## PRESIDENTIAL ADDRESS.

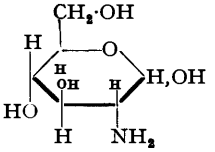
DELIVERED BEFORE THE CHEMICAL SOCIETY IN BIRMINGHAM ON APRIL 25TH, 1946.

# *Carbohydrate Components of Biologically Active Materials.*

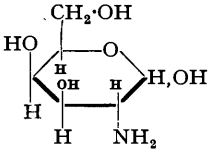
By W. N. HAWORTH, F.R.S.

THE inquiry into the structure and function of the carbohydrate components of biologically active products is being pursued by many workers. Some of the recent studies merit a more comprehensive review than I can give in the time at my disposal in this Presidential Address. In my present survey I propose to confine my attention to those carbohydrate components containing nitrogen as a substituent group. The number of sugar units which are attached to nitrogen in the complex molecule varies enormously in different cases, and other groups modifying the constituent sugar units are frequently present. Often these are high polymeric forms of carbohydrates occurring in combination with proteins or fats, and a classification of them includes mucopolysaccharides, mucoproteins, and mucolipoids originating in a wide variety of cells and especially in bacterial cells. A number of these products have been the subject of study here and elsewhere and a brief classification of many of them is given in the appendix to this Address.

In general these products are characterised by the presence of an amino-hexose unit, either chitosamine or chondrosamine; they have a low but significant protein content, and in aqueous solution the mucopolysaccharides show a high viscosity and tend to become insoluble in water after separation and dehydration. Our work here has revealed the identity of these amino-hexoses, and configurationally they are recognised respectively as 2-amino-glucose and 2-amino-galactose or as D-glucosamine (I) and D-galactosamine (II).



(I.)

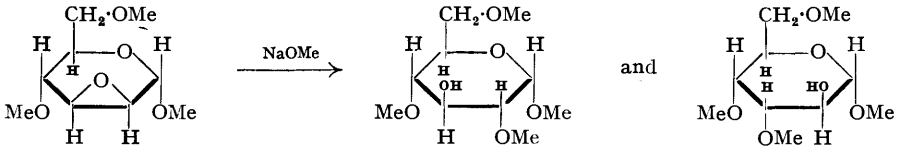


(II.)

These constitutional problems have long evaded solution because direct replacement of the amino-groups is accompanied by Walden inversion and treatment with nitrous acid leads to the anhydro-sugars such as chitose. It was suggested that chitosamine was 2-amino-mannose, whilst chondrosamine could be regarded either as 2-amino-talose or 2-amino-galactose.

Biological activity is closely related to stereochemical arrangement, and it was important to refer these unique and characteristic units to the parent configurational sugar types. Experimentally this has been achieved in my laboratories by the synthesis of glucosamine and galactosamine derivatives through the anhydro-sugars which have been intensively studied over a number of years. A purely physical confirmation is provided by *X*-ray methods which Cox and Jeffrey applied in these laboratories. The isomorphism of the hydrochloride and hydrobromide of α-D-glucosamine enabled direct methods of calculation of electron densities to be employed and, independently of any previous stereochemical assumption, the atomic positions were determined with high precision (*Nature*, **1939**, **143**, 984).

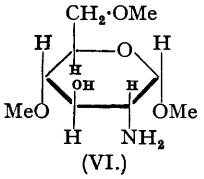
The chemical proof is founded on the initial observation that the derivative of 2 : 3-anhydro-mannoside (III) opens up its ethylene oxide ring in two ways with sodium methoxide to give both the 2 : 4 : 6-trimethyl methylglucoside (IV) and the 3 : 4 : 6-trimethyl methylaltroside (V), the attachment of the entering methyl group from the sodium methoxide being accompanied always by a Walden inversion at the position of the new attachment:



(III.) (IV.) (V.)

