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

Akio KANAI*^{†1}, Asako SATO*, Jun IMOTO*[‡] and Masaru TOMITA*^{†‡}

*Institute for Advanced Biosciences, Keio University, Tsuruoka, 997-0017, Japan, †Department of Environmental Information, Keio University, Fujisawa, 252-8520, Japan, and ‡Bioinformatics Program, Graduate School of Media and Governance, Keio University, Fujisawa, 252-8520, Japan

Using a stem-loop RNA oligonucleotide (19-mer) containing an AUG sequence in the loop region as a probe, we screened the protein library from a hyperthermophilic archaeon, *Pyrococcus furiosus*, and found that a flavin-dependent thymidylate synthase, *Pf*-Thy1 (*Pyrococcus furiosus* thymidylate synthase 1), possessed RNA-binding activity. Recombinant *Pf*-Thy1 was able to bind to the stem-loop structure at a high temperature (75 °C) with an apparent dissociation constant of 0.6 μ M. A similar stem-loop RNA structure was located around the translation start AUG codon of *Pf*-Thy1 RNA, and gel-shift analysis revealed that *Pf*-Thy1 could also bind to this stem-loop structure. *In vitro* translation analysis using chimaeric constructs containing the stem-loop sequence in their *Pf*-Thy1 RNA and a luciferase reporter gene

indicated that the stem-loop structure acted as an inhibitory regulator of translation by preventing the binding of its Shine–Dalgarno-like sequence by positioning it in the stem region. Addition of *Pf*-Thy1 into the *in vitro* translation system also inhibited translation. These results suggested that this class of thymidylate synthases may autoregulate their own translation in a manner analogous to that of the well characterized thymidylate synthase A proteins, although there is no significant amino acid sequence similarity between them.

Key words: archaea, *in vitro* translation, *Pyrococcus furiosus*, RNA-binding protein, stem-loop structure, thymidylate synthase.

INTRODUCTION

ThyA (thymidylate synthase A) is a well-known enzyme that catalyses the conversion of dUMP and 5,10-methylenetetrahydrofolate to dTMP and dihydrofolate [1,2]. In 2002, an alternative thymidylate synthase, Thy1 (also known as ThyX), was reported to be able to synthesize dTMP in the presence of the co-enzyme flavin [3]. Thy1 was initially identified as a protein complementing a thymidine-requiring mutant of *Dictyostelium discoideum* [4]. ThyA and Thy1 have neither amino acid sequence similarity nor structural similarity [5,6]. The complete genome data of many organisms also indicate that the *Thy1* gene is limited to microbial genomes lacking the *ThyA* gene, suggesting that these two enzymes form two distinct classes of thymidylate synthases [3]. However, ThyA protein is also known as an RNA-binding protein [7,8]. Purified ThyA protein has been shown by gel-shift assay to bind its own mRNA, and *in vitro* translation analysis has indicated that ThyA protein inhibits translation [7,8]. These results strongly suggest that expression of the *ThyA* gene is autoregulated by its own product at the translation level.

In our previous research [9], we used an expression cloning method to screen a genomic DNA library to identify DNA- or RNA-binding proteins of the hyperthermophilic archaeon *Pyrococcus furiosus*. We found that one gene product, which we named FAU-1 (*P. furiosus* AU-binding) protein, demonstrated the strongest binding activity of all the genomic library-derived proteins tested against an AU-rich RNA sequence, r(A-U)₁₀ [9]. Analysis of RNA-binding specificity led us to conclude that the FAU-1 protein bound specifically to the AU-rich sequence in the loop region of a certain stem-loop structure. In recent years, it has been well accepted that RNA secondary structures, including stem-loop structures, are involved in many stages of

gene regulation, such as transcription, splicing, translation and degradation [10]. In particular, a stem-loop RNA structure located in the vicinity of the translation start AUG codon appears to be the key regulator for translation [10].

In the present work, we isolated the gene encoding the thymidylate-synthase-complementing protein [*Pf*-Thy1 (*P. furiosus* thymidylate synthase 1)] by screening the protein library of *P. furiosus* using a stem-loop RNA oligonucleotide containing an AUG sequence in the loop region as a probe. We demonstrate here that *Pf*-Thy1 also binds to the stem-loop structure located in the vicinity of the translational start codon AUG in its own mRNA. *In vitro* analysis indicated that the stem-loop structure of *Pf*-Thy1 mRNA might work as a translational repressor. These results were consistent with previous findings [7,8] that another thymidylate synthase, ThyA, is a translational repressor that binds to its own mRNA, suggesting that the translation could be regulated in an analogous manner.

