Title

Reconstructed ancestral enzymes suggest long-term cooling of Earth's photic zone since the Archean

Journal

Proceedings of the National Academy of Sciences of the United States of America, 114(18)

ISSN

0027-8424

Authors

Garcia, Amanda K
Schopf, J William
Yokobori, Shin-Ichi
et al.

Publication Date

2017-05-02

DOI

10.1073/pnas.1702729114

Peer reviewed

Reconstructed ancestral enzymes suggest long-term cooling of Earth's photic zone since the Archean

Amanda K. Garcia[a,b,1], J. William Schopf[a,b,c,d,1], Shin-ichi Yokobori[e], Satoshi Akanuma[f], and Akihiko Yamagishi[e]

[a]Center for the Study of Evolution and the Origin of Life, University of California, Los Angeles, CA 90095; [b]Department of Earth, Planetary, and Space Sciences, University of California, Los Angeles, CA 90095; [c]Molecular Biology Institute, University of California, Los Angeles, CA 90095; [d]University of Wisconsin Astrobiology Research Consortium, Madison, WI 53706; [e]Department of Applied Life Sciences, Tokyo University of Pharmacy and Life Sciences, 1432-1 Horinouchi, Hachioji, Tokyo 192-0392, Japan; and [f]Faculty of Human Sciences, Waseda University, 2-579-15 Mikajima, Tokorozawa, Saitama 359-1192, Japan

Contributed by J. William Schopf, March 16, 2017 (sent for review February 16, 2017; reviewed by David J. Bottjer, L. Paul Knauth, and Donald R. Lowe)

Paleotemperatures inferred from the isotopic compositions ($\delta^{18}O$ and $\delta^{30}Si$) of marine cherts suggest that Earth's oceans cooled from 70 ± 15 °C in the Archean to the present ∼15 °C. This interpretation, however, has been subject to question due to uncertainties regarding oceanic isotopic compositions, diagenetic or metamorphic resetting of the isotopic record, and depositional environments. Analyses of the thermostability of reconstructed ancestral enzymes provide an independent method by which to assess the temperature history inferred from the isotopic evidence. Although previous studies have demonstrated extreme thermostability in reconstructed archaeal and bacterial proteins compatible with a hot early Earth, taxa investigated may have inhabited local thermal environments that differed significantly from average surface conditions. We here present thermostability measurements of reconstructed ancestral enzymatically active nucleoside diphosphate kinases (NDKs) derived from light-requiring prokaryotic and eukaryotic phototrophs having widely separated fossil-based divergence ages. The ancestral environmental temperatures thereby determined for these photic-zone organisms—shown in modern taxa to correlate strongly with NDK thermostability—are inferred to reflect ancient surface-environment paleotemperatures. Our results suggest that Earth's surface temperature decreased over geological time from ∼65–80 °C in the Archean, a finding consistent both with previous isotope-based and protein reconstruction-based interpretations. Interdisciplinary studies such as those reported here integrating genomic, geologic, and paleontologic data hold promise for providing new insight into the coevolution of life and environment over Earth history.

ancestral sequence reconstruction | enzyme thermostability | nucleoside diphosphate kinase | phototroph | Precambrian

Understanding of the interrelated evolution of life and environment requires knowledge both of biology and the associated physical setting, the latter traditionally provided by paleontological and geological evidence. Although in recent decades the known fossil record has been extended well into the Archean (1, 2), significant aspects of the development of Earth's environment—most notably, changes in day length and surface temperature (3)—remain to be fully explored. We here focus on the history of Earth's surface temperature, poorly resolved for the formative Precambrian seven-eighths of Earth history due to a lack of paleoclimate proxies like those available for the more recent ∼550 Ma of the Phanerozoic (e.g., ref. 4).

To date, the most complete temperature record for the Precambrian is that inferred from oxygen and silicon isotope compositions of marine cherts, cryptocrystalline siliceous rocks typically chemically precipitated from associated pore waters. $\delta^{18}O$ and $\delta^{30}Si$ measurements of such sediments from the 3,500–3,200 Ma Barberton greenstone belt register a notably hot 55–85 °C marine environment for the Archean that, coupled with measurements in younger cherts, suggest a geologically long-term decrease to the present ∼15 °C (5–7). This interpretation, however, has engendered skepticism due to uncertainties associated with possible changes in oceanic isotopic compositions (e.g., refs. 8–10), diagenetic and/or metamorphic resetting of the reported isotopic signatures by exchange with groundwater and/or recrystallization, respectively (e.g., refs. 11–14), and ambiguities regarding the depositional environment of some of the cherts analyzed that, if hydrothermal, would not be indicative of global surface conditions (e.g., ref. 13). To resolve such questions, an independent line of evidence is necessary.

