



OPEN

Ancestral sequence reconstruction produces thermally stable enzymes with mesophilic enzyme-like catalytic properties

Ryutaro Furukawa, Wakako Toma, Koji Yamazaki & Satoshi Akanuma

Enzymes have high catalytic efficiency and low environmental impact, and are therefore potentially useful tools for various industrial processes. Crucially, however, natural enzymes do not always have the properties required for specific processes. It may be necessary, therefore, to design, engineer, and evolve enzymes with properties that are not found in natural enzymes. In particular, the creation of enzymes that are thermally stable and catalytically active at low temperature is desirable for processes involving both high and low temperatures. In the current study, we designed two ancestral sequences of 3-isopropylmalate dehydrogenase by an ancestral sequence reconstruction technique based on a phylogenetic analysis of extant homologous amino acid sequences. Genes encoding the designed sequences were artificially synthesized and expressed in *Escherichia coli*. The reconstructed enzymes were found to be slightly more thermally stable than the extant thermophilic homologue from *Thermus thermophilus*. Moreover, they had considerably higher low-temperature catalytic activity as compared with the *T. thermophilus* enzyme. Detailed analyses of their temperature-dependent specific activities and kinetic properties showed that the reconstructed enzymes have catalytic properties similar to those of mesophilic homologues. Collectively, our study demonstrates that ancestral sequence reconstruction can produce a thermally stable enzyme with catalytic properties adapted to low-temperature reactions.

Enzymes, the proteinaceous biocatalysts produced by living organisms, have various superior properties relative to the inorganic catalysts that are often used in the chemical industry^{1–3}. The advantages of enzymes include their (1) high rate of catalytic reaction, (2) high enantioselectivity that allows almost 100% discrimination between mirror-image isomers, and (3) low environmental impact owing to their biodegradability. However, enzymes with required properties are not always found in nature because naturally-occurring enzymes have adapted to the host's specific growth environment and also the competing selection pressures exerted by the cellular environment over the course of evolution^{4,5}. For instance, enzymes derived from thermophilic organisms have excellent heat resistance, but often show only very low activity at lower temperatures. By contrast, cold-active enzymes isolated from psychrophilic organisms are often less stable but catalytically more active at low temperatures compared with their mesophilic and thermophilic counterparts^{6–11}, although certain cold-active enzymes in nature are also very stable^{12,13}. Cold-active enzymes are useful for low temperature processes that preserve thermally instable compounds. In addition, the use of cold-active enzymes in industrial processes sometimes results in lower energy costs^{9–11}. For effective use, it is often desirable to create enzymes that have both high thermal stability and strong catalytic activity over a wide range of temperatures, including low and moderate temperatures. At present, however, we do not understand the effects of amino acid substitutions remote from the active site, many of which influence activity and thermostability. Therefore, effective methodology is needed to create thermally stable enzymes with efficient catalytic activity at moderate or even low temperatures without relying on enzyme-design principles^{14,15}.

A conventional way to design enzymes with improved properties is the “rational design method,” in which the amino acid sequence is modified based on principles underlying the stability and function of proteins, and detailed tertiary structural information for the target enzyme^{16,17}. Nevertheless, modification of an enzyme's

Faculty of Human Sciences, Waseda University, 2-579-15 Mikajima, Tokorozawa, Saitama 359-1192, Japan. ✉email: akanuma@waseda.jp

properties remains a challenging task because we still do not have a comprehensive understanding of how proteins physically adapt to high or low temperatures. Directed evolution is another typical way to produce modified enzymes^{18,19}. An advantage of this approach is that modified biomolecules with desired properties can be feasibly obtained in the absence of either information on tertiary structure or design principles. Meanwhile, exhaustive investigation of the diversified gene population (library) is needed to select the desired property, a process that often involves considerable time and labor.

In ancestral sequence reconstruction (ASR), inferred ancestral amino acid sequences are created by using homologous amino acid sequences as input information^{20–26}. ASR can be a very powerful tool for accessing amino acid substitutions remote from the active site, which influence activity and thermostability^{27,28}. This method does not require knowledge about the three-dimensional structure or design principles of the protein. In addition, public databases contain a growing amount of protein sequences, thereby providing a necessary resource for studying ancestral sequence inference. The design method consists of three steps: (1) inference of an ancestral sequence based on a comparison of homologous amino acid sequences; (2) artificial synthesis of a gene encoding the inferred amino acid sequence; and (3) expressing the gene in a host organism such as *Escherichia coli*. Proteins reconstructed in this way are often highly thermostable^{29–32}. Moreover, by substituting putative ancestral amino acids into natural enzymes, the thermostability of a natural protein is highly likely to be increased without compromising the catalytic activity^{33–37}.

The enzyme 3-isopropylmalate dehydrogenase (IPMDH) is involved in the leucine biosynthetic pathway and encoded by the *leuB* gene. The three-dimensional structure of IPMDH from several species has been determined and, with the exception of enzymes from some hyperthermophilic species³⁸, IPMDH exists as a homodimer consisting of two identical subunits^{39,40}. As is true for many other enzymes, IPMDH shows a trade-off between thermostability and low-temperature activity⁴¹. Thermophilic IPMDHs have high thermal stability and low activity at low temperatures, whereas mesophilic enzymes are less thermally stable but highly active at moderate temperatures. IPMDHs are potentially useful for processes involving regeneration of NADH from NAD⁺.

Prior to this study, many mutant enzymes with enhanced catalytic activity at low temperatures had already been obtained from a thermostable IPMDH from the thermophilic bacterium, *Thermus thermophilus* by directed evolution. These studies showed that the catalytic activity at low temperature can be improved without largely affecting the thermal stability^{42–44}. Because we wanted to test a protein design approach other than directed evolution, we applied ASR to IPMDH. An ancestral IPMDH corresponding to the last common ancestor of *Bacillus* species has been reconstructed by others⁴⁵. In the present study, we built a phylogenetic tree of extant IPMDHs and their evolutionarily related enzymes, and designed two amino acid sequences of ancestral IPMDH based on the node of the bacterial common ancestor. The catalytic activity and thermostability of the two ancestral IPMDHs were analyzed, showing that these enzymes are highly thermally stable relative to a thermophilic IPMDH and have catalytic properties similar to those of mesophilic IPMDHs. Thus, this work exemplifies how an ancestral design method can create enzymes that simultaneously have thermophilic enzyme-like thermal stability and mesophilic enzyme-like catalytic properties.

Results

Inference of ancestral IPMDH sequences. As the first step in the reconstruction of ancestral sequences of IPMDH, we generated a multiple sequence alignment of 594 IPMDHs and evolutionarily related proteins from extant species. Using IQ-TREE⁴⁶ with a model of rate homogeneity, we built a maximum likelihood tree (Fig. 1; Fig. S1) that was rooted between archaeal isocitrate dehydrogenases and bacterial isocitrate dehydrogenases. The tree shows that IPMDHs and homoisocitrate dehydrogenases diverged from bacterial isocitrate dehydrogenases. Archaeal IPMDHs diverged earlier from bacterial isocitrate dehydrogenases relative to homoisocitrate dehydrogenases and major bacterial IPMDHs, consistent with the previously published IPMDH tree⁴¹. After diverging from archaeal IPMDHs, homoisocitrate dehydrogenases diverged multiple times. The complex group of IPMDHs subsequently diverged from homoisocitrate dehydrogenases, and included archaeal IPMDHs, bacterial IPMDHs, and tartrate dehydrogenase. Lastly, the bacterial IPMDHs group diverged to form major bacterial IPMDHs and thermophilic bacterial IPMDHs.

Two different ancestral sequences at the node corresponding to the last common ancestor of major bacterial IPMDHs were inferred from the tree by using IQ-TREE⁴⁶ and CodeML in PAML⁴⁷, and named ancIPMDH-IQ and ancIPMDH-ML, respectively (Fig. 2; Table S1). The two ancestral IPMDHs share 96.4% amino acid sequence identity. Comparison of the amino acid sequences of ancIPMDH-IQ and ancIPMDH-ML with that of *T. thermophilus* IPMDH showed that, respectively, 64.0% and 63.4% of 322 aligned residues are identical (Fig. 2). When the sequences of ancIPMDH-IQ and ancIPMDH-ML were compared with that of *Bacillus subtilis* IPMDH, 63.7% and 62.9% of 361 aligned residues are identical, respectively (Fig. 2). The ancestral IPMDH sequences have 60.3%, 60.7% and 59.8% amino acid sequence identity with psychrophilic IPMDHs from *Shewanella violacea*, *Pseudomonas haloplanktis* and *Psychrobacter cryohalolentis*, respectively. Genes encoding the ancIPMDH-IQ and ancIPMDH-ML amino acid sequences were artificially synthesized and expressed in *E. coli* using a Novagen pET protein expression system, and the two ancestral enzymes were purified by successive chromatography with HiTrap-Butyl and ResourceQ columns and then characterized.

Oligomeric structure of the ancestral IPMDHs. Although most naturally-occurring IPMDHs exist as homodimers consisting of two identical subunits^{39,40}, homotetrameric IPMDHs are observed in some hyperthermophilic species³⁸. We therefore investigated the oligomeric state of the ancestral enzymes by using analytical gel filtration (Fig. 3). The elution profile showed that ancIPMDH-IQ migrates as a single molecular species with a retention volume corresponding to the molecular weight expected for a dimer. The elution profile of

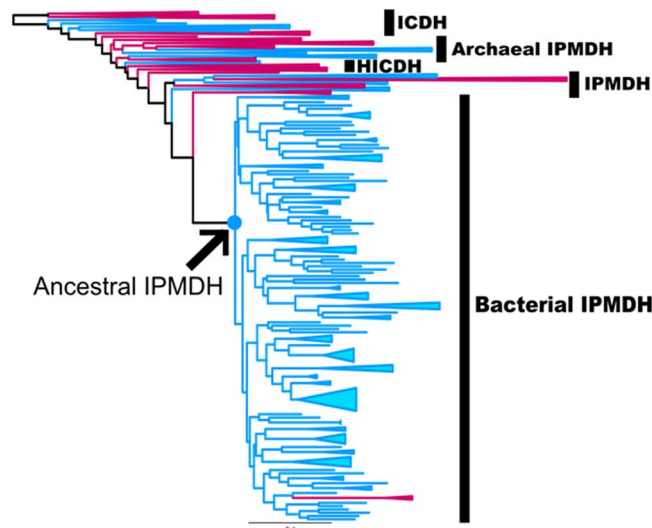


Figure 1. Phylogenetic tree used to infer ancestral IPMDH sequences. The arrow marks the node corresponding to the position of the ancestral IPMDH proteins. Red branches indicate archaeal sequences; blue branches indicate bacterial sequences. For the complete tree, see Fig. S1. The scale bar represents 0.4 substitutions per site.

ancIPMDH-ML suggested that this enzyme is predominantly a dimer, but a small elution peak corresponding to a tetramer was also observed.

