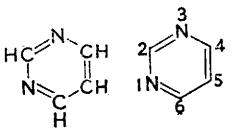
SOME ASPECTS OF PYRIMIDINE AND PURINE CHEMISTRY

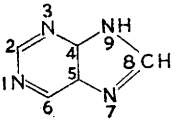
By B. LYTHGOE, M.A., PH.D.

(LECTURER IN ORGANIC CHEMISTRY, UNIVERSITY OF CAMBRIDGE)

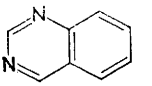
THE chemistry of pyrimidine (I), and of condensed systems containing the nucleus, such as purine (II), quinazoline (III), pteridine (IV), and the hypothetical *iso*alloxazine (V), has been developed largely because certain derivatives are important constituents of living organisms. Amongst natural pyrimidines are orotic acid (2 : 6-dihydroxypyrimidine-4-carboxylic acid), which occurs in milk, and vitamin B_1.[1] The naturally occurring purines such as uric acid, and the adenine and caffeine groups, were among the earliest investigated natural products ; [2] more recently coenzymes such as the co-dehydrogenases and co-phosphorylases have been identified as



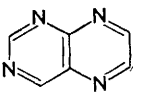
(I.)



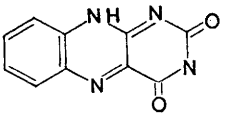
(II.)



(III.)



(IV.)



(V.)

adenine derivatives.[3] The alkaloid vasicine is a quinazoline compound, and vitamin B_2 (riboflavin) [1] contains the *iso*alloxazine system. It is only since 1940 that the chemistry of the pteridine system has become known ; [4] its derivatives occur in insect wing-pigments, and representatives such as pteroylglutamic acid (VI) have vitamin activity, and function in hæmatopoiesis. The complex cell-constituents known as nucleic acids have a biological significance scarcely inferior to that of proteins and polysaccharides ; they are built up from pyrimidine and purine compounds.

In view of the part played in cellular economy by such compounds, it is not surprising that pyrimidine derivatives should be known which have important pharmacological activity. The barbiturate hypnotics, which

[1] See H. R. Rosenberg, " The Chemistry and Physiology of the Vitamins ", New York, 1942.

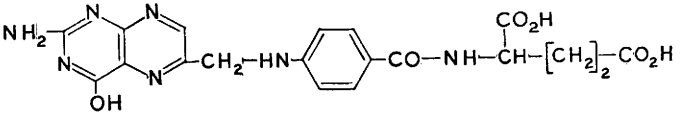
[2] See Mayer and Jacobson, " Lehrbuch der Organischen Chemie ", Berlin and Leipzig, 1923.

[3] See B. Lythgoe, *Ann. Reports*, 1945, **42**, 175.

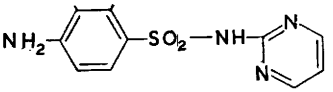
[4] See J. C. E. Simpson, *ibid.*, 1946, **43**, 250 ; M. Gates, *Chem. Reviews*, 1947, **41**, 63 ; T. D. Spies, *Ann. Rev. Biochem.*, 1947, **16**, 387.

have been known for many years, are used very widely in clinical practice. Sulphadiazine (VII) and its 4-methyl and 4 : 6-dimethyl derivatives are outstandingly valuable amongst sulphonamides [5] in the treatment of bacterial infections. During the war years an intensive study of pyrimidine derivatives was carried out with the object of obtaining new antimalarials ; some representatives with marked activity were found, though none with the outstanding value of the diguanide derivative " Paludrine ", discovered as a sequel to this work.[6] The antithyroid activity of thiouracil (6-hydroxy-2-mercaptopyrimidine) and the action of alloxan in causing diabetes through the destruction of the pancreatic islets are recently investigated topics.[7]

On the chemical side, a major part of our knowledge of the group consists of synthetic methods. This term is intended to convey not only the methods by means of which the pyrimidine nucleus can be built up, but



(VI)



(VII)

also the numerous transformations which substituent groups can undergo ; much of this knowledge has been gained in the work on the various biologically important compounds mentioned above. Apart from synthetic and biochemical work there has been little important activity in the purine field during the last twenty years. Although the systematic chemistry of the pyrimidine nucleus has never been given the attention which it deserves, sufficient information has now been accumulated to make a logical treatment possible ; the main deficiencies concern the simpler compounds ; for example, pyrimidine itself has probably never been obtained in amounts of more than a gram, and its 5-hydroxy-derivative is unknown. The practice of interpreting the behaviour of related heterocyclic systems in terms of modern theories of tautomerism and resonance has made progress in recent years, and there are signs that pyrimidine compounds are amenable to similar treatment. The way in which substituents interact with each other and with the nucleus presents some unusual features, and in hydroxy- and

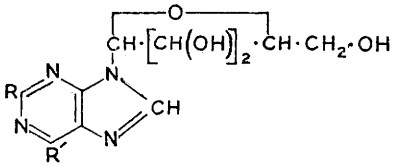
[5] See E. H. Northey, " The Sulphonamides and Allied Compounds ", London and New York, 1948.

[6] See K. C. Blanchard, *Ann. Rev. Biochem.*, 1947, **16**, 595.

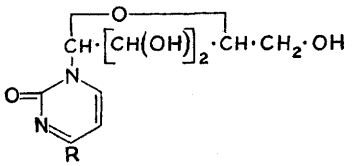
[7] See W. T. Salter, *ibid.*, 1945, **14**, 570 ; C. F. Cori and G. T. Cori, *ibid.*, 1946, **15**, 203 ; D. Stetten, *ibid.*, 1947, **16**, 136.

amino-derivatives interesting structural problems arise. These aspects will be given prominence in the present survey.

Relationship of Pyrimidines and Purines to Nucleic Acids.—Some of the recent work on pyrimidine and purine derivations owes its initiation to the biological importance of nucleic acids and related compounds. Some aspects of the biochemical investigations on nucleic acids have been collected in a recent symposium ; [8] of these only one or two can be mentioned here. O. T. Avery, C. M. McCarty, and M. McLeod [9] have shown that the type-transforming factor of pneumococci is a nucleic acid of the deoxyribose type. J. Brachet and T. Caspersson [10] independently brought forward evidence implicating nucleic acid in protein synthesis. Nucleic acids are present in combination with proteins in self-duplicating systems such as chromosomes and viruses ; [11] C. D. Darlington [12] has described the acids as having the function of " the molecular midwife of all reproductive particles ". The association of nucleic acids with proteins in these particles is especially significant in view of the fact that closely related nucleotides and dinucleotides act in conjunction with specific proteins in intact phosphorylase and dehydrogenase systems.



(VIII.)



(IX.)

The nucleic acids [13] are macromolecules ; they are polynucleotides, built up from nucleotide units in a way not yet fully understood. There are two main types, namely ribonucleic acids which are cytoplasmic constituents, and deoxyribonucleic acids ; it is the latter type which occurs in cell nuclei. The nucleotides obtained by the breakdown of both types appear to be the 3′-phosphoryl derivatives of compounds known as nucleosides ; the latter are *N*-glycosides of purine or pyrimidine derivatives. Those from ribonucleic acid contain ribose as the sugar component, and are guanosine (VIII ; R = NH_2, R′ = OH), adenosine (VIII ; R = H, R′ = NH_2), cytidine (IX ; R = NH_2), and uridine (IX ; R = OH). The nucleosides from deoxyribonucleic acid are probably similarly constituted, but the sugar component is 2-deoxyribose, and thymine (5-methyluracil) is present instead of uracil as the aglycone in one of them.

Clarification of the structures of the nucleoside units was clearly a necessary preliminary to constitutional work on the macromolecular acids. Much work directed to this end was done before 1940, using degradative

[8] Nucleic Acid. *Symposia Soc. Exp. Biol.*, 1947, Vol. I.
[9] *J. Exp. Med.*, 1944, **79**, 137. [10] Ref. 8, p. 127 ; *ibid.*, p. 207.
[11] See A. E. Mirsky, *Adv. Enzymol.*, 1943, **3**, 1.
[12] Ref. 8, p. 267. [13] See J. M. Gulland, ref. 8, p. 1.

methods which have been reviewed elsewhere.[14] Synthetic approaches [15] have recently made it possible to prepare the four ribonucleosides from compounds of known structure, and as a consequence their constitutions can now be regarded as established. The extension of similar methods to nucleotides and dinucleotides should in the future help to throw light on the way in which the nucleotide units are united in the polynucleotide molecule. It is probable that the internucleotidic linkages in the latter are of the phosphate ester type, but the molecular position linked in this way to the sugar hydroxyl at position 3′ of an adjacent nucleotide is not yet known. When this has been ascertained, the problems remaining for solution are similar to those arising in other macromolecules: whether a straight chain or branched structure is present; the relative number and disposition of the different nucleoside units; the identity of end groups; and so on. It would be digressing too far from pyrimidine and purine chemistry to consider here the various chemical and enzymatic degradation methods by which these problems are now being studied, but one aspect may be mentioned. The advances made recently in the investigation of protein structures owe much to the development of convenient methods for the separation and determination of the component amino-acids, and the solution of similar problems in the field of pyrimidines, purines, and their nucleosides would clearly be of equal value for constitutional research on nucleic acids. These problems are now being studied, and the techniques emerging seem of considerable promise. A few years ago it was assumed, as a result of earlier inexact work, that equimolecular amounts of each of the four nucleosides were present in ribonucleic acid (the "statistical tetranucleotide" concept), and, in the absence of criteria for their identification, it was impossible to say with confidence whether we should speak of one ribonucleic acid or of many. The preliminary results from the newer methods of estimation indicate that samples of acid from different sources may contain different proportions of the four nucleosides.[16] It seems, in fact, that the acids will eventually be found to have a manifold and highly organised nature.

