BASIC PRINCIPLES IN NUCLEIC ACID CHEMISTRY

Edited by

PAUL O. P. TS'O

Division of Biophysics
Department of Biochemical and Biophysical Sciences
The Johns Hopkins University
Baltimore, Maryland

VOLUME II

ACADEMIC PRESS *New York and London 1974*
A Subsidiary of Harcourt Brace Jovanovich, Publishers

ACADEMIC PRESS, INC.
111 Fifth Avenue, New York, New York 10003

United Kingdom Edition published by
ACADEMIC PRESS, INC. (LONDON) LTD.
24/28 Oval Road, London NW1

Library of Congress Cataloging in Publication Data

Ts'o, Paul On Pong, Date
Basic principles in nucleic acid chemistry.

Includes bibliographical references.
1. Nucleic acids. I. Title.
QD433.T77 547'.596 72-13612
ISBN 0-12-701902-2 (v. 2)

PRINTED IN THE UNITED STATES OF AMERICA

BASIC PRINCIPLES IN NUCLEIC ACID CHEMISTRY

VOLUME II

CONTRIBUTORS

WILLIAM BAUER

D. M. BROWN

C. ALLEN BUSH

HENRYK EISENBERG

PAUL O. P. TS'O

JEROME VINOGRAD

LIST OF CONTRIBUTORS

Numbers in parentheses indicate the pages on which the authors' contributions begin.

WILLIAM BAUER* (265), Department of Chemistry, University of Colorado Boulder, Colorado

D. M. BROWN (1), University Chemical Laboratory, Cambridge, England

C. ALLEN BUSH (91), Department of Chemistry, Illinois Institute of Technology, Chicago, Illinois

HENRYK EISENBERG (171), Polymer Department, The Weizmann Institute of Science, Rehovot, Israel

PAUL O. P. TS'O (305), Division of Biophysics, Department of Biochemical and Biophysical Sciences, The Johns Hopkins University, Baltimore, Maryland

JEROME VINOGRAD (265), Division of Biology and Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California

* *Present Address*: Department of Microbiology, State University of New York, *Stony* Brook, New York.

PREFACE

About one hundred years ago, a young Swiss physician, Friedrich Miescher, published the first paper on "nuclein" (or nucleohistone in current terminology) and thus launched chemical research on nucleic acids. Nearly twenty-five years ago, nucleic acid was identified as the physical basis of genes, and since that time the quest for knowledge on genes rightfully has become a major thrust in modern biological research. In fact, the tremendous progress in nucleic acid research has raised the possibility that advancements in this field may exert a profound influence on the future of man.

We, as researchers in nucleic acid chemistry, have prepared this multivolume treatise in honor of this historic event: the centennial anniversary of the discovery of nucleic acid. Our view is that progress in nucleic acid chemistry has been substantial and sufficient to justify an attempt to formulate certain basic principles in this field. We hope that these basic principles will not only endure the test of time but will serve as a foundation for further advancement in nucleic acid research as well. Not only have we critically examined the achievements of the past, we have also contemplated the future: the momentum of nucleic acid research and its contribution and influence on the destiny of man. Knowledge of nucleic acid chemistry will be utilized more extensively than ever in biomedical research areas such as cell biology, differentiation, microbiology, virology, oncology, genetic therapy, and genetic engineering. Hopefully, this treatise will serve as reference and resource material for many workers in biomedical research and as teaching material for instructors in institutions of higher learning.

In following the approach of Volume I, the first four chapters in Volume II are written by scholars who have expert knowledge in a particular area of research in nucleic acid chemistry. These are Chapter 1, Chemical Reactions of Polynucleotides and Nucleic Acids; Chapter 2, Ultraviolet Spectroscopy, Circular Dichroism, and Optical Rotatory Dispersion; Chapter 3, Hydrodynamic and Thermodynamic Studies; and Chapter 4, Circular DNA. Chapter 5, Dinucleoside Monophosphates, Dinucleotides, and Oligonucleotides, describing the current knowledge and concepts of nucleic acid chemistry at this level of complexity is written by the editor.