Thermal stability. The thermal stability of ancIPMDH-IQ and ancIPMDH-ML was estimated by circular dichroism (CD) measurement. We monitored the change in ellipticity at 222 nm for a 5 μ M protein solution in 20 mM potassium phosphate buffer (pH 7.6), 1 mM EDTA. The temperature-induced unfolding curves shown in Fig. 4 were normalized by assuming a linear temperature dependence between the baselines of the fully native and the fully unfolded states. The unfolding curves of IPMDHs from the thermophilic bacterium *T. thermophilus*, the mesophilic bacterium *B. subtilis*, and the eukaryotic microorganism *Saccharomyces cerevisiae*, obtained in previous studies⁴¹, were included in Fig. 4 for comparison.

The unfolding mid-point temperature (T_m) was used to compare the thermal stability of the proteins (Table 1). The T_m values of ancIPMDH-IQ and ancIPMDH-ML were 88 and 90 $^{\circ}$ C, respectively, which were slightly higher than that (86 $^{\circ}$ C) of *T. thermophilus* IPMDH. Figure 4 also shows that the thermal stabilities of the two ancestral enzymes were much higher than those of extant mesophilic IPMDHs from *B. subtilis* ($T_m = 48$ $^{\circ}$ C) and *S. cerevisiae* ($T_m = 43$ $^{\circ}$ C).

Temperature dependence of specific activity. We measured the enzymatic activity of the ancestral IPMDHs by assessing the oxidative decarboxylation of D-3-isopropylmalate (D-3-IPM) to produce 2-oxoisocaproate, using NAD^+ as the electron acceptor. Figure 5a shows the temperature dependence of the specific activities of the ancestral IPMDHs, as well as those of IPMDHs from thermophilic *T. thermophilus* and two mesophilic species (*B. subtilis* and *S. cerevisiae*) for comparison. Although the ancestral IPMDHs were slightly less catalytically active relative to the mesophilic IPMDHs, they were more active than the thermophilic IPMDH at temperatures ranging from 25 to 50 $^{\circ}$ C. Remarkably, the specific activities of ancIPMDH-IQ and ancIPMDH-ML were, respectively, 3.1- and 2.1-fold greater than that of *T. thermophilus* IPMDH at 25 $^{\circ}$ C (Table 1). At 70 $^{\circ}$ C, the specificity activities of the ancestral IPMDHs were greater than those of the mesophilic IPMDHs but smaller than that of the thermophilic enzyme.

Figure 5b shows the relative activities at various temperatures compared with the activity at the optimal temperature for each enzyme. Generally, cold-adapted enzymes maintain ~ 10 – 20% of their maximum activity at low temperatures and their thermal inactivation precedes the thermal unfolding⁴⁸. The thermophilic *T. thermophilus* IPMDH functioned optimally at 90 $^{\circ}$ C under the reaction condition employed and the activity was reduced to 0.5% of the maximum activity at 25 $^{\circ}$ C; by contrast, the mesophilic IPMDHs from *B. subtilis* and *S. cerevisiae* functioned optimally at 50 $^{\circ}$ C and maintained, respectively, 15% and 20% of their maximum activities at 25 $^{\circ}$ C. ancIPMDH-IQ and ancIPMDH-ML functioned optimally at 70–80 $^{\circ}$ C (Table 1) and maintained, respectively, 7% and 4% of their maximum activities at 25 $^{\circ}$ C. Thus, the ancestral IPMDHs were better adapted to low reaction temperatures compared to the thermophilic IPMDH as judged by the temperature profiles of the relative activities.

The activation energy (E_a) for the specific activity of the ancestral IPMDHs, as well as *T. thermophilus*, *B. subtilis* and *S. cerevisiae* IPMDHs, was calculated from the slope of the Arrhenius plot for the temperature range 25–70 $^{\circ}$ C (Fig. 5c). A smaller E_a value indicates that the reaction rate has less temperature dependency, and is thus at least partially associated with the higher catalytic activity of a mesophilic enzyme at low temperature relative to its thermophilic counterpart. Indeed, the mesophilic IPMDHs had smaller E_a values (*B. subtilis* IPMDH, 66 kJ/

	1		58
ancIPMDH-IQ	M-MTYKIAVLPDGGIGPEVVAEAVKVL	EAVAE-KYGLEFEFEEALVGGAAIDATGTP	LPE
ancIPMDH-ML	M-MTYKIAVLPDGGIGPEVVAEAVKVL	EAVAE-KFGLEFEFEEALVGGAAIDATGTP	LPE
<i>T. thermophilus</i> IPMDH	M---KVAVLPDGGIGPEVTEAALKVLR	ALDE-AEGLGLAYEVFPFGGAAIDAFGE	PFPE
<i>B. subtilis</i> IPMDH	M--KKRIALLPDGGIGPEVLESATDVL	KVAE-RFNHEFEFEYGLVGGAAIDHHN	PLPE
<i>S. cerevisiae</i> IPMDH	MSAPKKIVVLPGDHVQEIETAEAIKVL	KAISDVRSNVKFDNFENHLVGGAAIDATG	VPLPD
	59		117
ancIPMDH-IQ	ETLEVCKESDAVLLGAVGGPKWDNLP	PDLRPERGLLKLRLKELGFANLRPAK-V	YYPALVD
ancIPMDH-ML	ETLEVCKQSDAVLLGAVGGPKWDNLP	PDLRPERGLLKLRLKALGFANLRPAK-V	YYPALVD
<i>T. thermophilus</i> IPMDH	PTRKGVEEAEAVLLGAVGGPKWDGL	PRKIRPETGLLSLRKSQLGFANLRPAK-V	FPGLER
<i>B. subtilis</i> IPMDH	ETVAACKNADAILLGAAGGPKWDQNL	SELRPEKGLLSIRKQLDGFANLRPVK-V	FEESLSD
<i>S. cerevisiae</i> IPMDH	EALASKKADAVLLGAVGGPKWG--	TGSVRPEQGLLKIRKELQYANLRPCNF	ASDSL
	118		174
ancIPMDH-IQ	ASPLKPEVVEGVDILVVRELTGGIYFG	QPRGI---EEGNERAFDTMVTTRSEIER	IARVA
ancIPMDH-ML	ASPLKPEVVEGVDILVVRELTGGIYFG	QPRGI---DKGNERAFDTMVTTRSEIER	IARVA
<i>T. thermophilus</i> IPMDH	LSPLKEEIARGVDLIVRELTGGIYFG	EPGRM----SEAEAWNTERYSKPEVER	VARVA
<i>B. subtilis</i> IPMDH	RSPLKKEYIDNVDFVIVRELTGGIYFG	QPSKRYVNTGEQEVADTLFYKREIER	VERI
<i>S. cerevisiae</i> IPMDH	LSPIKPGFAGKTDFFVVRELVGGIYFG	KRKE---DDGDGVAWDSEQYTVPEVQ	ITRMA
	175		231
ancIPMDH-IQ	FELARKRRK--KVTSDKANVLESSQL	WREVVTE-VAKEYPDVELEHMYVDNC	AMQLVRN
ancIPMDH-ML	FELARKRRK--KVTSDKANVLESSQL	WREVVTE-VAKEYPDVELEHMYVDNC	AMQLVRN
<i>T. thermophilus</i> IPMDH	FEAARKRRK--HVVSDKANVLEVGFE	WRKTVEE-VGRGYPDVALEHQYVDAM	AMHLVRS
<i>B. subtilis</i> IPMDH	FKMAAARKG--KVTSDKANVLESSRL	WREVAED-VAQEFDPVLEHMLVDNA	AMQLIYA
<i>S. cerevisiae</i> IPMDH	AFMALQHEPPLPIWSDKANVLAASSRL	WRKTVEETIKNEFTLKVQHQLIDSA	AMILVKN
	232		284
ancIPMDH-IQ	PKQFD-VIVTENMFGDILSDAAML	TGSLGMLPSASLGD-----KVGLYEPV	HGSAPDI
ancIPMDH-ML	PRQFD-VIVTENMFGDILSDAAML	TGSLGMLPSASLGD-----KVGLYEPV	HGSAPDI
<i>T. thermophilus</i> IPMDH	PARFD-VVVVGNIFGDI	LSDLASVLPGSLGMLPSASLGR-----	GTPVFEPVHGSAPDI
<i>B. subtilis</i> IPMDH	PNQFD-VVVVGNIFGDI	LSDLASVLPGSLGMLPSASLSS-----	SGLHLFEPVHGSAPDI
<i>S. cerevisiae</i> IPMDH	PTHLNGIIIISNMFGDILSDAASV	IPGSLGMLPSASLASLPDKNTAFGL	YEPCHGSAPDL
	285		343
ancIPMDH-IQ	AGGGIANPIAATLSAAMMLRYSFGMEE	AAEAIEKAVEKVLAEGYRTADIAQGG	SK-LVST
ancIPMDH-ML	AGGGIANPIAATLSAAMMLRYSFGMEE	AAEAIEQAVEKVLAEGYRTADIAQGG	SK-LVST
<i>T. thermophilus</i> IPMDH	AGKGIANPTAATLSAAMMLHAFGLV	ELARKVEDAVAKALLETP-PPDL--	GGSA--GT
<i>B. subtilis</i> IPMDH	AGKGMANPFAATLSAAMMLRTSFGLEE	EAKAVEDAVNKVLASGKRTRDL--	ARSEEFSS
<i>S. cerevisiae</i> IPMDH	P-KNKVNPIATLSAAMMLKLSLNL	PEEGKAIEDAVKKVLDAGIRTGDL--	GGSN--ST
	344	358	
ancIPMDH-IQ	KEMGDAIVERLEKAV-----		
ancIPMDH-ML	KEMGDAIVERLESRV-----		
<i>T. thermophilus</i> IPMDH	E---AFTATVLRHLA-----		
<i>B. subtilis</i> IPMDH	Q---AITEEVKAAIMSENTISNV		
<i>S. cerevisiae</i> IPMDH	TEVGDAVAEEVKKILA-----		

