Some Aspects of the Chemistry of Nucleotides.

THE TILDEN LECTURE, DELIVERED BEFORE THE CHEMICAL SOCIETY ON FEBRUARY 18TH, 1943.

By J. MASSON GULLAND, M.A., D.Sc., Ph.D.

IN a lecture to the Society five years ago (Gulland, J., 1938, 1722) an attempt was made to review in outline the chemical knowledge which had at that time accumulated on the nucleic acids and related nucleotides. To-day it is far beyond the scope of a lecture to encompass the fields of research in which nucleotides play a part, since one of the striking developments of the intervening period has been the recognition of the widespread and fundamental intervention of nucleotides and polynucleotides (nucleic acids) into biological processes, as outlined below.*

The demonstration of the presence of deoxypentose nucleic acids in chromosomes and of pentose nucleic acids (nucleotides) in nucleoli, and the establishment of the existence of cycles of these acids in these tissues during cell divisions (Caspersson, *Skand. Arch. Physiol.*, Suppl., 1936, **73**, No. 8; *J. Roy. Microscop. Soc.*, 1940, **60**, 8; Caspersson and Schultz, *Proc. Nat. Acad. Sci.*, 1940, **26**, 507) have given new conceptions of these processes and of the nature of the structures involved in them (Darlington, *Nature*, 1942, **149**, 66). The recognition of deoxypentose nucleic acids as highly anisometric molecules or micelles (Signer, Caspersson, and Hammarsten, *Nature*, 1938, **141**, 122; Greenstein and Jenrette, *J. Nat. Cancer Inst.*, 1940, **1**, 77, 91) with a nucleotide spacing of the same magnitude as that between the amino-acid side chains of an extended polypeptide has given rise to the belief that the nucleic acid and polypeptide molecules run lengthwise along the chromosomes, united through their phosphoric acid groups and basic side chains respectively (Astbury and Bell, *Nature*, 1938, **141**, 477; *Cold Spring Harbor Symposia Quant. Biol.*, 1938, **6**, 109, Schmidt, *Naturwiss.*, 1938, **26**, 413).

The viruses are consistently found to be nucleoproteins as each is isolated and characterised, the majority so far being derivatives of pentose nucleic acids (Hoagland, *Ann. Rev. Biochem.*, 1943, **12**, 615; Stanley, *ibid.*, 1940, **9**, 545). As nucleoproteins, they have a general chemical relationship with the chromosomes, and it is notable that they also share the unique faculty of self-perpetuation, and thus stand on the threshold between the living and the never-having-lived. This attribute is specially marked in the case of the viruses, some of which have been separated from the living cell as crystalline, homogeneous compounds, which can be submitted to investigation much as other chemical substances, but when re-introduced into other living cells produce symptoms of their presence identical with those induced before their original segregation.

Allied to the phenomena just described are the relationships of nucleic acids with tumour cells. Koller (*Nature*, 1943, **151**, 244) concludes from a cytological analysis of such cells that all chromosome abnormalities and the increased rate of division itself can be explained by the assumption of a qualitative change in the nucleic acid synthesis due to alteration in the heterochromatic region of the chromosomes. The ultra-violet absorption spectra of the cytoplasm of active tumour cells show accumulations of pentose nucleotides in amounts greater than those occurring in the corresponding healthy cells (Caspersson, Nyström, and Santesson, *Naturwiss.*, 1941, **29**, 29). Claude (*J. Exp. Med.*, 1935, **61**, 27, 41; 1937, **66**, 59; *Amer. J. Cancer*, 1937, **30**, 742; 1939, **37**, 59; *Science*, 1937, **85**, 294; 1938, **87**, 467; 1939, **90**, 213) has isolated from Rous chicken sarcoma extracts an extremely active, tumour-producing fraction which consists of a pentose-containing nucleoprotein in association with a phospholipoid component. He (*Proc. Soc. Exp. Biol. Med.*, 1938, **39**, 398; *Science*, 1940, **91**, 77; see also *Trans. New York Acad. Sci.*, 1942, **4**, 79) has also isolated from chick and mammalian embryos a similar fraction having no tumour-producing activity, and the question arises " whether the main constituents of the purified chicken tumour fraction represent inert elements existing also in normal cells; or whether the substance found in normal chick embryo tissue represents a precursor of the tumour principle, which could assume, under certain conditions, the self-perpetuating properties of the tumour agent." Similar phospholipoid–nucleoprotein complexes are found in embryonic tissue (Taylor, Sharp, Beard, and Beard, *J. Infect. Diseases*, 1942, **71**, 115), in embryonic and adult brain (Taylor, Sharp, and Woodhill, *Science*, 1943, **97**, 226), in adult liver (Claude, *Science*, 1940, **91**, 77; *Cold Spring Harbor Symposia Quant. Biol*, 1941, **9**, 263), and in equine encephalomyelitis virus (Taylor, Sharp, Beard, Finkelstein, and Beard, *J. Infect. Diseases*, 1941, **69**, 224; Taylor, Sharp, Beard, and Beard, *ibid.*, 1943, **72**, 31).

Nucleic acids play essential rôles in embryonic development in several directions. Working with the sea-urchin egg, Brachet (*Compt. rend. Soc. Biol.*, 1931, **108**, 813, 1167; 1940, **133**, 90; *Arch. Biol.*, 1929, **39**, 677; 1933, **44**, 519; 1937, **48**, 529; 1940, **51**, 151) formulated the idea that during development the pentose nucleic acid which is present at first is changed into deoxypentose nucleic acid. Needham (" Biochemistry and Morphogenesis," 1942, p. 635, Cambridge Univ. Press) has summarised the ways in which nucleic acid metabolism may proceed, in four types of embryonic behaviour: (1) echinoderms, worms, molluscs, arthropods; simple transformation from stored pentose nucleic acid to deoxypentose nucleic acid: (2) amphibia; transformation, but synthesis of some deoxypentose nucleic acid: (3) teleostean and selachian fishes; store of purine and pyrimidine bases, followed by a synthesis of both forms of nucleic acid: (4) birds and reptiles; complete synthesis of both pentose and deoxypentose nucleic acids from non-cyclic precursors.

Caspersson and Schultz, using their ultra-violet absorption spectra technique, have obtained evidence of a

* It is intended that the two lectures should be read in conjunction,

relationship between acid metabolism and gene reproduction in the salivary chromosomes and the egg cytoplasm of *Drosophila melanogaster* (*Nature*, 1938, **142**, 294) and in cells where rapid division is occurring, including embryonic cells (*Nature*, 1939, **143**, 602; *Proc. Nat. Acad. Sci.*, 1940, **26**, 507).

