Volume 10 Number 22 1982 Nucleic Acids Research

Generalized structures of the 5S ribosomal RNAs

Nicholas Delihas and Janet Andersen

Department of Microbiology, School of Medicine, State University of New York, Stony Brook, NY 11794, USA

Received 25 June 1982; Revised and Accepted 20 September 1982

ABSTRACT

The sequences of 5S ribosomal RNAs from a wide-range of organisms have been compared. All sequences fit a generalized 5S RNA secondary structural model. Twenty-three nucleotide positions are found universally, i.e., in 5S RNAs of eukaryotes, prokaryotes, archaebacteria, chloroplasts and mitochondria. One major distinguishing feature between the prokaryotic and eukaryotic 5S RNAs is the number of nucleotide positions between certain universal positions, e.g., prokaryotic 5S RNAs have three positions between the universal positions PuU_{40} and G_{44} (using the E. coli numbering system) and eukaryotic 5S RNAs have two. The archaebacterial 5S RNAs appear to resemble the eukaryotic 5S RNAs to varying degrees depending on the species of archaebacteria although all the RNAs conform with the prokaryotic "rule" of chain length between PuU_{40} and G_{44}. The green plant chloroplast and wheat mitochondrial 5S RNAs appear prokaryotic-like when comparing the number of positions between universal nucleotides. Nucleotide positions common to eukaryotic 5S RNAs have been mapped; in addition, nucleotide sequences, helix lengths and looped-out residues specific to phyla are proposed. Several of the common nucleotides found in the 5S RNAs of metazoan somatic tissue differ in the 5S RNAs of oocytes. These changes may indicate an important functional role of the 5S RNA during oocyte maturation.

INTRODUCTION

The nucleotide sequences of 5S ribosomal RNAs from a wide cross-section of eukaryotic organisms representative of the metazoa (vertebrates and arthropods), protozoa, plants, algae and fungi have been determined (1). With the recent elucidation of several plant 5S RNA structures (2,3) and the correction of an algal 5S RNA sequence (4), it is now apparent that eukaryotic 5S RNAs have common nucleotides in particular regions of the molecule. Furthermore, the prokaryotic (eubacteria), archaebacterial, and organelle 5S RNAs share several of these positions; these we term universal positions. These positions are distributed throughout the 5S molecule and reveal a remarkable structural constancy in the compared 5S RNA sequences. The number of nucleotide positions between several of the universal nucleotides appears to be specific to either the prokaryotic or eukaryotic 5S

0305-1048/82/1022-7323$2.00/0

RNAs and this property forms a major distinguishing feature between the 5S RNAs of prokaryotes and eukaryotes. Organelle 5S RNAs share the prokaryotic "rules" for chain lengths between universal positions whereas the archaebacterial 5S RNAs have properties of both the prokaryotic and eukaryotic 5S RNAs.

In this paper, we propose a generalized 5S RNA structure showing the universal positions and discuss the phylogenetic and functional aspects of this generalized structure. Regional structural features (group-specific signatures) of the eukaryotic 5S RNAs that appear to have taxonomic significance are also presented.

DEFINITION OF TERMS USED

Universal positions are those positions that are found in all classes of 5S RNAs, i.e., from prokaryotes, eukaryotes, archaebacteria, and organelles even though they may not be present in all known sequences. *Invariant positions* are those found in all compared 5S RNAs. *Common positions* are those that are found in 90% or more of the eukaryotic 5S RNAs from somatic cells of mature organisms. Universal and common positions have been mapped using the base-paired regions of the generalized 5S RNA secondary structural model as points of reference. *Group-specific signatures* are those regions of the eukaryotic 5S RNA that have common structural properties among related organisms.

GENERALIZED STRUCTURE OF THE 5S RIBOSOMAL RNA AND UNIVERSAL NUCLEOTIDE POSITIONS

Fig. 1 represents the generalized 5S RNA structure derived by comparative analysis of over 60 sequences of 5S RNAs from eukaryotes,

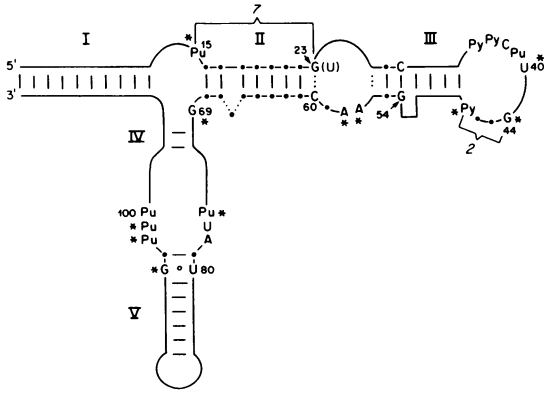


Fig. 1. Generalized structure for the 5S ribosomal RNA. The numbering system corresponds to that of the E. coli 5S RNA. Nucleotide positions shown are universal positions. Positions marked with asterisks designate total invariance. The dots represent constant chain lengths between universal positions found in most compared 5S RNAs; however the chain lengths between Pu_{15} and G_{23} and G_{44} and Py_{47} are always 7 and 2, respectively. Pu represents purine; Py represents pyrimidine.

prokaryotes, archaebacteria, and organelles (Table 1). The model is representative of the E. coli 5S RNA sequence and is shown in a secondary structural form containing five double-helical regions as proposed by Studnika et al. (63) for prokaryotic 5S RNAs. Nishikawa and Takemura (29) and Luehrsen and Fox (4) have proposed a similar model for 5S RNAs from eukaryotes. This model essentially represents the Fox and Woese base-pairing scheme originally derived by comparative sequence analysis of 5S RNAs (64) but includes extensions of the base-paired regions. All compared 5S RNA

Table 1

Sources of Eukaryotic 5S RNAs Used for the Determination of Common Positions of 5S RNA

