J. Mol. Evol. 3, 151—159 (1974)

The Nucleotide Sequence of Chicken 5S Ribosomal RNA

Norman R. Pace, Thomas A. Walker, and Bernadette Pace*

National Jewish Hospital and Research Center and Department of Biophysics and Genetics, University of Colorado Medical Center, Denver, Colorado 80206

Raymond L. Erikson

Department of Pathology, University of Colorado, Medical Center, Denver, Colorado 80220

Received January 9, 1974

Summary. The nucleotide sequence of the 5S ribosomal RNA of somatic cells of the chicken (*Gallus gallus*) differs from that of other vertebrates thus far examined. Besides base substitutions, alterations include two nucleotide deletions and two additions.

Key words: Nucleotide Sequence — 5S *r*RNA — Eukaryote — Phylogeny.

Introduction

The nucleotide sequences have been determined for 5S ribosomal RNA from tissue culture lines derived from several representatives of the order *Mammalia*, including human (Forget and Weissman, 1967), mouse (Williamson and Brownlee, 1969) and a marsupial (Averner and Pace, 1972). All of these 5S *r*RNA nucleotide sequences are identical and it is therefore likely that the 5S *r*RNA nucleotide sequence is invariant throughout the *Mammalia*. In the interest of determining the extent of evolutionary drift in 5S *r*RNA structure throughout the vertebrates, we have examined that of *Xenopus*, an amphibian, and chicken (*Gallus gallus*). Our results with 5S *r*RNA from somatic cells of *Xenopus* are in complete agreement with those published recently by Brownlee and his colleagues (Brownlee *et al.*, 1972) and Wegnez *et al.* (1972). In this communication we consider the 5S *r*RNA isolated from primary fibroblast cultures from chicken embryos. Chicken 5S *r*RNA is structurally distinct from that of mammals or *Xenopus*. Besides base substitutions, alterations include two nucleotide deletions and two additions. Evolutionary drift in 5S *r*RNA is compared to that of the cytochromes *c*.

* Present address: Department of Biology, Regis College, Denver, Colorado 80221.

Materials and Methods

Isotopic Labeling and Purification of RNA

12-day old embryos were aseptically removed from eggs (obtained from Kimber Farm), minced, and dispersed by incubation with 0.25 per cent trypsin in medium TD, which consisted of 0.14 M NaCl, 0.005 M KCl, 0.001 M NaH_2PO_4, and 0.05 M Tris. HCl, pH 7.5. Cells released by trypsin were decanted, recovered by centrifugation and resuspended in minimum essential medium (Earle's base, Temin-modified, obtained from Schwarz) containing 5% calf serum. 10^7 cells in 10 ml of medium were distributed to each of several 100 mm plastic dishes and grown under 5% CO_2 at 39°. After 24 h of growth the medium was exchanged for one similar to that above but containing only 2.5×10^{-5} M phosphate and including 50 μc of [^{32}P] orthophosphate per ml as well as calf serum to a final concentration of 10%. After further growth to confluence, cells were harvested from plates by scraping, washed once with phosphate-buffered 0.8% NaCl solution (Averner and Pace, 1972) and resuspended in 0.1 M NaCl, 0.001 M sodium EDTA, 0.05 M Tris. HCl, pH 7.5, or 0.01 M Tris. HCl, pH 7.2, 0.001 M $MgCl_2$. Bulk cytoplasmic RNA was purified by two extractions with phenol. Water-saturated phenol (two volumes) was added to cell suspensions in the magnesium-containing buffer and then, after mixing, 0.2 M sodium EDTA was added to the suspension to a final concentration of 0.01 M. We have observed that if cells are treated with chelating agents prior to phenol treatment, then the yield of 5′-terminal di- and triphosphorylated 5S *r*RNA is markedly lower than if EDTA is added subsequently to phenol. We presume that addition of EDTA to cell suspensions leads to disruption of cytoplasmic ribosomes and/or release of 5S *r*RNA, which then would be available to phosphatase activities. Following phenol extraction, bulk RNA was recovered by precipitation from 75% ethanol and the high molecular weight RNA species were separated from 5S *r*RNA and *t*RNA by precipitation with 1 M sodium chloride (Zubay, 1962). 5S *r*RNA then was purified by electrophoresis through, and elution from, 8% polyacrylamide gels as detailed previously (Pace and Pace, 1971).

