PROGRESS IN

Nucleic Acid Research and Molecular Biology

Volume 18

PROGRESS IN
Nucleic Acid Research and Molecular Biology

edited by

WALDO E. COHN

Biology Division
Oak Ridge National Laboratory
Oak Ridge, Tennessee

Volume 18

1976

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 NW1

LIBRARY OF CONGRESS CATALOG CARD NUMBER: 63-15847

ISBN 0-12-540018-7

PRINTED IN THE UNITED STATES OF AMERICA

List of Contributors

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

R. BRIMACOMBE (1), *Max-Planck-Institut für Molekulare Genetik, Berlin-Dahlem, Germany*

VOLKER A. ERDMANN (45), *Max-Planck-Institut für Molekulare Genetik, Abt. Wittmann, Berlin-Dahlem, Germany*

R. A. GARRETT (1), *Max-Planck-Institut für Molekulare Genetik, Berlin-Dahlem, Germany*

DAVID R. KEARNS (91), *Department of Chemistry, Revelle College, University of California, San Diego, La Jolla, California*

K. H. NIERHAUS (1), *Max-Planck-Institut für Molekulare Genetik, Berlin-Dahlem, Germany*

E. PALEČEK (151), *Institute of Biophysics, Czechoslovak Academy of Sciences, 612 65 Brno, Czechoslovakia*

BERNARD PULLMAN (215), *Institut de Biologie Physico-Chimique, Laboratoire de Biochimie, Théorique associé au C.N.R.S., Paris, France*

ANIL SARAN (215), *Institut de Biologie, Physico-Chimique, Laboratoire de Biochimie, Théorique associé au C.N.R.S., Paris, France*

H. G. WITTMANN (1), *Max-Planck-Institut für Molekulare Genetik, Berlin-Dahlem, Germany*

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 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 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 *N*-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 homo-

logous 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 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.

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 $(A_3,C_2)_n$, 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., $(A)_n \cdot (dT)_{12\text{-}18}$

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 $(A)_n \cdot 2(U)_m$
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.:

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

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.†

* 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.

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 Scientific Publ. Co., Amsterdam, 1973, and Supplement No. 1, *BBA* **429**, 1 (1976).
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
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 NRC 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

The Transfer RNAs of Cellular Organelles
W. E. BARNETT, L. I. HECKER AND S. D. SCHWARTZBACH

Mechanisms in Polypeptide Chain Elongation on Ribosomes
E. BERMEK

Mechanism of Action of DNA Polymerases
L. M. S. CHANG

Initiation of Protein Synthesis
M. GRUNBERG-MANAGO

Integration vs. Degradation of Exocellular DNA: An open Question
P. F. LURQUIN

The Messenger RNA of Immunoglobulin Chains
B. MACH

Bleomycin, an Antibiotic Removing Thymine from DNA
W. MÜLLER AND R. ZAHN

Protein Biosynthesis
S. OCHOA

Vertebrate Nucleolytic Enzymes and Their Localization
D. SHUGAR AND H. SIERAKOWSKA

Regulation of the Synthesis of Aminoacyl-tRNAs and tRNAs
D. SÖLL

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

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

Transfer RNA in RNA Tumor Viruses
L. C. WATERS AND B. C. MULLIN

Structure and Functions of Ribosomal RNA
R. ZIMMERMANN

Contents

* For Addendum to article, see p. 323.

Premelting Changes in DNA Conformation

E. PALEČEK

Quantum-Mechanical Studies on the Conformation of Nucleic Acids and Their Constituents

BERNARD PULLMAN AND ANIL SARAN

The Ribosome of *Escherichia coli*

R. Brimacombe,
K. H. Nierhaus,
R. A. Garrett and
H. G. Wittmann

Max-Planck-Institut für Molekulare Genetik Berlin-Dahlem, Germany

I. Introduction

The properties of ribosomes from *Escherichia coli* have been far more widely studied than those from any other organism. This review is therefore restricted to a discussion of the *E. coli* ribosome, except where interesting conclusions can be drawn by a comparison with corresponding studies on other organisms. The review is focused on two aspects, namely, structure and function. The structural aspects are considered in three sections concerning, respectively, the primary structure of the ribosomal proteins and ribosomal RNA, the interaction between proteins and RNA, and the topographical arrangement of the proteins. The section on function is subdivided into considerations of initiation of protein synthesis, elongation and termination. In general, the literature cited refers to the most recent work on a particular topic, rather than to the older findings,

which are covered in a number of recent reviews dealing in more detail with various aspects of the ribosome problem. These reviews are quoted in their appropriate place in the text.