In recent years, molecular biology has advanced understanding of the history of life and its environment, supplementing traditional geology- and paleontology-based approaches. Initially conceived as "chemical paleogenetics" by Pauling and Zuckerlandl (15), ancestral sequence reconstruction (ASR) estimates the monomeric sequences of ancient biomolecules by use of standard evolutionary statistics (e.g., maximum likelihood and Bayesian) applied to phylogenies constructed from the genomics of extant descendants. Based on the assumption that molecular sequence determines biological function and that the ancient molecules would have necessarily been adapted to and functional in their surroundings, the reconstructed molecules should evidence past environmental conditions.

The basal presence in molecular phylogenies of hyperthermophilic bacterial and archaeal lineages suggests an early evolved tolerance to high environmental temperatures (16). Previous ASR studies have evaluated this possibility by inferring the thermostabilities of ancestral reconstructed rRNA and protein amino acid sequences (17, 18), or by measuring the temperature-dependent functionality of experimentally "resurrected" ancient proteins (19–23). Most notably, Gaucher et al. (20) used the thermostabilities of experimentally reconstructed bacterial elongation factors to infer variations in

Significance

Geological evidence suggesting that Earth's oceans cooled from ∼55–85 °C in the Archean (∼3,500 Ma) to the present has met with skepticism due to possible geochemical alteration of the rocks analyzed and uncertainties about their depositional environment. Determination of the thermostability of experimentally reconstructed ancestral enzymes provides independent means to assess paleotemperature. Because some previously analyzed taxa may have inhabited atypically high-temperature environments (e.g., deep-sea hydrothermal vents), we have restricted our analyses to ancestral enzymes reconstructed from photic-zone cyanobacteria and land plants. Our findings indicate a cooling of Earth's surface temperature from ∼75 °C in the Archean (∼3,000 Ma) to ∼35 °C in the Devonian (∼420 Ma), consistent with previous geological and enzyme-based results.

Author contributions: J.W.S., S.-i.Y., S.A., and A.Y. designed research; A.K.G. and S.-i.Y. performed research; S.A. contributed new reagents/analytic tools; A.K.G., J.W.S., S.-i.Y., S.A., and A.Y. analyzed data; and A.K.G., J.W.S., S.-i.Y., S.A., and A.Y. wrote the paper.

Reviewers: D.J.B., University of Southern California; L.P.K., Arizona State University; and D.R.L., Stanford University.

The authors declare no conflict of interest.

[1]To whom correspondence may be addressed. Email: agarcia9@ucla.edu or schopf@ess.ucla.edu.

paleotemperature over Precambrian time, data producing a temperature trend similar to that suggested by the isotopic compositions of marine cherts (5–7) and interpreted to evidence a gradual cooling from 60 to 70 °C in the Archean.

The thermostabilities exhibited by reconstructed biomolecules can reflect the range of temperatures only of their immediate ancient surroundings (21). Thus, accurate assessment of variations over time of Earth's surface temperature requires evaluation of the reconstructed molecular constituents of widespread inhabitants of near-surface environments, thereby excluding the inclusion of data from thermophilic microorganisms prevalent in localized submarine or terrestrial hydrothermal settings. In this study, we accomplish this goal by including reconstructed ancestral proteins of phototrophs only, light-requiring organisms that, like their extant descendants, would have been restricted to photic-zone near-surface environments. Possible inhabitants of terrestrial hot springs are excluded by restricting our study to land plants and marine cyanobacteria and green algae.

The biomolecule selected to be reconstructed for this study is nucleoside diphosphate kinase (NDK), an enzyme that catalyzes the transfer of a phosphate group from nucleoside triphosphate to nucleoside diphosphate, is virtually ubiquitous among extant organisms, and the thermostability of which has been shown previously to correlate strongly with organismal growth temperature (22, 23). Thus, measurement of the thermostabilities of reconstructed ancestral NDK from cyanobacteria, green algae, and land plants, groups and subgroups of which have temporally widely spaced fossil-record-indicated divergence ages, should provide a firm basis for analysis of the history of Earth's surface temperature over geological time. Such data would afford an independent set of data by which to assess the paleotemperature trend previously inferred from mineralogical isotope-based evidence (5–7).

