A mutant chaperonin that is functional at lower temperatures enables hyperthermophilic archaea to grow under cold-stress conditions.

Running title: Mutation that improves cold adaptation of hyperthermophile

Le Gao, a Tadayuki Imanaka, c Shinsuke Fujiwara a, b

Research Center for Intelligent Bio-Materials, Graduate School of Science and Technology, Kwansei-Gakuin University, Sanda, Hyogo, Japan a, Department of Bioscience, Graduate School of Science and Technology, Kwansei-Gakuin University, Sanda, Hyogo, Japan b and College of Life Science and Biotechnology, Ritsumeikan University c
Thermococcus kodakarensis grows optimally at 85°C and possesses two chaperonins, cold-inducible CpkA and heat-inducible CpkB, which are involved in adaptation to low and high temperatures, respectively. The two chaperonins share high sequence identity (77%), except in their C-terminal regions. CpkA, which contains tandem repeats of a “GGM” motif, shows highest ATPase activity at 60°C-70°C, whereas in CpkB shows highest activity at temperatures higher than 90°C. To clarify the effects of changes in ATPase activity on chaperonin function at lower temperatures, various CpkA variants were constructed by introducing single point mutations into the C-terminal region. The CpkA variant, in which Glu530 was replaced with Gly (CpkA-E530G), showed increased ATPase activity, with greatest activity at 50°C. The efficacy of the CpkA variants against denatured indole-3-glycerol-phosphate synthase (TrpC\textsubscript{Tk}), which is a CpkA target, was then examined \textit{in vitro}. CpkA-E530G was more effective at facilitating the refolding of chemically unfolded TrpC\textsubscript{Tk} at 50°C than wild-type CpkA. The effect of \textit{cpkA-E530G} on cell growth was then examined by introducing \textit{cpkA-E530G} into the genome of \textit{T. kodakarensis} KU216 (\textit{pyrF}). The mutant strain, DA4 (\textit{pyrF, cpkA-E530G}), grew as well at 60°C as the parental KU216 strain. By contrast, DA4 grew more vigorously than KU216 at 50°C. These results suggested that the CpkA-E530G mutation prevented cold denaturation of proteins under cold-stress conditions, thereby enabling cells to grow in cooler environments. Thus, a single base pair substitution in a chaperonin...
IMPORTANCE

Thermococcus kodakarensis possesses two group II chaperonins, cold-inducible CpkA and heat-inducible CpkB, which are involved in adaptation to low and high temperatures, respectively. CpkA might act as an “adaptive allele” to adapt to cooler environments. Here, we compared the last 20 amino acids within the C-terminus of chaperonins and found a clear correlation between the CpkA-type chaperonin gene copy number and growth temperature. Furthermore, we introduced single mutations into the CpkA C-terminal region to clarify its role for cold adaptation, showing that a single base substitution allowed the organism to adapt to a lower temperature. The present data suggest that hyperthermophiles have evolved by obtaining mutation in chaperonin so that they adapt to a colder environment.

INTRODUCTION

Chaperonin, also known as heat shock protein 60 (HSP60), belongs to an evolutionarily conserved protein family that enables host cells to survive under stressful conditions, including restricted temperature, high/low salinity, and high/low hyperosmotic pressure. Chaperonins are classified into two subfamilies, group I and group II, on the basis of sequence homology and
structural differences (14). The thermosome is a group II chaperonin found in hyperthermophilic archaea, in which it plays a key role in thermal adaptation. Unlike bacterial chaperonins, which function as homotetradecamers, archaeal chaperonins form a heterohexadecamer comprising two back-to-back rings; each of these rings comprises eight heterogeneous chaperonin subunits. Similar to bacterial chaperonins, archaeal chaperonin monomers comprise three domains: the apical domain (which binds unfolded proteins), the intermediate domain (which acts as a hinge, thereby allowing the movement of the apical domain as well as the transition between trans and cis conformations needed for chaperonin function), and the equatorial domain (which is responsible for the ATPase and refolding activities in the central cavity of the ringed complex). However, thermosomes do not have co-chaperonins, such as GroES. Rather, there is a conserved protrusion from the apical domain, which may play a role in guiding unfolded proteins. Phylogenetic analyses reveal that bacterial chaperonin genes have undergone many rounds of duplication to generate these diverse structures (20). The hyperthermophilic archaean, *Thermococcus kodakarensis*, possesses two gene copies of a chaperonin, *cpkA* and *cpkB* (16, 24). CpkB is a classic thermosome, which is an essential heat-inducible factor that enables the cells to survive at high temperatures (> 85°C). A relative of *T. kodakarensis*, called *Pyrococcus sp*, has only one chaperonin, which is an orthologue of CpkB; therefore, this organism has a narrow, but high, growth temperature range (80 - 100°C). However, CpkA is an atypical archaeal group II chaperonin, which is cold-inducible at the stationary phase of
cell growth and it is not found in *Pyrococcus* sp. A cpkA disruptant strain showed a poor cell growth at 60°C, but no significant growth defect at 85°C and 93°C (15). Indole-3-glycerol-phosphate synthase (TrpC_{Tk}) was identified as one of specific target proteins for CpkA in *T. kodakarensis*. The refolding of partially cold denatured TrpC_{Tk} was accelerated by the addition of CpkA but not by adding CpkB (17).

