

INTRON AND RNA SPLICING IN ARCHAEA

Yoh-ichi Watanabe^{1*} and Shigeo Yoshinari^{1,2}

¹Department of Biomedical Chemistry, Graduate School of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, 113-0033, Japan

²Present address: Department of Biological Sciences, State University of New York at Buffalo, Buffalo, NY 14260, USA.

*Corresponding author: Department of Biomedical Chemistry, Graduate School of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, 113-0033, Japan, Fax: +81-3-5841-3444, e-mail: ywatanab@m.u-tokyo.ac.jp

Abstract

In Archaea, almost all introns in pre-tRNA, pre-rRNA, and pre-mRNA are spliced through two common steps by protein enzymes: cleavage of the precursor with splicing endonuclease, and ligation of the exons with RNA ligase. We found the first examples of archaeal pre-mRNA splicing and cleavage of the pre-mRNA with a novel subclass of archaeal splicing endonuclease. We further solved the novel tertiary structure of the splicing endonuclease, and revealed that the lineage-specific insertion of amino acid residues in the endonuclease expands the recognition of the substrate precursor RNA. We also discuss the possible involvement of tRNA splicing and its machinery in the origin of the tRNA molecule.

(Keywords)

intron, RNA splicing, Archaea, splicing endonuclease, RNA ligase, bulge-helix-bulge motif

Introduction

In bacteria and eukaryotes, some types of RNA splicing occur by precursor RNAs themselves, or with the aid of trans-factors including RNA-protein complexes [1]. However, in Archaea, with a few exceptions of group II introns found in Euryarchaeota [2], almost all of the introns in pre-tRNA [3], pre-rRNA [4], and pre-mRNA [5] are thought to be spliced through two steps catalyzed by protein-based enzymes: cleavage at exon-intron boundaries in precursor RNA with splicing endonuclease [6], and ligation of the exon fragments with RNA ligase [7].

The splicing endonuclease recognizes a specific structure in the substrate precursor RNA bulge-helix-bulge (BHB) motif, composed of a 4-basepair central helix flanked by two 3-base bulge loops at the 3' sides together with basepairs, and cleaves the specific sites in the "bulge" region [8, Fig. 1A]. The recognition of the substrate with the enzyme is independent of the other part of the substrate, including the majority of the exon region [9]. Archaeal splicing endonucleases are classified into three subclasses based on subunit structures: homo-tetrameric enzymes, homo-dimeric enzymes, and hetero-tetrameric (alpha2beta2) enzymes [6]. Archaeal splicing endonuclease shares a common ancestor with eukaryotic tRNA splicing endonuclease [10].

The recently identified archaeal RNA ligase [11] shares a common ancestor with a bacterial RNA repair enzyme [12] and a human tRNA ligase [13]. Archaeal RNA ligase joins the 3' end of the 5' exon

with a 2',3'-cyclic phosphate residue and the 5' end of the 3' exon without a phosphate residue, and uses the phosphate residue at the 3' end of the 5' exon as the linkage between the exons [14].

Discovery of archaeal pre-mRNA splicing, and identification of structural elements in archaeal splicing endonuclease for expansion of substrate RNA recognition

In 2002, we identified the first examples of introns in archaeal protein coding genes [15]. Further analysis revealed that pre-mRNA fragments were cleaved at the exon-intron boundaries with heteromeric crenarchaeal splicing endonuclease [5]. Notably, some of the exon-intron boundaries in orthologous genes in related species may not form strict BHB motifs, including mismatches in the helices and/or bulge loops with fewer or more than 3 bases (Fig. 1B) [5,15,16]. Conventional euryarchaeal