Metazoa (somatic tissue only)		**Green Plant and Algae**	
Human KB cell	(5)		
Rat liver	(6)	Triticum vulgare (wheat embryos)	(2)
Xenopus laevis	(7)	Spinacea oleracia (spinach)	(3)
Xenopus borealis	(7)		
Iguana iguana	(8)	Chorella pyrenoidosa (green alga)	(4)
Lingula anatina	(9)	Scenedesmus quardricula (green alga)	(4)
Notophthalmus viridescens (newt)	(10)		
Misgurnus fossilis	(11)	**Fungi**	
Terrapene carolina (turtle)	(12)	Neurospora crassa	(26)
Salmo giardneri (rainbow trout)	(13)	Aspergillus nidulans	(27)
Drosophila melinogaster (invertebrate)	(14)		
Bombyx mori (invertebrate)	(15)	Schizosaccharomyces pombe	(28)
Acyrthasiphon magnoliae (aphid, invertebrate)	(16)	Torulopsis utilis	(29,30)
Caenorhabditis elegans (nematode)	(17)	Saccharomyces cerevisiae	(31)
		Pichia membranaefacience	(32)
Oocyte and Embryonic Tissues		Phycomyces blakesleeanus	(33)
Xenopus borealis (oocyte)	(7)		
Xenopus laevis (oocyte)	(7)	**Prokaryotic Organisms**	
Misgurnus fossilis (loach oocyte)	(11)	Bacillus lichenformis	(45)
Artemia salina (shrimp cysts)	(34)	Bacillus stearothermophilus	(46)
Gallus gallus (chick embryo)	(35)	Bacillus subtilis	(46)
Lytechinus variegatus (sea urchin embryos)	(36)	Bacillus megaterium	(50)
		Bacillus brevis	(58)
		Lactobacillus viridescens	(47)
Protozoa and Other Eukaryotic Protists		Micrococcus lysodeikticus	(56)
Acanthamoeba castellanii (amoeba)	(18)	Clostridium pasteurianum	(49)
Tetrahymena thermophila (ciliate)	(19)	Escherichia coli	(51)
Crithidia fasciculata (flagellate)	(20)	Beneckea harveyi	(48)
Euglena gracilis (photosynthetic flagellate)	(21)	Photobacter	(52)
		Salmonella typhimurium	(55)
Cryptthecodinium cottnii (dino-flagellate)	(22)	Serratia marcescens	(55)
Dictyostelium discoideum (slime mold)	(23)	Pseudomonas fluorescens	(61)
		Proteus vulgaris	(62)
Physarum polycephalum (slime mold)	(24)	Thermus thermophilus	(53)
Chlamydomonas reinhardii (photo-synthetic flagellate)	(25)	Thermus aquaticus	(54)
		Mycoplasma capricolum	(57)
Organelles		Anacystis nidulans	(59)
		Synechococcus lividus	(60)
Chloroplast		**Archaebacteria**	
		Halobacterium cutirubrum	(41)
Lemna minor (duckweed)	(37)	Halococcus morrhuae	(42)
Nicotiana tabacum (tobacco)	(38)	Sulfolobus acidocaldarius	(43)
Spinacia oleracea (spinach)	(39)	Thermoplasma acidophilum	(44)
Mitochondria			
Triticum vulgare (wheat embryo)	(40)		

sequences fit the generalized 5S RNA structural model.

This model is supported by S1 and cobra venom nuclease cleavage sites on eukaryotic 5S RNAs (65,66). Chemical modification of eukaryotic 5S RNAs with kethoxal (67) and psoralen derivatives (68) as well as modification studies with the E. coli 5S RNA (69-75) provide data which are generally consistent with the consensus model. Furthermore, intergenomic comparations of 5S gene structures in Neurospora crassa (26) indicate that the putative helical structures are conserved.

Using the generalized secondary structural model, total free energies ($\Delta G°$) (76) have been calculated for the 5S RNAs from representative organisms (Table 2). The $\Delta G°$ values for three representative E. coli tRNAs are also shown. The range of $\Delta G°$ values indicates that the generalized secondary structural model is stable for 5S RNAs from a wide range of organisms.

There are 23 universal positions that have been identified; these are distributed throughout the molecule (Fig. 1). Seventeen of these positions are in single-stranded regions and include all of the universal adenine residues. Positions that we find totally invariant in all compared 5S RNA sequences (excluding those from oocyte and embryonic RNAs) are shown with an asterisk. There is also a total invariance in chain length in two regions of the molecule, i.e., there are always seven positions between the universal positions Pu_{15} and G_{23} and two positions between G_{44} and Py_{47}. The 23 universal positions were derived by a comparison of the common eukaryotic positions found in the 5S RNAs of prokaryotes, archaebacteria, and organelles (Figs. 2 and 3).

GENERALIZED STRUCTURE OF EUKARYOTIC 5S RNA

Representative structures of eukaryotic 5S ribosomal RNAs are shown in Fig. 4. The positions that are common in the known eukaryotic 5S RNA sequences are indicated in Fig. 2. (In the discussion of the eukaryotic 5S RNAs, all positions referred to are those shown in Figure 2.) These positions were deduced from a comparison of the 5S RNAs from the eukaryotic organisms listed in Table 1 and are representative of those in 90% or more of the eukaryotic 5S RNAs whose sequences were compared. The 5S RNAs that were compared were those from somatic tissues of mature eukaryotic organisms or the major 5S RNA species transcribed in the lower eukaryotes.

In the eukaryotic 5S RNA structure (Fig. 2), the 5' ends are purines while the 3' ends are pyrimidines. More than half of the common positions are purines. All of the conserved adenine residues are in single-stranded positions. Conserved G·C bonds are localized at the ends of several helices

Table 2

Total Free Energies of Representative 5S RNAs

Source of 5S RNA	ΔG°*	Number of A·U and G·C base pairs	Percentage of G·C
E. coli	-52.65	30	80%
L. viridescens	-23.50	27	59%
S. lividus	-41.70	29	69%
Spinach chloroplast	-40.40	29	59%
Spinach cytoplasmic	-50.80	31	71%
Euglena gracilis	-50.70	36	61%
Phycomyces blakesleeanus	-33.75	33	52%
Torulopsis utilis	-22.30	33	58%
Bombyx mori	-53.60	33	67%
Human KB cell	-58.10	31	81%
Sulfolobus acidocaldarius	-98.45	39	74%

E. coli tRNAs	ΔG°	Number of A·U and G·C base pairs	Percentage of G·C
Serine 1	-25.95	23	70%
Alanine 1A	-33.80	18	89%
Glycine 1	-34.80	20	75%

* The range of ΔG° for the 5S RNAs can be attributed to the total number of base pairs, the percent of G·C base pairing and the positioning of G·C base pairs relative to the 5' end of the molecule and to adjacent base pairs (76). ΔG° values also include G·U base pairs.

and flanking the looped-out region of helix III. The number of base residues between common nucleotide positions is constant (Fig. 2).

SOME EXCEPTIONS TO THE GENERALIZED STRUCTURE OF EUKARYOTIC 5S RNAs

Some eukaryotic 5S RNAs deviate from the generalized structure presented in Fig. 2. Some differences may constitute sequencing problems; however, in the structures we considered, we have assumed that this is not the case.

An insertion of one nucleotide between G_{66} (equivalent to universal G_{69})

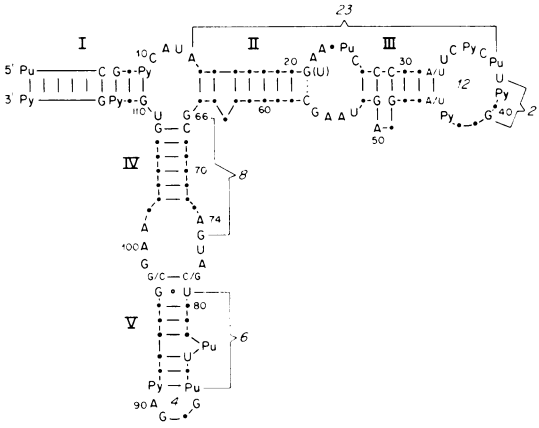


Fig. 2. Generalized eukaryotic 5S RNA structure showing common positions found in more than 90% of compared eukaryotic 5S RNAs. The numbers between these positions represent common chain lengths found in eukaryotic 5S RNAs.

