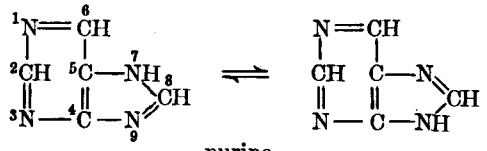
CHAPTER XVI

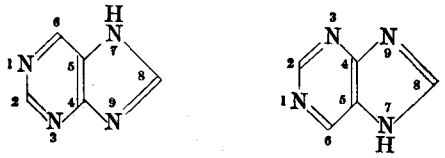
PURINES AND NUCLEIC ACIDS

§1. Introduction. Purine is the parent substance of a group of cyclic diureides and was used by E. Fischer to name systematically the naturally occurring derivatives. Purine exists in two tautomeric forms, and its structure consists of a pyrimidine ring fused to an imidazole ring. In the earlier literature, the formula of purine was written as follows (the method of numbering is also shown):



purine

These formulæ have been written as in A, but more recently, the practice is to write the nitrogen of the pyrimidine ring at the top as in B (*cf.* diazines,



A B

§10. XII). In this book, formula A is used (B is A turned upside down; there is no change in numbering, and so the reader can readily translate A into B).

§2. Uric acid. Guano (birds' excrement found on islands near the western coast of South America) contains up to about 25 per cent. uric acid; about 90 per cent. of snakes' excrement is ammonium urate. Small amounts of uric acid are also present in human urine; it was first discovered by Scheele (1776) in urinary calculi.

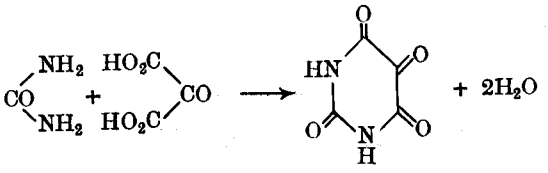
Liebig and Wöhler (1834) showed that the molecular formula of uric acid is $C_5H_4O_3N_4$. These authors also found, in 1838, that the oxidation of uric acid with nitric acid gives alloxan and urea in equimolecular proportions.

$$C_5H_4O_3N_4 + H_2O + [O] \xrightarrow{HNO_2} C_4H_2O_4N_2 + NH_2{\cdot}CO{\cdot}NH_2$$

Structure of alloxan, $C_4H_2O_4N_2$. When hydrolysed with alkali, alloxan produces one molecule of urea and one molecule of mesoxalic acid.

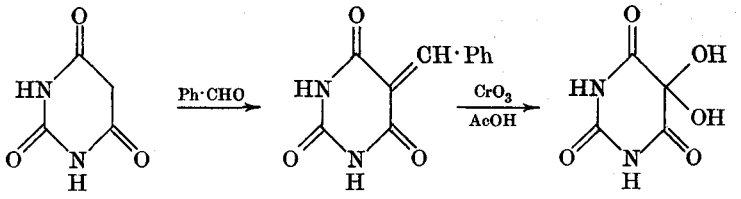
$$C_4H_2O_4N_2 + 2H_2O \xrightarrow{KOH} NH_2{\cdot}CO{\cdot}NH_2 + CO_2H{\cdot}CO{\cdot}CO_2H$$

Since alloxan contains no free amino or carboxyl groups, the products of hydrolysis suggest that alloxan is mesoxalylurea; this cyclic structure has been confirmed by the direct union of urea and mesoxalic acid to give alloxan (Liebig and Wöhler, 1838).



alloxan

Alloxan, as its monohydrate, is conveniently prepared from barbituric acid as follows (see also §13a. XII):

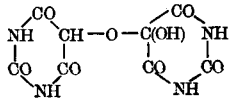


barbituric acid

alloxan monohydrate

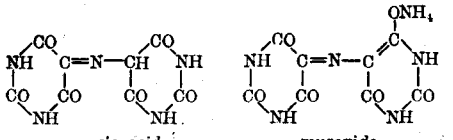
Alloxan is a strongly acid compound (in the *enol* form); it crystallises with four molecules of water of crystallisation. Three of these are readily lost on warming, but the fourth is lost only when the monohydrate is heated to 150°. Because of this, it is believed that the fourth molecule of water is not water of crystallisation but water of constitution (*cf.* chloral hydrate, Vol. I).

Alloxan stains the skin purple (due to the formation of murexide). The 5-oxime of alloxan is violuric acid (§13b. XII), and when reduced with zinc and hydrochloric acid, alloxan forms dialuric acid (§13b. XII). When alloxan is reduced with hydrogen sulphide, the product is *alloxantin*. According to Tipson *et al.* (1951), however, if excess of hydrogen sulphide is used, the product is dialuric acid only. Alloxantin is produced by reducing alloxan (one molecule) with half a molecule of hydrogen sulphide, or by mixing aqueous solutions of alloxan and dialuric acid.



alloxantin

When heated with ammonia in ethanolic solution, alloxantin forms *murexide*, which is the ammonium salt of *purpuric acid* (an unstable compound).



purpuric acid murexide

Murexide is soluble in water, giving a purple solution which turns blue on the addition of alkali. Purpuric acid slowly hydrolyses in solution to form alloxan and uramil.

When uric acid is oxidised with an aqueous suspension of lead dioxide, the products are allantoin and carbon dioxide (Liebig and Wöhler, 1838). These products are obtained in quantitative yield if the oxidation is carried out with alkaline permanganate (Behrend, 1904).

$$C_5H_4O_3N_4 + H_2O + [O] \rightarrow C_4H_6O_3N_4 + CO_2$$

Structure of allantoin, $C_4H_6O_3N_4$ (Baeyer, 1861–1864). When hydrolysed with alkali, allantoin forms two molecules of urea and one molecule of glyoxylic acid.

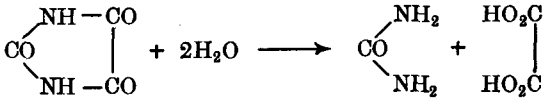
$$C_4H_6O_3N_4 + 2H_2O \longrightarrow 2NH_2{\cdot}CO{\cdot}NH_2 + CHO{\cdot}CO_2H$$

The formation of these hydrolytic products suggests that allantoin is the diureide of glyoxylic acid.

On oxidation with nitric acid, allantoin forms urea and *parabanic acid* in equimolecular proportions.

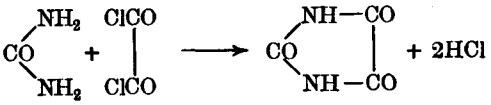
$$C_4H_6O_3N_4 + [O] \xrightarrow{HNO_2} NH_2{\cdot}CO{\cdot}NH_2 + C_3H_2O_3N_2$$

Now parabanic acid, on hydrolysis, gives urea and oxalic acid, and since there are no free amino or carboxyl groups present in the molecule, this suggests that parabanic acid is oxalylurea.



parabanic acid

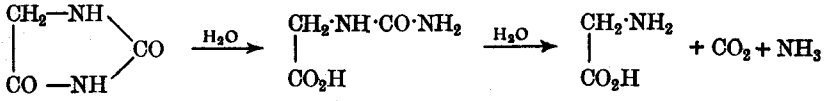
This structure has been confirmed by synthesis, *e.g.*, oxalyl chloride condenses with urea to form parabanic acid (Bornwater, 1912).



Thus, from the above facts, it can be seen that allantoin contains the parabanic acid nucleus joined to a molecule of urea. The point of the attachment is deduced from the following experimental evidence. When reduced with concentrated hydriodic acid at 100°, allantoin forms urea and *hydantoin*.

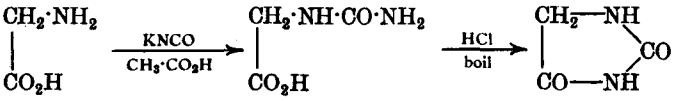
$$C_4H_6O_3N_4 + 2[H] \xrightarrow{HI} NH_2{\cdot}CO{\cdot}NH_2 + C_3H_4O_2N_2$$

Hydantoin, on controlled hydrolysis, gives *hydantoic acid* (*ureido-acetic acid*) and this, on further hydrolysis, gives glycine, ammonia and carbon dioxide. These results suggest that hydantoin is glycollylurea.

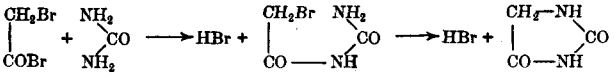


hydantoin hydantoic acid

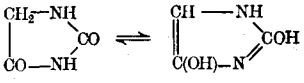
This structure for hydantoin has been confirmed by synthesis, *e.g.*, West (1918).



Hydantoin, m.p. 216°, may also be prepared by the electrolytic reduction of parabanic acid, or by the action of bromoacetyl bromide on urea.

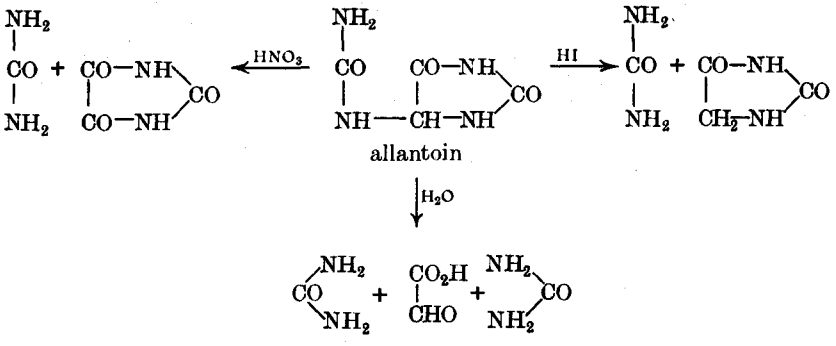


Hydantoin behaves as a tautomeric substance; the enol form is acidic and forms salts.

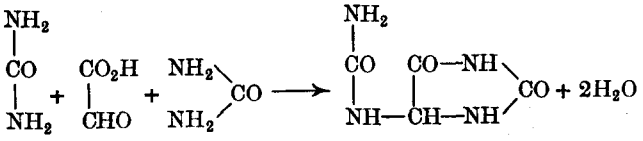


Hydantoin is oxidised to parabanic acid by bromine water.

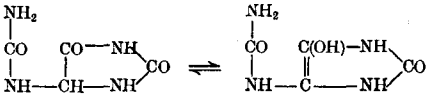
Thus the following structure for allantoin would account for all of the foregoing results:



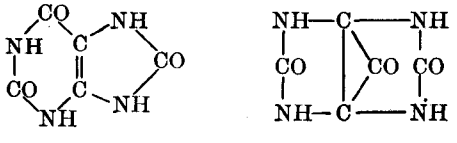
This has been confirmed by synthesis by heating urea with glyoxylic acid at 100° (Grimaux, 1876).



Examination of the structure of allantoin shows that it contains an asymmetric carbon atom; hence two optically active forms are possible. Both forms have been obtained, and they have been found to racemise rapidly in solution; the racemisation probably occurs *via* enolisation (*cf.* §8 iii. II).



In the formation of allantoin from uric acid by oxidation, one carbon atom is lost from the latter as carbon dioxide. The problem, then, is to fit one carbon atom into the allantoin structure. At the same time, the structure thus given to uric acid must also include the alloxan skeleton in order to account for the formation of this compound. Two structures that were proposed which both agreed with the facts known at the time were by Medicus (1875) and by Fittig (1878).

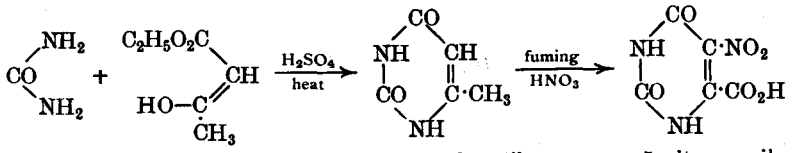


Medicus formula Fittig formula

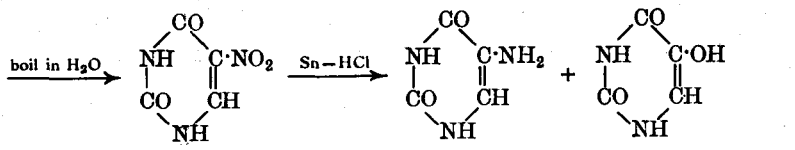
Fischer (1884) prepared two isomeric monomethyluric acids; one gave methylalloxan and urea on oxidation with nitric acid, and the other gave

alloxan and methylurea. Fittig's formula, which is symmetrical, can give rise to only *one* monomethyluric acid; hence this structure is untenable. On the other hand, the Medicus formula satisfies the existence of at least two isomeric monomethyl derivatives: one methyl group in the pyrimidine nucleus (at position 1 or 3) would produce methylalloxan and urea, and a methyl group in the imidazole nucleus (at position 7 or 9) would produce alloxan and methylurea (Fischer showed that the two monomethyluric acids were the 3- and 9-derivatives). Examination of the Medicus formula shows that it admits the possibility of four monomethyl, six dimethyl and four trimethyl derivatives. All of these have been prepared by Fischer and his co-workers, thus giving powerful support to the Medicus formula. Proof of the Medicus formula lies in the synthesis of uric acid; three syntheses are given here.

(i) Behrend and Roosen (1888) carried out the first unambiguous synthesis (see also §15. XII).

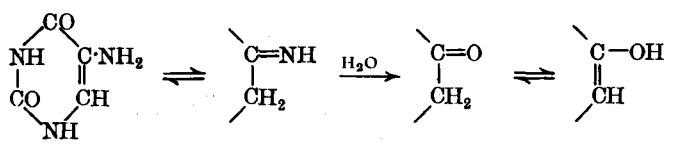


urea ethyl aceto-acetate 4-methyluracil 5-nitrouracil -4-carboxylic acid

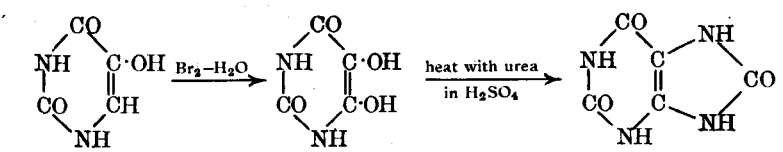


5-nitrouracil 5-aminouracil 5-hydroxyuracil

In this reduction, some of the aminouracil is converted into hydroxyuracil. The mechanism of this change is not certain, but a possibility is as follows:



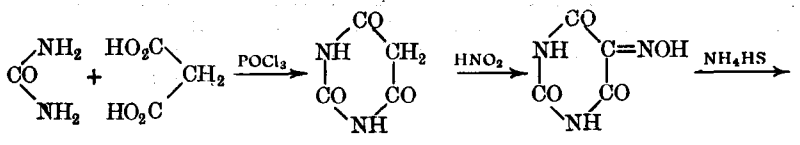
The reaction product was treated with nitrous acid, thereby converting the 5-aminouracil present into 5-hydroxyuracil; then the synthesis proceeded as follows:



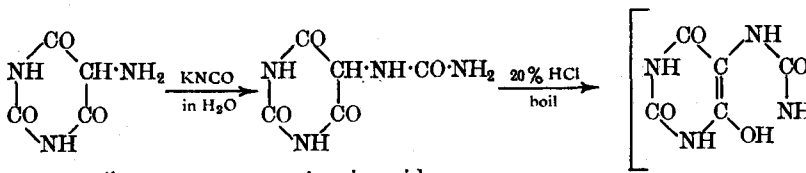
5-hydroxyuracil 4:5-dihydroxyuracil uric acid

(ii) Baeyer's synthesis (1863), completed by Fischer (1895). Baeyer arrived at ψ-uric acid and knew that uric acid contained one molecule of water less than this, but was unable to remove it to form uric acid. His failure was due to the fact that ψ-uric acid is not dehydrated by the usual

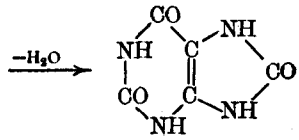
dehydrating agents; Fischer succeeded by fusion with anhydrous oxalic acid, and also obtained better results by boiling ψ-uric acid with 20 per cent. hydrochloric acid.



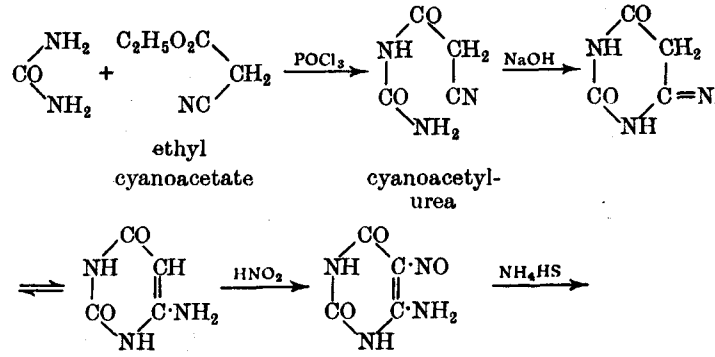
barbituric acid violuric acid



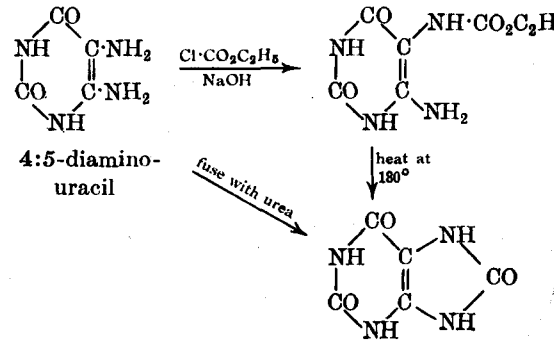
uramil ψ-uric acid



(iii) Traube's synthesis (1900) is the most important method, since it can be used to prepare any purine derivative; it is also the basis of various commercial methods for preparing the purines synthetically.

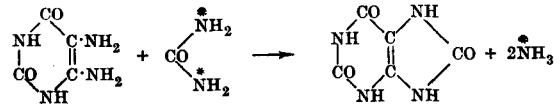


ethyl cyanoacetate cyanoacetyl-urea

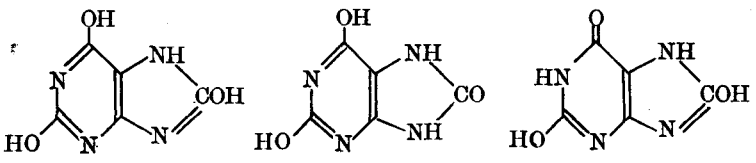


4:5-diamino-uracil

Clusius *et al.* (1953), using urea labelled with ^{15}N, have shown that the two nitrogen atoms in the diaminouracil are retained on fusion with urea.



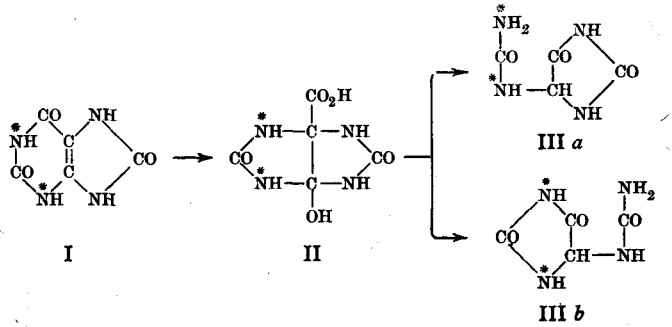
Uric acid is a white crystalline powder which is insoluble in the ordinary organic solvents. It behaves as a weak dibasic acid, forming two series of salts (*e.g.*, monosodium and disodium urate).



2:6:8- 2:6- 2:8-

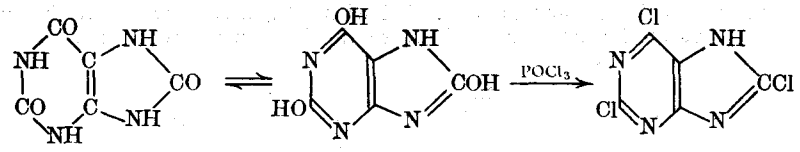
It thus appears that the tri-enol form (2 : 6 : 8-trihydroxypurine) is unlikely; this leaves three possible di-enol forms, 2 : 6-, 2 : 8- and 6 : 8-. Which of these di-enol forms is the one that forms the disodium salt still appears to be uncertain. Fischer thought that the di-enol form is the 2 : 6-. Evidence that may be quoted to support this is that in this arrangement the pyrimidine ring will be " aromatic " and so stabilised by resonance. There is, however, a certain amount of evidence which suggests the 2 : 8- di-enol form (*cf.* §§13a, 15).

It is also interesting to consider the path followed in the oxidation of uric acid to allantoin. Behrend (1904) suggested that the alkaline permanganate oxidation of uric acid (I) gives allantoin (III*a* and *b*) *via* the symmetrical intermediate II. Cavalieri *et al.* (1948) have carried out this oxidation using uric acid labelled with ^{15}N at N_1 and N_3, and found that the allantoin produced had this isotopic nitrogen distributed uniformly among all the four nitrogen atoms. This is in keeping with the intermediate formation of II.

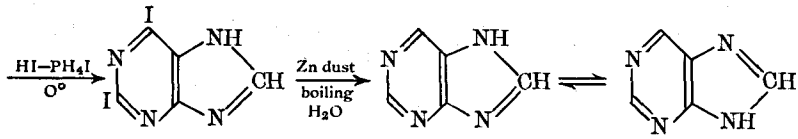


I II III *a* III *b*

§3. Purine. When uric acid is treated with phosphoryl chloride, 2 : 6 : 8-trichloropurine is obtained (uric acid behaves as the tri-enol in this reaction). This trichloro compound is a very important intermediate in the synthesis of purine derivatives, and a point worth noting is that the reactivities of the chlorine atoms are $6 > 2 > 8$. **Purine,** m.p. 217°, may be prepared from uric acid as follows:



2:6:8-trichloropurine



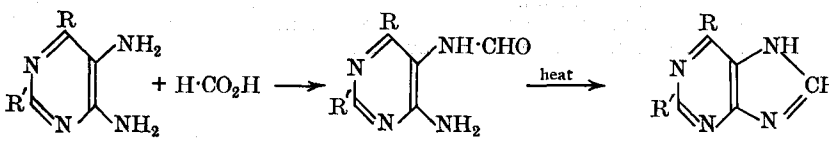
2:6-di-iodopurine purine

Purine is a fairly strong base and forms salts with acids; it has been found to occur naturally as its 9-D-ribofuranoside, *nebularine* (Löfgren *et al.*, 1953).

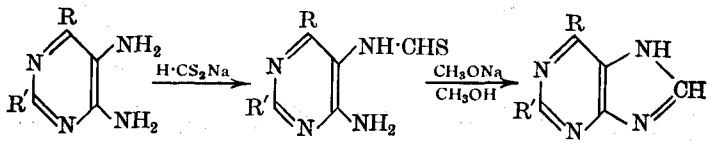
PURINE DERIVATIVES

§4. Synthesis of purines. Before describing some individual purine derivatives, let us first consider some general methods of synthesising purines. Fischer (1897, 1898) prepared various purines starting from 2 : 6 : 8-trichloropurine. There are, however, two general synthetic methods in which the pyrimidine ring is synthesised first and then the imidazole ring " built up " on this, or *vice versa*.

(i) *Traube's method.* This consists of synthesising a 4 : 5-diaminopyrimidine (see §14. XII) and then condensing with formic acid to produce the imidazole ring; the formyl derivative is ring-closed by heating alone or by heating its sodium salt.

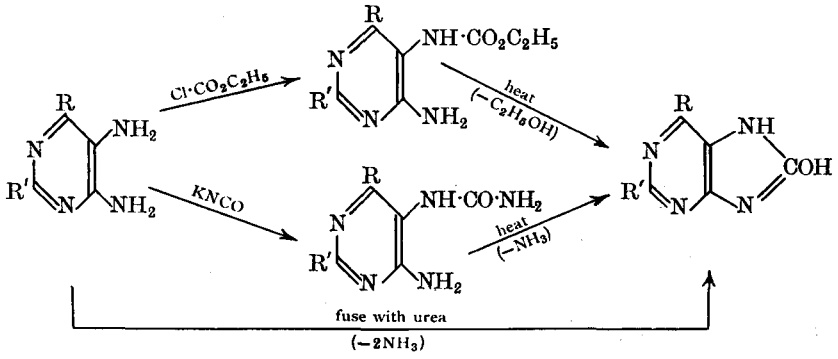


This synthesis leads to the preparation of purines that are unsubstituted in position 8. This type of purine may also be prepared by heating a 4 : 5-diaminopyrimidine with dithioformic acid in the presence of sodium hydroxide solution, and then heating the product with a methanolic solution of sodium methoxide.



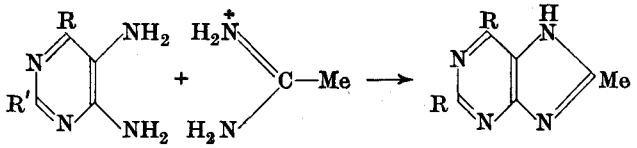
8-Hydroxypurines may be prepared by using ethyl chloroformate instead of formic acid. Alternatively, the diaminopyrimidine may be boiled with potassium *iso*cyanate and the product, a ureidopyrimidine, ring-closed by

heating. Finally, diaminopyrimidines may be fused with urea to produce 8-hydroxypurines.

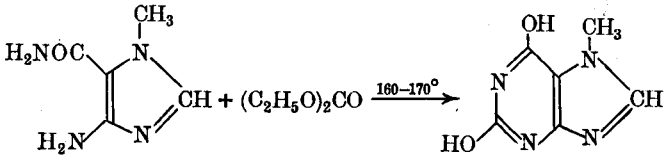


o-Aminohydroxypyrimidines may be used instead of *o*-diaminopyrimidines (*cf.* Baeyer's synthesis of *ψ*-uric acid, §2).

Bergmann *et al.* (1961) have prepared 8-substituted purines by condensing 4,5-diaminopyrimidines with amidine salts, *e.g.*,

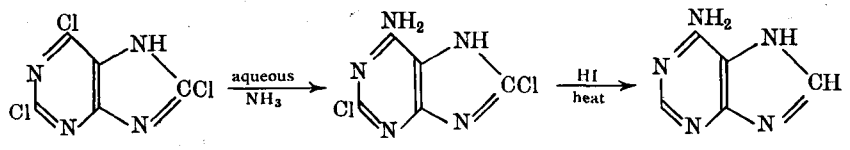


(ii) A less frequently used synthesis of purines starts with the imidazole derivative, *e.g.*, 7-methylxanthine from 4-amino-1-methylimidazole-5-carbonamide (Sarasin *et al.*, 1924):



§5. Adenine (6-aminopurine), d. 365°, occurs in the pancreas of cattle and in tea extract. Its general reactions showed that adenine was a purine, and its structure was established by synthesis.

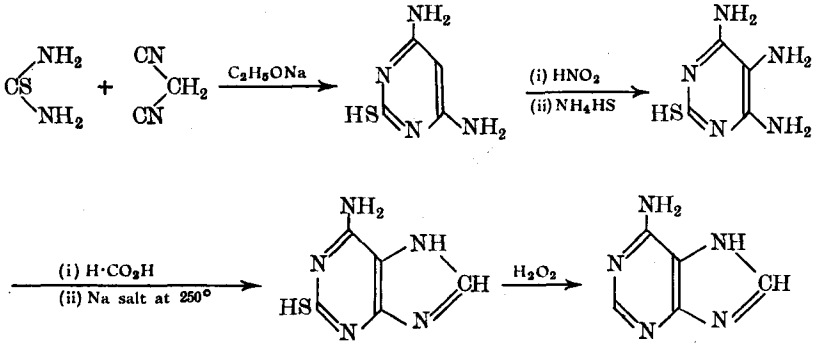
(i) Fischer (1897) (see also §6).



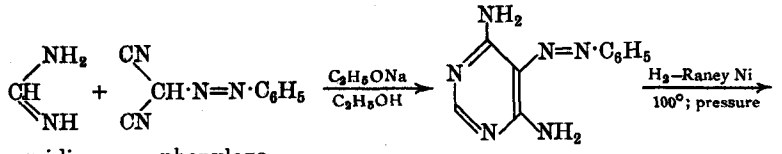
2:6:8-trichloro-purine

adenine

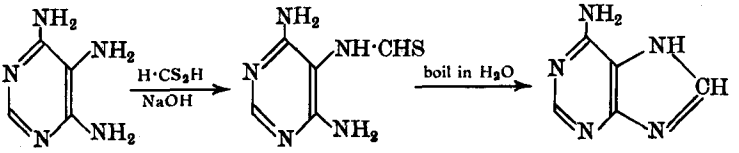
(ii) Traube (1904).



(iii) Todd *et al.*, (1943).

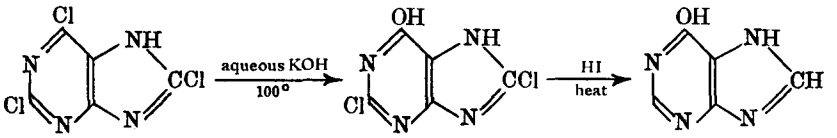


formamidine phenylazomalononitrile



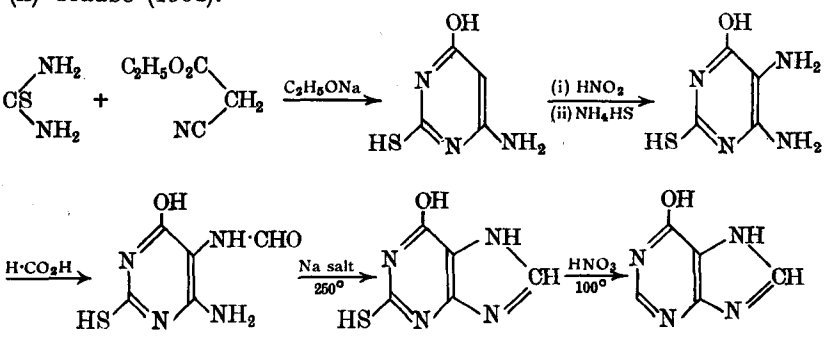
§6. **Hypoxanthine** (6-hydroxypurine), d. 150°, occurs in tea extract and in animal tissues. Its formation by the action of nitrous acid on adenine establishes its structure, and this is confirmed by synthesis.

(i) Fischer (1897, 1898).

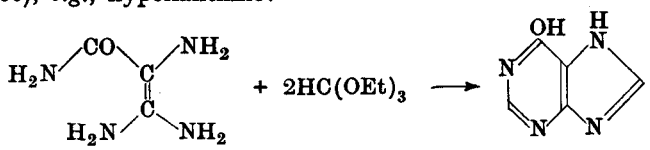


hypoxanthine

(ii) Traube (1904).

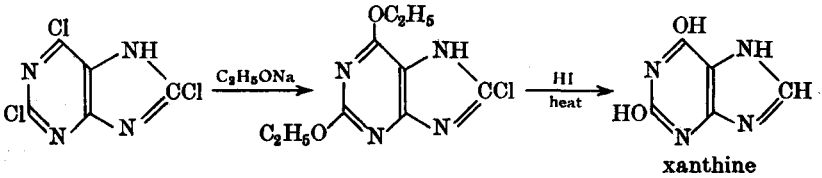


A new useful synthesis of hypoxanthines and adenines involves the condensation between 1,2,2-trimethylaminoacrylamide and ortho-esters (Richter *et al.*, 1960), *e.g.*, hypoxanthine:



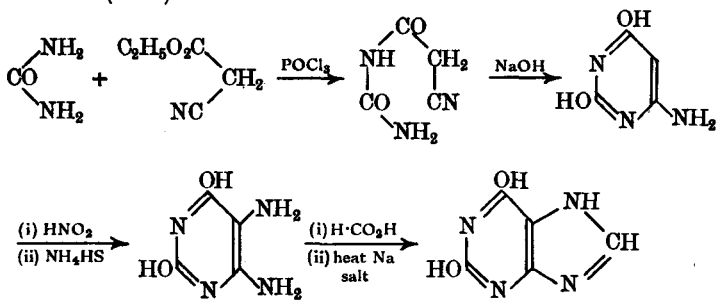
§7. Xanthine (2 : 6-dihydroxypurine), d. above 150°, occurs in tea extract and in animal tissues. When oxidised with potassium chlorate in hydrochloric acid solution, xanthine forms alloxan and urea; these products show the relationship of xanthine to uric acid, and its structure has been established by synthesis.

(i) Fischer (1898) (see also §10).



xanthine

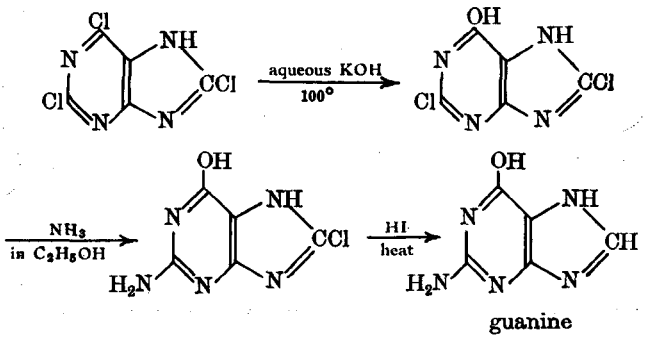
(ii) Traube (1900).



Xanthine is the parent substance of a number of compounds (see later).

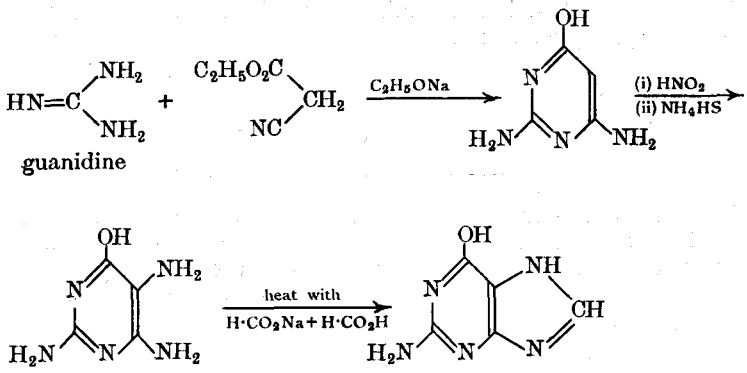
§8. Guanine (2-amino-6-hydroxypurine), d. 360°, occurs in the pancreas of cattle, in guano and in certain fish scales. Its structure is shown by the fact that it gives xanthine on treatment with nitrous acid; this conversion is also effected by boiling guanine with 25 per cent. hydrochloric acid (Fischer, 1910) (see also §13b).

(i) Fischer (1897).



guanine

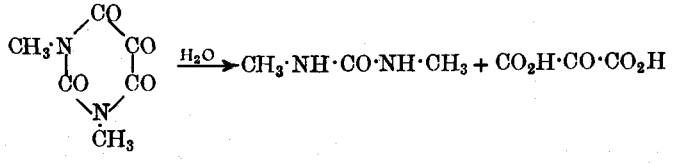
(ii) Traube (1900).



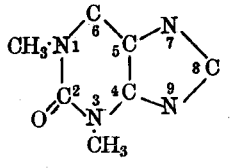
XANTHINE BASES

Three important methylated xanthines that occur naturally are caffeine, theobromine and theophylline. All three have been prepared from uric acid by Fischer and all have been synthesised by means of the Traube method.

§9. Caffeine (1 : 3 : 7-trimethylxanthine), m.p. 235–237°, occurs in tea, coffee, etc. Its molecular formula is $C_8H_{10}O_2N_4$, and its relationship to uric acid is shown by the fact that on oxidation with potassium chlorate in hydrochloric acid, caffeine gives dimethylalloxan and methylurea in equimolecular proportions. The structure of the former product is established by its conversion into *sym.*-dimethylurea and mesoxalic acid on hydrolysis, and is confirmed by synthesis from these two compounds.

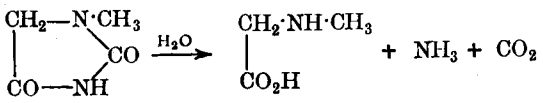


These results indicate that caffeine and uric acid have the same skeleton structure; at the same time the positions of two methyl groups and one oxygen atom in caffeine are also established. Thus the problem now is to ascertain the positions of the remaining methyl group and oxygen atom. The following skeleton structure for caffeine summarises the above information; the third methyl group is at either position 7 or 9, and the remaining oxygen atom at 6 or 8.