By using ammonia in methyl alcohol instead of sodium methoxide for the ring fission it was found that the entering amino-group at position 2 in the formula corresponding to (IV) gave a crystalline derivative identical with a reference product of chitosamine, and therefore this product was

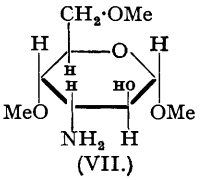
configurationally identified as a 2-amino-glucose (VI), and not 2-amino-mannose which would have the amino-group in the upper plane and not in the lower one. The two products (VI) and (VII) of the reaction with ammonia were therefore the following (Lake and Peat, and Peat and Wiggins, *J.*, 1938, 1417, 1810) :



(VI.)

2-Amino-glucose derivative.

and

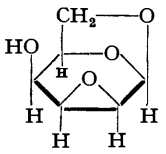


(VII.)

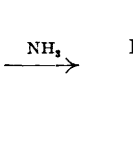
3-Amino-altrose derivative.

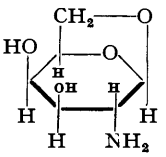
The entering anion attaches itself to the carbonium cation as the three-membered ring opens in either direction, and the Walden inversion occurs at the point of union of the anion.

During the current year a similar reaction has been successfully applied in the investigation of chondrosamine, a more recondite but equally important amino-hexose (James, F. Smith, Stacey, and Wiggins, *J.*, 1946, 625). Here the initial anhydro-sugar employed for the purpose was the dianhydride of talose (VIII) in which an anhydro-link was in positions 2 : 3, and another at 1 : 6, this latter not being concerned in the critical reaction :



(VIII.)

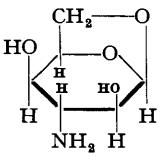




(IX.)

2-Amino-galactose anhydride.

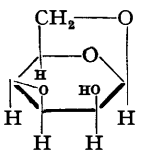
and



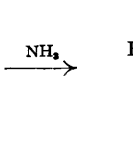
(X.)

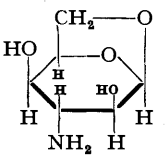
3-Amino-idose anhydride.

In confirmation of the conclusion that (X) was 3-amino-idose, a similar experiment was conducted by using the 3 : 4-1 : 6-dianhydrotalose (XI) as the initial material. Here the entering amino-group at position 3 furnished 3-amino-idose anhydride identical with (X), and also 4-amino-mannose anhydride (XII) where the Walden inversion had occurred at position 4.



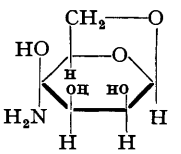
(XI.)





Identical with (X).

and



(XII.)

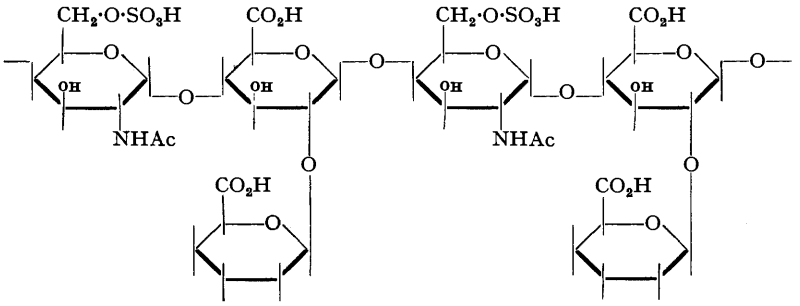
The 2-amino-galactose-1 : 6-anhydride (IX) was readily converted into 2-amino-galactose hydrochloride by using concentrated hydrochloric acid to open the 1 : 6-ring. Its identity with the naturally occurring chondrosamine hydrochloride was confirmed by an *X*-ray comparison of the two specimens. Inversion at the 1-position in this case would be immaterial in that mutarotation occurs in any event at this reducing position, and likewise it could not be apparent at the 6-position. The isolation of crystalline 2-amino-galactose hydrochloride from this sequence of changes is an achievement of the constitutional synthesis of chondrosamine, which is the analogue of glucosamine occurring in cartilaginous tissue and therefore a constituent unit in chondroitin sulphuric acid.

