combination of the effect of the intraring C–O bonds being shorter than the C–C bonds and to the induced strain from the macro-ring structure of CHA.

The orientation of the C_6–O_6 bond can be defined by the $C_4C_5C_6O_6$ conformation angle. A comparison of this angle as it occurs in different structures containing the glucose residue indicated there are at least two preferred orientations in the crystalline state. The 180 ± 12° orientation is found in all of the glucose residues in both CHA and cellobiose[24] while the −60 ± 2° orientation (O_6 staggered between atom O_5 and C_4) is found in the α-D-glucosamine hydrogen halides[27] and in sucrose.[26] The determining factor between these two orientations must certainly be the intermolecular hydrogen-bonding schemes within the crystals.

(*iv*) *The Geometry of the α-Glucosidic Linkage.* CHA is the first molecule studied which contains the α-(1,4)-glucosidic linkage as found in starch. The principal geometric feature of this linkage is the formation of a hydrogen bond (2.852 Å.) between atoms O_2 and O_3 of each pair of contiguous glucose residues. Another feature is the large valence angle (119.1°) for the linkage oxygen.

(*v*) *The Analysis of the Rigid-Body Thermal Motion.* The r.m.s. amplitudes along the principal directions of the anisotropic thermal motion for each atom are given in Table VIII. The analysis of the thermal motion of the atoms in terms of rigid-body motion[28] reveals the translational vibrations arise from the motion of the whole CHA molecule, while the largest rotational vibrations arise from the libration of the glucose residues about the O_1–O_4 axes.

The translational vibration tensors and the direction cosines of the principal translational directions are given in Table IX. The motion of all three glucose units coincides in magnitude and direction with the motion of CHA. The directions of motion are approximately parallel to the inertia axes of CHA which are nearly parallel to the cell axes. The largest displacement are along the *y*-axis.

In order to calculate the rotational vibration, the CHA molecule was treated as an ellipsoid of revolution because it has two nearly equal principal moments of inertia. Little of the thermal motion can be attributed

(24) R. A. Jacobson, J. A. Wunderlich, and W. N. Lipscomb, *Acta Cryst.*, 14, 598 (1961).
(25) A. Hordvik, *Acta Chem. Scand.*, 15, 16 (1961).
(26) G. A. Brown and H. A. Levy, *Science*, 147, 1038 (1963).
(27) G. A. Jeffrey and S. C. Chu, private communication, 1964.
(28) D. W. J. Cruickshank, *Acta Cryst.*, 9, 754 (1956).

TABLE VIII

ANALYSIS OF ATOM ANISOTROPIC THERMAL MOTION

THE R.M.S. AMPLITUOES (ANGSTROM UNITS TIME 10**3) ALONG THE PRINCIPAL OIRECTIONS, Q_i, OF THE ANISOTROPIC THERMAL MOTION AND THE OIRECTION COSINES (TIMES 10**3) OF THE PRINCIPAL AXES RELATIVE TO THE ORTHOGONALIZEO CELL AXES

