RESEARCH ARTICLE

Comprehensive mutagenesis to identify amino acid residues contributing to the difference in thermostability between two originally thermostable ancestral proteins

Satoshi Akanuma[1]*, Minako Yamaguchi[2], Akihiko Yamagishi[2]

1 Faculty of Human Sciences, Waseda University, Tokorozawa, Saitama, Japan, 2 Department of Applied Life Sciences, Tokyo University of Pharmacy and Life Sciences, Hachioji, Tokyo, Japan

* akanuma@waseda.jp

Abstract

Further improvement of the thermostability of inherently thermostable proteins is an attractive challenge because more thermostable proteins are industrially more useful and serve as better scaffolds for protein engineering. To establish guidelines that can be applied for the rational design of hyperthermostable proteins, we compared the amino acid sequences of two ancestral nucleoside diphosphate kinases, Arc1 and Bac1, reconstructed in our previous study. Although Bac1 is a thermostable protein whose unfolding temperature is around 100°C, Arc1 is much more thermostable with an unfolding temperature of 114°C. However, only 12 out of 139 amino acids are different between the two sequences. In this study, one or a combination of amino acid(s) in Bac1 was/were substituted by a residue(s) found in Arc1 at the same position(s). The best mutant, which contained three amino acid substitutions (S108D, G116A and L120P substitutions), showed an unfolding temperature more than 10°C higher than that of Bac1. Furthermore, a combination of the other nine amino acid substitutions also led to improved thermostability of Bac1, although the effects of individual substitutions were small. Therefore, not only the sum of the contributions of individual amino acids, but also the synergistic effects of multiple amino acids are deeply involved in the stability of a hyperthermostable protein. Such insights will be helpful for future rational design of hyperthermostable proteins.

Introduction

Elucidation of the mechanisms underlying protein thermostability is an important issue not only for understanding the amino acid sequence-structure-stability relationship, but also for developing proteins that can be used in industrial processes [1–4]. Comparative structural analyses between thermophilic and mesophilic protein structures have provided insights into the molecular mechanisms responsible for the thermostability of thermophilic proteins [5,6]. The mechanisms suggested by those studies include the formation of intra- or inter-molecular ion-pairs and

ion-pair networks in the native structure [7–9] and also in the denatured state [10], increased hydrophobicity and improved packing in the hydrophobic core [11–14], shorter loops [15,16], improved subunit-subunit interactions [3,17,18] and entropic advantages due to the increased flexibility of native structures [19]. However, we have yet to reach a comprehensive understanding of the mechanism of protein thermostability [20]. For example, one paper reported that no consistent trends were found between the amino acid composition of proteins and their stabilities [16], while another paper pointed out that small non-polar amino acids are more frequent in thermostable proteins [21]. Furthermore, in many cases, these factors would only partially explain the differences in thermostability between thermophilic proteins and their mesophilic homologues. A major hindrance in identifying all of the amino acid residues involved in determining the differences in thermostability between a pair of homologous proteins is the fact that many mutations that do not affect thermal stability have accumulated during evolution.

Previously, we inferred ancestral amino acid sequences of nucleoside diphosphate kinases (NDKs) possessed by the last archaeal common ancestor and the last bacterial common ancestor [22]. NDK is an enzyme that catalyzes the transfer of γ-phosphate of a nucleoside triphosphate to a nucleoside diphosphate. We built phylogenetic trees by comparing the amino acid sequences of NDKs possessed by extant archaea and bacteria, and reconstructed several archaeal and bacterial common ancestral NDKs. Among the reconstructed NDKs, the archaeal common ancestor Arc1 and the bacterial common ancestor Bac1 exist as homo-hexamers with a protomer of 139 amino acid residues, and only 12 amino acid residues differ between their amino acid sequences (Fig 1A). The unfolding midpoint temperature of Arc1 (113˚C) is 14˚C higher than that (99˚C) of Bac1 at pH 7.0. Crystal structures of Arc1 and Bac1 have been reported (PDB codes: 3VVT and 3VVU). Ramachandran plots computed by the PROCHECK module of SAVES (https://saves.mbi.ucla.edu/) showed that only one residue (Val113) existed in disallowed regions in both Arc1 and Bac1 structures (S1 Fig). Superposition of the structure of Arc1 upon that of Bac1 (Fig 1B) yielded a root mean square deviation of alpha carbons (C_α rmsd) of 0.61 Å for the 139 aligned residues, as calculated using PDBeFold ver. 2.59 (https://www.ebi.ac.uk/msd-srv/ssm/). Reduced non-polar accessible surface areas, and increased numbers of inter-subunit ion pairs and hydrogen bonds were proposed as structural features that likely contribute to the high thermal stabilities of Arc1 and Bac1 based on the comparison of their crystal structures with that of the NDK from the mesophilic *Dictyostelium discoideum* (see reference [22] for more details).

In this study, in order to understand the structural mechanisms by which originally very thermostable proteins become more stable, we first created twelve Bac1 mutants in which one amino acid was replaced by the residue found at the same position in the amino acid sequence of Arc1. Three mutants indeed showed an enhanced unfolding midpoint temperature of greater than 2˚C at pH 6.0, pH 7.0 and pH 7.6. Moreover, combinations of the three beneficial amino acid substitutions further improved the thermal stability of Bac1, thus producing a mutant that displayed an unfolding midpoint temperature quite similar to that of Arc1 at pH 7.6. Therefore, only three of the 12 amino acids that differ between Arc1 and Bac1 play a major role in the difference in the thermostabilities of the two proteins. However, the other nine amino acids in Arc1 contribute somewhat to its greater thermostability although their individual contributions are small. Finally, we compared the tertiary structures of Arc1 and Bac1 and discuss the structural mechanisms for the enhanced stability of the Bac mutant bearing the three amino acid substitutions.

Materials and methods

Site-directed mutagenesis

The gene encoding the ancestral NDK corresponding to the last bacterial common ancestor had been cloned into pET21c [22]. Site-directed mutagenesis was carried out on the Bac1 gene
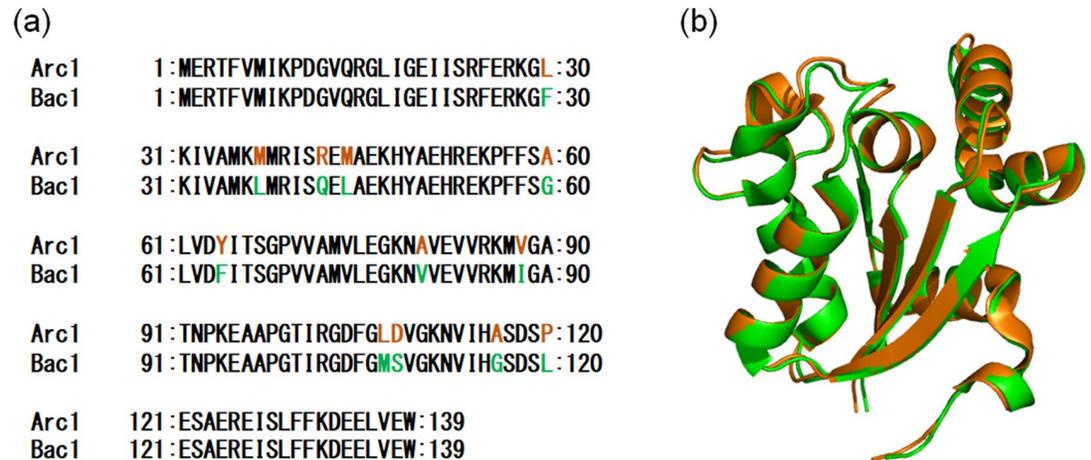
Fig 1. Comparison of the amino acid sequences and tertiary structures of the ancestral NDKs. (a) Pair-wise amino acid sequence alignment of Arc1 and Bac1. Inconsistent amino acid residues are colored. (b) Superimposed structures of Arc1 (orange) and Bac1 (green). Superimposition was performed with PDBeFold ver. 2.59 (https://www.ebi.ac.uk/msd-srv/ssm/) and the superimposed structures were visualized with PyMOL (https://pymol.org/2/).

https://doi.org/10.1371/journal.pone.0258821.g001

using the splicing-by-overlap extension (SOE) PCR method [23]. SOE PCR consists of two rounds of PCR using the outer primer pairs (T7_forward: 5'- TAATACGACTCACTATAGG-3' and T7_reverse: 5'-GCTAGTTATTGCTCAGCGG-3') that are common to all reactions and the inner mutagenic primer pairs (S1 Table). A pair of mutagenic primers have complementary sequences at their 3′-ends. T7_forward and a mutagenic reverse primer were used for amplification of the 5′ gene segment (PCR #1). A mutagenic forward primer and T7_reverse primer were used for amplification of the 3′ gene segment (PCR #2). Both amplifications used pET21c carrying the gene encoding Bac1 as the template. The two amplified fragments generated in PCR #1 and PCR #2 were purified using the FastGene Gel/PCR Extraction Kit (Nippon Genetics) and then mixed. They were used for PCR amplification again in the presence of T7_forward and T7_reverse primers (PCR #3). All PCR amplifications were performed in a reaction mixture containing 1× KOD -plus- PCR buffer (Toyobo), 1 mM MgSO$_4$, 0.2 mM each of dNTPs, 25 µM each of a pair of primers, and 1.0 unit of KOD -plus- DNA polymerase. The time-temperature program was as follows: step 1, 95˚C, 3 min; step 2, 95˚C, 30 s; step 3, 55˚C, 30 s; step 4, 68˚C, 1 min. Steps 2–4 were repeated 25 times. After PCR #3, the product was again purified using the FastGene Gel/PCR Extraction Kit (Nippon Genetics).