We have made some attempt to determine the molecular arrangement of this chondroitin sulphate. This is perhaps the most readily accessible sulphur-containing carbohydrate that occurs naturally, since it forms one of the major constituents of cartilaginous tissue from the trachea, nasal septa, aorta, etc. The first study on the nature of this complex acid carbohydrate was made by Schmiedeberg (*Arch. exp. Path. Pharm.*, 1889, **28**, 355) who gave a general idea of its type of structure. Levene, from his extensive investigations, considered that a derived sulphate-free substance " chondroitin " was a tetrasaccharide composed of two glucuronic acid residues conjugated with two units of chondrosamine, the whole then being esterified in the natural complex with sulphuric acid. The work of Bray, Gregory, and Stacey (*Biochem. J.*, 1944, **38**, 142) shows that this conception must now be modified. The Birmingham workers

prepared a methylated degraded chondroitin which was sulphate-free, and, from the hydrolysis of this, derivatives of *d*-glucuronic acid (including 2 : 3 : 4-trimethyl β-methyl-*d*-glucuronoside) and of chondrosamine (including *N*-acetyl 3 : 4 : 6-trimethyl methylchondrosaminide) were identified. About one-half of the D-glucuronic residues were attached glycosidically as terminal residues, so that it would appear that the repeating unit is a trisaccharide, and that the structure may be the branched chain type shown in Figure 1. In this, the precise points of linkage other than the glycosidic attachments have yet to be determined.

Karl Meyer and his colleagues (*Cold Spring Harbor Symposia Quant. Biol.*, 1938, 6, 99) examined the complexes which chondroitin sulphate forms with proteins and concluded that the $NH_2$ groups of the protein combine with the $SO_3H$ and $CO_2H$ groups of the carbohydrate. They

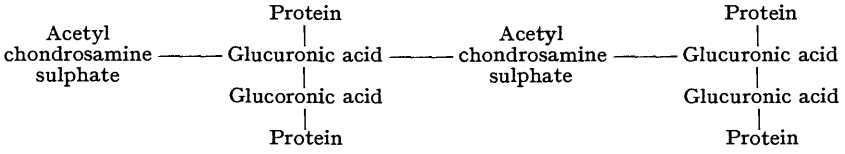
*Figure* 1.




*Scheme for Chondroitin Sulphuric Acid.*

suggested that hyaline cartilage is a salt of chondroitin sulphate with a protein of the gelatine type, and they succeeded in combining edestin with chondroitin sulphate to give fibres with properties closely resembling those of the elastic fibres of natural connective tissues. An idea of the structural complexity of the natural mucoprotein is suggested in Figure 2.

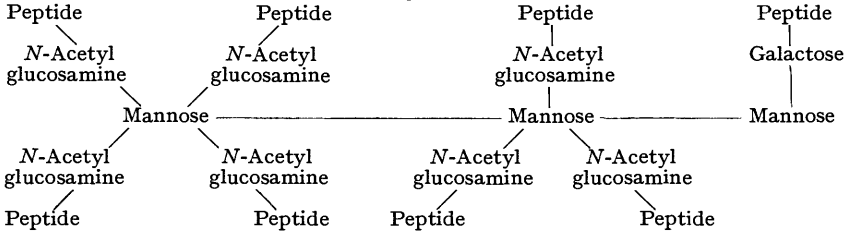
*Figure* 2.




*Chondroitin from Cartilage.*

A remarkable carbohydrate containing a high proportion of *N*-acetyl glucosamine residues has been studied by Stacey and Woolley (*J.*, 1940, 184; 1942, 550). This is the carbohydrate complex in ovomucoid from egg white. It was possible to saponify ovomucoid (which contains 80% of a protein constituent) with sodium hydroxide and to methylate simultaneously the carbohydrate residue with methyl sulphate to give an *N*-acetyl methyl derivative. Examination of the hydrolysis products of this revealed the remarkable result that by glycosidic attachment seven *N*-acetyl glucosamine units and one galactose unit radiate from a central core of three mannose units as in Figure 3.

*Figure* 3.




*Ovomucoid.*

In this representation the carbohydrate portion of the structure is shown not as a repeating unit but as the interior of one whole molecule and would thus be of relatively low molecular

weight. It would be equally possible to rearrange the peptide links to represent a much larger molecule of which the carbohydrate core is a repeating unit.