ATOM	Q	RMS	OCX	OCY	OCZ	ATOM	Q	RMS	OCX	OCY	OCZ
G1C1	1	170	802	314	-508	G2O6	1	207	34	83	996
	2	177	491	138	860		2	277	-499	865	-55
	3	230	-341	939	43		3	324	866	495	-71
G1C2	1	162	954	184	236	G3C1	1	150	923	-364	-123
	2	189	-283	301	911		2	213	-279	-855	438
	3	208	-97	936	-339		3	233	265	370	891
G1C3	1	146	-485	-149	862	G3C2	1	171	817	-447	365
	2	171	-873	27	-487		2	192	-551	-417	723
	3	223	-49	989	143		3	239	171	792	586
G1C4	1	153	984	-155	83	G3C3	1	151	-156	-167	973
	2	172	-47	226	973		2	160	-987	53	-149
	3	209	169	962	-215		3	223	27	984	174
G1C5	1	136	797	-48	-602	G3C4	1	133	802	-116	585
	2	180	604	66	794		2	147	-592	-25	806
	3	246	-2	997	-82		3	226	79	993	89
G1C6	1	153	-525	43	850	G3C5	1	140	-67	202	977
	2	175	-824	223	-521		2	194	-967	-254	-14
	3	317	212	974	82		3	231	-245	946	-212
G1O2	1	166	684	246	686	G3C6	1	164	-323	103	941
	2	211	-718	395	574		2	231	-869	-426	-252
	3	242	130	885	-447		3	290	-374	899	-227
G1O3	1	158	-641	399	655	G3O2	1	174	268	-457	848
	2	195	-739	-92	-667		2	201	-961	-191	201
	3	284	206	912	-354		3	287	-70	868	491
G1O4	1	151	919	84	-385	G3O3	1	165	440	-118	890
	2	177	372	-510	776		2	188	-898	-32	440
	3	197	131	856	499		3	266	23	993	120
G1O5	1	148	-527	-290	799	G3O4	1	142	913	-40	-407
	2	181	-850	193	-491		2	171	406	-31	913
	3	243	12	937	348		3	200	49	999	12
G1O6	1	175	-165	35	986	G3O5	1	145	-279	118	953
	2	249	-945	281	-168		2	195	-937	252	243
	3	400	282	959	14		3	232	-211	961	-180
G2C1	1	160	844	-253	473	G3O6	1	188	23	77	997
	2	175	-536	-369	759		2	255	-901	-429	54
	3	226	17	894	447		3	382	-432	900	-59
G2C2	1	145	572	-236	785	G1W	1	212	-507	442	740
	2	196	-820	-136	557		2	253	-394	645	-655
	3	223	25	962	271		3	382	767	624	153
G2C3	1	157	262	80	962	G3W	1	218	301	414	859
	2	167	-930	-244	273		2	251	241	839	-489
	3	216	-256	967	-10		3	354	923	-354	-153
G2C4	1	133	917	-63	-394	K+	1	238	67	334	940
	2	161	396	25	918		2	264	-971	-196	138
	3	214	48	998	-48		3	291	-231	922	-311
G2C5	1	127	-351	13	936	WK	1	305	-80	986	-146
	2	161	-820	478	-314		2	451	652	-60	-756
	3	244	451	879	157		3	709	754	156	638
G2C6	1	143	81	-135	988	W5	1	314	268	807	-526
	2	219	-667	729	154		2	459	465	-586	-664
	3	306	741	671	31		3	707	844	67	532
G2O2	1	158	441	-229	868	W6	1	228	972	-237	0
	2	223	-675	-722	152		2	332	-237	-972	0
	3	246	-591	653	473		3	494	0	0	1000
G2O3	1	165	311	117	943	AC-O1	1	272	715	-565	-412
	2	224	-876	421	236		2	358	488	-18	872
	3	241	370	899	-233		3	496	500	825	-263
G2O4	1	142	958	-35	-286	AC-O2	1	271	973	-72	-218
	2	174	179	-706	685		2	304	206	-149	967
	3	197	226	707	670		3	457	102	986	131
G2O5	1	173	-646	47	762						
	2	191	-595	593	-542						
	3	240	477	804	355						

TABLE IX

THE TRANSLATIONAL VIBRATION TENSORS

THE OIRECTION COSINES OF THE PRINCIPAL AXES OF T ARE RELATIVE TO THE ORTHOGONALIZEO CELL AXES. R.M.S. VALUES ARE IN ANGSTOM UNITS.

	T*10**-2 A**2			RMS	OCX	OCY	OCZ
G1	(2.50	-0.03	0.40)	0.13	-0.455	0.003	0.891
	(	4.30	-0.02)	0.16	-0.890	-0.022	-0.455
	(		1.93)	0.21	-0.019	0.999	-0.013
G2	(2.85	0.35	-0.08)	0.14	0.109	-0.067	0.992
	(	4.54	0.14)	0.17	-0.975	0.188	0.120
	(		1.89)	0.21	0.194	0.980	0.045
G3	(2.56	0.30	0.08)	0.14	-0.132	0.035	0.991
	(	4.29	-0.05)	0.16	-0.977	0.163	-0.136
	(		1.87)	0.21	0.166	0.986	-0.013
CHA	(4.14	0.23	0)	0.16	0	0	1.000
	(	6.56	0)	0.20	-0.996	0.092	0
	(		2.53)	0.26	0.092	0.996	0

