The Occurrence and Significance of the Pentose Sugars in Nature, and their Relationship to the Hexoses.

THE HUGO MÜLLER LECTURE, DELIVERED BEFORE THE CHEMICAL SOCIETY ON MARCH 4TH, 1948.

By E. L. HIRST, M.A., D.Sc., F.R.S.

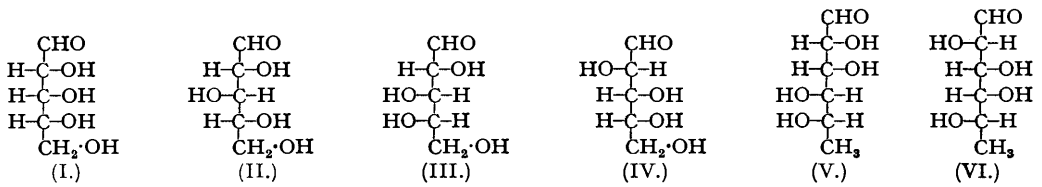
MODERN organic chemistry impinges in so many ways on botanical sciences that any chemist who is honoured by an invitation to deliver the Hugo Müller Lecture must be aware of many fields of investigation, each of which could serve as a subject for review fully compatible with the conditions attached to the lectureship. Even when the scope of inquiry is limited to the chemistry of the carbohydrates there still remains a wide range of problems which involve, at one and the same time, fundamental problems in both botany and chemistry. There are, for example, the many aspects of the problem of photosynthesis, and the multifarious questions relating to the structure and transformation products of the starches. These subjects, however, have been authoritatively reviewed within the past year or two, and it seemed opportune to use the present occasion for the discussion of certain problems connected with the pentose group of sugars. These questions, although clearly of great and increasing importance to both chemists and biologists, have not as yet passed beyond the initial stages of investigation, and, in consequence, it will be necessary to lay stress more on the nature of the problems which confront us than on clear cut solutions and generalisations.

The group of substances which are to be the main subject of discussion includes the four pentoses—xylose, lyxose, ribose, and arabinose—together with two methylpentoses, rhamnose and fucose. Considerable historical interest is attached to this group in view of the essential part its members played in the early development of carbohydrate chemistry when they were used in the elucidation of the stereochemistry of the hexoses. But at that time the main interest was centred on the hexoses themselves in view of the obvious importance of the many well-known natural products based on glucose, galactose, mannose, and fructose, and, in consequence, the pentoses were regarded as little more than chemical curiosities, despite the established occurrence of arabinose and xylose in certain plant materials. Later, when special attention was paid to the nature of the ring systems present in sugars, the problems presented by the pentoses were found to be somewhat simpler in character than those encountered in the hexose series, and it will be recalled that the occurrence of the pyranose ring in a stable carbohydrate derivative was, in fact, first demonstrated for the normal form of β-methylxyloside.[1]

In recent years, however, the widespread occurrence of the pentose sugars in Nature, and their importance in metabolism, have become increasingly manifest, and greater interest is now being taken in the study of these substances in their own right and not merely as aids or adjuncts to the chemistry of the hexoses. They occur, for example, as essential components of some members of the B group of vitamins and in certain of the co-enzymes present in systems responsible for oxidation and reduction. D-Ribose and its 2-deoxy-derivative are the carbohydrate components of the nucleic acids, the study of which is being so vigorously pursued today on account of their intimate connection with the chemistry of living processes. Pentose residues are encountered also in an astonishing variety of complex polysaccharides which act as food reserve materials in plant seeds. They are found in the cell walls of plants, and one of them plays a mysterious part in the chemical changes which result in lignification. It has been observed also that the new membrane formed in the course of cell division shows initially reactions which are characteristic of pectic materials, and the results of recent investigations suggest that pectic materials usually, and perhaps invariably, contain pentose constituents. It may well be therefore that pentose sugars play an essential part in this phase of cell division. As another example of their versatility it may be recalled that certain polysaccharides which contain pentose sugar residues in their molecular structure may function as water retaining materials of use to the germinating seed. In view of this wide range of uses and properties the biochemical importance of this group is unquestionable, and it appears appropriate to review briefly the ways in which the pentoses occur in Nature and to indicate such generalisations as have become apparent concerning their structure and their biochemical relationships.

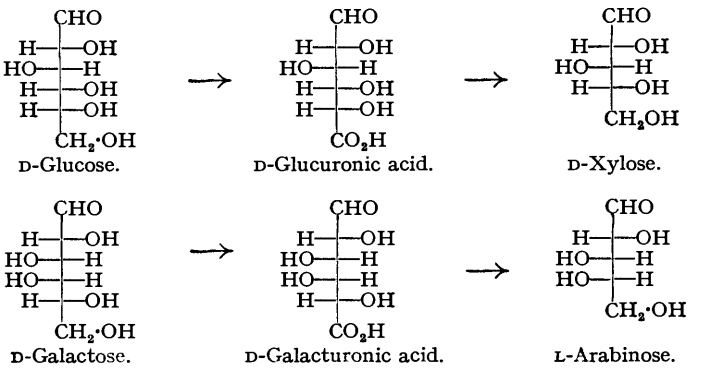
There are four possible aldopentoses and all of them are known; namely, arabinose, ribose,

xylose, and lyxose. Each of them can exist in enantiomorphic D- and L-forms, but of these eight sugars we shall be concerned mainly with D-ribose (I), D-xylose (II), L-arabinose (III),



and D-arabinose (IV), together with the two methylpentoses L-rhamnose (V) and L-fucose (VI), the other modifications being rarely, if ever, encountered amongst natural products. A detailed study of the chemistry of these sugars has shown that they give rise to derivatives of the aldehyde form, and of both the pyranose and the furanose ring form. In general their chemical properties are very similar to those of the corresponding hexoses, but they appear on the whole to be somewhat more reactive. So far as our present knowledge extends, each of these sugars tends to occur in natural products in one preferred ring form. For example, ribose and arabinose are encountered in the furanose ring form, whilst xylose and rhamnose invariably occur as pyranose sugars.