Biogenetic Aspects.—Cellular activities are so dependent upon purine and pyrimidine derivatives that the fundamental question of their origin is provoked: what powers of synthesising them do various organisms possess, and what compounds are involved as intermediates in these biosyntheses?

Experiments using compounds marked with isotopic nitrogen (^{15}N) have now shown that in mammals and birds very considerable powers of synthesis are present. In the rat, tissue purines and pyrimidines are probably mainly of endogenous origin; thus although dietary adenine is used to some extent for nucleic acid synthesis, guanine is not used in the same way.[17]

[14] See B. Lythgoe, *Ann. Reports*, 1944, **41**, 200.

[15] B. Lythgoe and A. R. Todd, ref. 8, p. 15; A. R. Todd, *J.*, 1946, 647.

[16] H. S. Loring, G. L. Ordway, and J. G. Pierce, *J. Biol. Chem.*, 1948, **176**, 1123; E. Vischer and E. Chargaff, *ibid.*, p. 715.

[17] G. B. Brown, P. M. Roll, and A. A. Plentl, *Fed. Proc.*, 1946, **6**, 517; A. A. Plentl and R. Schoenheimer, *J. Biol. Chem.*, 1944, **153**, 203.

F. W. Barnes and R. Schoenheimer [18] have found that when $^{15}NH_3$ is administered to rats and pigeons, the marked nitrogen atom is incorporated rapidly into the tissue nucleic acids and the excreted uric acid; that is, the purines and pyrimidines, like the proteins, are in a state of continual degradation and resynthesis.

It had been suggested in the past that the ureide systems in purines such as uric acid might originate in urea, and that amino-acids such as histidine might provide a source of the iminazole ring of the purine nucleus. The work with isotopically marked materials has refuted these suggestions, and points to the view that purines are built up *in vivo* from small metabolic units. In pigeons and humans, uric acid is synthesised in part from glycine, the atoms of which are incorporated as N_7 of the glyoxaline nucleus, and C_5 and C_4 of the pyrimidine nucleus respectively; glycine is similarly utilised for the synthesis of the nucleic acid purines in yeast. We now know something of the origin of the remaining carbon atoms of uric acid; pigeons incorporate the labelled atoms of $H\cdot{}^{13}CO_2H$ and $CH_3\cdot{}^{13}CO_2H$, but not that of $^{13}CO_2$, into both the ureide systems (*i.e.*, C_2 and C_8); C_6 of the pyrimidine ring is probably derived from CO_2.[19] For the success of studies of this kind convenient synthetic methods are needed to prepare labelled pyrimidine and purine compounds from the available isotope source, and convenient degradative methods are necessary so that at the conclusion of the biochemical work the locus of entry of the isotopes into the heterocyclic molecules may be determined: an interesting example of the dependence of biochemical work on earlier chemical studies.

For securing a more detailed insight into the processes of nucleic acid biosynthesis at least two other techniques are available. One of these, the study of isolated enzyme systems, may be illustrated by the work of H. M. Kalckar.[20] He has shown that in mammalian tissue a mechanism exists for the phosphorolytic synthesis of purine nucleosides from the purine aglycones. Rat liver contains an enzyme which catalyses the reversible change

Inosine + inorganic phosphate $\rightleftharpoons$ hypoxanthine + ribose-1 phosphate

By using the reverse reaction a physiological synthesis of inosine and guanosine has been effected. A related enzyme has been isolated from muscle by S. P. Colowick and W. H. Price,[21] and it is probable that enzymes capable of synthesising pyrimidine nucleosides may also be present in mammals.

The second method attempts to explore the synthetic pathway between the small metabolic units and the nucleic acid end-products by a study of micro-organisms which lack the enzymic equipment required to carry out one or more of the necessary synthetic steps. Work of this kind is expanding rapidly at the present time, and the results obtained to date should

[18] *Ibid.*, 1943, **151**, 123.

[19] J. C. Sonne, J. M. Buchanan, and A. M. Delluva, *ibid.*, 1948, **173**, 69, 81; R. Abrams, E. Hammarsten, and D. Shemin, *ibid.*, p. 429.

[20] Ref. 8, p. 38.

[21] *Fed. Proc.*, 1946, **5**, 130.

perhaps be regarded as suggestive rather than conclusive until the complete picture emerges in sharper definition. They do, however, justify us in taking it as a working hypothesis that in nucleic acid biosynthesis there is a pathway whereby sugar-free pyrimidines and purines are first synthesised from as yet unidentified intermediates and then converted into nucleosides; in the first of these steps, and possibly in both, vitamins of the B-group are implicated as coenzymes or coenzyme precursors.

On this hypothesis the fact that certain micro-organisms require for growth purines such as adenine, or pyrimidines such as thymine or uracil, is interpreted as indicating that the equipment necessary to synthesise these aglycones has been lost, although the power to transform them into nucleosides remains. One very interesting case of this type is an induced mutant of *E. coli* which requires *p*-aminobenzoic acid for growth. This requirement can be dispensed with if an adequate supply of thymine, purines, and methionine is provided; growth under these conditions is subculturable and is not inhibited by sulphonamides.[22] In view of the competitive relationship known to exist between *p*-aminobenzoic acid and sulphonamides, and of other circumstantial evidence such as the fact that yeast grown in the presence of sulphonamides has a reduced nucleic acid content,[23] the behaviour of *E. coli* appears significant. It is suggested that *p*-aminobenzoic acid or some related compound functions as a coenzyme in the synthesis of purines and thymine; there is some evidence that the active form of *p*-aminobenzoic acid may be a compound of the pteroylglutamic acid type.[24] W. Shive and his co-workers [25] point out that if sulphonamides act by inhibiting purine synthesis, the intermediate compound whose further transformation is prevented may accumulate in the culture medium. Under such conditions *E. coli* synthesises a base, 4-aminoglyoxaline-5-carboxyamide, and it was suggested that *p*-aminobenzoic acid or a related compound functions in converting this base, or one of its close relatives, into purine compounds. A physiological synthesis of this kind would be paralleled by chemical syntheses of the pyrimidine ring of Type III, which are discussed in a later section.

Other micro-organisms are known which fail to grow satisfactorily when pyrimidines such as uracil and cytosine are provided, but respond well to the intact nucleosides uridine and cytidine; *Neurospora* mutants provide examples of this class.[26] This behaviour may mean that the organisms in question lack the power to convert the pyrimidines into their glycosides; alternatively it might be that they are derived from types in which uracil and cytosine are not normal intermediates in nucleic acid biosynthesis. Thymidine is required for growth by several *Lactobacilli*; the aglycone thymine is ineffective. In a strain of *L. lactis* which requires vitamin B_{12}

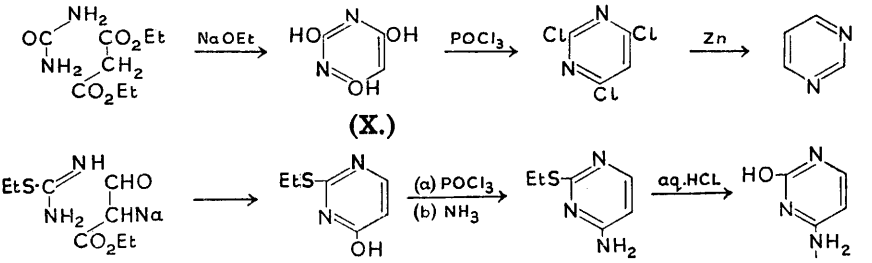
[22] J. O. Lampen, R. R. Roepke, and M. J. Jones, *J. Biol. Chem.*, 1946, **164**, 789.
[23] W. H. Schopfer, *Experientia*, 1946, **2**, 188.
[24] See D. D. Woods, *Ann. Rev. Biochem.*, 1947, **16**, 613.
[25] *J. Amer. Chem. Soc.*, 1947, **69**, 726.
[26] H. K. Mitchell and M. B. Houlahan, *Fed. Proc.*, 1947, **6**, 506; H. S. Loring and J. G. Pierce, *J. Biol. Chem.*, 1944, **153**, 61.

(the anti-pernicious anæmia factor of liver [27]) for growth, this factor can be dispensed with if adequate amounts of thymidine are supplied, and this has been interpreted as indicating that vitamin B_{12} functions as a coenzyme in the conversion of thymine into thymidine.[28]

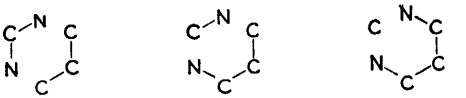
It is clear that the views tentatively expressed above may require some revision or expansion as further experimental evidence accumulates, but the methods described obviously have an important contribution to make to biogenetic studies. They hold out the promise that it may ultimately become possible to exercise some control over the activities of the cell nucleus, with important implications for the treatment of bacterial and virus-borne infections, perhaps even of cancer.

Synthetic Methods

Pyrimidines.—The earlier methods for the preparation of pyrimidine compounds are given in Mayer and Jacobson's text-book,[2] and need, therefore, no detailed discussion here. Most of them involve two separate stages, namely a synthesis of the ring-system, followed by a process in which the substituent groups are transformed into those present in the desired compound. These processes can be illustrated by the following preparations of pyrimidine itself from barbituric acid (X), and of cystosine (6-amino-2-hydroxypyrimidine) (XI) :



The principal transformations open to various substituents in the nucleus will be mentioned later ; the methods available for the formation of the nucleus can be classified into three main types, according to the distribution of nitrogen atoms in the two components used :

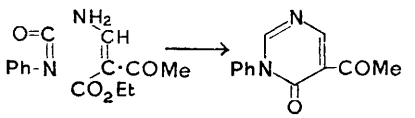


Type **I.** Type **II.** Type **III.**

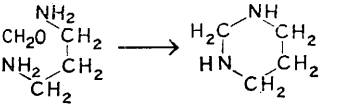
[27] E. Lester Smith, *Nature*, 1948, **162**, 144 ; E. L. Rickes *et al.*, *Science*, 1948, **107**, 396.