EXPERIMENTAL

Screening of RNA-binding activity from a protein library of *P. furiosus*

A protein library of *P. furiosus* DSM3638 was constructed using an expression pRSET A plasmid vector (Invitrogen, Carlsbad, CA, U.S.A.). Genomic DNA isolated from a culture of *P. furiosus* was partially digested with *Sau*3AI, and the resulting DNA fragments were fractionated by electrophoresis on a 0.6% agarose gel. Fragments with molecular size of approx. 7 kb were prepared from the gel and ligated into a *Bam*HI-cleaved pRSET A plasmid vector. Recombinant proteins expressed in *Escherichia coli* strain DH5 α were prepared as described previously [9], with slight

Abbreviations used: ds, double-stranded; FAM, 6-carboxyfluorescein; FAU-1, *Pyrococcus furiosus* AU-binding; Luc, luciferase; *Pf*-Thy1, *Pyrococcus furiosus* thymidylate synthase 1; SD, Shine–Dalgarno; ss, single-stranded; ThyA, thymidylate synthase A.

¹ To whom correspondence should be addressed (email akio@sfc.keio.ac.jp).

modifications. We made 40 protein pools, termed P1–P40, each pool consisting of 20 independent *E. coli* lysates expressing each of the above-mentioned plasmids. A total of 800 independent lysates were expected to cover the *P. furiosus* proteome (2 Mb). These extracts were screened by gel-shift assay using an S-5 probe, 5'-FAM-r(AGAUAGGGGAUGCCUAUCU)-3' (where FAM is 6-carboxyfluorescein).

Construction of the *Pf*-Thy1 expression vector and purification of His-tagged recombinant proteins

Genomic DNA of *P. furiosus* DSM3638 prepared by using a GNOME kit (BIO101, La Jolla, CA, U.S.A.) was partially digested with Sau3AI. The resulting DNA fragments were fractionated by electrophoresis on a 0.7% agarose gel. Fragments approx. 15 kb in size were prepared from the gel and used as templates in PCR cloning. The *PF0540* (*Pf*-Thy1) gene [NCBI (National Center for Biotechnology Information) protein database accession number NP_578269] was cloned into the pET-23b expression vector (Novagen, Madison, WI, U.S.A.) via PCR using the site-specific primers 540S, 5'-d(AATCATATGGTTCG-TGTTACGCTCGTTAACT)-3' and 540A, 5'-d(TAATCGAGG-CTCTACTTTCCGCAACTTTTCC)-3'. These oligonucleotides were designed to include NdeI and XhoI sites at the 5' and 3' termini respectively. The resulting vector, termed pET-0540, encoded a full-length Thy1 protein with a His₆ tag at its C-terminal end. The nucleotide sequences of the DNA inserts were determined and were confirmed to be identical with those in the NCBI database (accession number AE010177).

Recombinant *Pf*-Thy1 was prepared in the same way as described previously [11]. Briefly, *E. coli* strain BL21(DE3) was transformed with the plasmid pET-0540. Transformants growing exponentially in LB (Luria–Bertani) medium containing 50 µg/ml ampicillin at 37 °C were supplemented with 0.4 mM IPTG (isopropyl β-D-thiogalactoside). After 14–16 h of further growth at 30 °C, the cells were harvested by centrifugation (5000 g for 5 min at 4 °C) and the recombinant proteins were extracted by sonication (30 s) in buffer A [20 mM Tris/HCl (pH 8.0), 5 mM imidazole, 500 mM NaCl and 0.1% (v/v) Nonidet P40]. These extracts were heat treated at 85 °C for 15 min to destroy the endogenous ribonuclease-related proteins from the *E. coli*, and then centrifuged at 12000 g for 10 min at 4 °C to separate the cell debris from the extract. The recombinant *Pf*-Thy1 protein was purified on a Ni²⁺-Sepharose column, according to the manufacturer's instructions (Amersham Biosciences, Piscataway, NJ, U.S.A.). The eluted protein peaks were pooled and the buffer was exchanged for buffer B [50 mM Tris/HCl (pH 8.0), 1 mM EDTA, 0.02% Tween 20, 7 mM 2-mercaptoethanol and 10% (v/v) glycerol] by gel filtration using a HiTrap desalting column (Amersham Biosciences).

Gel-shift assay

The 5'-end-FAM-labelled oligoribonucleotides were chemically synthesized by Hokkaido System Science Co. (Hokkaido, Japan). Binding reactions containing S-5 (or SL-1W) oligonucleotide (0.3–1.0 µM for the gel-shift assay and 5 nM for the *K_d* determination) and either 0.05–5 µM purified *Pf*-Thy1 protein or approx. 2 µg of extract from *E. coli* expressing the *P. furiosus* genomic DNA were incubated at 75 °C or at various temperatures (as indicated below) for 15 min in 20 µl of RNA-binding buffer [10 mM Tris/HCl (pH 7.5), 50 mM NaCl, 0.5 mM EDTA, 0.5 mM MnCl₂, 5% (v/v) glycerol and 1 mM dithiothreitol]. To examine the effect of two bivalent metal ions on RNA-binding activity, we used 0.05–2 mM MgCl₂ or MnCl₂. The RNA–protein complexes were analysed by electrophoresis on a 6% non-