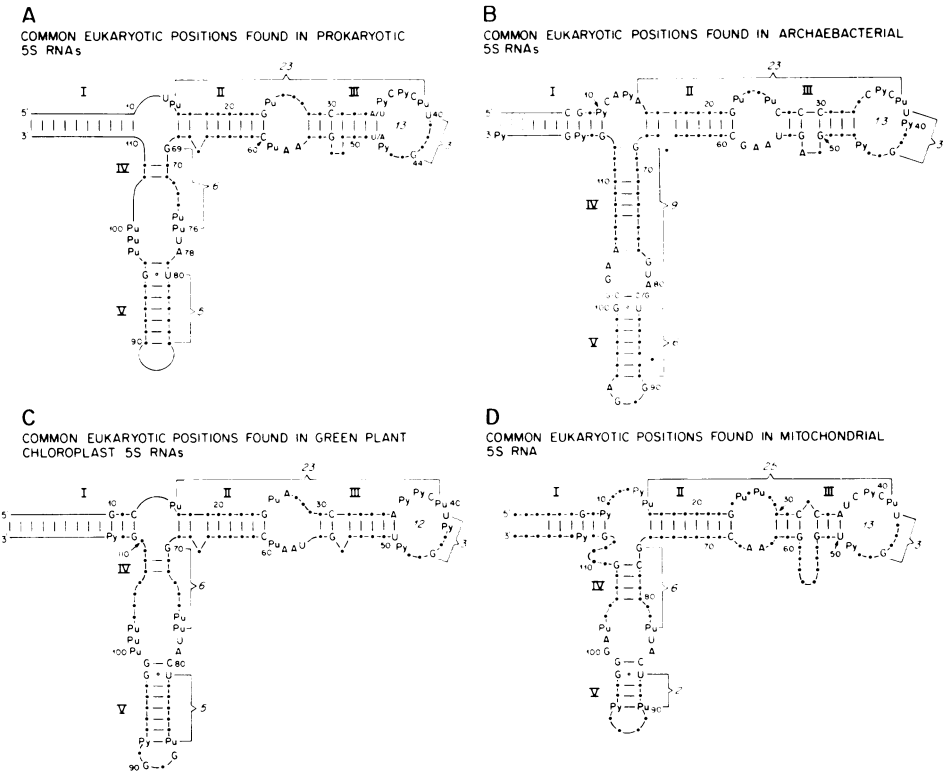


Fig. 3. A comparison of (A) prokaryotic, (B) archaebacterial and (C,D) organelle 5S RNAs showing the common eukaryotic positions found in the RNAs of these groups. The generalized chain lengths between common positions are indicated.

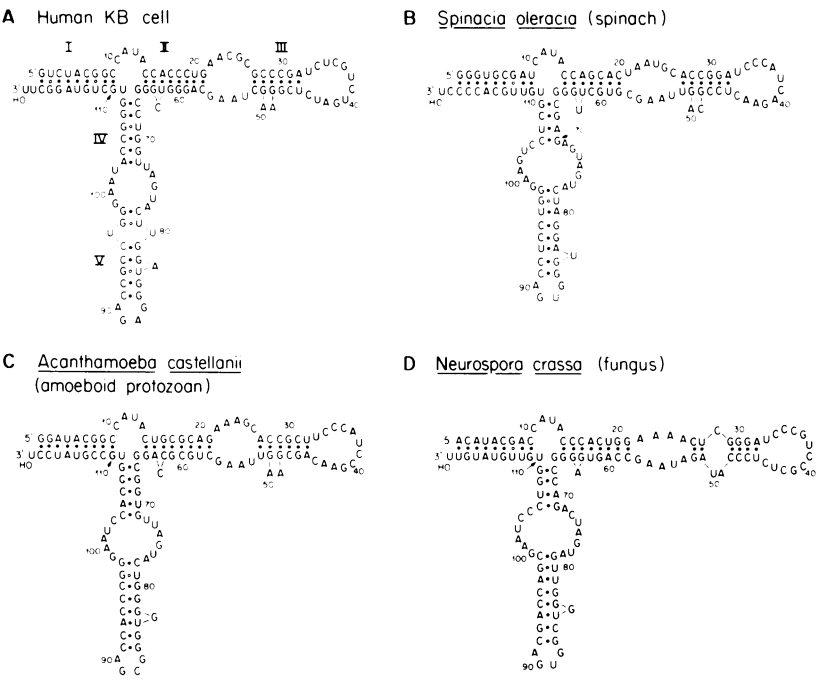


Fig. 4. Representative eukaryotic 5S RNA sequences drawn in the generalized secondary structural model.

and the sequence $A_{74}GUA$ is present in several yeast 5S RNAs (Torulopsis, Saccharomyces and Pichia) (29-32) and appears to be in the single-stranded region between helix IV and $A_{74}GUA$. This is the only known example of divergence from a uniform chain length between positions G_{66} and A_{74} of eukaryotic 5S RNAs.

Euglena 5S RNA (21) has an A substitution at position 33 in place of the common U. A contiguous helix of five base pairs can be formed in helix III and a resultant 10 membered single-stranded loop bounded by helix III.

Lingula 5S RNA (9) terminates at the 3' end with an adenine residue as opposed to the common pyrimidine.

The sequences of two species of 5S RNA from Chlamydomonas have been determined (25). Species I (the major species) has the sequence CAAA in place of the common $C_{10}AUA$ and GCCC in place of the common $A_{22}AXPuC$. Species II has a U in place of C_{10}.

The 5S RNA from the dinoflagellate, Crypthecodinium cohnii (22) diverges

from the generalized structure in two regions of the molecule: the 3' side of helix IV has an insertion of three nucleotides and the common eukaryotic position Pu_{37} is U.

Mispairing in helical regions are found in 5S RNAs of Crithidia (20) (helix II), Phycomyces (33) (helix II), Chlamydomonas (25) (helices I and II) and Lingula (9) (helix I).

The probability of seeing a change in the common positions or the regional lengths of the eukaryotic 5S RNA shown in Fig. 2 is low since a constancy in structure prevails throughout a wide range of organisms. Thus, observed deviations from the common structure in both sequence and chain length as described above is useful in taxonomic classification.

5S RNA REGIONAL HOMOLOGIES FOR RELATED EUKARYOTES

A comparison of eukaryotic 5S RNAs within kingdoms, phyla and divisions reveals nucleotide sequences, secondary structural features and looped-out residues that appear to be common to the 5S RNAs of related organisms. The phylogenetic consistency of these structures strengthens the proposal of the generalized model for the 5S RNA. Regional homologies in eukaryotic 5S RNAs are discussed using the parameters of this generalized structure.