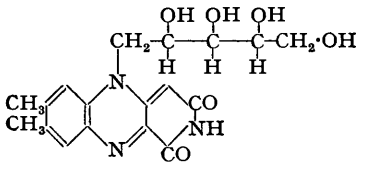
Inspection of these formulæ reveals some interesting stereochemical relationships. D-Xylose is stereochemically very similar to D-glucose, and so is L-arabinose to D-galactose. In both pairs the hexose sugar could be transformed into the corresponding pentose by oxidation at C_6 to the uronic acid, followed by decarboxylation. Now it is significant that in the group of plant gums and mucilages one very frequently finds D-galacturonic acid associated with D-galactose and L-arabinose, and D-glucuronic acid with D-xylose and D-glucose. It must, therefore, be regarded as probable that these pentoses arise in Nature by a mechanism of this type. Nevertheless a detailed examination of the structures into which the various residues are incorporated shows that the transformations cannot take place at the polysaccharide level, but must involve preliminary degradation to monosaccharides after which new structures are built up. Much of the experimental work about to be mentioned had its origin in attempts to find an answer to this question of the origin of the pentoses.



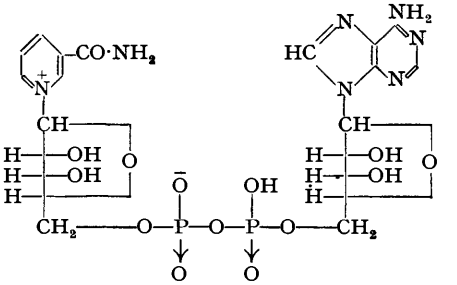
Among the pentose sugars, D-ribose occupies a special position in that it is the sugar present in nucleic acids, in certain co-enzymes and vitamins. For example it is found in a very unusual form, namely as the reduced ribityl residue, in riboflavin (6 : 7-dimethyl-9-D-ribityl*iso*alloxazine). Here we are concerned only with the chemistry of the sugar portion of the molecule, and it is significant that despite the absence of ring structure in the sugar the stereochemistry of the sugar residue is of great biological importance. Synthesis of analogous flavins containing other sugar residues gives rise to substances of quite different biological activity.

D-Ribofuranose residues are encountered in the diphosphopyridine nucleotide (co-enzyme I) [2] and in the nucleic acid group. The structures assigned have been arrived at on evidence based on degradations and methylation results. In some substances the nature of the ring system has been definitely ascertained, but it is clear that the chemistry of the nucleotides and nucleic acids requires much further exploration. Rapid advances are now being made as the result of

the synthetic work of Professor A. R. Todd and his colleagues at Cambridge, and recently these workers have succeeded in carrying through for the first time a synthesis of a naturally occurring

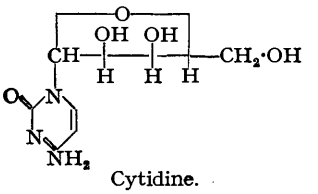


Riboflavin.



Diphosphopyridine nucleotide.

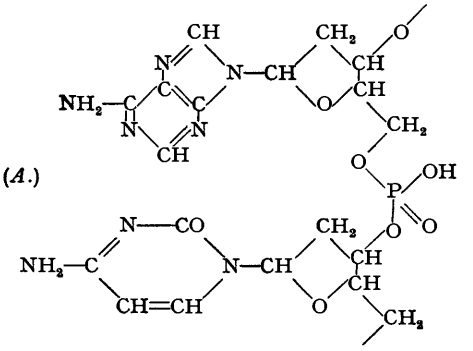
pyrimidine nucleoside (cytidine).[3] This involved the difficult feat of introducing the D-ribose residue in its furanose form, and now that this has been accomplished the way is open for further syntheses of nucleosides. Work in this field, it should be noted, involves all the difficulties inherent in the chemistry of those very labile substances, D-ribose and 2-deoxy-D-ribose, as well as those encountered in handling nitrogen glycosides.



Cytidine.

Some idea of the general nature of the polymeric nucleic acids has been gained by a combination of chemical and *X*-ray data (Astbury).[4] The diagram (*A*) indicates Astbury's views on the structure of nucleic acids based on deoxyribose. Here the mode of linkage of successive units is necessarily at C_5 and C_3 of the deoxyribose residues, but the problem is much more difficult for the ribose nucleic acids in which each sugar residue possesses an additional OH group. The evidence is conflicting and difficult to interpret owing largely to uncertainties concerning the rates of hydrolysis of phosphate ester groupings attached to the various OH groups, but according to Astbury the observed *X*-ray spacings tend rather to favour linkages through C_5 and C_3 (or perhaps C_5 and C_2).

(*A*.)

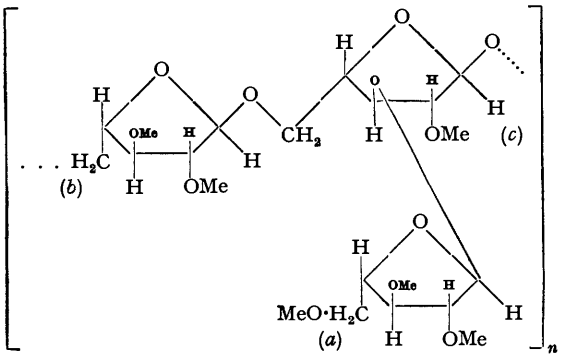


(Adenine.)

(Cytosine.)