Results and Discussion

rRNA and NDK Phylogenies. Ancestral NDK reconstructions were targeted to represent the molecular phylogeny-determined last common ancestors of cyanobacteria (oxygenic photosynthetic prokaryotes; ∼3,100–2,700 Ma), nostocaleans (later-evolved cyanobacteria characterized by morphologically distinctive cells; ∼2,300–2,100 Ma), Viridiplantae (green algae and land plants; ∼850–700 Ma), and Embryophyta (land plants only; ∼460–440 Ma). The estimated ∼2,900 Ma age for the origin of oxygen-producing cyanobacteria precedes the earliest widespread occurrence of geochemical evidence of oxygenic photosynthesis (e.g., banded iron formations and subaerial red beds) and the subsequent ∼2,400 Ma Great Oxidation Event, the "GOE" (24, 25), suggesting that older Archean stromatolitic mat-forming microbes may have been anoxygenic phototrophs (compare ref. 26). The origin of nostocalean cyanobacteria at ∼2,100 Ma is based on microfossils interpreted to be akinetes (27, 28), resting cells characteristic of the Nostocales as are heterocysts, differentiated trichomic thick-walled cells that shielded the oxygen-labile N_2-fixing nitrogenase enzyme system from the increase of ambient oxygen at the GOE (29). The earliest strong fossil evidence of green algae places the origin of the Viridiplantae in the Late Proterozoic (30), and that of isolated trilete spores dates the divergence of land plants as perhaps as early as the Mid-Ordovician (31).

Following Akanuma et al. (22) and due to concerns that NDK sequences alone might not accurately depict evolutionary relationships, we used 16S rRNA or 18S rRNA to construct two separate phylogenetic trees, 16S for prokaryotic cyanobacteria and nostocaleans and 18S for eukaryotic Viridiplantae and Embryophyta (Fig. 1 *A* and *B*). To further evaluate the eukaryotic ancestral NDKs, we constructed a third tree (Fig. 1*C*) based on NDK from the same Viridiplantae taxa used to construct the 18S rRNA tree.

In some respects the eukaryotic 18S rRNA (Fig. 1*B*) and NDK (Fig. 1*C*) trees differ. Most such variation occurs in the clustering of angiosperm taxa, differences that are relatively insignificant due to the lack of resolution for short angiosperm branch lengths in both trees. More significant variation occurs in the topological differences among major Viridiplantae lineages. Phylogenetic relations depicted in the 18S tree that are supported by previously published plant phylogenies (e.g., ref. 32) include the monophyletic clustering of chlorophyte green algae and monocotylendonous angiosperms. These relations are not shown by the NDK tree in which the Chlorophyta is paraphyletic and monocots are split into two distinct monophyletic groups. The most significant departure of the 18S tree (Fig. 1*B*) from accepted plant phylogenies is the occurrence of a gymnosperm sequence within the angiosperms rather than being a sister branch as in the NDK tree (Fig. 1*C*). The unsupported placement of this deep-branching sequence within the Embryophyta may be a source of significant sequence inaccuracy for our reconstruction of the Embryo18S NDK.

Ancestral NDK Sequences. To calculate the ancestral amino acid sequences of the NDKs to be reconstructed, we used a maximum likelihood (ML) algorithm shown to yield reliable reconstructions and thermostability inferences comparable or even preferable to more computationally expensive Bayesian methods (21, 23, 33, 34), specifically CODEML of the Phylogenetic Analysis by Maximum Likelihood (PAML) program package that incorporates a homogeneous amino acid substitution model (35). Ancestral NDK sequences were estimated using prokaryotic and eukaryotic NDK and rRNA reference trees. The amino acid sequences of six ancestral NDKs were determined, two from ancestral cyanobacterial and nostocalean nodes within the 16S rRNA prokaryote tree (NDKs referred to here, respectively, as "Cyano16S" and "Nosto16S"); two from ancestral Viridiplantae and Embryophyta nodes of the 18S rRNA eukaryote tree ("Viridi18S" and "Embryo18S"); and two from the eukaryotic NDK tree ("ViridiNDK" and "EmbryoNDK").

Experimental NDK Thermostability Measurements. Genes coding for the six ancestral NDKs were synthesized and cloned into expression vectors, expressed in *Escherichia coli*, and purified. NDKs from extant cyanobacteria (*Nostoc punctiforme*, *Synechocystis sp. PCC6803*, and *Thermosynechococcus elongatus*), a green alga (*Chlamydomonas reinhardtii*), and three land plants (*Physcomitrella patens, Oryza sativa,* and *Picea sitchensis*) were also constructed to provide taxon-specific calibration of NDK thermostability and organismal environmental temperature (compare ref. 22).