[28] E. E. Snell, E. Kitay, and W. S. McNutt, *J. Biol. Chem.*, 1948, **175**, 473 ; L. D. Wright, H. R. Skeggs, and J. W. Huff, *ibid.*, p. 475 ; W. Shive, J. M. Ravel, and R. E. Eakin, *J. Amer. Chem. Soc.*, 1948, **70**, 2614.

Examples of Type I are involved in the syntheses of pyrimidine, uracil, and cytosine already mentioned. The nitrogenous (left-hand) components can also be compounds such as thiourea, guanidine, and amidines. The second (right-hand) component may be ethyl malonate, ethyl cyanoacetate, malononitrile, a β-diketone, a β-keto-ester, or an $\alpha\beta$-unsaturated ketone. This type of synthesis was, and remains, the most versatile of all the methods. An example of Type II is the preparation of the derivative (XII) from phenyl *iso*cyanate and aminomethyleneacetoacetic ester; none of the methods of this type achieved much practical importance. Type III is illustrated by the formation of hexahydropyrimidine (XIII) from formaldehyde and 1 : 3-diaminopropane; such syntheses also had little practical value.



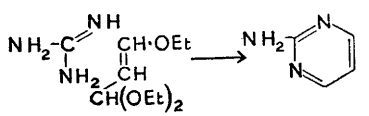
(XII.)



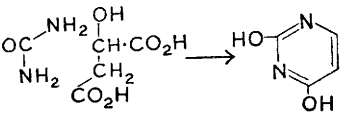
(XIII.)

These three types of synthesis have all been used and extended in more recent work and have proved adequate for most needs. No revolutionary methods have been introduced, but it is possible that in the future methods for the synthesis of hydropyrimidines [29] may play a more important part than formerly; their dehydrogenation to true pyrimidines has not yet been given much attention. Some of the more useful extensions of the general methods will now be mentioned.

The value of Type I methods has been increased by the use of components containing aldehyde groups or their equivalent. Such components



(XIV.)



(XV.)

allow a direct synthesis of derivatives containing unsubstituted 4 or 6 positions, which formerly had to be prepared by indirect methods. Components of the type mentioned include the nitro- and chloro-derivatives of malondialdehyde, which give 5-nitro- and 5-chloro-pyrimidines; [30] β-ethoxyacraldehyde acetal, which serves in place of malondialdehyde for the preparation of 2-aminopyrimidine (XIV),[31] and malic acid and its methyl derivative, which are used in the presence of fuming sulphuric acid and give direct and

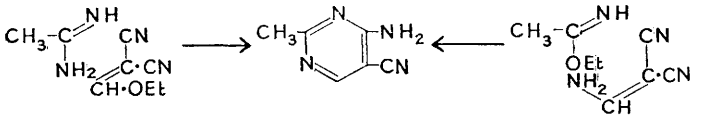
[29] Cf. M. Senkus, *J. Amer. Chem. Soc.*, 1946, **68**, 1611.

[30] W. J. Hale and H. C. Brill, *ibid.*, 1912, **34**, 82; R. O. Roblin, P. S. Winnek, and J. P. English, *ibid.*, 1942, **64**, 567.

[31] R. W. Price and A. Moos, *ibid.*, 1945, **67**, 207.

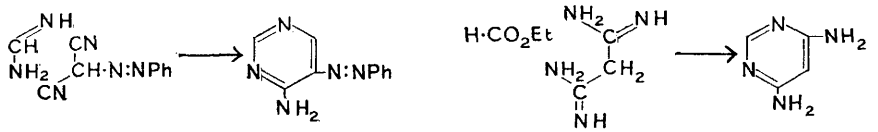
easy access to uracil (XV), *iso*cytosine (2-amino-6-hydroxypyrimidine), and thymine.[32]

Syntheses of Type I have been used extensively in the preparation of pyrimidine compounds related to vitamin B_1; these compounds belong to the class in which position 6 is unsubstituted, and accordingly ethoxymethylene derivatives were found useful as starting materials. Thus the important nitrile (XVI), which is used in F. Bergel and A. R. Todd's synthesis of the vitamin,[33] was prepared from acetamidine as shown below.[34] It has also been obtained by a synthesis of Type II from acetiminoether and aminomethylenemalononitrile ; [35] this represents one of the few important modern applications of this type of synthesis.



(XVI) (XVII)

Recently, formamidine has been introduced as a component for syntheses of Type I ; its interest lies in the access which it gives to pyrimidine derivatives containing an unsubstituted position 2, which are otherwise difficult to prepare. 4 : 6-Dihydroxypyrimidine can be obtained with its aid, but not 4 : 6-diaminopyrimidine, since with malononitrile the initial reaction takes place at the reactive methylene group ; the resulting aminomethylene-



(XVIII) (XIX)

malononitrile (XVII) then undergoes a Type II reaction with a second molecule of formamidine to give 4-amino-5-cyanopyrimidine.[36] With benzeneazomalononitrile, formamidine condenses normally, and the resulting 4 : 6-diamino-5-benzeneazopyrimidine (XVIII) has been used as the starting material for a very convenient synthesis of adenine (p. 192). 4 : 6-Diaminopyrimidine (XIX), formerly obtained by indirect methods from barbituric acid, is a compound of value for the synthesis of adenine nucleosides. It

[32] D. Davidson and O. Baudisch, *ibid.*, 1926, **48**, 2379 ; R. O. Roblin *et al.*, *ibid.*, 1940, **62**, 2002 ; 1946, **68**, 2339.

[33] A. R. Todd and F. Bergel, *J.*, 1937, 364.

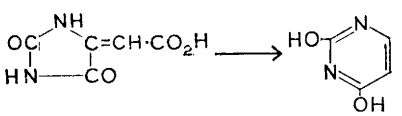
[34] R. Grewe, *Z. physiol. Chem.*, 1936, **242**, 89.

[35] O. Hromatka, D.R.-P., 670,635.

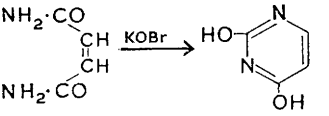
[36] G. W. Kenner, B. Lythgoe, A. R. Todd, and A. Topham, *J.*, 1943, 388 ; J. Baddiley, B. Lythgoe, and A. R. Todd, *ibid.*, p. 386.

has now become readily accessible by a synthesis of Type III in which ethyl formate and malondiamidine are condensed together.[37]

More direct methods of preparing compounds such as sulphamethazine are now available. Sulphamethazine was first synthesised [38] by a procedure involving interaction of 2-aminopyrimidine with an arylsulphonyl chloride. It has since been prepared directly by a Type I synthesis from sulphaguanidine and acetylacetone.[39]



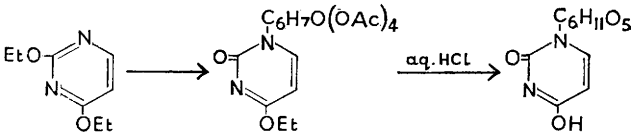
(XX.)



(XXI.)

Finally, two rearrangement reactions for the preparation of uracil derivatives deserve mention. The hydantoin (XX), which is best prepared from the product of a Wöhler synthesis on aspartic acid, undergoes ring-enlargement on treatment with hot alkali, giving orotic acid.[40] The preparation of uracil (XXI) by a Hofmann rearrangement of maleic diamide is not important as such, but it forms the basis of a method for the synthesis of condensed pyrimidine systems to be mentioned later.

Pyrimidine Nucleosides.—In 1930 G. E. Hilbert and T. B. Johnson [41] reported a preparation of 3-glucosidouracil by heating acetobromoglucose and 2 : 6-diethoxypyrimidine, followed by removal of the protecting acetyl and ethoxy-groups by acid hydrolysis :



A number of pyranose analogues of uridine and cytidine have been obtained by this method ; the preparation of one of the natural nucleosides awaited only a method for the preparation of the furanose acetobromo-compound of D-ribose. This has recently been provided by G. A. Howard, B. Lythgoe, and A. R. Todd,[42] whose synthesis of cytidine is shown on page 191 :

[37] G. W. Kenner, B. Lythgoe, A. R. Todd, and A. Topham, *J.*, 1943, 574.

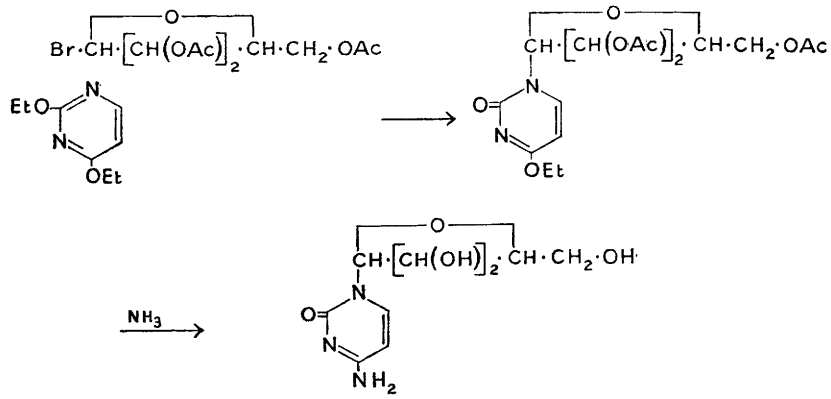
[38] W. T. Caldwell, E. C. Kornfeld, and C. K. Donell, *J. Amer. Chem. Soc.*, 1941, **63**, 2188.