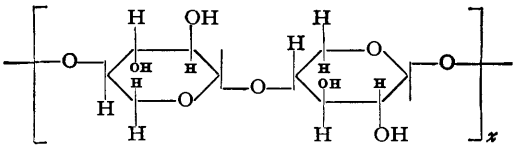
In the nucleic acids and nucleotides, ribose is not accompanied by other pentose sugars; these, however, occur very widely in plant products, notably in gums, mucilages, and pectic materials. The pentose sugars here involved are L-arabinose and D-xylose, and ribose is absent. Associated with the pentoses one frequently finds uronic acids (D-glucuronic, D-galacturonic) and hexose sugars, especially D-galactose, D-mannose, together with the methylpentoses L-rhamnose and L-fucose. The polysaccharide, araban,[5] which is found as a component of the pectic

material of many plants (for example ground nuts, sugar beet) may be taken as a simple example. It can be extracted from finely powdered ground nut,[6] after removal of the oil (by alcohol) and protein (by 10% aqueous sodium chloride), by treatment with dilute sodium hydroxide. This gives a mixture of pectic acid and araban which is extracted with alcohol, giving a crude araban still severely contaminated with pectic acid The pure araban is obtainable by regeneration of the free polysaccharide from its acetate. These operations are mentioned in order to illustrate one of the main difficulties encountered in this field of study, namely the isolation of pure homogeneous material suitable for structural investigations. The araban gives on methylation the dimethyl derivative, which in turn undergoes hydrolysis to 2 : 3 : 5-trimethyl, 2 : 3-dimethyl, and 2-methyl L-arabinose in equimolecular proportions. The ease of hydrolysis indicates that all the residues are furanose. These observations enable us to gain considerable insight into the general type of structure present in the araban, although they are insufficient to establish a single unique formula. The number of end groups (1 in 3) shows that a highly branched structure is present, and one of the very few structural formulæ which will fit the observations is shown below.[6]



(*a*) Gives 2 : 3 : 5-trimethyl, (*b*) gives 2 : 3-dimethyl, (*c*) gives 2-methyl L-arabinose.

At this stage one of the most important polysaccharides containing xylose may be briefly considered. This is the xylan or wood gum intimately associated with cellulose in woody tissues and found abundantly in wood, straw, esparto grass, oat-hulls, etc. It is extracted by alkali, but as normally isolated it contains residues of L-arabinose (3—8%) as well as of D-xylose. The main features of its structure are clear, since on methylation and hydrolysis it yields 2 : 3-dimethyl D-xylose, revealing thereby that it contains chains of 1 : 4-linked xylopyranose residues which are structurally analogous with the glucose chains in cellulose.[7] The molecular weight is fairly high, probably not less than 20,000. Much has still to be learned concerning the



fine structure. At one time the evidence seemed to indicate that L-arabofuranose residues terminated the chains of D-xylose units in xylan,[8] but it has recently been observed that the proportion of arabinose is variable, depending on the mode of extraction, and it has been found possible, by using only very gentle methods of purification, to obtain an undegraded xylan which contains no arabinose residues.[9] It seems likely, therefore, that xylan is composed solely of xylose units and that the arabinose normally associated with it occurs as an araban polysaccharide of the type we have just considered. Further evidence in favour of this view is the observation that, besides arabinose, the crude xylan contains D-galactose and D-galacturonic acid, typical of pectic materials.

It is only rarely, however, that plant polysaccharides are found containing one sugar residue only. In general, the structures display considerably greater complexity as will be apparent from Table I which indicates the sugar and uronic acid residues present in some typical gums and mucilages. When so many sugars are involved it becomes all the more difficult to be certain that

the material selected for examination is a single homogeneous substance. Fortunately, however, some of the gums (for example, that exuded by the damson tree) appear to possess an astonishingly invariable composition which facilitates examination. On the other hand there appear to be several types of gum arabic, differing in their arabinose content and possibly in other respects also.

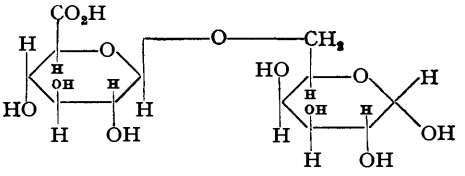
TABLE I.

Gum.	D-Glucuronic acid.	D-Galacturonic acid.	D-Galactose.	D-Mannose.	D-Xylose.	L-Arabinose.	L-Rhamnose.	L-Fucose.
Almond	+		+		+	+		
Arabic	+		+			+	+	
Cherry	+		+	+	+	+		
Cholla		+	+		+	+	+?	
Damson	+		+	+	+	+		
Egg plum	+		+		+	+		
Grape-fruit	+		+			+		
Lemon	+		+			+		
Mesquite	+*		+			+		
Orange	+		+			+		
Purple plum	+		+		+	+		

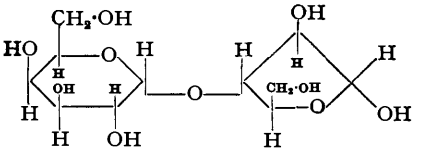
Mucilage.	D-Glucuronic acid.	D-Galacturonic acid.	D-Galactose.	D-Mannose.	D-Xylose.	L-Arabinose.	L-Rhamnose.	L-Fucose.
Brassica alba		+	+		+	+	+	
Lepidium sativum		+	+		+	+	+	
Linseed		+	+		+	+	+	
Locust bean			+	+				
Lucerne seed			+	+				
Plantago		+	+		+	+	+	
Quince seed		+			+	+		
Traga-canth		+	+		+	+		+
Ulmus fulva		+	+				+	

* Methoxy-derivative.

The presence of several different residues introduces additional complications into structural determinations, since questions concerning the order of arrangement of residues arise as with proteins. Progress has been made, however, by taking advantage of the fact that the different portions of the gum molecule are removed by hydrolysis at different rates. Usually the arabofuranose residues, which form a kind of outer layer, come off first, leaving a degraded gum containing the stable nucleus of hexose and uronic acid residues. Considerable information can, therefore, be gained by comparing the hydrolysis products of the fully methylated derivatives



Aldobionic acid from gum arabic.
(6-β-D-Glucuronosido-D-galactose.)
(VII.)