Fischer (*Nature*, 1939, **144**, 113; *Chem. Prod.*, 1940, **3**, 79; *Acta Physiol. Scand.*, 1942, **3**, 54) has isolated from cow embryo a nucleoprotein, containing both pentose and deoxypentose nucleotides, which is highly effective in promoting the growth of explanted cells.

Damage to cells apparently evokes the production of pentose nucleotides : the cell-free filtrates from yeast cells damaged by ultra-violet light or mechanically have been shown by Loofbourow and his collaborators to contain a growth-promoting factor for yeast, and similar results were obtained with the cells of chick and mammalian embryos. The active factor, which is regarded as a physiological response to injury, is apparently associated with adenine (either alone or in conjunction with guanine), pentose and phosphorus, but not with uracil or deoxypentose, and is thus a pentose nucleotide derivative (Loofbourow and Heyroth, *Nature*, 1934, **133**, 909; Sperti, Loofbourow, and Dwyer, *ibid.*, 1937, **140**, 643; *Stud. Inst. Divi Thomae*, 1937, **1**, 163; Sperti, Loofbourow, and Lane, *Science*, 1937, **86**, 611; Loofbourow, Cook, and Stimson, *Nature*, 1938, **142**, 573; Loofbourow, Cook, Dwyer, and Hart, *ibid.*, 1939, **144**, 553; Loofbourow, Cueto, and Lane, *Arch. exp. Zellforsch.*, 1939, **22**, 607; Loofbourow, Dwyer, and Morgan, *Stud. Inst. Divi Thomae*, 1939, **2**, 137; Loofbourow and Dwyer, *Nature*, 1939, **143**, 725; 1940, **145**, 185; Loofbourow, *Biochem. J.*, 1942, **36**, 737; Cook, Cronin, Kreke, and Walsh, *Nature*, 1943, **152**, 474).

By means of ultra-violet photomicrography, Mitchell (*Brit. J. Exp. Path.*, 1942, **23**, 285, 296, 309) investigated the disturbance of nucleic acid metabolism produced by doses of X- and γ-radiations, and found that an increased ultra-violet absorption by cell cytoplasm, which occurs after irradiation, is caused by an accumulation of pentose nucleotides, which contain adenine and some other as yet unidentified light-absorbing groups.

Finally, it is well known that a pentose nucleotide, muscle adenylic acid (adenosine-5-phosphate), either as such or associated with other nucleotide-like compounds, occurs in cell cytoplasm as the essential co-enzymes of glycolysis, phosphoric acid transfer, hydrogen transfer and amino-acid oxidation.

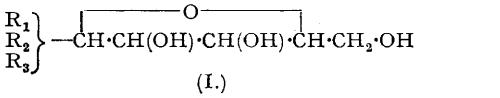
Underlying these spectacular biological advances and forming a unifying background, are the many, diverse chemical problems presented by the nucleotides and polynucleotides, and it is proposed to discuss some of these in this lecture.

The Synthesis of Nucleotides.

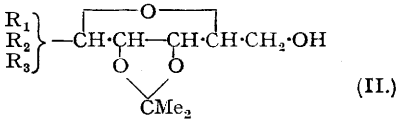
The synthesis of a nucleotide may be regarded conveniently as falling into two stages, the union of the nitrogenous base with the sugar by a glycosidic link to form a nucleoside, and the esterification of the sugar hydroxyl by the phosphoryl group. Methods for the synthesis of the purine nucleosides (Gulland and Story, J., 1928, 259) and the pyrimidine nucleosides (Hilbert, *J. Amer. Chem. Soc.*, 1937, **117**, 330; Hilbert and Rist, *J. Biol. Chem.*, 1937, **117**, 371) were described before the war, but the logical completion of these pieces of research, the synthesis of the four nucleosides themselves, has not been reached, partly no doubt owing to the difficulty of obtaining a sufficient quantity of the necessary sugar, ribofuranose, but also because of the general brake put on academic research by the war. Todd and his collaborators (J., 1943, 383, 386, 571, 574) have recently evolved a new mode of synthesising purine nucleosides, which offers greater scope in the production of analogues of the naturally occurring compounds than do the older syntheses.

Nevertheless, the second stage in the synthesis of nucleotides, the esterification by phosphoryl, has been studied in some detail by re-phosphorylation of the pentose nucleosides prepared by dephosphorylation of the natural nucleotides. These are therefore partial syntheses.

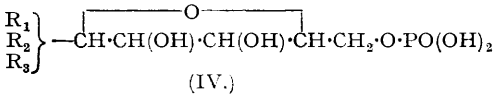
In proceeding along such lines the organic chemist thinks first in terms of protecting those hydroxyl groups of the sugar which are not to be esterified by radicals which can subsequently be removed by some gentle treatment. Adopting this procedure, Levene and Tipson (*J. Biol. Chem.*, 1934, **106**, 113; 1935, **111**, 313; 1937, **121**, 131) converted uridine (I, R_1), inosine (I, R_2) and adenosine (I, R_3) into the respective 5-phosphoesters (IV) by blocking the hydroxyls at C_2 and C_3 with the *iso*propylidene radical (II), phosphorylating the product with phosphorus oxychloride and pyridine (III), and removing the protective group by means of hot dilute acid; fortunately this reagent attacks the glycosidic and phosphoester links only slowly when the phosphoryl esterifies the hydroxyl at C_5, so that the *iso*propylidene radical can be hydrolysed preferentially In a similar fashion these authors (*J. Biol. Chem.*, 1937, **121**, 131) and also Bredereck, Berger, and Ehrenberg (*Ber.*, 1940, **73**, 269) phosphorylated 2 : 3-diacetyl adenosine and obtained adenosine-5-phosphate by removing the acetyl groups.



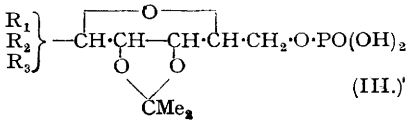
(I.)



(II.)

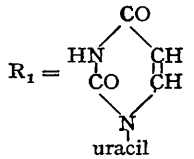


(IV.)

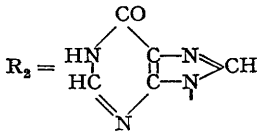


(III.)