CpkA and CpkB are highly homologous (77%), except for the C-terminal regions. This region of CpkA harbors tandem repeats of a “GGM” motif. Such coding tandem repeats are usually involved in protein-protein interactions (19). For example, in *Escherichia coli*, the C-terminal region of GroEL (which is also rich in “GGM” motifs) is involved in ATP hydrolysis and protein binding (11, 28, 31, 48, 49). Langer et al. found that removing the 50 C-terminal amino acid residues of GroEL resulted in the loss of ATPase activity (28). McLennan et al. found that although the 16 C-terminal residues of GroEL were dispensable for *E. coli* cell growth, the truncated protein could not complement the groEL_16 mutation in *E. coli* (34). Machida et al. reported that the C-terminal 23 residues of GroEL (which are hydrophilic) were not required for chaperonin function (31). However, when the last six residues (526-KNDAAD-531) were substituted with a neutral (neither hydrophobic nor hydrophilic) sequence (526-GGGAAG-531), or with a hydrophobic sequence (526-IGIAAI-531), chaperonin function was defective both *in vitro* and *in vivo*. Suzuki et al. used a special microscopic method called Zero-mode waveguides to show that the integrity of the C-terminal of the chaperonin
region facilitates the transition from the first to the second rate-limiting step (45). Studies of the C-terminal region of group II chaperonins have just begun. Luo et al. mutated the C-terminal 15–25 residues of the Pyrococcus furiosus chaperonin, which is highly homologous to CpkB, in an attempt to modify its thermostability and activity (30). An EK-rich motif (528-EKEKEKEGEK-537) was the key domain responsible for increased stability at around 100°C. Zhang et al. found that flexible interwoven termini determining the thermal stability of thermosomes, especially in beta subunit which has more charged amino acids than alpha subunit (51). However, the function of the CpkA C-terminal region remains unknown. Here, we found a clear correlation between the CpkA-type chaperonin gene copy number and growth temperature. We introduced single mutations into the CpkA C-terminal region to increase the number of “GGM” motifs to clarify its role during the evolution of T. kodakarensis.

MATERIALS AND METHODS

Microorganisms, media, and growth conditions. The microorganisms (strains) used in this study are listed in Table 1. T. kodakarensis strains KU216 (∆pyrF) and DA4 (∆pyrF, cpkA-E530G) were pre-cultivated anaerobically overnight at 85°C in 100 ml of MA-YT medium containing the following (per liter): 30.4 g of Marine Art SF reagent, 5 g of yeast extract (Nacalai Tesque, Kyoto, Japan), 5 g of tryptone (Nacalai Tesque), and 5 g of pyruvate (Wako Pure Chemical Industries, Ltd.,
Osaka, Japan). The culture was added to 20 ml of MA-2YTP medium (30.4 g/L of Marine Art SF reagent, 10 g/L of yeast extract, 10 g/L of tryptone, and 10 g/L of pyruvate.) and incubated at a given temperature. All growth studies were carried out three times. To make solid medium, 1% Gelrite (Wako, Osaka, Japan) was mixed with 2 ml/l polysulfide solution (10 g of Na2S 9H2O and 3 g of sulfur flowers in 15 ml of H2O). E. coli DH5α strains were routinely cultivated at 37°C in Lysogeny broth (LB) medium. Ampicillin (50 µg/ml) was added to the medium to select transformants.

Alignment of protein sequences and modeling of protein structure. All chaperonin (amino acid) sequences used for sequence alignment were identified and retrieved from the public National Center for Biotechnology Information (NCBI) database (1). The C-terminal 20 amino acids of each chaperonin were then aligned using ClustalW2, supported by the European Bioinformatics Institute. A three-dimensional (3D) structural octamer of full-length CpkA was built in MODELLER 9.12 (10) using the chaperonin alpha subunit of Thermococcus strain KS-1 (PDB ID: 1Q2V) as a template. The protein sequence of CpkA was retrieved from the NCBI sequence database. The construction generated three models, and the one with the lowest discrete optimized protein energy score was selected.

DNA manipulation. DNA was isolated and handled according to the standard techniques described by Sambrook and Russell (40). A small-scale preparation of plasmid DNA was prepared
from *E. coli* cells using the GenElute™ Plasmid Miniprep Kit (Sigma-Aldrich, MO). DNA sequencing was performed using the BigDye-Terminator Cycle-Sequencing Ready Reaction Kit, version 3.1, and a model 3130 capillary DNA sequencer (Applied Biosystems, CA). The plasmids and primers used are listed in Tables 1 and 2, respectively.

**Construction of expression plasmids encoding the cpkA variants.** pCPAE (25) was used as a template. Briefly, expression plasmids for the *cpkA* variants (E530G, E530M, Q533G, Q533M, P538G, P538M, D545G, and D545M) were generated by site-directed mutagenesis. To construct pCPAE-E530G, PCR was performed using primers cpkA-E530G-Fw and cpkA-E530G-Rv and the amplified DNA was ligated into pCPAE. *E. coli* DH5α cells were transformed with the ligated DNA and plasmid DNA was extracted from positive transformants. Plasmid construction was confirmed by nucleotide sequencing with the primer cpkA-singlechecking (Table 2). Mutant plasmids pCPAE-E530M, -Q533G, -Q533M, -P538G, -P538M, -D545G, and -D545M were constructed using primers cpkA-E530M-Fw and cpkA-E530M-Rv, cpkA-Q533G-Fw and cpkA-Q533G-Rv, cpkA-Q533M-Fw and cpkA-Q533M-Rv, cpkA-P538G-Fw and cpkA-P538G-Rv, cpkA-P538M-Fw and cpkA-P538M-Rv, cpkA-D545G-Fw and cpkA-D545G-Rv, and cpkA-D545M-Fw and cpkA-D545M-Rv, respectively. The plasmids were then transformed into *E. coli* BL21 (DE3) RIL cells.

**Protein expression and purification.** Recombinant CpkA and its variants were expressed in *E.
coli BL21(DE3)RIL at 25 °C as previously described (25). CpkA and its variants remained in the soluble fraction after heat treatment at 85°C for 30 min. Further purification was carried out by anion-exchange chromatography (MonoQ HR 5/5; GE Healthcare, Tokyo, Japan), followed by gel filtration on a Superdex 200 column (GE Healthcare). Recombinant TrpC<sub>Tk</sub> (indole-3-glycerol-phosphate synthase; InGPS) was expressed and purified as previously described (47).

