

THE CRYSTAL STRUCTURE OF α -CHITIN
(POLY-*N*-ACETYL-D-GLUCOSAMINE)*

BY DIEGO CARLSTRÖM†, M.D.

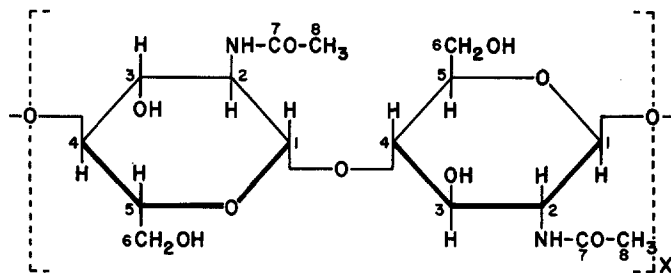
(From the Department of Biology, Massachusetts Institute of Technology, Cambridge)

(Received for publication, May 23, 1957)

INTRODUCTION

Chitin is an important polysaccharide of the animal kingdom, making up a major part of the skeletal and connective tissues of invertebrates belonging to several phyla, such as arthropods, molluscs, annelids, and brachiopods (see Rudall, 1955). Two crystallographic modifications, referred to as the α and the β forms, have been reported (Lotmar and Picken, 1950). The former is the common arthropod chitin, whereas the β modification has only been reported from a few sources. Among chitin-containing invertebrate tissues, the insect cuticles are by far the best studied, and it seems clear that in this case chitin and proteins together form a layered complex at the molecular level (Fraenkel and Rudall, 1947).

Chemically, chitin is a polymer of *N*-acetyl-D-glucosamine, the structural formula of which can be written as a multiple of two acetyl glucosamine residues (chitobiose) linked together by 1-4- β -glucosidic bonds.



It is thus closely related to cellulose, the only difference being that in chitin the aminoacetyl side groups replace the hydroxyl groups attached to the second carbon atoms of cellulose.

* This work was supported in part by Grant G-2007 from the National Science Foundation, under the supervision of Dr. Richard S. Bear.

† The author's permanent address is Department of Medical Physics, Karolinska Institutet, Stockholm, Sweden. The investigation was carried out during tenure of a Rockefeller Foundation Fellowship.

During an x-ray diffraction survey of chitin-containing invertebrate tissues it was found that the crystallographic data obtained were in disagreement with the commonly accepted structure of chitin proposed by Meyer and Pankow in 1935. A structure determination was therefore undertaken, since a detailed analysis should not only give a better basis for further ultrastructural work on invertebrate connective tissues but also have some importance for the understanding of the crystal structure of cellulose. It is of interest to note that previous crystal structure determinations on carbohydrates have been concerned with the simple sugars, glucose and sucrose, and the only polysaccharide analyzed in detail has been cellulose (Andress, 1929; Meyer and Misch, 1937).

Material and Methods

The best oriented and most detailed fiber patterns were obtained from the apodemes of the lobster (*Homarus americanus*), the chitin of which was purified by treatment in 10 per cent HCl for 3 hours in order to remove the calcium carbonate, and by boiling in 5 per cent KOH for 48 hours in order to get rid of the proteins making up about 60 per cent of the dry weight. After thorough washing the samples were dried under light tension in the fiber direction. The nitrogen content of the resulting material, as determined by micro-Kjeldahl methods, was 6.67 per cent or slightly less than the theoretical value (6.89 per cent). The water content of samples dried at room humidity was found to vary between 7 and 12 per cent, but no appreciable enlargement of the x-ray spacings was associated with this uptake of water. Therefore, in what follows the structure is developed as an anhydrous one. The density of vacuum-dried samples as determined by the flotation method was found to be $1.425 \text{ g} \cdot \text{cm}^{-3}$, in good agreement with earlier values (1.415, Meyer and Pankow, 1935; 1.40 to 1.42, Lotmar and Picken, 1950).

The fiber patterns were usually recorded in a cylindrical camera of 11.4 cm. diameter, with Ni-filtered Cu radiation. By varying the orientation of the fiber axis, reflections up to $37^\circ\theta$ could be observed. For accurate determination of the unit cell dimensions, suitable calibration substances (NaCl or Si) were deposited on the specimens.

The estimation of the intensities was made mainly by direct comparison of fiber patterns recorded at various exposure times (1 to 156 hours) under otherwise constant conditions. Allowance was made for the differences in reflection areas caused by arcing. After the usual corrections for polarization and Lorentz factors, the integrated relative intensities of the reflections selected for the Fourier syntheses of the (100) projection, were found to be in good agreement with corresponding corrected intensities obtained from powder patterns by microdensitometer recordings.

The structure determination was supported by use of optical diffraction, for which masks of the three main projections of each trial model were prepared by photographic procedures. The atomic scattering factors for C, N, and O were given the same weight and the intensities of the optical transforms obtained therefrom were visually compared with the observed x-ray intensities. A detailed description of the optical diffraction methods will be given elsewhere (Wyckoff *et al.*, 1957.)

The polarized infrared absorption in the region 650 cm^{-1} to $4,000 \text{ cm}^{-1}$ was made on well oriented sheets of purified material, 20 to 30μ thick. This part of the investigation was carried out by Dr. E. R. Blout, whose cooperation was greatly appreciated.

