

SCIENTIFIC REPORTS

OPEN

Contribution of protein Gar1 to the RNA-guided and RNA-independent rRNA: Ψ -synthase activities of the archaeal Cbf5 protein

Ryosuke Fujikane^{1,2,4}, Isabelle Behm-Ansmant¹, Anne-Sophie Tillault^{1,5}, Christine Loegler¹, Valérie Igel-Bourguignon¹, Evelyne Marguet², Patrick Forterre^{2,6}, Christiane Branlant¹, Yuri Motorin^{1,3} & Bruno Charpentier¹

Archaeal RNA:pseudouridine-synthase (PUS) Cbf5 in complex with proteins L7Ae, Nop10 and Gar1, and guide box H/ACA sRNAs forms ribonucleoprotein (RNP) catalysts that insure the conversion of uridines into pseudouridines (Ψ s) in ribosomal RNAs (rRNAs). Nonetheless, in the absence of guide RNA, Cbf5 catalyzes the *in vitro* formation of Ψ_{2603} in *Pyrococcus abyssi* 23S rRNA and of Ψ_{55} in tRNAs. Using gene-disrupted strains of the hyperthermophilic archaeon *Thermococcus kodakarensis*, we studied the *in vivo* contribution of proteins Nop10 and Gar1 to the dual RNA guide-dependent and RNA-independent activities of Cbf5 on 23S rRNA. The single-null mutants of the *cbf5*, *nop10*, and *gar1* genes are viable, but display a thermosensitive slow growth phenotype. We also generated a single-null mutant of the gene encoding Pus10, which has redundant activity with Cbf5 for *in vitro* formation of Ψ_{55} in tRNA. Analysis of the presence of Ψ s within the rRNA peptidyl transferase center (PTC) of the mutants demonstrated that Cbf5 but not Pus10 is required for rRNA modification. Our data reveal that, in contrast to Nop10, Gar1 is crucial for *in vivo* and *in vitro* RNA guide-independent formation of Ψ_{2607} (Ψ_{2603} in *P. abyssi*) by Cbf5. Furthermore, our data indicate that pseudouridylation at orphan position 2589 (2585 in *P. abyssi*), for which no PUS or guide sRNA has been identified so far, relies on RNA- and Gar1-dependent activity of Cbf5.

Pseudouridine (Ψ) is the most abundant RNA modification (for review^{1–3}). This modified nucleoside is generated by site-specific C5-ribosyl isomerization of uridine (U) in pre-existing RNA transcripts. Formation of Ψ in RNAs can increase base stacking⁴ and rigidify the sugar-phosphate backbone^{1,5,6}. More specifically, Ψ can stabilize RNA 3D conformation by water bridge formation, as demonstrated in tRNAs^{7,8}. A cluster of Ψ s can modulate the conformation of a large RNA domain as shown for helix 69 – H69 – in rRNA⁹. It is now firmly established that some evolutionarily conserved Ψ s in tRNAs and rRNAs are required for accurate function of these non-coding RNAs (ncRNAs) in translation (for review², and^{10–20}). Similarly, some of the Ψ s in U-rich small nuclear RNAs (UsnRNAs) are important for pre-mRNA splicing^{2,21,22}. Moreover, recent transcriptome-wide maps of pseudouridylation detected a large number of Ψ s in eukaryotic mRNAs^{23,24} but their precise function is not yet firmly established.

The conversion of uridine (U) into Ψ is catalyzed by specific RNA: Ψ -synthases (PUS) (for review²⁵). Five families of PUS have been defined based on protein sequence conservation with each of the bacterial TruA, TruB, TruD, RluA, and RsuA proteins (for review²⁵). Comparative analysis of eukaryotic genomes revealed the presence of members of the same five families of PUS in Eukarya. Studies on archaea identified a rather limited

¹Université de Lorraine, CNRS, UMR 7365 Ingénierie Moléculaire et Physiopathologie Articulaire (IMoPA), F-54500, Nancy, France. ²Institute for Integrative Biology of the Cell (I2BC), CEA, CNRS, Université Paris-Sud, Université Paris-Saclay, F-91198, Gif-sur-Yvette cedex, France. ³Université de Lorraine, CNRS, INSERM, UMS2008 IBSLor, F-54500, Nancy, France. ⁴Present address: Fukuoka Dental College, Department of Physiological Sciences and Molecular Biology, Section of Cellular and Molecular Regulation, 2-15-1 Tamura, Sawara-ku, Fukuoka, 814-0193, Japan. ⁵Present address: Department of Biological Sciences, University of Lethbridge, Lethbridge, Alberta, Canada. ⁶Institut Pasteur, Département de Microbiologie, 25 rue du Dr Roux, F-75005, Paris, France. Correspondence and requests for materials should be addressed to B.C. (email: bruno.charpentier@univ-lorraine.fr)

number of putative genes encoding PUS, while in addition to PUS homologous to bacterial PUS, a sixth family of PUS was detected in eukaryotes. This sixth family of PUS has an archaeal representative – Pus10 of the archaeon *Pyrococcus abyssi*^{26,27}.

In archaea and eukaryotes, the catalysts for Ψ synthesis are either stand-alone PUS proteins or ribonucleoprotein particles with PUS activities that are referred to as box H/ACA small RNPs (sRNPs) in archaea and as box H/ACA small nucleolar RNPs (snoRNPs) in eukaryotes (for review^{25,28,29}). In archaea, box H/ACA sRNPs are formed by association of distinct H/ACA sRNAs with a constant set of four proteins including the TruB-like PUS aCBF5, the ribosomal protein L7Ae, and proteins aNOP10 and aGAR1 (thereafter named Cbf5, L7Ae, Nop10, and Gar1, respectively)^{30,31}. In absence of sRNA, protein complexes can be formed in which Nop10 and Gar1 interact independently with the catalytic domain of Cbf5^{32–34}. The H/ACA sRNAs act as guide RNAs ensuring the specificity of the RNP enzyme. Each H/ACA sRNA contains at least one hairpin-tail structure with a large internal loop, which forms base-pair interactions with a target RNA sequence. Two duplexes are then formed on each side of the unpaired U to be modified by Cbf5. Cbf5 interacts together with the box H/ACA sRNA and proteins Nop10 and Gar1^{31,32,35–37}. The Gar1 protein is dispensable for total modification of substrate RNAs in single-turnover reactions³¹, but it enhances the kinetics of the reaction^{30,31,38,39}. By modulating the formation of open and closed states of the catalytic center, Gar1 facilitates substrate RNA loading, its efficient placement and the release of the modified RNA^{38,40–42}.

The 2'-O-methylations and U to Ψ conversions in RNAs have been suggested to play a role in the capacity of archaea to grow at high temperatures. Current knowledge of RNA modification in archaea showing a relatively low number of PUS encoding genes compared to other organisms raises the question of the enzymatic mechanism required for the generation of the rather large number of Ψ s present in RNAs of this kingdom. The use of both classical protein catalysts and H/ACA sRNP catalysts and the multi-functionality and wide specificity of PUS in archaea⁴³ may limit the number of enzymes involved in archaeal RNA modification. However, as some of the information in this field was obtained through *in vitro* assays, complementary *in vivo* experiments are needed to elucidate this puzzling question. For instance, Cbf5, which is the only representative of the TruB family in archaea and acts as the catalytic subunit of the H/ACA sRNPs^{30,31}, may also function as a stand-alone enzyme. *In vitro* assays using the *P. abyssi* Cbf5 showed that it is able to catalyze Ψ_{55} formation in the loop of the T Ψ C arm of elongator tRNAs^{26,44–49}. One argument against this hypothesis is the finding that Cbf5 is not essential for the generation of this modification *in vivo* in *Haloferax volcanii*, in which this activity is carried out by the Pus10 PUS^{26,50–54}. Nevertheless, the possibility that the tRNA: Ψ_{55} -synthase activity of Cbf5 is used in other archaeal species cannot be excluded. Furthermore, Cbf5 may have a stand-alone activity in modification of rRNA. Indeed, despite of a deep computer analysis, we failed to identify any H/ACA sRNA able to guide formation of the two Ψ s respectively detected at positions 2585 and 2603 in the *P. abyssi* in 23S rRNA⁵⁵, and we found that recombinant Cbf5 in presence of proteins Nop10 and Gar1 can modify *in vitro* residue U₂₆₀₃ in a transcribed 23S rRNA fragment that can fold into a stem-loop structure showing some similarity with the T Ψ C arm of tRNAs^{48,55}. As this T Ψ C-like-loop structure has not been predicted in classical 23S rRNA folding models, the possible RNA-independent activity of Cbf5 in archaeal 23S rRNA modification still remains to be demonstrated *in vivo*. Of note, the previous data did not reveal any possible activity of either Cbf5 alone or of the Cbf5–Nop10–Gar1 protein complex at position U₂₅₈₅, raising the question of the catalyst used at this position⁵⁵.

Here, we describe successful isolation of *T. kodakarensis* single knockout mutants $\Delta cbf5$, $\Delta nop10$, and $\Delta gar1$ of the ORFs encoding the Cbf5, Nop10, and Gar1 components of H/ACA sRNPs, as well as a Pus10 single knockout mutant $\Delta pus10$. Obtaining these mutants indicates that each of these proteins taken separately is not essential for viability. However, growth of *T. kodakarensis* at high temperatures was markedly slowed down. The null mutants were used to identify the catalysts responsible for *in vivo* pseudouridylation at the orphan positions U₂₅₈₉ (U₂₅₈₅ in *P. abyssi*) and U₂₆₀₇ (U₂₅₈₅ in *P. abyssi*) in 23S rRNA. Altogether, our data reveal a general function for Gar1 in the RNA guide-dependent and RNA-independent activities of Cbf5.

