THE STRUCTURE OF CHITIN

N. E. DWELTZ

Department of Physics, University of Madras (India)

(Received February 22nd, 1960)

SUMMARY

A new structure is proposed for the polysaccharide chitin, $[C_8H_{13}O_5N]_n$. X-ray diffraction fibre photographs of lobster tendon reveal a well-defined orthorhombic lattice with the space group $P2_12_12_1$ and cell dimensions of $a = 4.69$ Å, $b = 19.13$ Å, and $c = 10.43$ Å (fibre repeat). It is suggested that the unit cell contains two polysaccharide chains running in opposite directions and four asymmetric N-acetylglucosamine units which gives a density in close agreement with the observed experimental value of 1.42 g/ml. The structure was arrived at by a trial and error method, with the aid of an optical diffractometer for testing the diffraction patterns of the various possibilities suggested by stereochemical considerations. The structure finally arrived at, is in good agreement with infra-red absorption data, and has no short contacts. The chains are separated by a distance of 4.69 Å perpendicular to the "plane" of the sugar rings and consequently the NH and CO groups in the neighbouring aminoacetyl side chains are hydrogen bonded, the plane of the amide groups in the "trans" configuration being almost perpendicular to the fibre axis. The hydroxymethyl side chains containing a hydroxyl group are also hydrogen bonded, the OH of one being linked to the oxygen of a similar group from a neighbouring chain running in the opposite direction. The short hydroxyl group is intrahydrogen bonded to the oxygen of the amide group in the same asymmetric unit. The calculated intensities for the proposed set of coordinates are in fairly good agreement with observed X-ray intensities visually estimated from fibre photographs.

INTRODUCTION

Chitin, like cellulose, is a straight chain polysaccharide and is a condensation product of N-acetylglucosamine with the formula $[C_8H_{13}O_5N]_n$; it occurs both as well oriented fibres, as in the mandibular tendon of lobster, and in the form of sheets, as in the cuticles of arthropods. From X-ray diffraction photographs of lobster tendon MEYER AND PANKOW[1] deduced that the unit cell is orthorhombic, with $a = 9.40$ Å, $b = 10.46$ Å (fibre repeat) and $c = 19.25$ Å. Observing odd order meridional reflections in their normal beam diffraction photographs, they rejected the possibility of a two-fold screw axis along the fibre axis. Further, from the absence of a first order reflection of 19.25 Å, and also the possible absence in general of $00l$ with l odd, they were led to the two possible space groups, namely $P222_1$, and $P2_122_1$ for the structure. Of these, MEYER AND PANKOW[1] favoured the latter, which was also confirmed by

DARMON AND RUDALL[2] from their x-ray studies on doubly-oriented cuticles with the $\{00l\}$ planes lying nearly parallel to the surface, which definitely revealed the absence of the 100 reflection of spacing 9.40 Å.

Postulating an analogy between cellulose and chitin, and using the known chemical configuration of N-acetylglucosamine, MEYER AND MARK[3]—also MEYER AND WEHRLI[4]—arrived at the spatial formula for chitin in which the pyranose rings are linked by a 1,4 type of bonding (Fig. 1). This was confirmed by the work of BERGMANN, ZERVAS AND SILBERWEIT[5]. MEYER AND PANKOW[1] arranged the chains thus built up in their lattice so that the chains were separated by about 9.6 Å along *c*

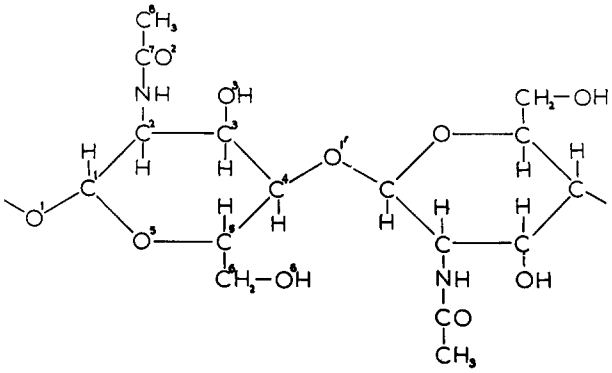

Fig. 1. The spatial configuration of side groups in chitin showing the two pyranose rings which constitute the repeating unit along the fibre axis.

in the direction of the plane of the sugar rings, and by about 4.65 Å along *a* in the direction perpendicular to them. With this arrangement of the chains conforming to the space group $P2_122_1$ there are 4 chains, and therefore 8 pyranose rings, contained in the unit cell. However, the space group requires only 4 equivalent points in the general position. To overcome this difficulty MEYER AND PANKOW[1] introduced the concept of the chitobiose unit, analogous to the cellobiose unit in cellulose, in which they postulated that the two adjacent pyranose rings along a chain, which together formed a fibre period of 10.46 Å along the *b* axis, were similar but not crystallographically identical in that they do not pass into each other by any symmetry operation. Then, the four chitobiose units which occurred in the unit cell of the above dimensions led to a calculated density value of 1.4 g/ml which agrees well with the measured value of 1.415 g/ml.

RUDALL[6] suggested that due to the presence of the amino-acetyl side chains of chitin, C=O . . . H–N hydrogen bonds could be formed of the type common in proteins. In attempting to make hydrogen bonds of this kind between amide groups of adjacent chains, RUDALL[6] found two possible different types of arrangement depending on whether the amide groups were in the *trans* or *cis* configuration. He then postulated that both configurations were present in the structure depending upon the positioning of the side groups with respect to the 2-fold and screw axis elements in the lattice. RUDALL[6] believed that this was supported by the interpretation of the infra-red absorption data obtained by DARMON AND SUTHERLAND[7]. This does not, however, seem to be necessarily true (see DISCUSSION (b)).

It is commonly found in most high polymers occurring in a crystalline form that

the residues or units in a chain take up crystallographically equivalent positions, either on an integral or non-integral helix. It was therefore rather surprising that the two pyranose rings in the chitin chain were not equivalent crystallographically in MEYER AND PANKOW's structure[1]. It is of course not unusual for the unit cell of a crystal to contain more than one set of equivalent groups as in this case (8 units, forming two sets of four equivalent units). However, there appeared to be a *prima facie* case for a re-examination of the structure owing to this feature, and further scrutiny has justified the doubt. Thus, the absence of a 2-fold screw axis along the fibre (*b*) axis was deduced from the occurrence of weak horizontal streaks along the meridian at positions corresponding to the odd-order reflections. This deduction is, however, not valid, for such streaks joining nearly non-meridional reflections are found in various fibres which, however, do not have an intensity peak at the exact meridian. Also, as will be shown in the next section, the 030 reflection can be proved to be non-meridional. In fact, it is found that a unit cell with half the above size is sufficient to index all the observed reflections and in this case there are just 4 equivalent groups in the orthorhombic unit cell. The details are given in the next section. The full structure has also been worked out on the basis of this new unit cell and it is discussed in the later sections.

