# Correspondence

## Prediction of a common fold for all four subunits of the yeast tRNA splicing endonuclease: implications for the evolution of the EndA/Sen family

Janusz M. Bujnicki*, Leszek Rychlewski

First published online 1 December 2000

tRNA maturation involves pre-tRNA splicing in all three domains of life. In Eubacteria, chloroplasts and mitochondria introns are self-splicing, but in Archaea and Eukaryota they are removed by the stepwise action of three protein enzymes: an endonuclease, a ligase and a phosphotransferase [1]. In the first step the pre-tRNA is cleaved at its two splice sites, generating two tRNA half molecules and the linear intron with 5′-OH and 3′-cyclic $PO_4$ ends. The Archaeal EndA endonuclease (a dimer in *Haloferax volcanii* and a homotetramer in *Methanococcus jannaschii*) recognizes and cuts pre-tRNA within the conserved bulge–helix–bulge motif. On the other hand the heterotetrameric Sen enzyme from yeast uses a measuring mechanism to determine the position of the splice sites relative to the conserved domain of pre-tRNA [2]. Trotta et al. [3] reported that the Sen2 and Sen34 subunits of the yeast enzyme contain the catalytic domain similar to all Archaeal enzymes. The Sen15 and Sen54 subunits were predicted to be non-homologous, since they showed no sequence similarity to each other and to the 'catalytic' subunits. The presence of a weakly conserved 17 amino acid (aa) sequence segment 'X' at the C-termini of all subunits was interpreted as essential for the quaternary structure formation, but not indicative of common origin [3–5].

The crystal structure of the homotetrameric EndA endonuclease from *M. jannaschii* revealed an isologous dimer of homodimers, in which the C-terminal 'catalytic' domains from two monomers participate in the cleavage reaction and the other two, along with four N-terminal domains, stabilize the quaternary structure of the protein [4]. Lykke-Andersen and Garrett [5] suggested that the dimeric endonuclease from *H. volcanii* arose from a duplicated *M. jannaschii*-type enzyme, in which the 'spare' active site in one of the repeats degenerated. This prediction has been recently validated by the X-ray crystallography of the related *Archaeaoglobus fulgidus* homodimeric endonuclease (Hong Li, personal communication). Similarly, in the yeast enzyme Sen2 and Sen34 were predicted to carry out catalysis and Sen15 and Sen54 to stabilize the quaternary structure of the endonuclease, with Sen54 handling the interaction with the mature domain of tRNA [3]. However, the unknown origin and presumed 'non-homology' of Sen15 and Sen54 has been a major obstacle in using the Archaeal enzymes as a model to predict the detailed structure and analyze the structure–function relationship of the eukaryotic enzyme.

In the absence of significantly similar sequences in the nr database, as reported by others and confirmed by us using PSI-BLAST [6], we sought to identify the homologs of Sen15 and Sen54 among structurally similar proteins by the sequence-to-structure threading method. We submitted the sequences of interest to secondary and tertiary structure prediction servers via the IIMCB Bioinformatics Unit Metaserver