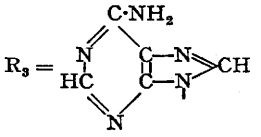
The following symbols are used throughout the lecture :—



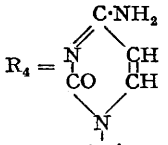
uracil



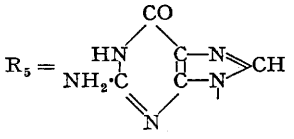
hypoxanthine



adenine



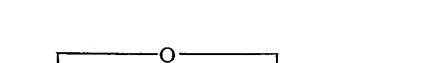
cytosine



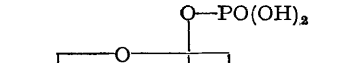
guanine

These nucleotides with phosphoryl at C_5 are, however, isomers of those obtained by hydrolysis of yeast ribonucleic acid, but Bredereck and his collaborators (*Ber.*, 1940, **73**, 269) obtained cytidylic acid (VI) itself, in the form of the brucine salt only, by blocking the hydroxyl at C_5 of cytidine with the trityl group (V), phosphorylating one of the free hydroxyls, and removing the protective trityl group. Esterification presumably occurred selectively at C_3, not at C_2, because the brucine salt isolated was identical with that of cytidylic acid obtained by alkaline hydrolysis of yeast ribonucleic acid; it is assumed that the phosphoryl group in natural cytidylic acid esterifies the hydroxyl at C_3.

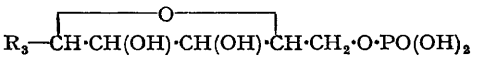
In all these syntheses the position of entry of the phosphoryl group is influenced by the presence of protecting groups previously introduced into the nucleoside, but in 1937 Jachimowicz (*Biochem. Z.*, 1937, **292**, 356) showed that phosphorylation of unprotected adenosine (I, R_3) itself with phosphorus oxychloride and pyridine yielded adenosine-5-phosphate (VII). This is to be expected, in view of the greater reactivity of the hydroxyls of primary, as compared with secondary, alcoholic groups of sugars towards acylating agents.



(V.)

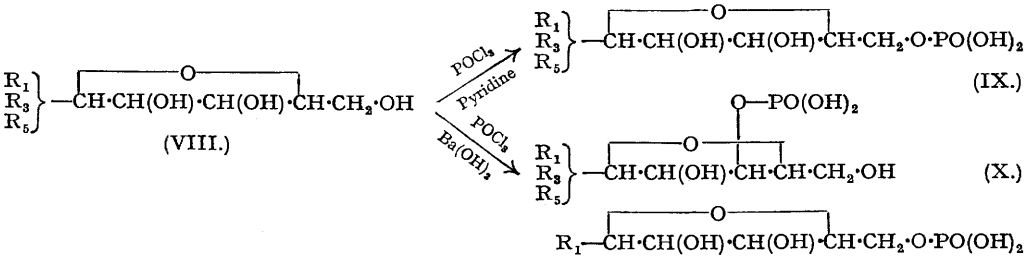


(VI.)



(VII.)

Meanwhile it seemed possible that a mere change of the conditions of phosphorylation might determine the point of entry of the phosphoryl without reliance on protective groups, and this indeed proved to be the case. Almost simultaneously with the synthesis of cytidylic acid by Bredereck, Berger and Ehrenberg, a description was given by Gulland and Hobday (J., 1940, 746) of the phosphorylation of guanosine (VIII, R_5) and uridine (VIII, R_1), a purine and a pyrimidine nucleoside, with phosphorus oxychloride (*a*) in pyridine and (*b*) in saturated aqueous baryta. In pyridine both nucleosides yielded the 5-phospho-esters (IX); this was clear because the ester from uridine was identical with uridine-5-phosphate (Levene and Tipson, *loc. cit.*), and that from guanosine exhibited the greater stability of the phosphoryl towards acid hydrolysis which characterises the 5-phospho-esters of the pentose nucleosides as compared with the 3-phospho-esters. These results were thus in line with Jachimowicz's synthesis of adenosine-5-phosphate. On the other hand, phosphorylation of guanosine with phosphorus oxychloride and baryta yielded guanosine-3-phosphate (X) identical with guanylic acid prepared from yeast ribonucleic acid; adenosine (VIII, R_3) in the same circumstances behaved similarly and gave only adenosine-3-phosphate (X), but no 5-phosphate (Barker and Gulland, J., 1942, 231). Uridine under these conditions formed a mixture of uridine-5-phosphate (IX) and uridine-3-phosphate (X), identical with uridylic acid from yeast ribonucleic acid; these were separated by fractionation of their brucine salts.



Thus, each of the four natural ribose-nucleotides has been prepared from the nucleosides by a part-synthesis, together with three of the four isomeric 5-phosphorylated nucleotides, and methods have been established for introducing the phosphoryl radical at will into either of the two positions which it is known to occupy in Nature.

In speculations on the reasons for the divergent results of these two methods of phosphorylation, it seems possible that phosphorylation of the hydroxyl in the 5-position is simply the interaction of the more reactive and presumably more exposed primary alcoholic group with the acid chloride, the pyridine performing its usual functions in reactions of this type.

In the phosphorylations with baryta, however, the type of reaction may be different. It is evident from consideration of the dissociation constants of sugar hydroxyl groups (Hirsch and Schlags, *Z. physikal. Chem.*, 1929, **141**, *A*, 387; Stearn, *J. Physical Chem.*, 1931, **35**, 2226; Urban and Shaffer, *J. Biol. Chem.*, 1932, **94**, 697; Urban and Williams, *ibid.*, 1933, **100**, 237) and of the titration curves of the nucleotides (Levene and Simms, *ibid.*, 1925, **65**, 519) that partial dissociation of at least one of the sugar hydroxyls of the nucleosides occurs at 0° in saturated baryta solution, the conditions of the phosphorylation. It thus seems likely that phosphorylation of this salt is taking place, and in agreement with this, Fawaz and Zeile (*Z. physiol Chem.*, 1940, **263**, 175) and Neuberg and Pollak (*Ber.*, 1910, **43**, 2060) failed to phosphorylate glucose with phosphorus oxychloride when barium or calcium hydroxide was replaced by calcium carbonate.

It might be thought that the primary alcoholic group would again be the more reactive, but in the relatively small number of cases where attempts have been made to ascertain which hydroxyls are concerned in salt formation, this is not always the case. Percival and his collaborators have shown that methylated sugars or their derivatives may be isolated when methylation is performed on the unstable complexes which are formed with potassium hydroxide by a variety of carbohydrates; from the constitutions of these methylated sugars it is deduced that in some cases primary, in others secondary, hydroxyls of the original sugars were the more active in complex formation with alkali (Percival, J., 1934, 1160; 1935, 648; Percival and Ritchie, J., 1936, 1765; Percival and Heddle, J., 1938, 1690). It seems, therefore, that the relative acidic strengths of the hydroxyl groups in the nucleosides are the factors which determine the position of the ester group in phosphorylation of the nucleotides in aqueous baryta, and that the reaction is analogous with a Schotten–Baumann acylation.