RESULTS

Determination of an Approximate Structure from Stereochemical Concepts and x-Ray Data

The Unit Cell Dimensions and Space Group.—The unit cell was found to be orthorhombic and the length of the axes (accuracy ± 0.03 Å) are tabulated below along with earlier published values.

(a)	(b)	(c)	Author
<i>A</i>	<i>A</i>	<i>A</i>	
11.58	10.44	19.42	Gonell, 1926
9.40	10.46	19.25	Meyer and Pankow, 1935
—	10.27	—	Lotmar and Picken, 1950
4.76	10.28	18.85	This work

Among the reflections listed by Meyer and Pankow, the only two having an odd h index, the faint (312) and (332) reflections, can be accounted for as (215) and (235), respectively, relative to their cell. This finding and the indexing of about 30 reflections not reported earlier, indicated the shorter length of the a axis here chosen. The interplanar spacings, relative intensities, and indices of all measured reflections are found in Table I, which also gives their $\sin^2\theta$ for comparison with the calculated $\sin^2\theta$ for all possible reflections within the observed range.

The space group could not be determined unequivocally from the x-ray diffraction patterns, since the a axis dimension is almost an exact even multiple of the c axis dimension. This prevented examination for systematic absences of the ($h00$) reflections. Extinctions of ($0k0$) and ($00l$), for k and l odd (the presence of a faint (010) reflection will be discussed later), suggested two sets of mutually perpendicular twofold screw axes. The space group could thus be either $P 2_12_1$ or $P 2_12_12_1$. Difficulties encountered in attempting to pack the polysaccharide chains in the unit cell according to $P 2_12_1$ and the final agreement between the derived model and x-ray data determined the choice of $P 2_12_12_1$.

The General Arrangement of the Molecular Chains.—The density calculated on the basis of four acetyl glucosamine residues per unit cell is $1.462 \text{ g}\cdot\text{cm}^{-3}$. Considering the fibrous texture of the material the agreement with the observed value ($1.425 \text{ g}\cdot\text{cm}^{-3}$) can be regarded as satisfactory. From chemical investigations it is well established that the pyranose rings in chitin are connected by 1-4- β -glucosidic linkages. The β linkage is also indicated by the infrared absorption, in which a peak at 892 cm^{-1} is seen. According to Barker *et al.* (1954) this band is characteristic of β -D-glucopyranose anomers. Since it is probable that the pyranose rings in chitin should have the Sachse C1 chair

form, found in D-glucose (Reeves, 1950; McDonald and Beevers, 1950, 1952) and glucosamine hydrobromide (Cox and Jeffrey, 1939), the 10.28 Å length of the fiber axis requires two chitobiose units in the unit cell. Either of the space groups given above demand that these chitobiose chains must run in opposite directions. The *a* axis dimension (4.76 Å) is close to that of the thickness of a pyranose residue, which indicates that the pyranose rings should be almost parallel to the (100) plane.

The Spatial Arrangement of the Pyranose Rings.—The arrangement of the pyranose rings forming the polysaccharide chains was derived by using the D-glucose configuration given by McDonald and Beevers (1950, 1952) to obtain the 10.28 Å repeat in accordance with the usual stereochemical requirements. Among several arrangements considered, the “straight” chain proposed by Meyer and Misch (1937) for the structure of native cellulose (see Fig. 1 *a*) was found to be sterically impossible (see discussion). Only one arrangement of the pyranose rings (Fig. 1 *b*) seems to have all the desired properties and in addition offers the possibility of formation of an intramolecular hydrogen bond.

The Aminoacetyl Side Groups and Chain Interrelations.—From the space available it is evident that the bulky aminoacetyl side group attached to the second carbon atom (*cf.* glucosamine hydrobromide, Cox and Jeffrey, 1939) also should determine the interrelation between the two chains in the unit cell. The configuration of the side group and the shift between the two chains had therefore to be considered simultaneously. (The shift between the chains is taken as the displacement in the *b* axis direction between an O_{1,4} in one chain and the closest O_{1,4} in the other.) The interatomic distances within each side group were taken as equal to those in *N*-acetyl glycine (Carpenter and Donohue, 1950). By varying the shape of the side group, space-filling models having a shift between the two chains from 0.4 Å to 3.2 Å could be constructed.

Optically derived Fourier transforms provided additional criteria for model selection. With this procedure it was found that only models having the pyranose rings nearly parallel to the (100) plane and almost planar aminoacetyl side groups perpendicular to the (100) plane (and accordingly a small shift between the chains, 0.4 to 1.1 Å) gave optical diffractograms having intensity distributions similar to the observed x-ray intensities.

