PROGRESS IN
Nucleic Acid Research and Molecular Biology

edited by

J. N. DAVIDSON
Department of Biochemistry
The University of Glasgow
Glasgow, Scotland

•

WALDO E. COHN
Biology Division
Oak Ridge National Laboratory
Oak Ridge, Tennessee

Volume 12

1972

ACADEMIC PRESS
New York and London

List of Contributors

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

EKKEHARD K. F. BAUTZ (129), *Institute for Molecular Genetics, University of Heidelberg, Heidelberg, Germany*

JACQUES R. FRESCO (1), *Department of Biochemical Sciences, Frick Chemical Laboratory, Princeton University, Princeton, New Jersey*

MEHRAN GOULIAN (29), *Department of Medicine, University of California San Diego, La Jolla, California*

NATHAN DE GROOT (189), *Department of Biological Chemistry, The Hebrew University of Jerusalem, Jerusalem, Israel*

YEHUDA LAPIDOT (189), *Department of Biological Chemistry, The Hebrew University of Jerusalem, Jerusalem, Israel*

M. LASKOWSKI, SR. (161), *Laboratory of Enzymology, Roswell Park Memorial Institute, Buffalo, New York*

ROBERT B. LOFTFIELD (87), *Department of Biochemistry, The University of New Mexico, School of Medicine, Albuquerque, New Mexico*

A. J. LOMANT (1), *Department of Biochemical Sciences, Frick Chemical Laboratory, Princeton University, Princeton, New Jersey*

SUSUMU NISHIMURA (49), *National Cancer Center Research Institute, Chuo-ku, Tokyo, Japan*

Preface

Volume 12 of *Progress in Nucleic Acid Research and Molecular Biology* includes seven essays covering a wide spectrum in the nucleic acid field. We believe that our readers will find the contributions of topical interest and importance. They follow our usual pattern of attempting to present "essays in circumscribed areas" in which recent developments in particular aspects of the field of nucleic acids and molecular biology are discussed by workers provided with an opportunity for more personal expression of points of view that may be individualistic and perhaps even controversial. We have not attempted to define or restrict any author's approach to his chosen subject, and have confined our editing to ensuring maximum clarity to the reader, whom we envisage to be a person himself active in or concerned with the general field of nucleic acids or molecular biology. Needless to say, we do not necessarily share all the opinions or concepts of all the authors and accept no responsibility for them. We seek rather to provide a forum for discussion and debate, and we welcome further suggestions from readers as to how this end may best be served. Indeed, we should like again to remind readers that we wish them to write to us with their comments.

Abbreviations and symbols used for nucleic acids and their derivatives are now well established by the authority of the Combined Commission on Biochemical Nomenclature (CBN) of the International Union of Biochemistry (IUB) and the International Union of Pure and Applied Chemistry (IUPAC). Those pertinent to our subject are not usually listed at the beginning of each chapter, but are listed on the following pages.

In this volume, in the interests of conserving space and reducing costs, we have made an innovation by simplifying the contractions for the titles of some of the most commonly cited journals. A list of these is appended to the section on abbreviations and symbols. We hope that this new usage will be acceptable to our authors and readers alike.

J. N. D.
W. E. C.

Abbreviations and Symbols

All contributors to this Series are asked to use the terminology (abbreviations and symbols) recommended by the IUPAC–IUB Commission on Biochemical Nomenclature (CBN) and approved by IUPAC and IUB, and the Editors endeavor to assure conformity. These Recommendations have been published in many journals (*1*) and compendia (*2*) in four languages and are available in reprint form from the NAS–NRC Office of Biochemical Nomenclature (OBN), as stated in each publication, and are therefore considered to be generally known. Those used in nucleic acid work, originally set out in section 5 of the above Recommendations (*1*) and recently revised and expanded (*2*, *3*), are given in condensed form (I–V) below for the convenience of the reader.

I. Bases, Nucleosides, Mononucleotides

1. *Bases* (in tables, figures, equations, or chromatograms) are symbolized by Ade, Gua, Hyp, Xan, Cyt, Thy, Oro, Ura; Pur = any purine, Pyr = any pyrimidine, Base = any base. The prefixes S–, H_2, F–, Br, Me, etc., may be used for modifications of these.

