Supplementary Figures and Table

The RNA-splicing endonuclease from the euryarchaeon *Methanopyrus kandleri* is a heterotetramer with constrained substrate specificity

Ayano Kaneta[1], Kosuke Fujishima[2], Wataru Morikazu[1], Hiroyuki Hori[1] and Akira Hirata[1]*

Author affiliation:

[1]Department of Materials Science and Biotechnology, Graduate School of Science and Engineering, Ehime University, Bunkyo-cho, Matsuyama, Ehime 790-8577, Japan.

[2] Earth-Life Science Institute, Tokyo Institute of Technology, Meguro-ku, Tokyo, 152-8550 Japan.

***To whom correspondence should be addressed:**

Akira Hirata

Department of Materials Science and Biotechnology,

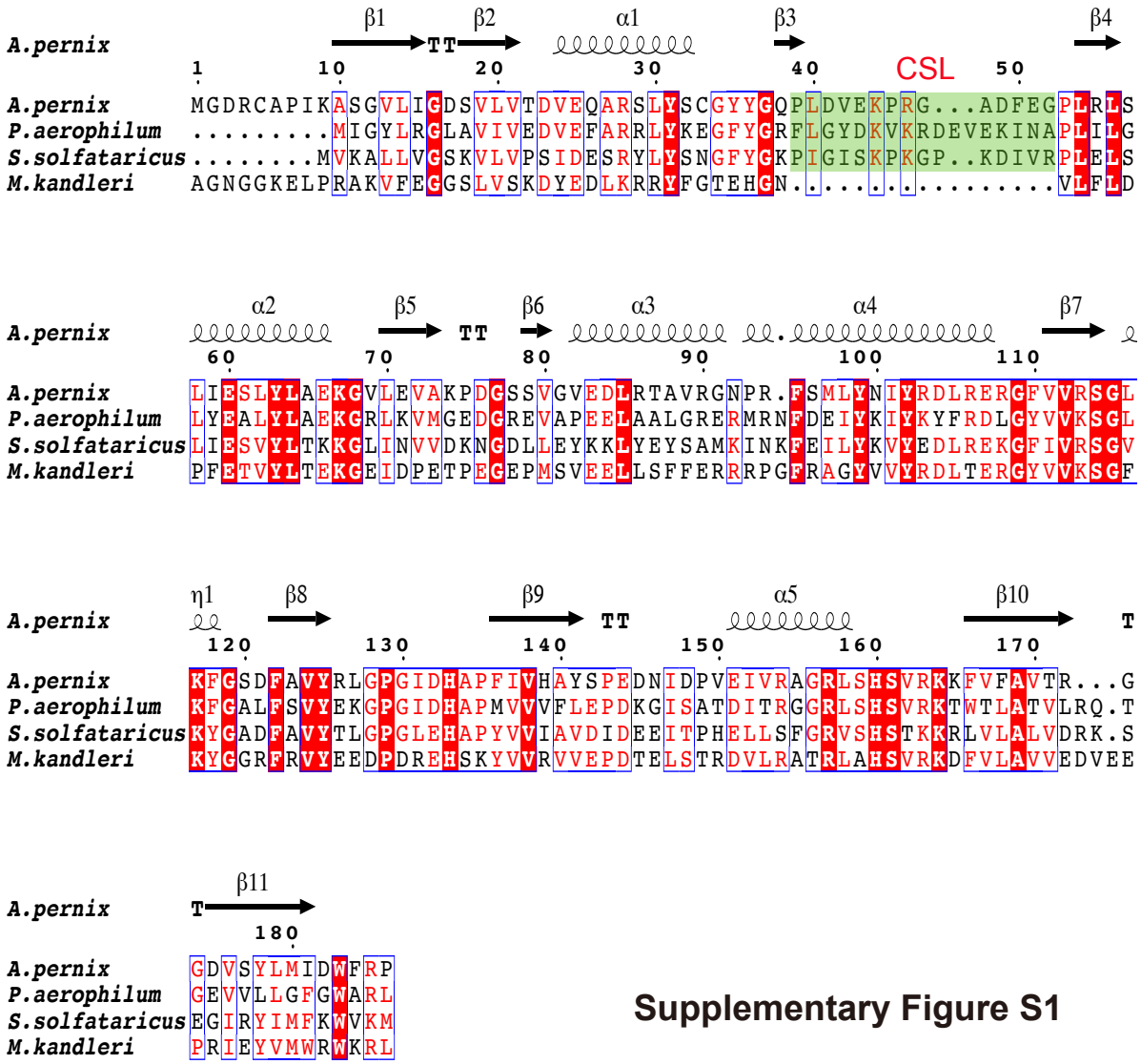
Graduate School of Science and Engineering, Ehime University,

