105

DNA Systematics

Volume I: Evolution

Editor
S. K. Dutta, Ph.D.
Professor
Departments of Botany,
Genetics and Human Genetics,
and Oncology
Howard University
Washington, D.C.

Chapter 4

ANALYSIS OF SMALL RNA SPECIES: PHYLOGENETIC TRENDS

Mirko Beljanski and Liliane Le Goff

CRC Press, Inc.
Boca Raton, Florida

Chapter 4

ANALYSIS OF SMALL RNA SPECIES: PHYLOGENETIC TRENDS

Mirko Beljanski and Liliane Le Goff

TABLE OF CONTENTS

I. Introduction 82

II. Transfer RNAs 82
- A. tRNA Genes 83

III. 5S Ribosomal RNA 85
- A. Secondary Structure of 5S rRNA 85
- B. 5S rRNA Gene 85
- C. The Biological Function of 5S rRNA 86
- D. A Common Ancestral Gene for tRNA and 5S rRNA? 87

IV. 5.8S Ribosomal RNA 87

V. 7S RNA Structure 88
- A. Biological Activity of 7S RNA 88

VI. Small RNAs Transcribed from Plasmid 88

VII. Small RNAs in Cells Infected with Virus 89
- A. Retrovirus Associated Small RNAs 90
- B. Viroid RNA Molecules 90

VIII. Low-Molecular-Weight Nuclear RNAs 90
- A. Phylogenetic Trends of USn RNAs 91

IX. Other Small RNAs in Prokaryotes and Eukaryotes 91

X. Small RNA Molecules and Oligoribonucleotides as Primers for DNA Replication, Transcription, and Translation 92
- A. Small RNAs in the Translation Process 93
- B. Free Intervening Sequences from Genes and Pseudogenes as Active Biological Molecules? 94
- C. Small RNAs: Molecular Selfreplication 95
- D. Transcription of Small Cellular RNAs into DNA 95

XI. Concluding Remarks 96

References 99

I. INTRODUCTION

Genetic information in eukaryotes and prokaryotes is stored in DNA molecules whose specific segments are transcribed by appropriate polymerases into various types of RNAs: messenger RNA (mRNA), ribosomal RNA (rRNA), transfer RNA (tRNA), and many other small RNAs (50 to 360 nucleotides). Information in viruses is stored either in DNA or RNA strands. A large variety of small RNAs is present in the nucleus of eukaryotic cells and/or in the cytoplasm of mammals, plants, viruses, and bacteria. The variety of these RNAs, as well as their "constant" or supposed engagement in various biological processes, raises the question of their possible phylogenetic relationship. Phylogenetic measurement, extensively accomplished for rRNAs from different species, has also been studied for various small RNAs. Thus, sequence data of cytoplasmic 5S rRNAs from metazoan somatic cells have been used to analyze the phylogenetic relationships. Many small RNAs have been characterized by sequence analysis, and by physical means, for some of them by proposed or established biological activity. Thus, purine-rich small RNAs stimulate the translation of mRNA[1] while others induce the transformation of bacteria.[2]

The widespread existence of endo- and exonucleases in all living organisms forces one to consider the participation of these enzymes in producing small RNAs which are more or less rich in purine or pyrimidine nucleotides. Some of these enzymes are involved in the formation of mature "mosaic" messenger RNAs, rRNAs, tRNAs, etc., which leaves intervening sequences "unused".

The over-production of rRNAs before cell division and their disappearance as soon as cell multiplication begins[3,4] raise the question of the function of so many extra copies of rRNAs. They may be the target for specific nucleases which deliver different small RNAs, as in the case of 5.8S rRNA emergence from prokaryotic 23S rRNA. Small RNAs interfere through base pairing with large DNA and/or RNAs during synthesis of cell constituents, cell multiplication, and differentiation.[2,5]

Small RNAs can be synthesized *de novo* from ribonucleoside-5′-triphosphates by *Escherichia coli* Qβ replicase, an enzyme which also replicates Qβ RNA template, as well as satellite RNA of Qβ virion. Polynucleotide phosphorylase (PNPase) from bacteria can also synthesize RNA from ribonucleoside-5′-diphosphates.[6] These facts and the observation that small RNAs from different origins can be transcribed into DNA, contribute to modifying notions about the origin and flow of information in cells. For these and other reasons, scientists have thus studied small RNAs and have visualized their possible participation in the creation of new genes and/or pseudogenes. In this chapter, we shall attempt to give an overview of many different small RNAs by describing their chemical and physical properties as well as their evident, or possible biological role. The phylogenetic trends between these RNAs will be evaluated on the basis of sequence data since at the present time the biological functions of the majority of small RNAs are unknown.

II. TRANSFER RNAs

Present in all living organisms, transfer RNAs are small molecules (#75 nucleotides). They bind activated amino acids, recognize a three-base "codon" (in mRNA) which specifies a particular amino acid, and deliver that amino acid to be incorporated into the polypeptide chain. Alanine transfer RNA ($tRNA_{ala}$) was the first RNA to be sequenced,[7] and its secondary structure has been proposed (cloverleaf diagram) (Figure 1). The importance of the structure, conformation, and interaction of tRNAs with other molecules was recently reviewed.[8] There are about 40 functionally different kinds of tRNAs which seemingly have a complicated molecular geometry and which appear to be of very ancient ancestry.[9] All tRNAs contain

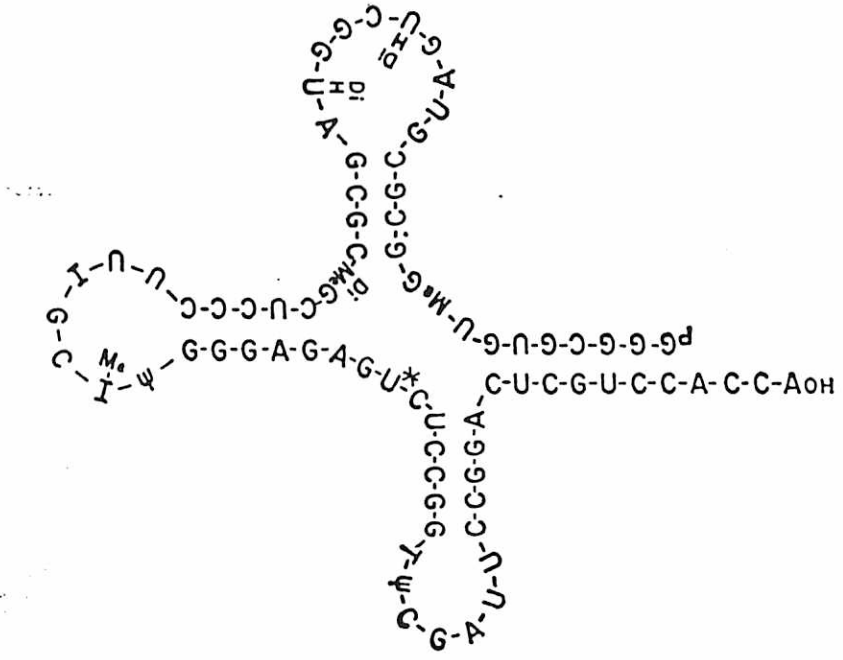

FIGURE 1. Schematic representation of conformation of the alanine tRNA with short double-stranded regions. (From Holley, R. W. et al., *Science*, 47, 1462, 1965. With permission.)

tRNAs: (1) the presence of unusual bases — pseudouracil, dihydrouracil, hypoxanthine, ribothymine, thiopyrimidines, and various methylated purines; (2) terminal sequences — CCA; and (3) several "invariant" and "semivariant" residues located in the same relative position.[10] The importance and function of modified nucleotides in tRNAs have been reviewed.[11] The sequences of about 120 tRNAs from different biological systems have been determined.[12] There exist tRNAs which initiate protein synthesis; in eukaryotes one tRNA carries methionine and in prokaryotes *N*-formyl methionine. Since their discovery,[13] the biological activities of tRNAs have been extensively studied in vitro and in vivo.[14]

A. tRNA Genes

Eukaryote and prokaryote tRNA genes are transcribed from DNA as independent units by RNA polymerase III first as precursor transcripts[15-19] (Figure 2). The tRNA precursor cannot be amino-acylated in vitro.[20] The $tRNA_{tyr}$ transcription unit was studied in particular detail because the corresponding gene was synthesized in vitro by chemical and enzymatic means and introduced under in vivo conditions where it functioned correctly.[21,22] Cloned tRNA genes, injected into the large oocyte nucleus of *Xenopus*[23] can also generate correct tRNAs. The number of tRNA genes varies according to species. The haploid yeast genome contains about 300 tRNA genes,[24] *Drosophila*, 600,[25] and *Xenopus*, 8000.[26]

In yeast mitochondria (mt), tRNA genes are clustered,[27] while in HeLa cells and *Xenopus* they appear to be dispersed.[28,29] mt DNA encodes tRNAs which are different from those transcribed in the nucleus (UU in the D stem replaces GU base pairs).[30] Since the degree of homology of mammalian mt tRNAs with non-mt species is between 30 and 50%, the integration of mitochondrial sequences into tRNA evolutionary trees is difficult, although mt tRNA genes are highly conserved.[31] Clustered or dispersed tRNA genes are several times larger than one tRNA length.[32] Several steps in the splicing of the precursor transcripts are required to obtain mature tRNA. The -CCA sequence present in all mature tRNAs is not

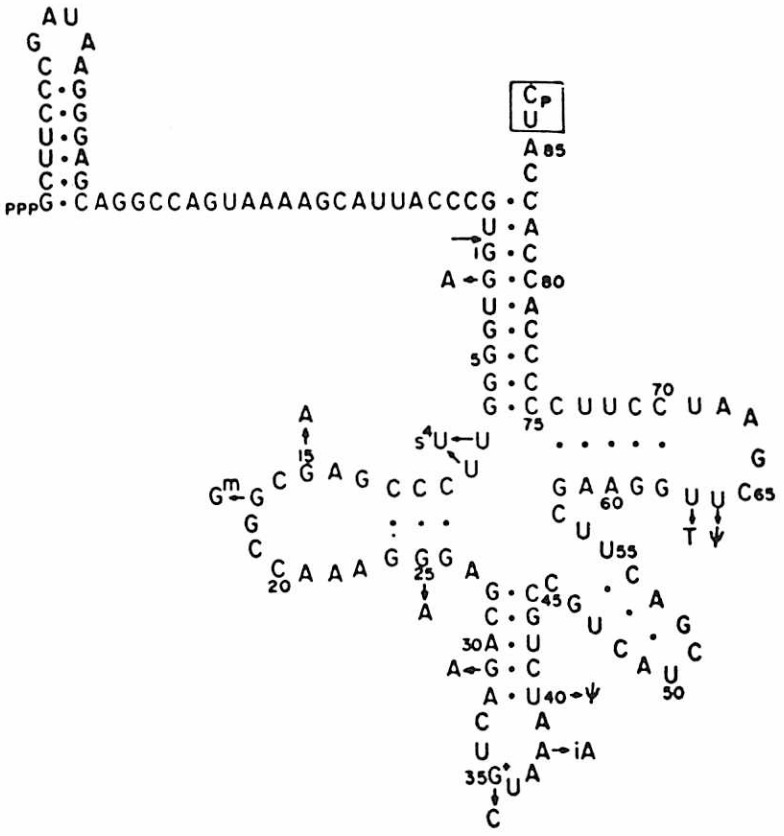

FIGURE 2. Nucleotide sequence of a precursor to *E. coli* $tRNA^{tyr}$. (From Altman, S. and Smith, J. D., *Nature (London) New Biol.*, 233, 35, 1971. With permission.)

not present but appear to be in *E. coli*.[17,34] They are added later by a nucleotidyl-transferase. The intervening sequences which are transcribed within tRNA genes differ according to the tRNA species and all contain high A + U base composition (A + U/G + C = 2/1).[12]

In prokaryotes tRNA genes are clustered and transcribed in "in vivo" conditions as multicistronic precursor molecules. Transcription starts at the first tRNA gene in a gene cluster and terminates at the last gene. Subsequently, the 5′ and 3′ flanking sequences are removed, liberating mature tRNA.[35]

In retroviruses tRNA molecules are present in large quantities.[36-42] About 200 tRNA molecules are selected from the host cell and included in a virion particle. Most tRNAs bound with 18S or 28S rRNA are unable to accept amino acids unless they are first released by heat treatment. This indicates at least one set of common sequences between tRNAs and rRNAs.