II. Ribosomal Components

A. Ribosomal Proteins

An essential prerequisite for studies on ribosomes at the molecular level is the isolation and characterization of all ribosomal proteins. Isolation of the proteins has been achieved mainly by combination of several methods, *viz.* separation of the ribosomal subparticles by zonal centrifugation, prefractionation by salt treatment if necessary, column chromatography on carboxymethyl- or phosphocellulose, and gel filtration on Sephadex. A description of these various isolation procedures has been reviewed recently (*1*).

Two-dimensional polyacrylamide gel electrophoresis shows (*2*) that the small subparticle of the *E. coli* ribosome contains 21 proteins, designated S1 to S21, and the large subparticle 34 proteins, L1 to L34.[1] The molecular weights of all the proteins fall within the range from 6000 to 32,000, with the exception of the 30 S protein S1, which has a much higher molecular weight of 65,000. The individual molecular weights are listed in a review by Wittmann (*1*). Most of the ribosomal proteins are basic in nature, and this is reflected in their amino-acid compositions and isoelectric points. Many proteins are very rich (up to 34%) in basic amino acids (*3–5*); protein S21, for instance, has 20% arginine and protein L33 has 21% lysine; in consequence, the isoelectric points of the proteins are very high: pH 10 or more for about 70% of the proteins. Only three proteins (S6, L7 and L12) are mildly acidic, and these have isoelectric points of approximately pH 5 (*6*).

The separation of tryptic peptides on a preparative scale and amino-acid analyses have so far been made for almost 50% of the ribosomal proteins (*7, 8,* and Wittmann-Liebold, unpublished results). Information about the primary structure of ribosomal proteins can also be very efficiently and relatively quickly obtained by means of a protein sequenator. In this way, up to 40–60 amino-acid residues from the N-terminal regions have so far been sequenced for each of 18 proteins from the small subparticle, and 14 proteins from the large (*8, 9*). A few

[1] The proteins are numbered according to their positions on two-dimensional electropherograms, number first from left to right (i.e., plus to minus, first dimension) and then from top to bottom (i.e., plus to minus, second dimension).

proteins (e.g., S5, S18, L7 and L11) have blocked N termini, and therefore cannot be analyzed by the sequenator.

The complete primary structure has been determined for ten proteins: S4 (*10*), S5 (*11*), S6 (*12*), S8 (*13*), S9 (*14*), S18 (*15*), L7 and L12 (*16*), L25 (*17, 17a*) and L29 (*18*). Sequences of 15 others are well on the way to completion. The total number of amino-acid residues sequenced to date is about 2500, i.e., 30% of the approximately 8000 amino acids contained in the *E. coli* ribosomal proteins. This large effort is justified for the following reasons.

1. The ribosome, and in particular the *E. coli* ribosome, is the only cell organelle for which a thorough understanding at the molecular level of both structure and function can be expected in the near future. The detailed determination of the spatial distribution of ribosomal proteins, for example, requires knowledge of the amino-acid sequences of the proteins, leading to an understanding of the chemical and physical basis of RNA–protein interaction in the ribosomal particle.

2. Ribosomal proteins are altered in many bacterial mutants, e.g., those resistant to antibiotics inhibiting protein biosynthesis. In *E. coli* ribosomes, ten proteins from the small subparticle (S2, S4, S5, S6, S7, S8, S12, S17, S18 and S20) and three from the large (L4, L6 and L22) have been found to be altered by mutation. In many of the altered proteins, single amino-acid replacements, clustered in "hot spots," have been found and localized. In some mutants of protein S4, more drastic differences, leading to completely different C-terminal regions as a result of frame shifts, have been determined and analyzed. These studies have been reviewed elsewhere (*19*) and are not described here.

3. The fact that ribosomes consist of numerous proteins raises the following questions. First, are the proteins present in ribosomes of a given organism, e.g., *E. coli*, completely different from each other, or do there exist regions of identical or similar amino-acid sequences among them? Second, how did the ribosomal components change during evolution, i.e., are there identical or similar regions in ribosomal proteins from organisms belonging to different classes, e.g., bacteria, plants and animals?