Construction of expression plasmids

The amplified DNA encoding each Bac1 mutant was digested with *Nde*I and *Bam*HI (New England Biolabs) and then purified by agarose gel electrophoresis. The purified DNA was ligated into the *Nde*I-*Bam*HI site of pET21c and then used to transform *Escherichia coli* JM109. Plasmid DNA was prepared from each transformant after cultivating overnight in 2 mL of Luria-Bertani (LB) medium supplemented with 150 µg/ml of ampicillin at 37˚C. The resulting plasmid DNA was subjected to DNA sequencing to confirm the nucleotide sequence of the region encoding the Bac1 mutant.

Protein preparation

For preparation of the Bac1 mutants, *E. coli* Rosetta2 (DE3) was transformed with each of the expression plasmids. Each transformant was cultured overnight at 37˚C in 2 L of LB medium

supplemented with 150 μg/ml of ampicillin. Overexpression was induced by the Overnight Express Autoinduction System (Novagen-Merck). The cells were then harvested, resuspended in 20 mM Tris, pH 7.5, 1 mM EDTA, and disrupted by sonication. The cell lysates were centrifuged at 60,000 x g for 20 minutes. The resulting supernatant was individually heat-treated at 75˚C for 20 min and then centrifuged again at 60,000 x g for 20 min. To purify the protein, the supernatant was filtered and then subjected to HiTrapQ (Cytiva) column chromatography. The adsorption buffer was 20 mM Tris, pH 7.5, 1 mM EDTA, and the elution buffer was 20 mM Tris, pH 7.5, 1 mM EDTA, 1 M NaCl. The fraction containing the Bac1 mutant was collected and dialyzed overnight against 20 mM Tris, pH 8.8, 1 mM EDTA, followed by ResourceQ (Cytiva) column chromatography. The adsorption buffer was 20 mM Tris, pH 8.8, 1 mM EDTA, and the elution buffer was 20 mM Tris, pH 8.8, 1 mM EDTA, 1 M NaCl. The fractions that were homogeneous judging from the results of Coomassie blue staining after SDS-poly-acrylamide gel electrophoresis were used for subsequent analysis.

Quantification of protein concentration

Protein concentrations were determined using the A_{280} values of the protein solutions. The molar absorption coefficient at 280 nm for each Bac1 mutant was calculated as reported by Pace and colleagues [24], who modified the procedure described by Gill and von Hippel [25].

Thermal stability measurement

The T_m value of each protein was estimated from the thermal denaturation curve obtained by monitoring the change in ellipticity at 222 nm using a J-1100 spectropolarimeter (Jasco) equipped with a programmable temperature controller and a pressure-proof cell compartment that prevented the aqueous solution from bubbling and evaporating at high temperatures. Protein solutions were diluted to a final concentration of 20 μM with 20 mM potassium phosphate (pH 6.0 or 7.6), 50 mM KCl, 1 mM EDTA. A 0.1 cm path-length cell was used. The temperature was increased at a rate of 1.0˚C/min.

Activity measurement

The enzymatic reaction catalyzed by Arc1, Bac1 and its mutants was assayed at 70˚C by monitoring the increase in the amount of the product ATP using the Kinase-Glo Plus Luminescent Kinase Assay kit (Promega). The assay solution consisted of 50 mM HEPES (pH 8.0), 25 mM KCl, 10 mM $(NH_4)_2SO_4$, 2.0 mM $Mg(CH_3COO)_2$, 1.0 mM dithiothreitol, 1.0 mM ADP and 2.5 mM GTP. One enzyme unit equaled 1 μmol ATP formed per min. The Michaelis constant values (K_ms) for the substrate ADP and the catalytic rate constant (k_{cat}) were calculated based on the steady-state kinetic data with an assay solution of 50 mM HEPES (pH 8.0), 25 mM KCl, 10 mM $(NH_4)_2SO_4$, 2.0 mM $Mg(CH_3COO)_2$, 1.0 mM dithiothreitol, 2.5 mM GTP, with ADP at concentrations between 50 and 1000 μM. The kinetic parameters were calculated by nonlinear least-square fitting of the steady-state velocity data to the Michaelis-Menten equation using the Enzyme Kinetics module of SigmaPlot Ver. 13 (Systat Software).

Results

Mutagenesis of residues in the bacterial common ancestral NDK, Bac1

A pairwise alignment of the amino acid sequences of Arc1 and Bac1 (Fig 1A) showed that 127 out of 139 aligned residues are identical between the two proteins. Therefore, only 12 sites were occupied by different amino acids in the two proteins. To identify amino acid substitutions that enhance the thermal stability of Bac1, an amino acid in Bac1 was substituted with

the residue found at the same position in Arc1. Accordingly, we created twelve mutants of Bac1.

The genes encoding the mutant proteins were PCR amplified and then overexpressed in *E. coli* Rosetta2 (DE3). Each mutant protein was purified to homogeneity by successive column chromatography steps and then used for temperature-induced unfolding experiments (S2 Fig). The unfolding experiments were first performed at pH 6.0, pH 7.0 and pH 8.0. Although all of the proteins showed cooperative two-state transitions at pH 6.0 and pH 7.0, G60A and S108D showed atypical unfolding curves at pH 8.0 (S3 Fig). In particular, S108D seemed to undergo a structural change in which the content of helical structure increases, rather than denaturing into a random coil, as the temperature rose. Therefore, the thermal unfolding experiments were also performed at pH 7.6. The unfolding midpoint temperatures (T_m) of Arc1, Bac1 and the 12 mutants are listed in S2 Table and compared in Fig 2. Three mutants (S108D, G116A, L120P) had significantly increased T_ms of greater than 2°C compared to Bac1 at pH 6.0, pH 7.0 and pH 7.6. The mutant with the highest T_m value among the mutants was G116A, which exhibited 5–7°C greater T_ms than those of Bac1. The T_m value of S108D was 6°C greater than that of Bac1 at pH 7.6, whereas the T_ms of the mutant was only 3°C greater than those of Bac1 at pH 6.0 and pH 7.0. Therefore, the negative charge of the aspartate residue's side chain is thought to be responsible for the increase in thermostability of S108D. The T_m values of L120P were also higher than those of Bac1 by 2–3°C between pH 6.0 and pH 7.6. Conversely, no mutants showed decreased T_ms of greater than 2°C between pH 6.0 and 7.6. At pH 8.0, four mutants (L37M, Q42R, L44M, I88V) showed decreased T_ms compared to Bac1.

To understand the effects of amino acid substitutions on catalytic properties, the kinetic parameters of Arc1, Bac1 and the 12 mutants were obtained from steady-state kinetic experiments using various concentrations of ADP and 2.5 mM GTP (Fig 3 and S2 Table). The specific activity measurements of Arc1 and Bac1 as a function of temperature revealed that both ancestral NDKs showed the highest specific activity at 70°C under the condition employed (S4 Fig). Therefore, it is reasonable to expect that all of the Bac1 mutants also showed the highest specific activity at 70°C. Accordingly, the steady-state kinetic experiments were performed at 70°C. It should be noted that the twelve mutated positions are not in the active site. The hexamers of Arc1 and Bac1 have six active sites, respectively, and each active site consists of residues in a single subunit. In addition, the inter-subunit contacts are robust even at unfolding temperature (see Discussion). Therefore, the effect of subunit dissociation on activity was not considered. To ensure high affinity with the substrate at a very high temperature, hyperthermophilic enzymes should require smaller K_m values at a given temperature than those of their less stable homologues. However, Arc1 had a K_m value for ADP (430 μM) that is 2.4 times more unfavorable than that (180 μM) of Bac1. The overall enzyme efficiency is customarily expressed as k_{cat}/K_m. The k_{cat}/K_m value (4.1 μM^{-1} s^{-1}) of Arc1 was two times better than that of Bac1 (2.0 μM^{-1} s^{-1}) due to the k_{cat} value (1800 s^{-1}) of Arc1 that was 4.9 times greater than that (370 s^{-1}) of Bac1. Six mutants (L37M, Q42R, L44M, M107L, G116A, L120P) had smaller k_{cat}/K_m values compared to Bac1. Conversely, F30L, G60A, F64Y, V80A, I88V and S108D had greater k_{cat}/K_m values compared to Bac1. When compared to the kinetic parameters of Arc1, all of the Bac1 mutants showed more favorable K_m values but smaller k_{cat} and k_{cat}/K_m values.

