

T. L. V. ULBRICHT

**PURINES
PYRIMIDINES
AND
NUCLEOTIDES
and the Chemistry
of Nucleic Acids**

The Commonwealth and International Library of Science, Technology, Engineering and Liberal Studies

Author's Preface

THE study of Nucleic Acids today combines the disciplines of biochemistry, organic and physical chemistry, crystallography and genetics, in what has come to be called 'molecular biology'. A number of specialized monographs on various aspects of this subject exist, but there is not any book which attempts to give a simple introduction to the basic chemistry of purines and pyrimidines and their derivatives. On the other hand, textbooks of heterocyclic chemistry give an inadequate account of these compounds, and in general textbooks of organic chemistry they are virtually neglected.

It is the modest aim of the present volume to provide such an introduction, and it is hoped that it will prove useful to senior undergraduates taking courses in heterocyclic and natural products chemistry. If the book errs on the side of brevity, that is intentional, and it is hoped that readers will avail themselves of the references given to monographs, review articles and a few original papers. The names of the more important individual compounds have been printed in bold type.

It remains for me to record my thanks to Sir Robert Robinson for his criticism and encouragement; to Dr. R. F. Curtis, who read the whole book, and Dr. D. M. Brown, who read Chapters IV and V, for their helpful and forthright comments. I am grateful to my wife for her invaluable help with the Index.

Note by the Editor

The volumes in the Advanced Section of this Course in Organic Chemistry have an independent value as short monographs by experts in the respective fields treated. They are therefore offered in advance of the General Sections of the Course, even though this implies a greater prior knowledge on the part of the student.

Contents

	PAGE
AUTHOR'S PREFACE	vii
INTRODUCTION	1
I. GENERAL CHEMISTRY OF PURINES AND PYRIMIDINES	4
1. General Character of Purines and Pyrimidines	4
2. Substitution by Electrophilic Reagents	6
3. Substitution by Nucleophilic Reagents	10
4. Rearrangements	19
5. Reduction	21
6. Addition Reactions	23
7. Acylation and Alkylation	25
8. Oxidation	27
References	29
II. SYNTHESIS OF PURINES AND PYRIMIDINES	30
1. Pyrimidines	30
2. Purines	34
References	37
III. NUCLEOSIDES	38
1. Structure	38
2. Synthesis	43
3. Mechanism and Stereochemistry of Nucleoside Synthesis	48
4. Analogues	50
References	52

IV. NUCLEOTIDES	53
1. Ribonucleotides	53
2. Deoxyribonucleotides	57
3. Other Nucleotides	58
4. Biosynthesis of Mononucleotides	61
References	63
V. NUCLEIC ACIDS	64
1. Structure of DNA	64
2. Structure of RNA	70
3. Biosynthesis of Nucleic Acids	72
References	72
INDEX	75

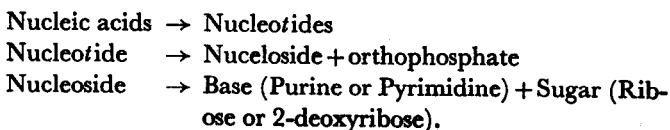
Introduction

THE reason for devoting a volume in this series of textbooks on organic chemistry to purines, pyrimidines and their derivatives, is their importance in Nature as well as their intrinsic chemical interest. Considerable attention is devoted to these compounds in current chemical and biochemical research.

Work on the **nucleic acids**, polymeric materials originally isolated from cell nuclei, was initiated in the nineteenth century by Miescher, Altmann and Kossel, and important contributions in the early part of this century were made by Levene and others, and the structure of the major bases, and of the sugars, was established.

After the war there were initially two major developments, both of which were greatly helped by the application of new chromatographic techniques. Improved methods of isolation and careful analytical work by Chargaff and others established that nucleic acids were polymers of high molecular-weight, and that in DNA, the ratio of the base content of adenine to thymine, and of guanine to cytosine, was equal to one.