denaturing polyacrylamide gel. The quantities of RNA–protein complexes were evaluated by scanning the fluorescent image using an FX Pro computerized image analyser (Bio-Rad Laboratories, Hercules, CA, U.S.A.). For competition analysis, an excess of unlabelled oligonucleotides (5, 10 or 50 molar excess) was added to the reaction mixture. The sequences of the oligoribonucleotides used were as follows: S-5, 5'-FAM-r(AGAUAGGGGAUGCCUAUCU)-3'; SL-1W, 5'-FAM-r(GGGGGUAAG-AAAUGCUUCGUGUACGCUC)-3'; 7RNAS, 5'-r(ACGUC-AG)-3'; 7RNAA, 5'-r(CUGACGU)-3'; S-5(DNA), 5'-d(AGATAGGGGATGCCTATCT)-3'; 7DNAS, 5'-d(ACGTCAG)-3'; 7DNAA, 5'-d(CTGACGT)-3'.

Construction of *Pf*-Thy1–Luc (luciferase) chimaeric vectors

Recombinant pSL-1W–Luc, pSL-2M–Luc, pSL-3M–Luc, pSL-4M–Luc and pSL-5M–Luc vectors (see Figure 1 and Table 1) under the control of a phage T7 promoter were created in the plasmid pGL3-Basic vector (Promega, Madison, WI, U.S.A.). Briefly, five pairs of oligonucleotides [SL-1S and SL-1A; SL-2S and SL-2A; SL-3S and SL-3A; SL-4S and SL-4A; and SL-5S and SL-5A] were annealed and ligated into the NheI–HindIII site of the pGL3-Basic vector. The resulting vectors that expressed mRNA consisted of the 5' untranslated region and the first 7-amino-acid segment of the *Pf*-Thy1 gene, the 11-amino-acid linker sequence, and the firefly Luc gene in-frame. The nucleotide sequences of the inserts were determined. The sequences of the oligodeoxyribonucleotides used were as follows: SL-1S, 5'-d(CTAGTAAT-ACGACTCACTATAGGGAAGTGGGGGTAAGAAATGGT-TCGTGTTACGCTCGT)-3'; SL-1A, 5'-d(AGCTACGACGTA-ACACGAACCATTTCTTACCCCCCACTTCCCTATAGTGAG-TCGTATTA)-3'; SL-2S, 5'-d(CTAGTAATACGACTCACTATA-GGGAAGTGGATGGTTCGTGTTACGCTCGT)-3'; SL-2A, 5'-d(AGCTACGAGCGTAACACGAACCATCCACTTCCCTATAGTGAGTCGTATTA)-3'; SL-3S, 5'-d(CTAGTAATACGACTCA-CTATAGGGAAGTGGGGGTAAGAAATGGTTCGTGTTAC-CCTCGT)-3'; SL-3A, 5'-d(AGCTACGAGGGTAACACGAAC-CATTTCTTACCCCCCACTTCCCTATAGTGAGTCGTATTA)-3'; SL-4S, 5'-d(CTAGTAATACGACTCACTATAGGGAAGTG-GGGGGTAAGAAATGGCGCGTGTACGCTCGT)-3'; SL-4A, 5'-d(AGCTACGAGCGTAACACGCGCCATTTCTTACCCCC-CACTTCCCTATAGTGAGTCGTATTA)-3'; SL-5S, 5'-d(CTA-GTAATACGACTCACTATAGGGAAGTGGGGGTAAGAAA-TGGTTCGTGACACGCTCGT)-3'; SL-5A, 5'-d(AGCTACG-AGCGTGTCACGAACCATTTCTTACCCCCCACTTCCCTAT-AGTGAGTCGTATTA)-3'.

In vitro transcription, *in vitro* translation and Luc assay

A DNA fragment containing the T7 RNA polymerase promoter and the entire *Pf*-Thy1–Luc chimaeric region was amplified by PCR from each of the *Pf*-Thy1–Luc chimaeric vectors (pSL-1W–Luc, pSL-2M–Luc, pSL-3M–Luc, pSL-4M–Luc and pSL-5M–Luc) using vector-specific oligonucleotides: RV primer 3, 5'-d(CTAGCAAAATAGGCTGTCC)-3', and RV primer 4, 5'-d(GACGATAGTCATGCCCGCG)-3'. The resulting DNA fragment was purified by agarose gel electrophoresis and used as a template for *in vitro* transcription. The *in vitro* transcription was performed using a MEGAscript kit (Ambion, Austin, TX, U.S.A.), and the resulting RNA was purified by lithium chloride precipitation, phenol/chloroform extraction and ethanol precipitation in accordance with the manufacturer's instructions.

