X-Ray Diffraction Studies on the Structure of α-Chitin

M. Abdul Haleem and K. D. Parker

Astbury Department of Biophysics, University of Leeds

(Z. Naturforsch. **31 c**, 383—388 [1976]; received February 25/April 2, 1976)

Structure of α-Chitin

The structure for α-chitin originally proposed by Carlström is generally considered to be the true structure. However, it fails to account for several remaining problems of chitin structure which include a complete explanation of the infrared spectrum and also the different properties of α- and β-chitins. We have reexamined the structure by X-ray diffraction using automatic rigid subunit least-squares refinement and also the difference Fourier method. The *R*-value was reduced from 0.31 to 0.22 by a number of small modifications to the structure proposed by Carlström. Consideration of symmetry suggests that two distinct types of statistical modifications could be present in the structure, both of which would allow complete intersheet hydrogen bonding between O_6H groups within the general framework of Carlström structure. The X-ray results give some support to one of these modifications. However, it is predictable that both would affect the X-ray diffraction only marginally so that a clear proof lies beyond the present sensitivity of the method. The proposed extra hydrogen bonding in α-chitin provides the first clear explanation for the different properties of these kinds of chitin.

Introduction

Chitin, a polysaccharide of N-acetyl-D-glucosamine is widely distributed in nature[1]. It occurs in three forms namely α, β, γ[2]. The structure of α-chitin was proposed by Carlström[3] with orthorhombic unit cell of dimension $a = 4.76$ Å, $b = 10.28$ Å (fibre repeat) and $c = 18.85$ Å. Carlström[4] criticized the structure proposed by Dweltz[10] on stereochemical grounds.

The present work describes:

1. Extension and refinement of α-chitin[3] using more modern and powerful techniques now available.
2. Determination of coordinates representing the standard configuration of glucopyranose ring[5] and its peptide chain[6] in the unit cell[3] and determination of various positions of O_6 atom.
3. Refinement of structure by leastsquare method[14].

Several programs were written in Algol 60 code and KDF9 computer was used throughout this work.

Measurement of Intensities

The protein free and decalcified tendon fibre of about 1.5 cm length was mounted with durofix on a stainless U-shaped steel under tension. The X-ray diffraction pattern of the sample was studied using the camera designed by Langridge *et al.*[7]. Small pin holes were used in the camera in order to obtain sharply defined X-ray photographs. The X-ray beam so obtained was weak and exposure time was long. A Hilger and Watts microfocus tube was used. Nickel filtered CuK_a radiation of wave length 1.54 Å was used throughout this work. The diffraction pattern was recorded on Kodak X-ray film. The films were developed for 7 min in Kodak developer with agitation rinsed, fixed for 13 min and dried in air. Fibres were tilted at the appropriate angle to record a meridional reflection. An X-ray diffraction photograph is shown in Fig. 1 *. The intensities were measured by Joyce Loebl microdensitometer. The area of individual reflections was taken by tracing the reflection. All the reflections in four quadrants were traced. The mean area over the four quadrants was taken for each reflection. The arc length in reciprocal space can be given as $D' \gamma$ where D' = distance from the centre to the arc and γ = angular spread of the reflection. Therefore the area under the curve must be multiplied by D' to relate it to the integrated intensity. The intensities were calculated using the relationship

$$I_0 = A \times D'$$

where A = mean area under the curve,
D' = distance of the reflection from the centre of X-ray diffraction photograph.

The overlap of reflections were resolved graphically such that sum of ordinates of the resolved

---

Request of reprints should be sent to M. A. Haleem, Department of Biochemistry (Biophysics Unit), University of Karachi, *Karachi-32*, Pakistan.

* Fig. 1 see Plate on page 382 b.

peaks were equal to unresolved peak. The observed intensities were corrected by Lorentz polarization factors.

Determination of coordinates

Coordinates representing standard bond length and bond angles for glucopyranose ring[5], peptide group[6] and O_6 atom were determined in the unit cell for the structure proposed by Carlström[3]. The detailed mathematical derivation involved in this procedure are given elsewhere[13]. The coordinates determined are called R – C system.

Refinement by leastsquare method[14]

The detailed mathematical derivations are given elsewhere[13]. The following parameters were used.