At the same time, chemical investigations, mainly by Todd and co-workers, led to the synthesis of nucleosides and nucleotides and the determination of the structure of the nucleotides obtained by chemical and enzymatic hydrolysis of nucleic acids. This made it possible to formulate nucleic acids as 3',5'-linked linear polynucleotides. The relationship between the various derivatives is indicated by the following hydrolytic sequence:



With this work as a basis, Watson and Crick were led to their brilliant interpretation of the X-ray crystallographic studies of polymeric DNA by Wilkins and others, and to propose the specifically hydrogen-bonded double-stranded helical structure for DNA.

However, the natural occurrence of purines and pyrimidines is not limited to nucleic acids. The methylated xanthines occur in the plants from which tea, coffee and cocoa are prepared, and uric acid was isolated as long ago as 1776 from gall-stones and urine. Purines and pyrimidines also occur in certain antibiotics—nebularine (purine), puromycin (6-dimethylaminopurine) and amicitin (cytosine)—and in numerous other derivatives.

Uric acid occupies a key position in that the fundamental work on purines by Emil Fischer was centred on it. At this time—the final decades of the last century—Pinner and others were beginning work on pyrimidines, the nucleic acid bases were isolated, and Traube developed his famous synthesis of purines from pyrimidines. Thereafter, interest in purines was maintained by the intriguing complexity of the oxidation of uric acid, but little work on pyrimidines, apart from the isolated efforts of T. B. Johnson at Yale, was done until the discovery that Vitamin B₁, thiamine, was a pyrimidine derivative. Subsequently, effective anti-malarial ('daraprim') and sulphonamide ('sulphadiazine') drugs were based on the pyrimidine nucleus.

Since the war, the realization of the central role of nucleic acids in the metabolism of all living matter has led to the synthesis of many analogues of nucleic acid derivatives as anti-metabolites, for the chemotherapy of cancer. Success in therapy has been limited but the investigation of the mode of action of active compounds has greatly deepened our understanding of fundamental biochemical processes.

References

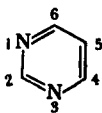
1. P. A. LEVENE and L. W. BASS, *Nucleic Acids*, Chemical Catalog Co., New York, 1931. This gives an interesting account of the older work.

2. E. CHARGAFF and J. N. DAVIDSON (Eds.), *The Nucleic Acids*, vols. 1 and 2, 1955; vol. 3, 1960. Academic Press, New York. The first two volumes provided a valuable, comprehensive summary of the accumulated knowledge available at that time. Vol. 1 (mainly devoted to chemistry) is still useful for reference; vol. 2 (biochemistry) is completely out of date.
3. A. M. MICHELSON, *The Chemistry of Nucleosides and Nucleotides*, Academic Press, 1963. Though not without faults, this comprehensive book (600 pages) is now the main work of reference in this field.
4. D. M. BROWN and T. L. V. ULBRICHT, *The Nucleic Acids*, in *Comprehensive Biochemistry*, vol. 8, ed. by M. FLORKIN and E. H. STOTZ, Elsevier, 1963. A chapter of about 100 pages, with the main emphasis on nucleosides, nucleotides and polynucleotides. The same volume contains a short section by M. H. F. WILKINS on The Three-Dimensional Configuration of the DNA molecule.

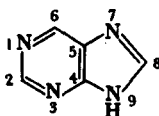
General Chemistry of Purines and Pyrimidines

Introduction

In this book, pyrimidine (I) and purine (II) are numbered as shown. This is the numbering used in *Beilstein*, in most important reference books, in all biochemical journals, and in most chemical journals. Unfortunately, *Chemical Abstracts* adopted a different numbering for the pyrimidine ring a few years ago, thus destroying the correspondence in the numbering of the two rings. The inconsistency of pyrimidine nomenclature in this journal renders the subject index virtually useless, and it is necessary to rely on the formula index.



(1)



(2)

The names of the more important individual compounds, many which have trivial names, are printed in heavy type.