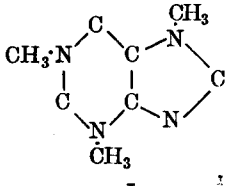


Position of the methyl group. As we have seen above, the oxidation of caffeine gives dimethylalloxan and methylurea. Fischer, however, also iso-

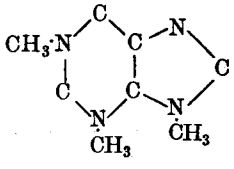
lated another oxidation product which, on hydrolysis, gave *N*-methylglycine, carbon dioxide and ammonia. Thus this third oxidation product must be *N*-methylhydantoin:



It therefore follows that caffeine contains two ring structures, that of dimethylalloxan and that of methylhydantoin. The following two skeleton structures for caffeine are both possible, since each could give the required oxidation products. Actually, the isolation of methylurea suggests I or II;



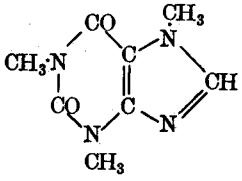
I



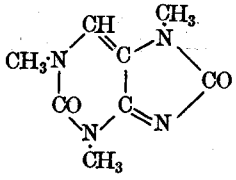
II

the isolation of methylhydantoin confirms these possibilities. Finally, Fischer isolated a fourth oxidation product, *viz.*, *sym.*-dimethyloxamide, $CH_3{\cdot}NH{\cdot}CO{\cdot}CO{\cdot}NH{\cdot}CH_3$. Examination of I and II shows that only I can give rise to the formation of this oxamide, and so I is the skeleton of caffeine.

Position of the oxygen atom. In view of what has been said above, we see that there are now two possible structures for caffeine which fit the facts equally well:

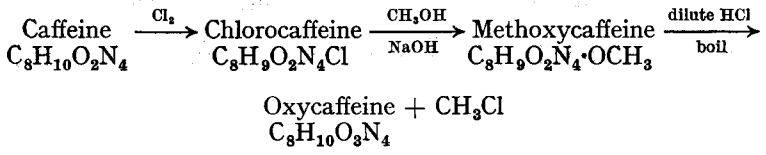


III

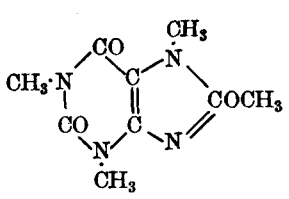


IV

By analogy with uric acid, III would appear the more likely one; this, however, is not proof. Fischer showed that III is caffeine as follows.

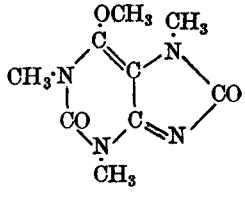


Fischer then showed that oxycaffeine was identical with a trimethyluric acid, since on methylation with methyl iodide in the presence of aqueous sodium hydroxide, oxycaffeine was converted into tetramethyluric acid. Thus methoxycaffeine is either V or VI, and oxycaffeine VII or VIII.



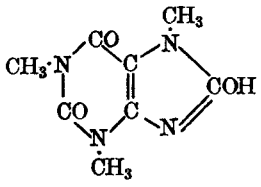
V

or



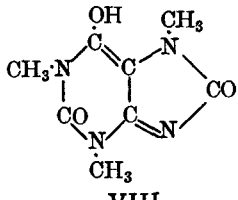
VI

methoxycaffeine



VII

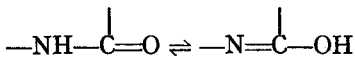
or



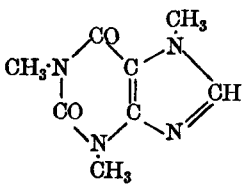
VIII

oxycaffeine

When oxycaffeine, *as its silver salt,* is heated with methyl iodide, it is converted into a mixture of tetramethyluric acid (which contains four *N*-methyl groups) and methoxycaffeine (which contains three *N*-methyl groups and one methoxyl group). The simultaneous formation of these two products suggests that oxycaffeine is a tautomeric substance, *i.e.*, it contains the *amido-imidol triad system*:

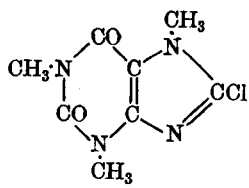


Now this triad system can exist only in the imidazole nucleus in oxycaffeine, since neither nitrogen atom in the pyrimidine nucleus is attached to a hydrogen atom (VII can give rise to the above tautomeric system, whereas VIII cannot). Thus the methoxyl group in methoxycaffeine is in the imidazole nucleus, and consequently the chlorine atom in chlorocaffeine is also in this nucleus; hence caffeine is IX and chlorocaffeine is X.



IX

caffeine

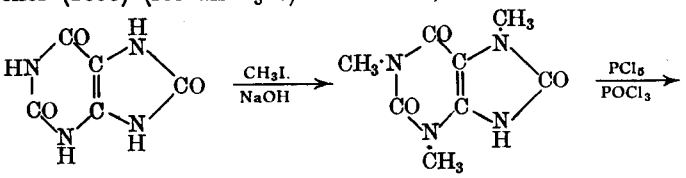


X

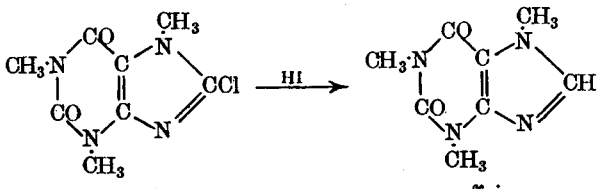
chlorocaffeine

This structure for caffeine has been confirmed by various syntheses, *e.g.*,

(i) Fischer (1899) (see also §10).

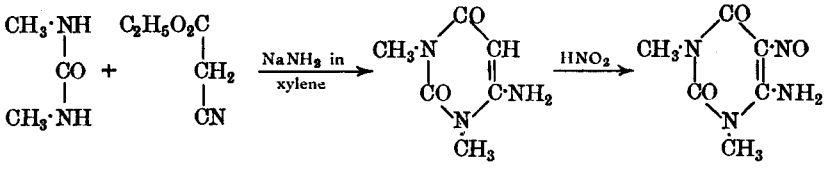


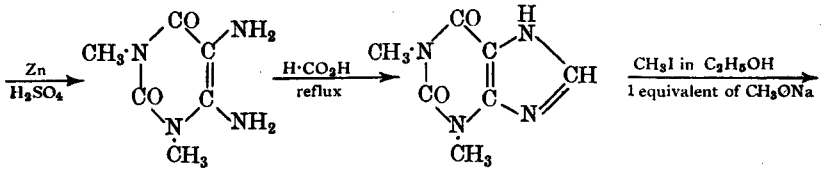
uric acid — 1:3:7-trimethyluric acid



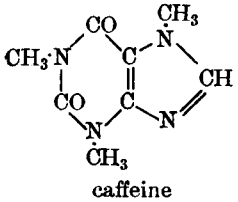
chlorocaffeine — caffeine

(ii) A commercial synthesis based on Traube's method is as follows:



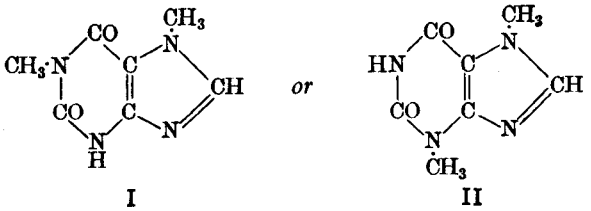


theophylline



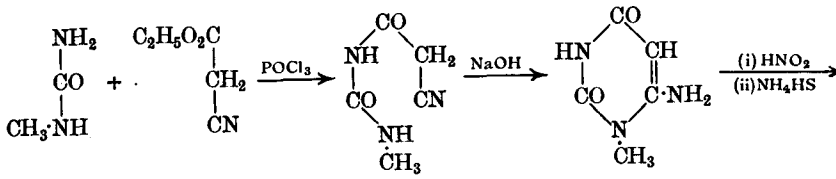
caffeine

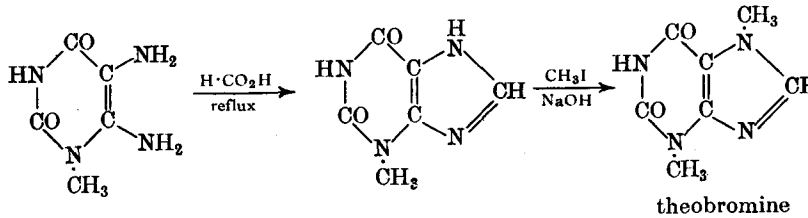
§10. Theobromine (3 : 7-dimethylxanthine), m.p. 337°, occurs in cocoa beans, tea, etc. The structure of theobromine has been deduced from the fact that, on oxidation with potassium chlorate in hydrochloric acid, it gives methylalloxan and methylurea, and also that it is converted into caffeine when its silver salt is heated with methyl iodide. Thus theobromine is either I or II.



I — II

The position of the methyl group in the pyrimidine nucleus has been shown to be 3 (*i.e.*, structure II) by synthesis using Traube's method.

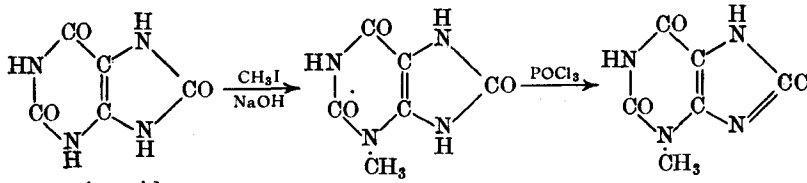




theobromine

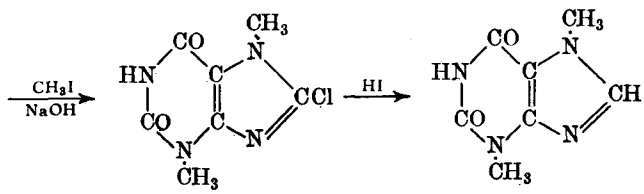
The product formed by the condensation between methylurea and ethyl cyanoacetate contained no free amino-group; thus the condensation must occur as shown (and not by the carbethoxyl group with the methylimino-group of the methylurea).

Fischer (1899) also prepared theobromine from uric acid as follows:



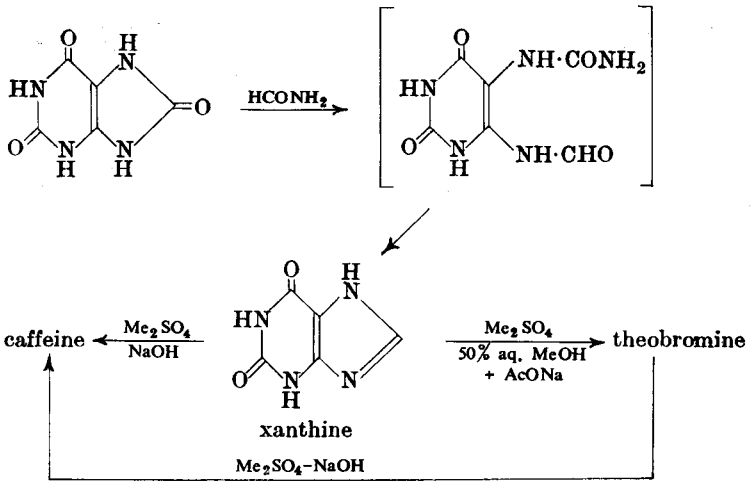
uric acid

3-methyluric acid



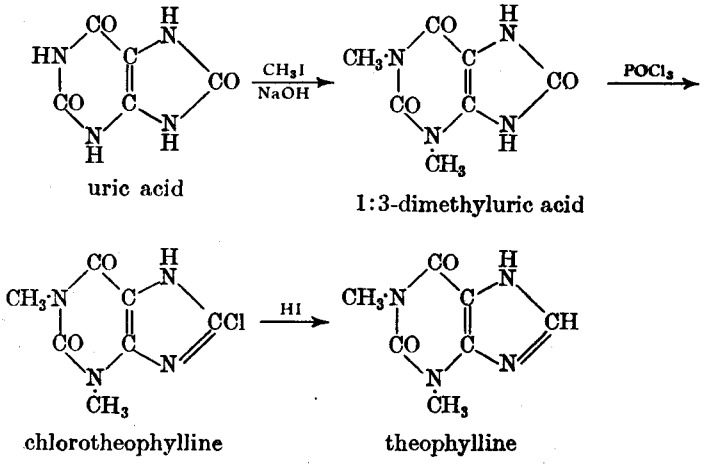
It should be noted that in this synthesis a mixture of phosphorus pentachloride and phosphoryl chloride cannot be used; this mixture replaces the oxygen atom (*i.e.*, the hydroxyl group) at position 6 and not at 8.

The simplest method of preparing xanthine (§7), caffeine (§9) and theobromine from uric acid is probably that of Bredereck (1950):

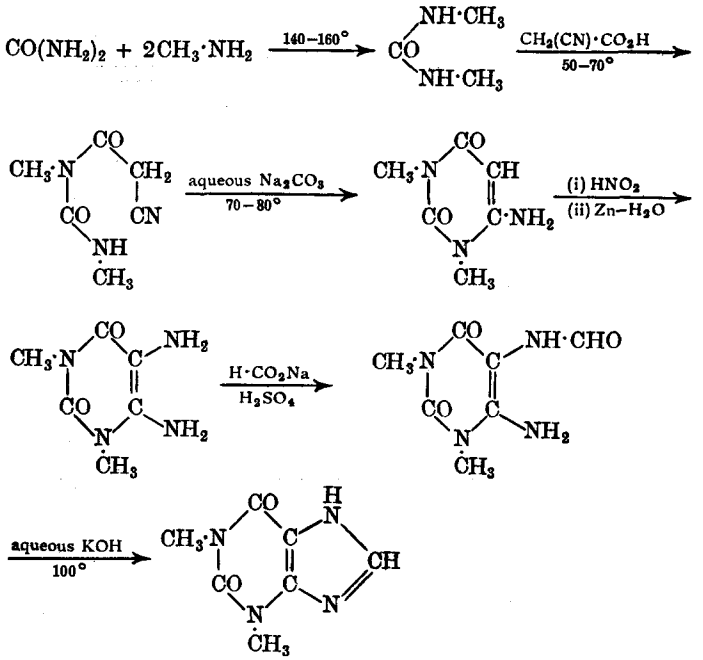


§11. Theophylline (1 : 3-dimethylxanthine), m.p. 269–272°, occurs in tea. Its structure has been deduced from the fact that it is converted into caffeine on methylation, and that it forms dimethylalloxan and urea on oxidation. Thus theophylline is 1 : 3-dimethylxanthine, and this structure has been confirmed by synthesis.

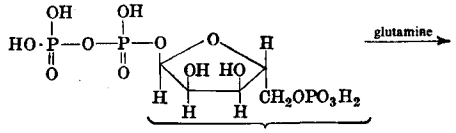
(i) Fischer (1899).



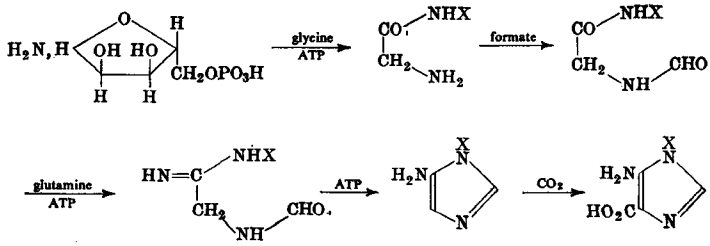
(ii) Theophylline has also been synthesised commercially by means of the Traube method (*cf.* caffeine, §9).

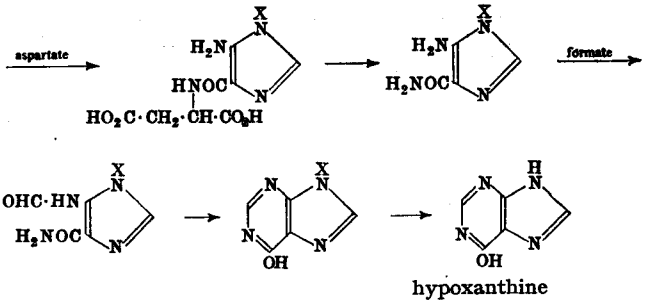


§11a. Biosynthesis of purines. Most of the work on the biosynthesis of purines has been carried out on uric acid by means of enzymes from bird liver. Sonne *et al.* (1946, 1948), working with the following labelled compounds (^{13}C), showed that carbon dioxide supplies C_6, formic acid C_2 and C_8, and glycine C_4, C_5 and N_7. Thus all the carbon atoms in uric acid are accounted for. It has also been shown that the carbon atoms in hypoxanthine are derived from the same precursors as those in uric acid. Furthermore, it was also shown that in the hypoxanthine biosynthesis in liver extracts, N_7 is derived from glycine, N_3 and N_9 are derived from the amide nitrogen of glutamine (§2. XIII; number 23 in the list of amino-acids) and N_1 is derived from aspartic acid (number 19 in list). According to Buchanan *et al.* (1948–) and Greenberg *et al.* (1951), hypoxanthine is produced from inosine-5′-phosphate (the nucleotide of hypoxanthine; see §13d). Inosine-5′-phosphate is believed to be biosynthesised as follows from ribose 5-phosphate 1-pyrophosphate (ATP is the co-enzyme adenosine triphosphate):



This fragment is X in the following reactions





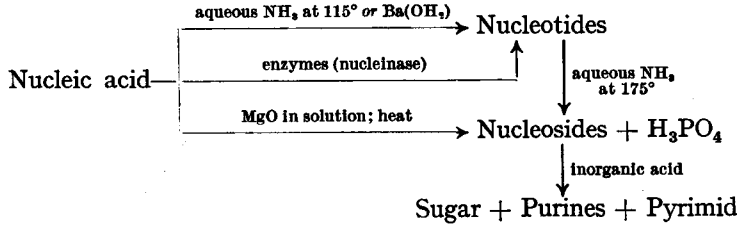
Oró (1961) has shown that adenine and the purine precursors 4-amino-imidazole-5-carboxamidine and formamidine are formed spontaneously from hydrogen cyanide in water–ammonia systems under conditions assumed to have existed on primitive Earth (*cf.* §18. XIII). Oró has also suggested a mechanism for the formation of adenine from hydrogen cyanide under the above conditions.

NUCLEIC ACIDS

§12. Introduction. Nucleoproteins are one of the classes of conjugated proteins (§7 B. XIII); the nucleic acid part is the prosthetic group, and the protein part consists of protamines and histones. These latter compounds are basic and form salt-like compounds, the nucleoproteins, with the nucleic acids. On careful hydrolysis, nucleoproteins are broken down into the nucleic acid and protein.

Genes are the units of inheritance, and there is now a great deal of evidence to show that a gene is a nucleic acid molecule.

§13. Structure of the nucleic acids. Nucleic acids are colourless solids, all of which contain the following elements: carbon, hydrogen, oxygen, nitrogen and phosphorus. The following chart shows the nature of the products obtained by hydrolysis under different conditions.



Hayes (1960) has shown that ribonucleic acids (§13a) may be rapidly and quantitatively degraded to ribonucleosides by refluxing with 50 per cent. aqueous formamide.

§13a. Sugars. Only two sugars have been isolated from the hydrolysates of nucleic acids; both are pentoses: D(−)-ribose and 2-deoxy-D(−)-ribose.

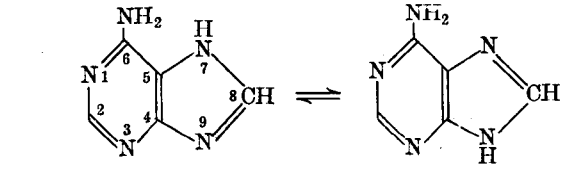
$$\underset{\text{ribose}}{CHO{\cdot}(CHOH)_3{\cdot}CH_2OH} \qquad \underset{\text{2-deoxyribose}}{CHO{\cdot}CH_2{\cdot}(CHOH)_2{\cdot}CH_2OH}$$

The nucleic acids are classified according to the nature of the sugar present: the *pentose nucleic acids* or *ribonucleic acids* (R.N.A.), and the *deoxypentose nucleic acids* or *deoxyribonucleic acids* (D.N.A.). Ribonucleoproteins are

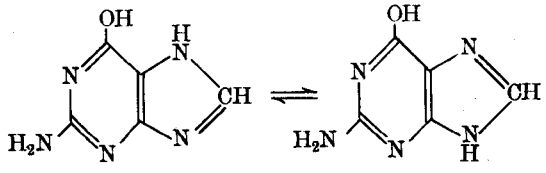
found mainly in the cytoplasm of the cell, whereas deoxyribonucleoproteins are found mainly in the cell nucleus. D(−)-Ribose is the pentose of yeast, liver and pancreas R.N.A.s; 2-deoxy-D(−)-ribose occurs in thymus D.N.A. Nucleic acids also occur in plant and animal viruses. The principal function of the nucleic acids appears to be in protein synthesis. Evidence has been obtained that the D.N.A.s act as carriers of genetic continuity (see §18. XIII).

Aldridge (1960) has shown that the addition of indium chloride solution to acetate-buffered solutions of nucleic acids in the presence of sodium chloride produces a precipitate containing indium and nucleic acid. Furthermore, by varying the concentration of the sodium chloride, it is possible to precipitate either the R.N.A. or the D.N.A. from aqueous solution.

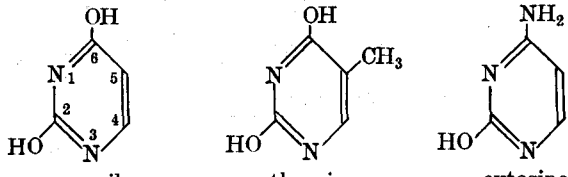
§13b. Bases. Until very recently, only two purines had been isolated from nucleic acids, adenine and guanine. In 1958, however, Littlefield *et al.* found 2-methyladenine and 6-methylaminopurine in R.N.A.s from several sources, and Adler *et al.* (1958) and Dunn *et al.* (1958) have shown that 2-methylamino- and 2-dimethylaminoguanine are widespread in R.N.A.s.



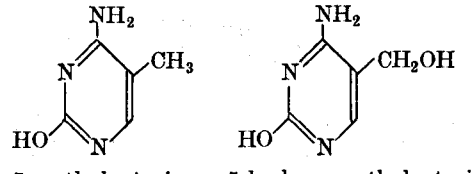
adenine



guanine



uracil thymine cytosine



5-methylcytosine 5-hydroxymethylcytosine

1-Methylguanine has also been found in minute quantities in R.N.A.s (Amos *et al.*, 1958). On the other hand, five pyrimidine bases have been isolated: uracil, thymine, cytosine, 5-methylcytosine and 5-hydroxymethylcytosine. Both types of nucleic acids (R.N.A.s and D.N.A.s) contain adenine and guanine. Cytosine also occurs in both types of nucleic acids, but uracil

occurs only in R.N.A.s. 5-Methylcytosine has been found to be a fairly common minor constituent of D.N.A.s, and Amos *et al.* (1958) have shown that it occurs in minute quantities in R.N.A.s (*cf.* methylguanine above). Also, thymine was believed to occur only in D.N.A.s, but Littlefield *et al.* (1958) have found it in R.N.A.s from several sources. 5-Hydroxymethylcytosine has been found in certain D.N.A.s (Wyatt *et al.*, 1952).

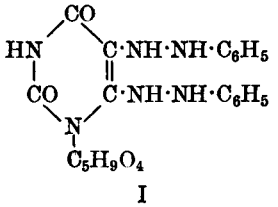
Angell (1961) has shown, from infra-red studies, that in the *solid* state and in ribose and deoxyribose nucleosides derived from these bases, adenine exists in the amino form, cytosine and guanine exist in the keto-amino form and uracil in the diketo form.

Combination of a base (either a purine or pyrimidine) with a sugar (ribose or deoxyribose) gives rise to a **nucleoside,** *e.g.*, adenosine (ribose + adenine), guanosine (ribose + guanine), cytidine (ribose + cytosine), uridine (ribose + uracil), thymidine (deoxyribose + thymine).

Combination of a nucleoside with phosphoric acid produces a **nucleotide,** *i.e.*, nucleotides are nucleoside phosphates, *e.g.*, adenylic, guanylic, cytidylic and uridylic acids. It might be noted here that the term nucleotide is now used to embrace a large group of compounds composed of the phosphates of *N*-glycosides of heterocyclic bases, and the pyrophosphates and polyphosphates containing one or more nucleosides. The term nucleotide also includes the nucleic acids themselves.

The problem now is to ascertain how these various units are linked in nucleosides and nucleotides.

§13c. Structure of nucleosides. Hydrolysis of nucleotides with aqueous ammonia at 175° under pressure gives nucleosides and phosphoric acid; thus in nucleosides the base is linked directly to the sugar. Furthermore, since nucleosides are non-reducing, the " aldehyde group " of the sugar cannot be free, *i.e.*, nucleosides are *glycosides* (*cf.* §24. VII). The next problem is to decide which atom of the base is joined to C_1 of the sugar. Let us first consider the pyrimidines. Cytidine, on treatment with nitrous acid, is converted into uridine; it therefore follows that the sugar residue is linked in the *same* position in both of these nucleosides. The point of linkage cannot be 1 or 6, since cytidine has a *free* amino-group at position 6 and consequently there cannot be a hydrogen atom on N_1. Also, since uridine forms a 5-bromo derivative, C_5 must be free (Levene *et al.*, 1912). When uridine is treated with an excess of bromine, followed by the addition of phenylhydrazine, a uridine derivative is obtained which contains *two* phenylhydrazino radicals. This compound was given structure I since work by Levene (1925) showed

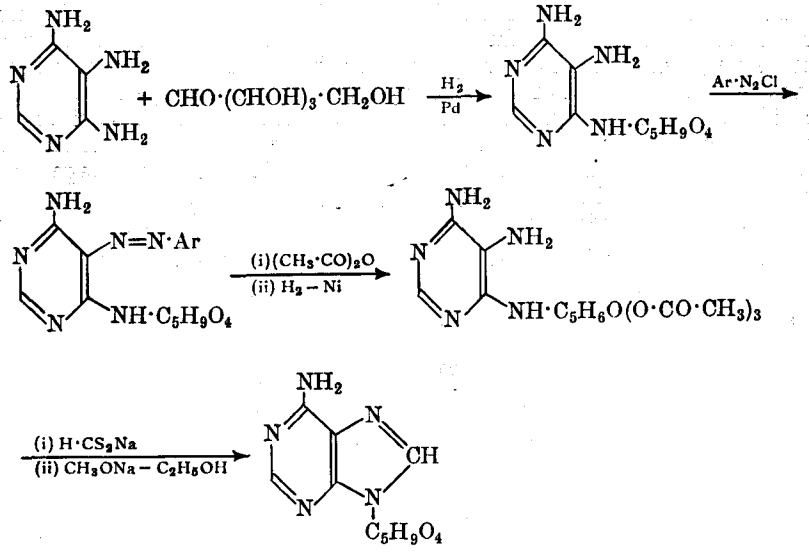


I

that this type of compound can be obtained only if uracil is substituted in position 3 and positions 4 and 5 are free. Thus the sugar is attached to N_3. In a similar way, it has been shown that the other pyrimidine nucleosides (ribosides and deoxyribosides) have the sugar residue linked at N_3. Todd *et al.* (1947) have synthesised uridine and cytidine, and thereby have confirmed the linkage at N_3. This linkage has also been confirmed by the X-ray analysis of cytidine (Furberg, 1950).

Now let us consider nucleosides containing purine bases. Adenosine has

a free amino-group at position 6; therefore the sugar cannot be at C_6 or N_1 (*cf.* cytidine). Similarly, since guanosine has a free amino-group at position 2, the sugar cannot be at C_2 or N_3. Now Levene found that the two purine ribosides are equally readily hydrolysed by dilute acids and by the same enzyme. He therefore assumed that the sugar residue is linked at the same place in both nucleosides. On this basis, only positions 7, 8 and 9 are possible points of attachment. Position 8 was then excluded since this point would involve a carbon–carbon bond, a linkage which would be very stable, whereas nucleosides are very readily hydrolysed by dilute acids (see also below). Thus positions 7 or 9 are free. This is supported by the following evidence (Levene, 1923). When guanosine is treated with nitrous acid, xanthosine is produced and this, on methylation with diazomethane followed by hydrolysis, gives theophylline (1 : 3-dimethylxanthine). Thus positions 1 and 3 are free in guanosine, and so the sugar must be attached at position 7 or 9. The evidence so far does not permit a decision to be made between these two positions since the system (in the imidazole nucleus) is tautomeric. It should be noted that had the sugar residue been attached to C_8, then a *trimethylxanthine* would have been obtained instead of theophylline (*cf.* above). The ultraviolet absorption spectrum of guanosine is very similar to that of 9-methylguanine and differs from that of 7-methylguanine; hence it appears likely that guanosine is the 9-guanine glycoside (Gulland *et al.*, 1936, 1938). Todd *et al.* (1947, 1948) have synthesised guanosine and adenosine in which the sugar is known to be in the 9-position, and showed that their synthetic compounds are identical with the natural products; *e.g.*, the synthesis of adenosine.

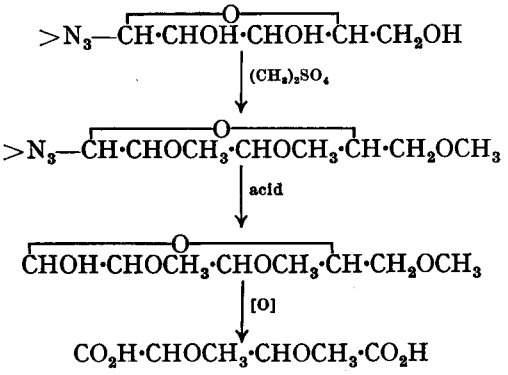


adenosine

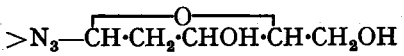
It might be noted, in passing, that glycosides are compounds formed by the linking of a sugar (at C_1) with a COH group. Thus the nucleosides are, strictly speaking, not glycosides; they should be called ribosyl-pyrimidines and ribosyl-purines.

The final problem to be elucidated in connection with the structure of nucleosides is the nature of the ring in the sugar residue and the type of linkage (α or β). Degradative experiments have shown that the sugar is

present as the furanose form, *e.g.*, methylation of a pyrimidine riboside, followed by hydrolysis, gives a trimethylribose which, on oxidation, forms dimethyl*meso*tartaric acid. This product shows that the ribose ring is furanose; had the ring been pyranose, then the final product would have been trimethoxyglutaric acid (*cf.* §§7a, 7b. VII).

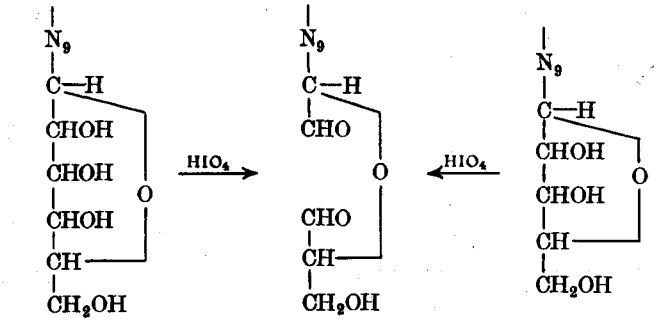


Deoxyribose has also been shown to be of the furanose type, *e.g.*, Lythgoe *et al.* (1950) found that pyrimidine deoxyribosides consume a negligible amount of periodic acid; this agrees with the 2-deoxyribofuranose structure since, in this state, the molecule does not contain two adjacent hydroxyl groups (*cf.* §7g. VII).

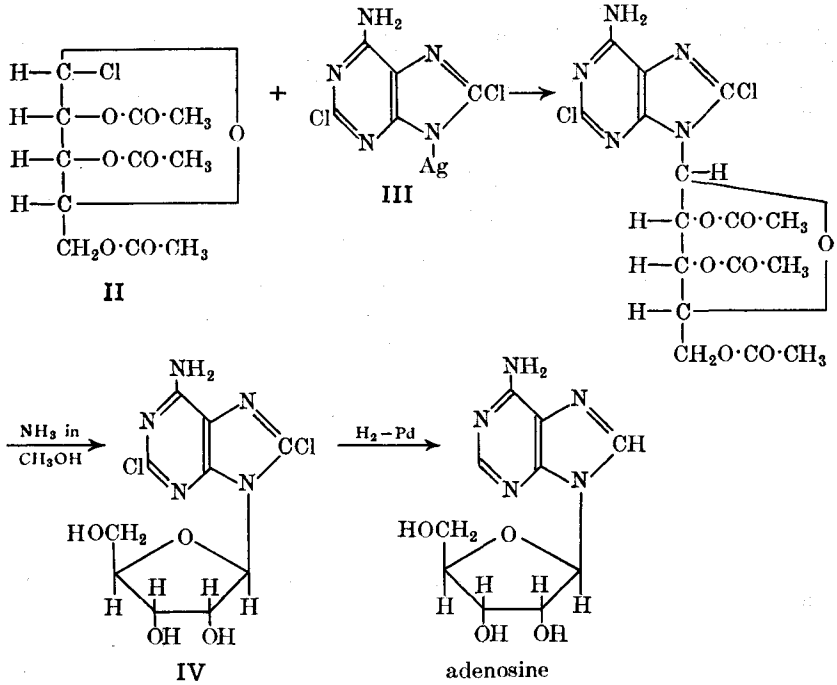


These results have been confirmed by other work (see below).

The configuration of the furanoside link has been shown to be β- by various means, *e.g.*, Todd *et al.* (1947) oxidised adenosine with periodic acid, and showed that the product is identical with that from the oxidation of 9-β-D-mannopyranosidyladenine (a synthetic compound). This proves that



the sugar residue is at position 9, has the furanose structure and that the linkage is β-. Similar experiments with other ribonucleosides suggest that all these compounds have a β-configuration. Also, Todd *et al.* (1946–1948) have synthesised adenosine, guanosine, cytidine and uridine, and thereby

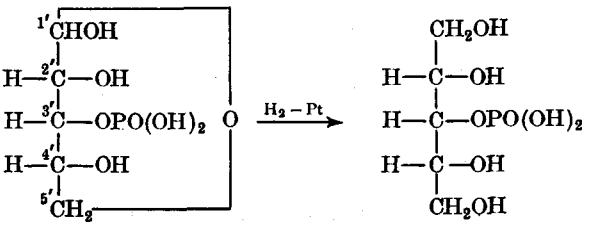


confirmed the β-configuration; *e.g.*, adenosine has been synthesised as follows (Todd *et al.*, 1948). Acetochloro-D-ribofuranose, II (*cf.* §24. VII), is condensed with the silver salt of 2 : 8-dichloroadenine, III, and the product deacetylated with a methanolic solution of ammonia to give 2 : 8-dichloro-9-β-ribofuranosyladenine, IV. IV, on catalytic reduction (palladium), is converted into adenosine.

Furberg (1950) has shown by means of the X-ray analysis of cytidine that the sugar residue is attached to N_3 and is β-D-ribofuranoside. Since other ribonucleosides exhibit the same general pattern, it is inferred that all are furanosides with the β-configuration. Manson *et al.* (1951), from absorption spectra measurements, have shown that deoxyribonucleosides also exist in the β-configuration.

It will be noted from the foregoing account that the sugar residue is attached to a nitrogen atom in the base. Recently, however, Davis *et al.* (1957) and Cohn *et al.* (1959) have isolated a new nucleotide from, *e.g.*, yeast R.N.A., which is unique in that it appears to be a *C*-glycoside. This linkage in the nucleoside is *not* broken by acid, and the evidence obtained so far suggests the nucleoside has a 5-substituted uracil structure.

§13d. Structure of nucleotides. When nucleotides are carefully hydrolysed, ribose monophosphate may be isolated from the products; thus the phosphoric acid is attached to the sugar residue in nucleotides. Examination of the nucleoside structures shows that the point of attachment may be 2′, 3′ or 5′ in the ribose molecule, and 3′ or 5′ in the deoxyribose molecule. On reduction with hydrogen in the presence of platinum, ribose phosphate is converted into an optically inactive phosphoribitol (Levene *et al.*, 1932, 1933). This product can be optically inactive only if the phosphate residue is attached to the *centre* hydroxyl group of the ribose molecule, *i.e.*, at the 3′-position.

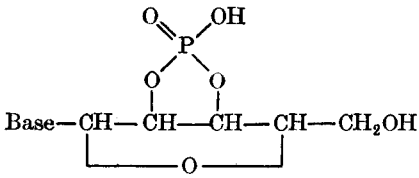


It should be remembered that the furanose structure occurs only when the sugar is in the form of a glycoside; on hydrolysis, the furanose sugar first liberated immediately changes into the stable pyranose form (see §7f. VII).