Table 1 lists the measured thermostability of the reconstructed NDK enzymes. The temperature-induced denaturation both of reconstructed and extant NDKs was monitored by circular dichroism (CD) spectroscopy at pH 6.0 and 7.6, their thermostability being defined as the midpoint unfolding temperature (T_m) evidenced by their two-state denaturation curves. Although all NDKs produced clear two-state curves at pH 6.0, several yielded atypical denaturation curves at pH 7.6 (Table 1) that nevertheless exhibited CD signal shifts coinciding within experimental error at the T_ms measured for the same NDK at pH 6.0.

The plausibility of the reconstructed enzymes accurately representing their original ancestral configuration can be assessed by comparison of the thermostability of NDKs reconstructed from the same ancestral node in topologically differing phylogenetic trees. For example, the Viridiplantae ancestral NDKs calculated from the 18S (Fig. 1*B*) and NDK (Fig. 1*C*) trees, Viridi18S and ViridiNDK, coincide at all but 8 of 152 amino acid residues and have essentially identical T_m values (81.7 ± 0.4 and 81.2 ± 0.5 °C, respectively; Table 1), notable similarities that support the accuracy of their reconstruction despite the differences in tree topologies used to determine their sequences. Such similarities, however, are not exhibited by the Embryophyta NDKs calculated from the two eukaryotic trees, Embryo18S and EmbryoNDK, reconstructed enzymes that differ at 17 of 153 residues and have T_m values that diverge by ∼15 °C. These differences seem likely to be a