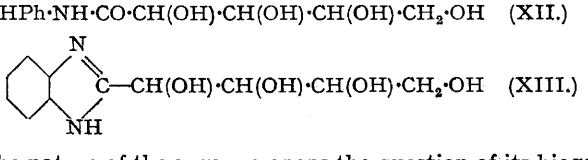
The Nature and Biogenesis of the Pentose of Nucleotides.

During the twenty years ending in 1912 there were many speculations as to the nature of the pentose in inosinic acid and the nucleic acid of yeast and pancreas. The ramifications of the story are complex, but *d*-xylose, *d*-lyxose and *dl*-arabinose were all proposed at various dates, on evidence which cannot always be lightly dismissed, although certainly not conclusive. In 1909, however, Levene and his collaborators isolated a crystalline sugar from inosine, prepared from inosinic acid, and declared this to be *d*(-)ribose. Immediately afterwards this claim was extended to the sugar of guanylic acid from liver, from pancreas and from yeast nucleic acid, and to the adenosine prepared from that acid.

For the past thirty years this conclusion—*d*(-)ribose—has been accepted. There has indeed been a tendency to assume, with no proof, that the sugar of all pentose-containing nucleotides is necessarily *d*-ribose, although examination has been insufficient to permit a statement that this is, or is not, the case; it is of interest in this connection that Loring (*J. Biol. Chem.*, 1939, **130**, 251) has recorded the isolation from the pentose nucleic acid of tobacco mosaic virus of an isomeric uridylic acid, differing from the normal acid in the physical properties of its brucine salt; this may possibly be due to the presence of a different pentose in the virus nucleotide. A review (Barker and Gulland, J., 1943, 625) of the literature, however, suggested that it was desirable to confirm Levene's conclusions, particularly because the nucleotides are the only natural occurrence of the sugar ribose. The pentose nucleic acid of yeast was selected for examination in the first instance, and the sugar components were oxidised to the aldonic acids (XI) and identified as the phenylhydrazides (XII) and as the benziminazole derivatives (XIII) (Moore and Link, *J. Biol. Chem.*, 1940, **133**, 293), prepared by condensation with *o*-phenylenediamine. The results obtained confirm Levene's claim that the sugar of the four nucleotides of yeast nucleic acid is *d*(-)ribose (XIV), and the examination is being extended to the pentoses of other nucleotides and nucleic acids. One point arising from this work is worth mention. *l*-Lyxose (XIX) was isolated in small amount, in the form of the benziminazole derivative, from the hydrolysis products of a commercial sample of yeast nucleic acid, and it is believed that this is the first time that this sugar has been identified in naturally occurring compounds. It is possible that it is not so rare, but that the difficulty of crystallising it or its derivatives has hitherto prevented its recognition; if so, then the benziminazole technique should prove useful, since the separation of the sugar in this way presents little difficulty.

$$CO_2H\cdot CH(OH)\cdot CH(OH)\cdot CH(OH)\cdot CH_2\cdot OH \quad \text{(XI.)}$$

$$NHPh\cdot NH\cdot CO\cdot CH(OH)\cdot CH(OH)\cdot CH(OH)\cdot CH_2\cdot OH \quad \text{(XII.)}$$



(XIII.)

This confirmation of the nature of the sugar re-opens the question of its biogenesis and the form in which it occurs in the nucleotides.

As part of a wider theory of the biogenesis of naturally occurring sugars, Robinson (*Nature*, 1927, **120**, 44, 656) put forward the suggestion that the nucleotides of yeast nucleic acid are really 3-phospho-derivatives of the common sugar *d*-xylose (XV) and that on dephosphorylation (by chemical hydrolysis) Walden inversion

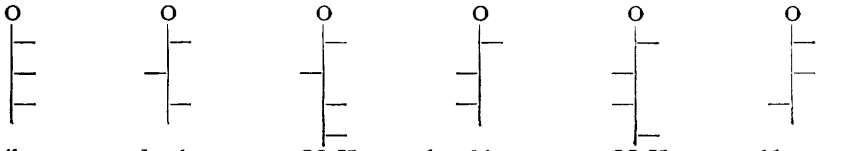
takes place at C_3, yielding the nucleosides as ribosides. He drew support for this suggestion from the possible analogy with the inversion observed (Phillips, J., 1925, **127**, 2552; Kenyon, Phillips, and Taylor, J., 1933, 173) when sulphonic esters are hydrolysed. Levene (*Nature*, 1927, **120**, 620) did not accept Robinson's suggestion. There the matter apparently rested, probably because there was at that time insufficient evidence to proceed further. This has now accumulated from three standpoints and unfortunately does not support the hypothesis, at any rate in its original form.

First, knowledge of the mechanism of the Walden inversion in the hydrolysis of carbohydrate esters has been greatly extended since 1927, and it is known that when inversion occurs during the hydrolysis of the ester, *e.g.*, toluenesulphonyl esters, it is accompanied practically without exception by the formation of an anhydro-sugar, in which is involved an adjacent hydroxyl originally in the *trans*-position to the ester group. Further hydrolysis of this anhydro-ring to reconstitute a fully hydroxylated sugar results in inversion at one or other of its two carbon atoms; the sugars thus formed have the hydroxyl groups once more in the *trans*-position relative to each other. It is evident, therefore, that *d*-ribose, with three *cis*-hydroxyls, is unlikely to be formed by chemical hydrolysis of an ester of *d*-xylose.

Second, an essential condition for the occurrence of a Walden inversion is that a true replacement should occur, that is, that the bond must be broken which joins the asymmetric atom to the group being displaced. This always happens when, for example, a halogen is being replaced by a hydroxyl or an amino-group, but when an ester group is replaced by hydroxyl this condition may, or may not, be fulfilled. Thus, inversion is rare in the hydrolysis of acyl esters, which are therefore regarded as undergoing fission as RO·COR, but is common with sulphonyl esters, which are therefore hydrolysed as $RO{\cdot}SO_2R$. Robinson's suggestion, based on a supposed similarity between phosphoric esters and sulphonic esters, would imply fission of the phosphoric esters as $RO{\cdot}PO(OH)_2$ in order to allow inversion. The most clean-cut evidence in the mechanism of hydrolysis of esters is by the use of water containing an excess of heavy oxygen. Polanyi and Szabo (*Trans. Faraday Soc.*, 1934, **30**, 508) hydrolysed amyl acetate by this means and obtained amyl alcohol containing no heavy oxygen, and Datta, Day, and Ingold (J., 1939, 838) obtained a similar result in the hydrolysis of methyl hydrogen succinate; consequently the isotopic oyxgen atoms in each case passed to the acid and not the alcoholic hydrolysis product, and the reactions occurred as R·O·COR. In the same way, Herbert and Blumenthal (*Nature*, 1939, **144**, 248) hydrolysed trimethyl phosphate, and, since the resulting methyl alcohol was free from isotopic oxygen, the carbon–oxygen bond was not broken, and consequently inversion could not have occurred in the case of a phosphoryl ester of an optically active alcohol instead of methyl alcohol. Thus phosphoric esters resemble acyl esters and undergo chemical hydrolysis without inversion.