[39] F. L. Rose and G. Swain, *J.*, 1945, 689.

[40] H. K. Mitchell and J. F. Nyc, *J. Amer. Chem. Soc.*, 1947, **65**, 1382.

[41] *Ibid.*, 1930, **52**, 2001 ; G. E. Hilbert and E. F. Jansen, *ibid.*, 1936, **58**, 60 ; G. E. Hilbert, *ibid.*, 1937, **59**, 330.

[42] *J.*, 1947, 1052.



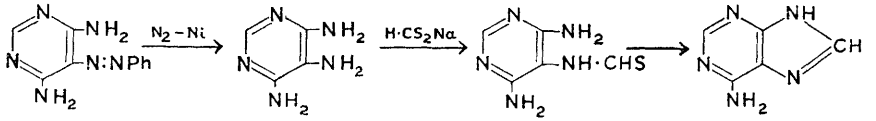
Purine Derivatives and Purine Nucleosides.—General methods of purine synthesis have undergone some improvements as a result of the interest aroused by the problems involved in the synthesis of purine nucleosides.

(a) *From* 4 : 5-*diaminopyrimidines.* W. Traube's [43] method of synthesising purine derivatives from these intermediates has long been regarded as the most valuable and general method for obtaining such compounds. It will be recalled that in this method a 4-aminopyrimidine is converted through its 5-nitroso-derivative into the 4 : 5-diaminopyrimidine, which is then cyclised by treatment with formic acid at elevated temperatures. One drawback to this method was the inaccessibility of the starting materials required for the preparation of purines unsubstituted in position 2; thus adenine had to be obtained by an unsatisfactory desulphurisation procedure from 2-mercaptoadenine.

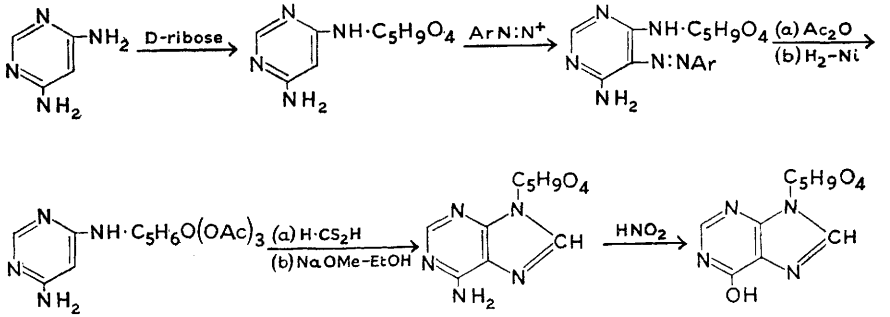
An examination of this and other steps in Traube's method has been made, the object of which was to adapt the latter for the synthesis of adenine and hypoxanthine glycosides; for this purpose mild procedures would be imperative in order to avoid the risk of hydrolysing sensitive intermediates. As a result, the following improvements have been introduced. (i) By methods indicated in an earlier section pyrimidines bearing a hydrogen atom at position 2 have become accessible. (ii) The value of 2-methylthiopyrimidines for purine synthesis has been demonstrated; after completion of the synthesis the methylthio-group can be replaced by hydrogen by means of Raney nickel containing adsorbed hydrogen.[44] (iii) A very satisfactory alternative to the introduction of the 5-amino-group by nitrosation and reduction is to employ coupling with a diazonium compound; the 5-arylazo-group introduced in this way is readily reduced to a 5-amino-group.[45] (iv) 4 : 5-Diamino-compounds are most easily cyclised to purines by treatment with sodium dithioformate to give 4-amino-5-thioformamido-compounds, which lose hydrogen sulphide readily to give the required purine. When in this cyclisation amino- and alkylamino-groups at 4 and 6 compete

[43] *Annalen*, 1904, **331**, 64; cf. C. O. Johns, *J. Biol. Chem.*, 1911, **9**, 161.
[44] G. A. Howard, B. Lythgoe, and A. R. Todd, *J.*, 1945, 556.
[45] B. Lythgoe, A. R. Todd, and A. Topham, *J.*, 1944, 315.

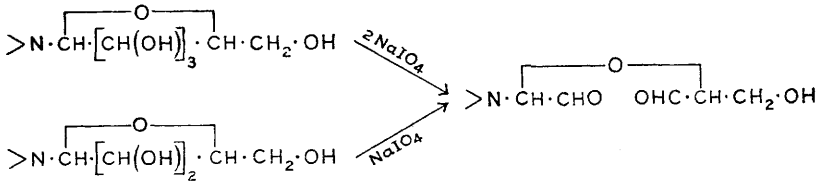
for reaction with the thioformamido-group, the superior reactivity of the alkylamino-group leads to the formation of a 9-alkyladenine derivative.[46] One outcome of this work was the convenient synthesis of adenine [47] shown below, which starts from 4 : 6-diamino-5-benzeneazopyrimidine (p. 189). It has been used for the preparation of adenine containing isotopic nitrogen.[17]



B. Lythgoe and A. R. Todd [15] have reviewed the way in which the experience gained in the above work has been applied to the synthesis of adenine and hypoxanthine nucleosides. The following example of the preparation of the pyranose analogues of adenosine and inosine illustrates the methods used, although a number of variations in the general procedure, each possessing advantages in particular cases, are possible.



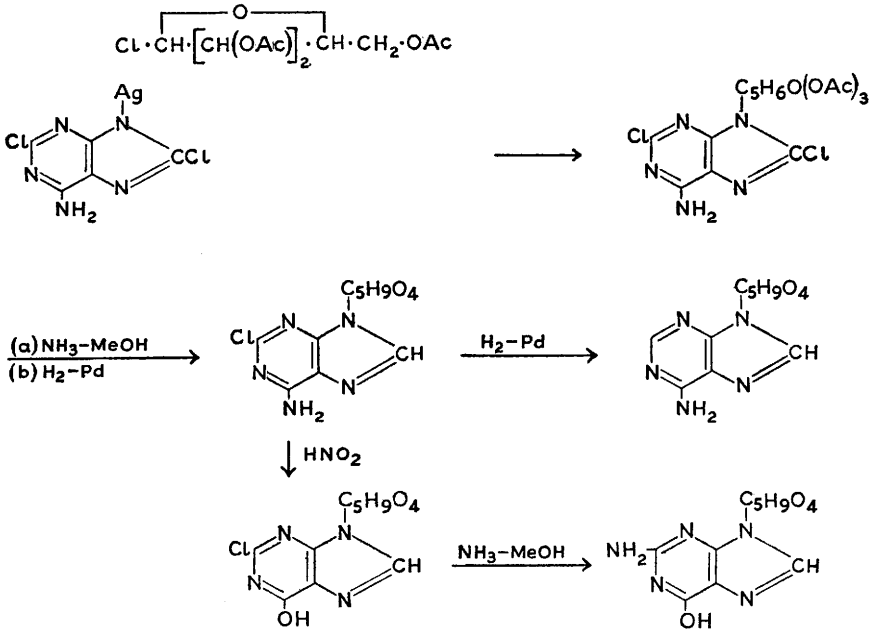
A particularly valuable feature of this method of synthesis is that the product must contain the sugar residue at N_9 of the purine system and not at N_7. Since 9-D-mannopyranosidoadenine prepared in this way is degraded by periodate to the same dialdehyde that is obtained from the degradation of adenosine, a direct chemical proof is available that the latter is a 9-glycoside :



This unambiguous synthesis of 9-glycosides can be extended to compounds of the furanose series, but so far no naturally occurring compound has been

[46] J. Baddiley, B. Lythgoe, D. McNeil, and A. R. Todd, *J.*, 1943, 383.
[47] J. Baddiley, B. Lythgoe, and A. R. Todd, *J.*, 1943, 386.

obtained by its application. However, by making use of the method of preparation of acetohalogenoribofuranoses employed in the synthesis of cytidine mentioned earlier, and by applying the method of E. Fischer and B. Helferich [48] for the introduction of a sugar residue into a preformed purine skeleton, the naturally occurring purine nucleosides adenosine and guanosine have been synthesised.[49] The method used is shown below.

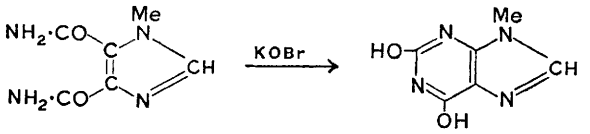


The glycosides prepared in this way, *e.g.*, adenine glucopyranoside, are undoubtedly 9-glycosides, since the same adenine glucopyranoside is obtained by the unambiguous synthetic route mentioned above. This second method of synthesis has the advantage that it allows the correlation of the glycosidic configurations in guanosine and adenosine; these and the natural pyrimidine ribonucleosides are now known to belong to the β-series.

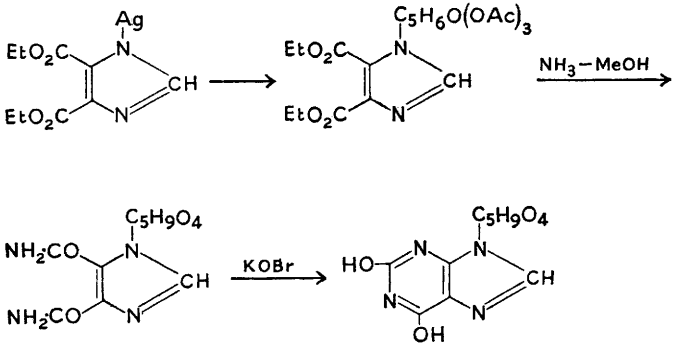
(b) *Syntheses from glyoxaline derivatives.* The purine syntheses discussed above start with the pyrimidine ring preformed and proceed to build up the glyoxaline nucleus. This order of events can be reversed, as first shown by J. Sarasin and E. Wegmann in 1924; [50] they used a synthesis of Type III for building up the pyrimidine nucleus on to a preformed glyoxaline nucleus, so obtaining 7-methylxanthine (xanthine is 2 : 6-dihydroxypurine). Xanthine itself has been obtained similarly by a synthesis of the pyrimidine nucleus belonging to Type II,[51] and very recently, R. H. Baxter and F. S.