Refinement of the Structure

The (100) and (001) projections were considered first, since the shift between the chains and smaller adjustments of the side groups should not appreciably alter the basal projection. A model having completely planar aminoacetyl side groups and a 0.5 Å shift between the chains was found to give the most satisfactory optical diffraction patterns. These could be further improved when the positions of the hydroxyl groups attached to the sixth carbon atoms were considered. By stepwise, small rotations of the chains around their fiber axis the

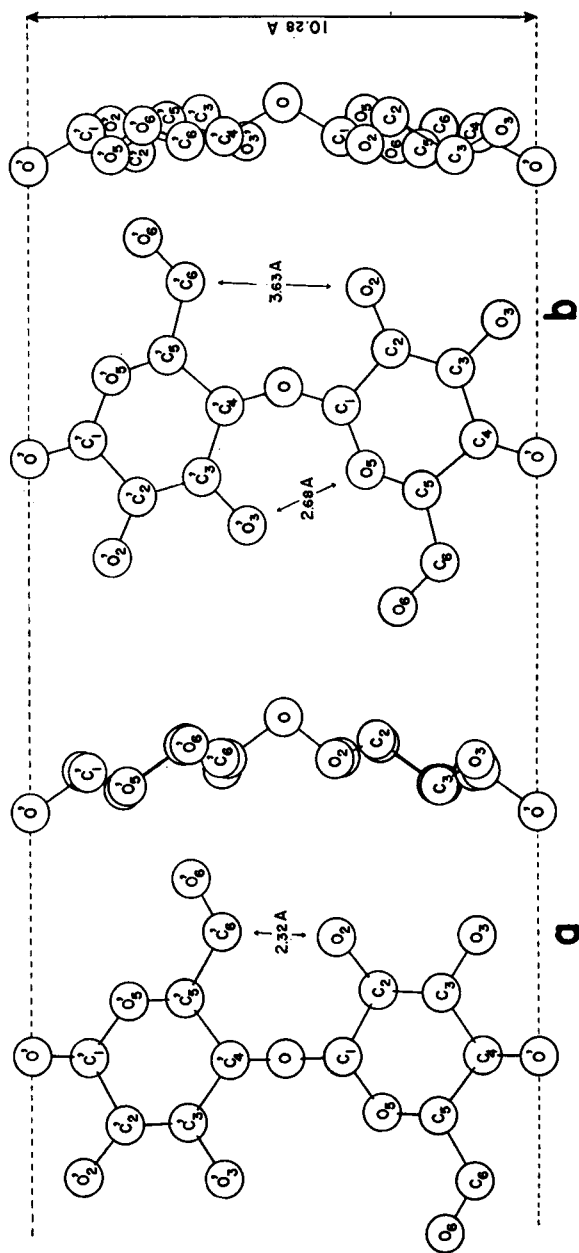


FIG. 1. Spatial arrangement of 1-4- β -linked glucopyranose rings as derived from the D-glucose structure and a repeating distance along the fiber axis of 10.28 Å. Fig. 1 a shows the "straight" chain configuration earlier proposed for cellulose, resulting in a very short O₂-C₆ distance. Fig. 1 b gives the glucopyranose chain here chosen for chitin, which not only has a reasonable O₂-C₆ separation but also a O₆-HO₃ distance suitable for an intramolecular hydrogen bond.

intensities of the optically derived Fourier transforms of the (010) projection could be brought into agreement with the observed intensities of the equatorial reflections on the x-ray diagrams.

The optical Fourier transforms of the three main projections of the best model can be seen in Fig. 2, which also gives the x-ray diffraction pattern for comparison.

A Fourier synthesis of the (100) projection, the only one in which atom overlap is not severe, was carried out optically according to Wycoff *et al.* (1957). The phases were determined by introducing an artificial "heavy atom" at one center of symmetry per cell, on the optical diffraction mask of the model. Because of superpositions of reflections occurring on the x-ray diffraction patterns, the observed amplitudes, F_{obs} , for some reflections had to be derived indirectly. The total intensity of an observed composite reflection was thus divided into parts roughly corresponding to the contribution of each single reflection, as derived from the Fourier transforms. In order to get the magnitude of the zero-order term, $F_{(000)}$, all F_{obs} values were put on an absolute scale by comparing them with F values calculated for 20 reflections. From the F_{obs} and their optically derived phases a mask of the weighted reciprocal lattice was prepared, with all F values increased so that the largest negative amplitude became zero (Fig. 3 *a*). The resulting optical synthesis ("image"), showing the distribution of the electron density, plus heavy peaks at one center of symmetry per unit cell, is seen in Fig. 3 *b*, where the (100) projection of the model is superimposed for comparison. The low resolution of the x-ray diffraction patterns of chitin (the cut-off occurring at spacings of about 1.8 Å) restricts the resolution of the Fourier synthesis. It was, therefore, not possible to obtain reliable indications for further refinements of the structure. However, the good general agreement between the model and the areas of high electron density was taken as a confirmation that the model was correct in its main features. A drawing of the chitin structure thus derived is given in Fig. 4, and the atomic parameters are given in Table II.

Infrared Absorption

The infrared data of the present studies are essentially in agreement with those already published by Darmon and Rudall (1950). However, in correlating their findings with x-ray data these authors ran into difficulties, since they adopted the then current structure of chitin. These difficulties are now resolved.