GROUP-SPECIFIC SIGNATURES

The helices shown in Fig. 2 exhibit group-specific specificities (with the exception of helix I). Helix II of most eukaryotic 5S RNAs has a length of eight base-pairs and a looped-out residue at position 63 (Fig. 5). The exception is the structure of plant 5S RNAs where helix II has seven base-pairs; a uracil substitution at position 21 negates the formation of the eighth base pair. The two base-pairs to the left of the looped-out position of helix II as shown in Fig. 5 are consistent within related groups of organisms. Whereas metazoan and plant 5S RNAs have $\overset{C}{G}\overset{C}{G}$ base-pairs, the protozoan 5S RNAs have $\overset{C}{G}\overset{Py}{Pu}$ pairing. The fungal 5S RNAs vary in sequence in this region. The characteristic looped-out position of helix II is C in metazoan and most protozoan 5S RNAs while it is U for the RNAs of plants and algae. The looped-out position of the higher fungal 5S RNAs in helix II is a purine.

Helix III reveals several uniform features in base-pairing and looped-out residues in metazoan, plant and protozoan 5S RNAs (Fig. 6). The two looped-out residues at positions 49 and 50 are flanked by G•C pairs on both sides. The twelve membered loop bounded by helix III is generally closed by A•U base-pairing. No phylogenetic group-specific pattern is evident in the structure of helix III; however, A_{49} found in metazoan and

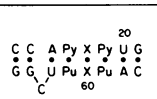
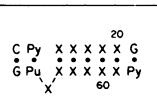
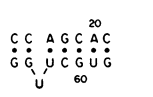
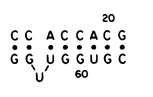
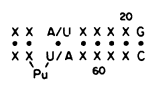
PHYLUM	HELIX II STRUCTURE
METAZOA	
PROTOZOA	
GREEN PLANTS	
GREEN ALGAE	
HIGHER FUNGI	

Fig. 5. Group-specific signatures for helix II. X denotes a variable position.

protozoan 5S RNAs is C in several plant 5S RNAs. The structure for helix III of the higher fungal 5S RNAs is variable due to an insertion around position 49 in some fungal 5S RNAs. This reflects the large divergence of organisms in this group.

The twelve-membered loop bounded by helix III has nucleotide positions specific to phyla (Fig. 7). The purine at position 37 is G in metazoan and fungal 5S RNAs, A in RNAs of the plants and algae and it is either G or A in

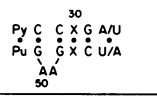
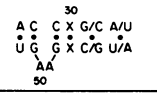
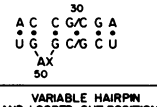
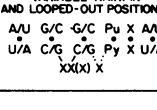
PHYLUM	HELIX III STRUCTURE
METAZOA	
PROTOZOA	
GREEN PLANTS AND ALGAE	
HIGHER FUNGI	

Fig. 6. Group-specific signatures for helix III.

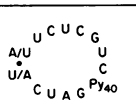
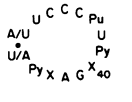
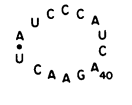
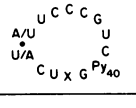
PHYLUM	SEQUENCE
METAZOA	
PROTOZOA	
GREEN PLANTS AND ALGAE	
HIGHER FUNGI	

Fig. 7. Consensus nucleotide sequences of the twelve membered loop bounded by helix III.

the protozoa. The protozoa also exhibit heterogeneity at positions 40 and 43.

The number of base-pairs in helix IV varies (Table 3). At the two extremes of helix length, helix IV of the metazoan 5S RNAs has six base pairs while that of plant and algae RNAs have four. In helix V, secondary structures with looped-out residues that are common in 5S RNAs of related organisms can be formed (Fig. 8). In this helical region, all protozoan and fungal 5S RNAs have a looped-out guanine residue four positions from the common position U_{79} (equivalent to universal position U_{80}). Although extensive nucleotide substitutions in this region are prevalent in the fungal and protozoan 5S RNAs, a uniform secondary structural scheme prevails. Due to a pyrimidine substitution at position 93, the plant 5S RNAs form a different helix which involves a uracil residue looped-out at the fifth position from U_{79}. The algal 5S RNA was drawn to conform to the plant model for helix V. In the model shown for the metazoan 5S RNAs, a looped-out adenine residue can be formed at four positions from U_{79}. This helical model also includes a U-U mispairing. An alternate structure for helix V of metazoan 5S RNAs can also be formed (Fig. 8). The stabilities of both the generalized and alternate structures are approximately the same. We favor the generalized structure for metazoan 5S RNAs since it corresponds with the base-pairing in the other eukaryotic structures for helix V.

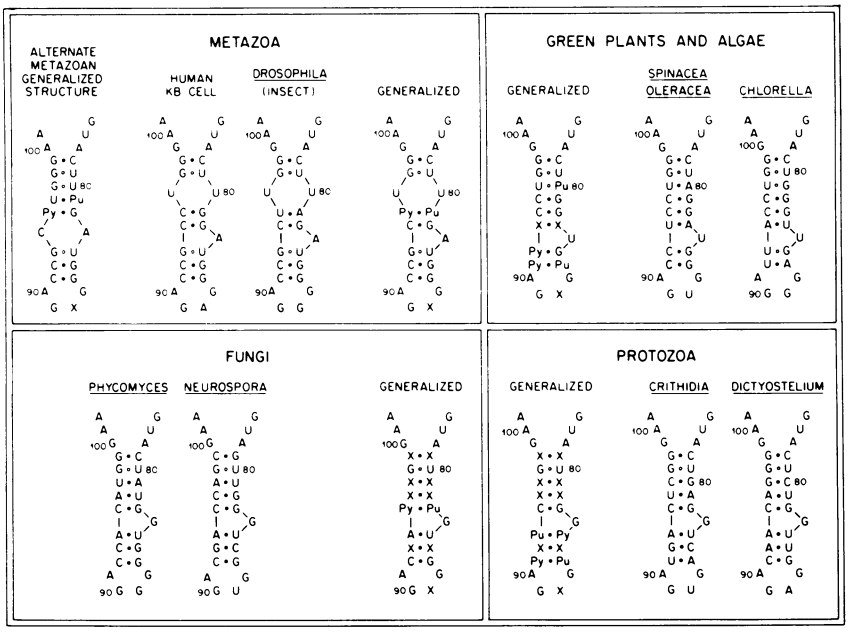


Fig. 8. Generalized helix V structures.

