An evolutionally conserved archaeal KEOPS complex functions in DNA repair

Pengju Wu, Qi Gan, Xuemei Zhang, Yunfeng Yang, Yuanxi Xiao, Qunxin She, Jinfeng Ni, Qihong Huang[#], Yulong Shen[#]
CRISPR and Archaea Biology Research Center, State Key Laboratory of Microbial Technology, Microbial Technology Institute, Shandong University, 266237, Qingdao, China
[#]Correspondence: yulgshen@sdu.edu.cn, huangqihong@sdu.edu.cn

Abstract

Double-stranded DNA break (DSB) repair is a fundamental process for all cellular life. Recently, KEOPS, a multiple-subunit protein complex that is conserved in eukaryotes and archaea and primarily involved in N^6-threonylcarbamoyl adenosine (t^6A) modification of tRNAs is reported to participate in homologous recombination in yeast. To functionally characterize archaeal KEOPS (aKEOPS), we conducted genetic and biochemical analyses of its encoding genes in the hyperthermophilic archaeon *Saccharolobus islandicus*. We show that aKEOPS possesses five subunits, Pcc1, Pcc1-like (or Gon7-like), Kae1, Bud32 and Cgi121, just as eukaryotic KEOPS. Pcc1-like has physical interactions with Kae1 and Pcc1 and can mediate the monomerization of the dimeric subcomplex (Kae1-Pcc1-Pcc1-Kae1), suggesting that Pcc1-like is a functional homolog of the eukaryotic Gon7 subunit. Strikingly, none of the genes encoding aKEOPS subunits, including Pcc1 and Pcc1-like, can be deleted in the wild type and in a t^6A modification complementary strain constructed, indicating that aKEOPS complex is essential in multiple cellular processes in this archaeon. Moreover, knock-down of the subunit genes leads to increase or decrease in the sensitivity of the cells to hydroxyurea and ultraviolet radiation both of which are DNA damage agents. These results indicated that aKEOPS plays an important role in DNA repair. *In vitro*, archaeal Cgi121 possesses dsDNA-binding activity which relies on its tRNA 3' CCA tail binding module. Our study indicates that DNA repair is an original intrinsic function of the evolutionarily conserved complex and reveals a possible link between two functions of the complex, t^6A modification and DSB repair.

Introduction

DNA double-strand break (DSB) is the most severe type of DNA damage which can result in chromosomal rearrangements and eventually cell death if not appropriately repaired (1). In eukaryotes, DSBs can be repaired by two major pathways, non-homologous end joining (NHEJ) and homologous recombination (HR) (2). DNA end processing is a key step for the two repair pathways. In NHEJ, Ku70/80 heterodimer binds to the breaks, which prevents DNA end resection by nucleases, whereas in HR, the end resection is required (3). In eukaryotic HR, DNA end processing is initiated by Mre11-Rad50-Xrs2/Mre11-Rad50-NBS1(MRX/N) in conjunction with Sae2/CtIP, yielding a short 3'-overhang ssDNA (4). This intermediate is further processed in two distinct pathways, mediated by Exo1/EXO1 or Dna2/DNA2-Sgs1/BLM, generating a long 3'-tail of single-stranded DNA (ssDNA) (5-8). Then, RPA binds to the ssDNA and recruits recombinase Rad51 for strand invasion subsequent processes of HR (9). Protein phosphorylation is critical for the regulation of DNA end resection. ATM kinases regulate the initiation of end resection via phosphorylation of the histone variant H2AX which recruits additional ATM, MDC1, and many other proteins. ATM also phosphorylates

Sae2/CtIP and MRN-interaction proteins stimulate the end resection (10-13). In addition, several kinases (ATM, ATR, CDK, and DNA-PK) phosphorylate RPA, attenuating the end resection primarily via inhibition of BLM (14). Although the end processing steps in eukaryotic HR pathway have been extensively investigated, novel proteins have been continuingly identified (15,16). Interestingly, recently KEOPS (Kinase putative Endopeptidase and Other Proteins of Small size protein complex) (17,18), a conserved multiple-functional complex, is reported to be involved in DNA end resection promoting HR repair in yeast (19). The DSB end resection ability and HR efficiency were reduced in the KEOPS deletion mutants (17,19) suggesting that KEOPS is a novel important player in DSB end resection or its regulation.

KEOPS is comprised of five subunits in eukaryotes, Gon7, Pcc1, Kae1, Bud32, and Cgi121 (17,18). Importantly, among these subunits, Kae1 is one of the sixty or so most conserved proteins across all domains of life (20) and Bud32 is a kinase/ATPase and shares a structural similarity to Rio kinases (21). The primary function of KEOPS is the N^6-threonylcarbamoyl adenosine (t^6A) modification at A37 of all tRNAs carrying an NNU anti-codon. This tRNA modification is critical for the stabilization of the anti-codon loop structure and the base pairing of the anticodon with the corresponding codon and for the enhancement of translation fidelity (22,23). Lack of the t^6A modification can cause global translation impact, including ribosome assembly defects and translational ambiguity in baker's yeast and a serious autosomal-recessive disease Galloway-Mowat syndrome in humans (24,25). Additionally, in budding yeast, the complex is also involved in telomere maintenance (17). Although the mechanism of KEOPS complex in t^6A modification is known relatively clear, the precise role of KEOPS complex in DSB repair and telomere maintenance is largely unexplored. Interestingly, the functions of KEOPS in HR and telomere maintenance are independent of its t^6A modification function (19,26).

Like in other information processing processes, archaea harbour eukaryotic-like proteins of the DNA recombination pathway, such as Mre11, Rad50, RadA (Rad51 homolog), and RPA (SSB) (27). In fact, many of the archaeal proteins, especially Mre11 and Rad50 and their complexes have served as excellent models for elucidating the structural and functional mechanisms for the eukaryotic counterparts (28-30). In thermophilic archaea, all key proteins involved in HR are essential (31,32), indicating that HR plays fundamental role in DNA repair for the survival of these archaea. Intriguingly, all the subunits of KEOPS complex except Gon7 are also found and highly conserved in archaea and eukaryotes (21,33). Similarly, archaeal KEOPS (aKEOPS) has been used as a model for the structural and biochemical studies of eukaryotic KEOPS (34-37). Nevertheless, so far genetic analysis has been studied only in a halophilic euryarchaeon *Haloferax volcanii* for some of the KEOPS subunits (38). Interestingly, in this euryarchaeon, the fused *kae1-bud32* gene was shown to be essential as was *cgi121*, which is dispensable in yeast

(38). In contrast, *pcc1* (encoding the putative dimerizing unit of KEOPS) was not essential. Deletion of the *pcc1* led to pleiotropic phenotypes, including decreased growth rate and increased cellular DNA content (38), implying a failure in chromosome segregation and/or DNA repair. Whether the DSB repair function is conserved or not in archaea and eukaryotes and its relationship with other functions of KEOPS, tRNA modification and telomere maintenance, are unclear up to now.

In this work, we carried out *in vivo* and biochemical analysis of KEOPS subunit in *Saccharolobus islandicus* (formally *Sulfolobus islandicus*) REY15A, a thermophilic model crenarchaeon close to the ancestral eukaryotes. We show that one Pcc1 homolog, Pcc1-like, is the fifth subunit and functional Gon7 homolog of aKEOPS. We further show that each of the five genes encoding aKEOPS subunits (including the gene annotated as *pcc1-like*) is essential in the wild-type and a strain complemented with a bacterial t^6A modification system, suggesting that KEOPS has additional essential functions in *S. islandicus*. Genotoxic sensitivity analysis of the *cgi121* knock-down strains reveals that aKEOPS is also involved in DNA damage repair, indicating that another basic function of eukaryotic KEOPS is also conserved in archaea. Further biochemical analysis suggests that the Pcc1-Kae1 subcomplex and Cgi121 are the main dsDNA binding proteins, and one of the key conserved residues of Cgi121 involved in tRNA 3' CCA tail binding also contributes to dsDNA binding, revealing a possible link between two functions of the complex, t^6A modification and DSB repair. Finally, a model for the evolutional scenario of t^6A modification systems is proposed.

Materials and methods

Strains and growth conditions

S. islandicus strain REY15A (E233S) (Δ*pyrEF*Δ*lacS*) (hereafter E233S) and its transformants were cultured as described previously (39). D-arabinose (0.2% w/v) was used for cell growth to induce protein overexpression controlled by different araS promoters.

Construction of plasmids and strains

The strains and vectors used and constructed in this study are listed in Supplementary Table S1 and S2. Oligonucleotides for spacers, P_{araS} mutant promoters, fluorescence-labelled ssDNA, and primers for PCR and RT-qPCR are listed in Supplementary Table S3. For the construction of *Thermotoga maritima* t^6A biosynthesis system knock-in strain TsaKI, constitutive expression vector pSeGP was constructed by replacing the *araS*-SD sequence with the promoter of *S. islandicus* glutamate dehydrogenase (P_{gdhA}) at the site between SphI and NdeI of pSeSD (40). The gene for the TsaB, TsaC, TsaD, or TsaE was inserted into the NdeI and SalI site of pSeGP individually. Then, the cassettes

for the TsaB, TsaC, TsaD, and TsaE were amplified and ligated into pRSF-Duet1 sequentially in the order of TsaB-TsaD-TsaE-TsaC. The homologous arms (L-arm and R-arm) of *amyα* (SiRe_1098) were ligated via gene splicing by overlap extension (SOE) PCR with KpnI/XhoI sites in between and inserted into the SphI/BlnI site of pUC19. Next, the cassette of *Tm*TsaB-TsaD-TsaE-TsaC was inserted into the KpnI/XhoI site between the two homologous arms. Finally, the fragment containing the homologous arms and the *Tm*TsaB-TsaD-TsaE-TsaC was amplified and inserted into the SphI/BlnI site of the pGE vector (41), yielding a vector pGE-*Tm*TsaBDEC-KI. To construct pGE-*Tm*TsaD-KI, the spacer was designed to target the endo-β-mannanase (SiRe_2238) and the fragment containing the homologous arms and *Tm*TsaD was inserted into the SphI/BlnI site of pGE. For construction of the vectors for the gene knock-down, the corresponding spacers were inserted into the mini-CRISPR locus in pGE (41). For the construction of the knock-out strain, the corresponding spacers were inserted into the mini-CRISPR locus and the fragment containing the homologous arms was inserted into the SphI/XhoI site of pGE. To construct vectors for protein overexpression in *S. islandicus*, the gene fragments were amplified by PCR, digested with the restriction enzymes, and inserted into the NdeI and SalI site of pSeSD (42). To construct vectors for protein co-overexpression in *S. islandicus*, another multiple clone site (XhoI/NheI) with an *araS* promoter was introduced into pSeSD yielding pSeSD-2ParaS for plasmid construction.

To construct the knock-in strain TsaKI harbouring *T. maritima* t^6A biosynthesis system, the vector pGE-*Tm*TsaBDEC-KI was transformed into E233S. The transformants were screened on uracil free media and verified by PCR using flanking primers (Table S3). The vector pGE-*Tm*TsaD-KI was transformed into *amyα::Tm*TsaB/TsaC/TsaE yielding the strain TsaKI containing the whole *T. maritima* t^6A biosynthesis system. The vectors for knock-down, knock-out, and overexpression of KEOPS genes were transformed into E233S to obtain the corresponding strains.

Growth curve and dot assay

The growth curves of the *S. islandicus* strains were measured as follows. Each strain was pre-cultured in a flask containing 25 ml medium at 75℃ for 24 h until the culture reached an OD_{600} of 0.4–0.6. Cells were collected and resuspended in 25 ml ddH_2O and then inoculated into 30 ml medium to an estimated OD_{600} of 0.04 and cultured at 75℃ in a 100 ml flask with shaking. The cell densities were measured at the indicated time intervals. Each strain was analysed in duplicate, and the measurements were performed three times independently. For the growth observation on plates, an overnight-grown culture was diluted to an estimated OD_{600} of 0.2. Cells were diluted in ten-fold gradience and spotted on plates with or without DNA damage agents. Plates were incubated at 75℃ for 4–12 days before photos were taken. For the ultraviolet (UV)

treatment, the plates were treated with 254 nm UV light using a CL-1000 Ultraviolet crosslinker (UVP, USA), covered with black plastic bags during the first two days, and cultivated for further 4–7 days.

Analysis of t^6A modification and transcription by RT-qPCR

Bulk tRNA were isolated according to the reported method (43) using hot acid-saturated phenol with slight modification. In particular, the cells were grown in 25 ml medium at 75℃ to OD_{600} of 0.8 before collection. The isolated bulk tRNA were analyzed according to the reported protocol (44). Briefly, tRNA (5 μg) was boiled for 2 min and cooled down on ice immediately. Then the tRNA were digested with 10 U Nuclease P1 (New England Biolabs, MA, USA) at 37℃ overnight followed by dephosphorylation using 10 U bacteria alkaline phosphatase (Thermofisher Scientific, USA) at 65℃ for 2 h. The ribonucleotides were analyzed by Ultimate 3000 HPLC (Thermofisher Scientific) using Zorbax SB-C18 (250 x 4.6 mm, 5 μM) reverse-phase column (Agilent). Ribonucleotides were separated by linear gradient elution from 2 % to 12.5 % buffer B at a flow rate of 1.5 ml/min. The compositions of the HPLC buffers were as follows: (A) 250 mM NH_4Ac, pH 6.5; (B) 40 % acetonitrile.

For RT-qPCR analysis, when the cultures were grown to an OD_{600} of 0.8, cells of 2 ml cultures were collected by centrifugation. Total RNA was isolated with 1 ml Trizol (TransGene Biotech, Beijing, China) followed by chloroform and isopropanol extraction. The RNA was precipitated by 75 % ethanol and dissolved in 50 μl DEPC water. RNA (500 ng) was used for cDNA synthesis using an Evo M-ML kit (Accurate Biology, Hunan, China) according to the manufacturer's instructions. RT-qPCR was conducted by CFX96 Real Time PCR detection system (Bio-Rad) using a SYBR Green Pro Taq HS kit (Accurate Biology) according to the manufacturer's instructions.

Electrophoretic mobility shift assay (EMSA)

The EMSA assay was performed in a 20 μl reaction mixture containing 10 nM fluorescence-labelled probes (Supplementary Table S3) and the indicated proteins (0–20 μM) in a binding buffer (50 mM Tris-HCl pH 8.0, 100 μM ATP). The reactions were incubated at 37℃ for 30 min. The products were separated by 8 % PAGE in 0.5 × TBE at 4℃. The gels were imaged with Amersham ImageQuant 800 (Cytiva, USA).

Protein expression and purification

For protein expression, the vectors were transformed into *E. coli* BL21(DE3)-Codon plus-RIL strain. No-tagged or hexa-histidine (6×His)-tagged *Sis*Kae1 (Kae1 from *S. islandicus*) and *Sis*Bud32 were expressed using pET15b. His-tagged *Sis*Pcc1-like, *Tko*Cgi121, and *Tko*Cgi121I75E were expressed using pRSF-Duet1. MBP- and His-tagged *Sis*Cgi121 and *Sis*Pcc1 were expressed

using pMALc2X. No-tagged *Sis*Kae1 were co-expressed with His-tagged *Sis*Pcc1 using pRSF-Duet1. His-tagged *Sis*Pcc1-like was co-expressed with Flag-tagged *Sis*Kae1 and *Sis*Pcc1 using pRSF-Duet1, respectively. The expression was induced by the addition of 0.1 mM IPTG to 1.0 L of LB media at OD_{600} 0.4-0.6 and the cells were cultivated at 37℃ for 4 h. The proteins were purified using nickel affinity chromatography (Ni-NTA) coupled to anion exchange (HiTrap Q FF) and size exclusion chromatography (Superdex 200 Increase 10/300) in a buffer containing 50 mM Tris-HCl pH 8.0, 5 % glycerol, and 50–500 mM NaCl depending on different proteins.

For protein expression and purification in *S. islandicus*, strains carrying the constructed pSeSD-based vectors were cultivated in sucrose medium. Expression was induced by the addition of 0.2 % D-arabinose to 1.0 L cultures at an OD_{600} of 0.2 and the cells were cultured at 75℃ for 24 h. His-tagged *Sis*Pcc1 and *Sis*Kae1 were expressed separately. His-tagged *Sis*Bud32 and no tagged *Sis*Kae1 were co-expressed by cloning *kae1-bud32* operon into pSeSD. The subsequent protein purification was the same as described above.

Pull down assay

His-tag pull down was performed using 10 μM His-*Sis*Pcc1-like protein mixed with 12 μM no-tagged *Sis*Kae1, or 10 μM His-*Sis*Kae1 mixed with 12 μM no-tagged *Sis*Bud32, in 200 μl binding buffer (50 mM NaCl, 50 mM Tris-HCl pH 8.0). The mixture was incubated with 50 μl pre-equilibrated Ni-NTA nickel-charged affinity resin. Binding was allowed to proceed for 10 min at 4℃ with mild shaking, followed by three times washes with the wash buffer (binding buffer containing 10 mM imidazole). The resin was resuspended in the elute buffer (binding buffer containing 400 mM imidazole) and the supernatant was analyzed by 15 % SDS-PAGE and Coomassie blue staining.