3-D-Galactosido-L-arabofuranose.
(VIII.)

of the original and the degraded gum, points of junction frequently becoming apparent in this way. These results give also the nature of the linkages whereby the various residues are bound in the polysaccharide. Sometimes it is possible to go further than this if conditions can be found for hydrolysis of the gum (or its methylated derivative) to oligosaccharides, for example aldobionic acids or disaccharides, the structures of which reveal at once how two of the residues are mutually

linked. Often, too, further information can be gained by a study of the oxidation of the gum by periodic acid and the identification of residues not attacked by the reagent.

In Table II some results are given for the hydrolysis of methylated derivatives of original and degraded gums. In the case of gum arabic it will be noted that an aldobionic acid (VII)

TABLE II.

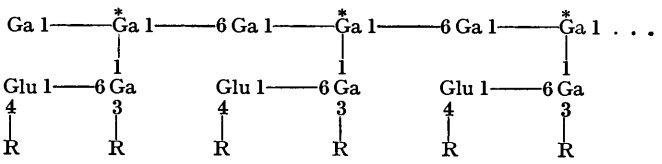
Hydrolysis Products from Methylated Gums.

Gum.	Methylated sugars.	Mode of attachment of residues in gum molecule.
Gum arabic [10]	L-Arabinose, 2, 3, 5	Al
	L-Arabinose, 2, 5	A^1_3
	D-Galactose, 2, 3, 4, 6	G_1
	D-Galactose, 2, 4	$6G^1_3$
	L-Rhamnose, 2, 3, 4	Rhl
	D-Glucuronic acid, 2, 3	Gl^1_4
Degraded gum arabic [11]	D-Galactose, 2, 3, 4, 6	Gl
	D-Galactose, 2, 3, 4	G^1_6
	D-Galactose, 2, 4	$6G^1_3$
	D-Glucuronic acid, 2, 3, 4	Gll
Damson gum [12]	L-Arabinose, 2, 3, 5	Al
	L-Arabinose, 2, 3	A^1_5
	D-Xylose, 2, 3, 4	Xl
	D-Galactose, 2, 4, 6	G^1_3
	D-Galactose, 2, 4	$6G^1_3$
	D-Galactose, 2	$^4_6G^1_3$
	D-Galactose, 4	$^3_6G^1_2$
	D-Mannose (dimethyl)	M?
	D-Glucuronic acid, 2, 3, 4	Gll
	D-Glucuronic acid, 2, 3	Gl^1_4
Degraded damson gum [13]	D-Galactose, 2, 3, 4, 6	Gl
	D-Galactose, 2, 3, 4	G^1_6
	D-Galactose, 2, 4, 6	G^1_3
	D-Galactose, 2, 4	$6G^1_3$
	D-Xylose, 2, 3, 4	Xl
	D-Mannose (dimethyl)	M?
	D-Glucuronic acid, 2, 3, 4	Gll
	D-Glucuronic acid, 2, 3	Gl^1_4
Cherry gum [14]	L-Arabinose, 2, 3, 5	Al
	L-Arabinose, 2, 5	A^1_3
	D-Mannose (positions?)	M?
	D-Galactose, 2, 4, 6	G^1_3
	D-Galactose, 2, 4	$6G^1_3$
	D-Glucuronic acid, 2, 3, 4	Gll
	D-Glucuronic acid, 2, 3	Gl^1_4
	(Methylated D-xylose)	X?
Egg plum (degraded) [15]	D-Galactose, 2, 3, 4, 6	Gl
	D-Galactose, 2, 3, 4	G^1_6
	D-Galactose, 2, 4, 6	G^1_3
	D-Galactose, 2, 4	$6G^3_1$
	D-Glucuronic acid (?)	Gl?
	(Aldobionic acid is 6-glucuronsidogalactose)	
Mesquite [16]	L-Arabinose, 2, 3, 5	Al
	L-Arabinose, 3, 5	A^1_2
	D-Galactose, 2, 4	$6G^1_3$
	D-Glucuronic acid, 2, 3, 4	Gll
	(Also a dimethyl methylhexoside)	
Tragacanthic acid [17]	D-Xylose, 2, 3, 4	Xl
	D-Xylose, 3, 4	X^1_2
	L-Fucose, 2, 3, 4	Fl
	D-Galacturonic acid, 2, 3	Gal^1_4
	D-Galacturonic acid (monomethyl)	Gal?

(Gum tragacanth contains also a neutral araban.)

[A = L-Arabofuranose, G = D-Galactopyranose, M = D-Mannopyranose, X = D-Xylopyranose, Rh = L-Rhamnopyranose, F = L-Fucopyranose, Gl = D-Glucuronic acid (pyranose), Gal = D-Galacturonic acid (pyranose)].

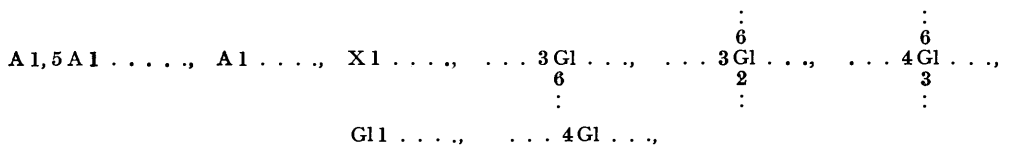
having the structure of 6-glucuronosidogalactose [18] reveals part of the gum's molecular architecture. The arabinose residues are present either as end groups united only by their reducing groups to the rest of the gum (A1 . . .) or doubly involved by links at C_1 and C_3 (. . . 3A1 . . .). Strangely enough, damson gum has an entirely different aldobionic acid residue (2-glucuronosidomannose), and its non-terminal arabofuranose residues are of the type . . . 5A1 . . . Returning for a moment to gum arabic and continuing the account of F. Smith's work on this polysaccharide, one may note that further insight into its structure was gained by the discovery of a disaccharide (3-galactosidoarabinose) [19] (VIII) as a product of partial hydrolysis. Armed with these various items of information Smith has been able to give the general outline of the structure and to eliminate all but a few closely similar alternative formulations.[20] The structure suggested is shown in (IX) where the backbone or main chain of galactose residues with 1:3 and 1:6 linkages—a common feature in plant gum structures—is apparent together with the various groupings attached as side chains.