We wish to thank the following reviewers for their helpful suggestions in

the preparation of this volume: Chapter 1, Freidrich Cramer, Max–Planck–Institut für Experimentelle Medizin, Göttingen, Germany; Chapter 2, Ignacio Tinoco, Jr., University of California at Berkeley; Chapter 3, John E. Hearst, University of California at Berkeley; Chapter 4, James C. Wang, University of California at Berkeley; Chapter 5, Charles Cantor, Columbia University. The valuable assistance of Carl Barrett and Jean Conley in the preparation of the Subject Index and the help from many colleagues and from the staff of Academic Press are gratefully acknowledged.

Onward to Volume III!

PAUL O. P. TS'O

CONTENTS OF OTHER VOLUMES

Volume I

Volume III

CONTENTS

4. Circular DNA

William Bauer and Jerome Vinograd

5. Dinucleoside Monophosphates, Dinucleotides, and Oligonucleotides

Paul O. P. Ts'o

1

CHEMICAL REACTIONS OF POLYNUCLEOTIDES AND NUCLEIC ACIDS

D. M. BROWN

I. Reactivity of Polynucleotides

A. Introduction

Ostensibly the subject of this chapter is the susceptibility of the nucleic acids toward attack by chemical reagents. This is indeed an important matter. It would, for example, be of great interest to know the principles which govern the reactivity of bases in DNA toward certain carcinogens or mutagens. What effect does nucleotide sequence or secondary structure have? What effect does tertiary structure in ribosomal, messenger, and transfer RNA have, in corresponding reactions? The fact is that very little is known. No *ab initio* estimates of steric hindrance in polynucleotide reactions have been made, nor is it likely that these could be usefully made at present, though the nature of ordered polynucleotide structures and the forces involved in stabilizing them have received much attention [1]. As a result much of the chapter will be devoted to a discussion of a variety of reagents important in nucleotide chemistry, and on the mechanistic principles which inform their reactions. Emphasis will be placed on those classes of compounds whose reactions with polynucleotides have been studied.

We may begin by introducing a generalization to the effect that the reactivity of bases or of the internucleotide linkages in polynucleotides will only differ from those of the corresponding monomeric systems *in rate*. This may appear a sufficiently obvious and naive suggestion as to warrant no mention, but, for example, it is not long since the view was seriously held that the difference between DNA and RNA in their susceptibilities to base hydrolysis was due to the secondary structure in the former. The generalization is introduced because it provides the justification for the view that a study of the chemistry of the monomeric species (be it base, nucleoside, or nucleotide) is a necessary and valid approach to the chemistry of the same residue in a polymeric system. The rate difference may of course be, and often is, very large. Thus, for example, reactions which depend on a specific approximation of the reacting species—photodimerization of pyrimidines—may be much faster (or the amount of product at equilibrium greater) in a polynucleotide duplex. On the other hand, reactions of the same polynucleotide with external reagents may be exceedingly slow.

The assumption behind many recent studies has been that tentative conclusions can be drawn in respect of secondary and tertiary structure on the basis of comparative rate data. This will be discussed in more detail later in the chapter.

B. Types of Nucleotides

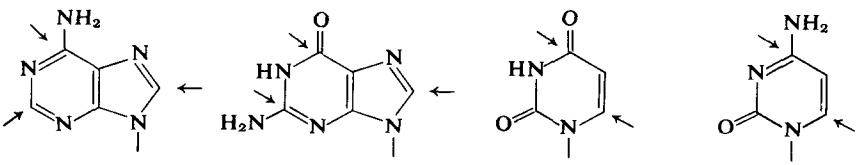
The discussion will be confined, in the main, to the chemistry of the base residues adenine, guanine, cytosine, and uracil as it is seen in nucleosides

and nucleotides or suitable analogs. The natural modified bases, e.g., dihydrouracil, the thiopyrimidines, the 5-substituted pyrimidines, pseudouridine, and certain alkylated purines will be treated, though in less detail. The discussion will be limited to certain classes of reagents which are considered to be important from the standpoint of biological activity or use in polynucleotide structure elucidation. For this reason reactions that involve vigorous conditions are not included. A number of topics included here are dealt with more exhaustively elsewhere [2].