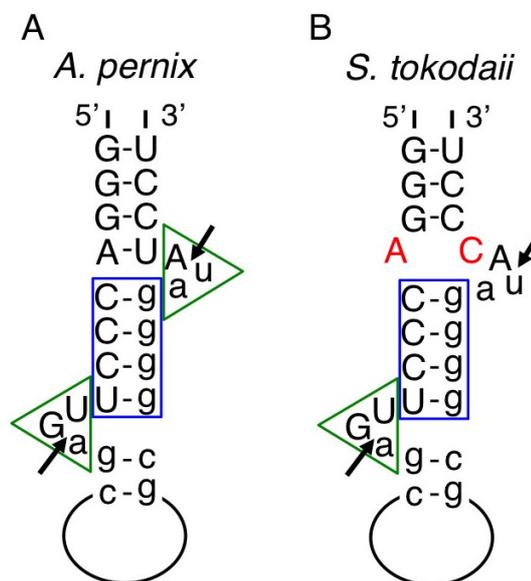


Fig 1. BHB motifs predicted in archaeal *cbf5* pre-mRNAs [15]. A, a conventional BHB motif in *Aeropyrum pernix* *cbf5* pre-mRNA. B, a relaxed BHB motif in *Sulfolobus tokodaii* *cbf5* pre-mRNA. Rectangles in blue represent the central 4-bp helices. Triangles in green represent the 3-base bulge loops. Note that in *S. tokodaii* pre-mRNA, the C residue at the 3' side of the 3' bulge (indicated in red) cannot form a Watson-Crick basepair with the A residue (also in red) at the 5' side of the central helix. Arrows represent cleavage sites with the splicing endonuclease confirmed *in vitro* [5,18]. Upper case, exon regions; lower case, intron regions.

homo-tetrameric and homo-dimeric splicing endonucleases cannot cleave such relaxed forms of the BHB motif [6]. To reveal the structural elements that expand the substrate recognition in the crenarchaeal heteromeric enzyme, we determined the crystal structures of two crenarchaeal enzymes, and found that they take a hetero-tetrameric ($\alpha_2\beta_2$) subunit structure [17,18]. Moreover, we showed that the specific insertion of amino acid sequences in the crenarchaeal enzymes, rather than the subunit structure of the crenarchaeal enzymes, is important for substrate recognition [17,18]. Transplantation of the insertion sequences into a euryarchaeal homo-dimeric enzyme expanded substrate recognition [19,20, Yoshinari, S. and Watanabe, Y. unpublished results]. A specific lysine residue in the inserted sequence is important for the substrate recognition/catalysis of the enzyme [18,19,20]. Although the nanoarchaeal enzyme with the expanded substrate recognition does not have the insertion at the corresponding position [21], a novel insertion at the alternative position may have a similar role [18].

tRNA splicing and the origin of tRNA

It has been proposed that tRNA was born by a fusion of two half RNA molecules which included a single stem-loop flanked by single strand regions at both the 5' and 3' sides [22]. Interestingly, some of the eukaryotic pre-tRNAs including introns in the anticodon loops have autocatalytic hydrolysis at or near the position of exon-intron boundaries [23], suggesting that a pre-tRNA molecule with an intron may be related to the prototypical tRNA. It would be interesting to investigate the autocatalytic hydrolysis of archaeal pre-tRNAs. Furthermore, many bacteria have homologs of archaeal and eukaryotic (t)RNA ligase, although bacteria do not have tRNA splicing driven by splicing endonucleases and RNA ligases [7], suggesting that this type of RNA ligase may exist in the last common ancestor of the three domains of life, and may have played a role in forming a prototype tRNA from two tRNA half molecules.

Acknowledgments

Our studies were partly supported by the metabolic protein group of the National Project on Protein Structural and Functional Analyses, and Grants-in-Aid of the Ministry of Science, Sports and Culture, Japan to YW and SY.