It should be noted that phylogenetically related vertebrates show more similar tRNA-rRNA interaction than do phylogenetically distant vertebrates.[43] tRNA serving as a primer for the reverse transcriptase of genomic RNA of a type C retrovirus binds to 28S rRNA of its host.[43] Nucleotide sequences characteristic of these tRNAs also play a role in the binding of tRNA to 5S rRNA.[44]

Retrovirus tRNA-host rRNA hybridization shows that nucleotide sequence of rRNAs binding the tRNAs are different among various vertebrates. Similar results were obtained with retrovirus genome RNAs, thus indicating a certain degree of nucleotide recognition between retrovirus genome RNAs and rRNAs with tRNA

III. 5S RIBOSOMAL RNA

5S rRNA was first discovered as a component of a 50S ribosomal subunit of *E. coli*,[45] and thereafter it was further analyzed.[46] The sequence of that 5S rRNA was determined by Sanger et al. in 1968,[47] and since then many prokaryotic and eukaryotic 5S rRNAs have been sequenced. Mammalian 5S rRNAs are considered to be practically identical.[48] This RNA contains about 120 to 121 nucleotides but no methylated or otherwise modified nucleotides. 5S rRNA is transcribed by RNA polymerase III from multiple structural genes. The in vitro synthesized product was identical to 5S rRNA transcribed from plasmid DNA which contains the *X. laevis* oocyte 5S gene.[49] In vitro hybridization of transcribed 5S rRNA with DNA fragments obtained by Eco RI digestion of calf thymus DNA shows that it originates in DNA. Calf thymus chromatin used as an in vitro template does not deliver 5S rRNA.[49]

The number of nucleotide substitutions between 5S rRNAs from vertebrates and invertebrates varies to a significant degree. The 5S rRNA sequences are identical between different genera but are distant from those which belong to the different order. Nucleotide sequences of 5S rRNA from the sponge which is the simplest form of multicellular animals are closely related to invertebrates. Tunicate has affinity to vertebrates.[50] The phylogenetic positions of 5S rRNA from these two species among metazoans were derived from the 5S rRNA sequences by a computer analysis based on the maximum parsimony principle. This method utilizing information contained in the sequence data differs from the conventional taxonomy.[50]

A. Secondary Structure of 5S rRNA

To obtain information about evolutionary changes, it is important not only to compare sequences of 5S rRNAs, but also to establish the secondary structure of this molecule from different species. The archaebacterial 5S rRNA secondary structure resembles typical bacterial 5S rRNA more than the eukaryotic 5S rRNAs do.[51] These observations support the view that the secondary structure of the 5S rRNA remains the "same" in the corresponding site of the stem and has been conserved throughout evolution, although the primary structure of these RNAs has undergone changes.[30]

Recently, with 17 prokaryotic 5S rRNAs used for comparison, a general model for the secondary structure of 5S rRNA has been proposed.[52] In this model eight double helical regions seem to be present and involved in the function of 5S rRNA (Figure 3).

B. 5S rRNA Gene

The genes of 5S rRNA are located in clusters at the telomeres of practically all chromosomes. DNA containing the oocytes 5S rRNA genes has been isolated from *X. laevis* and characterized. One 5S gene is part of a larger repeating unit in the DNA.[53] Sequences of the purified 5S DNA of *X. laevis* have been determined,[54] and the primary structure of the repeating unit in oocyte 5S DNA has been described.[48,55] The spacer sequences differ in base composition and internal repetitiveness. The 5S rRNA genes appear to evolve as a unit. The spacer sequence which is part of the repeating unit in a tandem gene cluster is not part of the gene.[56]

In the 5S DNA, the sequence coding for 5S rRNA alternates with a "spacer" region about 6 times as long.[56] It has been shown that in a cloned oocyte 5S DNA fragment from *X. borealis*, three 5S rRNA genes are separated by about 80 nucleotides.[57] RNA polymerase III binds to a "control region" that differs from promoter regions in prokaryotes. The enzyme correctly transcribes in vitro 5S rRNAs from cloned 5S DNA fragments. Although *Xenopus* and mammalian 5S rRNAs represent each a single homogeneous RNA of 120 nucleotides, they differ by 8 nucleotides.[54] The 5S rRNA sequences of all mammals that

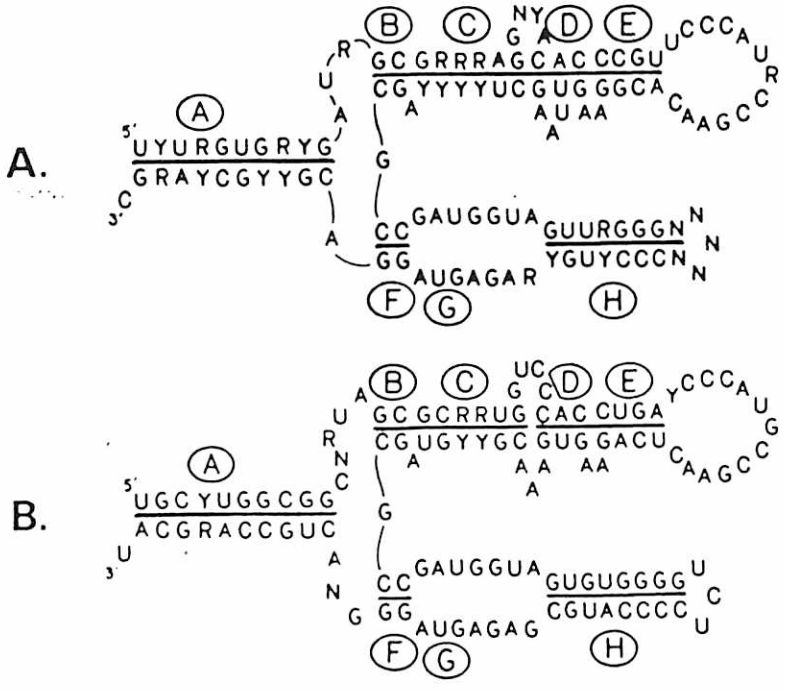

FIGURE 3. Schematic representation of conformation of 5S rRNAs from Gram positive bacteria (A) and from Gram negative bacteria (B). (From Studnicka, G. M. et al., *Nucleic Acids Res.*, 9, 1885, 1981. With permission.)

5S rRNA sequence in the chicken differs from the mammalian sequence by three base substitutions, two nucleotide additions, and two deletions. The sequence of 5S rRNA from *Bacillus megaterium* is related to that of *B. stearothermophilus* but differs by 23 base replacements and is 3 nucleotides shorter.[59]

The nucleotide sequence of 5S rRNA of vertebrates is apparently conserved throughout evolution. 5S DNA from different species and mouse satellite DNA have undergone frequent unequal homologous or nonhomologous crossing over during evolution.[56] It has been suggested that spacers and particularly repetitive sequences are "critically important" to both the stability and the evolutionary flexibility of the multigene family. The introduction of the development of repetitive spacers would tend to enhance the overall duplication/deletion rate in a multigene family.[48] According to these authors, there is a greater divergence of spacers than of gene sequences in homologous 5S gene clusters of different *Xenopus* species. This suggests that the somatic and oocyte genes present in the thousands of copies in *Xenopus* have evolved essentially as separate lines. The difference between 5S rRNA from *Xenopus* somatic and oocyte genes is 6 nucleotides.[48] The *Xenopus* somatic type 5S rRNA does not comprise more than 3% of the genes within 5S DNA. The 5S DNA of chicken embryo fibroblasts delivers an RNA copy of 121 base pairs and spacers of an approximate overall average length of 750 pairs.[60]

C. The Biological Function of 5S rRNA

Reconstitution experiments have demonstrated the importance of 5S rRNA as a structural component in the 50S ribosomal subunit.[45,61] 50S ribosomes that reconstitute without 5S rRNA are void of several proteins and show greatly reduced biological activity. All prokaryote 5S rRNAs tested were active in reconstituting *B. stearothermophilus* 50S ribosomal subunits,

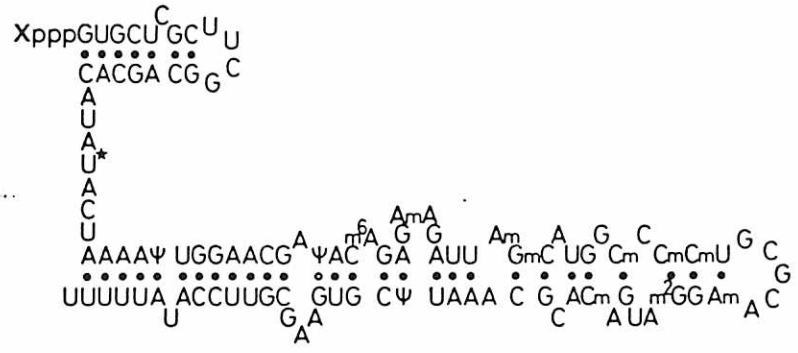

FIGURE 4. Primary and secondary structure of 5.8S rRNA from the silkgland of *Bombix mori*. (From Fujiwara, H. et al. *Nucleic Acids Res.*, 10, 2415, 1982. With permission.)

Although the participation of 5S rRNA as a structural component of 50S ribosomes has been demonstrated, no particular activity was assigned to this RNA. However, it has been shown that 5S rRNA isolated and purified from rabbit reticulocyte ribosomes exhibits a strong inhibitory effect on the translation of mRNA in a cell-free system. Used in nanogram quantities, 5S rRNA provokes a substantial polysomal breakdown in the presence of ATP and polysomes.[63,64] It was suggested that 5S rRNA may play a role in the peptidyl-transferase activity, where the RNA in question may serve as an intermediate acceptor of the growing peptide chain.[65] It was also assumed that 5S rRNA contributes to the movement of the ribosome relative to mRNA.[66] So far, the biological activity of 5S rRNA has been rather poorly established in comparison to analytical data accumulated for its primary and secondary structure.