2. *Ribonucleosides* (in tables, figures, equations, or chromatograms) are symbolized, in the same order, by Ado, Guo, Ino, Xao, Cyd, Thd, Ord, Urd (Ψrd), Puo, Pyd, Nuc. Modifications may be expressed as indicated in (1) above. Sugar residues may be specified by the prefixes r (optional), d (=deoxyribo), a, x, l, etc., to these, or by two three-letter symbols, as in Ara-Cyt (for aCyd) or dRib-Ade (for dAdo).

3. *Mono-, di-, and triphosphates of nucleosides* (5′) are designated by NMP, NDP, NTP. The N (for "nucleoside") may be replaced by any one of the nucleoside symbols given in II-1 below. 2′-, 3′-, and 5′- are used as prefixes when necessary. The prefix d signifies "deoxy." [Alternatively, nucleotides may be expressed by attaching P to the symbols in (2) above. Thus: P-Ado = AMP; Ado-P = 3′-AMP.] cNMP = cyclic 3′:5′-NMP.

II. Oligonucleotides and Polynucleotides

1. Ribonucleoside Residues

(a) Common: A, G, I, X, C, T, O, U, Ψ, R, Y, N (in the order of I-2 above).

(b) Base-modified: sI or M for thioinosine = 6-mercaptopurine ribonucleoside; sU or S for thiouridine; brU or B for 5-bromouridine; hU or D for 5,6-dihydrouridine; i for isopentenyl; f for formyl. Other modifications are similarly indicated by appropriate *lower-case* prefixes (in contrast to I-1 above) (*2*, *3*).

(c) Sugar-modified: prefixes are d, a, x, or l as in I-2 above; alternatively, by *italics* or **boldface** type (with definition) unless the entire chain is specified by an appropriate prefix. The 2′-*O*-methyl group is indicated by *suffix* m (e.g., -Am- for 2′-*O*-methyladenosine, but -mA- for *N*-methyladenosine).

(d) Locants and multipliers, when necessary, are indicated by superscripts and subscripts, respectively, e.g., -m_2^6A- = 6-dimethyladenosine; -s^{4}U- or -^{4}S- = 4-thiouridine; -ac^{4}Cm- = 2′-*O*-methyl-4-acetylcytidine.

(e) When space is limited, as in two-dimensional arrays or in aligning homologous sequences, the prefixes may be placed *over the capital letter*, the suffixes *over the phosphodiester symbol.*

2. Phosphoric Acid Residues [left side = 5′, right side = 3′ (or 2′)]

(a) Terminal: p; e.g., pppN . . . is a polynucleotide with a 5′-triphosphate at one end; Ap is adenosine 3′-phosphate; C > p is cytidine 2′:3′-cyclic phosphate (*1*, *2*, *3*).

(b) Internal: hyphen (for known sequence), comma (for unknown sequence); unknown sequences are enclosed in parentheses. E.g., pA-G-A-C(C_2,A,U)A-U-G-C > p is a sequence with a (5′) phosphate at one end, a 2′:3′-cyclic phosphate at the other, and a tetranucleotide of unknown sequence in the middle. (Only codon triplets are written without some punctuation separating the residues.)

3. Polarity, or Direction of Chain

The symbol for the phosphodiester group (whether hyphen or comma or parenthesis, as in 2b) represents a 3′-5′ link (i.e., a 5′ . . . 3′ chain) unless otherwise indicated by appropriate numbers. "Reverse polarity" (a chain proceeding from a 3′ terminus at left to a 5′ terminus at right) may be shown by numerals or by right-to-left arrows. Polarity in any direction, as in a two-dimensional array, may be shown by appropriate rotation of the (capital) letters so that 5′ is at left, 3′ at right when the letter is viewed right-side-up.

4. Synthetic Polymers

The complete name or the appropriate group of symbols (see II-1 above) of the repeating unit, enclosed in parentheses if complex or a symbol, is either (a) preceded by "poly," or (b) followed by a subscript "n" or appropriate number. *No space follows "poly"* (*2*,*5*).