References

1. Watson, J.D. Baker, T.A., Bell, S.P., Gann, A., Levine, M., and Losick, R. *Molecular Biology of the Gene* (7th Edition), Benjamin Cummings, USA, 2013
2. Dai, L., and Zimmerly, S. ORF-less and reverse-transcriptase-encoding group II introns in archaeobacteria, with a pattern of homing into related group II intron ORFs. *RNA* 9, 14-9 (2003)
3. Daniels, C.J., Gupta, R. and Doolittle, W.F. Transcription and excision of a large intron in the tRNA^{Trp} gene of an archaeobacterium, *Halobacterium volcanii*. *J. Biol. Chem.* 260, 3132-3134 (1985)
4. Kjerms, J., and Garrett R.A. An intron in the 23s ribosomal-rna gene of the archaeobacterium *Desulfurococcus mobilis* *Nature* 318, 675-677 (1985)
5. Yoshinari, S., Itoh, T., Hallam, S.J., DeLong, E.F., Yokobori, S., Yamagishi, A., Oshima, T., Kita, K., and Watanabe, Y. Archaeal pre-mRNA splicing: a connection to hetero-oligomeric splicing endonuclease. *Biochem. Biophys. Res. Commun.* 346, 1024-1032 (2006)
6. Calvin, K., and Li, H. RNA-splicing endonuclease structure and function. *Cell. Mol. Life Sci.* 65, 1176-1185 (2008)
7. Popow, J., Schleiffer, A., and Martinez, J. Diversity and roles of (t)RNA ligases. *Cell. Mol. Life Sci.* 69, 2657-2670 (2012)
8. Marck, C., and Grosjean H. Identification of BHB splicing motifs in intron-containing tRNAs from 18 archaea: evolutionary implications. *RNA* 9, 1516-1531 (2003)
9. Thompson, L.D., and Daniels, C.J. A tRNA^{Trp} intron endonuclease from *Halobacterium volcanii*. Unique substrate recognition properties. *J. Biol. Chem.* 263, 17951-17959 (1988)
10. Bujnicki, J.M., and Rychlewski, L. Prediction of a common fold for all four subunits of the yeast tRNA splicing endonuclease: implications for the evolution of the EndA/Sen family. *FEBS Lett.* 486, 328-329 (2000)
11. Englert, M., Sheppard, K., Aslanian, A., Yates, J.R. 3rd, and Söll, D. Archaeal 3'-phosphate RNA splicing ligase characterization identifies the missing component in tRNA maturation. *Proc. Natl. Acad. Sci. U S A.* 108, 1290-1295 (2011)
12. Tanaka, N., and Shuman, S. RtcB is the RNA ligase component of an *Escherichia coli* RNA repair operon. *J. Biol. Chem.* 286, 7727-7731 (2011)
13. Popow, J., Englert, M., Weitzer, S., Schleiffer, A., Mierzwa, B., Mechtler, K., Trowitzsch, S., Will, C.L., Lüthmann, R., Söll, D., and Martinez, J. HSPC117 is the essential subunit of a human tRNA splicing ligase complex. *Science* 331, 760-764 (2011)
14. Zofalova, L., Guo, Y., and Gupta, R. Junction phosphate is derived from the precursor in the tRNA spliced by the archaeon *Haloferax volcanii* cell extract. *RNA* 6, 1019-1030 (2000)
15. Watanabe, Y., Yokobori, S., Inaba, T., Yamagishi, A., Oshima, T., Kawarabayasi, Y., Kikuchi, H., and Kita, K. Introns in protein-coding genes in Archaea. *FEBS Lett.* 510, 27-30 (2002)
16. Yokobori, S., Itoh, T., Yoshinari, S., Nomura, N., Sako, Y., Yamagishi, A., Oshima, T., Kita, K., and Watanabe, Y. Gain and loss of an intron in a protein-coding gene in Archaea: the case of an archaeal RNA pseudouridine synthase gene. *BMC Evol. Biol.* 9, 198 (2009)
17. Yoshinari, S., Shiba, T., Inaoka, D.K., Itoh, T., Kurisu, G., Harada, S., Kita, K., and Watanabe, Y. Functional importance of crenarchaea-specific extra-loop revealed by an X-ray structure of a heterotetrameric crenarchaeal splicing endonuclease. *Nucleic Acids Res.* 37, 4787-4798 (2009)
18. Okuda, M., Shiba, T., Inaoka, D.K., Kita, K., Kurisu, G., Mineki, S., Harada, S., Watanabe, Y., and Yoshinari, S. A conserved lysine residue in the crenarchaea-specific loop is important for the crenarchaeal splicing endonuclease activity. *J. Mol. Biol.* 405, 92-104 (2011)
19. Hirata, A., Kitajima, T., and Hori H. Cleavage of intron from the standard or non-standard position of the precursor tRNA by the splicing endonuclease of *Aeropyrum pernix*, a hyper-thermophilic Crenarchaeon, involves a novel RNA recognition site in the Crenarchaea specific loop. *Nucleic Acids Res.* 39, 9376-9389 (2011)
20. Hirata, A., Fujishima, K., Yamagami, R., Kawamura, T., Banfield, J.F., Kanai, A., and Hori, H. X-ray structure of the fourth type of archaeal tRNA splicing endonuclease: insights into the evolution of a novel three-unit composition and a unique loop involved in broad substrate specificity. *Nucleic Acids Res.* 40, 10554-10566 (2012)
21. Mitchell, M., Xue, S., Erdman, R., Randau, L., Söll, D., and Li, H. Crystal structure and assembly of the functional Nanoarchaeum equitans tRNA splicing endonuclease. *Nucleic Acids Res.* 37, 5793-5802 (2009)
22. Bernhardt, H.S. The RNA world hypothesis: the worst theory of the early evolution of life (except for all the others)(a). *Biol. Direct.* 7, 23 (2012)
23. van Tol, H., Gross, H.J., Beier, H. Non-enzymatic excision of pre-tRNA introns? *EMBO J.* 8, 293-300 (1989)