UNIT CELL AND SPACE GROUP DETERMINATION

X-ray diffraction photographs of fibrous chitin, obtained from the mandibular tendon of a lobster, were taken with a micro-camera with a very fine collimator,

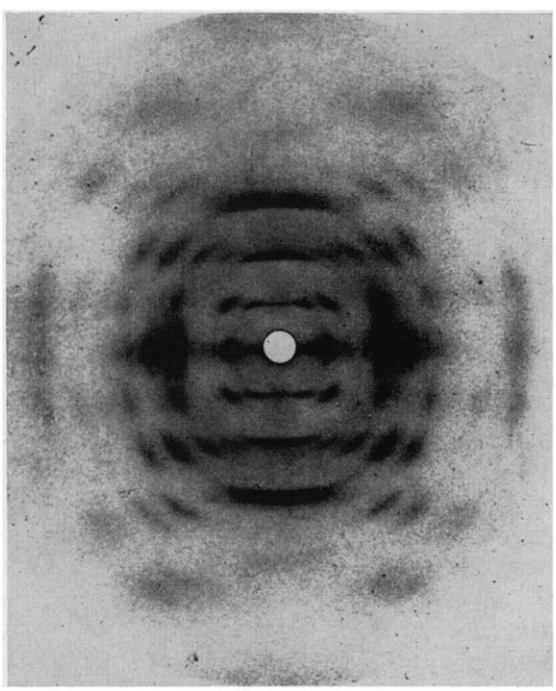

Fig. 2. X-ray diffraction photograph of chitin (enlarged to correspond to D = 4 cm) with X-ray beam normal to fibre axis.

using Ni filtered Cu $K\alpha$ radiation and with a fibre-film distance of 1.5 cm (flat film technique). This is shown in Fig. 2. The fibre was purified by treating it in the cold with dilute caustic potash, and then with dilute hydrochloric acid, each for about half an hour. It was then washed thoroughly in distilled water and dried with ether.

The Bragg spacings d of all the observed reflections were accurately measured for all the layers and are tabulated in Table I. From the spacings, the values of $(d^*)^2 = \lambda^2/d^2$ were calculated for all reflections. The fibre repeat c was measured as a mean of the layer-line spacings of the various observed layers and led to a value of $c = 10.43$ Å. Using this, the ξ^2 value of each reflection was calculated from the formula $\xi^2 = (d^*)^2 - (l\lambda/c)^2$. All these values are tabulated in Table I.

By a method of trial and error the smallest possible unit cell was deduced. As mentioned above, the spacings of the layer lines gave directly the fibre repeat $c = 10.43$ Å. The other two spacings a and b were obtained by trial and error from the

TABLE I

DATA OBTAINED FROM MICRO-CAMERA PHOTOGRAPHS

	Bragg spacing d(Å)	*Observed d^{*2}*	*Observed ξ^2*	*Calculated ξ^2*	*Indices h k l*	*Observed intensity I_0*
Equator	9.542	0.026	0.026	0.0260	020	19
	4.595	0.113	0.113	0.1145	110	100
	4.257	0.131	0.131	0.1340	120	0.5
	3.761	0.168	0.168	0.1665	130	12.5
	3.328	0.215	0.215	0.2120	140 (060)*	2
	2.311	0.445	0.445	0.4385	210 (200, 220)	12
First layer	9.071	0.029	0.007	0.0065	011	1
	6.940	0.049	0.027	0.0260	021	4
	4.257	0.131	0.109	0.1080	101 (041, 111)	7
	3.914	0.155	0.133	0.1340	121	1
	3.183	0.235	0.213	0.2120	141	2.5
	2.820	0.299	0.277	0.2705	151 (071)	1
	2.544	0.367	0.345	0.3420	161	2.5
	2.288	0.454	0.432	0.4320	201 (211, 171)	6
Second layer	5.056	0.093	arc	—	002 (012)	12
	4.514	0.117	0.026	0.0260	022	0.5
	3.970	0.151	0.060	0.0585	032	4
	3.445	0.200	0.110	0.1080	102	6
	3.031	0.259	0.170	0.1665	132	5
	2.580	0.357	0.269	0.2705	152 (072)	1.5
	2.157	0.511	0.423	0.4265	172 (082)	1
Third layer	3.358	0.211	arc	—	013**	50
	3.157	0.239	0.028	0.0260	023	3
	2.772	0.310	0.113	0.1145	113 (103)	6
	2.544	0.367	0.170	0.1665	133 (053)	6
	2.237	0.475	0.278	0.2705	153 (063)	1.5
Fourth layer	2.237	0.475	0.109	0.1080	104 (114)	2.5
	2.133	0.523	0.173	0.1665	134 (054)	1

* Other possible indices for the reflections are also given in brackets.
** Was split into two non-meridional reflections in the inclined fibre photograph.

equation $\xi^2 = h^2a^{*2} + k^2b^{*2}$. The task was comparatively easy, since trial values of a and b were taken to be $b = 19.25$ Å, corresponding to the spacing of the structure of MEYER AND PANKOW[1], and $a = 4.7$ Å which is half of MEYER AND PANKOW's value for a. The final dimensions were $a = 4.69$ Å, $b = 19.13$ Å and $c = 10.43$ Å (fibre repeat). (Note the change in designation of the axes from that adopted by MEYER AND PANKOW[1].)

There is, however, one very faint reflection present in the general equatorial streak in the diffraction photographs with spacing $d = 5.51$ Å, which does not fit in with this unit cell (or with the one given by MEYER AND PANKOW[1]). Consequently it has been omitted, as it is probably due to some radiation artifact or some impurity in the fibre used, which cannot be easily removed. Incidentally, this very weak reflection has not been observed by MEYER AND PANKOW.

Thus the unit cell is half that given by MEYER AND PANKOW. Consequently the space-group has also to be redetermined. There seems to be clear evidence for the presence of a two-fold screw axis parallel to the fibre axis. This means that odd order meridional reflections should be absent. Now the first order meridional reflection is very, very faint and is in fact a long streak bridging the $\{011\}$ reflections on either side of the meridian. This type of streak does not indicate the existence of a meridional reflection unless there is an intensification at the meridian, for it could arise from the occurrence of a cylindrical lattice (as is found in the 9-Å layer of collagen) or from disorder, etc. Actually, no intensification was found on the meridian even in a tilted fibre photograph.

Similarly, there is a very strong meridional arc on the third layer in the normal beam picture, which has to be explained. A tilted fibre X-ray diffraction photograph was therefore taken with the fibre inclined at the equi-inclination angle of the third layer, namely 13°. The photograph (Fig. 3) clearly showed that the arc is split up into two non-meridional reflections which, however, still overlap, with the indices 013. Similar tilted fibre photographs taken for the fourth and second layer equi-inclination angles revealed strong intensification of the meridional reflections 004 and 002. No fifth order meridional reflection 005 was observed in the former. Thus reflections of the type $00l$ with l odd seem to be absent and hence there must be a screw axis parallel to the fibre axis.

Considering next the other two 2-fold axes parallel to a and b, it is seen from Table I that among $h00$ only 200 probably occurs, while among reflections of the type $0k0$, 020 and 060 are recorded while 010, 030 and 050 seem to be absent. It is therefore reasonable to assume that there is a 2-fold screw axis along $b = 19.1$ Å, but along a the evidence is not clear since reflections with $h > 2$ are beyond the range of the observed diffraction pattern. Thus, the two possible space groups are $P22_12_1$, and $P2_12_12_1$, all space-groups containing reflection or glide planes being impossible since the compound is asymmetric.