Results

Construction of RNA: Ψ -synthases and H/ACA RNP components gene deletion mutants. We performed sequence analysis of the *T. kodakarensis* genome⁵⁶ in the aim of identifying ORFs for the PUS Cbf5 and Pus10, and for protein components L7Ae, Nop10, and Gar1 of the H/ACA RNPs. The protein-protein BLAST program using the amino acid sequences of *P. abyssi* PUS and H/ACA sRNP protein components as queries identified a single ortholog of Cbf5 (TK1509), Pus10 (TK0903), L7Ae (TK1311), Nop10 (TK1101), and Gar1 (TK2286) in *T. kodakarensis* genome (Fig. 1A and Supplementary data, Tables S1 and S2). To analyze how those proteins contribute to pseudouridylation of RNAs *in vivo*, we disrupted their genes using the gene targeting method (Supplementary data, Fig. S1)^{57,58}. The gene disruption vectors harboring the selectable marker *pyrF* (TK2276), which encodes orotidine-5'-monophosphate decarboxylase, and 1 kb of the 5' and 3' flanking regions of each target gene were constructed as described in Methods. They were used to transform the *T. kodakarensis* strain KU216 showing uracil auxotrophy due to a non-functional *pyrF* locus⁵⁹. Transformants harboring integration of the plasmid into the chromosome through homologous recombination were selected from the KU216 strain for uracil autotrophy. The intermediary *pyrF*⁺ strains were then selected by 5-FOA resistance to obtain gene knockout mutant strains⁵⁸. Using this strategy, we obtained colonies for the single knockout mutants $\Delta cbf5$, $\Delta pus10$, $\Delta nop10$, and $\Delta gar1$, indicating that, taken separately, none of these genes were essential for cell viability. However, this was not the case for the ORF encoding the ribosomal protein L7Ae, for which we did not succeed in isolating a viable disrupted mutant.

In order to validate the deletion of desired ORFs in mutant *T. kodakarensis* strains and to exclude potential off-target integration of *pyrF* gene used in knockout strategy, we performed whole genome sequencing and analysis of generated *T. kodakarensis* mutants. Fragmented genomic DNA extracted from each mutant was converted to the library and sequenced on Illumina HiSeq1000. Sequencing reads (~0.6–1 mln/library, 20–30x

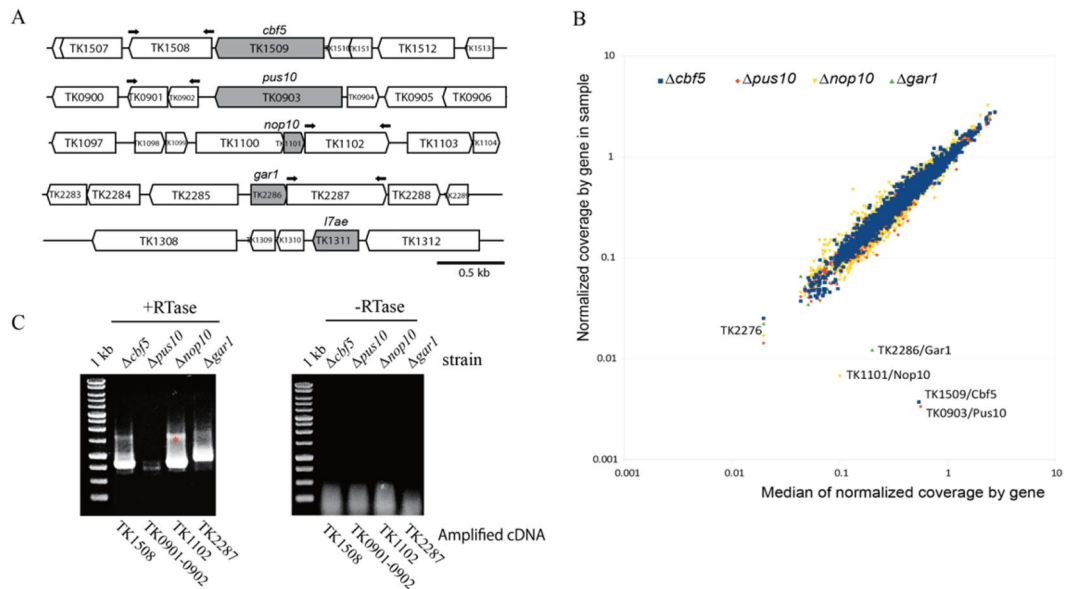


Figure 1. Construction of disruption mutants. **(A)** Schematic diagram of the loci targeted for disruption. The target genes are in gray. The black arrows indicate the primers for RT-PCR analysis (panel C and Table S3). **(B)** Normalized reads' coverage as a function of the median coverage observed for all four deleted *T. kodakarensis* strains. Colors correspond to the coverage observed for Δ TK1509/*cbf5* (blue), Δ TK0903/*pus10* (red), Δ TK1101/*nop10* (yellow) and Δ TK2286/*gar1* (green) strains. Identity of outlier points is indicated. TK2276 (*pyrF*) is affected in all four strains analyzed. **(C)** Disruption of each targeted ORFs does not impair transcription of their respective downstream contiguous ORF. RT-PCR reactions (gel + RTase) were performed with oligonucleotide primers (Table S3) complementary to the coding sequence of the target gene's adjacent ORFs (indicated below each lane). The RT-PCR products were fractionated by 0.8% agarose gel electrophoresis. The bands correspond to RNA amplification since no amplified products were obtained in a reaction mixture lacking reverse transcriptase (-RTase). The asterisk indicates non-specific amplification.

mean coverage) were aligned to the reference genome and coverage was calculated by each *T. kodakarensis* ORF. Figure 1B shows the normalized gene coverage versus the median value for all four sequenced strains. As clearly shown on the graph, only a single point for each strain demonstrates strong deviation from the median, those correspond to the expected ORFs deleted in the respective strain. As anticipated, ORF TK2276 (alias gene *pyrF*) is absent in all four strains used. More detailed information on the deletions is given in Supplementary data (% of covered positions by gene and sequencing reads aligned in the regions of expected deletions in Fig. S2). Altogether these data unambiguously confirm that only one ORF in each *T. kodakarensis* mutant was affected by deletion and no undesired off-target integration/popping out of *pyrF* gene occurred.

Since in archaeal genomes, many genes are organized in clusters or in operons⁶⁰, it was possible that the individual ORF disruption affected the structure of the operon or of the cluster. As a consequence, the transcription of downstream ORFs or genes may also be affected. To check for this possibility, we isolated total RNAs from the null mutant cells and performed RT-PCR to detect mRNA produced from the downstream gene. Bands corresponding to transcribed products from the downstream genes were detected (Fig. 1C), indicating that the gene deletions did not affect transcription of the downstream genes. Hence, the observed phenotype for each null mutant was expected to be only related to the lack of the disrupted ORF function.

Growth phenotype of the deletion mutant strains. The isolated mutant strains were cultivated at 85 °C, which is the optimal growth temperature for *T. kodakarensis*⁶¹, as well as at 75 °C and 90 °C. The number of growing cells as a function of time was counted with Thoma's cell counter under an optical microscope (Fig. 2). At 75 °C and at 85 °C, the mutant strains Δ *pus10*, Δ *nop10*, and Δ *gar1* showed visible growth retardation compared to the wild type strain and the final yields of cells were only three-fourths for Δ *gar1* and two-thirds for Δ *pus10* and Δ *nop10*. Remarkably, in Δ *pus10* and Δ *nop10* strains, cellular aggregation was observed after 4 hours cultivation (Supplementary data, Fig. S3). Aggregation of archaeal cells has already been observed in *Sulfolobus* after UV irradiation^{62,63}, and in *Halobacterium* after heat shock or divalent ion treatment⁶⁴. This may suggest that the disruption of *nop10* or *pus10* genes causes some cellular stress leading to cellular aggregation. As proteins Nop10 and Gar1 are known to contribute to Cbf5 activity and as no apparent phenotype was associated with disruption of the *cbf5* gene at suboptimal and optimal growth temperatures, one possible explanation is that Nop10 and Gar1 have additional cellular functions that are not connected with RNA modification.

All the mutant strains showed significant growth defects at 90 °C. These data suggest that pseudouridines synthesized by Cbf5, Pus10, and the H/ACA sRNPs contribute to rRNA and/or tRNA stabilization at high temperatures and/or are even important for ribosome biogenesis and activity.

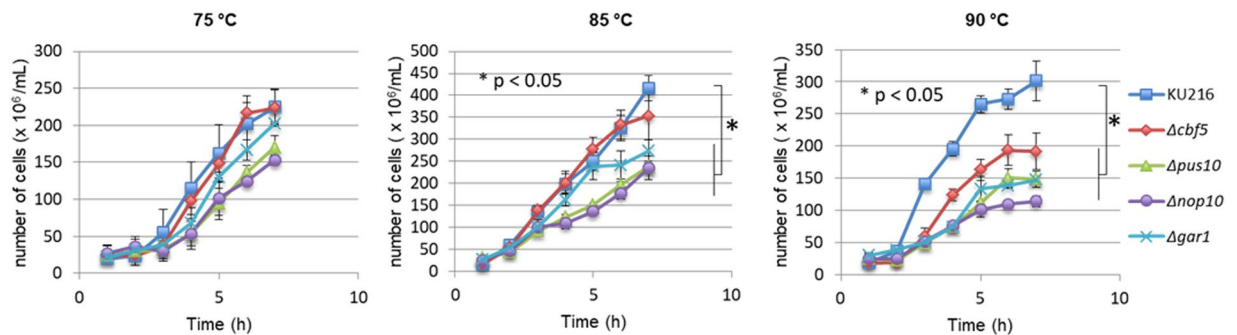


Figure 2. Analysis of the effect of gene disruption on cell growth. The various *T. kodakarensis* wild type (KU216) and mutant strains were grown in ASW-YT medium. Cultures were carried out at the optimal growth temperature of 85 °C, or at 75 °C and 90 °C, and cell growth was monitored by counting the number of cells with Thoma's cell counter during the course of the culture. The mean values obtained from three independent experiments and the standard deviations are shown with bars. Significance values of * $p < 0.05$ relative to KU216. square: KU216, diamond: $\Delta cbf5$, triangle: $\Delta pus10$, circle: $\Delta nop10$, and cross: $\Delta gar1$.

Cbf5 but not Pus10 is required for $\Psi_{2589(2585)}$ and $\Psi_{2607(2603)}$ formation in 23S rRNA. We next conducted phenotypic analysis of the null mutants by looking for the presence of Ψ s in 23S rRNA using the CMCT-RT approach. We focused this more extensive analysis on the Domain V region of 23S archaeal rRNA spanning residues 2588 to 2612 (Fig. 3A). This region contains three Ψ residues at positions 2589, 2592, and 2607. Pseudouridine at position 2607 (2603 in *P. abyssi*) can be formed in a RNA guide-independent way⁵⁵, while Ψ_{2589} (Ψ_{2585} in *P. abyssi*) still have no associated RNA guide and complex Cbf5–Nop10–Gar1 is not capable of forming Ψ_{2585} *in vitro*⁵⁵. RNA guide-dependent residue Ψ_{2592} (Ψ_{2588} in *P. abyssi*) was used as control.