D. A Common Ancestral Gene for tRNA and 5S rRNA?

The possible origin of 5S rRNA and tRNA has been considered from the evolutionary point of view.[67] First, precursor tRNA molecules are about the same length as the 5S rRNA.[17] Second, 5S rRNA and tRNA have a similar nucleotide composition[68] and an homology in the secondary structure.[52] Third, many tRNAs exhibit a high degree of sequence similarity (60% homology) with *E. coli* 5S rRNA. The first half of the tRNA sequence exhibits homology with the latter half of the 5S rRNA sequence. An extensive homology exists between several of the tRNAs and the 5S rRNAs from *E. coli*, human KB carcinoma cells, and *Pseudomonas fluorescens*. All these data lead to a hypothetical model for the origin of tRNA and 5S rRNAs in a common ancestral gene.[67]

IV. 5.8S RIBOSOMAL RNA

The 5.8S rRNA has been found as a component of the large ribosomal subunit of prokaryotes and eukaryotes.[69,70] In prokaryotes, 5.8S rRNA appears to evolve from the 5′ end of 23S rRNA.[70] In eukaryotes, 5.8S rRNA (160 nucleotides) has been also found in the large ribosomal subunit where it is base paired with the 28S rRNA molecule.[69,71] The nucleotide sequences of 5.8S rRNA (Figure 4) from over 10 very different species have been reported.[72] The 5.8S rRNA of the silkworm *Bombyx mori* is at the 5′ terminal sequence several nucleotides longer than those of other organisms. Since this RNA is base paired with 28S rRNA it would be fascinating to discover how the 5.8S rRNA interact with the 28S rRNA containing the hidden break.[72]

In yeast at all events 5.8S rRNA seems to derive from 7S RNA through enzymatic cleavage.[69] Yeast 5.8S rRNA has been sequenced. It is capable of binding *E. coli* 5S rRNA binding proteins L18 and L25. The binding complexes have ATPase and GTPase activities.[73]

Although it has been suggested that the eukaryotic 5.8S rRNA may play a role in protein synthesis,[73] its biological function remains to be clearly established. This might help to determine to which extent 5.8S rRNA was involved in molecular evolution as well as in the evolution of different species.

V. 7S RNA STRUCTURE

7S RNA was first isolated from eye lens[74] and was thereafter found as a component in normal higher eukaryotic cells: human, rat, murine, chicken,[75] and invertebrates. 7S RNA in RNA viruses[76] is derived from the host. Fingerprints of 7S RNA containing 295 nucleotides (GC content = 60%) from various species and viruses showed an extensive homology. The structure of 7S RNA appears to have been conserved throughout evolution, although there is no consensus about its subcellular localization. It is worthwhile noting that the 5′ ends of 7S RNA and that of La 4.5S RNA have extensive (60 to 70%) homologies.[77] About 90 nucleotides of the 5′ ends of 7S RNA are homologous with Alu-DNA sequences which are a class of repetitive DNA first found in mammals. These sequences are rich in GC bases (63%). In addition, 45 nucleotides at the 3′ end of 7S RNA have been found to be homologous with Alu-DNA sequences. One 300 nucleotide long Alu family sequence contains two binding sites for 7S RNA.[77] The significance of this interaction is not understood. 7S RNA from Novikoff hepatoma has been sequenced.[78] The human 7S L rRNA is heterogeneous within the population of one cell line.[81] This suggests that the 7S L rRNA is encoded by several nonidentical genes.

A. Biological Activity of 7S RNA

Different biological activities of 7S RNA have been suggested. Thus, 7S RNA isolated from eye lens[74] is capable of binding activated amino acids. This RNA, rich in GC pairs (G + C/A + U = 1,9), does not contain pseudouridine and differs from transfer RNA in its size and nucleotide composition. No further investigations have been carried out with this RNA to determine its participation in protein or peptide synthesis. Purified 7S RNA from chicken embryos plays a direct role in heart tissue differentiation and development.[82,83] Since a portion of 7S RNA sequence is homologous to sequences at the origin of DNA replication, it was suggested that 7S RNA may play a role in DNA replication. This may be possible and it should be feasible to demonstrate experimentally whether or not before acting as primer 7S RNA undergoes degradation in order to furnish smaller molecules whose size range is similar to that of RNA primers.

It has been suggested that the 7S L rRNA might be involved in the transport of mRNA from the nucleus into the cytoplasm.[81] Also 7S L sequenced cytoplasmic RNA appears to be involved as a constitutive and indispensable part of the signal recognition particle.[84] Since 7S RNAs do not seem to encode proteins, they probably play a structural role; however, the biological function of 7S RNAs remains to be firmly established in vitro and in vivo.

VI. SMALL RNAS TRANSCRIBED FROM PLASMID

Plasmid DNAs have been shown to code for several low molecular weight RNAs.[85,86] The in vitro transcription by RNA polymerase starts at two sites on the plasmid DNA (Col El DNA).[85] The transcript which starts at one of the sites may be cleaved by RNase H. It is involved as primer in DNA replication. Transcription from the other site on DNA leads to RNA I which inhibits the formation of the RNA primer that exhibits its effect only on homologous DNA template. The RNA I, containing 108 nucleotides, is without effect on initiation of transcription, elongation of RNA chains, or the processing of performed precursors. RNA I is transcribed from DNA and may inhibit the formation of a RNA-DNA hybrid which is a substrate for RNase H.

Plasmid containing mini-cells synthesize the RNA of about 4S to a considerable extent.[87] Multiple drug-resistant plasmid NR1 codes for about 10 small RNAs ranging from 60 to 120 nucleotides. They have been characterized by RNase T1 fingerprinting. Some hybridize with RNI DNA digested with restrition endonuclease, indicating that they originate in "the resistance transfer factor region of the plasmid genome." Several of these small RNAs are associated with DNA fragments that contain origins of replication.[86] One stable species of these RNAs migrates at the position of host tRNA. This component contains about 75 nucleotides carrying terminal-CCA residues at 3′. No modified bases have been detected. Thus, this small RNA is not tRNA. Since it has a-CCA residue, one should determine if it is capable of binding amino acids. Two 5S rRNA have been identified. They have no similarity with *E. coli* 5S rRNA.[86] Two RNA species contain 60 to 65 nucleotides possessing pUUAAGp at 5′, which indicates that they were not primary transcripts; they possibly resulted through nuclease action on larger RNAs (absence of 5′-PPP). It has been suggested that they might serve as primers for plasmid DNA replication. Curiously enough, a small RNA of about 60 nucleotides has been found solely in the nucleus of plasmid-infected cells. Using a constructed plasmid DNA carrying the 32 nucleotide intervening sequence, it was possible to demonstrate that this small RNA is complementary to an intervening sequence.[88] Sixty nucleotide RNA does not possess a 5′-terminal cap. In contrast to the nucleus, the cytoplasm of uninfected cells contains several small RNA species: 185, 160, 87, 77, and 73 nucleotides. The most abundant is an RNA with 185 nucleotides. These RNAs would seem to act in the translational process. However, direct proof is still unavailable.

VII. SMALL RNAs IN CELLS INFECTED WITH VIRUS

Low-molecular-weight RNAs have been found in cells infected with the SV40 virus, Epstein-Barr virus, vesicular stomatitis virus,[89,90] and defective interfering particles of the vesicular stomatitis virus.[91] In vesicular stomatitis virions, three small RNAs (28, 42, and 70 nucleotides) have been synthesized in vitro in addition to the 47 nucleotide leader sequence.[92] The 47 nucleotide RNA has been sequenced, and three other small RNAs contain (P) ppAA at their 5′ terminal. These RNAs are not polyadenylated.

Two species of small RNAs have been found as primary transcripts of the adenovirus 2 Ad-2 genome.[93] Two other species of RNAs (140 bases) appear to be degradation products of viral RNAs. Ad-2 viral small RNAs appear not to be involved in processing adenovirus hnRNA.

A small RNA (50 nucleotides) found in cells co-infected with a standard vesicular stomatitis virus and its interfering particles[94] has sequences complementary to the genome of the defective particles. It does not possess poly A residues and its function is unknown. Apparently its synthesis correlates with the replication of the DNA of defective interfering particles. In contrast, RNA species that have been found in cells infected only with the vesicular stomatitis virus range in size from about 46 bases to 12 kb.[95] Forty-six bases RNA synthesized in vitro has the following sequences:[96]

5′(pp)pACGAAGACCACAAAACCAGAU

AAAAAAUAAAAACCACAAGAGGG(U)C_{OH}3

"This sequence is identical to the sequence at the 5′ end of the infectious vesicular stomatitis virus RNA and is complementary to the sequence of the 3′OH terminus of this defective interfering particle genome RNA."[96] It should be emphasized that purines are largely in excess over pyrimidines (Pu/Py = 3.0) in this small RNA. Purine-rich RNAs (obtained by degradation of *E. coli* rRNA with pancreatic RNase) and containing 46 nucleotides are actively involved in hematopoiesis, stem cell genesis, and differentiation in animals or humans.[2,5,97]

Small RNAs are encoded by some viruses and expressed during late lytic infection. Small cytoplasmic RNA from cells infected with SV40 virus is 65 nucleotides in length and contains pyrimidine rich sequences (U = 43%; C = 33%; G = 12%; A = 12%). Several clusters contain 5 to 7 pyrimidines, one has 11. This RNA is homologous with the SV40 early mRNAs.[98] It may be involved in the control of late transcription. It should be stressed that pyrimidine rich small RNAs inhibit protein in vitro biosynthesis.[1]

A. Retrovirus Associated Small RNAs

Several small RNAs represent a discrete subclass of the complement of small host RNAs.[99] The 4.5S RNA has been found associated with Moloney leukemia virus and spleen focus-forming virus. These RNAs contain many extra uridylate residues at their 3′ termini.[100] The 4.5S RNA from the spleen focus-forming virus contains more than 30 components that vary with the length of the poly U.[100] The 4.5S RNA molecules are provided by the host cells.