IMPORTANCE IN TAXONOMIC CLASSIFICATION OF GROUP-SPECIFIC SIGNATURES

In addition to overall homology based on sequence comparisons group-specific signatures of 5S RNAs can be used as a parameter to determine or to "fine tune" taxonomic relationships. This is especially useful when the overall homology may by chance indicate a high homology between two unrelated organisms. For example, Euglena gracilis 5S RNA has 73% and 74% sequence homology to Crithidia fasiculata (a trypanosomatid protozoan) and

Table 3

Lengths of Helix IV in Eukaryotic 5S RNAs

Metazoa	6
Protozoa	4-5
Green Plants and Algae	4
Higher Fungi	4-5

human KB cell 5S RNAs, respectively (21). However, both the Euglena and Crithidia 5S RNAs have the sequence $G_{41}AUU$ as well as the protozoan-specific looped-out positions in helices II and V. The human KB cell 5S RNA differs from the RNA of Euglena with respect to these group-specific signatures.

The Euglena 5S RNA is unusual in having a 10 membered single-stranded loop bounded by helix III which the RNA of Crithidia does not have. This unique feature may offer the possibility of fine tuning the classification of Euglena and finding a closer relative than Crithidia.

In addition to the Euglena 5S RNA, the 5S RNA from another photosynthetic protist, Chlamydomonas reinhardii (25) has the characteristic protozoan structure of helix V. On the other hand, the 5S RNAs of Chlorella pyrenoidosa, a single-cell algal organism, and Scenedesmus quadricula (4) a planktonic colony former, have a structure for helix V characteristic of the green plants.

Chlamydomonas reinhardii 5S RNAs have several unusual structural features (25). The deviation from the common eukaryotic sequences in several regions exhibited by those 5S RNAs offers the potential for finding close relatives to this organism among the flagellated protozoa.

Based on comparisons of overall homologies of 5S RNAs, the slime mold Physarum polycephalum 5S RNA has been considered to be more closely related to protozoa and animals than to the fungi (24). Physarum 5S RNA has the characteristic structure of most protozoan 5S RNAs in helix II (i.e., the looped-out residue 63 and the first two base-pairs of the helix).

Several yeast 5S RNAs (namely from Torulopsis, Saccharomyces and Pichia) (29-32) have a G insertion between positions 73 and 74. This is the only known divergence in chain length in this region of the eukaryotic 5S RNA. The 5S RNA from another yeast, Schizosaccharomyces pombe does not have this insertion (28). This divergence suggests that Torulopsis, Saccharomyces and Pichia should be grouped together and placed on a separate branch from Schizosaccharomyces on the yeast phylogenetic tree.

No 5S RNA sequences are available from the red or brown algae. This would be an interesting group of organisms to analyze by 5S RNA phylogeny. The red algae (Rhodophyta) are of special interest in view of a proposed phylogenetic linkage between these organisms and the Ascomycetes (higher fungi) (77).

CHANGES IN COMMON POSITIONS IN 5S RNAs OF OOCYTES, EMBRYONIC TISSUES AND MULTIPLE 5S RNAs FROM AN ORGANISM

The generalized eukaryotic 5S RNA structure (Fig. 2) shows the common

nucleotides derived by a comparison of sequences only from the major species of 5S RNAs of somatic tissues from mature eukaryotic organisms. 5S RNAs of oocytes differ from those of somatic tissues in several positions. In addition, minor species of 5S RNAs differ in sequence from the major species in several of the lower eukaryotes. Some of these nucleotide changes are concentrated in common positions. For example, at position 12 the somatic 5S RNAs from both Xenopus borealis (7) and Misgurnus fossilis (loach) (11) have the common uracil residue. The oocyte 5S RNAs from these two organisms have a cytosine at that position. In comparing the two 5S ribosomal genes and the gene products from Neurospora crassa (26), a substitution (U→C) at position 12 was found in the minor species. The two species of Chlamydomonas reinhardii (25) 5S RNA differ both in positions 10 and 12. In this case, the minor species has the common U at position 12 whereas the major 5S RNA species is unusual in having an A residue at this position.

The sequence of the 5S RNA from chick embryo fibroblast cells has been determined (35), although the sequence of the RNA from mature organisms has not. The 5S RNA from chick embryo fibroblast cells differs in the common eukaryotic positions A_{13}(A→C) and $A_{22}A_{23}$(A_{22}→G, A_{23}→U).

There is a common adenine residue at position 55 in the somatic 5S RNA from Xenopus borealis, but the oocyte 5S RNA has a uracil at that position (Fig. 9). The somatic and oocyte 5S RNAs also differ at the common position 53 in Xenopus and Misgurnus. The sequence of the 5S RNA from shrimp cysts (dormant eggs of Artemia saliva) (34) has a cytosine at position 55 instead of the common adenine. (The sequence of the 5S RNA from the mature organism has not been reported).

Although not a common position, the looped-out residue at position 63 of helix II changes in Misgurnus 5S RNA from somatic and oocyte cells (C→U). The 5S RNA from embryos of sea urchins (which are in a cleavage state during embryonic development) deviates from the phylogenetically consistent cytosine residue of the metazoan 5S RNAs by having an adenine residue in the looped-out position 63 of helix II. Other differences between oocyte and somatic 5S RNAs involve mispairing in helix II. Changes in positions such as U_{79} are found in Xenopus oocyte 5S RNAs. U_{79} is of particular interest since it is also a universal position.

Since the universal and/or common positions are those that are stable in a wide range of eukaryotic organisms, base substitutions found in some of these positions in oocyte 5S RNAs (or 5S genes) suggest an important functional significance. It has been noted that Xenopus oocyte 5S RNA genes

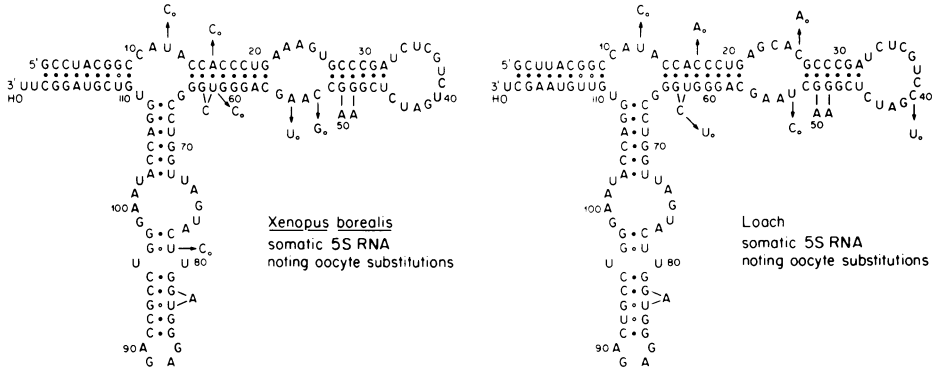


Fig. 9. Somatic and oocyte 5S RNAs.

which have substitutions in the intragenic transcriptional control region at positions 53, 55 (a universal position) and 62 compete only one-fourth as well as somatic 5S genes in an in vitro transcriptional system (80). A summary of changes in common positions of eukaryotic 5S RNAs found in 5S RNAs of oocytes or embryonic tissues is shown in Table 4.