Until recently, it was believed that the 3′-position was the only one occupied by the phosphate radical. Emden *et al.* (1929) claimed to have isolated a 5′-phosphate (from muscle nucleic acid). Carter and Cohn (1949) isolated two isomeric adenylic acids from the alkaline hydrolysates of R.N.A.s, and called them "*a*" and "*b*" adenylic acids. These authors, in 1950, also isolated two isomers of guanylic, uridylic and cytidylic acids. Carter and Cohn found that one of their adenylic acids was identical with adenosine-3′-phosphate, but the other was not the same as the 5′-compound of Emden. These authors therefore believed that their two isomers were the 2′- and 3′-phosphate. Todd *et al.* (1952) synthesised adenosine-2′- and 3′-phosphate, and showed that their synthetic compounds were identical with the "*a*" and "*b*" acids obtained by Carter and Cohn, but were not able to say which was which. Loring *et al.* (1952) showed that the "*a*" and "*b*" cytidylic acids resist oxidation by periodic acid, and hence it follows that they must be the 2′- and 3′-phosphates (but there is no indication from this which isomer is the 2′- and which is the 3′-); had one isomer been the 5′-compound, then it would have been oxidised by periodic acid (the two hydroxyls on 2′ and 3′ are free and adjacent). A study of the solubility, acidity and absorption spectra of these two cytidylic acids led Loring *et al.* to suggest that the "*a*" acid is the 2′-phosphate. This conclusion has been supported by Harris *et al.* (1953) from their study of the infra-red spectra of these compounds. Todd *et al.* (1954) have synthesised deoxycytidine-3′- phosphate, and comparison of its infra-red spectra and other properties with cytidine phosphates provides strong evidence that "*b*" cytidylic acid is cytidine-3′-phosphate, and therefore that "*b*" uridylic acid is uridine-3′- phosphate. Brown *et al.* (1955) have shown that hydrazine splits "*a*" and "*b*" cytidylic acid to give ribose 2- and 3-phosphate respectively. "*b*" Uridylic acid yields the same ribose phosphate obtained from "*b*" cytidylic acid. Thus the "*a*" and "*b*" isomers of these nucleotides are the 2′- and 3′-phosphates, respectively, of the ribonucleosides.

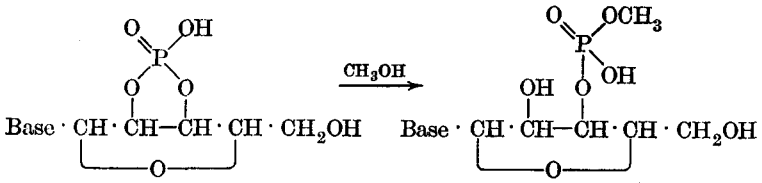
Experiments using enzymic hydrolysis of nucleic acids have shown that these acids also contain 5′-phosphoester links. Cohn *et al.* (1951) have isolated the 5′-phosphates of adenosine, guanosine, uridine and cytidine. These authors have also shown that the nucleotides in calf thymus D.N.A. are 3′- and 5′-phosphates (position 2′ is not possible since this is a CH_2 group).

Thus, according to the foregoing evidence, the phosphate radical can occupy the positions 2′, 3′ and 5′ in ribonucleotides, and 3′ and 5′ in deoxyribonucleotides. These, however, by no means exhaust the possible positions of the phosphate radical. Todd *et al.* (1951) have identified *cyclic* nucleoside phosphates (2′ : 3′-) from the enzymic hydrolysates of R.N.A.s. If these cyclic esters are actually present in nucleic acids, then the 2′- and 3′-phosphates obtained by hydrolysis may arise by the opening of the cyclic compound (either the 2′- or 3′-ester will be obtained). Todd *et al.* (1953)



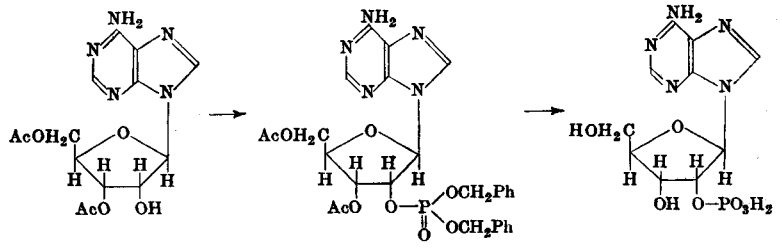
have also isolated thymidine-3′ : 5′- diphosphate and deoxycytidine-3′ : 5′-diphosphate from herring sperm deoxyribonucleic acid.

Heppel *et al.* (1955) have shown that these cyclic esters are converted into the 3′-phosphate in the presence of methanol or ethanol and ribonuclease provided the base is a cytosine or a uracil residue, *e.g.*,

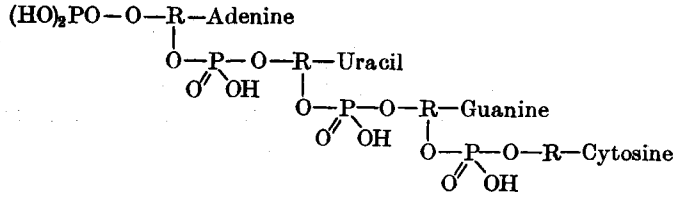


Barker *et al.* (1955) have shown that this reaction occurs only if the alcohol contains a primary alcoholic group, and suggest that if such a reaction is concerned in the biosynthesis of ribopolynucleotides from simpler units, then this requirement (*i.e.*, the primary alcoholic group) might explain why only 3′ : 5′-diester links are present in these polynucleotides (see §13e).

Nucleotides have been synthesised in various ways, *e.g.*, Levene *et al.* (1937) synthesised adenosine-5′-phosphate from 2′,3′-*O*-*iso*propylideneadenosine. This was phosphorylated with phosphoryl chloride in pyridine, followed by careful hydrolysis with acid to remove the *iso*propylidene residue. 2′- and 3′-phosphates are more difficult to synthesise because of their ready interconversion. Todd *et al.* (1954) synthesised adenosine-2′-phosphate by phosphorylating 3′,5′-di-*O*-acetyladenosine in the 2′-position with dibenzylphosphochloridate [$(PhCH_2O)_2POCl$] and removing the benzyl groups (as toluene) by hydrogenation (Pd), and finally removing the acetyl groups by treatment with alkali. Under these conditions no phosphate migration is possible.



§13e. Nucleic acids. Having obtained evidence about the structure of nucleotides, we must now consider the problem concerning their linkage to form nucleic acids. In the early work, when a nucleic acid, obtained by drastic alkaline purification, was subjected to hydrolysis, the products were four molecules of phosphoric acid, four molecules of sugar, two purine molecules and two pyrimidine molecules, *e.g.*, yeast ribonucleic acid gave four molecules of phosphoric acid, four molecules of ribose and one molecule each of adenine, guanine, cytosine and uracil. On this and other evidence (see v below) Levene (1926) was led to propose the "tetranucleotide" theory, *e.g.* (R = ribose):



This simple structure for nucleic acids has, however, been shown to be incorrect by more recent work, *e.g.*,

(i) It has been found that alkaline methods of purification degrade nucleic acids; thus the molecular weight varies with the methods used for the isolation of the acid. Furthermore, fractionation experiments on both R.N.A.s and D.N.A.s (from the same and from different sources) have shown that these nucleic acids are complex mixtures (see also iv).

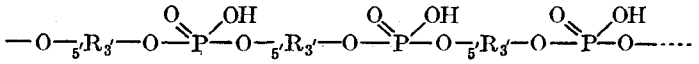
(ii) Various methods for determining molecular weights, *e.g.*, diffusion and the ultracentrifuge, have shown that the molecular weights of the nucleic acids are very high; they range from about 10^5 to 10^7, *e.g.*, a value of about 2×10^6 has been found for the R.N.A. from tobacco mosaic virus.

(iii) X-ray studies have shown that D.N.A.s are composed of two polynucleotide chains wound as spirals round a common axis but head in opposite directions (Wilkins *et al.*, 1953; Watson *et al.*, 1953). The two chains are held together by hydrogen bonds, thus producing a long, thin, relatively rigid molecule. Less is known about the structure of R.N.A.s, but according to Gierer (1957, 1958), the R.N.A. from tobacco mosaic virus is in the form of a flexible, moderately coiled chain.

(iv) The analysis of the hydrolysates of nucleic acids, particularly by chromatography, has shown that the acids from different sources have different chemical compositions (*cf.* i). According to Chargaff (1950), not one specimen of a nucleic acid gave analysis results corresponding to a tetranucleotide; thus the " statistical tetranucleotide " theory is untenable. Chargaff found that in D.N.A.s, the sum of the total purine nucleotides is equal to that of the pyrimidine nucleotides, and that the molar ratios of adenine to thymine, and of guanine to cytosine (or its analogues) are unity. Chargaff *et al.* (1954) also found the same regularities in R.N.A.s, with uracil taking the place of thymine. Chargaff estimated the nucleotide content from spectral data (as well as by some of the earlier methods), and pointed out that the regularities are not usually observed with *purified* samples of pentose nucleic acids, but only when, *e.g.*, whole cells are subjected to hydrolysis.

(v) Levene *et al.* (1926), from electromeric titration experiments, concluded that R.N.A.s show four primary and one secondary phosphate dissociation for each set of four phosphorus atoms present. On this evidence, and on the results of analysis, Levene put forward his tetranucleotide theory (see above). More refined titration experiments, however, have shown that R.N.A.s exhibit only three primary and one secondary phosphate dissociation (Gulland *et al.*, 1944). These latter findings are also supported by methylation experiments (Anderson *et al.*, 1949).

(vi) Various structures have been proposed for the nucleic acids, *e.g.*, Todd (1952) has suggested the following for deoxyribonucleic acids:



D.N.A.

The deoxyribose units are in the furanose form and attached to the phosphate molecule by the $C_{3'}$ and $C_{5'}$ hydroxyl groups; the base is attached to $C_{1'}$ of the sugar. The structure of the R.N.A.s is less certain, but the linkages are believed to be similar to those of the D.N.A.s. All work so far reinforces Todd's suggestion that both types of nucleic acid are 3',5'-linked polynucleotides.

Since the nucleic acids are complex mixtures, the problem of determining nucleotide sequence is very difficult indeed. One method has been the use of enzymes, but used alone, this method has yielded little information. Enzymic methods, however, have been successful in synthesising large polynucleotides, *e.g.*, Kornberg *et al.* (1958) have carried the biosynthesis of D.N.A. by means of enzymes. On the other hand, chemical methods which have been developed are giving some information. The most thoroughly studied stepwise degradation method is the one which depends on the periodate oxidation of the 2',3'-glycol system in the terminal nucleoside residue of a polynucleotide chain (Todd *et al.*, 1953). This method may be used for R.N.A.s, but the absence of the 2'-hydroxyl group in D.N.A.s precludes its use for these acids.

Jones *et al.* (1956) have also developed a chemical method for the specific degradation of deoxyribonucleic acids. These authors have found that on treatment with mercaptoacetic acid ($CH_2SH{\cdot}CO_2H$), purines are removed and replaced by carboxymethylthio groups. By this means it is possible to obtain information on the relative positions of purines and pyrimidines. Thus the results have shown that in calf-thymus deoxyribonucleic acids there are regions in which at least three pyrimidine nucleotides occur in adjacent positions.

A combination of enzymic and chemical methods appears to be the most successful technique. This type of approach was developed by Whitfield (1954) and has been improved by later workers (*inter alia*, Cohn *et al.*, 1961); the method can be applied to end-group analysis or to the analysis of short-chain fragments. Verwoerd *et al.* (1961) have introduced a method involving the use of hydroxylamine followed by enzymic treatment. The hydroxylamine displaces uracil and cytosine nuclei in the nucleic acid, and it has been shown that the enzyme (which normally hydrolyses the acid) does not break the chain at the sites where uracil has been removed.

READING REFERENCES

Fischer, Synthesen in der Puringruppe, *Ber.*, 1899, **32**, 435.

Stewart, *Recent Advances in Organic Chemistry*, Longmans, Green. Vol. I (1931, 6th ed.). Ch. 13. The Purine Group.

Lythgoe, Some Aspects of Pyrimidine and Purine Chemistry, *Quart. Reviews* (*Chem. Soc.*), 1949, **3**, 181.

Gilman (Ed.), *Advanced Organic Chemistry*, Wiley (1938, 1st ed.). Vol. II. Ch. 11. The Chemistry of Pyrimidines, Purines and Nucleic Acids.

Levene and Bass, *Nucleic Acids*, Chemical Catalogue Co. (1931).

Davidson, *The Biochemistry of the Nucleic Acids*, Methuen (1953, 2nd ed.).

Tipson, The Chemistry of the Nucleic Acids, *Advances in Carbohydrate Chemistry*, Academic Press, Vol. I (1945).

Schlenk, Chemistry and Enzymology of Nucleic Acids, *Advances in Enzymology*, Interscience Publishers, 1949, **9**, 455.

Rodd (Ed.), *Chemistry of Carbon Compounds*, Elsevier. Vol. IVC (1960). Ch. XX. Purines and Related Ring Systems. Ch. XXI. Nucleosides, Nucleotides and Nucleic Acids.

Avison and Hawkins, The Role of Phosphoric Esters in Biological Reactions, *Quart. Reviews* (*Chem. Soc.*), 1951, **5**, 171.

Baddiley and Buchanan, Recent Developments in the Biochemistry of Nucleotide Coenzymes, *Quart. Reviews* (*Chem. Soc.*), 1958, **12**, 152.

Fairley, Nucleic Acids, Genes, and Viruses, *J. Chem. Educ.*, 1959, **36**, 544.

Roth, Ribonucleic Acid and Protein Synthesis, *J. Chem. Educ.*, 1961, **38**, 217.
Todd, Some Current Problems in Polynucleotide Chemistry, *Proc. Chem. Soc.*, 1961, 187.
Verwoerd *et al.*, Specific Partial Hydrolysis of Nucleic Acids in Nucleotide Sequence Studies, *Nature*, 1961, **192**, 1038.
Spencer *et al.*, Determination of the Helical Configuration of Ribonucleic Acid Molecules by X-Ray Diffraction, *Nature*, 1962, **194**, 1014.

CHAPTER XVII

VITAMINS

§1. Introduction. In addition to oxygen, water, proteins, fats, carbohydrates and certain inorganic salts, a number of organic compounds are also necessary for the life, growth and health of animals (including man). These compounds are known as the " accessory dietary factors " or *vitamins*, and are only necessary in very small amounts. Vitamins cannot be produced by the body and hence must be supplied. Vitamin D, however, may be supplied in food or may be produced in the skin by irradiation (ultraviolet) of sterols.

Many vitamins have now been isolated and their structures elucidated. As each vitamin was isolated, it was named by a letter of the alphabet, but once its structure had been established (or almost established), the vitamin has generally been renamed (see text).

The vitamins have been arbitrarily classified into the " fat-soluble group " (vitamins A, D, E and K), and the " water-soluble group " (the remainder of the vitamins).

A number of vitamins have already been dealt with in various chapters dealing with natural products with which these particular vitamins are closely associated chemically, *viz.* vitamins A_1 and A_2 (§7. IX), vitamin C (§11. VII) and the vitamin D group (§§6, 6a, 6b. XI). This chapter is devoted to a number of other vitamins (see the reading references for further information).

From the point of view of chemical structure, there is very little common to the various vitamins, but from the point of view of chemical reactions, many of the water-soluble vitamins have one feature in common, and that is their ability to take part in reversible oxidation–reduction processes. Thus they form a part of various co-enzymes (see §17. XIII), *e.g.*, nicotinamide is present in co-enzyme I (diphosphopyridine nucleotide; DPN), and in co-enzyme II (triphosphopyridine nucleotide; TPN); phosphorylated pyridoxal is the co-enzyme of transaminases; riboflavin in flavin adenine nucleotide (FAD); pantothenic acid in co-enzyme A; etc.

VITAMIN B COMPLEX

§2. Introduction. Eijkman (1897) found that birds developed polyneuritis when fed with polished rice, and were cured when they were given rice polishings. Then Grijns (1901) found that rice polishings cured beriberi in man (beriberi in man corresponds to polyneuritis in birds; it is a form of paralysis). Grijns suggested that the cause of this paralysis was due to some " deficiency " in the diet, and this was confirmed by Funk (1911, 1912), who prepared a concentrate of the active substance from rice polishings. Funk believed that this active substance was a definite chemical compound, and since he separated organic bases when he prepared his concentrate, he named his " deficiency compound " a *vitamine*. It was then found that " vitamine B " was a complex mixture, and when a number of " vitamines " were obtained that contained no nitrogen, the name *vitamin* was retained for them. The name vitamin B is now reserved for the complex mixture of vitamins in this group.

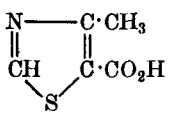
§3. Vitamin B_1, thiamine (aneurin). Thiamine is one member of the water-soluble vitamin B complex, and is in the thermolabile fraction; it is the absence of thiamine which is the cause of beriberi in man; thus this vitamin is the antineuritic factor (hence the name *aneurin*). Rice polishings and yeast have been the usual sources of thiamine; eggs are also a rich source.

Thiamine is obtained crystalline in the form of its salts; the chloride hydrochloride has been shown to have the molecular formula $C_{12}H_{18}ON_4Cl_2S$ (Windaus *et al.*, 1932); this salt is isolated in the form of its hemihydrate, d. 248–250°. When treated with a sodium sulphite solution saturated with sulphur dioxide at room temperature, thiamine is decomposed quantitatively into two compounds which, for convenience, we shall label A and B (R. R. Williams *et al.*, 1935).

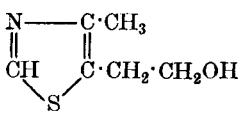
$$C_{12}H_{18}ON_4Cl_2S + Na_2SO_3 \rightarrow \underset{\text{A}}{C_6H_9ONS} + \underset{\text{B}}{C_6H_9O_3N_3S} + 2NaCl$$

Compound A, C_6H_9ONS. This compound shows basic properties, and since it does not react with nitrous acid, it was inferred that the nitrogen atom is in the tertiary state. The functional nature of the oxygen atom was shown to be alcoholic, *e.g.*, when A is treated with hydrochloric acid, a hydroxyl group (one oxygen atom and one hydrogen atom) is replaced by a chlorine atom. Furthermore, since the absorption spectrum of the chloro derivative is almost the same as that of the parent (hydroxy) compound, this suggests that the hydroxyl group is in a side-chain. The sulphur did not give the reactions of a mercapto compound nor of a sulphide; in fact, the stability (*i.e.*, unreactivity) of this sulphur atom led to the suggestion that it was in a heterocyclic ring. This conclusion was confirmed by the fact that A has an absorption spectrum characteristic of a thiazole (§5. XII).

R. R. Williams *et al.* (1935) found that oxidation of A with nitric acid gives the compound $C_5H_5O_2NS$, which can also be obtained by the direct oxidation of thiamine with nitric acid. This latter reaction had actually been carried out by Windaus *et al.* (1934), but these workers had not recognised the presence of the thiazole nucleus. Williams *et al.* showed that this oxidation product was a monocarboxylic acid, and found that it was identical with 4-methylthiazole-5-carboxylic acid, I, a compound already described in the literature (Wöhmann, 1890). From this it follows that A has a side-chain of two carbon atoms in place of the carboxyl group in I

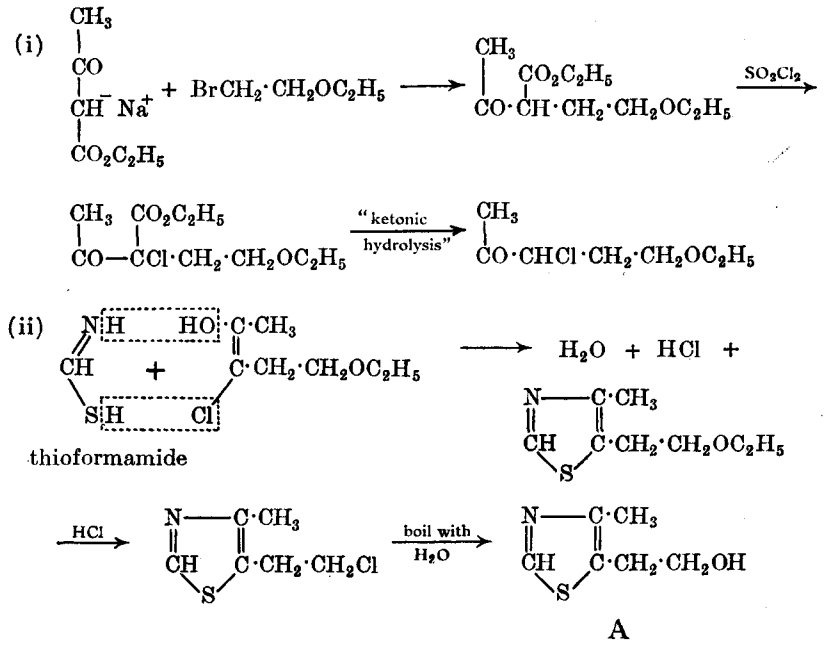


I



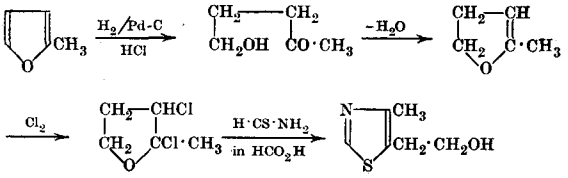
II

(one carbon atom is lost when A is oxidised to I). Since it is this side-chain which must contain the alcoholic group, the side-chain could be either —CH_2·CH_2OH or —CHOH·CH_3. Either of these could lose a carbon atom to form a carboxyl group directly attached to the thiazole nucleus. The second alternative, —CHOH·CH_3, was excluded by the fact that A does not give the iodoform test, and that A is not optically active (the second alternative contains an asymmetric carbon atom). Thus A was given structure II, and this has been confirmed by synthesis (Clarke *et al.*, 1935).



The hydrochloride of this compound is identical with that of the product obtained from thiamine (by fission), and also gives I on oxidation with nitric acid.

Londergan *et al.* (1953) have synthesised A from 2-methylfuran as follows:

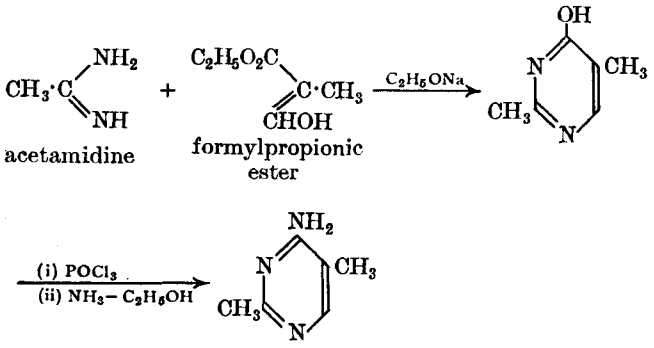


Compound B, $C_6H_9O_3N_3S$. This was shown to be a sulphonic acid, *e.g.*, when heated with water under pressure at 200°, B gives sulphuric acid; it also forms sodium sulphite when heated with concentrated sodium hydroxide solution. On treatment with nitrous acid, B evolves nitrogen; thus B contains one or more amino-groups. Analysis of the product showed that one amino-group is present in B (the product contained only one hydroxyl group). Furthermore, since the evolution of nitrogen was slow, and the reaction of B with benzoyl chloride was also slow, this suggests that B contains an amidine structure (Williams *et al.*, 1935). Williams *et al.* (1935) then heated B with hydrochloric acid at 150° under pressure, and obtained

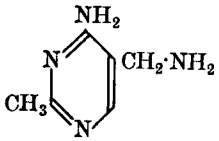
$$\underset{\text{B}}{C_6H_9O_3N_3S} + H_2O \xrightarrow{HCl} \underset{\text{C}}{C_6H_8O_4N_2S} + NH_3$$

compound C and ammonia. The formation of ammonia indicates the replacement of an amino-group by a hydroxyl group. This type of reaction is characteristic of 2- and 6-aminopyrimidines; it was therefore inferred that B is a pyrimidine derivative (*cf.* §14. XII). This is supported by the fact that the ultraviolet absorption spectrum of compound C was similar to that of synthetic 6-hydroxypyrimidines; thus B is probably a 6-aminopyrimidine.

When B is reduced with sodium in liquid ammonia, a sulphonic acid group is eliminated with the formation of an aminodimethylpyrimidine (Williams, 1936). Comparison of the ultraviolet absorption spectrum of this product with various synthetic compounds showed that it was 6-amino-2 : 5-dimethylpyrimidine, and this was confirmed by synthesis (Williams *et al.*, 1937).

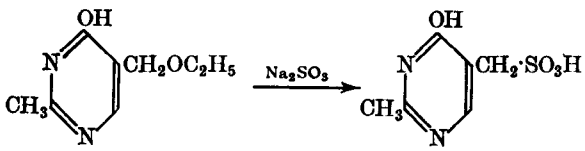


Thus B is 6-amino-2 : 5-dimethylpyrimidine with one hydrogen atom (other than one of the amino-group) replaced by a sulphonic acid group. When thiamine is treated with sodium in liquid ammonia, one of the products is the diamino derivative D, $C_6H_{10}N_4$ (Williams *et al.*, 1937). Compound D was identified as 6-amino-5-aminomethyl-2-methylpyrimidine by comparison



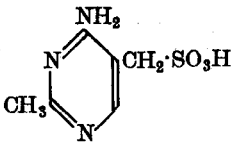
D

with the absorption spectra of methylated aminopyrimidines of known structure (Williams *et al.*, 1937). This is confirmed by the synthesis of Grewe (1936); Williams *et al.* had arrived at their conclusion independently of Grewe's work (see below for this synthesis). Thus, in compound D, there is an amino-group instead of the sulphonic acid group in B. Williams therefore concluded that the sulphonic acid group (in B) is joined to the methyl group at position 5. This was confirmed (in 1937) by treating 5-ethoxymethyl-6-hydroxy-2-methylpyrimidine (see the synthesis described for thiamine) with sodium sulphite, whereby 6-hydroxy-2-methylpyrimidyl-5-methanesulphonic acid was obtained, and this was shown to be identical with compound C.



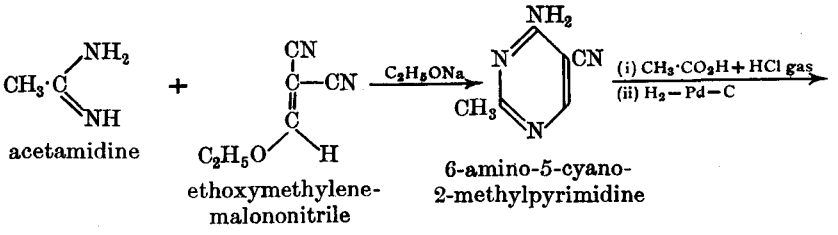
C

Thus B has the following structure:

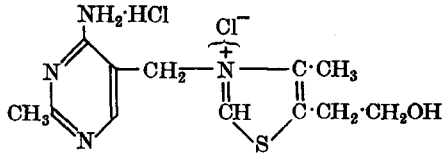


B

This structure is confirmed by synthesis (Grewe, 1936; Andersag *et al.*, 1937).

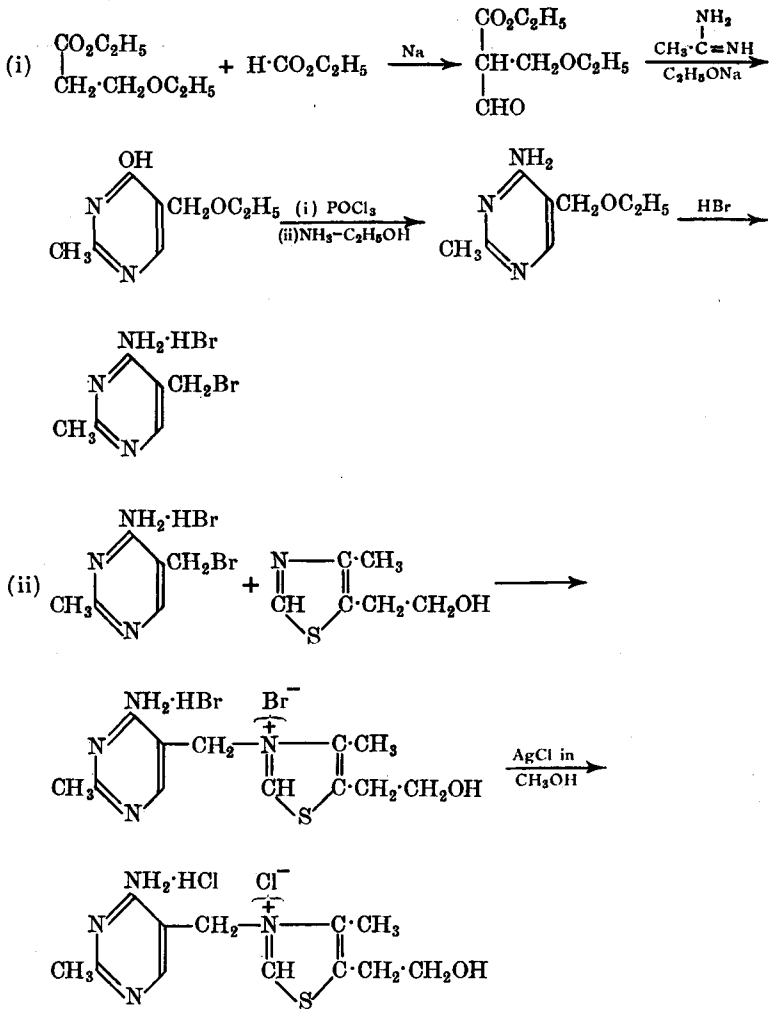


The final problem is: How are fragments A and B united in thiamine? As we have seen, the sulphonic acid group in B is introduced during the fission of thiamine with sodium sulphite; thus the point of attachment of fragment B is at the CH_2 group at position 5. To account for the formation of compound D, fragment B must be linked to the nitrogen atom of fragment A; in this position, the nitrogen atom of the thiazole ring is in a quaternary state, and so accounts for the chloride hydrochloride of thiamine. Had B been connected to A through a carbon atom of the latter, it would not be easy to account for the ready fission of this carbon–carbon bond by means of sodium and liquid ammonia, nor for the fact that thiamine does not form a *dihydrochloride.* Thus the chloride hydrochloride of thiamine is



thiamine chloride hydrochloride

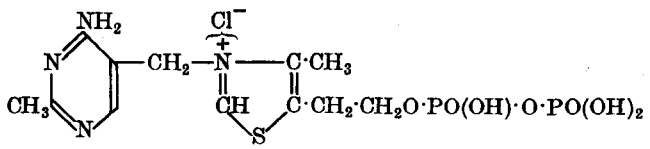
This structure has been confirmed by synthesis, *e.g.*, that of Williams *et al.* (1936, 1937).



§4. **Co-carboxylase.** This is the co-enzyme of *carboxylase*, and has been shown to be the pyrophosphate of thiamine (Lohmann *et al.*, 1937). Carboxylase, which requires the co-enzyme for action (see §15. XIII), breaks down pyruvic acid, formed in alcoholic fermentation, to acetaldehyde and carbon dioxide.

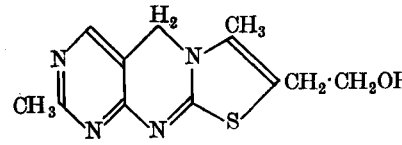
$$CH_3{\cdot}CO{\cdot}CO_2H \xrightarrow{\text{carboxylase}} CH_3{\cdot}CHO + CO_2$$

Co-carboxylase is



§5. **Thiochrome** was isolated from yeast by Kuhn *et al.* (1935); it is a yellow basic solid and its solutions show a blue fluorescence. Thiochrome

is also formed by the oxidation of thiamine with alkaline potassium ferricyanide (Todd *et al.*, 1935); it has also been synthesised by Todd *et al.* (1936).



thiochrome

§6. Vitamin B_2, riboflavin (lactoflavin), $C_{17}H_{20}O_6N_4$. Riboflavin is a water-soluble, thermostable vitamin which occurs in the vitamin B complex. It is necessary for growth and health, and occurs widely distributed in nature, *e.g.*, in yeast, green vegetables, milk, meat, etc. Chemically, vitamin B_2 is closely related to the yellow water-soluble pigments known as the *flavins* (*iso*alloxazines), and since it was first isolated from milk, vitamin B_2 is also known as *lactoflavin*.

Riboflavin is a bright yellow powder, m.p. 292°, showing a green fluorescence; it is soluble in water and in ethanol, but is insoluble in chloroform and other organic solvents.

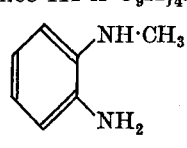
When exposed to light, lactoflavin in sodium hydroxide solution forms mainly lumi-lactoflavin, $C_{13}H_{12}O_2N_4$ (this is soluble in chloroform). Lumi-lactoflavin, on boiling with barium hydroxide solution, is hydrolysed to one molecule of urea and one molecule of the barium salt of a β-ketocarboxylic acid, I, $C_{12}H_{12}O_3N_2$ (Kuhn *et al.*, 1933, 1934). The nature of this acid is shown by the fact that, on acidification of the barium salt, the free acid immediately eliminates carbon dioxide to form the compound, II, $C_{11}H_{12}ON_2$. This compound showed the properties of a lactam, and on vigorous hydrolysis by boiling with sodium hydroxide solution, it forms one molecule of glyoxylic acid and one molecule of the compound $C_9H_{14}N_2$ (III).

$$\underset{\text{lactoflavin}}{C_{17}H_{20}O_6N_4} \xrightarrow[\text{light}]{\text{NaOH}} \underset{\text{lumi-lactoflavin}}{C_{13}H_{12}O_2N_4} \xrightarrow{Ba(OH)_2}$$

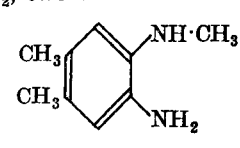
$$CO(NH_2)_2 + \underset{\text{I}}{[C_{12}H_{12}O_3N_2]} \xrightarrow[-CO_2]{\text{acid}}$$

$$\underset{\text{II}}{C_{11}H_{12}ON_2} \xrightarrow{\text{NaOH}} CHO{\cdot}CO_2H + \underset{\text{III}}{C_9H_{14}N_2}$$

The structure of III was elucidated as follows (Kuhn *et al.*, 1934). Preliminary tests showed that III was an aromatic diamino compound. Then it was found that it gave a blue precipitate with ferric chloride, and since this reaction is characteristic of monomethyl-*o*-phenylenediamine, it suggests that III contains the following nucleus, IV. The molecular formula of IV is $C_7H_{10}N_2$, and since III is $C_9H_{14}N_2$, two carbon and four hydrogen atoms



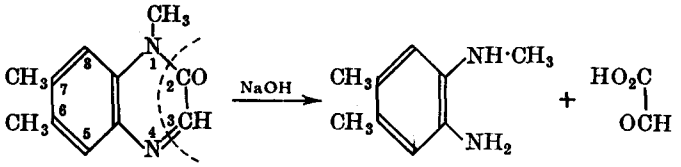
IV



III

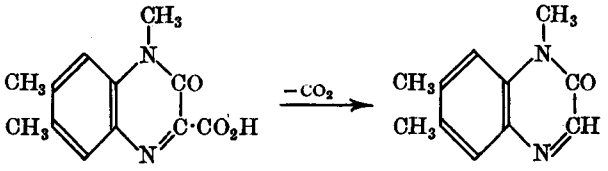
must be accounted for. This can be done by assuming the presence of an ethyl group or of two methyl groups in the benzene ring. Kuhn *et al.*

carried out a series of synthetic experiments and showed that III has the structure given, *N*-methyl-4 : 5-diamino-*o*-xylene. Kuhn then proposed II as the structure of the precursor of III, since this would produce the required products of hydrolysis.



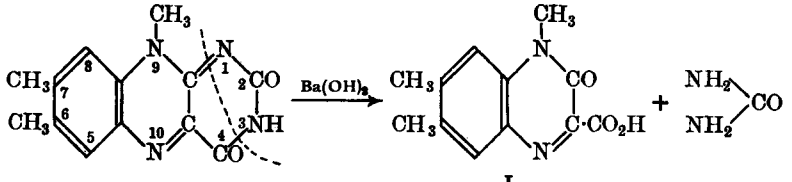
II III

II could therefore have been produced from the β-ketocarboxylic acid I



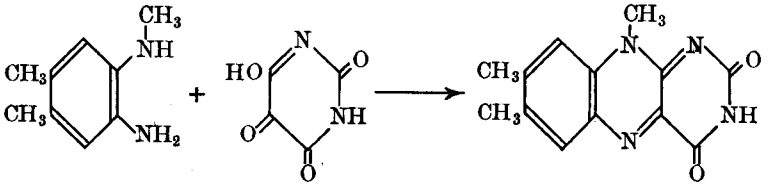
I II

Since I and a molecule of urea are obtained from lumi-lactoflavin, the latter could be 6 : 7 : 9-trimethyl*iso*alloxazine (6 : 7 : 9-trimethylflavin).



lumi-lactoflavin I

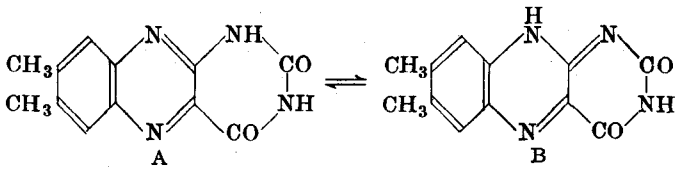
This structure for lumi-lactoflavin has been confirmed by synthesis (Kuhn *et al.*, 1934). *N*-Methyl-4 : 5-diamino-*o*-xylene is condensed with alloxan hydrate (§2. XVI) in aqueous solution at 50–60°.



