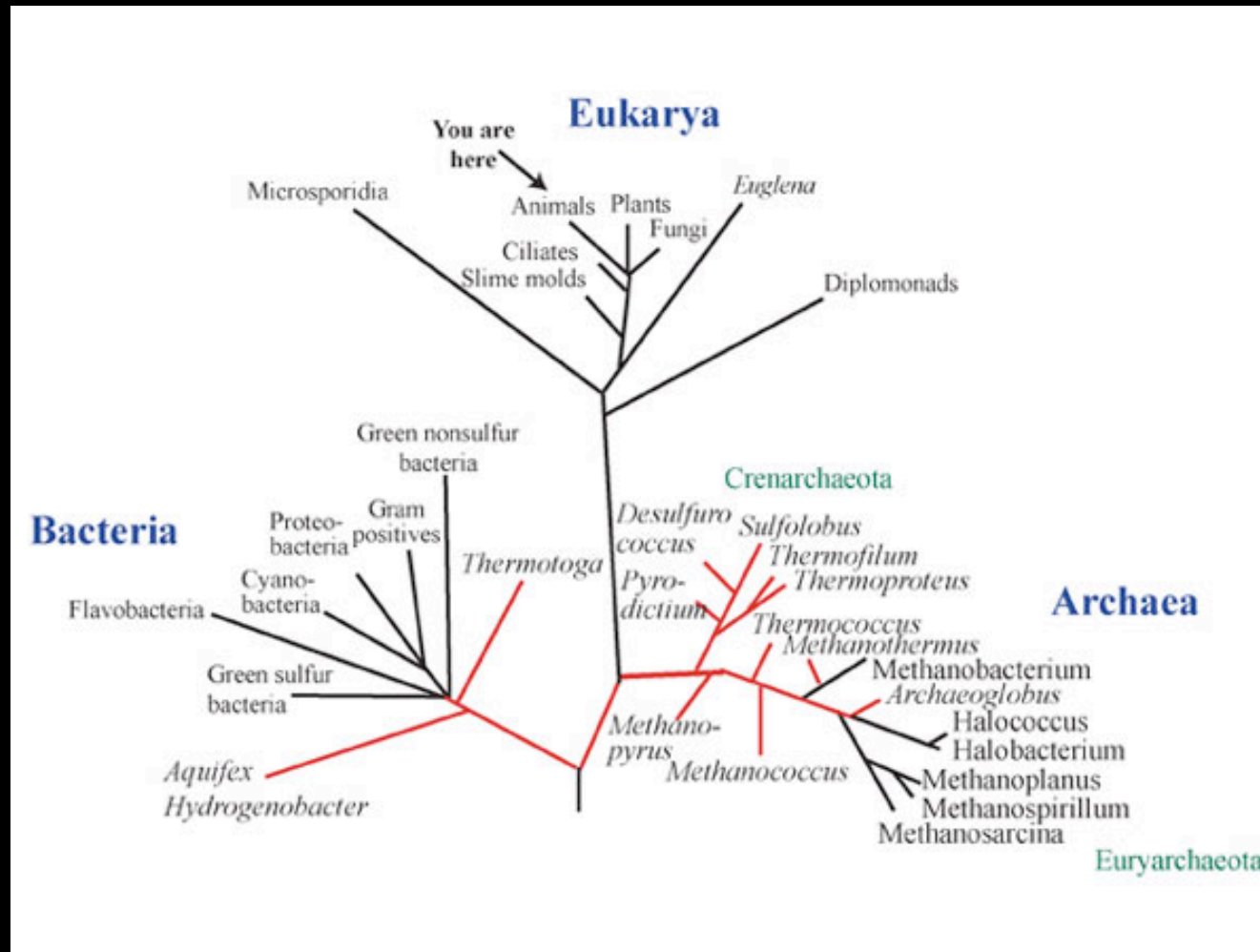


System Biology of Archaeal RNAs and RNA-Binding Proteins

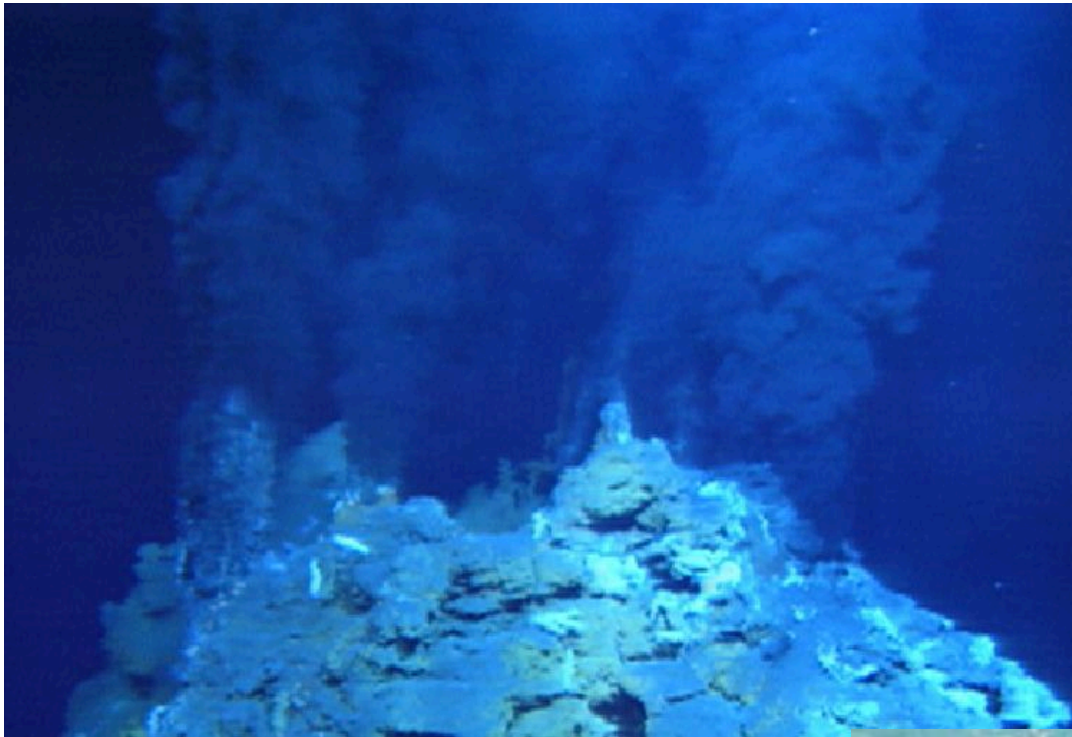
Akio Kanai and Junichi Sugahara

Institute for Advanced Biosciences, Keio University, Japan

Universal phylogenetic tree



Those lines highlighted in red lead to organisms that are heat-loving.



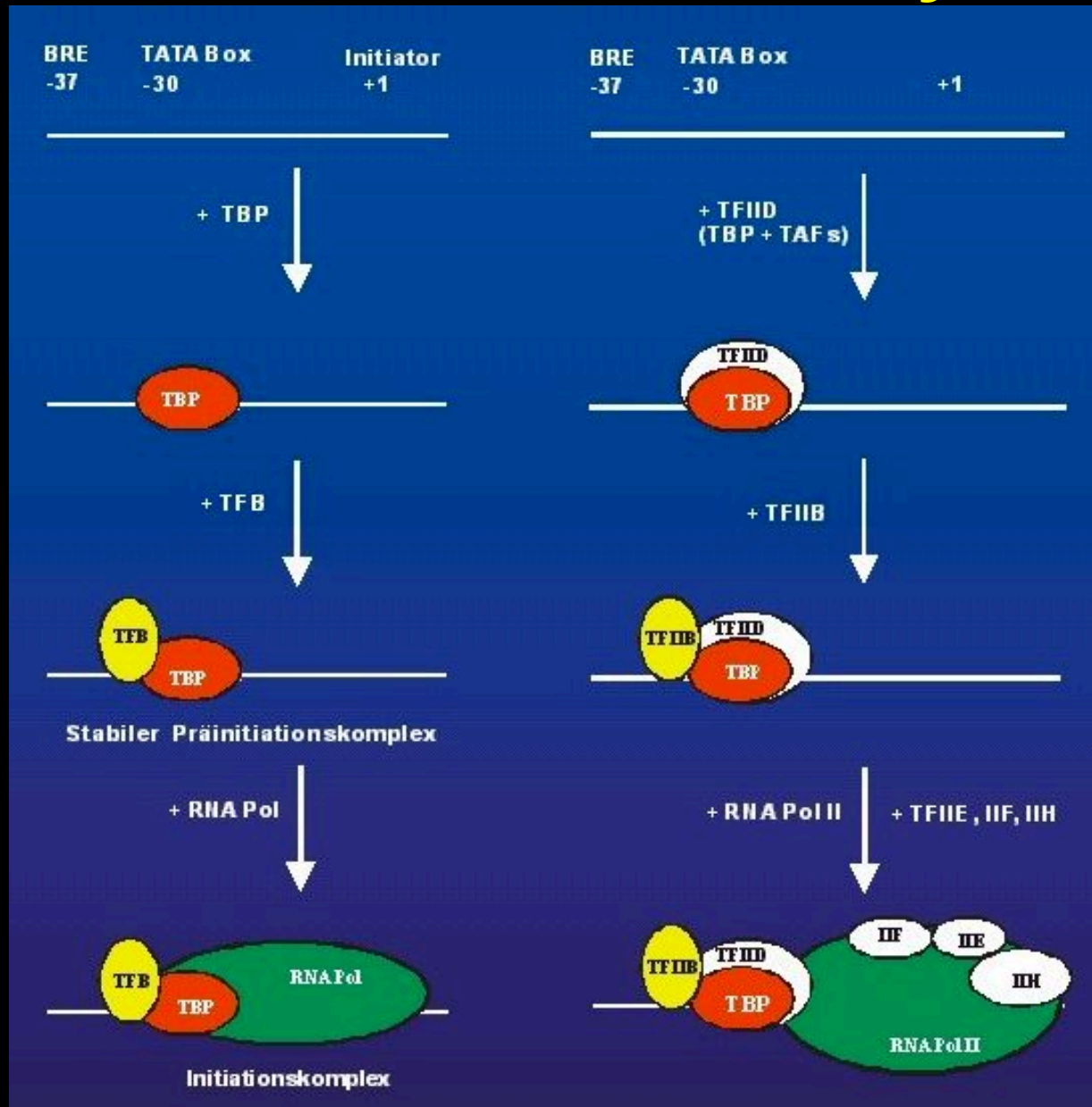
**Beginning
of life ?**

**Deep-sea
vents**



Archaea

Eukarya



Pyrococcus furiosus

Hyperthermophilic archaeon

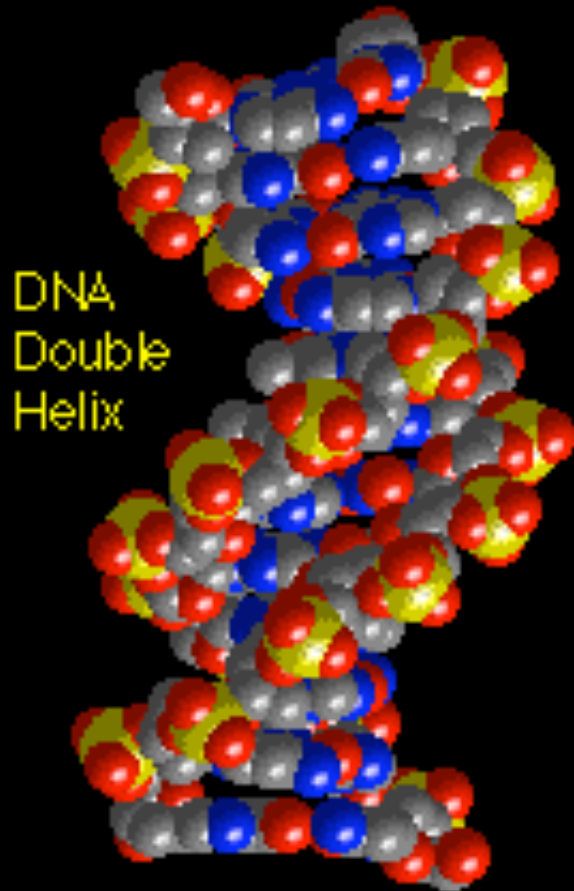
- Grows in a deep sea around at 100 °C
- Heatstable proteins

Complete genome sequence

- Genome size: 1908 kb
- # of ORF: 2065



Central Dogma



DNA



RNA



Protein

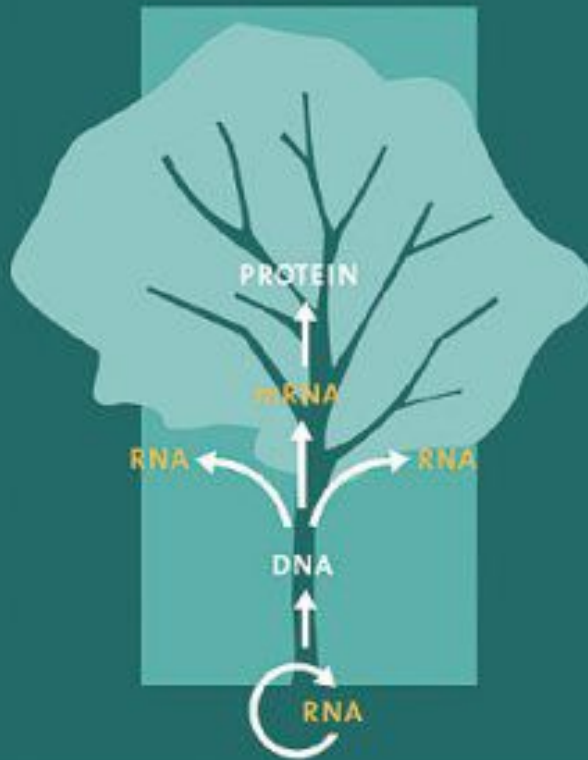
**Why are RNAs are
so important?**

Why are RNAs are so important?