[48] *Ber.*, 1914, **47**, 210.
[49] J. Davoll, B. Lythgoe, and A. R. Todd, *J.*, 1948, 967; *idem*, in the press.
[50] *Helv. Chim. Acta*, 1924, **7**, 713.
[51] W. E. Allsebrook, J. M. Gulland, and L. F. Story, *J.*, 1942, 232.

Spring [52] have employed the Hofmann rearrangement method (p. 190) to prepare derivatives of xanthine and alloxazine; derivatives of the pteridine system had been obtained in the same way earlier. Their methods can be illustrated by the following preparation of 9-methylxanthine:



The fact that this cyclisation leads to 9- and not to 7-substituted xanthines permits its application to the synthesis of 9-glycosidoxanthines; thus by using the preparation of acetochlororibofuranose mentioned earlier, the nucleoside xanthosine, identical with material from the deamination of natural guanosine, was obtained: [53]



The synthetic methods mentioned in this and the previous section clearly represent a considerable advance, and in the near future we may expect to see this work extended to clarify the structures of the naturally occurring purine deoxyribosides.

The Chemical Behaviour of Pyrimidines

Since the behaviour of pyrimidine derivatives has not been reviewed for many years, the following pages will be devoted to mentioning briefly the more important transformations to which the different nuclear substituents can be subjected (since these play such an important part in preparative work in the series), and to a discussion of the various theoretical points of interest.

Three prominent features of pyrimidine chemistry are:

(i) In simple derivatives, containing alkyl, aryl, or nitro-groups, or halogen atoms, but no hydroxy- or amino-groups, the nucleus has aromatic character, and behaves like that of pyridine.

[52] *J.*, 1945, 229; 232; 1947, 378.

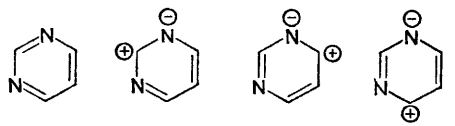
[53] G. A. Howard, A. C. McLean, G. T. Newbold, F. S. Spring, and A. R. Todd, in the press.

(ii) Nuclear substituents vary in their behaviour according to the position which they occupy. At position 5 the properties of a group can be loosely described as similar to those which it normally possesses when attached to an aromatic nucleus ; at 2, 4, and 6 marked deviations from the normal behaviour are observed. The contrast is parallel to that between β-substituted pyridines and their α- and γ-isomers.

(iii) The aromatic behaviour mentioned in (i) diminishes progressively as hydroxy- or amino-groups are introduced into positions 2, 4, and 6. This effect is seen in uracil and barbituric acid, into which substituents are readily introduced at position 5, even by mild reagents such as diazonium compounds ; simple pyrimidines such as those mentioned in (i) appear to be very resistant to electrophilic substitutions. As the simpler compounds are much less well known than the highly hydroxylated or aminated members, a rather distorted impression of pyrimidine chemistry has grown up, much as if the behaviour of benzene were known only through the reactions of compounds like phloroglucinol. The groups which give rise to this atypical behaviour (OH, SH, NH_2) have been termed, somewhat loosely, " tautomeric " substituents ; the structural problems arising when they are present will be discussed later.

Each of these three features will now be considered separately in more detail.

The Analogy with Pyridine.—The pyrimidine nucleus can be regarded as a resonance hybrid of the structure (XXII—XXV) with four equivalent

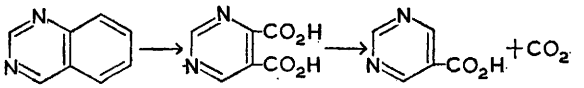


(XXII) (XXIII) (XXIV) (XXV)

structures. In the classical structure (XXII) the nitrogen atoms are doubly linked, and their electronic effects, which are the same as that of the doubly-linked nitrogen in the pyridine nucleus, co-operate since the key atoms are situated 1 : 3 to each other. Pyrimidine would thus be expected to have aromatic characteristics similar to those of pyridine or nitrobenzene,[54] with the differences from benzene even more strongly marked. There should be a marked electron-deficiency at positions 2, 4, and 6, and a similar, though smaller, deficiency at 5, which, although insulated from the ring-nitrogens, will be affected by induction. The general electron-deficiency is borne out by the fact that when quinazoline is oxidised it is the pyrimidine nucleus rather than the benzene nucleus which survives ; the resulting 4 : 5-dicarboxylic acid gives pyrimidine-5-carboxylic acid on being heated.

[54] For a discussion of the analogy between pyridine and nitrobenzene, see T. W. J. Taylor and W. Baker, " Sidgwick's Organic Chemistry of Nitrogen ", Oxford, 1937, Chap. XVIII.

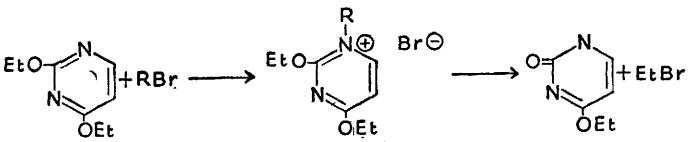
These changes are analogous to those whereby quinoline is degraded to nicotinic acid.



It is the stability of the nucleus towards oxidising agents which allows pyrimidinecarboxylic acids to be prepared by oxidising alkylpyrimidines. On the other hand such stability is often accompanied in aromatic nuclei by lability to reducing agents ; the ready reduction of pyridine is well known. Pyrimidine has not been reduced catalytically, but 4-methylpyrimidine is reduced by sodium and alcohol with the formation of 1 : 3-diaminobutane. This is presumably due to the instability under these conditions of the hexahydro-compound first formed, which contains a carbon atom (C_2) directly united to two nitrogen atoms.

Another effect of the electron deficiency at positions 2, 4, and 6 is to encourage their attack by nucleophilic reagents. Thus 4-methylpyrimidine can be aminated with sodamide (the Tschitschibabin reaction, familiar in the pyridine series), giving 2-amino- and 2 : 6-diamino-derivatives.[55] Nuclear substitution by electrophilic reagents, such as nitric acid and halogens, on the other hand, would be expected to be difficult, as it is in pyridine, and indeed, substitutions of this kind have only been recorded where one or more " tautomeric " substituents are present in the nucleus ; the discussion of such reactions is accordingly deferred till later in this review. When they do take place, the substituent is invariably introduced at position 5, the position where the electron density is least reduced by the ring-nitrogens. It would be interesting to know at what positions radical substituting agents attack the nucleus, but such substitutions have not yet been attempted.

One of the most characteristic reactions in heterocyclic compounds similar to pyridine is the conversion of their quaternary compounds into the dihydro-derivatives known as pseudo-bases, a reaction which reflects the susceptibility of the positions α or γ to the positively charged nitrogen atom to attack by negative ions. The quaternary salts of the pyrimidine series are not well known, but a phenomenon analogous to pseudo-base formation exists in the behaviour of 2 : 6-diethoxypyrimidine with alkyl and glycosyl halides.[41] These compounds react with the elimination of ethyl halide and the formation of *N*-substituted ketoethoxydihydropyrimidines, a reaction clearly due to the attack at position 2 by the halide ion of the quaternary pyrimidinium compounds (XXVI) :



(XXVI.)

As already mentioned, these analogies with pyridine lose some of their

[55] E. Ochiai and M. Karii, *J. Pharm. Soc. Japan*, 1939, **59**, 18.

validity where two or more hydroxyl or amino-groups are present in the pyrimidine nucleus ; thus in uracil and barbituric acid the nuclear stability towards oxidising agents is completely lost. Vestiges of the behaviour of the parent compound, however, remain ; the 2, 4, and 6 positions retain their electrophilic character, as shown by the reactivity of halogen atoms located there ; and substitution by electrophilic reagents is still confined to position 5.

Properties of Substituent Groups.—In this section the properties associated with some of the substituent groups will be surveyed with the object of bringing out the contrast between their behaviour at position 5 and that at 2, 4, and 6, which was mentioned on p. 195. Since the behaviour at position 5 is of the kind normally found in groups attached to aromatic nuclei, no special comment is needed concerning them. Groups at positions 2, 4, and 6 show, in general, an abnormal or reactive behaviour which is due to resonance with or within the nucleus. There is one aspect of this reactive behaviour which has very important practical consequences ; the replaceability of groups such as Cl, OEt, and SEt, to mention only the more important, is used extensively in the preparation of those substituted pyrimidines which cannot be obtained directly by building up the pyrimidine nucleus. Chloro-, alkoxy- and alkylthio-compounds therefore occupy a key position in preparative operations, for which they are well fitted by virtue of their ready accessibility. Chloro-compounds are obtained by heating the corresponding hydroxy-compounds with phosphoryl chloride ; it has recently been found that where this reaction is difficult it can often be promoted by the addition of dimethylaniline.[56] Alkylthio-derivatives are mainly of value in cases where the alkylthio-group is present at position 2. Such compounds can be obtained in two ways ; *S*-alkylthioureas can be condensed with a second component in a synthesis of Type I ; or thiourea can be used, and the resulting 2-thiol compound afterwards subjected to *S*-alkylation. Both these methods are convenient, since Type I syntheses with thiourea or its *S*-alkyl compounds proceed very readily in comparison with those where the less basic urea is employed.