$P1$ = Rotation of molecular chain about screw axis.
$P2$ = Shift of molecular chain along its length.
$P3$ = Rotation of O_6 atom about $C_5 - C_6$ bond (Φ').
$P4$ = Tempterature factor.
$P5$ = Scaling factor.
$P6$ = Rotation of peptide group about an axis through C_2 parallel to OX.
$P7$ = Rotation of $C_5 - C_6 \setminus O_6$ group about an axis through C_5 parallel to OX.
$P8$ = Rotation of peptide group about an axis through C_2 parallel to OY.
$P9$ = Rotation of $C_5C_6 \setminus O_6$ group about an axis through C_5 parallel to OY.

Φ[9, 14] and R-value[8–9] were calculated after each cycle. The observed values of the composite reflections can be divided in proportion according to the calculated values. If $|Fc|_i^2$ represents the calculated structure factor for the ith reflection of the group containing I' reflections, I_0 represents the observed intensity, then the intensity according to the calculated structure factors can be given by

$$|Fc|_i^2 \times I_0 / \sum_{I'} |Fc|_i^2 .$$

If $U_1, U_2 \ldots U_n$ are the initial values of the parameters and $\Delta U_1 \ldots \Delta U_n$ correction to be applied. If U_m is a good approximation then $U'_m = U_m + \Delta U_m$ ($m = 1 \ldots n$) can be taken as better approximation. Because the series has been truncated by neglecting second and higher terms in ΔU_m[14], the calculations must be repeated using these approximate values. For example

$$\text{cycle I: } U_1' = U_1 + \Delta U_1 , \quad \text{cycle II: } U_1'' = U_1' + \Delta U_2$$

and so on. The iterative process is complete when there is no significant change in the parameters in the next successive cycles. At this stage the value of Φ can be taken as a minimum value. ΔU values of the parameters obtained in a cycle were used to calculate corrected values. The corrected values of the parameters were used in order to modify coordinates of R – C model for these new parameters. These coordinates were then used for calculations of the structure factors using isotropic temperature factor and scaling factor.

In the space group $P2_12_12_1$ the origin is normally chosen half-way between pairs of non intersecting twofold screw axis. The use of this origin complicates the expression for the electron density and structure factors. If the origin is taken on the two fold screw axis then the space group coordinates must be transformed into projection coordinates by substracting 1/4 from the fractional z components. This brings the origin on one of the screw axis parallel to a which means the 010 projection has a centre of symmetry. Using the new coordinates for the equivalent positions the electron density in the unit cell can be given by

$$\underset{X_0 Z_0}{\Delta\varrho = \Delta\varrho} = 1/A \sum_h \sum_l \Delta F \cos 2\pi \left(h X_0 - \frac{h-l}{4} \right) \cos 2\pi \left(l Z_0 - \frac{l-h}{4} \right)$$

where A = area of the unit cell,
$X_0 Z_0$ = are the new coordinates which represent crystalographic axes,
ΔF = difference between observed and calculated structure factors ($Fo - Fc$).

The calculated structure factors (Fc) were adjusted to the shifted unit cell by multiplication by $\exp 2\pi i l/4$. This has the obviously necessary effect of making all structure factors real. The maps were calculated at an interval of 1/10 in the X-axis direction and 1/40 in the Z-axis direction. Contours in the $\Delta\varrho$ map were drawn at an equal intervals. The molecule was plotted after translating in the Z-axis direction by $-1/4$.

Results

The values of the parameters after 84th and 127th cycle are given in Table I. The quantity Φ was reduced to 1.470. A difference synthesis map at $\Phi = 1.470$ was also plotted (Fig. 2). In this map scaling factor (K) was applied to Fc's. In the map there is a negative peak at (0, 4/40) and positive peak at (–6/10, 7/40). Moreover they are placed roughly in positions accesible to the O_6 atom. In order to account for the possibility of two coexisting positions of O_6 atoms it was decided to calculate structure factors with O_6 at these two positions separately and

Table I. Refinement by least-square method.