- ◎ RNA Splicing, RNA processing, RNA editing
- ◎ So many non-coding RNAs
- ◎ Ribozymes
- ◎ Anti-sense RNAs
- ◎ RNAi
- ◎ Aptamer
- ◎ Coenzymes (NAD, FAD)
- ◎ ATP, GTP
- ◎ RNA primers for DNA replication
- ◎ RNA localization & Development

The RNA World

THIRD EDITION



EDITED BY Raymond F. Gesteland
Thomas R. Cech
John F. Atkins

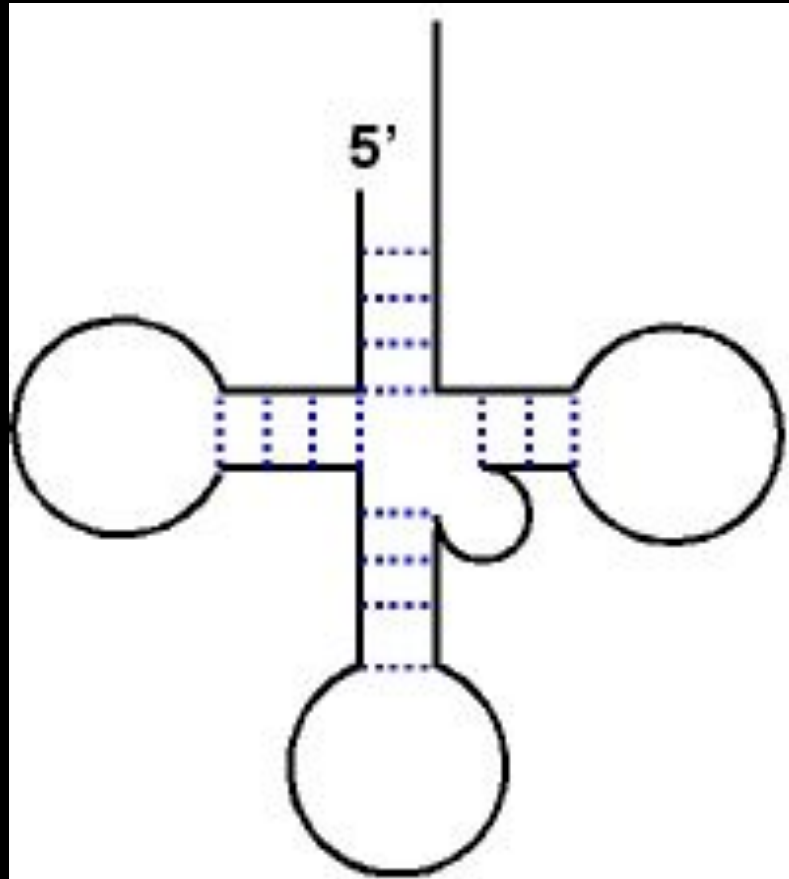
**Cold Spring Harbor
Laboratory Press
(2006)**

© RNAs
(non-coding RNA)

© RNA-binding protein
(enzymes)

tRNA

Used as a
primer for
DNA synthesis
(retrovirus)



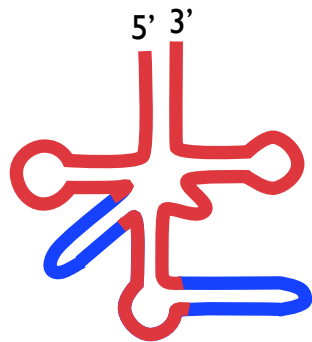
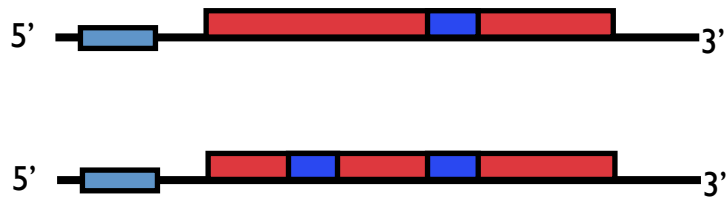
Involved in
Translation

RNA molecule

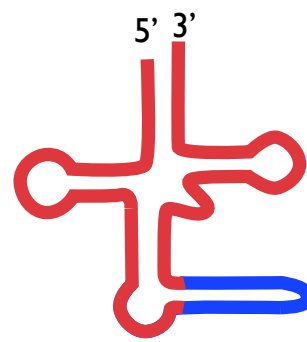
Unique type of tRNAs found in the genome of archaeal species

Intron-containing tRNA

Promoter



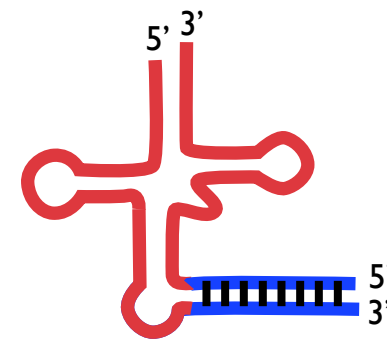
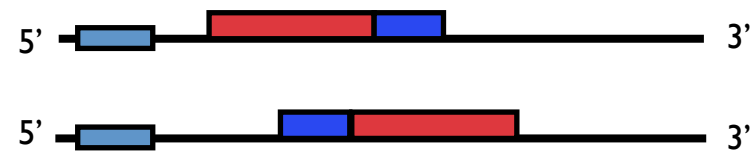
Multiple intron type



Single intron type

Split tRNA

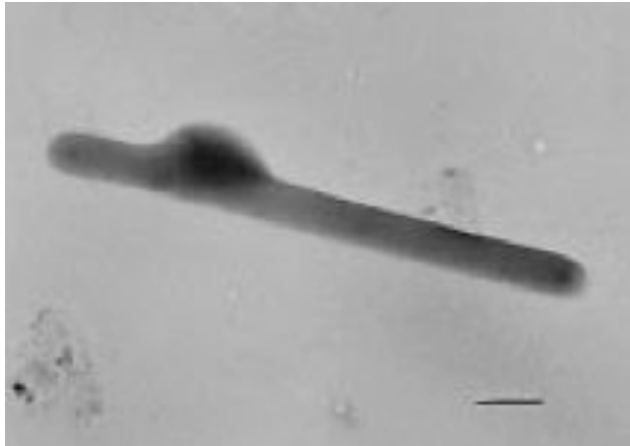
Promoter



→ **SPLITSdb system: A comprehensive prediction of archaeal tRNAs**

Sugahara, J., Kikuta, K., Fujishima, K., Yachie, N., Tomita, M. and Kanai, A.
(2008) *Molecular Biology and Evolution* 25(12): 2709-2716.

Several tRNA genes are **missing** in the Hyperthermoacidophilic archaeon *Caldivirga maquilingensis*



Itoh et al. *Int J Syst Bacteriol* (1999)

Temperature 85°C

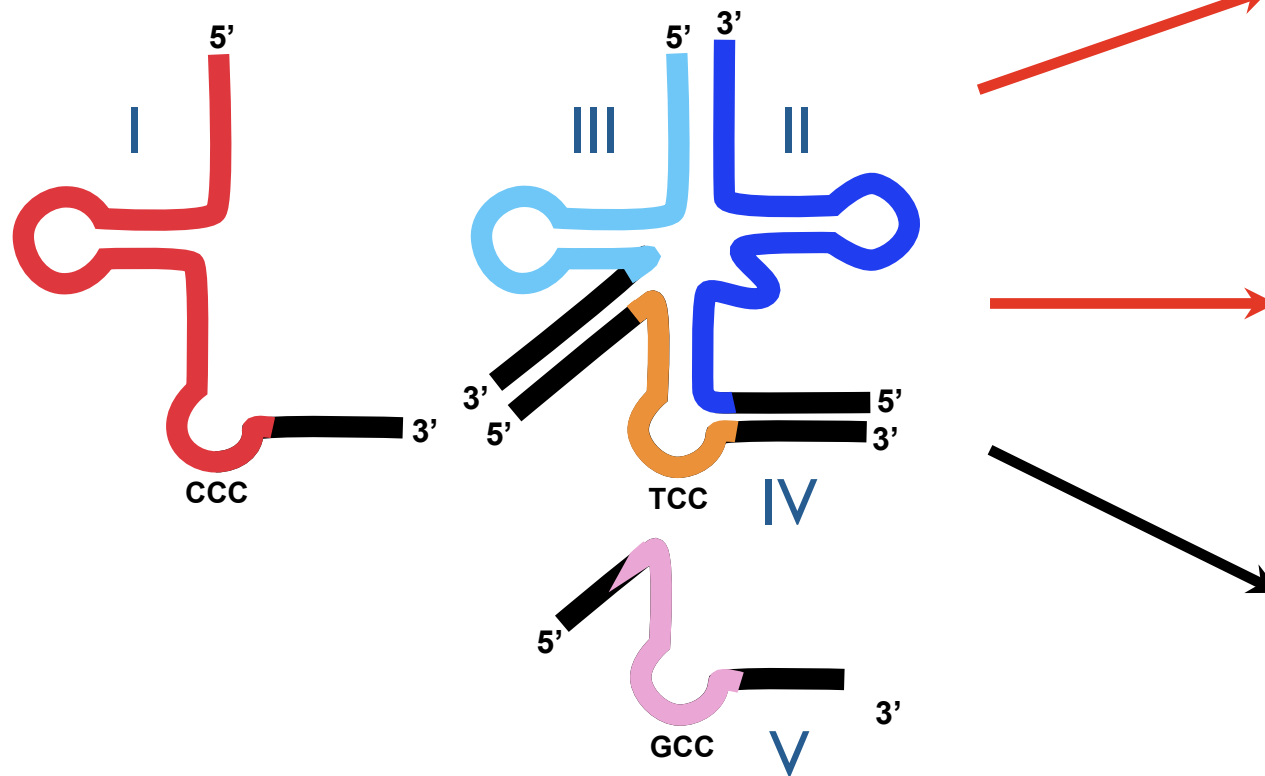
pH 3.7 - 4.2

Genome size ~2Mb

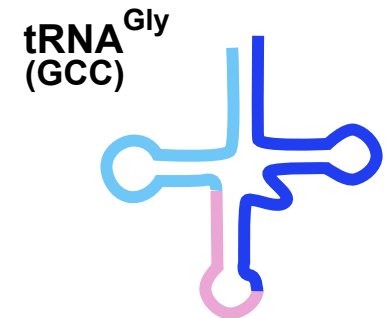
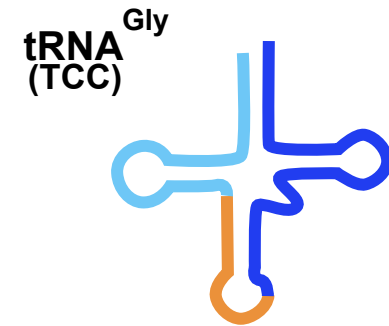
2067 genes

