

Collection of published 5S and 5.8S RNA sequences and their precursors

Volker A. Erdmann

Max-Planck-Institut für Molekulare Genetik, Abt. Wittmann, Ihnestr. 63-73, 1000 Berlin-Dahlem (33), GFR

Received 20 December 1978

The previous collection (1) of mature 5S and 5.8S RNA sequences as well as those of their precursors are updated. This summary does not include those earlier publications in which the oligonucleotide composition, but not the sequences of 5S RNAs has been reported. For this information the reader is referred to reference 1.

The possible structures and functions of prokaryotic 5S and 5.8S RNAs are discussed in two other reviews (2,3).

I would also like to thank those colleagues who have sent me their pre- or reprints on small ribosomal RNA sequences in 1978.

Prokaryotic 5S RNA Sequences

Abbreviation	RNA Source	Reference Number
A.N.	<u>Anacystis nidulans</u> 1405/1 Katz/Allen (Blue-green Alga)	4
B.L.	<u>Bacillus licheniformis</u> S 244	5, 6, 8
B.M.	<u>Bacillus megaterium</u> KM	7
B.S.(a)	<u>Bacillus stearothermophilus</u> 1439 FV	8
B.S.(b)	<u>Bacillus stearothermophilus</u> (strain not given)	9
B.S.(b)	<u>Bacillus stearothermophilus</u> 799	10
B.Su.	<u>Bacillus subtilis</u> 168	6, 8
B.Q.	<u>Bacillus Q</u>	6
C.P.	<u>Clostridium pasteurianum</u> ATCC 6013	11
E.C.(a)	<u>Escherichia coli</u> MRE600	12 - 16
E.C.(b)	<u>Escherichia coli</u> CA265	12 - 15

PROKARYOTIC 5S RNA SEQUENCES

	1	10	20	30	40	50	60	70	80	90	100	110	120
H.C.	<u>P</u> UUA <u>AGCGG</u> CC <u>CAU</u> AGCGGUGGGUUA <u>CUC</u> CGGUA <u>CC</u> CAUCCGGAACGGAAGAAAGCCGCGCCUGCGU <u>UCC</u> GGUCAGUA <u>CUG</u> GAGUGCGAGCCUUGGGA <u>AAU</u> CCGGU <u>U</u> CGCGCCUA <u>C</u> U <u>OH</u>												
P.V.	<u>P</u> UGU <u>CUG</u> CGCC <u>CAU</u> AGCCAGUGGUC <u>CA</u> CCUGA <u>UCC</u> CAU <u>GG</u> CCGAACUC <u>AG</u> AAAGUGAAACGU <u>U</u> AGCGCCGA <u>U</u> GAU <u>GU</u> CGU <u>U</u> CGCGU <u>U</u> CC <u>CAU</u> UGAGAGUAGGGA <u>AA</u> CU <u>U</u> CGCGG <u>CAU</u> <u>OH</u>												
P.	<u>P</u> UCU <u>U</u> GGG <u>G</u> AC <u>CAU</u> AGCGGU <u>U</u> AGGACCC <u>CA</u> CCUGA <u>UCC</u> U <u>U</u> GGCC <u>CA</u> CU <u>AG</u> UAGUAAAGCCGCGGA <u>U</u> GGUAGU <u>GU</u> GGGGU <u>U</u> CGCC <u>CAU</u> UGAGAGUAGGGA <u>AA</u> CU <u>U</u> CGCGG <u>CAU</u> <u>OH</u>												
P.F.	<u>P</u> UGU <u>U</u> GU <u>U</u> AGCAG <u>U</u> AGUGG <u>CAU</u> UGGAAC <u>CU</u> CA <u>UCC</u> CA <u>UCC</u> CGGA <u>CU</u> AGAGCGGUA <u>AA</u> AGG <u>CAU</u> CGCGG <u>CAU</u> GGUAGU <u>GU</u> GGGGU <u>U</u> CGCC <u>CAU</u> UGA <u>U</u> CGA <u>CU</u> AGAG <u>CAU</u> AGAG <u>CAU</u> <u>OH</u>												
T.A.	<u>P</u> AAUCCCGCC <u>CUU</u> AGCGGCGG <u>U</u> GAAGACCC <u>GU</u> CC <u>CAU</u> UCCGAA <u>U</u> CGGAAGGAAAGCCGCGCC <u>CAU</u> GGU <u>CA</u> CU <u>U</u> GGGACCGCGG <u>GU</u> CC <u>U</u> GGAGAGUAGGUG <u>U</u> CGCGGGG <u>U</u> <u>OH</u>												