Figure 2. Multiple sequence alignment of ancIPMDH-IQ, ancIPMDH-ML, *T. thermophilus* IPMDH, *B. subtilis* IPMDH, and *S. cerevisiae* IPMDH. Numbers above the sequences are those of the ancestral enzymes. Residues conserved among the five sequences are highlighted in green. V277 and H278, the two residues mutated in the ancestral sequences, producing ancIPMDH-IQ-VAHG and ancIPMDH-ML-VAHG, are boxed.

mol; *S. cerevisiae* IPMDH, 64 kJ/mol) as compared with the *T. thermophilus* enzyme (93 kJ/mol). The ancestral IPMDHs had further reduced E_a values (ancIPMDH-IQ, 50 kJ/mol; ancIPMDH-ML, 59 kJ/mol) as compared with the mesophilic IPMDHs. Overall, the temperature dependency of the specific activity of the ancestral IPMDHs resembles that of the mesophilic IPMDHs more closely than that of the thermophilic enzyme.

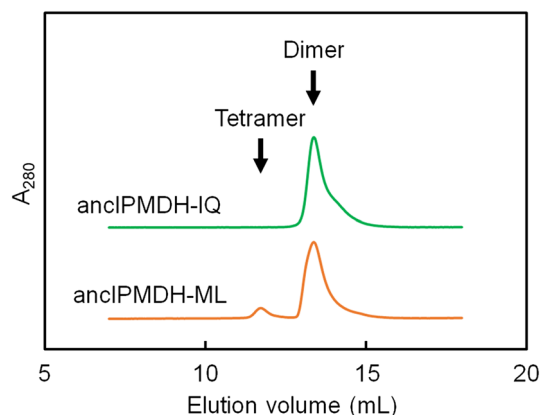


Figure 3. Analytical gel filtration using Superdex200 Increase resin. Proteins were applied at 5 μ M in 20 mM potassium phosphate, pH 7.6, 150 mM KCl, and 1 mM EDTA with a flow rate of 0.7 ml/min. Elution positions corresponding to the dimer and hexamer are indicated. A_{280} , absorbance at 280 nm.

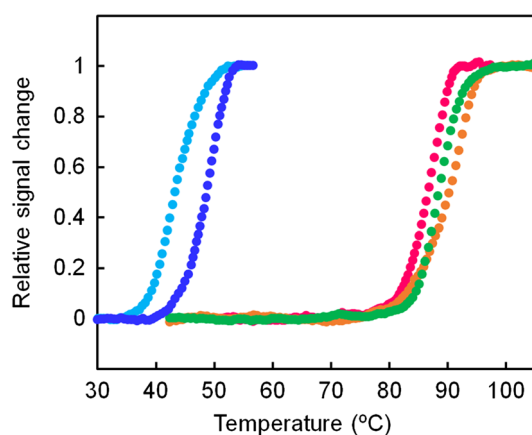


Figure 4. Thermal melting curves of the ancestral and extant IPMDHs. The change in ellipticity at 222 nm was monitored as a function of temperature. The scan rate was 1.0 $^{\circ}$ C/min. The samples comprised 5.0 μ M protein in 20 mM potassium phosphate (pH 7.6), 0.5 mM EDTA. Each experiment was conducted in duplicate with identical melting profiles within experimental error. The plots were normalized with respect to the baseline of the native and denatured states. Green, ancIPMDH-IQ; orange, ancIPMDH-ML; magenta, *T. thermophilus* IPMDH; blue, *B. subtilis* IPMDH; cyan, *S. cerevisiae* IPMDH.

Enzyme	T_m ($^{\circ}$ C) ^a	T_{opt} ($^{\circ}$ C) ^b	Specific activity (U/mg) ^c	
			25 $^{\circ}$ C	70 $^{\circ}$ C
ancIPMDH-IQ	88	75	2.76 \pm 0.02	38.5 \pm 0.5
ancIPMDH-ML	90	77	1.87 \pm 0.04	39.1 \pm 0.1
ancIPMDH-IQ-VAHG	83	67	3.56 \pm 0.13	29.6 \pm 0.5
ancIPMDH-ML-VAHG	85	68	2.74 \pm 0.10	22.4 \pm 0.4

Table 1. Unfolding midpoint temperatures, optimum catalytic temperatures and specific activities of ancIPMDH-IQ, ancIPMDH-ML and their mutant enzymes. ^a T_m values were estimated from the data shown in Figs. 2 and S5. ^b T_{opt} values are the optimum temperatures for catalytic activity estimated from the data shown in Figs. 5b and 6b. ^cAverage values and standard errors were calculated from three independent measurements.

Kinetic parameters. We estimated the kinetic parameters of the catalytic activity of the ancestral IPMDHs from steady-state experimental data obtained at 25, 40 and 70 $^{\circ}$ C (Table 2). We previously reported that changes in the specific activity of IPMDH correlate mainly to the change in K_m for the coenzyme NAD^+ and not to the change in K_m for the substrate (D-3-IPM)⁴⁴. In addition, amino acid substitutions that modulate the coenzyme-binding pocket affect the temperature dependency of the catalytic activity of *T. thermophilus* IPMDH.

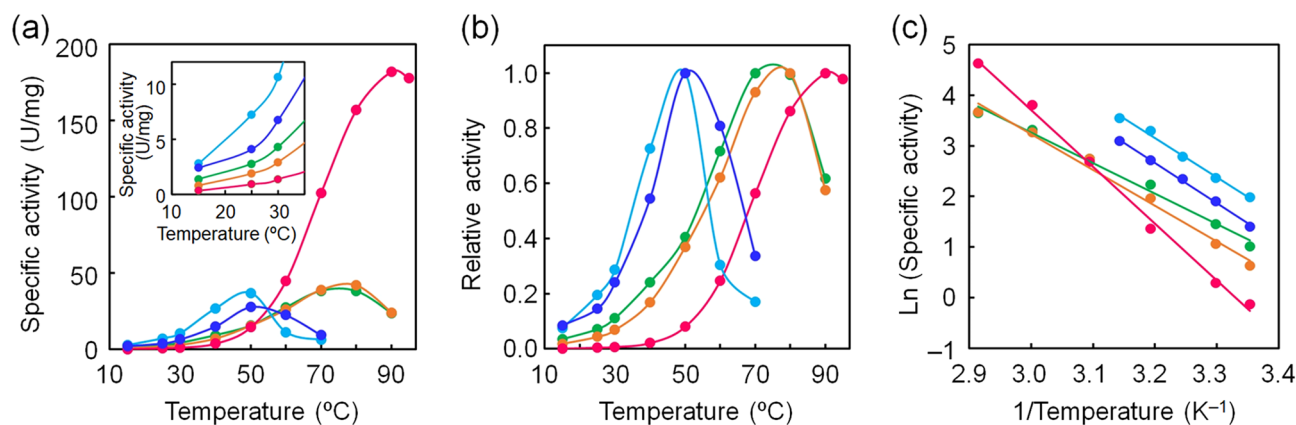


Figure 5. Specific activities of the ancestral and extant IPMDHs. **(a)** Plot of specific activity as a function of temperature. The assay solution was composed of 50 mM HEPES (pH 8.0), 100 mM KCl, 5 mM MgCl₂, 0.2 mM D-3-IPM, 5 mM NAD⁺, and 0.1–1.0 μM protein. Each value is the average of three measurements. The specific activities of both ancIPMDH-IQ and ancIPMDH-ML were greater than that of *T. thermophilus* enzymes by a factor of more than three and two, respectively, below 30 °C (inset). **(b)** Plot of relative activity as a function of temperature. Shown are the relative values of specific activity at various temperatures compared with the activity at the optimal temperature for each enzyme. **(c)** Arrhenius plot of the specific activities of the ancestral and extant IPMDHs. Activation energy (E_a) was calculated from the slope of each plot. E_a : ancIPMDH-IQ, 50 kJ/mol; ancIPMDH-ML, 59 kJ/mol; *T. thermophilus* IPMDH, 93 kJ/mol; *B. subtilis* IPMDH, 66 kJ/mol; *S. cerevisiae* IPMDH, 64 kJ/mol. Green, ancIPMDH-IQ; orange, ancIPMDH-ML; magenta, *T. thermophilus* IPMDH; blue, *B. subtilis* IPMDH; cyan, *S. cerevisiae* IPMDH.