B. Viroid RNA Molecules

The viroid of the potato spindle tuber disease (PSTV) is an RNA that represents a covalently closed ring of 359 ribonucleotides,[101] which is approximately the size of the tetrahymena intervening sequence.[102] These latter authors have described "unit-length and linear (+)-strand" RNAs that seem to be intermediates in viroid RNA replication. Does the mature circular viroid involve end-to-end joining of a linear intermediate or does it arise like tetrahymena intervening sequence RNA cyclization via cleavage-ligation at an internal point in a linear molecule?[102] Modified nucleotides have not been detected in any complete or partial PSTV fragment. A stretch of 18 purines, mainly adenosines,[101] exists. At least 60% of PSTV is represented by sequences in the DNA of several *Solanaceous* host species. Linear PSTV RNA molecules may be converted into circular molecules (circle) by an RNA ligase purified from wheat germ,[102] in which circles are indistinguishable from circles extracted from infected plants. Non-*Solanaceous* species contain few or no sequences related to only a small portion of PSTV. It has been suggested that PSTV originates in genes in normal *Solanaceous* plants.[103] The mechanism of infection of plants by PSTV RNA is not completely understood.[104]

VIII. LOW-MOLECULAR-WEIGHT NUCLEAR RNAs

Eukaryotic cells contain, structually and functionally, small RNAs differing from other types of known cellular RNAs.[105-107] In HeLa cells some of the small RNA (Sn RNAs) components are located mainly in the nucleoplasm and one mainly in the nucleoli.[106] All USn RNAs: U1, U2, U3, U4, U5, and U6 exist in RNP particles. The analysis of several of these RNAs revealed the presence of a 5′ cap (a base $m^{2,2,7}$-methylguanosine) that differs from the 5′ cap of mRNA (m^7-G).[77] This explains why Sn RNAs are not translated in the cytoplasm. Among Sn RNAs, U3 RNA is particularly localized in the nucleus and has not been found in any other site in the cell. It may play a role in transcription of the rRNA from the rDNA templates by maintaining stable open gene complexes.[77] Detailed analysis of Sn RNAs has been described.[80] The U RNAs contain a large number of more or less clustered uridine residues. U5 RNA is very enriched in uridine (35%). RNase Tl fingerprints of the purified U3 RNA, U2, and U1 are practically identical (# 300 nucleotides) in HeLa cells, human normal fibroblasts, and Novikoff hepatoma cells. These uridine rich low molecular weight RNAs appear to have been conserved throughout evolution.[108]

The sequence of 18 nucleotides close to the 5′ end of Ul RNA containing 165 nucleotides is complementary to the nucleotides from intron transcript adjacent to the splice point.[109] Some nucleotide sequences of U1 RNA and U2 RNA are expected to form a hybrid with intron-exon borders.[110-112] These RNAs and other U RNAs may play a role in the splicing

of premessenger RNAs. However, the results obtained in vitro have not been demonstrated *in situ*, where U RNAs exist as an RNA-protein complex.[113] U2 nuclear RNA hybridizes to heterogeneous nuclear RNAs but not nucleolar RNA. The nucleolus contains the precursors of ribosomal RNAs. Using a human cloned U2 DNA probe, it has been demonstrated that U2 RNA is paired with complementary sequences in heterogeneous nuclear RNA in vivo.[114] Both U2 and U1 RNAs have complementary sequences with intron-exon borders in mRNAs precursor molecules.[110] It is interesting to recall here that pseudogenes, complementary to the small nuclear RNAs U1, U2, and U3, are dispersed and abundant in the human genome. Three pseudogenes — U1.101, U2.13, and U3.5 sequences[115] — are flanked by short direct repeats of 16 to 19 bp.[116]

Although USn RNAs have been purified and sequenced, their precise biological function has not been assigned.

A. Phylogenetic Trends of USn RNAs

The genes of U1, U2, and U3 RNAs appear to have been highly conserved throughout evolution. The genes of these RNAs have been isolated from human genome. This genome appears to contain more pseudogenes for small RNAs than real genes. Notable similarities have been found between sequences of the U1 RNA of *Drosophila* and U1 RNA sequences from other origins.[117] There are similarities between the USn RNAs of vertebrates and invertebrates, as far as primary structure is concerned. Dinoflagellates considered as eukaryotes contain U1-U6 Sn RNA which have a differing molecule center.[80]

IX. OTHER SMALL RNAs IN PROKARYOTES AND EUKARYOTES

Several small RNAs isolated from *Alcaligenes faecalis* bacteria, sediment on the sucrose gradient at 3.9S, 4.3S, 5.5S, and 6.3S.[118] 5.5S rRNA particularly studied differs by its base ratio (G + C/A + U = 2.0) and hyperchromicity (38%) from transfer RNAs (G + C/A + U = 1.35 and hyperchromicity 28%) and from rRNAs. The 5.5S RNA possesses the capacity to accept amino acids in the presence of each of the four ribonucleoside-5′-triphosphates and bacterial enzymes. It may play a role in peptide formation.[119] Showdomycin-resistant *E. coli* cells excrete into the culture medium a 6S RNA (G + A/C + U = 1.82) that acts as a transforming agent and can be in vitro and in vivo transcribed into DNA.[120]

A small stable 10S RNA has been identified and characterized from *E. coli*.[121] Ribonuclease P, unique among all the RNA processing enzymes, contains an RNA moiety that is required for its function.[122,123] Two RNAs are present in this enzyme. One of them, termed M2, is identical or similar to 10S RNA from *E. coli*. Also, ribonuclease P from yeast and *Bacillus subtilis* contains both protein and RNA components.[124] Using a cloned segment of DNA for complementation it has been shown that it codes for an RNA species of 340 bases. Sequence analysis of RNA indicates that this RNA is highly G + C rich.[125] It originates from DNA. It would be interesting to determine if GC rich 5.5S RNA from *Alcaligenes faecalis* originates from 10S RNA bound to RNAse.

An RNA termed 2RNA has been isolated and purified from whole cell soluble RNA from *E. coli*. The molecular weight of this RNA is significantly higher than that of 5S RNA used as a control.[126] The 2RNA may be similar to 5.5S of RNA found in *A. faecalis*.[118]

La 4.5S RNA (80 to 100 bases) is small RNA that emerges in *E. coli* from the progenitor of 23S rRNA through fragmentation at the 3′ end.[127] This finding is in agreement with previously reported data on the presence of small RNAs in *E. coli*[128] and *Agrobacterium tumefaciens*.[129,130] La 4.5S RNA has been shown to be a group of hydrogen RNAs bounded to poly A containing nuclear of cytoplasmic RNA present in cultured Chinese hamster ovary cells.[131] This RNA from different cells has been sequenced[80,132] (Figures 5 and 6). It is less conserved than USn RNAs. La 4.5S RNA (96 residues) terminates with pppGp at its 5′ end

```
          10          20          30          40          50
pppGCCGGUAGUG  GUGGCGCACG  CCGGUAGGAU  UUGCUGAAGG  AGGCAGAGGC
       U

          60          70          80          90
AGAGGGAUCA  CGAGUUCGAG  GCCAGCCUGG  GCUACACAUU  UUUU_OH
```

FIGURE 5. Nucleotide sequence of La 4.5 RNA of mouse and hamster. (From Harada, F. and Kato, N., *Nucleic Acids Res.*, 8, 1273, 1980. With permission.)

```
          10          20          30          40        U 50
pppGGCUGGAGAG  AUGGC(UC)AGC  CGUUAAAGGC  UAGGCUCACA  ACCAAAAAUA

          60          70          80          90          98
UAAGAGUUCG  GUUCCCAGCA  CCCACGGCUG  UCUCUCCAGC  CACCUUUU(U)_OH
```

FIGURE 6. Nucleotide sequence of La 4.5 IRNA of Novikoff hepatoma (modified). (From Busch, H. et al. *Annu. Rev. Biochem.*, 51, 617, 1982. With permission).

and with a short oligo (U) sequence of variable length at its 3′ end. La 4.5S RNA contains some regions of sequence that have been found in other small RNAs, all transcribed by RNA polymerase III.

A new class of small RNA molecules, tcRNAs (translational control RNA), associated with the function of mRNAs has been found in differentiating muscle.[133] Among them two species have been purified and sequenced. A 102 nucleotide species ($tcRNA_{102}$) has the poly A binding properties. It contains 40% uracil and 5% cytosine. Another RNA ($tcRNA_{89}$) is less rich in uracil than $tcRNA_{102}$, but contains 8% of cytosine residues. Both RNAs are bound to mRNAs and are transcribed from DNA. The authors have suggested the role of mRNA-associated small RNAs in translational control. One might postulate that $tcRNA_{102}$ containing 40% of uracil would be involved in inhibiting the translation process as it has been shown for oligoribonucleotides which are particularly rich in uracil (47%).[1]

A small RNA, termed unique membrane RNA, has been purified from rat liver microsomes and characterized.[134] This RNA (3.14S) differs from tRNA and 5S rRNA and is not the result of degradation of high molecular weight RNA. It accepts amino acids rather poorly and is not present in ribosomes. Its biological function has not been defined.

Small RNAs have also been found in Ehrlich ascites tumor cells. Electrophoresis on polyacrylamide gel has shown that five of them migrate more slowly than 5S rRNA and three components migrate between 5S and 4S RNA. All these RNAs are localized in the nucleus and account for 0.7 to 2.9% of total nuclear RNAs.[135] These small RNAs have a stability comparable to rRNAs. Their synthesis is more or less efficiently inhibited with the actinomycin D that inhibits the transcription of rRNAs and 5S rRNA from DNA. It remains to be seen whether RNAs are generally the primary transcript and what their biological significance may be.

X. SMALL RNA MOLECULES AND OLIGORIBONUCLEOTIDES AS PRIMERS FOR DNA REPLICATION, TRANSCRIPTION, AND TRANSLATION

Neither prokaryotic nor eukaryotic polymerases have been reported to catalyze *de novo* and in vitro DNA synthesis.[2,136-140] Several lines of evidence demonstrate that oligoribonucleotides and small RNAs are both necessary to initiate DNA replication.[1,2,138,140-143] This variety of RNAs does modify the rhythm of DNA replication and transcription.

What is the origin of such active, small RNAs? They may appear during transcription of

DNA but may also be generated by ribonucleases which degrade giant molecules of various RNAs. It is difficult to collect large quantities of such naturally occurring small RNAs. To get around this difficulty, purified ribosomal RNAs (23S + 16S) was degraded from showdomycin resistant *E. coli* and wild type with different ribonucleases including pancreatic RNase. All RNases used contained phosphatase activity which removes the 2′ or 3′ phosphate from terminal nucleotide. A family of RNA fragments of 25 to 55 nucleotides is to be found in an RNase digest. They have been selected by their size and characterized by base ratio (purines/pyrimidines ranging from 5.8 to 0.75). Purine-rich RNA fragments interact in vitro with DNA from different origins but do not promote the in vitro synthesis of DNAs from various sources. RNA fragments bind to DNA by hydrogen bonds. This hybrid can be characterized.[2,137,144] Only those RNA fragments that are rich in purine nucleotides are excellent in vitro primers for the conversion of ϕX174 and λ DNA single strands to replicative form.[137] These data have been confirmed by isolation of the primer:

$$pppAGGGCGAAAAACCGUCU_{A_{102}A}GGGCGAUGG$$

which is required for conversion of phage fd single strand DNA to replicative form.[145] In cited cases, RNA primers contain approximately twice as many more purine bases (G + A/C + U = 2.2 to 2.4) complementary to pyrimidines present at the origin of the phage DNA replication site: TGCTCCCCCAACTTGOH(3′ end).[146,147] The fact that purine-rich RNA fragments isolated from bacterial ribosomal RNAs act as excellent primers for phage DNA replication[137] is explained by the observation that these fragments bind to the complementary bases of phage DNA.[2]

An RNA primer for DNA replication by mitochondrial DNA polymerase is generated by RNase H on a precursor transcript A transcript of 700 nucleotides synthesized from a promoter 508 bases upstream of the origin.[148] A second smaller transcript is copied from the opposite strand of the template and its length (100 nucleotides) is complementary with the promoter proximal region of the primer.[149] There is an RNA of 200 nucleotides which binds to the nascent transcript in order to prevent hybrid formation (RNA primer DNA), but not transcription.[148]

A. Small RNAs in the Translation Process

Participation of small RNAs and/or oligoribonucleotides in the translation of mRNAs into proteins has been postulated[141,150] and thereafter confirmed by experimental data.[143,151,152] Oligoribonucleotides have been found attached to initiation factors involved in protein biosynthesis. Such an RNA molecule called "i-RNA"[143] is rich in adenine (46%) and poor in guanine (7%).[153] This "i-RNA" found in reticulocytes is present in ribosomes from various sources. Although it participates in translation of mRNAs, it lacks specificity for an mRNA from a given species. A small "translating control RNA" rich in uridine, inhibits in vitro the translation of mRNA.[152] Small RNAs, isolated from newborn rats, alter the translation of mRNA in a system of wheat germ.[143] Such an RNA, containing poly U stretches, inhibits the translation of homologous or hererologous mRNA.[154] RNA without poly U appears to activate the translation.