MBP-tag pull down was conducted using 10 μM MBP-*Sis*Cgi121-His mixed with 12 μM His-*Sis*Bud32 in 200 μl binding buffer. The mixture was incubated with 50 μl pre-equilibrated amylose resin. Binding was allowed to proceed for 10 min at 4℃ with mild shaking, followed by wash in binding buffer for three times. The resin was resuspended in elute buffer (binding buffer containing 10 mM maltose) and the supernatant was analyzed by 15 % SDS-PAGE and Western blotting.

Size exclusion chromatography (SEC) analysis

SEC analysis of single proteins or combinations of the subunits was conducted using AKTA basic 10 FPLC system (GE healthcare) and a Superdex 200 Increase 10/300 SEC column. Chromatography was performed in a running buffer (200 mM NaCl, 50 mM Tris-HCl pH 8.0, and 5 % glycerol). For complex reconstitution, 30 μM *Sis*Pcc1-like was mixed with 20 μM *Sis*Kae1/*Sis*Pcc1-like, or 6 μM *Sis*Pcc1-like mixed with 30 μM *Sis*Kae1, in 600 μl running buffer at 60℃

for 30 min before loading onto the column.

Bioinformatic analysis

BLAST and genome context analysis was harnessed to identify Sua5 and KEOPS homolog sequences in archaea and eukaryotes (45). The evolutionary distances were calculated using the *p*-distance method (46). For phylogenetic analysis of Pcc1 and Gon7, 126 homolog sequences from 82 representative species of each phylum of archaea and eukaryotes were aligned by MAFFT v7 (47,48). The evolutionary history was inferred by using the Maximum Likelihood method based on the Le_Gascuel_2008 model (49). Initial tree (s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using a JTT model, and then selecting the topology with superior log likelihood value. A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories (+G, parameter = 2.8806)). The tree was drawn to scale, with branch lengths measured in the number of substitutions per site. All positions with less than 95% site coverage were eliminated. Evolutionary analyses were conducted in MEGA7 and prettified by iTOL (https://itol.embl.de/) (50).

Results

All the genes encoding aKEOPS subunits are essential for cell viability in *S. islandicus*

The KEOPS in eukaryotes is comprised of five subunits, Gon7, Pcc1, Kae1, Bud32, and Cgi121 (17,18). Except for Gon7 that has only been identified in eukaryotes thus far, all subunits of the eukaryotic KEOPS complex are highly conserved in archaea, (17,21,33). Interestingly, many archaea harbour two Pcc1 paralogs denoted Pcc1 and Pcc1-like, which are encoded by SiRe_1278 and SiRe_1701, respectively, in *Saccharolobus islandicus* REY15A (51). Archaeal Pcc1 shares a higher sequence similarity to the eukaryotic Pcc1 than the Pcc1-like does. To address the role of KEOPS in archaea (aKEOPS), we attempted to knock out each of its encoding genes in the crenarchaeon using the genome-editing method based on the endogenous CRISPR-Cas systems (41). These included the two *pcc1* genes described above and those coding for Kae1, Bud32, and Cgi121. We found that none of these genes could be deleted after more than three attempts (Fig. S1), consistent with a previous report that they are essential in the *S. islandicus* M16.4 strain (52). In bacteria, the genes involved in t^6A synthesis could not be knocked out in nearly all species, possibly due to their essential role in the t^6A modification (53). In budding yeast, deletion of the KEOPS subunit genes resulted in a t^6A modification defect and severe slow growth (54). Therefore, this raised a question if the observed essentiality of KEOPS genes in *S. islandicus* could be attributed to the t^6A modification

function or other fundamental ones that are essential for cell viability such as HR repair in Sulfolobales (31,32).

aKEOPS genes are still essential in the *S. islandicus* strain harboring an active bacterial t^6A modification system

In bacteria, TsaC and the threonylcarbamoyl (TC) transferase complex TCT complex (TsaB/TsaD/TsaE) function in the t^6A biosynthesis (55-57). Thus, we tested if the bacterial TSA system could functionally complement the t^6A modification function of aKEOPS in *S. islandicus*. A knock-in plasmid was constructed carrying *Tm*TsaB/TsaD/TsaE/TsaC, a tandem array of four genes, each of which was fused to the constitutive promoter of glutamate dehydrogenase (P_{gdhA}) of the crenarchaeon. After transformation and colony screening, genotypes having only two and three of the genes, *amyα::Tm*TsaE/TsaC and *Tm*TsaB/TsaE/TsaC, respectively, were obtained, suggesting recombination between their P_{gdhA} promoters (586 bp) could have occurred during the transformation and gene editing processes (Fig. S2A and S2B). Then, the fourth gene *Tm*TsaD was inserted into a gene encoding the endo-β-mannanase in the *amyα::Tm*TsaB/TsaE/TsaC strain, yielding a strain containing the complete *T. maritima* t^6A modification system (named as TsaKI) (Fig. 1A). The *T. maritima* t^6A modification system is active in the TsaKI strain as demonstrated in *cgi121* knocked-down experiments (Fig. 2A and 2B, and next section). Nevertheless, the t^6A level in TsaKI was similar to that in E233S (data not shown), implying that the t^6A level is already at the upper limit in E233S. The growth of TsaKI was slightly slower than the wild-type E233S (Fig. 1B). It is not clear whether this could be due to disruption of the endo-β-mannanase gene now.

To assess if TsaKI could be suitable for DNA repair study, we performed genotoxic sensitivity analysis of TsaKI in comparison with E233S. Both strains were cultured to the early log phase in sucrose broth from which samples were taken for preparing cell suspensions in a series of dilution. Two assays were then conducted for testing their sensitivity to DNA damage agents. Firstly, the cell suspensions were dotted on the sucrose plates supplemented with the DNA damage agent hydroxyurea (HU), and secondly, after dotting onto drug-free plates, the dotted plates were treated with UV radiation. As shown in Fig. 1C, the growth difference between TsaKI and the wild type strain was the same on the plates with or without DNA damage treatment, indicating that knock-in of the *T. maritima* t^6A modification system did not have any influence on DNA repair in *S. islandicus*.

Then, the KEOPS knock-out plasmids were introduced into TsaKI along with E233S. We found that similar transformation frequencies were obtained for all KEOPS knock-out plasmids in both strains, which was much lower than that of pSeSD, a reference plasmid, suggesting that these transformants could

carry escape mutations that inactivate the CRISPR immune responses. Indeed, analysis of their genotypes revealed that all transformants were wild type (Fig. S1), indicating that KEOPS subunits have other essential function *in vivo* in addition to the t^6A modification.

Knock-down of *cgi121* leads to reduction of the t^6A level that can be complemented by the bacterial modification system

Since all KEOPS subunits are essential either in E233S or TsaKI, suggesting KEOPS has additional functions besides t^6A modification, we performed gene knock-down analysis to confirm this and further characterize these essential functions in *S. islandicus*. A previous transcriptomic study showed that the transcripts of *pcc1*, *pcc1-like*, and *bud32* were quite low (FPKM~60-100), while those of *kae1* and *cgi121* were at a moderate transcription level (FPKM~400-800) (58). In addition, CRISPR knockdown in this archaeon is based on Type III-B CRISPR-Cas systems, and a prerequisite for the gene knockdown experiment is the presence of a protospacer-flanking sequence in target genes matching 5'-GAAAG-3', the 5'-handle tag derived from the CRISPR repeat sequences (41). However, such a motif is lacking in the coding sequence of *cgi121*. Therefore, we harnessed two methods for gene knock-down: CRISPR-Cas-based mRNA cleavage for *pcc1*, *pcc1-like*, *kae1*, and *bud32*, and replacement of the native promoters with mutant P_{araS} ($P_{araS}24$ or $P_{araS}38$) (59) for *cgi121* and the *kae1-bud32* operon (Fig. S2C and S2D). Proteins can only slightly expressed in the presence of sucrose and induced (about 20-fold) in the presence of arabinose with the mutant promoters (59).

As shown in Fig. 2A, the transcriptional levels of *kae1*, *bud32*, and *cgi121* in the $P_{araS}38$ (left and middle) replacement strains were reduced to 40-70 % of those of the wild type in E233S and TsaKI in sucrose medium. A severe growth retardation was observed in the *cgi121* knock-down E233S strain in sucrose medium (Fig. 2A, right), whereas the retardation disappeared when the strain was cultivated in arabinose medium (Fig. S3). In contrast, the *cgi121* knock-down TsaKI strain still exhibited growth retardation, although less severe, indicating that the *T. maritima* t^6A modification system could not fully rescue the growth retardation (Fig. 2A, right). These results suggest that Cgi121 possesses a fundamental function independent of t^6A modification. Interestingly, knock-down of *kae1-bud32* in E233S had no effect, but slightly stimulated the cell growth in the TsaKI strain.

To study whether the t^6A modification level was affected in the knock-down strains, we estimated their t^6A modification levels by HPLC. Consistent with the growth phenotypes, *kae1-bud32* knock-down in E233S or TsaKI did not reduce the t^6A levels as compared with that in normal E233S (Fig. 2B); in contrast, the *cgi121* knock-down led to a dramatic reduction in t^6A modification level in the knock-down E233S strain. However, the t^6A modification level in the *cgi121*

knock-down TsaKI strain did not change significantly as compared with that in E233S (Fig. 2B). The results indicated that Cgi121 is critical for t^6A modifications and that the *T. maritima* t^6A system produced sufficient t^6A modification to support the normal cell growth in the *cgi121* knock-down and perhaps also other TsaKI strains.

The aKEOPS complex is involved in DNA repair in *S. islandicus*

To reveal potential roles of archaeal KEOPS in DNA repair, we performed genotoxic sensitivity analysis of different strains to DNA damage agents. As shown in Fig. 2C, on the untreated plates, the cell growth of the *kae1-bud32* knock-down E233S strain was not affected, whereas the *kae1-bud32* knock-down TsaKI strain grew better than that of the TsaKI strain. Moreover, *cgi121* knock-down E233S and TsaKI strains showed growth retardation, consistent with the growth curves (Fig. 2A). We found that neither the *kae1-bud32* operon knock-down E233S strain nor the *kae1-bud32* operon knock-down TsaKI strain showed any difference with various DNA damage agent treatment (Fig. 2C). Strikingly, both *cgi121* knock-down E233S and *cgi121* knock-down TsaKI strains exhibited reduction in the cell sensitivity to HU (Fig. 2C), indicating that the resistance resulted from *cgi121* knock-down was independent of its t^6A function.

We also conducted knock-down of *kae1*, *bud32*, *pcc1*, and *pcc1-like* by the CRISPR-Cas mediated targeting method (Fig. S4A). We attempted to assess the transcription levels of these genes by qPCR, however, the results were not consistent, perhaps due to low levels of transcripts of these genes. Nevertheless, we found that the sensitivity of presumably *bud32* knock-down strain to 25 J/m^2 UV and HU was reduced while the sensitivity of other knock-down strains to diverse DNA damage agents increased (Fig. S4B). These results suggested that each subunit of the *Sulfolobus* KEOPS complex is involved in DNA repair, but they may play different roles in this process.

Overexpression of KEOPS subunits in *S. islandicus*

To further investigate the role of KEOPS subunits in DNA repair and to get more insight into the involved mechanisms, the subunits were overexpressed alone or in different combinations in *S. islandicus*. We found that overexpression of either Kae1 or Bud32, the two core components of KEOPS complex, inhibited the cell growth, with the latter showing more profound growth inhibition than the former (Fig. 3A). Interestingly, overexpression of a kinase/ATPase inactive mutant Bud32D134A did not have this inhibition, suggesting that the growth retardation was due to its kinase/ATPase activity (Fig. 3A). When Kae1 and Bud32 were co-overexpressed in a native gene organization manner (Fig. 3C), no growth retardation was observed. However, cell growth was inhibited if the expression of the two genes was controlled by two separate P_{araS} promoters

and induced by arabinose (Fig. 3A and 3C). In a previous study, it was shown that the transcript of *bud32* was several folds lower than that of *kae1 (58)*. This could be attributed to the difference in the stability of the two transcripts. Alternatively, an unknown mechanism could exist to maintain the ratio of Kae1 and Bud32 *in vivo.* Additionally, we also analyzed the effect of the expression of Pcc1, Pcc1-like, and Cgi121, and co-expression of Cgi121 and Bud32 on the cell growth. When overexpressed alone in *S. islandicus*, the proteins of Pcc1, Pcc1-like, and Cgi121 could not be detected by Western blotting. However, when Cgi121 and Bud32 were co-expressed, both proteins were clearly detected (Fig. 3B), indicating that Cgi121 and Bud32 may stabilize with each other *in vivo*. Interestingly, the growth of Cgi121 and Bud32 co-overexpression strain had no difference with the strain carrying the control plasmid pSeSD. Taken together, our results imply that Kae1 and Bud32 and the kinase/ATPase activity of Bud32 have significant physiological functions and that Kae1 and Cgi121 have regulatory roles for the activity of Bud32. The results also suggest that the levels of KEOPS subunits, especially Bud32, Kae1, and Cgi121 must be maintained in an appropriate proportion in *S. islandicus* for their stability and proper functions.

aKEOPS subunits exhibit dsDNA binding activity

It was reported that tRNA recruitment and recognition is the key step in t^6A modification, in which Cgi121 is responsible for tRNA recruitment while other subunits, Pcc1, Kae1, and Bud32, contribute to specific tRNA recognition (37). Interestingly, reconstructed *S. cerevisiae* KEOPS complex exhibits ssDNA and dsDNA binding activity *in vitro* (19). To probe possible roles and mechanism of aKEOPS in DNA repair, we analyzed the DNA binding activity of each subunit and subcomplexes. We found that Cgi121 and Bud32 were able to bind to dsDNA (Fig. 4A), although the binding activity of Bud32 was much weaker than that of Cgi121. Intriguingly, Kae1 and Pcc1 in combination, not Kae1 or Pcc1 alone, also exhibited dsDNA binding activity (Fig. 4B). It was shown that Cgi121 is an ssRNA binding protein which can specifically recognize tRNA 3' CCA tail (37). Since 6×His-tagged *Sis*Cgi121 and *Sso*Cgi121 (Cgi121 from *S. solfataricus*) were always degraded when they were expressed in *S. islandicus* or *E. coli* (data not shown), a maltose-binding protein (MBP) tag (40 kDa) was fused to *Sis*Cgi121 for purification. To exclude the potential effect of the MBP tag on DNA binding ability of *Sis*Cgi121, we also purified the wild type and the tRNA 3' CCA tail binding deficient mutant I75E of *Tko*Cgi121 for comparison. Interestingly, the wild type *Tko*Cgi121 was able to bind to dsDNA while the I75E mutant lost the binding ability. The I75E mutant, not *Tko*Cgi121, and MBP-*Sis*Cgi121 had weak ssDNA binding activity at a concentration of 20 μM, which is likely nonspecific binding (Fig. 4C and 4D). The results revealed that Ile75, which is indispensable for tRNA 3' CCA tail binding in *Tko*Cgi121 (37), is critical for dsDNA binding.

Pcc1-like is the fifth subunit of KEOPS complex in archaea

Four subunits of KEOPS, Pcc1, Kae1, Bud32, and Cgi121, are highly conserved in archaea and eukaryotes. Although the fifth subunit Gon7 were identified in yeast and human (17,18,33), it is still unclear whether aKEOPS contains a counterpart of Gon7. Our genetic work revealed that both *pcc1* and *pcc1-like* were essential in *S. islandicus* (Fig. S1), indicating that their functions *in vivo* are not redundant. Because both Pcc1 and Gon7 are small proteins, it is reasonable to ask whether Pcc1-like is the fifth subunit in archaea. To investigate this, we conducted affinity purification and mass spectrum analysis by overexpression of the tagged subunits in *S. islandicus* REY15A. As shown in Table 1, both Pcc1 and Pcc1-like were co-purified with Kae1 and Kae1-Bud32 from the Kae1 and Kae1-Bud32 overexpression cells, respectively, while Kae1 and Pcc1-like were co-purified with Pcc1 from the Pcc1 overexpression cells. Intriguingly, Sua5, which catalyzes the synthesis of t^6A intermediate threonylcarbamoyl-AMP, was also co-purified with Kae1 and Kae1-Bud32 from their overexpression cells, respectively, indicating that Sua5 might interact with KEOPS complex *in vivo*. *In vitro* pull down analysis showed that Pcc1-like was able to interact with Pcc1 and Kae1, respectively (Fig. S5A). Therefore, we provide strong biochemical evidences that Pcc1-like was the fifth subunit of Sulfolobales KEOPS.