TABLE X

DIRECTION COSINES OF THE O1 TO O4 LINES

	OCX	OCY	OCZ
G1	-0.948	0.320	0.005
G2	-0.196	0.981	-0.003
G3	0.779	0.628	-0.003

TABLE XI

THE ROTATIONAL VIBRATION TENSORS

THE OIRECTION COSINES OF THE OMEGA TENSOR ARE RELATIVE TO THE ORTHOGONALIZEO CELL AXES. THE ORIGIN OF THE OMEGA LIBRATIONS FOR EACH GLUCOSE RESIOUE WAS ASSUMEO TO BE THE M10-POINT OF THE O1-O4 LINES. R.M.S. VALUES ARE IN OEGREES AND THOSE ENCLOSEO IN () OENOTE NEGATIVE EIGENVALUE SOLUTION WHICH WERE REPLACEO BY ZERO.

	OMEGA*OEG**2			RMS	OCX	OCY	OCZ
G1	(22.73	-4.32	2.36)	(0.00)	-0.056	0.236	0.970
	(	7.92	-2.18)	2.56	0.286	0.935	-0.211
	(		-2.47)	4.92	0.957	-0.268	0.107
G2	(8.54	-1.99	-1.22)	(0.00)	0.160	0.087	0.983
	(	15.15	-1.01)	2.86	-0.952	-0.251	0.177
	(		-3.31)	3.97	-0.261	0.965	-0.034
G3	(18.18	5.91	-1.26)	0.39	0.077	-0.022	0.997
	(	11.30	-0.21)	2.82	-0.493	0.868	0.058
	(		0.25)	4.65	0.867	0.496	-0.056
CHA	(1.49	0	0)	(0.00)	0	0	1.000
	(	1.49	0)	1.22	0	1.000	0
	(		-1.49)	1.22	1.000	0	0

TABLE XII

INTERMOLECULAR OISTANCES

ALL INTERMOLECULAR OISTANCES LESS THAN 3.5 ANGSTOMS AND THEIR ESTIMATEO STANOARO OEVIATIONS, S, ARE LISTEO. ATOM 1 (COOROINATES XYZ) IS ONE OF THE ATOMS IN THE BASIC SET NS=1. ATOM 2 (COOROINATES UVW) IS A SYMMETRY RELATEO ATOM. NS REFERS TO ONE OF THE FOLLOWING SYMMETRY TRANSFORMATIONS.

NS	TRANSFORMATION EQUATIONS			NS	TRANSFORMATION EQUATIONS		
1	U=X	V=Y	W=Z	6	U=0.5-X	V=0.5+Y	W=1-Z
2	U=-x	V=-Y	W=Z	7	U=0.5-X	V=Y-0.5	W=-Z
3	U=X	V=Y	W=Z-1	8	U=0.5-X	V=Y-0.5	W=1-Z
4	U=X	V=Y	W=Z+1	9	U=1-X	V=-Y	W=Z
5	U=0.5-X	V=0.5+Y	W=-Z				