Salt-washed ribosomes of dormant and developing *Artemia salina* embryos contain two distinct factors involved in the translation of mRNAs into proteins.[1] Both compounds are oligoribonucleotides. The inhibitor (6000 mol wt) is rich in pyrimidines (47% U, 11% A, 26% C, 16% G), sensitive to RNase A, and resistant to RNase Tl. The activator (9000 mol wt) is rich in guanine (33% U, 10% A, 6% C, 51% G), sensitive to RNase Tl, and resistant to RNase A.

Bases	Inhibitor (%)	Activator (%)
U	47	33
A	11	10
C	26	6
G	16	51
Purine/pyrimidine	0.37	1.6

The activator of translation counteracts the effect of the inhibitor. It has been found in the same fractions of developing embryos. There is little activator in undeveloped cysts. The amount of these oligoribonucleotides appears to depend on hydration of RNases.[1] Whether such RNases degrade some ribosomal or nuclear RNAs present in excess in the studied material or whether they appear during degradation of mRNAs and splicing of giant precursors RNAs has yet to be determined.

B. Free Intervening Sequences from Genes and Pseudogenes as Active Biological Molecules?

In eukaryotes, intervening sequences (introns) which interupt the coding regions of many genes specifying various RNAs are transcribed as part of precursor RNAs and then removed by the splicing process. This process consists of the excision of introns and thereafter ligation of the functional units, exons. It was reported that the intervening sequence of the tetrahymena ribosomal RNA-precursor is excised as a linear molecule which cyclizes itself in the absence of the enzyme.[155] During cyclization, the AAAUAG fragment is lost. Such a purine rich fragment may be involved in the regulation of DNA replication[2,144] or in protein biosynthesis.[1]

Two classes of genomic sequence have recently been implicated in the formation of processed genes. These are multigene families that correspond to the middle repetitive Alu-genes[156] and the multiple Sn RNA genes that encode the ubiquitous small nuclear RNAs prevalent in human and other genes.[116] Members of both gene families are transcribed into small RNAs of generally unknown function.[157] When treated with the restriction endonuclease Alu 1 gene provides two fragments, one of # 170 bp and one of # 120 bp. At least half of the 300 nucleotides and the 300 nucleotide irDNA (inverted repeated DNA) belong to a single sequence family termed the Alu family.[158]

In Chinese hamster cells each poly(A)-terminated nuclear RNA has one molecule of low molecular weight RNA (100 nucleotides) hydrogen bounded to it.[131] This small RNA may interact with sequences transcribed from the Alu family of interspersed repeated DNA.[159] These authors have suggested that the Alu family of interspersed repeated sequence may function as origins of DNA replication in mammalian cells. This remains to be confirmed by direct experimental data. Alu-like elements dispersed throughout the mammalian genome are substrates for the Alu restriction endonuclease.[160] They range in length from 135 bp in the mouse to 300 in man. Alu-like elements display a high degree of sequence homology with the small cytoplasmic 7S and 4.5S RNAs.[156] The 4.5S RNA, which has been found in mouse and hamster but not in human cells, binds through hydrogen bonds to nuclear and cytoplasmic poly(A)$^+$ RNA.[131] The sequence homology is observed for 120 bases starting from the 5′ end of the Alu monomer.[161]

Pseudogenes are partial duplicates of structural genes. In most cases, they appear not to be transcribed into RNAs, although there is no definite certainty about this. Pseudogene sequences have been characterized for gene encoding small nuclear RNAs U1, U2, U3.[162] They also have been produced through an RNA intermediate.[116] During the processing of this intermediate, some RNA sequences are lost. Their possible function and that of small nuclear RNAs encoded by pseudogene sequences have to be determined. "Processed genes and Alu-like elements are heritable and inserted into DNA in cells. 1% of mammalian DNA

contains sequences expressed as proteins, the remaining 99% may be derived and maintained by a continuous flux of such elements."[163]

C. Small RNAs: Molecular Selfreplication

The Qβ virus which infects *E. coli* has been used as a model for studying the mechanism of molecular selfreplication. The isolated Qβ replicating enzyme has the capacity to reproduce viral RNA in a cell-free system.[164] This single-stranded RNA of 4500 nucleotides possesses infectious potential. A noninfectious "satellite" RNA containing 220 nucleotides purified from *E. coli* cells has also been used as template and in vitro is replicated by the Qβ replicase. The nucleotide sequence of the relevant parts of Qβ RNA has been analyzed.[165] In this small RNA, the sequence-CCC has been identified[166] as the recognition site for interaction with Qβ replicase. A remarkable property of Qβ replicase is that it is capable of synthesizing RNA from four ribonucleoside-5′-triphosphates, even in the absence of RNA template. This enzyme accepts only few natural RNAs and some synthetic ribopolymers as template for RNA in vitro synthesis.[167,168]

RNA synthesis always starts with the incorporation of GTP. The 3′ terminating nucleotide of the template does not need to be cytidylic acid. However, cytidylic acid polymers are accepted as template only when they contain at their 3′ end a cytidylic acid sequence or more than 5 nucleotides.[168] Randomly mixed ribopolymers containing cytidylic acid poly(C) are accepted as active template with Qβ replicase.[169]

Transforming 6S RNA excreted by showdomycin-resistant *E. coli* differs from other RNAs species found in the wild type strain. This small RNA is purine rich (G + A/C + U = 1.82) in comparison to RNAs from the wild type of *E. coli* (G + A/C + U = 1.0). In Cs_2SO_4 gradient centrifugation, 6S RNA sediments exclusively in the region of RNA and does not contain DNA nor does it hybridize with DNA from the same strain. It was unexpected to observe that 6S RNA can interfere in the in vitro activity of polynucleotide phosphorylase (PNPase) in the exclusive presence of all four ribonucleoside-5′-diphosphates (XDPs) in the incubation mixture.[170] PNPase^6 from the wild strain synthesizes, from equal amounts of XDPs, a poly AGUC whose ratio is 1:1:1:1. In the presence of 6S RNA, the rate of synthesis is increased several-fold and the base ratio of the synthesized product G + A/C + U = 1.75 is practically identical to that found for 6S RNA (G + A/C + U = 1.80). RNase A abolishes template activity of 6S RNA. XTPs cannot replace XDPs. Rifampicin is without effect on the activity of PNPase, while it inhibits that of DNA-dependent RNA polymerase. Crude soluble extracts of 250-fold purified enzyme gives the same results. Ribosomal RNAs, tRNAs, polyovirus RNAs, and poly AGUC are not used as template by PNPase. On sucrose gradient, the ^{14}C-synthesized product made in the presence of 6S RNA, sediments at a position differing from that of the ^{14}C polymer-synthesized in the absence of 6S RNA.[171]

PNPase from showdomycin-resistant *E. coli*, incubated in the presence of all four XDPs each used in equal amounts, synthesizes in vitro AGUC polymer in which the amount of purine bases is twice that of pyrimidines. This shows that modified PNPase may synthesize an RNA that does not hybridize with DNA. PNPase from *E. coli* wild strain is capable of synthesizing small purine rich RNA (G + A/C + U = 1.5 to 1.8) from XDPs at 70°C, acting as primers in in vitro replication of DNAs from bacteria and mammals, but it has a low effect on phage DNAs under the same experimental conditions.[170] Also, purine rich RNA synthesized by PNPase under normal conditions but in the presence of RNase A acts as primer for replication of some DNAs.[171] Thus, it may be imagined that PNPase and RNase, both capable of behaving as "thermoresistant enzymes," might have played a decisive role in synthesizing different RNA primers or even templates for the formation of cell constituents.

D. Transcription of Small Cellular RNAs into DNA

Interest in the conversion of RNA molecules into DNA has recently increased, since this process of transcription seems to be the only one that may explain the creation of pseudo-

genes.[163] It has been suggested that small RNA molecules transcribed first on DNA may carry the information back to chromosomal DNA but in different segments of DNA than those from which they originate. This concept implies the existence in eukaryotes and prokaryotes of an enzyme, reverse transcriptase, capable of transcribing small RNAs into DNA.

Reverse transcriptase has been discovered in animal tumor viruses,[172] and several reports have been published on transcription of RNA templates into complementary DNA by enzymes found in chick embryos,[173] in bacteria,[120,170,174,175] in plants,[176] in monkey placenta,[177] in normal human lymphocytes,[178] and in fungi.[3,4] 28S ribosomal RNA from *Drosophila melanogaster* can be transcribed into DNA-like material by a DNA polymerase I purified to about 90%. The size of the ^{3}H-DNA product is about 4S, but it may be longer.[179] 25S ribosomal RNA transcription was found to occur to some extent in the absence of the added oligonucleotide primer.[180]

A free RNA-bound "reverse transcriptase", isolated from *E. coli* extracts, synthesizes in vitro DNA, providing that all four ribonucleoside-5′-triphosphates are present in the incubation mixture. Both template RNA and synthesized ^{3}H-DNA have a size of about 6S. ^{3}H-DNA is complementary to template RNA.[120] It is possible to distinguish the RNA-bound enzyme from known DNA-dependent DNA polymerase on several criteria. A particular purine-rich small RNA (6S) excreted by showdomycin-resistant *E. coli* has a transforming potential.[128,181] When added to the culture medium of a tumorogenic strain of *A. tumefaciens*, it is possible to isolate a series of transformants which have partially or totally lost their tumor potential. Different biochemical changes have been observed in these transformants. Hybridization data showed that the RNA has penetrated into recipient bacteria and one copy at least was found in the DNA of complete transformants. In vitro this RNA can be transcribed into complementary DNA. It possesses selfpairing regions and probably a 3′ OH looped terminus required for transcription. When these different small RNAs are injected into *Datura stramonium* inverted stem section under axenic conditions and in the presence of auxin, they lead to the appearance of characteristic Crown-gall tumors[176,182-184] (Figures 7 and 8).