Enzyme	$K_m^{D-3-IPM}$ (μM) ^a	K_m^{NAD} (μM) ^a	k_{cat} (s ⁻¹) ^a
25 °C			
ancIPMDH-IQ	n.d. ^b	45 ± 6	1.7 ± 0.1
ancIPMDH-ML	n.d. ^b	37 ± 4	1.2 ± 0.0
ancIPMDH-IQ-VAHG	n.d. ^b	77 ± 4	2.0 ± 0.0
ancIPMDH-ML-VAHG	n.d. ^b	33 ± 2	1.6 ± 0.0
<i>T. thermophilus</i> IPMDH	n.d. ^b	2.2 ± 0.3	0.37 ± 0.01
<i>B. subtilis</i> IPMDH	n.d. ^b	110 ± 10	2.8 ± 0.1
<i>S. cerevisiae</i> IPMDH	n.d. ^b	130 ± 10	5.1 ± 0.1
40 °C			
ancIPMDH-IQ	5.3 ± 0.2	67 ± 7	4.6 ± 0.1
ancIPMDH-ML	4.7 ± 0.4	54 ± 6	4.3 ± 0.1
ancIPMDH-IQ-VAHG	5.0 ± 0.2	230 ± 12	8.5 ± 0.1
ancIPMDH-ML-VAHG	5.0 ± 0.3	240 ± 40	5.0 ± 0.3
<i>T. thermophilus</i> IPMDH	2.3 ± 0.8	12 ± 1	2.4 ± 0.1
<i>B. subtilis</i> IPMDH	0.78 ± 0.11	420 ± 20	9.4 ± 0.2
<i>S. cerevisiae</i> IPMDH	4.5 ± 0.7	650 ± 70	20 ± 1
70 °C			
ancIPMDH-IQ	10 ± 2	960 ± 50	22 ± 0
ancIPMDH-ML	10 ± 2	730 ± 80	22 ± 1
ancIPMDH-IQ-VAHG	13 ± 3	4,100 ± 500	34 ± 2
ancIPMDH-ML-VAHG	11 ± 2	3,300 ± 400	24 ± 1
<i>T. thermophilus</i> IPMDH	3.7 ± 1.4	210 ± 10	79 ± 3

Table 2. Kinetic constants of the ancestral IPMDHs, their mutants, and extant IPMDHs. ^a K_m and k_{cat} were calculated from steady-state experiments performed at 25, 40, or 70 °C with an assay buffer containing 50 mM HEPES (pH 8.0), 100 mM KCl, 5 mM MgCl₂, and various concentrations of D-3-IPM and NAD⁺. The values and standard errors were obtained by nonlinear least-square fitting of the steady-state velocity to the Michaelis–Menten equation using the Enzyme Kinetics module of SigmaPlot version 13.0, from Systat Software, Inc., San Jose California USA, www.systatsoftware.com. ^b K_m values for D-3-IPM at 25 °C could not be determined because appropriately designed steady-state kinetic experiments would include reactions that use very small amounts of D-3-IPM, which would be consumed quickly and preclude accurate velocity measurements.

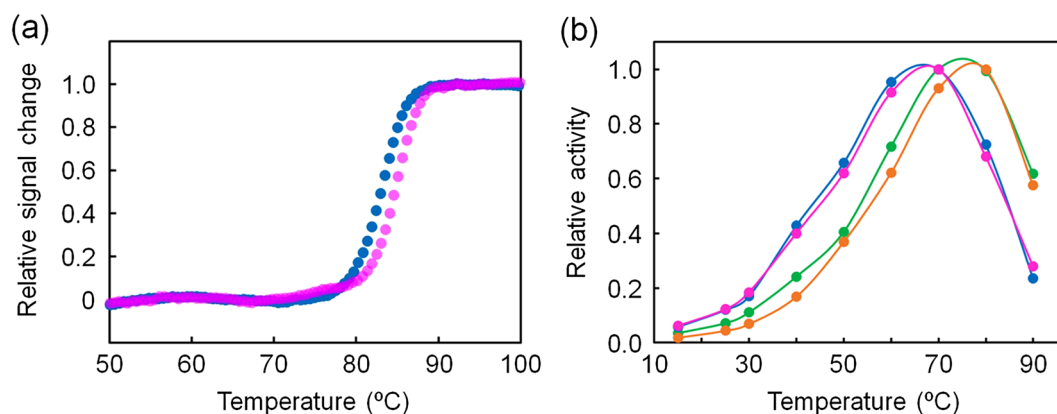


Figure 6. Characterization of the mutants of the ancestral IPMDHs. **(a)** Thermal melting curves of ancIPMDH-IQ-VAHG (blue) and ancIPMDH-ML-VAHG (magenta). The change in ellipticity at 222 nm was monitored as a function of temperature. The scan rate was 1.0 °C/min. The samples were comprised of 5.0 μ M protein in 20 mM potassium phosphate (pH 7.6), 0.5 mM EDTA. Each experiment was conducted in duplicate with identical melting profiles within experimental error. The plots were normalized with respect to the baseline of the native and denatured states. **(b)** Plot of relative activities of ancIPMDH-IQ (green), ancIPMDH-ML (orange), ancIPMDH-IQ-VAHG (blue) and ancIPMDH-ML-VAHG (magenta) as a function of temperature. Shown are the relative values of specific activity at various temperatures compared with the activity at the optimal temperature for each enzyme.

Indeed, although the $K_m^{D-3-IPM}$ values of the ancestral IPMDHs were less favorable than those of the thermophilic and mesophilic IPMDHs at 40 °C, the variation observed among the $K_m^{D-3-IPM}$ values was smaller than that observed among the K_m^{NAD} values of the ancestral and extant IPMDHs. In addition, the ancestral IPMDHs had significantly worse K_m^{NAD} values as compared with the *T. thermophilus* enzyme at 25, 40 and 70 °C; however, the k_{cat} values of ancIPMDH-IQ and ancIPMDH-ML were, respectively 4.6- and 3.2-fold greater than that of *T. thermophilus* IPMDH at 25 °C. At 40 °C, the ancestral IPMDHs also showed greater k_{cat} values relative to the *T. thermophilus* enzyme. At 70 °C, the k_{cat} values of the ancestral IPMDHs were only 28% of the k_{cat} value of *T. thermophilus* IPMDH. It has been reported that adaptation of an enzyme to low temperature is accompanied by less favorable K_m and improved k_{cat} values^{49–51}. Therefore, the catalytic properties of the ancestral IPMDHs are better adapted to low temperature as compared with the thermophilic enzyme.

In a comparison of the kinetic parameters of the ancestral IPMDHs with those of the mesophilic enzymes, the ancestral enzymes had smaller k_{cat} values, but had improved K_m^{NAD} values at 25 and 40 °C.

Attempt at further adaptation of ancestral IPMDHs to low-temperature reaction. We previously created *T. thermophilus* IPMDH mutants in which one or a few amino acid(s) located within a 12-Å distance from the active site was/were replaced by the residue(s) found at the same position(s) in the amino acid sequence of the *E. coli* IPMDH⁴¹. One of the mutants, which contained two amino acid substitutions, Val272 → Ala and His273 → Gly, had 7.6-fold greater specific activity at 25 °C, and a T_m value only 2 °C lower than that of the wild-type enzyme⁴¹. Both residues are located near the active site of the thermophilic wild-type enzyme: the side chain of Val272 is directly opposite the active site, whereas the side chain of His273 interacts with the adenine of NAD⁺. Because these residues are conserved in the amino acid sequences of ancIPMDH-IQ and ancIPMDH-ML (Val277 and His278; Fig. 2), we replaced Val277 and His278 with Ala and Gly, respectively, in both ancestral IPMDHs by site-directed mutagenesis to produce ancIPMDH-IQ-VAHG and ancIPMDH-ML-VAHG. The genes encoding the mutants were then expressed in *E. coli*, and the proteins were purified to homogeneity.

The structural properties of ancestral IPMDHs and their mutants were investigated by using CD spectroscopy to assess their secondary structure (Fig. S2). The far-UV CD spectra of ancIPMDH-IQ and ancIPMDH-ML as well as those of ancIPMDH-IQ-VAHG and ancIPMDH-ML-VAHG were quite similar, indicating that they have almost same amounts and types of secondary structure. Thermal unfolding experiment of the mutants showed that the T_m values of ancIPMDH-IQ-VAHG and ancIPMDH-ML-VAHG were 83 and 85 °C, respectively (Table 1; Fig. 6a). Thus, the T_m values of the mutants were slightly lower than those of the parent ancestral enzymes, but still much higher than those of the mesophilic *B. subtilis* and *S. cerevisiae* IPMDHs.

Enzymatic activity measurements indicated that the mutants had slightly improved specific activities at 25 °C (Table 1): the activities at 25 °C of ancIPMDH-IQ-VAHG (3.6 U/mg) and ancIPMDH-ML-VAHG (2.7 U/mg) were, respectively, 1.3- and 1.5-fold greater than those of ancIPMDH-IQ (2.8 U/mg) and ancIPMDH-ML (1.9 U/mg). Therefore, the extent to which the amino acid substitutions enhanced the specific activity at low temperature was substantially smaller than when the amino acid substitutions were introduced into *T. thermophilus* IPMDH.

The relative activities of ancIPMDH-IQ-VAHG and ancIPMDH-ML-VAHG as a function of temperature were compared with those of ancIPMDH-IQ and ancIPMDH-ML (Fig. 6b; Table 1), and showed that the optimal temperatures for ancIPMDH-IQ-VAHG and ancIPMDH-ML-VAHG shifted to low temperatures. In addition,

ancIPMDH-IQ-VAHG and ancIPMDH-ML-VAHG maintained greater relative activities at 15–60 °C. Therefore, the mutants were further adapted to low reaction temperatures compared to their parent ancestral IPMDHs.

The kinetic parameters of ancIPMDH-IQ-VAHG and ancIPMDH-ML-VAHG were also estimated by using steady-state kinetic data obtained at 25, 40 and 70 °C (Table 2). The $K_m^{D-3-IPM}$ values of the two mutants at 40 and 70 °C were, within the margin of error, the same as those of the parent ancestral IPMDHs. Although the K_m^{NAD} values of the mutants at 25 °C were similar to or slightly worse than those of the ancestral IPMDHs, the mutants had significantly worse K_m^{NAD} values as compared with the ancestral enzyme at 70 °C. Therefore, the substitutions unfavorably affected interaction with the coenzyme to a larger extent at higher temperatures. Lastly, the two amino acid substitutions only slightly improved the k_{cat} values of the ancestral IPMDHs at 25, 40 and 70 °C (Table 2). When the Val272 → Ala and His273 → Gly substitutions were introduced into *T. thermophilus* IPMDH, by contrast, the k_{cat} value was substantially improved by a factor of 9.5 at 25 °C⁴¹. Taken together, these data show that identical amino acid substitutions may not have the same effect in different proteins even if the proteins are homologous.

Discussion

Reconstruction of ancestral sequences based on a comparison of extant homologous protein sequences is a way to infer ancient protein sequences that were plausibly present in extinct species^{20–23}. In combination with an empirical analysis of the reconstructed proteins, this method has been previously used to characterize the physical properties of ancestral proteins, including thermal stability and substrate specificity^{29,30,52–58}. The sequence reconstruction approach has also been used to investigate the evolution of protein folds⁵⁹ and to identify key amino acid residues in a metabolic enzyme complex⁶⁰. It also allows us to deduce information on the history of life on Earth and to estimate long-term changes in the environment of the biosphere^{30,58,61}.