ATOM	NS	ATOM	OIST	S	ATOM	NS	ATOM	OIST	S
G1C1	6	G3O3	3.499	.010	G1W	1	G2O5	3.493	.011
						1	G3C6	3.481	.015
G1C3	4	G1O6	3.458	.012		1	K+	2.675	.011
						7	G1O5	2.821	.011
G1C6	2	AC-O2	3.37	.03		7	G1O6	3.212	.014
						8	G1O2	2.784	.012
G1O2	4	G1O6	3.464	.011					
	6	G1W	2.784	.012	G3W	1	K+	2.725	.010
						5	G3O5	2.882	.011
G1O3	4	G1O6	2.720	.012		5	G3O6	3.251	.015
	6	W5	3.06	.02		6	G3O2	2.792	.012
G1O5	5	G1W	2.821	.011	K+	1	G2C6	3.387	.011
						1	G2O5	2.848	.007
G1O6	1	AC-O1	2.83	.04		1	G2O6	2.934	.010
	2	AC-O2	2.85	.03		1	G1W	2.677	.011
	3	G1C3	3.458	.012		1	G3W	2.725	.010
	3	G1O2	3.464	.011		1	WK	2.934	.023
	3	G1O3	2.720	.012		3	G2O2	2.741	.007
	5	G1W	3.212	.014					
					WK	1	K+	2.934	.025
G2C3	4	G2O6	3.430	.013		3	W5	2.89	.04
						9	WK	3.19	.04
G2C6	1	K+	3.387	.011					
	1	AC-O1	3.41	.04	W5	4	WK	2.89	.04
						6	G3O2	3.36	.02
G2O2	4	G2O6	3.332	.012		8	G1O3	3.06	.02
	4	K+	2.741	.007		9	W5	3.040	.04
G2O3	4	G2O6	2.717	.012	W6	1	AC-O1	3.24	.06
						1	AC-O2	3.39	.06
G2O5	1	G1W	3.493	.011		1	AC-C	2.88	.08
	1	K+	2.848	.007		1	AC-ME	3.31	.10
						1	W7	3.02	.07
G2O6	1	K+	2.934	.010					
	1	AC-O1	2.94	.03	W7	1	W6	3.02	.07
	1	AC-ME	3.15	.07		4	AC-O1	2.86	.05
	3	G2C3	3.430	.013		4	AC-O2	2.74	.05
	3	G2O2	3.332	.012		4	AC-C	2.54	.06
	3	G2O3	2.717	.012					
					AC-O1	1	G1O6	2.82	.04
G3C3	4	G3O6	3.451	.012		1	G2C6	3.41	.04
						1	G2O6	2.94	.03
G3C6	1	G1W	3.481	.015		1	W6	3.24	.06
						3	W7	2.86	.05
G3O2	4	G3O6	3.406	.013					
	8	G3W	2.792	.012	AC-O2	1	G3O6	2.70	.03
	8	W5	3.358	.023		1	G4C6	3.37	.03
						1	G4O6	2.85	.03
G3O3	4	G3O6	2.731	.011		1	W6	3.39	.06
	8	G1C1	3.499	.010		3	W7	2.74	.05
G3O5	7	G3W	2.882	.011	AC-C	1	W6	2.88	.08
						3	W7	2.54	.06
G3O6	1	AC-O2	2.70	.03					
	1	AC-ME	2.93	.07	AC-ME	1	G2O6	3.15	.07
	3	G3C3	3.451	.012		1	G3O6	2.93	.07
	3	G3O2	3.406	.013		1	W6	3.31	.09
	3	G3O3	2.731	.011					
	7	G3W	3.251	.015					

to the rigid-body rotational vibration of CHA. However, as Tables X and XI indicate, the individual glucose residues can be treated as rigid-body units to explain the thermal motion of atoms. In each case the largest libration (r.m.s. amplitude approximately 4.5°) is about the O_1–O_4 line.

In solving for the magnitudes of libration of the CHA molecule and the individual glucose units, negative minimum eigenvalues were encountered. These negative eigenvalues were set to zero, since their magnitude has no physical meaning. We interpreted this situation as indicating a small libration relative to the other eigenvalues. We feel that our results reflect correctly the relative relationship of the rotational vibrations, although the absolute values are subject to systematic error.

(vi) The Packing of the Molecules. The CHA molecules form rigid channel structures which are stacked

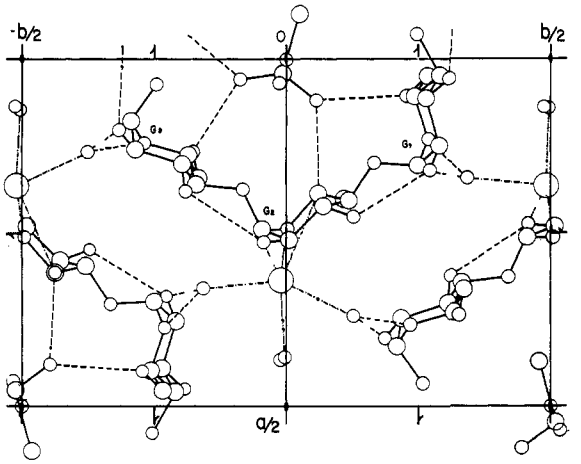

Figure 3. The projection of the cyclohexaamylose–potassium acetate complex onto (001). A possible configuration for the disorder of the acetate ions is shown.