PROKARYOTIC 5S RNA PRECURSORS

	1	10	20	140	150	160	170	179
P.B.Su.	<u>p</u> UGAGAA <u>C</u> AC <u>U</u> CA <u>U</u> CA <u>U</u> U <u>U</u> U <u>G</u> <u>B.</u> <u>s</u> ubtilis 5S RNA..... <u>A</u> AGCU <u>U</u> AA <u>CC</u> AG <u>CU</u> CA <u>U</u> AGAG <u>CU</u> GG <u>U</u> GU <u>U</u> U <u>U</u> U <u>U</u> U <u>U</u> U <u>U</u> U <u>G</u> <u>OH</u>							
P.E.C.	<u>p</u> AUU , <u>p</u> UU and <u>p</u> U have been found at the 5`end of <u>E. coli</u> 5S RNA							

The single underlined sequences are tentative. The double underlined nucleotides occur in less than one mole per mole 5S RNA. Nucleotide written directly under another nucleotide in the sequence indicates that it may also be found in this position. For abbreviation of organisms and literature references see opposite page.

Prokaryotic 5S RNA Sequences

Abbreviation	RNA Source	Reference Number
H.C.	<u>Halobacterium cutirubrum</u> N.R.C. 34001	17
P.V.	<u>Proteus vulgaris</u> (strain not given)	16
P	<u>Photobacter</u> 8265	18
P.F.	<u>Pseudomonas fluorescens</u> ATCC 13430	19
T.A.	<u>Thermus aquaticus</u> ATCC 25104	20
<u>Prokaryotic 5S RNA Precursors</u>		
P.B.Su.	<u>Bacillus subtilis</u> 168	21,22
P.E.C.	<u>Escherichia coli</u> 217 (sud-1)	23