	*P*1	*P*2	*P*3	*P*4	*P*5	*P*6	*P*7	*P*8	*P*9	Φ	*R*
84th cycle	9.2°	−0.038 Å	135.7°	0.40883	0.01698	4.6°	−9.0°	−7.3°	4.7°	1.567	0.23
127th cycle	9.2°	−0.027 Å	143.3°	0.38268	0.01693	7.3°	−14.7°	−7.0°	3.1°	1.470	0.225

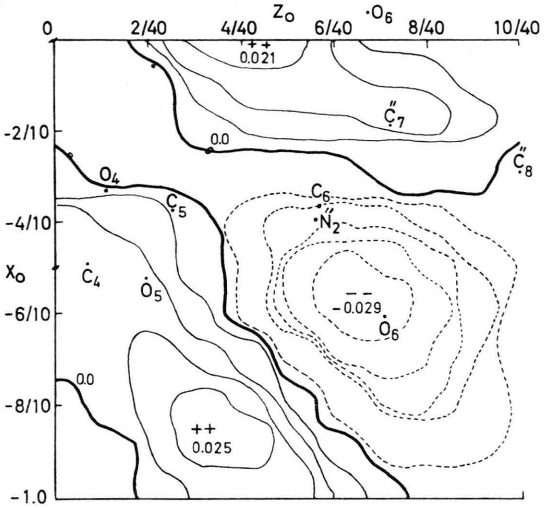

Fig. 2. Difference synthesis map (127th cycle, *K*-value = 0.0169; contour interval = 0.005/*K* e/Å²).

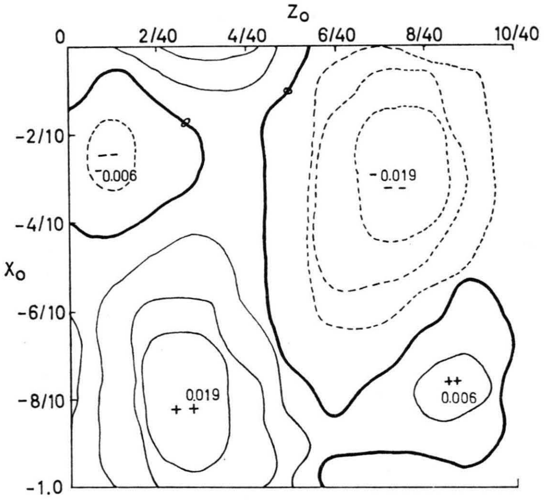

Fig. 3. Difference synthesis map (26th cycle, *K*-value = 0.0171, contour interval = 0.005/*K* e/Å²).

then to divide by 2. These two chains ($C_5-C_6\diagdown O_6$) were rotated about an axis parallel to x and parallel to y through C_5. The starting position of the O_6 atom of the first residue was rotated about C_5-C_6 by an angle 143.3° (previous value). The starting position of the second O_6 atom was found by rough geometry 300° to give the best position in the $\Delta\varrho$ plot. This value was used for further refinement by a new leastsquare program. The $\Delta\varrho$ map is shown in Fig. 3. The *R*-value of 0.22 ($\Phi=1.388$) is obtained after 26 cycles.

Discussions

The first object in the present calculations is to add information on the structure of α-chitin. Marchessault *et al.*[11] suggested the presence of a rotational isomer of α-chitin. The present calculations show that the calculated structure factors are not in general harmony with the observed values suggesting that the presence of this rotational isomer is insignificant. In view of this no further consideration has been made even though a small proportion of some specimens of α-chitin might possess this conformation. The *R*-value of 0.53 found for this structure is so high as to rule out from further consideration. Carlström[3] tried to find the suitable parameters in order to improve the optical transform. In contrast the present work used the least square method[14]. Seven geometrical parameters, scaling factor and temperature factor were used to refine the structure. The *R*-value reported by Ramakrishnan *et al.*[12] is 0.42 using 4 geometrical parameters. The *R*-value is reduced to 0.24 in 63 cycles (from the R−C model). The Φ-value after 63 cycles is found to be 1.566. Further refinement gives an *R*-value of 0.23 and $\Phi=1.47$ in 127 cycles. At this stage one parameter *i. e.* rotation of $C_5C_6\diagdown O_6$ about an axis through C_5 gives a large change of angle which may be unacceptable. However after 63 cycles this value is only −9°. The bond angle of $C_5C_6\diagdown O_6$ is found to be 123.9° (in cellobiose it is 115.3°)[15]. The difference synthesis map (Fig. 2) after 127 cycles of refinement shows significant improvement over R−C model particularly peak at the main sugar residue is reduced. However, the difference synthesis ($\Delta\varrho$) map after 127 cycles shows two approximately equal peaks of ($1.5\,\Delta e$) electrons and of opposite sign. The negative peak lies near the O_6 atom. The $\Delta\varrho$ map (Fig. 3) is still more improved by further refinement of dividing the O_6 atom between the two