The choice between the two space-groups was, however, readily made by considering the possible structural configurations. It will be seen that in the case of the space-group $P22_12_1$ (Fig. 4(a)), the chains running in opposite directions would be in the same plane parallel to the 19.1-Å axis, and also have the plane of their pyranose rings in this plane as well. As a result the side-chains belonging to the adjacent chains would either have to abut or interlock with each other. When the longer side-chains abut there will definitely be undesirable short contacts. When they are made to

interlock the orientation of the chains are such that in no position do they give rise to suitable cross-chain hydrogen bonding between hydroxymethyl groups of adjacent chains parallel to the plane of the pyranose rings, as is obtained in Fig. 4(c) with the space group $P2_12_12_1$. Even if the chains are rotated about the screw axes parallel to them no suitable configuration is obtained for cross-chain hydrogen bonding. This cannot be achieved with the space group $P22_12_1$ as the pyranose rings are separated by 4.7 Å in a direction perpendicular to the plane of the rings. In these respects, the space group $P2_12_12_1$ is superior in that there is a position in which adjacent

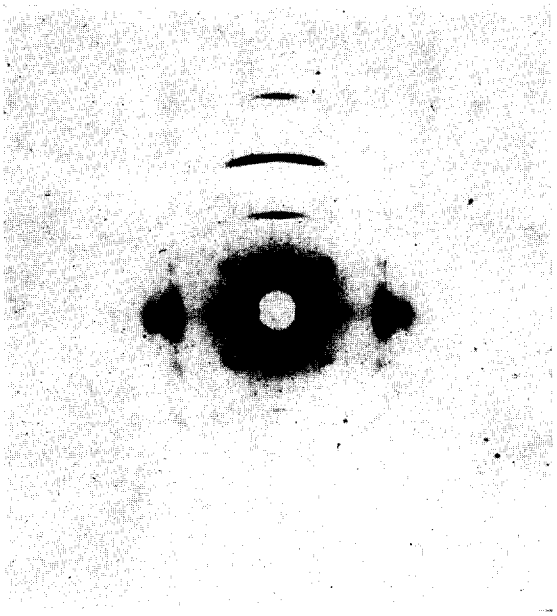

Fig. 3. X-ray diffraction photograph of chitin, with fibre inclined through the equi-inclination angle for the third layer namely 13° from the normal position of Fig. 2. Note the splitting of the meridional arc in the third layer and the occurrence of a meridional spot in the fourth layer.

chains can be readily hydrogen bonded (as we shall see in a later section) to give a structure that is stereochemically very satisfactory. This comes about because of the fact that along the 19.1-Å direction, adjacent chains are not in the same plane but are staggered by a half-unit translation along the *a* axis. Finally it may be said that the absence of the 100 reflection is consistent with the space group $P2_12_12_1$ although this alone would not demand the space group assigned.

Consequently, chitin possesses an orthorhombic structure with cell parameters $a = 4.70$ Å, $b = 19.1$ Å and $c = 10.4_3$ Å (fibre period), and with the space group $P2_12_12_1$, having four molecules per unit cell.

ESTIMATION OF INTENSITIES OF REFLECTIONS

To determine the intensities of the various reflections, X-ray fibre photographs of chitin were taken using a cylindrical camera of radius 3 cm, and Ni-filtered Cu $K\alpha$ radiation. The usual multiple film technique was adopted, although only two films

were used, and the exposure was made with the fibre stationary. The intensities of the reflections on both films were then visually estimated with the aid of a set of standard spots which were recorded using the intense reflection 020 from the chitin fibre itself.

To obtain the integrated intensity of each reflection, the observed value of the peak intensity was multiplied by half the spread of the reflection along its corresponding powder ring.

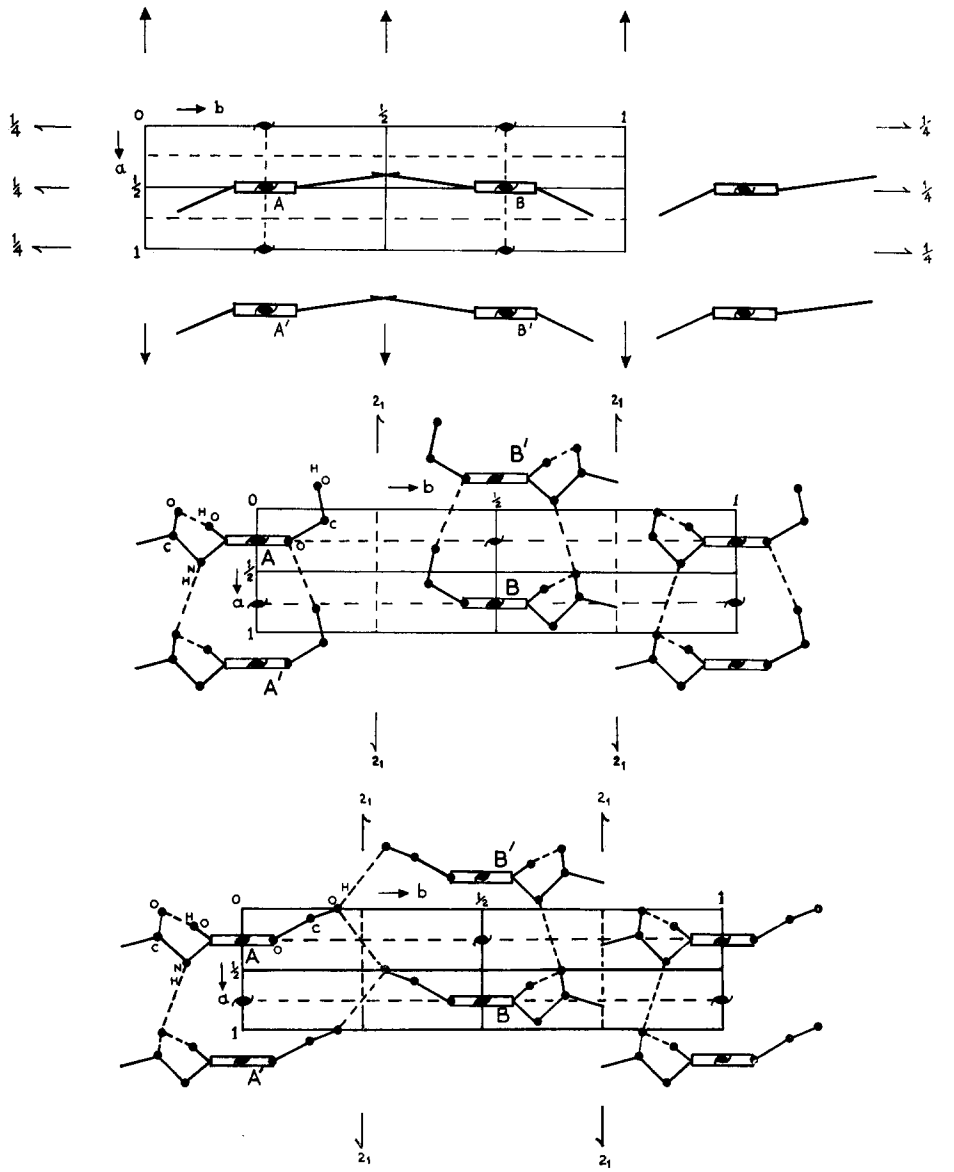

Fig. 4(a). A schematic *c*-projection of one half of the unit cell along the *c*-axis in the case of the space group $P22_12_1$; (b) A schematic *c*-projection of one half of the unit cell along the *c*-axis in the case of the space group $P2_12_12_1$ and with the parameter $u = 0$; (c) A schematic *c*-projection of one half of the unit cell along the *c*-axis in the case of the space group $P2_12_12_1$ and with the parameter $u = -0.25$.