Methylation (methyl sulphate) of this synthetic product gives a tetramethyl compound identical with the product obtained by the methylation of the natural lumi-lactoflavin.

Side-chain of lactoflavin

Exposure of a *neutral* solution of lactoflavin to light produces *lumichrome*, $C_{12}H_{10}O_2N_4$ (Karrer *et al.*, 1934). Analytical work similar to that described for lumi-lactoflavin showed that the structure of lumichrome is 6 : 7-dimethylalloxazine (A).

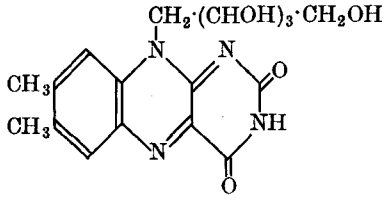


A B

lumichrome

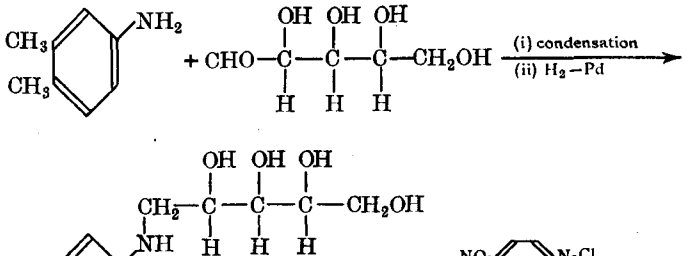
The *iso*alloxazine (structure B) is a tautomer of the alloxazine (structure A); B does not exist as such, but this structure is fixed when there is a substituent at position 9 (see also §25. XII). Stern *et al.* (1934) have shown that the absorption spectra of compounds containing a 9-substituent are different from those in which the mobile 9-hydrogen atom is present. Also, in the latter case, the alloxazine structure (A) predominates.

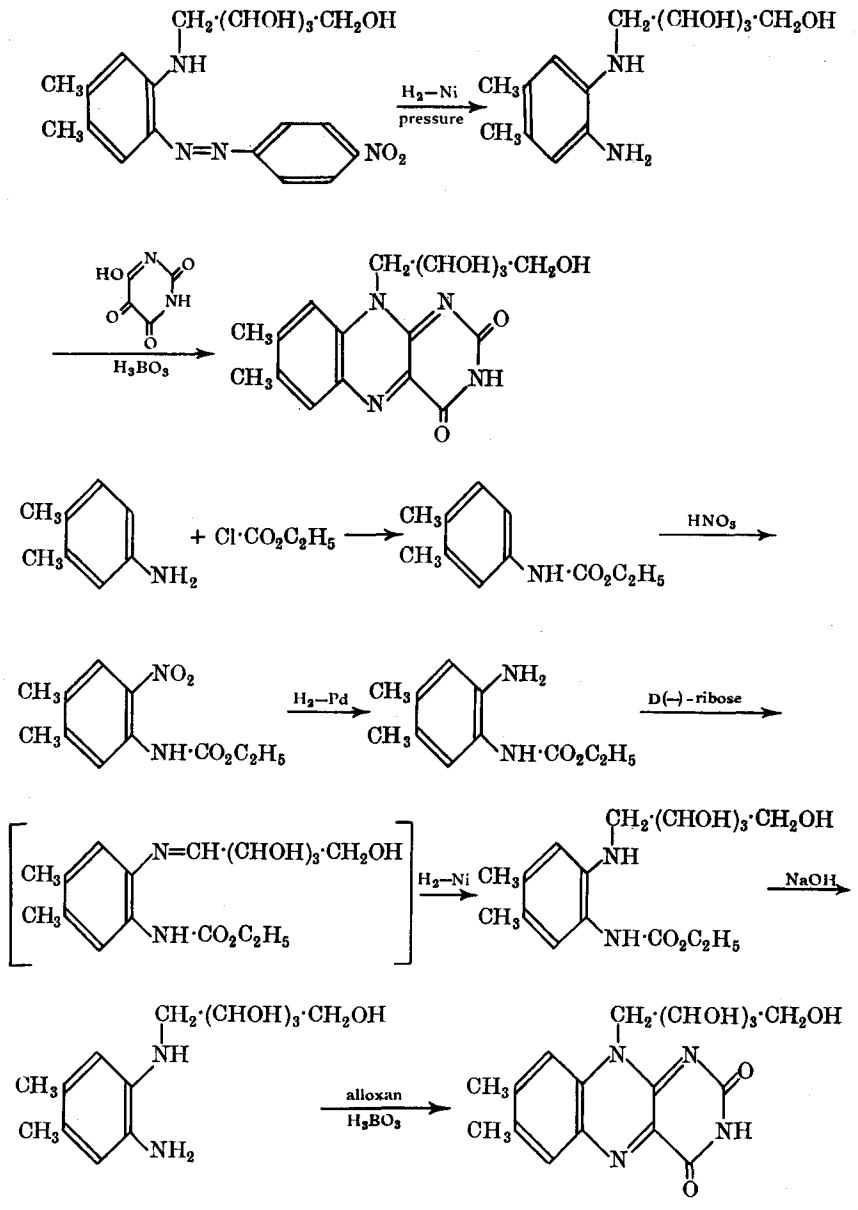
Thus lumichrome is lumi-lactoflavin with a hydrogen atom instead of a methyl group at position 9. This suggests that lactoflavin contains a side-chain (of five carbon atoms) attached to N_9. The Zerewitinoff procedure shows that lactoflavin contains five active hydrogen atoms; thus the molecule contains four hydroxyl groups (one active hydrogen atom is the hydrogen of the NH group at position 3). The presence of these four hydroxyl groups is supported by the fact that the silver salt of lactoflavin (the silver atom replaces the hydrogen of the NH group) forms a tetra-acetate. Thus the side-chain is a tetra-hydroxy derivative, and so a possible structure for lactoflavin is:



lactoflavin

This side-chain contains three asymmetric carbon atoms, and so there are eight optically active forms possible. Which configuration is actually present was solved by synthesising a number of pentose derivatives, and it was finally shown by Karrer *et al.* (1935) that the configuration is that of D(−)-ribose. The following syntheses are due to Karrer *et al.* (1935).





Thus lactoflavin is 6 : 7-dimethyl-9-[D-1′-ribityl]-*iso*alloxazine. Of all the pentoses (and hexoses) used, only the compound from D-ribose shows growth-promoting properties. For this reason vitamin B_2 (lactoflavin) is also known as riboflavin. More recently, however, it has been found that L-lyxoflavin occurs naturally; it has been synthesised and shows some vitamin activity (Folkers *et al.*, 1951).

Biosynthetic experiments have established that riboflavin is formed from purine precursors (McNutt, 1954, 1956; Goodwin *et al.*, 1954; Plaut *et al.*, 1959). It has also been shown that the dimethylbenzene system is derived from acetate

units (Plaut, 1954; Birch *et al.*, 1957; Goodwin *et al.*, 1958). Cresswell and Wood (1960) have synthesised riboflavin by a method which has possible implications in the biosynthesis of this vitamin.

§7. **Pantothenic acid,** $C_9H_{17}O_5N$, is a chick antidermatitis factor, and is also capable of promoting the growth of yeast and of bacteria; it has been isolated from many sources, *e.g.*, liver, kidney, yeast, etc.

Pantothenic acid shows the reactions of a monocarboxylic acid, *e.g.*, it can be esterified to form monoesters (R. J. Williams *et al.*, 1939). The application of the method for determining active hydrogen atoms shows that pantothenic acid contains two hydroxyl groups, and since the acid condenses with benzaldehyde (to form a benzylidene derivative) and with acetone (to form an *iso*propylidene derivative), this suggests that the two hydroxy groups are in either the 1 : 2- or 1 : 3-position (*cf.* §§8, 9. VII). When warmed with dilute hydrochloric acid, pantothenic acid is hydrolysed into compounds I and II. Investigation of I showed that it was β-alanine

$$\underset{}{C_9H_{17}O_5N} \xrightarrow{HCl} \underset{\text{I}}{C_3H_7O_2N} + \underset{\text{II}}{C_6H_{10}O_3}$$

(actually present as the hydrochloride, $\overset{-}{Cl}\{\overset{+}{H_3N}\cdot CH_2\cdot CH_2\cdot CO_2H$). On the other hand, when hydrolysed with alkali, pantothenic acid forms β-alanine (I) and the salt of an acid which, on acidification, spontaneously forms the lactone II. Thus the free acid of II is probably a γ- or δ-hydroxycarboxylic acid; also, since the rate of lactonisation is fast, II is more likely a γ-lactone than a δ-lactone (*cf.* §7c. VII). As pointed out above, pantothenic acid contains two hydroxyl groups. One of these has now been accounted for, and so the problem is to find the position of the second one. This was shown to be α- by the fact that the sodium salt of the acid of the lactone II gives a canary-yellow colour with ferric chloride (a test characteristic of α-hydroxyacids), and also by the fact that II, on warming with concentrated sulphuric acid, liberates carbon monoxide (a test also characteristic of α-hydroxyacids). Thus II is most probably the γ-lactone of an α-hydroxyacid (R. J. Williams *et al.*, 1940).

II was shown to contain one active hydrogen atom, and the application of the Kuhn–Roth methyl side-chain determination (§3. IX) showed the presence of a *gem*-dimethyl group (Stiller *et al.*, 1940); the presence of this group is confirmed by the formation of acetone when the lactone II is oxidised with barium permanganate. Thus a possible structure for II is α-hydroxy-β : β-dimethyl-γ-butyrolactone:

$$\underset{\text{II}}{CH_2\text{—}C(CH_3)_2\text{—}CHOH\text{—}CO} \equiv C_6H_{10}O_3$$

This has been confirmed as follows. Treatment of the lactone with methylmagnesium iodide, followed by hydrolysis, gives a trihydric alcohol which, on oxidation with lead tetra-acetate, gives acetone and an aldehyde. This aldehyde, on oxidation with silver oxide, gave a compound III, which was shown to be β-hydroxy-α : α-dimethylpropionic acid. The foregoing reactions may be formulated as follows:

$$\underset{\text{II}}{CH_2\cdot C(CH_3)_2\cdot CHOH\cdot CO} \xrightarrow[\text{(ii) } H_2O]{\text{(i) } CH_3MgI}$$

$$CH_2OH{\cdot}C(CH_3)_2{\cdot}CHOH{\cdot}C(OH)(CH_3)_2 \xrightarrow{(CH_3{\cdot}CO_2)_4Pb} CH_3{\cdot}CO{\cdot}CH_3 + CH_2OH{\cdot}C(CH_3)_2{\cdot}CHO \xrightarrow{Ag_2O} CH_2OH{\cdot}C(CH_3)_2{\cdot}CO_2H$$

III

Examination of II shows that it contains one asymmetric carbon atom. The lactone, **pantolactone** (the acid is known as **pantoic acid**), obtained from pantothenic acid is lævorotatory, and the structure assigned to it has been confirmed by synthesis (Stiller *et al.*, 1940).

$$(CH_3)_2CH{\cdot}CHO + CH_2O \xrightarrow{K_2CO_3} (CH_3)_2C(CH_2OH)(CHO) \xrightarrow[(ii)\ KCN]{(i)\ NaHSO_3}$$

*iso*butyraldehyde formalin

$$(CH_3)_2C(CH_2OH)(CHOH{\cdot}CN) \xrightarrow[\text{boil}]{HCl} (CH_3)_2C(CH_2{-}O)(CHOH{-}CO)$$

(±)-lactone

The (±)-lactone (as the sodium salt of the acid) was resolved with quinine hydrochloride, and the (—)-form was identical with the lactone obtained from pantothenic acid.

In pantothenic acid, the nitrogen atom is not basic. Also, since hydrolysis of pantothenic acid produces a free amino-group (in β-alanine), this suggests that the group —CO·NH— is present, *i.e.*, pantothenic acid is an amide. Thus the hydrolysis may be formulated:

$$CH_2OH{\cdot}C(CH_3)_2{\cdot}CHOH{\cdot}CO{\cdot}NH{\cdot}CH_2{\cdot}CH_2{\cdot}CO_2H$$

pantothenic acid

$$\downarrow HCl$$

$$[CH_2OH{\cdot}C(CH_3)_2{\cdot}CHOH{\cdot}CO_2H] + NH_2{\cdot}CH_2{\cdot}CH_2{\cdot}CO_2H$$

$$\downarrow \text{lactonises}$$

$$CH_2{\cdot}C(CH_3)_2{\cdot}CHOH{\cdot}CO \quad (O\text{ bridges } CH_2 \text{ and } CO)$$

This interpretation of the results has been proved by the synthesis of pantothenic acid. Stiller *et al.* (1940) warmed pantolactone (synthesised as described above) with the ethyl ester of β-alanine, and removed the ester group by hydrolysis with a cold solution of barium hydroxide.

$$CH_2{\cdot}C(CH_3)_2{\cdot}CHOH{\cdot}CO + NH_2{\cdot}CH_2{\cdot}CH_2{\cdot}CO_2C_2H_5$$

$$\downarrow \text{warm}$$

$$CH_2OH{\cdot}C(CH_3)_2{\cdot}CHOH{\cdot}CO{\cdot}NH{\cdot}CH_2{\cdot}CH_2{\cdot}CO_2C_2H_5$$

$$\downarrow Ba(OH)_2$$

$$CH_2OH{\cdot}C(CH_3)_2{\cdot}CHOH{\cdot}CO{\cdot}NH{\cdot}CH_2{\cdot}CH_2{\cdot}CO_2H$$

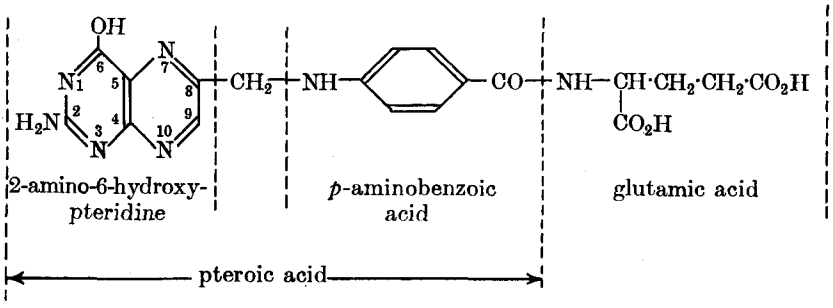
A better yield of pantothenic acid is obtained by warming the dry lactone with the dry sodium salt of β-alanine (R. J. Williams *et al.*, 1940).

§8. Folic acid complex. A number of micro-organisms need certain concentrates (prepared from natural sources) for growth; several active principles have been shown to be necessary. In addition to the above property, some of these active principles also exhibit other effects, *e.g.*, the prevention of certain types of anæmia in chicks. The following compounds have been described as constituents of the folic acid complex.

(i) Folic acid. It has been suggested that folic acid from animal sources is different from that from vegetable sources.
(ii) *Lactobacillus casei* factor (three forms).
(iii) *S. lactis* R factor.
(iv) Vitamin B_c factor (this now identified as liver *L. casei* factor).
(v) Vitamin B_c conjugate.
(vi) Vitamins B_{10}, B_{11} and factors R, S and U.
(vii) Vitamin M.

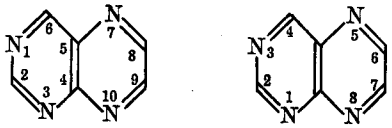
It is possible that some of these are identical; names have been given by different workers to the active substances that they have isolated (see, *e.g.*, iv).

Angier *et al.* (1946) have shown that liver *L. casei* factor (also called vitamin B_c) is:



Fermentation *L. casei* factor contains three glutamic acid residues; yeast vitamin B_c conjugate contains seven glutamic acid residues.

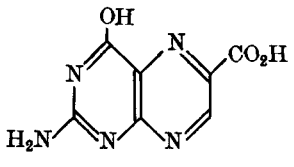
§8a. Structure of *L. casei* factors (Angier *et al.*, 1946). The alkaline hydrolysis of the fermentation *L. casei* factor, in the absence of oxygen, formed two molecules of D-glutamic acid and the DL-form of liver *L. casei* factor. On the other hand, the alkaline hydrolysis of the fermentation *L. casei* factor, in the presence of air, gave two substances, I and II. I was



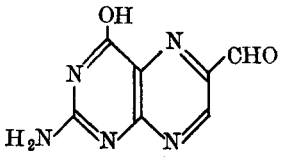
A pteridine B

shown to be a monocarboxylic acid, and the examination of its ultraviolet absorption spectrum led to the conclusion that it was a pteridine derivative (A is the system of numbering used here; B is an alternative system of

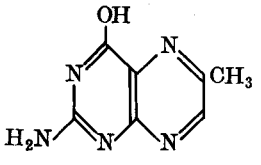
numbering frequently used in American publications). A further examination of compound I showed that it also contained one hydroxyl and one amino-group. Oxidation of I with chlorine water, followed by hydrolysis with hydrochloric acid, produced guanidine, $NH{=}C(NH_2)_2$, as one of the products. The formation of this compound suggests that the amino-group is at position 2. Finally, I was shown to be 2-amino-6-hydroxypteridine-8-carboxylic acid by synthesis.



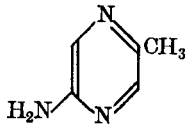
I



IV



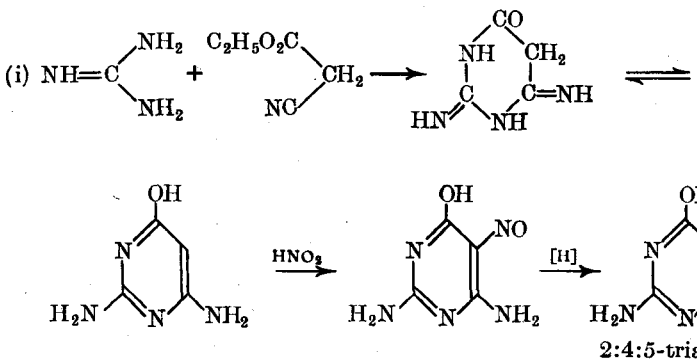
V



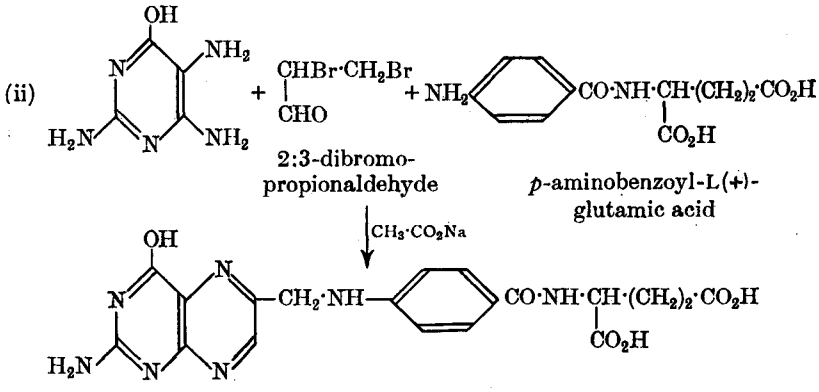
VI

The reactions of compound II showed that it was a primary aromatic amine, and on hydrolysis it gave one molecule of *p*-aminobenzoic acid and three molecules of glutamic acid.

Hydrolysis of the fermentation *L. casei* factor with sulphurous acid gave an aromatic amine, III, and an aldehyde, IV. III, on hydrolysis, gave one molecule of *p*-aminobenzoic acid and three molecules of glutamic acid, *i.e.*, II and III are identical. When the aldehyde IV was allowed to stand in dilute sodium hydroxide solution in the absence of air, compound I and another compound, V, were produced. V, on vigorous hydrolysis, gave 2-amino-5-methylpyrazine, VI. From this it was concluded that V is 2-amino-6-hydroxy-8-methylpteridine, and IV is 2-amino-6-hydroxypteridine-8-aldehyde. Consideration of this evidence led to the suggestion that the liver *L. casei* factor has the structure given in §8; this has been confirmed by synthesis, *e.g.*, that of Angier *et al.* (1946).

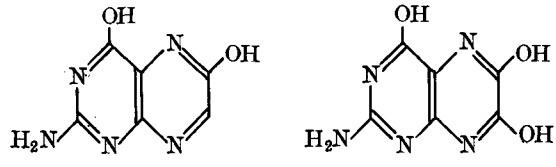


2:4:5-triamino-6-hydroxypyrimidine



liver *L. casei* factor

It might be noted, in passing, that the **pterins** are pigments of butterfly wings, wasps, etc.; they were first isolated from butterfly wings.



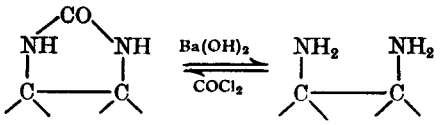
xanthopterin leucopterin

§9. **Biotins** (vitamin H). Bios, an extract of yeast, was shown to be necessary for the growth of yeast (Wildiers, 1901). It was then found that bios consisted of at least two substances (Fulmer *et al.*, 1922), and two years later, Miller showed that three substances were present in bios. The first of these was named Bios I, and was shown to be *meso*inositol (Eastcott, 1928; see also §13). The second constituent, named Bios IIA, was then shown to be β-alanine (Miller, 1936) or pantothenic acid (Rainbow *et al.*, 1939). The third substance, named Bios IIB, was found to be identical with *biotin*, a substance that had been isolated by Kögl *et al.* (1936) as the methyl ester from egg-yolk. Subsequently other factors present in bios have been isolated, *e.g.*, pyridoxin (see §10) and nicotinic acid (§11).

Biotin is a vitamin, being necessary for the growth of animals. In 1940, du Vigneaud *et al.* isolated from liver a substance which had the same biological properties as biotin. Kögl *et al.* (1943) named their extract from egg-yolk α-biotin, and that from liver β-biotin. Both compounds have the same molecular formula $C_{10}H_{16}O_3N_2S$.

β-Biotin (Bios IIB or biotin), m.p. 230–232°, behaves as a saturated compound (the usual tests showed the absence of an ethylenic double bond). β-Biotin forms a monomethyl ester $C_{11}H_{18}O_3N_2S$ which, on hydrolysis, gives an acid the titration curve of which corresponds to a monocarboxylic acid; thus the formula of β-biotin may be written $C_9H_{15}ON_2S{\cdot}CO_2H$. When heated with barium hydroxide solution at 140°, β-biotin is hydrolysed to carbon dioxide and diaminocarboxylic acid $C_9H_{18}O_2N_2S$ which, by the action of carbonyl chloride, is reconverted into β-biotin (du Vigneaud *et al.*, 1941). These reactions suggest that β-biotin contains a cyclic ureide structure. Furthermore, since the diaminocarboxylic acid condenses with phenanthraquinone to form a quinoxaline derivative, it follows that the two amino-

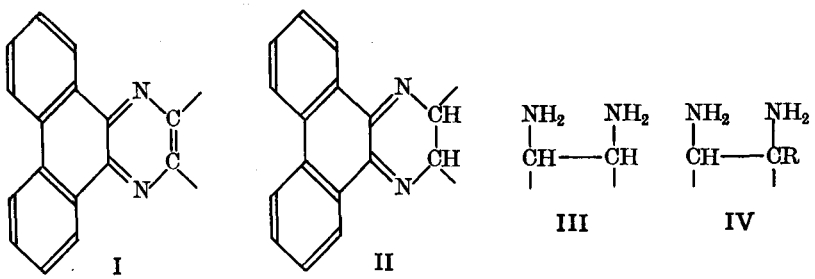
groups are in the 1 : 2-positions (*cf.* §19. XII), and thus the cyclic ureide is five-membered. Hence we may write the foregoing reactions as follows:



β-biotin diamino-compound

When this diaminocarboxylic acid is oxidised with alkaline permanganate, adipic acid is produced (du Vigneaud *et al.*, 1941). One of the carboxyl groups in adipic acid was shown to be that originally present in β-biotin as follows. When the carbomethoxyl group of the methyl ester of β-biotin was replaced by an amino-group by means of the Curtius reaction (ester → hydrazide → azide → urethan → NH_2; see Vol. I), and the product hydrolysed with barium hydroxide solution, a triamine was obtained which did not give adipic acid on oxidation with alkaline permanganate (du Vigneaud *et al.*, 1941, 1942). It was therefore inferred that β-biotin contains a $-(CH_2)_4{\cdot}CO_2H$ side-chain (*n*-valeric acid side-chain).

The absorption spectrum of the quinoxaline derivative (formed from phenanthraquinone and the diaminocarboxylic acid) showed that it was a quinoxaline, I, and not a dihydroquinoxaline, II; thus the diaminocarboxylic could be III but not IV.



It therefore follows that the *n*-valeric acid side-chain cannot be attached to a carbon atom joined to an amino-group.

The nature of the sulphur atom in β-biotin was shown to be of the thioether type (*i.e.*, C—S—C) since:

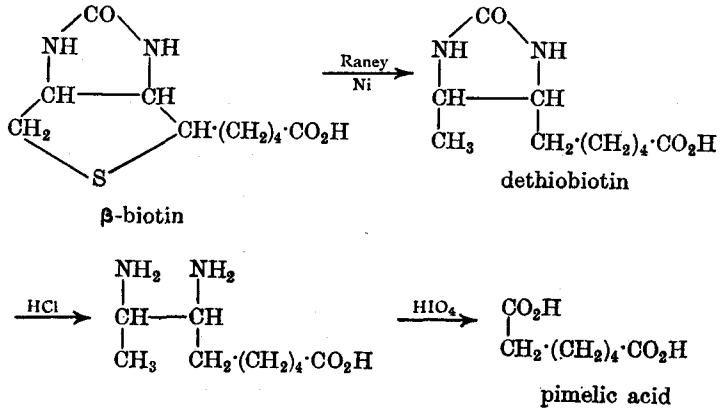
(i) Oxidation of β-biotin with hydrogen peroxide produced a sulphone.

(ii) When the methyl ester of β-biotin was treated with methyl iodide, a sulphonium iodide was formed.

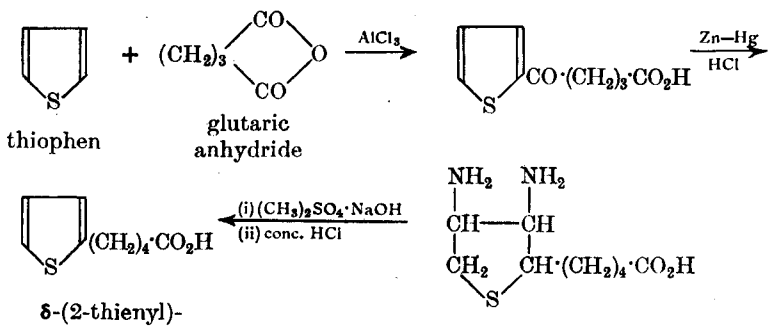
As we have seen, β-biotin does not contain a double bond; hence, from its molecular formula, it was deduced that β-biotin contains two rings (du Vigneaud *et al.*, 1941; Kögl *et al.*, 1941). The sort of argument that may be used is as follows. The molecular formula of β-biotin is $C_{10}H_{16}O_3N_2S$. The carboxyl group may be regarded as a substituent group, and so the parent compound will be $C_9H_{16}ON_2S$. Also, since two NH groups are present, these may be replaced by CH_2 groups; thus the parent compound is $C_{11}H_{18}OS$. The CO group may be replaced by a CH_2 group and the sulphide atom also by a CH_2 group. This gives a compound of formula $C_{13}H_{22}$ which has the same "structure" as β-biotin. Now the formula $C_{13}H_{22}$ corresponds to the general formula C_nH_{2n-2}, and this, for a *saturated* compound, corresponds to a system containing two rings.

When heated with Raney nickel, β-biotin formed *dethiobiotin* by elimination of the sulphur atom (this is an example of the *Mozingo reaction*, 1943).

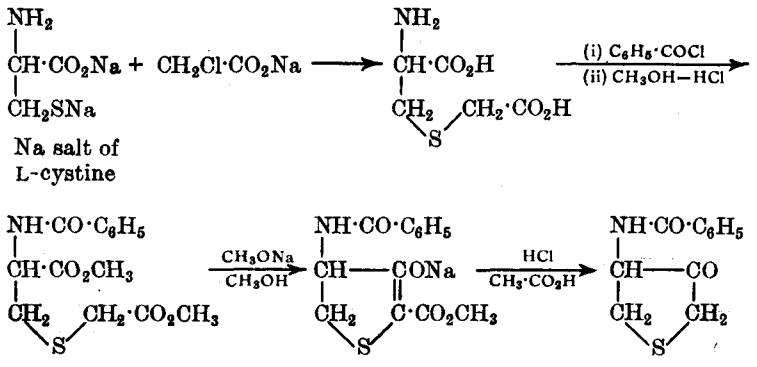
Dethiobiotin, on hydrolysis with hydrochloric acid, gave a diaminocarboxylic acid which, on oxidation with periodic acid, gave pimelic acid (du Vigneaud *et al.*, 1942). These results can be explained by assuming that the sulphur atom is in a five-membered ring and the *n*-valeric acid side-chain is in the position shown.

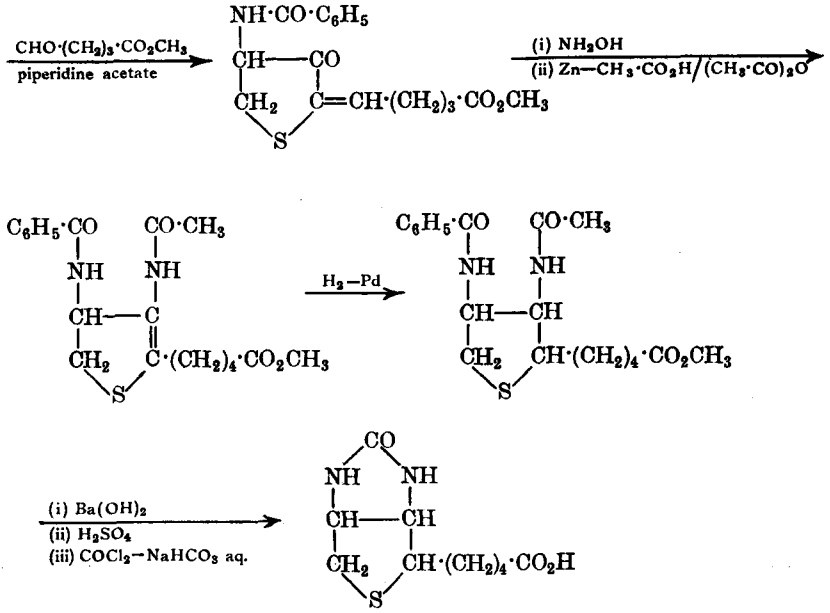


Further evidence for this structure is given by the fact that the exhaustive methylation of the diaminocarboxylic acid (produced from β-biotin), followed by hydrolysis, gave δ-(2-thienyl)-valeric acid (du Vigneaud *et al.*, 1942); the structure of this compound was confirmed by synthesis.



The above structure for β-biotin has been confirmed by synthesis (Harris *et al.*, 1943, 1944).





Two racemates were isolated, one of which was (±)-β-biotin; this was resolved *via* its esters with (−)-mandelic acid.

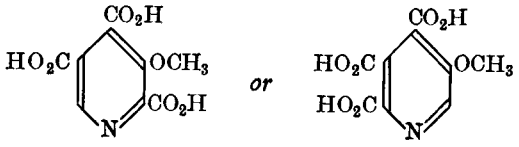
Examination of the β-biotin formula shows the presence of three asymmetric carbon atoms; the rings are fused in the *cis*-position in β-biotin and the orientation of the side-chain is also *cis*, as shown by X-ray analysis (Traub, 1956).

The structure of α-biotin is uncertain.

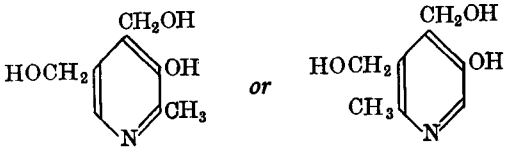
§10. Pyridoxin (Adermin, vitamin B_6), $C_8H_{11}O_3N$, is obtained from rice bran and yeast; it cures dermatitis in rats. Pyridoxin behaves as a weak base, and the usual tests showed the absence of methoxyl and methylamino-groups. Application of the Zerewitinoff method showed the presence of three active hydrogen atoms. When treated with diazomethane, pyridoxin formed a monomethyl ether which, on acetylation, gave a diacetyl derivative (Kuhn *et al.*, 1938). It therefore appears that the three oxygen atoms in pyridoxin are present as hydroxyl groups, and since one is readily methylated, this one is probably phenolic. This conclusion is supported by the fact that pyridoxin gives the ferric chloride colour reaction of phenols. Thus the other two hydroxyl groups are alcoholic.

Examination of the ultraviolet absorption spectrum of pyridoxin showed that it is similar to that of 3-hydroxypyridine. It was therefore inferred that pyridoxin is a pyridine derivative with the phenolic group in position 3. Since lead tetra-acetate has no action on the monomethyl ether of pyridoxin, this leads to the conclusion that the two alcoholic groups are not on adjacent carbon atoms in a side-chain (Kuhn *et al.*, 1939). When this methyl ether is *very carefully* oxidised with alkaline potassium permanganate, the product is a methoxypyridinetricarboxylic acid, $C_9H_7O_7N$. This acid gave a blood-red colour with ferrous sulphate, a reaction which is characteristic of pyridine-2-carboxylic acid; thus one of the three carboxyl groups is in the 2-position. When the methyl ether of pyridoxin was oxidised with alkaline permanganate under the usual conditions, the products were carbon dioxide and the anhydride of a dicarboxylic acid, $C_8H_5O_4N$; thus these two carboxyl groups are in the *ortho*-position. Furthermore, since this anhydride, on

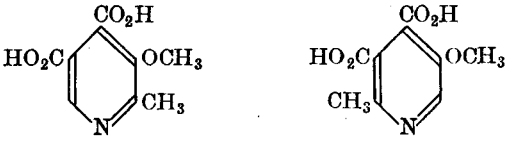
hydrolysis to its corresponding acid, did not give a red colour with ferrous sulphate, there is no carboxyl group in the 2-position. It therefore follows that, on decarboxylation, the tricarboxylic acid eliminates the 2-carboxyl group to form the anhydride; thus the tricarboxylic acid could have either of the following structures.



Now pyridoxin methyl ether contains three oxygen atoms (one as methoxyl and the other two alcoholic); it is therefore possible that two carboxyl groups in the tricarboxylic acid could arise from two CH_2OH groups, and the third from a methyl group, *i.e.*, pyridoxin could be either of the following:

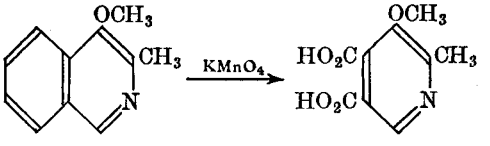


A decision between the two structures was made on the following evidence. When pyridoxin methyl ether was oxidised with barium permanganate, the product was a dicarboxylic acid, $C_9H_9O_5N$, which did not give a red colour with ferrous sulphate; thus there is no carboxyl group in the 2-position. Also, since the dicarboxylic acid formed an anhydride and gave a phthalein on fusion with resorcinol, the two carboxyl groups must be in the *ortho*-position. Furthermore, analysis of both the dicarboxylic acid and its anhydride showed the presence of a methyl group. Thus the structure of this dicarboxylic acid is either I or II.

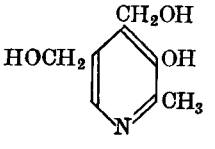


I II

Kuhn *et al.* (1939) showed that the anhydride was that of I from its formation by the oxidation of 4-methoxy-3-methyl-*iso*quinoline (a synthetic compound of known structure).