This work is of importance and may throw some light on problems connected with the occurrence of prosthetic carbohydrate groups in proteins, considered particularly from the immunological point of view.

One of the most fascinating and important fields of study of mucopolysaccharides containing amino-sugars and uronic acids is the group of bacterial polysaccharides, especially those which form the protective capsular material of the cell. The main reason for lack of rapid progress in this field is the expense and difficulty of obtaining sufficiently large amounts of authentic standardised material for chemical study. Some idea of the problems involved can be gathered from the recent enumeration by Boyd (" Fundamentals of Immunology ", **1943**, Interscience Publishers) of more than 40 types of *Pneumococcus*, each possessing a different and well characterised capsular polysaccharide. Complex carbohydrates from these pneumococcus and other bacterial cells usually contain a certain proportion of nitrogen as shown in the table, which records some of our own findings.

*Nitrogen Content of Polysaccharides of Bacterial Origin.*

| Polysaccharide from : | Nitrogen (%). |
|---|---|
| Type I *Pneumococcus* ..... | 5·5 |
| Type IV *Pneumococcus* ..... | 6·0 |
| *Pneumococcus* Somatic (Groups) ..... | 5·5 |
| *Bact. dysenteriæ* (Shiga) ..... | 2—3 |
| *Bact. typhosum* ..... | 2—3 |

Frequently this nitrogen owes its origin to the presence of constituent hexosamine units, although where the percentage is low, *e.g.*, 0·3%, there is usually found a small amount of a protein or a nucleic acid prosthetic group. The two nucleic acids, the ribo- and the deoxyribo-types account for the presence of pentoses and deoxypentoses in some bacterial polysaccharides, for very rarely are pentoses found to occur as monosaccharide units. A notable exception is in the case of two specific polysaccharides present in the tubercle bacillus which contains the rare pentose, D-arabinose. In addition, both of these polysaccharides contain hexosamine units, so that they present suitable types for study more especially since organisms causing tuberculosis can be grown on a standard and entirely synthetic medium in comparatively large quantities. We have recently been able to record some progress on the antigenic as well as on the haptenic polysaccharide constituents of this important pathogenic cell.

Probably no other single organism has been the subject of so much investigation as *Mycobacterium tuberculosis*. It was discovered and characterised by Koch in 1889 and the unique nature of the organism was soon recognised. In addition to the peculiar bacteriological properties of the four types of *M. tuberculosis* (human, bovine, avian, and cold-blooded) the chemical properties of their major constituents, polysaccharides, fats, and nucleoproteins have proved to be of unusual interest.

Although *M. tuberculosis* is readily killed by heat, it is relatively resistant to chemical attack, and this property has been ascribed in some measure to the presence of a waxy protective layer possessed by the organism. Recent American evidence (electron microscope) has confirmed an earlier opinion that the envelope is not uniform but has a granular texture, and although the question as to whether the tubercle organism possesses a true capsule is controversial (Dubos, " The Bacterial Cell ", Harvard University) we have shown that it does at least possess a surface layer. The waxy substances of which the layer is composed have been extensively investigated, more particularly by R. J. Anderson and his school (*Chem. Reviews*, **1941**, **29**, 225). He has found that the lipoids occurring in these acid-fast bacteria are built on quite a different plan from those found in plants and animals and that their composition is more complex. Lipoids contained in five different strains of the human tubercle bacillus have been examined, together with those from the bovine and avian types, the so-called leprosy bacillus, and the non-pathogenic timothy glass bacillus, all grown on the Long synthetic medium. The interesting discovery was made that the fat of the human strain is not a glyceride, and the ordinary definition of a fat will need extension to include these new observations. The disaccharide trehalose takes the place of glycerol as the water-soluble component of the fat, whilst the fatty acids include new and hitherto unknown members such as tuberculostearic (10-methylstearic) and phthioic acids (see below).

Similarly, the phosphatides reveal new combinations in which the glycerophosphoric acid

occurring in plants and animals is partly replaced by a phosphorylated polysaccharide or glycoside; choline and aminoethyl alcohol appear to be absent.