Pcc1-like eliminates Pcc1-mediated dimerization of Pcc1-Kae1 subcomplex *in vitro*

Thus far, there is no study on the subunit interaction of KEOPS from crenarchaea. A linear subunit organization of Pcc1:Kae1:Bud32:Cgi121 was proposed for the KEOPS complex of a euryarchaeon *Methanococcus jannaschii* and budding yeast *S. cerevisiae* (21,60). To determine the inter-subunit relationship of the crenarchaeal KEOPS, each subunit (except for Pcc1) of the KEOPS from *S. islandicus* was expressed in and purified from *E. coli* and pull-down assay was performed. As shown in Fig. S5B and S5C, Bud32 was pulled down by His-tagged Kae1 and MBP-tagged Cgi121, respectively, verifying that there are interactions between Bud32 and Kae1 as well as between Bud32 and Cgi121. Since Pcc1 is insoluble when it is expressed in *E. coli* alone, we co-express Pcc1 and Kae1 in *E. coli*. His-Pcc1 is able to pull down Kae1, confirming that Pcc1 interacts with Kae1 (Fig. S5A, middle). Therefore, linear subunit organization of Pcc1:Kae1:Bud32:Cgi121 should also be the case for KEOPS from crenarchaea.

The whole KEOPS complex of *S. cerevisiae* exists as a dimeric structure mediated by a Pcc1 dimer, each monomer of which interacts with a separate Kae1 molecule (36). Gon7 interacts with Pcc1 and abolishes this dimerization (60). To understand whether archaeal Pcc1-like shares the same property with eukaryotic Gon7, we examined subcomplex formation among Pcc1, Pcc1-like,

and Kae1 by size exclusion chromatography (SEC). As shown in Fig. 5A, the peak of Kae1-Pcc1 subcomplex is corresponding to a 90 kDa complex, which was close to a 2:2 subcomplex with the expected molecular mass of 96 kDa. Kae1-Pcc1-like subcomplex was eluted at a peak of 66 kDa corresponding to a Kae1:2 Pcc1-like subcomplex with the predicted molecular mass of 58 kDa when Pcc1-like was excessed over Kae1 (Fig. 5B), indicating that Pcc1-like could not mediate the dimerization of Kae1-Pcc1-like subcomplex (ca. 96 kDa). When Kae1 was excessed, the stoichiometry of Kae1-Pcc1-like subcomplex became 1:1 with an observed mass of 42 kDa close to the expected mass of 48 kDa (Fig. 5B). Notably, when Pcc1-like was added into the mixture of 2Kae1:2Pcc1 subcomplex, the 2:2 peak disappeared and a 51 kDa peak was formed (Fig. 5A), suggesting a 1:1:1 stoichiometry of Kae1:Pcc1:Pcc1-like. It should be noted that the elution volume of Kae1:Pcc1:Pcc1-like subcomplex was between those of Kae1:2 Pcc1-like and Kae1:Pcc1-like, but not close to that of Kae1:2 Pcc1-like, suggesting that there was a putative conformational change after the ternary subcomplex was formed. These results suggest that Pcc1 mediates the dimerization of Kae1-Pcc1 while Pcc1-like disrupts the 2Kae1:2Pcc1 subcomplex and stabilizes the Kae1-Pcc1 subcomplex, the same as eukaryotic Pcc1 and Gon7 in eukaryotes (37,60). In *M. jannaschii*, there was a small conformational change when the 2:2 Kae1-Pcc1 subcomplex changed to a 1:2 one, implying a transition state during Pcc1-like evolution (36). Our structural superposition shows that Pcc1-like matches with yeast and human Gon7 proteins with two β sheets and one α helix at the interaction surface of Gon7-Pcc1 (Fig. 5C)(61), reinforcing that Pcc1-like is a potential Gon7 homolog. Hence, we propose that archaea Pcc1-like is the putative functional and structural homolog of Gon7.

Pcc1-like is present in most archaeal superphyla forming a lineage with Gon7 homologs

To get more insight into the relationship between the Gon7 homolog and Pcc1-like, we searched for the distribution of the homologs of Sua5 and KEOPS subunits in all archaeal and representative eukaryotic phyla. It is confirmed that only Sua5, Pcc1, Kae1, and Bud32, the minimal components for t^6A modification, are conserved in all archaea and eukaryotes (Fig. 6A). Consistently with the previous report, no sequence homolog of Gon7 is found in archaea as a single domain protein (33). However, in rare cases, Gon7 is present as a domain of a proteasome-activating nucleotidase, such as WP_015321675.1 from *Natronococcus occultus* SP4, WP_007139789.1 from *Halobiforma lacisalsi* AJ5, and ELY65446.1 from *Natronococcus jeotgali* DSM 18795, or MutS (e.g., WP_008364526.1 from *Halorubrum litoreum*) (Fig. 6B) in some halophilic archaea. Similar to Sulfolobales, most archaea have two Pcc1 paralogs (Fig. 6A). Pcc1 is more widely distributed than Pcc1-like. The distribution of Pcc1-like is also found in most superphyla with the least appearance in the DPANN superphylum (Fig. 6A). Sequence alignment and

phylogenetic analysis of Pcc1, Pcc1-like, and Gon7 reveals that Pcc1 exhibits a higher evolution rate, and archaeal Pcc1-like and some Gon7 homologs form a separated lineage which is distal to both archaeal Pcc1 and eukaryotic Pcc1 (Fig. 6B). These results support that one of the two Pcc1 paralogs in archaea, Pcc1-like, is the functional Gon7 homolog of eukaryotes.

Discussion

DNA homologous recombination (HR) pathway is one of most complicated and the mechanism has not been fully elucidated. In this study, we demonstrate that KEOPS has a t^6A modification-independent function that is essential in *S. islandicus*, a thermophilic crenarchaeon. To circumvent the possible essentiality of archaeal KEOPS, we introduced a thermophilic bacteria-derived t^6A modification system (53) to *S. islandicus* and performed genetic analysis of the KEOPS. We show that the bacterial t^6A modification system is active in *S. islandicus* can complement the t^6A modification function of KEOPS. Using this constructed knock-in strain TsaKI, we demonstrate that all the KEOPS subunits cannot be deleted and knock-down of different subunits lead to increase or decrease in genotoxic sensitivity, indicating that aKEOPS is involved in DNA repair independent of t^6A modification function. This is similar to KEOPS in yeast, where the KEOPS stimulates DNA end resection during recombination and telomere maintenance, independently of t^6A modification (19,26).

Among the DNA repair pathways of thermophilic archaea such as Sulfolobales and Thermococales, only the HR pathway is essential and most genes involved in HR cannot be knocked out (31,32,52,62). Considering this, the pathway in which KEOPS involved in *S. islandicus* is most likely HR repair. In halophilic euryarchaeon *H. volcanii*, *kae1*, *bud32*, and *cgi121* are essential for cell viability, whereas deletion of *pcc1* resulted in cells with higher nucleic acid content (38). The authors did explain the cause of this observation, but we now assume this could be due to presence of higher genome copy number which may indicate that KEOPS may be involved in DNA repair *H. volcanii* is highly polyploid with approximately 20 copies of genomes per cell and is capable of origin-free replication depending on DNA recombination (63,64). Therefore, in addition to t^6A modification, *H. volcanii* KEOPS might also contribute to this polyploidy species-specific, HR-dependent replication mode. In humans, knock-down of genes encoding KEOPS subunits induces the DDR signalling and subsequent apoptosis (65). Therefore, involvement in DSB resection could be a conserved and ancestral function of KEOPS. We found that knock-down of *cgi121* and *bud32* individually reduced the sensitivity to HU while knock-down of other subunit genes increased the sensitivity to diverse DNA damage agents. On the other hand, overexpression of different subunits, their mutant, or co-expression of combinations of subunits had distinct effect on cell growth. The phenotypes resemble those in the yeast to some extent. It was reported that the activity sites of Bud32 plays a critical role in HR repair in *S.*

cerevisiae, and the kinase activity of *S. cerevisiae* Bud32 was inhibited by Kae1 and activated by Cgi121 *in vitro* (19,21,66). These results suggest that there is a sophisticated regulatory mechanism of the KEOPS complex. The mechanism how KEOPS participates in this HR and how this function is coordinated with the t^6A modification will be an intriguing subject for further investigation.

A mystery of KEOPS is the molecular basis to explain why KEOPS has such seemingly un-related three functions, t^6A, HR repair, and telomere maintenance in yeast. Up to now, mechanisms have been elucidated for the t^6A modification of KEOPS except for the precise roles of Gon7 and Pcc1 in the process. Cgi121 was reported to be a tRNA 3' CCA tail and ssRNA binding protein for the recruitment of tRNA to KEOPS during t^6A modification (37). Interestingly, telomeres in most eukaryotes are composed of TT[A/T]GGG repeats, whose antisense strands and corresponding telomerase RNAs are CCC[T/A]AA repeats (67). Therefore, the nucleic acids CCA binding ability of Cgi121 may be a link among the three functions of eukaryotic KEOPS. Our study revealed that Cgi121 is the only subunit that binds to a long dsDNA individually. Strikingly, mutation at the tRNA binding residue of *Tko*Cgi121 led to a defect in long dsDNA binding ability, suggesting that the mechanisms of KEOPS in t^6A modification and DNA repair are related. The evolutionary transition from circular to linear chromosomes, which harbour telomeric structures, brought more challenges to genome stability in eukaryotic cells. This could be the reason why multiple DNA repair proteins, especially classical HR proteins and KEOPS also participate in telomere maintenance.

We identified a Pcc1 paralog in archaea, Pcc1-like, that is likely the functional homolog of Gon7. As other subunits of aKEOPS, *pcc1-like* can be deleted neither. Pcc1-like interacts with both Pcc1 and Kae1. It inhibits the dimerization of Kae1-Pcc1 and stabilize the subcomplex *in vitro*, just like Gon7. Additionally, bio-informational analysis shows that the Pcc1-Pcc1-like heterodimer interface resembles that of the eukaryotic Pcc1-Gon7 heterodimer. Interestingly, through sequence alignment and phylogenetic analysis, we found that the evolution rate of Pcc1 and Pcc1-like was higher than other KEOPS subunits and Pcc1-like lineage seems to have a closer relationship with Gon7 lineage than archaeal and eukaryotic Pcc1 in the phylogenetic tree (Fig. 6B), which supports that Pcc1-like is a functional homolog of Gon7. The fifth subunit Gon7 also functions in t^6A modification and DNA repair in *S. cerevisiae* and *H. sapiens* (19,25,33,54). However, because its structure is partially disordered, the mechanism of Gon7 in these functions is still unknown (25,60). Structurally, Gon7, interacts with Pcc1 and stabilizes the KEOPS forming a 1:1:1:1:1 complex (25,60). The precise role of Pcc1-like/Gon7 in HR repair and t^6A modification needs further study. To this end, Pcc1-like and aKEOPS as the whole complex from thermophilic archaea could be used as excellent structural and functional models for KEOPS study because of their stability.

Finally, based on studies on the composition and function of the t^6A modification systems, we proposed a model for the evolution of the KEOPS complex and other TC-transferase or its complexes (Fig. 7). The ancestral TC-transferase could be a single protein Kae1/TsaD, from which mitochondrial TC-transferase Ori7 and bacterial TsaD originated independently (54,68). Then, ATPase Bud32/TsaE and the dimerization module Pcc1/TsaB join to form an initial KEOPS complex, which contains the minimum components and is essential for t^6A modification (53,54). Meanwhile, Cgi121 come into the initial complex of archaea-eukaryotes lineage to recruit tRNA (37). At a later stage, Pcc1-like appears via a gene duplication, stabilizes the complex, and improves the efficiency of t^6A modification (Fig. S6) (25). Finally, due to the accessory role of Pcc1-like in t^6A modification, Pcc1-like exhibits the highest mutation rates among the subunits and evolved into Gon7 in eukaryotes. During the later stages as the additional subunits join, KEOPS gains functions other than t^6A modification, such as HR repair, before the separation of archaea and eukaryotes. In eukaryotes, KEOPS is also involved in telomere maintenance which also needs a process of DNA recombination, like other HR repair proteins such as Mre11 and Rad50 (69,70).

Acknowledgements

This work was supported by the National Key Research and Development Program of China (No. 2020YFA0906800), the National Natural Science Foundation of China (No. 31970546 and 31670061 to YS, 31900055 to QH, 31970119 to JN, and 31771380 to QS) and the State Key Laboratory of Microbial Technology. We would like to thank all the lab members of the CRISPR and Archaea Biology Research Centre for helpful discussions.

References

1. Ohnishi, T., Mori, E. and Takahashi, A. (2009) DNA double-strand breaks: their production, recognition, and repair in eukaryotes. *Mutation research*, **669**, 8-12.
2. Paull, T.T. (2021) Reconsidering pathway choice: a sequential model of mammalian DNA double-strand break pathway decisions. *Curr Opin Genet Dev*, **71**, 55-62.
3. Cejka, P. and Symington, L.S. (2021) DNA End Resection: Mechanism and Control. *Annu Rev Genet*, **55**, 285-307.
4. Lammens, K., Bemeleit, D.J., Mockel, C., Clausing, E., Schele, A., Hartung, S., Schiller, C.B., Lucas, M., Angermuller, C., Soding, J. *et al.* (2011) The Mre11:Rad50 structure shows an ATP-dependent molecular clamp in DNA double-strand break repair. *Cell*, **145**, 54-66.
5. Gravel, S., Chapman, J.R., Magill, C. and Jackson, S.P. (2008) DNA helicases Sgs1 and BLM promote DNA double-strand break resection. *Genes Dev*, **22**, 2767-2772.
6. Mimitou, E.P. and Symington, L.S. (2008) Sae2, Exo1 and Sgs1 collaborate in DNA double-strand break processing. *Nature*, **455**, 770-774.
7. Zhu, Z., Chung, W.H., Shim, E.Y., Lee, S.E. and Ira, G. (2008) Sgs1 helicase and two nucleases Dna2 and Exo1 resect DNA double-strand break ends. *Cell*, **134**, 981-994.

8. Cannavo, E., Cejka, P. and Kowalczykowski, S.C. (2013) Relationship of DNA degradation by Saccharomyces cerevisiae exonuclease 1 and its stimulation by RPA and Mre11-Rad50-Xrs2 to DNA end resection. *Proc Natl Acad Sci U S A*, **110**, E1661-1668.
9. Krejci, L., Altmannova, V., Spirek, M. and Zhao, X. (2012) Homologous recombination and its regulation. *Nucleic acids research*, **40**, 5795-5818.
10. Baroni, E., Viscardi, V., Cartagena-Lirola, H., Lucchini, G. and Longhese, M.P. (2004) The functions of budding yeast Sae2 in the DNA damage response require Mec1- and Tel1-dependent phosphorylation. *Mol Cell Biol*, **24**, 4151-4165.
11. Wang, H., Shi, L.Z., Wong, C.C., Han, X., Hwang, P.Y., Truong, L.N., Zhu, Q., Shao, Z., Chen, D.J., Berns, M.W. *et al.* (2013) The interaction of CtIP and Nbs1 connects CDK and ATM to regulate HR-mediated double-strand break repair. *PLoS Genet*, **9**, e1003277.
12. Bai, Y., Wang, W., Li, S., Zhan, J., Li, H., Zhao, M., Zhou, X.A., Li, S., Li, X., Huo, Y. *et al.* (2019) C1QBP Promotes Homologous Recombination by Stabilizing MRE11 and Controlling the Assembly and Activation of MRE11/RAD50/NBS1 Complex. *Mol Cell*, **75**, 1299-1314 e1296.
13. Paiano, J., Wu, W., Yamada, S., Sciascia, N., Callen, E., Paola Cotrim, A., Deshpande, R.A., Maman, Y., Day, A., Paull, T.T. *et al.* (2020) ATM and PRDM9 regulate SPO11-bound recombination intermediates during meiosis. *Nat Commun*, **11**, 857.
14. Soniat, M.M., Myler, L.R., Kuo, H.C., Paull, T.T. and Finkelstein, I.J. (2019) RPA Phosphorylation Inhibits DNA Resection. *Mol Cell*, **75**, 145-153 e145.
15. Broderick, R., Nieminuszczy, J., Baddock, H.T., Deshpande, R., Gileadi, O., Paull, T.T., McHugh, P.J. and Niedzwiedz, W. (2016) EXD2 promotes homologous recombination by facilitating DNA end resection. *Nat Cell Biol*, **18**, 271-280.
16. Wang, X., Dong, Y., Zhao, X., Li, J., Lee, J., Yan, Z., Yang, S., Wu, W., Hou, X., Liu, G. *et al.* (2021) Rtt105 promotes high-fidelity DNA replication and repair by regulating the single-stranded DNA-binding factor RPA. *Proc Natl Acad Sci U S A*, **118**.
17. Downey, M., Houlsworth, R., Maringele, L., Rollie, A., Brehme, M., Galicia, S., Guillard, S., Partington, M., Zubko, M.K., Krogan, N.J. *et al.* (2006) A genome-wide screen identifies the evolutionarily conserved KEOPS complex as a telomere regulator. *Cell*, **124**, 1155-1168.
18. Kisseleva-Romanova, E., Lopreiato, R., Baudin-Baillieu, A., Rousselle, J.C., Ilan, L., Hofmann, K., Namane, A., Mann, C. and Libri, D. (2006) Yeast homolog of a cancer-testis antigen defines a new transcription complex. *EMBO J*, **25**, 3576-3585.
19. He, M.H., Liu, J.C., Lu, Y.S., Wu, Z.J., Liu, Y.Y., Wu, Z., Peng, J. and Zhou, J.Q. (2019) KEOPS complex promotes homologous recombination via DNA resection. *Nucleic acids research*, **47**, 5684-5697.
20. Koonin, E.V. (2003) Comparative genomics, minimal gene-sets and the last universal common ancestor. *Nat Rev Microbiol*, **1**, 127-136.
21. Mao, D.Y., Neculai, D., Downey, M., Orlicky, S., Haffani, Y.Z., Ceccarelli, D.F., Ho, J.S., Szilard, R.K., Zhang, W., Ho, C.S. *et al.* (2008) Atomic structure of the KEOPS complex: an ancient protein kinase-containing molecular machine. *Mol Cell*, **32**, 259-275.
22. Srinivasan, M., Mehta, P., Yu, Y., Prugar, E., Koonin, E.V., Karzai, A.W. and Sternglanz,

R. (2011) The highly conserved KEOPS/EKC complex is essential for a universal tRNA modification, t6A. *EMBO J*, **30**, 873-881.