Third, the partial syntheses of the nucleotides of yeast ribonucleic acid from the corresponding nucleotides, described at the beginning of this lecture, also show that inversion does not occur during their dephosphorylation. The nucleotides were dephosphorylated by hydrolysis and the resulting nucleosides were rephosphorylated by phosphorus oxychloride. In the latter reaction, between a hydroxyl and an acid chloride, the link between asymmetric carbon and oxygen may be assumed not to be broken (consequently no inversion), and since the synthetic acids were identical with the natural nucleotides, it follows that hydrolysis of the nucleotides must have occured without inversion.

It is clear, then, that the *d*-ribose isolated after hydrolysis of yeast ribonucleic acid and its component nucleotides is present as such in the original molecule, and the same may apply to *l*-lyxose, but with less certainty, since it is not known in what circumstances it was originally present. It is opportune here to consider the biogenesis of these sugars. Several speculations have been advanced to explain the formation of *d*-ribose, which is of special interest in view of its wide-spread occurrence but specific limitation to nucleotides, and of the marked biological importance of these compounds.

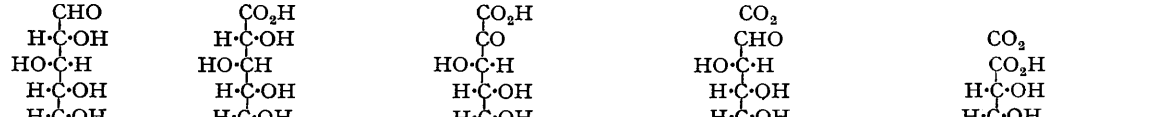


d-ribose	*d*-xylose	CO_2H *d*-glucuronic acid	*l*-arabinose	CO_2H *d*-galacturonic acid	*l*-lyxose
(XIV.)	(XV.)	(XVI.)	(XVII.)	(XVIII.)	(XIX.)

It has long been suggested that the naturally occurring pentoses have their origin in the hexuronic acids, which decompose into pentose and carbon dioxide (see, for example, Hirst, J., 1942, 70; Hirst and Jones, *Chem. and Ind.*, 1937, **56**, 724). This reaction has not been demonstrated in an isolated enzyme system, although Salkowski and Neuberg (*Z. physiol. Chem.*, 1902, **36**, 261) have shown that glucuronic acid is decarboxylated to *d*-xylose by bacteria isolated from putrefying flesh. That the reaction is of wider significance, however, is highly probable from the frequent occurrence of xylose (XV) of the *dextro*-series, arising from *d*-glucuronic acid (XVI), but of arabinose (XVII) of the *lævo*-series which would be formed from *d*-galacturonic acid (XVIII); these uronic acids occur naturally, and the related hexoses are frequently accompanied by the corresponding pentoses. If all pentoses are formed in this straightforward manner, the precursor of *d*-ribose would be *d*-alluronic or *d*-taluronic acid, and of *l*-lyxose would be *l*-mannuronic or *d*-guluronic acid, none of

which is found in Nature. Consequently, if this theory of the formation of pentoses is to be generally accepted, it is necessary to assume that Walden inversion occurs in the formation of *d*-ribose and *l*-lyxose, at C_3 of *d*-glucuronic and *d*-galacturonic acids respectively. In this connection it may be relevant to note that so far ribose is known to occur naturally only when phosphorylated at C_3 (VI and X) (guanylic, adenylic, cytidylic and uridylic acids of pentose nucleic acids) or at C_5 [muscle adenylic acid (VII) and the adenine co-enzymes]; the two exceptions, uric acid riboside (Newton and Benedict, *J. Biol. Chem.*, 1922, **54**, 595; Falconer and Gulland, J., 1939, 1369) and crotonoside (*iso*guanine riboside; Cherbuliez and Bernhard, *Helv. Chim. Acta*, 1932, **15**, 464, 978; Spies and Drake, *J. Amer. Chem. Soc.*, 1935, **57**, 774; Spies, *ibid.*, 1939, **61**, 350; Falconer, Gulland, and Story, J., 1939, 1784), can probably be regarded as dephosphorylated nucleotides. A possible hypothesis might then be that inversion is a necessary consequence of the enzymic phosphorylation at C_3 of these uronic acids. In that case the formation of muscle adenylic acid and its derivatives would be secondary to the production of ribose-3-phosphate, and the enzymic mechanism for this stage from adenosine, is well established (Ostern and Terszakowec, *Z. physiol. Chem.*, 1937, **250**, 155, and subsequent papers), whereas the enzymic phosphorylation of ribose nucleosides at C_3 is not known. Against this hypothesis is the fact that the conversion of 3-phosphoglyceric acid into 2-phosphoglyceric acid by phosphoglyceromutase apparently proceeds without inversion (Kiessling and Schuster, *Ber.*, 1938, **71**, 123, footnote 9); other examples of enzymic phosphorylation of the hydroxyl at an asymmetric carbon atom are lacking.

A route to *d*-ribose from glucose (XX), which has many attractions, has been advanced by Dickens (*Nature*, 1936, **138**, 1057; *Biochem. J.*, 1938, **32**, 1626, 1645), who assumes that the first stage is esterification to glucose-6-phosphate. An enzyme prepared from Lebedew yeast maceration fluid converts this ester by way of phosphogluconic acid (XXI) and 2-ketophosphohexonic acid (XXII) into pentose-5-phosphoric acid (XXIII) and thence by further oxidation into phosphorylated hydroxy-carboxylic acids of lower carbon content (XXIV). The pentose-phosphoric acid has not as yet been isolated as a pure compound, but the presence of pentose has been indubitably demonstrated. In this series of reactions it might be anticipated that the pentose formed would have been *d*-arabinose-5-phosphoric acid (XXIII), but in concurrent experiments it was demonstrated that *d*-ribose-5-phosphate not only underwent further oxidation by the enzyme, but was also fermented considerably more readily than *d*-arabinose-5-phosphoric acid or *d*-xylose-5-phosphoric acid. The theory was therefore advanced that *d*-ribose-5-phosphoric acid may originate from hexoses by a process of phosphorylation, oxidation and decarboxylation, inversion occurring at C_3 of glucose. If this mechanism is accepted, a possible mode of formation of the sugar moiety of adenylic acid and its derivatives seems clear, and the 3-phosphonucleotides of pentose nucleic acids might be secondary products arising through an enzymic phosphorylation which has not as yet been demonstrated satisfactorily *in vitro*.