These integrated intensities were then corrected by the Lorentz and polarization factors, which together are given for a general level by the expression

$$\frac{1 + \mathrm{Cos}^2 2\theta}{\mathrm{Sin}\ 2\theta} \cdot \frac{\mathrm{Cos}\ \theta}{[\mathrm{Sin}^2 \alpha - \mathrm{Sin}^2 \theta]^{1/2}}$$

where, α is the angle between the normal to the reflecting planes ($h\,k\,l$) and the fibre axis, and θ is the Bragg angle.

These corrected intensities were then suitably scaled to make the strongest observed reflection, namely 110 with $d = 4.59$ Å, equal to 100. These final values of the observed intensities I_0 are given in Table III.

DETERMINATION OF THE STRUCTURE

Models of the chitin chains were built up to scale in conformity with the chemical formula of chitin, and with the 1,4 type of pyranose ring linkage through an oxygen atom. The bond distances and bond angles were made in accordance with standard values, while the directions of some of the bonds coming out of the ring as well as the "chair" shaped configuration of the ring itself were maintained to be nearly the same as those found in the structure of glucosamine hydrobromide worked out by COX AND JEFFREY[8]. The side chains were attached as indicated by the chemical formula, while the orientation of the longer side chain containing the amide group was arranged such that the amide group held in the *trans* configuration was planar, and with its plane perpendicular to the fibre axis. This planarity of the amide group is now pretty well established, while the particular orientation was adopted since the infra-red absorption data of DARMON AND RUDALL[2] indicate it. The hydroxymethyl or shorter side group was not rigidly fixed but was allowed to assume any convenient orientation. The chains, thus built up, were then set up in the orthorhombic unit cell so as to satisfy the symmetry of the space group $P2_12_12_1$. The chains themselves were kept with their length along the screw axes 2_1 parallel to the c axis, which is required by its chemical structure (Fig. 1). Then, the other two 2-fold screw axes pass midway between the chains.

Fig. 4(b) shows schematically the projection of one half of the unit cell along the c axis. There are two chains in the unit cell running in opposite directions. If the chain nearer the origin of the lattice (A in Fig. 4(b)) is assumed to be coming up, the other near the middle of the cell (marked B) is going down. The chains A and A′ separated along the a-axis are similar, and are a unit translation apart. It was then possible to bond the amide groups of such adjacent chains, the N–H of one chain bonding with the C=O of a similar chain, the hydrogen bond distance being 2.8 Å, while the associated bond angle is within the usual range. The O–H group attached to the ring directly could not be bonded to any group in another adjacent chain. However, it could be internally hydrogen bonded with the oxygen atom of the amide group attached to the same chain, the O–H . . .O=C bond distance being within the permitted limit of 2.8 Å.

The bonding of the hydroxymethyl group was examined next. Now since the chains running in opposite directions parallel with the c axis have to coincide with the screw axes, the number of possible orientations for each chain is limited. The two possible degrees of freedom permitted for each chain are (1) a relative displace-

ment along the screw axis parallel to the c axis, and (2) a (slight) rotation of the chains about this axis. The latter is limited since the a-parameter is short and a large rotation would cause the atoms of the long side-chains containing the amide groups to approach each other and cause short contacts. Consequently, rotations of the two chains were not considered and only the effect of displacements of the chains along the screw axes were examined in detail.

Initially the oxygen atom connecting two pyranose rings, namely O'_1, was adjusted to be in a special position, namely in the ab plane containing the origin. For convenience let this position be represented by a parameter $u = 0$, expressed as a fractional coordinate of the translation of the atom O'_1 along the positive c axis. Having thus fixed the position of chain A, the position of chain B was fixed by the operation of the screw axis parallel to the a axis. It was then found that the hydroxymethyl group of chain A could not hydrogen bond with the hydroxymethyl group of chain B. Bonding of this group was then tried between similar chains A and A', and it was found that it was possible to bond the hydroxymethyl group of chain A' with the oxygen atom in the ring of chain A as shown in Fig. 4(b). The O–H . . .O hydrogen bond distance was slightly too large, but it could be brought to a reasonable value by slightly distorting some bond angles. This arrangement was stereochemically not very satisfactory since unusual hydrogen bonding was encountered, and further

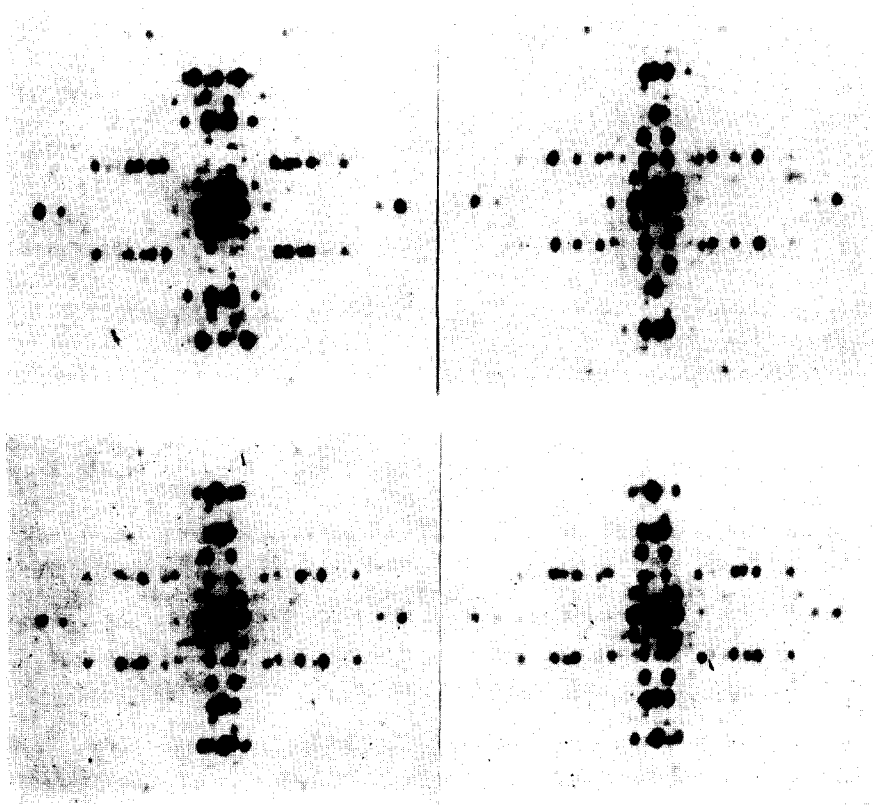

Fig. 5. Optical transforms of the a-projection of four unit cells of the structure having the parametres (a) $u = -0.08$; (b) $u = -0.14$; (c) $u = -0.16$; (d) $u = -0.18$.

there was no stabilizing cross-hydrogen bonding between chains A and B. However, the coordinates of all the carbon, nitrogen and oxygen atoms in one asymmetric unit were obtained as accurately as possible from the model. The intensities of all the reflections within the observed sin θ range were calculated and compared with the observed intensities.

Fairly good agreement was obtained for the zero level, but for higher levels the agreement was far from good as large discrepancies between calculated and observed intensities were found. This indicated that in projection the structure was almost correct, *i.e.*, that the coordinates along the fibre axis were evidently in error while the x and y coordinates were almost correct. The correctness of the projection indicated that no rotation of the chains about the fibre axis was required and that a shift along the fibre axis was necessary for getting good agreement with higher layer data.