**Thermostability of the CpkA and its variants.** The secondary structures of chaperones were measured using a far-UV circular dichroism (CD) spectroscopy method in wavelength scan mode (220 nm to 260 nm) at 20°C using a J-820 spectrometer (JASCO, Japan) as described previously (17). The thermostability of CpkA and its variants was measured by far-UV CD spectroscopy using a J-820 spectrometer (JASCO, Japan) set in temperature scan mode. Chaperones were dissolved in 25 mM Tris-HCl (pH 7.8) buffer containing 300 mM KCl (final concentration, 50 μg ml<sup>-1</sup>) and the change in absorption at 222 nm was monitored from 30 to 100°C. The temperature scan rate was 1°C min<sup>-1</sup>.

**Enzyme assay to examine ATPase activity.** ATPase activity was measured by monitoring the amount of phosphate released from ATP by chaperones. The temperature dependency of the ATPase activity of CpkA and its variants was examined by incubating 80 nM of each chaperonin in HKM buffer (25 mM, HEPES-NaOH, 100 mM KCl, 5 mM MgCl<sub>2</sub> pH 7.5) at different temperatures (20, 9
40, 60, 80, and 93°C). The reaction mixtures were preheated to the test temperature for 5 min prior to the addition of 2 mM ATP to trigger the reaction. After 10 minutes, the reaction was terminated by cooling on ice and the amount of released phosphate was measured using a Biomol Green Kit (Enzo life Science, NY) (21, 32). One unit of ATPase activity was defined as the amount of enzyme that liberated 1 nmol of inorganic phosphate from ATP in 1 min.

**Examination of chaperonin activity.** Purified TrpC\(_{Tk}\) was chemically unfolded and used as a substrate for chaperonins in a refolding assay as previously described (17). TrpC\(_{Tk}\) (16 μM) was denatured on ice for 3 days in 50 mM Tris-HCl buffer (pH 8.0) containing 7 M urea. Then, 2 μl of unfolded TrpC\(_{Tk}\) (8 μM) was added to 198 μl of 100 mM HKM buffer (100-fold dilution) containing 2 mM ATP in the presence or absence of wild-type or variant CpkA. The solutions were incubated for 30 min at 50 or 60°C and the refolding rates determined by monitoring the amount of InGPS activity at 50 or 60°C as previously described (17). One unit of InGPS activity was defined as the amount of activity that resulted in the formation of 1 μmol of InGPS in 1 min.

**Plasmids used to construct mutant strain DA4.** The gene disruption strategy has been described previously (41). Briefly, plasmid pUD2-cpkA-E530G was first constructed by amplifying PCR fragments from the pUD2-cpkAfr plasmid using primer pair cpkA-E530G-Fw and cpkA-E530G-Rv. The ligated PCR products were then used to transform *E. coli* DH5α cells. Plasmid DNA was extracted from transformants and the nucleotide sequence of the *cpkA* region was
confirmed by sequencing using primer cpkA-singlechecking (Table 2). The resulting plasmid was
named pUD2-cpkA-E530G and used to construct mutant strain DA4.

Transformation of *T. kodakarensis*. To construct mutant strain DA4 (*pyrF*, *cpkA-E530G*),
KU216 was used as the parental strain and plasmid pUD2-cpkA-E530G was introduced into strain
KU216. Transformation was performed as described in our previous work (17). Genotype of positive
candidates were confirmed by sequencing. The correctly constructed mutant was named strain DA4.

Growth profile of DA4 and KU216. The growth of strains DA4 and KU216 was monitored at
OD$_{660}$ in an automatic Bioshaker (BR-43FH; TAITEC, Japan). Cells were first pre-cultured
overnight at 85°C in MA-YT medium followed by culture in MA-2YTP medium with 1%
inoculating ratio. The cells were then cultured at given test temperatures (50, 60, or 85°C) for several
days.

Competitive cell growth assay. KU216 and DA4 were grown in MA-YT medium in separate
overnight cultures at 85°C. The two strains were then inoculated into fresh MA-2YTP medium (20
ml) at KU216:DA4 ratios of 1:1, 10:1, 10$^2$:1, and 10$^3$:1. The mixtures were then cultured at 50°C for
2 days. The mixed culture was then inoculated into another 20 ml of fresh MA-2YTP medium and
incubated at 50°C for another 2 days. Finally, the cells were spread on the MA-2YTP solid medium
and incubated at 85°C overnight. Next, 106 colonies were picked at each KU216:DA4 ratio and the
genotype of the cells were checked. The genotype of cells harboring wild-type *cpkA* was confirmed
by PCR using primers SNP-E530-Fw and SNP-E530-Rv, whereas the genotype of cells harboring

cpkA-E530G was confirmed using primers SNP-E530-Fw and SNP-E530G-Rv. To increase the

specificity of the PCR, an extra mismatched nucleotide was introduced into the 3’ end of primers

SNP-E530-Rv and SNP-E530G-Rv (Table 2) according to the principles described in (29); this was

done to avoid amplification of mismatches caused by the tandem repeats within the C-terminal

region of the cpkA gene. The PCR conditions were as follows: initial denaturation for 2 min at 94°C,

followed by 20 cycles of denaturation at 94°C (15 s), annealing at 62°C (30 s), and extension at

68°C (1 min). GoTaq® Green Master Mix (Promega, WI) was applied to perform the PCR described

above. Genomic DNA was extracted from ten colonies randomly selected at each KU216:DA4 ratio

to test the validity of the PCR above. Briefly, the cpkA gene region was enriched in first

PCR-amplification using primers cpkA-Fw and cpkA-Rv. The genotypes were then confirmed by

sequence analysis with the primer cpkA-singlechecking (Table 2).