ASR may help to engineer proteins and enzymes that are potentially useful for industrial processes because the ancestral design may create enzymes with desired traits that are not found in naturally-occurring enzymes⁶². For example, Ito and coworkers engineered an ancestral L-amino acid oxidase by using the phylogeny-based method⁶³. Unlike naturally-occurring L-amino acid oxidases, the ancestral oxidase had broad substrate selectivity and high productivity via a prokaryotic heterologous expression system. As a result, the ancestral oxidase will be potentially useful for yielding pure D-amino acids from D,L-mixtures by enzymatic deracemization. Gillam and coworkers also engineered highly functional thermostable vertebrate CYP3 P450 ancestor by ASR⁶⁴.

Herein, we have described the production of enzymes that have thermophilic enzyme-like thermal stability and mesophilic enzyme-like catalytic properties by ASR. Although the relationship between thermal stability and low-temperature catalytic activity varies in naturally occurring enzymes, incompatibility between thermal stability and low-temperature activity are sometimes found^{41,50,65,66}. The incompatibility is often problematic in the industrial use of enzymes^{6,67}. For example, the high catalytic activities of cold-adapted enzymes at low temperatures are often suitable for food and beverage processing, but the concurrent thermolability of the cold-adapted enzymes often limits their use in such applications because the enzymes may be inactivated by small changes in temperature^{14,68}. Thus, the creation of enzymes with great thermal stability and catalytic efficiency at moderate temperature is highly desirable. Several examples have demonstrated that the incompatibility between thermal stability and moderate or even low temperature activity can be avoided^{12,13,44,69,70}. The GH42 cold-active β -galactosidase from the psychrophilic bacterium *Marinomonas* ef1 combines cold activity with unusual thermostability⁷¹. Moreover, phylogenetic analyses indicate a close relationship with thermophilic β -galactosidases, suggesting that this enzyme evolved from a thermostable scaffold. Those examples encouraged us to create enzymes that simultaneously show high activity at low temperatures and thermostability by ASR. The thermal stability of both ancestral IPMDHs designed in this study was slightly greater than that of the thermophilic IPMDH. High thermal stabilities have been similarly observed in many other reconstructed ancestral proteins^{29,55–57,72}. Moreover, the ancestral enzymes had enhanced low-temperature activities relative to the thermophilic enzyme. Therefore, our current study exemplifies how the reconstruction of ancestral sequence can serve as an effective methodology to create thermally stable enzymes with higher low-temperature activity as compared with a thermophilic enzyme although their activities are lower than those of their mesophilic homologues.

The difference in thermal stability and magnitude of catalytic activity between thermophilic and mesophilic proteins has sometimes been argued on the basis of protein topology and oligomerization state. However, change in protein topology is not the case for IPMDH because all of IPMDHs with known tertiary structure share the same protein topology. Although homotetrameric structures are observed in some hyperthermophilic IPMDH, many naturally-occurring IPMDHs self-associate as dimeric structures. Analytical gel filtration suggested that ancIPMDH-IQ and ancIPMDH-ML exist purely or predominantly as homodimers (Fig. 3), a quaternary structure found for most extant IPMDHs. Therefore, the oligomerization state seems not to be responsible for the thermal stability and catalytic properties of the ancestral enzymes. The difference in thermal stability between thermophilic and mesophilic proteins has also been argued on the basis of their amino acid compositions. It has been also shown that the unfolding temperature of a protein reflects the habitat temperature of its host^{56,73}. By comparing the amino acid sequences of a set of well-conserved proteins, Shakhnovich and coworkers found that the total content of Ile, Val, Tyr, Trp, Arg, Glu and Leu correlates well with the habitat temperature of organisms⁷⁴. Accordingly, the content of these seven types of amino acids (%IVYWREL) in a protein is thought to correlate with its thermal stability. We therefore calculated the %IVYWREL values of the ancestral and extant IPMDHs (Table S1), which showed that the thermophilic *T. thermophilus* IPMDH has a higher %IVYWREL value (42.9%) as compared with the mesophilic IPMDHs (*B. subtilis* IPMDH, 40.3%; *S. cerevisiae* IPMDH, 37.2%). The %IVYWREL values of the psychrophilic *S. violacea*, *P. haloplanktis* and *P. cryohalolentis* IPMDHs are 39.0%, 36.0% and 38.8%, respectively. Thus, this correlation roughly holds for the extant IPMDHs. The %IVYWREL values of

ancIPMDH-IQ (45.3%) and ancIPMDH-ML (44.4%) were greater than that of *T. thermophilus* IPMDH, which is consistent with their experimentally determined T_m values being higher than that of the thermophilic enzyme.

In large part, thermally stable enzymes are generally less active at moderate temperatures than are their less stable homologues^{41,50,65,66}. However, ancIPMDH-IQ and ancIPMDH-ML were slightly more thermally stable and more catalytically active at 25 °C as compared with the *T. thermophilus* IPMDH. It has been pointed out that the catalytic efficiency of an enzyme at low temperature correlates with the localized flexibility of the active site^{50,65,75,76}. In addition, we previously showed that enhancing the volume of the coenzyme-binding pocket can improve catalytic efficiency of *T. thermophilus* IPMDH at low temperatures⁴⁴. Because no crystal and/or solution structures of the ancestral IPMDHs are available, we explored structural factors contributing to the enhanced low-temperature activity of the ancestral enzymes by using the crystal structures of *T. thermophilus* IPMDH (PDB code: 1HEX)⁷⁷ as a guide. Figures S3a and S3b show that 17 residues are located in the NAD⁺-binding pocket and involved in binding to the coenzyme. These 17 residues are also conserved in the ancIPMDH-IQ and ancIPMDH-ML sequences (Fig. S3c), indicating that the ancestral and thermophilic enzymes share the same side chains in the coenzyme-binding pocket. In addition, the flanking sequences of the 17 residues involved in NAD⁺ binding are relatively well conserved among *T. thermophilus* and ancestral IPMDHs, and no increase in the number of glycines is observed in these regions of the ancestral enzymes (Fig. S3c), although a higher number of glycines would increase the backbone flexibility and therefore induce improved catalytic activity at low temperature. Therefore, the enhanced low-temperature activity of the ancestral enzymes is likely to be induced by a subtle change in backbone conformation. Our steady-state kinetic experiment showed that the ancestral IPMDHs have a more unfavorable K_m^{NAD} than does *T. thermophilus* IPMDH. It is possible that the ancestral IPMDHs have an expanded binding pocket for NAD⁺ that weakens the packing of the bound coenzyme, thereby improving their low-temperature catalytic activity.

Residues 277 and 278 in the ancestral IPMDHs (residues 272 and 273 in *T. thermophilus* IPMDH), which were mutated in ancIPMDH-IQ-VAHG and ancIPMDH-ML-VAHG, are located near the NAD⁺ binding pocket (Figs. S3a and S3b). The non-polar side chain of residue 277 is directly opposite to the NAD⁺ binding pocket but its main chain oxygen interacts with both bound D-3-IPM and NAD⁺. Moreover, the side chain of residue 278 directly interacts with the adenine of NAD⁺. Because the two mutated residues in the ancestral IPMDH mutants both have side chains of smaller volume than those that were replaced, the mutations plausibly expand the binding pocket for NAD⁺. In addition, substitution with a glycine would increase the flexibility of the peptide backbone and therefore contribute to the improved low-temperature activity and slightly worsened thermal stability.

The remaining important question involves how the ancestral sequences behave differently from the extant sequences and confer the observed differences in behavior. However, based on the experimental data provided herein, it is difficult to predict which specific interactions contribute to the smaller E_a values and high T_m values. Therefore, further structural and mechanistic analyses, including docking and molecular dynamics simulations, and even quantum mechanics/molecular mechanics methods, are necessary to reveal how ancestral enzymes achieved great thermal stability and mesophilic enzyme-like catalytic properties.

In conclusion, reconstruction of ancestral sequences led to the production of thermally stable IPMDHs whose unfolding midpoint temperatures were approximately 90 °C, as reported in previous studies^{31,55–57}. In addition, the ancestral enzymes had catalytic properties similar to those of mesophilic enzymes—namely, a smaller E_a for the catalytic reaction, a larger K_m for the coenzyme, higher specific activity, and higher catalytic turnover at low temperature^{49–51}. Therefore, the present data show that ASR may be a way to create thermally stable enzymes whose catalytic properties are better adapted to low reaction temperatures. Thus, the present work provides another example of the utility of ASR. However, the T_m of ancIPMDH-IQ and ancIPMDH-ML (88 and 90 °C) were still lower than that of IPMDH from the hyperthermophilic archaeon *Sulfolobus tokodaii* (96 °C)³⁸, although the specific activities of the ancestral enzymes (4.3 and 2.9 U/mg) were much greater than that of the hyperthermophilic enzyme at 30 °C (0.88 U/mg). In addition, the catalytic efficiency of the ancestral enzymes remained lower than those of the mesophilic IPMDHs below 50 °C. An ancestral IPMDH corresponding to the last common ancestor of *Bacillus* species, which has been reconstructed by others⁴⁵, had a T_m value of 65.3 °C, which is approximately 25 °C lower than those of ancIPMDH-IQ and ancIPMDH-ML, although its catalytic turnover at 70 °C was higher than those of our ancestral enzymes. Therefore, reconstruction of an ancestral IPMDH corresponding to another node of the tree used to infer ancIPMDH-IQ and ancIPMDH-ML sequences is likely to create enzymes that have a different balance of thermal stability and catalytic efficiency. Moreover, the application of a directed evolution technique^{18,19} to the ancestral IPMDHs designed in this study would help to further improve their thermal stability and catalytic efficiency at low temperature. Such studies will be the focus of future research.

Materials and methods

Phylogenetic analysis and ancestral sequence inference. Generally, ASR consists of the following steps. Homologous amino acid sequences of the target protein are first collected from databases and subjected to multiple sequence alignment. A phylogenetic tree is then built and the amino acid sequence of an ancestral node is inferred. The next step is reconstruction of the inferred ancestral amino acid sequence by genetic engineering: the inferred sequence is reverse-translated into a nucleotide sequence, and the resulting gene is artificially synthesized. Lastly, the inferred amino acid sequence is synthesized by expressing the gene in a host organism such as *E. coli*.