This shift was then applied to the chain A, the oxygen atom O'_1 being moved down along the negative direction of c by 0.25 of the translation c. This configuration is represented by the parameter $u = -0.25$, and is the maximum shift which must be considered, all possible configurations being contained in the range 0.25 u. Bonding between the hydroxymethyl groups of chain A and A' with the hydroxymethyl group in chain B now became possible, the type of bonding being as shown in Fig. 4(c), where the O–H . . .O hydrogen bonds are inclined at about 45° to the planes of the sugar rings. The hydrogen bond angles are better than before though the bond distance is slightly large, around 2.9 Å. This structure is more satisfactory than the former in that cross hydrogen bonding between all chains is now possible. The intensities of all reflections within the observed range of sin θ were again calculated and checked against the observed intensities. More satisfactory agreement was found for some levels though the agreement was not complete.

This indicated that the z coordinates were still in error, and a change in the parameter u was required. The parameter was therefore varied continuously between $u = 0$ and $u = -0{,}25$ and the condition for the most satisfactory hydrogen bonding of hydroxymethyl groups of adjacent chains was examined. It was found that hydrogen bonding was possible within fairly wide limits of translation of the chains, from about $u = -0.08$ to $u = -0.25$, with the best hydrogen bonding around $u = -0.14$. It was then decided to try out four different configurations represented by the parameters $u = -0.08$, $u = -0.14$, $u = -0.16$ and $u = -0.18$. Values of $|u| > 0.18$ were not considered as $u = -0.25$ did not give very satisfactory agreement with observed data.

In order to reduce the amount of work, optical transforms of the a-projection of the four different structures were taken. These were obtained using a diffractometer constructed in the laboratory. In this way all the general reflections hkl could not be obtained, but only reflections of the type okl. However, these were sufficient to decide the correct position of the chains from among the four configurations decided upon. This was because most of the low angle reflections observed were of the type okl, and also because many of them were in the high intensity group with clearly-marked ratios among their various individual intensities.

Four a-projection masks were prepared for the four different structures mentioned earlier; the optical transforms corresponding to the four values of u are shown in Fig. 5.

The transforms obtained with the parameters $u = -0.08$ and $u = -0.18$

indicated ratios among the intensities which were rather different from those in the x-ray diffraction photographs. For instance in Fig. 5(a) with the chains having the parameter $u = -0.08$, the third layer was very weak in the transform, whereas a strong x-ray reflection 013 is actually recorded. In addition to this the second and fourth order meridional reflections were also absent in the transform indicating that the structure was considerably in error. In the transform obtained with the parameter $u = -0.18$ (Fig. 5(d)) the general appearance was better in that the fourth order meridional reflection was observed, and also the reflection 013 was not too weak. Here again the ratios of the intensities were not quite correct, the intensity of 014 being too high. This was again an indication that the structure was in error.

In the transform obtained with the chains having the parameter $u = -0.16$ (Fig. 5(c)) the general appearance was better than for $u = -0.18$, in that 004 was stronger while 014 had decreased. Again 013 was stronger than before though not as strong as it should be, and the magnitude of the ratios among other reflections had also improved indicating that the correct configuration was a little lower probably at $u = -0.14$.

The transform with $u = -0.14$ (Fig. 5(b)) gave the best results of all, the agreement with observed x-ray diffraction data being good for all levels. The 014 reflection had almost fallen to zero, while the magnitude of 013 was the highest recorded over the entire range studied. The ratio of 011 to 021 was also fairly near the experimental value. The actual magnitudes of the different ratios, were, however, not quite correct indicating that while the positions of the main chain in the lattice were roughly correct, the coordinates of the atoms in the side chains required slight alteration.

An excellent feature of this configuration with parameter $u = -0.14$ was that the O–H . . .O hydrogen bonds between the hydroxymethyl groups of adjacent chains were the shortest possible (2.8 Å), and further the O-H group was also directed almost straight to the oxygen atom of the adjacent group.

Before calculating the intensities for all reflections, the amide group was tilted very slightly so as to improve the internal hydrogen bond distance to 2.6 Å. An optical transform taken with this minor change indicated slight improvement thus confirming the shorter bond distance.

Refinement of the structure was stopped at this stage and the calculation of intensities for all reflections was carried out. Since the agreement for the zero level was found to be good, no attempt was made to rotate the chains about the screw axis. Thus in the *c*-projection the plane of the rings is perpendicular to the *a* axis. The final fractional coordinates of one asymmetric unit are given in Table II omitting the hydrogen atoms. These final coordinates are expected to be accurate to about $\pm$ 0.5 Å for atoms in the side-groups and much less for atoms in the pyranose ring. A more accurate estimate of the coordinates, as well as the deviation of the various atoms in the molecule from these coordinates is not possible at this stage. Further refinements in the coordinates is possible by adopting the least square refinement techniques, but this is not likely to alter the sterochemical configuration of the present molecular arrangement.

COMPARISON OF CALCULATED AND OBSERVED INTENSITIES

The intensities of all reflections up to the fourth layer were calculated, so as to cover the observed range of $\sin\theta$ using the standard formula for the space group $P2_12_12_1$

given in the International Tables for Crystallography and also the standard scattering factors for carbon, nitrogen and oxygen atoms given by VIERVOLL AND ÖGRIM[9]. In these calculations the contributions of the hydrogen atoms have been ignored. The calculated intensities were corrected for a temperature effect, by applying the usual temperature factor for intensities, namely $\exp -2B(\sin\theta/\lambda)^2$. An isotropic

TABLE II

DESCRIPTION OF THE STRUCTURE OF CHITIN; FRACTIONAL COORDINATES OF ATOMS IN ONE N-ACETYLGLUCOSAMINE RESIDUE

Unit Cell: Orthorhombic; $a = 4.69$ Å; $b = 19.13$ Å; $c = 10.43$ Å (fibre repeat). Space-group: $P2_12_12_1$. Set of equivalent points: x,y,z; $\frac{1}{2} - x, \bar{y}, \frac{1}{2} + z$; $\frac{1}{2} + x, \frac{1}{2} - y, \bar{z}$; $\bar{x}, \frac{1}{2} + y, \frac{1}{2} - z$.

Atom	*x*	*y*	*z*
C_1	0.324	0	0.232
C_2	0.223	— 0.063	0.190
C_3	0.276	— 0.063	0.046
C_4	0.164	0	— 0.002
C_5	0.250	0.063	0.046
C_6	0.084	0.136	0
C_7	0.170	— 0.175	0.201
C_8	0.255	— 0.251	0.254
N	0.383	— 0.120	0.206
O_1	0.250	0	0.360
O_2	— 0.044	— 0.167	0.179
O_3	0.155	— 0.107	— 0.077
O_5	0.250	0.063	0.190
O_6	0.250	0.200	0

temperature factor was assumed, and it was found that a value of $B = 1.5$ Å^2 agreed best with the observed decay of intensities with increasing Bragg angle. All the calculated intensities were then corrected for the temperature effect with this value, and these corrected intensities I_c are given in Table III.

The intensities I_c were then scaled so that the most intense reflection, namely 110 with spacing 4.59 Å, had a value 100. The other calculated intensities were similarly scaled, and these values I_c^* are also given in Table III.

From the $\sin\theta$ values in Table III it will be noticed that many calculated reflections *hkl* have Bragg angles quite close to one another. Since in x-ray diffraction photographs of fibres, the reflections are diffuse and have a definite spread of intensity, allowance has to be made for the possible overlap of reflections. Consequently, the intensities of overlapping reflections have been added together wherever it occurs. These total intensities have also been given in the column I_c^*, bracketing those reflections which have contributed to the sum. Comparison of calculated and observed intensities indicate good agreement for all levels.