Combination of the beneficial mutations

Fig 2 shows that no single mutation was sufficient to improve the thermal stability of Bac1 to the same level as that of Arc1. It is reasonable to assume that combining multiple beneficial
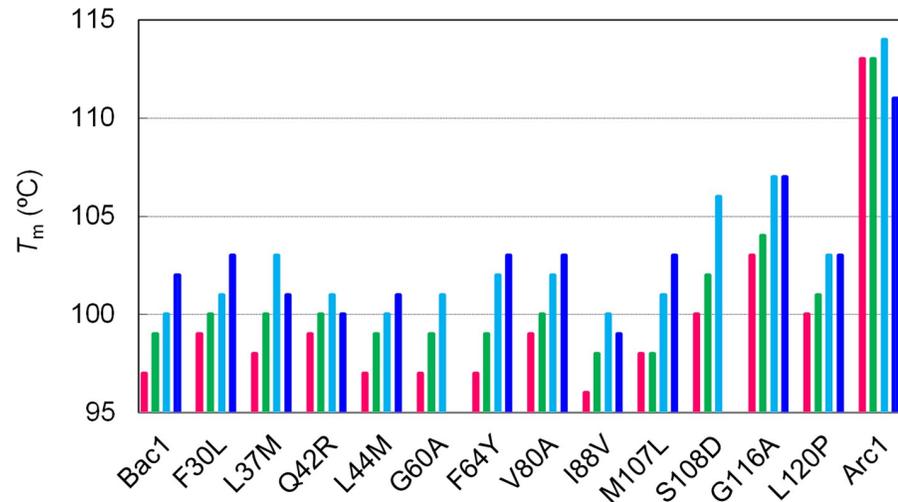
Fig 2. T_ms of Arc1, Bac1 and its mutants at pH 6.0 (magenta), pH 7.0 (green), pH 7.6 (cyan) and pH 8.0 (blue).
The T_m values were estimated from the data shown in S2 Fig. T_m values are not presented for G60A, S108D, S108D/G116A and S108D/G116A/L120P at pH 8.0 because atypical unfolding curves were observed for the proteins at pH 8.0 (S3 Fig).

https://doi.org/10.1371/journal.pone.0258821.g002

substitutions in a protein would further improve its thermostability if the effects of the substitutions are not conflicting. Therefore, we simultaneously introduced S108D and G116A substitutions into Bac1. Temperature-induced unfolding curves of the resulting double mutant (S108D/G116A) at pH 6.0, pH 7.0 and pH 7.6 were compared to those of Arc1, Bac1 and G116A (Fig 4). Similar to the S108D mutant, S108D/G116A showed atypical unfolding curves at pH 8.0. The T_m value (105˚C) of S108D/G116A at pH 6.0 was 8˚C and 2˚C higher than those of Bac1 and G116A, respectively, but 8˚C lower than that of Arc1. At pH 7.0, T_m (107˚C) of S108D/G116A was 8˚C and 3˚C higher than those of Bac1 and G116A, respectively, and 6˚C lower than that of Arc1. At pH 7.6, the T_m (112˚C) of S108D/G116A was 12˚C and 5˚C higher than those of Bac1 and G116A, respectively, and only 2˚C lower than that of Arc1. Thus, a greater synergistic effect of the combination of the two amino acid substitutions was observed at pH 7.6 than at pH 6.0 and pH 7.0.
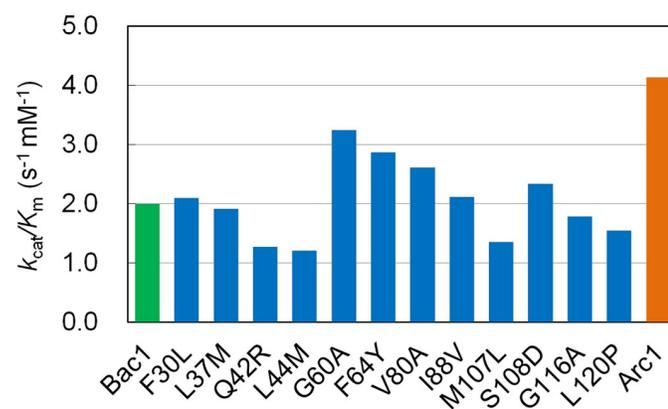

Fig 3. The k_{cat}/K_m values of Arc1, Bac1 and its mutants at 70˚C. K_m for ADP, and k_{cat} were calculated by nonlinear least-square fitting of the steady-state kinetic data to the Michaelis-Menten equation using the Enzyme Kinetics module of SigmaPlot Ver. 13 (Systat Software, Richmond) and listed in S2 Table.
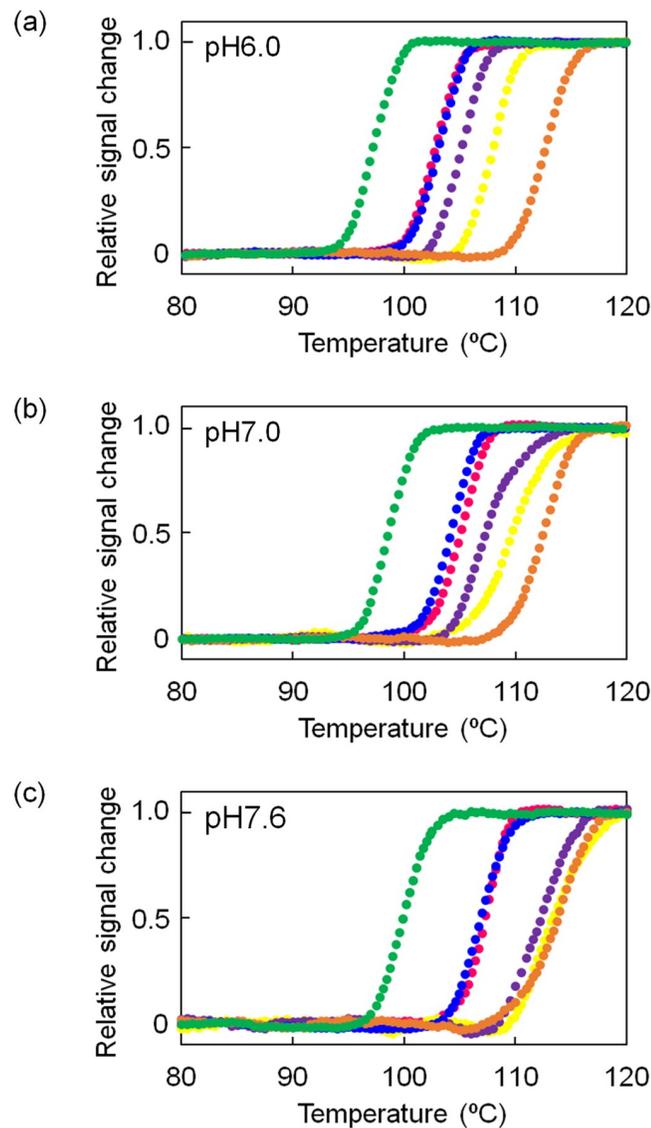
https://doi.org/10.1371/journal.pone.0258821.g003

Fig 4. Temperature-induced unfolding curves of Arc1, Bac1 and its mutants at pH 6.0 (a), pH 7.0 (b) and pH 7.6 (c). The change in ellipticity at 222 nm was monitored as a function of temperature. The scan rate was 1.0˚C/min. The samples contained 20 μM protein in 20 mM potassium phosphate (pH 6.0 or 7.6), 50 mM KCl, 0.5 mM EDTA. The plots were normalized with respect to the baseline of the native and denatured states. Orange, Arc1; green, Bac1; blue, G116A; purple, S108D/G116A; yellow, S108D/G116A/L120P; magenta, Bac1-mu9.

We also introduced the L120P substitution into the S108D/G116A mutant because the substitution also increased the T_m of Bac1 by 2–3˚C. Fig 4 shows that the resulting S108D/G116A/L120P mutant had further improved thermostability. Moreover, the T_m value (113˚C) of the resulting S108D/G116A/L120P mutant was close to that (114˚C) of Arc1 at pH 7.6. Thus, although twelve amino acids are different between the sequences of Arc1 and Bac1, the difference in T_m of 14˚C between the two ancestral NDKs can be almost eliminated at pH 7.6 by only three amino acid substitutions.

The kinetic parameters of S108D/G116A and S108D/G116A/L120P at 70˚C are provided in S2 Table. The K_m values of the double and triple mutants were between those of Bac1 and Arc1. The k_{cat} values of both mutants were slightly greater than that of Bac1 but much smaller

than that of Arc1. The mutants had k_{cat}/K_m values that are smaller than that of Bac1. Thus, unlike thermostability, the catalytic efficiency was not improved by the combination of the two or three amino acid substitutions.