Nuclease Digestions and Oligonucletide Analysis

The purified 5S *r*RNA was completely digested by either T1 ribonuclease (Sankyo) or pancreatic ribonuclease (Sigma). The resulting oligonucleotides were resolved by two-dimensional electrophoresis on cellulose acetate at pH 3.5; this was followed by transfer to, and electrophoresis on, DEAE paper in 7% formic acid, all as detailed by Sanger *et al.* (1965), except that the pH 3.5 buffer consisted of 0.003 M disodium EDTA, 7 M urea, and sufficient glacial acetic acid to adjust the buffer to pH 3.5 at 25°. After radioautography to locate the digestion products, individual oligonucleotides were characterized by secondary digestion with either T1 ribonuclease or pancreatic ribonuclease and, in some cases, venom phosphodiesterase (Schwarz/Mann) or U_2 ribonuclease, all as detailed previously (Averner and Pace, 1972).

Results

The distribution of chicken 5S *r*RNA and *t*RNA in 8% polyacrylamide gels is shown in Fig. 1. After elution and pooling of the indicated gel slices, the purified 5S *r*RNA was digested by either T1 or pancreatic ribonuclease and digestion products were resolved by two-dimensional electrophoresis. Radioautograms of oligonucleotide fingerprints are presented in Fig. 2. These fingerprints of chicken 5S *r*RNA show considerable similarity to equivalent analyses of mammalian 5S *r*RNA, but differences are apparent.

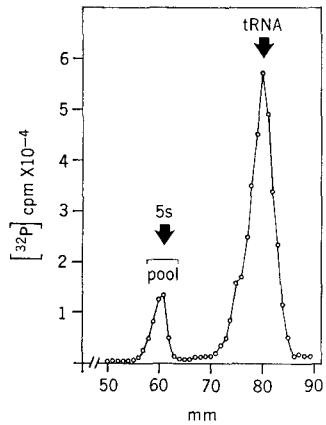

Fig. 1. Preparative polyacrylamide gel electrophoresis of chicken 5S *r*RNA. Salt-soluble, ^{32}P-labeled RNA, purified from primary tissue cultures as described in the text, was resolved by electrophoresis through 8% polyacrylamide gels, and 1 mm slices of the gel were monitored in a scintillation spectrometer for radioactive content by Cerenkov emission

The molar distributions of isolated oligonucleotides and their sequences or compositions, as determined by secondary digestions, are listed in Table 1. Spots in the radioautograms not numbered or mentioned in Table 1 are inconstant features and presumably result from contaminating RNA, incomplete digestion or oversplitting of oligonucleotides.

Certain oligonucleotides, identified in Table 1 with an asterisk, are found in chicken but not mammalian 5S *r*RNA. The structure of these oligonucleotides fortunately is such that the complete nucleotide sequence of chicken 5S *r*RNA may be assessed, if we assume that the molecule is homologous with mammalian 5S *r*RNA, and that there are no sequence displacements. Although sequence displacements in the two 5S *r*RNA molecules may occur, they must be highly specific in order to escape detection in catalogs of both T1 and pancreatic ribonuclease-generated oligonucleotides; these displacements are therefore extremely improbable. The differences between chicken and mammalian 5S *r*RNA are identified as follows:

Position 2. Chicken T1 oligonucleotide T14 differs from, but is homologous to, T1 oligonucleotide No. 56 from mammalian cells (Forget and Weissman, 1967), whose sequence is –U–C–U–A–C–G–. Venom phosphodiesterase digestion of alkaline phosphatase-treated chicken T14 yields two moles of C and one of U; the sequence of T14 must be –C–C–U–A–C–G. This also is evident in oligonucleotide A21, which is pG–C–, the 5′ terminus of the molecule.

Positions 13 and 17–20. Chicken oligonucleotide T20 differs significantly from its mammalian counterpart (T55 of Forget and Weissman, 1962) in