However, a few near-meridional reflections are still found to have a larger observed intensity than the values calculated for them. This phenomenon must be attributed to the fact that the fibres have a small spread in their orientation with respect to the fibre axis. Since the Lorentz factor is very sensitive to small changes in ξ near the meridian (actually $\longrightarrow \infty$ for $\xi = 0$), the observed intensity will be larger,

TABLE III

CALCULATED AND OBSERVED INTENSITIES

	h k o	*Calculated sin θ*	I_c	I_c*		*Observed sin θ*	I_o
Zero layer	020	0.081	4790	21.0		0.081	19.0
	040	0.160	29	0			
	110	0.168	22802	100.0		0.168	100.0
	120	0.181	135	0.6		0.181	0.5
	130	0.205	481	2.1		0.205	12.5
	140	0.232	213	0.9			
	060	0.240	243	1.1	1.7	0.240	2.0
	150	0.259	128	0.6			
	160	0.291	68	0.3			
	080	0.321	0	0			
	170	0.325	78	0.3			
	200	0.329	3061	13.4			
	210	0.334	14	0.1	15.9	0.334	12.0
	220	0.339	548	2.4			
	230	0.350	104	0.5			
	180	0.360	12	0	2.9	0.365	2.5
	240	0.366	539	2.4			
	250	0.385	28	0.1			
	190	0.396	46	0.2			
	0100	0.401	52	0.2	1.1	0.408	1.5
	260	0.407	76	0.3			
	1100	0.429	82	0.4			
	270	0.432	0	0			
	280	0.459	21	0.1			
	290	0.488	19	0.1			

	h k 1	*Calculated sin θ*	I_c	I_c*		*Observed sin θ*	I_o
First layer	011	0.085	189	0.8		0.085	1.0
	021	0.111	1639	7.2		0.111	4.0
	031	0.141	491	2.2			
	041	0.177	1772	7.8			
	101	0.181	7	0	12.2	0.181	7.0
	111	0.185	998	4.4			
	121	0.197	372	1.6		0.198	1.0
	051	0.214	63	0.3			
	131	0.217	263	1.2			
	141	0.242	356	1.6		0.242	2.5
	061	0.252	69	0.3			
	151	0.273	4	0			
	071	0.290	177	0.8		0.280	1.0
	161	0.303	524	2.3		0.305	2.5
	081	0.329	194	0.9			
	171	0.333	130	0.6			
	201	0.337	187	0.8	4.0	0.338	6.0
	211	0.340	235	1.1			
	221	0.347	131	0.6			
	231	0.358	59	0.3			
	181	0.368	188	0.8		0.368	1.0
	091	0.368	7	0			
	241	0.373	21	0.1			
	251	0.392	19	0.1			
	191	0.403	33	0.2	0.4	0.403	1.0
	0101	0.407	13	0.1			
	1101	0.439	3	0			

TABLE III (continued)

	h k 2	*Calculated sin θ*	I_c	I_{c^*}		*Observed sin θ*	I_o
Second layer	002	0.148	197	0.9	4.3	0.148	12.0
	012	0.153	777	3.4			
	022	0.168	27	0.1		0.170	0.5
	032	0.191	709	3.1		0.194	4.0
	042	0.218	72	0.3			
	102	0.221	136	0.6	1.5	0.224	6.0
	112	0.225	32	0.2			
	122	0.236	97	0.4			
	052	0.249	56	0.2			
	132	0.252	381	1.7	3.0	0.255	5.0
	142	0.273	251	1.1			
	062	0.282	12	0.1			
	152	0.299	68	0.3			
	072	0.317	417	1.8		0.310	1.5
	162	0.327	314	1.4		0.330	1.5
	082	0.353	12	0.1			
	172	0.357	240	1.1		0.357	1.0
	202	0.361	11	0			
	212	0.363	8	0			
	222	0.370	142	0.6			
	232	0.380	16	0.1			
	182	0.389	483	2.1	6.1	0.390	4.0
	092	0.390	743	3.3			
	242	0.395	23	0.1			
	192	0.423	14	0.1			
	0102	0.427	99	0.4			
	1102	0.458	119	0.5			

	h k 3	*Calculated sin θ*	I_c	I_{c^*}		*Observed sin θ*	I_o
Third layer	013	0.226	2831	12.4		0.226	50.0
	023	0.236	1095	4.8		0.240	3.0
	033	0.252	37	0.2			
	043	0.274	12	0.1			
	103	0.276	54	0.2			
	113	0.279	453	2.0	3.1	0.278	6.0
	123	0.288	183	0.8			
	053	0.299	882	3.9			
	133	0.301	104	0.5	6.0	0.303	6.0
	143	0.319	365	1.6			
	063	0.327	205	0.9			
	153	0.341	69	0.3	1.8	0.345	1.5
	073	0.358	145	0.6			
	163	0.366	9	0			
	083	0.390	10	0			
	173	0.394	127	0.6			
	203	0.397	118	0.5			

even after correcting for the Lorentz factor corresponding to the mean value of ξ. If allowance is made for this fact, the agreement between observed and calculated intensities can be considered to be fairly good even for these reflections.

TABLE III (continued)

	$h\,k\,4$	Calculated $\sin\theta$	I_c	I_{c^*}		Observed $\sin\theta$	I_o
Fourth layer	004	0.296	1266	5.6		*	
	014	0.299	892	3.9		0.297	4.0
	024	0.307	157	0.7			
	034	0.319	636	2.8		0.320	1.5
	044	0.336	24	0.1			
	104	0.339	322	1.4			
	114	0.341	223	1.0	2.9	0.340	2.5
	124	0.348	112	0.5			
	054	0.357	64	0.3			
	134	0.359	100	0.4	1.2	0.361	1.0
	144	0.375	114	0.5			
	064	0.381	2	0			

* Not observed due to fibre orientation.

DISCUSSION

(a) *Description of the structure*

The structure in its final form consists of two chains running in opposite directions with screw axes lying along the chain directions. There are four asymmetric N-acetylglucosamine residues per unit cell. The 1,4 type of bonding through an oxygen atom links the adjacent rings, which are "chair" shaped while the "planes" of the rings are perpendicular to the short repeat $a = 4.69$ Å. Fig. 6 represents the a-projection of the final structure, while Fig. 7 represents the c-projection with only half the unit cell along the fibre axis c projected. That is, it is a projection of just two asymmetric units which are approximately at the same height and belong to adjacent chains running in opposite directions. From Figs. 6 and 7 it can be seen that the plane of the amide group is almost parallel with the ab-plane, and is very slightly inclined to it.

A list of bond distances and associated bond angles are given in Table IV for one asymmetric unit of the structure. The bond distances are expected to be correct to about $\pm$ 0.05 Å, while the bond angles may not deviate by more than $\pm 3^\circ$ from the listed values. The table also includes the bond distances and bond angles of the three hydrogen bonds present per single asymmetric unit of the structure.

(b) *Comparison with infra-red data*

It is interesting to note that the orientation of the amide group in the proposed structure is in good agreement with the infra-red absorption data published by DARMON AND RUDALL[2]. These workers found absorption frequencies indicating that the C=O and NH groups participate in intergroup hydrogen bonding of the type N–H . . . O=C, and that the strength of these hydrogen bonds were comparable with those in polyamides and proteins. They found that the N–H stretching band at 3265 cm^{-1} and the C=O stretching band at 1656 cm^{-1} both exhibit perpendicular dichroism, as also the N–H deformation band at 1560 cm^{-1}. Now the latter band can be shown to arise from the deformation of the N–H bond in the plane of the amide group, and it is thus apparent from the above data that the plane of the amide group as well as the C=O and N–H bonds themselves should be roughly

perpendicular to the chain axis. The hydrogenic stretching absorption frequency has also probably overlapped with the N–H stretching frequency around 3250 cm^{-1} to give strong perpendicular dichroism in this vicinity.