23. Murphy, F.V.t., Ramakrishnan, V., Malkiewicz, A. and Agris, P.F. (2004) The role of modifications in codon discrimination by tRNA(Lys)UUU. *Nat Struct Mol Biol*, **11**, 1186-1191.
24. Thiaville, P.C., Legendre, R., Rojas-Benitez, D., Baudin-Baillieu, A., Hatin, I., Chalancon, G., Glavic, A., Namy, O. and de Crecy-Lagard, V. (2016) Global translational impacts of the loss of the tRNA modification t(6)A in yeast. *Microb Cell*, **3**, 29-45.
25. Arrondel, C., Missoury, S., Snoek, R., Patat, J., Menara, G., Collinet, B., Liger, D., Durand, D., Gribouval, O., Boyer, O. *et al.* (2019) Defects in t(6)A tRNA modification due to GON7 and YRDC mutations lead to Galloway-Mowat syndrome. *Nat Commun*, **10**, 3967.
26. Liu., He, M.H., Liu, J.C., Lu, Y.S., Peng, J. and Zhou, J.Q. (2018) Yeast KEOPS complex regulates telomere length independently of its t(6)A modification function. *J Genet Genomics*, **45**, 247-257.
27. White, M.F. (2011) Homologous recombination in the archaea: the means justify the ends. *Biochem Soc Trans*, **39**, 15-19.
28. Hopfner, K.P., Craig, L., Moncalian, G., Zinkel, R.A., Usui, T., Owen, B.A., Karcher, A., Henderson, B., Bodmer, J.L., McMurray, C.T. *et al.* (2002) The Rad50 zinc-hook is a structure joining Mre11 complexes in DNA recombination and repair. *Nature*, **418**, 562-566.
29. Williams, R.S., Moncalian, G., Williams, J.S., Yamada, Y., Limbo, O., Shin, D.S., Groocock, L.M., Cahill, D., Hitomi, C., Guenther, G. *et al.* (2008) Mre11 dimers coordinate DNA end bridging and nuclease processing in double-strand-break repair. *Cell*, **135**, 97-109.
30. Sung, S., Li, F., Park, Y.B., Kim, J.S., Kim, A.K., Song, O.K., Kim, J., Che, J., Lee, S.E. and Cho, Y. (2014) DNA end recognition by the Mre11 nuclease dimer: insights into resection and repair of damaged DNA. *EMBO J*, **33**, 2422-2435.
31. Grogan, D.W. (2015) Understanding DNA Repair in Hyperthermophilic Archaea: Persistent Gaps and Other Reasons to Focus on the Fork. *Archaea*, **2015**, 942605.
32. Huang, Q., Liu, L., Liu, J., Ni, J., She, Q. and Shen, Y. (2015) Efficient 5'-3' DNA end resection by HerA and NurA is essential for cell viability in the crenarchaeon Sulfolobus islandicus. *BMC Mol Biol*, **16**, 2.
33. Wan, L.C., Maisonneuve, P., Szilard, R.K., Lambert, J.P., Ng, T.F., Manczyk, N., Huang, H., Laister, R., Caudy, A.A., Gingras, A.C. *et al.* (2017) Proteomic analysis of the human KEOPS complex identifies C14ORF142 as a core subunit homologous to yeast Gon7. *Nucleic acids research*, **45**, 805-817.
34. Perrochia, L., Crozat, E., Hecker, A., Zhang, W., Bareille, J., Collinet, B., van Tilbeurgh, H., Forterre, P. and Basta, T. (2013) In vitro biosynthesis of a universal t6A tRNA modification in Archaea and Eukarya. *Nucleic acids research*, **41**, 1953-1964.
35. Perrochia, L., Guetta, D., Hecker, A., Forterre, P. and Basta, T. (2013) Functional assignment of KEOPS/EKC complex subunits in the biosynthesis of the universal t6A tRNA modification. *Nucleic acids research*, **41**, 9484-9499.

36. Wan, L.C., Pillon, M.C., Thevakumaran, N., Sun, Y., Chakrabartty, A., Guarne, A., Kurinov, I., Durocher, D. and Sicheri, F. (2016) Structural and functional characterization of KEOPS dimerization by Pcc1 and its role in t6A biosynthesis. *Nucleic acids research*, **44**, 6971-6980.
37. Beenstock, J., Ona, S.M., Porat, J., Orlicky, S., Wan, L.C.K., Ceccarelli, D.F., Maisonneuve, P., Szilard, R.K., Yin, Z., Setiaputra, D. *et al.* (2020) A substrate binding model for the KEOPS tRNA modifying complex. *Nat Commun*, **11**, 6233.
38. Naor, A., Thiaville, P.C., Altman-Price, N., Cohen-Or, I., Allers, T., de Crecy-Lagard, V. and Gophna, U. (2012) A genetic investigation of the KEOPS complex in halophilic Archaea. *PLoS One*, **7**, e43013.
39. Huang, Q., Lin, Z., Wu, P., Ni, J. and Shen, Y. (2020) Phosphoproteomic Analysis Reveals Rio1-Related Protein Phosphorylation Changes in Response to UV Irradiation in Sulfolobus islandicus REY15A. *Front Microbiol*, **11**, 586025.
40. Peng, Deng, L., Mei, Y., Jiang, D., Hu, Y., Awayez, M., Liang, Y. and She, Q. (2012) A synthetic arabinose-inducible promoter confers high levels of recombinant protein expression in hyperthermophilic archaeon Sulfolobus islandicus. *Appl Environ Microbiol*, **78**, 5630-5637.
41. Li, Y., Pan, S., Zhang, Y., Ren, M., Feng, M., Peng, N., Chen, L., Liang, Y.X. and She, Q. (2016) Harnessing Type I and Type III CRISPR-Cas systems for genome editing. *Nucleic acids research*, **44**, e34.
42. Peng, N., Ao, X., Liang, Y.X. and She, Q. (2011) Archaeal promoter architecture and mechanism of gene activation. *Biochem Soc Trans*, **39**, 99-103.
43. Porat, J. and Bayfield, M.A. (2020) Use of tRNA-Mediated Suppression to Assess RNA Chaperone Function. *Methods Mol Biol*, **2106**, 107-120.
44. Gehrke, C.W. and Kuo, K.C. (1989) Ribonucleoside analysis by reversed-phase high-performance liquid chromatography. *J Chromatogr*, **471**, 3-36.
45. Saitou, N. and Nei, M. (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol*, **4**, 406-425.
46. Nei, M. and Kumar, S. (2000) *Molecular evolution and phylogenetics*. Oxford University Press, Oxford ; New York.
47. Katoh, K., Rozewicki, J. and Yamada, K.D. (2019) MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief Bioinform*, **20**, 1160-1166.
48. Kuraku, S., Zmasek, C.M., Nishimura, O. and Katoh, K. (2013) aLeaves facilitates on-demand exploration of metazoan gene family trees on MAFFT sequence alignment server with enhanced interactivity. *Nucleic acids research*, **41**, W22-28.
49. Le, S.Q. and Gascuel, O. (2008) An improved general amino acid replacement matrix. *Mol Biol Evol*, **25**, 1307-1320.
50. Kumar, S., Stecher, G. and Tamura, K. (2016) MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol Biol Evol*, **33**, 1870-1874.
51. Guo, L., Brugger, K., Liu, C., Shah, S.A., Zheng, H., Zhu, Y., Wang, S., Lillestol, R.K., Chen, L., Frank, J. *et al.* (2011) Genome analyses of Icelandic strains of Sulfolobus islandicus, model organisms for genetic and virus-host interaction studies. *J Bacteriol*, **193**, 1672-1680.

52. Zhang, C., Phillips, A.P.R., Wipfler, R.L., Olsen, G.J. and Whitaker, R.J. (2018) The essential genome of the crenarchaeal model Sulfolobus islandicus. *Nat Commun*, **9**, 4908.
53. Thiaville, P.C., El Yacoubi, B., Kohrer, C., Thiaville, J.J., Deutsch, C., Iwata-Reuyl, D., Bacusmo, J.M., Armengaud, J., Bessho, Y., Wetzel, C. *et al.* (2015) Essentiality of threonylcarbamoyladenosine (t(6)A), a universal tRNA modification, in bacteria. *Mol Microbiol*, **98**, 1199-1221.
54. Wan, L.C., Mao, D.Y., Neculai, D., Strecker, J., Chiovitti, D., Kurinov, I., Poda, G., Thevakumaran, N., Yuan, F., Szilard, R.K. *et al.* (2013) Reconstitution and characterization of eukaryotic N6-threonylcarbamoylation of tRNA using a minimal enzyme system. *Nucleic acids research*, **41**, 6332-6346.
55. Luthra, A., Swinehart, W., Bayooz, S., Phan, P., Stec, B., Iwata-Reuyl, D. and Swairjo, M.A. (2018) Structure and mechanism of a bacterial t6A biosynthesis system. *Nucleic acids research*, **46**, 1395-1411.
56. Missoury, S., Plancqueel, S., Li de la Sierra-Gallay, I., Zhang, W., Liger, D., Durand, D., Dammak, R., Collinet, B. and van Tilbeurgh, H. (2018) The structure of the TsaB/TsaD/TsaE complex reveals an unexpected mechanism for the bacterial t6A tRNA-modification. *Nucleic acids research*, **46**, 5850-5860.
57. Luthra, A., Paranagama, N., Swinehart, W., Bayooz, S., Phan, P., Quach, V., Schiffer, J.M., Stec, B., Iwata-Reuyl, D. and Swairjo, M.A. (2019) Conformational communication mediates the reset step in t6A biosynthesis. *Nucleic acids research*, **47**, 6551-6567.
58. Sun, M., Feng, X., Liu, Z., Han, W., Liang, Y.X. and She, Q. (2018) An Orc1/Cdc6 ortholog functions as a key regulator in the DNA damage response in Archaea. *Nucleic acids research*, **46**, 6697-6711.
59. Peng , Xia, Q., Chen, Z., Liang, Y.X. and She, Q. (2009) An upstream activation element exerting differential transcriptional activation on an archaeal promoter. *Mol Microbiol*, **74**, 928-939.
60. Zhang, W., Collinet, B., Graille, M., Daugeron, M.C., Lazar, N., Libri, D., Durand, D. and van Tilbeurgh, H. (2015) Crystal structures of the Gon7/Pcc1 and Bud32/Cgi121 complexes provide a model for the complete yeast KEOPS complex. *Nucleic acids research*, **43**, 3358-3372.
61. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Zidek, A., Potapenko, A. *et al.* (2021) Highly accurate protein structure prediction with AlphaFold. *Nature*, **596**, 583-589.
62. Fujikane, R., Ishino, S., Ishino, Y. and Forterre, P. (2010) Genetic analysis of DNA repair in the hyperthermophilic archaeon, Thermococcus kodakaraensis. *Genes Genet Syst*, **85**, 243-257.
63. Breuert, S., Allers, T., Spohn, G. and Soppa, J. (2006) Regulated polyploidy in halophilic archaea. *PLoS One*, **1**, e92.
64. Hawkins, M., Malla, S., Blythe, M.J., Nieduszynski, C.A. and Allers, T. (2013) Accelerated growth in the absence of DNA replication origins. *Nature*, **503**, 544-547.
65. Braun, D.A., Rao, J., Mollet, G., Schapiro, D., Daugeron, M.C., Tan, W., Gribouval, O., Boyer, O., Revy, P., Jobst-Schwan, T. *et al.* (2017) Mutations in KEOPS-complex genes

cause nephrotic syndrome with primary microcephaly. *Nat Genet*, **49**, 1529-1538.

66. Hecker, A., Lopreiato, R., Graille, M., Collinet, B., Forterre, P., Libri, D. and van Tilbeurgh, H. (2008) Structure of the archaeal Kae1/Bud32 fusion protein MJ1130: a model for the eukaryotic EKC/KEOPS subcomplex. *EMBO J*, **27**, 2340-2351.
67. Greider, C.W. (1990) Telomeres, telomerase and senescence. *Bioessays*, **12**, 363-369.
68. Hecker, A., Leulliot, N., Gadelle, D., Graille, M., Justome, A., Dorlet, P., Brochier, C., Quevillon-Cheruel, S., Le Cam, E., van Tilbeurgh, H. *et al.* (2007) An archaeal orthologue of the universal protein Kae1 is an iron metalloprotein which exhibits atypical DNA-binding properties and apurinic-endonuclease activity in vitro. *Nucleic acids research*, **35**, 6042-6051.
69. Zhong, Z.H., Jiang, W.Q., Cesare, A.J., Neumann, A.A., Wadhwa, R. and Reddel, R.R. (2007) Disruption of telomere maintenance by depletion of the MRE11/RAD50/NBS1 complex in cells that use alternative lengthening of telomeres. *J Biol Chem*, **282**, 29314-29322.
70. Jiang, W.Q., Zhong, Z.H., Henson, J.D., Neumann, A.A., Chang, A.C.M. and Reddel, R.R. (2005) Suppression of alternative lengthening of telomeres by Sp100-mediated sequestration of the MRE11/RAD50/NBS1 complex. *Molecular and Cellular Biology*, **25**, 2708-2721.