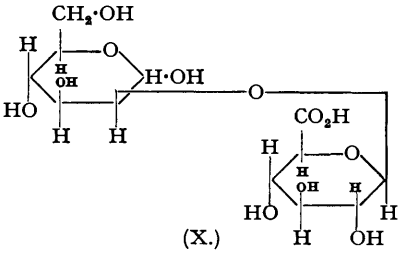
(IX.) Smith's formula for gum arabic.

Ga = D-Galactopyranose, Glu = D-glucuronic acid. The unspecified linkages at Ga* are either 3 and 6, or 6 and 3 respectively, but precise assignment is not yet possible. The R groups are (1) A1, where A is L-arabofuranose, (2) Rh1, where Rh is L-rhamnopyranose, and (3) Ga 1—3A1. The individual points of attachment of the various R groups are at present uncertain.

Damson gum is decidedly more complicated, as will be seen from Table II, both with respect to the degraded and to the undegraded gum, and methylation studies have revealed that the following residues are incorporated in the molecular structure :



together with mannose residues of as yet unknown linkage (where A = L-arabofuranose, X = D-xylopyranose, G = D-galactose, and Gl = D-glucuronic acid, linked as shown). Here again the pentose residues are present in the outer structure of the branched molecule and are readily removed. It is probable also that the main backbone of the molecule is a chain of D-galactose residues with 1 : 3- and 1 : 6-linkages. The aldobionic acid (X), which comprises



(X.)

Aldobionic acid from damson gum/cherry gum.

one of the most stable structural elements is 2-glucuronosidomannose, which is present also in cherry gum. No detailed formula for these gums can, however, be advanced at present, and it is clear that new methods of attack will be required before the problem can be fully solved.

Information is gradually being accumulated concerning other gums, and although it is clear that they vary considerably in structure, the possession of easily removable pentose residues is a general feature. Some of the observations are summarised in Table II which gives the hydrolysis products isolated from some methylated gums and the points of attachment of the corresponding sugar residues in the gum.

Mesquite gum [16] is a particularly interesting case. This contains arabinose, galactose, and a uronic acid containing a methoxyl residue. On gentle hydrolysis only the labile arabofuranose residues are eliminated. Two aldobionic acids, both of them methylated, have been isolated (6-glucuronosidogalactose and 4-glucuronosidogalactose). The main structure to which the pentose residues are attached is a branched chain of galactose and methylated glucuronic acid residues, with the latter linked to C_3 and C_6 of galactose residues.

Gum tragacanth,[17] the exudate of *Astragalus* species (leguminosæ), is a complex mixture, comprising an araban and an acidic polysaccharide tragacanthic acid. The latter contains residues of the pentose, D-xylose (some attached as end groups), and the methylpentose, L-fucose, which is more commonly found in seaweeds. This again occurs, partly at least, as an

TABLE III.

Hydrolysis Products from Methylated Mucilages

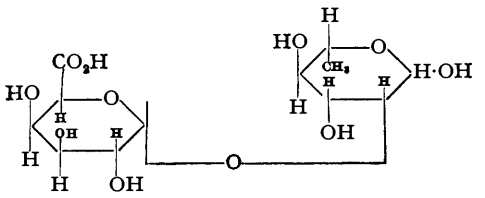
Mucilage.	Methylated sugars.	Mode of attachment of residues in mucilage molecule.
Salep mannan [21]	D-Mannose, 2, 3, 6	M^1_4
	D-Mannose, 2, 3, 4, 6	M1
Gum gatto [22] (locust bean) ...	D-Galactose, 2, 3, 4, 6	G1
	D-Mannose, 2, 3, 6	M^1_4
	D-Mannose, 2, 3	$6M^1_4$
Lucerne seed [23]	D-Galactose, 2, 3, 4, 6	G1
	D-Galactose, 2, 4, 6	G^1_3
	D-Mannose, 3, 4	$6M^1_2$
Plantago lanceolata [24] (seed)	D-Galactose, 2, 3, 4, 6	G1
	D-Galactose, 2, 4, 6	G^1_3
	D-Xylose, 2, 3, 4	X1
	D-Xylose, 2, 4	X^1_3
	D-Xylose, 2, 3	X^1_4
	D-Xylose, 2	$4X^1_3$
	D-Xylose, 3	$4X^1_2$
	D-Xylose unmethylated	$^3_4X^1_2$
(Also methylated derivatives of L-arabinose, L-rhamnose, and D-galacturonic acid.)		
Plantago ovata [25] (seed)	L-Arabinose, 2, 3, 5	A1
	D-Xylose, 2, 3, 4	X1
	D-Xylose, 2, 4	X^1_3
	D-Xylose, 2	$4X^1_3$
	D-Xylose, 3	$4X^1_2$
	D-Xylose, unmethylated	$^3_4X^1_2$
(Methylated D-galacturonic acid and methylated L-rhamnose, combined as galacturonosido-2-rhamnose.)		
Ulmus fulva [26] (bark)	D-Galactose, 2, 3, 4, 6	G1
	D-Galactose, 2, 4, 6	G^1_3
	D-Galactose, 2, 3, 6	G^1_4
	L-Rhamnose, 3, 4	Rh^1_2
	L-Rhamnose, 4	$3Rh^1_2$
	L-Rhamnose, unmethylated	$^3_4Rh^1_2$
	D-Galacturonic acid, 2, 3, 4	Gal 1
	D-Galacturonic acid, 2, 3	Gal $\frac{1}{4}$

end group. The uronic acid is D-galacturonic, and the occurrence of this in place of D-glucuronic acid stresses the peculiarity of this gum and aligns it rather with the plant mucilages.