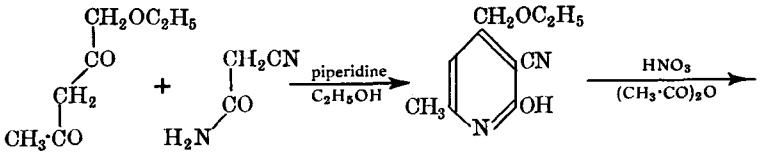


Hence, on the foregoing evidence, pyridoxin is

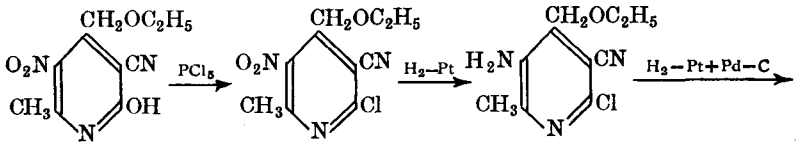


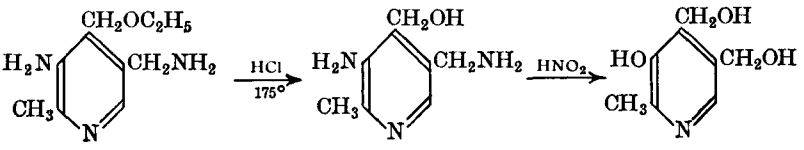
pyridoxin

This structure has been confirmed by synthesis, *e.g.*, that of Harris and Folkers (1939):



ethoxyacetyl-acetone cyano-acetamide

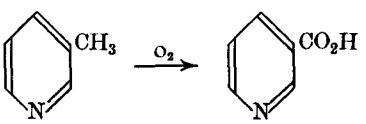




pyridoxin

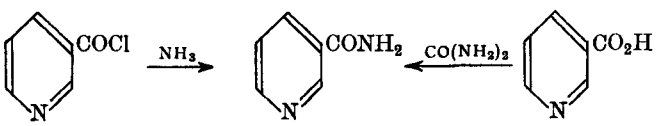
§11. Nicotinic acid and nicotinamide. These two compounds have been shown to be the human pellagra-preventing (P.P.) factor. Nicotinamide is part of the co-enzymes codehydrogenase I and II, which play a part in many biological oxidations.

Nicotinic acid (*Niacin*) was first prepared by the oxidation of nicotine (§21. XIV). This is now used as a commercial method; another commercial method for the preparation of nicotinic acid is the vapour-phase oxidation of 3-methylpyridine (β-picoline) in the presence of a vanadium and iron catalyst.



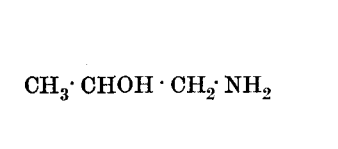
Still another commercial method is the oxidation of quinoline to quinolinic acid, which is then decarboxylated to nicotinic acid (see also §21. XIV).

Nicotinamide, m.p. 131°, is manufactured by various methods, *e.g.*, by the action of ammonia on nicotinyl chloride, or by heating nicotinic acid with urea in the presence of a molybdenum catalyst.

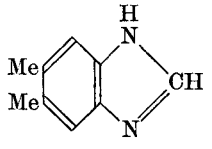


§12. Vitamin B_{12}, Cyanocobalamin. This is the anti-pernicious anæmia factor, and has been isolated from liver extract. Folic acid (§8) also has anti-anæmic properties. Vitamin B_{12} has been obtained as a red crystalline substance (Folkers *et al.*, 1948; Smith *et al.*, 1948, 1949), and the elements present have been shown to be C, H, O, N, P, Co; this vitamin is the first natural product found to contain cobalt. The cobalt has been shown to

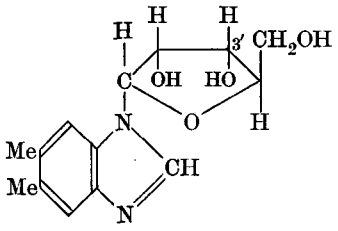
be attached to a cyano group. The hydrolysis of vitamin B_{12} with hydrochloric acid under different conditions produces ammonia, 1-aminopropan-2-ol (I), 5 : 6-dimethylbenzimidazole (II), 5 : 6-dimethylbenzimidazole-1-α-D-ribofuranoside (III) and the 3′-phosphate of III (Folkers *et al.*, 1949, 1950; Todd *et al.*, 1950). Compound IV (a succinimide derivative) has also been isolated by the chromic acid oxidation of hydrolysed vitamin B_{12} (Folkers, 1955).



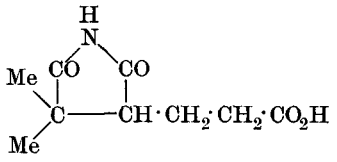
I



II

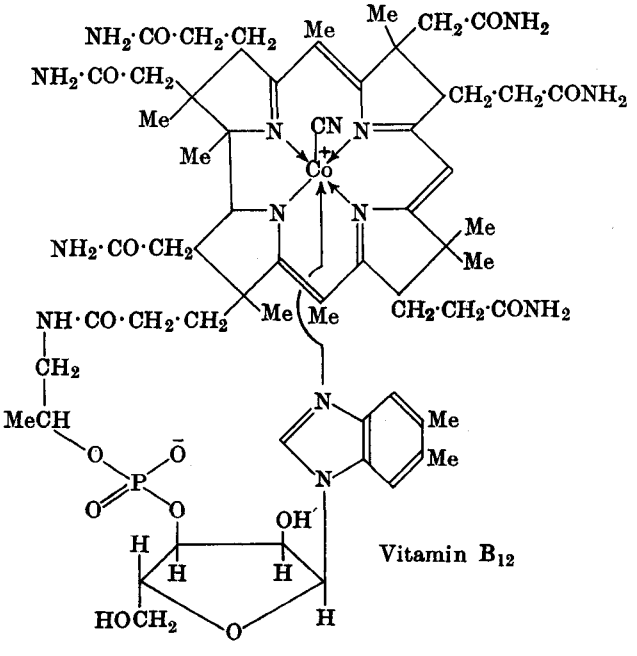


III



IV

Other work has shown that six amido groups are present in the molecule. Also, alkaline hydrolysis of vitamin B_{12} gives a mixture consisting mainly of a penta- and a hexacarboxylic acid, in both of which the nucleotide fragment is absent. As the result of a detailed X-ray analysis of the hexacarboxylic acid, vitamin B_{12} has been assigned the structure shown.



Vitamin B_{12}

A point of interest is that the arrangement of the four pyrrole nuclei is somewhat similar to that in the natural porphin derivatives such as hæm and chlorophyll (§§2, 7. XIX).

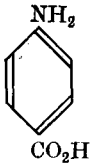
A number of vitamin B_{12} compounds have now been isolated which differ only in the nature of the basic component of the nucleotide. The remainder of the molecule, which is referred to as **Factor B,** is common to all the members of the vitamin B_{12} group. A partial synthesis of vitamin B_{12} (starting from factor B) has now been carried out by Bernhauer *et al.* (1960).

§13. Other compounds of the vitamin B complex. Three other compounds which have definitely been isolated from the vitamin B complex are:

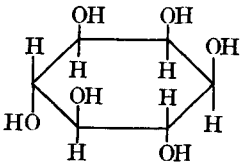
(i) *p*-Aminobenzoic acid; this is a growth factor for bacteria.

(ii) *meso*Inositol (m.p. 225–226°). This is a growth factor in animals, and its configuration has been elucidated by Posternak (1942; see also §11 iv. IV).

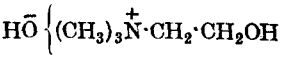
(iii) Choline. The absence of this compound leads to the formation of a fatty liver in animals.



p-amino-benzoic acid



*meso*inositol



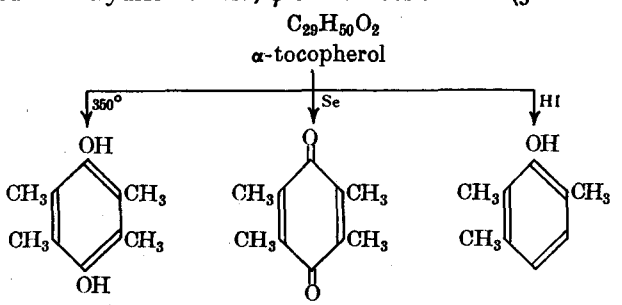
choline

Other vitamins of the vitamin B complex that have been said to exist are vitamins B_3, B_4, B_5, B_{10}, B_{11}, B_{13}, B_{14} and others.

VITAMIN E GROUP

§14. Introduction. Vitamin E is the anti-sterility factor; it occurs in seed germ oils. It is now known that there are three closely related compounds comprising " vitamin E "; all three are biologically active, and are known as α-, β- and γ-tocopherol. The main source of α- and β-tocopherol is wheat germ oil; the γ-compound is obtained from cotton seed oil. Wheat germ oil was first subjected to chromatographic analysis to remove sterols, etc., and then the α- and β-tocopherols were purified by conversion into their crystalline allophanates (see §12. XII) or 3 : 5-dinitrobenzoates. Hydrolysis of these derivatives gave the tocopherols as pale yellow oils.

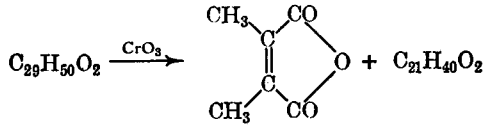
§15. α-Tocopherol, $C_{29}H_{50}O_2$. When α-tocopherol is heated at 350°, duroquinol is obtained (Fernholz, 1937). On the other hand, when heated with selenium, α-tocopherol forms duroquinone (McArthur *et al.*, 1937). Finally, when heated with hydriodic acid, ψ-cumenol is formed (John *et al.*, 1937).



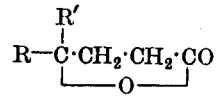
duroquinol duroquinone ψ-cumenol

The formation of these products led to the suggestion that α-tocopherol was the monoether of duroquinol; the possibility that it might be the diether was ruled out by the fact that α-tocopherol forms an allophanate, which indicates the presence of one free hydroxyl group. This monoether structure was shown to be incorrect by the fact that the ultraviolet absorption spectra of various monoethers of duroquinol were different from that of α-tocopherol (Fernholz, 1938).

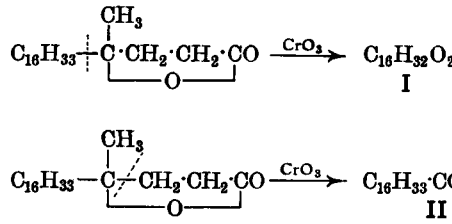
Oxidation of α-tocopherol with chromic acid forms dimethylmaleic anhydride and a compound $C_{21}H_{40}O_2$.



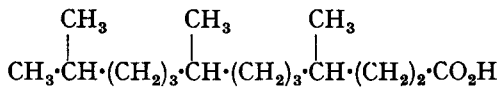
This latter compound was shown to be an optically active saturated lactone. This lactone was then shown to be derived from a γ-hydroxyacid in which the hydroxyl group is tertiary, *e.g.*, the acid lactonised immediately its salt was acidified, and also could not be oxidised to a keto-acid. Thus the structure of this lactone may be written (R + R′ = 17C):



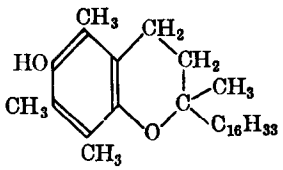
Now α-tocopherol acetate, on oxidation with chromic acid, forms an acid, $C_{16}H_{32}O_2$, I, and a ketone, $C_{18}H_{36}O$, II. Both of these compounds must be produced by the oxidation of the lactone at different points in the chain. Fernholz therefore suggested that if in the lactone R = $C_{16}H_{33}$ and R′ = CH_3, then the products I and II can be accounted for; thus:



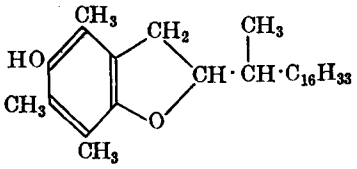
Fernholz then showed that the acid (I) contained methyl groups (*cf.* §3. IX), and was led to propose a structure based on the isoprene unit, *viz.*



The evidence obtained so far indicates the presence of a substituted benzene ring and a long side-chain in α-tocopherol. When the monoethers of duroquinol (see above) were oxidised with silver nitrate solution, the action took place far more slowly than for α-tocopherol when oxidised under the same conditions. Furthermore, whereas the former compounds were oxidised to duroquinone, the latter compound gave a red oil which appeared to have approximately the same molecular weight as α-tocopherol (Fernholz, 1938). Since duroquinone is not split off during this oxidation, it suggests that the side-chain is connected to the aromatic ring by a carbon bond as well as an ether link. In this case α-tocopherol is either a chroman or coumaran derivative:



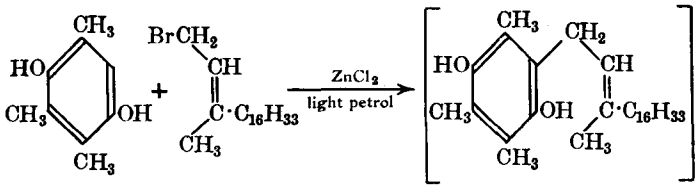
chroman structure

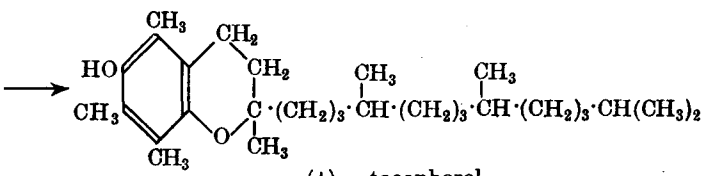


coumaran structure

According to Fernholz, the oxidation products are best explained on the chroman structure. This has been supported by ultraviolet absorption measurements of α-tocopherol (John *et al.*, 1938).

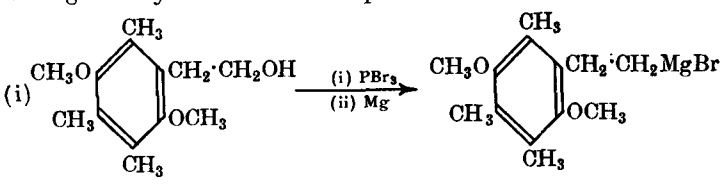
Karrer *et al.* (1938) have synthesised (±)-α-tocopherol by condensing trimethylquinol with phytyl bromide (§30. VIII).

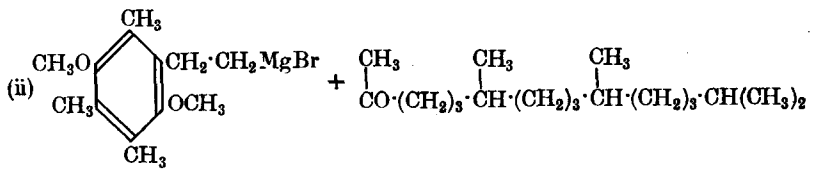


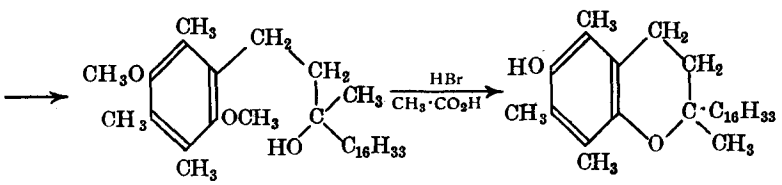


(±)-α-tocopherol

This synthesis, however, is not completely unambiguous, since phenols may condense with allyl compounds to form coumarans. Smith *et al.* (1939) have shown that $\gamma:\gamma$-disubstituted halides form only chromans, and since phytyl bromide is a halide of this type, this strengthens the course of the synthesis given above. Finally, Smith *et al.* (1942) have carried out an unambiguous synthesis of α-tocopherol as follows:



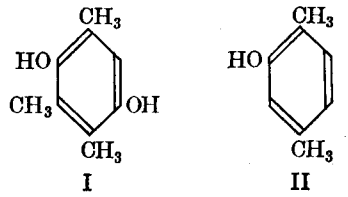




(±)-α-tocopherol

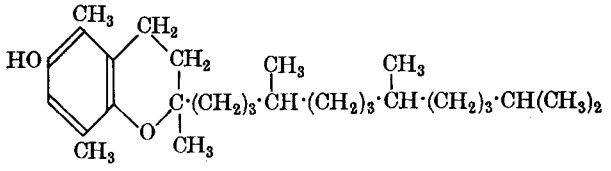
Smith *et al.* prepared the methyl ketone by ozonolysis of phytol, and also by oxidation of phytol with chromic acid.

§16. β-Tocopherol, $C_{28}H_{48}O_2$. This formula differs from that of α-tocopherol by CH_2. Thermal decomposition of β-tocopherol gives trimethylquinol, I, and heating with hydriodic acid *p*-xylenol, II (John *et al.*, 1937).



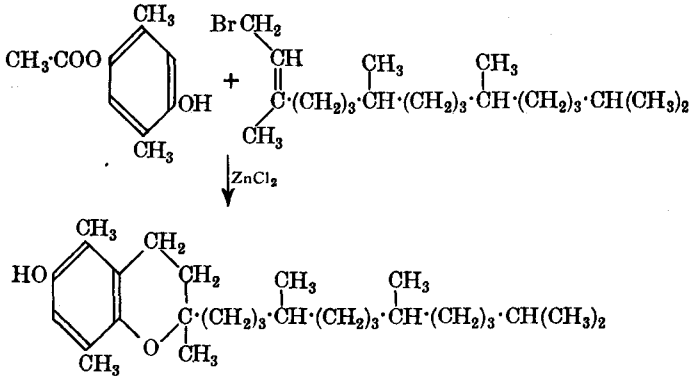
I II

When oxidised with chromic acid, β-tocopherol gives the same lactone ($C_{21}H_{40}O_2$) as that obtained from α-tocopherol. Thus the only difference between the two tocopherols is that the α-compound has one more methyl group in the benzene ring than the β-; hence the latter is

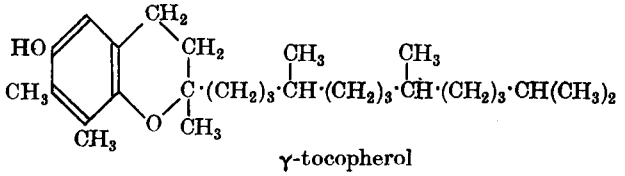


β-tocopherol

This has been confirmed by synthesis, starting from the monoacetate of *p*-xyloquinol and phytyl bromide.

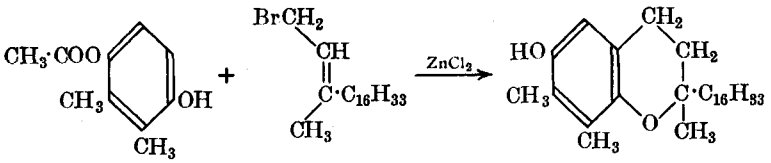


§17. γ-Tocopherol, $C_{28}H_{48}O_2$. This is isomeric with β-tocopherol; the only difference is the positions of the two methyl groups in the benzene ring, *e.g.*, when heated with hydriodic acid, γ-tocopherol gives *o*-xyloquinol. Thus γ-tocopherol is

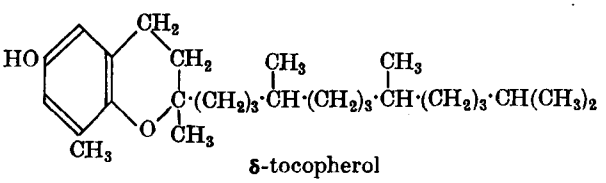


γ-tocopherol

This structure has been confirmed by synthesis, starting from the monoacetate of *o*-xyloquinol and phytyl bromide.



§18. δ-Tocopherol, $C_{27}H_{46}O_2$. This was isolated from soya bean oil by Stern *et al.* (1947); it is a yellow oil, and is inactive physiologically. The structure of δ-tocopherol is

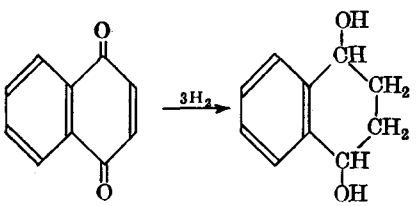


δ-tocopherol

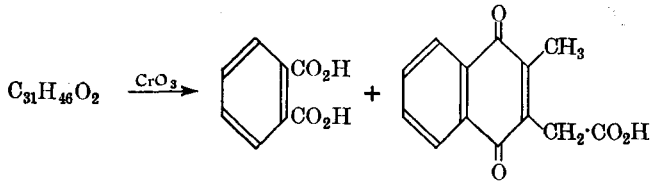
VITAMIN K GROUP

§19. Introduction. Dam *et al.* (1939) and Doisy *et al.* (1939) isolated vitamin K from alfalfa, and called it vitamin K_1 to distinguish it from a substance called vitamin K_2 which had been isolated from putrefied fish meal by Doisy *et al.* (1939). Both are antihæmorrhagic vitamins; they are connected with the enzymes involved in blood clotting, a deficiency of them lengthening the time of blood clotting. Kegel *et al.* (1962) have obtained chemical evidence for the presence of vitamin K_1 in extracts from spinach chloroplasts.

§20. Vitamin K_1 (α-phylloquinone), $C_{31}H_{46}O_2$, is a light yellow oil. The redox potential of vitamin K_1 is very similar to that of 1 : 4-quinones (Karrer *et al.*, 1939), and its absorption spectrum is very similar to that of 2 : 3-disubstituted 1 : 4-naphthaquinones (McKee *et al.*, 1939). Thus vitamin K_1 appears to be a 1 : 4-naphthaquinone derivative, and this is in keeping with the fact that the vitamin is very sensitive to light and to alkalis. Now the catalytic hydrogenation of vitamin K_1 causes the addition of four molecules of hydrogen (McKee *et al.*, 1939); the product is a colourless compound. Since it is known that three molecules of hydrogen are added when 1 : 4-naphthaquinone is reduced under these conditions, the addition of a fourth molecule of hydrogen to the vitamin suggests the presence of an ethylenic double bond in a side-chain.

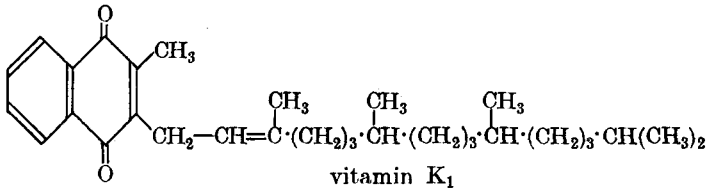


When subjected to reductive acetylation (*i.e.*, acetylated under reducing conditions), vitamin K_1 is converted into the diacetate of dihydrovitamin K_1 (Binkley *et al.*, 1939). This diacetate is difficult to hydrolyse; this is a property characteristic of 2 : 3-disubstituted 1 : 4-naphthaquinones. When oxidised with chromic acid, vitamin K_1 gives phthalic acid, but when the oxidation is carried out under controlled conditions, the product is a compound with the molecular formula $C_{13}H_{10}O_4$. This latter compound was subsequently shown to be 2-methyl-1 : 4-naphthaquinone-3-acetic acid (Binkley *et al.*, 1939).



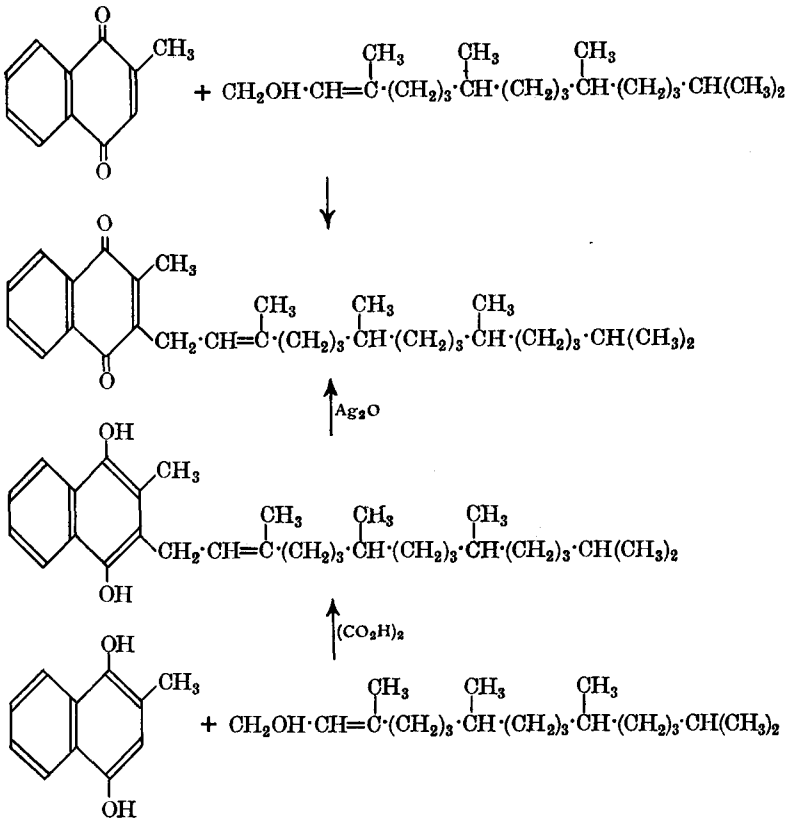
Thus the presence of the 1 : 4-naphthaquinone structure is confirmed, and at the same time these products show that one ring is unsubstituted and that the other (the quinonoid ring) has substituents in the 2- and 3-positions.

When the diacetate of dihydrovitamin K_1 (see above) was subjected to ozonolysis, a compound $C_{18}H_{36}O$ was obtained, which was then shown to be identical with the ketone produced by the oxidation of phytol (McKee *et al.*, 1939; *cf.* Smith's synthesis of α-tocopherol, §15). Hence, on the evidence obtained above, vitamin K_1 is 2-methyl-3-phytyl-1 : 4-naphthaquinone.



vitamin K_1

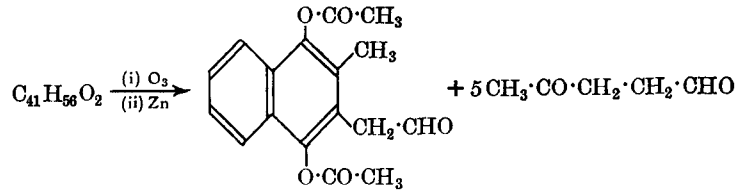
This structure has been confirmed by synthesis: Almquist *et al.* (1939) obtained vitamin K_1 by condensing 2-methyl-1 : 4-naphthaquinone with phytol; Fieser *et al.* (1939) obtained a better yield by heating 2-methyl-1 : 4-naphthaquinol with phytol in dioxan solution in the presence of anhydrous oxalic acid, and then oxidising the product, dihydrovitamin K_1, with silver oxide in ether.



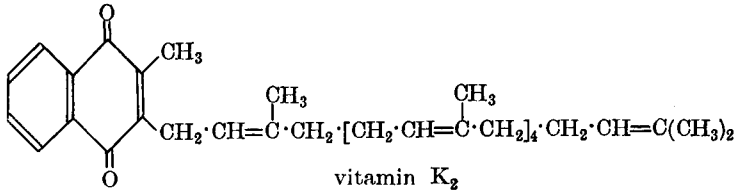
Wendler *et al.* (1954) have obtained vitamin K_1 in good yield by condensing the 1-acetyl derivative of 2-methyl-1 : 4-naphthaquinol with phytol in the presence of boron trifluoride.

§21. Vitamin K_2, $C_{41}H_{56}O_2$, is a yellow solid, m.p. 54°; it is less potent than vitamin K_1. It was shown to contain a 1 : 4-naphthaquinone nucleus by the facts that it is sensitive to light and to alkalis, and that it has an absorption spectrum similar to that of vitamin K_1 (McKee *et al.*, 1939). When catalytically reduced, vitamin K_2 adds on nine molecules of hydrogen, and since three of these are absorbed by the naphthaquinone nucleus (see §20), it therefore suggests that there is a side-chain present which contains six double bonds. Furthermore, since vitamin K_2 does not form an adduct with maleic anhydride, no conjugation is present (McKee *et al.*, 1939). That these six double bonds are ethylenic is shown by the fact that on reductive acetylation, vitamin K_2 forms the diacetate of dihydrovitamin K_2, which can add on six molecules of bromine.

The oxidation of vitamin K_2 with permanganate produces phthalic acid; therefore one ring is unsubstituted. On the other hand, when ozone is passed into a solution of vitamin K_2 in acetic acid, and the product then treated with zinc dust in ether, 1 : 4-diacetoxy-2-methylnaphthalene-3-acetaldehyde is produced. At the same time there is obtained lævulaldehyde in a yield of 93 per cent. calculated on the basis that one molecule of vitamin K_2 can produce five molecules of the aldehyde.



Acetone is also formed in this reaction, and is obtained in a yield of 56 per cent. based on the assumption that one molecule of acetone is produced from one molecule of vitamin K_2 (McKee *et al.*, 1940). On this evidence, it has been suggested that vitamin K_2 is 3-difarnesyl-2-methyl-1 : 4-naphthaquinone (Binkley *et al.*, 1940).



vitamin K_2

§22. Other compounds possessing antihæmorrhagic properties. It has been shown that simple 1 : 4-naphthaquinones have blood-clotting properties. 2-Methyl-1 : 4-naphthaquinone is more active than either vitamin K_1 or K_2 (Fernholz *et al.*, 1939); it is therefore used instead of the natural vitamins. *Phthiocol* (3-hydroxy-2-methyl-1 : 4-naphthaquinone) is also an active compound, and is water-soluble. It is also interesting to note that many quinones other than 1 : 4-naphthaquinones have also been found to be active, *e.g.*, some *p*-benzoquinones.

READING REFERENCES

Vitamins, A Survey of Recent Knowledge, Medical Research Council Report (1932).
Rosenberg, *Chemistry and Physiology of the Vitamins,* Interscience Publishers (1942).
Vitamins and Hormones, Academic Press (1943–).
Stewart and Graham, *Recent Advances in Organic Chemistry,* Longmans, Green. Vol. III (1948, 7th ed.). Ch. 2. Vitamins.
Robinson, *The Vitamin B Complex,* Chapman and Hall (1951).
Harris, *Vitamins in Theory and Practice,* Cambridge University Press (1955, 4th ed.).
Rodd (Ed.), *Chemistry of Carbon Compounds,* Elsevier. Vol. IVC (1960). Ch. XXII. Pteridines, Alloxazines, Flavins.
Structure of Vitamin B_{12}.
(i) Dorothy Crowfoot Hodgkin *et al.*, *Nature,* **1955, 176,** 325; **1956, 178,** 64.
(ii) Todd *et al.*, *Nature,* **1955, 176,** 330.
(iii) *Chem. Soc. Special Publ.*, No. 3, 1955.

CHAPTER XVIII

CHEMOTHERAPY

§1. Introduction. The term *chemotherapy* was introduced by Ehrlich (1909), and it now appears to be used in the sense of the treatment of diseases due to bacterial invasion by chemical compounds which destroy the micro-organisms without affecting, to any material extent, the tissues (of the host). Many compounds, *e.g.*, formaldehyde, phenol, iodine, etc., are also active in destroying bacteria. These compounds, however, are applied *externally*, and tend to destroy the tissues; thus they are not included under the heading of therapeutic agents, but are known as *disinfectants*.

The first compounds to be used by Ehrlich (1907) were organic dyes. From then onwards, organic compounds of diverse chemical structures have been used in chemotherapy. It has now been found that a given compound is specific in its toxicity towards a particular micro-organism. The relationship between chemical structure and chemotherapeutic action is extremely complicated, but some progress has been made in this field.

Compounds which exert various physiological effects of therapeutic value are collectively known as *drugs*. The ideal requirement of a drug is that, on administration (to the host), it should be localised at the site where it is required. In practice, however, no drug behaves in this way, but tends to distribute itself anywhere in the tissues of the host. Another difficulty is that cells, which were originally susceptible to a particular drug, may acquire a tolerance (resistance) to that drug. In some cases it has been found that the drug actually reverses its original action, *i.e.*, it stimulates the cell instead of inhibiting it.

There have been three approaches to the problem of finding a drug to combat a particular disease:

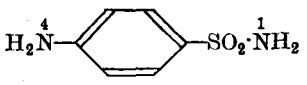
(i) The method of trial and error. This involves the trial of all kinds of compounds, natural and synthetic.

(ii) The method requiring a knowledge of the cell system, and then synthesising compounds which interfere with it.

(iii) The method in which one starts with a compound known to have some of the required activity (this information has been gained from the previous methods), and then to vary the structure of the molecule systematically. This method has, so far, proved to be the most fruitful.

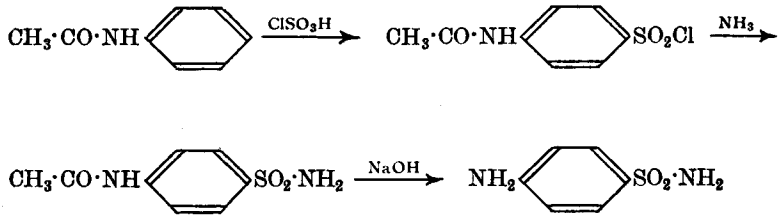
§2. Sulphonamides. Sulphanilamide (*p*-aminobenzenesulphonamide) and its derivatives have great antibacterial powers; sulphanilamide itself is widely used in medicine against " cocci infections "—streptococci, gonococci and pneumococci. Research in the sulphonamide field was stimulated by the discovery of Domagk (1934) that prontosil (see below) had a curative effect when injected into mice infected with streptococci.

The system of numbering is as follows: substituents of the amide group of sulphanilamide are called N^1-substituents, and substituents of the amino-group are called N^4-substituents.

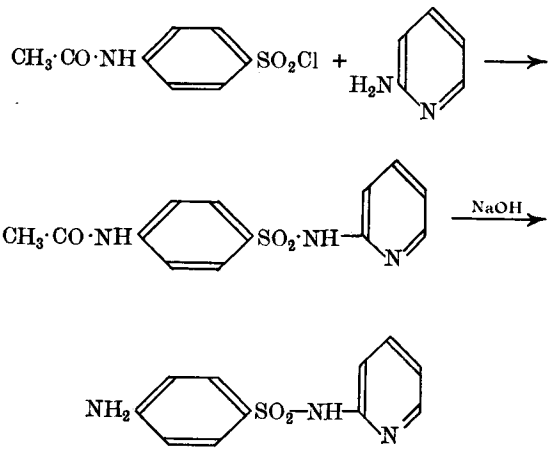


sulphanilamide

Sulphanilamide may be prepared from acetanilide:

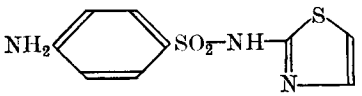


Sulphapyridine (N^1-2-pyridylsulphanilamide) was the first drug to effect cures of pneumonia; it is more potent than sulphanilamide. It may be prepared as follows:



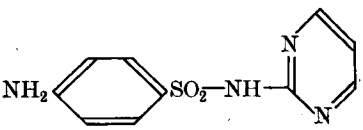
This compound was introduced under the trade name of *M and B* 693.

Sulphathiazole (N^1-2-thiazolylsulphanilamide) is more potent than *Sulphapyridine* and less toxic; it is used mainly in severe infections. It is



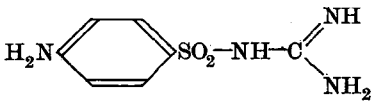
prepared in the same way as *Sulphapyridine* except that 2-aminothiazole is used instead of 2-aminopyridine.

Sulphadiazine (N^1-2-pyrimidylsulphanilamide; *Sulphapyrimidine*) is less toxic than *Sulphathiazole*; it is the most widely used of the " sulpha " drugs, its main use being for mild infections. It is prepared in the same way as the previous compound, except that 2-aminopyrimidine is used in this case.

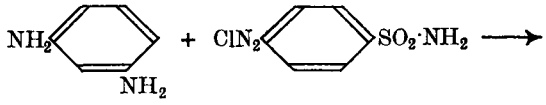


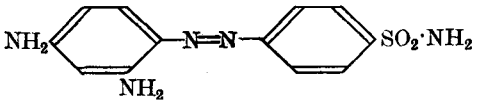
Sulphamezathine (N'-2(4 : 6-dimethylpyrimidyl)-sulphanilamide) is also used for general purposes.

Sulphaguanidine, since it is only slightly absorbed in the intestinal tract, can therefore be be given in relatively large doses in the treatment of bacillary dysentery.