In this study, a total of 727 species, 591 bacteria and 136 archaea, were selected from species whose genome sequences were publicly available. The 591 bacterial species included 45 thermophiles, 112 mesophiles and 4 psychrophiles. The remaining 430 species were found in environmental samples and have not yet been cultured and therefore their actual growth temperatures are unknown. In general, we selected species that have been well

studied because they are likely to have a high degree of genomic sequence accuracy. The names and taxonomic classification of the selected species are listed in the Supplementary Dataset. Amino acid sequence data for the selected organisms were downloaded from the National Center for Biotechnology Information (NCBI, <https://www.ncbi.nlm.nih.gov/>) to construct an in-house database comprising all protein sequences from the 727 organisms. This in-house database was used for subsequent BLAST searches.

The amino acid sequences of 3-isopropylmalate dehydrogenase (IPMDH), homoisocitrate dehydrogenase (HICDH), and isocitrate dehydrogenase (ICDH) from *T. thermophilus* (accession numbers: AAS81211.1, AAS81354.1, and AAS81514.1, respectively), two IPMDHs from *Methanococcus maripaludis* (CAF30095.1, CAF30436.1), and HICDH from *Pyrococcus horikoshii* (BAA30836.1) were downloaded from the NCBI database. Using these six sequences as query sequences, we searched for homologous amino acid sequences in the aforementioned in-house database. As a result, 913 bacterial and 291 archaeal sequences were collected. After removing duplicate sequences, the remaining 729 sequences were aligned by using MAFFT ver.7.3⁷⁸. Next, sequences that differed significantly from the others in length were excluded and the remaining 594 sequences (435 bacterial and 159 archaeal sequences) were re-aligned by MAFFT, incorporating secondary structural information of IPMDH, HICDH and ICDH (PDBID: 1WPW, 1OSI, 1X0L, 5HN3, 2D1C, 2DHT, and 1ZOR), and then manually corrected to generate the multiple sequence alignment that was subsequently used for tree building.

Well-aligned regions were selected from the final alignment by using the automated1 mode of trimAl⁷⁹. IQ-TREE ver.1.6.9⁴⁶, in conjunction with the LG + R10 amino acid substitution model, was used to build the phylogenetic tree. The LG + R10 model was selected as the optimal amino acid substitution model by ModelFinder⁸⁰. From the resulting phylogenetic tree, we inferred ancestral sequences by using IQ-TREE with the LG + R10 model and CodeML in PAML⁴⁷ with the LG + Gamma (eight-class) model. We also used GASP⁸¹ to estimate the location of gaps in the ancestral sequences. The ancestral sequence predicted by IQ-TREE was named ancIPMDH-IQ, and that predicted by CodeML was named ancIPMDH-ML.

Construction of expression plasmids for ancestral and extant mesophilic IPMDHs. The inferred amino acid sequences of the ancestral IPMDHs were reverse-translated to generate nucleotide sequences encoding the ancestral amino acid sequences. The codon usage was optimized for an *E. coli* expression system. The nucleotide sequences were artificially synthesized by Eurofins Genomics (Tokyo) and then cloned into the *NdeI*-*BamHI* site of plasmid pET23a(+) (Merck, Tokyo). For comparison, mesophilic IPMDHs from *B. subtilis* and *S. cerevisiae* were prepared and characterized. The respective genes were cloned into the *NdeI*-*BamHI* site of plasmid pET21c(+) (Merck, Tokyo) and expressed in *E. coli*.

Site-directed mutagenesis. Genes encoding ancestral IPMDH mutants containing Val277 → Ala and His278 → Gly substitutions were amplified by using the splicing-by-overlap-extension PCR method⁸². To amplify the mutagenized genes, the PCR reaction solution contained 1 × PCR buffer for KOD-plus DNA polymerization, 1 mM MgSO₄, dNTPs at 0.2 mM each, synthetic oligonucleotides at 0.4 μM each, 1.0 unit of KOD-plus DNA polymerase (Toyobo, Osaka), and 0.2 ng/μl of template DNA. The oligonucleotides used were 5'-GCG CTG ACC CAC CAG CCG GTT CAT ATA AG-3' and 5'-CTT ATA TGA ACC GGC TGG GTC AGC GC-3' for the mutagenesis of ancIPMDH-IQ, and 5'-GCG CAC TGC CAC CAG CGG GCT CAT AG-3' and 5'-CTA TGA GCC CGC TGG CAG TGC GC-3' for the mutagenesis of ancIPMDH-ML. The following time-temperature program was used: step 1, 95 °C, 3 min; step 2, 95 °C, 30 s; step 3, 55 °C, 30 s; step 4, 68 °C, 1 min; and steps 2–4 were repeated 25 times. The amplified DNAs were then digested with *NdeI* and *BamHI* (New England Biolabs, Tokyo) and cloned into the *NdeI*-*BamHI* site of pET23a(+) (Merck, Tokyo).

Enzyme purification. To produce the ancestral and extant mesophilic IPMDHs, *E. coli* Rosetta2 (DE3) was transformed with the respective expression plasmids. Each transformant was cultured in Luria-Bertani medium supplemented with 150 μg/mL of ampicillin. Overnight Express Autoinduction system 1 reagent (Merck, Tokyo) was used to induce gene expression. This induction system allows the induction of protein expression without the need to monitor cell density during culturing and without adding a conventional induction reagent such as isopropyl β-D-1-thiogalactopyranoside. After cultivation overnight at 37 °C (ancestral IPMDHs) or 30 °C (mesophilic IPMDHs), the cells were harvested by centrifugation, and then disrupted by sonication followed by centrifugation at 60,000 × g for 20 min. To purify the ancestral IPMDHs, the supernatants were individually heat-treated at 70 °C for 20 min to precipitate proteins originating from *E. coli*. After centrifugation again at 60,000 × g for 20 min, the supernatants were successively passed through HiTrap-Butyl and ResourceQ (Cytiva, Tokyo). To purify the mesophilic IPMDHs, the soluble fractions of the cell lysates were successively passed through HiTrap-Q, HiTrap-Butyl and ResourceQ (Cytiva, Tokyo). The purity of the enzymes used in this study was >95% as judged by the results of SDS-polyacrylamide gel electrophoresis coupled with Coomassie Blue staining.

The concentrations of protein solutions were determined from the OD₂₈₀ values as described by Pace and colleagues⁸³, who modified the procedure described by Gill and von Hippel⁸⁴.

Analytical gel filtration. The oligomeric state of the ancestral IPMDHs was determined by analytical gel filtration using a Superdex200 Increase column (Cytiva, Tokyo). Aliquots of protein (0.1 ml) were loaded onto the column and elution was performed at 25 °C with a flow rate of 0.7 ml/min and 20 mM potassium phosphate, pH 7.6, 150 mM KCl, 1 mM EDTA. The eluent was monitored by absorbance at 280 nm.

Thermal stability measurement. CD measurements were carried out using a J-1100 spectropolarimeter (Jasco, Hachioji). The enzymes were dissolved in 20 mM potassium phosphate buffer (pH 7.6), 1 mM

EDTA at a concentration of 5 μM and placed in a sample cell with a 0.1-cm path length. The temperature of the enzyme solution was increased at a rate of 1.0 $^{\circ}\text{C}/\text{min}$ by using a programmable temperature controller. Thermal denaturation was observed by monitoring the change in ellipticity at 222 nm, which reflects the abundance of α -helices in a protein.

Enzyme activity measurements. Specific activity was measured at various temperatures in an assay solution comprised of 50 mM HEPES (pH 8.0), 100 mM KCl, 5 mM MgCl_2 , 0.2 mM D-3-IPM and 5.0 mM NAD^+ by monitoring the increase in absorbance at 340 nm, which reflects the generation of NADH, a product of the reaction catalyzed by IPMDH. One enzyme unit was defined as the formation of 1 μmol of NADH per min.

Michaelis constant (K_m) values for D-3-IPM, the substrate of IPMDH, were determined in steady-state experiments with an assay solution containing 50 mM HEPES (pH 8.0), 100 mM KCl, 5 mM MgCl_2 , 5 mM NAD^+ , and various concentrations of D-3-IPM. To determine the values of K_m for the coenzyme NAD^+ and k_{cat} , the concentration of NAD^+ was varied while the D-3-IPM concentration was fixed (0.2 mM). Protein concentrations were 20 nM (25 $^{\circ}\text{C}$), 6.0 nM (40 $^{\circ}\text{C}$) or 2.0 nM (70 $^{\circ}\text{C}$). To obtain the kinetic parameters, the observed steady-state velocity data were fitted to the Michaelis–Menten equation by using the Enzyme Kinetics module of SigmaPlot version 13.0, from Systat Software, Inc., San Jose California USA, www.systatsoftware.com.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Received: 16 June 2020; Accepted: 31 August 2020