3 Bunkyo-cho, Matsuyama, Ehime 790-8577, Japan.

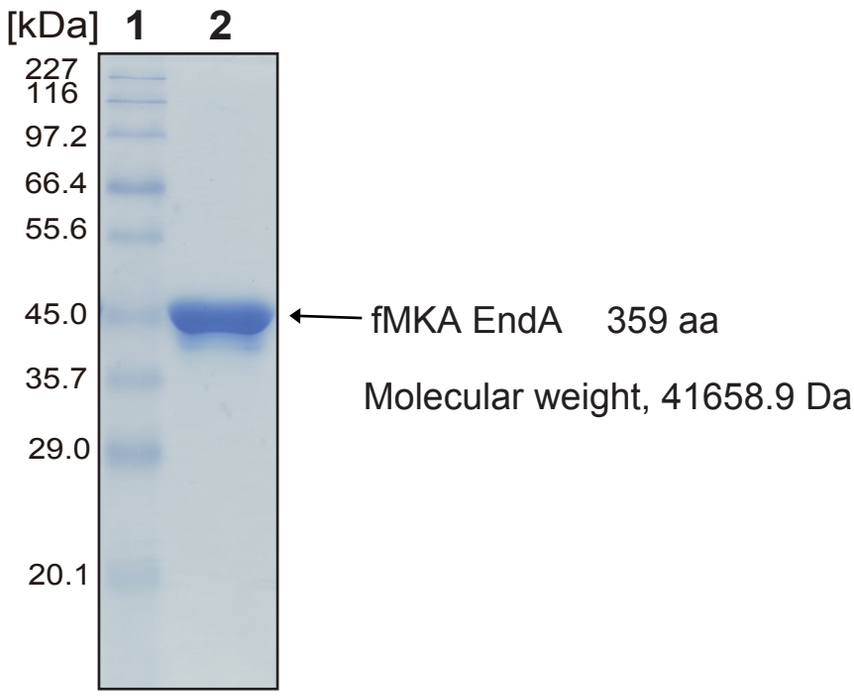
Phone: 81-89-927-9919

Fax: 81-89-927-9919

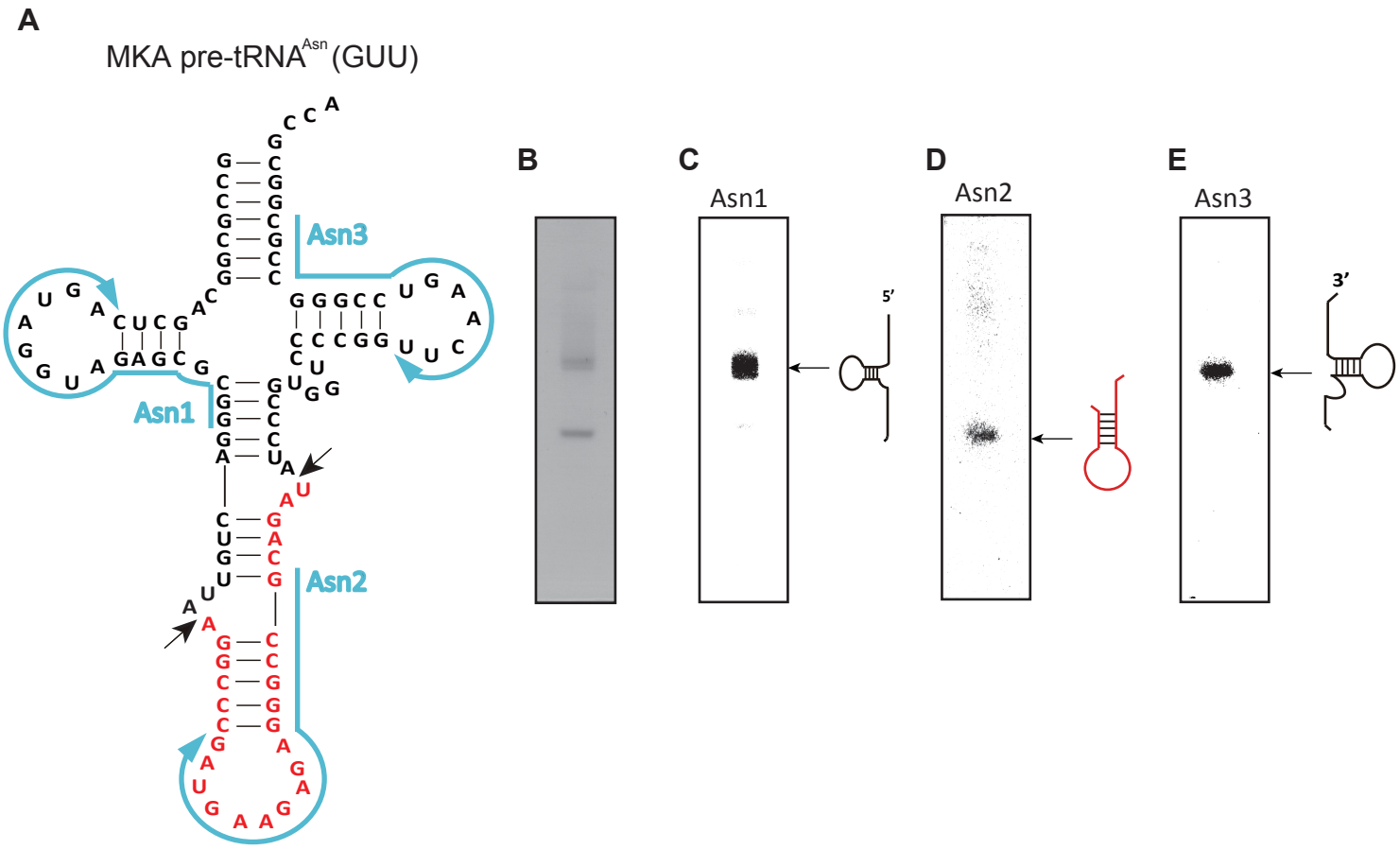
E-mail: ahirata@ehime-u.ac.jp



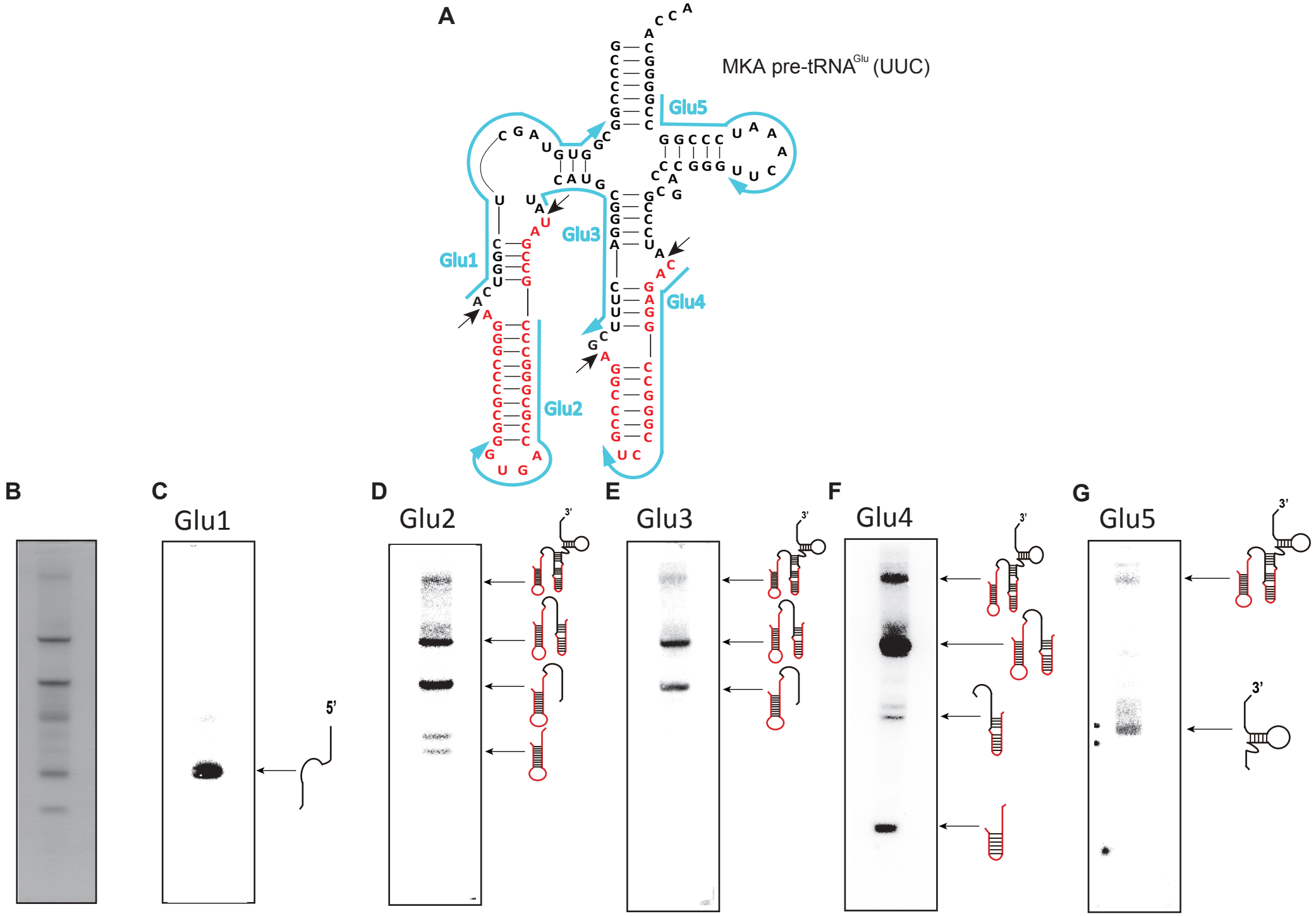
Supplementary Figure S1



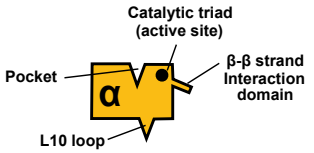
Supplementary Figure S2



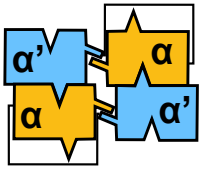