The second feature of interest in the structure is the internal hydrogen bond O–H . . . O=C which, as can be seen from Figs. 6 and 7, is approximately parallel to the fibre axis. DARMON AND RUDALL[2] have observed an absorption band at 3485 cm^{-1} which has parallel dichroism. This band has been attributed by them to be a weakly bonded O–H vibration. This frequency or the one at 3445 cm^{-1} (the latter more likely) can thus be associated with the hydrogenic stretching frequency of a weakly bonded O–H . . . O=C hydrogen bond with parallel dichroism. The band at 2750 cm^{-1} may also be attributed to this bond. Thus the internal hydrogen bond is also in agreement with infra-red data.

The third feature of interest in the structure is the lateral O–H . . . O hydrogen bond between the hydroxymethyl groups of the shorter side-chains of main adjacent chains running in opposite directions. From the projections given in Fig. 9 and 10 it is clearly seen that this hydrogen bond is perpendicular to the fibre axis. Now

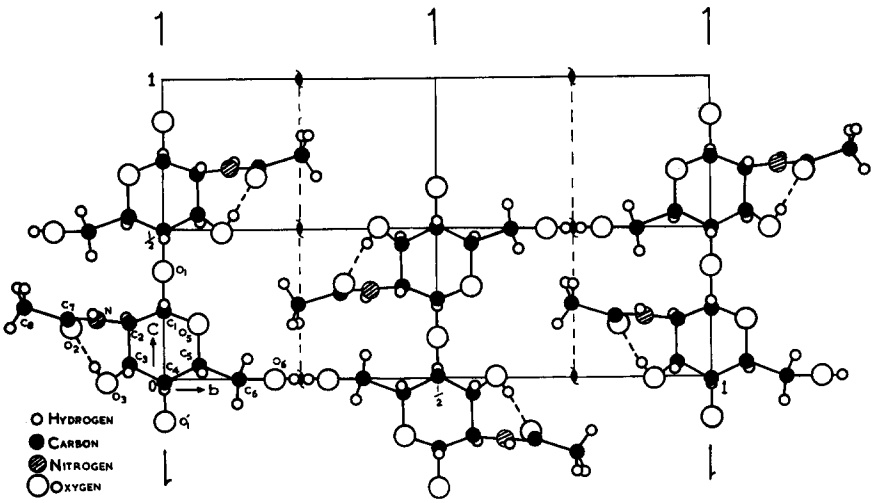

Fig. 6. The *a*-projection of the final structure.

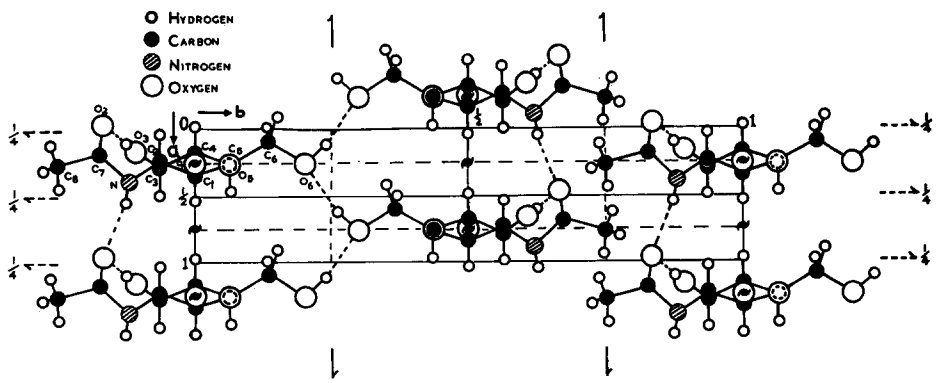

Fig. 7. The *c*-projection (fibre) of one half of the unit cell along the *c*-axis of the final structure.

TABLE IV

BOND DISTANCES AND BOND ANGLES

Bond	*Bond distance (Å)*	*Atoms in bond*	*Angle (degrees)*
C_1–C_2	1.48	O_5–C_1–O_1	112
C_2–C_3	1.54	C_2–C_1–O_1	112
C_3–C_4	1.48	O_5–C_1–C_2	112
C_4–C_5	1.47	C_1–C_2–C_3	112
C_5–O_5	1.50	C_2–C_3–C_4	112
C_5–C_3	1.58	C_3–C_4–C_5	112
C_3–O_3	1.30	C_4–C_5–O_5	111
C_6–O_6	1.50	C_5–O_5–C_1	116
O_5–C_1	1.47	O_5–C_5–C_6	109
O_1–C_1	1.48	C_4–C_5–C_6	109
C_2–N	1.42	C_5–C_6–O_6	118
N–C_7	1.47	C_4–C_3–O_3	109
C_7–C_8	1.54	C_2–C_3–O_3	112
C_7=O_2	1.10	C_3–C_2–N	107
		C_1–C_2–N	109
		C_2–N–C_7	111
		N–C_7–C_8	128
		N–C_7=O_2	120
		C_8–C_7=O_2	112
Hydrogen bond distance		*Hydrogen bond angles*	
N–H . . . O_2'	2.82	C_7–N (H) . . . O_2'	115
O_3–H . . . O_2	2.58	C_6–O_6(H) . . . O_6'	97
O_6–H . . . O_6'	2.80	C_3–O_3(H) . . . O_2	97

DARMON AND RUDALL[2] have observed an absorption band at 3100 cm^{-1} with strong perpendicular dichroism. But they have attributed this to a second N–H stretching mode. This, however, is very unlikely as the frequency is far below the usual values assigned to the N–H stretching mode. On the other hand, it is conveniently within the normal range of an O–H . . .O hydrogenic stretching frequency. From the magnitude of this absorption frequency it is indicated that these cross hydrogen bonds have a fairly high bond energy, around 7 kcal/mole. Thus the O–H . . . O hydrogen bond perpendicular to the fibre axis is also in agreement with infra-red data.

The CH_2 group in this hydroxymethyl side-chain is oriented such that the CH_2 plane is roughly parallel to the fibre axis (Figs. 6 and 7), with each of the C–H bonds equally inclined to the *ab* plane. In other words the bisector of the H–C–H angle of this group is in the *ab* plane and is perpendicular to the fibre axis. Consequently, the symmetric stretching frequency of the CH_2 group should show perpendicular dichroism. The antisymmetric stretching frequency should show parallel dichroism, and the wagging frequency of the CH_2 plane should show perpendicular dichroism. Now according to TSUBOI[10] the symmetric frequency of the CH_2 group in polysaccharides (cellulose) is about 2850 cm^{-1}, the antisymmetric stretching frequency for the CH_2 group is about 2960 cm^{-1}, the CH_2 wagging frequency around 1250 cm^{-1} and the CH_2 rocking frequency around 1030 cm^{-1}. An examination of these bands (assuming them to be roughly the same for chitin) in the absorption curve provided by DARMON AND RUDALL[2] reveals parallel dichroism at 2960 cm^{-1} indicating that the antisymmetric stretching frequency is parallel to the fibre axis; perpendicular dichroism at 2840 cm^{-1} indicating that the symmetric stretching frequency of the

CH_2 group is perpendicular to the fibre axis; and perpendicular dichroism at 1270 cm^{-1} indicating that the wagging frequency of the CH_2 plane is perpendicular to the fibre axis. Also strong parallel dichroism at 1030 cm^{-1} in the spectrum indicates that the in-plane rocking frequency of the CH_2 plane is parallel to the fibre axis. This very close agreement of the infra-red absorption data for the various modes of vibration of the CH_2 group thus confirms the orientation of the hydroxymethyl side-group containing the CH_2 group. A complete list of infra-red absorption data is given in Table V.