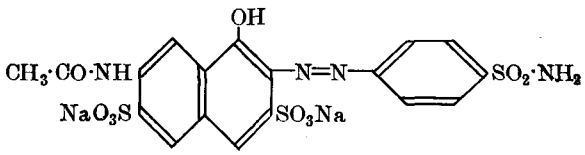
Prontosil (4-sulphonamido-2′ : 4′-diaminoazobenzene) was the first sulphonamide to be used in medicine. It is prepared by diazotising sulphanilamide and then coupling with *m*-phenylenediamine.



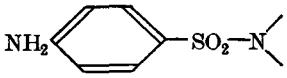


It was suggested that *Prontosil* broke down in the body to sulphanilamide; this led to the discovery that the latter compound is very active against bacteria.

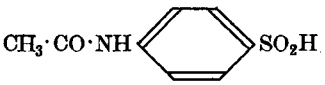
Prontosil S is more soluble than *Prontosil.*



Mechanism of action of the sulphonamides. It appears that the antibacterial activity of the sulphonamides is associated with the group



Some compounds containing slight variations from this structure are also active, *e.g.,*

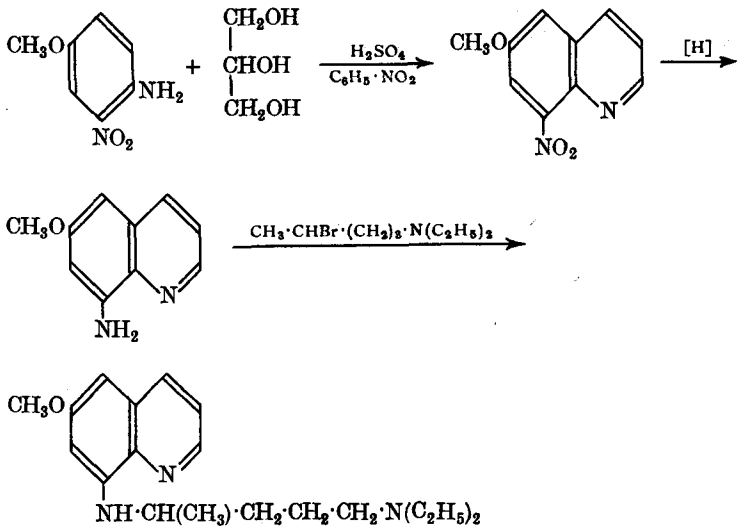


Compounds in which the amino-group is *ortho* or *meta* to the sulphonamido-group are either less active or completely inactive.

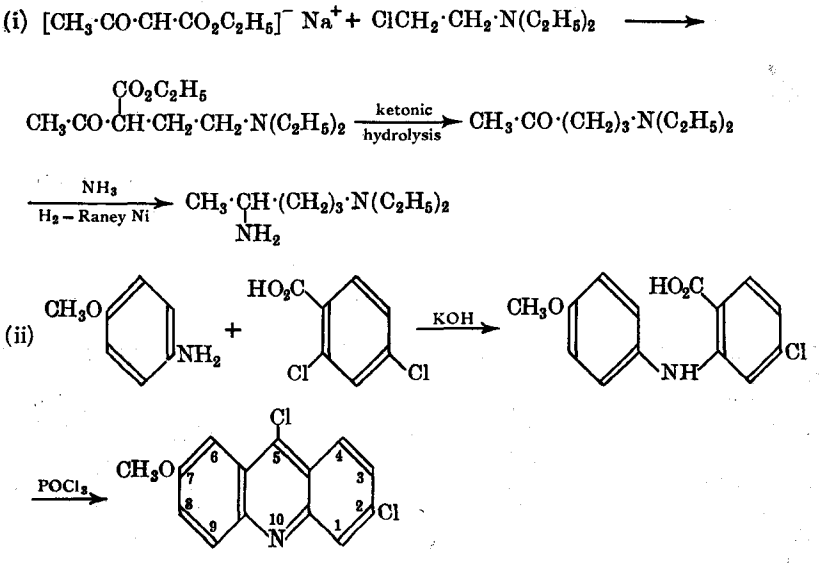
p-Aminobenzoic acid is an essential growth factor for most bacteria susceptible to the sulphonamides. The theory of action is that, owing to the similarity in structure, bacteria absorb a sulphonamide " by mistake ", and once this compound is ingested, the bacteria cease to grow in numbers (Woods, 1940). Thus the sulphonamides are not bactericidal but bacteriostatic.

§3. Antimalarials. Quinine (§25b. XIV) was originally the only drug known to be effective against malaria. Now there is a number of synthetic compounds used for this purpose, *e.g.*, *Plasmoquin, Mepacrine, Proguanil.*

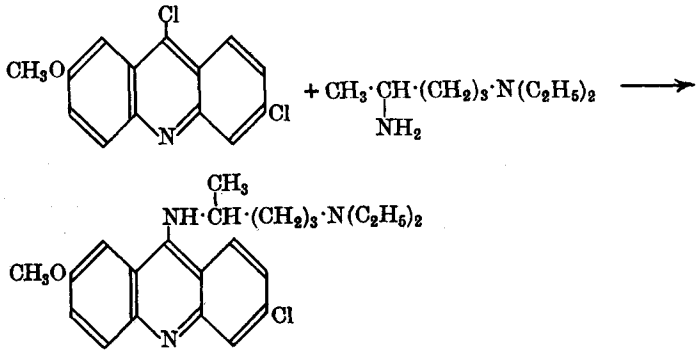
Plasmoquin (*Pamaquin*) is 8-(4′-diethylamino-1′-methylbutylamino)-6-methoxyquinoline. One preparation that has been described for this compound is the condensation between 4-bromo-1-diethylaminopentane and 8-amino-6-methoxyquinoline, the latter being prepared from 4-amino-3-nitroanisole by means of the Skraup synthesis (see Vol. I).



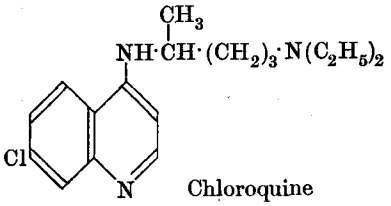
Mepacrine (*Atebrin, Quinacrine*) is 2-chloro-5-(4′-diethylamino-1′-methylbutylamino)-7-methoxyacridine. It is better than quinine, and it has been prepared as follows:



(iii)

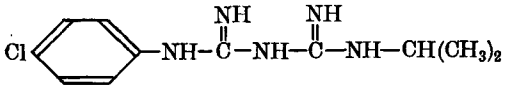


Mepacrine has certain unpleasant side-effects (such as producing a yellow colour in the skin, nausea, etc.), and a drug superior to both quinine and *Mepacrine* is *Chloroquine* (*Aralen*).



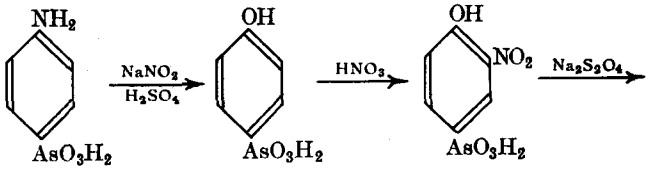
Chloroquine

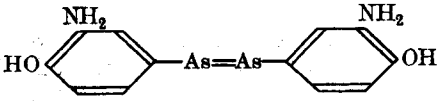
Proguanil (*Paludrine*) is N^1-*p*-chlorophenyl-N^5-*iso*propyldiguanide. It is superior to *Mepacrine* and *Chloroquine*, and appears to be the best antimalarial known at the present time.



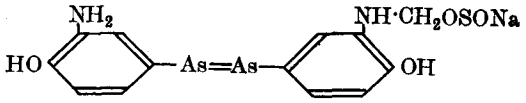
§4. Arsenical drugs. A particularly important use of arsenical drugs is in the treatment of syphilis.

Arsphenamine (*Salvarsan*, "606") was first introduced by Ehrlich (1909); it is 3 : 3′-diamino-4 : 4′-dihydroxyarsenobenzene, and may be prepared as follows:



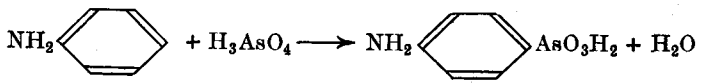


Arsphenamine is an unstable compound; it is stable as its dihydrochloride which, however, cannot be used as such but must be converted into the soluble sodium salt. Ehrlich (1912) overcame this difficulty by preparing **neoarsphenamine** (*Neosalvarsan*), a soluble compound, which may be produced by condensing arsphenamine with sodium formaldehydesulphoxylate, $CH_2OH{\cdot}SO_2Na$.

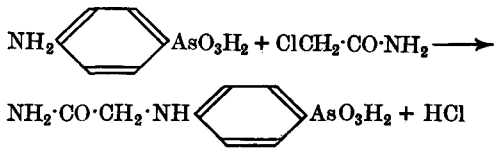


neoarsphenamine

Atoxyl is the sodium salt of *p*-arsanilic acid (*p*-aminophenylarsonic acid); it is used in the treatment of sleeping sickness. *p*-Arsanilic acid may be prepared by heating aniline with arsenic acid at 200° (*cf.* sulphanilic acid, Vol. I).



Tryparsamide is the sodium salt of *N*-phenylglycineamide-*p*-arsonic acid; it is less toxic than *Atoxyl*, and may be prepared by refluxing the latter with chloroacetamide.



§5. **Antibiotics.** Many micro-organisms produce within themselves chemical substances which, when excreted, interfere with the growth or metabolism of other micro-organisms. Such compounds are known as *antibiotics*, and need be present only in low concentration to bring about this antibiotic action. Antibiotics are thus chemotherapeutic agents.

In 1929, Fleming discovered a mould of the *Penicillium* species which inhibited the growth of certain bacteria. This observation was investigated later by a number of workers and culminated in the isolation of the active principle *penicillin*. At the same time, research along this line led to the isolation of many other antibiotics.

§6. **The penicillins.** Penicillin is the name given to the mixture of natural compounds having the molecular formula $C_9H_{11}O_4N_2SR$, and differing only in the nature of R. There are at least five natural penicillins.

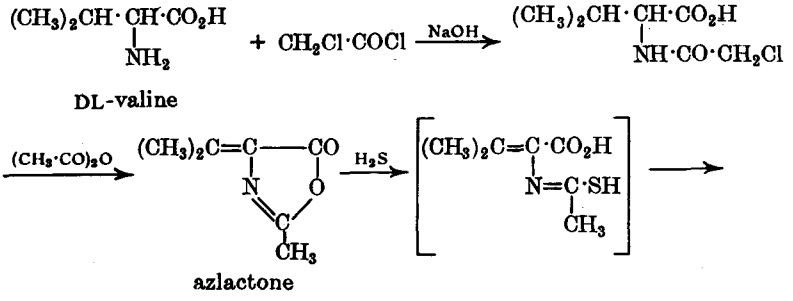
Chemical Name	Other Names	R
Pent-2-enylpenicillin . .	Penicillin-I or F	$-CH_2{\cdot}CH{=}CH{\cdot}CH_2{\cdot}CH_3$
Benzylpenicillin	Penicillin-II or G	$-CH_2{\cdot}C_6H_5$
p-Hydroxybenzylpenicillin .	Penicillin-III or X	$-CH_2{\cdot}C_6H_4{\cdot}OH(1:4)$
n-Heptylpenicillin . . .	Penicillin-K	$-(CH_2)_6{\cdot}CH_3$
n-Amylpenicillin	Dihydro-F-penicillin	$-(CH_2)_4{\cdot}CH_3$

Commercial preparations of penicillin contain one or more of the penicillins in varying proportions. It has been found that the addition to the culture medium of various compounds containing a benzyl group, *e.g.*, phenylacetic acid, phenylacetamide, etc., increases the total yield of penicillin, and also the proportion of benzylpenicillin. Similarly, the addition of compounds containing the *p*-hydroxybenzyl group to the culture medium increases the proportion of *p*-hydroxybenzylpenicillin. On the other hand, by adding various compounds to the culture medium, a number of " unnatural " penicillins have been prepared (see §6b).

§6a. Structure of the penicillins. The penicillins are all strong monobasic acids, *e.g.*, they form salts. The penicillins are hydrolysed by hot dilute inorganic acids; one carbon atom is eliminated as carbon dioxide, and two products are obtained in equimolecular amounts, one being an amine, *penicillamine*, and the other an aldehyde, *penilloaldehyde*. All the penicillins give the same amine, but different aldehydes; it is the latter which contain the R group.

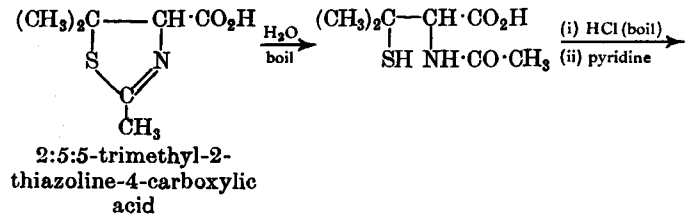
$$C_9H_{11}O_4N_2SR + 2H_2O \xrightarrow{HCl} CO_2 + C_5H_{11}O_2NS + C_3H_4O_2NR$$

D-Penicillamine, $C_5H_{11}O_2NS$. Since penicillamine gives the indigo colour reaction with ferric chloride, a test characteristic of cysteine, this suggests that the amine is probably a substituted cysteine. The structure of penicillamine was proved to be D-β : β-dimethylcysteine by synthesis, *e.g.*,

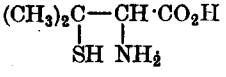


DL-valine

azlactone

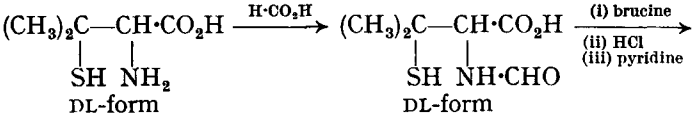


2:5:5-trimethyl-2-thiazoline-4-carboxylic acid

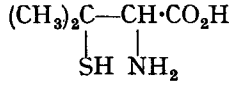


DL-penicillamine

The racemic amine was resolved as follows: the amine was converted into the formyl derivative, which was then resolved by means of brucine. D-Penicillamine was obtained after removal of the formyl group by hydrolysis.



DL-form DL-form



D-penicillamine

This was found to be identical with the natural penicillamine.

When treated with diazomethane, penicillin is converted into its methyl ester and this, on treatment with an aqueous solution of mercuric chloride, gives the methyl ester of penicillamine. Thus the carboxyl group in penicillamine is the carboxyl group in penicillin itself.

Penilloaldehyde. On vigorous hydrolysis, all the penilloaldehydes give a substituted acetic acid and aminoacetaldehyde. Thus the penilloaldehydes are acylated derivatives of aminoacetaldehyde.

$$R{\cdot}CO{\cdot}NH{\cdot}CH_2{\cdot}CHO + H_2O \longrightarrow R{\cdot}CO_2H + NH_2{\cdot}CH_2{\cdot}CHO$$

This structure has been confirmed by synthesis:

$$R{\cdot}COCl + NH_2{\cdot}CH_2{\cdot}CH(OC_2H_5)_2 \longrightarrow$$

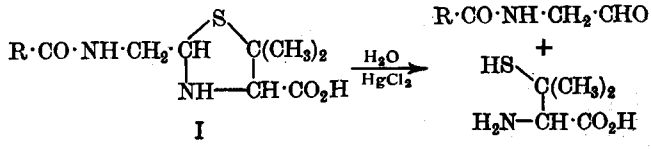
$$R{\cdot}CO{\cdot}NH{\cdot}CH_2{\cdot}CH(OC_2H_5)_2 \xrightarrow{HCl} R{\cdot}CO{\cdot}NH{\cdot}CH_2{\cdot}CHO$$

As pointed out above, the acid hydrolysis of penicillin gives penicillamine, penilloaldehyde and carbon dioxide. The formation of this molecule of carbon dioxide gave rise to the belief that it is formed by the ready decarboxylation of an unstable acid. Such an acid is a β-keto-acid, and so a possible explanation is that penilloaldehyde-carboxylic acid (penaldic acid) is formed as an intermediate in the hydrolysis of penicillin (see also below):

$$R{\cdot}CO{\cdot}NH{\cdot}\underset{\displaystyle CO_2H}{\underset{|}{CH}}{\cdot}CHO \longrightarrow CO_2 + R{\cdot}CO{\cdot}NH{\cdot}CH_2{\cdot}CHO$$

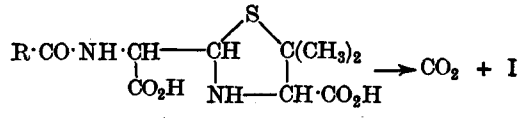
penaldic acid

The problem now is: How are the two fragments, penicillamine and penilloaldehyde, combined in penicillin? The hydrolysis of penicillin with dilute alkali or with the enzyme penicillinase produces *penicilloic acid* (a dicarboxylic acid), which readily eliminates a molecule of carbon dioxide to form *penilloic acid*. This suggests that a carboxyl group is in the β-position with respect to a negative group (*cf.* above). Penilloic acid, on hydrolysis with aqueous mercuric chloride, gives penicillamine and penilloaldehyde. This hydrolysis is characteristic of compounds containing a thiazolidine ring (*cf.* §5b. XII). Thus penilloic acid could be I, since this structure would give the required products.



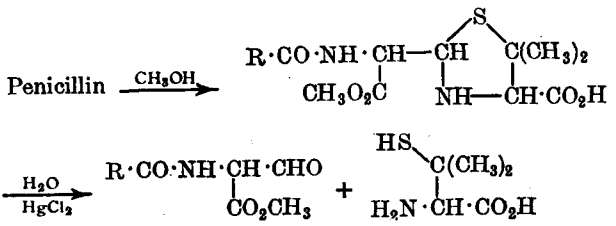
I

Hence, if I is penilloic acid, then penicilloic acid would be II.

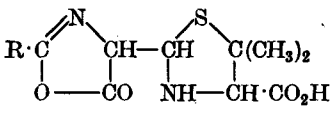


II

This structure (II) is supported by the fact that the treatment of penicillin with methanol gives methyl penicilloate which, on hydrolysis with aqueous mercuric chloride, gives methyl penaldate (see also above) and penicillamine.

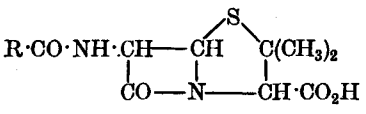


On the basis of the foregoing evidence, two structures are possible for penicillin, *viz.* III and IV.



III

oxazolone structure

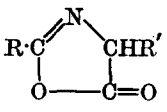


IV

β-lactam structure

It was not possible to decide between these two on chemical evidence alone, since penicillin readily undergoes molecular rearrangements, *e.g.*, on treatment with dilute acid, penicillin rearranges to penillic acid. It was therefore necessary to examine the molecule by physical methods (thereby leaving the molecule intact).

(i) Infra-red spectra studies showed the presence of two double bonds; these were exclusive of the C=O group in the carboxyl group in penicillin. The examination of the infra-red spectra of a number of oxazolones (these contain two double bonds, C=O and C=N) showed that this ring structure could not account for the absorption maxima obtained for penicillin. Thus structure III is untenable. On the other hand, it was found from an

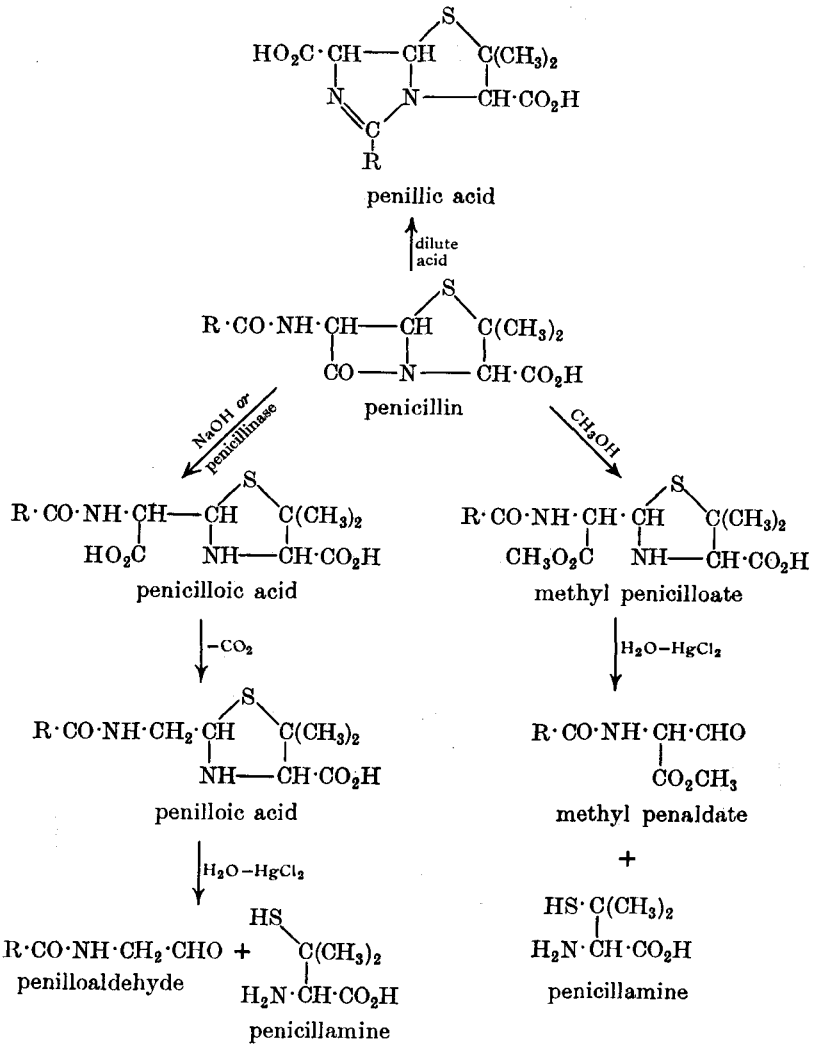


oxazolone

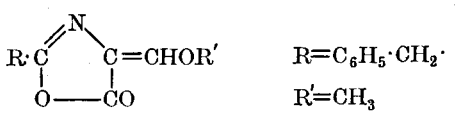
examination of the spectra of a number of amides that an amide structure could account for the spectrum of penicillin; thus IV is the probable structure of penicillin.

(ii) The X-ray analysis of the sodium, potassium and rubidium salts of benzylpenicillin showed the presence of a β-lactam ring; thus IV is the structure of penicillin.

Using this structure, we can now formulate the chemical reactions described above.



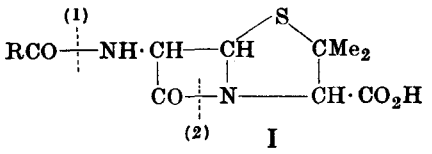
Penicillin has been synthesised by condensing synthetic D-penicillamine with a suitably substituted oxazolone containing a potential aldehyde group, *e.g.*,



§6b. " Synthetic " penicillins. It has been found that most strains of staphylococci are highly sensitive to penicillin, but after a time these strains become resistant. This result has been shown to be due to the fact that these resistant strains produce the enzyme penicillinase which converts penicillin into the inactive penicilloic acid (see §6a).

Of all the natural penicillins, benzylpenicillin (penicillin G) is still the best. It has been recently found that different types of penicillin are produced by *Penicillium chrysogenum* when the cultural conditions are changed.

Batchelor *et al.* (1959) isolated pure 6-aminopenicillanic acid from fermentation liquors to which *no* precursors had been added. This acid had already been synthesised by Sheehan (1958); it is the amino-compound (I) with the RCO group removed.

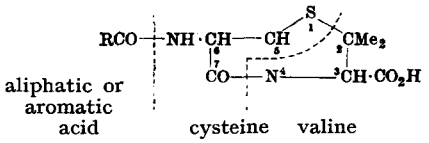


I

It has also been shown that (1) is the site of action of the enzyme penicillin amidase (Rolinson *et al.*, 1960; Claridge *et al.*, 1960) and, as mentioned above, (2) is the site of action of penicillinase.

Many " synthetic " penicillins have now been prepared (by the method described in §6). α-Aminobenzylpenicillin (Rolinson *et al.*, 1961) has been synthesised and shows considerable activity against many organisms against which benzylpenicillin is not very effective. 6-Aminopenicillanic acid itself has also been used as the starting point of many new penicillins either chemically or by means of amidases.

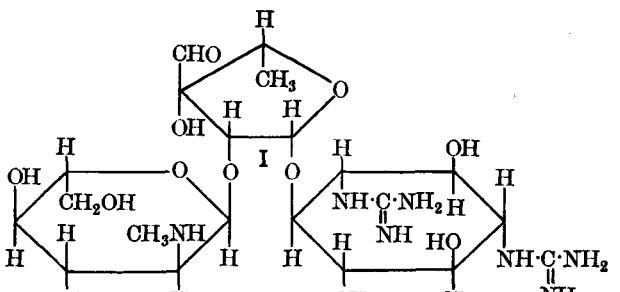
§6c. Biosynthesis of penicillins. This has been studied and much progress has been made; the structure of penicillin can be dissected into an acid, cysteine, and valine.



(*a*) *Side-chain precursors* (R·CO). Various aliphatic and aromatic acids have been used (see above and §6).

(*b*) *Precursors of the thiazolidine-β-lactam ring system.* The use of labelled compounds has shown that (i) L-cystine (or cysteine), and (ii) L-valine are precursors of penicillin. Bentley *et al.* (1961) have also shown that malonate functions as a part-precursor of penicillic acid.

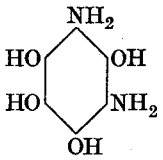
§7. Streptomycin. Streptomycin was isolated by Waksman *et al.* (1944) from cultures of *Streptomyces griseus*. This antibiotic is very effective in the treatment of tuberculosis, meningitis and pneumonia. Streptomycin is a solid with a lævorotation, and its structure has been shown to be composed of the three units streptose, I, *N*-methyl-L-glucosamine, II, and streptidine, III.



II III

The following is a very brief account of the evidence that led to this structure for streptomycin. The molecular formula was shown to be $C_{21}H_{39}O_{12}N_7$. Three nitrogen atoms are strongly basic (the molecule forms a trihydrochloride), and on mild acid hydrolysis, streptomycin gives one molecule of streptidine, $C_8H_{18}O_4N_6$, and one molecule of streptobiosamine, $C_{13}H_{23}O_9N$ (Folkers *et al.*, 1945).

Streptidine (unit III), on oxidation with potassium permanganate, gave two molecules of guanidine (Peck *et al.*, 1946); thus two guanido groups are present in streptidine. Streptidine, on alkaline hydrolysis, gave **streptamine** and ammonia (Brink *et al.*, 1945). Streptamine was shown to be

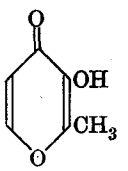


streptamine

a diaminotetrahydroxy*cyclo*hexane, and examination of the oxidation products of dibenzoylstreptamine with periodic acid led to the suggestion that streptidine is 1 : 3-diguanido-2 : 4 : 5 : 6-tetrahydroxy*cyclo*hexane (Carter *et al.*, 1946). Streptidine has been synthesised from streptamine (Wolfrom *et al.*, 1948). Since streptidine is not optically active, the configuration of the molecule must be *meso*, with the two guanido groups *cis* (see unit III).

***N*-Methyl-L-glucosamine** (unit II). When streptomycin is treated with methanolic hydrogen chloride (methanolysis), and then subjected to acid hydrolysis followed by acetylation, the penta-acetate of *N*-methyl-L-glucosamine is obtained; the parent compound is obtained by hydrolysis. The structure of *N*-methyl-L-glucosamine was confirmed by synthesis from L-arabinose (Kuehl *et al.*, 1946, 1947).

Streptose (unit I). The streptose fragment has not been isolated from streptomycin by degradation. It appears to be too unstable, but its structure was elucidated by various degradative experiments, *e.g.*, the alkaline



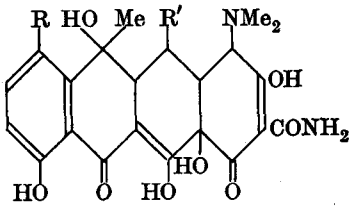
maltol

hydrolysis of streptomycin gives **maltol** (Schenck *et al.*, 1945), and this is produced by the conversion of a *furanose ring* into γ-pyrone.

Streptobiosamine (units I and II). Analytical work showed that this compound was a disaccharide, and from it was isolated *N*-methyl-L-glucosamine (see above). The formation of maltol and other analytical work led to the structure (I + II) for streptobiosamine, and then the points of attachment between streptobiosamine and streptidine were found, and so led to the structure given above for streptomycin (Kuehl *et al.*, 1947, 1948).

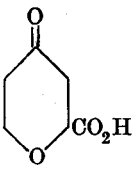
§7a. **Aureomycin** and **Terramycin.** Aureomycin was isolated from cultures of *Streptomyces aureofaciens*, and is used in the treatment of typhoid

fever, etc. Terramycin was isolated from cultures of *Streptomyces rimosus*, and is very effective in the treatment of trachoma. The structures of these antibiotics are (Woodward *et al.*, 1952):

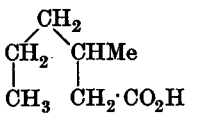


Aureomycin: R = Cl; R′ = H
Terramycin: R = H; R′ = OH

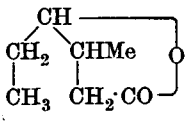
§8. **Patulin.** This has been obtained from various moulds. It is an optically inactive solid, and it inhibits Staphylococci and coliforms. The molecular formula of patulin is $C_7H_6O_4$; it is a neutral substance and forms a monoacetate. Hydrolysis of patulin with acid produces one molecule of formic acid and a small yield (10 per cent.) of tetrahydro-γ-pyrone-2-carboxylic acid (I). Catalytic reduction followed by further reduction with hydrogen iodide and red phosphorus gives 3-methylhexoic acid (II) and the lactone of 4-hydroxy-3-methylhexoic acid (III) [Birkinshaw *et al.*, 1943].



I

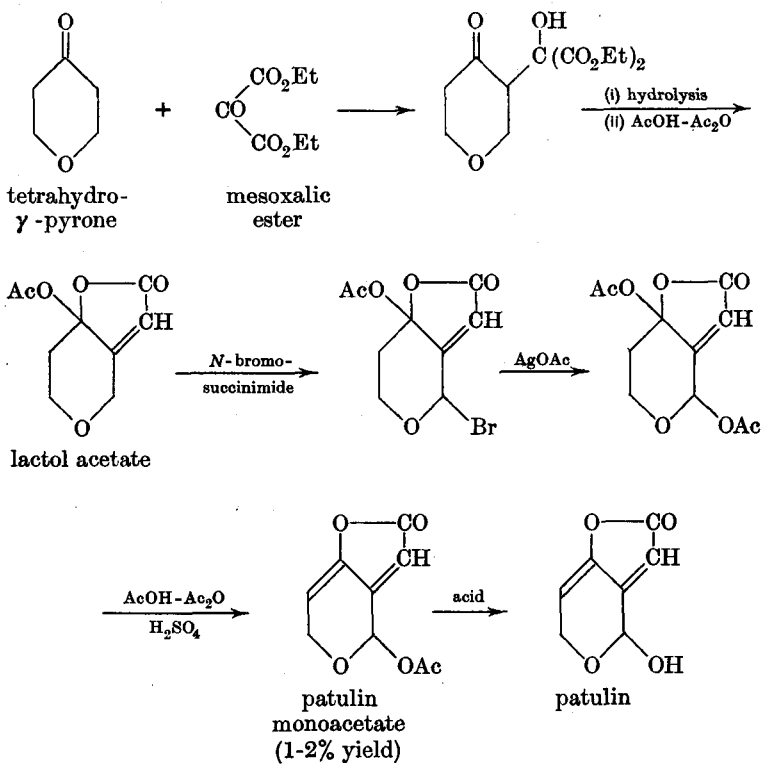


II



III

Woodward *et al.* (1949, 1950) have synthesised patulin as follows:



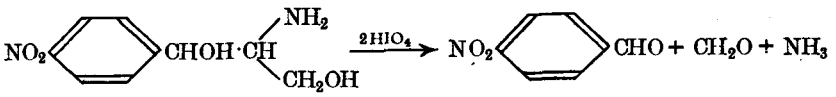
The monoacetate (obtained above) was shown to be identical with that obtained from patulin.

§9. **Chloramphenicol** (Chloromycetin). Chloramphenicol is a lævo-rotatory compound that is produced by *Streptomyces venezuelæ* (Carter *et al.*, 1948); it is very effective in the treatment of typhoid fever, etc.

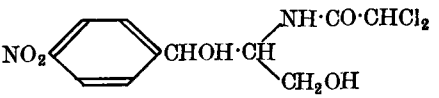
The molecular formula of chloramphenicol is $C_{11}H_{12}O_5N_2Cl_2$, and its absorption spectrum is similar to that of nitrobenzene. The presence of a nitro-group was shown by the reduction of chloramphenicol with tin and hydrochloric acid, followed by diazotisation and then coupling to give an orange-red precipitate with 2-naphthol (Rebstock *et al.*, 1949). When catalytically reduced (palladium), chloramphenicol gives a product which has an absorption spectrum similar to that of *p*-toluidine, and the solution contains ionic chlorine. The hydrolysis of chloramphenicol with acid or alkali produces dichloroacetic acid and an optically active base, $C_9H_{12}O_4N_2$. This base was shown to contain a primary amino-group, and when treated with methyl dichloroacetate, the base reformed chloramphenicol (Rebstock *et al.*, 1949).

Chloramphenicol is converted into a diacetyl derivative on treatment with acetic anhydride in pyridine; the base obtained from chloramphenicol forms a triacetyl derivative on similar treatment. Thus chloramphenicol probably contains two hydroxyl groups. When the base is treated with periodic acid, two molecules of the latter are consumed with the formation of one molecule each of ammonia, formaldehyde and *p*-nitrobenzaldehyde.

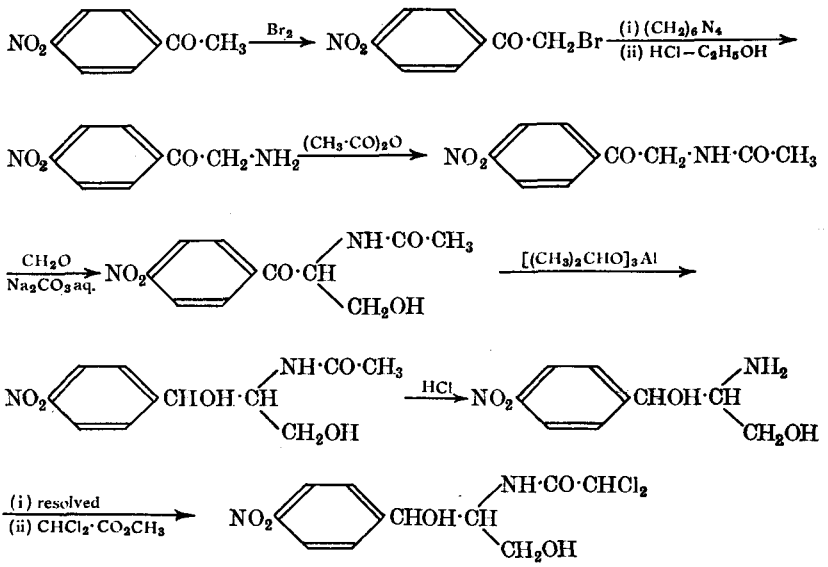
These products may be accounted for if the base is assumed to be 2-amino-1-*p*-nitrophenylpropane-1 : 3-diol (Rebstock *et al.*, 1949).



Thus chloramphenicol will be



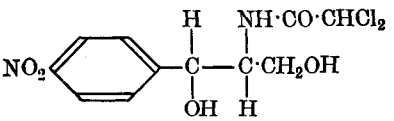
This structure has been confirmed by synthesis, *e.g.*, that of Long *et al.* (1949).



(–)-chloramphenicol

This structure has also been confirmed by crystallographic studies (Dunitz, 1952).

Chloramphenicol and the base contain two asymmetric carbon atoms; thus there are two possible pairs of enantiomorphs. Comparison of the properties of the base with those of norephedrine and nor-*ψ*-ephedrine (§7. XIV) showed that the configuration of the base was similar to that of nor-*ψ*-ephedrine (Rebstock *et al.*, 1949). Thus chloramphenicol is D-(—)-*threo*-2-dichloroacetamido-1-*p*-nitrophenylpropane-1 : 3-diol.



It is interesting to note that chloramphenicol is the first natural compound found to contain a nitro-group; the presence of the $CHCl_2$ group is also most unusual.