A preliminary answer to the question of homology among the various *E. coli* ribosomal proteins, as well as between proteins from *E. coli* and other organisms, can be given by immunological, electrophoretic and chromatographic methods (reviewed in *19, 20*). However, a more direct and detailed approach to this question is a comparison of the amino-acid sequences of the proteins. As mentioned above, approximately 2500 amino-acid positions have to date been determined in *E. coli* ribosomal proteins, and these sequences, comprising about 30% of the total number

of amino acids in the *E. coli* ribosome, are distributed among 32 of the proteins. No regions longer than five amino acids common to different proteins have been found (*9*), with the exception of two proteins (L7 and L12) that have identical primary sequences and differ only in the presence of an acetyl group at the N terminus of protein L7 (*16*). These direct results are in very good agreement with the conclusions drawn from immunological studies on proteins isolated from *E. coli* ribosomes (*21*).

A comparison of the amino-acid sequences of ribosomal proteins from different organisms has up to now been limited to *E. coli* on the one hand and to only two different bacteria (*Bacillus stearothermophilus* and *Halobacter cutirubrum*) on the other. Yaguchi *et al.* (*22*) compared the N-terminal regions (15 amino acids) of 21 proteins isolated from *B. stearothermophilus* 30 S subparticles with the sequences of *E. coli* proteins. Similarly, Higo and Loertscher (*23*) made a comparison of five *E. coli* and *B. stearothermophilus* 30 S proteins (25–30 amino acids each). The first authors found, for each *E. coli* 30 S protein (with the exception of S1), a homologous protein from *B. stearothermophilus* ribosomes. The degree of homology, depending on the particular protein, varied from 25 to 90% of the regions sequenced. These sequence homologies support earlier structural and functional comparisons of the proteins isolated from these organisms (*24, 25*).

It remains to be seen, by comparison of protein sequences from more organisms, whether a correlation exists between the degree of conservation of the primary structure during evolution and other properties of the proteins, e.g., ability to bind to RNA, or presence in an active ribosomal site. The hypothesis, that functionally important proteins are mainly conserved, is supported by the finding that the structures of proteins L7 and L12, which have important functions during initiation, elongation and termination (reviewed in *26*), have been relatively strongly conserved during evolution. This has been shown by comparing corresponding regions of the protein chains from *E. coli, B. stearothermophilus* and *H. cutirubrum* (*27*). Further, immunological and biochemical studies indicate that proteins related to *E. coli* proteins L7 and L12 are present in ribosomes from yeast and rat liver (*28, 29*).

B. Ribosomal RNA

1. Primary Structure

The primary structures of the three ribosomal RNAs have been extensively studied during the last ten years. The sequence of 5 S RNA was

completed some years ago (*30*), and that of the 16 S RNA is about 90% completed (*31–33*). Sequencing of the 23 S RNA has advanced to the stage where large sections of the 3′ and 5′ regions of the molecule are known, as well as several sections in the central region (*34*).

The current state of the 16 S RNA sequence is given in Fig. 1, which also indicates those regions whose sequences are not yet certain. The sequences of some sections were determined by Santer and Santer (*35*). They obtained results very similar to those cited above (*31–33*), and confirmed many of the connections between the sections. Moreover, Uchida *et al.* (*36*), from analyses of the T1 ribonuclease oligonucleotides, confirmed almost all the oligonucleotide sequences reported by Fellner *et al.* (*32*); very few changes were incorporated into the latest sequence [(*33*) and Fig. 1]. The 3′-terminal oligonucleotide was sequenced by Shine and Dalgarno (*37*) and confirmed by Ehresmann *et al.* (*38*) and by Noller and Herr (*39*).

A number of interesting properties of the ribosomal RNAs can be derived from the RNA base-sequence determinations. (a) The 16 S RNA contains about 1600 nucleotides, and the 23 S RNA about 3200. (b) The overall base composition of the 16 S RNA is 31.8% guanine, 24.8% adenine, 23.3% cytosine and 20.0% uracil. (c) A low level of heterogeneity of approximately 1% of the nucleotides in the 16 S RNA was found. These tend to be clustered at a few points along the sequence, and comparable levels of heterogeneity were found both in 5 S RNA [reviewed by Monier (*40*)] and 23 S RNA (*34*). (d) Although substantial duplication of sequence was found within the 5 S RNA (*30*), no such repeated sequences have yet been detected in either the 16 S RNA or the 23 S RNA. Moreover, a comparison of the T1 ribonuclease oligonucleotides of 16 S and 23 S RNAs revealed no similarities, and the methylated regions and the 5′ and 3′ termini are completely different. (e) Several palindromes were detected in the 16 S RNA sequence, i.e., the sequences are the same whether they are read from the 5′ or 3′ end. These occur in sections L, R, I′, I, K, C'_2 and D, and some of the longer examples are indicated in Fig. 1. Moreover, some base-paired hairpin loops contain 2-fold axes of symmetry.