Supplementary Figure S3

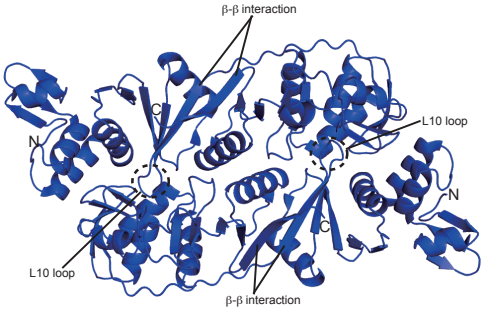


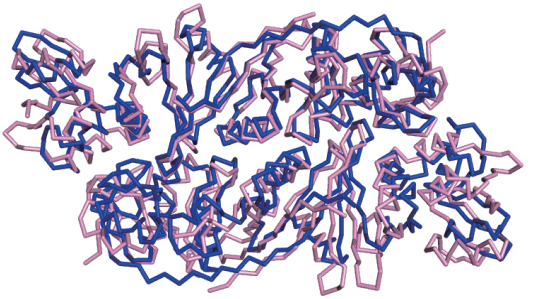
Supplementary Figure S4



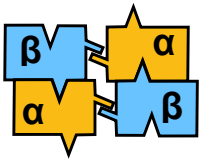
A α_2 (AFU EndA)

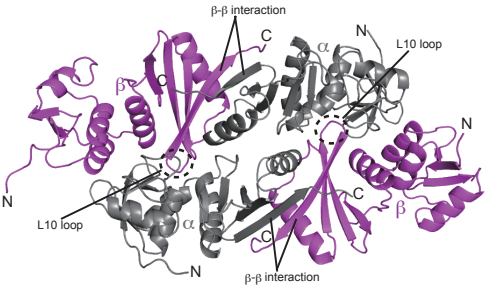


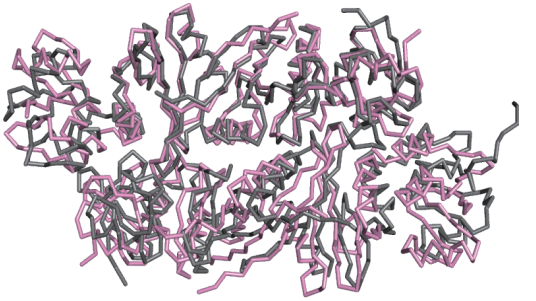




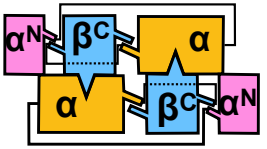
B $(\alpha\beta)_2$ (APE EndA)