203

RESULTS

Sequence alignment and temperature-dependent distribution of chaperonins. Although

CpkA and CpkB share high overall sequence identity (77%), they harbor unique differences within

the C-terminal region. To examine the role of the C-terminal region, we compared the last 20 amino

acids within the C-terminus of chaperonins expressed by microorganisms at different growth
temperatures (Fig. 1). The aligned sequences are shown in Fig. 1A. The C-terminal regions of chaperonins fell into three general clusters: “CpkA-type”, “CpkB-type” and “C-type”. The CpkA-type chaperonin had a mildly hydrophobic and flexible C-terminus comprised primarily of GGM motifs, whereas the CpkB-type chaperonin had a more hydrophilic C-terminal region comprising glutamic acid (Glu [D]), serine (Ser [S]), lysine (Lys [K]), and aspartic acid (Asp [E]). More interestingly, there were some “promiscuous” chaperonins that fell between these two clusters (“C-type”; Fig.1A). C-type chaperonin contained more glycine (Gly [G]), alanine (Ala [A]) and methionine (Met [M]) residues, but also retained a considerable number of charged or polar amino acids. In other words, two major types of C-terminal region were identified: one showing increased hydrophobicity and flexibility (CpkA-type), and the other showing increased electrical charge (CpkB-type). Some hyperthermophiles such as *Pyrococcus furiosus* and *Aeropyrum pernix* harbor only CpkB-type chaperonins. CpkB-type chaperonins seem indispensable for their preferred high growth temperature. The copy number of the CpkA-type chaperonin tended to increase as the optimal cell growth temperature decreased. This tendency was not unique to archaea (Fig. 1B). Extreme psychrophilic bacteria such as *Psychromonas ingrahamii* which shows maximal growth at 15°C or lower for optimal growth and 0°C or lower for minimum growth harbored two CpkA-type chaperonins and one C-type chaperonin, whereas *Psychrobacter cryohalolentis* which grows at 22°C optimally harbored only a CpkA-type chaperonin. Therefore, the appearance of CpkA-type
chaperonins appears to be related to low temperature adaptation. Microorganisms harboring the CpkA-type chaperonin might adapt to lower growth temperatures more readily than those from which it is absent.

**Variant construction and protein stability.** The C-terminal tail of GroEL is rich in GGM motifs; however, five of the residues in the tail region of CpkA are neither Gly nor Met (these residues are Pro529, Glu530, Gln533, Pro538, and Asp545). Therefore, to make this region resemble that of GroEL, we replaced Glu530, Gln533, Pro538, and Asp545 with Gly or Met. Pro529 was not replaced because substitution of this residue with Met or Gly would not result in successive “GGM” motifs. The variant proteins were then expressed in *E. coli* BL21 (DE3) RIL and purified as described in the “Materials and methods” section. The molecular mass of the chaperonins was around 60 kD (Supplemental Fig. 1A). CpkA and its variants showed almost identical CD profiles within the far-UV region (220 nm to 260 nm) at 20°C, suggesting that all possessed very similar secondary structures (Supplemental Fig 1B). Next, we examined the thermostability of each purified protein. The thermal denaturation curves of the purified proteins were determined by monitoring the change in the CD value at 222 nm as the temperature increased from 20 to 100°C at a rate of 5°C min⁻¹. All spectra were corrected to allow for the contribution of the buffer to the signal (Fig. 2A–H). All of the chaperonins showed a major unfolding event at around 90°C. The thermal unfolding plots of CpkA-E530M, -Q533M, -P538M, and -D545G were very similar to those of wild-type CpkA.
CpkA-E530G and -Q533M showed slightly lower thermal stability, while CpkA-P538G and -D545M showed slightly higher stability than wild-type CpkA.

**Effect of mutations on ATPase activity.** Next, we examined the temperature dependency of the ATPase activity of each variant. The results are shown in Fig. 3. CpkA showed the highest ATPase activity between 60 and 70°C. Fig. 3A shows that CpkA-E530G showed optimal ATPase activity at 50°C; the value was almost 4-fold higher than that of CpkA at the same temperature. CpkA-E530M showed greatest activity at 70°C, although the value was almost half that of wild-type CpkA. The ATPase activity of CpkA-E530M at 50°C was a little lower than that of wild-type CpkA.

CpkA-Q533G showed greatest activity at 50°C, although it was similar to that of wild-type CpkA. The optimum activity of CpkA-Q533M was displayed at 80°C, and was 4-fold higher than that of wild-type CpkA at this temperature (Fig. 3B). CpkA-P538G showed a similar temperature-dependent profile to CpkA, although its ATPase activity was almost 2-fold lower. CpkA-P533M showed optimal activity at 70°C, with a value similar to that of wild-type CpkA (Fig. 3C). Finally, CpkA-D545G showed highest activity at 60°C, the value being equal to that of wild-type CpkA. The optimal activity of CpkA-D545M was detected at 70°C, with a value equal to that of wild-type CpkA (Fig. 3D).

**Effect of mutations on chaperonin activity.** InGPS from *T. kodakarensis* (TrpCtk) is a target protein for CpkA (17). Because the refolding of unfolded TrpCtk requires the assistance of CpkA,
we performed an in vitro refolding assay using chemically unfolded TrpC$_{Tk}$ as the substrate. Purified TrpC$_{Tk}$ (25.4 kD; Supplemental Fig. 1A) was denatured in 7 M urea on ice and then refolded at either 50 or 60°C in the presence or absence of wild-type or mutant CpkA. The recovered InGPS activity is shown in Fig. 4. At 60°C, the InGPS activity recovered by CpkA and CpkA-E530G was almost the same (0.66 and 0.67 U/mg, respectively), but was only 28% of the native activity (2.39 U/mg). The InGPS activity recovered by CpkA-D545G (0.63 U/mg) was a little lower than that recovered by CpkA and CpkA-E530G. The InGPS activity of refolded TrpC$_{Tk}$ (assisted by these three chaperonins) was almost 7-fold higher than that of spontaneously refolded TrpC$_{Tk}$ (0.1 U/mg). CpkA-E530M, -Q533G, -Q533M, -P538G, -P538M, and -D545M were also able to assist the refolding of TrpC$_{Tk}$, but the recovered InGPS activity was less than that for CpkA, CpkA-D545G, and D545M. There were significant differences between the chaperonin activities at 50°C. When TrpC$_{Tk}$ was refolded in the absence of CpkA, the specific activity of the refolded TrpC$_{Tk}$ was 0.06 U/mg (about 10% of the native activity (0.58 U/mg)). This increased to 0.25 U/mg (43% of native activity) in the presence of CpkA. Moreover, refolded TrpC$_{Tk}$ showed 1.5-fold higher activity (0.38 U/mg) in the presence of CpkA-E530G than in the presence of CpkA. Refolded TrpC$_{Tk}$ also showed slightly higher activity (0.3 U/mg) in the presence of CpkA-E530M than in the presence of CpkA. In general, the chaperonin activity of the variants was consistent with their ATPase activity (Fig. 3); i.e., a higher ATPase activity equated to a higher chaperonin activity and vice versa. The InGPS activity
of refolded TrpC was not significantly higher in the presence of CpkA-Q533G, and -D545G than in the presence of wild-type CpkA. Moreover, CpkA-P538G and -P538M were less effective than CpkA in terms of refolding of TrpC; their ATPase activities were also much lower than that of CpkA.