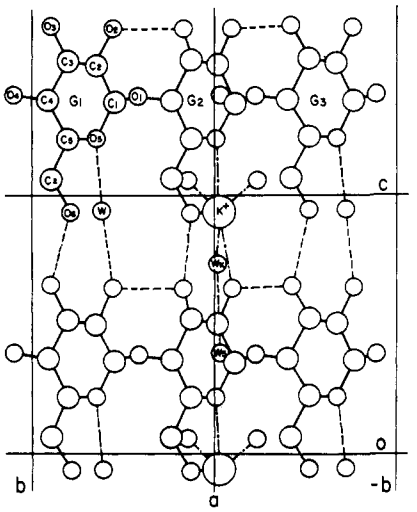

Figure 4. The projection of the cyclohexaamylose–potassium acetate structure onto a concentric cylinder, viewed from the inside. The axis of the cylinder is coincident with the twofold axis. The glucose units are named G_1, G_2, and G_3. The naming of the atoms within each glucose, as shown in the upper left, is in accordance with the usual convention used by carbohydrate chemists. The linkage oxygen atom between G_1 and G_2 can be referred to as G_1O_1 or G_2O_4 depending on the context of the discussion. Water molecules are denoted by W. G_1W, G_3W, and W_K refer to the water molecules associated with G_1, G_3, and K^+. A fourth water molecule competing statistically for the K^+ position is called G_2W. The other water molecules are not easily identified by an association with the various structural classes; hence, they are labeled W_5, W_6, and W_7. The atoms of the acetate group, not shown in this figure, are labeled AcO_1, AcO_2, AcC, and AcMe, the latter referring to the methyl end of the molecule.

parallel in an open orthorhombic framework. Adjacent molecules along a channel are connected by six O_3 to O_6 hydrogen bonds. The water molecules G_1W, G_3W, and G_2W along with K^+ link atoms O_5 and O_2 in the adjacent CHA molecules and contribute to the stability of the channel structure (see Figure 4 and Table XII). These atoms also cooperate to connect adjacent channels to form the framework of the crystal. Each carbohydrate channel is cross linked to four adjacent inverted channels and each K^+ is coordinated to three channels. Figures 3, 4, and 5 illustrate these interconnections. A distorted octahedral coordination

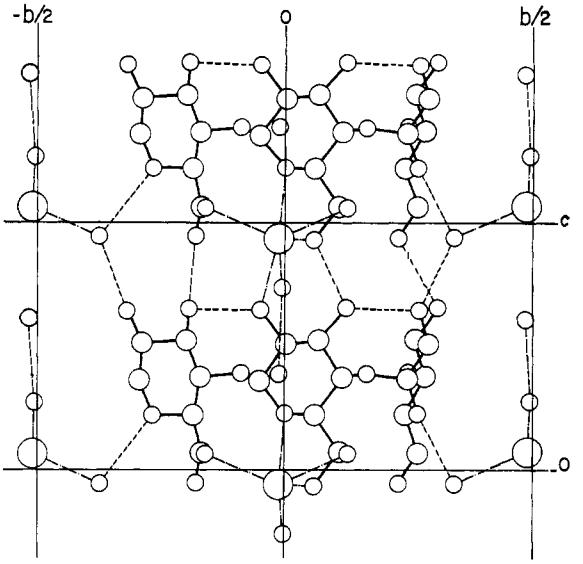

Figure 5. The projection of cyclohexaamylose onto (001).

about K^+ consists of atoms G_2O_5, G_2O_6, and G_2O_2 from one channel, atoms G_1W and G_3W from two adjacent channels, and atom W_K.

Atoms W_K and W_5 fill the superfluous cavity between channels. Their lateral restrictions are at best very weak hydrogen bonds. They are attached to K^+ in a pendulum-like fashion; W_5 is loosely hydrogen bonded to W_K which is coordinated to K^+. This interchannel cavity is the most plausible location for the extra Ac^-. The anisotropic diffuseness of W_K and W_5 would tend to support this idea. If present, Ac^- must contribute statistically to the scattering from this area and hence some of its scattering is probably included into the parameters of W_K and W_5. The occupancy factor for this ion is about 0.25, which substantially reduces its scattering power. Any orientational disorder or mobility of the ion would further decrease the probability of observing its scattering.