Combination of S108D or G116A substitution with other substitutions

As mentioned above, the S108D and G116A substitutions had the most positive impact on the thermostability of Bac1. It is reasonable to expect that multiple amino acid substitutions at sites that are close to each other are more likely to show a synergistic effect than those that are far apart. In the tertiary structures of Bac1 and Arc1, the amino acid residue at position 108 is located at the subunit boundary (Fig 5) and is in close proximity to the amino acid at position 30 of a neighboring subunit and the amino acid at position 107 in the same subunit (Fig 6A). We therefore constructed a mutant protein in which two amino acid substitutions (F30L and M107L substitutions) were combined with the S108D substitution. The T_m value of the resulting F30L/M107L/S108D mutant was 4˚C higher than that of the S108D mutant at pH 7.0 (S5 Fig). When the F30L substitution was individually introduced into Bac1, its T_m was increased by only 1˚C. Moreover, the M107L substitution decreased the T_m of Bac1 by 1˚C. Thus, although the individual contributions of the F30L and M107L substitutions to the improvement in thermostability were small or negative, they showed a synergetic effect with the S108D substitution at pH 7.6.

In the structure of Bac1, Gly116 is located on a β-strand and surrounded by residues 84, 88, 101, and 114 (Fig 6B). In Arc1, the side chain of Ala116 is pointed toward the interior core. The same amino acids occupy positions 84, 101 and 114 between the two ancestral proteins. However, an isoleucine occupies position 88 in Bac1, whereas a valine occupies this position in Arc1. We therefore created one more mutant by introducing the I88V substitution into the G116A mutant. However, the resulting I88V/G116A mutant showed the same T_m value as that of G116A at pH 7.0 (S5 Fig). In other words, regardless of whether residue 116 was Gly or Ala, the I88V substitution did not improve the thermostability of the proteins.

Combinatorial effects of amino acid substitutions whose individual contributions to the thermostability are small

While three of the amino acid substitutions (S108D, G116A and L120P substitutions) substantially enhanced the thermostability of Bac1, the individual effects of the remaining nine amino
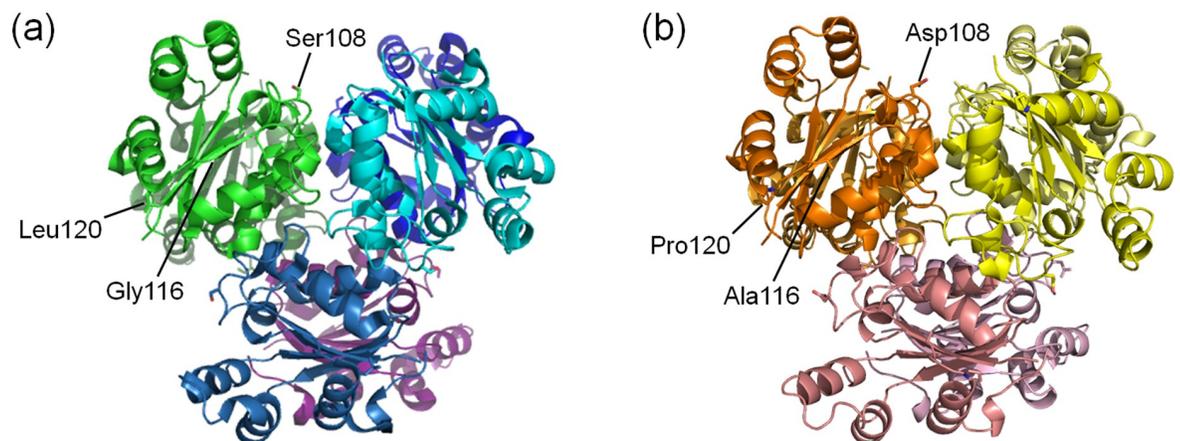

Fig 5. Hexameric structures of Bac1 (a) and Arc1 (b). The subunits of the proteins are each colored differently. Residues at positions 108, 116 and 120 in a subunit of each hexamer are indicated.
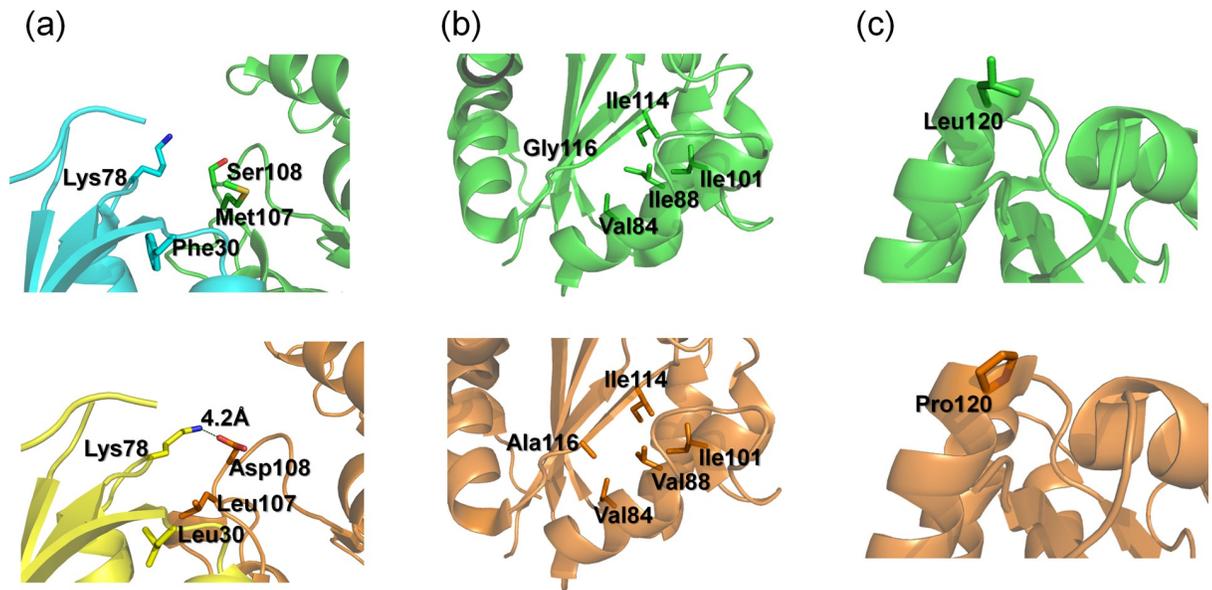
https://doi.org/10.1371/journal.pone.0258821.g005

Fig 6. (a) Local structures around position 108 of Bac1 (upper panel) and Arc1 (lower panel). Two subunits are colored differently. In Arc1, the carboxyl group of Asp108 forms a middle-range (4.2 Å) ion-pair with the amino group of Lys78 in the adjacent subunit. (b) Structures around position 116. In Bac1 (upper panel), Gly116 is located within a hydrophobic core and is surrounded by aliphatic residues including Val84, Ile88, Ile101 and Ile114. In Arc1 (lower panel), Val and Ala occupy positions 88 and 116, respectively. (c) Leu120 in Bac1 (upper panel) and Pro120 in Arc1 (lower panel) are each located at the N-terminal end of an α-helix.

acid substitutions on the thermostability of Bac1 were small. Therefore, we initially predicted that the remaining nine amino acid substitutions would not improve the thermostability of Bac1 even if the nine amino acid substitutions were introduced into Bac1 together. To test this prediction, we replaced residues 30, 37, 42, 44, 60, 64, 80, 88 and 107 in Bac1 by the amino acids found at the same positions in Arc1. The resulting mutant, named Bac1-mu9, was over-expressed in *E. coli* and then used for temperature-induced unfolding measurements at pH 6.0, pH 7.0, pH 7.6 and pH 8.0 (Fig 4 and S2 Table). Contrary to the expectation, the T_m value of Bac1-mu9 was greater than that of Bac1 by 6˚C at pH 6.0 and pH 7.0, by 7˚C at pH 7.6, and by 4˚C at pH 8.0. At pH 7.0 in particular, the magnitude of the increase in T_m by the combination of amino acid substitutions was greater than the sum of the magnitude of the increase or decrease in T_m by the individual amino acid substitutions.

The kinetic parameters of Bac1-mu9 at 70˚C are presented in S2 Table. The K_m value of Bac1-mu9 was more unfavorable than that of Arc1, although the difference is within the margin of error. Bac1-mu9 had a greater k_{cat} value than that of Bac1, but its k_{cat}/K_m value is smaller than that of Bac1.