The conventions of II-2b are used to specify known or unknown (random) sequence, e.g.,

polyadenylate = poly(A) or $(A)_n$, a simple homopolymer;

poly(3 adenylate, 2 cytidylate) = poly(A_3C_2) or $(A_3,C_2)_n$, a *random* copolymer of A and C in 3:2 proportions;

poly(deoxyadenylate-deoxythymidylate) = poly[d(A-T)] or poly(dA-dT) or $(dA\text{-}dT)_n$ or $d(A\text{-}T)_n$, an *alternating* copolymer of dA and dT;

poly(adenylate, guanylate, cytidylate, uridylate) = poly(A,G,C,U) or $(A,G,C,U)_n$, a random assortment of A, G, C, and U residues, proportions unspecified.

The prefix copoly or oligo may replace poly, if desired. The subscript "n" may be replaced by numerals indicating actual size.

III. Association of Polynucleotide Chains

1. *Associated* (e.g., H-bonded) chains, or bases within chains, are indicated by a *center dot* (not a hyphen or a plus sign) separating the *complete* names or symbols, e.g.:

poly(A)·poly(U) or $(A)_n\cdot(U)_m$

poly(A)·2 poly(U) or $(A)_n\cdot 2(U)_m$

poly(dA-dC)·poly(dG-dT) or $(dA\text{-}dC)_n\cdot(dG\text{-}dT)_m$;

also, "the adenine·thymidine base-pair" or "A·T base-pair" in text.

2. *Nonassociated* chains are separated by the plus sign, e.g.:

$2[\text{poly(A)}\cdot\text{poly(U)}] \xrightarrow{\Delta} \text{poly(A)}\cdot 2\ \text{poly(U)} + \text{poly(A)}$ (II-4a)

or $2[A_n\cdot U_m] \xrightarrow{\Delta} A_n\cdot 2U_m + A_n$ (II-4b).

3. Unspecified or unknown association is expressed by a comma (again meaning "unknown") between the completely specified residues.

Note: In all cases, each chain is completely specified in one or the other of the two systems described in II-4 above.

IV. Natural Nucleic Acids

RNA	ribonucleic acid or ribonucleate
DNA	deoxyribonucleic acid or deoxyribonucleate
mRNA; rRNA; nRNA	messenger RNA; ribosomal RNA; nuclear RNA
D-RNA; cRNA	"DNA-like" RNA; complementary RNA
mtDNA	mitochondrial DNA
tRNA	transfer (or acceptor or amino acid-accepting) RNA; replaces sRNA, which is not to be used for any purpose
aminoacyl-tRNA	"charged" tRNA (i.e., tRNA's carrying aminoacyl residues); may be abbreviated to AA-tRNA
alanine tRNA or $tRNA^{Ala}$, etc.	tRNA normally capable of accepting alanine, to form alanyl-tRNA
alanyl-tRNA or alanyl-$tRNA^{Ala}$	The same, with alanyl residue covalently attached. [*Note:* fMet = formylmethionyl; hence $tRNA^{fMet}$ or $tRNA_f^{Met}$]

Isoacceptors are indicated by appropriate subscripts, i.e., $tRNA_1^{Ala}$, $tRNA_2^{Ala}$, etc.

V. Miscellaneous Abbreviations

P_i, PP_i	inorganic orthophosphate, pyrophosphate
RNase, DNase	ribonuclease, deoxyribonuclease

Others listed in Table II of Reference 1 may also be used without definition. No others, with or without definition, are used unless, in the opinion of the editors, they increase the ease of reading.

Enzymes

In naming enzymes, the recommendations of the IUB Commission on Enzymes, approved by IUB in 1964 (*4*), are followed as far as possible. At first mention, each enzyme is described *either* by its systematic name *or* by the equation for the reaction catalyzed, followed by its EC number in parentheses. Subsequent mention may use a trivial name. Enzyme names are not to be abbreviated except when the substrate has an approved abbreviation (e.g., ATPase, but not LDH, is acceptable).