No doubt many such intermediate substances will be encountered, but on the whole it will be seen from Tables I and III that D-galacturonic acid tends to be characteristic of the plant mucilages. Some of these, for example salep mannan [21] (which in structure recalls ivory-nut mannan,[27] having chains of 1 : 4-linked β-D-mannose residues), the mannogalactan gum gatto (locust-bean mucilage),[22] and the mannogalactan of lucerne seed,[23] contain no pentose residues. It is clear, therefore, that the mucilaginous properties of these substances are not necessarily bound up with their pentose components. Mannogalactans are indeed of common occurrence (guar, Kentucky coffee bean, honey locust, foenugreek, gum tragon, to name only a few), but, nevertheless, mucilages in general are much more complex, and there is a pronounced tendency for D-galactose, D-galacturonic acid, and L-arabinose residues to occur together. The simultaneous occurrence of D-xylose and D-galactose is less common, but is notable in the case

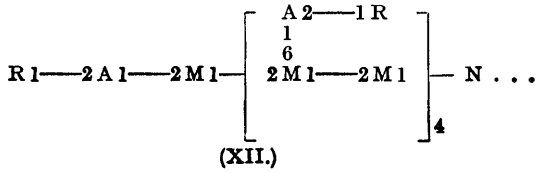
of the *Plantago* mucilages (Table III). These mucilages are remarkable for the complexity of their structure, containing galactose, galacturonic acid, xylose, and possibly both arabinose and methylpentose residues. The xylose residues are combined in the mucilage in the most diverse ways. Examination of the methylated mucilage from the seeds of *Plantago lanceolata* [24] has revealed that xylose residues are present linked through C_1 (end groups), through C_1 and C_3, through C_1 and C_4, through C_1, C_3, and C_4, through C_1, C_2, and C_4, and through C_1, C_2, C_3, and C_4. Somewhat similar results were obtained with *P. arenaria*,[28] the mucilage of which gives rise to an unusual aldobionic acid which is a D-galacturonosido-D-xylose, whilst *P. ovata* yields D-galacturonosido-2-L-rhamnose.[25]

Many more examples of mucilages could be given in which xylose, arabinose, galactose, and galacturonic acid residues occur together (*e.g.*, cress seed, white mustard seed), but it must

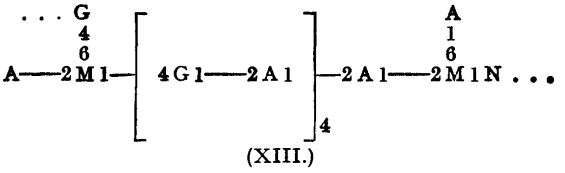


Aldobionic acid from slippery-elm and flax-seed mucilage.
(XI.)

suffice to remark upon a tendency for the pentose residues to be replaced wholly or in part by methyl pentose (L-rhamnose). This is seen in flax-seed mucilage [29] (residues, D-xylose, L-rhamnose, L-galactose, and D-galacturonic acid, with D-galacturonosido-2-L-rhamnose as the aldobionic acid), where the occurrence of L-galactose is noteworthy and is reminiscent of agar. L-Rhamnose is found also as a major constituent of the mucilage from the bark of *Ulmus fulva* [26] (slippery elm), where it occurs with D-galactose, D-xylose, and D-galacturonic acid. The aldobionic acid is D-galacturonosido-2-L-rhamnose (XI), and this mucilage is remarkable for containing residues of L-rhamnose in which every available hydroxyl group is involved in linkages with other residues. Very few examples of such complete substitution are known, but it occurs also in the *Plantago* mucilages (xylose) and in D-mannose residues present in the polysaccharide constituent of ovomucoid.



(XII.)



(XIII.)

We have already noticed an exception to the rule that one only of the enantiomorphs of each sugar is favoured in Nature in the occurrence of L-galactose in certain mucilages in place of the normal D-galactose. An even more remarkable case has recently come to light in Heidelberger's study of the polysaccharides associated with *Mycobacterium tuberculosis*.[30] From the somatic part of the bacterium a polysaccharide has been isolated which contains D-arabinose. It is found in this important biological polysaccharide together with L-rhamnopyranose, D-mannopyranose, and D-glucosamine (all normal stereochemically with the exception of the D-arabinose). An indication (XII) of the type of structural unit present has been given by Sir Norman Haworth and Professor Stacey.[31] D-Arabinose residues are found also in a polysaccharide which is associated with the waxes of *M. tuberculosis*, the accompanying residues in this case being D-galactopyranose, D-mannopyranose, and D-glucosamine (XIII).[32] Clearly this appearance of D-arabofuranose residues in so active a bacterium is highly significant, and the discovery may lead to important developments in our knowledge of metabolic processes.

These examples show how widely distributed in Nature the pentose sugars are and how important a part they play as component structural units of many complex polysaccharides. Their mode of occurrence reveals such a bewildering variety of links both with themselves and with other sugar residues that generalisations are almost impossible to discern at the present stage of investigation. Some regularities do nevertheless emerge. There are, for example, preferred types of ring structure assumed by the various pentoses in combining to give polysaccharides. For D-xylose this is the pyranose modification, and no evidence has yet been obtained of the occurrence of xylofuranose residues in hemicelluloses, gums, or mucilages. But these xylopyranose residues can take part in polysaccharide formation in an astonishing variety of ways, sometimes as end groups linked to other residues only through the reducing groups, sometimes doubly linked through C_2, C_3, or C_4 in addition to C_1 or trebly linked, involving a branched chain, through C_1, C_3, and C_4. Most frequently the residue to which the xylose is directly linked is another reducing sugar, usually xylose, glucose, or galactose, but sometimes xylose is found as one of the residues in an aldobionic acid, the other portion being glucuronic or galacturonic acid.