Discussion

Several methods for creating more thermostable proteins have been proposed. Two conventional methods to create proteins with enhanced thermostability are rational design [9,14,26,27] and directed evolution [3,28–30]. As alternative approaches, consensus design and ancestral sequence reconstruction (ASR) have also been developed to create thermostable proteins. Consensus design has been used in conjunction with multiple amino acid sequence alignments (MSAs) of homologous proteins [31–33]. Its theoretical basis is that amino acids that contribute to a protein's stability have a higher probability of being selected during the evolutionary process than those that do not contribute to stability [34]. ASR is a way to design

amino acid sequences that are predicted to have been possessed by extinct species [35–39]. In conjunction with the topology of a phylogenetic tree and MSAs of homologous proteins, ASR has also created many thermostable proteins. The design of thermostable proteins by ASR assumes that the ancestral organisms were (hyper)thermophilic and therefore ancestral residues would be responsible for the thermostability of a protein to a much greater extent than non-ancestral residues [40]. Many studies have validated the reliability of consensus design and ASR as methods to create highly thermostable proteins [22,41–48], which could serve as industrially useful tools and good models for examining the structural mechanisms of protein thermostability.

More stable proteins are preferred as potential scaffolds for protein engineering [49,50]. Therefore, it is an attractive challenge to further improve the thermostability of an originally thermostable protein [9,51]. The two ancestral NDKs, Bac1 and Arc1, were created in our previous ASR experiments [22]. The amino acid sequences of the two ancestral proteins are very similar, differing only by 12 out of 139 amino acids. Although Bac1 is a very thermostable protein, Arc1 is much more thermostable and its T_m is greater than that of Bac1 by 14°C at pH 7.0. Therefore, comparison of Bac1 and Arc1 can provide insight into the structural mechanisms by which originally very thermostable proteins become more stable. The difference in thermostability between a pair of homologous proteins has sometimes been discussed on the basis of protein topology and oligomerization state. High-resolution three-dimensional structures of Bac1 and Arc1 have been determined at 2.2 Å resolution by X-ray crystallography [22], showing that both proteins are hexamers composed of six identical subunits each consisting of 139 amino acid residues (Fig 5). Moreover, significant similarity is found between the backbone structures of Bac1 and Arc1 (C_α rmsd = 0.61 Å; Fig 1B). Therefore, differences in protein topology or oligomerization state are not the basis of the differences in thermostability between Bac1 and Arc1.

From the thermostability analysis of the series of mutant proteins, we found that all of the proteins, except for Arc1, showed higher T_ms at alkaline pHs than at acidic pH (Fig 2 and S2 Table). Because the isoelectric points of Arc1 and Bac1 are both 6.40, the isoelectric point may not be related to the difference in pH dependence of thermal stability. Perhaps the negative charges of acidic amino acids contribute significantly to the thermostability of Bac1 and its mutants, whereas the positive charges of basic amino acids contribute more significantly to the thermostability of Arc1. Indeed, similar to Arc1, the Q42R mutant showed a lower T_m at pH 8.0 than at pH 7.6, although the extent of the decrease in T_m was smaller than that observed for Arc1. However, the data presented here are not necessarily sufficient to explain the differences in pH-dependent thermal stability. Further experimental and/or computational studies are therefore necessary to explain more precisely the structural features that are related to the pH dependence of thermal stability.

We also found that three (positions 108, 116, 120) out of twelve positions where different amino acids are present in Bac1 and Arc1 are primarily responsible for the difference in thermostability between the two ancestral proteins. Serine108 is located on a surface loop that contacts an adjacent subunit (Figs 5A and 6A). The S108D substitution increased the T_m of Bac1 by 6°C at pH 7.6, but only by 3°C at pH 6.0. The fact that the magnitude of the improvement in thermostability by the S108D substitution is dependent on pH suggests that the negative charge of the aspartate sidechain contributes to the enhanced thermostability. In the Arc1 tertiary structure, the carboxyl group of Asp108 is 4.2 Å away from the positively charged sidechain of Lys78 in a neighboring subunit, forming a middle-range ion pair (Fig 6A). The ion pair is the only possible interaction involving the negative charge of the aspartate side chain. Since the ancestral NDK is a homo-hexamer, the presence of aspartate instead of serine at position 108 generates six more inter-subunit ion pairs. The contribution of an increased number

of inter-subunit ion-pairs to greater protein thermostability has been proposed previously [52,53]. In addition, because residues 78 and 108 are both located in flexible loops, the formation of an ion pair would make the loops more rigid, thus contributing to improved thermostability. Less flexible surface regions have also been pointed out as a key feature associated with great stability of thermophilic proteins [16].

In order to investigate whether the dissociation of subunits occurs during denaturation, we performed denaturation experiments of Bac1 at different protein concentrations. If the subunits are dissociated upon thermal denaturation, then their stability would increase as the protein concentration increases. As shown in S6 Fig, the T_m was 99°C when the protein concentration was 12 μM. In contrast, when the protein concentration was 48 μM, the T_m (98°C) was slightly smaller. Therefore, subunit dissociation associated with degeneration was unlikely to have occurred. Nevertheless, the formation of the inter-subunit ion-pairs in the natural state can enthalpically contribute to the stability of the protein if the ion-pairs are lost in the denatured state.

Another interpretation may be also possible for the molecular mechanism of enhanced thermostability by the S108D substitution. Fig 7 compares the distribution of electrostatic potential in Bac1 and Arc1 hexamers computed by the Adaptive Poisson-Boltzmann Solver (APBS) software [54]. The area around residue 108 is completely positive in Bac1, while the electrostatic potential is shifted toward neutral in Arc1. Therefore, localized charge repulsion may be occurring in Bac1. The presence of a negatively charged aspartate residue at position 108 might mitigate the repulsion between the local positive charges and thus contribute to the greater stability of Arc1 and the S108D mutant compared to Bac1.

The G116A substitution had the greatest impact on the thermostability of Bac1. According to the tertiary structure of Bac1, Gly116 is located on a β-strand and surrounded by the hydrophobic residues Val84, Val88, Ile101 and Ile114 (Fig 6B). It is well accepted that hydrophobicity within a protein's interior core is one of the major driving forces that stabilize the native structures [55]. The contribution of the hydrophobic residues involved in the formation of the hydrophobic core of a hyperthermophilic protein to its extreme thermostability has been suggested by a study using ribonuclease HII from *Thermococcus kodakaraensis* [11]. Arc1 has an alanine instead of a glycine at position 116 (Fig 6B). The additional methyl group of the alanine residue would increase the hydrophobicity of the interior core and thus enhance the thermostabilities of Arc1 and the G116A mutants of Bac1.
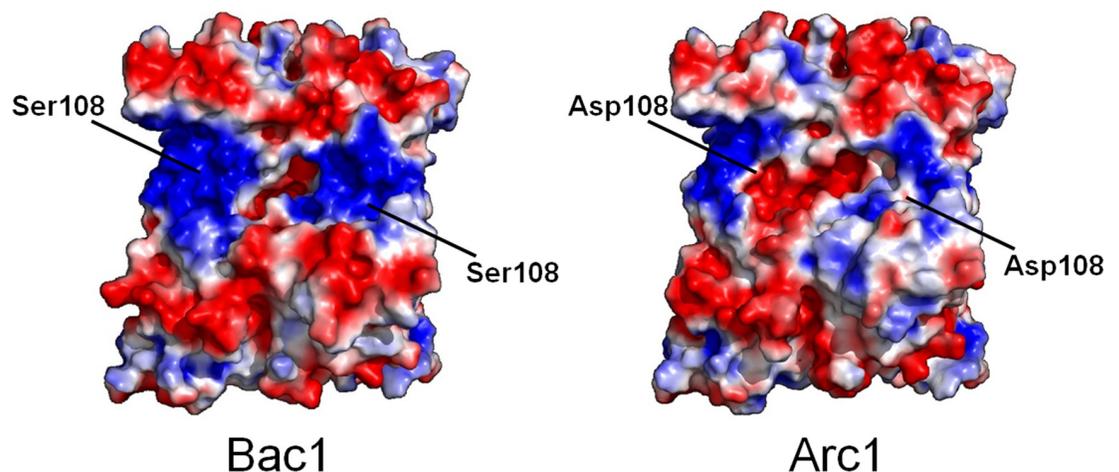

Fig 7. Electrostatic potentials on the surfaces of Bac1 and Arc1 hexamers. Electrostatic potentials were calculated and visualized using APBS [54] integrated into PyMOL (https://pymol.org/2/). Red and blue indicate negative and positive potentials, respectively. Ser108 or Asp108 are indicated in each hexameric structure.

https://doi.org/10.1371/journal.pone.0258821.g007

However, it cannot simply be expected that an increase in the interior hydrophobicity of a protein improves its thermostability. Because more hydrophobic amino acids have larger side chains, increasing the hydrophobicity of an interior core requires the introduction of larger amino acids, which often causes structural stress. It has been shown that the difference in stability between a pair of proteins can be attributed to both hydrophobic interactions and packing density with an equivalent energetic magnitude [56]. Therefore, improving a protein's stability by increasing the hydrophobicity of the interior core would be successful only when the amino acid substitutions also improve the packing density of the core. The addition of one methyl group to the interior core of Bac1 by the G116A substitution may have been the best way to improve the balance of hydrophobicity and packing density of the core.