One can easily isolate the endogeneous RNA-bound RNA-dependent DNA polymerase from the microsomal pellet fraction of the eukaryotic fungus *Neurospora crassa*.[3,4] Here also, free template RNA and the synthesized DNA sedimented on sucrose density gradients at approximately 5 to 6S. ^{3}H-DNA is complementary to template RNA as judged by hybridization data. It is quite probable that the eukaryotic organism may well possess RNA-bound "reverse transcriptase". It will be of importance to determine which small RNAs and/or intervening sequences might be transcribed into complementary DNA in eukaryotes. If such data can be accurately established using different species, this would strongly support the concept of pseudogenes formation, as recently suggested.[163] In this respect it has been recently reported that *Drosophila melanogaster* contains nonvirales particles containing 4S, 4.5S, 5S, and 6S molecules of RNA as the major constituents. In addition, these particles contain particle-bound reverse transcriptase.[185] This observation strongly supports several findings reported many years ago[3,4,120] and discussed here.

XI. CONCLUDING REMARKS

The majority of small RNAs from eukaryotes, prokaryotes, and virions described here have been sequenced and their secondary structure proposed. Other small RNAs not sequenced but characterized by chemical and physical means have been studied mostly for their biological activities. Many small RNAs appear to originate from DNA as direct transcripts, while others emerge from RNA precursor transcripts or even from mature RNAs. One common feature of small RNAs is that they bind with relative efficiency to large RNAs (rRNAs, mRNAs, viral genome RNAs), and DNAs through hydrogen bonding. This process

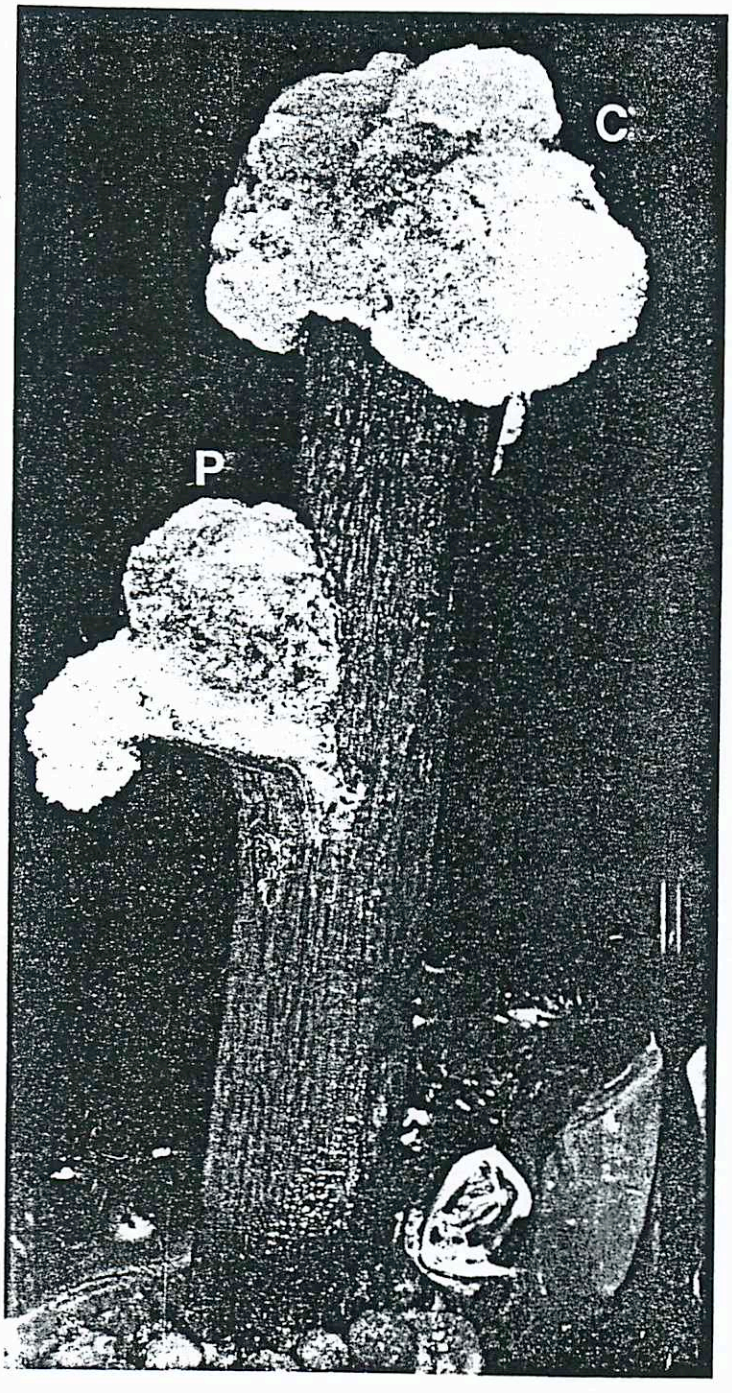

FIGURE 7. Demonstration of the oncogenic capacity of the tumor-inducing small-size RNA from *A. tumefaciens*. Tumor obtained by inoculation of RNA into a fragment of inverted stem of *Datura stramonium* cultured in vitro. C, auxinic callus tissue; P, overgrowth tissue obtained with RNA. (From Le Goff, L. et al., *Can. J. Microbiol.*, 22, 694, 1976. With permission.)

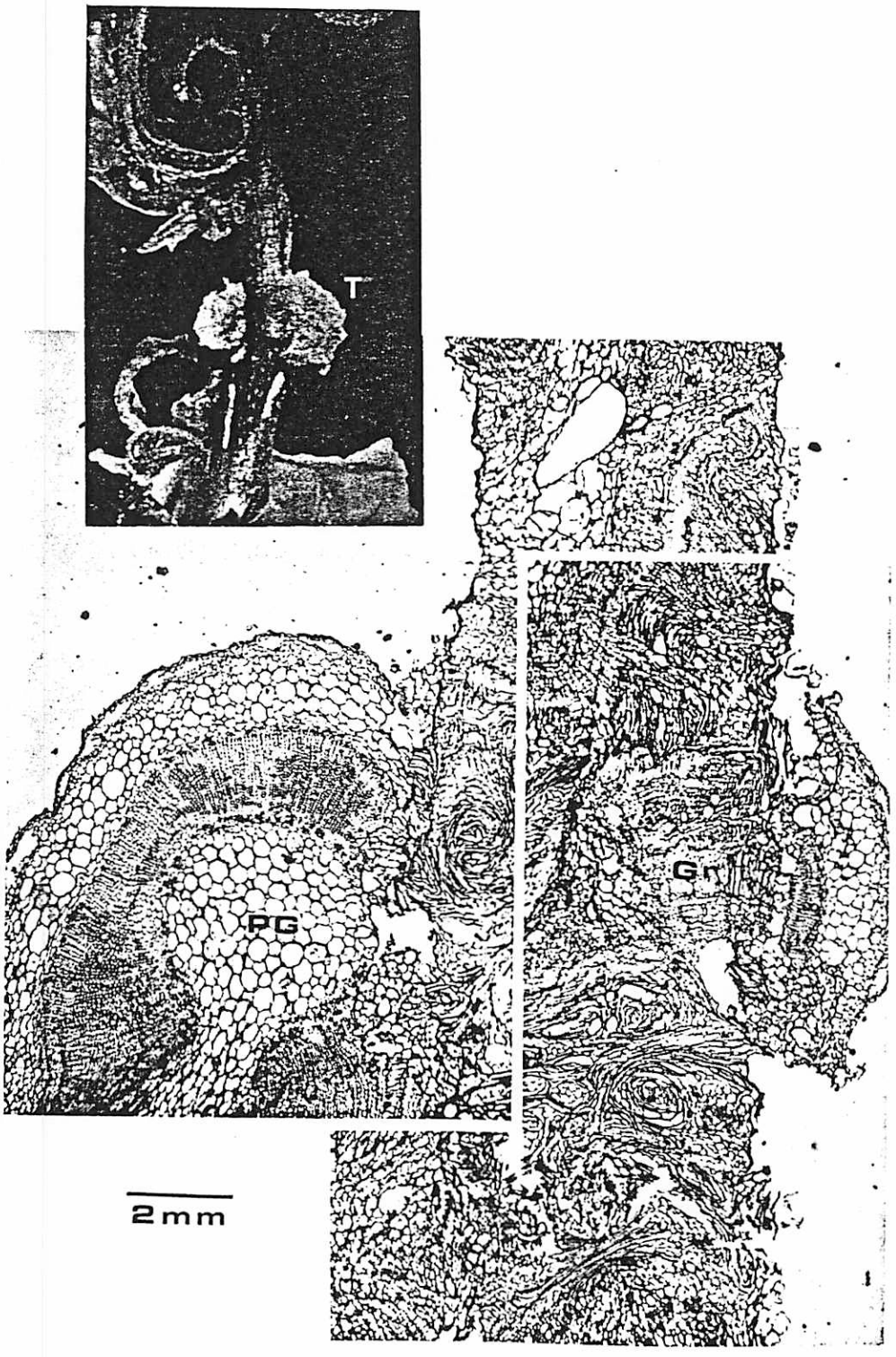

FIGURE 8. Demonstration of the oncogenic capacity of the tumor-inducing small-size RNA from *A. tumefaciens*. Development of secondary tumor (T) after grafting of the primary overgrowth tissue. Schematic section of a tumor fragment (Gr) engrafted upon stem of healthy plant (PG). (From Aaron-Da Cunha, M. I. et al., *C. R. Soc. Biol.*, 169(3), 755, 1975. With permission.)

necessarily implies sequence homology in certain regions. Comparison of sequences suggests that there exist relative phylogenetic trends between several small and large RNAs, as well as between different species of small RNAs themselves. From this, one may conclude that small RNAs have largely been conserved throughout evolution of various species; understanding of the biological activities of the RNAs should greatly elucidate this conservation process.

Over the last decade, few scientists have been involved in studying the extent to which small RNAs can in vitro and *in situ* be transcribed into DNA. Since the recent discovery of pseudogenes, this problem has come to the fore with particular urgency, partly because some regions of pseudogenes are transcribed into small RNAs. It is at present postulated but not proven as far as we know, that small RNAs, via transcription to DNA, are the only candidates for the creation of new genes or pseudogenes. Also, it should not be forgotten that some small RNAs can be in vitro synthesized *de novo* in the absence of DNA or RNA template. However, it remains difficult to ascertain that this could happen *in situ* in different biological systems. The role of such RNAs may have been of importance during the early stages of the appearance of RNAs or DNAs, i.e., at the creation of life.

As shown here, many small RNAs have been characterized by excellent analytical methods. However, their respective biological roles in the cells of different species need to be demonstrated accurately in order to elucidate their possible phylogenetic relationship and their part in the life of the cell.