L-Arabinose is almost equally versatile in its modes of combination, but this sugar is normally encountered in the five-membered ring form, and in the polysaccharides in which it occurs it is present exclusively in the form of arabofuranose residues. This generalisation holds also for the D-variety of the sugar which is encountered in the furanose form as a constituent of the bacterial polysaccharides. It may be noted in passing that the occurrence of arabinose in both the D- and the L-modification is of special interest in view of the close stereochemical relationships between arabinose and galactose which is also found in Nature in both the D- and the L-form. On the other hand xylose and the correspondingly related hexose, glucose, display no such tendency to occur in more than one of the possible enantiomorphic forms, and this applies also to rhamnose and fucose. In the polysaccharides, residues of L-arabinose are linked sometimes exclusively to other arabinose residues as in the highly branched molecule of the polysaccharide araban, which possesses arabofuranose units linked through C_1 in three different ways, namely through C_1 (end group), through C_1 and C_5 (unbranched union), and through C_1, C_2, and C_5 (branched chain) respectively, but more frequently they are attached to other sugar residues, particularly galactose. In the plant gums there are arabinose residues present as end groups linked only through C_1, residues forming part of unbranched chains with links involving C_1 and C_5, and C_1 and C_3, and residues at which branching of the chain occurs with links involving three points of junction. The arabinose units are mostly, if not entirely, in the outer portions of the gum molecule, and we have at least a hint from the little that is so far known of the structural chemistry of the gums that many of these polysaccharides contain as backbone a chain of galactopyranose residues linked through their 1, 6, and 1 : 3 positions to which are attached uronic acid and pentose residues in the form of side chains.

Ribose and its 2-deoxy-derivative occupy a somewhat isolated position in our review. They occur in the D-forms, and so far as is known always in the furanose ring modification. They are the special sugars of the important group of nucleic acids, and seem to disdain admixture with their commoner relatives, with the result that little or nothing can be said regarding possible phytochemical inter-relationships with either the pentose or the hexose series of sugars. One faint hint may perhaps be found in the frequent occurrence of the 2-deoxy-form of ribose. The carbon atom involved here is the labile C_2 which can so readily undergo a Walden inversion and there may be, therefore, just a possibility that D-ribose is related *via* D-arabinose to L-galactose as its ultimate source.

The methyl pentoses, L-rhamnose and L-fucose, offer equally puzzling problems. Rhamnose occurs as the L-form, and so far as is known at present always as a pyranose residue. Its stereochemical pattern places it with L-mannose, but despite much speculation nothing is known with certainty regarding its biochemical relationship to other sugars. It is found in some polysaccharides united through its second carbon atom to the reducing group of galacturonic acid. It is curious, but whether significant or not cannot be estimated at present, that mannose which is closely related stereochemically to rhamnose is found in both cherry gum and damson gum united through its C_2 position to a uronic acid residue. L-Rhamnose also occurs doubly linked to other units through the C_1 and C_2 positions and trebly through C_1, C_2, and C_3, as for example in slippery-elm mucilage. Very little, however, is yet known of the range and variety of the linkages in which rhamnose can participate in plant mucilages, or of the mode of junction of fucose units in seaweed polysaccharides, and it is only too clear that much exploratory work remains to be done in these fields.

One general conclusion which can be drawn with reasonable certainty from the knowledge

gained concerning the molecular structures of the gums, pectins, and mucilages is that the transformation into pentosans of polysaccharides built up of hexose residues cannot take place in the plant at the polysaccharide level by way of oxidation at the C_6 position of the hexose residues with formation of uronic acid units, which on decarboxylation would yield pentose residues. The structures of the galactans, for example, differ so fundamentally from those of the gums and polyuronides based on D-galacturonic acid, and the latter are so far removed structurally from the arabans and from the L-arabinose residues present in the gums, that it is necessary to regard each polysaccharide as built up unit by unit by its own enzyme system. No such simple change, at the polysaccharide level, as galactan $\longrightarrow$ polyuronide $\longrightarrow$ araban is possible. The position rather seems to be analogous to that observed for starch in the course of the work of Hanes and Peat, where the transformation of amylose into amylopectin involves degradation followed by re-synthesis. It is certain, however, that for the gums and mucilages the changes necessary are much more drastic, involving the complete disintegration of one polysacharide, which may perhaps be starch, with transformation of hexoses, by way of the uronic acids, into pentoses, followed by re-synthesis. In view of the frequency with which D-galactose, D-galacturonic acid, and L-arabinose are encountered together in this group it is difficult to avoid the conclusion that L-arabinose does in fact stand in close phytochemical relationship to D-galacturonic acid and D-galactose (but we have to confess complete ignorance of the biochemical inter-relationships of D-glucose and D-galactose). Nothing is known, moreover, of the nature of the enzyme systems responsible for the synthesis of the gums and mucilages, but in view of the mass of evidence which has been accumulating in recent years concerning the importance and versatility of hexose-1 phosphates and phosphorylase enzyme systems in the synthesis of starch and glycogen (Cori, Hanes, Peat) and various disaccharides (Hassid), it would not be surprising to find similar mechanisms in operation here also.

If it is true that few generalisations can yet be made covering the occurrence and the mode of linkage of the various pentose sugars in natural products, still less is it possible to draw general conclusions concerning the functional significance of these substances in the economy of Nature. At first sight it might be thought that the bewildering variety of sugar residues and the complexity of their arrangement in the gums might denote a haphazard piecing together of unwanted molecular species, yet such evidence as is available has led my colleague Dr. J. K. N. Jones to the view that the composition of the individual gums is so specific and invariable that a knowledge of the sugars involved and of their quantitative relationships, such as may now be obtained rapidly by the use of the paper chromatogram, may suffice to identify the botanical species or even the variety of the plant which produced the gum.