Leucine 120 of Bac1 is located at the N-terminus of the sixth α-helix (Fig 6C). Leucine residues have a relatively high helix propensity, whereas prolines have the lowest helix propensity and are considered to be helix breakers [57]. However, it has been pointed out that the introduction of one or a few proline residue(s) into a loop region decreases the flexibility of the denatured state, making the denatured state entropically unfavorable and thus stabilizing the folded structure [58,59]. The greater thermostabilities of Arc1, which has a proline at position 120, and the L120P mutant compared to Bac1 indicate that, as suggested before [60], the presence of a proline at the end of the α-helix can contribute to enhanced protein thermostability.

In this study, we found that three amino acid residues are primarily responsible for the difference in the thermostability between Bac1 and Arc1. In order to explore whether the structural mechanisms for the great thermostability of Arc1 suggested in this study are also used in NDKs from extant hyperthermophiles, we compared the amino acid sequences of extant mesophilic, thermophilic and hyperthermophilic NDKs (S7 Fig). Table 1 shows that all NDKs from hyperthermophiles have an alanine at position 116, while all NDKs from mesophiles have a glycine. Thus, for residue 116, Arc1 has acquired high thermostability by a mechanism that is also employed by the hyperthermophilic NDKs. In contrast, the amino acid residues at positions 108 and 120 do not parallel those found in NDKs from extant hyperthermophiles; there was no tendency for the NDKs from hyperthermophiles to have aspartate and proline at positions 108 and 120, respectively. Rather, a hyperthermophilic NDK from *Sulfurisphaera tokodaii* has a serine at position 108 and some hyperthermophilic NDKs have a positively or negatively charged amino acid residue instead of a proline at position 120. In addition, a mesophilic NDK from *Dictyostelium discoideum* has an aspartate at position 108 and the *Methanothermobacter thermautotrophicus* NDK, whose T_m is 17˚C lower than that of Bac1, has a proline at position 120. Thus, the structural mechanisms that confer high thermostability on Arc1 are not always employed by the hyperthermophilic NDKs found in nature.

Conclusion

Three out of twelve amino acid residues that differ between Bac1 and Arc1 are likely to be the primary determinants for the difference in the thermostability of the two proteins. Our mutagenesis experiments on Bac1 and comparisons of the tertiary structures around the three positions of Bac1 and Arc1 suggest that the formation of inter-subunit ion pairs, elimination of electronic potential bias, improvement of hydrophobicity and packing effects of the interior core, and introduction of a proline at the terminus of an α-helix contributed to further improving the thermostability of the already thermostable Bac1. However, the remaining nine amino acid residues also contribute to the difference in thermostability between the two ancestral proteins. Therefore, a combination of many factors, some of which have very small individual effects, are responsible for the greater thermostability of Arc1 compared to Bac1. The last point

Table 1. Optimal growth temperatures (OGT) of eleven microorganisms, T_ms of their NDKs and amino acid residues at positions corresponding to those at 108, 116 and 120 in the ancestral NDKs.

Source [a]	Accession number [b]	OGT (°C)	T_m (°C) [c]	Amino acid		
				108	116	120
Pho	O58429	98	111	E	A	K
Ape	Q9Y9B0	95	108	D	A	P
Mja	Q58661	85	101	T	A	E
Afu	O29491	83	100	D	A	P
Sto	Q976A0	80	107	S	A	E
Tth	Q72GQ0	75	99	T	G	L
Mth	O26358	65	82	E	A	P
Tac	Q9HJ59	57	99	G	A	P
Bsu	P31103	37	57	F	G	L
Eco	P0A763	37	56	S	G	V
Ddi	P22887	22	62	D	G	V

[a] The eleven microorganisms are: Pho, *Pyrococcus horikoshii* (strain ATCC 700860/DSM 12428/JCM9974/NBRC 100139/OT-3); Ape, *Aeropyrum pernix* (strain ATCC 700893/DSM11879/JCM 9820/NBRC 100138/K1); Mja, *Methanocaldococcus jannaschii* (strain ATCC 43067/DSM 2661/JAL-1/JCM 10045/NBRC 100440); Afu, *Archaeoglobus fulgidus* (strain ATCC 49558/VC-16/DSM 4304/JCM 9628/NBRC 100126); Sto, *Sulfurisphaera tokodaii* (strain DSM 16993/JCM 10545/NBRC 100140/7); Tth, *Thermus thermophilus* (strain ATCC BAA-163/DSM 7039/HB27); Mth, *Methanothermobacter thermautotrophicus* (strain ATCC 29096/DSM 1053/JCM 10044/NBRC 100330/Delta H); Tac, *Thermoplasma acidophilum* (strain ATCC 25905/DSM 1728/JCM 9062/NBRC 15155/AMRC-C165); Bsu, *Bacillus subtilis* (strain 168); Eco, *E. coli* (strain K12); Ddi, *Dictyostelium discoideum* (Slime mold).

[b] The accession number for each species obtained from *UniProt* Knowledgebase (https://www.uniprot.org/).

[c] The T_m values of the extant microbial NDKs are from reference 22.

https://doi.org/10.1371/journal.pone.0258821.t001

reminds us that not only the sum of the contributions of individual amino acids, but also the synergistic effects of multiple amino acids are deeply involved in the stability of a protein.

Supporting information

S1 Fig. Ramachandran plots for the dimeric structures of Arc1 (PDB ID: 3VVT) and Bac1 (PDB ID: 3VVU) computed by the PROCHECK module of SAVES (https://saves.mbi.ucla.edu/). The most favored regions are represented by red; additional allowed regions are represented by yellow; generously allowed regions are represented by light yellow; disallowed regions are represented by white. Glycine residues are shown as triangles and the other residues are shown as squares. In each plot, only Val113 exists in disallowed regions.
(TIF)

S2 Fig. Temperature-induced unfolding curves of Arc1, Bac1 and its mutants at pH 6.0, pH 7.0, pH 7.6 and pH 8.0. The change in ellipticity at 222 nm was monitored as a function of temperature. The scan rate was 1.0°C/min. The samples comprised 20 μM protein in 20 mM potassium phosphate, 50 mM KCl, 0.5 mM EDTA. The plots were normalized with respect to the baseline of the native and unfolded states. Orange filled circles, Arc1; green filled circles, Bac1; cyan open circles, F30L; blue open circles, L37M; yellow open circles, Q42R; brown filled circles, L44M; black filled circles, G60A; magenta open circles, F64Y; gray filled circles, V80A; cyan filled circles, I88V; yellow filled circles, M107L; magenta filled circles, S108D; blue, G116A; purple filled circles, L120P.
(TIF)

S3 Fig. Atypical unfolding curves observed for G60A, S108D, S108D/G116A and S108D/G116A/L120P at pH 8.0. The raw data for the change in ellipticity at 222 nm are presented as

a function of temperature. The scan rate was 1.0˚C/min. The samples comprised 20 μM protein in 20 mM potassium phosphate (pH 8.0), 50 mM KCl, 0.5 mM EDTA.
(TIF)

S4 Fig. Temperature dependence of the specific activities of Arc1 and Bac1. The specific activities were measured using a reaction mixture consisting of 50 mM HEPES (pH 8.0), 25 mM KCl, 10 mM $(NH_4)_2SO_4$, 2.0 mM $Mg(CH_3COO)_2$, 1.0 mM dithiothreitol, 1.0 mM ADP and 2.5 mM GTP. One enzyme unit equaled 1 μmol ATP formed per min. Each value is the average of at least three replicas.
(TIF)

S5 Fig. Temperature-induced unfolding curves of Bac1 mutants with single or multiple amino acid substitution(s). The change in ellipticity at 222 nm was monitored as a function of temperature. The scan rate was 1.0˚C/min. The samples comprised 20 μM protein in 20 mM potassium phosphate (pH 7.0), 50 mM KCl, 0.5 mM EDTA. The plots were normalized with respect to the baseline of the native and denatured states. Magenta, S108D; red, F30L/M107L/S108D; blue, G116A; cyan, I88V/G116A.
(TIF)

S6 Fig. Temperature-induced unfolding curves of Bac1 at two different protein concentrations. Bac1 was dissolved in 20 mM potassium phosphate, pH 7.0, 50 mM KCl, 1 mM EDTA. The protein concentration was 12 μM (light green) or 48 μM (dark green). The scan rate was 1.0˚C/min. The plots were normalized with respect to the baseline of the native and denatured states.
(TIF)

S7 Fig. Multiple amino acid sequence alignment of eleven microbial NDKs. Positions corresponding to those at 108, 116 and 120 of the ancestral NDKs are indicated above the sequences. The sequences were aligned with MAFFT (https://mafft.cbrc.jp/alignment/server/) and visualized using ClustalX (http://www.clustal.org/clustal2/).
(TIF)

S1 Table. Mutagenic primers used in this study.
(DOCX)

S2 Table. The numerical data for T_m, K_m, k_{cat} and k_{cat}/K_m of Arc1, Bac1 and its mutants.
(DOCX)

Author Contributions

Conceptualization: Satoshi Akanuma.