1. General Character of Purines and Pyrimidines

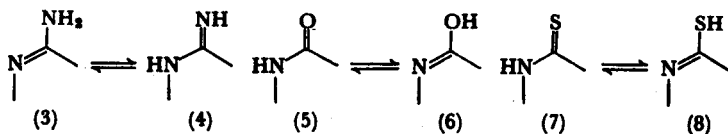
Pyrimidine, also known as *m*-diazine, may be regarded as derived from benzene by substitution of two meta —CH= groups by —N= . As a result of the electronic effects of these two nitrogen

atoms, which reinforce each other, positions 2, 4 and 6 are even more electron-deficient than the α and γ positions in pyridine. The same applies to positions 2, 6 and 8 in purine. Thus, halogen atoms in these positions will undergo substitution reactions with water, ammonia, thiols and alkoxides; methyl groups condense with aldehydes, and are easily oxidized to carboxylic acids which are readily decarboxylated. Position 5 in pyrimidine is less electron-deficient—it corresponds to the β position in pyridine, and electrophilic substitution is possible. The presence of electron-releasing groups in positions 2, 4 and 6 in pyrimidine activates position 5 to electrophilic attack, and at the same time, deactivates the other positions to nucleophilic attack; similarly, such groups at positions 2 and 6 in purine make possible electrophilic substitution at position 8. Conversely, a strong electron-withdrawing group at position 5 in pyrimidine, accentuates the electron-deficiency at the other positions.

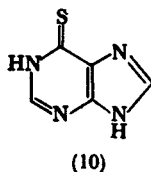
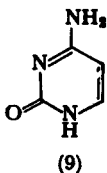
With the reservations which the above remarks imply, it may nevertheless be said that both ring systems show marked aromatic properties, including resistance to oxidation. The solubilities and melting-points of derivatives are strikingly dependent on the nature of the substituents. Alkyl, aryl, alkoxyl, alkylthio, halogen and similar derivatives are liquids or low-melting solids which are readily soluble in organic solvents. On the other hand, compounds containing one amino, hydroxy or mercapto substituent are water-soluble, but become less soluble and higher-melting the more such substituents they contain, owing to strong inter-molecular hydrogen bonding.

One of the most important features of purines and pyrimidines is the existence of tautomerism in derivatives containing $-\text{NH}_2$, $-\text{OH}$ and $-\text{SH}$ groups. Comparison of the ultra-violet and infra-red spectra of these compounds with those of the unambiguously synthesized methyl derivatives corresponding to the different tautomers, has established that amino compounds exist in the amino (3), not the imino (4) form, whereas hydroxy compounds exist in the lactam (5) rather than the lactim (6) and mercapto compounds in the thione (7), not the thiol (8) forms.

Throughout this book, the correct tautomeric forms will be given in structural formulae, but, because the nomenclature is simple and



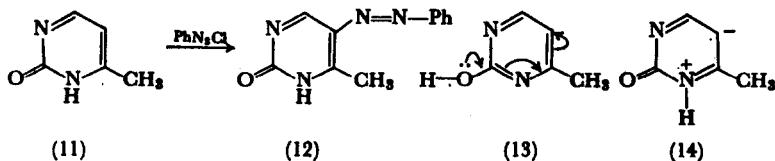
unambiguous, compounds will be named as amino, hydroxy and mercapto derivatives of the aromatic ring, e.g., 2-amino-6-hydroxypyrimidine, cytosine (9), 6-mercaptapurine (10), etc.



2. Substitution by Electrophilic Reagents

As already explained, electrophilic substitution requires the presence of electron-releasing groups. The magnitude of their activating effect is usually in the order OH , $\text{NH}_2 > \text{SH}$, SR , $\text{OR} > \text{alkyl}$, though this may be modified by a change in pH.

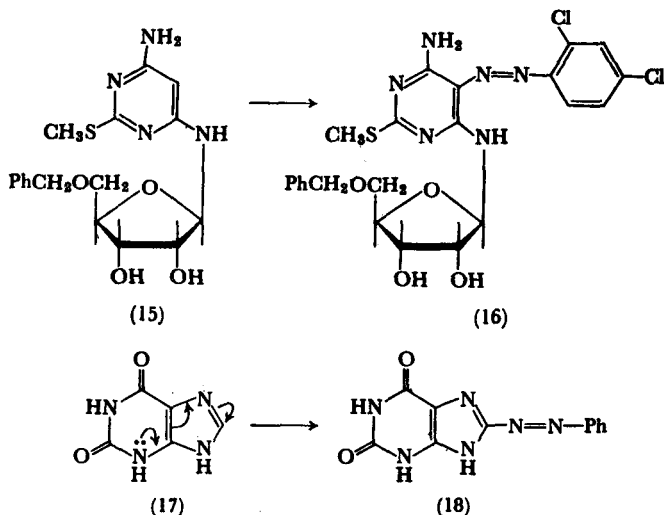
An important reaction in pyrimidines is the introduction of nitro, nitroso and azo groups at position 5, since reduction gives