PROKARYOTIC 5S RNAs

The prokaryotic 5S RNAs have primary and secondary structural features close to those of the eukaryotic 5S RNAs (Fig. 3A). Common eukaryotic positions can be found in prokaryotic 5S RNAs as shown in Figure 3A. The alignment of these residues was determined from the location of the putative positions relative to the helical regions (e.g. G_{69} is the first residue on the 3' side adjacent to helix II). As in the determination of the eukaryotic common positions, these positions were noted if they could be found in 90% or more of the prokaryotic 5S RNAs that were compared.

Three regions of the 5S RNA reveal a demarkation between the prokaryotic and eukaryotic molecules that involve differences in chain lengths:

1. All prokaryotic 5S RNAs have three positions between the universal nucleotides U_{40} and G_{44} and all eukaryotic 5S RNAs have two. Thus, the generalized structure of the 5S RNA of prokaryotes has a thirteen membered loop bounded by helix III while the structure for the 5S RNA of eukaryotes has twelve.
2. Prokaryotic 5S RNAs have six positions between the universal positions G_{69} and $Pu_{76}UA$ and most

Table 4
Nucleotide Changes in Common Eukaryotic or Conserved Metazoan Positions found in 5S RNAs of Oocytes or Embryonic Tissues

source of 5S RNA		loach somatic	loach oocyte	X. laevis somatic	X. laevis oocyte	X. borealis somatic	X. borealis oocyte	sea° urchin embryo	shrimp° cysts	chick° embryo
expect	position									
U	12*	U	C			U	C			
A	13*									C
A	16+					A	C			
Py	17+	C	A							
A	22*									G
A	23*									U
C	26*	C	A					U	U	
conserved b.p. with 47	30			C	U					
G	31+								A	
Py	40+	C	U						A	
C	46+								U	
conserved b.p. with 30	47			G	A					
Py	53*	U	C	C	G	C	G			
A	55*			A	U	A	U		C	
G	56*			G	A				A	
U	62					U	C			
C	63+	C	U					A		
U	79*			U	C	U	C			
G	93+									U

* Common Eukaryotic positions in somatic 5S RNAs
\+ Conserved positions in Metazoan somatic 5S RNAs
° 5S RNA from mature organism has not been sequenced

eukaryotic 5S RNAs have eight. (Some yeast 5S RNAs have nine.)

3. Between the universal positions U_{80} and the last nucleotide on the 5' side of helix V (see Fig. 3A) prokaryotic 5S RNAs have a chain length of five (or less) whereas the eukaryotic 5S RNAs uniformly have six positions. In addition, between the universal positions U_{80} and G_{96} prokaryotic 5S RNAs have sixteen (or less) residues and the eukaryotic 5S RNAs have seventeen.

Other major differences between prokaryotic and eukaryotic 5S RNAs include:

1. The relative length of helix IV and
2. The lack of a uniform length and sequence in the single-stranded regions analogous to the eukaryotic sequences $C_{10}AUA$ and $G_{87}XGA$. On the other hand, the prokaryotes have the conserved sequence $PyCG_{44}AAC$ in their 5S RNAs, a sequence that is variable in eukaryotic RNAs. A summary of differences between prokaryotic and eukaryotic 5S RNAs is shown in Table 5.

PROKARYOTIC NATURE OF ORGANELLE 5S RIBOSOMAL RNAs

Chloroplast 5S RNA

Known nucleotide sequences of chloroplast 5S RNAs from the green plants appear to be nearly identical (37-39). The chloroplast 5S RNAs have structural properties common to the prokaryotic 5S RNAs and different from those of the eukaryotes (Fig. 3C). The chloroplast 5S RNAs have the following prokaryotic features:

1. Three nucleotide positions between the universal positions U_{40} and G_{44}.
2. Six positions between residues G_{69} and $Pu_{76}UA$.
3. Five positions between the universal U_{80} and the last nucleotide on the 5' side of helix V.
4. The sequence $PyCG_{44}AAC$.

In addition, long stretches of nucleotide sequences are homologous to the 5S RNA of the cyanobacterium Anacystis nidulans (37, 59). However, there appears to be a striking homology between the 5S RNAs of the plant

Table 5

Major Differences Between Prokaryotic and Eukaryotic 5S RNAs

	Prokaryotic 5S RNA	Eukaryotic 5S RNA
Chain Length Between Universal Positions		
U_{40} and G_{44}	3	2
G_{69} and Pu_{76} UA	6	8 (9)
$G \cdot U_{80}$ base pair	$\leq$16	17
Nucleotide Sequence		
First Single-Stranded region near 5' side	Variable in Length and Sequence	C_{10} AUA
Sequence surrounding universal G_{44}	$PyCG_{44}AAC$	$PyXG_{40}XXPy$
Sequences in loop between helices IV and V	$PuPu_{76}UA$	$AGUA_{76}$
	$GXPuPuPu_{100}$	$GG/cGA_{100}A$
Length of Helix IV	2	4-6
Looped-out Position in Helix V	absent	present

chloroplasts and that of the cyanobacterium Synechococcus lividus (60). Unique structural features of the green plant chloroplast 5S RNAs are a deletion between positions 34 and Pu_{39} (E. coli numbering system) and base substitutions at positions 31, 52 and 53 which result in an extended helix beyond four contiguous base pairs in helix III proximal to the hairpin loop bounded by the helix. These properties of helix III and the single-stranded loop bounded by it are not found in prokaryotic, eukaryotic, archaebacterial or mitochondrial 5S RNAs with the exception of the 5S RNA of the cyanobacterium Synechococcus lividus.

The green plant chloroplast 5S RNAs do not have eukaryotic features such as the sequence $C_{10}AUA$, a looped-out residue in helix V and the single-stranded sequences $A_{74}GUA$ and $GA_{100}A$ (Fig. 3C).

Mitochondrial 5S RNA

The complete nucleotide sequence of the wheat mitochondrial 5S RNA has been determined and a secondary structural model has been proposed for this

RNA (40). This model conforms with the generalized structure presented in Fig. 1. Although the wheat mitochondrial 5S RNA has several unique structural features, it has most of the properties of prokaryotic 5S RNAs (Fig. 3D):

1. Three positions between the universal positions U_{40} and G_{44}.
2. Six positions between the universal G_{69} and $PuUA_{78}$.
3. The conserved prokaryotic sequence $PyCG_{44}AAC$.

Helix V of the mitochondrial 5S RNA is truncated [as is helix V of the Mycoplasma capricolum 5S RNA (57)]. Common features of eukaryotic 5S RNAs such as the sequences $C_{10}AUA$, $A_{74}GUA$, and $GA_{100}A$ are not present. However, this 5S RNA does have four base-pairs in helix IV in the model proposed by Spencer et al (40).