		Second base of codon							
		T		C		A		G	
First base of codon									
	T	Phe (TTT)	1	Ser (TCT)	1	Tyr (TAT)	1	Cys (TGT)	1
Phe (TTC)		1	Ser (TCC)	1	Tyr (TAC)	1	Cys (TGC)	1	C
Leu (TTA)		1	Ser (TCA)	1	End (TAA)	Stop	End (TGA)	Stop	A
Leu (TTG)		1	Ser (TCG)	1	End (TAG)	Stop	Trp (TGG)	1	G
C	Leu (CTT)	1	Pro (CCT)	1	His (CAT)	1	Arg (CGT)	1	T
	Leu (CTC)	1	Pro (CCC)	1	His (CAC)	1	Arg (CGC)	1	C
	Leu (CTA)	1	Pro (CCA)	1	Gln (CAA)	1	Arg (CGA)	1	A
	Leu (CTG)	1	Pro (CCG)	1	Gln (CAG)	1	Arg (CGG)	1	G
A	Ile (ATT)	1	Thr (ACT)	1	Asn (AAT)	1	Ser (AGT)	1	T
	Ile (ATC)	1	Thr (ACC)	1	Asn (AAC)	1	Ser (AGC)	1	C
	Ile (ATA)	1	Thr (ACA)	1	Lys (AAA)	1	Arg (AGA)	1	A
	Ile (ATG)	1	Thr (ACG)	1	Lys (AAG)	1	Arg (AGG)	1	G
G	Val (GTT)	1	Ala (GCT)	1	Asp (GAT)	1	Gly (GGT)	1	T
	Val (GTC)	1	Ala (GCC)	1	Asp (GAC)	1	Gly (GGC)	0	C
	Val (GTA)	1	Ala (GCA)	0	Glu (GAA)	1	Gly (GGA)	0	A
	Val (GTG)	1	Ala (GCG)	0	Glu (GAG)	0	Gly (GGG)	0	G

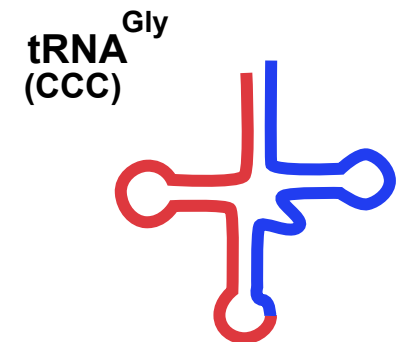
Tri-split tRNA is a transfer RNA made from three transcripts



Tri-split tRNA



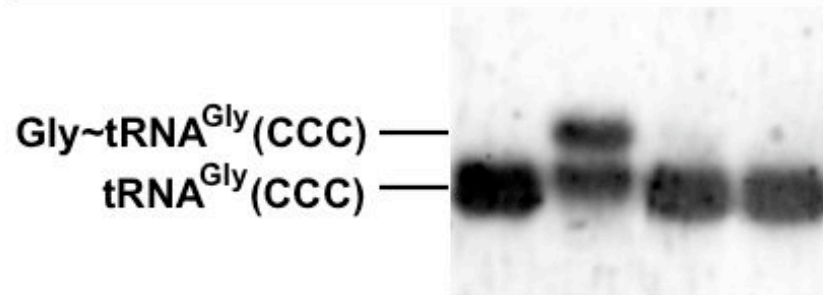
Split tRNA



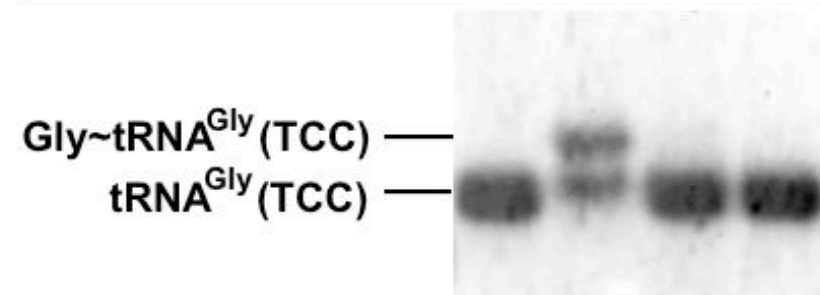
Fujishima, K., Sugahara, J., Kikuta, K., Hirano, R., Sato, A., Tomita, M. and Kanai, A. (2009) *Proceedings of the National Academy of Sciences. U. S. A.* 106(8): 2683-2687.

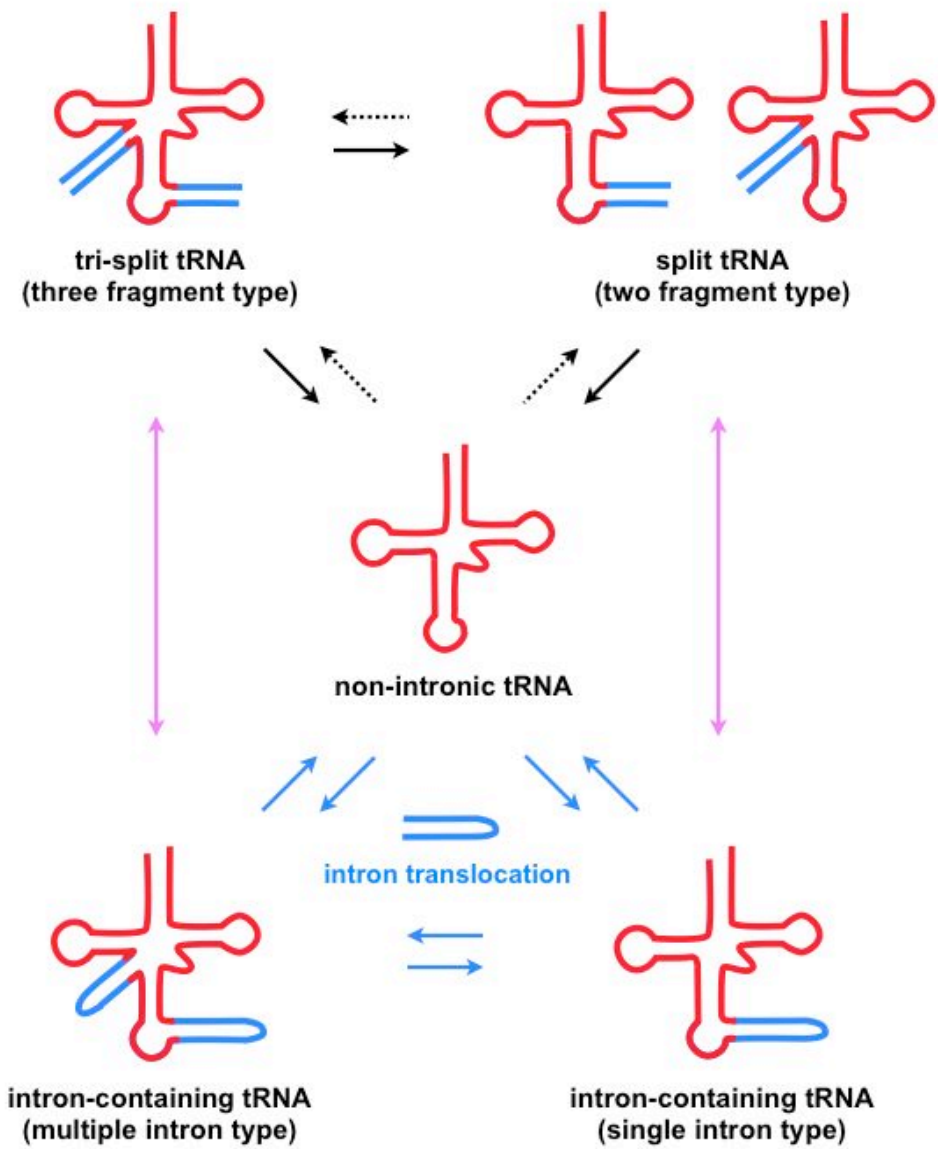
In vitro Aminoacylation assay of mature tRNA (Glys)

GlyRS		+	+	+
Glycine	+	+		
Alanine			+	



GlyRS		+	+	+
Glycine	+	+		
Alanine			+	





Sequence Evidence in the Archaeal Genomes that tRNAs Emerged Through the Combination of Ancestral Genes as 5' and 3' tRNA Halves

Kosuke Fujishima^{1,2}, Junichi Sugahara^{1,2}, Masaru Tomita^{1,2,3}, Akio Kanai^{1,2,3*}

1 Institute for Advanced Biosciences, Keio University, Tsuruoka, Yamagata, Japan, **2** Systems Biology Program, Graduate School of Media and Governance, Keio University, Fujisawa, Japan, **3** Department of Environment and Information, Keio University, Fujisawa, Japan

February 2008, Vol. 3 Issue 2, e1622

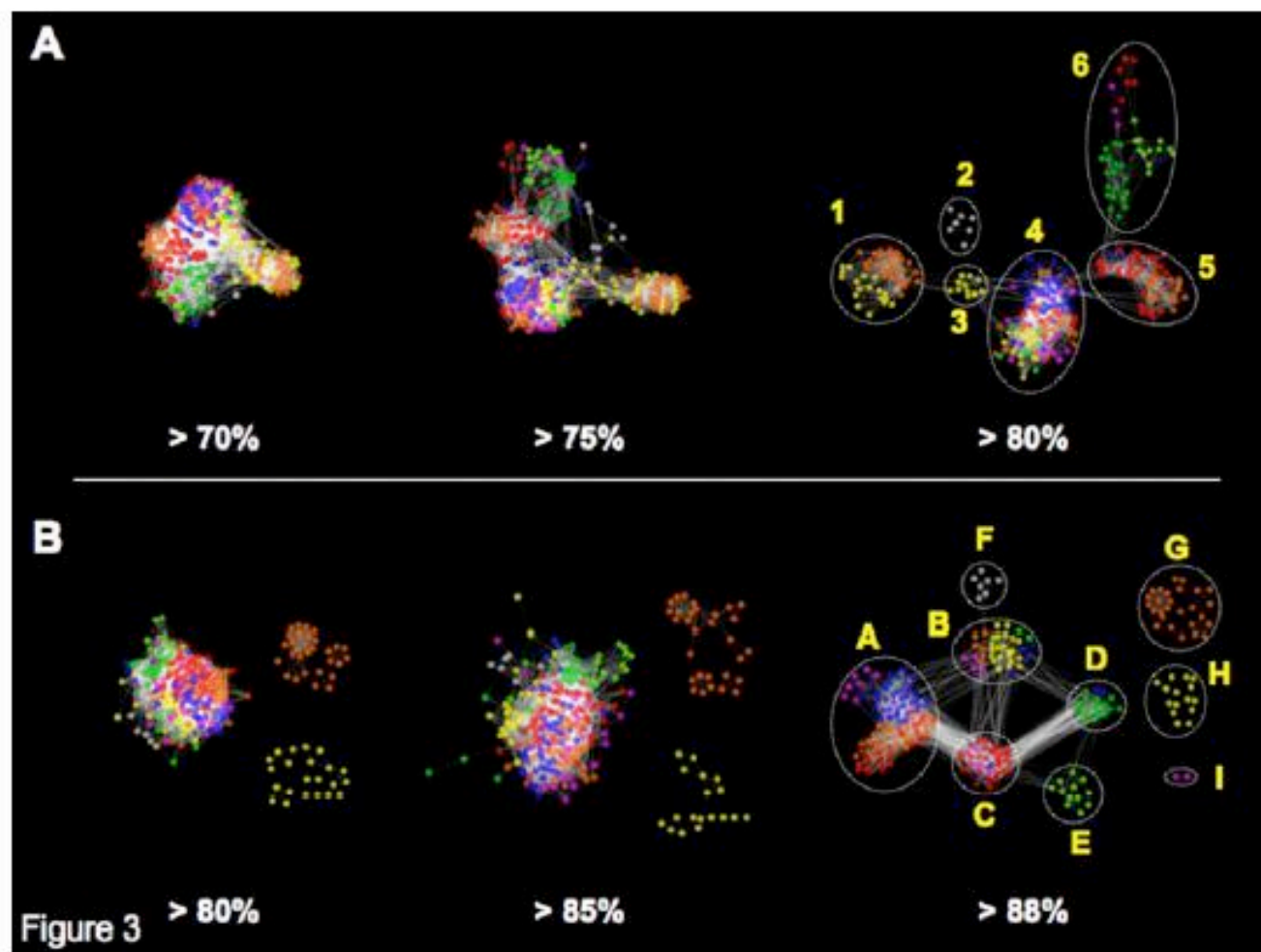


Figure 3. Network analysis based on the sequence similarities of 5' and 3' tRNA halves. A total of 296 mature tRNA sequences from 7 archaeal species (Neq, Sso, Ape, Pae, Mka, Pfu, Mja) were artificially split into 5' and 3' halves at the anti-codon region. Each node (colored dot) represents a tRNA half, and its color indicates the charged amino acid's chemical properties (DE, mid-green; MNQ, light green; RKH, blue; FWY, purple; AGP, red; ILV, orange; CST, yellow; iMet, gray). Nodes are linked by a white line (edge) when the sequence similarity is above the threshold. (A) Network created by set of 5' half sequences with thresholds of >70%, >75%, and >80%. The sequences are classified into 6 clusters (1–6) at a threshold of >80%. (B) Network created by set of 3' half sequences with thresholds of >80%, >85%, and >88%. The sequences are classified into 9 clusters (A–I) at a threshold of >88%.
doi:10.1371/journal.pone.0001622.g003