Total RNAs extracted from the various strains were treated with CMCT as described in Methods and alkaline resistant Ψ –CMC adducts were detected in the region analyzed by primer extension analysis, at the expected positions: nucleotides U_{2589} , U_{2592} , and U_{2607} (corresponding to U_{2585} , U_{2588} , and U_{2603} in *P. abyssi*) (Fig. 3A,B, lanes 3 and 4). This analysis complements our previous study on *P. abyssi* 23S rRNA⁵⁵, and reveals the presence of a highly conserved modification at these positions in both *T. kodakarensis* and *P. abyssi* 23S rRNAs. None of the three modifications occurred in the absence of Cbf5 whereas all of them were present in the *pus10*-null strain (Fig. 3B, compare lanes 7 and 8 with lanes 19 and 20). These data demonstrated that Cbf5 and not Pus10 activity is required *in vivo* for Ψ formation at these three positions.

Nop10 is dispensable but Gar1 is crucial for $\Psi_{2607(2603)}$ formation in 23S rRNA. The *P. abyssi* heterotrimer complex Cbf5–Nop10–Gar1 was previously found to modify *in vitro* residue U_{2603} when it is contained in an RNA substrate composed of a 22-nucleotide long fragment of 23S rRNA⁵⁵. Surprisingly, in the present *in vivo* study, we observed that the disruption of ORFs *nop10* or *gar1* in *T. kodakarensis* had a different effect on 23S rRNA modification at the corresponding position U_{2607} . The modification Ψ_{2607} was still detected in mutant $\Delta nop10$, but was no longer detected in mutant $\Delta gar1$ (Fig. 3B, compare lane 12 with lane 16). These data show that Ψ_{2607} formation *in vivo* requires the presence of Gar1 but does not require the presence of Nop10. In contrast, both Ψ_{2589} and Ψ_{2592} were absent in all three strains with deletions of *cbf5*, *nop10*, or *gar1* genes (Fig. 3B).

In a previous work, we showed that the *P. abyssi* Pab40 H/ACA sRNA guides modification of 23S rRNA at position U_{2588} *in vitro*⁵⁵. Modification of the equivalent nucleotide U_{2592} in *T. kodakarensis* 23S rRNA is likely to be also guided by an H/ACA sRNA since the gene sequence for a Pab40 homolog (Tko1) is present in the *T. kodakarensis* genome⁵⁵. The absence of Nop10 or Gar1 has the same effect as the absence of Cbf5, since it completely stopped the formation of Ψ_{2592} in 23S rRNA (Fig. 3B, lanes 11 and 12, and lanes 15 and 16). These data are in agreement with the implication of proteins Nop10 and Gar1 in the activity of a conserved H/ACA RNP catalyzing $\Psi_{2592(2588)}$ formation.

To further pursue our investigation of the role of the auxiliary proteins, we compared single-turnover activities of the free Cbf5 protein, and the Cbf5–Nop10 (CN), Cbf5–Gar1 (CG), and Cbf5–Nop10–Gar1 (CNG) protein mixtures (Fig. 3C), for *in vitro* modification of an excess of the 23S rRNA fragment mini-23S–2607 (Fig. 3A). Formation of Ψ_{2607} was observed upon incubation of the labeled mini-23S–2607 RNA with Cbf5–Gar1 and Cbf5–Nop10–Gar1, but not with Cbf5 alone or in association with Nop10. These *in vitro* data confirm that Gar1, but not Nop10, is required for the non-RNA-guided activity of Cbf5 at position 2607(2603) in 23S rRNA.

Contribution of Gar1 to H/ACA RNP enzyme activity. As the function of Gar1 in H/ACA sRNP activity had only been studied using short RNA fragments, we tested its activity on a long RNA substrate mimicking the 23S rRNA peptidyl transferase center (PTC). We engineered a DNA template allowing transcription of a 143-nucleotide long RNA fragment containing a large portion of the PTC domain (Fig. 4A). This substrate RNA designated 23S–143nt carries the U at position 2690(2685) in 23S rRNA that is targeted *in vivo* by the Pab91 sRNP, which was well characterized in our previous studies^{31,35,55,65}. We verified that this substrate was modified *in vitro* at the target site by the *P. abyssi* Pab91 LCN enzyme under single-turnover conditions albeit less efficiently than the LCNG enzyme (Fig. 4B). However, the modification turnover was considerably diminished in the absence of Gar1 (Fig. 4C,D). The complex formed between the reconstituted Pab91 sRNP LCN and the

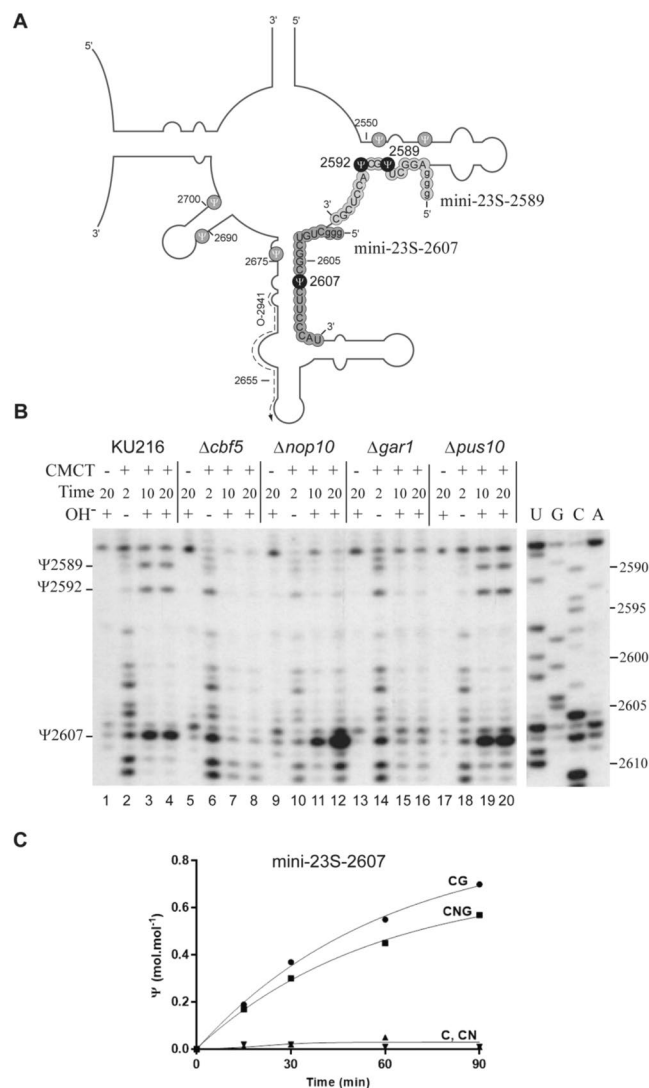


Figure 3. The effect of gene disruption on Ψ formation in the domain V of *T. kodakarensis* 23S rRNA. **(A)** Ψ s identified in this work in the *T. kodakarensis* 23S rRNA, which are conserved in the *P. abyssi* 23S rRNA are indicated by black dots. The grey dots correspond to other Ψ s detected in *P. abyssi* 23S rRNA in a previous work⁵⁵. The secondary structure of the *T. kodakarensis* 23S rRNA is adapted from Piekna *et al.*¹⁶. The sequence of the two substrates mini-23S-2607 and mini-23S-2589, synthesized by *in vitro* transcription and used *in vitro* in Panel C (mini-23S-2607) and in Fig. 5 (mini-23S-2589), is highlighted. **(B)** Identification of Ψ s in the *T. kodakarensis* 23S rRNA by the CMCT-RT method. As indicated at the top of each lane, total RNAs extracted from the various *T. kodakarensis* strains were treated in the absence (–) or the presence of CMCT (+), for 2, 10, and 20 min as described in Methods. The CMCT treatment was (+) or was not (–) followed by an alkaline treatment at pH 10.4. The positions of Ψ s were identified by primer extension analysis using oligonucleotide O-2941 (represented by the dashed arrow) as described in Methods. Digital image of the autoradiogram was obtained by scanning the x-ray film. The presence of Ψ s is revealed by the appearance of alkali-resistant RT stops. Lanes U, G, C, and A correspond to the RNA sequencing ladder. The positions of nucleotides in 23S rRNAs are indicated to the left of the autoradiogram. The two panels correspond to a cropping of two sections of the same autoradiogram. The full-length gel is presented in Supplementary Figure S4. **(C)** Time course analysis of *in vitro* modification by recombinant proteins in RNA substrate mini-23S-2607. The RNA substrate was radiolabeled by the incorporation of [α -³²P]CTP during *in vitro* transcription. The RNA was incubated with different protein combinations: Cbf5 alone (C), Cbf5 and Nop10 (CN), Cbf5, and Gar1 (CG), and a set of the three proteins (CNG). At each time point, an aliquot of the reaction mixture was analyzed by 2D-TLC. The radioactivity was quantified by PhosphorImager analysis. The quantities of Ψ nucleotides formed are expressed in moles per mole of mini-23S-2607.

radiolabeled 23S-143nt RNA substrate (complex CII') was analyzed by electrophoretic mobility shift analysis³¹. Binding of this long substrate is efficient as high amounts of complex CII' are formed (Fig. 4E, lane 4). The steady amounts of complex CII' observed after 60 min of incubation (Fig. 4E, lane 5), are consistent with the absence of an efficient turnover of the RNP enzyme lacking Gar1.