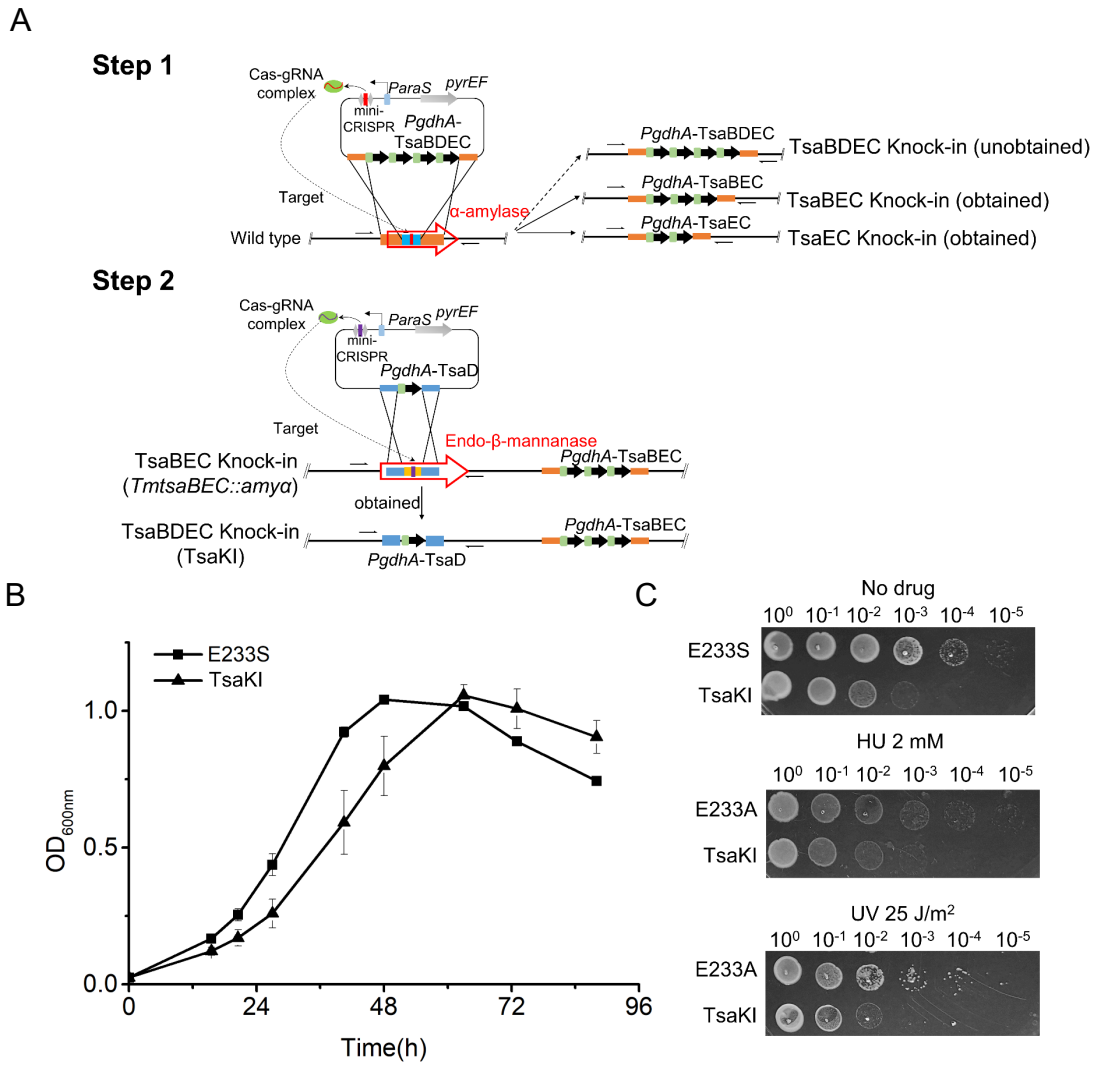

Figure 1. Construction of knock-in strain TsaKI harboring the *T. maritima* t^{6}A modification system. (**A**) Schematic for the construction of TsaKI. Step 1, plasmid was constructed to knock in all the four genes of the subunits of *Tm* t^{6}A modification system, *TsaB/D/E/C*, at the *α-amylase* locus of *S. islandicus* genome using the endogenous CRISPR-Cas-based method. Strains containing three (*TsaB/E/C*) and two (*TsaE/C*) genes of the subunits were obtained. Step 2, *TsaD* was knocked in at the locus coding for the β-mannanase using the obtained *Tm*TsaBEC knock-in strain, yielding TsaKI. Green rectangle, the promoter of *gdhA* (*PgdhA*). Black arrow, the genes of *TsaB/C/D/E*. (**B**) Growth curves of the wild type (E233S, filled square) and TsaKI (filled triangle). (**C**) Genotoxic analysis of the wild type (E233S) and TsaKI. The cultures were diluted in 10X gradience, dotted, and grown on the medium containing hydroxyurea (HU, 2 mM) or treated with ultraviolet (UV, 25 J/m^2) before spread onto the normal medium. The plates were incubated for 5 days before photos were taken.

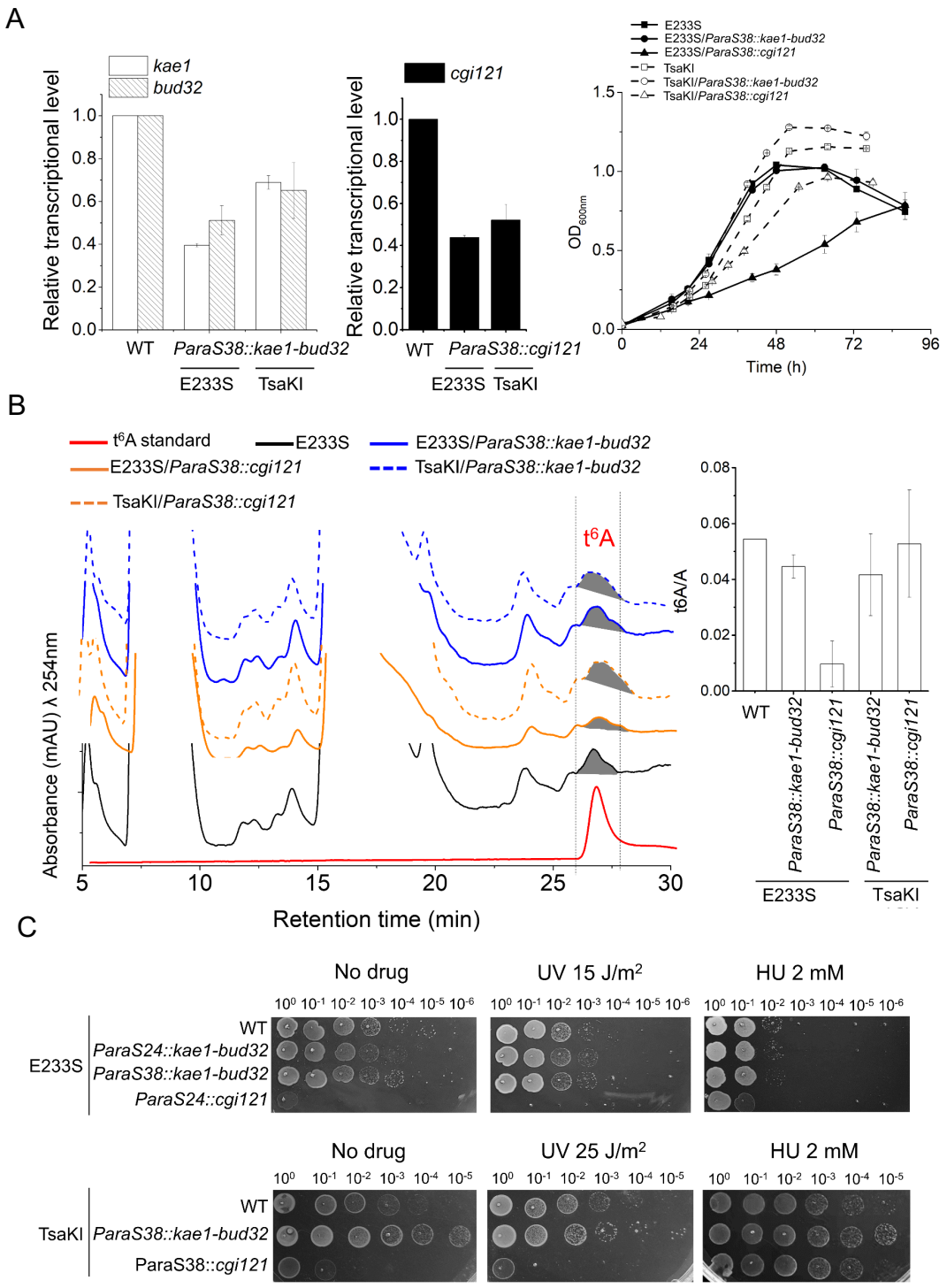

Figure 2. Knock-down analysis of genes encoding *Sis*KEOPS subunits in E233S and TsaKI. (**A**) RT-qPCR and growth analysis of *kae1-bud32* and *cgi121* knock-down strains with (TsaKI) and without the bacterial system (E233S). (**B**) HPLC analysis of t^6A levels in bulk tRNA from the wild type strain E233S (black), *kae1-bud32* knock-down in E233S (blue), *cgi121* knock-down in E233S (orange), *kae1-bud32* knock-down in TsaKI (dashed blue), and *cgi121* knock-down in TsaKI (dashed orange). The t^6A standard is shown in red. (**C**) Genotoxic sensitivity analysis of *kae1-bud32* and *cgi121* knock-down strains in E233S or TsaKI. The strains were gradiently diluted and dotted on the media containing hydroxyurea (HU, 2 mM) or treatment with ultraviolet (UV, 25 J/m^2). The plates without treatment or treated with UV were incubated for 5 days, whereas those treated with HU were incubated for 12 days.

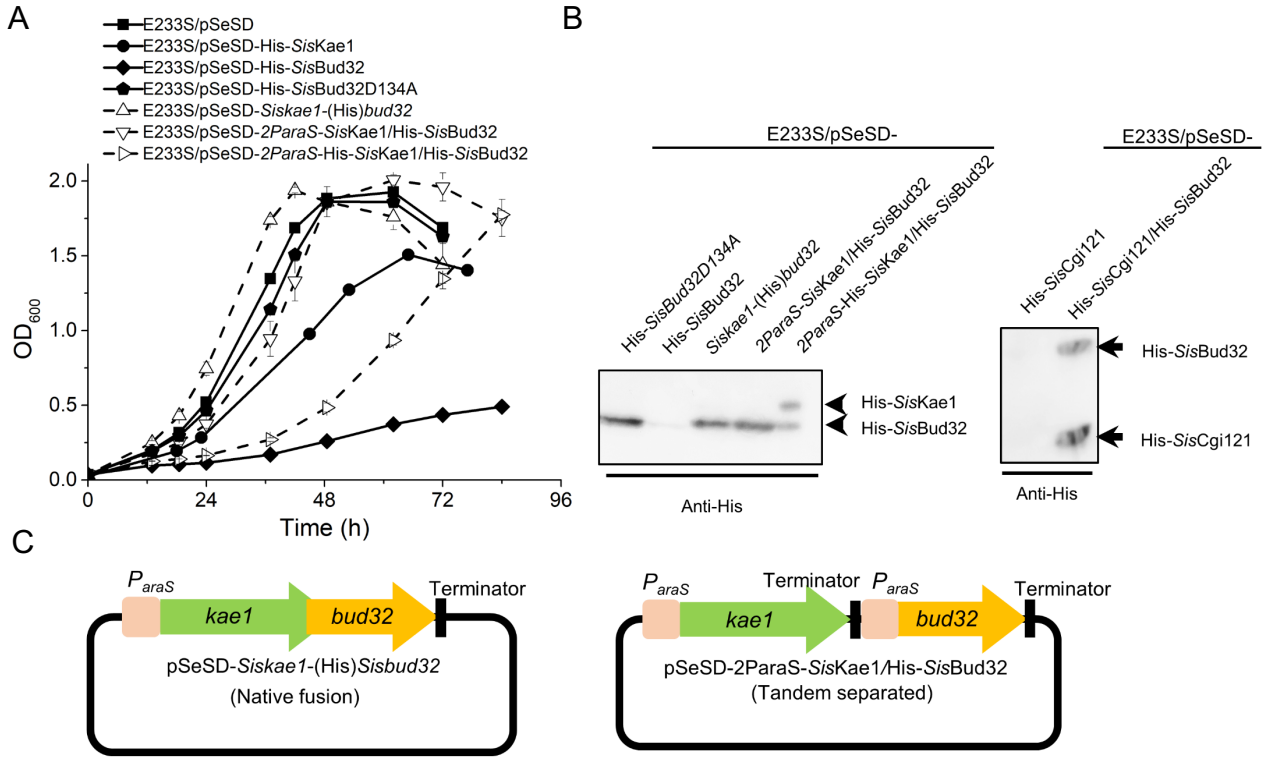

Figure 3. Effects of overexpression of KEOPS subunit on the cell growth. (**A**) Overexpression of Kae1 and/or Bud32. The overexpression strains were cultured in arabinose medium at an initial OD_{600}~0.05. The OD values were measured every 6 or 12 h. The growth curves were draw based on the averages of three independent cultures. Error bars indicate standard deviation. E233S/pSeSD was used as a control. (**B**) Western blotting analysis of the expression of Kae1, Bud32, and Cgi121. The cells were collected at the exponential phase and lysed by boiling for 10 min. The protein levels were detected by Western blotting using anti-His antibody. (**C**) Schematic of two kinds of two for *Sis*Kae1 and *Sis*Bud32 co-expression. *Kae1* and *bud32* have a 40 bp overlap in Native fusion plasmid as in the genome. In the Tandem separated plasmid, *kae1* and *bud32* expression are controlled by two independent araS promoters.

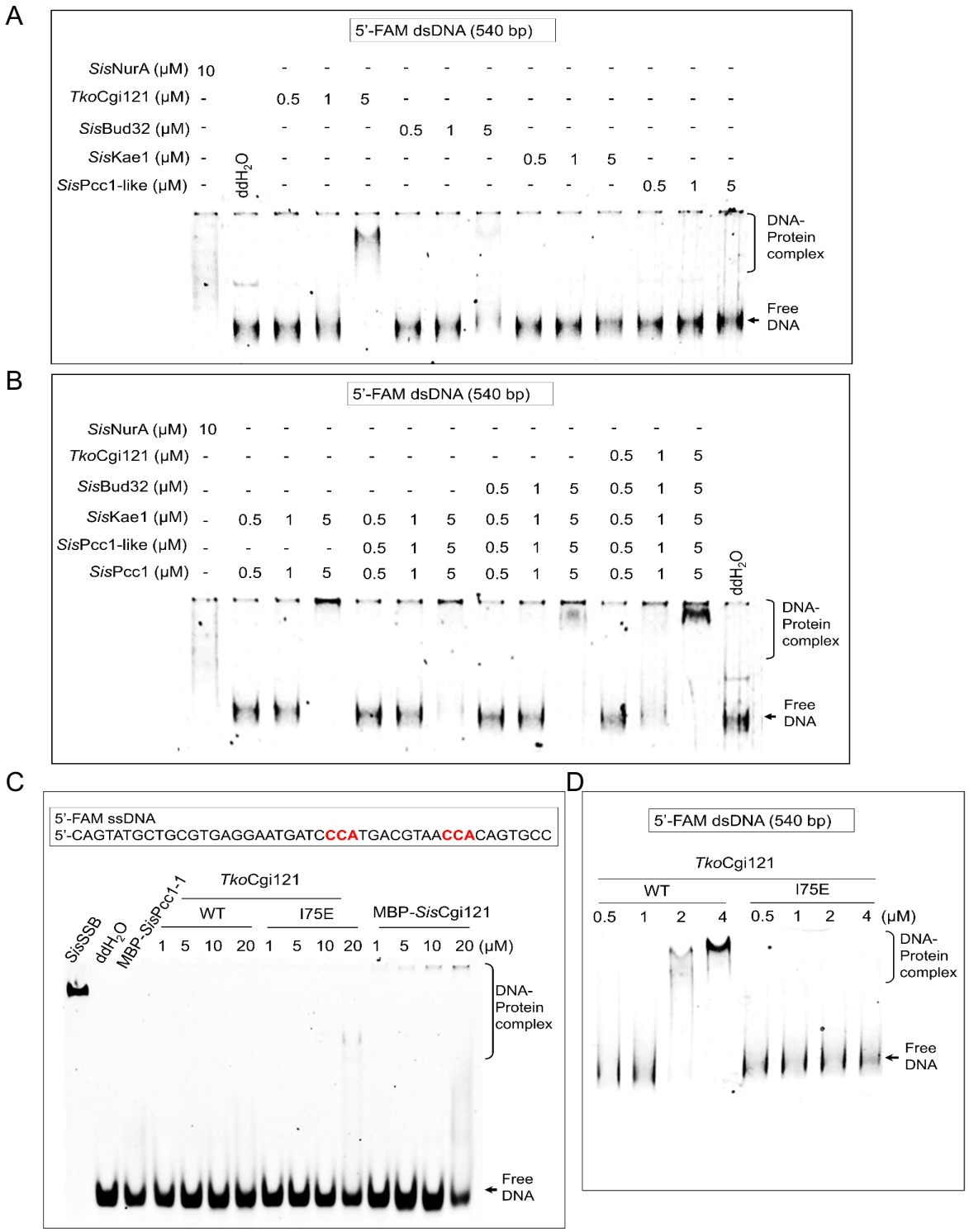

Figure 4. Analysis of the DNA binding ability of archaeal KEOPS. (**A**) Binding for dsDNA of individual subunit. (**B**) Binding for dsDNA of the combination of subunits FAM-labeled dsDNA (10 nM) was incubated with samples of the indicated subunits (0.5, 1, 5 μM) at 37℃ for 30 min. *Sis*NurA (10 μM) and ddH_2O were used as a positive and a negative control, respectively. (**C**) *Tko*Cgi121 (WT), *Tko*Cgi121(I75E), and *Sis*Cgi121 for ssDNA binding. Purified Cgi121 (1, 5, 10, 20 μM) was incubated in a reaction mixture containing 10 nM of FAM-labeled ssDNA at 37℃ for 45 min. *Sis*SSB (5 μM) was used as a positive control and MBP-tagged *Sis*Pcc1 (20 μM) was used as negative control. (**D**) dsDNA binding activity of *Tko*Cgi121 (WT) and the CCA-3' RNA binding site mutant (I75E). The purified proteins (0.5, 1, 2, 4 μM) were incubated in a reaction mixture containing 10 nM of FAM-labeled dsDNA (540 bp) at 37℃ for 30 min. The samples were loaded in native PAGE gels. The gels were scanned using Amersham ImageQuant 800 after electrophoresis. The detailed reaction conditions are described in the Materials and Methods.