Growth profiles of strains KU216 and DA4 and the growth competition assay. To observe
the effect of the E530G mutation on T. kodakarensis in vivo, we constructed strain DA4 (pyrF, cpkA-E530G) and compared its growth profile with that of the parental strain KU216. At 85°C, DA4 showed a little higher growth rate during log phase than KU216 (Fig. 5A). At 60°C, DA4 showed a slightly shortened lag period in the growth curve (Fig. 5B). However, at 50°C, the lag period of DA4 was much shorter than that of KU216. Strain DA4 started to propagate 30 hours earlier and reached the stationary growth phase nearly 80 hours earlier than strain KU216 (Fig. 5C). These results indicate that strain DA4 has a growth advantage in a cold-stressed environment.

To examine how dominant the growth of DA4 was at 50°C, we performed a co-cultivation experiment. KU216 and DA4 were cultured separately to OD660 ~ 0.6. Then, the two cultures were mixed at different KU216:DA4 ratios (1:1, 10:1, 10^2:1, and 10^4:1) and cell growth was monitored at 50°C. When growth entered log phase (after about 2 days), the cultures were inoculated into fresh medium and cultured for a further 2 days at 50°C. The cultures were then spread onto solid medium and incubated at 85°C overnight until colonies appeared. Finally, 106 colonies were picked at each
KU212:DA4 ratio and genotyped. The final results are shown in Fig. 6. At a ratio of 1:1, DA4 was the dominant (~87%) population. However, the percentage of DA4 gradually decreased as the DA4:KU216 ratio decreased (83.5% at a ratio of $10^{-1}:1$, 76.5% at $10^{-2}:1$, and 64.5% at $10^{-4}:1$, respectively). It is striking that DA4 was the dominant throughout, even when the starting population was $10^4$-fold less than that of KU216.

**DISCUSSION**

The duplication and divergence of α and β chaperonin subunit genes such as *cpkA* and *cpkB* in *T. kodakarensis* were thought to have occurred in a common ancestor; the α subunit was subsequently lost from the *Pyrococcus* lineage (2). However, the results of the present study suggest another possibility. The natural habitat of *Thermococcus* is a solfatara, which fluctuates between middle (around solfatara, ~60°C) and high temperatures (at the center of solfatara, ~100°C); therefore, these thermophiles have evolved an extra chaperonin that allows them to adapt to a colder environment. Living in an environment that moves between temperature extremes facilitates natural selection; therefore, such thermophile can evolve quickly (9). Thus, *Thermococci* adapted to lower temperatures by duplicating the chaperonin subunit α (*CpkA*). CpkA acts as an “adaptive allele” and enables *T. kodakarensis* to adapt to its changing habitat. In the absence of *cpkA* gene, *T. kodakarensis* was hard to grow at 60°C (15).
Cold tolerance and cold adaptation are related, but different, behaviors. Cold tolerance usually occurs in response to a temporary drop in temperature, whereupon cold shock proteins (CSPs) enable their hosts to survive. Cold adaptation refers to a longer period of low temperature, triggering a response by not only CSPs but also by many non-stress proteins. However, these non-stress proteins are unstable and incorrectly folded under cold-stress conditions. Chaperonin plays an important role in protein folding and refolding; however, CpkB is not suited to this task under cold-stress conditions because it has low ATPase activity below 60°C (17). Thus, CpkA evolved as a unique cold-inducible chaperonin. CpkA and CpkB differ at the C-terminus, which forms part of equatorial domain. The C-terminal region of CpkA has characteristics similar to those of *E. coli* GroEL in terms of amino acid sequence. CpkA, a cold-inducible thermosome, was classified as a CpkA-type chaperonin belonging to group I chaperonins (which include GroEL of *E. coli*) (Fig. 1A).

By contrast, MM_1798, which encodes a mesophilic chaperonin, was classified as a CpkB-type chaperonin along with those thermosomes from *Pyrococcus*. Moreover, the growth temperature-dependent distribution of chaperonins (Fig. 1B) and the growth profile of the *cpkA* disruptant (*T. kodakarensis* strain DA1) suggest that CpkA-type chaperonins play a key role in cold adaptation. An extra C-type chaperonin in microorganisms might respond to other situations; an example is the TCP1 (CCT)-containing chaperonin in eukaryotic cells, which has up to eight different subunits that interact with different substrates (7). The wide variation in the C-terminal
regions of chaperonins enables microorganisms to adapt to changing environments.