EUKARYOTIC 5S RNA SEQUENCES

	1	10	20	30	40	50	60	70	80	90	100	110	120
B. B.	P	A	G	U	C	G	A	U	A	C	A	C	A
C. (a)	<u>P</u>	<u>P</u>	<u>G</u>	<u>C</u>	<u>C</u>	<u>A</u>	<u>U</u>	<u>C</u>	<u>C</u>	<u>C</u>	<u>C</u>	<u>C</u>	<u>U</u>
C. (b)	<u>P</u>	<u>P</u>	<u>G</u>	<u>C</u>	<u>C</u>	<u>A</u>	<u>U</u>	<u>C</u>	<u>C</u>	<u>C</u>	<u>C</u>	<u>C</u>	<u>U</u>
C. P.	<u>P</u>	<u>P</u>	<u>P</u>	<u>A</u>	<u>U</u>	<u>C</u>	<u>C</u>	<u>C</u>	<u>A</u>	<u>U</u>	<u>C</u>	<u>C</u>	<u>U</u>
D. H.	<u>P</u>	<u>P</u>	<u>P</u>	<u>G</u>	<u>C</u>	<u>C</u>	<u>A</u>	<u>U</u>	<u>C</u>	<u>C</u>	<u>C</u>	<u>C</u>	<u>U</u>
D. B.	<u>P</u>	<u>P</u>	<u>P</u>	<u>G</u>	<u>C</u>	<u>C</u>	<u>A</u>	<u>U</u>	<u>C</u>	<u>C</u>	<u>C</u>	<u>C</u>	<u>U</u>
H. L.	<u>P</u>	<u>P</u>	<u>P</u>	<u>G</u>	<u>C</u>	<u>C</u>	<u>A</u>	<u>U</u>	<u>C</u>	<u>C</u>	<u>C</u>	<u>C</u>	<u>U</u>
K. B.	<u>P</u>	<u>P</u>	<u>P</u>	<u>G</u>	<u>C</u>	<u>C</u>	<u>A</u>	<u>U</u>	<u>C</u>	<u>C</u>	<u>C</u>	<u>C</u>	<u>U</u>
Re.	<u>P</u>	<u>P</u>	<u>P</u>	<u>G</u>	<u>C</u>	<u>C</u>	<u>A</u>	<u>U</u>	<u>C</u>	<u>C</u>	<u>C</u>	<u>C</u>	<u>U</u>
Ry.	<u>P</u>	<u>P</u>	<u>P</u>	<u>G</u>	<u>C</u>	<u>C</u>	<u>A</u>	<u>U</u>	<u>C</u>	<u>C</u>	<u>C</u>	<u>C</u>	<u>U</u>
R. T.	<u>P</u>	<u>P</u>	<u>P</u>	<u>G</u>	<u>C</u>	<u>C</u>	<u>A</u>	<u>U</u>	<u>C</u>	<u>C</u>	<u>C</u>	<u>C</u>	<u>U</u>
S.	<u>P</u>	<u>P</u>	<u>P</u>	<u>G</u>	<u>C</u>	<u>C</u>	<u>A</u>	<u>U</u>	<u>C</u>	<u>C</u>	<u>C</u>	<u>C</u>	<u>U</u>
To.	<u>P</u>	<u>P</u>	<u>P</u>	<u>G</u>	<u>C</u>	<u>C</u>	<u>A</u>	<u>U</u>	<u>C</u>	<u>C</u>	<u>C</u>	<u>C</u>	<u>U</u>
Tu.	<u>P</u>	<u>P</u>	<u>P</u>	<u>G</u>	<u>C</u>	<u>C</u>	<u>A</u>	<u>U</u>	<u>C</u>	<u>C</u>	<u>C</u>	<u>C</u>	<u>U</u>
W. E.	<u>P</u>	<u>P</u>	<u>P</u>	<u>G</u>	<u>C</u>	<u>C</u>	<u>A</u>	<u>U</u>	<u>C</u>	<u>C</u>	<u>C</u>	<u>C</u>	<u>U</u>
X. L. S.	<u>P</u>	<u>P</u>	<u>P</u>	<u>G</u>	<u>C</u>	<u>C</u>	<u>A</u>	<u>U</u>	<u>C</u>	<u>C</u>	<u>C</u>	<u>C</u>	<u>U</u>
X. L. O.	<u>P</u>	<u>P</u>	<u>P</u>	<u>G</u>	<u>C</u>	<u>C</u>	<u>A</u>	<u>U</u>	<u>C</u>	<u>C</u>	<u>C</u>	<u>C</u>	<u>U</u>
X. M. S.	<u>P</u>	<u>P</u>	<u>P</u>	<u>G</u>	<u>C</u>	<u>C</u>	<u>A</u>	<u>U</u>	<u>C</u>	<u>C</u>	<u>C</u>	<u>C</u>	<u>U</u>
X. M. O.	<u>P</u>	<u>P</u>	<u>P</u>	<u>G</u>	<u>C</u>	<u>C</u>	<u>A</u>	<u>U</u>	<u>C</u>	<u>C</u>	<u>C</u>	<u>C</u>	<u>U</u>

Single underlined sequences are tentative. Double underlined nucleotides or 5' phosphates occur in less than one mole per mole 5S RNA. ? underneath the 3' terminal U of the D.B. sequence indicates that it has not clearly been identified as uridine. X in W.E. is not certain; could be occupied by one or more unknown nucleotides. For abbreviations of organisms and literature references see opposite page.