References*

1. *JBC* **241,** 527 (1966); *Bchem* **5,** 1445 (1966); *BJ* **101,** 1 (1966); *ABB* **115,** 1 (1966), **129,** 1 (1969); and elsewhere.†
2. *EJB* **15,** 203 (1970); *JBC* **245,** 5171 (1970); *JMB* **55,** 299 (1971); and elsewhere.†
3. "Handbook of Biochemistry" (H. A. Sober, ed.), 2nd ed. Chemical Rubber Co., Cleveland, Ohio, 1970, Section A and pp. H130–133.
4. "Enzyme Nomenclature," Elsevier Publ. Co., New York, 1965. [Revision under construction.]
5. "Nomenclature of Synthetic Polypeptides," *JBC* **247,** 323 (1972); *Biopolymers* **11,** 321 (1972); and elsewhere.†

* Contractions for names of journals follow.

† Reprints of all CBN Recommendations are available from the Office of Biochemical Nomenclature (W. E. Cohn, Director), Biology Division, Oak Ridge National Laboratory, Box Y, Oak Ridge, Tennessee 37830, USA.

Abbreviations of Journal Titles

Journals	*Abbreviations used*
Annu. Rev. Biochem.	ARB
Arch. Biochem. Biophys.	ABB
Biochem. Biophys. Res. Commun.	BBRC
Biochemistry	Bchem
Biochem. J.	BJ
Biochim. Biophys. Acta	BBA
Cold Spring Harbor Symp. Quant. Biol.	CSHSQB
Eur. J. Biochem.	EJB
Fed. Proc.	FP
J. Amer. Chem. Soc.	JACS
J. Bacteriol.	J. Bact.
J. Biol. Chem.	JBC
J. Chem. Soc.	JCS
J. Mol. Biol.	JMB
Proc. Nat. Acad. Sci. U.S.	PNAS
Proc. Soc. Exp. Biol. Med.	PSEBM

Some Articles Planned for Future Volumes

Transcription and Translation in Mitochondria
W. E. BARNETT

Properties of Dissected tRNA Molecules
A. A. BAYEV AND A. D. MIRZABEKOV

Gene Regulation in Higher Cells
R. J. BRITTEN

Synthesis and Function of the —CCA Terminus of Transfer RNA
M. DEUTSCHER

Reactions of Formaldehyde with Nucleic Acids
M. Y. FELDMAN

RNA-Dependent DNA Polymerase
M. GREEN

RNA and Immunity
A. A. GOTTLIEB

Initiation Factors in Protein Synthesis
M. GRUNBERG-MANAGO

Immunogenic Polynucleotides
L. D. HAMILTON

Mammalian RNA Polymerase
SAMSON T. JACOB

X-Ray Diffraction Studies of Nucleic Acids
R. LANGRIDGE

Supercoiled DNA Molecules in Biological Structures
M. F. MAESTRE

Mechanism in Polypeptide Synthesis on Ribosomes
H. MATTHAEI AND E. BERMEK

The Influence of Steroid Hormones on Genetic Expression
G. C. MUELLER AND B. VONDERHAAR

Elucidation of Nucleic Acid Structure by Resonance Spectroscopy
W. D. PHILLIPS AND C. C. MCDONALD

Transmissible Drug Resistance
M. H. RICHMOND

Deoxycytidylate Aminohydrolase in DNA Synthesis
E. SCARANO

The Stereochemistry of Actinomycin Binding to DNA
H. M. SOBELL

Initiation of Protein Synthesis
R. E. THACH

Structure and Function of a Viral RNA
C. WEISSMANN

Erratum

Volume 11

page 307, 4th line. Sentence should read: Yoshida *et al.* (*4a*) located a 2-thiouridine derivative at position 1 of the anticodon in a glutamic acid tRNA which bound GAA, but not GAG (see footnote *b* to Table II).

