PROGRESS IN
Nucleic Acid Research and Molecular Biology

edited by

WALDO E. COHN

Volume 21

PROGRESS IN
Nucleic Acid Research and Molecular Biology

edited by

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

Volume 21

1978

ACADEMIC PRESS
New York San Francisco London
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 N.V1 7DX

LIBRARY OF CONGRESS CATALOG CARD NUMBER: 63–15847

ISBN 0–12–540021–7

PRINTED IN THE UNITED STATES OF AMERICA

List of Contributors

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

CHANDER P. BAHL (101), *Section of Biochemistry, Molecular and Cell Biology, Cornell University, Ithaca, New York*

W. EDGAR BARNETT (143), *Biology Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee*

ENGIN BERMEK (63), *Biofizik Kürsüsü, Tip Fakültesi, Istanbul Üniversitesi, Istanbul, Turkey*

L. I. HECKER* (143), *Biology Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee*

SUSAN D. MORGAN (181), *Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut*

SARAN A. NARANG (101), *Division of Biological Sciences, National Research Council of Canada, Ottawa, Ontario, Canada*

A. A. PREOBRAZHENSKY (1), *A. N. Bakh Institute of Biochemistry, Academy of Sciences of the USSR, Moscow, USSR*

S. D. SCHWARTZBACH (143), *Biology Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee*

DIETER SÖLL (181), *Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut*

A. S. SPIRIN (1, 39), *Institute of Protein Research, Academy of Sciences of the USSR, Poustchino, Moscow Region, USSR*

RAY WU (101), *Section of Biochemistry, Molecular and Cell Biology, Cornell University, Ithaca, New York*

* Present address: Chemical Carcinogenesis Group, Frederick Cancer Research Center, P. O. Box B, Frederick, Maryland.

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 Editor endeavors to assure conformity. These Recommendations have been published in many journals *(1, 2)* and compendia *(3)* in four languages and are available in reprint form from the 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 first Recommendations *(1)* and subsequently revised and expanded *(2, 3)*, are given in condensed form (I–V) below for the convenience of the reader. Authors may use them without definition, when necessary.

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; Bt_2cAMP = dibutyryl cAMP, etc.

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 6-methyladenosine).

(d) Locants and multipliers, when necessary, are indicated by superscripts and subscripts, respectively, e.g., -m^6_2A- = 6-dimethyladenosine; -s^4U- 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)*; p < A is adenosine 3′ : 5′-cyclic phosphate.

(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 parentheses, 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.

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 $(A_3,C_2)_n$, an *irregular* 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, e.g., $A_n \cdot dT_{12-18}$.

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 $A_n \cdot 2U_m$

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

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

$2[\text{poly(A)} \cdot \text{poly(U)}] \rightarrow \text{poly(A)} \cdot 2\,\text{poly(U)} + \text{poly(A)}$

or $2[A_n \cdot U_m] \rightarrow A_n \cdot 2U_m + A_n$.

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

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
hnRNA	heterogeneous 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 $tRNA^{fMet}$, identical with $tRNA_f^{Met}$]

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
t_m (not T_m)	melting temperature (°C)

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 editor, they increase the ease of reading.

Enzymes

In naming enzymes, the 1972 recommendations of the IUPAC-IUB Commission on Biochemical Nomenclature (CBN) *(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 *or* by the recommended trivial name, followed by its EC number in parentheses. Thereafter, a trivial name may be used. 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" (G. Fasman, ed.), 3rd ed. Chemical Rubber Co., Cleveland, Ohio, 1970, 1975, Nucleic Acids, Vols. I and II, pp. 3–59.
4. "Enzyme Nomenclature," Elsevier Scientific Publ. Co., Amsterdam, 1973, and Supplement No. 1, *BBA* **429,** (1976).

* 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.

5. "Nomenclature of Synthetic Polypeptides," *JBC* **247**, 323 (1972); *Biopolymers* **11**, 321 (1972); and elsewhere.*

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
Nature, New Biology	Nature NB
Nucleic Acid Research	NARes
Proc. Nat. Acad. Sci. U.S.	PNAS
Proc. Soc. Exp. Biol. Med.	PSEBM
Progr. Nucl. Acid Res. Mol. Biol.	This Series

* 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.

Some Articles Planned for Future Volumes

Chromatin Structure and Function
P. Chambon

Transcription Units for mRNA Production in Eukaryotic Cells and Their DNA Viruses
J. E. DARNELL, JR.