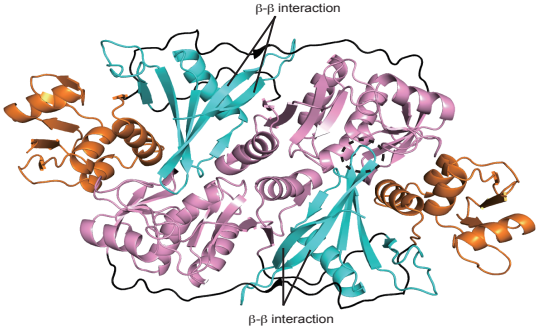


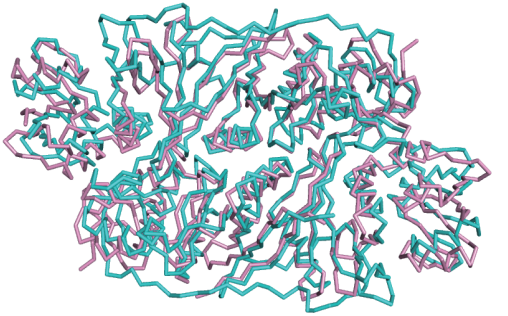




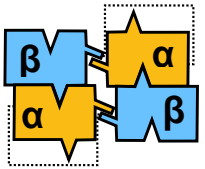
C ε_2 (ARMAN-2 EndA)