Again, the waxes of these acid-fast bacteria are quantitatively important and are novel in composition since they are mainly esters in which optically active hydroxy-acids are combined with carbohydrates instead of with the usual alcohols. The wax from the human strain contains a specific polysaccharide, not found in waxes from other strains which often contain trehalose, and in the leprosy bacillus wax the alcohol is glycerol. Noteworthy among the optically active hydroxy-acids concerned in this combination in the human strain is mycolic acid $C_{88}H_{176}O_4$—a remarkably high molecular weight—and probably the property of acid-fastness in the bacilli is attributable to this component. For these and other outstanding researches we are indebted to R. J. Anderson and his co-workers.

As indicated below we have elucidated part of the structure of a polysaccharide obtained from the somatic portion of the cell after removal of the waxy layer. The polysaccharide behaved as a haptene and reacted with immune serum from tuberculous patients in high dilution (1 : 2,000,000). Some of the fatty acids are of profound interest since on injection into animals they are able to stimulate giant-cell and tubercle formation.

Recently a structural investigation of the complex mixture of fatty acids was undertaken by Sir Robert Robinson and N. Polgar (*J.*, 1945, 389). Phthioic acid, one of the specific cellular stimulants responsible for tubercle formation, has been shown to have the following structure :

$$CH_3{\cdot}[CH_2]_3{\cdot}CHMe{\cdot}[CH_2]_6{\cdot}CHMe{\cdot}[CH_2]_9{\cdot}CHMe{\cdot}CH_2{\cdot}CO_2H$$

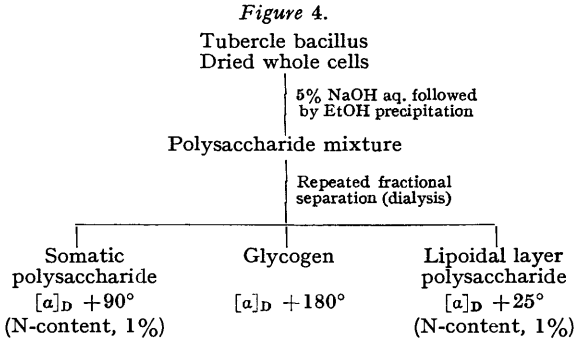
*Phthioic* (3 : 13 : 19-*Trimethyltricosanoic*) *Acid.*

The literature contains many references to polysaccharides which have been isolated from the cells and also from the culture medium remaining after the organisms have been grown. From the latter source, polysaccharides were found in close association with deoxyribonucleic acid and " tuberculin " protein, the latter being the active agent in the well-known specific skin test. The problem of separation of the major components of tuberculin, using the Tiselius apparatus, was investigated in great detail by Seibert and her co-workers (*Chem. Reviews*, **1944**, **34**, 107).

A variety of specific polysaccharide fractions has been isolated from the whole bacterial cell by various workers; the fractions have widely different properties and serological activities, though in many cases there is reason to doubt the homogeneity of certain fractions. Tuberculosis glycogen has been separated from the polysaccharides of the cell, and measurements of its molecular size were made in the ultra-centrifuge (Chargaff and Moore, *J. Biol. Chem.*, 1944, **155**, 493) although the quantity available was insufficient to permit of an end-group assay. The glycogen assumed a characteristic red colouration with iodine, but it was noted that enzymic degradation with amylase proceeded more slowly than in the case of liver glycogen.

Here with my colleagues Stacey and Kent we have investigated the carbohydrate content of the outer lipoidal layer of the bacillus as well as that of the inner or somatic region. Two distinctive polysaccharides of different rotatory power are recognisable as well as the more common polysaccharide, glycogen. The two former are haptens which react in a dilution of well over one in a million with serum from tuberculous patients. They appear similar to those isolated by Heidelberger and his colleagues by other methods (*J. Biol. Chem.*, **1937**, **118**, 79). The scheme of isolation, shorn of all but essential details, is outlined in Figure 4.