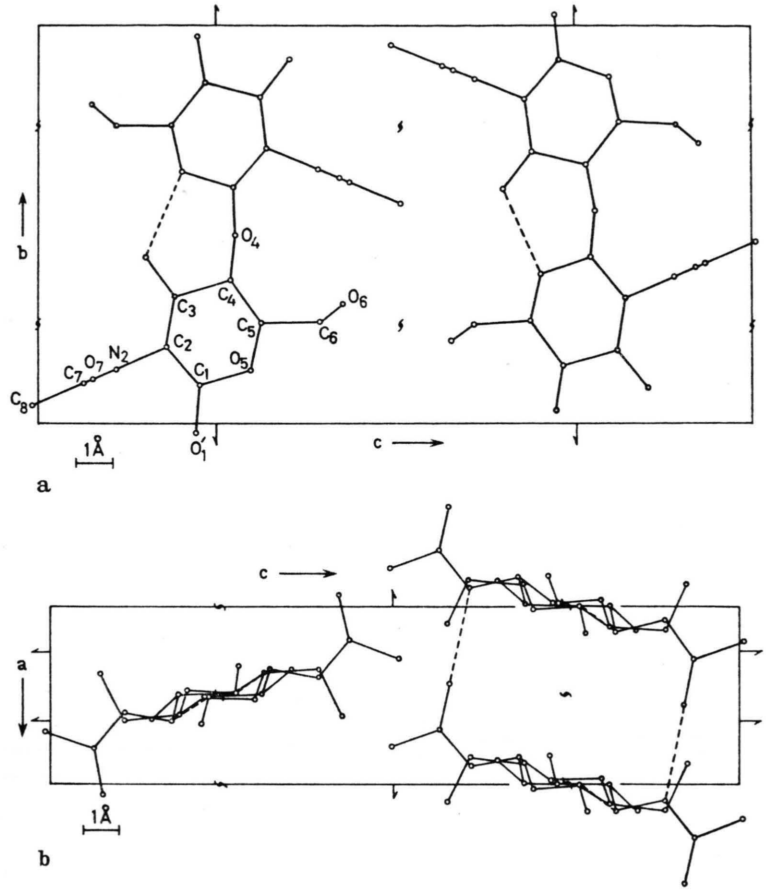

Fig. 4. The α-chitin structure (84th cycle). a. Projection on bc plane; b. projection on ac plane.

positions ($\Phi' = 143.3°$, $300°$). The *R*-value of 0.22 ($\Phi = 0.1388$) is obtained after 26 cycles.

These calculations suggest that the model using two positions of the O_6 atom is worth further consideration. This split structure is in agreement with the idea that the O_6 atoms are arranged somewhat at random in order to form intermolecular H-bonding between the chains. Moreover this is just the kind of structure required to make possible statistical structure of model closely examined.