TABLE V

INFRA-RED ABSORPTION DATA ACCORDING TO DARMON AND RUDALL

Absorption frequency (cm^{-1})	*Observed dichroism*	*Dichroism expected from present structure*	*Assignment*
3485	//	//	O–H stretching
3445	//	//	O–H . . . O=C hydrogenic stretching
3265	⊥	⊥	N–H stretching
3100	⊥	⊥	O–H . . . O hydrogenic stretching
2960	//	//	CH_2 antisymmetric stretching
2840	⊥	⊥	CH_2 symmetric stretching
1656	⊥	⊥	C=O stretching
1560	⊥	⊥	N–H in plane deformation
1270	⊥	⊥	CH_2 wagging frequency
1030	//	//	CH_2 rocking frequency

The earlier structure proposed by MEYER AND PANKOW[1] had no stabilizing hydrogen bonds along the 19.25 Å axis, and at the same time it left all hydroxyl groups unbonded. Both these conditions are stereochemically undesirable. In both respects the present structure is more satisfactory. Further it is in fairly good agreement both with x-ray diffraction data and infra-red absorption data. The most interesting aspect of this structure determination is the fact that the configuration which from the accepted criteria of stereochemistry may be said to be the most satisfactory, is the one which is actually taken up.

The results of this investigation confirm the unit cell parameters and the orthorhombic space group $P2_12_12_1$ proposed by CARLSTRÖM[11]. However, they have led to a straight chain structure, instead of the buckled structure adopted by CARLSTRÖM, and also to somewhat different types of hydrogen bonds. We feel that the present structure should be considered to be more satisfactory for the following reasons.

In the structure proposed by CARLSTRÖM the chains are buckled, with the result that the hydroxymethyl side-groups cannot be hydrogen bonded, the distance between the oxygen atoms of these side groups in adjacent chains being 3.3 Å. This renders the structure unstable since the chains are unbonded along the long 19 Å axis. On the other hand, in the attempt to make this hydrogen bond as strong as possible, we found that the relative disposition of the two chains in the unit cell thus obtained also gave the best agreement with the X-ray diffraction pattern.

Now the infra-red absorption spectrum of DARMON AND RUDALL[2] indicates no absorption band at 3640 cm^{-1} corresponding to the unbonded O–H stretching frequency, but only bands at lower wave numbers indicating that the hydroxyl groups in the structure are all hydrogen bonded. Now if we assume that due to hydrogen bonding the O–H stretching frequency has shifted to lower wave numbers, and has overlapped either with the true N–H stretching frequency observed at

3265 cm^{-1} or with its strongly dichroic companion at 3100 cm^{-1}, both of which show perpendicular dichroism, then lateral hydrogen bonding between adjacent chains becomes essential, the hydrogen bonds being perpendicular to the fibre axis.

Now the reason for the buckling of the chains in CARLSTRÖM's structure seems to be three-fold. In the first place, bonding is achieved between the short hydroxyl group of one asymmetric residue and the ring oxygen of the next. If this bond really exists, it must be very strong to force the chains to buckle. But the small frequency shift from 3640 cm^{-1} to the observed band at 3485 cm^{-1} indicates that the hydrogen bonding is fairly weak. Furthermore, the fact that the ring oxygen is almost saturated and is not sufficiently negative indicates that this hydrogen bond if present would be rather weak. This is also confirmed by the fact that in the structure of sucrose by BEEVERS, MCDONALD *et al.*[12], the ring oxygen does not take part in hydrogen bonding, while in the structure of α-glucose by MCDONALD AND BEEVERS[13] the ring oxygen only bonds weakly with the hydrogen atom attached to the oxygen atom O_1. On the other hand the hydroxyl group can be conveniently hydrogen bonded to the oxygen of the amide group in the same asymmetric unit which is sufficiently negative to take part in double hydrogen bonding. This bonding is also parallel to the fibre axis and is rather weak, and thus accounts for the absorption band at 3485 cm^{-1} having parallel dichroism.

The second reason for buckling of the chains is the short contact between the nitrogen atom of one residue and the carbon atom C'_6 of the residue above. Now in the present straight chain structure, this interatomic distance is 3.1 Å, which is really not a short contact since a number of instances of such short interatomic distances between C and N atoms are known. This distance can in fact be increased to 3.4 Å by making slight alterations in the interatomic angles and thereby slightly changing the atomic coordinates of the various atoms in the side chain. This, however, calls for refinement of the structure which is reserved for further work.

Thirdly, buckling of chains was resorted to in order to remove the steric interference between the hydrogen atoms on the carbon atoms C_1 and C'_4. In the present straight chain structure the distance between these hydrogen atoms (see Fig. 6) is about 2.3 Å. A distance larger than 2.0 Å between the hydrogens seems to be generally permissible. In fact a van der Waals contact distance of 2.3 Å between the hydrogens of CH_2 groups occurs in the structure of diglycine hydrochloride (HAHN AND BUERGER[14]). Thus, no serious steric interference could be considered to exist in the straight chain configuration proposed in this paper.

ACKNOWLEDGEMENT

The author wishes to express his gratitude to Professor G. N. RAMACHANDRAN for his guidance, his many valuable suggestions and for the keen interest evinced by him in this work.

REFERENCES

1 K. H. MEYER AND G. W. PANKOW, *Helv. Chim. Acta*, 18 (1935) 589.
2 S. E. DARMON AND K. M. RUDALL, *Discussions Faraday Soc.*, 9 (1950) 251.
3 K. H. MEYER AND H. MARK, *Ber.*, 61 (1928) 1936.
4 K. H. MEYER AND H. WEHRLI, *Helv. Chim. Acta*, 20 (1937) 353.

[5] M. BERGMANN, L. ZERVAS AND E. SILBERWEIT, *Naturwissenschaften*, 19 (1931) 20.
[6] K. M. RUDALL, *Progr. Biophys. and Biophys. Chem.*, 1 (1950) 39.
[7] S. E. DARMON AND G. B. B. M. SUTHERLAND, *Nature*, 164 (1949) 440.
[8] E. G. COX AND C. A. JEFFREY, *Nature*, 143 (1939) 894.
[9] H. VIERVOLL AND O. ÖGRIM, *Acta Cryst.*, 2 (1949) 277.
[10] M. TSUBOI, *J. Polymer Sci.*, 25 (1957) 159.
[11] D. CARLSTRÖM, *J. Biophys. Biochem. Cytol.*, 3 (1957) 669.
[12] C. A. BEEVERS, T. R. R. MCDONALD, J. H. ROBERTSON AND F. STERN, *Acta Cryst.*, 5 (1952) 689.
[13] T. R. R. MCDONALD AND C. A. BEEVERS, *Acta Cryst.*, 5 (1952) 654.
[14] T. HAHN AND M. J. BUERGER, *Z. Krist.*, 108 (1957) 419.