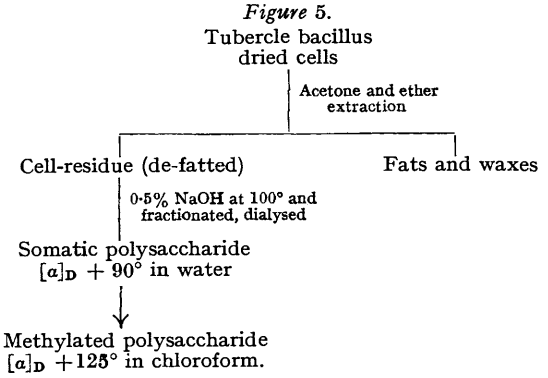
*Figure* 4.




It was apparent that an alternative route to the isolation of the somatic polysaccharide lay through the extraction by ether and acetone of all materials of the lipoidal or waxy layer. The

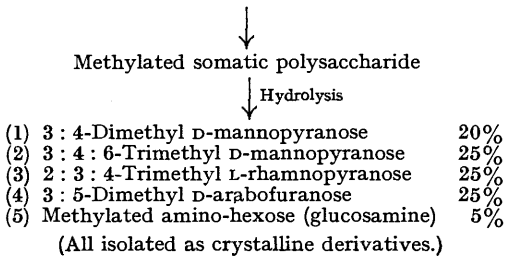
de-fatted cell-residue was then passed through the same procedure as that outlined in the above scheme. There was then isolated what appeared to be only the somatic polysaccharide ($[\alpha]_D$ + 90° in water); the methyl ether derivative, having $[\alpha]_D$ + 125° in chloroform, gave a quantitative separation of hydrolysis products, all of them in the form of crystalline compounds.

*Figure* 5.




The haptenic polysaccharide obtained by vigorous saponification methods from the somatic part of the cell after removal of the waxy layer (Figure 5) also contains an amino-sugar, mannose, L-rhamnose, and a D-arabinose constituent, and this structure of the polysaccharide has been investigated in some detail. The methyl ether, $[\alpha]_D$ + 125°, obtained in good yield, is readily hydrolysed with methanolic hydrogen chloride. From the hydrolysis mixture the constituents as shown in Figure 6 were characterised and identified.

*Figure* 6.



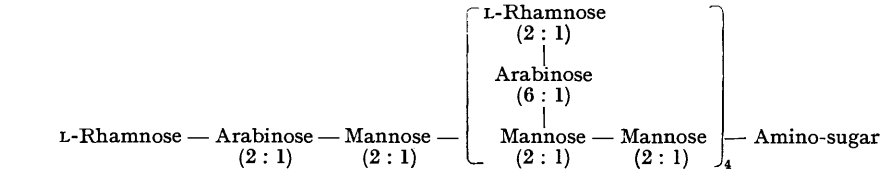

It is possible to speculate on the complete formulation of this polysaccharide; some interesting points are clear and the possible structural types can be suggested. The molecule is evidently highly branched since some of the mannopyranose units are linked through the 1 and 2 positions, while others are linked through the 1, 2, and 6 positions.

The whole of the L-rhamnose residues are in the pyranose form and constitute terminal units linked glycosidically to the rest of the molecule. The existence of L-rhamnose in this complex has not previously been recognised, and the points of union of each of the sugar units are given for the first time. The arabinose units are of great interest. Other workers (Heidelberger and Anderson) have shown that tuberculosis polysaccharides contain D-arabinose units, but I am able to announce the first recorded isolation of D-arabofuranose from this or indeed any other natural source. The arabofuranose units are all linked through the 1 and 2 positions. Preliminary results seem to indicate that the amino-sugar closely resembles D-glucosamine, so that this important amino-hexose plays a ubiquitous role in the composition of many carbohydrates of biologically active materials.