II. Base Modification by Nucleophilic Species

A. Hydrolysis

The positions available for nucleophilic displacements and addition reactions in the common purine and pyrimidine bases are those indicated by arrows:



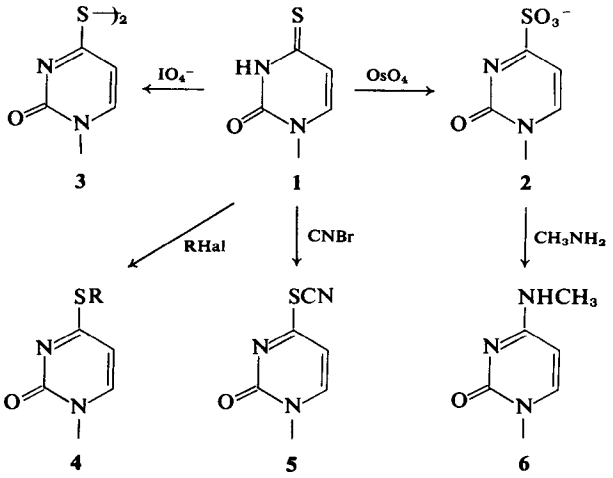
It is worth noting here that nucleophilic reactions may often be subject to electrophilic catalysis. For instance, hydrolysis or aminolysis displacement reactions may proceed much more rapidly on the protonated heterocycle than on the free base [3,4]. Alkylation, too, has marked effects on the hydrolytic stability of nucleosides. The ease with which purines are hydrolyzed, with opening of the 5-membered imidazole ring, varies enormously. Attack at C-8 is much facilitated by alkylation on N-7 in nucleosides (Section III,B). Reactions involving exchange of C-8 also occur and, for example, α-diketones may afford pteridines [4a].

Addition of nucleophilic reagents to the C-6 position in the pyrimidines leads to a variety of results which have important biological consequences. This will be discussed in Section II,C, but in broadest outline the initial consequence of nucleophilic attack at C-6 is activation of C-4 to nucleophilic and of C-5 to electrophilic substitution.

The hydrolysis of adenine and of guanine to hypoxanthine and xanthine, respectively, either acid- or base-catalyzed, appear to be reactions so slow as not to be significant [5]. Cytidine is very slowly deaminated to uridine in strong base [6]. It is also hydrolyzed in a reaction having a rate maximum at pH 4–5 [7,8]. This reaction is dependent on buffer ion catalysis, for example, by phosphate and citrate. The hydrolysis may result in part from direct attack

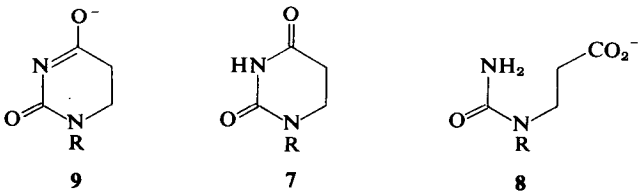
by water at C-4 of the pyrimidine ring with general acid–base catalysis, but it is clear that a more complex mechanism is also involved, which is discussed below. In any event, the neutral hydrolysis of cytidine to uridine, though slow, is sufficiently fast to be biologically significant.

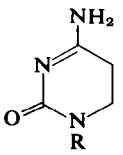
The displacement of sulfur from 2- and 4-thionucleosides can be effected, the order of reactivity being 2-thio < 4-thio [3]. The reaction rate is much increased if the thiono group is first modified. Thus osmium tetroxide [9], or



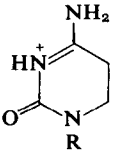
periodate oxidation [10], alkylation [3], or treatment with cyanogen bromide [11], lead to reactive intermediates (**1**–**5**, respectively) which can undergo hydrolysis or aminolysis (to **6**). Thus the thiouridine residue in yeast tRNA is converted to a uridine residue via **5** with retention of acceptor activity and its conversion to a cytidine or N_4-methylcytidine residue via **2** and **5** has also been demonstrated [10,12].

Dihydrouridines (**7**) suffer ring cleavage in strong base to give the corresponding β-ureidopropionate (**8**). The reaction would be much faster but for the formation at high pH values of the nonproductive anion (**9**); λ_{max}, 230 nm) [13,14]. When this cannot form, as in the case of 1,3-dimethyl-5,6-dihydrouracil (**10**; R = CH_3), hydrolysis is rapid [15].