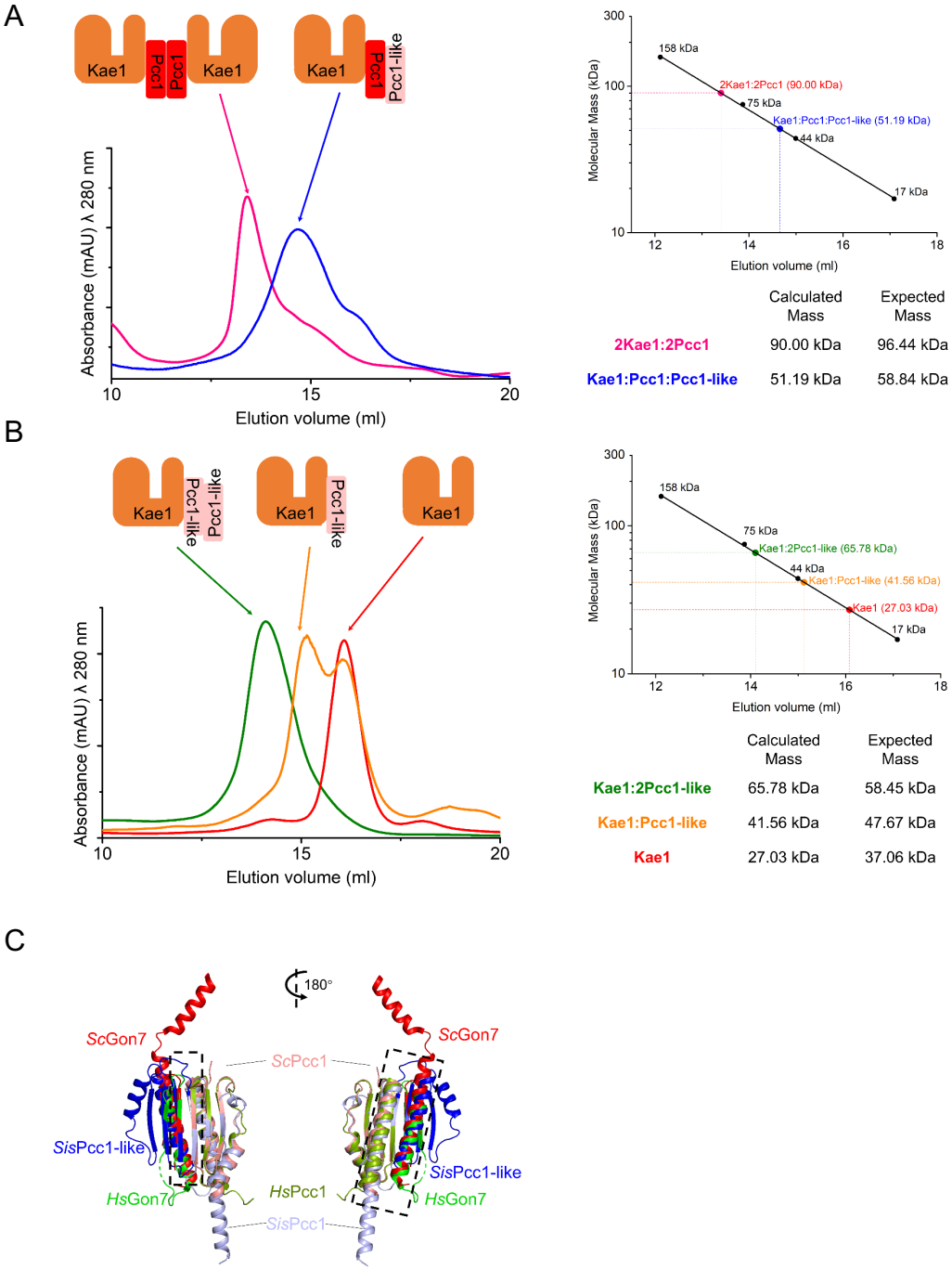

Figure 5. Subunit interaction and complex formation of Pcc1, Pcc1-like, and Kae1 revealed by pull down and size exclusion chromatography (SEC). (**A**) SEC profiles of Kae1/Pcc1 (pink) and Kae1/Pcc1+Pcc1-like (blue). (**B**) SEC profiles of Kae1/Pcc1-like (Pcc1-like excessed, green), Kae1/Pcc1-like (Kae1 excessed, orange), and Kae1 (red). The molecular masses of subcomplexes were calculated based on the protein standards (**A** and **B**, right). (**C**) Structural modeling of *S. islandicus* Pcc1:Pcc1-like and its superposition with *Saccharomyces cerevisiae* (Sc) Pcc1:Gon7 and *Homo sapiens* (*Hs)* Pcc1:Gon7. The modeling was performed using the online tool AlphaFold2_advanced (https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/beta/AlphaFold2_advanced.ipynb) based on the structures of the *Pyrococcus furiosus* homologs (61). Structure alignment was conducted using the cealign (CE algorithm) in PyMOL.

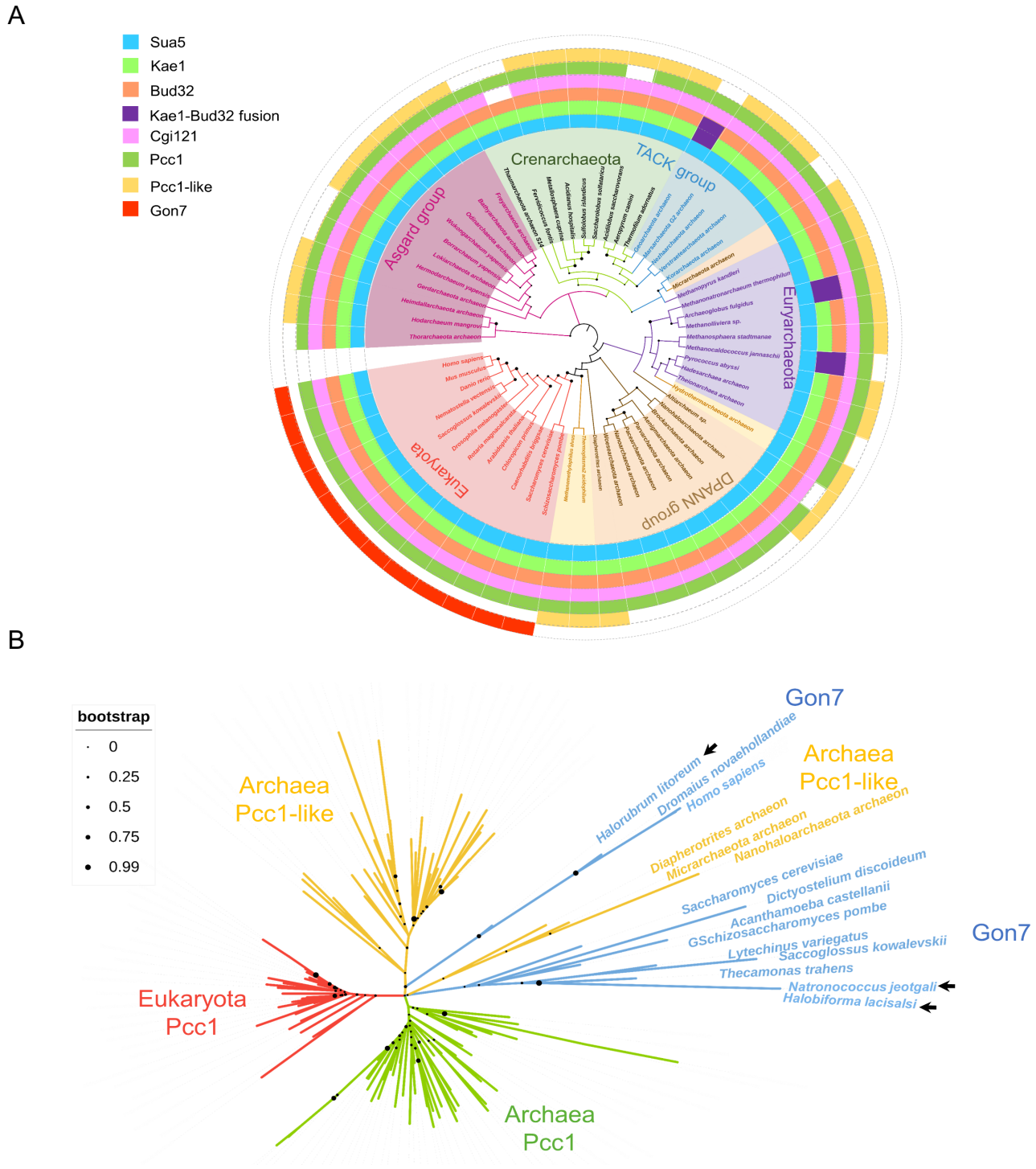

Figure 6. Distribution and phylogenetic analysis of Sua5 and KEOPS subunits in archaea and eukaryotes. (**A**) Distribution of Sua5 and KEOPS subunits. The phylogenetic tree was built based on the amino acid sequences of 59 Kae1 proteins from representative species of each archaea phylum and eukaryotes using the Neighbor-Joining method. The homologs were identified by protein-protein BLAST and protein-nucleotide BLAST in NCBI (https://blast.ncbi.nlm.nih.gov). If no homolog of KEOPS subunits was found by BLAST, genome context analysis was conducted using ribosome genes, which co-occurred with most KEOPS subunits, or XTP dephosphorylase, which co-co-occurred with the Asgard cgi121, as the references. (**B**) Phylogenetic analysis of Pcc1 and Gon7 homologs. The unrooted maximum likelihood tree of Pcc1, Pcc1-like, and Gon7 homologs was constructed based on 126 sequences from 82 species of archaea and eukaryotes. Filled circles at the nodes indicate bootstrap values of 1000 replicates in different sizes. The arrows indicate Gon7 sequences as a domain of the proteasome-activating nucleotidases (WP_015321675.1 from *Natronococcus occultus* SP4, WP_007139789.1 from *Halobiforma lacisalsi* AJ5, and ELY65446.1 from *Natronococcus jeotgali* DSM 18795), or MutS (WP_008364526.1 from *Halorubrum litoreum*) of halobacteria of euryarchaea. The sequences are available upon request.

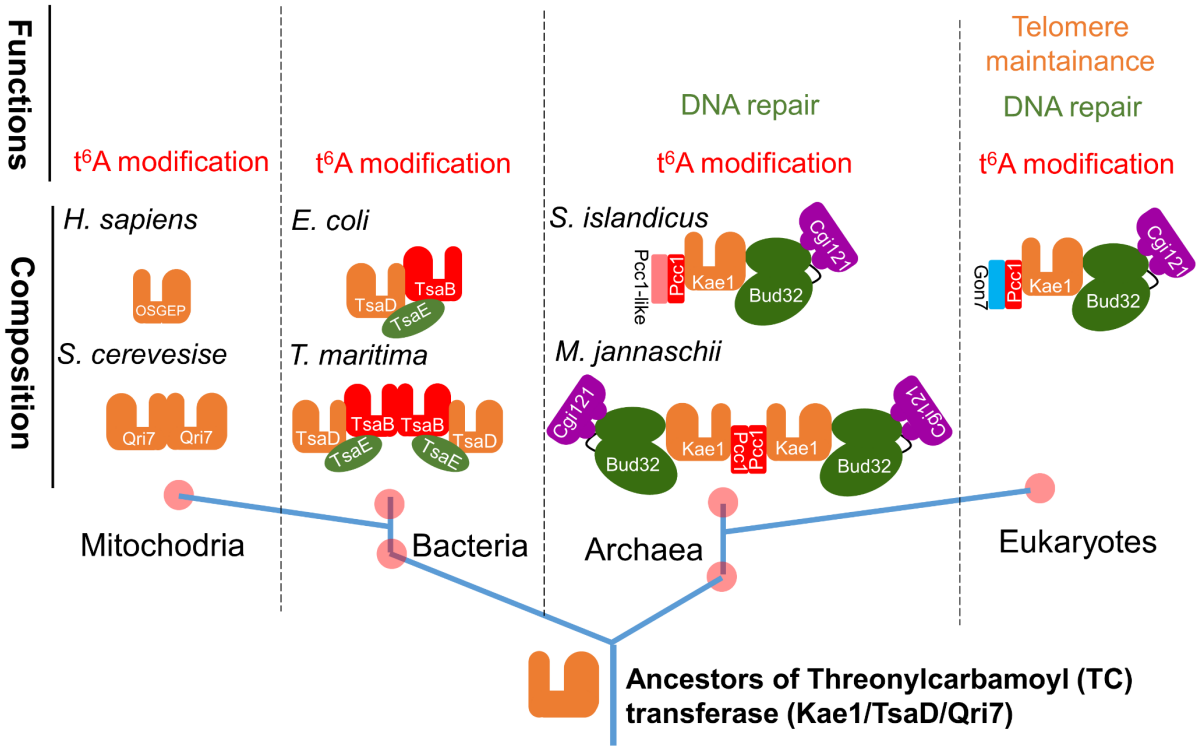

Figure 7. A proposed model for the evolution of TC-transferase and its complex. The ancestral TC-transferase could be Kae1/TsaD, from which mitochondrial TC-transferase and bacterial TsaD originated functioning for TC transfer. Later, the ATPase Bud32/TsaE and the dimerization module Pcc1/TsaB joined to form complexes, which are comprised of the minimum components and essential for t^6A modification. Meanwhile, Cgi121 joined into the initial complex of archaea-eukaryotes linkage to recruit tRNA. In the later stage, Pcc1-like might appear via a gene duplicate to stabilize the complex. Since Pcc1-like worked as a regulator in t^6A modification, it exhibited the highest mutation rate among the subunits and evolved to Gon7 in eukaryotes. As these more subunits integrated into the complex, KEOPS gained additional functions other than t^6A modification, such as DNA repair (HR repair) before the separation of archaea and eukaryotes. In eukaryotes, KEOPS also participates telomere maintenance which requires DNA recombination processes fulfilled by multiple HR repair protein.

Table 1. Sua5 and KEOPS subunits identified from the co-purified proteins of various KEOPS subunits expressed in *S. islandicus* by mass spectrometry.

Bait proteins	Components	Mass (Da)	Score	Sequences*	Coverage
His-Kae1					
	Sua5	39611	236	7 (7)	86%
	Kae1	36323	26851	35 (31)	33%
	Pcc1	10523	103	3 (3)	51%
	Pcc1-like	9138	124	2 (1)	28%
His-Pcc1					
	Kae1	36323	205	3 (3)	11%
	Pcc1	10523	198	3 (2)	35%
	Pcc1-like	9138	34	1 (1)	13%
Kae1/ His-Bud32					
	Sua5	39611	7543	25 (24)	72%
	Kae1	36323	1126	17 (15)	45%
	Bud32	26899	5229	13 (12)	46%
	Pcc1	10523	42	2 (1)	34%
	Pcc1-like	9138	15	2 (1)	22%

* The total numbers of sequences matched. The number of matches above the significance threshold is shown in brackets.