		Second base of codon												
		T			C			A			G			
		5'	3'		5'	3'		5'	3'		5'	3'		
First base of codon	T	Phe (TTT)			Ser (TCT)			Tyr (TAT)			Cys (TGT)			T
		Phe (TTC)	4	A	Ser (TCC)	1	H	Tyr (TAC)	4	A	Cys (TGC)	1	B	C
		Leu (TTA)	1	G	Ser (TCA)	1	H	End (TAA)			End (TGA)			A
		Leu (TTG)	1	G	Ser (TCG)	1	H	End (TAG)			Trp (TGG)	4	A	G
	C	Leu (CTT)			Pro (CCT)			His (CAT)			Arg (CGT)			T
		Leu (CTC)	1	G	Pro (CCC)	4	A	His (CAC)	4	B	Arg (CGC)	4	A	C
		Leu (CTA)	1	G	Pro (CCA)	5	A	Gln (CAA)	6	E	Arg (CGA)	4	A	A
		Leu (CTG)			Pro (CCG)	5	A	Gln (CAG)	6	E	Arg (CGG)			G
	A	Ile (ATT)			Thr (ACT)			Asn (AAT)			Ser (AGT)			T
		Ile (ATC)	4	A	Thr (ACC)	4	B	Asn (AAC)	4	*	Ser (AGC)	1	*	C
		Ile (ATA)	4	A	Thr (ACA)	4	B	Lys (AAA)	4	A	Arg (AGA)	4	A	A
		Ile eMet	4	B	Thr (ACG)	4	B	Lys (AAG)	4	A	Arg (AGG)	4	A	G
		iMet	2	F										
	G	Val (GTT)			Ala (GCT)			Asp (GAT)			Gly (GGT)			T
		Val (GTC)	5	A	Ala (GCC)	5	A	Asp (GAC)	6	D	Gly (GGC)	5	A	C
		Val (GTA)	5	A	Ala (GCA)	5	A	Glu (GAA)	6	D	Gly (GGA)	5	C	A
Val (GTG)		5	A	Ala (GCG)	5	A	Glu (GAG)	6	D	Gly (GGG)	5	C	G	

Figure 5. Representation of *N. equitans* codon table filled with 5' and 3' tRNA halves. The table is filled with the group IDs (see Fig. 3) corresponding to each of the 5' and 3' tRNA halves in *N. equitans* (Neq). The anti-codon corresponding to the 6 split tRNAs is shown in red. An asterisk indicates that a sequence does not have a similar sequence above the threshold (5' half, 80%; 3' half, 88%).
doi:10.1371/journal.pone.0001622.g005

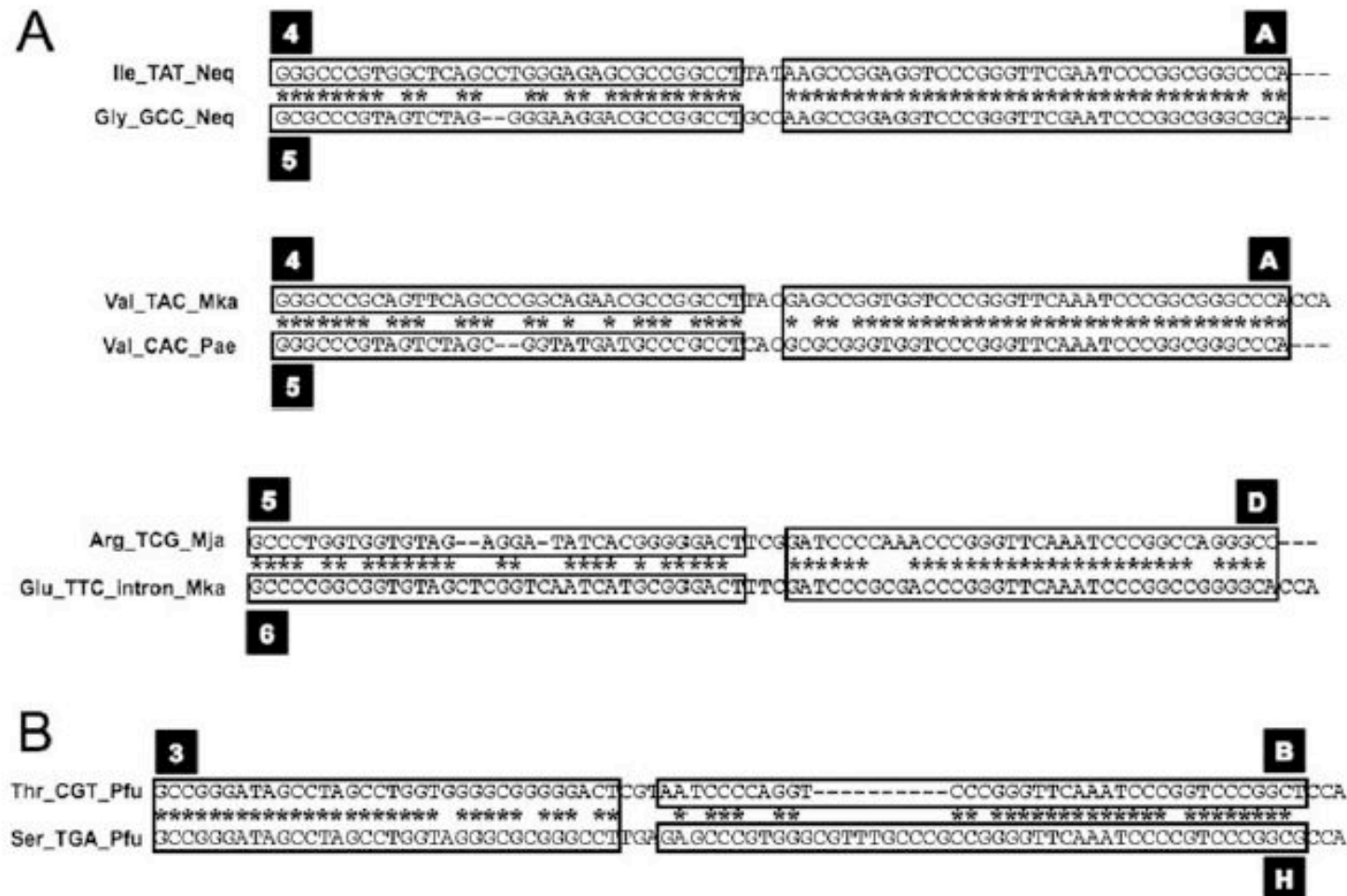


Figure 6. Examples of tRNA sequences explained by specific combinations of 5' and 3' tRNA halves. (A) Examples of tRNA sequences with different 5' half sequences but a common 3' half sequence. (B) Example of tRNA sequences with different 3' half sequences but a common 5' half sequence. Asterisk denotes matching nucleotides in the aligned tRNA sequences. Boxes delineate sequences belonging to the same group in the sequence similarity network (Fig. 3). The corresponding group ID is labeled in a black box.
doi:10.1371/journal.pone.0001622.g006

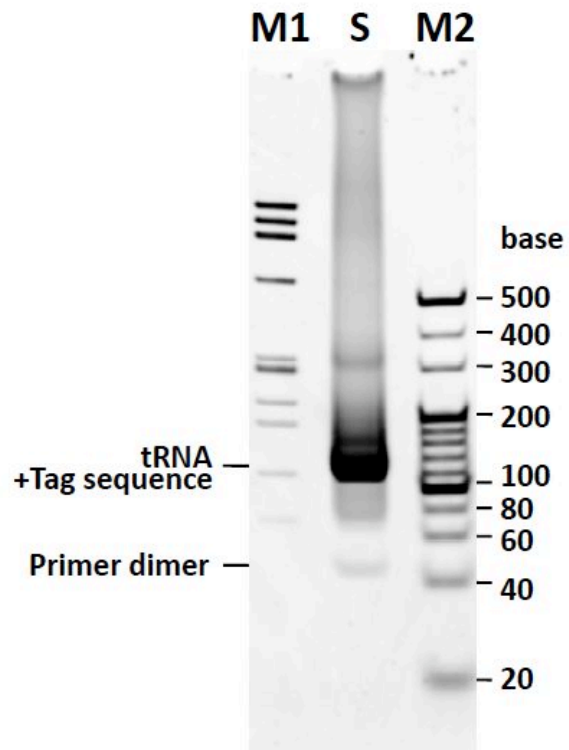
More example !

tRNA Meta Genome



Prep. of cDNA library from small RNA fraction

→ Analyzing DNA sequence
using next-generation sequencing technologies
(on going)



M1: ϕ x174/*Hae*III
M2: 20bp DNA Ladder
S : cDNA library (using small RNAs)
~400ng/lane

10-20% Polyacrylamide Gel
(Gel red staining)

Pyrococcus furiosus

Hyperthermophilic archaeon

- Grows in a deep sea around at 100 °C
- Heatstable proteins

Complete genome sequence

- Genome size: 1908 kb
- # of ORF: 2065



○ Systematic Analysis of RNA-binding Proteins

(1) Molecular Biology

Kanai, A. *et al. Biochemical J.* (2003)

Kanai, A. *et al. Biochemical J.* (2006)

Kanai, A. *et al. RNA* (2009)

(2) Bioinformatics

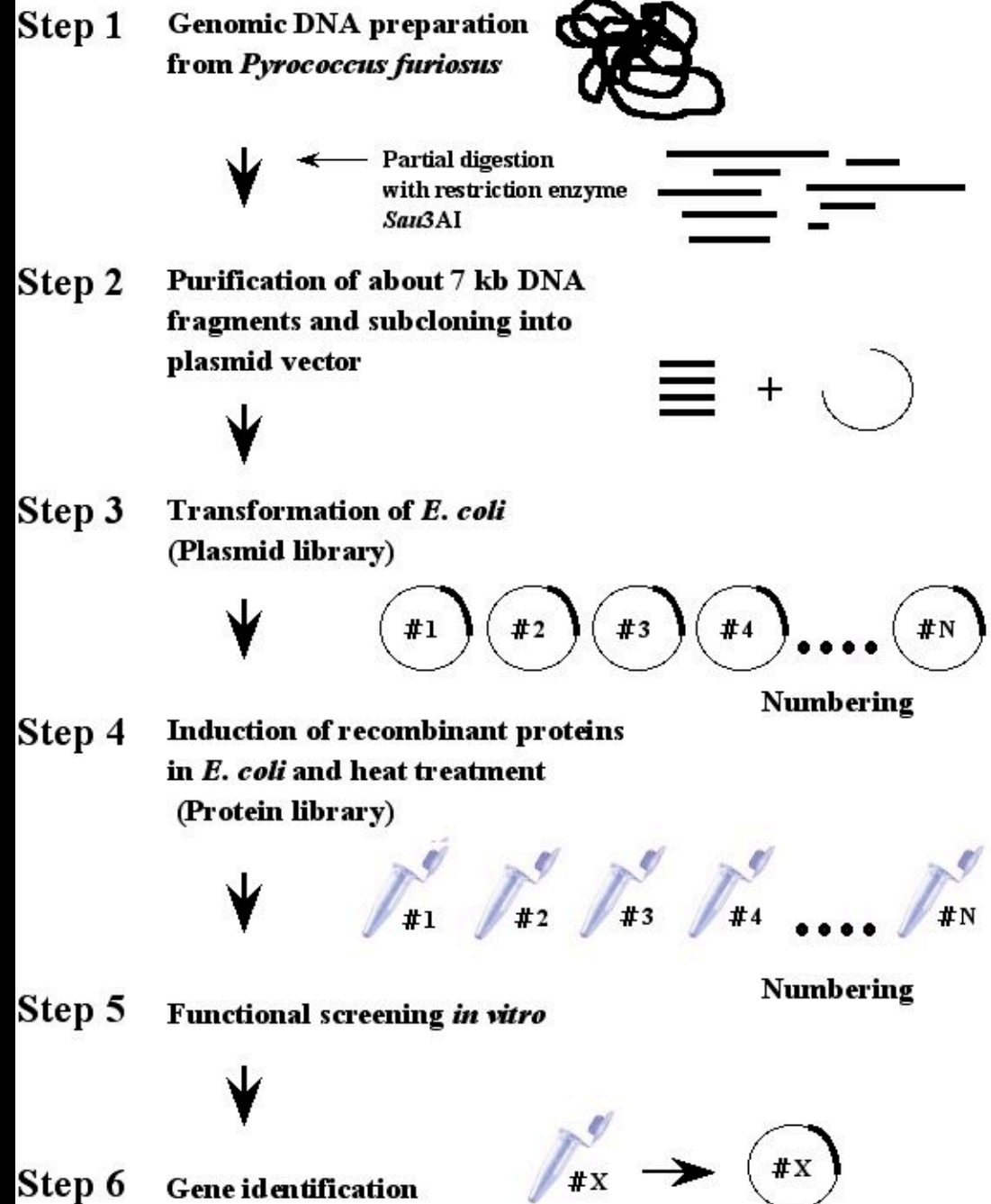
Fujishima, K. *et al. DNA Res.* (2007)