Here, we constructed eight CpkA variants harboring single amino acid substitutions in the C-terminal region, leading to an increase in the number of “GGM” motifs. Purified CpkA-E530G showed increased ATPase and chaperonin activity at 50°C in vitro. Cold adaptation of T. kodakarensis is a multi-loci controlled phenomenon, which is fine-tuned by chaperonins (15), transcription factors (22), the composition of the cell membrane (33), and DEAD-box RNA helicase (35). It is thought that beneficial mutations accumulate incrementally via each of these mechanisms and act synergistically to enable adaptation to cold environments. However, when we introduced cpkA-E530G into the genome of T. kodakarensis, we were surprised to see that a single base pair substitution enabled T. kodakarensis to adapt to a temperature of 50°C; indeed, it showed a shorter lag phase comparing to its parental strain (KU216) (Fig. 5C). We believe that the increased ability to adapt to colder conditions correlates with the improved chaperonin activity of CpkA-E530G at 50°C.

The C-terminal region of GroEL also plays a role in ATPase activity (28) and substrate binding (49). CpkA has a similar substrate preference to GroEL (17). Here, we showed that the ATPase and chaperonin activities of CpkA were dependent on the C-terminal region (Figs. 3 and 4). However, because the high flexibility of the C-terminal region means that it cannot be observed within the crystal structure of α subunit of the chaperonin from Thermococcus sp. KS-1 (PDB ID: 1Q2V) (43), it is difficult to discuss the structure-function relationships of CpkA, which is an orthologue of the
chaperonin in *Thermococcus* sp. KS-1 (99% identical). We did however construct a 3D structural model of a full-length CpkA octamer in MODELLER 9.12 using 1Q2V as a template. The results showed that the C-terminal tail of CpkA is fully extended to form the floor of a cleft within the octamer structure (Supplemental Fig. 2A), and that Glu530 faces a short loop between β1 and β2, which is adjacent to the catalytic residue Asp64 (Supplemental Fig. 2B). The Cryo-electron microscopic study using chaperonins from *Acidianus tengchongensis* also revealed a flat bottom shape of chaperonin (cpn-α complex) (51). According to the results presented herein, it is not so arbitrary to say that the location of the amino acids within the C-terminal region is the primary factor affecting chaperonin function. Replacement of Glu530, which is ~20 Å from the functional ATPase domain of CpkA (Supplemental Fig. 2B), by either Gly or Met plays an important role in increasing chaperonin activity at low temperatures. This distance is also an ideal mutation area in the directed evolution experiment (3). We also found that particular amino acids within the C-terminal region play another important role in cold adaptation. For example, Gly enhanced ATPase activity at 50°C to a greater extent than Met. Pro, Ser, and Ala are the most common amino acids in tandem repeats, whereas Met, Ile, and Trp are largely absent (19). Replacing Gly with polar residues (E530G, Q538G, and D545G) increased ATPase activity at 50°C, whereas replacing Met maintained ATPase activity above 60°C. Functional movement of the catalytic core responsible for ATPase activity determine the catalytic rate and substrate binding of the chaperonin α-subunit of *Thermococcus* KS-1.
Glu530 substitution with Gly in CpkA increased ATPase activity and would remove any possible blemish between intermediate domain and equatorial domain so that the functional movement would easily occur compared to wild type (Supplemental Fig. 2B). Pro is usually thought of as a “turn maker” or “helix breaker”, and dictates the local structure of a protein by providing extra rigidity (i.e., resulting in lower entropy). Replacing Pro538 with either Gly or Met (CpkA-P538G or CpkA-P538M) would alter the structure of the peptide backbone and push the C-tail in the wrong direction, resulting in reduced ATPase activity. We also tried to replace both Glu530 and Asp545 with Gly (CpkA-E530G/D545G); however, this double-mutant (CpkA-E530G/D545G) did not show any additional increase in ATPase activity at 50°C. An alternative evolutionary path would be to increase the copy number of CpkA-type chaperonin genes, thereby increasing chaperonin activity at lower temperatures; indeed, this is the evolutionary pathway taken by some mesophiles and psychrophiles (Fig. 1B).

The final cell yield of DA4 and KU216 were almost similar when the growth entered the stationary phase (Fig. 5A–C). However, DA4 showed a shorter lag phase than KU216 at 50°C, suggesting that CpkA-E530G in DA4 could perform its function on target proteins from the very start of the growth cycle, whereas KU216 would take much longer to accumulate correctly folded proteins. In summary, we identified two growth temperature-dependent trends (correlation between the
CpkA-type chaperonin gene copy number and growth temperature) in group II chaperonins, both of which are driven by the initial duplication and subsequent mutation of a chaperonin gene. The tandem repeats within the C-terminal region of CpkA are a mutational hotspot (13). Here, we showed that a single base pair substitution (E530G) in the C-terminal region of *T. kodakarensis* CpkA allowed the organism to adapt to a lower growth temperature (50°C). Because chaperonins regulate protein-protein interactions within cells, especially under stress conditions (36), changes in the activity of chaperonins will lead to changes in cell properties, such as cold adaptation. GroEL has a marked effect on the refolding of proteins called mutators, which are involved in DNA repair (27); thus, GroEL might indirectly increase the mutation rate of the host cell. Because GroEL and CpkA have a similar substrate bias (17), we would expect the role of CpkA in *T. kodakarensis* to resemble that of GroEL. Further *in vivo* studies should aim to derive a mesophilic lineage from hyperthermophiles to better understand cold adaptive mechanisms in archaea.

**ACKNOWLEDGMENTS**

This study was mainly supported by Grant of Japan Society for the Promotion of Science (JSPS) KAKENHI (26292045). Bioinformatic analysis was supported by a Grant for Individual Special Research, provided by Kwansei Gakuin University.


ingrahamii sp. nov., a novel gas vacuolate, psychrophilic bacterium isolated from Arctic

Bakermans C, Ayala-del-Rio HL, Ponder MA, Vishnivetskaya T, Gilichinsky D,
Thomashow MF, Tiedje JM. 2006. Psychrobacter cryohalolentis sp. nov. and
Psychrobacter arcticus sp. nov., isolated from Siberian permafrost. Int J Syst Evol
Microbiol 56: 1285-1291.

Bigotti MG, Clarke AR. 2008. Chaperonins: The hunt for the Group II mechanism. Arch
Biochem Biophys 474: 331-339.