All of the *in vitro*-transcribed RNAs (1 µg) were incubated either in *E. coli* S30 extract at 37 °C or in rabbit reticulocyte lysate at 30 °C for 60 min in 25 µl of reaction mixture, according to the manufacturer's instructions (Promega). Luc activity was

Figure 1 Sequences and possible RNA secondary structures of oligonucleotides used in this study

(a) S-5 and (b) SL-1W oligonucleotides were used in the gel-shift assay. The SD-like sequence of wild-type probe (SL-1W) for *Pf*-Thy1 RNA is boxed. (c) SL-2M, (d) SL-3M, (e) SL-4M and (f) SL-5M were mutants of the wild-type stem-loop structure, SL-1W (b), that was located at the translation initiation site of *Pf*-Thy1 RNA. These RNA secondary structures, except for SL-2M, were predicted using the Mfold algorithm and correspond to the reporter constructs shown in Table 1. AUG sequences are indicated by circles. The calculated minimum free energies (in kJ/mol) of the wild-type and mutant oligonucleotides predicted by Mfold were as follows (at 37°C and 1 M NaCl conditions): (a) S-5, -34; (b) SL-1W, -17; (d) SL-3M, -37; (e) SL-4M, -13; and (f) SL-5M, -9.2.

determined by using the Luciferase Assay System (Promega) and quantified by an automated luminometer (Lumat LB9507; Berthold Technologies, Bad Wildbad, Germany). In this system, measurements were linear for more than eight orders of magnitude. In some cases, purified *Pf*-Thy1 was added to the reaction mixture.

RESULTS AND DISCUSSION

Expression cloning and identification of the gene for the RNA-binding protein *Pf*-Thy1

The *E. coli* extracts that expressed *P. furiosus* genes (40 pools, 800 independent clones) were screened by gel-shift assay using an S-5 probe, which was designed as a stem-loop structure containing an AUG sequence in the loop region (Figure 1). In our previous research, we used a series of simply repeated probes, such as r(A-U)₁₀ and r(G-C)₁₀ [9]. After the cloning of genes for RNA-binding proteins to those probes, however, we found that almost all RNA-binding proteins preferentially bound to certain RNA secondary structures, including stem-loop structures. Therefore, in the present research, we used the artificial S-5 probe, the sequence of which is predicted to form a stable RNA secondary structure [minimum free energy = -34 kJ/mol under 1 M NaCl conditions at 37°C] by the Mfold program [12]. Genome-level

Figure 2 Multiple sequence alignment of *Pf*-Thy1 and related proteins

Sequences from: GS, *G. sulfurreducens*; DD, *D. desulfuricans*; PF, *P. furiosus*; and PA, *P. abyssi*. Sequence alignment was done with ClustalW [19]. Identical amino acids are indicated by asterisks, and partly conserved amino acids are indicated by dots (two-dot regions show a higher degree of similarity). Gaps (—) were inserted to maximize the number of amino acid matches. The two conserved regions are boxed.

screening was used to obtain three positive pools containing RNA-binding activity. Of these positive pools, one plasmid pool, P7, showing the strongest binding activity to the S-5 probe at 75°C, was analysed further. We isolated one plasmid clone, P7-134, in the second screening from the P7 plasmid pool. We found that plasmid clone P7-134 contained an insert of approx. 2 kb (a DNA deletion in the plasmid insert may have occurred during genomic library preparation), and we determined the whole nucleotide sequence for this insert. The sequence was identical with positions 1030–2931 in section 52 of 173 of the complete *P. furiosus* genome (accession number AE010177). We found a region that contained two open reading frames, PF0540 and PF0541. We subcloned each of the two genes into the pET-23b expression vector, made recombinant proteins and confirmed that only protein PF0540 possessed RNA-binding activity (results not shown).

From nucleotide sequence analysis of the genomic clone (plasmid P7-134), we deduced that the complete amino acid sequence of protein PF0540 consisted of 244 amino acids (Figure 2) with a calculated molecular mass of 29 kDa. After searching the NCBI database by using the PSI-BLAST program (<http://www.ncbi.nlm.nih.gov/BLAST/>), we found that protein PF0540 had 88% identity and 98% similarity with the recently identified flavin-dependent thymidylate synthase Thy1 (or ThyX) from *P. abyssi* (NCBI protein database accession number Q9UZ51) [3]. Weak sequence similarity (40% and 44% identity respectively) was observed between the protein PF0540 and the predicted alternative thymidylate synthases from two bacteria, *Geobacter sulfurreducens* (NCBI protein database accession number NP_954147) and *Desulfovibrio desulfuricans* (ZP_00129777). Thus we concluded that protein PF0540 is a flavin-dependent thymidylate synthase from *P. furiosus*. We refer to this protein as *Pf*-Thy1 hereafter.