Other hypotheses involve the aldol condensation of lower sugars, as, for example, the condensation of glycollaldehyde with *d*-glyceraldehyde phosphate, or with dihydroxyacetone phosphate, followed by an immediate rearrangement to the aldose series, postulated by Gardner (*J. Org. Chem.*, 1943, **8**, 111).

It is clear that the problem of the biogenesis of ribose and of the nucleotides can gain little more from speculation and can be advanced only by experimental investigation.

*The Structure of Nucleic Acids.**

No matter from what sources nucleic acids are isolated, and whether animal or vegetable, present knowledge allows them to be classified as belonging to either of two main types, pentose nucleic acids and deoxypentose nucleic acids. This conclusion is based on the structures of the four nucleotides into which each type of acid may be broken by suitable means, the former type by alkaline hydrolysis, the latter by enzymic fission (see Gulland, J., 1938, 1722).

Until relatively recently it was generally assumed that the molecule of each type of acid consists of a tetranucleotide, that is, a molecule formed by union of one of each of the four nucleotides. The reasons for this belief arose chiefly from five causes (see Gulland, J., 1938, 1728, for a more detailed account). First, it was often assumed that the major features of the structures of the acids of both types were the same, and that in consequence conclusions reached from experiments on one acid could be applied directly to considerations of the other type. Second, Jones's ("Nucleic Acids," 1920, p. 40, Longmans Green and Co., London) quantitative hydrolysis of the pentose nucleic acid of yeast showed that it contained equimolecular amounts of purine and pyrimidine nucleotides, and this, in conjunction with the isolation of the two purine and two pyrimidine bases themselves after drastic degradation, apparently clinched the tetranucleotide hypothesis. Third, and this was a powerful factor, Myrbäck and Jorpes (*Z. physiol. Chem.*, 1935, **237**, 159) recorded for the pentose nucleic acid of yeast a molecular weight corresponding roughly to that of a single tetranucleotide. Fourth,

* Space restrictions prohibit a discussion in greater detail than is given here.

data on the titration of both types of acid accumulated independently by Hammarsten, Levene, Makino, Bredereck, and ourselves showed that the acids exhibited approximately either 4 or 5 phosphoric acid dissociations. The presence of five acidic groups, believed to be four primary and one secondary, could only be explained on the basis of a tetranucleotide having an open-chain structure, and this could have arisen by hydrolysis of the other alternative, a cyclic tetranucleotide with four primary acidic dissociations, each deriving from a doubly esterified phosphoryl radical; the latter conception not only fell into line with the results of the molecular weight determinations of Myrbäck and Jorpes, but also agreed with the results of enzymic fission. Fifth, specific enzymes, phosphomonoesterase and phosphodiesterase, were shown by Takahashi (*J. Biochem. Japan*, 1932, **16**, 463) and by Gulland and Jackson (J., 1938, 1492) to fail to cause dephosphorylation of the pentose nucleic acid of yeast when acting separately, whereas a mixture of the enzymes did so; such results favoured the cyclic tetranucleotide.

Despite this evidence pointing to the simple tetranucleotide, it is now clear that the true molecular sizes of both types of nucleic acid are larger, in some cases much larger, and that they vary considerably with the type and source of the acid and also with the method of isolation and treatment during purification; alkali is specially effective in reducing molecular size, undoubtedly because it causes hydrolysis.

Considering deoxypentose nucleic acid first, Signer, Caspersson, and Hammarsten (*Nature*, 1938, **141**, 122) found the molecular weight of the sodium salt of the acid from thymus to be between 500,000 and 1 million from studies of viscosity and double refraction of flow, that is, about 1500 to 3000 nucleotides, since the average weight of a nucleotide is about 330; the length of these particles was 300 times their width. Astbury and Bell (*ibid.*, p. 747) confirmed this size as being approximately 2000 nucleotides from *X*-ray studies of stretched films of the sodium salt. Schmidt, Pickels, and Levene (*J. Biol. Chem.*, 1939, **127**, 251) obtained sedimentation data showing that the nucleic acid of thymus, prepared by different methods, was a mixture of substances ranging in molecular weight from about 1500 to over 1 million, the molecular size being dependent on the mode of preparation. Tennent and Vilbrandt (*J. Amer. Chem. Soc.*, 1943, **65**, 424), also using the ultracentrifuge, found a molecular weight of 500,000 for the sodium salt of the acid from thymus. Characteristic properties of these large molecules are the gel-formation and streaming birefringence shown by aqueous solutions of the sodium salts; the particles are highly asymmetric in shape, very long compared with their width.

On the other hand, using a drastic alkaline method of purification, Bredereck and Jochmann (*Ber.*, 1942, **75**, 395) obtained from the deoxypentose nucleic acid of spleen an acid which may have been a true tetranucleotide; it was soluble in water, unlike the more complex material, had approximately the correct analytical composition and molecular weight for a compound composed of one molecule of each of the nucleotides, had five acidic dissociations, and did not exhibit streaming birefringence or form a gel.

Turning next to the pentose nucleic acids, Loring (*J. Biol. Chem.*, 1939, **128**, Sci. Proc. **33**, 61) used the Northrop–Anson diffusion method of analysis and found molecular sizes of 116 nucleotides for the acid of tobacco mosaic virus and 56—88 nucleotides for the acid of yeast. Later Cohen and Stanley (*J. Biol. Chem.*, 1942, **144**, 589) carried out a detailed study of the physical properties of the acid of tobacco mosaic virus, isolated by heat denaturation of the virus protein and thus avoiding the use of alkali; it is therefore possible that this material was, or approached closely to, the true native nucleic acid of the virus. They studied electrophoresis, partial specific volume, diffusion, sedimentation, viscosity, electron microscopy, osmotic pressure, and optical properties. The freshly isolated acid had an approximate particle weight of 300,000 and was highly asymmetric in shape with axial ratio about 60; it decomposed spontaneously to form still asymmetric, birefringent particles with molecular weight about 61,000. Cold alkali converted each of these particles into molecules with a weight of 15,000 and a length–width ratio of 10. Cohen and Stanley state that there is at present no evidence that a unit smaller than *M* 15,000 and larger than a nucleotide exists as a fundamental unit of the virus nucleic acid. Fischer, Böttger, and Lehmann-Echternacht (*Z. physiol. Chem.*, 1941, **271**, 246) found the molecular weight of the pentose nucleic acid of yeast to be 10,350 by a dialysis method. According to our own measurements (Fletcher, Gulland, Jordan, and Dibben, J., 1944, 30), the molecular weight of the acid from yeast ranges between 10,280 (32 nucleotides) and 23,250 (70 nucleotides), depending on the commercial source from which the samples were originally derived and hence possibly on their treatment during extraction.

