166–167 Nucleic Acids Research, 2000, Vol. 28, No. 1

5S ribosomal RNA database Y2K

Maciej Szymanski, Miroslawa Z. Barciszewska, Jan Barciszewski and Volker A. Erdmann[1,*]

Institute of Bioorganic Chemistry of the Polish Academy of Sciences, Noskowskiego 12, 61704 Poznan, Poland and [1]Institut für Chemie, Freie Universität Berlin, Thielallee 63, 14195 Berlin, Germany

Received September 29, 1999; Revised October 5, 1999; Accepted October 13, 1999

ABSTRACT

This paper presents the updated version (Y2K) of the database of ribosomal 5S ribonucleic acids (5S rRNA) and their genes (5S rDNA), http://rose.man/poznan.pl/5SData/index.html . This edition of the database contains 1985 primary structures of 5S rRNA and 5S rDNA. They include 60 archaebacterial, 470 eubacterial, 63 plastid, nine mitochondrial and 1383 eukaryotic sequences. The nucleotide sequences of the 5S rRNAs or 5S rDNAs are divided according to the taxonomic position of the source organisms.

CURRENT STUDIES ON 5S rRNA

Bacterial ribosomes (70S) consist of two unequal subunits, 30S (small) and 50S (large) (1–3). The 50S subunit contains 34 different proteins (L1–L34), a 23S rRNA of 2904 nt and a 5S rRNA and has recently been crystallised and its structure has been solved to 5 Å resolution (4–7). Ribosomal 5S RNA, a 120 nt long RNA of molecular weight 40 000 is found in virtually all ribosomes with the exception of mitochondria of some fungi, higher animals and most protists (8,9). However, recent data showed that 5S rRNA is a true organellar species in mitochondrial fractions purified from mammalian cells (10). It is located in the central protuberance of the large ribosomal subunit near the peptidyl transferase and factor-binding sites (1–3). Since the discovery in 1963 as a component of the *Escherichia coli* ribosome (11) a large amount of sequence data has been also collected for 5S RNA (12–14). However, we are still far away from a detailed knowledge of the tertiary structure and detailed function of 5S rRNA, although the last year resulted in some new very important structural data (15,16). The crystal structures of a 62 nt domain of *E.coli* 5S ribosomal RNA and the duplex dodecamer encompassing an internal loop E have been determined at 3.0 and 1.5 Å, respectively (15). Also the solution structure of a 42 nt derivative of *E.coli* 5S rRNA which includes the loops D and E has been determined by nuclear magnetic resonance spectroscopy (16). It was demonstrated that this portion of the 5S rRNA is a double helical region with several irregularities. Recently, great progress has been observed in crystallisation of short fragments of 5S rRNA. The X-ray structures of single domain E and of a helix E octamer and heptamer from *Thermus flavus* 5S rRNA have been solved at atomic resolution (17).

In prokaryotes and organelles, 5S rRNAs are synthesised as a part of single long transcript, together with 16S and 23S rRNAs. Eukaryotic 5S rRNAs of cytoplasmic ribosomes are usually encoded by separate genes arranged in tandem arrays of repeating units. Their number varies significantly up to several thousands in vertebrates and plants. The 5S rRNA genes are transcribed by polymerase III which is strongly inhibited by p53 (18) and depends strongly in eukaryotic cells on the binding of a transcription factor IIIA (TF IIIA) to the internal control region of 5S rRNA genes (19). There is also evidence for direct interactions of upstream regulatory elements and a new independent upstream promoter element centred about –17 to –20 (20). In *Xenopus* somatic cells histone H1 effects the transcription repression of oocyte type 5S rRNA genes, without altering the transcription of the somatic type 5S rRNA genes. This means that the locations of positioned nucleosomes on somatic and oocyte 5S rDNAs differ significantly, resulting in a differing accessibility of the TF IIIA binding site in the two nucleosomes and the binding of TF IIIA to the oocyte nucleosomes is achieved by nucleosome repositioning (21,22). One of the remarkable features of TF IIIA is that it is capable of specific binding to the 5S rRNA gene and the gene product with high affinity and specificity, although the three-dimensional structures of RNA and DNA are clearly different. A minimal RNA fragment that is sufficient for TF IIIA binding includes a truncated/mutated helix I, helix II and helix V, as well as structurally intact loops A and E (23,24). The X-ray structure of a TFIIIA–DNA complex shows how zinc fingers have been deployed to bind to separated promoter elements (25). The presence of the TF IIIA protein alters the UV-induced photoproducts on DNA and also reduces nucleotide excision repair (26,27).