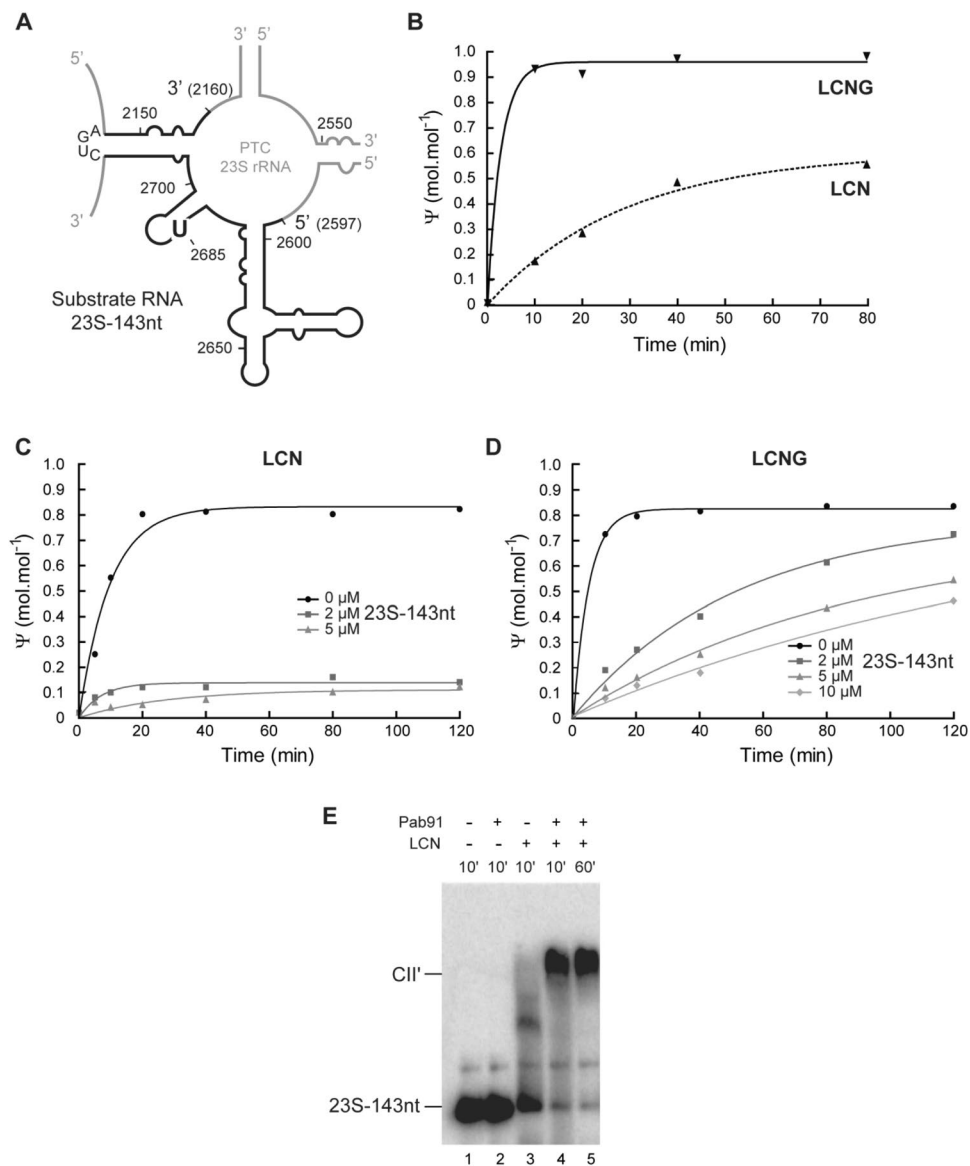


Figure 4. *In vitro* activity of *P. abyssi* Pab91 RNP enzyme for formation of Ψ_{2685} in a 143-nucleotide-long rRNA substrate. **(A)** The RNA substrate mimicking the peptidyl transferase center (PTC) region of the 23S rRNA. The secondary structure model of the *P. abyssi* PTC is represented. The U at position 2685 targeted by the Pab91 RNP is indicated. The region of RNA substrate 23S-143nt is delineated by the black line. The loop CUGA replaced a large portion of 23S rRNA corresponding to domains I, II, III, IV, VI and is used as a linker between the 5' and 3' parts of domain V. **(B)** Single-turnover activity of the *P. abyssi* Pab91 LCN and LCNG RNP enzymes for modification of radiolabeled 23S-143nt at 65 °C. The ³²P was introduced in the phosphodiester bond preceding U₂₆₈₅ by a splinted ligation (detailed in Material and Methods). **(C-D)** Multiple-turnover activity of the LCN **(C)**, and LCNG **(D)** enzymes for modification of radiolabeled 22-U substrate RNA at 65 °C. The unlabeled substrate RNA 23S-143nt was added in a 4-fold (2 μ M), 10-fold (5 μ M), or 20-fold (10 μ M) excess over the RNP (~0.5 μ M). A control was performed in absence of the unlabeled RNA (0 μ M). **(E)** Electrophoretic mobility shift analysis (EMSA) of the binding of the radiolabeled substrate 23S-143nt with the Pab91 sRNP assembled with the L7Ae, Cbf5, and Nop10 mix. Incubation was performed at 65 °C during 10 and 60 minutes.

Modification at position U2589(2585) requires an RNA component and the H/ACA RNP proteins including Gar1. At this stage of the study, the modification $\Psi_{2589(2585)}$ in 23S rRNA remained enigmatic. CMCT-RT analysis of rRNAs extracted from the Δ *cbf5* mutant indicated that Cbf5 is implicated in the formation of this modification (Fig. 3B, lanes 7 and 8), and that both Nop10 (Fig. 3B, lanes 11 and 12) and Gar1 (Fig. 3B, lanes 15 and 16) are required for this activity. The modification could either result from the non-RNA-guided activity of Cbf5 in complex with the two auxiliary proteins, or from the RNA-guided activity of Cbf5 present in a H/ACA sRNP enzyme with a still unknown RNA guide. In order to distinguish between these two possibilities, we tested (Fig. 5) the *in vitro* activity of the recombinant Cbf5 in the modification of the mini-23S-2589 RNA, whose sequence includes residues 2583 to 2599 of the 23S rRNA (Fig. 3A). Incubation of the radiolabeled

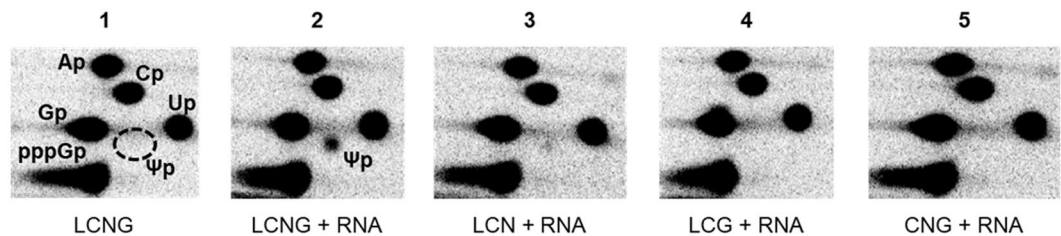


Figure 5. *In vitro* activity tests suggesting that Cbf5's RNA dependent activity modifies orphan position $U_{2589(2585)}$. The mini-23S-2589 RNA fragment was radiolabeled by incorporation of $[\alpha\text{-}^{32}\text{P}]\text{GTP}$ during *in vitro* transcription. This substrate was incubated with different combination of the four recombinant proteins from *T. kodakarensis* L7Ae (L), Cbf5 (C), Nop10 (N), and Gar1 (G) and total RNAs extracted from a S100 fraction from KU216 cells (+RNA). After digestion with RNase T2, the amount of Ψ formation was estimated by two-dimension thin layer chromatography as described in Methods. The control reaction (panel 1) was performed in the presence of LCNG but without the addition of total RNAs. The plates corresponding to each panel were exposed together on a same screen of the PhosphorImager. The full-length image is presented in Supplementary Figure S5.

mini-23S-2589 with the combination of the four recombinant proteins Cbf5, Nop10, Gar1, and L7Ae did not allow formation of Ψ . However, we observed partial conversion of U_{2589} into Ψ ($\Psi \sim 0.14 \text{ mol}^{-1} \cdot \text{mol}^{-1}$) when total RNAs extracted from *T. kodakarensis* cells were added to the protein mix. This activity requires the presence of the full set of proteins Cbf5, Nop10, Gar1, and L7Ae. We hypothesize that modification at position U_{2589} in 23S rRNA is generated by an RNP that is assembled by association of proteins L7Ae, Cbf5, Nop10, and Gar1 with a still unknown guide sRNA present in the *T. kodakarensis* total RNAs.

Discussion

Genetic manipulation by gene disruption and functional complementation is an extremely valuable tool for the study of metabolic processes in a living cell. Such methods have been widely applied in bacterial or eukaryotic model organisms and have enabled the characterization of numerous modification enzymes acting on RNAs in *E. coli* and *S. cerevisiae*^{66,67}. However, archaeal organisms were not included in this thorough-full characterization since targeted gene deletion in most of those organisms is still a rather difficult task. Genetic manipulations have only been used in few archaeal species such as Thermococcales, Sulfolobales, Methanogens, and Halophiles⁶⁸⁻⁷². In the RNA modification research field, gene deletion has only been performed in halophilic *Haloferax volcanii* for tRNA modification enzymes⁷³, for 2'-O-methylation guide box C/D RNA⁷⁴ and for several PUS including Cbf5 and Pus10⁵⁰. Establishment of the *pyrF* KU216 *T. kodakarensis* strain and the use of *pyrF* as a marker to select uracil auxotrophic cells enabled the development of a strategy for targeted gene disruption in this sulfur reducing, strictly anaerobic hyperthermophilic organism^{57-59,75}. By applying this system to *T. kodakarensis*, we generated the first disruption mutants of the two PUS enzymes Cbf5 and Pus10, and of the Gar1 and Nop10 components of the catalytic H/ACA RNPs of a hyperthermophilic archaea. These mutants make it possible to analyze the activity of the modification machinery at high temperatures and to study the effect of RNA modifications in these peculiar growth conditions. Up to now, the *in vitro* specificity and turnover of the H/ACA sRNP enzymes from *P. abyssi* have been shown to rely on high temperatures³⁹ and *in vivo* data have not been documented.

In contrast to eukaryotes⁷⁶, but in agreement with observations made in halophilic archaeon *H. volcanii*⁵⁰, the *T. kodakarensis* gene encoding Cbf5 homolog is dispensable. However, deletion of *cbf5* led to a marked decrease in growth rate of *T. kodakarensis* cells at 90 °C (Fig. 2). It was previously observed that deletions of genes encoding modification enzymes belonging to the same PUS family as Cbf5, e.g. *truB* in *E. coli*^{67,77} or in the thermophilic eubacterium *Thermus thermophilus*⁷⁸, are not lethal either, but are associated with thermosensitive phenotypes. The thermosensitive slow growth phenotype of $\Delta cbf5$ mutant can be rationalized by considering that lack of modification of one subset of RNAs modified by Cbf5, e.g. tRNAs and/or rRNAs, results in reduced RNA stability or reduced ribosome activity at temperatures higher than the optimal growth temperature. It is also conceivable that Cbf5 has a function that is not related to pseudouridylation but that contributes to growth in these conditions. Surprisingly, deletions of the H/ACA RNP components Gar1 and Nop10 led to a slow growth phenotype at all the temperatures tested, i.e. 75 °C, 85 °C and 90 °C (Fig. 2). These data suggest that these proteins may have additional functions not associated with Cbf5 and H/ACA RNPs.