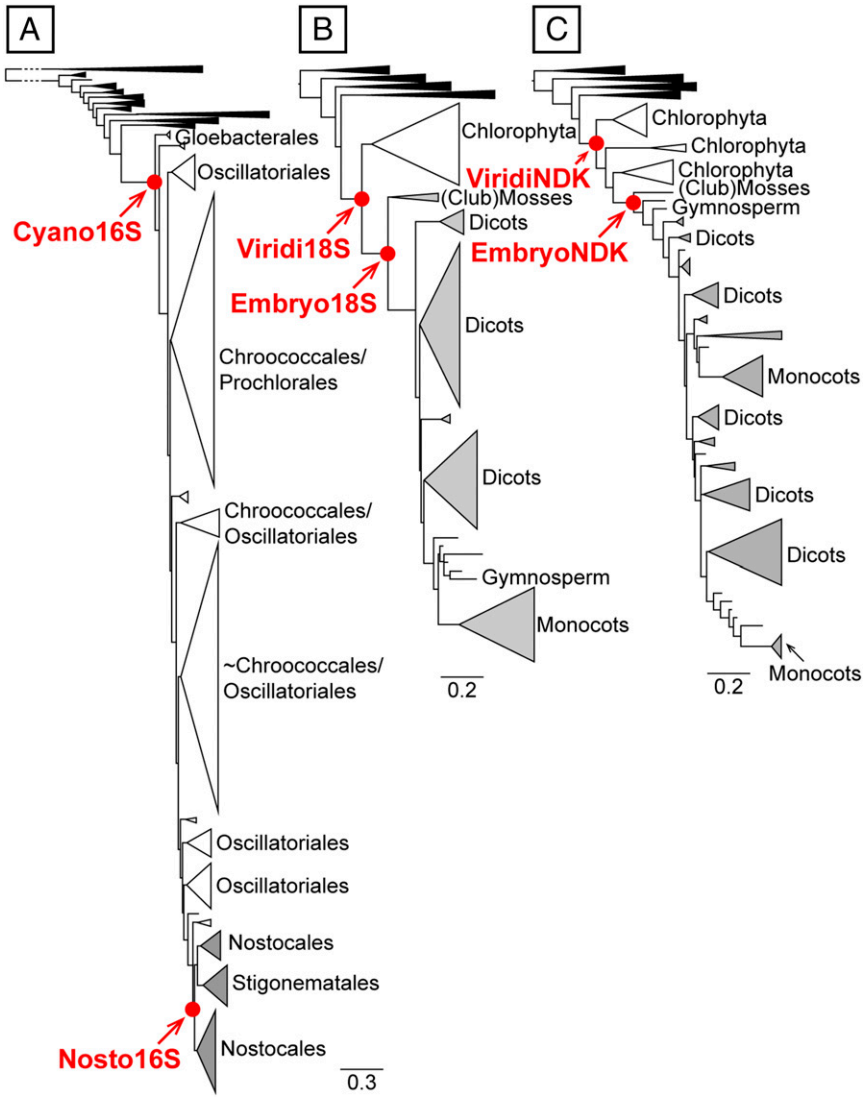

Fig. 1. Phylogenetic trees used to evaluate ancestral NDK sequences. Nodes from which ancestral NDK sequences were calculated are indicated by red arrows. (Scale bars indicate the number of substitutions per site.) (*A*) An ML 16S rRNA tree of cyanobacteria based on the NDK sequences of outgroups (archaea and bacteria, black-filled), nonnostocalean cyanobacteria (white-filled), and nostocaleans (gray-filled). (*B*) An ML 18S rRNA tree of Viridiplantae, based on the NDK sequences showing non-Viridiplantae outgroups (black-filled), nonembryophyte Viridiplantae (white-filled), and Embryophyta (gray-filled). (*C*) An ML NDK tree constructed using NDK sequences of the Viridiplantae taxa used to construct *B*, the various parts colored like those in *B*.

result of the misplacement discussed above of a deep-branching gymnosperm sequence within the Embryophyta 18S rRNA tree (Fig. 1*B*). Because this may be a significant source of sequence inaccuracy for Embryo18S, we have greater confidence that the T_m of ∼64 °C measured for EmbryoNDK more accurately represents the thermostability of ancestral Embryophyta NDK. Such disparities highlight the need for the testing of multiple phylogenetic models in ASR studies.

Table 1. Reconstructed ancestral NDK divergence ages, thermostabilities, and environmental temperature ranges inferred from Fig. 3

Reconstructed NDK enzyme	Median divergence age, Ma	T_m ±1 SD, °C		Environmental temperature range, °C
		pH 6.0	pH 7.6	
Cyano16S	∼2,900	100.1 ± 0.4	n.d.*	64–82
Nosto16S	∼2,200	77.8 ± 0.1	n.d.*	36–56
Viridi18S	∼775	81.7 ± 0.4	82.4 ± 1.2	24–61
ViridiNDK	∼775	81.2 ± 0.5	83.4 ± 0.5	24–62
Embryo18S	∼450	80.0 ± 0.3	n.d.*	24–58
EmbryoNDK	∼450	64.3 ± 0.6	n.d.*	20–41

*n.d.: T_m not determined at pH 7.6 due to atypical denaturation curves.

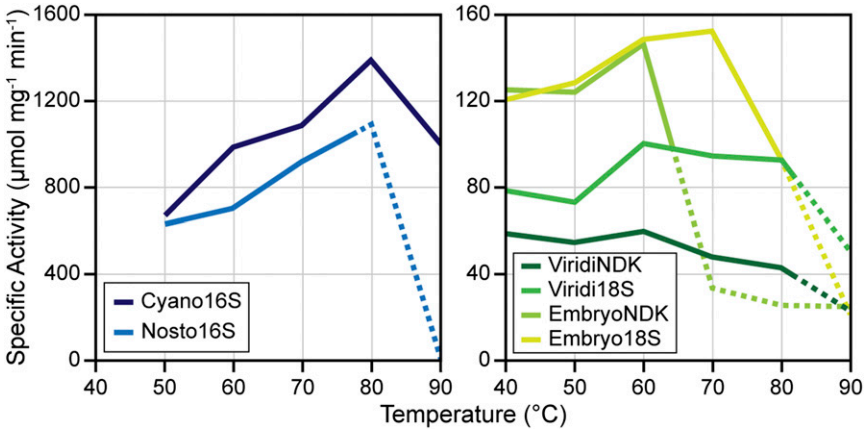

Fig. 2. Temperature dependence of ancestral NDK specific activity measured by the ATP produced, each value being the average of three or more measurements and the dotted lines spanning activity values measured at temperature intervals greater than the T_ms analyzed by circular dichroism spectroscopy. Data are not available for the specific activity of Cyano16S at temperatures greater than its T_m (~100 °C).

NDK Enzymatic Activity. All of the enzymes were assayed for the production of ATP by the NDK-catalyzed transfer of phosphate from GTP to ADP, specific activities measured in triplicate at 10 °C intervals, typically from 40 to 90 °C (Fig. 2). Consistent with the plausible accuracy of the reconstructions, all NDKs, ancestral and extant, were enzymatically active with the experimentally induced loss of their specific activity generally coinciding with their denaturation as inferred from measured T_m values. The order of magnitude difference between the specific activities of the reconstructed prokaryotic and eukaryotic NDK (Fig. 2) is characteristic both of the ancestral and extant NDKs studied here, perhaps evidencing an inheritably lower catalytic activity of the eukaryotic enzyme.