Functional Aspects of the Interaction of Chemical Carcinogens with Nucleic Acids
D. GRUNBERGER AND I. B. WEINSTEIN

Ribonucleotide Reductase
F. D. HAMILTON

Genetically Controlled Variation in the Conformation of Enzymes
G. B. JOHNSON

Mechanism of Interferon Action
P. LENGYEL AND G. SEN

N-Glycosylases and Endonucleases for Apurinic/Apyrimidinic Sites in DNA
T. LINDAHL

Mitochondrial Nucleic Acids of Yeast
M. RABINOWITZ AND J. LOCKER

Ribosome Turnover in Eukaryotic Cells and Tissues
J. F. SCOTT

Physical Structure, Chemical Modification and Functional Role of the Acceptor Terminus of tRNA
M. SPRINZL AND F. CRAMER

Naturally Occurring Nucleoside Antibiotics
R. SUHADOLNIK

The Biochemical and Microbiological Action of Platinum Compounds
A. J. THOMSON AND J. J. ROBERTS

Patterns of Nucleic Acid Synthesis in *Physarum polycephalum*
G. TURNOCK

Structure and Functions of Ribosomal RNA
R. ZIMMERMANN

Contents

Synthetic Oligodeoxynucleotides for Analyses of DNA Structure and Function

RAY WU, CHANDER P. BAHL, AND SARAN A. NARANG

The Transfer RNAs of Eukaryotic Organelles

W. EDGAR BARNETT, S. D. SCHWARTZBACH, AND L. I. HECKER

Regulation of the Biosynthesis of Aminoacid:tRNA Ligases and of tRNA

SUSAN D. MORGAN AND DIETER SÖLL

Informosomes and Their Protein Components: The Present State of Knowledge

A. A. PREOBRAZHENSKY
AND
A. S. SPIRIN

A. N. Bakh Institute of Biochemistry
Academy of Sciences of the USSR, Moscow, and
Institute of Protein Research
Academy of Sciences of the USSR
Poustchino, Moscow Region, USSR

I. Introduction

In 1964, free ribonucleoprotein particles containing nonribosomal RNA (presumably mRNA) were discovered in the cytoplasm of animal cells (*1*; see also 2); they were called *informosomes*. Later, mRNA · protein complexes (messenger ribonucleoproteins) were released from polyribosomes of animal cells (*3,4*). In 1965, the presence of ribonucleoprotein particles containing hnRNA in animal nuclei was shown (*5, 6*). Recently, analogous ribonucleoprotein particles of the informosome type have been found in cytoplasm, polyribosomes, and nuclei of higher plants (*7–9*). Thus it has been established that mRNA at all stages of its lifetime in eukaryotic cells is present in the form of ribonucleoprotein complexes and, consequently, always bound to certain proteins.

It is very likely that these ribonucleoprotein particles, different in their cellular localization, differ also in their protein composition and functional specialization. Nonetheless, their common feature is that they contain either mRNA or mRNA precursor, i.e., informational RNA in the broad sense of the word. They also display a number of very characteristic common group properties: (a) an RNA/protein ratio of about 1 : 3 and a corresponding buoyant density in CsCl of about 1.4 g/cm^3; (b) a linear correlation between the sedimentation coefficients of the particles and the sedimentation coefficients of the RNA contained in them; (c) a high sensitivity to ribonucleases despite the large amount of protein; (d) resistance to removal of Mg^{2+} from the medium and to EDTA treatment; etc. (see review in *10*).

Since the discovery of informosomes and the other ribonucleoprotein particles of the informosome type, several review papers have appeared on this subject (*2, 10–16*). In the past few years, many new experimental results have been published, the most important point being that the protein moiety of the ribonucleoprotein particles was investigated with especial attention. In this review, the most recent data on ribonucleoprotein particles of the informosome type, in particular the data on their protein components, are surveyed.

II. Free Cytoplasmic Informosomes

A. Universal Distribution and Physicochemical Properties

Initially, informosomes were found in the cytoplasm of differentiating fish embryo cells (*1*). Later studies showed that informosomes are universal particles intrinsic to all cells of the eukaryotic organisms studied (see reviews, *2, 10, 12*).

Informosomes have two universal physicochemical properties.

1. The sedimentation coefficient of informosomes depends on the sedimentation coefficient of the RNA contained within them, exceeding it by 2–2.5 times *(17–22)*. A consequence of this is the fact that, where mRNA is heterogeneous in molecular size, informosomes display a heterogeneous sedimentation distribution, whereas informosomes are homogeneous if only one main class of mRNA is present.

2. Informosomes have comparatively low buoyant density values in CsCl, from 1.46 g/cm^3 to 1.36 g/cm^3. Several density components of informosomes are often observed: one major component and some minor ones. The density of the major component for informosomes from the most varied sources is usually about 1.4 g/cm^3 *(1, 4, 7, 18, 19, 21, 23–38)*.

B. Informosomes Exist *in Vivo*

Shortly after informosomes were discovered, special proteins that can form complexes with isolated RNA *in vitro* were found in animal cell extracts. The artificial ribonucleoprotein particles formed did not differ from informosomes either in sedimentation characteristics (the ratio of sedimentation coefficient of particles to that of RNA) or in buoyant density *(39, 157)*. (A more detailed description of these proteins is given in Section V.) The finding of "informosome-forming" proteins posed the question whether informosomes indeed exist *in vivo* or were formed from free mRNA of the cytoplasm and these proteins during the cell disruption procedure.

The hypothesis of an artifactual origin of informosomes was disproved by experiments in which cells were disrupted in the presence of exogenous RNA competing for the informosome-forming proteins. The presence of such a competitor did not in any way affect the sedimentation and density characteristics of informosomes *(2, 26, 33)*. This evidence for the existence of informosomes in cells was confirmed by further comparison of the properties of the artificial complexes with those of informosomes isolated as such [*(40)*; see Section V].