2. Secondary Structure

The overall secondary structures of the isolated large ribosomal RNAs have been studied extensively by different spectroscopic methods and by spectrophotometric titration (see References *41–43* for review of the earlier work) 60–70% of the nucleotides, were estimated to be involved in base-pairing.

Mainly from the efforts of Cox (*44*) and Spencer and their co-workers, more detailed information has been obtained on the nature of the

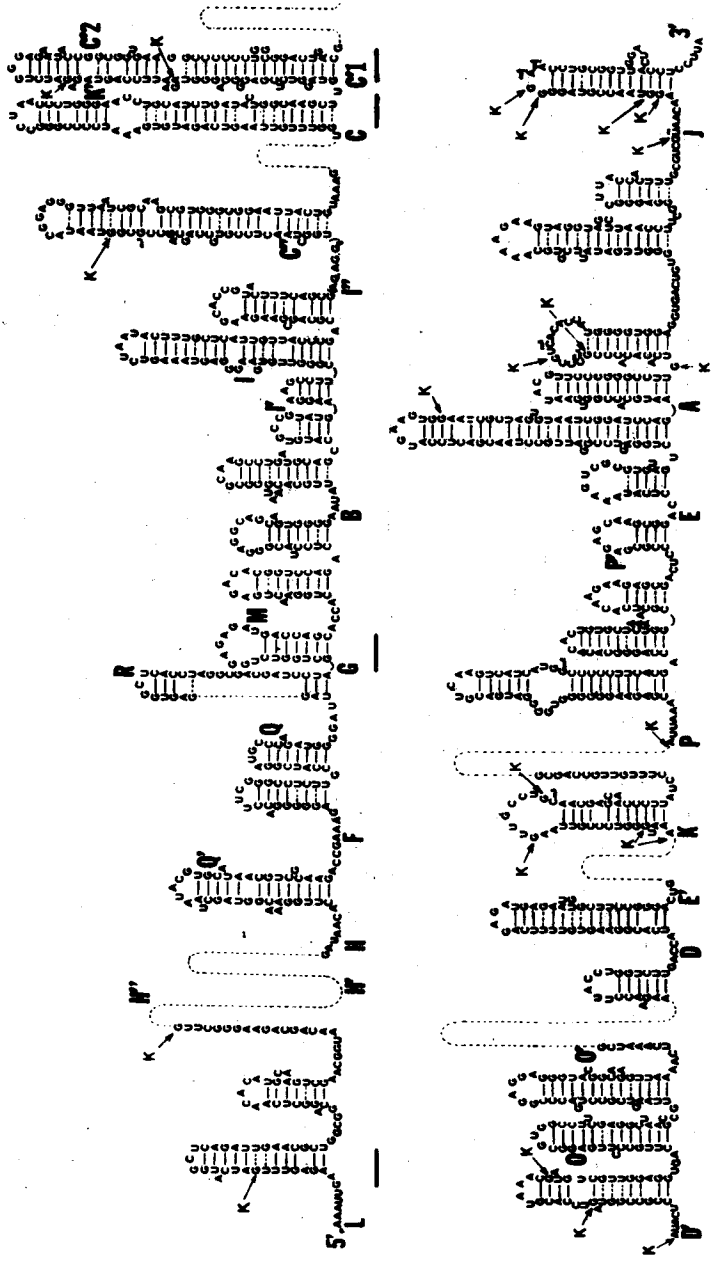

FIG. 1. The current sequence of the 16 S RNA with a theoretically evaluated base-pairing scheme [from Ehresmann *et al.* (33)]. -----, Regions for which the order of the oligonucleotides is still uncertain. Some "hairpins" that exhibit perfect, or near-perfect, palinfromes or symmetry about a 2-fold axis are underlined. Guanine residues that are readily modified by kethoxal in the 30 S subunit are indicated by a *K*.