By comparison of the amino acid sequences of these proteins, we found that two regions (termed Boxes 1 and 2) were conserved in the different species (Figure 2). The crystal structure of

the Thy1 protein from *Thermotoga maritima* (termed TM0449) has been reported, and the Thy1 protein has been shown to form a tetramer [6,13]. From homology modelling analysis for comparison between the Thy1 protein from *T. maritima* and *Pf*-Thy1, we found that each of the Box 1 regions in a monomer corresponded to the centre of the tetramer, making a large active pocket of FAD. Box 2 regions surrounded the Box 1 regions, suggesting the importance of these conserved regions (results not shown). We did not find, however, any known RNA-binding domain in the amino acid sequence of *Pf*-Thy1.

Characterization of RNA-binding activity of *Pf*-Thy1

To obtain large amounts of *Pf*-Thy1 for biochemical studies, we overexpressed the recombinant C-terminal His₆-tagged protein in *E. coli* strain BL21(DE3) and purified it to near homogeneity by heat treatment at 85 °C for 15 min followed by Ni²⁺-Sepharose column chromatography. The purified protein had molecular mass of approx. 29 kDa, as determined by SDS/PAGE (Supplementary Figure 1, <http://www.BiochemJ.org/bj/393/bj3930373add.htm>). This finding was in agreement with the values determined from the deduced amino acid sequences of the genes. Using the purified *Pf*-Thy1 protein, we carried out several experiments to determine the biochemical characteristics of its RNA-binding ability. First, RNA-binding activities with various amounts of *Pf*-Thy1 (between 0.05 and 5 µM) and various temperature conditions (0, 22, 37, 55, 75 and 95 °C) were examined by gel-shift assay (Figure 3). RNA-*Pf*-Thy1 complexes were observed, their frequency depending on the amount of *Pf*-Thy1 used, and optimum RNA-binding activity was achieved in the 75 °C experiment, suggesting that *Pf*-Thy1 is a heat-stable RNA-binding protein. We did not observe any RNA-binding activity at 95 °C, possibly because the secondary structure of the S-5 RNA probe became denatured at this temperature. On the basis of these observations, we determined the K_d of *Pf*-Thy1 for the S-5 probe at 75 °C ($K_d = 0.6$ µM). We also examined the effects of two bivalent metal ions (Mg²⁺ and Mn²⁺) on the RNA-binding activity and found that 0.5 mM Mn²⁺ enhanced the activity (Figure 4). In our previous report, we found that the RNase HII enzyme from *P. furiosus* requires Mn²⁺ ion for its activation [11]; together these results suggest that Mn²⁺ may be a common ion that is necessary for enzymatic activation in these archaea.

In order to examine RNA-binding specificity of the *Pf*-Thy1 protein, we performed a competitive gel-shift assay (Figure 5). For this purpose, excess amounts of the ribonucleotides {S-5 (stem-loop RNA), 7RNAS [ss (single-stranded) RNA] or 7RNASA [ds (double-stranded) RNA]}, or the deoxyribonucleotides [S-5 (stem-loop DNA), 7DNAS (ssDNA) and 7DNASA (dsDNA)] were added to the binding mixture for *Pf*-Thy1 protein, and the S-5 RNA-*Pf*-Thy1 complexes were analysed. Only S-5 RNA effectively blocked the formation of S-5 RNA-*Pf*-Thy1 complexes when a 10-fold molar excess of the competitor oligonucleotide was used. The results suggest that *Pf*-Thy1 bound to the stem-loop RNA structure.

Stem-loop structure at the translational start site of the *Pf*-Thy1 gene and the inhibitory effect of *Pf*-Thy1 on translation *in vitro*

Because the S-5 probe contained the AUG start codon in its loop region, and the other thymidylate synthase, ThyA, has been shown to bind a similar stem-loop structure located at the translation start site of its mRNA [2,8], we searched for possible RNA secondary structures in *Pf*-Thy1 mRNA by using an Mfold algorithm [12]. We found a similar stem-loop RNA structure containing a three-nucleotide bulge and a symmetrical internal loop structure located in the vicinity of the translation start AUG codon of *Pf*-Thy1 RNA

Figure 3 RNA gel-shift assay using purified *Pf*-Thy1

(a) Temperature-dependent RNA-binding activity of *Pf*-Thy1. The FAM-end-labelled oligonucleotide S-5 (0.3 µM) was incubated with 1 µM of purified *Pf*-Thy1 at various temperatures. The arrowhead indicates the position of the RNA-protein complex. Similar results were obtained in at least two independent experiments. (b) Binding affinity of *Pf*-Thy1 for the S-5 (5 nM) probe. Data for the curve were normalized to the saturation point for the *Pf*-Thy1 protein and the saturation point is defined as the protein concentration at which 100% of the RNA probes were shifted. The data represent the means $\pm$ S.D. for three separate experiments. The K_d value (0.6 µM) is equal to the protein concentration at which 50% of the maximal RNA-binding activity is shown.