With better oriented specimens the polarized infrared absorption spectra show an appreciably higher dichroism than was observed earlier. In general, the region between 1800 and 4000 cm^{-1} corresponds with the spectra given by Darmon and Rudall. The strong doublet around 3430 cm^{-1} has parallel dichroism and indicates weak hydrogen-bond vibrations approximately parallel to the fiber axis. The earlier speculation that one of the maxima in this absorption

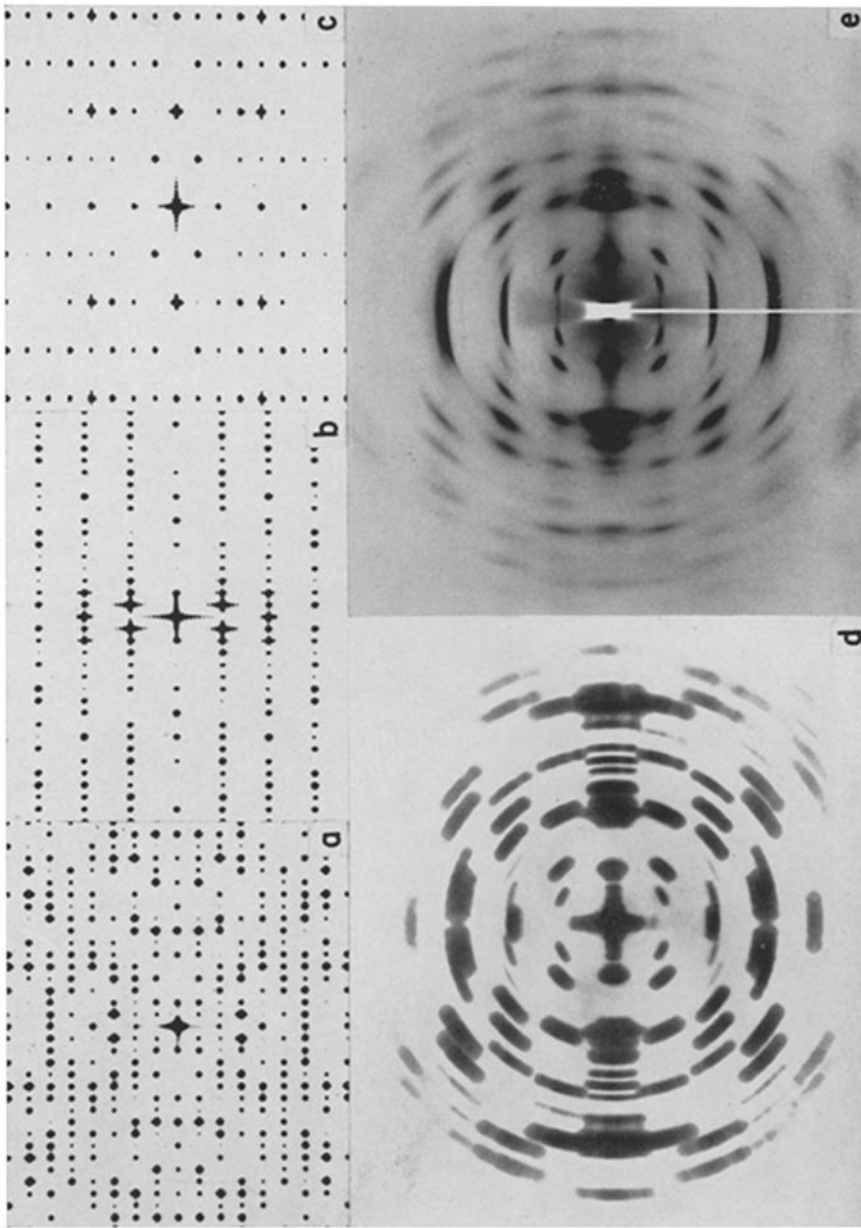


FIG. 2. Optically derived Fourier transforms (diffractograms) from the (100), (010), and (001) projections, (a), (b), and (c) respectively, of the best model of chitin. Fig. 2 d is the composite diffractogram of (a), (b), and (c), with the reflections from (b) transferred to the equator; the disorientation and the fading off have been introduced artificially in order to match the x-ray diffraction pattern, Fig. 2 e. The fiber axis is vertical, except for (b), where it is perpendicular to the plane of the figure.

band is due to free NH vibrations, finds no support from the structure here derived. Strong $\text{—CO}\cdots\text{HN}$ hydrogen bonds perpendicular to the fiber axis are confirmed by the perpendicular dichroism observed at 3260 and 3090 cm.^{-1} . With the new data (see Fig. 5, which shows the spectra from 1800 to 800 cm.^{-1}) there may be observed perpendicular dichroism of the C=O stretching frequencies at 1660 and 1625 cm.^{-1} (amide I), and weak perpendicular dichroism of the amide II band (NH deformation at 1560 cm.^{-1}). These indicate that the amide groups lie predominantly perpendicular to the fiber axis.