5S rRNA is the only known RNA species that binds ribosomal proteins before it is incorporated into the ribosomes both in prokaryotes and eukaryotes.

In eukaryotes, the 5S rRNA molecule binds only ribosomal protein L5, whereas in bacteria it interacts with three ribosomal proteins L5, L18 and L25. The 5S rRNA assembly to 23S rRNA requires proteins L18, and L5 but not L25, which already binds to 5S rRNA (28). The tertiary structure of L25 showed high similarity to tRNA anticodon-binding domain of glutaminyl-tRNA synthetase (29). A limited trypsinisation of eukaryotic ribosomes releases two peptides of 32 and 14 kDa suggesting that the N- and C-terminal ends of the L5 protein were the first to be hydrolysed and exposed on the surface of the ribosome (30). Protein L5 being a central component of the 5S rRNA export system interacts also with eukaryotic initiation factor 5A which also binds HIV-1 Rev (31). 5S rRNA binds the viral protein (32) and enhances the methionyl- and isoleucyl-tRNA synthetase activities by direct interactions with

*To whom correspondence should be addressed. Tel. +49 30 838 6402; Fax. +49 30 838 6403; Email: erdmann@chemie.fu-berlin.de

MetRS and $tRNA^{Met}$ in the macromolecular aminoacyl-tRNA synthetase complex (33). Recently 5S rRNA has been identified in the degradosome complex (34). It was demonstrated that it is complementary with 12 and 15 nt strings to the intron 1 sequences of cobrotoxin b and cobrotoxin genes (35).

It seems that in addition to 5S rRNA–protein interactions an important role is played by the contacts of 5S rRNA with 23S rRNA. Multiple cross-links from 5S rRNA to two distinct regions of the 23S rRNA were observed. The first and second regions were located at sites between nt 885 and 992 and 2272 and 2345 of 23S rRNA, respectively (36). Base-paired interaction between 5S rRNA (residues 91–110) in the large subunit and 18S rRNA in small subunit could contribute to the reversible association of the ribosomal subunits (37).

Thus 5S rRNA is an attractive model system for exploring fundamental issues of RNA conformation and RNA protein interaction due to its relatively small size, ease of preparation and its rich array of non-canonical base pairs (16).

To get a consistent picture of structure–function relationships of 5S rRNA, detailed knowledge concerning the primary structure of this RNA species from different sources is required.

THE DATA BANK

This edition of the database contains 1985 nt sequences of 5S rRNAs and 5S rDNAs published through September 1999. In comparison with the 1999 edition of the database (11), 96 entries are new. Most of them are partial sequences of plant 5S rRNA genes. The database entries use the format of the EMBL Nucleotide Sequence Data Bank. The 5S rDNA nucleotide entries contains the 5S rRNA coding sequence as well as information on the length of the original clone and location of the structural gene.

Files with the primary structure data and the nucleotide sequence alignments are available via the WWW at http://rose.man.poznan.pl/5SData/index.html . Any nucleotide sequence can be retrieved using the taxonomy browser or from an alphabetical list of organisms.

ACKNOWLEDGEMENTS

This work has been supported by the Deutsche Forschungsgemeinschaft (Gottfried Wilhelm Leibnitz Prize to V.A.E., the Sonderforschungsbereich 344-C8, the Deutsche Agentur fur Raumfahrtangelegenheiten GmbH, the Fonds der Chemischen Industrie e.V. and the Polish State Committee for Scientific Research.