The null mutant of the PUS Pus10 showed marked phenotypes with a decrease of the growth rate at all the temperatures tested and formation of aggregates in liquid culture (Supplementary data, Fig. S3). These data suggest Pus10 contributes more than Cbf5 to cell growth at high temperatures. However, in contrast to *H. volcanii*⁵⁰, Pus10 from *T. kodakarensis* is not essential for cell viability. The essential role of Pus10 in *H. volcanii* may be either due to its tRNA modification activity, i.e. the formation of tRNA Ψ_{54} and Ψ_{55} , one or both of these modifications may be crucial for the cell viability of this halophilic archaeon. Alternatively, in this organism, Pus10 could also modify other cellular RNAs or have a function that is not related to pseudouridylation.

In the future, it would be of interest to extend the analysis of the null mutants produced to compare the Cbf5 and Pus10 activities at positions 54 and 55 in elongator tRNAs. This would make it possible to determine whether Cbf5 is simply needed to compensate for occasional Pus10 activity defects, as the tRNA Ψ_{55} nucleotide is extremely important to stabilize the tertiary structure of tRNA⁷⁹. Isolation of the double mutant $\Delta cbf5 \Delta pus10$ would also be helpful for these studies.

The present work confirms the rather large substrate specificity of Cbf5 in archaea. This multifunctional PUS modifies cellular RNAs using two independent, but complementary mechanisms: (i) an activity linked to its inclusion in H/ACA sRNPs complexes together with L7Ae, Nop10, and Gar1 and one H/ACA sRNA^{30,31}, (ii) as a protein stand-alone enzyme acting alone or in complex with auxiliary proteins Nop10 and/or Gar1⁵⁵. We previously demonstrated that in addition to tRNA U₅₅, U₂₆₀₃ in *P. abyssi* 23S rRNA represents a possible second substrate for the non RNA-guided activity of Cbf5 in *P. abyssi*⁵⁵. Here, we found that the modification is conserved in *T. kodakarensis* 23S rRNA, i.e. at position U₂₆₀₇ (Fig. 3B). Modification at this position likely implies that at some stage of its maturation, the rRNA sequence mimics the folding of the tRNA TΨC arm. Our comparison of the Ψ status in wild-type, Δ*cbf5* and Δ*pus10* null mutants clearly demonstrates (Fig. 3B) that in *T. kodakarensis*, there is no redundancy between the Cbf5 and Pus10 activities for modification of U₂₆₀₇ in 23S rRNA. It only relies on Cbf5.

The present data also enlarge the catalog of sites modified by Cbf5 to position U₂₅₈₉ in 23S rRNA. Our observation that total RNA fraction from *T. kodakarensis* restores Ψ₂₅₈₉ formation *in vitro* supports a possible role for an RNA component likely acting as a guide RNA for the activity of a catalytic RNP enzyme. Analysis of the 23S rRNA extracted from each disrupted mutant strain demonstrated the implication of each of the three *cbf5*, *nop10*, and *gar1* genes encoding the H/ACA sRNP components for Ψ₂₅₉₂ (used as control) as well as for Ψ₂₅₈₉ formation (Fig. 3B). We did not succeed in isolating a *l7ae* null mutant. This is not surprising, as in addition to being a component of H/ACA RNPs, L7Ae is a ribosomal protein. Nevertheless, we showed that L7Ae is required for Ψ₂₅₈₉ formation *in vitro* (Fig. 5). This adds credit to the assumption that a fully active RNP enzyme could be formed when L7Ae and the three other protein components Cbf5, Nop10, and Gar1 of box H/ACA sRNP assembled with a guide RNA present in the total RNA fraction extracted from *T. kodakarensis* cells. In a previous work, we identified six H/ACA sRNAs by bioinformatics analysis of the *P. abyssi* genome and found that none of the *in vitro* sRNPs assembled with each of these RNA guides catalyzes Ψ₂₅₈₅ formation, i.e. modification of the U equivalent to position 2589 in *T. kodakarensis* 23S⁵⁵. To identify the guide RNA potentially involved in Ψ₂₅₈₉ formation, we fractionated *T. kodakarensis* total RNAs on acrylamide gel, eluted RNA fractions from gel slices and tested if they could complement L7Ae, Cbf5, Nop10, and Gar1 for Ψ₂₅₈₉ formation *in vitro*. An active fraction was indeed identified and its RNA content was analyzed by RNASeq. However, no new RNAs with the known features of archaeal box H/ACA sRNAs were detected suggesting that the RNA required for the Ψ₂₅₈₉₍₂₅₈₅₎ formation could correspond to a new type of guide RNA. Several RNA candidates need to be individually tested in *in vitro* modification assays before generating a null mutant of the gene sequence of the identified RNA for *in vivo* analysis.

We show that both *in vivo* and *in vitro*, unlike Nop10, Gar1 is absolutely indispensable for the RNA-independent activity of Cbf5 at position 2607(2603) in archaeal 23S (Fig. 3B,C). The work by R. Gupta's team already outlined differential roles of *Methanocaldococcus jannaschii* Nop10 and Gar1 for Ψ₅₅ formation in tRNA by Cbf5 *in vitro*⁴⁴. Differential roles of Nop10 and Gar1 were also observed in H/ACA RNP activity. Nop10 is a key component of RNP architecture and activity^{30,31,35}. Gar1 helps position the thumb to open and close the active site of Cbf5³⁸. However, Gar1 is dispensable for total modification of an excess of substrate RNA in single-turnover reaction catalyzed by H/ACA RNPs³¹. Nevertheless, under these conditions the presence of Gar1 enhances the kinetics of the reaction^{30,31,38,39}. Interestingly, the critical contribution of Gar1 to the turnover of an H/ACA RNP is also evidenced by the activity test performed *in vitro* in this study on the large 23S rRNA fragment of domain V (Fig. 4). Gar1 may thus play equivalent role in Cbf5 RNA-dependent and RNA-independent activities. In particular, it may stimulate the turnover of Cbf5 for rRNA Ψ₂₆₀₇₍₂₆₀₃₎ formation. As an extrapolation, we can also suggest that the RNA-independent activity relies on the presence of Gar1 to facilitate the binding of Cbf5 to the RNA and/or to allow catalytic activity of the enzyme. Structural information on the [Cbf5:substrate RNA] complex formed in the presence and absence of Gar1 and comparison with the structure of the [H/ACA RNP:substrate RNA] complex³⁸ might provide further information on the exact roles played by Gar1 in the RNA-dependent and RNA-independent activities of Cbf5.

Conclusion

We present here the analysis of several proteins involved in pseudouridylation of ribosomal rRNA in *T. kodakarensis*. We demonstrate that Cbf5 is responsible for Ψ formation at multiple positions in the large subunit rRNA. Cbf5 performs this function as the catalytic subunit of RNPs and modifies at least one position with an RNA guide-independent activity. Pus10 does not appear to be involved in rRNA modification. Gar1 contributes to each of the activities of Cbf5 characterized *in vivo*.

Methods

Strain and culture. Cells of the *T. kodakarensis* strain KU216, lacking *pyrF* gene (KU216 *pyrF*⁻) were cultivated under anaerobic conditions at the optimal growth temperature of 85°C in a nutrient-rich medium (ASW-YT) containing 0.5% yeast extract, 0.5% tryptone and artificial sea water as described previously^{58,59}. Other growth temperatures (75°C and 90°C) were also used to study the phenotypic effects of the deletions. ASW-CH medium was used for the selection based on uracil autotrophy as described previously⁵⁷.

Sequence analysis and homology search. *T. kodakarensis* putative PUS genes were identified in a homology search with NCBI BLAST⁸⁰, using the amino acid sequence of *Pyrococcus abyssi* PUS Pab0356 (*cbf5*) and Pab2391 (*pus10*), and H/ACA sRNP components Pab7213 (*nop10*), Pab3084 (*gar1*), Pab0460 (*l7ae*) as query. Sequence analysis of *T. kodakarensis* was performed using the NCBI genomic database (<http://www.ncbi.nlm.nih.gov>) and the BAGET sequence exploration tool (<http://archaea.u-psud.fr/bin/baget.dll/EXEC>). To ascertain the identity of the identified ORFs, sequences of *T. kodakarensis* proteins were compared with already known and biochemically characterized PUS and H/ACA sRNP proteins from *P. abyssi*, *P. furiosus*, or *Haloferax volcanii*. Only Pus10/Cbf5 showed slight similarity, and Nop10 and L7Ae shared limited similarity with IF2-gamma and

L30e, respectively (see Supplementary data, Table S1). Inspection of the entire genome of *T. kodakarensis* revealed no orthologs of any other potential proteins (Supplementary data, Table S2).

Construction of the disruption vectors. The plasmids used for gene disruption in *T. kodakarensis* were constructed as follows. DNA fragments containing the sequence of the ORF to be deleted along with about 1 kb of each 5' and 3' flanking regions were amplified by PCR, using specific primer pairs (Primers F and R in Supplementary data, Table S3) and genomic DNA from the KU216 strain as template. The PCR-amplified fragments were cloned into the pGEM-T easy vector (Promega). Each recombinant clone was confirmed by full sequencing of the fragments. Inverse PCR reactions were performed with InvF and InvR primer pairs (Supplementary data, Table S3) to amplify the flanking regions of each target gene along with the vector backbone. The PCR fragments were ligated by T4 DNA ligase, giving a continuous DNA fragment containing both flanking sequences without the target gene. The resultant DNA fragments were excised with the appropriate restriction enzymes and inserted into the pUD2 vector. This vector is a pUC118-based plasmid harboring the *pyrF* gene as a nutrient marker⁵⁸.

Construction of the gene deletion mutants of *T. kodakarensis*. The knockout strains were obtained with the gene targeting method for *T. kodakarensis*^{57,58}. KU216 *pyrF*⁻ cells were transformed with each of the disruption vectors. After transformation, the *pyrF*⁺ transformants showing uracil prototrophy were selected by cultivating the cells in the ASW-CH medium lacking uracil. Then the *pyrF*⁺ cells were plated on ASW-YT containing 0.85% (w/v) of 5-fluoroorotic acid (5-FOA) for selection of uracil auxotrophic cells produced by internal chromosomal recombination. The colonies were picked and cultivated in 5 mL ASW-YT. Genomic DNA of each culture was isolated by phenol/chloroform extraction, followed by isopropanol precipitation. The genotypes of isolated strains were validated by whole genome sequencing of each mutant.