A few examples will make clear the use of the chloro-, alkoxy- and alkylthio-compounds ; for others the reader is referred to a review by T. B. Johnson and D. Hahn in 1933,[57] and to the more recent literature citations. The most important replacements of chlorine atoms are by amino- or substituted amino-groups, usually by heating the chloro-compound with alcoholic ammonia or with the appropriate amine,[58] and by alkoxy-groups, by the action of sodium alkoxides. Cytosine [59] is prepared from uracil by the following route. Uracil $\longrightarrow$ 2 : 6-dichloropyrimidine $\longrightarrow$ (6-chloro-2-aminopyrimidine) + 2-chloro-6-aminopyrimidine $\longrightarrow$ 6-amino-2-ethoxypyrimidine $\longrightarrow$ 6-amino-2-hydroxypyrimidine. The syntheses of vitamin B_1

[56] Ref. 36 ; J. Baddiley and A. Topham, *J.*, 1944, 678.

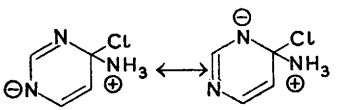
[57] *Chem. Reviews*, 1933, **13**, 193.

[58] F. H. S. Curd and F. L. Rose, *J.*, 1946, 343 ; J. P. English *et al.*, *J. Amer. Chem. Soc.*, 1946, **68**, 1039.

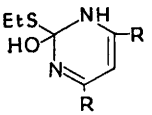
[59] G. E. Hilbert and T. B. Johnson, *ibid.*, 1930, **52**, 1154.

due to R. R. Williams and his co-workers [60] and to H. Andersag and K. Westphal [61] make use of ammonolyses of 4-chloropyrimidines. Ethoxyl groups are replaced by hydroxyl groups by treatment with dilute acids, or by amino-groups by the action of hot alcoholic ammonia ; these reactions are made use of in Hilbert's method for the synthesis of pyrimidine nucleosides (p. 190). Ethylthio-groups can similarly be replaced by hydroxyl, amino-, and substituted amino-groups.[62]

All these replacements, and many others, such as the exchange of amino-groups for hydroxyl groups on boiling with dilute acids, fall into a common pattern. In their reactive behaviour halogen atoms at positions 2, 4, and 6 are analogous to those in 1-chloro-2 : 4-dinitrobenzene or α- and γ-halogenopyridines. The halogen replacements are due to the electrophilic character of the nuclear carbons at these positions (p. 196) and their attack by the nucleophilic reagents (NH_3, $OEt^{\ominus}$, etc.) is facilitated by the low energy of transition states such as (XXVII). A similar explanation applies to the



(XXVII)



(XXVIII)

replacement of OEt and SEt groups in the presence of strong bases such as ammonia. It has been mentioned, however, that certain replacements, such as those of OEt, SEt, and NH_2 groups by OH require acid as the reagent ; this presumably acts by permitting a neutral transition state such as (XXVIII). A similar effect has been found in the replacement of halogen atoms by feebly basic arylamines ; the reaction can be conducted more easily in the presence of acid.[63] That this is not the case with more strongly basic amines is no doubt due to the latter undergoing salt formation.

(a) *Alkylpyrimidines.* The most striking property of methyl groups at positions 2, 4, and 6 is their reaction with benzaldehyde in the presence of zinc chloride to give styrylpyrimidines [64] (*e.g.*, XXIX) ; 5-methyl derivatives do not behave in this way. These reactions recall those of the methyl groups in α- and γ-methylpyridines, and no doubt the reactive entity is a zinc chloride complex, such as (XXX). In the light of the behaviour of the analogous methyl group in acetophenone it is perhaps not surprising that such methyl groups are capable of side-chain bromination [65] and

[60] *J. Amer. Chem. Soc.*, 1937, **59**, 1052. [61] *Ber.*, 1937, **70**, 2035.

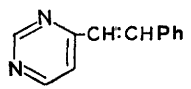
[62] Y. F. Chi and Y. S. Kao, *J. Amer. Chem. Soc.*, 1936, **58**, 772 ; F. H. Case and A. J. Hill, *ibid.*, 1929, **51**, 1590 ; F. H. S. Curd and F. L. Rose, *J.*, 1946, 343.

[63] C. K. Banks, *J. Amer. Chem. Soc.*, 1944, **66**, 1127 ; F. H. S. Curd *et al.*, *J.*, 1946, 343, 370.

[64] S. Gabriel and J. Colman, *Ber.*, 1903, **36**, 3383 ; E. Ochiai and M. Yanai, *J. Pharm. Soc. Japan*, 1938, **58**, 397.

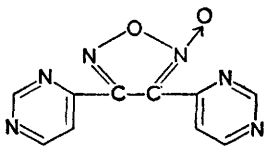
[65] C. C. Price, N. J. Leonard, and R. L. Whittle, *J. Org. Chem.*, 1945, **10**, 327.

the formation of pyrimidylfuroxans (XXXI) on treatment with nitric acid.



(XXIX)

(XXX)



(XXXI)

(b) *Halogenopyrimidines.* Members in which the halogen atom is present in position 5 have the expected inertness, paralleling that of β-chloropyridine or bromobenzene. The halogen can be replaced catalytically by hydrogen,[66] and exchange for a cyano-group on heating with cuprous cyanide in quinoline has been observed,[67] but exchange for amino- or alkoxy-groups is not generally practicable. The presence of two " tautomeric " groups in the nucleus modifies this behaviour ; thus 5-bromouracil and 5-bromocytosine can with some difficulty undergo ammonolysis to give 5-amino-derivatives.

The ready reactivity of chlorine atoms at the remaining positions has already been discussed. 4-Halogenopyrimidines react with phenylmagnesium bromide, or with benzene in the presence of aluminium chloride, to give 4-phenyl derivatives,[68] and they have been used in Ullmann reactions to give dipyrimidyl compounds.[69] The reductive replacement of halogen atoms at positions 2, 4, and 6 is very frequently used for preparing pyrimidines with unsubstituted positions. Recently this has been effected by catalytic methods ;[70] examples are given in the nucleoside syntheses mentioned on p. 193.

(c) *Nitropyrimidines.* 5-Nitropyrimidines are the only known representatives of this class ; in early work they were used extensively for the preparation of 5-amino-compounds. The nitro-group exerts an influence on groups situated *o-* and *p-* to it, just as in the benzene series. Thus the reactivity of the 6-chlorine atom in 2 : 6-dichloro-5-nitropyrimidine is so enhanced that it is replaced rapidly by ammonia at room temperature, and the amino-group of 5-nitro-2-aminopyrimidine is removed as ammonia on heating with aqueous alkali.

(d) *2-Mercapto- and 2-alkylthio-compounds.* The importance which these compounds have in preparative work, and some of their replacement reactions, have been noted earlier. It is also possible in some cases to

[66] W. Huber and H. A. Hölscher, *Ber.*, 1938, **71**, 87 ; J. P. English *et al.*, *J. Amer. Chem. Soc.*, 1946, **68**, 1039.

[67] C. C. Price, N. J. Leonard, and R. H. Reitsema, *ibid.*, p. 766.

[68] M. Anker and A. H. Cook, *J.*, 1941, 323 ; E. Ochiai, *J. Pharm. Soc. Japan*, 1940, **60**, 164.

[69] M. Yanai and T. Naito, *ibid.*, 1941, **61**, 99.

[70] R. O. Roblin, J. H. Williams, P. S. Winnek, and J. P. English, *J. Amer. Chem. Soc.*, 1940, **62**, 2002 ; H. J. Backer and A. B. Grevenstuk, *Rec. Trav. chim.*, 1942, **61**, 291.

replace alkylthio-groups at position 2 by the action of chlorine in aqueous solution,[71] but this method has little preparative value.

A useful reaction of the thiol group at position 2 in both pyrimidines and purines is the replacement by hydrogen which it undergoes when treated with hydrogen peroxide ; [72] it is used in Traube's synthesis of adenine and hypoxanthine, and is interesting enough to require comment, although it has never been examined very closely. It undoubtedly occurs by oxidation of the thiol to a sulphonic acid group, which is then split off by hydrolysis as sulphuric acid ; similar behaviour is observed in 2-mercaptoglyoxalines. The structural requirement is clearly adjacence of the thiol group to a doubly linked heterocyclic nitrogen, but the question arises as to why the sulphonic acids which are first formed should undergo such ready hydrolysis ; the electrophilic character which is the outstanding feature of such nuclear positions seems inadequate to provide an explanation. The answer is probably to be found in the observation of D. Ll. Hammick and his collaborators [73] that the ready decarboxylation of pyridine- and quinoline-2-carboxylic acids is due to the intervention of a rather stable pyridyl or quinolyl anion in which the negative charge is carried at position 2 of these heterocyclic systems ; they point out that the stability of such anions is probably due to their cyanide-ion-like structure. Positions 2, 4, and 6 in the pyrimidine nucleus are fully analogous to the 2-position in pyridine, and the ready decarboxylation of a carboxyl group at position 4 compared with that at position 5 has already been noted. It seems probable that the ready loss of the sulphonic acid group from position 2 in the pyrimidine nucleus is due to the operation of similar factors.

A desulphurisation method which is more convenient than that just mentioned applies to 2-alkylthio-compounds, which can be converted into derivatives with an unsubstituted position 2 by treatment with Raney nickel containing adsorbed hydrogen. This method which is an application of a procedure first used by J. Bougeault,[74] has recently been used to prepare adenine-9-glycosides from their 2-methylthio-derivatives.

Hydroxy- and Amino-compounds.—(a) *Hydroxypyrimidines.* Neither 5-hydroxypyrimidine nor any well-authenticated alkyl or aryl homologue is known. These compounds should like β-hydroxypyridine, have well-defined phenolic properties ; the only evidence supporting this view is the colour given by *iso*uracil [75] (2 : 5-dihydroxypyrimidine) with ferric chloride.