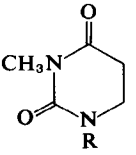




11



12

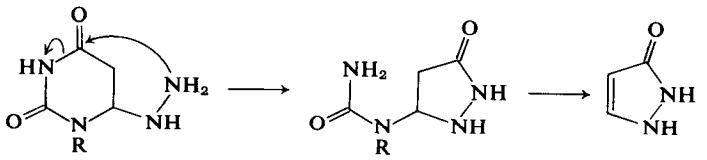


10

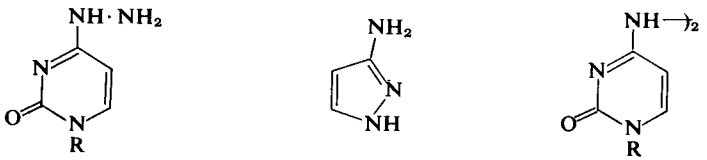
Dihydrocytosines (**11**) undergo extremely rapid hydrolysis under mild conditions. Thus the cation (**12**) has a $t_{1/2}$ = 15 min at 37°C [16]. Dihydrocytosine nucleosides and nucleotides can be made by catalytic hydrogenation over rhodium-on-alumina, but purity is difficult to achieve due to deamination [16,17]. Part of the latter results from overhydrogenation to the tetrahydro derivative [18]. No dihydrocytosine derivatives have been observed to occur naturally. Their sensitivity to hydrolysis may have prevented detection, but it is probable that the dihydrouridine residues in tRNA's are modified uridines and not artifacts of dihydrocytidine hydrolysis [19].

B. Reactions with Hydrazines

It has long been known that hydrazine hydrate cleaves uracil and cytosine but not the purines. The former gives rise to pyrazolone (**15**) and the latter to aminopyrazole (**16**) as the major product [20,21]. Mechanistically the reaction is believed to proceed as shown for the uracil case, the essential point being that nucleophilic addition to the activated 5,6-double bond, to give **13**, initiates the cleavage reaction; a ureidopyrazolidone intermediate (**14**) has been isolated in one instance [22]. Discussion of the further consequences of



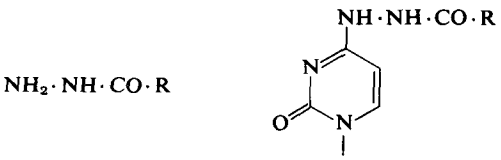
13 **14** **15**



17 **16** **18**

this cleavage reaction for polynucleotide degradation is deferred until later (Section IV,D,3). More dilute hydrazine solutions and those buffered in the

region of pH 8 are erratic reducing agents converting uridine to dihydrouridine [23]. The major reaction at pH 6 with cytosines is the displacement at N-4 generating an N_4-aminocytosine (**17**) and even the *N,N'*-bispyrimidinylhydrazine (**18**) [24], uridine and thymidine apparently being stable [25]. Other *N*-alkyl hydrazines also react in this way, but of particular interest are a series of *N*-acyl derivatives, *inter alia* **19** (R = CH_3, NH_2, $CH_2CO_2^-$, and 4-pyridyl). These do not react with uridine or thymidine, nor with the purine nucleosides, but with cytosine derivatives give the corresponding 4-hydrazides (**20**) [26–30]. The pH dependence (rate maxima in the pH range 4–5) suggests that the reaction may have value in polynucleotide modification, though the reactions are disadvantageously slow, $t_{1/2}$ = 50 hr with **19** (R = $CH_2NC_5H_5^+$). On the other hand, it is noted that ribonuclease



19 **20**

hydrolysis of cytidine-2',3'-phosphate after modification with **19** (R = $CH_2NC_5H_5^+$) is one-thirtieth of that of the parent cyclic phosphate. The application to polynucleotide sequence studies is obvious [31].

It is not clear whether the N_4-substitution products are the primary reaction products or whether an intermediate 5,6-double bond adduct is formed initially in these reactions. The significance of this point will become clearer when the related reaction with hydroxylamine is discussed.

C. Reactions with Hydroxylamines

Hydroxylamine resembles hydrazine in that, although it is a relatively weak base (pK_a, 6.5), it is an exceptionally powerful nucleophile. In this connection the so-called "α-effect" has been discussed extensively [32]. Alkoxyamines (pK_a of CH_3ONH_2 is 5.8), too, are very reactive, though less so than the parent base. Since it is important that reactions with nucleic acids should be conducted near neutrality, the low pK_a values of oxyamines allow a high concentration of the free base to be present in these circumstances. Hydrazine and hydroxylamine and their *N*-substituted derivatives are unstable. It is likely that many of their biological effects are to be ascribed to reactions resulting from radical production and not only from the chemical processes discussed here [33,34]. Alkoxyamines in general are much more stable.