Whole genome sequencing of *T. kodakarensis* DNA. Genomic DNA from *T. kodakarensis* mutant strains was extracted following the standard procedure⁸¹. About 100 ng of DNA in 130 μ l of water was fragmented using Covaris M200 AFA ultrasonicator to fragments of ~200 bp. Sonication efficiency was confirmed by capillary electrophoresis using HS DNA chip (Agilent). DNA fragments were converted to sequencing libraries using NEBNext[®] Ultra[™] II DNA Library Prep Kit, according to the manufacturer's recommendations. Amplified libraries were verified by capillary electrophoresis and quantified by fluorometry using QuBit. Sequencing was performed in SR50 and/or PE2x100 mode, using Illumina HiSeq1000. After trimming of the adapter sequences, clean reads were aligned to the reference genome (GCF_Tkodakarensis_ASM996/ATGC280) using Bowtie2 in *local-sensitive* mode. Counting of the reads' coverage was done using bedtools and the associated gff3 file.

Total RNA isolation. Growth of *T. kodakarensis* cells was stopped at the end of the exponential phase. Total RNA was isolated as described previously⁸². The cells were centrifuged and washed in 0.8x artificial sea water. The pellets (~10⁹ cells) were dissolved in extraction buffer (4M guanidine thiocyanate, 25 mM sodium citrate pH 7; 0.5% *N*-lauroylsarcosine and 0.1 M β -mercaptoethanol) and RNAs extracted by phenol/chloroform/isoamyl alcohol followed by ethanol precipitation, dissolved in H₂O and stored at -80 °C. For RNAs extracted from the S100 fraction, the cells were resuspended in sonication buffer (25 mM Tris-HCl pH 7.5, 5 mM MgCl₂, 25 mM KCl, 2 mM DTT, and 10% glycerol), treated by ultra sounds, the concentration of KCl was adjusted to 150 mM before centrifugation at 16,000 \times g for 30 min. The S100 extract was obtained after centrifugation at 100,000 \times g for 1 hour and stored in 20% glycerol at -80 °C. RNAs from the S100 fraction were extracted with phenol/chloroform/isoamyl alcohol followed by ethanol precipitation, dissolved in H₂O and stored at -80 °C.

Mapping of pseudouridine (Ψ) residues in 23S rRNA. The *N*-cyclohexyl-*N'*-(2-morpholinoethyl)-carbodiimide metho-*p*-toluenesulfonate (CMCT) modification protocol⁸² was adapted from⁸³. Briefly, 10 μ g of total RNAs were modified by CMCT, which preferentially reacts with U and Ψ residues. Extended incubation of the CMC-modified RNAs at alkaline pH (10.4) allows the hydrolysis of U-CMC adducts, which are less stable than the Ψ -CMC adducts. The positions of Ψ -CMC adducts were then detected by primer extension using avian myeloblastosis virus (AMV) reverse transcriptase (MP Biomedicals) and 5'-end labeled primer O-2941 (5'-CGTTCCTTAAATGGGTGA-3') complementary to 23S rRNA sequence. The synthesized cDNAs were fractionated by electrophoresis on a 7% sequencing gel which was analyzed by autoradiography.

Recombinant protein production. The recombinant proteins Cbf5, Gar1, Nop10, and L7Ae were produced in strain *E. coli* C41 (DE3) pRare2 as GST-fusion proteins from recombinant pGEX-6P1 plasmids, and purified as described previously^{31,39}. To ensure removal of contaminating nucleic acids from *E. coli*, the purification protocol includes incubation with polyethylenimine (PEI)³⁹. Circular dichroism (CD) spectroscopy was used to control the absence positive $\Delta\epsilon$ values in the spectral region above 240 nm, which is specific of CD spectra measured for RNAs³⁵ (Supplementary data, Fig. S6).

Production of RNA transcripts. DNA template used for *in vitro* transcription of the *P. abyssi* Pab91 guide RNA was obtained by PCR amplification with the forward primer generating the sequence of the T7 promoter³¹. The DNA fragments carrying the T7 RNA polymerase promoter sequence upstream of the sequence encoding the 23S rRNA sequence targeted by Pab91 and corresponding to the 22-U substrate RNA³⁹, or encoding the two 23S rRNA fragments mini-23S-2589 and mini-23S-2607 were obtained by annealing of complementary DNA oligonucleotides. The DNA template encoding the substrate rRNA 23S-143nt was generated by multiple PCR amplifications. In this construct nucleotide at position 2717 in *P. abyssi* 23S rRNA was connected to nucleotide 2143 by the 5'-CTGA-3' sequence. The obtained fragments were cloned in vector pCR2.2 and their sequences verified by sequencing.

Small fragments were transcribed using the MEGAscript T7 transcription kit (Invitrogen). For the other fragments, the transcription reactions were carried out as previously described^{31,65}. The RNA transcripts were purified on 8 M urea-PAGE, and stored in H₂O at -80 °C.

The substrate RNA 22-U used for Pab91 sRNP activity measurements was radiolabeled during transcription in presence of [α -³²P]CTP. The substrates mini-23S-rRNA-2589 and mini-23S-rRNA-2607 used to measure Cbf5 activity were radiolabeled during transcription in presence of [α -³²P]CTP and [α -³²P]GTP, respectively. Radiolabeling of the phosphodiester linkage 3' to the targeted U₂₆₈₅ in the substrate RNA 23S-143nt was obtained by splinted ligation of two *in vitro* transcribed RNA fragments of the 23S rRNA⁸⁴. Briefly, 10,000 cps of the 5' end radiolabeled donor RNA spanning 23S rRNA positions 2686 to 2160 and containing the CUGA sequence connecting position 2717 with position 2143 were mixed with 100 pmol of the acceptor RNA spanning positions 2603 to 2685, in the presence of 50 pmol of a DNA single-stranded splint, which was perfectly complementary to the sequence of the donor and acceptor RNAs, in 50 mM Tris-HCl pH 7.4 containing 10 mM of MgCl₂, 2 nmol of ATP, and 2 nmol of DTT. Hybridization of the two RNAs with the DNA splint was obtained by incubation at 75 °C for 4 min and a slow cooling to room temperature. The ligation reaction was achieved with 75 U of T4 DNA ligase (Fermentas) at 4 °C overnight. The DNA splint was digested by adding 3 U of DNase RQ1 (Promega) at 37 °C for 30 min. The radiolabeled ligation products were purified on a 8 M urea-PAGE (6%).

Assembly of the LCN and LCNG RNP enzymes. LCN Pab91 RNP enzyme assembly was obtained by mixing the sRNA Pab91 and the proteins L7Ae, Cbf5, and Nop10 as previously described⁶⁵. Protein Gar1 was added to this mix to assemble the LCNG RNP enzyme.

Analysis of substrate RNA binding to the Pab91 sRNP. Complex CII' was obtained by association of the sRNA present in the sRNP containing proteins L7Ae, Cbf5, and Nop10 (LCN) with radiolabeled substrate RNA, as described previously^{31,35,39}. The radioactivity present in the CII' complex or in the unbound fraction was revealed using a Phosphorimager (Typhoon 9410, Amersham Biosciences).

RNA:Ψ-synthase activity measurements in single- and multiple-turnover conditions. To test for the RNA:Ψ-synthase activity, we used an approach based on the nearest-neighbor analysis⁶⁵. Conditions for time course analysis of single-turnover reactions by reconstituted Pab91 sRNP particles LCN and LCNG were previously detailed^{31,35,39,48}. Briefly, ~4 pmol of unlabeled Pab91 sRNA and ~150 fmol of labeled 22-U substrate RNA were mixed with protein combination L7Ae-Cbf5-Nop10 or L7Ae-Cbf5-Nop10-Gar1 (200 nM each) at 65 °C. In these conditions, the concentration for the reconstituted sRNP was estimated to be of ~0.5 μM. For multiple turnover reactions, an excess of unlabeled substrate RNA 23S-143nt, ranging from 1 to 10 μM, was added to the reaction mix.

The guide RNA was omitted to perform single turnover reactions for the measurement of the RNA guide-independent activity of Cbf5. The radiolabeled substrates mini-23S-rRNA-2589 and mini-23S-rRNA-2607 (~150 fmol) were incubated with different combination of the recombinant proteins Cbf5, Nop10, Gar1, and L7Ae (200 nM each).

For all these assays, aliquots were collected at several time points and the reaction was stopped by phenol/chloroform extraction followed by ethanol precipitation. The recovered RNAs were digested with ~0.4 U of RNase T2 to generate 3'-monophosphate nucleotides that were fractionated by thin layer chromatography (TLC)⁸⁵. For the reactions performed on the RNA 22-U, one dimension chromatography was sufficient to fractionate the labeled nucleotides whereas two-dimensional TLC is needed for activity measurement with mini-23S-rRNA-2589 or mini-23S-rRNA-2607. The radioactivity of the 3'-monophosphate nucleotides was quantified with the Phosphorimager and quantified using the ImageQuant software Version 5.2 (Molecular Dynamics). The quantities of Ψs formed were determined by taking into account the total number of U nucleotides in the substrate RNA. The obtained values in moles of Ψ per mole (mol.mol⁻¹) of substrate RNA were analyzed with Prism software Version 5.04 (GraphPad).

Data Availability

All data generated or analyzed during this study are included in this published article (and its Supplementary Information files). Raw sequencing data are available in European Nucleotide Archive under accession number PRJEB28421.

References

- Charette, M. & Gray, M. W. Pseudouridine in RNA: what, where, how, and why. *IUBMB Life* **49**, 341–351 (2000).
- Ge, J. & Yu, Y. T. RNA pseudouridylation: new insights into an old modification. *Trends Biochem. Sci.* **38**, 210–218, <https://doi.org/10.1016/j.tibs.2013.01.002> (2013).
- Spenkuch, F., Motorin, Y. & Helm, M. Pseudouridine: still mysterious, but never a fake (uridine)! *RNA Biol.* **11**, 1540–1554, <https://doi.org/10.4161/15476286.2014.992278> (2014).
- Davis, D. R. Stabilization of RNA stacking by pseudouridine. *Nucleic Acids Res.* **23**, 5020–5026 (1995).
- Arnez, J. G. & Steitz, T. A. Crystal structure of unmodified tRNA(Gln) complexed with glutaminyl-tRNA synthetase and ATP suggests a possible role for pseudo-uridines in stabilization of RNA structure. *Biochemistry* **33**, 7560–7567 (1994).
- Newby, M. I. & Greenbaum, N. L. Investigation of Overhauser effects between pseudouridine and water protons in RNA helices. *Proc. Natl. Acad. Sci. USA* **99**, 12697–12702, <https://doi.org/10.1073/pnas.202477199> (2002).
- Davis, D. R. & Poulter, C. D. 1H-15N NMR studies of *Escherichia coli* tRNA(Phe) from *hisT* mutants: a structural role for pseudouridine. *Biochemistry* **30**, 4223–4231 (1991).
- Kierzek, E. *et al.* The contribution of pseudouridine to stabilities and structure of RNAs. *Nucleic Acids Res.* **42**, 3492–3501, <https://doi.org/10.1093/nar/gkt1330> (2014).
- Jiang, J., Aduri, R., Chow, C. S. & SantaLucia, J. Jr. Structure modulation of helix 69 from *Escherichia coli* 23S ribosomal RNA by pseudouridylations. *Nucleic Acids Res.* **42**, 3971–3981, <https://doi.org/10.1093/nar/gkt1329> (2014).