5-amino-pyrimidines, often required as intermediates in the synthesis of purines and of pteridines. The presence of one activating group in a pyrimidine is sufficient to make azo-coupling

possible, e.g., with 2-hydroxy-4-methylpyrimidine (11). If this had the lactim form, the activation by the hydroxy group could be written as (13), exactly as with phenol, but as it has the lactam form, activation may be due to contribution by resonance structures like (14).

When two activating groups are present, azo-coupling takes place readily and in good yield, though the use of a more reactive diazonium salt (an electron-withdrawing group in the *p*-position increases the electrophilic character of the diazonium cation) is sometimes advantageous. The value of this reaction lies in the

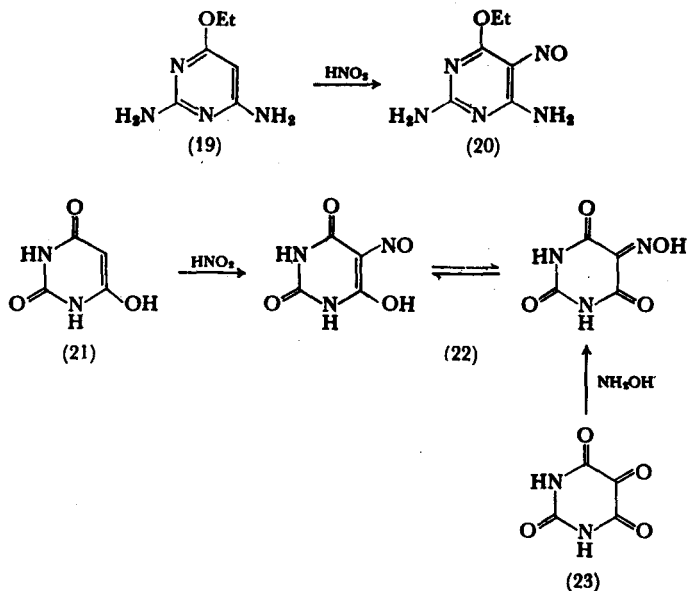


mildness of the conditions under which it is carried out, and is illustrated by its success with (15), a compound which, clearly, could not be nitrated because of the sensitivity of the sugar moiety.

Position 8, in purines in which there are activating groups in the pyrimidine ring, like **xanthine** (17), can also be coupled; this reaction is used as a test to see whether this position is substituted.

Nitrosation requires the presence of two or three activating groups, as in (19).

A special form of tautomerism is possible in these compounds, involving the nitroso group, which is particularly clearly revealed by the synthesis of violuric acid (22) either by nitrosating **barbituric acid** (21) or by the reaction of **alloxan** (23) with hydroxylamine.

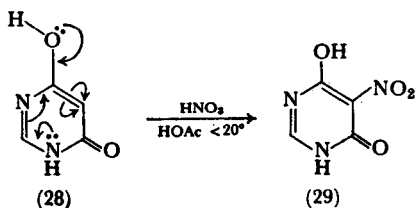
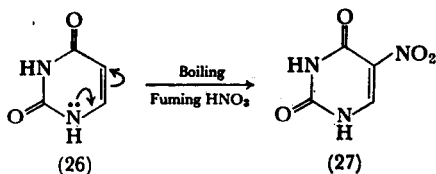
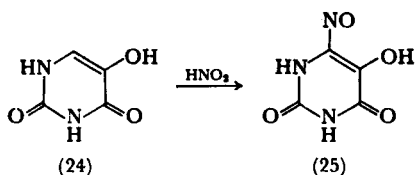


Most known nitrosopyrimidines have the nitroso group in position 5, but when three activating groups are present, as in isobarbituric acid (24), substitution in position 6 is possible, to give, in this case, isovioluric acid (25).