Feedback

○ Functional RNA elements

Expression Cloning of novel RNA-binding proteins

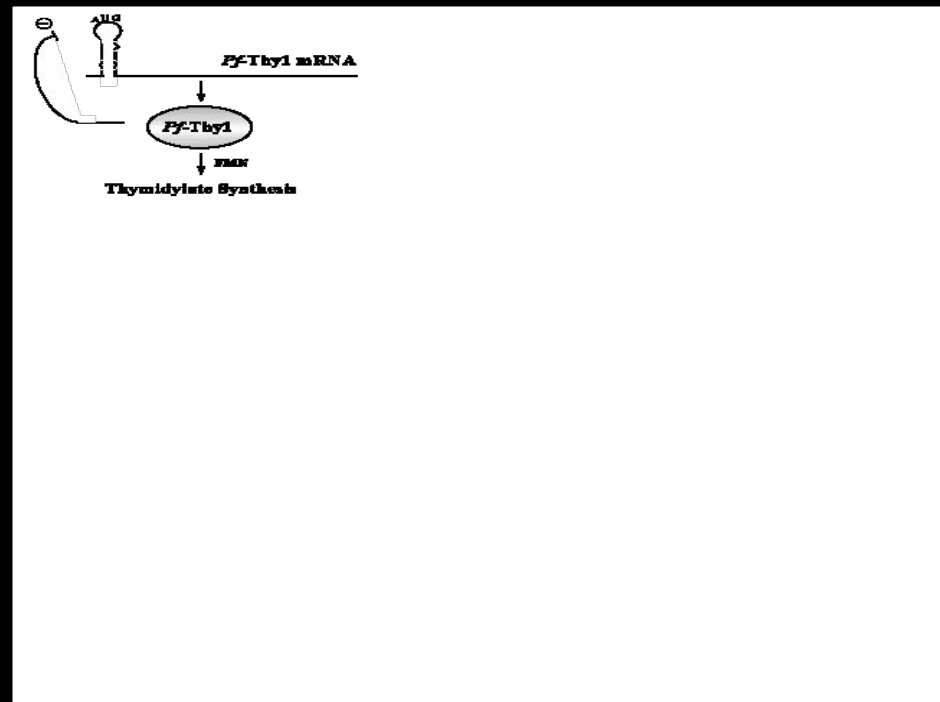
Kanai, A. *et al.* (2003)
Biochem J. 372, 253-261



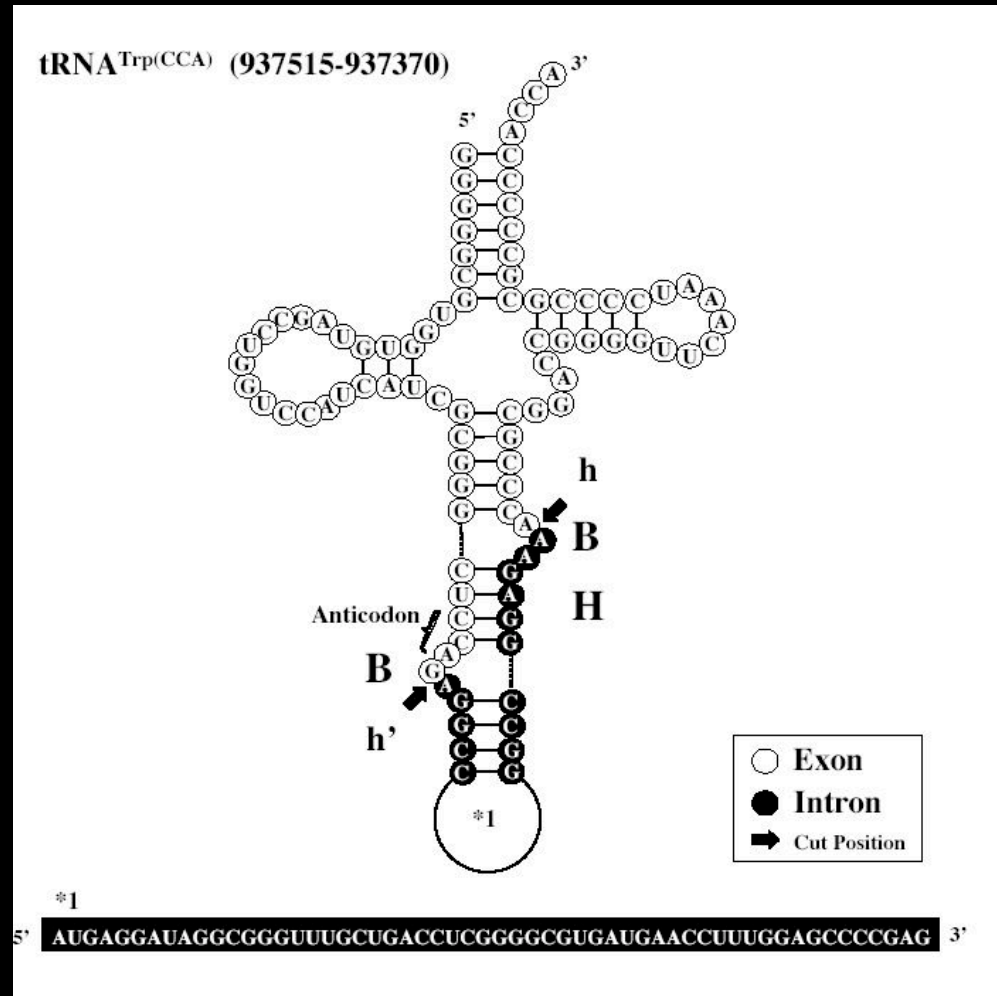
Biochem. J. (2006) 393, 373-379 (Printed in Great Britain) doi:10.1042/BJ20050608

Archaeal *Pyrococcus furiosus* thymidylate synthase 1 is an RNA-binding protein

Akio Kanai, Asako Sato, Jun Imoto and Masaru Tomita

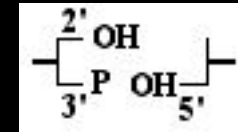


Pre-tRNA Splicing

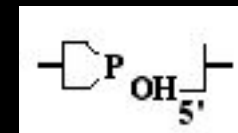


Structures of RNA terminals

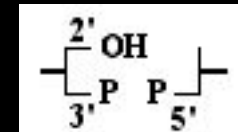
(1)



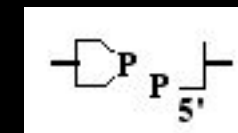
(2)



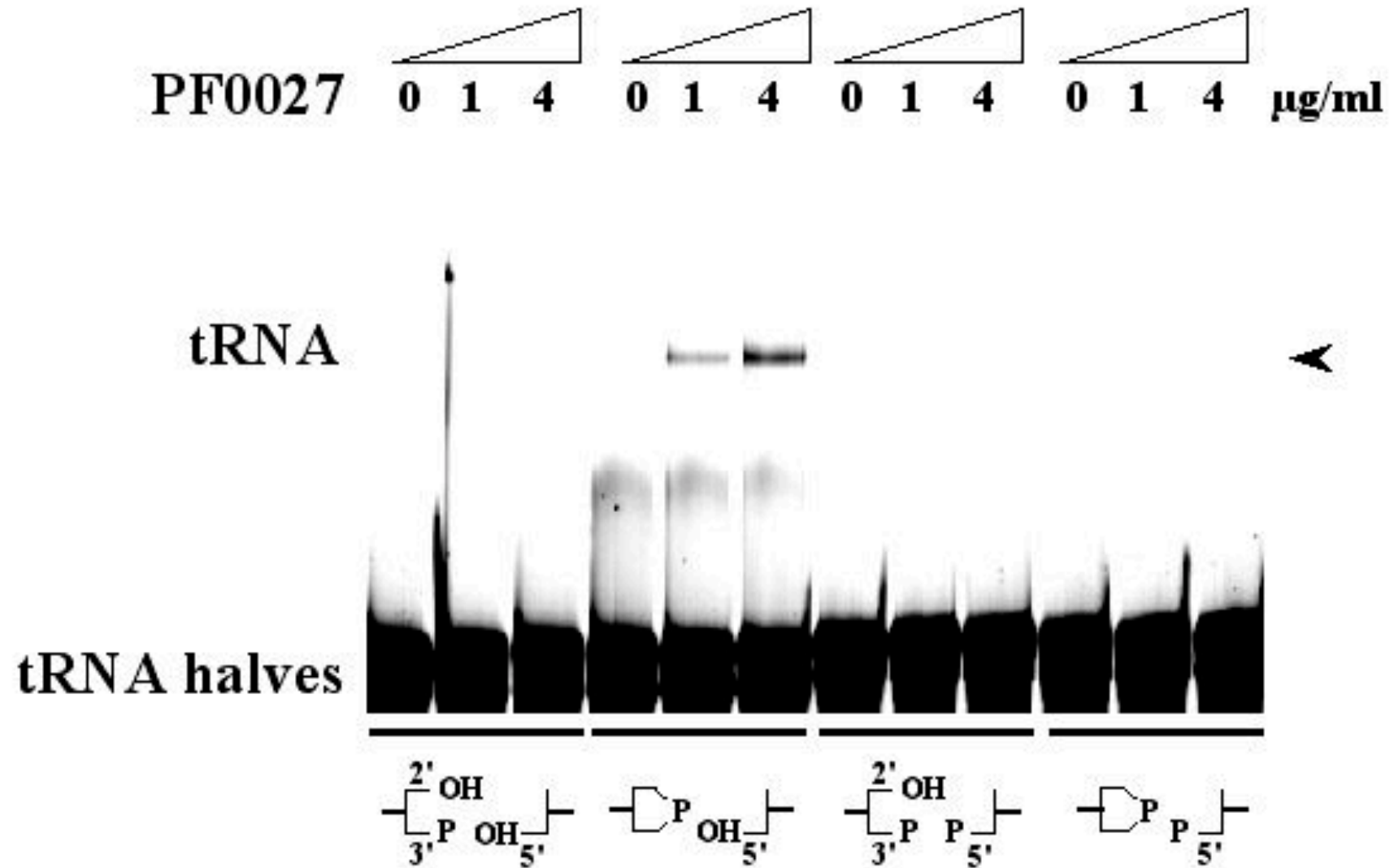
(3)