Ancestral Environmental Temperature. The T_m values listed in Table 1 show a marked overall decrease in the thermostability of the reconstructed NDKs with increasingly younger fossil-based divergence ages. Previous studies have demonstrated a strong correlation in extant archaea and bacteria between such thermostability and organismal environmental growth temperature, data that have provided the basis for construction of a calibration curve from which to infer paleoenvironmental temperatures (22, 23). The present work extends such studies to include prokaryotic cyanobacteria and eukaryotic green algae and land plants of the Viridiplantae, light-requiring phototrophs from which surface or near-surface photic-zone temperatures can be inferred.

In Fig. 3, the ranges of thermostabilities determined for the NDKs of extant archaea and bacteria (22), cyanobacteria, and Viridiplantae are plotted relative to their mean organismal environmental temperatures, using values for archaea and bacteria as an upper limit and those either for cyanobacteria (for Cyano16S and Nosto16S) or Viridiplantae (for ViridiNDK, Viridi18S, EmbryoNDK, and Embryo18S) as a lower limit. The calculated combined correlation for all taxa between NDK T_m and organismal environmental temperature is strong ($r = 0.92$), the two parameters consistently offset by ~20 °C, and is particularly strong for the cyanobacterial NDK ($r = 1.00$). For the Viridiplantae, however, the correlation is less robust ($r = 0.21$), possibly because thermophiles, including hot spring chlorophytes (36), were purposefully not included in this study. In addition, the geologically relatively recent (~775 Ma; Table 1) divergence of the Viridiplantae resulted in the absence of modern mesophilic taxa that, evidently like extant archaeans, bacteria, and cyanobacteria, were derived from ancestors adapted to an earlier, appreciably hotter Earth (Fig. 4).

NDK Paleotemperature Trend. In Fig. 4, paleoenvironmental temperature ranges inferred from the modern taxa calibration (Fig. 3) are plotted relative to the estimated geological age of lineage origination and the paleotemperature trend suggested by $\delta^{18}O$ and $\delta^{30}Si$ measurements of marine cherts (5, 7). The data from both our reconstructed NDK studies and the isotopic geochemical record suggest a general cooling of Earth's environment over geological time, the datasets exhibiting notable agreement for an estimated Archean temperature of ~65–80 °C.

The ASR methods used in our study incorporate two principal sources of uncertainty: the accuracy of the reconstructed ancestral

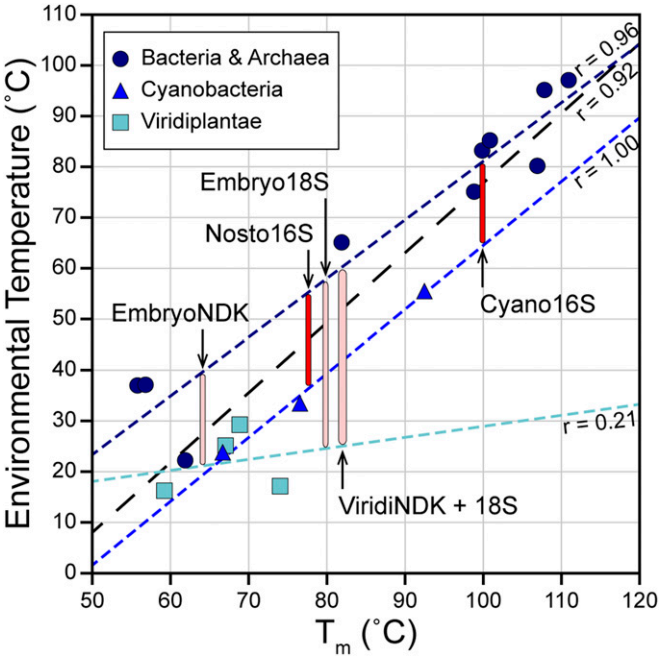

Fig. 3. Calibration curves showing the correlation between NDK T_ms and the organismal environmental growth temperatures of extant taxa of archaea and bacteria (22), cyanobacteria, and Viridiplantae. The best-fit linear regression lines for each group were constructed from the indicated data, the black-dashed line being the best fit through all available data and "r" specifying the calculated correlation coefficients. Environmental temperature ranges inferred from ancestral NDK T_ms are indicated by red bars for Cyano16S and Nosto16S and by pink bars for eukaryotic Viridi18S, ViridiNDK, Embryo18S, and EmbryoNDK.