Hydroxypyrimidines bearing up to three hydroxyl groups at positions 2, 4, and 6 are well known ; they include uracil, thymine, and barbituric acid. Their properties will be considered later ; here a few features of their behaviour may be noted which throw doubt on the view that their structures are those of simple hydroxyl derivatives. They show no phenolic

[71] J. M. Sprague and T. B. Johnson, *J. Amer. Chem. Soc.*, 1935, **57**, 2252 ; 1936, **58**, 423 ; 1938, **60**, 1622. [72] H. L. Wheeler, *J. Biol. Chem.*, 1907, **3**, 285.

[73] *Nature*, 1948, **162**, 73 ; *J.*, 1939, 809 ; K. Mislow, *J. Amer. Chem. Soc.*, 1947, **69**, 2559.

[74] J. Bougeault, E. Cattelain, and P. Chabrier, *Bull. Soc. chim.*, 1939, **6**, 34 ; 1940, **7**, 781 ; *Compt. rend.*, 1939, **208**, 657.

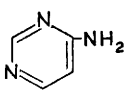
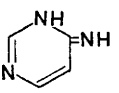
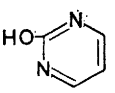
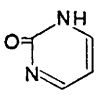
[75] J. Tafel and P. A. Houseman, *Ber.*, 1907, **40**, 3743.

behaviour. The " hydroxyl groups " are replaced by chlorine atoms on heating with phosphoryl chloride. They are difficult to acylate, and their acyl derivatives are readily hydrolysed. The action of alkylating agents on them varies with the compound and the reagent used. Thus uracil is converted into its N^1 : N^3-dimethyl derivative either by methyl sulphate and alkali, or by ethereal diazomethane [76]; 6-hydroxy-2-phenacylthio-4-methylpyrimidine gives the corresponding *O*-ether on treatment with sodium ethoxide and ethyl bromide. The *O*-ethers have the interesting property of rearranging on heating; 2 : 6-dimethoxypyrimidine gives N^1 : N^3-dimethyl-uracil.

(b) *Aminopyrimidines.* 5-Aminopyrimidine and its 4-methyl derivative have been prepared, but not examined closely. We are thus not in a position to say what properties are associated with an isolated 5-amino-group. Such knowledge of 5-amino-derivatives as exists comes from the well-known 4 : 5-diaminopyrimidines, which are important as intermediates from which purine and pteridine derivatives can be synthesised. Apart from the cyclisation reactions involved in these syntheses, only the amino-group at position 5 in these compounds shows normal reactivity. It alone can be acylated (*e.g.*, with acetic anhydride or sodium dithioformate) or made to undergo Wöhler syntheses (with potassium cyanate or alkyl *iso*thiocyanates). With aldose sugars too, *N*-glycoside formation takes place only at the amino-group at position 5. With nitrous acid it is not clear if this amino-group can give true diazonium compounds, since by secondary reaction with the 4-amino-group cyclisation to pyrimidotriazole derivatives takes place.[77]

In pyrimidines bearing amino-groups at 2, 4, or 6, anomalous properties arise which justify doubts as to their structures. They are difficult to acylate; reaction with dithioformates, which the 5-amino-group undergoes readily, fails completely. The reaction with aldose sugars, which has been examined in connection with the synthesis of purine nucleosides, is difficult, and requires special structural features for its success.[78] With nitrous acid, rather sluggish deamination occurs, apparently without the intervention of diazonium compounds. It has already been mentioned that deamination can also be effected by the action of hot acids.

(c) *Structure of amino- and hydroxy-compounds.* The anomalous properties of compounds containing " hydroxyl " or " amino- " groups at 2, 4, or 6 suggests that they may in reality be derivatives of imino- or keto-dihydropyrimidines (*e.g.*, XXXIII, XXXV). Similar problems, of course,

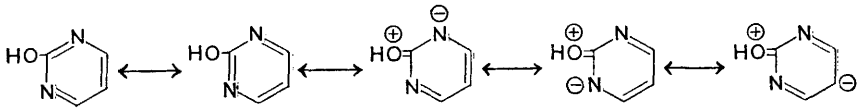





(XXXII.) (XXXIII.) (XXXIV.) (XXXV.)

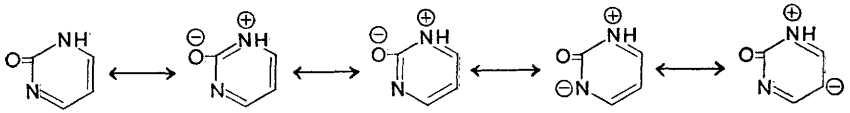
[76] F. H. Case and A. J. Hill, *J. Amer. Chem. Soc.*, 1930, **52**, 1536.
[77] R. O. Roblin *et al.*, *ibid.*, 1945, **67**, 290.
[78] J. Baddiley, B. Lythgoe, and A. R. Todd, *J.*, 1943, 571.

are met with in other heterocyclic systems, the so-called " tautomerism " of isatin and α- and γ-hydroxypyridines being long-known examples. In the following discussion it will be simplest to consider the hydroxy-derivatives of pyrimidines for illustration purposes, but much of what will be said applies also to the amino-compounds.

2-Hydroxypyrimidine can be written in two classical forms, the lactim (XXXVI) and lactam (XLI). These represent quite distinct compounds; but each of them possesses various resonance possibilities and is to be regarded as a hybrid (XXXVI—XL) and (XLI—XLV). When two such

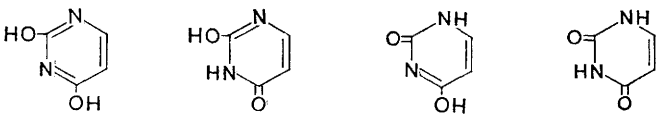


(XXXVI.) (XXXVII.) (XXXVIII.) (XXXIX.) (XL.)



(XLI.) (XLII.) (XLIII.) (XLIV.) (XLV.)

hydroxyl groups are present in the nucleus, as in uracil, four distinct forms are possible, each a hybrid derived from one of the classical structures (XLVI—XLIX). It is important to decide for any particular derivative



(XLVI.) (XLVII.) (XLVIII.) (XLIX.)

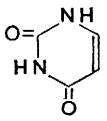
whether it shall be represented by the lactim or lactam structure, or whether one of these is present in the solid state and an equilibrium mixture in solution, with proportions controlled by the environmental conditions. No evidence has been obtained of the latter possibility, which would imply that these compounds were truly tautomeric.

The lactim–lactam question has been aired most extensively for uracil and barbituric acid. One approach to a problem of this kind is to use ultraviolet spectroscopy, a technique which has already shown its value in other branches of pyrimidine and purine chemistry. It is being increasingly used for the quantitative determination of small quantities of pyrimidine and

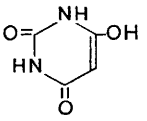
purine derivatives.[79] In the hands of T. Caspersson [80] it has proved of the utmost value for the detection and investigation of the behaviour of the nucleic acids present in chromosomes. H. M. Kalckar [81] has combined the use of specific enzymes for transformations such as adenine → hypoxanthine → xanthine → uric acid with ultra-violet measurements to give a " differential ultra-violet spectroscopy " which should be of great value in determining these derivatives in the presence of each other. On the constitutional side, the natural purine nucleosides were first diagnosed as 9-glycosides by ultra-violet spectroscopy,[82] and the spectrum of one of the fission products of vitamin B_1 gave the clue to its identity as a 4-aminopyrimidine derivative.[83]

In an attempt to clarify the structure of uracil by spectroscopic methods, J. E. Austin [84] has investigated the ultra-violet absorption of a number of its derivatives. She finds that uracil does not behave as the form (XLVI) since the spectrum is different from those of 2 : 6-dichloro- and 2 : 6-diethoxypyrimidine, which must possess this type of structure. Uracil differs also from 6-methoxy-3-methyluracil, which is a derivative of the form (XLVIII), and also from 3-methyluracil and 1 : 3-dimethyluracil; the last two compounds have the same spectrum, and the second of them is certainly, the first of them in all probability, a derivative of the form (XLIX). Uracil has the same spectrum as 1-methyluracil, in which the atomic arrangement at positions 1 and 6 is stabilised in the lactam form. Austin concludes, and to all appearances justifiably, that in solution in alcohol, in which the measurements were made, uracil is to be represented by the form (XLVII).

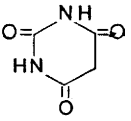
F. Arndt,[85] on the other hand, considers that the lactam structure is preferred by hydroxypyrimidines, except where such a structure would increase the energy of the system by removing the resonanace energy derived from the aromatic character of the ring; he formulates uracil as (L), but



(L.)



(LI.)



(LII.)

barbituric acid is given structure (LI), in which aromatic character is still possible by virtue of resonance, rather than structure (LII), where it is forbidden. Arndt attaches importance to the fact that with ethereal diazomethane uracil gives 1 : 3-dimethyluracil; if, as he maintains, this

[79] E. Vischer and E. Chargaff, *J. Biol. Chem.*, 1947, **168**, 781; J. F. Tinker and J. B. Brown, *ibid.*, 1948, **173**, 585; R. D. Hotchkiss, *ibid.*, 1948, **175**, 315.

[80] Ref. 8, p. 127. [81] *J. Biol. Chem.*, 1947, **167**, 429, 445, 461, 477.

[82] J. M. Gulland, E. R. Holiday, and T. F. Macrae, *J.*, 1934, 1639; J. M. Gulland and E. R. Holiday, *J.*, 1936, 765; J. M. Gulland and L. F. Story, *J.*, 1938, 259, 692.

[83] R. R. Williams, *J. Amer. Chem. Soc.*, 1935, **57**, 229; R. R. Williams, E. R. Buchman, and A. E. Ruehle, *ibid.*, p. 1093.