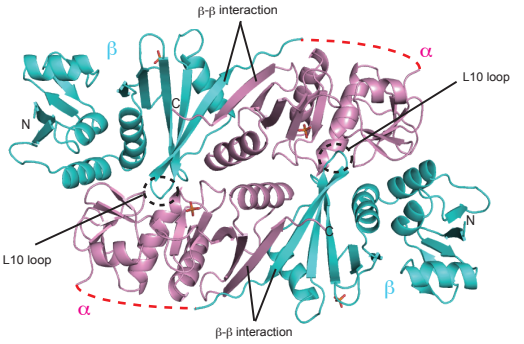




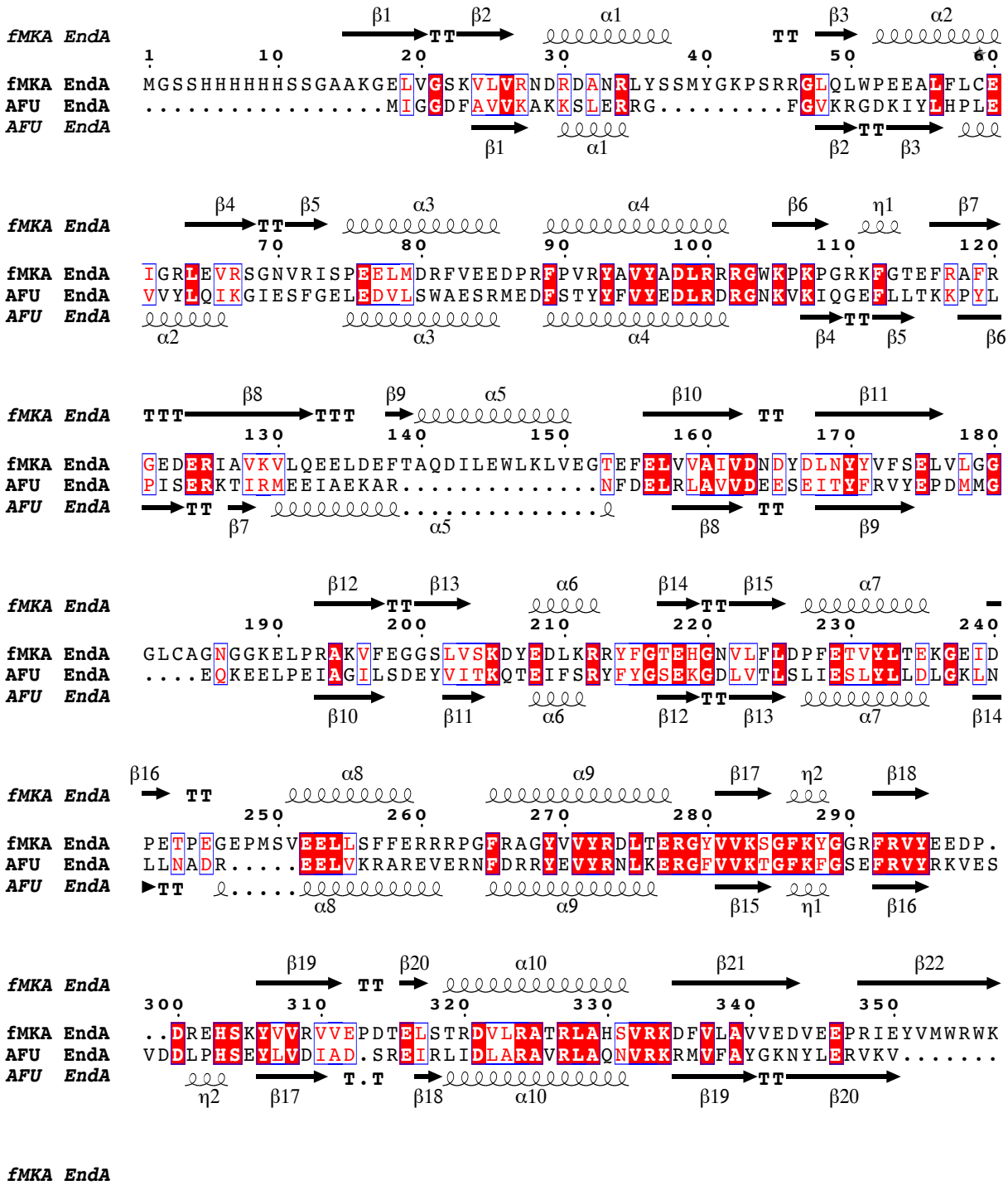


D $(\alpha\beta)_2$ (fMKA EndA)