Figure 4 Effect of bivalent ions Mg²⁺ and Mn²⁺ on RNA-binding activity of *Pf*-Thy1

RNA-binding activity was examined by gel-shift analysis. FAM-end-labelled oligonucleotide S-5 (0.3 µM) was incubated with 1 µM purified *Pf*-Thy1 at 75 °C in the presence of various concentrations (0–2 mM) of either MgCl₂ or MnCl₂. The arrowhead indicates the position of the RNA-protein complex. Similar results were obtained in at least two independent experiments.

Figure 5 Competitive RNA gel-shift assay

FAM-end-labelled oligonucleotide S-5 (0.3 μ M) was incubated with 1 μ M purified *Pf*-Thy1 in the presence of an excess of unlabelled oligonucleotides (5, 10 or 50 molar excess; —, control). 7RNAS, dsRNA consisting of two RNA oligonucleotides, 7RNAS and 7RNAA; 7DNAS, dsDNA consisting of two DNA oligonucleotides (7DNAS and 7DNAA). The arrowhead indicates the position of the RNA-protein complex. Similar results were obtained in at least two independent experiments.

Figure 6 *Pf*-Thy1 is bound to its own stem-loop structure, which is located at the translation initiation codon

RNA-binding activity was examined by gel-shift analysis. FAM-end-labelled oligonucleotide, which was either (a) SL-1W (1 μ M) or (b) S-5 (1 μ M), was incubated with 0–5 μ M of purified *Pf*-Thy1 at 75 °C. Similar results were obtained in at least two independent experiments.

itself (Figure 1b). As shown in Figure 6(a), gel-shift analysis revealed that *Pf*-Thy1 can also bind to this stem-loop structure (SL-1W probe), as found for the S-5 probe (Figure 6b), and the number of shifted bands increased with the amount of *Pf*-Thy1 used. Since Thy1 protein from *T. maritima* has been reported to form a tetramer [6,13], these bands may reflect the formation of oligomeric structures by *Pf*-Thy1. The SL-1W probe showed higher affinity for *Pf*-Thy1 protein than the S-5 probe, suggesting that the real K_d value of *Pf*-Thy1 protein is probably lower than 0.6 μ M when a suitable RNA probe is used. In addition, we found the same stem-loop structure in the vicinity of the translation start codon of Thy1 RNA from *P. abyssi* and similar structures from *G. sulfurreducens* and *D. desulfuricans* (results not shown), suggesting that these stem-loop structures play important roles in the regulation of Thy1 RNAs.

In an attempt to find out whether or not this stem-loop structure is actually involved in any regulatory function, we made a series of chimaeric constructs containing each of the wild-type or mutant stem-loop sequences and Luc reporter genes (Figure 1 and Table 1). The efficiency of translational initiation in the archaea is dependent on the SD (Shine–Dalgarno)-like sequence,

Table 1 *In vitro* translation analysis of *Pf*-Thy1–Luc chimaeric RNAs

Structural organization of the constructs and relative Luc activities are shown. Each of the *in vitro*-transcribed RNAs (1 μ g) was incubated for 60 min in either *Escherichia coli* S30 extract at 37 °C or rabbit reticulocyte lysate at 30 °C. Luc activity was determined as described in the Experimental section. Data shown are the means $\pm$ S.E.M. of three experiments. The activity derived from pSL-1W–Luc was normalized to 1.0, and the activities of the other constructs are expressed relative to this.

Constructs	Relative amount of luciferase activity	
	<i>E. coli</i> S30 extract	Rabbit reticulocyte lysate
pSL-1W–Luc	1.00	1.00
pSL-2M–Luc	0.22 $\pm$ 0.03	0.73 $\pm$ 0.08
pSL-3M–Luc	0.77 $\pm$ 0.04	0.95 $\pm$ 0.05
pSL-4M–Luc	0.35 $\pm$ 0.03	0.63 $\pm$ 0.03
pSL-5M–Luc	2.52 $\pm$ 0.50	0.59 $\pm$ 0.08

located approx. 10 nucleotides upstream of the start codon [14,15]. Since the 3' end of the 16 S rRNA sequence from *P. furiosus* is 5'-UCACCUCCU-3' (the complementary sequence is 5'-AGGAGGTGA-3'), the sequence between –12 and –4 from the AUG start codon in *Pf*-Thy1 RNA, 5'-GGGGGGTAA-3', may be used as the SD-like sequence. Because the SD-like sequence

is matched to half of the wild-type stem-loop structure SL-1W (Figure 1b), we changed the nucleotide sequences downstream of the AUG codon to make the mutants SL-3M, SL-4M and SL-5M, with the exception of SL-2M, which completely lacks the SD-like sequence. We also predicted their possible RNA secondary structures using the Mfold program, with the exception of the mutant SL-2M. For *in vitro* translation analysis, we used both *E. coli* S30 extract and a rabbit reticulocyte lysate system, because an archaea-specific *in vitro* translation system is not currently available. In Table 1, we normalized the activity derived from pSL-1W-Luc to 1.0 in both *E. coli* S30 extract and the rabbit reticulocyte lysate system for better comparisons of the mutants. The levels of Luc activity in the two systems, however, were quite different. From comparison of standard curves using recombinant Luc, when we used 1 µg of pSL-1W-Luc RNA, we calculated that the *E. coli* lysate system produced approx. 3 pg of Luc, whereas the rabbit reticulocyte lysate system produced approx. 130 ng of Luc (results not shown), indicating that the activity of the rabbit system is about 44 000 times that of the *E. coli* system.