Supplementary Figures and Tables for

An evolutionally conserved archaeal KEOPS complex functions in DNA repair

Pengju Wu, Qi Gan, Xuemei Zhang, Yunfeng Yang, Yuanxi Xiao, Qunxin She, Jinfeng Ni, Qihong Huang#, Yulong Shen#

CRISPR and Archaea Biology Research Center, State Key Laboratory of Microbial Technology, Microbial Technology Institute, Shandong University, 266237, Qingdao, China

#Correspondence: yulgshen@sdu.edu.cn, huangqihong@sdu.edu.cn

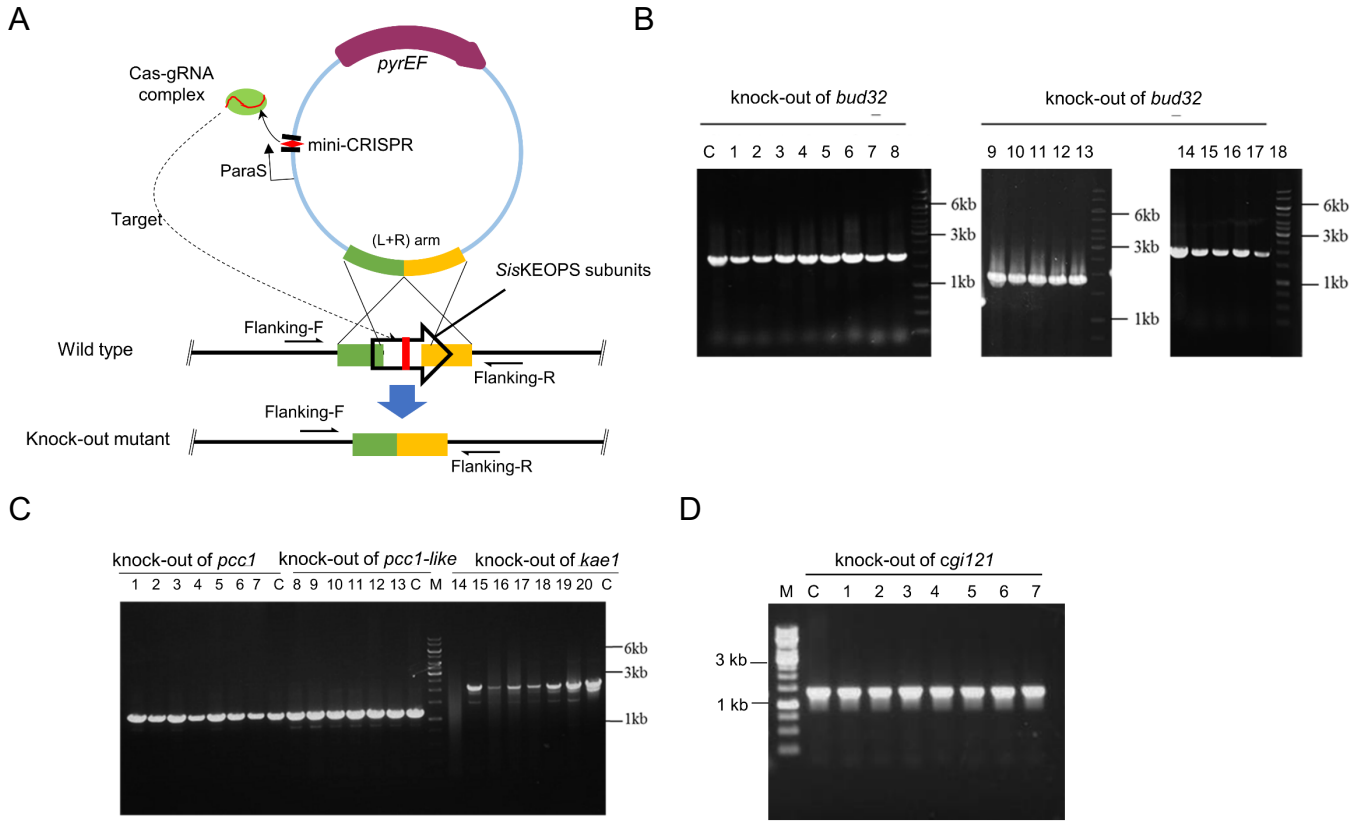

Figure S1. PCR analysis of colonies obtained in the *Sis*KEOPS gene knock-out experiments. (A) Schematic of the gene knock-out strategy. **(B)** PCR Screening for *bud32* knock-out colonies. **(C)** Screening for *pcc1*, *pcc1-like,* and *kae1* knock-out colonies **(D)** PCR screening for *cgi121* knock-out colonies. M, molecular size marker. C, control (E233S). The numbers indicate colonies picked from the plates after transformation.

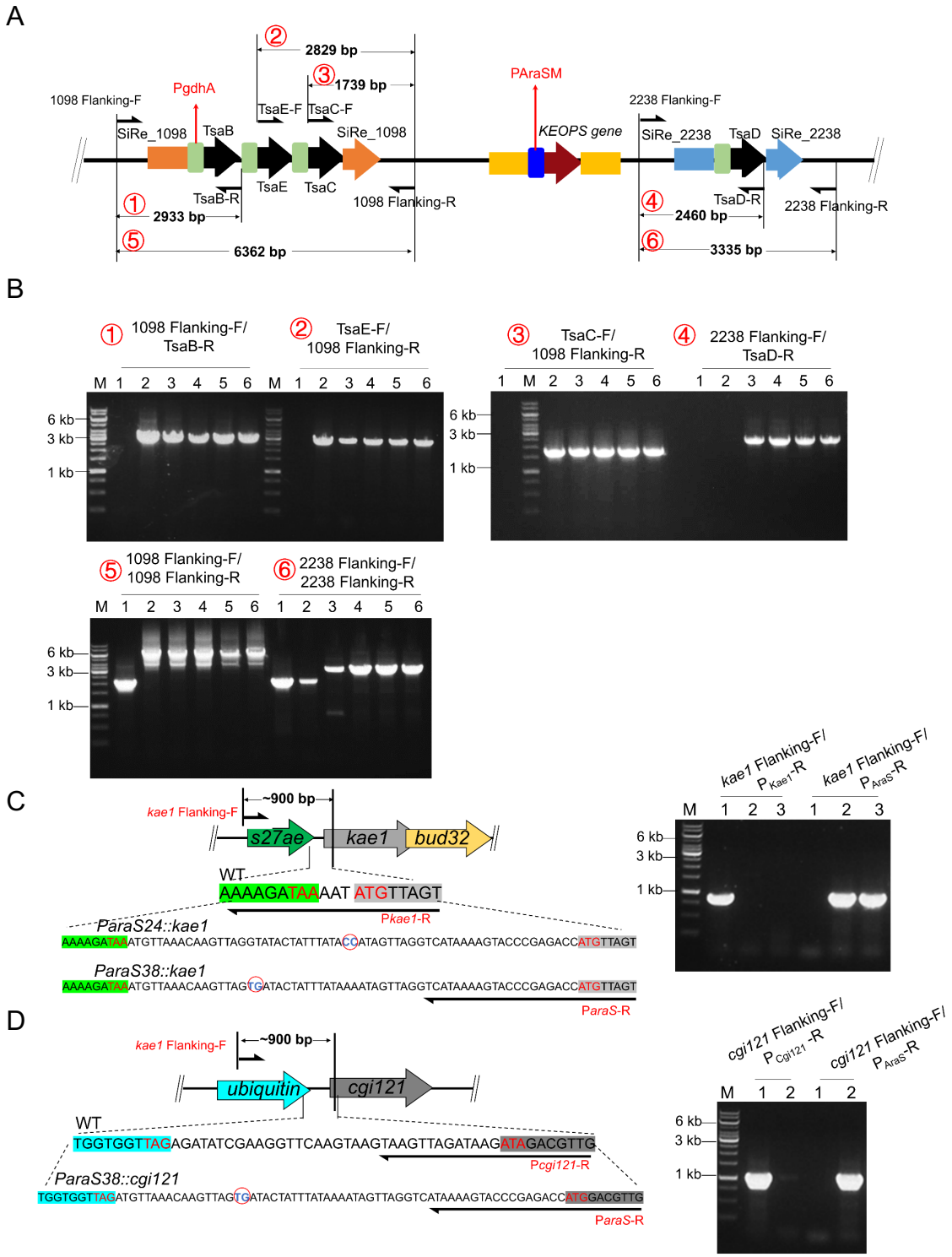

Figure S2. PCR verification of the insertion of the *TmTsaB/C/D/E* genes in E233S and the promoter replacement of the KEOPS genes in TsaKI. (**A**) Schematic for the primers and PCR products at the knock-in loci. The numbers in red indicated the PCR products of their corresponding numbers in (B). (**B**) Analysis of the PCR products at the *α-amylose* (*SiRe_1098*) and the *β-mannanase* (*SiRe_2238*) loci. Lanes 1, E233S; 2, E233S/*TmTsaBEC::amyα*; 3, TsaKI; 4, TsaKI/P_{araS}*24::kae1-bud32*.; 5, TsaKI/P_{araS}*38::kae1-bud32*. 6, TsaKI/P_{araS}*38::cgi121*. (**C**) Analysis of the PCR products at the loci of *kae1* promoters (right). The schematic for the promoter replacement of *kae1-bud32* is shown on the left. The start and stop codons of *kae1* (or *cgi121* in D) and its upstream gene are indicated in red. The two nucleotides in red circular are mutations of the wild type arabinose promoters. 1, TsaKI; 2, TsaKI/P_{araS}*24::kae1-bud32*.; 3, TsaKI/P_{araS}*38::kae1-bud32*. (**D**) Analysis of the PCR products at the loci of *kae1-bud32* promoters (right). The schematic for the promoter replacement is shown on the left. 1, TsaKI; 2, TsaKI/P_{araS}*38::cgi121*. M, marker.

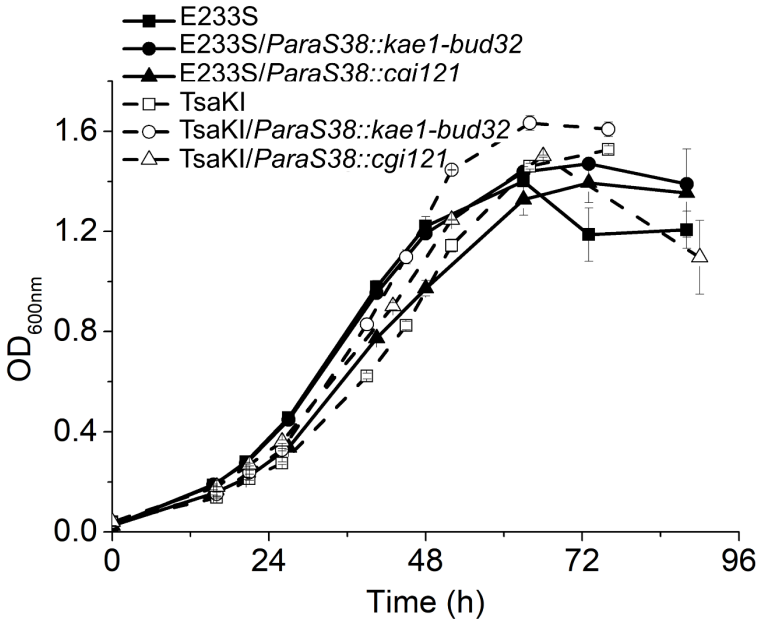

Figure S3. Growth of *kae1-bud32* and *cgi121* knock down strains in arabinose medium. *kae1-bud32* and *cgi121* knock-down strains containing the bacterial t^6A system (TsaKI) and without bacterial t^6A system (E233S) were cultured and transferred into arabinose medium for induction of *kae1-bud32* and *cgi121*, respectively. The OD values at 600 nm were measured every 6 or 12 h. The growth curves were draw based on the averages of three independent cultures. Error bars indicate standard deviation.

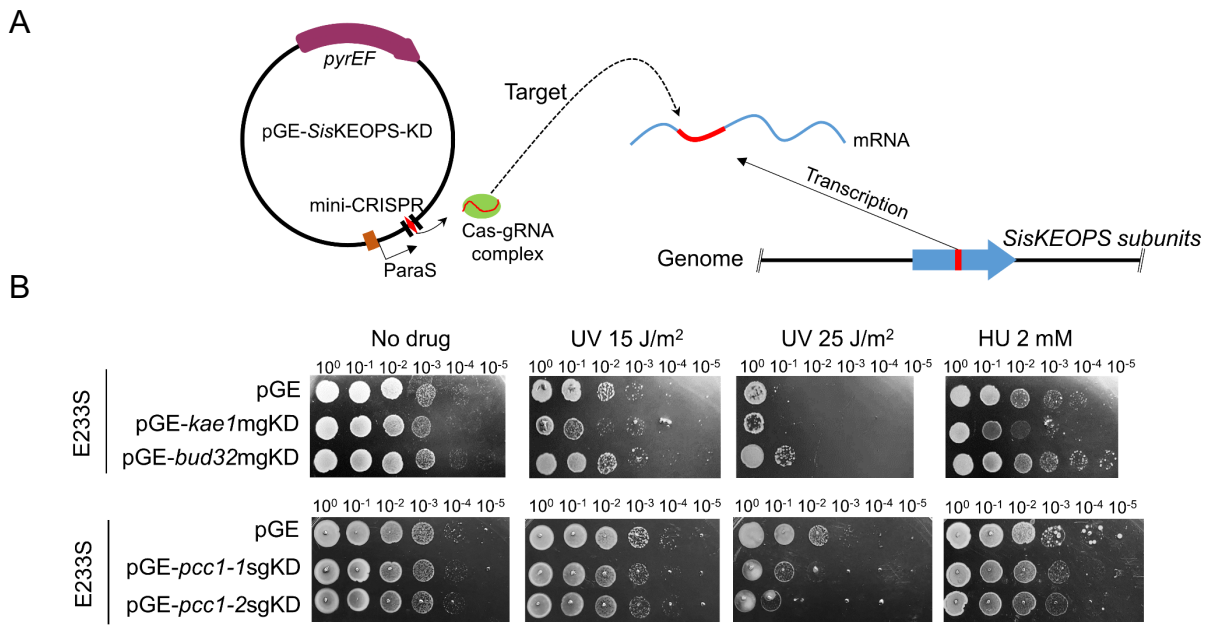

Figure S4. Genotoxic sensitivity analysis of gRNA-targeted knock down of genes encoding KEOPS subunits. (**A**) Schematic for the gRNA-targeted knock down strategy. (**B**) Genotoxic sensitivity analysis of the knock down strains. The cells were serially diluted and dotted on media containing hydroxyurea (HU, 2 mM) or treatment with ultraviolet (UV, 15 and 25 J/m^2). The plates were incubated for 5 days except that those treated with HU were incubated for 12 days.

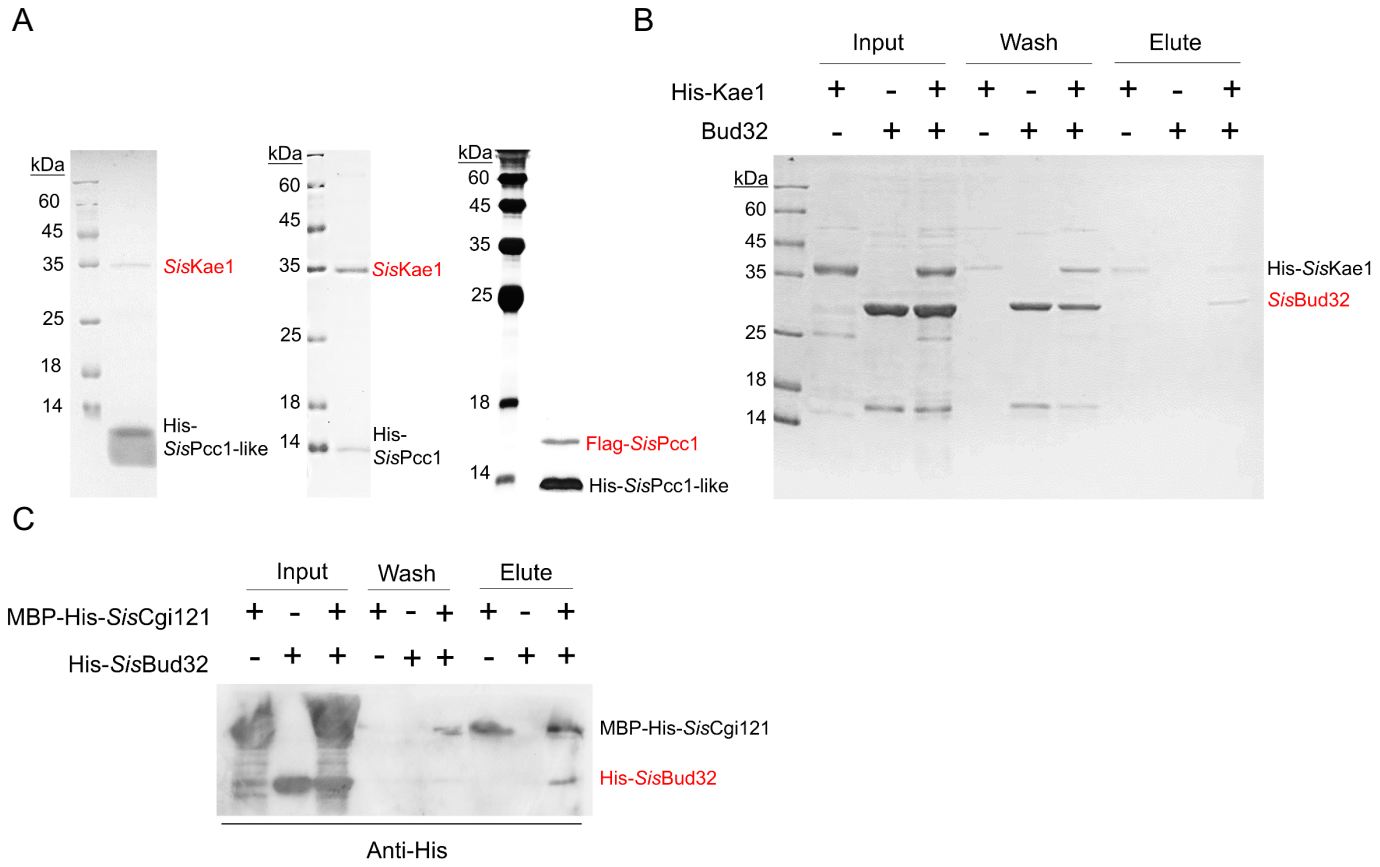

Figure S5. Analysis of the interactions between five aKEOPS complex subunits by pull down. (**A**) Pull-down assay using proteins expressed and purified from *E. coli*. The interaction between *Sis*Kae1 and *Sis*Pcc1-like was determined using individually purified proteins. The interactions between *Sis*Kae1 and *Sis*Pcc1, and *Sis*Pcc1 and *Sis*Pcc1-like were analyzed by identification of co-purified proteins from cells co-expressing the His-tagged proteins and the target proteins. The eluted samples were analyzed by SDS-PAGE. (**B**) His-*Sis*Kae1 was used as a bait to pull-down non-tagged *Sis*Bud32. The samples were analyzed by SDS-PAGE and Coomassie blue staining. (**C**) MBP-His-*Sis*Cgi121 was used as a bait to pull-down His-*Sis*Bud32 using amylose resin. The samples were analyzed by Western blotting. All the proteins were expressed in *E. coli* individually and purified by affinity chromatography.