Contents

Ultraviolet Photochemistry as a Probe of Polyribonucleotide Conformation

A. J. Lomant and Jacques R. Fresco

Some Recent Developments in DNA Enzymology

Mehran Goulian

Minor Components in Transfer RNA: Their Characterization, Location, and Function

SUSUMU NISHIMURA

The Mechanism of Aminoacylation of Transfer RNA

ROBERT B. LOFTFIELD

Regulation of RNA Synthesis

EKKEHARD K. F. BAUTZ

The Poly(dA-dT) of Crab

M. LASKOWSKI, SR.

The Chemical Synthesis and the Biochemical Properties of Peptidyl-tRNA

YEHUDA LAPIDOT AND NATHAN DE GROOT

Ultraviolet Photochemistry as a Probe of Polyribonucleotide Conformation[1,2]

A. J. LOMANT AND
JACQUES R. FRESCO

Department of Biochemical Sciences,
Frick Chemical Laboratory,
Princeton University,
Princeton, New Jersey

[1] This review was abstracted from the Ph.D. thesis of A. J. L., Princeton University, 1971. Work original with the authors was supported by grants from the National Institutes of Health (GM-07654), the National Science Foundation (GB-18865), and the American Heart Association. A. J. L. was a National Institutes of Health predoctoral trainee (GM-00962).

[2] Abbreviations: $\overset{*}{\text{U}}$ and $\overset{*}{\text{C}}$, photohydrates of uridine and cytidine, respectively, or of their nucleotide residues; $\widehat{\text{UU}}$ and $\widehat{\text{CC}}$, cyclobutane-type dimer photoproducts of uracil (or derivatives) and cytosine (or derivatives), respectively; $\widehat{\text{C-C}}$, $\widehat{\text{U-U}}$, and $\text{d}(\widehat{\text{U-U}})$, cyclobutane-type dimer photoproducts of C-C, U-U, and dU-dU; σ, cross section for formation of the photoproduct indicated by the subscript; s^4U, 4-thiouridine residue; fl^5U, 5-fluorouridine residue; br^5U, 5-bromouridine residue; m^7G, 7-methylguanosine residue. For other symbols, see the preface to this volume.

I. Introduction

Exposure of nucleic acids to small doses of near-ultraviolet radiation generally brings about chemical changes only in their pyrimidine moieties. The photochemistry of nucleic acids has been discussed in several reviews (*1–6*), but an analysis has not been made of the conformational determinants of the photoreactivity of pyrimidine residues. Yet, selective photochemical modification of complex polynucleotides appears to occur only under particular conditions of solvent, temperature, wavelength, and exposure. This review has been written, therefore, with the aim of providing a background for future refinement of ultraviolet photochemistry as a probe of nucleic acid conformation. Attention is directed primarily toward studies with poly*ribo*nucleotides. DNA photochemistry has been treated in the earlier reviews, and studies with model polydeoxynucleotides are few. On the other hand, an extensive literature on model polyribonucleotides has accumulated.

II. Mono- and Dinucleotides

A. Solvent-Addition Photoproducts

In 1949, Sinsheimer and Hastings (*7*) found conditions that favor the formation of a uracil photoproduct reversible by heat or acid. Subsequently, several workers (*8–11*) showed indirectly that the principal photoproduct obtained from uracil solutions irradiated with UV light of wavelength longer than 230 nm is a water adduct across the C5-C6 double bond, namely, 6-hydroxy-5,6-dihydrouracil. Equivalent photoreactions occur in the nucleoside and nucleotide. Miller and Cerutti (*12*) later unequivocally

confirmed the position of the hydroxyl group by reducing the nucleoside photoproduct with sodium borohydride, obtaining ribosylurea as the principal glycosyl compound.

In the presence of other nucleophiles, irradiation of uracil can lead to alternative photoproducts. Thus, HCN (*13*), ethanol (*14*), and amino acids (*15*) form adducts at the C5-C6 double bond. Addition is at the 5 position in the case of cysteine (yielding 5-*S*-cysteine-6-hydrouracil), and probably occurs there as well for the alcohol-addition products. Uracil in the excited state is also susceptible to attack by hydride ion, leading to dihydrouracil (*16*).

Numerous studies suggest that the water-addition photoproduct, or photohydrate, originates from the first excited singlet state, as hydrate formation is insensitive to the presence of nonnucleophilic triplet quenchers (*17–21*). On the basis of uracil photohydrate yield as a function of pH, Burr, Gordon, and Park (*22*) inferred that the excited singlet is first protonated before reacting with water to yield the photohydrate. Since the rate of uracil photohydration is independent of ionic strength, it seems that the reactive species of water is the neutral molecule, not the hydroxide ion.