Data curation: Satoshi Akanuma, Minako Yamaguchi.

Funding acquisition: Satoshi Akanuma, Akihiko Yamagishi.

Investigation: Satoshi Akanuma, Minako Yamaguchi.

Project administration: Satoshi Akanuma, Akihiko Yamagishi.

Supervision: Akihiko Yamagishi.

Writing – original draft: Satoshi Akanuma.

Writing – review & editing: Satoshi Akanuma, Akihiko Yamagishi.

References

1. Mozhaev VV, Berezin IV, Martinek K. Structure-stability relationship in proteins: fundamental tasks and strategy for the development of stabilized enzyme catalysts for biotechnology. CRC Crit Rev Biochem. 1988; 23(3):235–81. https://doi.org/10.3109/10409238809088225 PMID: 3069328

2. Amadei A, Del Galdo S, D'Abramo M. Density discriminates between thermophilic and mesophilic proteins. J Biomol Struct Dyn. 2018; 36(12):3265–73. https://doi.org/10.1080/07391102.2017.1385537 PMID: 28952426

3. Xu Z, Cen YK, Zou SP, Xue YP, Zheng YG. Recent advances in the improvement of enzyme thermostability by structure modification. Crit Rev Biotechnol. 2020; 40(1):83–98. https://doi.org/10.1080/07388551.2019.1682963 PMID: 31690132

4. Rigoldi F, Donini S, Redaelli A, Parisini E, Gautieri A. Review: Engineering of thermostable enzymes for industrial applications. APL Bioeng. 2018; 2(1):011501. https://doi.org/10.1063/1.4997367 PMID: 31069285

5. Chakravarty S, Varadarajan R. Elucidation of factors responsible for enhanced thermal stability of proteins: a structural genomics based study. Biochemistry. 2002; 41(25):8152–61. https://doi.org/10.1021/bi025523t PMID: 12069608

6. Razvi A, Scholtz JM. Lessons in stability from thermophilic proteins. Protein Sci. 2006; 15(7):1569–78. https://doi.org/10.1110/ps.062130306 PMID: 16815912

7. Tanaka T, Sawano M, Ogasahara K, Sakaguchi Y, Bagautdinov B, Katoh E, et al. Hyper-thermostability of CutA1 protein, with a denaturation temperature of nearly 150 degrees C. FEBS Lett. 2006; 580 (17):4224–30. https://doi.org/10.1016/j.febslet.2006.06.084 PMID: 16831434

8. Rahman RN, Muhd Noor ND, Ibrahim NA, Salleh AB, Basri M. Effect of ion pair on thermostability of F1 protease: integration of computational and experimental approaches. J Microbiol Biotechnol. 2012; 22 (1):34–45. https://doi.org/10.4014/jmb.1105.05055 PMID: 22297217

9. Irumagawa S, Kobayashi K, Saito Y, Miyata T, Umetsu M, Kameda T, et al. Rational thermostabilisation of four-helix bundle dimeric de novo proteins. Sci Rep. 2021; 11(1):7526. https://doi.org/10.1038/s41598-021-86952-2 PMID: 33824364

10. Yutani K, Matsuura Y, Naitow H, Joti Y. Ion-ion interactions in the denatured state contribute to the stabilization of CutA1 proteins. Sci Rep. 2018; 8(1):7613. https://doi.org/10.1038/s41598-018-25825-7 PMID: 29769700

11. Dong H, Mukaiyama A, Tadokoro T, Koga Y, Takano K, Kanaya S. Hydrophobic effect on the stability and folding of a hyperthermophilic protein. J Mol Biol. 2008; 378(1):264–72. https://doi.org/10.1016/j.jmb.2008.02.039 PMID: 18353366

12. Han M, Liao S, Peng X, Zhou X, Chen Q, Liu H. Selection and analyses of variants of a designed protein suggest importance of hydrophobicity of partially buried sidechains for protein stability at high temperatures. Protein Sci. 2019; 28(8):1437–47. https://doi.org/10.1002/pro.3643 PMID: 31074908

13. Sammond DW, Kastelowitz N, Himmel ME, Yin H, Crowley MF, Bomble YJ. Comparing Residue Clusters from Thermophilic and Mesophilic Enzymes Reveals Adaptive Mechanisms. PLoS One. 2016; 11 (1):e0145848. https://doi.org/10.1371/journal.pone.0145848 PMID: 26741367

14. Kimura N, Mochizuki K, Umezawa K, Hecht MH, Arai R. Hyperstable De Novo Protein with a Dimeric Bisecting Topology. ACS Synth Biol. 2020; 9(2):254–9. https://doi.org/10.1021/acssynbio.9b00501 PMID: 31951376

15. Nagi AD, Regan L. An inverse correlation between loop length and stability in a four-helix-bundle protein. Fold Des. 1997; 2(1):67–75. https://doi.org/10.1016/S1359-0278(97)00007-2 PMID: 9080200

16. Ang DL, Hoque MZ, Hossain MA, Guerriero G, Berni R, Hausman JF, et al. Computational Analysis of Thermal Adaptation in Extremophilic Chitinases: The Achilles' Heel in Protein Structure and Industrial Utilization. Molecules. 2021; 26(3). https://doi.org/10.3390/molecules26030707 PMID: 33572971

17. Nakka M, Iyer RB, Bachas LG. Intersubunit disulfide interactions play a critical role in maintaining the thermostability of glucose-6-phosphate dehydrogenase from the hyperthermophilic bacterium Aquifex aeolicus. Protein J. 2006; 25(1):17–21. https://doi.org/10.1007/s10930-006-0015-3 PMID: 16721657

18. Pang J, Allemann RK. Molecular dynamics simulation of thermal unfolding of Thermatoga maritima DHFR. Phys Chem Chem Phys. 2007; 9(6):711–8. https://doi.org/10.1039/b611210b PMID: 17268682

19. Liu Z, Lemmonds S, Huang J, Tyagi M, Hong L, Jain N. Entropic contribution to enhanced thermal stability in the thermostable P450 CYP119. Proc Natl Acad Sci U S A. 2018; 115(43):E10049–e58. https://doi.org/10.1073/pnas.1807473115 PMID: 30297413

20. Okafor CD, Pathak MC, Fagan CE, Bauer NC, Cole MF, Gaucher EA, et al. Structural and Dynamics Comparison of Thermostability in Ancient, Modern, and Consensus Elongation Factor Tus. Structure. 2018; 26(1):118-29.e3. https://doi.org/10.1016/j.str.2017.11.018 PMID: 29276038

21. Panja AS, Bandopadhyay B, Maiti S. Protein Thermostability Is Owing to Their Preferences to Non-Polar Smaller Volume Amino Acids, Variations in Residual Physico-Chemical Properties and More Salt-Bridges. PLoS One. 2015; 10(7):e0131495. https://doi.org/10.1371/journal.pone.0131495 PMID: 26177372

22. Akanuma S, Nakajima Y, Yokobori S, Kimura M, Nemoto N, Mase T, et al. Experimental evidence for the thermophilicity of ancestral life. Proc Natl Acad Sci U S A. 2013; 110(27):11067–72. https://doi.org/10.1073/pnas.1308215110 PMID: 23776221

23. Horton RM, Ho SN, Pullen JK, Hunt HD, Cai Z, Pease LR. Gene splicing by overlap extension. Methods Enzymol. 1993; 217:270–9. https://doi.org/10.1016/0076-6879(93)17067-f PMID: 8474334

24. Pace CN, Vajdos F, Fee L, Grimsley G, Gray T. How to measure and predict the molar absorption coefficient of a protein. Protein Sci. 1995; 4(11):2411–23. https://doi.org/10.1002/pro.5560041120 PMID: 8563639

25. Gill SC, von Hippel PH. Calculation of protein extinction coefficients from amino acid sequence data. Anal Biochem. 1989; 182(2):319–26. https://doi.org/10.1016/0003-2697(89)90602-7 PMID: 2610349

26. Ulmer KM. Protein engineering. Science. 1983; 219(4585):666–71. https://doi.org/10.1126/science.6572017 PMID: 6572017

27. Bommarius AS, Paye MF. Stabilizing biocatalysts. Chem Soc Rev. 2013; 42(15):6534–65. https://doi.org/10.1039/c3cs60137d PMID: 23807146

28. Arnold FH, Wintrode PL, Miyazaki K, Gershenson A. How enzymes adapt: lessons from directed evolution. Trends Biochem Sci. 2001; 26(2):100–6. https://doi.org/10.1016/s0968-0004(00)01755-2 PMID: 11166567