Eukaryotic 5S RNA Sequences

Abbreviation	Source	Reference Number
B.B.	Broad bean (<u>Vicia faba</u>)	24, 25
C (a)	Chicken (<u>Gallus gallus</u>), embryo fibroblast culture	26
C (b)	Chicken, embryo fibroblast culture	27
C.P.	<u>Chlorella pyrenoidosa</u> 211/8b	28
D.M.	<u>Drosophila melanogaster</u> F6 of KC	29
D.B.	Dwarf bean (<u>Phaseolus vulgaris</u>)	24, 25
H.L.	HeLa cells	30, 31
K.B.	KB cells	32, 33
Re.	Reptile (<u>Iguana iguana</u>)	34
Ry.	Rye (<u>Secale cereale</u> c.v. Lovaszpatonai)	24, 25
R.T.	Rainbow trout (<u>Salmo gairdneri</u> , RTG-2)	35
S.	Sunflower (<u>Helianthus annuus</u>)	24, 25
To.	Tomato (<u>Lycopersicum esculentum</u>)	24, 25
Tu.	Turtle (<u>Terrapene carolina</u> , TH-I line of heart cells)	36
W.E.	Wheat embryo (Thatcher variety)	37-39
X.L.S.	<u>Xenopus laevis</u> (somatic from kidney)	40-42
X.L.O.	<u>Xenopus laevis</u> (oocytes)	40-42
X.M.S.	<u>Xenopus mulleri</u> (somatic)	43
X.M.O.	<u>Xenopus mulleri</u> (oocytes)	43

EUKARYOTIC 5S RNA SEQUENCES

	1	10	20	30	40	50	60	70	80	90	100	110	120
Y. S. Ca. (a)	ppp	g	u	u	g	g	g	g	g	g	g	g	g
Y. S. Ca. (b)	ppp	g	u	u	g	g	g	g	g	g	g	g	g
Y. S. Ce.	ppp	g	u	u	g	g	g	g	g	g	g	g	g
Y. K. L.	ppp	g	u	u	g	g	g	g	g	g	g	g	g
Y. P. M.	ppp	g	u	u	g	g	g	g	g	g	g	g	g
Y. T. U.	ppp	g	u	u	g	g	g	g	g	g	g	g	g

EUKARYOTIC 5S RNA PRECURSORS

P. D. M. Drosophila melanogaster 5S RNA sequence plus at 3' end ... CGUCCACAACUUUUUU^{OH} 125 130 135

P. H. L. 5S RNA synthesized by isolated HeLa cell nuclei in vitro was found to terminate at the 3' end with CUU^{OH} (60%), CUUU^{OH} (20%) and CUUUU^{OH} (20%).

Single underlined sequences are tentative. Double underlined nucleotides or 5' phosphates occur in less than one mole per mole 5S RNA. For abbreviation of organisms and literature references see opposite page.

Eukaryotic 5S RNA Sequences

Abbreviation	Source	Reference Number
Y.S.Ca. (a)	Yeast (<u>Saccharomyces carlsbergensis</u>)	44
Y.S.Ca. (b)	Yeast (<u>Saccharomyces carlsbergensis</u>)	45
Y.S.Ce.	Yeast (<u>Saccharomyces cerevisiae</u>)	45, 46
Y.K.L.	Yeast (<u>Kluyveromyces lactis</u>)	45
Y.P.M	Yeast (<u>Pichia membranaefaciens</u>)	45
Y.T.U.	Yeast (<u>Torulopsis utilis</u>)	47

Eukaryotic 5S RNA Precursors

Abbreviation	Source	Reference Number
P.D.M.	<u>Drosophila melanogaster</u> KcO	48
P.H.L.	HeLa cells	49

Eukaryotic 5.8S RNA Sequences

Abbreviation	RNA Source	Reference Number
C.	Chicken (embryonic cells)	50
H.L.	HeLa cells	50
N.	Novikoff hepatoma ascites cells	51
R.T.	Rainbow trout (<u>Salmo gairdneri</u> , RTG-2)	52
T.	Turtle (heart cells CCL 50)	53
X.B.	<u>Xenopus borealis</u> (somatic)	54
X.L.	<u>Xenopus leavis</u> (somatic)	50, 54
Y.S.Ce.	Yeast (<u>Saccharomyces cerevisiae</u> A364A gal-1 ade-1 ura-1 his-7 lys-2 tyr-1 (ATCC 22 244))	55

EUKARYOTIC 5.8S RNA PRECURSORS

p.H.L. pUCG instead of pCG has also been found at the 5' end of HeLa cell 5.8S RNA.

p.X.L. Three different 5' nucleotides are reported: pUCG (40%), pCG (20%) and pG (40%).
 From DNA sequencing data the additional 5' and 3' nucleotide sequences in precursor 5.8S RNA were deduced as follows:
 5'end: GCGCGCGCGGACCCUCACAGCGGACGCCCGGUGGCCUCCGACCCGAAAGAAAACCCGACCGCGGCGGAGAGCUUCG...
 3'end:..GACGUCCAUCGCCCCCGCGGUGCCGCGCGG_{OH}

p.Y.S.Ca. The following additional sequence has been found at the 3' end: CCUUCUCAAACAUUCUGp

p.Y.S.Ce. pUAUUA and pAUUAUA have been found at the 5' end of this yeast 5.8S RNA.