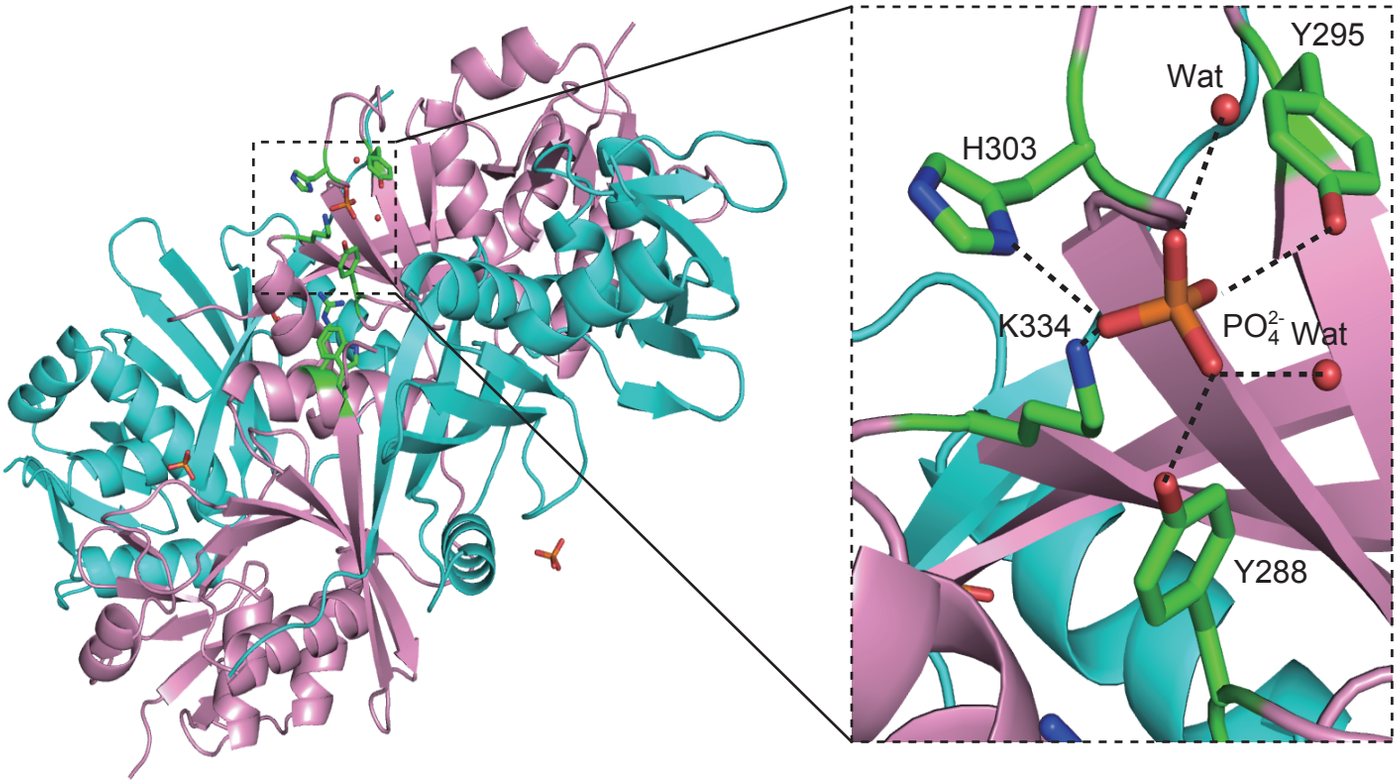




Supplementary Figure S5

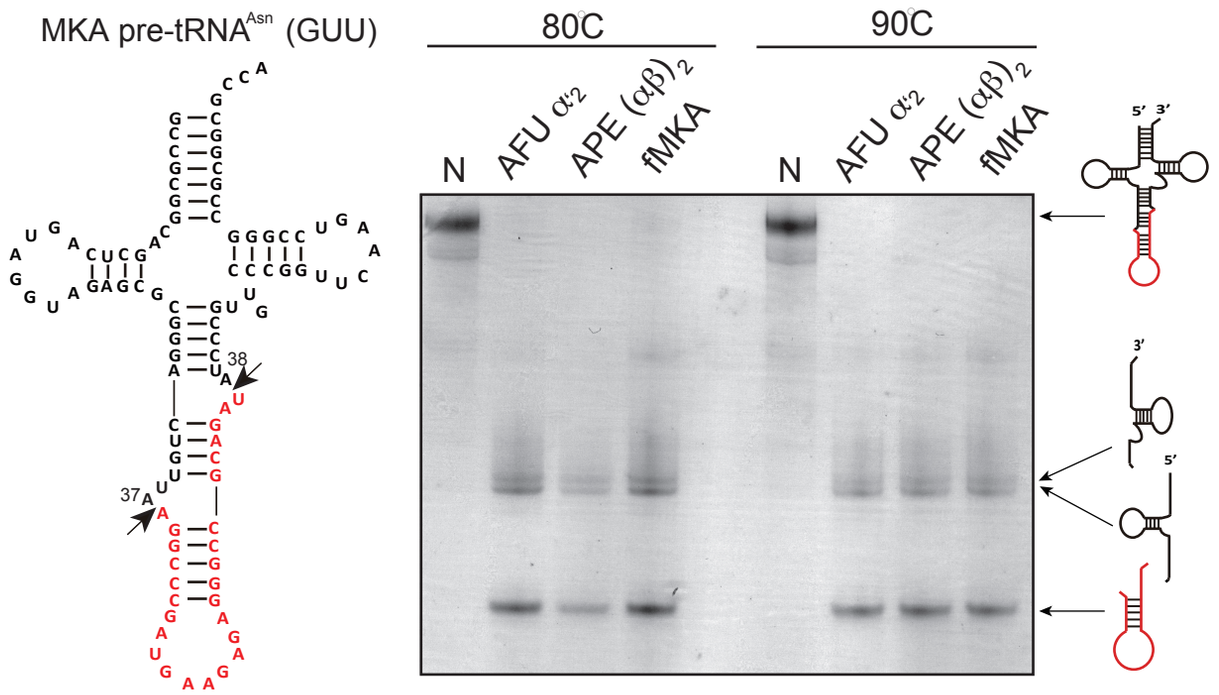


Supplementary Figure S6

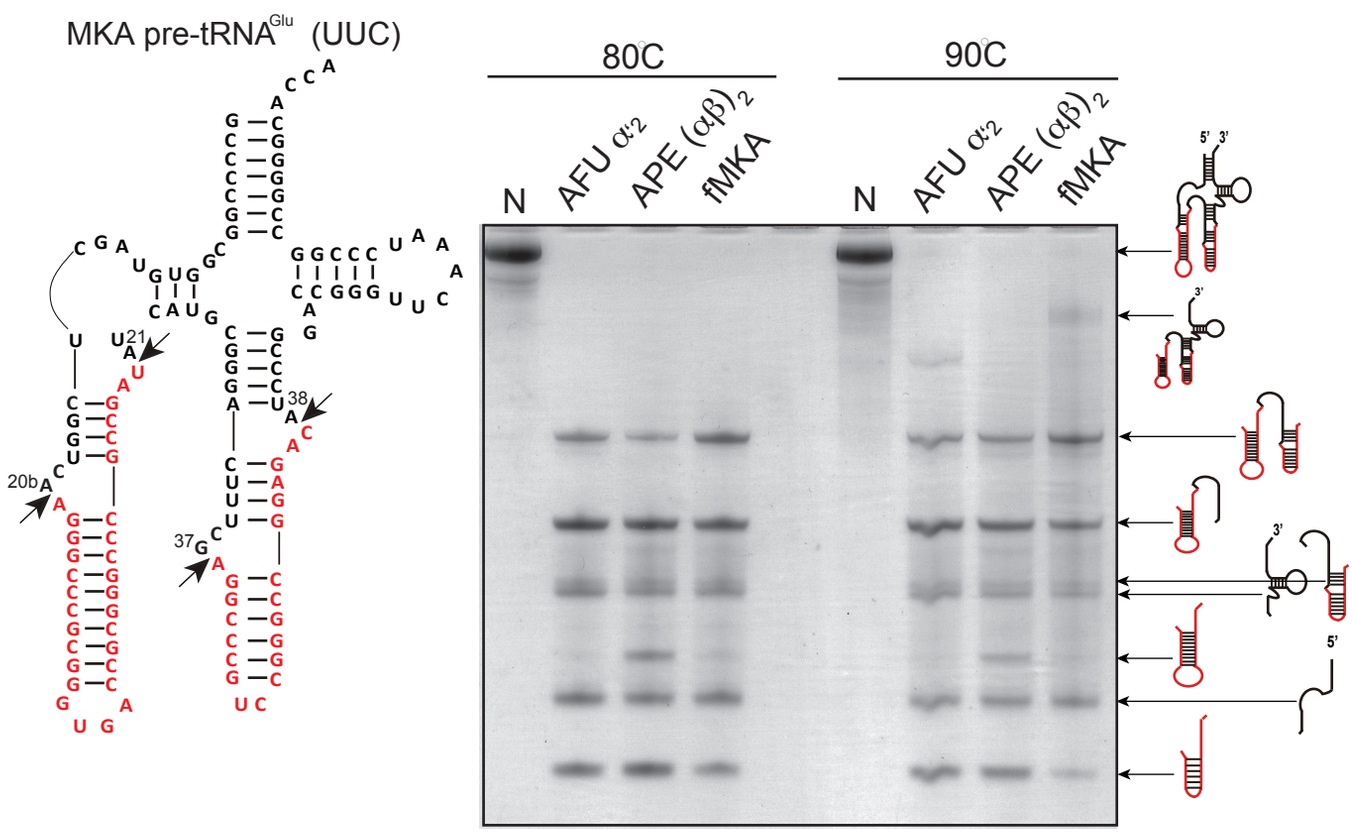


Supplementary Figure S7

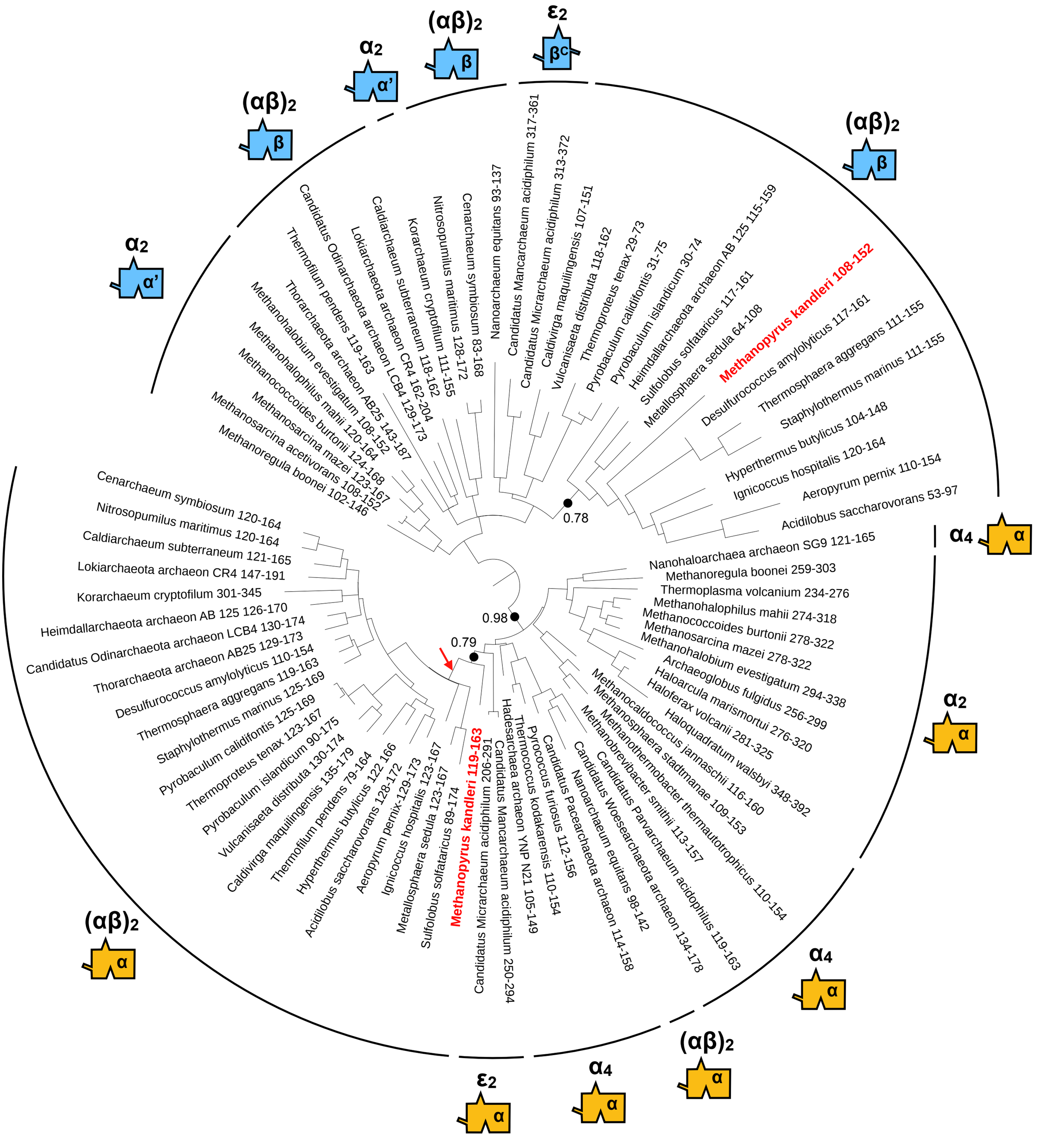
A



B



Supplementary Figure S8



Supplementary Figure S9

Supplementary Figure S1. Structural information-based amino acid sequence alignments of α subunit in EndAs from three Crenarchaeota (*Aeropyrum pernix*, *Pyrobaculum aerophilum* and *Sulfolobus solfataricus*) and one Euryarchaeota (*Methanopyrus kandleri*). The secondary structure of α subunit in *A. pernix* EndA are shown in top location along the alignment with high-level similarity. The amino acid sequences of Crenarchaea specific loop (CSL) conserved in Crenarchaeal EndAs are highlighted by green color. Identical and similar residues are boxed. This figure was created using ClustalW (48) and ESPript (49).

Supplementary Figure S2. SDS-PAGE (12.5%) of fMKA EndA purified by the two steps column chromatography works (see Methods). Lane 1: Molecular-weight marker; Lane 2: The numbers on the left side of gel indicate molecular weight (kDa). The arrow on the left side of gel shows the main band of purified fMKA EndA with its number of amino acid residues and calculated molecular weight (Da).

Supplementary Figure S3. Northern blot analysis of RNA fragments cleaved by fMKA EndA in MKA pre-tRNA^{Asn} (GUU). (**A**) Predicted secondary structure of MKA pre-tRNA^{Asn} (GUU) containing an intron (red) with BHB motif at the canonical position 37/38. Two black arrows indicate cleavage sites by fMKA EndA. The three long arrows (cyan) along with sequences of MKA pre-tRNA^{Asn} (GUU) show complementary DNA probes (Asn1, Asn2 and Asn3) used for Northern hybridization. (**B**) The cleaved products were separated by 15% PAGE/7 M urea. (**C**) Autoradiogram of the cleaved products containing 5´exon detected with Asn1. (**D**) Autoradiogram of the cleaved products containing an intron detected with Asn2. (**E**) Autoradiogram of the cleaved products containing 3´exon detected with Asn3. The detected cleavage products are shown using schematic models at the right-hand side of the gel.