The main differences between the spectra reported here and those of Darmon and Rudall lie in the three bands at 1420, 1377, and 1325 cm.^{-1} , attributable to

TABLE II
Atomic Parameters

Atom	<i>x</i>	<i>y</i>	<i>z</i>
	<i>A</i>	<i>A</i>	<i>A</i>
C ₁	2.38	0.39	4.38
C ₂	2.84	1.92	3.29
C ₃	2.08	3.23	3.62
C ₄	2.38	3.70	5.04
C ₅	1.99	2.54	6.01
C ₆	2.26	2.91	7.49
C ₇	3.33	1.00	1.16
C ₈	2.90	0.43	18.63
O ₁	3.06	10.02	4.08
N ₂	2.42	1.38	2.00
O ₃	2.70	4.15	2.71
O ₄	1.70	4.88	5.34
O ₅	2.69	1.41	5.69
O ₆	1.08	2.81	8.29
O ₇	4.56	1.08	1.47

carbon-hydrogen deformation modes. These bands all show rather strong parallel dichroism indicating that such groups lie essentially perpendicular to the fiber axis. The strong parallel dichroism shown by the group of bands lying between 1000 and 1210 cm.^{-1} is in agreement with earlier findings and can be attributed to C—O frequencies. However, the previously stated fact that two of the absorption bands at around 1015 and 1060 cm.^{-1} show perpendicular dichroism cannot be confirmed.

DISCUSSION

The crystal structure of chitin proposed here is quite different from the earlier Mayer and Pankow model, although both models have in common an orthorhombic unit cell containing chitobiose units running in opposite directions, with the planes of the pyranose rings more or less perpendicular to the

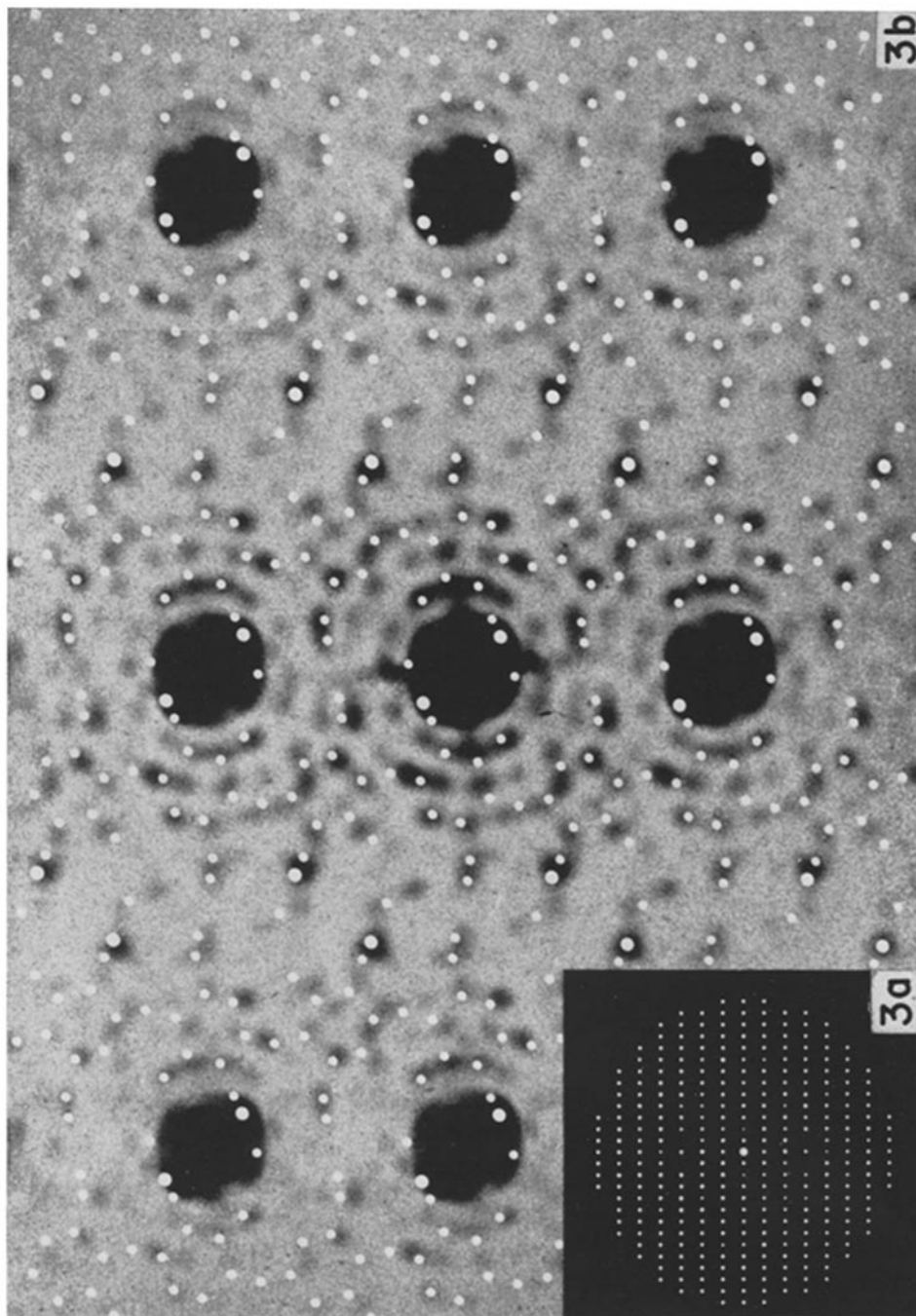
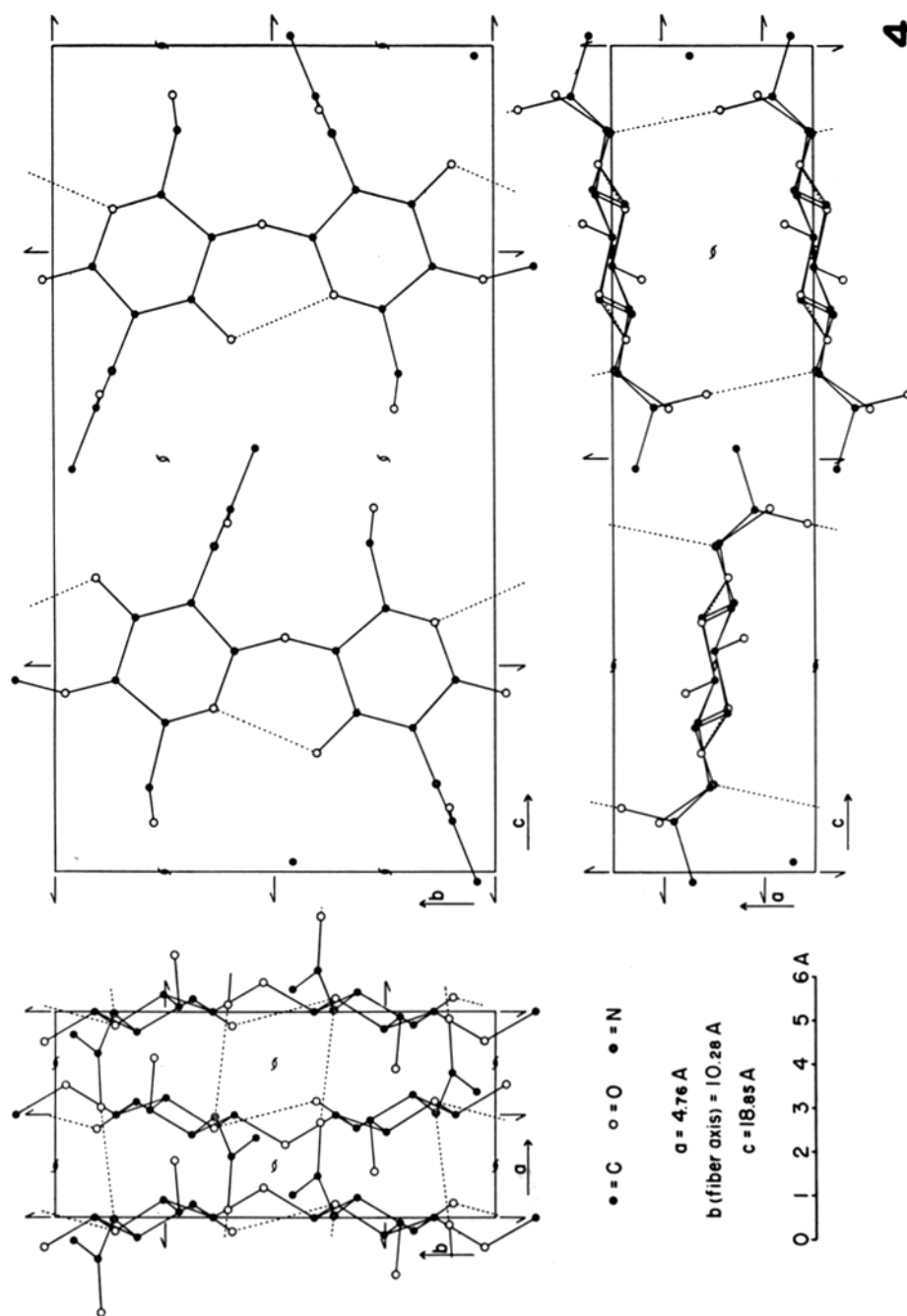