REFERENCES

1. **Lee-Huang, S., Sierra, J. M., Naranjo, R., Filipowitz, W., and Ochoa, S.,** Eukaryotic oligonucleotides affecting mRNA translation, *Arch. Biochem. Biophys.*, 180, 276, 1977.
2. **Beljanski, M.,** The regulation of DNA replication and transcription. The role of trigger molecules in normal and malignant gene expression, *Experimental Biology and Medicine*, Vol. 8, Wolsky, A., Ed., S. Karger, Basel, 1983.
3. **Dutta, S. K., Beljanski, M., and Bourgarel, P.,** Endogenous RNA-bound RNA-dependent DNA polymerase activity in *Neurospora crassa*, *Exp. Mycol.*, 1, 173, 1977.
4. **Dutta, S. K., Mukhopadhyay, D. K., and Bhattacharyya, J.,** RNase-sensitive DNA polymerase activity in cell fractions and mutants of *Neurospora crassa*, *Biochem. Genet.*, 18, 743, 1980.
5. **Beljanski, M., Plawecki, M., Bourgarel, P., and Beljanski, M.,** Leukocyte recovery with short-chain RNA-fragments in cyclophosphamide-treated rabbits, *Cancer Treat. Rep.*, 67, 611, 1983.
6. **Ochoa, S.,** Enzymic synthesis of polynucloetides. III. Phosphorolysis of natural and synthetic ribopolynucleotides, *Arch. Biochem. Biophys.*, 69, 119, 1957.
7. **Holley, R. W., Apgar, J., Everett, G. A., Madison, J. T., Marquisee, M., Merrill, S. H., Penswick, J. R., and Zamir, A.,** Structure of a ribonucleic acid, *Science*, 147, 1462, 1965.
8. **Sundaralingam, M.,** Nucleic acid principles and transfer RNA, in *Biomolecular Structure Conformation and Function*, Evol. Proc. Int. Symp., Vol. 1, Srinivasan, R., Subramanian, E., and Yathindra, N., Eds., Pergamon Press, Oxford, 1981, 259.
9. **Woese, C. R.,** *The Genetic Code: The Molecular Basis for Genetic Expression*, Harper & Row, New York, 1967.
10. **Nishimura, S.,** Modified nucleosides in tRNA, in *Transfer RNA: Structure, Properties and Recognition*, Schimmell, P. R., Söll, D., and Abelson, J. N., Eds., Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1979, 59.
11. **McCloskey, J. A. and Nishimura, S.,** Modified nucleosides in transfer RNA, *Acc. Chem. Res.*, 10, 403, 1977.
12. **Sprinzl, M., Grüter, F., and Gauss, D. H.,** Compilation of tRNA sequences, *Nucleic Acids Res.*, 5(Suppl.), r15, 1978.
13. **Hoagland, M. B., Zamecnik, P. C., and Stephenson, M. L.,** Intermediate reactions in protein biosynthesis, *Biochim. Biophys. Acta*, 24, 215, 1957.
14. **Zachau, H. G.,** Transfer ribonucleic acids, *Angew. Chem. Int. Ed.*, 8, 711, 1969.

15. **Burdon, R. H., Martin, B. T., and Lal, B. M.,** Synthesis of low molecular weight ribonucleic acid in tumor cells, *J. Mol. Biol.*, 28, 357, 1967.
16. **Darnell, J. E.,** Ribonucleic acids from animal cells, *Bacteriol. Rev.*, 32, 262, 1968.
17. **Altman, S. and Smith, J. D.,** Tyrosine tRNA precursor molecule polynucleotide sequence, *Nature (London) New Biol.*, 233, 35, 1971.
18. **Abelson, J.,** The organization of tRNA genes, in *Transfer RNA: Biological Aspects*, Söll, D., Abelson, J. N., and Schimmel, P. R., Eds., Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1980, 211.
19. **Mazzara, G. P. and McClain, W. H.,** tRNA synthesis, in *Transfer RNA: Biological Aspects*, Söll, D., Abelson, J. N., and Schimmel, P. R., Eds., Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1980, 3.
20. **O'Farrell, P. Z., Cordell, B., Valenzuela, P., Rutter, W. J., and Goodman, H. M.,** Structure and processing of yeast precursor tRNA containing intervening sequences, *Nature (London)*, 274, 438, 1978.
21. **Ryan, M. J., Brown, E. L., Belagaje, R., Khorona, H. G., and Fritz, H. J.,** Cloning of two chemically synthesized genes for a precursor to the Su^{3+} suppressor $tRNA^{Tyr}$, in *Transfer RNA: Biological Aspects*, Söll, D., Abelson, J. N., and Schimmel, P. R., Eds., Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1980, 245.
22. **Khorana, H. G.,** Total synthesis of a gene, *Science*, 203, 614, 1979.
23. **Cortese, R., Melton, D. A., Tranquilla, T., and Smith, J. D.,** Cloning of nematode tRNA genes and their expression in the frog oocyte, *Nucleic Acids Res.*, 5, 4593, 1978.
24. **Schweizer, E., Mackechnie, C., and Halvorson, H. O.,** Redundancy of ribosomal and transfer RNA genes in *Saccharomyces cerevisiae*, *J. Mol. Biol.*, 40, 261, 1969.
25. **Weber, L. and Berger, E.,** Base sequence complexity of the stable RNA species of *Drosophila melanogaster*, *Biochemistry*, 15, 5511, 1976.
26. **Clarkson, S. G. and Birnstiel, M. L.,** Clustered arrangement of tRNA genes of *Xenopus laevis*, *Cold Spring Harbor Symp. Quant. Biol.*, 38, 451, 1974.
27. **Martin, M. C., Rabinowitz, M., and Fukuhara, H.,** Yeast mitochondrial DNA specifies tRNA for 19 amino acids. Deletion mapping of the tRNA genes, *Biochemistry*, 16, 4672, 1977.
28. **Angerer, L., Davidson, N., Murphy, W., Lynch, P., and Attardi, G.,** An electron microscope study of the relative positions of the 4S and ribosomal RNA genes in HeLa cell mitochondrial DNA, *Cell*, 9, 81, 1976.
29. **Dawid, I. B., Klukas, C. K., Ohi, S., Ramirez, J. L., and Upholt, W. B.,** Structure and evolution of animal mitonchondrial DNA, in *The Genetic Function of Mitochondrial DNA*, Saccone, C. and Kroon, A. M., Eds., North Holland, Amsterdam, 1976, 3.
30. **Wittmann, H. G.,** Components of bacterial ribosomes, *Ann. Rev. Biochem.*, 51, 155, 1982.
31. **Cantatore, P., De Benedetto, C., Gadaleta, G., Gallerani, R., Kroon, A. M., Holtrop, M., Lanave, C., Pepe, G., Quagliariello, C., Saccone, C., and Sbisa, E.,** The nucleotide sequences of several tRNA genes from rat mitochondria: common features and relatedness to homologous species, *Nucleic Acids Res.*, 10, 3279, 1982.
32. **Clarkson, S. G., Birnsteil, M. L., and Purdon, I. F.,** Clustering of transfer RNA genes of *Xenopus laevis*, *J. Mol. Biol.*, 79, 411, 1973.
33. **Deutscher, M. P.,** Synthesis and functions of the -C-C-A terminus of transfer RNA, *Prog. Nucleic Acid Res. Mol. Biol.*, 13, 51, 1973.
34. **Daniel, V., Sarid, S., and Littauer, U. Z.,** Bacteriophage induced transfer RNA in *Escherichia coli*, *Science*, 167, 1682, 1970.
35. **Gruissem, W. and Seifart, K. H.,** Transcription of 5SRNA genes *in vitro* is feedback-inhibited by HeLa 5SRNA, *J. Biol. Chem.*, 257, 1468, 1982.
36. **Beaudreau, G. S., Sverak, L., Zischka, R., and Beard, J. W.,** Attachment of ^{14}C amino acids to BA1 strain A (myeloblastosis) avian tumor virus RNA, *Natl. Cancer Inst. Monogr.*, 17, 791, 1964.
37. **Bonar, R. A., Sverak, L., Bolognesi, D. P., Langlois, A. J., Beard, D., and Beard, J. W.,** Ribonucleic acid components of BA 1 strain A (myeloblastosis) avian tumor virus, *Cancer Res.*, 27, 1138, 1967.
38. **Travnicek, M.,** RNA with amino acid-acceptor activity isolated from an oncogenic virus, *Biochim. Biophys. Acta*, 166, 757, 1968.
39. **Erikson, E. and Erikson, R. L.,** Isolation of amino acid-acceptor RNA from purified avian myeloblastosis virus, *J. Mol. Biol.*, 52, 387, 1970.
40. **Erikson, E. and Erikson, R. L.,** Association of 4S ribonucleic acid with oncornavirus ribonucleic acids, *J. Virol.*, 8, 254, 1971.
41. **Sawyer, R. C. and Dahlberg, J. E.,** Small RNAs of Rous sarcoma virus: characterization of two dimensional polyacrylamide gel electrophoresis and fingerprint analysis, *J. Virol.*, 12, 1126, 1973.
42. **Waters, L. C. and Mullin, B. C.,** Transfer RNA in RNA tumor viruses, *Prog. Nucleic Acid Res. Mol. Biol.*, 20, 131, 1977.