enzyme sequences and the paleotemperatures inferred from the environmental temperature ranges of extant taxa (Fig. 3). As noted above, the plausibility of the reconstructed sequences can be evaluated by the functionality of the reconstructed enzymes and the degree of similarity of the measured thermostability of the same ancestral enzyme constructed based on differing phylogenetic trees. All NDKs reconstructed in our study are demonstrably enzymatically active, satisfying the first of these criteria. The second criterion, previously demonstrated to be met by archaea and bacteria (22) but not investigated for cyanobacteria, is shown to be satisfied here by the reconstructed NDKs of Viridiplantae for which the enzymes calculated both from 18S rRNA- and NDK-based phylogenies exhibit essentially identical thermostabilities (Table 1). And, although the difference of ∼15 °C between the T_m values of the reconstructed Embryo18S and EmbryoNDK enzymes (Table 1) signifies a need for further detailed study, it does not alter significantly the geologically long-term cooling trend indicated by the available data. Similarly, the ∼20–40 °C range of environmental temperature derived from the extant organism-based calibration curve (Fig. 3) does not overprint this first-order trend of gradually decreasing paleotemperature.

The hot, ∼65–80 °C Archean temperatures inferred from stable isotopes in marine cherts (5, 7), reconstructed archaeal and bacterial enzymes (20, 22, 23), and here from the reconstructed ancestral NDK of cyanobacteria are consistent with the upper environmental temperature limits of comparable modern taxa. For example, diverse extant hyperthermophilic prokaryotes, including methanogens, thrive in such settings (16) and some modern cyanobacteria inhabit thermal springs as hot as ∼75 °C (36). Given that methanogens were evidently a significant component of the pre-2,400 Ma biosphere (37) and that early cyanobacterial fossils are known solely from shallow marine sediments (26, 38–40), it is plausible and perhaps likely that thermophily was ecologically widespread during early Earth history. Interestingly, although eukaryotic algae originated much later in geological time than their prokaryotic precursors, the ∼40–60 °C upper temperature limit for extant taxa (36) is similarly consistent with the NDK-based temperature estimates derived here for ancestral Viridiplantae (Table 1).

The ASR methods used here to determine the thermostabilities of reconstructed ancestral NDKs of photic zone-inhabiting cyanobacteria, algae, and land plants provide a promising approach to analyses of the long-term temperature history of Earth's surface environment. Interdisciplinary studies such as this, combining the newly available data and techniques of genomic molecular biology with the more traditional findings of geology and paleontology, can be expected to play an increasingly significant role in understanding the interrelated evolution of life and its environment.

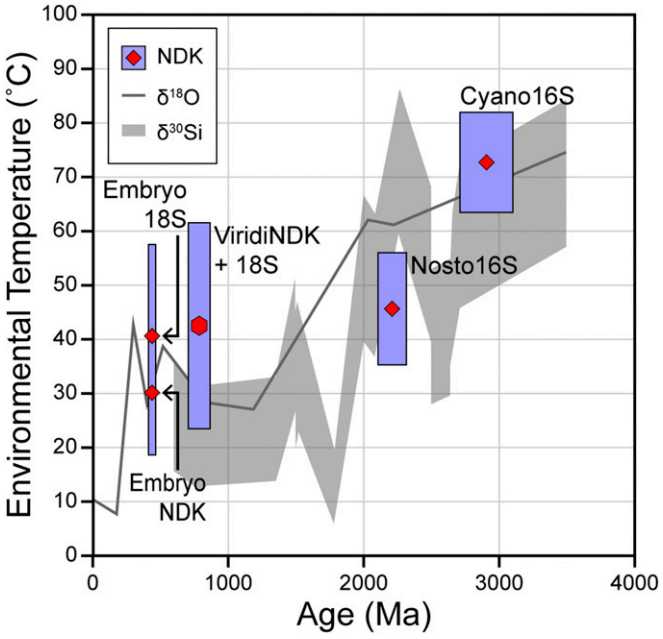

Fig. 4. Environmental temperature ranges inferred from reconstructed ancestral NDK T_ms plotted against fossil-record-indicated first appearance of the various groups. Paleotemperatures inferred from $\delta^{18}O$ (5) and $\delta^{30}Si$ (7) in marine cherts are included for comparison. Blue boxes show the inferred NDK-based temperature ranges (Fig. 3) and fossil-based age uncertainties, the red diamonds denoting temperature and age midpoint values for which ViridiNDK and Viridi18S have been combined due to the similarity of their T_ms.