FIG. 3. An enlarged diffraction mask, representing the weighted reciprocal lattice of the (100) projection, is seen in Fig. 3 *a*. This mask was used for the production of an optical image (Fourier syntheses) showing the distribution of electron density, Fig. 3 *b*. The areas of high electron density (black) correspond fairly well with the superimposed (100) projection of the model (white dots). The large black spots at one center of symmetry per unit cell arise from the introduction of "heavy atoms" in these positions.



4

FIG. 4. Drawing of the three main projections of the unit cell. The origin is halfway between three sets of non-intersecting, mutually perpendicular, twofold screw axes. Hydrogen bonds are represented by dotted lines.

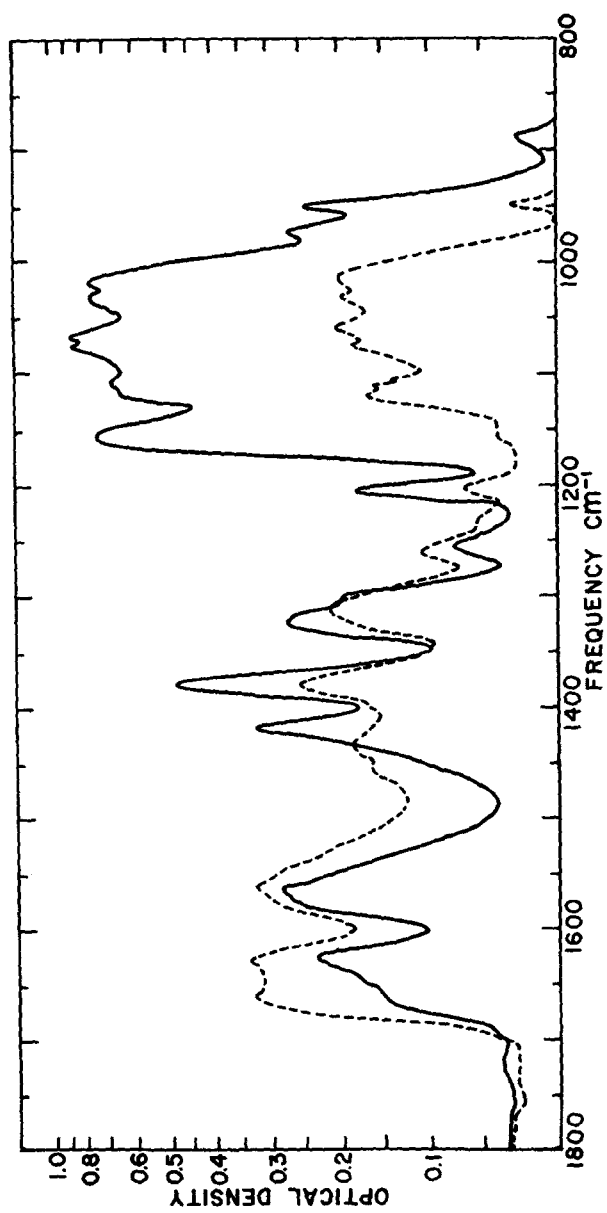


Fig. 5. Infrared absorption spectra in the region 1800 to 800 cm^{-1} of well oriented chitin, showing a strong dichroism; electric vibration direction parallel (full line) and perpendicular (dotted line) to the fiber axis.

a axis. A more detailed comparison between these two models cannot be made since the earlier structure was deduced without considerations of the side groups and was based on an incorrect unit cell and space group.

The new model of chitin shows, however, several striking similarities to that of native cellulose, the commonly accepted structure of which has been given by Meyer and Misch (1937), with minor improvements suggested by Frey-Wyssling (1955). The repeating distance along the fiber axis is essentially the same for chitin (10.28 Å) as for cellulose I (10.306 Å, Kiessig, 1950; 10.28 Å, Trillat and Legrand, 1954). Both unit cells contain four glucopyranose rings linked in two pairs by 1-4- β -glucosidic bonds. In spite of the obvious twofold screw axis in the fiber direction, the diffraction patterns of chitin and cellulose I both show a faint (010) reflection. The imperfections probably causing this forbidden reflection have been discussed for native cellulose (Meyer and van der Wyk, 1941); in view of the absence of (030), (050), and (070) in the chitin pattern it is felt that the faint meridional streak at (010) does not invalidate the twofold screw.

The arrangement of the pyranose residues in the polysaccharide chains of the Meyer and Misch cellulose model seems to be sterically unacceptable, since the $O_2-C'_6$ distance is much too short: 2.68 Å, from the atomic parameters given by Meyer and Misch, or 2.32 Å, from the known glucose configuration (see Fig. 1 *a*). Another feature of the Meyer and Misch arrangement, already pointed out by Hermans *et al.* (1943), is that hydrogens on C_1 and C'_4 interfere. The configuration of the polysaccharide chains of chitin here suggested (see Fig. 1 *b*) is indeed very similar to that proposed by Hermans *et al.* for the structure of native cellulose, and this arrangement has recently also been suggested for cellulose II (Petitpas and Mering, 1956). The position of the glucopyranose residues (when considering only a single chain) is more restricted in chitin than in cellulose because of the bulky aminoacetyl group protruding from the second carbon atom. The $O_2-C'_6$ distance (3.63 Å) given in Fig. 1 *b* was found to be just sufficient to make van der Waals contact between C_6 and the side group. Furthermore, this arrangement gives a reasonable $O_5-HO'_5$ hydrogen bond distance (2.68 Å), an acceptable glucosidic bond angle (107°), and a *b* axis repeating distance of 10.28 Å. For these reasons it is evident that only very small deviations from this arrangement are sterically justified.

The planarity of the aminoacetyl side groups can be attributed to resonance in the carbonyl group and is in agreement with earlier crystallographic findings (*cf.* the crystal structure of *N*-acetyl glycine (Carpenter and Donohue, 1950)). The rotation in the C_2-N_2 bond is limited so that the N_2 -hydrogen has to point approximately towards a carbonyl oxygen of an adjacent chain in the *a* axis direction. The favourable N_2H-O_7 distance found (2.69 Å) and indications from infrared studies furnish strong support for a hydrogen bond in this direction.