Yeast ribonucleic acid readily suffers complete fission to nucleotides by alkaline hydrolysis, but by regulating the conditions so as to avoid this full decomposition, Bredereck and Hoepfner (*Ber.*, 1942, **75**, 1086) obtained what they considered to be the tetranucleotide of this acid. This material had the properties which might be expected, *viz.*, solubility in water, a molecular weight of 1177 (calc., 1304), the correct analytical composition of a molecule built up of one of each of the four nucleotides, and five phosphoric acid dissociations. In addition, this so-styled tetranucleotide was stated to undergo smooth deamination to a deaminated tetranucleotide, which was still pentabasic when titrated to the phenolphthalein end-point, and which yielded xanthine (from guanine) on hydrolysis. Unfortunately for the claim of Bredereck and Hoepfner to have isolated a tetranucleotide unit, it has been shown (Fletcher, Gulland, and Jordan, J., 1944, 33) that theory demands that this deaminated tetranucleotide should have been hexabasic, since the hydroxyl group of xanthine titrates over the range pH 5—8. Further, a careful repetition by us of Bredereck and Hoepfner's hydrolysis experiment yielded material of higher average molecular weight and greater acidity than expected, and it seems necessary to stress our present opinion that both Bredereck and Hoepfner's and our material were mixtures of indefinite composition and that there is no evidence on these grounds for the existence of a pentose tetranucleotide.

The question of molecular size of nucleic acids, true molecular size in contrast to micellar size, is complicated

by a reversible aggregation or polymerisation observed in the case of deoxypentose nucleic acids. Indeed Pedersen (quoted by Svedberg and Pedersen, " The Ultracentrifuge," 1940, p. 443, Oxford Univ. Press) found that the sodium salt of this acid was not monodisperse and had an average molecular weight of the order of 200,000 in dilute solution, but the sedimentation constant was highly dependent on the concentration. With these colloidal aggregates the particle size may be dependent not only on the concentration of the nucleic acid, but also on the concentration of the buffer solution in which it is dispersed. The spontaneous fall in molecular weight of tobacco mosaic virus nucleic acid mentioned above may be a micellar disaggregation of an analogous type.

It is appropriate here to mention the work of Greenstein and Jenrette (*J. Nat. Cancer Inst.*, 1940, **1**, 77). Considerable structural viscosity and double refraction of flow were marked characteristics of all samples of deoxypentose nucleic acid examined from various sources, although the magnitudes of these properties varied appreciably even when these samples were prepared from the same source by the same method. The addition of certain salts to solutions of the sodium salt of this acid resulted in a parallel decrease of viscosity and intensity of streaming birefringence, and at a sufficiently high concentration of the added salt neither property could any longer be observed; guanidinium was the most effective cation, and iodide the most efficient anion. Both birefringence and viscosity were almost completely restored to the nucleic acid by removal of the added salt, either by dialysis or by precipitating the nucleic acid. Greenstein and Jenrette attributed these phenomena to a depolymerisation of large to small particles, which then repolymerised on removal of the salt. No molecular weights, however, were recorded for these so-called smaller particles, and consequently it is difficult to place an exact interpretation on the results, for which alternative explanations are possible.

Numerous investigators have studied the action on both types of nucleic acid of those enzymes which bring about hydrolytic fission to molecules greater than nucleotides without any simultaneous loss of groups or radicals (Dubos, *Science*, 1937, **85**, 549; Dubos and Thompson, *J. Biol. Chem.*, 1938, **124**, 501; Schmidt and Levene, *Science*, 1938, **88**, 172; Schmidt, Pickels, and Levene, *J. Biol. Chem.*, 1939, **127**, 251; Schmidt and Levene, *ibid.*, 1938, **126**, 423; Kunitz, *Science*, 1939, **90**, 112; *J. Gen. Physiol.*, 1940, **24**, 15; Allen and Eiler, *J. Biol. Chem.*, 1941, **137**, 757; Eiler and Allen, *Proc. Soc. Exp. Biol. Med.*, 1941, **46**, 436; Greenstein and Jenrette, *J. Nat. Cancer Inst.*, 1941, **1**, 845; **2**, 301; Fischer, Böttger, and Lehmann-Echternacht, *Z. physiol. Chem.*, 1941, **271**, 246; *J. pr. Chem.*, 1941, **158**, 79; Bolomey and Allen, *J. Biol. Chem.*, 1942, **144**, 113).

For each type of acid there appears to be a specific depolymerising enzyme (known at present as ribonuclease, and deoxyribonuclease or thymonucleodepolymerase respectively), wide-spread in both animal and plant tissues, which rapidly reduces the molecular size, and with it the viscosity and streaming birefringence when these are evident. These enzymes do not cause dephosphorylation. They are highly specific for the particular nucleic acid type, and act neither on a wide range of natural synthetic substances which have been offered as substrates, nor on the other type of nucleic acid; Bredereck and Müller (*Ber.*, 1939, **72**, 1429) believe, however, that one enzyme alone exists and that it acts on both acids. This high specificity has given rise to the suggestion that the natures of the linkages of the tetranucleotides in the two types of nucleic acid are essentially different. Whether in fact they are, or are not, different, a conclusion based on this line of argument alone seems to be scarcely tenable in view of the extremely selective specificity of various enzymes.

The final products in the case of the deoxypentose nucleic acid are apparently tetranucleotides; Fischer and his collaborators showed that the molecular size, measured by a method admittedly incapable of high accuracy, lies between 3·6 and 5·3 mononucleotides and that there is an increase in titratable acidity of one group for every four atoms of phosphorus, a change in the total number of acid groups from four to five. It is not clear, however, whether these tetranucleotides are the same as the material obtained through alkaline hydrolysis by Bredereck and Jochmann.

In the case of the pentose nucleic acid of yeast, Fischer and his collaborators recorded that the products of depolymerisation were dinucleotides with an average molecular weight of 690, whereas Allen and Eiler stated that they were tetranucleotides with an increase in titratable acidity of one group for every four atoms of phosphorus. Further, Allen and Eiler showed that this group is a secondary phosphoryl dissociation because of its titration range, and this implies that the enzymic hydrolysis occurs at a phosphoryl which already has one acidic dissociation. These divergences need further investigation, although it seems possible that Fischer's dinucleotides might be the results of further hydrolysis of the tetranucleotide.