The *R*-value for the structure after 84 cycles is 0.23 (Table I). At this stage there is no short contact between the peptide groups (residue II = $\bar{x}$, $\frac{1}{2}+y$, $\frac{1}{2}-z$ and residue III = $\frac{1}{2}-x$, $\bar{y}$, $\frac{1}{2}+z$). Interatomic distances were calculated for the residue II, III and IV ($1-x$, $\frac{1}{2}+y$, $\frac{1}{2}-z$). The distance between CH_3 (residue II) and O_3 (III) is found to be 3.41 Å (contact distance = 3.3 Å). The intra-atomic distance between O_5 and O_6 ($\Phi' = 135.7°$) is 2.95 Å (O–H H–O distance = 2.9 Å). It is worth noting that no atoms or groups appear to be in Van-der Waals contact in this structure. The closests is that above mentioned (C_8H_3(II) – O_3(III) of 3.41 Å interatomic). This is suggestive of a H-bond dominated structure. Fig. 4 represents the projection of the molecule after 84th cycle. The accuracy of the intensities measured by densitometer from fibre digram is not as good as that obtained from single crystals diffractogram. The variety of shapes and sizes of the photographic reflections makes it impossible to measure very accurately the integrated intensities of the reflections. The *R*-value reported here is for $P2_12_12_1$ symmetry. The *R*-value of 0.23 significantly better than 0.42 [12]. This gives only an over all structural fit. Details for individual parameters are not given with certainty, since large variations of any single parameter would not produce very large change in the *R*-value. α-chitin shows little interaction with water as indicated by Infrared and X-ray studies. This property is markedly different from that of β-chitin. Furthermore calculations of interatomic distances show that it is very difficult to introduce a water molecule between the chains. This agrees with Carlström's structure. In one of the structure proposed here the C_6O_6 groups are considerably removed from Carlström's position, so the problem requires reexamination. Since the IR spectrum shows that free OH groups are absent[11] we would expect to find H-bonds. Interchain bond between sheets will clearly give stability and the non swelling property. Examination shows that lateral H-bond can be formed with the R – C model (and also the refined models described in the present work). The positions selected are shown schematically in Fig. 5 for the projections of the structure. With the residues shown two bonding schemes are possible. These are energetically equivalent and may be described as O_6 in residue I (anti-Carlström) – O_6 in residue IV (*gauche-trans*)[16] or O_6 in dresidue I (*g t*)[16] – O_6 in residue IV (anti-Carlström). Fixing our attention on residue IV, the O_6 atom has a choice of adapting the *g t* (*gauche-trans*) configuration and bonding with residue I (which must take up anti-Carlström position) or it may take up the anti-Carlström position and bond with residue I (which must then be in *g t*). The reason for the choice of *g t* and anti-Carlström amongst a range of values which could possibly form a H-bond was that *g t* is favoured for a variety of reasons as a likely or favoured position for O_6 and although no evidence for the occurrence of anti-Carlström configuration has been suggested it looks good on skeletal models because, it has C_5 or C_6 in the mutually staggered configuration (N. B. Carlström position has $C_5C_6 \diagdown O_6$ in the relatively unfavoured eclipsed configurations).

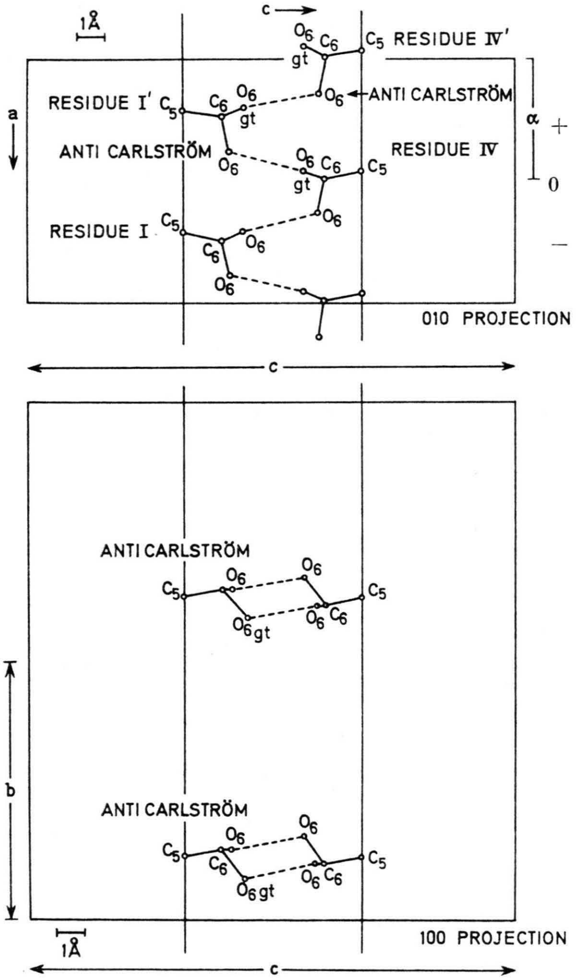

Fig. 5. Interlayer H-bonding in α-chitin. Positions selected are shown schematically for two projections of the structure. 1. O_6 in residue I (anti-Carlström) ... O_6 in residue IV (*gt*) *or* 2. O_6 residue I (*gt*)... O_6 in residue IV (anti-Carlström).