REFERENCES

1. Porse,B.T. and Garrett,R.A. (1999) *Cell*, **97**, 423–426.
2. Garrett,R. (1999) *Nature*, **400**, 811–812.
3. Pennisi,E. (1999) *Science*, **285**, 1343.
4. Agrawal,R.K. and Frank,J. (1999) *Curr. Opin. Struct. Biol.*, **9**, 215–221.
5. Moore,P.B. (1998) *Annu. Rev. Biomol. Struct.*, **27**, 35–58.
6. Ban,N., Nissen,P., Penczek,P., Hansen,J., Capel,M., Moore,P.B. and Steitz,T.A. (1999) *Nature*, **400**, 841–847.
7. Clemons,W.M., May,J.L.C., Wimberly,B.T., McCutcheon,J.P., Capel,M.S. and Ramakrishan,V. (1999) *Nature*, **400**, 833–840.
8. Gray,M.W., Burger,G. and Lang,B.F. (1999) *Science*, **283**, 1476–1481.
9. Spruyt,N., Delarbe,C., Gachelin,G. and Laudet,V. (1998) *Nucleic Acids Res.*, **26**, 3279–3285.
10. Magalhaes,P.J., Andreu,A.L. and Schon,E.A. (1998) *Mol. Biol. Cell*, **9**, 2375–2382.
11. Rosset,R. and Monier,R. (1963) *Biochim. Biophys. Acta*, **68**, 653–656.
12. Barciszewska,M., Erdmann,V.A. and Barciszewski,J. (1996) *Biol. Rev.*, **71**, 1–25.
13. Zheng,P., Burrows,C.J. and Rokita,S.E. (1998) *Biochemistry*, **37**, 2207–2214.
14. Szymanski,M., Barciszewska,M., Barciszewski,J. and Erdmann,V.A. (1999) *Nucleic Acids Res.*, **27**, 158–160.
15. Correll,C.C., Freeborn,B., Moore,P.B. and Steitz,T.A. (1997) *Cell*, **91**, 705–712.
16. Dallas,A. and Moore,P.B. (1997) *Structure*, **5**, 1639–1653.
17. Perbandt,M., Nolte,A., Lorentz,S., Bald,R., Betzel,C. and Erdmann,V.A. (1998) *FEBS Lett.*, **429**, 211–215.
18. Cairns,C.A. and White,R.J. (1998) *EMBO J.*, **17**, 3112–3223.
19. Shastry,B.S. (1996) *J. Cell Sci.*, **109**, 535–539.
20. Lee,Y., Wong,W.M., Guyer,D., Erkine,A.M. and Nazar,R.N. (1997) *J. Mol. Biol.*, **269**, 676–683.
21. Panetta,G., Buttinell,M., Flaus,A., Richmond,T.J. and Rhodes,D. (1998) *J. Mol. Biol.*, **282**, 683–697.
22. Tomaszewski,R., Mogielnicka,E. and Jerzmanowski,A. (1998) *Nucleic Acids Res.*, **26**, 5596–5601.
23. Theunissen,O., Rudt,F. and Pieler,T. (1998) *Eur. J. Biochem.*, **258**, 758–767.
24. Leontis,N.B and Westhof,E. (1998) *RNA*, **4**, 1134–1153.
25. Nolte,R.T., Conlin,R.M., Harrison,S.C. and Brown,R.S. (1998) *Proc. Natl Acad. Sci. USA*, **95**, 2938–2943.
26. Conconi,A., Liu,X., Koriazova,L., Ackerman,E.J., and Smerdon M.J. (1999) *EMBO J.*, **18**, 1387–1396.
27. North,M.T. and Allison,L.A. (1998) *J. Cell Biochem.*, **69**, 490–505.
28. Ostergaard,P., Phan,H., Johansen,L.B., Egebjerg,J., Ostergaard,L., Porse,B.T. and Garrett,R.A. (1998) *J. Mol. Biol.*, **284**, 227–240.
29. Stoldt,M., Wohnert,J., Gorlach,M. and Brown,L.R. (1998) *EMBO J.*, **17**, 6377–6384.
30. Lin,E., Liu,S.-R. and Lin,A. (1999) *J. Biochem.*, **125**, 1029–1033.
31. Schatz,O., Oft,M., Dascher,C., Schebesta,M., Rosorius,O., Jaksche,H., Dobrovnik,D., Bevec,D. and Hauber,J. (1998) *Proc. Natl Acad. Sci. USA*, **95**, 1607–1612.
32. Lam,W.-H., Seifert,J.M., Amberger,F., Graf,C., Auer,M. and Millar,D. (1998) *Biochemistry*, **37**, 1800–1809.
33. Ogata,K., Ohno,R., Morioka,S. and Terao,K. (1996) *J. Biochem.*, **120**, 869–880.
34. Bessarab,D.A., Kaberdin,V.R., Wei,C.-L., Liou,G.-G. and Lin-Chao,S. (1998) *Proc. Natl Acad. Sci. USA*, **95**, 3157–3161.
35. Chang,L.-S., Cho,Y.-C., Lin,S.-R., Wu,B.-N., Lin,J., Hong,E., Sun,Y.-J. and Hsiao,C.-D. (1997) *J. Biochem.*, **122**, 1252–1259.
36. Osswald,M. and Brimacombe,R. (1999) *Nucleic Acids Res.*, **27**, 2283–2290.
37. Azad,A.A., Failla,P. and Hanna,P.J. (1998) *Biochem. Biophys. Res. Commun.*, **248**, 51–56.