Supplementary Figure S4. Northern blot analysis of RNA fragments cleaved by

fMKA EndA in MKA pre-tRNA^{Glu} (UUC). (**A**) Predicted secondary structure of MKA pre-tRNA^{Glu} (UUC) containing double introns (red) with the HBh´ and BHB motifs at the non-canonical position 20b/21 and canonical position 37/38. Four black arrows indicate cleavage sites by fMKA EndA. The five long arrows (cyan) along with sequences of MKA pre-tRNA^{Glu} (UUC) show complementary DNA probes (Glu1, Glu2, Glu3, Glu4 and Glu5) used for Northern hybridization. (**B**) The cleaved products were separated by 15% PAGE/7 M urea. (**C**) Autoradiogram of the cleaved products containing a 5´exon detected with Glu1. (**D**) Autoradiogram of the cleaved products containing an intron with HBh´motif detected with Glu2. (**E**) Autoradiogram of the cleaved products containing 5 ´exon detected with Glu3. (**F**) Autoradiogram of the cleaved products containing an intron with BHB motif detected with Glu4. (**G**) Autoradiogram of the cleaved products containing 3´exon detected with Glu5. The detected cleavage products are shown using schematic models at the right-hand side of the gel.

Supplementary Figure S5. Structures and characteristics of four types of archaeal EndAs (AFU EndA (**A**), APE EndA (**B**), ARMAN-2 EndA (**C**) and fMKA EndA (**D**)). The subunit interactions are represented by cartoon models on the left side. The β–β interaction responsible for inter/intraunit formation, the L10 loop and pocket responsible for dimer/tetramer formation are highlighted. The catalytic triads are marked by green circles. The center panels show the ribbon models of EndAs. The superimposed structures (obtaining by aligning respective Cα atoms) of respective AFU EndA (navy), APE EndA (gray) and ARMAN-2 EndA (cyan) onto that of fMKA EndA (pink) are represented on the right side. The root mean square deviations (r.m.s.d) of AFU EndA-fMKA EndA, APE EndA-fMKA EndA and ARMAN-2 EndA-fMKA EndA were 2.2 Å (using 294 Cα atoms), 1.6 Å (using 161 Cα atoms) and 2.5 Å (using 179 Cα atoms), respectively.

Supplementary Figure S6. Structural information-based amino acid sequence alignments of fMKA EndA and AFU EndA. The location of secondary structure of

fMKA EndA is shown in parallel with that of AFU EndA with a high-level similarity. Identical and similar residues are boxed. This figure was created using ClustalW (48) and ESPript (49).

Supplementary Figure S7. Close-up view of the binding site of phosphate ion. The candidates for three catalytic residues (Y295, H303 and K334) are shown by green stick model in active site. The phosphate ion (PO_4^{2-}) and water (Wat) are depicted as stick (brown) and ball (red) models, respectively. The dotted lines show the hydrogen bonds.

Supplementary Figure S8. Splicing activity and specificity of AFU EndA, APE EndA, and fMKA EndA. (A) *Left*, predicted secondary structure of MKA pre-tRNA^{Asn} (GUU) containing the intron with the BHB motif at canonical position 37/38 (red). The two arrows show the cleavage sites. *Right*, splicing reaction of the MKA pre-tRNA^{Asn} (GUU) at 80°C and 90°C for 5 min. (B) *Left*, predicted secondary structure of MKA pre-tRNA^{Glu} (UUC) containing double introns consisting of a non-canonical HBh´ intron at 20b/21 and a canonical intron at 37/38 (red). The two arrows show the cleavage sites. *Right*, splicing reaction of the MKA pre-tRNA^{Glu} (UUC) at 80°C and 90°C for 5 min. Reaction mixtures were separated on 15% polyacrylamide/7 M urea gels. In each gel, “N” (first lane) indicates the control (no enzyme). The cleavage products are shown using schematic models at the right-hand side of the gel.

Supplementary Figure S9. Maximum likelihood tree of archaeal EndA proteins. Phylogenetic relationship of total 81 archaeal EndA α and β subunits from α_4, $(\alpha\beta)_2$, ε_2, and α´ unit from α'_2 are represented as an unrooted tree. aLRT SH-like branch supports were computed for the catalytic α unit/subunits (0.98) and groups including the MKA EndA α and β subunits (0.79 and 0.78, respectively), indicated by black circle and assigned values. MKA EndA subunits are highlighted in red, and the red arrow indicates the possible insertion timing of L1 loop.

References

(48) Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T.J., Higgins, D.G. and Thompson, J.D. (2003). Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Res* **31**, 3497-500.

(49) Gouet, P., Robert, X. and Courcelle, E. (2003). ESPript/ENDscript: Extracting and rendering sequence and 3D information from atomic structures of proteins. *Nucleic Acids Res* **31**, 3320-3.

Supplementary Table S1.

Primers list used for construction of the expression plasmid pFMKAEndA.

MKA_aF

Forward

5′-GGAGATATA**CATATG**TTGTGCGCCGGGAACGGG-3′

MKA_aR

Reverse

5′-CTCGAATTC**GGATCC**TTACAGCCTCTTCCACCGCCACATC-3′

MKA_bF

Forward

5′- GGAGATATA**CATATG**GCGGCGAAAGGCGAGTTAGTC -3′

MKA_bR

Reverse

5′-CTCGAATTC**GGATCCC**TAGAGAACTAGTTCTGAAAACACATAGTAAT TCAAGTC-3′

Del_F

Forward

5-′AGCAGCGGCGCGGCGAAAGGCGAGT-3′

Del_R

Reverse

5′-GCTGCTGCCCATGGTATATCTCCTTCTTAAAG -3′

GlyF

Forward

5′-GGCGGTGGCTTGGTGCGCCGGGAACGGGG -3′

GlyR

Reverse

5′-GAGAACTAGTTCTGAAAACACATAGTAATTCAAGTCA-3′

The bold show a restriction sites. Mutational sites colored red.