43. **Yang, W. K. and Hwang, D. L. R.,** *In vitro* selective binding of tRNAs to rRNAs of vertebrates, in *Transfer RNA: Biological Aspects*, Söll, D., Abelson, J. N., and Schimmel, P. R., Eds., Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1980, 517.
44. **Richter, D., Erdman, V. A., and Sprinzl, M.,** Specific recognition of GTψ C loop (loop IV) of tRNA by 50S ribosomal subunits from *E. coli*, *Nature (London) New Biol.*, 246, 132, 1973.
45. **Elson, D.,** A ribonucleic acid particle released from ribosomes by salt, *Biochim. Biophys. Acta*, 53, 232, 1961.
46. **Rosset, R. and Monier, R.,** A propos de la présence d'acide ribonucléique de faible poids moléculaire dans les ribosomes d' *E. coli*, *Biochim. Biophy. Acta*, 68, 653, 1963.
47. **Sanger, F., Brownlee, G. G., and Barrell, B. G.,** Sequence of 5S ribosomal RNA, *Structure, Function, and Transfer RNA 5SRNA*, Proc. Meet. Fed. Euro. Biochem. Soc., 4th ed., Froeholm, L. O., Ed., Academic Press, New York, 1968, 1.
48. **Fedoroff, N. V. and Brown, D. D.,** The nculeotide sequence of oocyte 5S DNA in *Xenopus laevis*. I. The AT-rich spacer, *Cell*, 13, 701, 1978.
49. **Furth, J. J., Ormsby, J., Su, C. Y., and Averyhart-Fullard, V.,** Transcription of calf thymus 5S ribosomal DNA and calf thymus satelite DNA by RNA polymerase III, *Fed Proc. Fed. Am. Soc. Exp. Biol.*, 41, 1294, 1982.
50. **Komiya, H., Hasegawa, M., and Takemura, S.,** Nucleotide sequences of 5SrRNA from sponge *Halichondria japonica* and tunicate *Halocynthia roretzi* and their phylogenetic positions, *Nucleic Acids Res.*, 11, 1964, 1983.
51. **Woese, C. R., Magrum, L. J., and Fox, G. E.,** Archaebacteria, *J. Mol. Evol.*, 11, 245, 1978.
52. **Studnicka, G. M., Eiserling, F. A., and Lake, J. A.,** A unique secondary folding pattern for 5SRNA corresponds to the lowest energy homologous secondary structure in 17 different prokaryotes, *Nucleic Acids Res.*, 9, 1885, 1981.
53. **Brown, D. D., Wensink, P. C., and Jordan, E.,** Purification and some characteristics of 5SDNA from *Xenopus laevis*, *Proc. Natl. Acad. Sci. U.S.A.*, 68, 3175, 1971.
54. **Brownlee, G. G., Cartwright, E. M., and Brown, D. D.,** Sequence studies of the 5SDNA of *Xenopus laevis*, *J. Mol. Biol.*, 89, 703, 1974.
55. **Miller, J. R., Cartwright, E. M., Brownlee, G. G., Fedoroff, N. V., and Brown, D. D.,** The nucleotide sequence of oocyte 5SDNA in *Xenopus laevis*. II. The GC-rich region, *Cell*, 13, 717, 1978.
56. **Brown, D. D. and Sugimoto, K.,** 5SDNAs of *Xenopus laevis* and *Xenopus mulleri:* evolution of a gene family, *J. Mol. Biol.*, 78, 397, 1973.
57. **Korn, L. J. and Brown, D. D.,** Nucleotide sequence of *Xenopus borealis* oocyte 5SDNA: comparison of sequences that flank related eukaryotic genes, *Cell*, 15, 1145, 1978.
58. **Ford, P. J. and Brown, R. D.,** Sequences of 5S ribosomal RNA from *Xenopus mulleri* and the evolution of 5S gene-coding sequences, *Cell*, 8, 485, 1976.
59. **Pribula, C. D., Fox, G. E., Woese, C. R., Sogin, M., and Pace, N.,** Nucleotide sequence of *Bacillus megaterium* 5SRNA, *FEBS Lett.*, 44, 322, 1974.
60. **Brownlee, G. G. and Cartwright, E. M.,** The nucleotide sequence of the 5SRNA of chicken embryo fibroblasts, *Nucleic Acids Res.*, 2, 2279, 1975.
61. **Nomura, M. and Erdmann, V. A.,** Reconstitution of 50S ribosomal subunits from dissociated molecular components, *Nature (London)*, 228, 744, 1970.
62. **Wrede, P. and Erdmann, V. A.,** *Escherichia coli* 5SRNA binding proteins L18 and L25 interact with 5.8SRNA but not with 5SRNA from yeast ribosomes, *Proc. Natl. Acad. Sci. U.S.A.*, 74, 2706, 1977.
63. **Wreschner, D. H.,** The role of ribosomal RNA in protein synthesis. Inhibition of translation by reticulocyte 5S ribosomal RNA, *FEBS Lett.*, 94, 139, 1978.
64. **Wreschner, D. H.,** Reticulocyte 5S ribosomal RNA inhibition of cell-free protein synthesis. Novel responses in ribosomal behaviour, *FEBS Lett.*, 94, 145, 1978.
65. **Raacke, I. D.,** A model for protein synthesis involving the intermediate formation of peptidyl-5SRNA, *Proc. Natl. Acad. Sci. U.S.A.*, 68, 2357, 1971.
66. **Kao, T. H. and Crothers, D. M.,** A proton-coupled conformational switch of *Escherichia coli* 5S ribosomal RNA, *Proc. Natl. Acad. Sci. U.S.A.*, 77, 3360, 1980.
67. **Mullins, D. W., Jr., Lacey, J. C., Jr., and Hearn, R. A.,** 5SrRNA and tRNA: evidence of a common evolutionary origin, *Nature (London) New Biol.*, 242, 80, 1973.
68. **Spirin, A. S. and Gavrilova, L. P.,** *The Ribosome*, Springer-Verlag, New York, 1969.
69. **Veldman, G. M., Brand, R. C., Klootwijk, J., and Planta, R. J.,** Some characteristics of processing sites in ribosomal procursor RNA of yeast, *Nucleic Acids Res.*, 8, 2907, 1980.
70. **Nazar, R. N.,** A 5.8SrRNA-like sequence in prokaryotic 23SrRNA, *FEBS Lett.*, 119, 212, 1980.
71. **Sankoff, D., Cedergren, R. J., and McKay, W.,** A strategy for sequence phylogeny research, *Nucleic Acids Res.*, 10, 421, 1982.
72. **Fujiwara, H., Kawata, Y., and Ishikawa, H.,** Primary and secondary structure of 5.8S rRNA from the silk gland of *Bombix mori*, *Nucleic Acids Res.*, 10, 2415, 1982.

73. **Lee, J. C. and Henry, B.,** Binding of rat ribosomal proteins to yeast 5.8S ribosomal ribonucleic acid, *Nucleic Acids Res.*, 10, 2199, 1982.
74. **Virmaux, N., Mandel, P., and Urban, P. F.,** Evidence of two soluble RNA types in eye lens, *Biochem. Biophys. Res. Commun.*, 16, 308, 1964.
75. **Erikson, E., Erikson, R. L., Henry, B., and Pace, N. R.,** Comparison of oligonucleotides produced by RNase TI digestion of 7SRNA from avian and murine oncornaviruses and from uninfected cells, *Virology*, 53, 40, 1973.
76. **Bishop, J. M., Levinson, W., Quintrell, N., Sullivan, D., Fanshier, L., and Jackson, J.,** Low molecular weight RNAs of Rous sarcoma virus. II. 7SRNA, *Virology*, 42, 927, 1970.
77. **Busch, H.,** The function of the 5′ cap of mRNA and nuclear species, *Perspect. Biol. Med.*, 19, 549, 1976.
78. **Li, W. Y., Reddy, R., Henning, D. H., Epstein, P. M., and Busch, H.,** Nucleotide sequence of 7SRNA. Homology to Alu DNA and La 4.5SRNA, *J. Biol. Chem.*, 257, 5136, 1982.
79. **Reddy, R., Henning, D., and Busch, H.,** The primary nucleotide sequence of U4RNA, *J. Biol. Chem.*, 256, 3532, 1981.
80. **Busch, H., Reddy, R., Rothblum, L., and Choi, Y. C.,** SnRNAs, SnRNPs and RNA processing, *Annu. Rev. Biochem.*, 51, 617, 1982.
81. **Ullu, E., Murphy, S., and Melli, M.,** Human 7SL RNA consists of 140 nucleotide middle-repetitive sequence inserted in an Alu sequence, *Cell*, 29, 195, 1982.
82. **Desphande, A. K., Niu, L. C., and Niu, M. C.,** Requirement of informational molecules in heart formation, in *The Role of RNA in Reproduction and Development*, Niu, M. C. and Segal, S. J., Eds., North-Holland/American Elsevier, New York, 1973, 229.
83. **Desphande, A. K. and Siddiqui, M. A. Q.,** Acetylcholinesterase differentiation during myogenesis in early chick embryonic cells caused by inducer RNA, *Differentiation*, 10, 133, 1978.
84. **Walter, P. and Blobel, G.,** Signal recognition particles contains a 7SRNA essential for protein translocation across the endoplasmic reticulum, *Nature (London)*, 229, 691, 1982.
85. **Tomizawa, J., Tateo, I., Selzer, G., and Som, T.,** Inhibition of ColEl RNA primer formation by a plasmid-specified small RNA, *Proc. Natl. Acad. Sci. U.S.A.*, 78, 1421, 1981.
86. **De Wilde, M., Davies, J. E., and Schmidt, F. J.,** Low molecular weight RNA species encoded by a multiple drug resistance plasmid, *Proc. Natl. Acad. Sci. U.S.A.*, 75, 3673, 1978.
87. **Roozen, K. J., Fenwick, R. G., Jr., and Curtiss, R.,** Synthesis of ribonucleic acid and protein in plasmid-containing minicells of *Escherichia coli* K12, *J. Bacteriol.*, 107, 21, 1971.
88. **Campos, R., Javanovich, S., and Villarreal, L. P.,** A small RNA complementarity to an intervening sequence produced late in SV40 infection, *Nature (London)*, 291, 344, 1981.
89. **Lerner, M. R., Andrews, N. C., Miller, G., and Steitz, J. A.,** Two small RNAs encoded by Epstein-Barr virus and complexed with protein are precipitated by antibodies from patients with systemic lupus erythematosus, *Proc. Natl. Acad. Sci. U.S.A.*, 78, 805, 1981.
90. **Colonno, R. J. and Banerjee, A. K.,** A unique RNA species involved in initiation of vesicular stomatitis virus RNA transcription *in vitro*, *Cell*, 8, 197, 1976.
91. **Emerson, S. V., Dierks, P. M., and Parson, J. T.,** *In vitro* synthesis of a unique RNA species by a T particle of vesicular stomatitis virus, *J. Virol.*, 23, 708, 1977.
92. **Testa, D., Chanda, P. K., and Banerjee, A. K.,** Unique mode of transcription *in vitro* by vesicular stomatitis virus, *Cell*, 21, 267, 1980.
93. **Mathews, M. B.,** Genes for VA-RNA in adenovirus 2, *Cell*, 6, 223, 1975.
94. **Rao, D. D. and Huang, A. S.,** Synthesis of a small RNA in cells coinfected by standard and defective interfering particles of vesicular stomatitis virus, *Proc. Natl. Acad. Sci. U.S.A.*, 76, 3742, 1979.
95. **Colonno, R. J. and Banerjee, A. K.,** Complete nucleotide sequence of the leader RNA synthesized *in vitro* by vesicular stomatitis virus, *Cell*, 15, 93, 1978.
96. **Semler, B. L., Perrault, J., Abelson, J., and Holland, J. J.,** Sequence of a RNA templated by the 3′ OH RNA terminus of defective interfering particles of vesicular stomatitis virus, *Proc. Natl. Acad. Sci. U.S.A.*, 75, 4704, 1978.
97. **Plawecki, M. and Beljanski, M.,** Comparative study of *Escherichia* endotoxin, hydrocortisone, and Beljanski-Leukocyte-Restorer activity in cyclophosphamide-treated rabbits, *Proc. Soc. Exp. Biol. Med.*, 168, 408, 1981.
98. **Alwine, J. C. and Khoury, G.,** Simian virus 40-associated small RNA: mapping on the simian virus 40 genome and characterization of its synthesis, *J. Virol.*, 36, 701, 1980.
99. **Sawyer, R. C., Harada, F., and Dahlberg, J. E.,** Virion-associated RNA primer for Rous sarcoma virus DNA synthesis: isolation from uninfected cells, *J. Virol.*, 13, 1302, 1974.
100. **Harada, F., Kato, N., and Noshino, H. O.,** Series of 4.5S RNAs associated with poly(A)-containing RNAs of rodent cells, *Nucleic Acids Res.*, 7, 909, 1979.
101. **Gross, H. J., Domdey, H., Lossow, C., Jank, P., Raba, M., Alberty, H., and Saenger, H. L.,** Nucleotide sequence and secondary structure of potato spindle tuber viroid, *Nature (London)*, 273, 203, 1978.