Nitration is carried out with fuming nitric acid in the case of **uracil** (26), and more usually with a nitric/sulphuric mixture, but with 4,6-dihydroxypyrimidines, in which one OH group must be genuinely enolic and which are more sensitive to oxidation, acetic acid is used as solvent. It will be noted that, in **uracil**, only N_3 can make electrons available at C_5 whereas in (28), both N_3 and the oxygen of the OH group can do so.

In the course of nitration, mercapto and alkylthio groups are usually converted to —OH , and if elevated temperatures are used, C-methyl groups may be oxidized to —COOH .

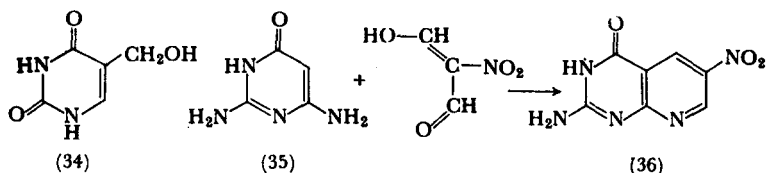
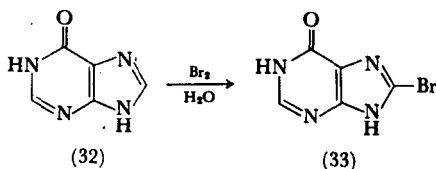
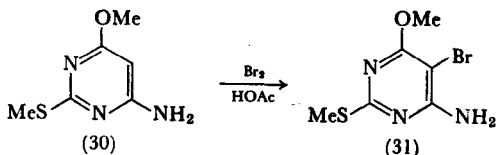
Direct halogenation, in water or acetic acid, is a reaction which proceeds more easily. For example, (30) cannot be nitrosated but



readily reacts with bromine in acetic acid. Similarly there are numerous purines, e.g., **hypoxanthine** (32), which can be halogenated in the 8 position.

A few pyrimidines have been sulphonated. Chlorosulphonation of **uracil** gives the 5-chlorosulphonyl derivative, but with **cytosine** the 5-sulphonic acid is isolated. Although cytosine is a weak base, it is apparently sufficiently basic to effect the hydrolysis of the chlorosulphonyl group.

A reaction of some interest is the condensation of pyrimidines with aldehydes. The reaction of **uracil** with formaldehyde is catalysed both by acid and base, and gives **5-hydroxymethyl-uracil** (34). Earlier attempts to synthesize this compound had



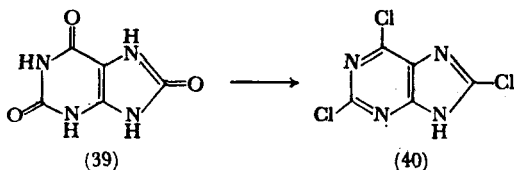
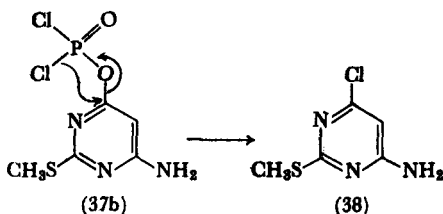
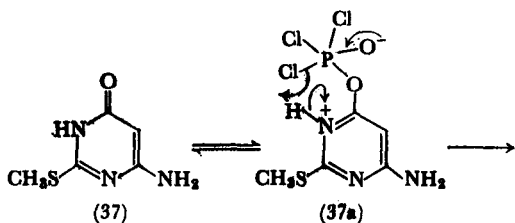
failed and it was concluded, mistakenly, that the product quickly lost formaldehyde. In aminopyrimidines, condensation may take place either at position 5 or on the amino group—2,4-diamino-6-hydroxypyrimidine (35) reacts with glucose to give 60% of the C-glucoside, and a double condensation takes place with nitro-malondialdehyde, forming the pyridopyrimidine (36).

3. Substitution by Nucleophilic Reagents

2-, 6- and 8-substituted purines and 2-, 4- and 6-substituted pyrimidines undergo a wide variety of nucleophilic displacements. These are best classified according to the nature of the incoming group.