[84] *Ibid.*, 1934, **56**, 2141. [85] *Rev. Fac. Sci. Istanbul*, 1944, *A*, **9**, 19.

reagent acts by replacing the most acidic hydrogens of a molecule by methyl groups without alteration of structure, then solid uracil would be a fully lactam compound. While perhaps no final verdict should be given at present, the difficulties inherent in chemical methods of deciding this sort of structural problem will be remembered, and it seems likely that a complete clarification may ultimately come from the application of physical techniques. Infra-red spectroscopy appears particularly promising for this purpose, especially if technical difficulties, such as those involved in the use of hydroxylic media, can be overcome.

In the case of barbituric acid good chemical reasons for preferring structure (LI) exist; they are provided by the contrast between barbituric acid and its 5-monosubstitution products on the one hand and the 5 : 5-disubstitution products on the other; the latter must, of course, be derived from structure (LII). Thus barbituric acid and its 5-ethyl derivative are strong acids (K_a, 1051 and 383 $\times$ 10^{-7}); the 5 : 5-diethyl derivative is very much weaker (K_a, 0·37 $\times$ 10^{-7}).[86] The feeble acidity of ethyl malonate suggests that the difference is not likely to be due to the possibility of direct C–H dissociation in the first two compounds, but rather that an enol system, as in (LI) contributes to their acidity. A similar point occurs in considering the bromo-derivatives. One of the two bromine atoms in 5 : 5-dibromobarbituric acid, which is the product of direct bromination of the parent acid, is readily displaced by ammonia or sodium acetate, giving 5-bromobarbituric acid; the bromine atom in the latter is hard to replace, and replacement of both halogens of the dihalogeno-compounds occurs only in isolated cases, for example, in condensation with phenylhydrazine to give alloxan phenylhydrazone, and in the formation of riboflavin from the 5 : 5-dichloro-compound.[87]

The 5-halogeno-5-alkylbarbituric acids, like the 5 : 5-dibromo-compounds, contain a reactive halogen, which is replaceable by ammonia to give 5-alkyluramils. Clearly there is a difference in type between the 5 : 5-dihalogeno-compounds and 5-bromo-5-alkyl compounds, which must both be derived from structure (LII), and 5-bromobarbituric acid; the last, and the parent acid, are presumably structures of the type (LI).

(d) *Influence of structure on properties of amino- and hydroxy-compounds.* Many of the properties of hydroxy- and amino-pyrimidines are to be regarded as consequences of their hybrid structures, irrespective of whether these are of the lactim or of the lactam type. Taking first the physical properties, it is known that the polyhydroxy-compounds behave quite differently from pyrimidine and its simple alkyl and halogeno-derivatives; they are soluble in water rather than in organic solvents, and have high melting points (above 300°). These properties are due to hydrogen-bonding favoured by the contribution of Zwitterion structures such as (XXXVIII—XL; XLII—XLV). The amino-compounds are usually more soluble in alcohol, melt lower, and are often volatile enough to be purified by sublimation; here the

[86] J. C. Wood, *J.*, 1906, **89**, 1831.
[87] M. Tishler and J. W. Wellman, U.S.P. 2,261,608.

hydrogen-bonding is weaker, because nitrogen is less electronegative than oxygen.

Removal of a proton from either lactim or lactam forms gives an anion with effective resonance possibilities ; the acidity of the hydroxy-compounds depends on this. Introduction of an amino-group, on the other hand, increases the basicity of the molecule ; thus 2-hydroxypyrimidine is amphoteric ; uracil has no basic properties ; cytosine has well-marked basic properties. The separation methods which are becoming of major importance in nucleic acid chemistry make use of these principles of solubility and acid–base behaviour. Cytosine and uridine are readily separated using ion-exchange resins ; [88] filter-paper chromatography and counter-current extraction methods using buffers permit the separation and identification of purine and pyrimidine derivatives.[79]

Hydroxyl and amino-groups at positions 2, 4, and 6 can influence profoundly the chemical properties of compounds containing them. An important example of this is the behaviour of pyrimidines towards electrophilic substituting agents ; it has already been noted that, in the absence of at least one " tautomeric " group, halogenation of pyrimidine derivatives has not been effected. It becomes possible when one such group is present, as in 2-aminopyrimidine,[89] and is easy in uracil, barbituric acid, and their amino-analogues. The way in which substitution is facilitated by the progressive introduction of hydroxyl or amino-groups will be clear from the structures (XL) and (XLV).

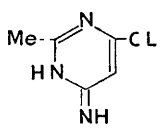
Nitration of the nucleus appears to be rather more difficult than halogenation, and only when two or more " tautomeric " groups are present has it been used for preparative purposes. Given these structural features it is possible to effect nuclear substitution by more weakly electrophilic reagents ; nitrosation and coupling with diazonium compounds take place at position 5, and some applications of these useful reactions have already been mentioned. There is little information as to the mechanism whereby these two substitutions occur in pyrimidine compounds, but the structural requirements necessary for their success are known to some extent.[45] With sodium nitrite and acetic acid the introduction of a nitroso-group requires " tautomeric " groups at both 4- and 6-positions. It is facilitated by a group such as SMe at position 2 which can act as an electron source in the conventional manner ; where less activating groups (H or Me) are present at position 2 in a 4 : 6-diaminopyrimidine the presence of mineral acid is required to promote the nitrosation. Pyrimidine derivatives will couple with reactive diazonium compounds derived from chloro- or nitro-anilines when " tautomeric " groups are present at either the 2 and 4 or at the 4 and 6 positions ; when three such groups are present, as in barbituric acid, coupling occurs even with such a mild reagent as benzenediazonium chloride. The coupling reaction presents some interesting and, as yet, unexplained features. Thymine reacts with diazosulphanilic acid, although no hydrogen is available for

[88] D. T. Elmore, *Nature*, 1948, **161**, 931 ; R. J. C. Harris and J. F. Thomas, *ibid.*
[89] J. P. English, *et al.*, *J. Amer. Chem. Soc.*, 1946, **68**, 453.

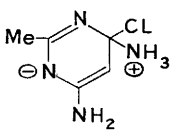
replacement at position 5. The observation [90] that both uracil and 1-methyluracil react under these conditions, whilst 3-methyluracil does not, probably has an important bearing on the lactam–lactim structure of uracil; it may be compared with Austin's data on the ultra-violet absorption of these compounds.

Among the other substitution reactions of hydroxypyrimidines, hydroxymethyl and chloromethyl groups can be introduced into the uracil molecule; as in the benzene series, the 5-chloromethyl compounds so obtained contain a replaceable halogen atom. The many substitution reactions which barbituric acid derivatives can undergo are probably not all of the same type as those described for the less heavily hydroxylated compounds. Among them are the introduction of an hydroxyl group at position 5 by use of hydrogen peroxide, the alkylation with bases and alkyl halides to give 5-alkyl- and 5 : 5-dialkyl-barbituric acids, and the formation of 5-arylidene derivatives when 1 : 3-dimethylbarbituric acid reacts with aromatic aldehydes. These reactions are more reminiscent of the behaviour of malonic acid derivatives than of true pyrimidine derivatives; much of the aromatic character of the nucleus is lost in barbituric acid.

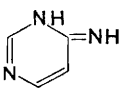
The loss of aromatic character resulting from the presence of hydroxyl groups is shown by the resonance formulations given on p. 202; it is already apparent when two such groups are present in the nucleus, and is then manifested by loss of stability to oxidising agents and by the addition reactions which can take place at the 4 : 5-double bond. Hypohalous and nitric acids both undergo addition to this link in thymine; oxidising agents probably act by adding two hydroxyl groups to it. The addition compounds so formed are unstable to alkali; this affords a method of degrading the nucleus. Addition reactions often intervene to complicate the substitution of uracil and its relatives; uracil and bromine water give 5 : 5-dibromo-4-hydroxydihydrouracil. In monoamino-compounds the resonance energy of the nucleus is sufficient to forbid the addition reactions, and, with aqueous halogens, only substitution occurs.



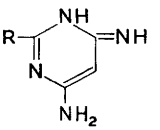
(LIII.)



(LIV.)



(LV.)



(LVI.)

One further effect of the presence of "tautomeric" groups in the pyrimidine nucleus calls for comment, namely their capacity for influencing decisively the behaviour of another group situated *m*- to them. The replacement of the halogen atoms by amino-groups in a di- or tri-halogenopyrimidine can be accomplished step by step, a fact of some importance in preparative work. Clearly the reactivity of the second halogen atom is depressed by the introduction of an amino-group *m*- to it. Taking 4 : 6-dichloro-2-

[90] T. B. Johnson and S. H. Clapp, *J. Biol. Chem.*, 1908, **5**, 49.

methylpyrimidine as an illustration, this effect can be explained by reference to formula (LIII) for the 4-chloro-6-amino-2-methylpyrimidine obtained by mild treatment with ammonia. The tendency of the latter to behave as the iminodihydro-form reduces the chances of forming the transition state (LIV) necessary for the replacement of the second halogen atom.

Similarly it has been suggested [78] that the failure of 4-aminopyrimidine derivatives to yield 4-glycosidamino-compounds on treatment with aldose sugars is due to their tendency to react in the iminodihydro-form (LV). In 4 : 6-diaminopyrimidines bearing at position 2 a " non-tautomeric " group such as H, Me, or SMe it is clear that both of the amino-groups cannot exist in the imino-form, so that one of them should show the characteristics of a true amino-group ; such compounds, *e.g.* (LVI ; R = H, Me, SMe), do in fact show the expected capacity for reaction. It is to be expected that in the future many other reactions of pyrimidine derivatives will find an interpretation along similar lines.