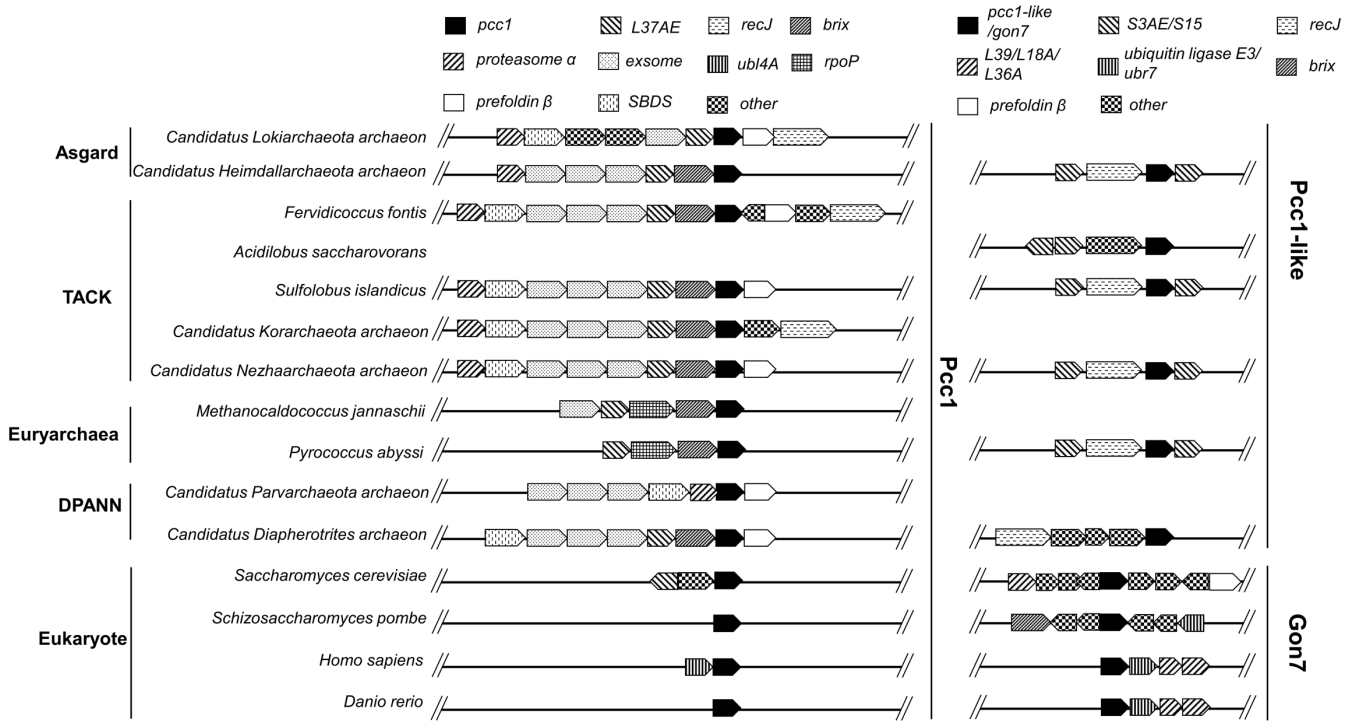

Figure S6. Genome context analysis of *pcc1*, *pcc1-like*, and *gon7* in archaea and eukaryotes. Shown are Local gene maps of representative species from each superphylum of archaea possessing one and two *pcc1* paralogs together with those of eukaryotic species. The gene *pcc1-like* possibly appeared via gene duplication of *pcc1* and was separated from *pcc-1* along with *recJ* during evolution.

Table S1. Strains used in this study

Strain	Properties	Source or reference
S. islandicus E233S	*S. islandicus* REY15A Δ*pyrEF* Δ*lacS*	Peng *et al.*, 2009, 2012
E233S/P_{araS}*24::kae1-bud32*	*kae1-bud32* promoter replaced with P_{araS}*24* in E233S	this study
E233S/P_{araS}*38::kae1-bud32*	*kae1-bud32* promoter replaced with P_{araS}*38* in E233S	this study
E233S/P_{araS}*24::cgi121*	*cgi121* promoter replaced with P_{araS}*24* in E233S	this study
E233S/*amyα::TmTsaB/TsaE/TsaC*	*TmTsaB/TsaE/TsaC* knocked in *amyα* (*SiRe_1098)* in E233S	this study
TsaKI	*TmTsaD* knocked in *mannanase* (*SiRe_2238*) in E233S/*amyα::TmTsaB/TsaE/TsaC*	this study
TsaKI/P_{araS}*38::kae1-bud32*	*kae1-bud32* promoter replaced with P_{araS}*38* in TsaKI	this study
E233S/P_{araS}*38::cgi121*	*cgi121* promoter replaced with P_{araS}*38* in TsaKI	this study

Table S2. Vectors used in this study.

Vectors	Properties or usages	Source or reference
pGE	*Sulfolobus-E. coli* shuttle vector containing mini-CRISPR and *pyrEF* for CRISPR-Cas based gene editing	Li *et al.*, 2016
pGE-*kae1*mg-KD	knock down of *kae1* with multi-gRNA	this study
pGE-*bud32*mg-KD	knock down of *bud32* with multi-gRNA	this study
pGE-*pcc1*sg-KD	knock down of *pcc1* with single-gRNA	this study
pGE-*pcc1-like*sg-KD	knock down of *pcc1-like* with single-gRNA	this study
pGE-*kae1*-KO	knock out of *kae1* with single-gRNA	this study
pGE-*bud32*-KO	knock out of *bud32* with single-gRNA	this study
pGE-*kae1-bud32*-KO	knock out of kae1-*bud32* with multi-gRNA	this study
pGE-*cgi121*-KO	knock out of *cgi121* with single-gRNA	this study
pGE-*pcc1*-KO	knock out of *pcc1* with single-gRNA	this study
pGE-*pcc1-like*-KO	knock out of *pcc1-like* with single-gRNA	this study
pGE-*cgi121*M52E	construction of *cgi121* mutant M52E with single-gRNA	this study
pGE-*cgi121*I64E	construction of mutant *cgi121* mutant I64E with single-gRNA	this study
pGE-*TmTsaBDEC*-KI	*tsaBDEC* knock in at *amyα* locus with single-gRNA	this study
pGE-*TmTsaD*-KI	*TmtsaD* knock in at *mannanase* locus with single-gRNA	this study
pSeSD	*Sulfolobus-E. coli* shuttle vector containing *araS*-SD promoter and MCS for proteins expression in E233S	Peng *et al.*, 2012
pSeSD-His-*Sis*Kae1	expression of C-terminal 6×His tagged *Sis*Kae1	this study
pSeSD-His-*Sis*Bud32	expression of C-terminal 6×His tagged *Sis*Bud32	this study
pSeSD-His-*Sis*Bud32D134A	expression of C-terminal 6×His tagged *Sis*Bud32D134A	this study
pSeSD-His-*Sis*Cgi121	expression of N-terminal 6×His tagged *Sis*Cgi121	this study
pSeSD-His-*Sis*Pcc1	expression of C-terminal 6×His tagged *Sis*Pcc1	this study
pSeSD-His-*Sis*Pcc1-like	expression of C-terminal 6×His tagged *Sis*Pcc1-like expression	this study
pSeSD-*Siskae1*-(His)*bud32*	expression of *Siskae1-bud32* operon with bud32 being tagged with C-terminal 6×His	this study
pSeSD-2ParaS	containing two *araS* promoter for protein expression	this study
pSeSD-2ParaS-*Sis*Kae1/His-*Sis*Bud32	Co-expression of *Sis*Kae1 and C-terminal 6×His tagged *Sis*Bud32	this study
pSeSD-2ParaS-His-*Sis*Kae1/His-*Sis*Bud32	Co-expression of C-terminal 6×His tagged *Sis*Kae1 and C-terminal 6×His tagged *Sis*Bud32	this study
pSeSD-2ParaS-His-*Sis*Cgi121/His-	N-terminal 6×His tagged *Sis*Cgi121 and C-terminal 6×His	this study

*Sis*Bud32	tagged *Sis*Bud32 coexpression vector	
pSeGP	*araS-SD* replaced with P*gdhA* in pSeSD	this study
pSeGP-His-*Tm*TsaB	constitutive expression of *Tm*TsaB	this study
pSeGP-His-*Tm*TsaC	constitutive expression of *Tm*TsaC	this study
pSeGP-His-*Tm*TsaD	constitutive expression of *Tm*TsaD	this study
pSeGP-His-*Tm*TsaE	constitutive expression of *Tm*TsaE	this study
pET22b-His-*Sis*Cgi121	expression of the N-terminal 6×His tagged *Sis*Cgi121	this study
pET22b-His-*Sis*Cgi121	expression of C-terminal 6×His tagged *Sis*Cgi121	this study
pET15b-*Sis*Kae1	expression of no tagged *Sis*Kae1	this study
pET15b-*Sis*Bud32	expression of no tagged *Sis*Bud32	this study
pET15b-His-*Sis*Bud32	expression of N-terminal 6×His tagged *Sis*Bud32	this study
pRSFDuet1-His-*Sis*Kae1	expression of N-terminal 6×His tagged *Sis*Kae1	this study
pRSFDuet1-Flag-*Sis*Bud32	N-terminal Flag tagged *Sis*Bud32 expression vector	this study
pRSFDuet1-His-*Sis*Pcc1	N-terminal 6×His tagged *Sis*Pcc1 expression vector	this study
pRSFDuet1-His-*Sis*Pcc1-like	N-terminal 6×His tagged *Sis*Pcc1-like expression vector	this study
pRSFDuet1-His-*Sso*Cgi121	expression of N-terminal 6×His tagged *S. solfataricus* Cgi121	this study
pRSFDuet1-His-*Tko*Cgi121	expression of N-terminal 6×His tagged *T. kodakarensis* Cgi121	this study
pRSFDuet1-His-*Tko*Cgi121I75E	expression of N-terminal 6×His tagged *Tko*Cgi121(I75E)	this study
pRSFDuet1-His-*Sis*Pcc1/*Sis*Kae1	co-expression of N-terminal 6×His tagged *Sis*Pcc1 and no tagged *Sis*Kae1	this study
pRSFDuet1-His-*Sis*Pcc1-like/Flag-*Sis*Kae1	co-expression of the N-terminal 6×His tagged *Sis*Pcc1-like and N-terminal Flag tagged *Sis*Kae1	this study
pRSFDuet1-Flag-*Sis*Pcc1/His-*Sis*Pcc1-like	co-expression of the N-terminal Flag tagged *Sis*Pcc1 and N-terminal 6×His tagged *Sis*Pcc1-like	this study
pMALc2X-MBP-His-*Sis*Cgi121	co-expression of the N-terminal MBP and C-terminal 6×His tagged *Sis*Cgi121	this study
pMALc2X-MBP-His-*Sis*Pcc1	co-expression of N-terminal MBP and C-terminal 6×His tagged *Sis*Pcc1	this study

Table S3. Oligonucleotide and promoter DNA sequences used in this study.

Oligonucleotides	Sequences (5'to3')
kae1-KO-Spacer	GGCGGAAATACAATCATAACTACCTTCTATAAAGGGAGGT
bud32-KO-Spacer	GCGGTGATCTAACAACTAACAATCTCATCCTAAGTTCTATA
kae1-bud32-KO-Spacer1	ACTATGAAAGGACAATATTAGAGGCTAAGATAATTTATAC
kae1-bud32-KO-Spacer2	GGCGGAAATACAATCATAACTACCTTCTATAAAGGGAGGT
kae1-bud32-KO-Spacer3	AGGAATATCTACCTCGTCTACTCTCCATCTAGGTCTTAT
cgi121-KO-Spacer	ACTATAATACAACTAGAAATAAGATAAAGAGTTCAACTAT
pcc1-KO-Spacer	ATAAAAACGAATTACAGGATATAATATATGATTCGATAAT
pcc1-like-KO-Spacer	ATTCAAACCAAGGTTGAAGGCAAAGAGTTAGAGATTGTAA
kae1-KD-Spacer1	TCAGAATGTATACTTTCTAAAGATCTTAGAAATATATGAA
kae1-KD-Spacer2	ATATCGTTATTTATAGAACTTAGGATGAGATTGTTAGTTG
kae1-KD-Spacer3	CACTGCAGGTACATTCACATCGTTTTTAAGCGCAGTATAA
bud32-KD-Spacer1	GGTCTTTGGCCTCAGTTGTTAAATACCCTATTTCAATATG
bud32-KD-Spacer2	CTACTGCTATATAATTAATATCATTTATACTAATATTAGC
bud32-KD-Spacer3	TATCCCTTTCGTTTGCCAATATGTAGGGTGGTTGATCTTT
pcc1-KD-Spacer	GGGATTTCTTAATTTTTACATATTTAGTATCTATCTTTTC
pcc1-like-KD-Spacer	TCGTTTGGCTTTGTAATTCTCATCTTTATTACAATCTCTA
kae1-bud32-PE-Spacer	AAGAAAAGATAAAATATGTTAGTACTGGGTATCGAATCTA
cgi121-PE-Spacer	TATCTAACTTACTTACTTGAACCTTCGATATCTCTAACCA
*cgi121*M52E/I64E-Spacer	TACTTCTCCTATTACCATATGAACAAATAAAGGATGCATT
1098-KI-Spacer	TTATTCTCAAGCTCTATTGGTAGCCATTTGAGAAATTCTA
2238-KI-Spacer	ATGCACCATCAGCACCGTAAACTAACGCTCCCCTCAGTAT
P_{araS}*24*	ATGTTAAACAAGTTAGGTATACTATTTATACCATAGTTAGGTCATAAAAGTACCCGAGACC
P_{araS}*38*	ATGTTAAACAAGTTAGTGATACTATTTATAAAATAGTTAGGTCATAAAAGTACCCGAGACC
P_{gdhA}(with 30 bp *gdhA* leader)	(-556)AACACTAATGAGAAAGT……CCGGGAAAACGAATTTATATTG**ATG**GAAGAAGTTCT TAGTTCGAGTCTGCAT (+30) *
5'-FAM labled ssDNA	CAGTATGCTGCGTGAGGAATGATCCCATGACGTAACCACAGTGCC
kae1-qPCR-F	CTCGGAGTAGGAATAGCAAAAGATCA
kae1-qPCR-R	CGTTTTAGTAGATCTCCAGGCTTCAT
bud32-qPCR-F	GAGGCTAAGATAATTTATACTGCGCTT

bud32-qPCR-R	CCTTAACTATCTCCCCTTCTATATATTC
cgi121-qPCR-F	GTCAAGTCTTACCTATCGTACCGTTT
cgi121-qPCR-R	CGGTAATAGGAGAAGTAAGAAAAACATAG
pcc1-qPCR-F	GTGAAGATAGAGATTTCAATTTATCCAGATAAT
pcc1-qPCR-R	CCTAGCTCTTGTAATTGATGGTGCA
pcc1-like-qPCR-F	TTACCAGAAGGAATGTCAATTCAAACCA
pcc1-like-qPCR-R	TTATAGCATCAAATGACGATTGAAGTGC
1715-qPCR-F	TAACGGTGTCAGAGATAACCTACAC
1715-qPCR-R	TGTTATGTCCTTTACTCTTATACCTACC
tbp-qPCR-F	CAACAGTTACGTTAGAGCAAAGTTTGG
tbp-qPCR-R	GTAACCTTGGGCTGTTCTAATCTGA

* The numbers in the brackets indicate the position of the last nucleotide at the upstream (-) or downstream (+) of the start codon (in red).

References

Peng N., Xia, Q., Chen, Z., Liang, Y.X., and She, Q. (2009). An upstream activation element exerting differential transcriptional activation on an archaeal promoter. Mol Microbiol 74, 928-939.

Peng, N., Deng, L., Mei, Y., Jiang, D., Hu, Y., Awayez, M., Liang, Y., She, Q. (2012). A synthetic arabinose-inducible promoter confers high levels of recombinant protein expression in hyperthermophilic archaeon Sulfolobus islandicus. Appl Environ Microbiol, 78(16), 5630-5637. doi:10.1128/AEM.00855-12

Li, Y., Pan, S., Zhang, Y., Ren, M., Feng, M., Peng, N., Chen, L., Liang, Y.X., and She, Q. (2016). Harnessing Type I and Type III CRISPR-Cas systems for genome editing. Nucleic Acids Res 44, e34.