Having established a H-bond between say residue I in *g t* and IV in anti-Carlström what we say about the other O_6 atoms forming the intersheet region? Neighbouring groups must be effected. For instance O_6 on residue I′ (Fig. 5) is now excluded from forming a H-bond with IV. It can turn to IV′ and form a H-bond. A chain reaction *could* set in and we could postulate that all O_6 atoms in the sheet discussed would fall into this pattern. This is a tempting speculation because it is an immediate reason for the shift ± *e* Å (Fig. 6) between sheets. One has to suppose that *g t*/anti-Carlström configuration contains a slight defect in its H-bonds which can become better after shift ± *e* Å. Before we can accept this we must also extend the pattern to O_6 atoms which occur at a remove of *b* from

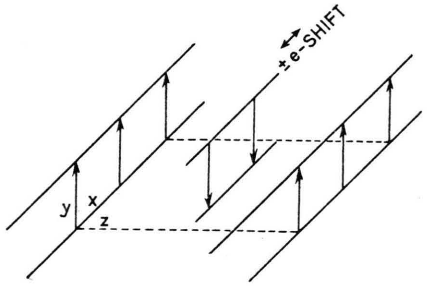

Fig. 6. ± e-shift. The chains in the second layer were placed at *a*/2 with respect to in the first layer the interlayer spacing being *c*/2. The chain direction was parallel to *oy* and the sheets to be formed parallel to the plane *xoy*. The shift is indicated in the diagram.

these already discussed. These are no longer restricted by the choice formed by group I and IV above and could take up the opposite choice. If so the tendency to shift by e would be in the opposite sense and cancel the effect of the first layer of O_6 atoms. It seems possible that the O_6 variability can be summed up by the hypothesis that there are two positions of the sheets for least energy respectively those with shift $+e$ and $-e$. If the sheets find themselves in either one or the other of these positions the H-bonds will snap down into a manner analogous to any bistable system.

There is one piece of direct experimental evidence on the X-ray photographs supporting the shift hypothesis. This is the greater breadth of certain reflections of the type 1 k l compared corresponding 0 k l. Turning to the least squares refinement of models in which two residues were allowed to vary independently we obtain further support for the hypothesis. The two residues model was introduced primarily to improve some residual density peaks on the Difference Fourier plots in which a certain success was achieved. However the final predictions by this procedure for the O_6 atoms of two independent residues may be compared with those predicted by model building alone.

Now gt has $\Phi' = 0$ and anti-Carlström has $\Phi = 110°$. However examinations of the model shows that H-bonds can be formed from the X-ray derived positions which are apparently equally good. In fact a whole range of angles is possible for the groups to form a satisfactory H-bond. The interatomic distances for the two possibilities discussed here have been calculated:

a. O_6 (gt) . . . 2.93 Å . . . O_6 (anti-Carlström);

b. O_6 ($\Phi' = 315°$) . . . 2.94 Å . . . O_6 ($\Phi' = 135.7°$).

Judging by the model alone it is estimated that the H-bond range for the angular variable Φ' of the O_6 atoms lies somewhere about.

Allowed for residue I'	*Allowed for residue IV*
60° (anti-Carlström 110°)	290°
135.7° – (least squares predicted)	315°
180°	gt

In view of the experimental errors and residual R-values obtained it is clear that no significance can be placed on the precise value of the Φ values quoted here from the least squares programs. At the most the results be regarded as pointer in support of the H-bonding rather than as proofs. This type of statistical model with this types of residue in a mixed structure could explain the observed frequencies at 1656 cm^{-1} and 1625 cm^{-1}. No detailed work has been made in this direction. For the statistical model more detailed calculations are required.

---

[1] Jo Conrad, Enc. Polym. Sci. **3**, 695 [1965].