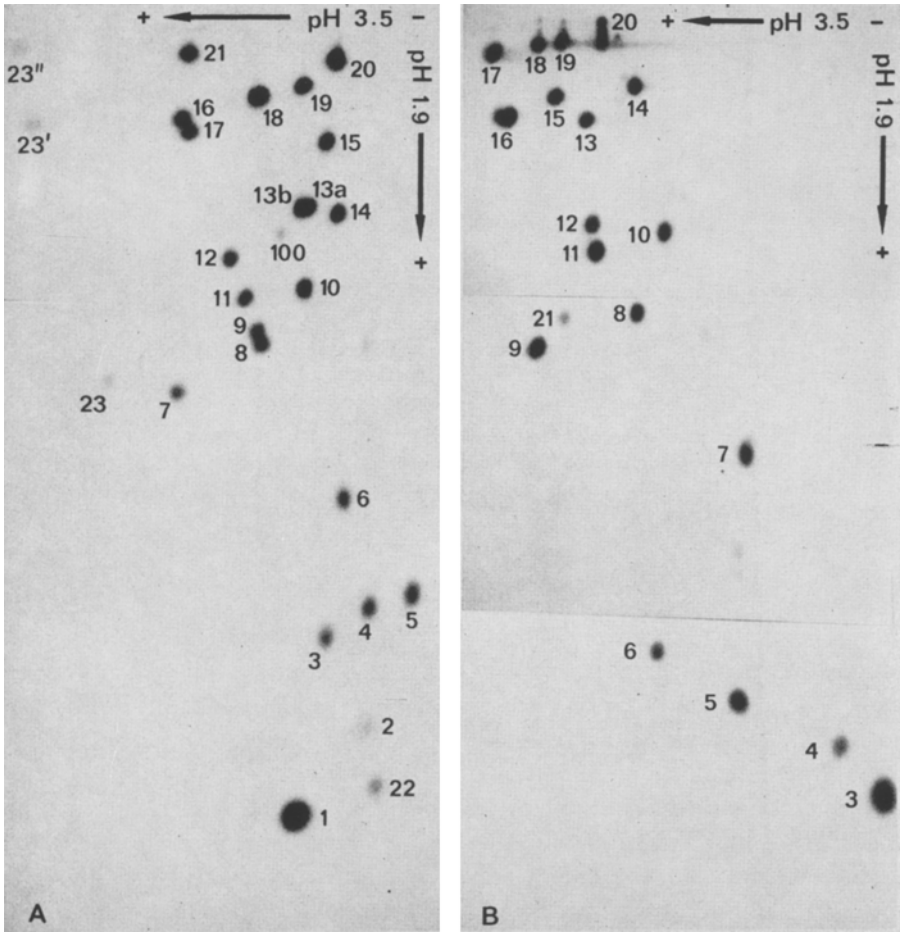

Fig. 2A and B. Two-dimensional electrophoresis of A T1, and B pancreatic ribonuclease digestion products of chicken 5S *r*RNA. Purified chicken 5S *r*RNA was digested with T1 or pancreatic ribonuclease, then digestion products were resolved first by electrophoresis at pH 3.5 on cellulose acetate followed by transfer to DEAE paper and electrophoresis in 7% formic acid, all as detailed by Sanger *et al.* (1965). After drying, the DEAE sheets were exposed to X-ray film for radioautography

possessing two additional C residues and lacking one A residue. U_2 ribonuclease digestion of T20 yields C–C–A, U–C–C–C–A and C–C–C–C–U–G. The sequence of the latter oligonucleotide was determined from the fragments of U_2 ribonuclease "overdigestion" (Sogin, 1972); these include C–U–G and U–G, but not C–G. The minimal deviation from the mammalian or *Xenopus* 5S *r*RNA sequences which is compatible with these yields and the pancreatic ribonuclease digestion products of oligonucleotide T20 (Table 1) substitutes a C residue at position 13 and inserts one C residue at

Table 1. Digestion products of chicken 5S *r*RNA

A. T1 RNase digestion products				B. Pancreatic RNase digestion products			
1 Spot No.	2 Sequence or Compositions	3 Exper. moles (No.)	4 Theor. moles (No.)	1 Spot No.	2 Sequence or Compositions	3 Exper. moles (No.)	4 Theor. moles (No.)
T1	G	13.2	15	A1	C	not determined	19
T2	CG	0.3	—	A2	U	12.9	14–15
T3	AG	1.1	1	A3	AC	4.4	4
T4	CAG	0.8	1	A4*	AAC	1.2	1
T5	CCCG	0.9	1	A5	GC	1.9	2
T6	AAG	0.8	1	A6	AU	1.0	1
T7	UG	1.4	1	A7	AAGC	0.8	1
T8	UCG	1.0	1	A8	GGC	0.9	1
T9	CUG	1.1	1	A9	GU	2.5	3
T10	C_2, U, G	1.1	1	A10	AGGC	0.9	1
T11	UAG	1.1	1	A11	GAU	2.2	2
T12	AUG	1.1	1	A12	AGU	1.3	1
T13a	C, U, AAG	0.9	1	A13	GGGC	1.1	1
T13b*	UAACG	0.9	1	A14	G_2, AAG, C	0.9	1
T14*	C, C, U, AC, G	0.9	1	A15	GGAU	1.2	1
T15	AAU, AC, C, G	0.9	1	A16	GGU	2.2	1
T16	UUAG	1.3	1	A17	GGGU	0.9	1
T17	U_2, C, G	1.0	1	A18	AG, G_2, U	1.1	1
T18	AU, C_2, U, G	2.2	2	A19	GGGAAU	1.1	1
T19*	AC, C_3, U_2, G	0.9	1	A20	G_3, AG, AC	0.9	1
T20*	Cn_{7-8}, AU, AC, U, G	1.0	1	A21*	*p*GC	0.7	1
T21	U_3, AC, G	1.1	1				
T22	CUU_{OH}	0.6	1				
T100	$CUUU_{OH}$	0.3					
T23	*p*G	0.8	1				
T23′	*pp*G						
T23″	*ppp*G						