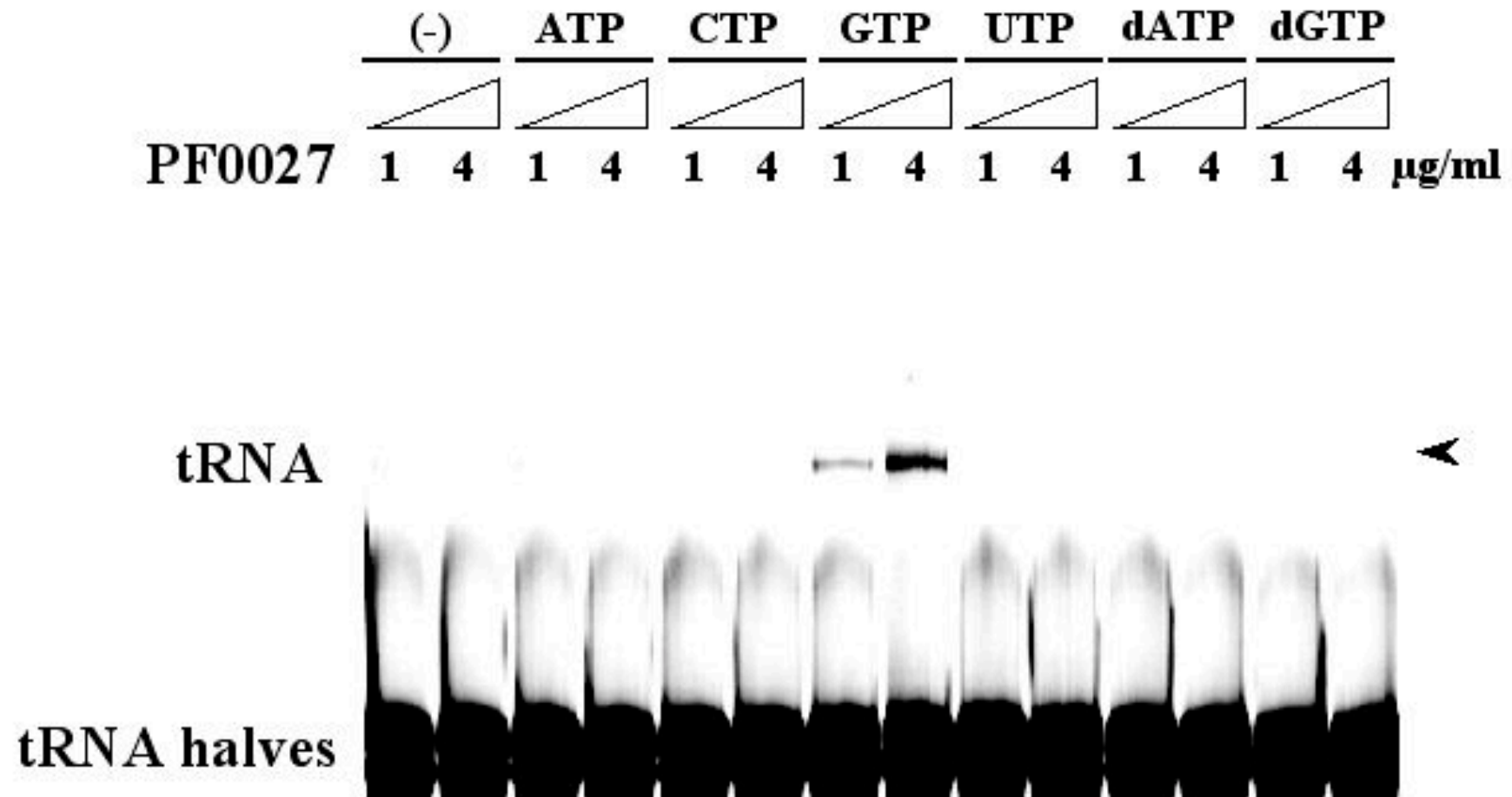
(4)



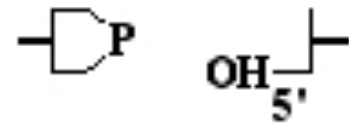
tRNA Ligation Activity (PF0027 protein)



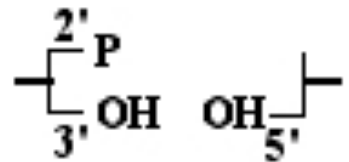
tRNA Ligation Activity (2)



Possible Model

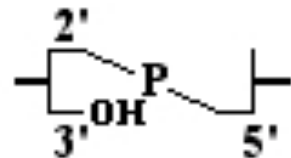


CPDase activity

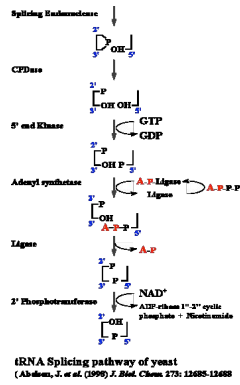


RNA ligase activity

(+GTP)



One Enzyme



In Yeast & Plant

tRNA ligase

Yeast, Plant

Yes

Human

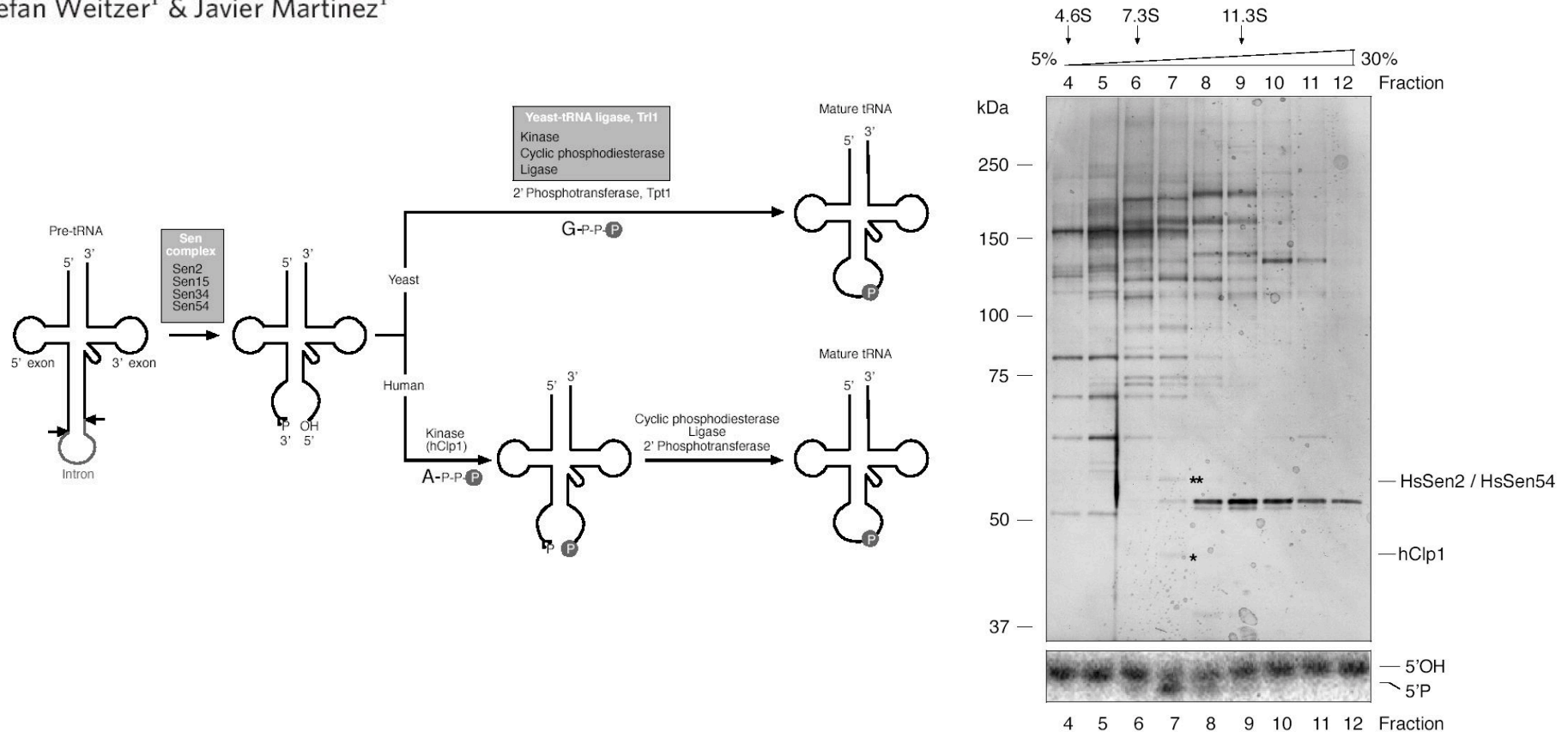
?

Archaea

?

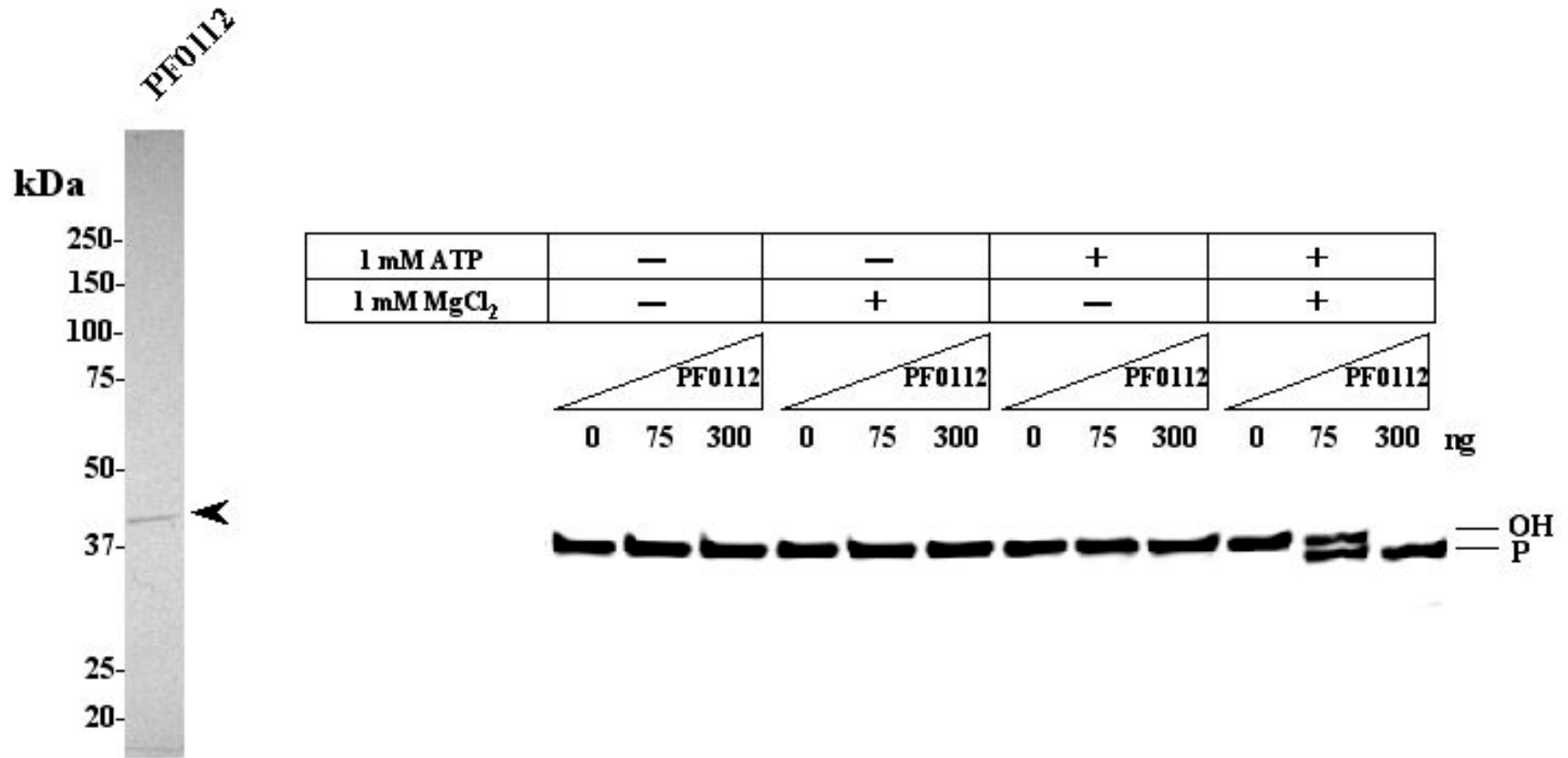
LETTERS

The human RNA kinase hClp1 is active on 3' transfer RNA exons and short interfering RNAs