10. Yarian, C. S. *et al.* Structural and functional roles of the N1- and N3-protons of psi at tRNA's position 39. *Nucleic Acids Res.* **27**, 3543–3549 (1999).
11. Durant, P. C. & Davis, D. R. Stabilization of the anticodon stem-loop of tRNA^{Lys,3} by an A+ -C base-pair and by pseudouridine. *J. Mol. Biol.* **285**, 115–131, <https://doi.org/10.1006/jmbi.1998.2297> (1999).
12. Lecointe, F. *et al.* Lack of pseudouridine 38/39 in the anticodon arm of yeast cytoplasmic tRNA decreases *in vivo* recoding efficiency. *J. Biol. Chem.* **277**, 30445–30453, <https://doi.org/10.1074/jbc.M203456200> (2002).
13. Baudin-Baillieu, A. *et al.* Nucleotide modifications in three functionally important regions of the *Saccharomyces cerevisiae* ribosome affect translation accuracy. *Nucleic Acids Res.* **37**, 7665–7677, <https://doi.org/10.1093/nar/gkp816> (2009).
14. Ejby, M., Sorensen, M. A. & Pedersen, S. Pseudouridylation of helix 69 of 23S rRNA is necessary for an effective translation termination. *Proc. Natl. Acad. Sci. USA* **104**, 19410–19415 (2007).
15. King, T. H., Liu, B., McCully, R. R. & Fournier, M. J. Ribosome structure and activity are altered in cells lacking snoRNPs that form pseudouridines in the peptidyl transferase center. *Mol. Cell* **11**, 425–435 (2003).
16. Piekna-Przybylska, D., Przybylski, P., Baudin-Baillieu, A., Rousset, J. P. & Fournier, M. J. Ribosome performance is enhanced by a rich cluster of pseudouridines in the A-site finger region of the large subunit. *J. Biol. Chem.* **283**, 26026–26036 (2008).
17. Liang, X. H., Liu, Q. & Fournier, M. J. rRNA modifications in an intersubunit bridge of the ribosome strongly affect both ribosome biogenesis and activity. *Mol. Cell* **28**, 965–977, <https://doi.org/10.1016/j.molcel.2007.10.012> (2007).
18. Liang, X. H., Liu, Q. & Fournier, M. J. Loss of rRNA modifications in the decoding center of the ribosome impairs translation and strongly delays pre-rRNA processing. *RNA* **15**, 1716–1728, <https://doi.org/10.1261/rna.1724409> (2009).
19. Badis, G., Fromont-Racine, M. & Jacquier, A. A snoRNA that guides the two most conserved pseudouridine modifications within rRNA confers a growth advantage in yeast. *RNA* **9**, 771–779 (2003).
20. Jack, K. *et al.* rRNA pseudouridylation defects affect ribosomal ligand binding and translational fidelity from yeast to human cells. *Mol. Cell* **44**, 660–666, <https://doi.org/10.1016/j.molcel.2011.09.017> (2011).
21. Wu, G., Yu, A. T., Kantartzis, A. & Yu, Y. T. Functions and mechanisms of spliceosomal small nuclear RNA pseudouridylation. *Wiley Interdiscip. Rev. RNA* **2**, 571–581, <https://doi.org/10.1002/wrna.77> (2011).
22. Wu, G. *et al.* Pseudouridines in U2 snRNA stimulate the ATPase activity of Prp5 during spliceosome assembly. *EMBO J.* **35**, 654–667, <https://doi.org/10.15252/embj.201593113> (2016).
23. Carlile, T. M. *et al.* Pseudouridine profiling reveals regulated mRNA pseudouridylation in yeast and human cells. *Nature* **515**, 143–146, <https://doi.org/10.1038/nature13802> (2014).
24. Schwartz, S. *et al.* Transcriptome-wide mapping reveals widespread dynamic-regulated pseudouridylation of ncRNA and mRNA. *Cell* **159**, 148–162, <https://doi.org/10.1016/j.cell.2014.08.028> (2014).
25. Hamma, T. & Ferre-D'Amare, A. R. Pseudouridine synthases. *Chem. Biol.* **13**, 1125–1135 (2006).
26. Roovers, M. *et al.* Formation of the conserved pseudouridine at position 55 in archaeal tRNA. *Nucleic Acids Res.* **34**, 4293–4301 (2006).
27. McCleverty, C. J., Hornsby, M., Spraggon, G. & Kreusch, A. Crystal structure of human Pus10, a novel pseudouridine synthase. *J. Mol. Biol.* **373**, 1243–1254, <https://doi.org/10.1016/j.jmb.2007.08.053> (2007).
28. Reichow, S. L., Hamma, T., Ferre-D'Amare, A. R. & Varani, G. The structure and function of small nucleolar ribonucleoproteins. *Nucleic Acids Res.* **35**, 1452–1464, <https://doi.org/10.1093/nar/gkl1172> (2007).
29. Watkins, N. J. & Bohnsack, M. T. The box C/D and H/ACA snoRNPs: key players in the modification, processing and the dynamic folding of ribosomal RNA. *Wiley Interdiscip. Rev. RNA* **3**, 397–414, <https://doi.org/10.1002/wrna.117> (2012).
30. Baker, D. L. *et al.* RNA-guided RNA modification: functional organization of the archaeal H/ACA RNP. *Genes Dev.* **19**, 1238–1248, <https://doi.org/10.1101/gad.1309605> (2005).
31. Charpentier, B., Muller, S. & Branlant, C. Reconstitution of archaeal H/ACA small ribonucleoprotein complexes active in pseudouridylation. *Nucleic Acids Res.* **33**, 3133–3144, <https://doi.org/10.1093/nar/gki630> (2005).
32. Hamma, T., Reichow, S. L., Varani, G. & Ferre-D'Amare, A. R. The Cbf5-Nop10 complex is a molecular bracket that organizes box H/ACA RNPs. *Nat. Struct. Mol. Biol.* **12**, 1101–1107 (2005).
33. Manival, X. *et al.* Crystal structure determination and site-directed mutagenesis of the *Pyrococcus abyssi* aCBF5-aNOP10 complex reveal crucial roles of the C-terminal domains of both proteins in H/ACA sRNP activity. *Nucleic Acids Res.* **34**, 826–839, <https://doi.org/10.1093/nar/gkj482> (2006).
34. Rashid, R. *et al.* Crystal structure of a Cbf5-Nop10-Gar1 complex and implications in RNA-guided pseudouridylation and dyskeratosis congenita. *Mol. Cell* **21**, 249–260, <https://doi.org/10.1016/j.molcel.2005.11.017> (2006).
35. Fourmann, J. B. *et al.* Comparative study of two box H/ACA ribonucleoprotein pseudouridine-synthases: relation between conformational dynamics of the guide RNA, enzyme assembly and activity. *PLoS One* **8**, e70313, <https://doi.org/10.1371/journal.pone.0070313> (2013).
36. Li, L. & Ye, K. Crystal structure of an H/ACA box ribonucleoprotein particle. *Nature* **443**, 302–307, <https://doi.org/10.1038/nature05151> (2006).
37. Liang, B. *et al.* Structure of a functional ribonucleoprotein pseudouridine synthase bound to a substrate RNA. *Nat. Struct. Mol. Biol.* **16**, 740–746, <https://doi.org/10.1038/nsmb.1624> (2009).
38. Duan, J., Li, L., Lu, J., Wang, W. & Ye, K. Structural mechanism of substrate RNA recruitment in H/ACA RNA-guided pseudouridine synthase. *Mol. Cell* **34**, 427–439, <https://doi.org/10.1016/j.molcel.2009.05.005> (2009).
39. Tillault, A. S. *et al.* RNA size is a critical factor for U-containing substrate selectivity and permanent pseudouridylated product release during the RNA:Psi-synthase reaction catalyzed by box H/ACA sRNP enzyme at high temperature. *Biochimie* **113**, 134–142, <https://doi.org/10.1016/j.biochi.2015.04.009> (2015).
40. Liang, B. *et al.* Long-distance placement of substrate RNA by H/ACA proteins. *RNA* **14**, 2086–2094, <https://doi.org/10.1261/rna.1109808> (2008).
41. Wang, P., Yang, L., Gao, Y. Q. & Zhao, X. S. Accurate placement of substrate RNA by Gar1 in H/ACA RNA-guided pseudouridylation. *Nucleic Acids Res.* **43**, 7207–7216, <https://doi.org/10.1093/nar/gkv757> (2015).
42. Yang, X. *et al.* Kinetic and thermodynamic characterization of the reaction pathway of box H/ACA RNA-guided pseudouridine formation. *Nucleic Acids Res.* **40**, 10925–10936, <https://doi.org/10.1093/nar/gks882> (2012).
43. Muller, S. *et al.* Deficiency of the tRNA^{Tyr}:Psi 35-synthase aPus7 in Archaea of the Sulfolobales order might be rescued by the H/ACA sRNA-guided machinery. *Nucleic Acids Res.* **37**, 1308–1322 (2009).
44. Gurha, P., Joardar, A., Chaurasia, P. & Gupta, R. Differential roles of archaeal box H/ACA proteins in guide RNA-dependent and independent pseudouridine formation. *RNA Biol.* **4**, 101–109 (2007).
45. Kamalampeta, R. & Kothe, U. Archaeal proteins Nop10 and Gar1 increase the catalytic activity of Cbf5 in pseudouridylating tRNA. *Sci. Rep.* **2**, 663, <https://doi.org/10.1038/srep00663> (2012).
46. Majumder, M., Bosmeny, M. S. & Gupta, R. Structure-function relationships of archaeal Cbf5 during *in vivo* RNA-guided pseudouridylation. *RNA* **22**, 1604–1619, <https://doi.org/10.1261/rna.057547.116> (2016).
47. Muller, S., Fourmann, J. B., Loegler, C., Charpentier, B. & Branlant, C. Identification of determinants in the protein partners aCBF5 and aNOP10 necessary for the tRNA:Psi55-synthase and RNA-guided RNA:Psi-synthase activities. *Nucleic Acids Res.* **35**, 5610–5624, <https://doi.org/10.1093/nar/gkm606> (2007).
48. Tillault, A. S. *et al.* Contribution of two conserved histidines to the dual activity of archaeal RNA guide-dependent and -independent pseudouridine synthase Cbf5. *RNA* **21**, 1233–1239, <https://doi.org/10.1261/rna.051425.115> (2015).