EUKARYOTIC 5.8S RNA PRECURSORS

Abbreviation	RNA Source	Reference Number
p.H.L.	HeLa cells	50
p.X.L.	<u>Xenopus leavis</u> (somatic)	54, 56
p.Y.S.Ca.	Yeast (<u>Saccharomyces carlsbergensis</u> , S-74)	57
p.Y.S.Ce.	Yeast (<u>Saccharomyces cerevisiae</u> S288 α mal gal-2)	58

Single underlined sequences are tentative.

REFERENCES

1. Erdmann, V.A. (1978). *Nucleic Acids Res.* 5, r1-r13.
2. Monier, R. (1974) in *Ribosomes*, Nomura, M., Tissières, A. and Lengyel, P., Eds., pp 141-168. Cold Spring Harbor Laboratory, New York.
3. Erdmann, V.A. (1976) in *Progress in Nucleic Acid Research and Molecular Biology*, Cohn, W.E., Ed., Vol. 18, pp 45-90. Academic Press, New York.
4. Corry, M.J., Payne, P.I. and Dyer, T.A. (1974). *FEBS Lett.* 46, 63-66.
5. Raué, H.A., Stoof, T.J. and Planta, R.J. (1975). *Eur. J. Biochem.* 59, 35-42.
6. Raué, H.A., Rosner, A. and Planta, R.J. (1977). *Molec. gen. Genet.* 156, 185-193.
7. Pribula, C.D., Fox, G.E. and Woese, C.R. (1974). *FEBS Lett.* 44, 322-323.
8. Marotta, C.A., Varricchio, F., Smith, I., Weissman, S.M., Sogin, M.L. and Pace, N.R. (1976). *J. Biol. Chem.* 251, 3122-3127.
9. Stanley, J.R. and Penswick, J.R. (1975). 9th FEBS Meeting, Abstract 421.
10. Zimmermann, J. and Erdmann, V.A. (1978). *Nucleic Acids Res.* 5, 2267-2288.
11. Pribula, C.D., Fox, G.E. and Woese, C.R. (1976). *FEBS Lett.* 64, 350-352.
12. Brownlee, G.G., Sanger, F. and Barrell, B.G. (1967). *Nature* 215, 735-736.
13. Brownlee, G.G. and Sanger, F. (1967). *J. Mol. Biol.* 23, 337-353.
14. Brownlee, G.G., Sanger, F. and Barrell, B.G. (1968). *J. Mol. Biol.* 34, 379-412.
15. Brownlee, G.G. (1972) in "Laboratory Techniques in Biochemistry and Molecular Biology" (eds. Work, T.S. and Work, E.) North-Holland Publishing Company, Amsterdam, p. 102.