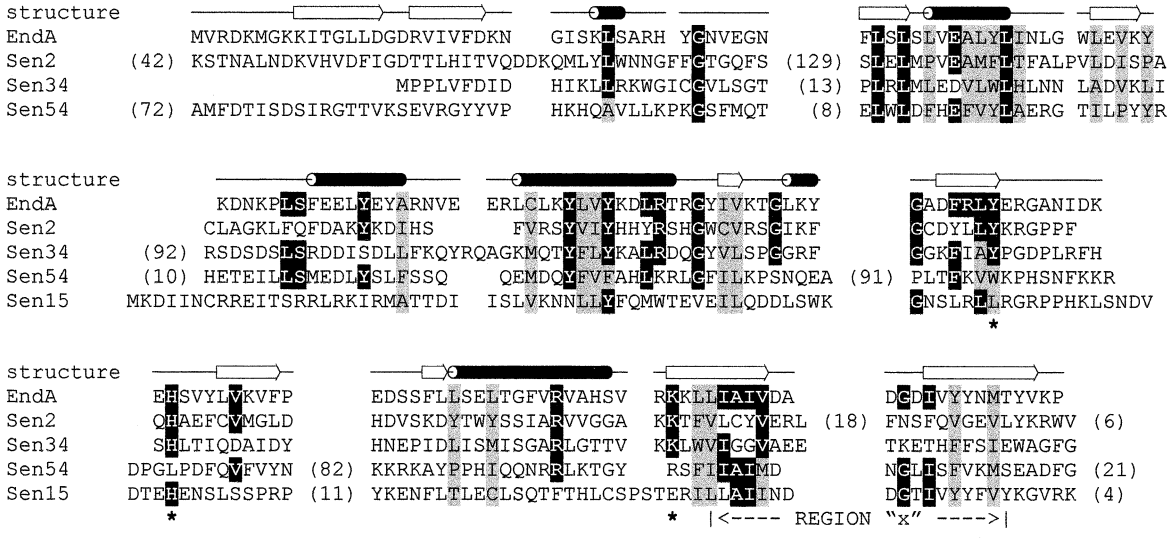


Fig. 1. The consensus threading-based alignment of *M. jannaschii* EndA (1a79) and four subunits of the *S. cerevisiae* Sen enzyme (gene identification numbers: Sen2p, 6323134; Sen34p, 6319322; Sen54p, 6325174; Sen15p, 6323704). The secondary structure of EndA is shown at the top. The first and last residues of the aligned region are shown preceding the first block and following the last block, respectively. The size of gaps between blocks is indicated by numbers in parentheses. Residues, which are identical in EndA and at least two subunits of Sen, are highlighted in black. Positions, where the hydrophobic or aromatic character of the side chain is conserved (aa L, I, V, M, F, Y, W, P), are shown in gray. The catalytic aa are labeled by stars, the position of the previously identified common 'region X' is indicated.

interface (http://bioinfo.pl/meta/). Surprisingly, the results revealed that the 128 aa sequence of Sen15 perfectly fits the structure of the C-terminal catalytic domain of *M. jannaschii* EndA (aa 75–171 of the PDB entry 1a79) according to the first hits reported by 3DPSSM [7], high score 0.000293; FUGUE [8], moderate score 5.47; and GENTHREADER [9], low score 0.199 (scoring systems of the individual servers are not normalized). Moreover, the full-length 467 aa sequence of Sen54 showed significant similarity to the EndA N-terminal domain (aa 1–83), but according only to SAMT-99 [10] (score 27.92). Inspection of sub-optimal alignments and secondary structure patterns predicted for all proteins suggested that Sen54 possessed several regions missing in other subunits. Many of these regions were also poorly conserved in putative Sen54 orthologs from other eukaryotes (not shown). When these 'insertions' were removed from the Sen54 sequence, all algorithms reported its similarity to the full-length EndA with very high scores (for instance FFAS [11] $=42.61$; GENTHREADER [9] $=1$; 3DPSSM [7] $=10^{-8}$; INBGU [12] $=124.7$). Such high structural similarity of closely associated proteins indicates that they may be evolutionarily related, especially since the structure of EndA does not resemble any of the ubiquitous folds believed to have a polyphyletic origin. The results of our analysis are summarized in Fig. 1.

Our findings suggest that, unlike previously proposed, both catalytic and non-catalytic subunits of eukaryotic tRNA splicing endonuclease share common origin with the Archaeal enzyme. We predict that the core of the yeast enzyme is essentially identical to its Archaeal counterpart, despite the fact that the Sen15 subunit apparently lost its N-terminal domain and Sen54 become enlarged by integration of elements presumably making up the distance-measuring domain. We failed to predict the three-dimensional structure of the yeast-specific elaborations of the common EndA fold. However we hope that our prediction of conserved and variable elements in all four subunits of the yeast enzyme will facilitate further structural, functional and phylogenetic studies of the Sen/EndA family.

*Acknowledgements:* We would like to thank Dr. Hong Li for providing us with unpublished material.

## References

[1] Lykke-Andersen, J., Aagaard, C., Semionenkov, M. and Garrett, R.A. (1997) Trends Biochem. Sci. 22, 326–331.
[2] Reyes, V.M. and Abelson, J. (1988) Cell 55, 719–730.
[3] Trotta, C.R., Miao, F., Arn, E.A., Stevens, S.W., Ho, C.K., Rauhut, R. and Abelson, J.N. (1997) Cell 89, 849–858.
[4] Li, H., Trotta, C.R. and Abelson, J. (1998) Science 280, 279–284.
[5] Lykke-Andersen, J. and Garrett, R.A. (1997) EMBO J. 16, 6290–6300.
[6] Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997) Nucleic Acids Res. 25, 3389–3402.
[7] Kelley, L.A., McCallum, C.M. and Sternberg, M.J. (2000) J. Mol. Biol. 299, 501–522.
[8] Shi, J., Blundell, T.L. and Mizuguchi, K. http://www-cryst.bioc.cam.ac.uk/~fugue/.
[9] Jones, D.T. (1999) J. Mol. Biol. 287, 797–815.
[10] Park, J., Karplus, K., Barrett, C., Hughey, R., Haussler, D., Hubbard, T. and Chothia, C. (1998) J. Mol. Biol. 284, 1201–1210.
[11] Rychlewski, L., Jaroszewski, L., Li, W. and Godzik, A. (2000) Prot. Sci. 9, 232–241.
[12] Fischer, D. (2000) Pac. Symp. Biocomput., 119–130.

*Corresponding author. Fax: (48)-22-668 5288.
E-mail: iamb@bioinfo.pl

*Bioinformatics Laboratory, International Institute of Molecular and Cell Biology, ul. ks. Trojdena 4, 02-109 Warsaw, Poland*

PII: S0014-5793(00)02322-X