Published online: 23 September 2020

References

- Chapman, J., Ismail, A. E. & Dinu, C. Z. Industrial applications of enzymes: Recent advances, techniques, and outlooks. *Catalysts* **8**, 238. <https://doi.org/10.3390/catal8060238> (2018).
- Choi, J. M., Han, S. S. & Kim, H. S. Industrial applications of enzyme biocatalysis: Current status and future aspects. *Biotechnol. Adv.* **33**, 1443–1454. <https://doi.org/10.1016/j.biotechadv.2015.02.014> (2015).
- Kirk, O., Borchert, T. V. & Fuglsang, C. C. Industrial enzyme applications. *Curr. Opin. Biotechnol.* **13**, 345–351. [https://doi.org/10.1016/s0958-1669\(02\)00328-2](https://doi.org/10.1016/s0958-1669(02)00328-2) (2002).
- Jemli, S., Ayadi-Zouari, D., Hlima, H. B. & Bejar, S. Biocatalysts: Application and engineering for industrial purposes. *Crit. Rev. Biotechnol.* **36**, 246–258. <https://doi.org/10.3109/07388551.2014.950550> (2016).
- Prasad, S. & Roy, I. Converting enzymes into tools of industrial importance. *Recent Pat. Biotechnol.* **12**, 33–56. <https://doi.org/10.2174/1872208311666170612113303> (2018).
- Sarmiento, F., Peralta, R. & Blamey, J. M. Cold and hot extremozymes: Industrial relevance and current trends. *Front. Bioeng. Biotechnol.* **3**, 148. <https://doi.org/10.3389/fbioe.2015.00148> (2015).
- Siddiqui, K. S. Some like it hot, some like it cold: Temperature dependent biotechnological applications and improvements in extremophilic enzymes. *Biotechnol. Adv.* **33**, 1912–1922. <https://doi.org/10.1016/j.biotechadv.2015.11.001> (2015).
- Aqvist, J. Cold adaptation of triosephosphate isomerase. *Biochemistry* **56**, 4169–4176. <https://doi.org/10.1021/acs.biochem.7b00523> (2017).
- Santiago, M., Ramirez-Sarmiento, C. A., Zamora, R. A. & Parra, L. P. Discovery, molecular mechanisms, and industrial applications of cold-active enzymes. *Front. Microbiol.* **7**, 1408. <https://doi.org/10.3389/fmicb.2016.01408> (2016).
- Collins, T. & Margesin, R. Psychrophilic lifestyles: Mechanisms of adaptation and biotechnological tools. *Appl. Microbiol. Biotechnol.* **103**, 2857–2871. <https://doi.org/10.1007/s00253-019-09659-5> (2019).
- Mangiagalli, M., Brocca, S., Orlando, M. & Lotti, M. The “cold revolution”. Present and future applications of cold-active enzymes and ice-binding proteins. *N Biotechnol.* **55**, 5–11. <https://doi.org/10.1016/j.nbt.2019.09.003> (2020).
- Fedoy, A. E., Yang, N., Martinez, A., Leiros, H. K. & Steen, I. H. Structural and functional properties of isocitrate dehydrogenase from the psychrophilic bacterium *Desulfotalea psychrophila* reveal a cold-active enzyme with an unusual high thermal stability. *J. Mol. Biol.* **372**, 130–149. <https://doi.org/10.1016/j.jmb.2007.06.040> (2007).
- Leiros, H. K. *et al.* Structure of phenylalanine hydroxylase from *Colwellia psychrerythraea* 34H, a monomeric cold active enzyme with local flexibility around the active site and high overall stability. *J. Biol. Chem.* **282**, 21973–21986. <https://doi.org/10.1074/jbc.M610174200> (2007).
- Yang, G. *et al.* Rational engineering of a cold-adapted α -amylase from the Antarctic ciliate *Euplotes focardii* for simultaneous improvement of thermostability and catalytic activity. *Appl. Environ. Microbiol.* **83**, e00449–e517. <https://doi.org/10.1128/aem.00449-17> (2017).
- Rigoldi, F., Donini, S., Redaelli, A., Parisini, E. & Gautieri, A. Review: Engineering of thermostable enzymes for industrial applications. *APL Bioeng.* **2**, 011501. <https://doi.org/10.1063/1.4997367> (2018).
- Bommarius, A. S. & Paye, M. F. Stabilizing biocatalysts. *Chem. Soc. Rev.* **42**, 6534–6565. <https://doi.org/10.1039/c3cs60137d> (2013).
- Ulmer, K. M. Protein engineering. *Science* **219**, 666–671 (1983).
- Chen, K. & Arnold, F. H. Engineering new catalytic activities in enzymes. *Nat. Catal.* **3**, 203–213. <https://doi.org/10.1038/s41929-019-0385-5> (2020).
- Arnold, F. H. & Moore, J. C. Optimizing industrial enzymes by directed evolution. *Adv. Biochem. Eng. Biotechnol.* **58**, 1–14. <https://doi.org/10.1007/bfb0103300> (1997).
- Thornton, J. W. Resurrecting ancient genes: Experimental analysis of extinct molecules. *Nat. Rev. Genet.* **5**, 366–375. <https://doi.org/10.1038/nrg1324> (2004).
- Gaucher, E. A., Kratzer, J. T. & Randall, R. N. Deep phylogeny—How a tree can help characterize early life on Earth. *Cold Spring Harb. Perspect. Biol.* **2**, a002238. <https://doi.org/10.1101/cshperspect.a002238> (2010).
- Merkel, R. & Sterner, R. Ancestral protein reconstruction: Techniques and applications. *Biol. Chem.* **397**, 1–21. <https://doi.org/10.1515/hsz-2015-0158> (2016).
- Akanuma, S. Characterization of reconstructed ancestral proteins suggests a change in temperature of the ancient biosphere. *Life (Basel)* **7**, 33. <https://doi.org/10.3390/life7030033> (2017).
- Akanuma, S. & Yamagishi, A. A strategy for designing thermostable enzymes by reconstructing ancestral sequences possessed by ancient life in *Biotechnology of Extremophiles: Advances and Challenges* (ed. Rampelotto, P. H.) 581–596 (Springer, 2016).
- Wheeler, L. C., Lim, S. A., Marqusee, S. & Harms, M. J. The thermostability and specificity of ancient proteins. *Curr. Opin. Struct. Biol.* **38**, 37–43. <https://doi.org/10.1016/j.sbi.2016.05.015> (2016).