READING REFERENCES

Gilman (Ed.), *Advanced Organic Chemistry*, Wiley. Vol. III (1953). (i) Ch. 5. Some Aspects of Chemotherapy. (ii) Ch. 6. Antibiotics.
Raiziss and Gavron, *Organic Arsenical Compounds*, Chemical Catalog Co. (1923).
Work and Work, *The Basis of Chemotherapy*, Oliver and Boyd (1948).
Northey, *The Sulphonamides and Allied Compounds*, Reinhold (1948).
Northey, Structure and Chemotherapeutic Activities of Sulphanilamide Derivatives, *Chem. Reviews*, 1940, **27**, 85.
Haynes, Physiologically Active Unsaturated Lactones, *Quart. Reviews* (*Chem. Soc.*), 1948, **2**, 46.
Cook, The Chemistry of the Penicillins, *Quart. Reviews* (*Chem. Soc.*), 1948, **2**, 203.
The Chemistry of Penicillin, Princeton University Press (1949).
Knox, A Survey of New Penicillins, *Nature*, 1961, **192**, 492.
Antibiotics, Oxford Press (2 volumes; 1949).
Robinson, *Antibiotics*, Pitman (1953).
Waksman (Ed.), *Streptomycin*, Williams and Wilkins Co. (1949).
Lemieux and Wolfrom, The Chemistry of Streptomycin, *Advances in Carbohydrate Chemistry*, Academic Press, 1948, **3**, 337.
Brink and Folkers, Some Aspects of Streptomycin and Other *Streptomyces* Antibiotics, *Advances in Enzymology*, Interscience Publishers, 1950, **10**, 145.
Birkinshaw, The Chemistry and Biochemistry of Streptomycin and Related Compounds, *J. Pharm. Pharmacol.*, 1951, **3**, 529.
Rebstock *et al.*, Chloramphenicol, *J. Amer. Chem. Soc.*, 1949, **71**, 2458, 2463.
Long *et al.*, Chloramphenicol, *J. Amer. Chem. Soc.*, 1949, **71**, 2469, 2473.
Brink and Harman, Chemistry of Some Newer Antibiotics, *Quart. Reviews* (*Chem. Soc.*), 1958, **12**, 93.
Rose, A Chemotherapeutic Search in Retrospect, *J.C.S.*, **1951**, 2770.
Barber, Chance and Design in the Search for New Drugs, *Chem. and Ind.*, **1955**, 1460.
Burger, Rational Approaches to Drug Structure, *J. Chem. Educ.*, 1956, **33**, 362.
Bracken, *The Chemistry of Micro-Organisms*, Pitman (1955).

CHAPTER XIX

HÆMOGLOBIN, CHLOROPHYLL AND PHTHALOCYANINES

§1. Introduction. Two of the most important compounds of the natural porphyrins are hæmoglobin and chlorophyll. The bile pigments, which are formed mainly in the liver, are degradation products of hæmoglobin. Hæmoglobin and chlorophyll act as catalysts (biological) in many biological processes.

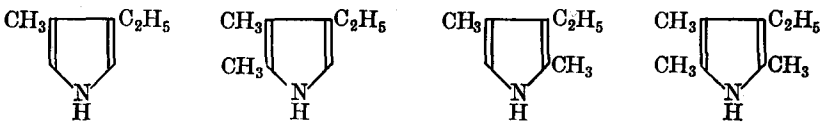
HÆMOGLOBIN

§2. Degradation products of hæmoglobin. Hæmoglobin occurs in all vertebrates (with certain exceptions) and in many invertebrates; it has also been found in certain strains of yeasts, moulds, etc. It is a chromo-protein (§7 B. XIII), the protein part being *globin* (94 per cent.), and the prosthetic group being *hæm* (6 per cent.). The composition of hæmoglobin varies slightly, depending on the species from which it is isolated; the variation occurs only in the globin part of the molecule. It is interesting to note that hæmoglobin was the first protein to be obtained in a crystalline form.

The way in which the globin part is bound to hæm has been the subject of much discussion. There appears to be agreement that the iron atom is bound to some part of the protein. The iron atom (bivalent) in hæm uses four co-ordination valencies in this molecule, and since iron has a co-ordination number of six, it is believed that it is these two valencies (which are perpendicular to the other four) that are joined to the globin molecule. Keilin (1960) has shown that only the basic nitrogen atoms in amino-acids can combine with hæm.

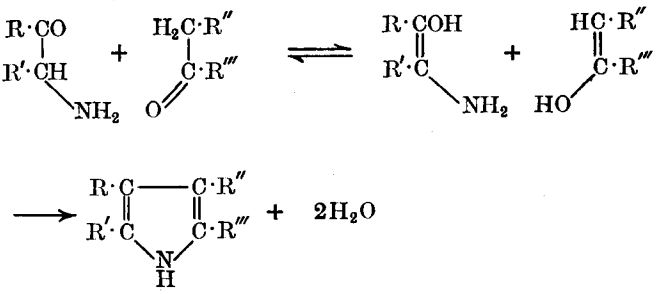
In the animal body, hæmoglobin readily combines with oxygen to form *oxyhæmoglobin*, and this, when treated with glacial acetic acid, forms *hæmatin*, $C_{34}H_{32}O_4N_4Fe^{III}{\cdot}OH$. The chloride of hæmatin is known as *hæmin*; its molecular formula is $C_{34}H_{32}O_4N_4Fe^{III}Cl$ (the chlorine is ionised, and the iron atom is in the ferric state). Hæmin may be prepared by warming blood with acetic acid and sodium chloride (Teichmann, 1853). The iron can be removed from hæmin, and replaced. The iron-free compounds are known as *porphyrins*, and the iron-containing compounds as *hæms*; the nature of the porphyrin depends on the conditions which are used to remove the iron atom from hæmin. When hæmin is reduced with sodium hyposulphite, the base *hæm* is produced in which the atom of iron is in the bivalent state; the molecular formula of hæm is $C_{34}H_{32}O_4N_4Fe$.

Since hæmin forms a diester with methanol, the molecule therefore contains two carboxyl groups. Also, since hæmin absorbs two molecules of hydrogen when catalytically reduced (palladium), two ethylenic double bonds are thus probably present in the molecule. When subjected to vigorous reduction with hydriodic acid and phosphonium iodide or hydriodic acid and acetic acid, hæmin is degraded into the four pyrrole derivatives opsopyrrole, I, hæmopyrrole, II, cryptopyrrole, III, and phyllopyrrole, IV. All four compounds have been synthesised by means of the Knorr pyrrole synthesis (1884, 1886); this is the condensation between an α-aminoketone

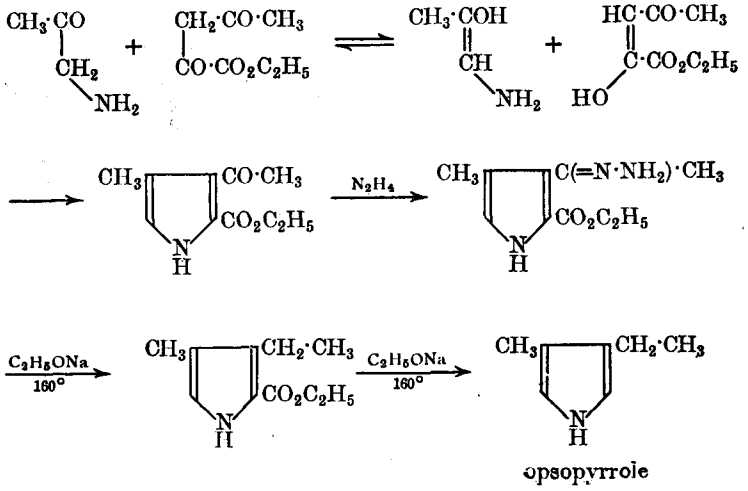


opsopyrrole I | hæmopyrrole II | cryptopyrrole III | phyllopyrrole IV

and a ketone containing an active methylene group, *i.e.*, a compound containing the group —CH_2·CO—. The mechanism of the reaction is not known; possibly the enol forms are involved, and so we may write the general reaction as follows:

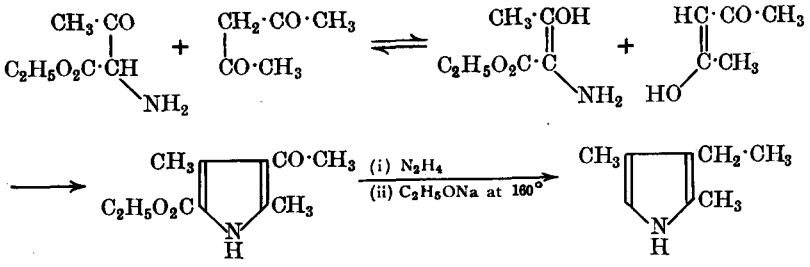


A detailed study of this reaction has shown that the yields depend on the nature of R, R′, R″ and R‴; when R′ and R″ are acyl or carbalkoxyl groups, the yields are usually very good. As examples of the Knorr synthesis, let us consider the preparation of opsopyrrole (I) and cryptopyrrole (III). Opsopyrrole may be synthesised by condensing aminoacetone with ethyl 2: 4-diketopentanoate, and then subjecting the product to the Wolff–Kishner reduction, *i.e.*, first converting the product into the hydrazone and then heating the latter with sodium ethoxide at 160°. By this means a keto-group is converted into a methylene group (see also Vol. I). By using an excess of sodium ethoxide, decarboxylation is also effected at the same time.



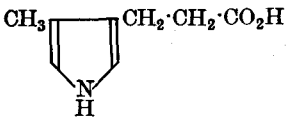
opsopyrrole

Cryptopyrrole may be prepared in a similar manner, starting from ethyl α-aminoacetoacetate and acetylacetone (penta-2 : 4-dione).

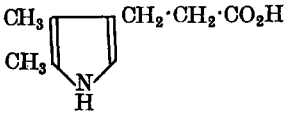


cryptopyrrole

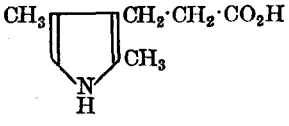
When reduced with tin and hydrochloric acid, hæmin is again degraded into four pyrrole derivatives, but in this case the products are all carboxylic acids in which each of the four pyrroles I–IV contains a carboxyl group attached to the ethyl group:



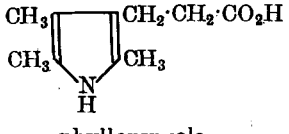
opsopyrrole-carboxylic acid
V



hæmopyrrole-carboxylic acid
VI

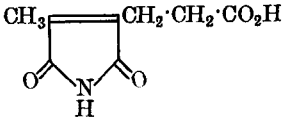


cryptopyrrole-carboxylic acid
VII

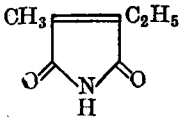


phyllopyrrole-carboxylic acid
VIII

When oxidised with chromic acid, hæmin gives two molecules of *hæmatinic acid* (IX). On the other hand, *mesoporphyrin* (see below) gives, on oxidation, two molecules of ethylmethylmaleimide (X).



hæmatinic acid
IX



ethylmethylmaleimide
X

The treatment of hæmin with iron dust and formic acid results in the removal of the iron atom and the formation of *protoporphyrin*, $C_{34}H_{34}O_4N_4$. The iron atom is also removed from hæmin by the action of hydrobromic acid in acetic acid, but in this case the product is *hæmatoporphyrin*, $C_{34}H_{38}O_6N_4$. If, however, hæmin is treated with hydriodic acid in acetic acid, the iron atom is again removed and *mesoporphyrin*, $C_{34}H_{38}O_4N_4$, is obtained.

Finally, when porphyrins containing two carboxyl groups are decarboxylated, the products obtained (after reduction, if necessary) are known as

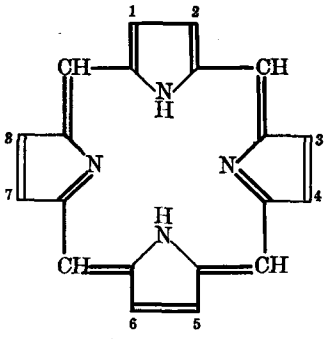
ætioporphyrins, e.g., when protoporphyrin is decarboxylated, and the product then reduced, the final product is *ætioporphyrin*, $C_{32}H_{38}N_4$, which is also a degradation product of chlorophyll. Thus hæmin and chlorophyll are closely related chemically.

The table summarises the reactions that have been discussed.

Compound	Reaction	Products
Hæmoglobin . .	Atmospheric oxidation	Oxyhæmoglobin
Oxyhæmoglobin .	$CH_3{\cdot}CO_2H$	Hæmatin
Oxyhæmoglobin .	$CH_3{\cdot}CO_2H$ + NaCl	Hæmin
Hæmin . . .	$Na_2S_2O_4$	Hæm
Hæmin . . .	HI + PH_4I	Osopyrrole, Hæmopyrrole, Cryptopyrrole and Phyllopyrrole
Hæmin . . .	Sn—HCl	Opsopyrrole-, Hæmopyrrole-, Cryptopyrrole- and Phyllopyrrole-carboxylic acids
Hæmin . . .	CrO_3—H_2SO_4	Hæmatinic acid
Mesoporphyrin .	CrO_3—H_2SO_4	Ethylmethylmaleimide
Hæmin . . .	Fe—$H{\cdot}CO_2H$	Protoporphyrin
Hæmin . . .	HBr—$CH_3{\cdot}CO_2H$	Hæmatoporphyrin
Hæmin . . .	HI—$CH_3{\cdot}CO_2H$	Mesoporphyrin
Porphyrin . . .	Decarboxylation (and then reduction, if necessary)	Ætioporphyrins

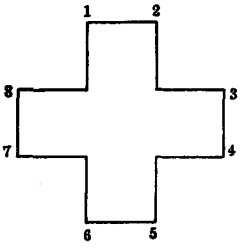
From the foregoing evidence (the molecular formula and the degradation products of hæmin), it is reasonable to infer that hæmin contains four substituted pyrrole nuclei linked together. The isolation of the pyrroles I–IV suggests that each of the four pyrrole nuclei contains a methyl group in the β-position. The isolation of the oxidation products IX and X (oxidation at the α-position), and of the reduction products I–VIII (appearance of a methyl group at the α-position), suggests that the pyrrole nuclei are linked at the α-positions *via* one carbon atom. The isolation of two molecules of IX suggests the presence of two propionic acid residues each in the β-position of two pyrrole nuclei (this would also account for the two carboxyl groups present in hæmin). The appearance of ethyl groups in I–IV on the reduction of hæmin could be explained by the presence of two vinyl groups in the β-position of two pyrrole nuclei (hæmin contains two ethylenic double bonds). A possible structure for hæmin is thus a ring structure containing four pyrrole nuclei linked at the α-positions *via* one carbon atom, with four β-positions occupied by methyl groups, two β'-positions by vinyl groups and the remaining two β'-positions by propionic acid residues. Küster (1912) was the first to propose that the four pyrrole nuclei formed a cyclic structure, and this has been proved correct by synthetic work; the porphyrins so obtained had the same absorption spectra as the natural porphyrins. At the same time, this synthetic work established the nature and the positions of the substituent groups.

The parent substance of all the compounds mentioned above is *porphin* (XI), and this may conveniently be written as XII (H. Fischer). In this porphin molecule there is an eighteen-membered ring containing a complete arrangement of conjugated double bonds. Thus many resonating structures contribute to this molecule, and consequently its stability will be great; this is observed in practice, *e.g.*, the molecule has a very large heat of combustion. Also, the resonance gives rise to the colour in porphin derivatives (see Ch. XXXI, Vol. I); porphin itself does not occur naturally. It has been shown, by analogy with the X-ray data on phthalocyanines (§9), that the



porphin

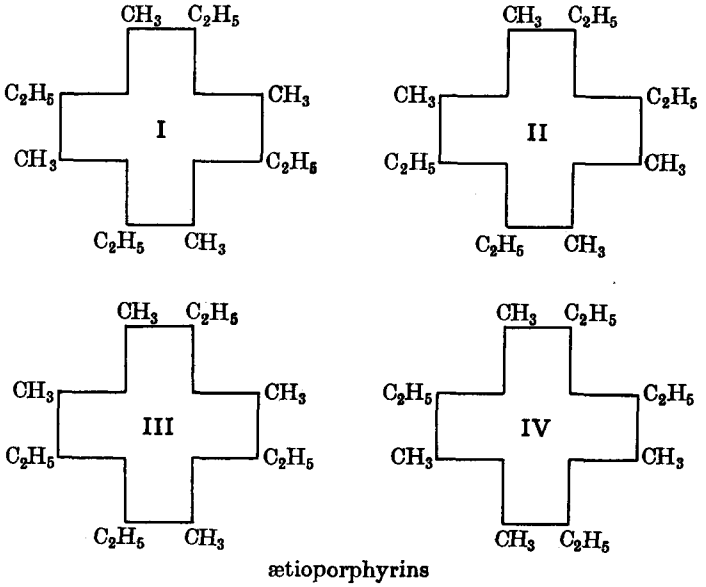
XI



XII

porphin molecule is planar, and this planar structure is also in agreement with magnetic measurements.

The ætioporphyrins, $C_{32}H_{38}N_4$, are derivatives of porphin in which the 3- and 4-positions of each pyrrole nucleus are substituted by methyl and ethyl groups. Four isomers are possible, and these are known as ætioporphyrin I, II, III and IV, respectively.

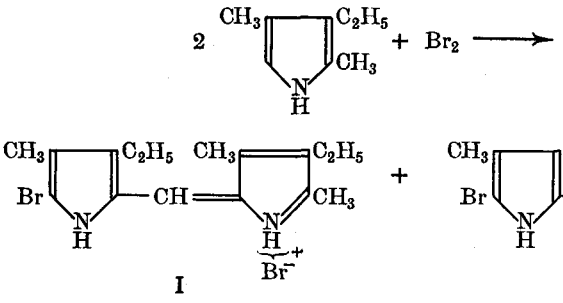


ætioporphyrins

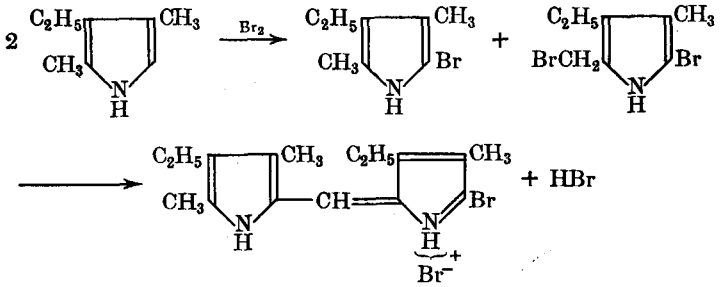
All four ætioporphyrins have been synthesised; the degradation of hæmin gives ætioporphyrin III.

§3. Synthesis of the porphyrins. The first step in the synthesis of porphyrins is the synthesis of the dipyrrylmethenes.

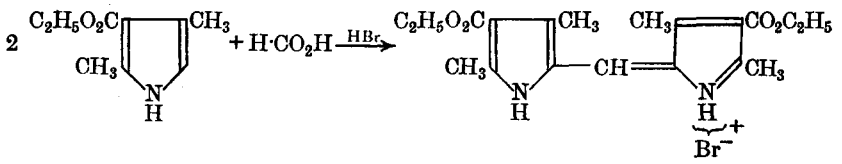
(i) Dipyrrylmethenes may be prepared by the bromination of a *2-methylpyrrole in which position 5 is vacant* (H. Fischer, 1915); at least two products are obtained, *e.g.*, cryptopyrrole gives compounds I and II.



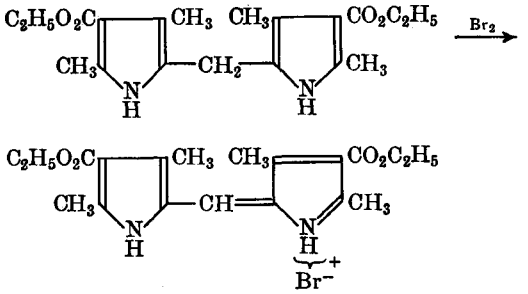
According to Corwin *et al.* (1944), the mechanism of this reaction is:



(ii) When pyrroles, in which the 5-position is vacant, are coupled by means of formic acid in the presence of hydrobromic acid, dipyrrylmethenes are produced (H. Fischer *et al.*, 1922); *e.g.*,

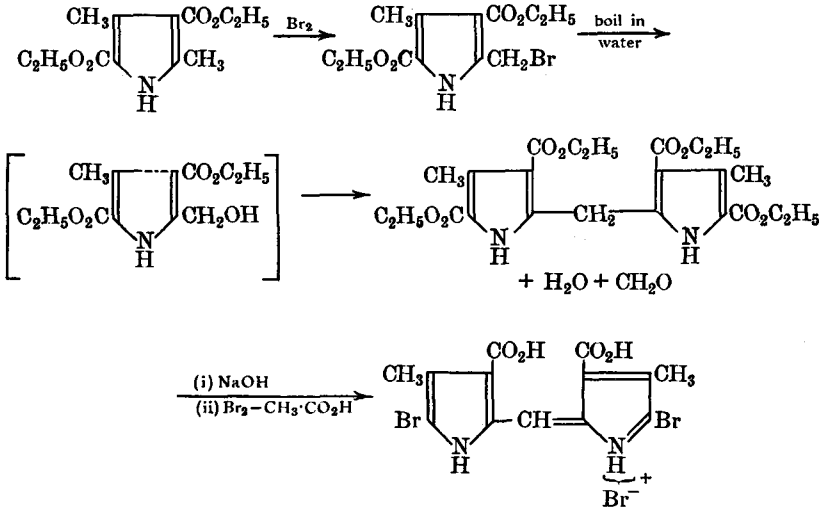


(iii) Piloty *et al.* (1914) showed that dipyrrylmethanes may be oxidised to the corresponding methenes by means of bromine, *e.g.*,

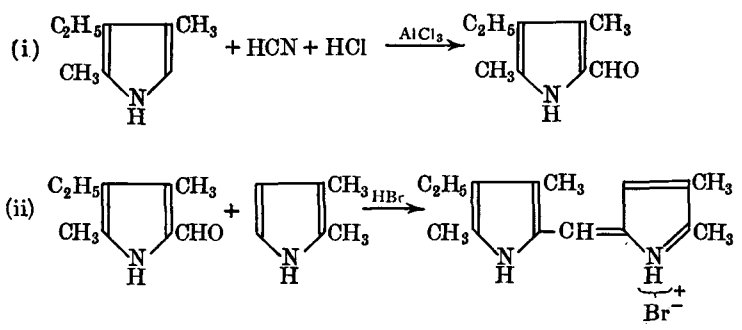


H. Fischer *et al.* (1923) modified the above procedure as follows. A dipyrrylmethane containing carbethoxyl groups was first prepared, this then hydrolysed and then treated with bromine in acetic acid. In this way the methane

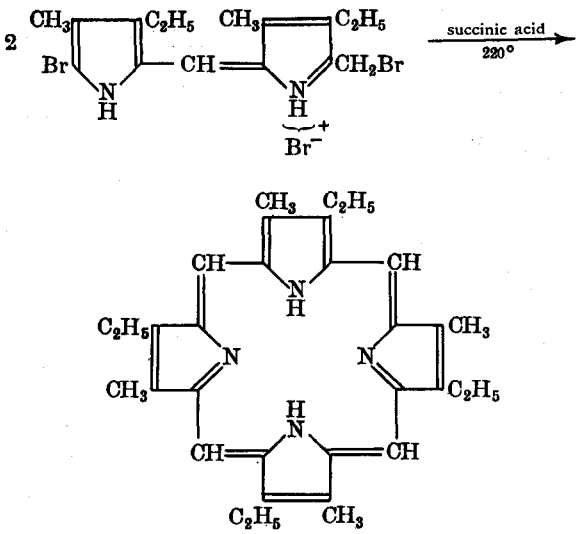
derivative is oxidised to the methene compound, but at the same time the carboxyl groups in position 5 : 5′ are replaced by bromine atoms, *e.g.*,



(iv) The foregoing methods (except i) lead to the formation of *symmetrical* dipyrrylmethenes. The preparation of *unsymmetrical* dipyrrylmethenes is best carried out as follows, using the Gattermann aldehyde synthesis (Piloty *et al.*, 1912, 1914; H. Fischer *et al.*, 1926); *e.g.*,

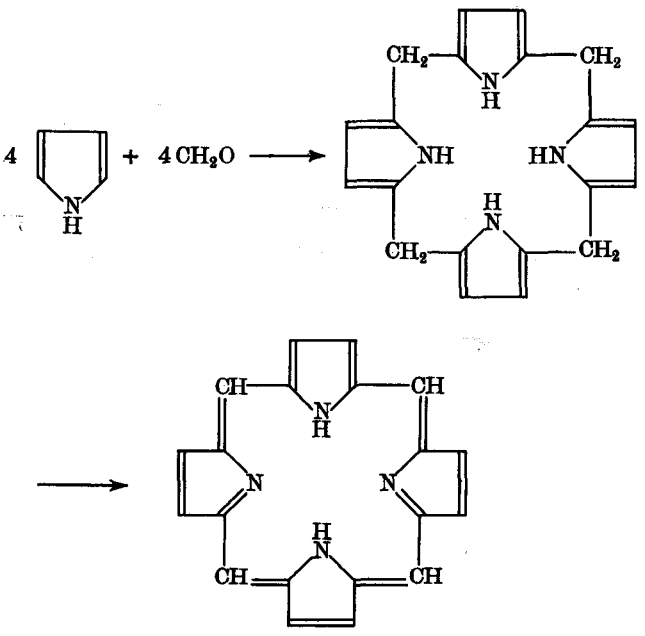


The dipyrrylmethenes are coloured solids. H. Fischer *et al.* (1926) then prepared porphyrins by condensing two molecules of a dipyrrylmethene by heating with succinic acid at 220°, *e.g.*, ætioporphyrin I. Porphin itself was synthesised by H. Fischer *et al.* (1935) by boiling pyrrole-2-aldehyde with



ætioporphyrin I

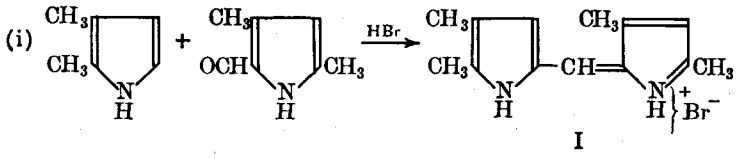
formic acid and ethanol. A later synthesis is by heating pyrrole with formaldehyde in the presence of a mixture of methanol and pyridine (Rothemund, 1936, 1939; Calvin *et al.*, 1943).



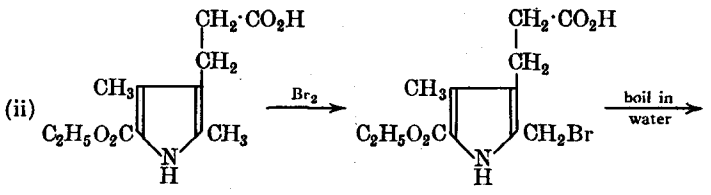
porphin

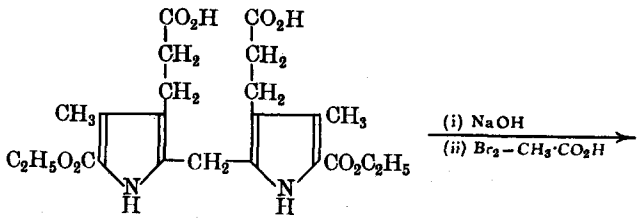
It should be noted that the two imino hydrogen atoms are replaced by the iron atom in the hæms, and the iron atom is covalently bound.

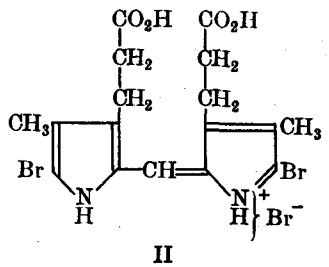
§4. Synthesis of hæmin (H. Fischer *et al.*, 1929).



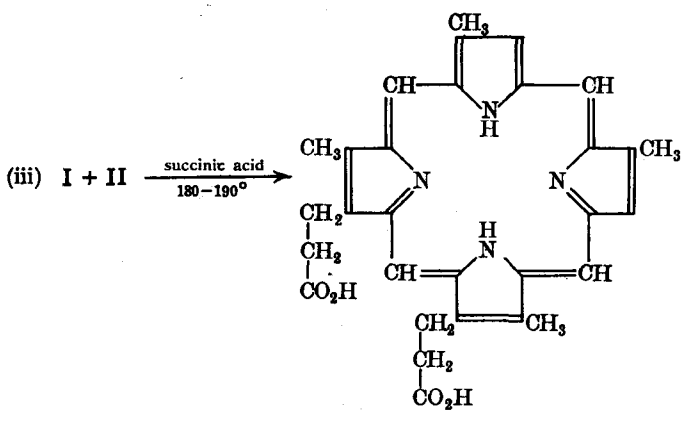
I



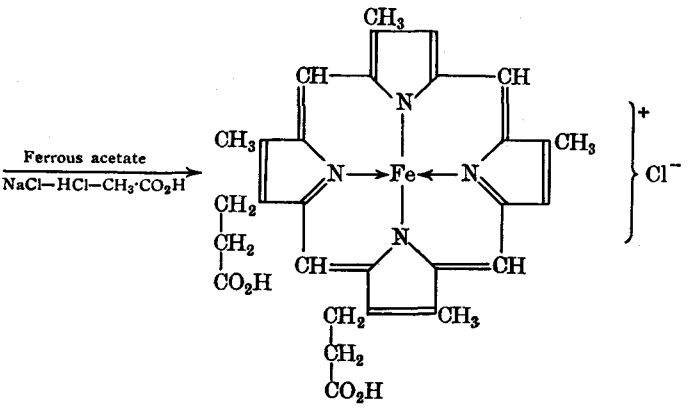




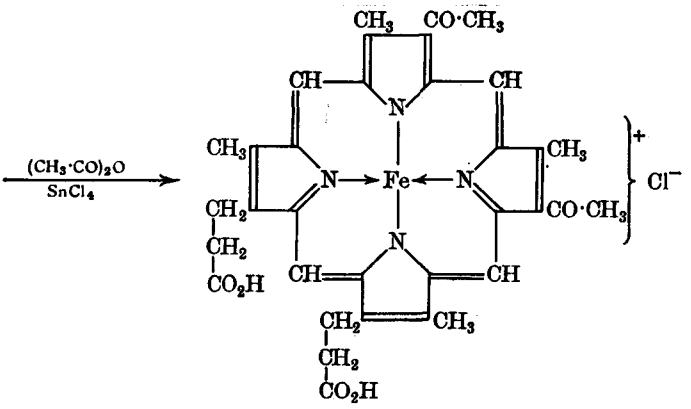
II



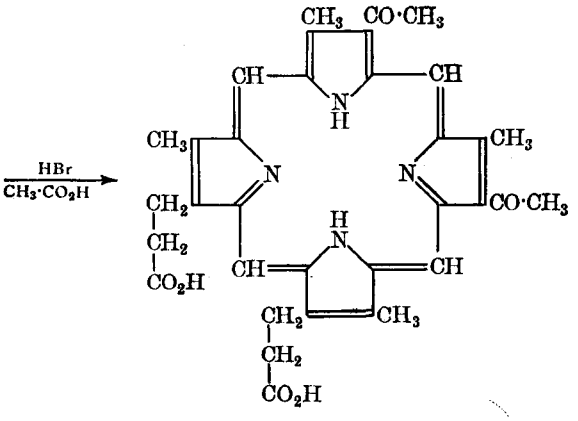
deuteroporphyrin



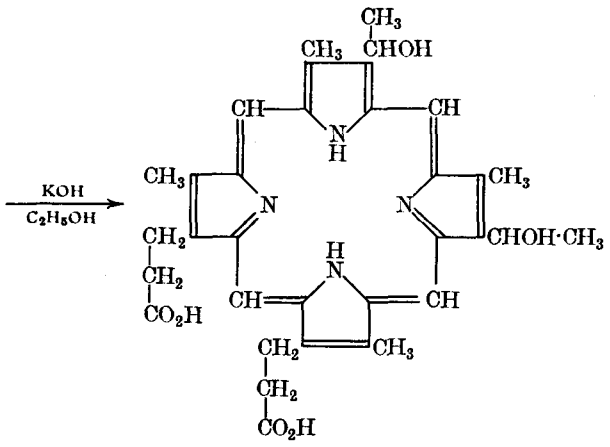
deuterohæmin



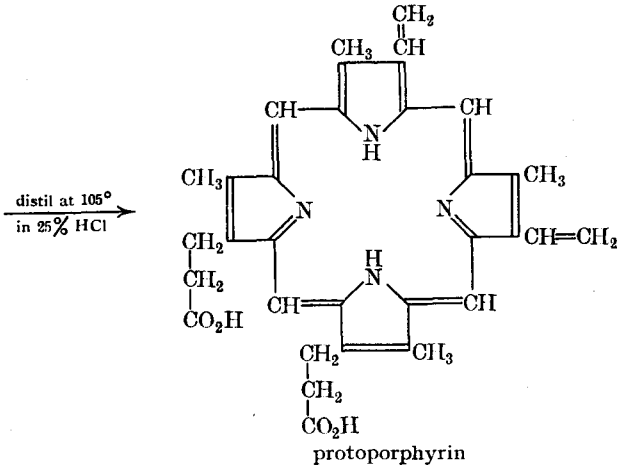
diacetyldeuterohæmin



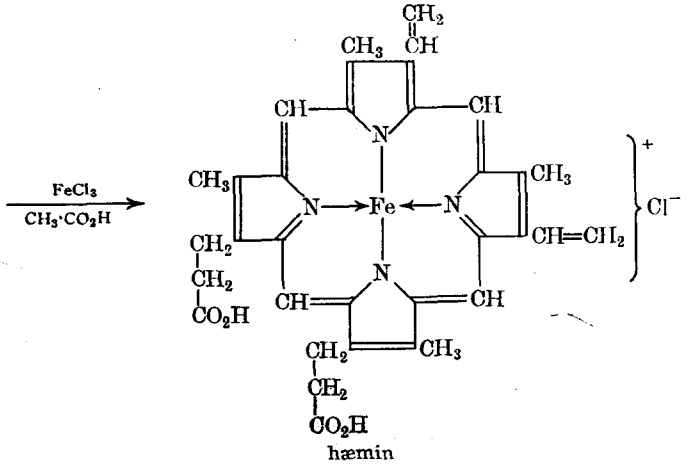
diacetyldeuteroporphyrin



hæmatoporphyrin



protoporphyrin



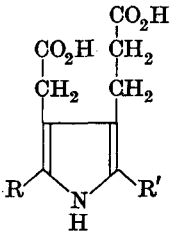
hæmin

It should be noted that the introduction of the iron atom into deuteroporphyrin to give deuterohæmin renders the pyrrole nuclei more reactive.

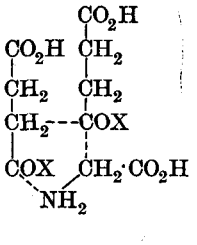
§4a. Biosynthesis of porphyrin. The progress made in this field is one of the outstanding examples of the use of isotopes. Tracer syntheses *in vivo* and *in vitro* and degradation methods have established the origin of all the carbon and nitrogen atoms in protoporphyrin (of hæm), and have also established the nature of the pyrrole precursors. These results are the outcome of a large volume of work, but in the following account only a few experiments have been mentioned. These indicate, to some extent, the lines of research pursued.

Bloch *et al.* (1945), using acetic acid labelled with deutero atoms, showed that deuterohæmin was produced. Thus at least the methyl carbon of acetic acid is involved in the biosynthesis of hæm. Then Shemin *et al.* (1950) and Neuberger *et al.* (1950) carried out experiments with $^{14}CH_3 \cdot CO_2H$ and $CH_3 \cdot {}^{14}CO_2H$, and showed that *both* carbon atoms of acetate participate in the synthesis of hæm. The latter authors also showed that with $^{14}CH_3 \cdot CO_2H$, about half of the radioactive tracer atom appeared in the two pyrrole nuclei carrying the vinyl radicals, and the other half in the two pyrrole nuclei carrying the propionic acid residues. When, however, $CH_3 \cdot {}^{14}CO_2H$ was used as the precursor, then about 20 per cent. of the tracer atom appeared in the vinyl pyrrole nuclei and 80 per cent. in the propionic acid pyrrole nuclei. In neither case of the labelled acetates was there any significant radioactivity in the methine carbon of the hæm. Thus the carbons of the methine bridges do not originate from acetate.