49. Zhou, J., Liang, B. & Li, H. Structural and functional evidence of high specificity of Cbf5 for ACA trinucleotide. *RNA* **17**, 244–250, <https://doi.org/10.1261/rna.2415811> (2011).
50. Blaby, I. K. *et al.* Pseudouridine formation in archaeal RNAs: The case of *Haloferax volcanii*. *RNA* **17**, 1367–1380, <https://doi.org/10.1261/rna.2712811> (2011).
51. Chatterjee, K. *et al.* The archaeal COG1901/DUF358 SPOUT-methyltransferase members, together with pseudouridine synthase Pus10, catalyze the formation of 1-methylpseudouridine at position 54 of tRNA. *RNA* **18**, 421–433, <https://doi.org/10.1261/rna.030841.111> (2012).
52. Gurha, P. & Gupta, R. Archaeal Pus10 proteins can produce both pseudouridine 54 and 55 in tRNA. *RNA* **14**, 2521–2527 (2008).
53. Joardar, A. *et al.* Role of forefinger and thumb loops in production of Psi54 and Psi55 in tRNAs by archaeal Pus10. *RNA* **19**, 1279–1294, <https://doi.org/10.1261/rna.039230.113> (2013).
54. Kamalampeta, R., Keffer-Wilkes, L. C. & Kothe, U. tRNA binding, positioning, and modification by the pseudouridine synthase Pus10. *J. Mol. Biol.* **425**, 3863–3874, <https://doi.org/10.1016/j.jmb.2013.05.022> (2013).
55. Muller, S. *et al.* Combined *in silico* and experimental identification of the *Pyrococcus abyssi* H/ACA sRNAs and their target sites in ribosomal RNAs. *Nucleic Acids Res.* **36**, 2459–2475 (2008).
56. Fukui, T. *et al.* Complete genome sequence of the hyperthermophilic archaeon *Thermococcus kodakaraensis* KOD1 and comparison with *Pyrococcus* genomes. *Genome Res.* **15**, 352–363 (2005).
57. Fujikane, R., Ishino, S., Ishino, Y. & Forterre, P. Genetic analysis of DNA repair in the hyperthermophilic archaeon. *Thermococcus kodakaraensis*. *Genes Genet. Syst.* **85**, 243–257 (2010).
58. Sato, T., Fukui, T., Atomi, H. & Imanaka, T. Improved and versatile transformation system allowing multiple genetic manipulations of the hyperthermophilic archaeon *Thermococcus kodakaraensis*. *Appl. Environ. Microbiol.* **71**, 3889–3899 (2005).
59. Sato, T., Fukui, T., Atomi, H. & Imanaka, T. Targeted gene disruption by homologous recombination in the hyperthermophilic archaeon *Thermococcus kodakaraensis* KOD1. *J. Bacteriol.* **185**, 210–220 (2003).
60. Glandsdorff, N. On the origin of operons and their possible role in evolution toward thermophily. *J. Mol. Evol.* **49**, 432–438 (1999).
61. Atomi, H., Fukui, T., Kanai, T., Morikawa, M. & Imanaka, T. Description of *Thermococcus kodakaraensis* sp. nov., a well studied hyperthermophilic archaeon previously reported as *Pyrococcus* sp. KOD1. *Archaea* **1**, 263–267 (2004).
62. Frols, S. *et al.* UV-inducible cellular aggregation of the hyperthermophilic archaeon *Sulfolobus solfataricus* is mediated by pili formation. *Mol. Microbiol.* **70**, 938–952 (2008).
63. Frols, S. *et al.* Response of the hyperthermophilic archaeon *Sulfolobus solfataricus* to UV damage. *J. Bacteriol.* **189**, 8708–8718 (2007).
64. Kawakami, Y., Ito, T., Kamekura, M. & Nakayama, M. Ca(2+)-dependent cell aggregation of halophilic archaeon. *Halobacterium salinarum*. *J. Biosci. Bioeng.* **100**, 681–684 (2005).
65. Charpentier, B., Fourmann, J. B. & Branlant, C. Reconstitution of archaeal H/ACA sRNPs and test of their activity. *Methods Enzymol.* **425**, 389–405, [https://doi.org/10.1016/S0076-6879\(07\)25016-5](https://doi.org/10.1016/S0076-6879(07)25016-5) (2007).
66. Huang, L. *et al.* Identification of two *Escherichia coli* pseudouridine synthases that show multisite specificity for 23S RNA. *Biochemistry* **37**, 15951–15957, <https://doi.org/10.1021/bi981002n> (1998).
67. Kinghorn, S. M., O'Byrne, C. P., Booth, I. R. & Stansfield, I. Physiological analysis of the role of *truB* in *Escherichia coli*: a role for tRNA modification in extreme temperature resistance. *Microbiology* **148**, 3511–3520, <https://doi.org/10.1099/0021287-148-11-3511> (2002).
68. Allers, T. & Mevarech, M. Archaeal genetics - the third way. *Nat. Rev. Genet.* **6**, 58–73, <https://doi.org/10.1038/nrg1504> (2005).
69. Farkas, J. A., Picking, J. W. & Santangelo, T. J. Genetic techniques for the archaea. *Annu. Rev. Genet.* **47**, 539–561, <https://doi.org/10.1146/annurev-genet-111212-133225> (2013).
70. Leigh, J. A., Albers, S. V., Atomi, H. & Allers, T. Model organisms for genetics in the domain Archaea: methanogens, halophiles, Thermococcales and Sulfolobales. *FEMS Microbiol. Rev.* **35**, 577–608, <https://doi.org/10.1111/j.1574-6976.2011.00265.x> (2011).
71. Wagner, M. *et al.* Expanding and understanding the genetic toolbox of the hyperthermophilic genus *Sulfolobus*. *Biochem. Soc. Trans.* **37**, 97–101, <https://doi.org/10.1042/BST0370097> (2009).
72. Blaby, I. K. *et al.* Towards a systems approach in the genetic analysis of archaea: Accelerating mutant construction and phenotypic analysis in *Haloferax volcanii*. *Archaea* **2010**, 426239, <https://doi.org/10.1155/2010/426239> (2010).
73. El Yacoubi, B. *et al.* A Gateway platform for functional genomics in *Haloferax volcanii*: deletion of three tRNA modification genes. *Archaea* **2**, 211–219 (2009).
74. Joardar, A., Gurha, P., Skariah, G. & Gupta, R. Box C/D RNA-guided 2'-O methylations and the intron of tRNA^{Trp} are not essential for the viability of *Haloferax volcanii*. *J. Bacteriol.* **190**, 7308–7313, <https://doi.org/10.1128/JB.00820-08> (2008).
75. Atomi, H., Imanaka, T. & Fukui, T. Overview of the genetic tools in the Archaea. *Front. Microbiol.* **3**, 337, <https://doi.org/10.3389/fmicb.2012.00337> (2012).
76. Zebarjadian, Y., King, T., Fournier, M. J. & Clarke, L. & Carbon, J. Point mutations in yeast CBF5 can abolish *in vivo* pseudouridylation of rRNA. *Mol. Cell. Biol.* **19**, 7461–7472 (1999).
77. Gutsell, N. *et al.* Deletion of the *Escherichia coli* pseudouridine synthase gene *truB* blocks formation of pseudouridine 55 in tRNA *in vivo*, does not affect exponential growth, but confers a strong selective disadvantage in competition with wild-type cells. *RNA* **6**, 1870–1881 (2000).
78. Ishida, K. *et al.* Pseudouridine at position 55 in tRNA controls the contents of other modified nucleotides for low-temperature adaptation in the extreme-thermophilic eubacterium *Thermus thermophilus*. *Nucleic Acids Res.* **39**, 2304–2318, <https://doi.org/10.1093/nar/gkq1180> (2011).
79. Nobles, K. N., Yarian, C. S., Liu, G., Guenther, R. H. & Agris, P. F. Highly conserved modified nucleosides influence Mg²⁺-dependent tRNA folding. *Nucleic Acids Res.* **30**, 4751–4760 (2002).
80. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410, [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2) (1990).
81. Lepage, E. *et al.* Molecular diversity of new Thermococcales isolates from a single area of hydrothermal deep-sea vents as revealed by randomly amplified polymorphic DNA fingerprinting and 16S rRNA gene sequence analysis. *Appl. Environ. Microbiol.* **70**, 1277–1286 (2004).
82. Motorin, Y., Muller, S., Behm-Ansmant, I. & Branlant, C. In *Methods Enzymol.* Vol. 425 (ed Gott, J.) 21–53 (2007).
83. Bakin, A. & Ofengand, J. Four newly located pseudouridylation residues in *Escherichia coli* 23S ribosomal RNA are all at the pseudotransferase center: analysis by the application of a new sequencing technique. *Biochemistry* **32**, 9754–9762 (1993).
84. Kirsch, W. C., Muller, J., Wombacher, R. & Helm, M. Optimizing splinted ligation of highly structured small RNAs. *RNA* **11**, 1909–1914, <https://doi.org/10.1261/rna.2170705> (2005).
85. Keith, G. Mobilities of modified ribonucleotides on two-dimensional cellulose thin-layer chromatography. *Biochimie* **77**, 142–144 (1995).

Acknowledgements

We thank Drs Haruyuki Atomi and Rie Matsumi at Kyoto University for their kind gift of KU216 strain and vector, and for technical advices on the gene-targeting method. Dr. Virginie Marchand (UMS2008 IBSLor, NGS core facility) for supervising library preparation and performing Illumina sequencing. This work was supported by the Agence Nationale de la Recherche (ANR-08-BLAN-0036-01), the French Ministère de l'Enseignement

Supérieur et de la Recherche (doctoral fellowship to A.-S.T.), the Centre National de la Recherche Scientifique, and the Université de Lorraine (UL).

Author Contributions

R.F., I.B.-A., C.B., Y.M., and B.C. conceived and designed the experiments; R.F., I.B.-A., A.-S.T., C.L., V.I.-B., and E.M. performed the experiments; R.F., I.B.-A., C.B., Y.M., and B.C. analyzed the data; P.F. contributed reagents/materials/analysis tools; B.C. wrote the manuscript.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-018-32164-0>.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018