26. Rouet, R. *et al.* Structural reconstruction of protein ancestry. *Proc. Natl. Acad. Sci. USA* **114**, 3897–3902. <https://doi.org/10.1073/pnas.1613477114> (2017).
27. Wilding, M., Hong, N., Spence, M., Buckle, A. M. & Jackson, C. J. Protein engineering: The potential of remote mutations. *Biochem. Soc. Trans.* **47**, 701–711. <https://doi.org/10.1042/bst20180614> (2019).
28. Wilding, M. *et al.* Reverse engineering: Transaminase biocatalyst development using ancestral sequence reconstruction. *Green Chem.* **19**, 5375–5380. <https://doi.org/10.1039/C7GC02343J> (2017).
29. Gaucher, E. A., Thomson, J. M., Burgan, M. F. & Benner, S. A. Inferring the palaeoenvironment of ancient bacteria on the basis of resurrected proteins. *Nature* **425**, 285–288. <https://doi.org/10.1038/nature01977> (2003).
30. Gaucher, E. A., Govindarajan, S. & Ganesh, O. K. Palaeotemperature trend for Precambrian life inferred from resurrected proteins. *Nature* **451**, 704–707. <https://doi.org/10.1038/nature06510> (2008).
31. Akanuma, 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.* **412**, 212–225. <https://doi.org/10.1016/j.jmb.2011.07.042> (2011).
32. Thomas, A., Cutlan, R., Finnigan, W., van der Giezen, M. & Harmer, N. Highly thermostable carboxylic acid reductases generated by ancestral sequence reconstruction. *Commun. Biol.* **2**, 429. <https://doi.org/10.1038/s42003-019-0677-y> (2019).
33. Watanabe, K., Ohkuri, T., Yokobori, S. & Yamagishi, A. Designing thermostable proteins: Ancestral mutants of 3-isopropylmalate dehydrogenase designed by using a phylogenetic tree. *J. Mol. Biol.* **355**, 664–674. <https://doi.org/10.1016/j.jmb.2005.10.011> (2006).
34. Shimizu, H. *et al.* Extremely thermophilic translation system in the common ancestor commonote: Ancestral mutants of glycyl-tRNA synthetase from the extreme thermophile *Thermus thermophilus*. *J. Mol. Biol.* **369**, 1060–1069. <https://doi.org/10.1016/j.jmb.2007.04.001> (2007).
35. Yamashiro, K., Yokobori, S., Koikeda, S. & Yamagishi, A. Improvement of *Bacillus circulans* beta-amylase activity attained using the ancestral mutation method. *Protein Eng. Des. Sel.* **23**, 519–528. <https://doi.org/10.1093/protein/gzq021> (2010).
36. Fukuda, Y. *et al.* Epistasis effects of multiple ancestral-consensus amino acid substitutions on the thermal stability of glycerol kinase from *Cellulomonas* sp. NT3060. *J. Biosci. Bioeng.* **121**, 497–502. <https://doi.org/10.1016/j.jbiosc.2015.09.011> (2016).
37. Hamuro, Y. *et al.* Characterization of a thermostable mutant of *Agaricus brasiliensis* laccase created by phylogeny-based design. *J. Biosci. Bioeng.* **124**, 623–629. <https://doi.org/10.1016/j.jbiosc.2017.07.004> (2017).
38. Sasaki, M., Uno, M., Akanuma, S. & Yamagishi, A. Random mutagenesis improves the low-temperature activity of the tetrameric 3-isopropylmalate dehydrogenase from the hyperthermophile *Sulfolobus tokodaii*. *Protein Eng. Des. Sel.* **21**, 721–727. <https://doi.org/10.1093/protein/gzn054> (2008).
39. Imada, K. *et al.* Three-dimensional structure of a highly thermostable enzyme, 3-isopropylmalate dehydrogenase of *Thermus thermophilus* at 2.2 Å resolution. *J. Mol. Biol.* **222**, 725–738 (1991).
40. Wallon, G. *et al.* Crystal structures of *Escherichia coli* and *Salmonella typhimurium* 3-isopropylmalate dehydrogenase and comparison with their thermophilic counterpart from *Thermus thermophilus*. *J. Mol. Biol.* **266**, 1016–1031. <https://doi.org/10.1006/jmbi.1996.0797> (1997).
41. Akanuma, S. *et al.* Establishment of mesophilic-like catalytic properties in a thermophilic enzyme without affecting its thermal stability. *Sci. Rep.* **9**, 9346. <https://doi.org/10.1038/s41598-019-45560-x> (2019).
42. Suzuki, T., Yasugi, M., Arisaka, F., Yamagishi, A. & Oshima, T. Adaptation of a thermophilic enzyme, 3-isopropylmalate dehydrogenase, to low temperatures. *Protein Eng.* **14**, 85–91 (2001).
43. Yasugi, M., Amino, M., Suzuki, T., Oshima, T. & Yamagishi, A. Cold adaptation of the thermophilic enzyme 3-isopropylmalate dehydrogenase. *J. Biochem.* **129**, 477–484 (2001).
44. Hayashi, S., Akanuma, S., Onuki, W., Tokunaga, C. & Yamagishi, A. Substitutions of coenzyme-binding, nonpolar residues improve the low-temperature activity of thermophilic dehydrogenases. *Biochemistry* **50**, 8583–8593. <https://doi.org/10.1021/bi200925f> (2011).
45. Hobbs, J. K. *et al.* On the origin and evolution of thermophily: Reconstruction of functional precambrian enzymes from ancestors of *Bacillus*. *Mol. Biol. Evol.* **29**, 825–835. <https://doi.org/10.1093/molbev/msr253> (2012).
46. Nguyen, L. T., Schmidt, H. A., von Haeseler, A. & Minh, B. Q. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **32**, 268–274. <https://doi.org/10.1093/molbev/msu300> (2015).
47. Yang, Z. PAML 4: Phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**, 1586–1591. <https://doi.org/10.1093/molbev/msm088> (2007).
48. Feller, G. Psychrophilic enzymes: From folding to function and biotechnology. *Scientifica (Cairo)* **512840**, 2013. <https://doi.org/10.1155/2013/512840> (2013).
49. Bjelic, S., Brandsdal, B. O. & Aqvist, J. Cold adaptation of enzyme reaction rates. *Biochemistry* **47**, 10049–10057. <https://doi.org/10.1021/bi801177k> (2008).
50. Siddiqui, K. S. & Cavicchioli, R. Cold-adapted enzymes. *Annu. Rev. Biochem.* **75**, 403–433. <https://doi.org/10.1146/annurev.biochem.75.103004.142723> (2006).
51. Smalas, A. O., Leiros, H. K., Os, V. & Willassen, N. P. Cold adapted enzymes. *Biotechnol. Annu. Rev.* **6**, 1–57 (2000).
52. Bridgham, J. T., Carroll, S. M. & Thornton, J. W. Evolution of hormone-receptor complexity by molecular exploitation. *Science* **312**, 97–101. <https://doi.org/10.1126/science.1123348> (2006).
53. Ortlund, E. A., Bridgham, J. T., Redinbo, M. R. & Thornton, J. W. Crystal structure of an ancient protein: Evolution by conformational epistasis. *Science* **317**, 1544–1548. <https://doi.org/10.1126/science.1142819> (2007).
54. Bridgham, J. T., Ortlund, E. A. & Thornton, J. W. An epistatic ratchet constrains the direction of glucocorticoid receptor evolution. *Nature* **461**, 515–519. <https://doi.org/10.1038/nature08249> (2009).
55. Butzin, N. C. *et al.* Reconstructed ancestral Myo-inositol-3-phosphate synthases indicate that ancestors of the *Thermococcales* and *Thermotoga* species were more thermophilic than their descendants. *PLoS ONE* **8**, e84300. <https://doi.org/10.1371/journal.pone.0084300> (2013).
56. Akanuma, S. *et al.* Experimental evidence for the thermophilicity of ancestral life. *Proc. Natl. Acad. Sci. USA* **110**, 11067–11072. <https://doi.org/10.1073/pnas.1308215110> (2013).
57. Akanuma, S., Yokobori, S., Nakajima, Y., Bessho, M. & Yamagishi, A. Robustness of predictions of extremely thermally stable proteins in ancient organisms. *Evolution* **69**, 2954–2962. <https://doi.org/10.1111/evo.12779> (2015).
58. Hart, K. M. *et al.* Thermodynamic system drift in protein evolution. *PLoS Biol.* **12**, e1001994. <https://doi.org/10.1371/journal.pbio.1001994> (2014).
59. Richter, M. *et al.* Computational and experimental evidence for the evolution of a (β_α)₈-barrel protein from an ancestral quarter-barrel stabilised by disulfide bonds. *J. Mol. Biol.* **398**, 763–773. <https://doi.org/10.1016/j.jmb.2010.03.057> (2010).
60. Holinski, A., Heyn, K., Merkl, R. & Sterner, R. Combining ancestral sequence reconstruction with protein design to identify an interface hotspot in a key metabolic enzyme complex. *Proteins* **85**, 312–321. <https://doi.org/10.1002/prot.25225> (2017).
61. Garcia, A. K., Schopf, J. W., Yokobori, S. I., Akanuma, S. & Yamagishi, A. Reconstructed ancestral enzymes suggest long-term cooling of Earth's photic zone since the Archean. *Proc. Natl. Acad. Sci. USA* **114**, 4619–4624. <https://doi.org/10.1073/pnas.1702729114> (2017).
62. Gumulya, Y. & Gillam, E. M. Exploring the past and the future of protein evolution with ancestral sequence reconstruction: The “retro” approach to protein engineering. *Biochem. J.* **474**, 1–19. <https://doi.org/10.1042/bcj20160507> (2017).
63. Nakano, S., Minamino, Y., Hasebe, F. & Ito, S. Deracemization and stereoinversion to aromatic D-amino acid derivatives with ancestral L-amino acid oxidase. *ACS Catal.* **9**, 10152–10158. <https://doi.org/10.1021/acscatal.9b03418> (2019).

64. Gumulya, Y. *et al.* Engineering highly functional thermostable proteins using ancestral sequence reconstruction. *Nat. Catal.* **1**, 878–888. <https://doi.org/10.1038/s41929-018-0159-5> (2018).
65. Feller, G. Protein stability and enzyme activity at extreme biological temperatures. *J. Phys. Condens Matter* **22**, 323101. <https://doi.org/10.1088/0953-8984/22/32/323101> (2010).
66. Fields, P. A. Review: Protein function at thermal extremes: Balancing stability and flexibility. *Comp. Biochem. Physiol. A Mol. Integr. Physiol.* **129**, 417–431 (2001).
67. Sterner, R. & Liebl, W. Thermophilic adaptation of proteins. *Crit. Rev. Biochem. Mol. Biol.* **36**, 39–106. <https://doi.org/10.1080/20014091074174> (2001).
68. Gerday, C. *et al.* Cold-adapted enzymes: From fundamentals to biotechnology. *Trends Biotechnol.* **18**, 103–107. [https://doi.org/10.1016/s0167-7799\(99\)01413-4](https://doi.org/10.1016/s0167-7799(99)01413-4) (2000).
69. Miller, S. R. An appraisal of the enzyme stability–activity trade-off. *Evolution* **71**, 1876–1887. <https://doi.org/10.1111/evo.13275> (2017).
70. Nguyen, V. *et al.* Evolutionary drivers of thermoadaptation in enzyme catalysis. *Science* **355**, 289–294. <https://doi.org/10.1126/science.aah3717> (2017).
71. Mangiagalli, M. *et al.* The co-existence of cold activity and thermal stability in an Antarctic GH42 β -galactosidase relies on its hexameric quaternary arrangement. *FEBS J.* <https://doi.org/10.1111/febs.15354> (2020).
72. Busch, F. *et al.* Ancestral tryptophan synthase reveals functional sophistication of primordial enzyme complexes. *Cell Chem. Biol.* **23**, 709–715. <https://doi.org/10.1016/j.chembiol.2016.05.009> (2016).
73. Gromiha, M. M., Oobatake, M. & Sarai, A. Important amino acid properties for enhanced thermostability from mesophilic to thermophilic proteins. *Biophys. Chem.* **82**, 51–67 (1999).
74. Zeldovich, K. B., Berezovsky, I. N. & Shakhnovich, E. I. Protein and DNA sequence determinants of thermophilic adaptation. *PLoS Comput. Biol.* **3**, e5. <https://doi.org/10.1371/journal.pcbi.0030005> (2007).
75. Truongvan, N., Jang, S. H. & Lee, C. Flexibility and stability trade-off in active site of cold-adapted *Pseudomonas mandelii* Esterase EstK. *Biochemistry* **55**, 3542–3549. <https://doi.org/10.1021/acs.biochem.6b00177> (2016).
76. Zavodszky, P., Kardos, J., Svingor & Petsko, G. A. Adjustment of conformational flexibility is a key event in the thermal adaptation of proteins. *Proc. Natl. Acad. Sci. USA* **95**, 7406–7411 (1998).
77. Hurley, J. H. & Dean, A. M. Structure of 3-isopropylmalate dehydrogenase in complex with NAD⁺: Ligand-induced loop closing and mechanism for cofactor specificity. *Structure* **2**, 1007–1016. [https://doi.org/10.1016/s0969-2126\(94\)00104-9](https://doi.org/10.1016/s0969-2126(94)00104-9) (1994).
78. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780. <https://doi.org/10.1093/molbev/mst010> (2013).
79. Capella-Gutierrez, S., Silla-Martinez, J. M. & Gabaldon, T. trimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**, 1972–1973. <https://doi.org/10.1093/bioinformatics/btp348> (2009).
80. Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A. & Jermini, L. S. ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nat. Methods* **14**, 587–589. <https://doi.org/10.1038/nmeth.4285> (2017).
81. Edwards, R. J. & Shields, D. C. GASP: Gapped ancestral sequence prediction for proteins. *BMC Bioinform.* **5**, 123. <https://doi.org/10.1186/1471-2105-5-123> (2004).
82. Horton, R. M. *et al.* Gene splicing by overlap extension. *Methods Enzymol.* **217**, 270–279 (1993).
83. Pace, C. N., Vajdos, F., Fee, L., Grimsley, G. & Gray, T. How to measure and predict the molar absorption coefficient of a protein. *Protein Sci.* **4**, 2411–2423. <https://doi.org/10.1002/pro.5560041120> (1995).
84. Gill, S. C. & von Hippel, P. H. Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem.* **182**, 319–326 (1989).

Acknowledgements

This work was supported by the Astrobiology Center Program of National Institutes of Natural Sciences (NINS) (Grant Number AB022003).

Author contributions

S.A. designed the study. R.F., W.T., K.Y. and S.A. conducted the experiments and the data analysis. R.F. and S.A. wrote the manuscript. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-72418-4>.

Correspondence and requests for materials should be addressed to S.A.

Reprints and permissions information is available at www.nature.com/reprints.

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



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

© The Author(s) 2020