Stefan Weitzer¹ & Javier Martinez¹

PF0112 Protein:

Human siRNA kinase (hClp1)-like enzyme



354 aa (40.1 kDa)

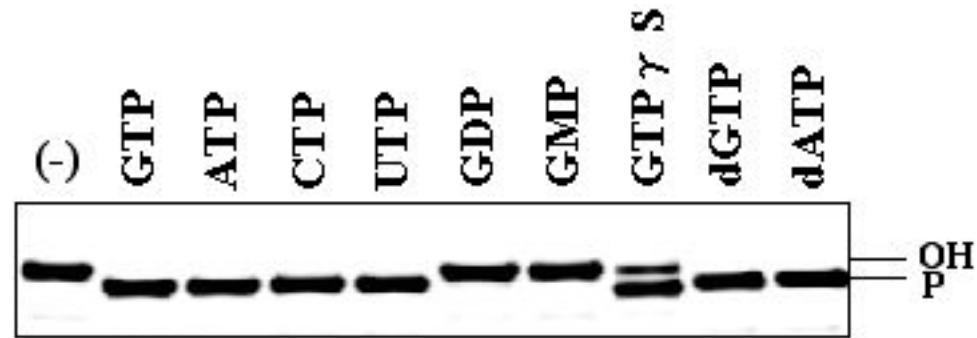
tRNA ligation mechanism

Yeast, Plant

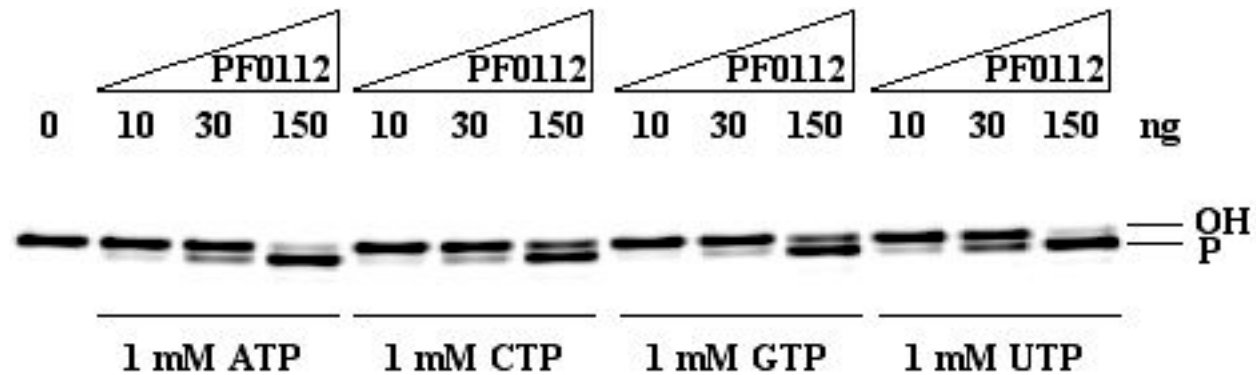
[Human

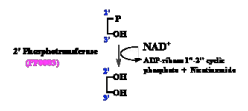
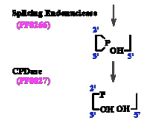
Archaea

(A) Oligo RNA Substrate



(B) Oligo DNA Substrate





Possible (RNA Splicing enzymes in archaea)
(770270)

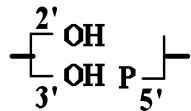
Archaeal RNA ligase is a homodimeric protein that catalyzes intramolecular ligation of single-stranded RNA and DNA

Christopher Torchia¹, Yuko Takagi¹ and C. Kiong Ho^{1,2,*}

¹Department of Biological Sciences and ²Department of Microbiology and Immunology, State University of New York at Buffalo, Buffalo, NY 14260, USA

Received July 07, 2008; Revised August 21, 2008; Accepted September 5, 2008

T4 Type RNA ligase



tRNA Type RNA ligase

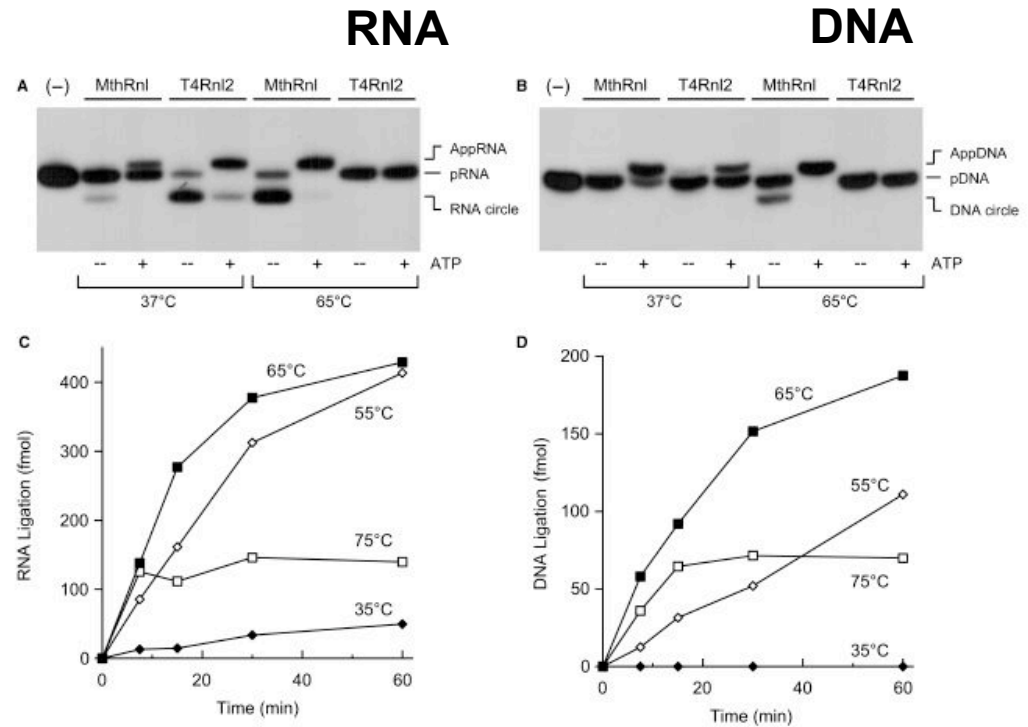
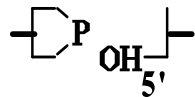
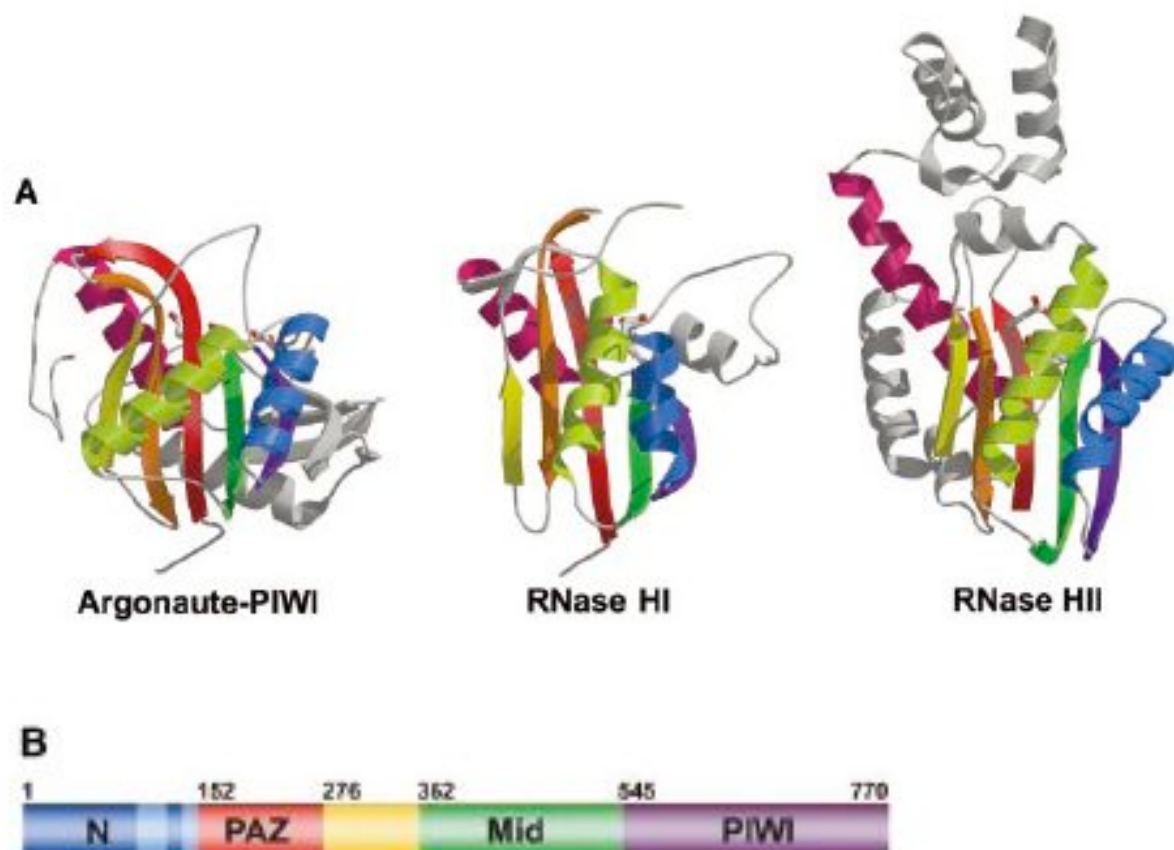


Figure 2. MthRnl can circularize single-stranded RNA and single-stranded DNA at elevated temperature. (A) RNA ligation. Standard ligation reactions containing 5 pmol MthRnl or T4Rnl2, with or without 1 mM ATP, were incubated for 30 min at either 37°C or 65°C as indicated. Reaction products were resolved by denaturing PAGE. An autoradiogram of the gel is shown. Position of pRNA, AppRNA and circularized RNA (RNA circle) are indicated. A control reaction, lacking enzyme, is shown in the lane indicated by (-). (B) DNA ligation. Identical to (A) except that the substrate for ligation was ³²P-labeled 24-mer DNA (pDNA). Reaction products, AppDNA and circularized DNA (cDNA) are indicated. (C) Kinetics and temperature-dependency. A reaction mixture (50 µl) containing 50 mM Tris-HCl (pH 6.5), 0.5 mM MgCl₂, 2.5 pmol pRNA and 10 pmol MthRnl was incubated at either 35°C, 55°C, 65°C or 75°C. Aliquots (10 µl) were withdrawn at the times indicated and quenched immediately with formamide-EDTA. The level of circular RNA product is plotted as a function of incubation time. (D) Identical to (C) except that the substrate for ligation was 24-mer DNA.