29. Arnold FH, Volkov AA. Directed evolution of biocatalysts. Curr Opin Chem Biol. 1999; 3(1):54–9. https://doi.org/10.1016/s1367-5931(99)80010-6 PMID: 10021399

30. Wang M, Si T, Zhao H. Biocatalyst development by directed evolution. Bioresour Technol. 2012; 115:117–25. https://doi.org/10.1016/j.biortech.2012.01.054 PMID: 22310212

31. Steipe B. Consensus-based engineering of protein stability: from intrabodies to thermostable enzymes. Methods Enzymol. 2004; 388:176–86. https://doi.org/10.1016/S0076-6879(04)88016-9 PMID: 15289071

32. Porebski BT, Buckle AM. Consensus protein design. Protein Eng Des Sel. 2016; 29(7):245–51. https://doi.org/10.1093/protein/gzw015 PMID: 27274091

33. Sternke M, Tripp KW, Barrick D. The use of consensus sequence information to engineer stability and activity in proteins. Methods Enzymol. 2020; 643:149–79. https://doi.org/10.1016/bs.mie.2020.06.001 PMID: 32896279

34. Steipe B, Schiller B, Pluckthun A, Steinbacher S. Sequence statistics reliably predict stabilizing mutations in a protein domain. J Mol Biol. 1994; 240(3):188–92. https://doi.org/10.1006/jmbi.1994.1434 PMID: 8028003

35. Thornton JW. Resurrecting ancient genes: experimental analysis of extinct molecules. Nat Rev Genet. 2004; 5(5):366–75. https://doi.org/10.1038/nrg1324 PMID: 15143319

36. Gaucher EA, Kratzer JT, Randall RN. Deep phylogeny—how a tree can help characterize early life on Earth. Cold Spring Harb Perspect Biol. 2010; 2(1):a002238. https://doi.org/10.1101/cshperspect.a002238 PMID: 20182607

37. Merkl R, Sterner R. Ancestral protein reconstruction: techniques and applications. Biol Chem. 2016; 397(1):1–21. https://doi.org/10.1515/hsz-2015-0158 PMID: 26351909

38. Akanuma S. Characterization of Reconstructed Ancestral Proteins Suggests a Change in Temperature of the Ancient Biosphere. Life. 2017; 7(3). https://doi.org/10.3390/life7030033 PMID: 28783077

39. Wheeler LC, Lim SA, Marqusee S, Harms MJ. The thermostability and specificity of ancient proteins. Curr Opin Struct Biol. 2016; 38:37–43. https://doi.org/10.1016/j.sbi.2016.05.015 PMID: 27288744

40. Akanuma S, Iwami S, Yokoi T, Nakamura N, Watanabe H, Yokobori S, et al. Phylogeny-based design of a B-subunit of DNA gyrase and its ATPase domain using a small set of homologous amino acid sequences. J Mol Biol. 2011; 412(2):212–25. https://doi.org/10.1016/j.jmb.2011.07.042 PMID: 21819994

41. Nikolova PV, Henckel J, Lane DP, Fersht AR. Semirational design of active tumor suppressor p53 DNA binding domain with enhanced stability. Proc Natl Acad Sci U S A. 1998; 95(25):14675–80. https://doi.org/10.1073/pnas.95.25.14675 PMID: 9843948

42. Lehmann M, Kostrewa D, Wyss M, Brugger R, D'Arcy A, Pasamontes L, et al. From DNA sequence to improved functionality: using protein sequence comparisons to rapidly design a thermostable consensus phytase. Protein Eng. 2000; 13(1):49–57. https://doi.org/10.1093/protein/13.1.49 PMID: 10679530

43. Bloom JD, Glassman MJ. Inferring stabilizing mutations from protein phylogenies: application to influenza hemagglutinin. PLoS Comput Biol. 2009; 5(4):e1000349. https://doi.org/10.1371/journal.pcbi.1000349 PMID: 19381264

44. Jäckel C, Bloom JD, Kast P, Arnold FH, Hilvert D. Consensus protein design without phylogenetic bias. J Mol Biol. 2010; 399(4):541–6. https://doi.org/10.1016/j.jmb.2010.04.039 PMID: 20433850

45. Gaucher EA, Thomson JM, Burgan MF, Benner SA. Inferring the palaeoenvironment of ancient bacteria on the basis of resurrected proteins. Nature. 2003; 425(6955):285–8. https://doi.org/10.1038/nature01977 PMID: 13679914

46. Butzin NC, Lapierre P, Green AG, Swithers KS, Gogarten JP, Noll KM. Reconstructed ancestral Myo-inositol-3-phosphate synthases indicate that ancestors of the Thermococcales and Thermotoga species were more thermophilic than their descendants. PLoS One. 2013; 8(12):e84300. https://doi.org/10.1371/journal.pone.0084300 PMID: 24391933

47. Akanuma S, Yokobori S, Nakajima Y, Bessho M, Yamagishi A. Robustness of predictions of extremely thermally stable proteins in ancient organisms. Evolution. 2015; 69(11):2954–62. https://doi.org/10.1111/evo.12779 PMID: 26404857

48. Busch F, Rajendran C, Heyn K, Schlee S, Merkl R, Sterner R. Ancestral Tryptophan Synthase Reveals Functional Sophistication of Primordial Enzyme Complexes. Cell Chem Biol. 2016; 23(6):709–15. https://doi.org/10.1016/j.chembiol.2016.05.009 PMID: 27291401

49. Finch AJ, Kim JR. Thermophilic Proteins as Versatile Scaffolds for Protein Engineering. Microorganisms. 2018; 6(4). https://doi.org/10.3390/microorganisms6040097 PMID: 30257429

50. Trudeau DL, Tawfik DS. Protein engineers turned evolutionists-the quest for the optimal starting point. Curr Opin Biotechnol. 2019; 60:46–52. https://doi.org/10.1016/j.copbio.2018.12.002 PMID: 30611116

51. Kotsuka T, Akanuma S, Tomuro M, Yamagishi A, Oshima T. Further stabilization of 3-isopropylmalate dehydrogenase of an extreme thermophile, Thermus thermophilus, by a suppressor mutation method. J Bacteriol. 1996; 178(3):723–7. https://doi.org/10.1128/jb.178.3.723-727.1996 PMID: 8550506

52. Cheung YY, Lam SY, Chu WK, Allen MD, Bycroft M, Wong KB. Crystal structure of a hyperthermophilic archaeal acylphosphatase from Pyrococcus horikoshii—structural insights into enzymatic catalysis, thermostability, and dimerization. Biochemistry. 2005; 44(12):4601–11. https://doi.org/10.1021/bi047832k PMID: 15779887

53. Karlström M, Stokke R, Steen IH, Birkeland NK, Ladenstein R. Isocitrate dehydrogenase from the hyperthermophile Aeropyrum pernix: X-ray structure analysis of a ternary enzyme-substrate complex and thermal stability. J Mol Biol. 2005; 345(3):559–77. https://doi.org/10.1016/j.jmb.2004.10.025 PMID: 15581899

54. Jurrus E, Engel D, Star K, Monson K, Brandi J, Felberg LE, et al. Improvements to the APBS biomolecular solvation software suite. Protein Sci. 2018; 27(1):112–28. https://doi.org/10.1002/pro.3280 PMID: 28836357

55. Gromiha MM, Pathak MC, Saraboji K, Ortlund EA, Gaucher EA. Hydrophobic environment is a key factor for the stability of thermophilic proteins. Proteins. 2013; 81(4):715–21. https://doi.org/10.1002/prot.24232 PMID: 23319168

56. Font J, Benito A, Torrent J, Lange R, Ribó M, Vilanova M. Pressure- and temperature-induced unfolding studies: thermodynamics of core hydrophobicity and packing of ribonuclease A. Biol Chem. 2006; 387(3):285–96. https://doi.org/10.1515/BC.2006.038 PMID: 16542150

57. Pace CN, Scholtz JM. A helix propensity scale based on experimental studies of peptides and proteins. Biophys J. 1998; 75(1):422–7. https://doi.org/10.1016/s0006-3495(98)77529-0 PMID: 9649402

58. Matthews BW, Nicholson H, Becktel WJ. Enhanced protein thermostability from site-directed mutations that decrease the entropy of unfolding. Proc Natl Acad Sci U S A. 1987; 84(19):6663–7. https://doi.org/10.1073/pnas.84.19.6663 PMID: 3477797

59. Kheirollahi A, Khajeh K, Golestani A. Rigidifying flexible sites: An approach to improve stability of chondroitinase ABC I. Int J Biol Macromol. 2017; 97:270–8. https://doi.org/10.1016/j.ijbiomac.2017.01.027 PMID: 28082225

60. Takano K, Higashi R, Okada J, Mukaiyama A, Tadokoro T, Koga Y, et al. Proline effect on the thermostability and slow unfolding of a hyperthermophilic protein. J Biochem. 2009; 145(1):79–85. https://doi.org/10.1093/jb/mvn144 PMID: 18977771