The proportions of the constituents actually isolated are as follows:

| | | | | | |
|---|---|---|---|---|---|
| D-Mannopyranose | linked | through | the | 1, 2, and 6 positions | : 4-parts. |
| ,, | ,, | ,, | | 1 and 2 ,, | : 5 ,, |
| D-Arabofuranose | ,, | ,, | | 1 and 2 ,, | : 5 ,, |
| D-Rhamnopyranose | ,, | ,, | | 1 position ,, | : 1 ,, |
| Partly methylated 2-amino-sugar, linkage unknown | | | | | : 1 ,, |

A possible formulation for a polysaccharide having these specifications is the following :

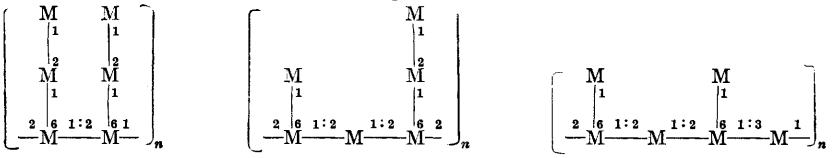



Graded hydrolysis or acetolysis methods are needed to decide the more exact positions of the constituent units.

There is a general structural similarity between this mannose-containing polysaccharide and mannans of microbiological origin, *e.g.* those from yeasts and moulds though these are less complex.

Some possible structures for the yeast mannan (Haworth, Heath, and Peat, *J.*, 1941, 833) are shown in Figure 7.

*Figure* 7.



The polysaccharide from the lipoidal layer (see Figure 4) was also submitted to constitutional study through its methyl ether. Hydrolysis led to the isolation and quantitative estimation of crystalline derivatives of the following components in their glycosidic forms :

| | Molecular ratio. |
|---|---|
| 2 : 3 : 5-Trimethyl methyl-D-arabofuranoside ..... | 2·05 |
| 3 : 5-Dimethyl ,, ,, ..... | 5·3 |
| 2 : 3 : 6-Trimethyl methyl-D-galactopyranoside ..... | 4·7 |
| 3 : 4-Dimethyl methyl-D-mannopyranoside ..... | 2·1 |
| Dimethyl amino-methylhexoside ..... | 1·0 |

From considerations analogous to those applied above in the case of the somatic polysaccharide a possible and provisional structural formula may be suggested for this carbohydrate.

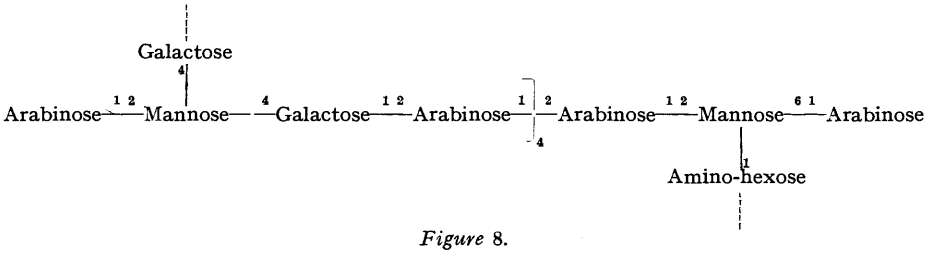
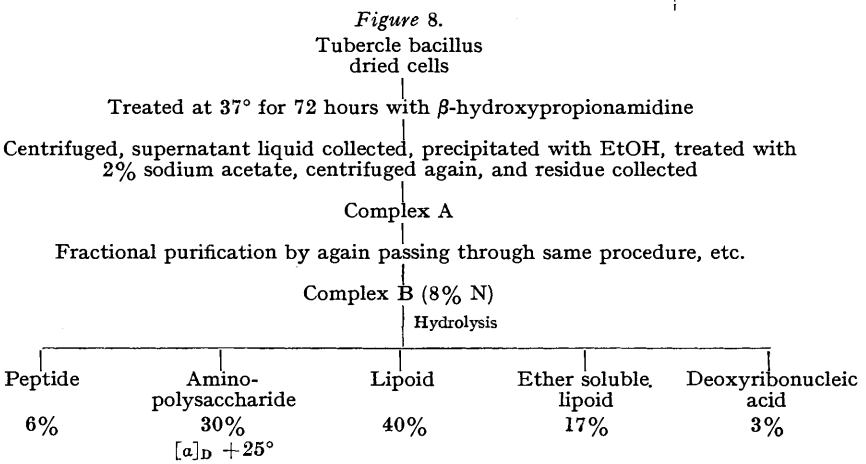



*Figure* 8.