Many of these gums are produced by the plant in response to a stimulus resulting from injury, and once formed they serve as a protective layer which seals off the living portion of the plant.[33] The movement of the gummy material can be traced, and its formation, in some cases at any rate, is said to coincide with cytoplasmic changes which involve the disappearance of starch granules. The quantity of gum is on occasions so great that it appears likely that the plant can mobilise the necessary carbohydrate materials at or near the point of injury. The inference is that the actual precursor of the gum may be glucose (or fructose), but no hint has yet been obtained of the mechanism of the changes which give rise to the arabinose, xylose, galactose, mannose, and uronic acid residues which are built up into the molecular structure of the gum. Furthermore, the particular functions served by the pentose residues remain quite obscure unless perhaps it is that their presence renders the protective mass less susceptible to the ordinary forms of attack by micro-organisms. There is, on the other hand, no convincing evidence that the plant gums themselves have their origin in the synthetic activities of micro-organisms which have invaded the plant tissues.

Little more is known concerning the part played in plant metabolism by the pentose residues present in the pectins and mucilages. Very often these materials appear to be part of the food reserve of the plant, but it is not clear what special function is to be attributed to the pentose constituents when these reserves are mobilised by an initial process of hydrolysis for use by the growing plant. One has to envisage indeed a wide variety of activities in which the pentose components appear to play an essential part. For example it is clear that very different metabolic roles are played by the pectic substances which occur in the middle lamella of flax fibres, by those present in the seed of the ground nut, and by the pentose residues in the mucilages which occur in the outer coat of quince or cress seeds and may be concerned in the maintenance of moist conditions round the seed by virtue of their power to absorb water. Very different again from any of these are the functions of the ribose or deoxyribose residues in the nucleic acids, and here the very special and highly specific occurrence of ribose as the carbohydrate of

this important group of substances presents one of the most baffling of problems, and one for which no solution can as yet be offered. The recent *X*-ray studies of Astbury and his colleagues on the structure of the macromolecules of the nucleic acids may perhaps give hints concerning the possible lines of inquiry here arising out of the conditions which have to be satisfied in the geometrical arrangement of the residues in the polynucleotides. It will be recalled that, in commenting on the relationship between the spacings in the proteins and those in the nucleic acids, Astbury referred to the flatness of the furanose ring of the ribose residue and its importance in the general building up of the long chains of nucleotides. But the chemist is unable to say why the dimensions of ribose residue or its configuration should be particularly favourable in this respect. Further investigations into the chemistry of ribose may give a clue, but for the present we must leave the problem there.

References.

[1] Hirst and Purves, *J.*, 1923, 1352.
[2] Von Euler and Schlenk, *Z. physiol. Chem.*, 1937, **246**, 64.
[3] Howard, Lythgoe, and Todd, *J.*, 1947, 1052.
[4] Astbury, *Symp. Soc. Exp. Biol.*, 1947, **1**, 66.
[5] Ehrlich and Schubert, *Biochem. Z.*, 1928, **203**, 343.
[6] Hirst and Jones, *J.*, 1938, 496; 1939, 452, 454, 1865; 1947, 1221.
[7] Hampton, Haworth, and Hirst, *J.*, 1929, 1739.
[8] Haworth, Hirst, and Oliver, *J.*, 1934, 1917.
[9] Chanda, Hirst, Jones, and Percival, unpublished.
[10] Smith, *J.*, 1940, 1035.
[11] Smith, *J.*, 1939, 744, 1724; Jackson and Smith, *J.*, 1940, 74, 79.
[12] Hirst and Jones, *J.*, 1946, 506.
[13] Hirst and Jones, *J.*, 1938, 1174; 1939, 1482.
[14] Jones, *J.*, 1947, 511.
[15] Hirst and Jones, *J.*, 1947, 1064; 1948, 120.
[16] White, *J. Amer. Chem. Soc.*, 1946, **68**, 275; 1947, **69**, 622, 2264; 1948, **70**, 367; Cunneen and Smith, *J.*, 1948, 1141, 1146.
[17] James and Smith, *J.*, 1945, 749.
[18] Challinor, Haworth, and Hirst, *J.*, 1931, 258.
[19] Smith, *J.*, 1939, 744.
[20] Jackson and Smith, *J.*, 1940, 79; Smith, *J.*, 1940, 1035.
[21] Klages and Niemann, *Annalen*, 1936, **523**, 224.
[22] Hirst and Jones, *J.*, 1948, 1278; F. Smith, *J. Amer. Chem. Soc.*, 1948, **70**, 3249.
[23] Hirst, Jones, and Walder, *J.*, 1947, 1225.
[24] Percival and Willox, unpublished; Mullan and Percival, *J.*, 1940, 1501.
[25] Laidlaw and Percival, unpublished.
[26] Anderson, *J. Biol. Chem.*, 1934, **104**, 163; Gill, Hirst, and Jones, *J.*, 1939, 1469; 1946, 1025; Hirst and Jones, unpublished.
[27] Lüdtke, *Annalen*, 1927, **456**, 201; Klages, *ibid.*, 1934, **509**, 159; 1934, **512**, 185.
[28] Nelson and Percival, *J.*, 1942, 58.
[29] Anderson and Crowder, *J. Amer. Chem. Soc.*, 1930, **52**, 3711; Anderson, *J. Biol. Chem.*, 1933, **100**, 249; Tipson, Christman, and Levene, *ibid.*, 1939, **128**, 609; Anderson and Lowe, *ibid.*, 1947, **168**, 289.
[30] Heidelberger and Menzel, *J. Biol. Chem.*, 1937, **118**, 79.
[31] Haworth, Kent, and Stacey, Abs. Chicago Meeting, Amer. Chem. Soc., 1946; *J.*, 1948, 1211.
[32] *Idem*, *J.*, 1948, 1220.
[33] Swarbrick, *J. Pomology*, 1927, **6**, 137.