Brock TD, Freeze H. 1969. Thermus aquaticus gen. n. and sp. n., a nonsporulating extreme

Diaz Arenas C, Cooper TF. 2013. Mechanisms and selection of evolvability: experimental

Eswar N, Webb B, Marti-Renom MA, Madhusudhan MS, Eramian D, Shen MY,
Protoc Bioinformatics Chapter 5: Unit 5 6.

Farr GW, Fenton WA, Horwich AL. 2007. Perturbed ATPase activity and not "close
confinement" of substrate in the cis cavity affects rates of folding by tail-multiplied GroEL.


residues 526 KNDAAD 531 in the flexible C-terminal region of the chaperonin GroEL are critical for substrate protein folding within the central cavity. J Biol Chem 283: 6886-6896.


FIGURE LEGENDS

FIG 1. Three types of chaperonin and their distribution in Prokaryotae. (A) Alignment of the C-terminal regions (last 20 amino acids) of chaperonins from Pyrococcus, Thermococcus, Methanococcus, Sulfolobus, Thermus, and E. coli. Chaperonins whose C-terminal regions do not belong to the CpkA or CpkB types are labeled “C-type”. (B) Distribution of chaperonins according to growth temperature. Prokaryotes include Aeropyrum pernix (39), Archaeoglobus fulgidus (18), Chloroflexus aurantiacus (37), Escherichia coli (26), Methanosarcina acetivorans (44), Methanococcoides burtonii (18), Methanocaldococcus jannaschii (18), Methanopyrus kandleri (18), Methanococcus mazei (38), Methanothermobacter thermautotrophicus (50), Methanothermobacter thermolithotrophicus (23), Picrophilus torridus (42), Psychromonas ingrahamii (5), Psychrobaculum aerophilum (18), Pyrococcus abyssi (18), Psychrobacter cryohalolentis (6), Pyrococcus furiosus (12),
**Pyrococcus horikoshii** (18), *Sulfolobus solfataricus* (18), *Sulfolobus tokodaii* (46), *Thermus aquaticus* (8), *Thermococcus kodakarensis* (4), *Thermoplasma acidophilum* (18), and *Thermoplasma volcanium* (18). The numbers in the middle column represent the number of CpkA-type, CpkB-type, and C-type genes. The circled numbers indicate the lowest growth temperature, and the arrows indicate the growth temperature ranges.

**FIG 2. Thermostability of CpkA and its variants.** Each sample contained 50 μg ml\(^{-1}\) of chaperonin dissolved in 25 mM Tris-HCl (pH 7.8) containing 300 mM KCl. Sample denaturation was monitored by measuring the change in absorption at 222 nm at temperatures from 30°C to 100°C. Data for each sample were plotted after subtracting data obtained at indicated temperatures. The denaturation of CpkA is indicated by dark dots in each frame, whereas each mutation is indicated by a white dot. (A) CpkA-E530G. (B) CpkA-E530M. (C) CpkA-Q533G. (D) CpkA-Q533M. (E) CpkA-P538G. (F) CpkA-P538M. (G) CpkA-D545G. (H) CpkA-D545M.

**FIG 3. ATPase activity of CpkA and its variants.** Temperature-dependent ATPase activity of CpkA (black triangle) and its variants (glycine substitutions are indicated by white diamonds and methionine substitutions are indicated by white circles) are shown in panels A to D. (A) CpkA, CpkA-E530G, and CpkA-E530M. (B) CpkA, CpkA-Q533G, and CpkA-Q533M. (C) CpkA,
FIG 4. Effect of chaperonins on TrpC refolding. The InGPS activity of refolded TrpC was measured at 60°C (A) and 50°C (B) in the absence or presence of chaperonins. Columns 1, 2, 3, 4, 5, 6, 7, 8, and 9 show the InGPS activity of refolded TrpC in the presence of CpkA, CpkA-E530G, CpkA-E530M, CpkA-Q533G, CpkA-Q533M, CpkA-P538G, CpkA-P538M, CpkA-D545G, and CpkA-D545M, respectively. Column 10 shows the InGPS activity of spontaneously refolded TrpC in the absence of chaperonin. Column 11 shows the activity of native TrpC. One unit (U) of specific activity is defined as the amount of activity resulting in the formation of 1 μmole of InGPS in 1 min. All analyses were conducted in triplicate and mean values are shown.

FIG 5. Growth profiles of KU216 and DA4. KU216 and DA4 were cultured in MA-2YTP medium at different temperatures and their growth profiles monitored at OD₆₆₀. (A) Growth profiles of KU216 (white circles) and DA4 (black circles) at 85°C. (B) Growth profiles of KU216 (white circles) and DA4 (black circles) at 60°C. (C) Growth profiles of KU216 (white circles) and DA4 (black circles) at 50°C.

FIG 6. Dominant growth of DA4 at lower temperatures. KU216 and DA4 cells were pre-cultured
separately at 85°C. The two cultures were then mixed at different KU216:DA4 ratios (1:1, 10:1, 10²:1, and 10³:1) and cultured at 50°C for 2 days. The cultures were then added to fresh growth medium and cultured for a further 2 days. The log phase cultures were spread onto solid medium and cultured overnight at 85°C. Over 100 colonies per KU216:DA4 ratio were randomly picked and genotyped by PCR using primer pairs SNP-E530-Fw/SNP-E530-Rv and SNP-E530-Fw/SNP-E530G-Rv. (A) Genotype of colonies confirmed by PCR and sequencing. Ten representative colonies from each KU216:DA4 ratio are shown. The results were confirmed by sequencing (shown at the bottom of gels). “K”, KU216 genotype; “D”, DA4 genotype. (B) Proportion of DA4 and KU216 cells in co-cultures at each KU216:DA4 ratio after culture at 50°C. DA4 and KU216 are indicated by black and white bars, respectively.
Fig 3
Table 1. Strains and plasmids used in this study