Shemin *et al.* (1945, 1946) carried out experiments with [^{15}N] glycine, and showed that all the nitrogen atoms in hæm are derived from this glycine. Shemin *et al.* (1950) also used $CH_2 \cdot NH_2 \cdot {}^{14}CO_2H$, and showed that the carboxyl group of glycine is *not* incorporated into protoporphyrin. On the other hand, Altman *et al.* (1948), using $^{14}CH_2 \cdot NH_2 \cdot CO_2H$, showed that the α-carbon atom of glycine *is* used in the protoporphyrin synthesis. This was confirmed by Shemin *et al.* (1950) who used $^{14}CH_2 \cdot {}^{15}NH_2 \cdot CO_2H$ and showed that for each nitrogen used for hæm synthesis, two α-carbon atoms of glycine were also incorporated into the molecule. Similar results were obtained by Neuberger *et al.* (1950) who also showed that the α-carbon atom of glycine is used in the formation of the methine bridge. Thus all the carbon atoms of protoporphyrin, except eight derived from the α-carbon of glycine, originate from acetate. Furthermore, a detailed study of the degradation products of the labelled protoporphyrins showed that it was very probable that the two sides of the pyrrole nuclei were synthesised from identical intermediates. It also seemed very reasonable that a *common*



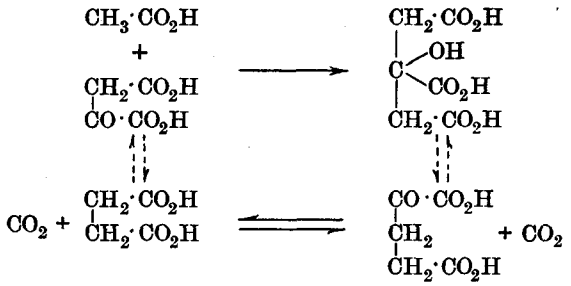
I



II

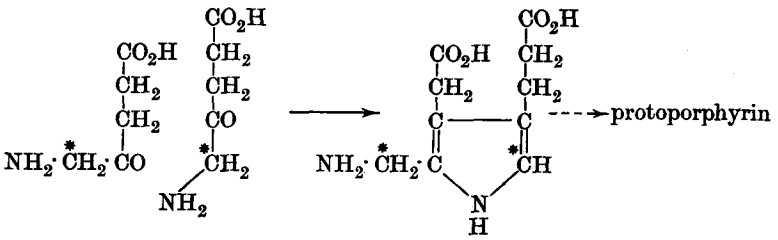
pyrrole of the type I was formed first. Also, consideration of the distribution of the radioactivity of the carbon atoms of the propionic acid residue and the (pyrrole) nuclear carbon to which it was attached led to the suggestion

that succinic acid was a precursor, and that two molecules of this, on condensation with one molecule of glycine, could form the common pyrrole (I). The tracer distribution of the labelled succinic acid could arise by acetate entering the *Krebs cycle* (§18. XIII). Two molecules of " active " succinate



(succinyl-coenzyme A) and one of glycine then forms the common precursor (see II). Shemin *et al.* (1952) tested this succinic acid hypothesis by using $^{14}CO_2H{\cdot}CH_2{\cdot}CH_2{\cdot}^{14}CO_2H$ and $CO_2H{\cdot}^{14}CH_2{\cdot}^{14}CH_2{\cdot}CO_2H$, and showed that hæm contained the labelled carbon.

In 1952, Westall isolated porphobilinogen from the urine of humans suffering from acute porphyria. Based on this, Shemin *et al.* (1953) now proposed that δ-aminolævulic acid can replace " active " succinate and glycine in porphyrin synthesis:



porphobilinogen

This pyrrole synthesis is supported by various experiments, *e.g.*, Shemin *et al.* (1954) used [δ^{14}C]δ-aminolævulic acid as precursor, and showed that half of the radioactivity is equally distributed among the four pyrrole nuclei and the other half is in the methine-bridge carbons. This distribution is in agreement with the equation given. Furthermore, Falk *et al.* (1953) have shown that porphobilinogen is the common precursor in porphyrin synthesis.

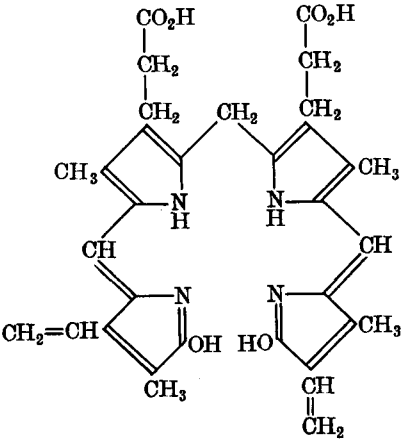
The problem of the conversion of porphobilinogen into protoporphyrin has still to be elucidated. There is evidence to show that porphobilinogen is first converted mainly into uroporphyrinogen III (this is ætioporphyrin III (§2) with Me = $\cdot CH_2{\cdot}CO_2H$ and Et = $\cdot CH_2{\cdot}CH_2{\cdot}CO_2H$) by certain enzymes, and then this compound is converted into protoporphyrin by enzymes. Decarboxylation of the acetic acid radicals would produce the methyl radicals (in protoporphyrin). The conversion of the propionic acid residues into vinyl radicals takes place by a series of steps; a possibility is:

$$—CH_2{\cdot}CH_2{\cdot}CO_2H \longrightarrow —CH{=}CH{\cdot}CO_2H \longrightarrow —CH{=}CH_2$$

§5. Bile pigments. Several pigments occur in bile, *e.g.*, bilirubin, mesobilirubin, etc.; the most important one is **bilirubin,** $C_{33}H_{36}O_6N_4$. On vigorous

oxidation, bilirubin gives hæmatinic acid; and on vigorous reduction, it gives cryptopyrrole and cryptopyrrolecarboxylic acid. When catalytically reduced, bilirubin gives mesobilirubin, $C_{33}H_{40}O_6N_4$, which, on reduction with hydriodic acid in acetic acid, forms, among other products, bilirubic acid, $C_{17}H_{24}O_3N_2$, and neobilirubic acid, $C_{16}H_{22}O_3N_2$. Finally, the reduction of bilirubic acid gives cryptopyrrolecarboxylic acid as the main product, and the reduction of neobilirubic acid gives hæmopyrrolecarboxylic acid. From this evidence it is reasonable to conclude that bilirubin contains the four pyrrole nuclei that occur in hæmoglobin. Furthermore, there is much evidence to show that bilirubin is a degradation product of hæmoglobin.

Since the absorption spectrum of bilirubin is not like that of a porphyrin, it is assumed that bilirubin has an *open-chain* structure. Further degradative and synthetic work has shown that bilirubin probably has the following structure.



bilirubin

CHLOROPHYLL

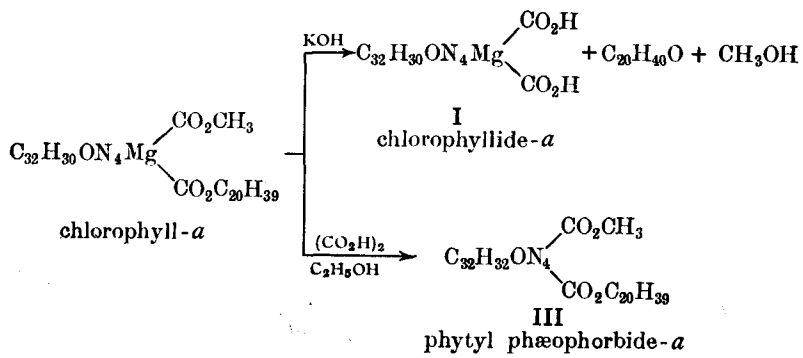
§6. Introduction. Chlorophyll is the green colouring matter of leaves and green stems, and its presence is essential for photosynthesis. Photosynthesis is the process in which light energy is used by plants to synthesise carbohydrates, proteins and fats. In green plants it is the chlorophyll which absorbs the light energy.

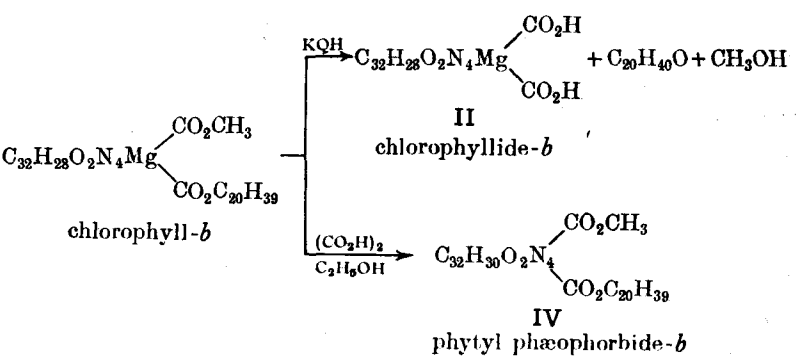
The name *chlorophyll* was given to the green pigment in leaves by Pelletier and Caventou (1818). There the matter rested until 1864, when Stokes showed, from spectroscopic evidence, that chlorophyll was a mixture. This paper apparently did not attract much attention, and it was not until Willstätter entered the field that any progress in the chemistry of chlorophyll was made.

When dried leaves are powdered and then digested with ethanol, a " crystalline " chlorophyll is obtained after concentration of the solvent. If, however, ether or aqueous acetone is used instead of ethanol, then the product is " amorphous " chlorophyll (Willstätter *et al.*, 1908). The extraction of chlorophyll is also accompanied by the extraction of two other pigments, carotene and xanthophyll (see Ch. IX). Willstätter *et al.* (1910) then showed that " crystalline " chlorophyll was produced during the extraction of chlorophyll by means of ethanol, a molecule of phytyl alcohol being replaced by ethanol under the influence of an enzyme, chlorophyllase (which is present in leaves). Nettle leaves are the main source for the extraction of chlorophyll on a large scale.

Willstätter *et al.* (1911) originally gave chlorophyll the molecular formula $C_{55}H_{72}O_6N_4Mg$, but in 1912 Willstätter *et al.* showed that chlorophyll, obtained from a wide variety of sources, was a mixture of two compounds, chlorophyll-*a* and chlorophyll-*b*. The separation was effected by shaking a light petrol solution of chlorophyll with aqueous methanol; chlorophyll-*a* remains in the light petrol, and chlorophyll-*b* passes into the aqueous methanol. Chlorophyll-*a* is a bluish-black solid, giving a green solution in organic solvents; chlorophyll-*b* is a dark green solid, also giving a green solution in organic solvents. The two components occur in proportions of approximately 3 of *a* to 1 of *b* in natural chlorophyll. Winterstein *et al.* (1933) have separated the two chlorophylls by means of chromatography (on sucrose as adsorbent). This technique has been improved by various workers (*inter alia*, Calvin *et al.*, 1962).

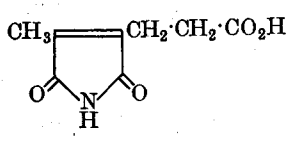
The molecular formulæ that have been assigned to chlorophyll-*a* and chlorophyll-*b* are $C_{55}H_{72}O_5N_4Mg$ and $C_{55}H_{70}O_6N_4Mg$, respectively (Willstätter, 1913); the two compounds have different absorption spectra (*cf.* Stokes, above). The hydrolysis of both chlorophylls with cold dilute potassium hydroxide solution gives one molecule of phytol, $C_{20}H_{40}O$ (see §30. VIII), one molecule of methanol, and one molecule of chlorophyllide-*a* (chlorophyllin-*a*), I, or chlorophyllide-*b* (chlorophyllin-*b*), II. Thus the chlorophylls are di-esters. When either chlorophyll is heated with an ethanolic solution of hydrated oxalic acid, the magnesium atom is replaced by two hydrogen atoms to produce phytyl phæophorbide-*a* (III) or *b* (IV; these phytyl phæophorbides are also known as phæophytins *a* and *b*, and " crystalline " chlorophyll is ethyl chlorophyllide). The foregoing reactions may be formulated as follows:



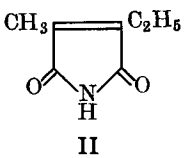


§6a. Nomenclature of the chlorophyll degradation products. Porphyrins are substituted porphins (see §2). Phyllins, phyllides and chlorophylls contain magnesium, whereas phorbins, phorbides and phytins are magnesium-free compounds, the magnesium atom having been removed and replaced by two hydrogen atoms. 7 : 8-Dihydroporphin is the nucleus of the *chlorin* series of compounds (tricarboxylic derivatives) which are derived from chlorophyll-*a*; *rhodins* are the corresponding compounds derived from chlorophyll-*b*. The introduction of the extra ring—two methylene groups across the 6 : γ-positions (see §7)—gives rise to the *phorbins*. The prefix phæo designates those compounds which have the same substituents that occur in chlorophyll. Chlorin itself is dihydroporphin, and the natural *red* porphyrin pigments are derivatives of porphin, whereas the *green* chlorophylls and their derivatives are derivatives of chlorin. Furthermore, examination of formulæ XIV and XV (in §7) shows that there is still complete conjugation in chlorin as in porphin (formula XI, §2). Chlorin has been synthesised by Linstead *et al.* (1955), and has been dehydrogenated to porphin.

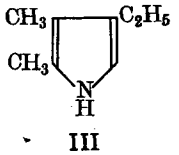
§7. Structure of chlorophyll-a. When phytyl phæophorbide-*a* is hydrolysed with boiling methanolic potassium hydroxide (30 seconds), the product is chlorin-*e*. This is a tricarboxylic acid (*e.g.*, it forms a trimethyl ester), and its molecular formula may thus be written as $C_{31}H_{33}N_4(CO_2H)_3$. Chlorin-*e*, on oxidation with chromic acid or with Caro's acid, gives hæmatinic acid, I, and ethylmethylmaleimide, II (Willstätter *et al.*, 1910). When chlorin-*e* is



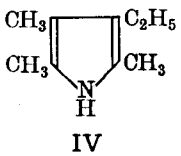
I



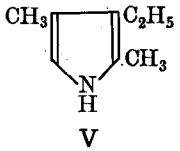
II



III

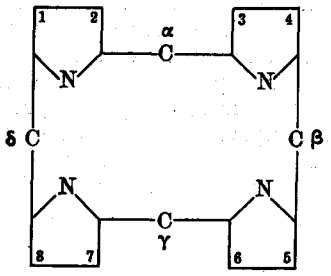


IV



V

reduced with hydriodic acid in acetic acid, hæmopyrrole, III, and phyllopyrrole, IV, are produced (Willstätter *et al.*, 1911). When phylloporphyrin (see below) is reduced under the same conditions, the products are now III, IV, and cryptopyrrole, V. From these results it is reasonable to infer that chlorophyll-*a* contains four pyrrole nuclei, each probably having a methyl group in the β-position (see II–V). It is also reasonable to suppose that at least one pyrrole nucleus contains a propionic acid residue in the β'-position (see I). It also appears likely that a vinyl group is present in the molecule (this would account for the presence of an ethyl group on reduction; at the same time, the presence of an ethyl group, as such, is not excluded). Furthermore, the isolation of I and II on oxidation (giving oxidation at the α-position), and of III and IV on reduction (the appearance of a methyl group at the α-position), can be interpreted as meaning that the four pyrrole nuclei are joined to each other at their α-positions *via* one carbon atom (*cf.* §2). Thus a possible skeleton structure for chlorin-*e* could be a cyclic one, VI; the positions of the various substituent groups cannot be assigned on the evidence obtained so far, *e.g.*, a methyl group at 1 and a propionic acid residue at 2 would produce the same oxidation product I had the positions

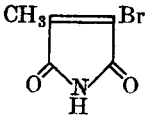


VI

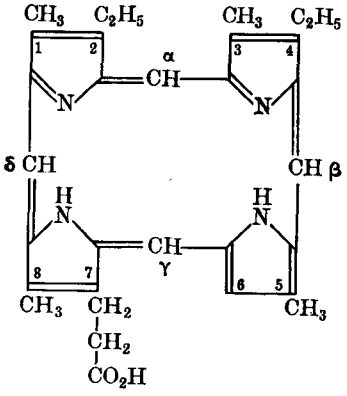
of the two groups been interchanged in VI. It is also necessary to fit a second carboxyl group into this structure (VI), since chlorophyll-*a* forms chlorophyllide-*a* on hydrolysis (the latter compound contains two carboxyl groups). Furthermore, since chlorophyllide-*a*, on further hydrolysis, forms chlorin-*e*, a tricarboxylic acid, some group must be present which can give rise to this third carboxyl group. Such a group could be a lactone; it must be *cyclic* since no carbon atoms are lost after the hydrolysis.

By the further degradation of chlorin-*e*, *e.g.*, heating in a sealed tube with ethanolic potassium hydroxide, various porphyrins are obtained. Three of these are pyrroporphyrin, rhodoporphyrin and phylloporphyrin.

Pyrroporphyrin, $C_{30}H_{33}N_4{\cdot}CO_2H$, has an absorption spectrum closely resembling that of mesoporphyrin (see §2); this agrees with the tentative skeleton structure VI proposed for chlorin-*e*. Pyrroporphyrin, on bromination followed by oxidation with chromic acid, gives bromocitraconimide, VII, as one of the products (Treibs *et al.*, 1928). It therefore follows that at least one of the pyrrole nuclei in pyrroporphyrin has a free β-position available for bromination. Synthetic work then showed that pyrroporphyrin has structure VIII (H. Fischer *et al.*, 1929, 1930, 1933); thus the positions of the four methyl groups and the position of the propionic acid group are now established.



VII



VIII

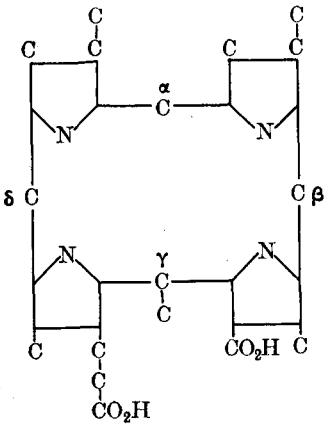
pyrroporphyrin

Rhodoporphyrin, $C_{30}H_{32}N_4(CO_2H)_2$, on heating with sodium ethoxide, readily loses *one* carboxyl group to form pyrroporphyrin (VIII). From a detailed study of the hæmin series, it was observed that a carboxyl group in a *side-chain* of a pyrrole nucleus was difficult to remove. Hence it is

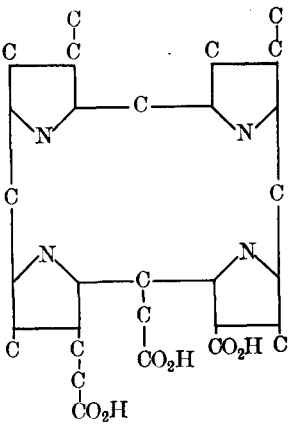
probable that the carboxyl group lost from rhodoporphyrin is attached *directly* to a pyrrole nucleus. The only position for this carboxyl group is at 6 (see structure VIII); elimination of the carboxyl group from rhodoporphyrin would then give one pyrrole nucleus with a free β-position (6), *i.e.*, pyrroporphyrin. Furthermore, comparison of the absorption spectra of rhodoporphyrin with compounds of known structure showed that the two carboxyl groups are in positions 6 and 7 (the latter is the propionic acid residue), and this was confirmed by the synthesis of rhodoporphyrin.

Phylloporphyrin, $C_{31}H_{35}N_4{\cdot}CO_2H$, contains one CH_2 group more than pyrroporphyrin, and may be converted into the latter by heating with sodium ethoxide. It therefore follows that the alkyl groups in both compounds occupy similar positions. Synthetic work then showed that phylloporphyrin contains a methyl group attached to the γ-methyne carbon atom (H. Fischer *et al.*, 1930, 1933).

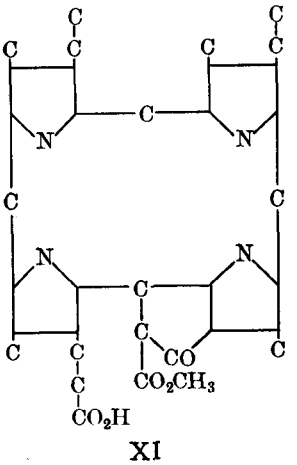
Consideration of the information obtained from the structures of the porphyrins described above shows that the skeleton structure IX is present in chlorin-*e*. Now chlorin-*e* contains three carboxyl groups and one more carbon atom than the structure shown in IX. The formation of a methyl



IX



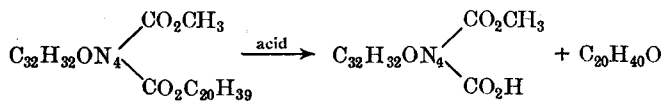
X



XI

group (at the γ carbon atom) could be explained by assuming a carboxyl group is attached as shown in structure X.

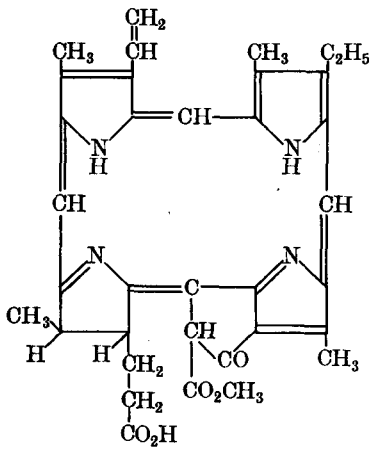
When phytyl phæophorbide-*a* (III, §6) is hydrolysed with acid, the phytyl group is removed to form phæophorbide-*a*.



phytyl phæophorbide-*a* phæophorbide-*a*

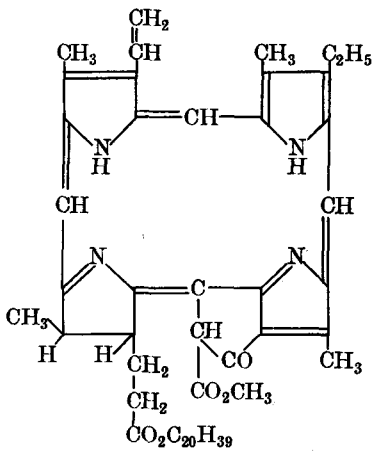
When phæophorbide-*a* is treated with hydriodic acid in acetic acid and followed by atmospheric oxidation, the product is phæophorphyrin-a_5. This, on further treatment with hydriodic acid in acetic acid, forms phylloerythrin, $C_{33}H_{34}O_3N_4$, by loss of the carbomethoxyl group; phylloerythrin has the same absorption spectrum as that of the porphyrins, and so the porphin structure is still present. Now both phæophorbide-*a* and phylloerythrin contain a keto group (as is shown by the formation of an oxime, etc.), and so when the carbomethoxyl group is hydrolysed, the elimination of carbon dioxide can be expected if the keto group is in the β-position with respect to the carboxyl group (produced on hydrolysis). Furthermore, the hydrolysis of phæophorbide-*a* with methanolic potassium hydroxide gives chlorin-*e*. In this reaction, apart from the hydrolysis of the carbomethoxyl group, the keto group is lost and a carboxyl group is introduced *without the loss of any carbon atoms*. This may be explained by assuming that this carboxyl group (the third one in chlorin-*e*) is produced by the fission of a *cyclic ketone*, and not from a lactone as suggested previously (see above). Thus a possible skeleton structure for phæophorbide-*a* is XI; if the ketone ring is opened, then the formation of X can be expected. Also, the hydrolysis of XI would produce a β-keto-acid, which can be expected to lose carbon dioxide readily to form phylloerythrin.

Phæophorbide-*a* can be reduced catalytically to its dihydro-derivative in which the keto group remains intact. This suggests the presence of a readily reducible double bond. Oxidation experiments on phæophorbide-*a* and dihydrophæophorbide-*a* showed the presence of one vinyl group in the



XII

phæophorbide-*a*

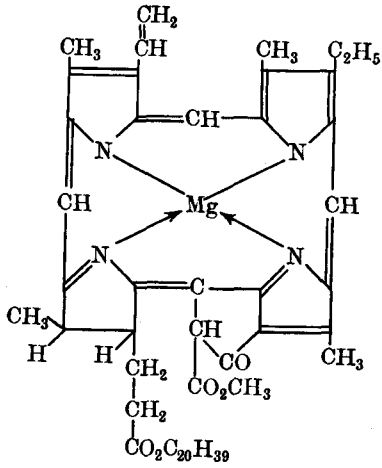


XIII

phytyl phæophorbide-*a*

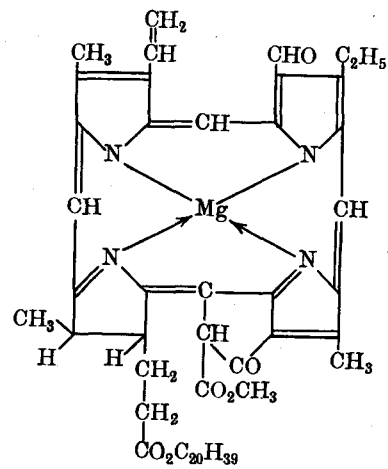
former. Furthermore, the existence of a vinyl group in the ester of chlorin-*e* was shown by the reaction with diazoacetic ester to give a *cyclo*propane derivative, which was isolated by the oxidation of the addition product (H. Fischer *et al.*, 1935; *cf.* §2a. XII). Thus one of the ethyl groups (see pyrroporphyrin, VIII) must have been a vinyl group before reduction. Further degradative and synthetic work by H. Fischer *et al.* (1934–1936) showed that phæophorbide-*a* is XII and that phytyl phæophorbide-*a* is XIII.

The replacement of the two imino hydrogen atoms in XIII by a magnesium atom would therefore give chlorophyll-*a*; this is XIV. Chlorophyll-*b* has been assigned structure XV.



XIV

chlorophyll-*a*



XV

chlorophyll-*b*

The total synthesis of chlorophyll-*a* has now been carried out by Woodward *et al.* (1960) and by Strell *et al.* (1960). Chlorin-e_6 trimethyl ester was synthesised, and since this had already been converted into chlorophyll-*a*, this constitutes a total synthesis.

A new chlorophyll, chlorobium chlorophyll, has been isolated from the culture *Chlorobium thiosulphatofilium* (Holt *et al.*, 1960).

Biosynthesis of chlorophyll. Although the steps are not clearly defined, Granick (1948, 1961) has produced evidence to show that chlorophyll is synthesised by green plants from protoporphyrin (*cf.* §4a).

PHTHALOCYANINES

§8. Preparation of the phthalocyanines. Phthalocyanines are a very important class of organic dyes and pigments; they are coloured blue to green. They were discovered by accident at the works of Scottish Dyes Ltd. in 1928. It was there observed that some lots of phthalimide, manufactured by the action of ammonia on molten phthalic anhydride in an iron vessel, were contaminated with a blue pigment. The structure and method of formation of this compound were established by Linstead and his co-workers (1934).

The phthalocyanines form metallic complexes with many metals, and the colour depends on the nature of the metal (copper, magnesium, lead, etc.); greener shades are obtained by direct chlorination or bromination. The

metal phthalocyanines are insoluble in water, and are used as pigments. They are made water-soluble by sulphonation, and these soluble salts are used as dyes.

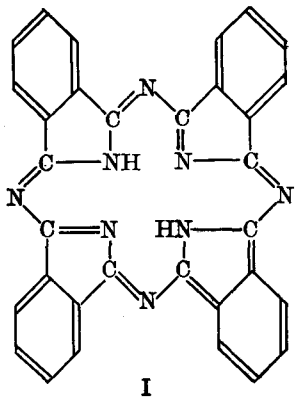
Metal phthalocyanines may be prepared as follows:

(i) By passing ammonia into molten phthalic anhydride or phthalimide in the presence of a metal salt.

(ii) By heating *o*-cyanobenzamides or phthalonitriles with metals or metallic salts.

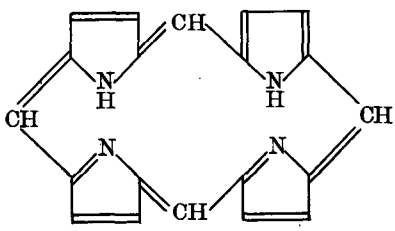
(iii) By heating phthalic anhydride or phthalimide with urea and a metallic salt, preferably in the presence of a catalyst such as boric acid.

Phthalocyanine, I, the parent substance of this group, may be prepared by heating phthalonitrile with a little triethanolamine. It can be seen from formula I that phthalocyanine contains four *iso*indole nuclei joined in a ring by means of nitrogen atoms. If we ignore the benzene nuclei, then we have four pyrrole nuclei linked by nitrogen atoms, a structure similar



I

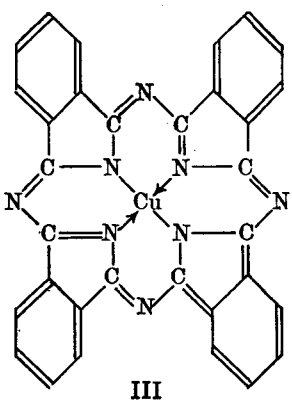
phthalocyanine



II

porphin

to the porphyrins, in which the pyrrole nuclei are linked by methyne groups (II is porphin; *cf.* §2). Both types of compounds are coloured, and both contain two imino hydrogen atoms which can be replaced to form metal complexes. Because of these similarities the phthalocyanines are often known as the tetra-azaporphyrins. The first commercial phthalocyanine pigment was **Monastral Fast Blue BS;** this is copper phthalocyanine (III).

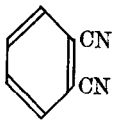


III

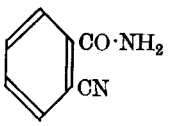
Monastral Fast Blue BS

§9. Structure of the phthalocyanines. Analysis showed that the phthalocyanines had an empirical formula $C_{32}H_{16}N_8M$, where M is a bivalent metal, *e.g.*, copper, magnesium, etc. The molecular weight determination of magnesium phthalocyanine by the ebullioscopic method with naphthalene as solvent showed that the empirical formula was also the molecular formula (Linstead *et al.*, 1934). This has been confirmed by means of X-ray measurements (Robertson, Linstead *et al.*, 1935).

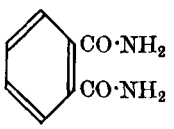
Linstead showed that the phthalocyanines can be obtained by reaction between a metal and phthalonitrile, I, *o*-cyanobenzamide, II, phthalamide, III, but *not* with, for example, terephthalonitrile, IV, homophthalonitrile, V, or *o*-xylylene dicyanide, VI. It is therefore reasonable to infer that in the



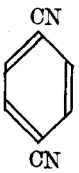
I



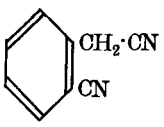
II



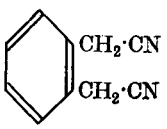
III



IV

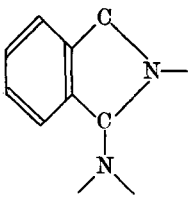


V

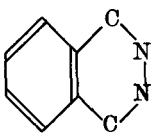


VI

formation of phthalocyanines, the two nitrile groups involved must be in the *ortho*-position. Thus there are probably four $C_8H_4N_2$ units, each having an *iso*indole structure, VII, or a phthalazine structure, VIII. VIII was



VII



VIII

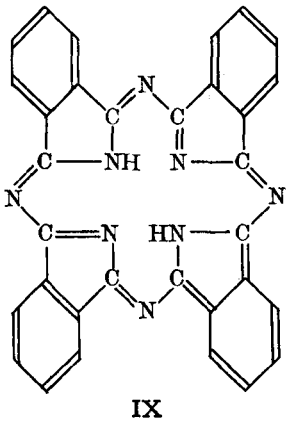
shown to be untenable since no phthalocyanine could be prepared from compounds containing this skeleton.

The oxidation of phthalocyanines with hot nitric acid, cold acid permanganate or ceric sulphate produces phthalimide and ammonium salts, the amount of phthalimide being that which would correspond to the presence of four *iso*indole units. The problem then is: How are these units joined together? The treatment of magnesium phthalocyanine with sulphuric acid replaces the magnesium atom by two hydrogen atoms.

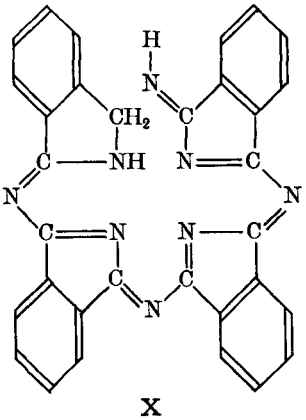
$$(C_8H_4N_2)_4Mg \xrightarrow{H_2SO_4} (C_8H_4N_2)_4H_2$$

This suggests that in metal phthalocyanines, the metal has replaced two *imino* hydrogen atoms. A reasonable structure for phthalocyanine is one in which the four *iso*indole units are joined through nitrogen atoms to form

a cyclic structure (IX). On the other hand, an open-chain structure could also be produced by joining four *iso*indole units through nitrogen atoms (X);



IX



X

in this case the molecular formula would be $(C_8H_4N_2)_4H_4$. It seems unlikely that X could be rejected on these grounds alone, since in a large molecule of this type it appears to be difficult to estimate the hydrogen with certainty (IX contains approximately 3·5 per cent. hydrogen, and X 3·9 per cent.). X, however, is unlikely, since phthalocyanine is a very stable substance; the presence of an imino group at the end of the molecule could be expected to render the compound unstable to, *e.g.*, acid reagents. Furthermore, the oxidation of phthalocyanine with ceric sulphate in dilute sulphuric acid proceeds according to the following equation (over 90 per cent. of the phthalimide has been isolated).

$$(C_8H_4N_2)_4H_2 + 7H_2O + [O] \longrightarrow 4C_8H_5O_2N + 4NH_3$$

This agrees with IX, but had the structure been X, then the molecule would have required *two* atoms of oxygen.

$$(C_8H_4N_2)_4H_4 + 6H_2O + 2[O] \longrightarrow 4C_8H_5O_2N + 4NH_3$$

Thus IX represents best the known properties of phthalocyanine. The two imino hydrogen atoms are replaceable by a bivalent metal, and the remaining two nitrogen atoms form co-ordinate links (see formula III, §8).

In metal phthalocyanines resonance is possible, and so all four nitrogen atoms linked to the metal atom would be equivalent. Phthalocyanines (with and without a central metal atom) have been examined by means of X-ray analysis (Robertson, 1936), and the results show that these compounds are large *flat* molecules with a centre of symmetry. The bond lengths of the C—N bonds indicate resonance, as do those of the benzene ring (all the lengths are equal). Robertson also showed that for nickel phthalocyanine, if the radius of the nickel atom be assumed, then the positions of the other atoms in the molecule are exactly those obtained by chemical evidence.

READING REFERENCES

Stewart and Graham, *Recent Advances in Organic Chemistry*, Longmans, Green. Vol. III (1948, 7th ed.). (i) Ch. 5. Some Natural Porphyrins and Related Compounds. (ii) Ch. 6. The Azaporphyrins.

Phthalocyanines.
(i) Linstead *et al.*, *J.C.S.*, **1934**, 1016; **1936**, 1745.
(ii) Robertson, *J.C.S.*, **1935**, 615; **1936**, 1195; **1937**, 219; **1940**, 36.
(iii) Dahlen, *Ind. Eng. Chem.*, 1939, **31**, 839.

Elderfield (Ed.), *Heterocyclic Compounds*, Wiley. Vol. I (1950). Ch. 6. Chemistry of Pyrrole and its Derivatives.

Fischer and Orth, *Die Chemie des Pyrrols*, Leipzig. Vol. II (Part I, 1937; Part II, 1940).

Gilman (Ed.), *Advanced Organic Chemistry*, Wiley (1943, 2nd ed.). (i) Ch. 16. The Chemistry of the Porphyrins. (ii) Ch. 17. Chlorophyll.

Lemberg and Legge, *Hæmatin Compounds and Bile Pigments*, Interscience Publishers (1949).

Rodd (Ed.), *Chemistry of Carbon Compounds*, Elsevier. Vol. IVB (1959). Ch. XIII. The Pyrrole Pigments.

Bentley, *The Natural Pigments*, Interscience (1960).

Maitland, Biogenetic Origin of the Pyrrole Pigments, *Quart. Reviews* (*Chem. Soc.*), 1950, **4**, 45.

Popják, Chemistry, Biochemistry, and Isotopic Techniques, Lectures, Monographs and Reports of the Royal Institute of Chemistry, 1955, No. 2.

Neuberger *et al.*, Biosynthesis of Porphyrins and Chlorophylls, *Nature*, 1961, **192**, 204.

Linstead, Discoveries among Conjugated Macrocyclic Compounds, *J.C.S.*, **1953**, 2873.

Willstätter and Stoll, *Investigations on Chlorophyll*, Science Press (1928).

Livingston, Physiochemical Aspects of Some Recent Work on Photosynthesis, *Quart. Reviews* (*Chem. Soc.*), 1960, **14**, 174.

Arnon *et al.*, Photoproduction of Hydrogen, Photofixation of Nitrogen and a Unified Concept of Photosynthesis, *Nature*, 1961, **192**, 601.