The hydroxyl groups attached to C₆ were found to have large rotational freedom, but it was clear that they could not occupy the same relative positions as in D-glucose, since O₆ then would come too close (2.63 Å) to the C₃ methyl group in the next following residue along the *b* axis direction. The optical Fourier transforms of several possible positions indicated that the O₆H groups should extend as far as possible from the chain axis. In models having the chains slightly rotated around the *b* axis in a direction opposite to the one finally suggested by the diffraction results, the O₆H groups in adjacent chains came close enough to suggest the presence of weak intermolecular hydrogen bonds forming a continuous network between the sixth hydroxyl groups. However, the optical transforms of the (010) projections of these models did not account properly for the equatorial x-ray reflections. The O₆H—O₆H distance suggested by the model accounting best for the diffraction is too large (3.3 Å) for such hydrogen bonds if no pronounced alterations are allowed in the intramolecular angles and bond distances originally assumed. In either case the forces holding the chains together in the *c* axis direction seem likely to be very weak, and thin lamellar aggregates of chains united in the *a* direction may be expected to have a certain degree of stability for use, along with protein layers, in cuticular membranes. As it is spatially impossible to insert water molecules into the full chitin crystallites, the introduction of water into fibers probably occurs at surfaces of free chains or lamellae.

The accuracy of the parameters given in Table II is difficult to estimate. It was found that the sensitivity of the optical method for the derivation of the Fourier transforms was in fact greater than necessary in this structure determination. Because of the small number of (*h*00), (*hk*0), and (*h*0*l*) reflections which were not superimposed on other ones in the x-ray diagrams, determination of an accurate fit between the optically derived transforms and the directly observed intensities of such reflections became rather uncertain. The accuracy of the *x* coordinates is consequently less than that of the *y* and *z* ones, which could be determined by comparison between transforms of the projections of various models and x-ray intensities of the relatively large number of (00*l*), (0*kl*), and (0*k*0) reflections. The errors in the relative atomic positions within each chain are probably much smaller than their absolute positions within the unit cell. The shift between the chains could be determined with an accuracy better than ±0.1 Å, but the rotation of the chains around the *b* axis remains uncertain by an amount of ±3°.

As the model now proposed accounts properly for the x-ray diffraction patterns as well as for the infrared data, and seems consistent with all known facts about the chemical and physical properties of chitin, it is felt that it is a very good approximation to the true crystalline structure.

I am very indebted to Professor R. S. Bear for having had the opportunity to work in the very stimulating atmosphere of his laboratories and for his constant interest and valuable criticism in the course of this work.

SUMMARY

A detailed model for the crystal structure of the fibrous polysaccharide chitin is proposed. The structure determination has been carried out by using an optical analogue instrument which proved to be an adequate and rapid tool for the derivation of Fourier transforms, signs of amplitudes and the production of optical Fourier syntheses. The new model of chitin accounts properly for known chemical and physical properties, including the infrared absorption as well as for x-ray data, but because of the limited resolution of the diffraction patterns it can only be regarded as a good approximation. The stereochemical configuration of the polysaccharide chains has certain implications for the structure of cellulose.

BIBLIOGRAPHY

- Andress, K. R., *Z. phys. Chem., Abt. B*, 1929, **2**, 380.
Barker, S. A., Bourne, E. J., Stacey, M., and Whiffen, D. H., *J. Chem. Soc.*, 1954, 171.
Carpenter, G. B., and Donohue, J., *J. Am. Chem. Soc.*, 1950, **72**, 2315.
Cox, E. G., and Jeffrey, G. A., *Nature*, 1939, **143**, 894.
Darmon, S. E., and Rudall, K. M., *Discussions Faraday Soc.*, 1950, **9**, 251.
Fraenkel, G., and Rudall, K. M., *Proc. Roy. Soc. London Series B*, 1947, **134**, 111.
Frey-Wyssling, A., *Biochim. et Biophysica Acta*, 1955, **18**, 166.
Gonell, H. W., *Z. physiol. Chem. (Hoppe-Seyler's)*, 1926, **152**, 18.
Hermans, P. H., de Booy, J., and Maan, C. J., *Kolloid-Z.*, 1943, **102**, 169.
Kiessig, H., *Z. Elektrochem.*, 1950, **54**, 320.
Lotmar, W., and Picken, L. E. R., *Experientia*, 1950, **6**, 58.
McDonald, T. R. R., and Beevers, C. A., *Acta Cryst.*, 1950, **3**, 394.
McDonald, T. R. R., and Beevers, C. A., *Acta Cryst.*, 1952, **5**, 654.
Meyer, K. H., and Misch, L., *Helv. Chim. Acta*, 1937, **20**, 232.
Meyer, K. H., and Pankow, G. W., *Helv. Chim. Acta*, 1935, **18**, 589.
Meyer, K. H., and van der Wyk, A. J. A., *Z. Elektrochem.*, 1941, **47**, 353.
Petitpas, T., and Mering, J., *Compt. rend. Acad. sc.*, 1956, **243**, 47.
Reeves, R. E., *J. Am. Chem. Soc.*, 1950, **72**, 1499.
Rudall, K. M., *Symp. Soc. Exp. Biol.*, (Great Britain), 1955, No. 9, 49.
Trillat, J.-J., and Legrand, C., *Bull. Soc. franç. Minér. et Crist.*, 1954, **77**, 302.
Wycoff, H. W., Bear, R. S., Morgan, R. S., and Carlström, D., *J. Opt. Soc. America*, 1957, in press.