The polysaccharide from the waxy layer was bound in a novel complex which had some interesting serological properties. This complex was possibly an antigen which, in addition to the polysaccharide (*a*), contained (*b*) an acid-fast lipoid material, (*c*) deoxyribonucleic acid, and (*d*) a small amount of a peptide constituent, and it could be obtained only from the intact cells. The complex was extracted at 37° as shown in Figure 8 by organic lipophilic bases (including urea and β-hydroxypropionamidine), further fractionated with organic bases, and finally isolated by being dried from the frozen state.

For the serological tests the materials, *e.g.*, Complexes A and B, were both separately coated on to a colloidal suspension of collodion particles and then tested against the sera of tuberculous and normal humans. The tests were carried out by Dr. Nassau of the Harefield Sanatorium and are being extended elsewhere. The collodion particles coated with Complex A and B give a good agglutination under appropriate conditions with sera from tuberculous patients, and already the results appear to show promise in diagnosis apart from any possible value in immunisation. It is of interest to note that no antigenic material could be obtained from defatted cells.

Other complex carbohydrates of high biological significance are the specific blood group factors, an account of which has been given by Stacey in "Advances in Carbohydrate Chemistry". Group A resembles the tuberculosis polysaccharides in possessing branched chain structures with glucosamine, mannose, and galactose, as constituent units. A specimen examined (Bray, Henry, and Stacey, *Biochem. J.*, 1946, 124) contained a methylpentose end-group, namely L-fucose, which thus differed significantly from the L-rhamnose end group in the tubercle polysaccharide described above.

It will be seen that many projects in the carbohydrate field still await the attention of workers interested in the problems of infectious disease processes, and it is worthy of mention that American investigators have isolated from the important antibiotic, streptomycin, the optical enantiomorph of glucosamine, namely the L-isomer.

## *Appendix.*

| | *Identified products of acid hydrolysis of carbohydrate portion.* |
|---|---|
| 1. *Containing hexosamine and hexuronic acid.* | |
| (a) Sulphate-free. | |
| (i) Hyaluronic (from vitreous humour, umbilical cord, synovial fluid, ovarian tumour, Group "A" hæmolytic *Streptococcus* and skin). | *N*-Acetyl D-glucosamine<br>Hexuronic acid (probably D-glucuronic acid) |
| (ii) Type I *Pneumococcus* specific polysaccharide. | Acetic acid<br>Hexosamine<br>D-Galacturonic acid |
| (b) Sulphate-containing. | |
| (i) Heparin. | Sulphuric acid<br>D-Glucuronic acid<br>D-Glucosamine (unacetylated) |
| (ii) Chondroitin sulphate. | *N*-Acetyl chondrosamine<br>Sulphuric acid<br>D-Glucuronic acid |
| (iii) Mucoitin sulphate (from gastric mucin). | *N*-Acetyl D-glucosamine<br>Hexuronic acid<br>Sulphuric acid |
| (iv) Hyaluronic acid sulphate (from the cornea). | Sulphuric acid<br>*N*-Acetyl hexosamine<br>Hexuronic acid |
| 2. *Containing hexosamine but no hexuronic acid.* | |
| (a) Chitin. | *N*-Acetyl D-glucosamine |
| (b) Bacterial. | |
| (i) *Pneumococcus* "C" (or Group) polysaccharide. | *N*-Acetyl amino-sugar<br>D-Glucose |
| (ii) Type IV *Pneumococcus* specific polysaccharide. | *N*-Acetyl amino-sugar<br>D-Glucose |
| (iii) Type XIV specific polysaccharide. | *N*-Acetyl D-glucosamine<br>D-Galactose |
| (c) Blood Group. | |
| (i) Blood Group A carbohydrate (from saliva, pepsin, urine, gastric mucin, pancreas, etc.). | *N*-Acetyl D-glucosamine<br>D-Mannose<br>D-Galactose<br>L-Fucose |
| (ii) Blood Group B and O carbohydrates from human saliva and gastric mucin. | Hexosamine |