(*vii*) *Trouble in the Canal Zone.* The disorder and the very large anisotropy of Ac^- and the water molecules located inside the CHA channels clearly indicate their pseudo-fluid nature. The intrachannel Ac^- anions are statistically disordered at the twofold special positions. The Ac^- molecular plane appears to be more or less perpendicular to the twofold axis. Each Ac^- is hydrogen bonded to four of the O_6 hydroxyl groups on CHA.

The water molecule W_6 is located on the twofold axis near the center of the CHA torus and at a normal packing distance above Ac^-. Water molecule W_7 is located near the level of the O_2 and O_3 hydroxyl groups on CHA, packed between W_6 and Ac^- in the next unit cell. It is disordered into positions off the twofold axis which are within hydrogen-bonding range of the acetate oxygens.

(*viii*) *The Case of Joint Occupancy.* Careful quantitative analysis for K^+ indicated that there are 3.08 K^+ per unit cell. This implies that the crystal is nonstoichiometric in K^+ in the sense that a general position requires four atoms per unit cell. The observed population parameter for the K^+ position using f_{K^+} is 0.885 ± 0.007. Since it is unreasonable that holes the size of

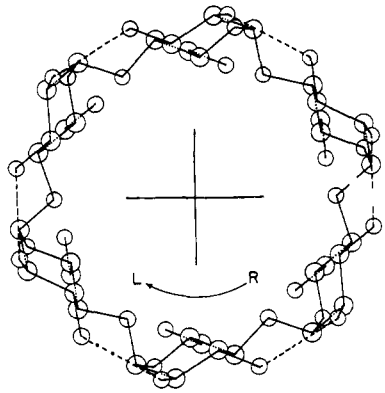

Figure 6. Projection of the proposed model for V-amylose along the helix axis.

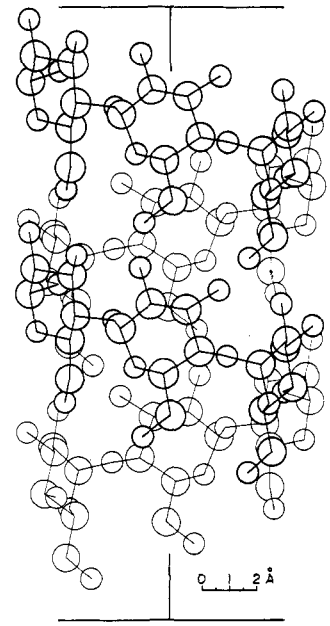

Figure 7. A view of the proposed model for helical V-amylose perpendicular to the helix axis.

K^+ ions would exist in the crystal, we expect that the remaining portion of this position is occupied by water (G_2W) which is hydrogen bonded to two CHA molecules in the same manner as G_1W and G_3W. (The hydrogen bond distances and the coordination distances are both about 2.8 Å.) Based upon this assumption we calculate 3.17 K^+ per unit cell from the X-ray data.

Only one K^+ per CHA is required as counterion for the Ac^- anion at the special position inside each CHA torus although there are two symmetry equivalent K^+ sites on the outside. Any additional K^+ incorporated in excess of one per CHA must depend on the accommodation of additional Ac^- somewhere else in the crystal framework. The crystal stability depends on the cross linking of adjacent CHA channels through the octahedral coordination about K^+. G_2W is not a stabilizing substitute in these cross linkages. Our analytical results, although not designed to test this point, seem to indicate that the ratio of K^+ to CHA in the crystals is relatively invariant to the amount of excess KAc in solution. Thus it seems that at least 3 K^+ per cell are needed for crystal formation but not much more is incorporated owing, in all probability, to the difficulty of accommodating the additional Ac^-.

(*ix*) *The Extrapolation to Amylose.* A postulated structure for the helical form of amylose is shown in Figures 6 and 7. It is a single-threaded sinistral helix composed of six α-D-glucose units per turn all in the C1 conformation linked α-(1,4) with hydrogen bonds between atoms O_2 and O_3 of each pair of contiguous glucose residues. An equivalent drawing of the dextral helix indicates that it is probably a less satisfactory model for amylose. Its elimination, however, as a plausible model cannot be accomplished by so simple an argument; we think this could be done experimentally using the (*hk*0) data from the butanol-precipitated amylose. The projections of the two helical models are different so they can be tested by structure factor calculations.