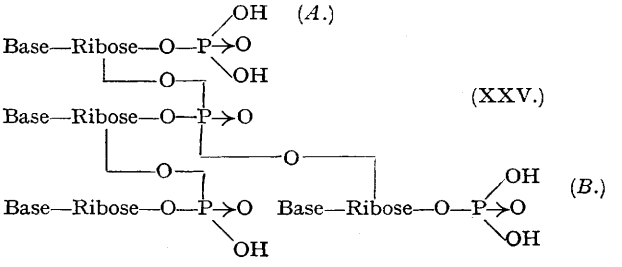
It is thus universally agreed that the molecules of nucleic acids are composed of large numbers of nucleotides, but it must be confessed that remarkably little is known with certainty of the inner structures of these molecules, the modes of combination of these nucleotides. For example, there is as yet no evidence as to whether the nucleic acids of each type are, or are not, identical in structure amongst themselves, irrespective of their source, or even whether there is any general nucleic acid pattern common to all those nucleic acids which belong to one type but are derived from different sources.

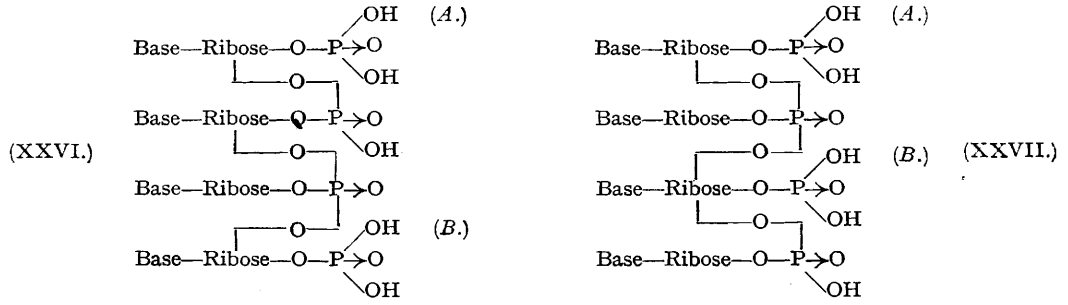
From time to time suggestions have been made that the molecules of deoxypentose nucleic acids and yeast ribonucleic acid are polytetranucleotides, with an assumption that each tetranucleotide unit of the polymer is identical with its neighbours, as regards both the sequence of its component nucleotides and its uniform content of one molecule of each of these; this postulate, however, has not been made for the pentose acid of tobacco mosaic virus, the most fully investigated of the virus acids. It should be realised that the conception of a molecule composed of polymerised tetranucleotides has grown from a mental superposition of the later demonstrations of high molecular weights on to the older ideas of a simple molecule containing one of each of the four

appropriate nucleotides; had the true molecular sizes been realised earlier, it is doubtful whether the conception would have gained such firm hold as is apparently the case.

At present the facts tend to favour the hypothesis of molecules composed of polymerised tetranucleotides in the cases of yeast ribonucleic acid and some examples, at any rate, of deoxypentose nucleic acids, but sound evidence is non-existent either to support the existence of a uniform tetranucleotide composed of one of each of the appropriate nucleotides or to indicate a regular sequence of nucleotides. Turning to pentose nucleic acids in general, there are contrary indications to a uniform content of the four nucleotides. For example, Graff and Maculla (*J. Biol. Chem.*, 1935, **110**, 71) showed that in the ribose acid from yeast the percentage of purines is almost that required for equimolecular proportions of the four nucleotides, but in tobacco mosaic virus nucleic acid Loring (*ibid.*, 1939, **130**, 251) found a 20% deficiency of purines, and in the allonucleic acid of pancreas the guanine nucleotides predominate (see Gulland, J., 1938, 1722).

Nevertheless, the conception of two uniform tetranucleotides for all nucleic acids, one containing pentose, the other deoxypentose, each containing one of each of the four appropriate nucleotides, can form a practical working hypothesis, which is helpful in the interpretation of experimental results, provided its limitations are fully realised.





In support of this hypothesis, in so far as it concerns yeast ribonucleic acid, it is appropriate to mention the results of work (Fletcher, Gulland and Jordan, J., 1944, **33**) on electrometric titrations of samples of this acid having molecular weights of approximately 10,000, 15,000, and 22,000. Comparison of the experimental curves with those constructed theoretically from a knowledge of the dissociation constants of the groups present showed that for each four atoms of phosphorus there were four acidic dissociations of phosphoryl radicals, three primary and one secondary. This result was fully confirmed by similar titration of the deaminated nucleic acid, which can be prepared under particular conditions without lowering of molecular weight. Basing these observations on the assumption of an acid composed of polymerised, uniform, hypothetical tetranucleotides containing one molecule of each nucleotide, it follows that one phosphoryl radical per tetranucleotide must be triply linked and be neutral; it is believed that this is that radical which forms part of uridylic acid. Two dissociations must be primary, associated with a phosphoryl group which is doubly linked to other radicals; and one phosphoryl group can be only singly linked to a ribose radical and must exhibit a primary and a secondary dissociation. Possible formulæ for the tetranucleotide which are in agreement with the observed facts are (XXV), (XXVI) and (XXVII); polymerisation to form the polytetranucleotide would occur either through the group (*A*) or (*B*), the other member of this pair remaining singly linked.

Mild alkaline hydrolysis of the nucleic acid liberates secondary phosphoryl dissociations, and this is confirmed by the hydrolysis of the deaminated acid in boiling water, one secondary phosphoryl dissociation then being liberated for each tetranucleotide; the product has a molecular weight of roughly the correct order for a deaminated tetranucleotide and does not undergo further hydrolysis under the conditions of the experiment. Consequently, when the tetranucleotide of the acid comes to be isolated, it should exhibit three primary and two secondary phosphoryl dissociations, not four and one respectively as always postulated hitherto. It also shows that the fission of the acid into tetranucleotides by chemical means hydrolyses a link involving a secondary acidic group of a phosphoryl radical; the same is true for the enzymic fission by ribonuclease, but it is unknown whether the resulting tetranucleotides are identical and the same phosphoryl radical is concerned in both kinds of hydrolysis.

A further point arising from this work is worth mention. Examination of the analyses in the literature

over many years shows that there is a marked tendency for the figures for phosphorus to be low. On the basis of a polymer of tetranucleotides constituted as hitherto believed, namely, tetranucleotides with four primary and one secondary dissociation united by means of the secondary acidic group, a detachment of phosphorus could only occur by disruption of the nucleic acid molecule, except for the removal of the singly linked terminal radical. No such disruption had apparently taken place, and the extent of deficiency of phosphorus was often greater than could be accounted for by removal of the terminal, singly linked phosphoryl group. On the basis of a polymer made up of the tetranucleotide now proposed, however, it is possible for up to one-quarter of the total phosphorus, as theoretically required, to be absent without affecting the main structure of the polymer, since each tetranucleotide bears one phosphoryl group which is singly linked and plays no part in maintaining the skeletal structure.

I wish to acknowledge my indebtedness to my colleagues Dr. G. R. Barker and Mr. D. O. Jordan, M.Sc., with whom I have discussed many points in this lecture.