Archaebacterial 5S RNAs

The archaebacteria occupy a unique phylogenetic nitch (78). Several archaebacterial 5S RNA sequences have been determined (41-44). These 5S RNAs generally have lengths between universal positions that are eukaryotic-like with the exception of the chain length between U_{40} and G_{44} which is prokaryotic-like. A schematic of the archaebacterial 5S RNA is shown in Figure 3B high-lighting the common eukaryotic positions found in these RNAs.

Sulfolobus acidocaldarium 5S RNA has ten contiguous base-pairs in helix IV and like the Thermoplasma acidophilum 5S RNA has no looped-out position in helix II. In Sulfolobus 5S RNA a contiguous helix comprising helices IV and V with only one mismatched base-pair has been proposed by Stahl et al (43). Sulfolobus 5S RNA has an anomolous structure in the region between helices IV and V in that it does not have the universal positions AU_{78} or Pu_{100}.

SUMMARY OF UNIVERSAL PROPERTIES OF THE 5S RNA

The universal 5S ribosomal RNA structure (Fig. 1) reveals the following common properties:

1. Universal nucleotide positions (23 positions) are found throughout the molecule.
2. No insertions or deletions between the universal positions Pu_{15}-G_{23} and G_{44}-Py_{47} are found in any compared sequences of 5S RNAs.
3. A consensus secondary structure having five double helical regions can be formed for all 5S RNAs.

4. All conserved adenine residues are in single-stranded regions in the generalized secondary structural model.

Possible functions of universal positions may involve the participation in multiple processes such as transcriptional control, 5S processing and/or translation. Such multiple involvement would provide a strong selective pressure against a molecule bearing mutations in these positions that would still be functional.

Oocyte 5S RNAs have base substitutions in several common eukaryotic positions (Fig. 9); some of these positions may play a role in transcriptional control. A control region near the center of the 5S genes of Xenopus oocytes (positions 50-83) has been identified for accurate initiation of 5S RNA synthesis (79,80). This control region has nucleotide substitutions in several of the common eukaryotic positions. In addition, the oocyte 5S RNA can also inhibit its own synthesis in an in vitro system, presumably by binding to a transcription factor (81).

The prevalence of conserved adenine residues in single-stranded regions of the 5S RNA may be associated with protein contact points. Although it is not a universal position, the looped-out adenine residue of helix II of the E. coli 5S RNA has been suggested as a recognition site for the binding of ribosomal protein L18 (82). In addition to the conservation of chain length between certain universal positions, the location of phylogenetically conserved looped-out positions in helices and a conserved secondary structure may be essential for protein interaction.

ACKNOWLEDGMENTS

We thank Daniel Bogenhagen, Ann Jacobson, John Vournakis, William Curtiss, and Robert A. Zimmermann for reading the manuscript and for suggestions. We also want to thank Karen Henrickson for the elaborate drawings. Supported by grant National Institutes of Health GM-20052.

Note added in proof. After this manuscript was submitted for publication, a paper by De Wachter et al., [Biochimie (1982) 64:311-329] was brought to our attention. The consensus generalized secondary structural model we use here is similar to the model used by De Wachter et al.

References

1. Erdmann, V.A. (1982) Nucleic Acids Res. 10:r93-r116.
2. MacKay, R.M., Spencer, D.F., Doolittle, W.F. and Gray, W.M. (1980) Eur. J. Biochem. 112:561-576.
3. Delihas, N., Andersen, J., Sprouse, H.M., Kashdan, M. and Dudock, B. (1981) J. Biol. Chem. 256:7515-7517.
4. Luehrsen, K.R. and Fox, G.E. (1981) Proc. Natl. Acad. Sci. USA 78:2150-2155.
5. Forget, B.G. and Weissman, S.M. (1969) J. Mol. Biol. 244:3148-3165.
6. Aoyama, K., Kidaka, S., Tanaka, T., Ishikawa, K. (1982) J. Biochem. 91: 363-367.
7. Ford, P.J. and Brown, R.D. (1976) Cell 8:485-493; and Bogenhagen, D.F. revise sequences (personal communication).
8. Roy, K.L. and Enns, L. (1976) J. Biol. Chem. 251:6352-6354.
9. Komiya, H., Shimizu, N., Kawakami, M. and Takemura, S. (1980) J. Biochem. 88:1449-1456.
10. Kay, B.K. and Gall, J.G. (1981) Nucleic Acids Res. 9:6457-6469.
11. Mashkova, T.D., Sorenkova, T.I., Mazo, A.M., Avdonina, T.A., Timofeyeva, M.Y. and Kisselev, L.L. (1981) Nucleic Acids Res. 9:2141-2151.
12. Roy, K.L. (1977) FEBS Lett. 80:266-270.
13. Roy, K.L. (1978) Can. J. Biochem. 56:60-65.
14. Benhamou, J. and Jordan, B.R. (1976) FEBS Lett. 62:146-149.
15. Komiya, H., Kawakami, M. and Takemura, S. (1981) J. Biochem. 89:717-722.
16. Kawata, Y. and Ishikawa, H. (1982) Nucleic Acids Res. 10:1833-1840.
17. Butler, M., Wall, S.M., Luehrsen, K.R., Fox, G.E., Hecht, R.M. (1981) J. Mol. Evol. 18:18-23.
18. MacKay, R.M. and Doolittle, W.F. (1981) Nucleic Acids Res. 9:3321-3334.
19. Luehrsen, K.R., Fox, G.E. and Woese, C.R. (1980) Current Microbiology 4:123-126.
20. MacKay, R.M., Gray, M.W. and Doolittle, W.F. (1980) Nucleic Acids Res. 8:4911-4917.
21. Delihas, N., Andersen, J., Andresini, W., Kaufman, L. and Lyman, H. (1981) Nucleic Acids Res. 9:6627-6633.
22. Hinnebusch, A.G., Klotz, L.C., Blanken, R.I. and Loeblich, A.R. (1981) J. Mol. Evol. 17:334-347.
23. Hori, H., Osawa, S., Iwabuchi, M. (1980) Nucleic Acids Res. 8:5535-5539.
24. Komiya, H. and Takemura, S. (1981) J. Biochem. 9:1577-1581.
25. Darlix, J.-L. and Rochaix, J.-D. (1981) Nucleic Acids Res. 9:1291-1299.
26. Selker, E.U., Yanofsky, C., Driftmier, K., Metzenberg, R.L. Alzner-DeWeerd, B., RajBhandary, U.L. (1981) Cell 24:819-828.
27. Piechulla, B., Hahn, U., McLaughlin, L.W. and Küntzel, H. (1981) Nucleic Acids Res. 9:1445-1450.
28. Mao, J., Appel, B., Schaack, J., Sharp, S., Yamada, H., Söll, D. (1982) Nucleic Acids Res. 10:487-500.
29. Nishikawa, K. and Takemura, S. (1974) J. Biochem. 76:935-947.
30. Tabata, S. (1980) Eur. J. Biochem. 110:107-114.
31. Miyazaki, M. (1974) J. Biochem. 75:1407-1410.
32. Miyazaki, M. (1977) Nucleic Acids Res. Supplement No. 3:153-156.
33. Andersen, J., Andresini, W., Delihas, N. (1982) J. Biol. Chem., 257:9114-9118.
34. Diels, L., DeBaere, R., Vandenberghe, A., DeWachter, R. (1981) Nucleic Acids Res. 9:5741-5744.
35. Brownlee, G.G. and Cartwright, E.M. (1975) Nucleic Acids Res. 2:2279-2288.