On the other hand, SD-like-sequence-dependent translation was observed in the *E. coli* system (Table 1). We noticed that the consensus *E. coli* SD sequence was 5'-AGGAGGT-3', and this sequence is 100% identical with the complementary sequence of the 3' end of 16 S rRNA from *P. furiosus*, suggesting that *E. coli* ribosomes may recognize the SD-like sequence in *Pf*-Thy1 RNA. In particular, the pSL-2M-Luc construct had dramatically reduced Luc activity, suggesting that the SD-like sequence is necessary for efficient translation. Interestingly, translation activities were also reduced in the construct pSL-3M-Luc, which lacks an internal loop, and in the construct pSL-4M-Luc, which introduces a larger loop, suggesting that this stem-loop structure may be important for regulation. Introducing a large internal loop directly after the SD-like sequence (as in the construct pSL-5M-Luc) resulted in an approx. 2.5-fold increase in Luc activity from that of pSL-1W-Luc. These results suggest that the stem-loop structure acts as an inhibitory regulator of translation by preventing binding of its SD-like sequence by positioning it in the stem region. In contrast, although its activity levels were higher than with the *E. coli* system, the rabbit reticulocyte lysate system showed no marked change among the constructs (Table 1).

We next examined the effect of purified *Pf*-Thy1 on the *E. coli* S30 *in vitro* translation system. The addition of 150 ng/µl (approx. 5 µM) *Pf*-Thy1 to the translation reaction mixture inhibited Luc synthesis by approx. 80% at 37°C (Figure 7). In contrast with this, BSA was not effective at inhibiting the *in vitro* translation of the same mRNA. These results suggest that *Pf*-Thy1 is an inhibitory factor for the *in vitro* translation system. However, we were unable to determine any distinct structure specificity of this *Pf*-Thy1 inhibition. So far, it has not been possible to set up an *in vitro* translation system for the hyperthermophilic archaea, and even a transfection system for this organism is not yet available, suggesting that a suitable archaeal reconstitution system is needed in order to substantiate this view.

Conserved mechanism of translational regulation by thymidylate synthase

In summary, the present study shows the following: *Pf*-Thy1 is an RNA-binding protein; there is a stem-loop structure at the translation start site of *Pf*-Thy1 mRNA, and *Pf*-Thy1 can bind to this structure; the stem-loop structure has an inhibitory effect on the *E. coli in vitro* translation system; and *Pf*-Thy1 also has an inhibitory effect on the *in vitro* translation system. This direct and circumstantial evidence strongly suggests that *Pf*-Thy1 con-

Figure 7 Inhibitory effect of *Pf*-Thy1 on translation of *Pf*-Thy1-Luc chimaeric RNA *in vitro*

In vitro-transcribed RNA (1 µg) from pSL-1W-Luc was incubated in *E. coli* S30 extract at 37°C for 60 min in the presence of either *Pf*-Thy1 (●) or BSA (△). Luc activity was determined as described in the Experimental section. Data shown are the means of three experiments, and the S.E.M. is less than 12% for each case. The activity derived from pSL-1W-Luc was normalized to 1.0 for either *Pf*-Thy1 or BSA.

trols the regulation of its own mRNA as an RNA-binding protein. This finding is consistent with the fact that another thymidylate synthase, ThyA, acts as an RNA-binding protein against its own mRNA [2,8], although it belongs to a different class of thymidylate synthases. Two human enzymes, DHFR (dihydrofolate reductase), a critical enzyme for *de novo* purine and thymidylate biosynthesis [16], and cSHMT (cytoplasmic serine hydroxymethyltransferase), an enzyme involved in dTMP metabolism, are mRNA-binding proteins [17]. It is also well known that all dNTPs are made from NTPs [18]. On the basis of these observations and our current work, we can suggest that regulation at the RNA level has roles in DNA synthesis that are more important than previously expected.

We thank Kenji Higashi and Nozome Yachie (Institute for Advanced Biosciences, Keio University) for their technical assistance with computer prediction of the RNA secondary structures. We also thank Dr Gota Kawai and Dr Rintaro Saito for their helpful discussions, and Ms Xiaojing Zhang for critical reading of the manuscript. This research was supported in part by grants from: the Project for Development of a Technological Infrastructure for Industrial Bioprocesses on Research and Development of New Industrial Science and Technology Frontiers; NEDO (New Energy and Industrial Technology Development Organization); METI (Ministry of Economy, Trade and Industry), Japan; a Grant-in-Aid for Scientific Research on Priority Areas and a Grant-in-Aid from the 21st Century COE (Center of Excellence) Program, entitled 'Understanding and Control of Life's Function via Systems Biology (Keio University)'; the Computer Simulation Project, Ministry of Education, Culture, Sport, Science and Technology, Japan; and Keio University.