**Archaeal enzymes vaguely
recognized DNA and RNA**

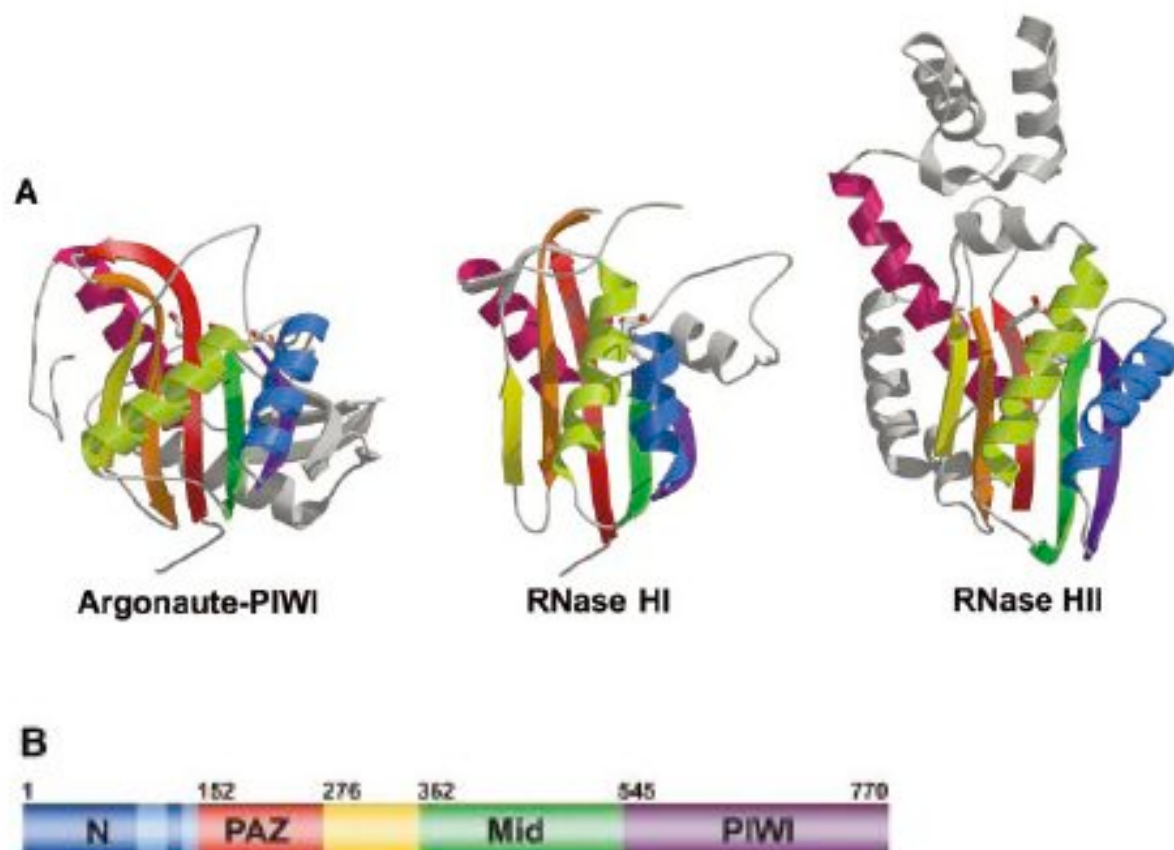
Crystal Structure of Argonaute and Its Implications for RISC Slicer Activity

Ji-Joon Song,^{1,2} Stephanie K. Smith,² Gregory J. Hannon,¹
Leemor Joshua-Tor^{1,2*}



Crystal Structure of Argonaute and Its Implications for RISC Slicer Activity

Ji-Joon Song,^{1,2} Stephanie K. Smith,² Gregory J. Hannon,¹
Leemor Joshua-Tor^{1,2*}



Summary of *Pf*-RNase HII- *Pf*-Ago chimeric proteins

	<i>RNase HII</i> β -strand	<i>Ago</i> α -helix
<i>Pf</i> -RNase HII WT	ILLYAD	AADVD ANRFASLIERE
M1	ILLYAD	AADVD ANRFASLIERE
M2	ILLYAD	AADVT ANRFASLIERE
M3	ILLYAD	RDGRIY ANRFASLIERE
M4	KLILAD	RDGRIY ANRFASLIERE
M5	ILLYAD	AADVD ANRFASLIERE
M6	ILLYAD	AADVD ANRFASLIERE
M7	ILLYAD	AADVD ANRFASLIERE
M8	ILLYAD	AADVD ANRFASLIERE
M9	ILLYAD	AADVT ANRFASLIERE
<i>Pf</i> -Ago1 WT	KLILAD	RDGRIY ANRFASLIERE



SDS-PAGE (RESOURCE-S Column Fractions)

Substrate: RNA/RNA

Substrate: RNA/DNA

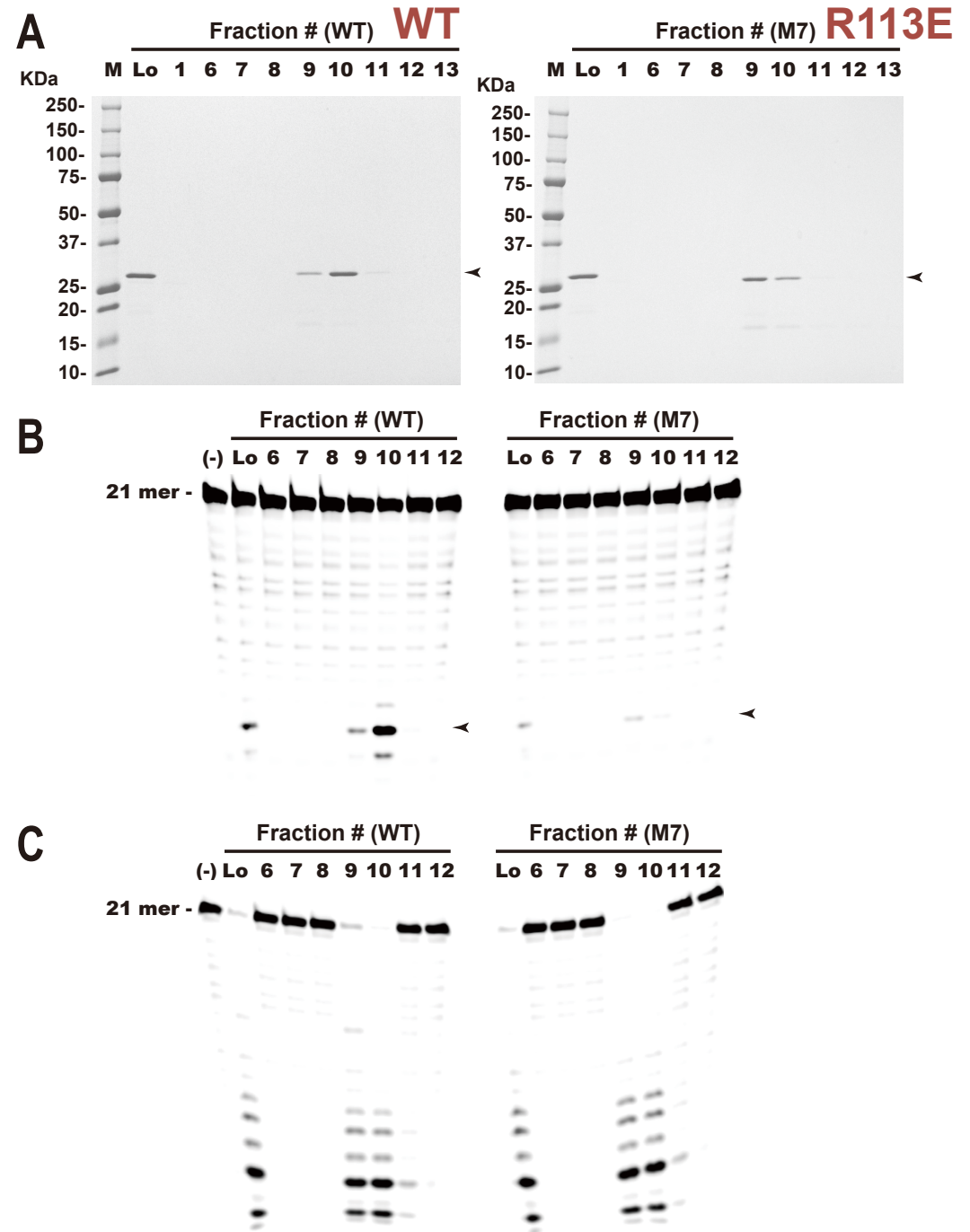


Figure 5 Kitamura *et al.*

Summary of *Pf*-RNase HII- *Pf*-Ago chimeric proteins

	<i>Pf</i> β -strand	<i>Pf</i> α -helix
<i>Pf</i> -RNase HII WT	ILLYAD	AADVDAANRFASIIERR
M1	ILLYAD	AADVDAANRFASIIERR
M2	ILLYAD	AADVDAANRFASIIERR
M3	ILLYAD	RDGRIITANRFASIIERR
M6	KLILAD	RDGRIITANRFASIIERR
M8	ILLYAD	AADVDAANRFASIIERR
M9	ILLYAD	AADVDAANRFASIIERR
M7	ILLYAD	AADVDAANRFASIIERR
M8	ILLYAD	AADVDAANRFASIIERR
M9	ILLYAD	AADVDAANRFASIIERR
<i>Pf</i> -Ago1 WT	KLILAD	RDGRIITANRFASIIERR

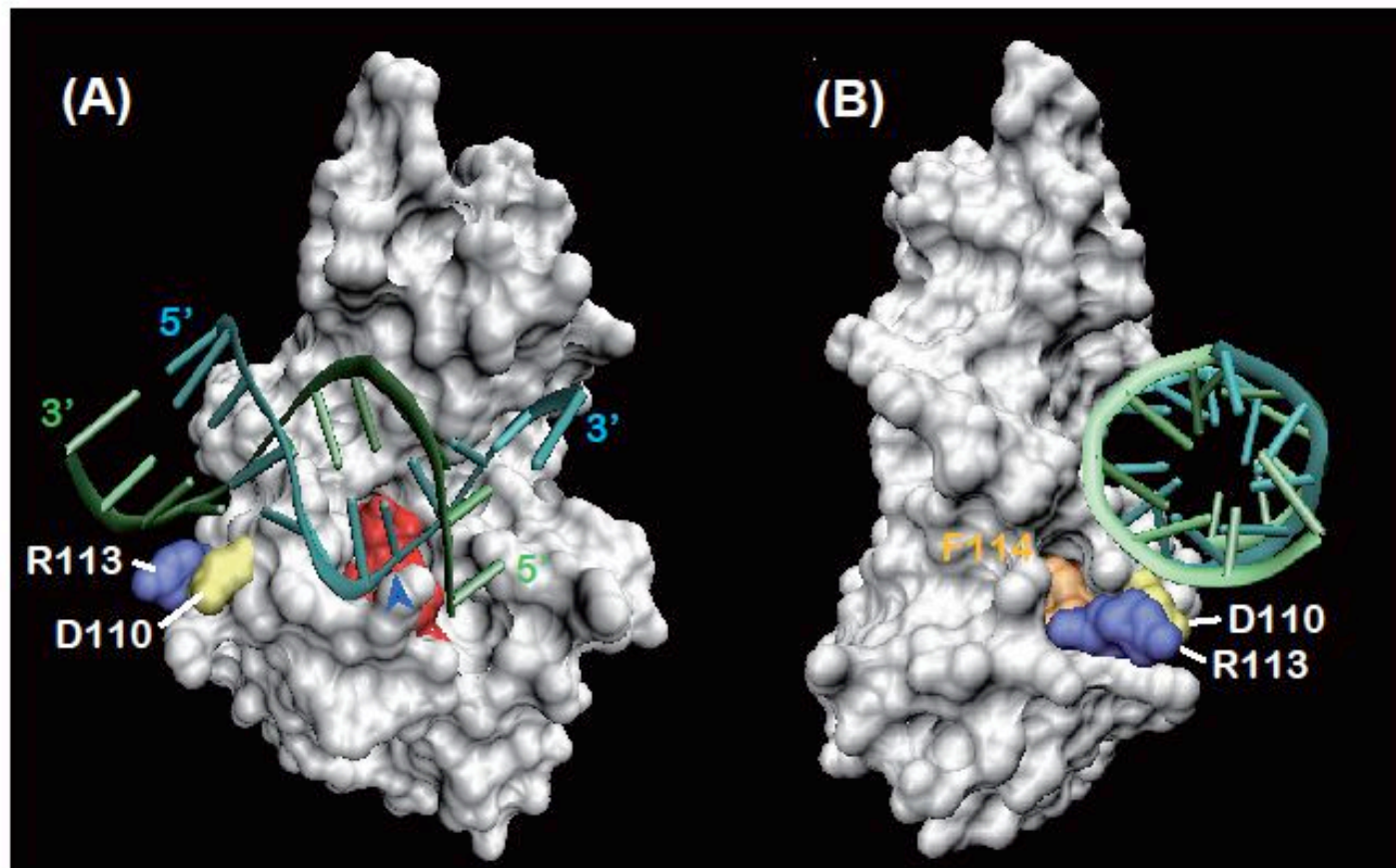


dsRNA cleavage activity



Substrate:
RNA/RNA

Possible structural model of *Pf*-RNase HII with a dsRNA



**Archaeal enzymes vaguely
recognized DNA and RNA**

Evolution of Biological System:

Simple → **Complex**

Non-Specific → **Specific**

RNA Group

A. Sato (Staff)

K. Aoyagi (Staff)

H. Kochiwa (Staff)

K. Fujishima (Post Doc)

M. Matsui (D1)

Y. Watanabe (D2)

J. Sugahara (D1)

A. Shinhara (M1)

K. Takane (M1)

K. Kikuta (M1)

N. Saito (M1)

H. Toyoshima (M1)

Y. Chiwata (B2)

K. Hamashima (B4)

A. Kanai (Staff)

S. Murakami (B3)

K. Morita (B3)

October, 2009

(Core members only)