Catalogs of A. T1, and B. pancreatic ribonuclease digestion products of chicken 5S *r*RNA. T1 and pancreatic ribonuclease digestion products were resolved by two-dimensional electrophoresis as described for Fig. 2, located by radioautography, excised from the DEAE paper, and their sequence or composition was determined as detailed in the text. In Column 1 the oligonucleotides are numbered. An asterisk (*) indicates that the oligonucleotide is not found among the digestion products of mammalian 5S *r*RNA; Column 2 gives the sequence or, where the sequence cannot be inferred from secondary digestions, the composition of the oligonucleotide; Column 3 gives the observed molar yield of the oligonucleotides; Column 4 gives the theoretical yield from the inferred structure of the molecule.

one of the positions 17–20. All of these latter positions are occupied by C residues and so the exact position of the insertion is impossible to discern. The sequence of oligonucleotide T20 therefore is C–C–A–U–C–C–C–A–C–C–C–C–U–G.

Position 21. Chicken oligonucleotide T13b is not present in the mammalian or *Xenopus* 5S *r*RNA molecule. Pancreatic ribonuclease digestion of T13b yields U, A–A–C and G, and U_2 ribonuclease releases U–A, U–A–A– and C–G. The sequence of T13b therefore is U–A–A–C–G, which is one residue longer than the homologous mammalian fragment (oligonucleotide T20 of Forget and Weissman, 1967). Consequently, we suggest that one U residue is inserted, relative to the mammalian or *Xenopus* sequences, at position 21.

Positions 26–27. Chicken 5S *r*RNA yields one less C–G fragment among the T1 ribonuclease digestion products than does mammalian 5S *r*RNA and it yields two less G–C fragments among the products of pancreatic ribonuclease digestion. One of the absent G–C fragments is accounted for by a difference in the molecules at position 93 (see below). The remaining G–C difference among the pancreatic ribonuclease digestion products and the C–G difference among the T1 ribonuclease products are both accounted for by deletion of one of the adjacent C–G pairs at positions 24–27. It is impossible to decide which of the C–G pairs is deleted in the chicken 5S *r*RNA molecule, relative to the mammalian 5S *r*RNA; the structural outcome of deleting either is the same. This deletion compensates for the addition of two residues (see above) and restores the reference frame of the chicken 5S *r*RNA sequence to that of the mammalian and *Xenopus* molecules. The close proximity of the deletions and additions in the chicken 5S *r*RNA implies that the molecules possess some sort of required linear reference frame. For example, 5S *r*RNA function might require precise alignment of complementary sequences.

Although we do not recover one mole of C–G per mole of chicken 5S *r*RNA digested with T1 ribonuclease, we do routinely observe 0.1–0.3 moles of the dinucleotide. This fractional yield of C–G conceivably could be evidence for some heterogeneity within the 5S *r*RNA population, but it also may present minor, non-specific scission by T1 ribonuclease. It is noteworthy that T1 ribonuclease digestion of 5S *r*RNA from KB cells (Forget and Weissman, 1967), *Potorous* cells (Averner and Pace, 1972) or mouse L-cells (our unpublished observations) also routinely yield slightly more than the expected one mole of C–G per mole of 5S *r*RNA. Furthermore, Brownlee *et al.* (1972) and Wegnez *et al.* (1972) have reported, and we have confirmed, that the 5S *r*RNA of somatic cells of *Xenopus laevis* also yields fractional quantities of C–G; the derived sequence predicts that none of this dinucleotide should appear among the digestion products.