REFERENCES

- Kornberg, A. and Baker, T. A. (1992) Biosynthesis of DNA precursors. In *DNA Replication*, 2nd edn, pp. 53–100. W.H. Freeman and Company, New York.
- Liu, J., Schmitz, J. C., Lin, X., Tai, N., Yan, W., Farrell, M., Bailly, M., Chen, T. and Chu, E. (2002) Thymidylate synthase as a translational regulator of cellular gene expression. *Biochim. Biophys. Acta* **1587**, 174–182.
- Myllykallio, H., Lipowski, G., Leduc, D., Filee, J., Forterre, P. and Liebl, U. (2002) An alternative flavin-dependent mechanism for thymidylate synthesis. *Science* **297**, 105–107.
- Dynes, J. L. and Firtel, R. A. (1989) Molecular complementation of a genetic marker in *Dictyostelium* using a genomic DNA library. *Proc. Natl. Acad. Sci. U.S.A.* **86**, 7966–7970.
- Leduc, D., Graziani, S., Lipowski, G., Marchand, C., Le Marechal, P., Liebl, U. and Myllykallio, H. (2004) Functional evidence for active site location of tetrameric thymidylate synthase X at the interphase of three monomers. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 7252–7257.

- 6 Murzin, A. G. (2002) Biochemistry. DNA building block reinvented. *Science* **297**, 61–62
- 7 Chu, E., Koeller, D. M., Casey, J. L., Drake, J. C., Chabner, B. A., Elwood, P. C., Zinn, S. and Allegra, C. J. (1991) Autoregulation of human thymidylate synthase messenger RNA translation by thymidylate synthase. *Proc. Natl. Acad. Sci. U.S.A.* **88**, 8977–8981
- 8 Chu, E., Voeller, D., Koeller, D. M., Drake, J. C., Takimoto, C. H., Maley, G. F., Maley, F. and Allegra, C. J. (1993) Identification of an RNA binding site for human thymidylate synthase. *Proc. Natl. Acad. Sci. U.S.A.* **90**, 517–521
- 9 Kanai, A., Oida, H., Matsuura, N. and Doi, H. (2003) Expression cloning and characterization of a novel gene that encodes the RNA-binding protein FAU-1 from *Pyrococcus furiosus*. *Biochem. J.* **372**, 253–261
- 10 Dandekar, T. (2002) *RNA Motifs and Regulatory Elements*, 2nd edn, Springer-Verlag, Berlin
- 11 Sato, A., Kanai, A., Itaya, M. and Tomita, M. (2003) Cooperative regulation for Okazaki fragment processing by RNase HII and FEN-1 purified from a hyperthermophilic archaeon, *Pyrococcus furiosus*. *Biochem. Biophys. Res. Commun.* **309**, 247–252
- 12 Zuker, M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* **31**, 3406–3415
- 13 Kuhn, P., Lesley, S. A., Mathews, I. I., Canaves, J. M., Brinen, L. S., Dai, X., Deacon, A. M., Elsliger, M. A., Eshaghi, S., Floyd, R. et al. (2002) Crystal structure of thy1, a thymidylate synthase complementing protein from *Thermotoga maritima* at 2.25 Å resolution. *Proteins* **49**, 142–145
- 14 Fuglsang, A. (2004) Compositional nonrandomness upstream of start codons in archaeobacteria. *Gene* **332**, 89–95
- 15 Sartorius-Neef, S. and Pfeifer, F. (2004) *In vivo* studies on putative Shine–Dalgarno sequences of the halophilic archaeon *Halobacterium salinarum*. *Mol. Microbiol.* **51**, 579–588
- 16 Chu, E., Takimoto, C. H., Voeller, D., Grem, J. L. and Allegra, C. J. (1993) Specific binding of human dihydrofolate reductase protein to dihydrofolate reductase messenger RNA *in vitro*. *Biochemistry* **32**, 4756–4760
- 17 Liu, X., Reig, B., Nasrallah, I. M. and Stover, P. J. (2000) Human cytoplasmic serine hydroxymethyltransferase is an mRNA binding protein. *Biochemistry* **39**, 11523–11531
- 18 Berg, J. M., Tymoczko, J. L. and Stryer, L. (2002) *Nucleotide biosynthesis*. In *Biochemistry*, 5th edn, pp. 693–714, W.H. Freeman and Company, New York
- 19 Thompson, J. D., Higgins, D. G. and Gibson, T. J. (1994) ClustalW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673–4680

Received 14 April 2005/19 September 2005; accepted 21 September 2005

Published as BJ Immediate Publication 21 September 2005, doi:10.1042/BJ20050608