<table>
<thead>
<tr>
<th>Strain or plasmid</th>
<th>Relevant characteristic(s)</th>
<th>Source or reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Strains</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E. coli</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DH5α</td>
<td>F–Φ80lacZΔM15 Δ(lacZYA-argF) U169 recA1 endA1 hsdR17 (rK-, mK+) phoA supE44 λ– thi-1 gyrA96 relA1</td>
<td>This study</td>
</tr>
<tr>
<td>BL21(DE3)pLysE</td>
<td>F–ompT hsdSB(rB–, mB–) gal dcm (DE3) pLysE (CamR)</td>
<td>This study</td>
</tr>
<tr>
<td><strong>T. kodakarensis</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KU216</td>
<td>pyrF</td>
<td>41</td>
</tr>
<tr>
<td>DA4</td>
<td>pyrF cpkA-E530G</td>
<td>This study</td>
</tr>
<tr>
<td><strong>Plasmids</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pUD2-cpkAfr</td>
<td>Plasmid used to replace the cpkA gene in <em>T. kodakarensis</em></td>
<td>15</td>
</tr>
<tr>
<td>pUD2-cpkA-E530G</td>
<td>pUD2-cpkAfr derivative; cpkA-E530G</td>
<td>This study</td>
</tr>
<tr>
<td>pCPAE</td>
<td>Expression vector for cpkA in <em>E. coli</em></td>
<td>25</td>
</tr>
<tr>
<td>pCPAE-E530G</td>
<td>pCPAE derivative; mutation, cpkA-E530G; expression vector for <em>E. coli</em></td>
<td>This study</td>
</tr>
<tr>
<td>pCPAE-E530M</td>
<td>pCPAE derivative; mutation, cpkA-E530M; expression vector for <em>E. coli</em></td>
<td>This study</td>
</tr>
<tr>
<td>pCPAE-Q533G</td>
<td>pCPAE derivative; mutation, cpkA-Q533G; expression vector for <em>E. coli</em></td>
<td>This study</td>
</tr>
<tr>
<td>pCPAE-Q533M</td>
<td>pCPAE derivative; mutation, cpkA-Q533M; expression vector for <em>E. coli</em></td>
<td>This study</td>
</tr>
<tr>
<td>pCPAE-P538G</td>
<td>pCPAE derivative; mutation, cpkA-P538G; expression vector for <em>E. coli</em></td>
<td>This study</td>
</tr>
<tr>
<td>pCPAE-P538M</td>
<td>pCPAE derivative; mutation, cpkA-P538M; expression vector for <em>E. coli</em></td>
<td>This study</td>
</tr>
<tr>
<td>pCPAE-D545G</td>
<td>pCPAE derivative; mutation, cpkA-D545G; expression vector for <em>E. coli</em></td>
<td>This study</td>
</tr>
<tr>
<td>pCPAE-D545M</td>
<td>pCPAE derivative; mutation, cpkA-D545M; expression vector for <em>E. coli</em></td>
<td>This study</td>
</tr>
<tr>
<td>Name</td>
<td>Sequence (5' to 3')</td>
<td></td>
</tr>
<tr>
<td>--------------------</td>
<td>---------------------</td>
<td></td>
</tr>
<tr>
<td>cpkA-E530G-Fw</td>
<td>CGCGGGCGCGGAGGCGACA</td>
<td></td>
</tr>
<tr>
<td>cpkA-E530G-Rv</td>
<td>GCCGTGGCCGCTGGCTGGCTT</td>
<td></td>
</tr>
<tr>
<td>cpkA-E530M-Fw</td>
<td>CGCGGGCGCGGAGGCGACA</td>
<td></td>
</tr>
<tr>
<td>cpkA-E530M-Rv</td>
<td>GCCGTGGCCGCTGGCTGGCTT</td>
<td></td>
</tr>
<tr>
<td>cpkA-Q533G-Fw</td>
<td>GCCGTGGCCGCTGGCTGGCTT</td>
<td></td>
</tr>
<tr>
<td>cpkA-Q533G-Rv</td>
<td>CGCGGGCGCGGAGGCGACA</td>
<td></td>
</tr>
<tr>
<td>cpkA-Q533M-Fw</td>
<td>GCCGTGGCCGCTGGCTGGCTT</td>
<td></td>
</tr>
<tr>
<td>cpkA-Q533M-Rv</td>
<td>CGCGGGCGCGGAGGCGACA</td>
<td></td>
</tr>
<tr>
<td>cpkA-P538G-Fw</td>
<td>GCCGTGGCCGCTGGCTGGCTT</td>
<td></td>
</tr>
<tr>
<td>cpkA-P538G-Rv</td>
<td>CGCGGGCGCGGAGGCGACA</td>
<td></td>
</tr>
<tr>
<td>cpkA-P538M-Fw</td>
<td>GCCGTGGCCGCTGGCTGGCTT</td>
<td></td>
</tr>
<tr>
<td>cpkA-P538M-Rv</td>
<td>CGCGGGCGCGGAGGCGACA</td>
<td></td>
</tr>
<tr>
<td>cpkA-D545G-Fw</td>
<td>GCCGTGGCCGCTGGCTGGCTT</td>
<td></td>
</tr>
<tr>
<td>cpkA-D545G-Rv</td>
<td>CGCGGGCGCGGAGGCGACA</td>
<td></td>
</tr>
<tr>
<td>cpkA-D545M-Fw</td>
<td>GCCGTGGCCGCTGGCTGGCTT</td>
<td></td>
</tr>
<tr>
<td>cpkA-D545M-Rv</td>
<td>CGCGGGCGCGGAGGCGACA</td>
<td></td>
</tr>
<tr>
<td>cpkA-singlechecking</td>
<td>GCCGTGGCCGCTGGCTGGCTT</td>
<td></td>
</tr>
</tbody>
</table>

The mutated nucleotides within the *cpkA* gene are indicated by solid underlining. Introduced restriction sites are indicated by dashed underlining.