Formation of a reversible photoproduct of cytosine and of its nucleoside and nucleotide derivatives was also first detected by Sinsheimer (*23*). Subsequent investigators (*24–28*) established that an addition of water across the C5-C6 double bond, analogous to the photohydration of uracil, occurs when cytosine is irradiated in dilute buffer at low temperature (4–6°C). Miller and Cerutti (*12*) reduced the cytidine photohydrate with sodium borodeuteride and obtained a rearranged alkylated glycosylurea with deuterium at a position consistent with the hydroxyl group of the photoproduct being originally at position 6. Because of the reduced stability of the photoadduct and its facile deamination to the uracil photohydrate, Grossman and Rodgers (*29*) sought to measure the formation of cytosine photohydrate by exchange of tritium from [5-^{3}H]cytosine with water after dehydration of the photoproduct [cf. (*30*)]. They reported quantitative release of radioactivity for each photohydrate formed from irradiated [5-^{3}H]CMP, a finding that is difficult to explain without invoking steric bias for the proton released. DeBoer and Johns (*31*) subsequently showed that nearly 15% of the tritium is lost rapidly, probably reflecting an exchange in the excited state. Consequently, the exchange technique is inadequate for determining cytosine photohydrate formation.

Alternative methods for the quantitation of photoproduct formation and for the measurement of the rates of photoproduct deamination and reversal were developed by Johns, LeBlanc, and Freeman (*32*), who employed high voltage electrophoresis to separate rapidly the photoproducts

($\overset{*}{\text{C}}$p and $\overset{*}{\text{U}}$p) and the starting material [^{32}P]Cp. They found that the extent of deamination of $\overset{*}{\text{C}}$p is independent of temperature, and is maximal at alkaline pH (minimum at pH 5). Photoproduct reversal to Cp is fastest at pH 5, however, whereas formation of $\overset{*}{\text{C}}$p is greatest at neutral pH (*33*). Using absorbance measurements, DeBoer, Klinghoffer, and Johns (*34*) carefully studied the reversal of cytosine, cytidine, and cytidylate photohydrates. Reversal is consistent with a model involving general base catalysis, proceeding at faster rates if buffers of higher pK are present. The reversal of $\overset{*}{\text{C}}$p results in part from internal catalysis by the phosphate charge.

The excited-state precursor of cytosine photohydrate is as yet undetermined. However, N^4-acetylcytosine fluoresces at room temperature and is subject to nucleophilic attack of the excited state analogous to that undergone by the uracil singlet. Photochemical reduction of N^4-acetylcytosine by hydride ion to N^4-acetyl-3,4,5,6-tetrahydrocytosine (*35*) probably proceeds from the excited singlet, so it is not unreasonable to consider the cytosine excited singlet as the photohydrate precursor.

B. Dimer Photoproducts of the Cyclobutane Type

The formation and identification of a cyclobutane photodimer of thymine was first reported by Beukers and Berends (*36*). Uracil dimers were subsequently isolated by the same technique of irradiation of frozen solutions (*37*). These dimers exhibit short-wavelength reversal, absorption spectra, and molecular structure properties similar to those of thymine photodimers (*10, 38, 39*). By analogy with the photosensitized dimerization of thymine and thymidine, uridine solutions containing acetone and irradiated at energies below the first excited singlet of uracil form dimers exclusively (*40*), so that the immediate precursor of $\widehat{\text{UU}}$ is the first excited triplet. Lamola and Mittal (*41*) had earlier suggested an involvement of the excited singlet as a direct precursor of the uracil photodimer on the basis of experiments in which uracil in acrylonitrile was irradiated in the presence of isoprene, a triplet quencher. Their result, however, may represent incomplete trapping of the triplet state energy.[4] Eisinger and Lamola (*42*) and others (*43*) have advanced a model for the photodimerization of thymine, orotic acid, and uracil in which excitation first yields the singlet excimer followed by intersystem crossing to the lower-lying triplet. Whillans,

[3] See Addendum.

[4] On the other hand, in highly concentrated solutions of *thymine*, where aggregates are formed, $\widehat{\text{TT}}$ probably does originate from an excited singlet (*135*).