16. Fischel, J.L. and Ebel, J.-P. (1975). *Biochimie* 57, 899-904.
17. Nazar, R.N. and Matheson, A.T. (1978). *J. Biol. Chem.* 253, 5464-5469.
18. Woese, C.R., Pribula, C.D., Fox, G.E. and Zablen, L.B. (1975). *J. Mol. Evol.* 5, 35-46.
19. DuBuy, B. and Weissman, S.M. (1971). *J. Biol. Chem.* 246, 747-761.
20. Nazar, R.N. and Matheson, A.T. (1977). *J. Biol. Chem.* 252, 4256-4261.
21. Sogin, M.L. and Pace, N.R. (1974). *Nature* 252, 598-600.
22. Sogin, M.L. and Pace, N. (1976). *J. Biol. Chem.* 251, 3480-3488.
23. Feunteun, J., Jordan, B.R. and Monier, R. (1972). *J. Mol. Biol.* 70, 465-474.
24. Payne, P.I., Corry, M.J. and Dyer, T.A. (1973). *Biochem. J.* 135, 845-851.
25. Payne, P.I. and Dyer, T.A. (1976). *Eur. J. Biochem.* 71, 33-38.
26. Pace, N.R., Walker, T.A. and Pace, B. (1974). *J. Mol. Evol.* 3, 151-159.
27. Brownlee, G.G. and Cartwright, E.M. (1975). *Nucleic Acid Res.* 2, 2279-2288.
28. Jordan, B.R., Galling, G. and Jourdan, R. (1974). *J. Mol. Biol.* 87, 205-225.
29. Benhamou, J. and Jordan, B.R. (1976). *FEBS Lett.* 62, 146-149.
30. Hatlen, L.E., Amaldi, F. and Attardi, G. (1969). *Biochem.* 8, 4989-5005.
31. Vigne, R. and Jordan, B.R. (1977). *J. Mol. Evol.* 10, 77-86.
32. Forget, B.G. and Weissman, S.M. (1967). *Science* 158, 1695-1699.
33. Forget, B.G. and Weissman, S.M. (1969). *J. Biol. Chem.* 244, 3148-3165.
34. Roy, K.L. and Enns, L. (1976). *J. Biol. Chem.* 251, 6352-6354.

-
35. Roy, K.L. (1978). *Can. J. Biochem.* 56, 60-65.
 36. Roy, K.L. (1977). *FEBS Lett.* 80, 266-270.
 37. Barber, C. and Nichols, J.L. (1978). *Can. J. Biochem.* 56, 357-364.
 38. Soave, C., Nucca, R., Sala, E., Viotti, A. and Galante, E. (1973). *Eur. J. Biochem.* 32, 392-400.
 39. Cunningham, R.S., Bonen, L., Doolittle, W.F. and Gray, M.W. (1976). *FEBS Lett.* 69, 116-122.
 40. Brownlee, G.G., Cartwright, E., McShane, T. and Williamson, R. (1972). *FEBS Lett.* 25, 8-12.
 41. Wegnez, M., Monier, R. and Denis, H. (1972). *FEBS Lett.* 25, 13-20.
 42. Ford, P.J. and Southern, E.M. (1973). *Nature New Biology* 241, 7-12.
 43. Ford, P.J. and Brown, R.D. (1976). *Cell* 8, 485-493.
 44. Hindley, J. and Page, S.M. (1972). *FEBS Lett.* 26, 157-160.
 45. Miyazaki, M. (1977). *Nucl. Acid Res. Special Supplement* No. 3, 153-156.
 46. Miyazaki, M. (1974). *J. Biochem.* 75, 1407-1407.
 47. Nishikawa, K. and Takemura, S. (1974). *J. Biochem.* 76, 935-947.
 48. Jacq, B., Jourdan, R. and Jordan, B.R. (1977). *J. Mol. Biol.* 117, 785-795.
 49. Yamamoto, M. and Seifart, K.H. (1978). *Biochemistry* 17, 457-461.
 50. Khan, N.S.N. and Maden, B.E.H. (1977). *Nucl. Acid Res.* 4, 2495-2505.
 51. Nazar, R.N., Sitz, T.O. and Busch, H. (1975). *J. Biol. Chem.* 250, 8591-8597.
 52. Nazar, R.N. and Roy, K.L. (1978). *J. Biol. Chem.* 253, 395-399.
 53. Nazar, R.N. and Roy, K.L. (1976). *FEBS Lett.* 72, 111-116.
 54. Ford, P.Y. and Mathieson, T. (1978). *Eur. J. Biochem.* 87, 199-214.
 55. Rubin, G.M. (1973). *J. Biol. Chem.* 248, 3860-3875.
-

56. Boseley, P.G., Tuyns, A. and Birnstiel, M.L. (1978).
Nucleic Acids Res. 5, 1121-1137.
57. DeJonge, D., Kastelein, R.A. and Planta, R.J. (1978).
Eur. J. Biochem. 83, 537-546.
58. Rubin, G.M. (1974). Eur. J. Biochem. 41, 197-202.