Position 93. Chicken oligonucleotide T19 is not found in mammalian 5S *r*RNA. The mammalian homolog of T19 is readily identified, since chicken 5S *r*RNA does not possess the adjacent T1 ribonuclease-generated

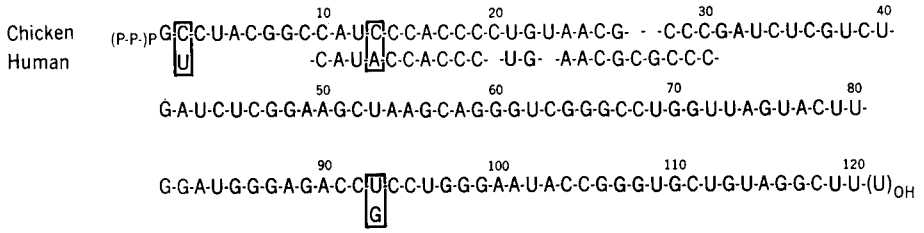

Fig. 3. The nucleotide sequence of chicken 5S *r*RNA. Differences from the mammalian 5S *r*RNA nucleotide sequence are indicated

mammalian fragments A–C–C–G and C–C–U–G. Substitution of U for G at position 93 accounts for chicken oligonucleotide T19.

All T1 and pancreatic ribonuclease fragments of chicken 5S *r*RNA except those noted are compatible with the mammalian 5S *r*RNA sequence. The chicken 5S *r*RNA sequence is shown in Fig. 3; the differences from the mammalian sequence are indicated.

Discussion

The 5S *r*RNA nucleotide sequences of the eukaryotes initially examined were all derived from mammals, including human (Forget and Weissman, 1967), mouse (Williamson and Brownlee, 1969), and a marsupial (Averner and Pace, 1972), and were found to have identical structures. This suggested that the 5S *r*RNA structure might be evolutionarily highly conservative. The availability of the nucleotide sequences of 5S *r*RNA from *Xenopus* and chicken bears out this suggestion. In all of these organisms considerable drift has occurred within the primary structures of the cytochromes *c*, which themselves are considered rather evolutionarily conservative (Nolan and Margoliash, 1968). Table 2 compares differences, among paired organisms, in the *m*RNA specifying the cytochromes *c* (calculated by the methods of Holmquist *et al.*, 1972; Jukes and Holmquist, 1972) and the 5S *r*RNA molecules. Apparently the primary structure of 5S *r*RNA is substantially more evolutionarily conservative than that of the *m*RNA defining the cytochromes *c*. The conservative nature of the 5S *r*RNA presumably reflects a rather rigorous requirement of the protein synthesizing machinery for a very particular structure of 5S *r*RNA. It may be that all or most of the nucleotides comprising the 5S *r*RNA molecule interact in a specific fashion with other components of the translation apparatus. That is, it would appear that little, if any, of the 5S structure is used as mere "spacer", whose detailed nucleotide sequence is of little consequence.

As with the 5S *r*RNA derived from somatic cells of other eukaryotes, we find no convincing evidence that structural heterogeneity exists within

Table 2. Structural divergence in genes specifying cytochrome *c* and 5S *r*RNA

	1 Cyto. *c* Amino acid residue changes	2 Cyto. *c* *m*RNA Probable base changes	3 % Divergence in cyto. *c* *m*RNA	4 5S *r*RNA base changes	5 % Divergence in 5S *r*RNA
Anuran: chicken	11	70	22.4	12	10
Anuran: marsupial	13	58	18.6	8	6.7
Chicken: marsupial	12	26	8.3	7	5.8
Marsupial: human	10	19	6.1	0	0

For the pairs of organisms considered, Column 1 gives the differences in amino acid residues in their cytochromes *c*. The anuran cytochrome *c* amino acid sequence is from bullfrog (S. K. Chan, O. F. Walasek, G. H. Barlow and E. Margoliash; cited in Dayhoff, 1972), and has been arranged to give maximum homology with other cytochromes *c* (Dayhoff, 1972). The marsupial sequence is from kangaroo (Nolan and Margoliash, 1966). The amino acid sequence of cytochrome *c* of chicken is after Chan and Margoliash (1966) and that of human after Matsubara and Smith (1962). Probable base changes between the cytochrome *c* *m*RNA of the organisms paired was calculated as described by Holmquist *et al.* (1972) and Jukes and Holmquist (1972). Column 3 lists the per cent base differences in cytochrome *c* *m*RNA between the paired organisms, considering 312 nucleotides in the *m*RNA (cytochrome *c* is composed of 104 amino acids). Column 4 lists observed differences in 5S *r*RNA from the paired organisms. The anuran sequence is that of *Xenopus laevis* (Brownlee *et al.*, 1972; Wegnez *et al.*, 1972) the marsupial 5S sequence is that of the rat kangaroo, *Potorous tridactylus* (Averner and Pace, 1972). Column 5 gives the per cent differences between the respective pairs of organisms.