36. Lu, A.L., Steege, D.A. and Stafford, D.W. (1980) Nucleic Acids Res. 8:1839-1853.
37. Dyer, T.A. and Bowman, C.M. (1979) Biochem. J. 183:595-604.
38. Takaiwa, F. and Sugiura, M. (1981) Mol. Gen. Genet. 182:385-389.
39. Delihas, N., Andersen, J., Sprouse, H.M. and Dudock, B. (1981) Nucleic Acids Res. 9:2801-2805.
40. Spencer, D.F., Bonen, L. and Gray, M.W. (1981) Biochem. 20:4022-4029.
41. Nazar, R.N. and Matheson, A.T. (1978) J. Biol. Chem. 253:5464-5469.
42. Luehrsen, K.R., Nicholson, D.E., Eubanks, D.C. and Fox, G.E. (1981) Nature 293:755-756.
43. Stahl, D.A., Luehrsen, K.R., Woese, C.R. and Pace, N.R. (1981) Nucleic Acids Res. 9:6129-6137.
44. Luehrsen, K.R., Fox, G.E., Kilpatrick, M.W., Walker, R.T., Domdey, H., Krupp, G. and Gross, H.J. (1981) Nucleic Acids Res. 9:965-970.
45. Raué, H.A., Stoff, T.J., and Planta, R.J. (1975) Eur. J. Biochem. 59:35-42.
46. Marotta, C.A., Varricchio, F., Smith I., Weissman, S.M., Sogin, M.L. and Pace, N.R. (1976) J. Biol. Chem. 251:3122-3127.
47. Alexander, L.J. and Stewart, T.S. (1980) Nucleic Acids Res. 8:979-987.
48. Luehrsen, K.R. and Fox, G.E. (1981) J. Mol. Evol. 17:52-55.
49. Pribula, C.D., Fox, G.E. and Woese, C.R. (1976) FEBS Lett. 64:350-352.
50. Pribula, C.D., Fox, G.E. and Woese, C.R. (1974) FEBS Lett. 44:322-323.
51. Brownlee, G.G. (1972) Determination of Sequences in RNA, Laboratory Techniques in Biochemistry and Molecular Biology. (North-Holland Pub. Co., London) Vol. 3 T.S. Work and E. Work, ed. p. 102.
52. Woese, C.R., Pribula, C.D., Fox, G.E. and Zablen, L.B. (1975) J. Mol. Evol. 5:35-46.
53. Kumagai, I., Digweed, M., Erdmann, V.A., Watanabe, K. and Oshima, T. (1981) Nucleic Acids Res. 9:5159-5162.
54. Nazar, R.N. and Matheson, A.T. (1977) J. Biol. Chem. 252:4256-4261.
55. Sogin, S.J., Sogin, M.L. and Woese, C.R. (1972) J. Mol. Evol. 1:173.
56. Hori, H., Osawa, S., Murao, K. and Ishikura, H. (1980) Nucleic Acids Res. 8:5423-5426.
57. Hori, H., Sawada, M. Osawa, S., Murao, K. and Ishikura, H. (1981) Nucleic Acids Res. 9:5407-5410.
58. Woese, C.R., Luehrsen, K.R., Pribula, C.D. and Fox, G.E. (1976) J. Mol. Evol. 8:143-153.
59. Corry, M.J., Payne, P.I. and Dyer, T.A. (1974) FEBS Lett. 46:63-66.
60. Delihas, N. Andresini, W., Andersen, J. and Berns, D. (1982) J. Mol. Biol., in press.
61. DuBuy, B. and Weissman, S.M. (1971) J. Biol. Chem. 246:747-761.
62. Fischel, J.-L. and Ebel, J.-P. (1975) Biochimie 57:899-904.
63. Studnicka, G.M., Eiserling, F.A. and Lake, J.A. (1981) Nucleic Acids Res. 9:1885-1904.
64. Fox, G.E. and Woese, C.R. (1975) Nature 256:505-507.
65. Nishikawa, K. and Takemura, S. (1977) J. Biochem. 81:995-1003.
66. Troutt, A., Savin, T.J., Curtiss, W.C., Celentano, J. and Vournakis, J.N. (1982) Nucleic Acids Res. 10:653-664.
67. Nishikawa, K. and Takemura, S. (1978) J. Biochem. 84:259-266.
68. Thompson, J.F., Wegnez, M.R. and Hearst, J.E. (1981) J. Mol. Biol. 147:417-436.
69. Bellemare, G., Jordan, B.R., Rocca-Serra, J. and Monier, R. (1972) Biochimie 54:1453-1466.
70. Delihas, N., Dunn, J.J. and Erdmann, V.A. (1975) FEBS Lett. 58:76-80.
71. Larrinua, I. and Delihas, N. (1979) FEBS Lett. 108:181-184.
72. Noller, H.F. and Herr, W. (1974) J. Mol. Biol. 90:181-184.
73. Noller, H.F. and Garrett, R.A. (1979) J. Mol. Biol. 132: 621-636.

74. Rabin, D. and Crothers, D.M. (1979) Nucleic Acids Res. 7:689-703.
75. Digweed, M. and Erdmann, V.A. (1981) Nucleic Acids Res. 9:3187-3198.
76. Tinoco, jun, I., Uhlenbeck, O.C. and Levine, M.D. (1971) Nature 230: 362-367, and Tinoco, jun, I. (1982) updated parameters (personal communication).
77. Kohlmeyer, J. (1975) BioScience 25:86-93.
78. Woese, C.R., Margrum, L.J. and Fox, G.E. (1978) J. Molec. Evol. 11: 245-252.
79. Bogenhagen, D.F., Sakonju, S. and Brown, D.D. (1980) Cell 19:27-35.
80. Wormington, W.M., Bogenhagen, D.F., Jordan, E. and Brown D.D. (1981) Cell, 24:809-817.
81. Pelham, H.R.B. and Brown, D.D. (1980) Proc. Natl. Acad. Sci. USA 77: 4170-7174.
82. Peattie, D. Noller, H.F., Douthwaite, S. and Garrett, R.A. (1981) Proc. Natl. Acad. Sci. USA 78:1331-1336.