[2] K. M. Rudall, Conformation of Biopolymers (G. N. Ramachandran, ed.), **Vol. 2**, p. 751, Academic Press, New York 1967.

[3] Diego Carlström, J. Biophys. Biochem. Cyt. **3**, 669 [1957].

[4] Diego Carlström, Biochim. Biophys. Acta **59**, 361 [1962].

[5] G. N. Ramachandran, C. Ramakrishnan and V. Sasiskharan, Aspects of Protein Structure (G. N. Ramachandran, ed.), p. 121, 1963.

[6] Tatsue Ooi, Roy A. Scott, Garret Vanderkooi, and Harold A. Scheraga, J. Chem. Phys. **46**, 11, 4410 [1967].

[7] R. Langridge, H. R. Wilson, C. W. Hooper, M. H. F. Wilkins, and L. D. Hamilton, J. Mol. Biol. **2**, 19 [1960].

[8] A. J. C. Wilson, Acta Crystallogr. **3**, 397 [1950].

[9] S. Arnott and A. J. Wonacott, J. Mol. Biol. **21**, 371 [1966].

[10] N. E. Dweltz, Biochim. Biophys. Acta **44**, 416 [1960].

[11] R. H. Marchessault, F. G. Pearson, and G. Y. Liang, Biochim. Biophys. Acta **45**, 499 [1960].

[12] C. Ramakrishanan and N. Prasad, Biochim. Biophys. Acta **261**, 123 [1972].

[13] M. A. Haleem, Ph. D. Thesis, University of Leeds 1971. 1971.

[14] S. Arnott and A. J. Wonacott, Polymer **7**, 157 [1966].

[15] R. A. Jacobson, J. A. Wunderlich, and W. N. Lipscomb, Acta Crystallogr. **14**, 598 [1961].

[16] M. Sundaralingam, Biopolymers **6**, 189 [1968].

M. Abdul Haleem and K. D. Parker, *X-Ray Diffraction Studies on the Structure of α-Chitin (p. 383)*

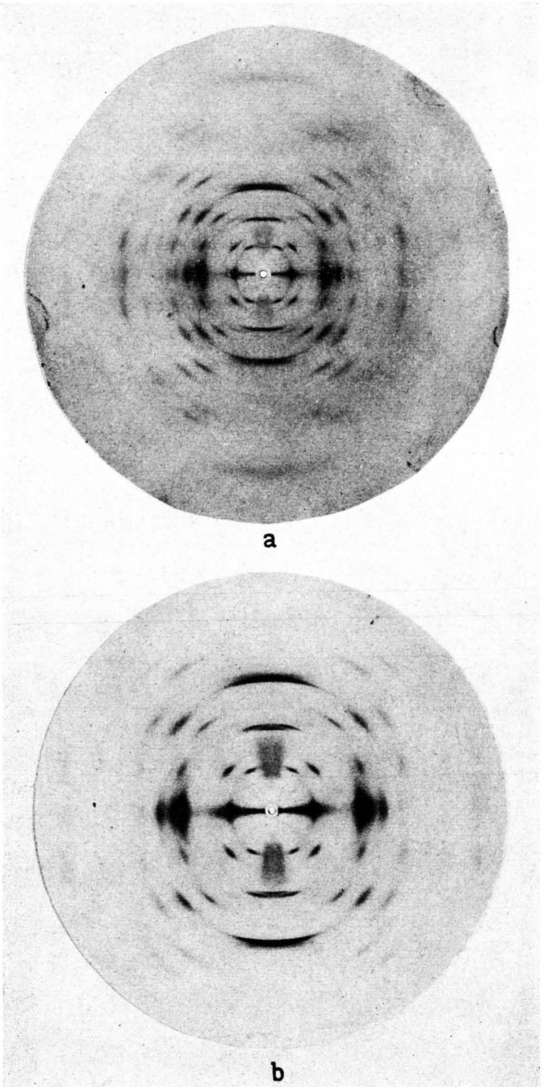

Fig. 1. X-ray diffraction photograph of α-chitin. (Protein free and decalcified tendon fibre from crab — Fibre normal to the X-ray beam). a. 0.15 mm lead apertures 5 cm apart, b. 0.25 mm lead apertures 5 cm apart.