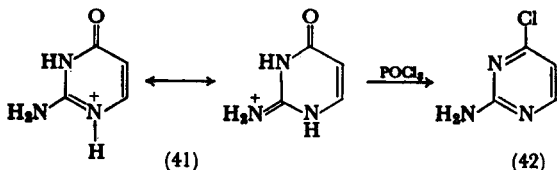
(a) *Substitution by Cl* ($-\text{OH} \rightarrow -\text{Cl}$).

This transformation is usually carried out with phosphorus oxychloride. (Phosphorus pentachloride has the disadvantage that, in pyrimidines, it may introduce chlorine into a free 5-position.) The yield is often improved by the addition of a base like dimethylaniline, e.g., in the chlorination of (37), and of **uric acid** (39).

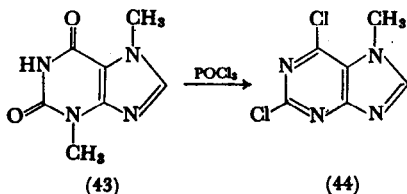


Amino groups exert a deactivating effect, presumably on the key step (37b \rightarrow 38), an intramolecular nucleophilic displacement. In the case of isocytosine the reaction is carried out in the presence of

sulphuric acid; protonation (41) overcoming the deactivating effect.



Methylated xanthines may react with phosphorus oxychloride with loss of N-methyl groups; for example, **theobromine** (43) yields 2,6-dichloro-7-methylpurine (44); a molecule of methyl chloride is eliminated in the first step.

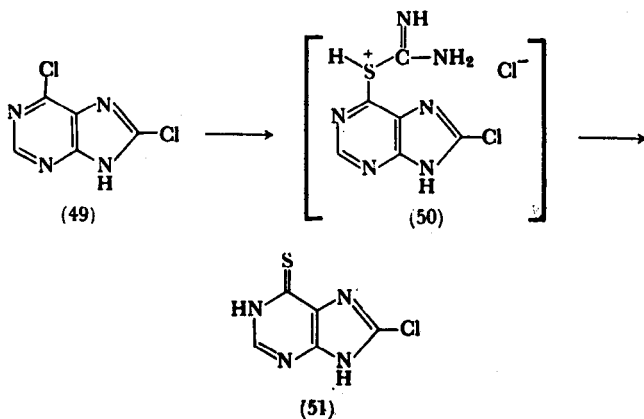
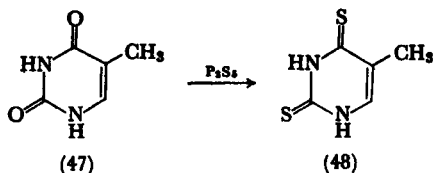
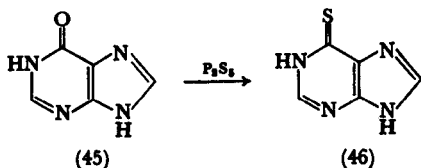


(b) *Substitution by SR* ($-\text{OH}$, $-\text{Cl} \rightarrow -\text{SH}$, $-\text{S}-\text{alkyl}$)

Direct replacement of $-\text{OH}$ by $-\text{SH}$ is effected using phosphorus pentasulphide, either in pyridine (**hypoxanthine** (45) to **6-mercaptapurine** (46)) or tetralin (**thymine** (47) to dithiothymine (48)). The mechanism of this reaction is not known, but must involve an addition, followed by an intramolecular nucleophilic displacement.

Chlorine may be replaced by SH using thiourea in ethanol; the intermediate thiuronium salt (50) is so readily hydrolysed that it is not usually isolated. With one mole of thiourea, 6,8-dichloropurine (49), gives (51), and with two moles of thiourea, 6,8-dimercaptapurine is obtained.

Alkali-metal hydrosulphides and alkylmercaptides (e.g., NaSEt) are also used to replace $-\text{Cl}$ by $-\text{SH}$ and by $-\text{S-alkyl}$ groups.



(c) *Substitution by alkoxides* ($-\text{Cl} \rightarrow -\text{OR}$)

This reaction proceeds readily, even in compounds containing amino groups, and can be carried out in a stepwise manner. Reaction of 2,4,6-trichloropyrimidine (52), with one, two, or three moles of a sodium alkoxide in the corresponding alcohol at