the 5S *r*RNA population purified from chicken somatic cells; certainly the majority of the 5S *r*RNA population is structurally homogeneous. The genes specifying these 5S *r*RNA sequences of course must be identical, and hundreds of copies of these identical genes probably are functional in each cell (Attardi and Amaldi, 1970). Since extensive structural change may occur through evolution in the 5S *r*RNA sequence, and yet the multiple gene copies within individual species are identical, it would appear that the structures of these multiple genes are not drifting independently.

It is noteworthy that chicken 5S *r*RNA, like that from Hela cells (Hatten *et al.*, 1969), may possess 5′ terminal di- or triphosphorylated residues. We have also recovered 5′ di- and triphosphorylated guanylate residues from T1 ribonuclease digests of *Xenopus* and mouse L-cell 5S *r*RNA (our unpublished observations) as well as chicken 5S *r*RNA. As Hatten *et al.* (1969) have noted, if this phosphorylation is not a posttranscriptional event, then at least the 5′ terminus of mature 5S *r*RNA in all these eukaryotes is never associated with excess precursor sequences, unlike in the prokaryotes (Monier *et al.*, 1969; Pace *et al.*, 1973).

Acknowledgments. We thank Dr. C. R. Woese for carrying out U_2 RNase analyses of certain oligonucleotides and Drs. M. Averner and H. V. Rickenberg for helpful discussions. This research was supported by United States Public Health Service Grants GM-20147 to N. R. Pace and CA-05164 and AI-06844 to R. L. Erikson. This is contribution 578 from the Department of Biophysics and Genetics.

References

Attardi, G., Amaldi, F.: Ann. Rev. Biochem. **39**, 183 (1970)
Averner, M. J., Pace, N. R.: J. Biol. Chem. **247**, 4491 (1972)
Brownlee, G. G., Cartwright, E., McShane, T., Williamson, R.: FEBS Lett. **25**, 8 (1972)
Chan, S. K., Margoliash, E.: J. Biol. Chem. **241**, 507 (1966)
Dayhoff, M. O. (Ed.): Atlas of protein sequence and structure, Vol. 5. Silver Springs, Md.: National Biomedical Research Foundation 1972
Forget, B. G., Weissman, S. M.: Sci. **158**, 1695 (1967)
Hatten, L. E., Amaldi, F., Attardi, G.: Biochem. **8**, 4989 (1969)
Holmquist, R., Cantor, C., Jukes, T. H.: J. Mol. Biol. **64**, 145 (1972)
Jukes, T. H., Holmquist, R.: J. Mol. Biol. **64**, 163 (1972)
Matsubara, H., Smith, E. L.: J. Biol. Chem. **237**, 3575 (1962)
Monier, R., Feunteun, J., Forget, B., Jordan, B., Reynier, M., Varricchio, F.: Cold Spring Harbor Symp. Quant. Biol. **34**, 139 (1969)
Nolan, C., Margoliash, E.: J. Biol. Chem. **241**, 1049 (1966)
Nolan, C., Margoliash, E.: Ann. Rev. Biochem. **37**, 727 (1968)
Pace, B., Pace, N. R.: J. Bacteriol. **105**, 12 (1971)
Pace, N. R., Pato, M. L., McKibbin, J., Radcliffe, C. W.: J. Mol. Biol. **75**, 619 (1973)
Sanger, F., Brownlee, G. G., Barrell, B. G.: J. Mol. Biol. **13**, 373 (1965)
Sogin, M. L.: Ph. D. Thesis, The University of Illinois, Urbana (1972)
Wegnez, M., Monier, R., Denis, H.: FEBS Lett. **25**, 13 (1972)
Williamson, R., Brownlee, G. G.: FEBS Lett. **3**, 306 (1969)
Zubay, G.: J. Mol. Biol. **4**, 347 (1962)

Dr. Norman R. Pace
Department of Biophysics and Genetics
University of Colorado
Medical Center
Denver, Colorado 80220, USA