C. Chemical Composition of Informosomes

The following facts indicate that informosomes contain RNA:

1. Informosomes become radioactive upon incubation of cells with radioactive uridine and other RNA precursors (this fact is established in all the papers cited above and here).

2. Informosomes are sensitive to ribonucleases *(7, 17, 29, 44)*.

3. RNA can be isolated from informosomes by standard extraction procedures *(17, 18, 29, 33, 42)*.

The following facts indicate the presence of protein in informosomes:

1. Informosomes become radioactive upon incubation of cells with radioactive amino acids *(1, 18, 23, 41, 42)*.

2. Informosomes are sensitive to proteolytic enzymes *(17, 23, 29, 33)*.

3. Informosomes are adsorbed on nitrocellulose filters under conditions in which neither RNA nor the bulk of proteins, but only ribonucleoproteins, are adsorbed *(17, 18, 43)*.

4. Their buoyant density in CsCl, 1.4 g/cm^3, is intermediate between those of RNA (1.9–2.0 g/cm^3) and the protein (1.2–1.3 g/cm^3).

5. A set of polypeptides can be directly extracted from informosome preparations (see below).

DNA is not contained in informosomes *(1, 41, 45, 46)*.

The presence of lipids in informosomes has not been shown *(12)*. The question of the presence of polysaccharides has not been studied.

If it is assumed that informosomes consist only of RNA and protein, one can estimate the protein/RNA ratio within the particles from the buoyant density in CsCl on the basis of the empirical formula *(2)* without isolating pure informosomes:

$$\text{protein } \% = (1{,}85 - \rho)/0.006$$

where ρ is the buoyant density of particles in CsCl. Hence, a buoyant density of 1.4 g/cm^3 indicates particles that are 75% protein and 25% RNA.

D. Nature of RNA within Informosomes

The RNA within free informosomes of the cytoplasm is mostly mRNA.

1. Numerous earlier results suggested this indirectly. Thus, a rapid labeling of RNA within informosomes by radioactive precursors was noted in all the papers cited above. The RNA of informosomes differed from both ribosomal and transfer RNA in sedimentation distribution *(17, 18, 22, 24, 29)*. In certain early embryos, the RNA of informosomes is synthesized in the absence of ribosomal RNA synthesis *(1, 7, 47)*. The RNA of informosomes has a DNA-like base composition *(29, 33, 34, 48)*. The RNA of informosomes is highly hybridizable with DNA *(42, 49)*.

2. The rapidly labeled RNA of informosomes from the animal cells infected with RNA viruses is viral mRNA *(19, 20, 28, 50)*. The rapidly labeled RNA of informosomes of cells infected with DNA viruses appears also to be viral mRNA *(23)*.

3. The messenger nature of RNA of informosomes was strongly supported by experiments in which a transition of this RNA into translating polyribosomes was observed *(20, 31, 43, 51–56a)*.

4. Preparations of informosomes, or RNA extracted from them, stimulate the incorporation of amino acids into proteins in a cell-free system of protein synthesis *(1, 29, 48)*. In the case of informosomes of rabbit and duck reticulocytes, the products synthesized in a cell-free system have been identified as globin chains *(22, 57–61)*. A protein identified as actin was synthesized in a cell-free system when using template RNA extracted from a highly purified preparation of informosomes of hen embryo muscle *(62)*. These observations offer the most direct evidence for the messenger nature of the RNA of informosomes.

E. Proteins of Informosomes. The Problem of Isolation of Informosomes in a Pure Form

While the RNA of informosomes can be rapidly labeled with radioactive precursors, the selective labeling of informosome proteins has not been achieved in the majority of cases. Therefore, to analyze the protein component of informosomes, the latter must first be isolated from all other cellular proteins and protein-containing structures.

A serious obstacle in isolating pure preparations of informosomes is the small amount of them in the cytoplasm. Their instability also hinders purification; during the isolation procedure, a partial deproteinization of the particles can take place as well as degradation of their RNA by ribonucleases. Therefore, to obtain pure preparations of informosomes, mild, very effective and simple procedures are required.

If the problem of isolation of informosomes from the bulk of soluble proteins with low sedimentation coefficients can be solved by centrifugation techniques or gel chromatography, reliable preparative methods providing effective separation of informosomes from ribosomes, ribosomal subparticles and polyribosomes have not been yet developed. The use of zonal gradient centrifugation for this purpose is limited, as for the majority of sources, the sedimentation coefficients of a significant fraction of informosomes overlap with those of the ribosomal material (*1, 2, 7, 17, 42*, and others).

Nevertheless, 20 S informosomes containing globin mRNA have been isolated from duck reticulocytes by repeated gradient centrifugation (*22*). Electrophoresis in the presence of sodium dodecyl sulfate revealed eight types of polypeptide chains with M_r's of 15,000, 15,500, 17,000, 19,000, 21,000, 22,000, 24,000 and 51,000. Unfortunately, the procedure used has a serious disadvantage: the preparation of infor-