The $C_4C_5C_6O_6$ conformation angle of 180° was used for the C_6O_6 bond orientation illustrated in Figures 6 and 7. The +60° orientation was also considered but appears to increase the packing diameter too much and is less favorable for the interstitial inclusion of water at the pocket sites along the helix. For both orientations of the C_6O_6 bond a hydrogen bond can be formed between atoms O_6 in one turn to atom O_2 in the next turn. The interior of the helix is lined with the CH groups of atoms C_3 and C_5 and by the glucosidic oxygen atoms. The glucose units probably librate with a r.m.s. amplitude of near 4.5° about their O_1O_4 axes analogous to CHA. This suggests that helical amylose is more like a long flexible coiled spring than a rigid rod.

Zaslow and Miller[29] have shown that the anhydrous form of V-amylose takes up one molecule of water per glucose at interstitial sites along the exterior of the helix at a vapor pressure of 24–26 mm. By analogy to CHA, the pocket positions between the O_5 in one turn and the O_2 in the next are the probable locations of the water molecules. Their addition to the helical framework increases the packing diameter to 13.7 Å.[10]

The proposed model is like a cylindrical screw with a left-handed thread. It has an 8-Å. pitch with a lead angle of 17.5°. The major and minor diameters are 13.4 and 11 Å. The crest and root contours of the thread are irregular and the packing diameter is close to 13 Å.

The idealized packing of cylindrical screws is a hexagonal closest packing with interlocking of the crests and roots of the screw threads. The standard screw thread has a twofold axis normal to thread axis so that its head and tail are indistinguishable. At this point the amylose helix departs from the screw analogy because the reducing end is easily distinguishable from the nonreducing end. Rundle and Edwards[30] found that alternate helices are directed oppositely for the monohydrated butanol-precipitated amylose. In addition, the irregularity of the contour along the crest of the thread suggests that a preferred rotational orientation can be expected.

Amylose exhibits several other crystalline forms. Senti and Witnauer,[31–33] working with stretched fibers, have been able to effect an interconversion between these various forms. The explanation of concomitant

(29) B. Zaslow and R. L. Miller, *J. Am. Chem. Soc.*, **83**, 4378 (1961).
(30) R. E. Rundle and F. C. Edwards, *ibid.*, **65**, 2200 (1943).
(31) F. R. Senti and L. P. Witnauer, *ibid.*, **68**, 2407 (1946).
(32) F. R. Senti and L. P. Witnauer, *ibid.*, **70**, 1438 (1948).
(33) F. R. Senti and L. P. Witnauer, *J. Polymer Sci.*, **9**, 115 (1952).

structural changes which occur during these transitions is an area open to conjecture; however, there is no reason to expect a change in conformation of the glucose residue.

Acknowledgments. We are very grateful to Professor D. French for suggesting the problem, providing samples of CHA, and discussing the chemistry of this series of carbohydrate compounds. We thank the editor and referees whose comments on the first draft of this article were helpful in improving the presentation of the material. We are indebted to the Corn Industries Research Foundation for support of computer usage for the crystal structure analysis, and to the Health Science Computer Center of the University of Maryland School of Medicine for assistance with the condensation of our tables.

The Configuration of Random Polypeptide Chains. I. Experimental Results

David A. Brant and Paul J. Flory

Contribution from the Department of Chemistry, Stanford University, Stanford, California. *Received March 6, 1965*

The dimensionless characteristic ratio $\langle r^2\rangle_0/n_p l_p^2$ of the measured mean square unperturbed end-to-end distance $\langle r^2\rangle_0$ to the number n_p of planar, trans peptide units multiplied by the square of the length l_p between successive α-carbons has been evaluated for four polypeptides. This ratio was deduced from intrinsic viscosities, molecular weights, and second virial coefficients. Measurements reported here on poly-β-benzyl-L-aspartate in m-cresol at 100°, on poly-L-glutamic acid in aqueous 0.3 M sodium phosphate at pH 7.85 and 37°, and on poly-L-lysine in aqueous 1.0 M sodium bromide at pH 4.54 and 37° yielded values of the characteristic ratio of 9.6, 8.8, and 8.6, respectively. From the data of Doty, Bradbury, and Holtzer for poly-γ-benzyl-L-glutamate in dichloroacetic acid at 25°, a value of 8.8 was calculated. No dependence of the unperturbed dimensions upon the solvents or amino acid side chains represented here is discernible within the estimated experimental uncertainty of ca. 10%.

The ordered configurations of synthetic polypeptides have attracted much interest in recent years.[1] In contrast, little attention has been devoted to the random coil form of the polypeptide chain. Interpretation of changes in molecular configuration accompanying protein denaturation[2-8] demands an adequate characterization of the denatured form. Denaturation usually involves disordering of the native molecule in some degree, and under certain conditions conversion to the random coil form is complete. The molecular interpretation of dimensional changes occurring in fibrous proteins[9] also requires a knowledge of the random polypeptide chain dimensions.

The dimensions of the poly(benzyl glutamate) random coil in dilute dichloroacetic acid solution have been evaluated from published data[10,11] by P. J. F.[12] and by Kurata and Stockmayer,[13] and the configuration of polysarcosine in dilute aqueous solutions has been investigated by Fessler and Ogston.[14] Only these few experimentally determined polypeptide random coil dimensions appear to be available in the literature, and theoretical investigations of polypeptide coil dimensions have been limited to free rotation treatments.[12,15] In view of the paucity of information on the configuration of random coil polypeptides, we have undertaken a systematic experimental study of the configuration of these molecules in dilute solution.

We shall focus attention in what follows on the *average unperturbed polypeptide chain dimensions* which depend only on short-range interactions within the polymer chain, *i.e.*, those determined by the covalent chemical bonding and internal rotational potentials, including interactions between near-neighbor nonbonded groups.[13,16] The unperturbed polymer chain is of particular interest inasmuch as it is amenable to theoretical treatment which permits analytical correlation of the average chain dimensions with the polymer structure. A detailed theoretical investigation of the random polypeptide chain is presented in the following paper.[17]

The unperturbed dimensions of polymer chains are subject in general to direct experimental determination in dilute solution in appropriate ideal or Θ-solvents in which the volume exclusion effect upon the coil dimensions is nullified.[13,16] Such ideal solvents are necessarily poor solvents for the polymer in question.

(1) P. Urnes and P. Doty, *Advan. Protein Chem.*, **16**, 401 (1961).
(2) J. A. Schellman, *Compt. rend. trav. lab. Carlsberg, Ser. chim.*, **29**, 223, 230 (1955).
(3) W. Kauzmann, *Advan. Protein Chem.*, 14, 1 (1959).
(4) H. A. Scheraga, *J. Phys. Chem.*, **64**, 1917 (1960).
(5) C. Tanford, *J. Am. Chem. Soc.*, **84**, 4240 (1962).
(6) B. H. Havsteen and G. P. Hess, *ibid.*, **85**, 791 (1963); B. H. Havsteen, B. Labouesse, and G. P. Hess, *ibid.*, **85**, 796 (1963).
(7) D. B. Wetlauffer, S. K. Malik, L. Stoller, and R. L. Coffin, *ibid.*, **86**, 508 (1964).
(8) J. F. Brandts, *ibid.*, **86**, 4291, 4302 (1964).
(9) L. Mandelkern, *Ann. Rev. Phys. Chem.*, **15**, 421 (1964).
(10) P. Doty, J. H. Bradbury, and A. M. Holtzer, *J. Am. Chem. Soc.*, **78**, 947 (1956).
(11) G. Spach, *Compt. rend.*, **249**, 543 (1959).
(12) P. J. Flory, *Brookhaven Symp. Biol.*, **13**, 89 (1960).
(13) M. Kurata and W. H. Stockmayer, *Fortschr. Hochpolymer. Forsch.*, **3**, 196 (1963).
(14) J. H. Fessler and A. G. Ogston, *Trans. Faraday Soc.*, 47, 667 (1951).
(15) W. G. Crewther, *J. Polymer Sci.*, **A2**, 123 (1964).
(16) P. J. Flory, "Principles of Polymer Chemistry," Cornell University Press, Ithaca, N. Y., 1953, Chapter XIV.
(17) D. A. Brant and P. J. Flory, *J. Am. Chem. Soc.*, 87, 2791 (1965).