Materials and Methods

rRNA and NDK Phylogenetic Tree Building. Following the methods of Akanuma et al. (22), NDK amino acid sequences were acquired from the National Center for Biotechnology Information (NCBI) Protein database for 170 taxa of extant cyanobacteria (including 31 nostocaleans) and 69 Viridiplantae (including 58 embryophytes). Cyanobacterial genomes include a single NDK gene. Of the multiple NDK genes occurring in the genomes of plants, the cytosolic Viridiplantae NDKI isoform (referred to here simply as "NDK") was selected for this study because of its similarity in sequence length and composition to cyanobacterial NDK. To serve as outgroups for the analyzed cyanobacteria and Viridiplantae, NDK sequences were obtained from the database for 193 extant taxa of prokaryotes and 24 eukaryotes, respectively. Cyanobacteria and Viridiplantae NDK sequences were aligned separately by Multiple Alignment Using Fast Fourier Transform (41) Version 7.222 and adjusted to correct for gap positions.

Complementary 16S rRNA or 18S rRNA nucleotide sequences from the same taxa used for the NDK alignments were obtained from the SILVA database, which for eight of the Viridiplantae were genus- rather than species-matched due to missing entries in SILVA (42). ML rRNA phylogenetic trees (Fig. 1 *A* and *B*) were constructed using randomized axelerated maximum likelihood (RAxML) (43) and the general time-reversible plus gamma model for cyanobacteria and LG + Gamma model for Viridiplantae; an NDK ML tree (LG + Gamma model) was also constructed using RAxML for the Viridiplantae (Fig. 1*C*). Best-fit evolutionary models were selected using jModelTest or ProtTest (44, 45).

Ancestral NDK Sequence Calculation. Ancestral NDK sequences were inferred by use of CODEML in the PAML program package (35). A 16S rRNA tree topology was used as a reference tree to infer the NDK sequences of the last common ancestors of cyanobacteria (Cyano16S) and nostocaleans (Nosto16S). The same procedure was performed using an 18S rRNA tree topology to infer the NDK sequences of the last common ancestors of Viridiplantae (Viridi18S) and Embryophyta (Embryo18S). Viridiplantae and Embrophyta ancestral NDK sequences were also inferred using an NDK tree topology (for ViridiNDK and EmbryoNDK).

Ancestral NDK Gene Construction. Ancestral NDK gene sequences were determined by reverse translation of the inferred amino acid sequences with optimal codon use for gene expression in *E. coli*. Cyano16S and Nosto16S gene constructs and the NDKs of three extant cyanobacterial taxa were prepared by Eurofins Scientific and ligated into pTAKN-2 (BioDynamics Laboratory Inc.). Viridi18S, ViridiNDK, Embryo18S, and EmbryoNDK gene constructs and the NDKs of four extant Viridiplantae taxa were prepared by GenScript and ligated into pET23a(+) (Novagen).

NDK Expression and Purification. pTAKN-2 constructs were digested with NdeI and BamHI (New England Biolabs), purified by agarose gel electrophoresis, and ligated to pET23a(+), a ligation step not necessary for GenScript constructs. *E. coli* Rosetta2 (DE3) (Novagen) cultures transformed with the expression plasmids were cultivated in Luria–Bertani medium supplemented with ampicillin (150 μg/mL). Gene expression was induced using Overnight Express Autoinduction system reagents (Novagen) at 37 °C. Cells were then disrupted by sonication and heated at 70 °C (for cyanobacterial proteins) or 60 °C (for Viridiplantae proteins) for 15 min to denature *E. coli* proteins. The heated suspensions were then centrifuged at 60,000 × *g*, and to purify the NDK, the resulting supernatants were chromatographed through HiTrapQ, ResourceQ, and Superdex 200 columns (GE Healthcare Biosciences). Protein homogeneity was confirmed by SDS/PAGE followed by Coomassie Brilliant Blue staining.

Thermostability and Enzyme Activity Analyses. Purified NDK protein concentrations were analyzed by spectrophotometric absorbance at 280 nm. For

thermostability analyses, each NDK protein was dissolved in 20 mM KPi (pH 6.0 or 7.6), 50 mM KCl, and 1 mM EDTA to a final concentration of 20 μM. Thermal denaturation experiments were performed in a 0.1-cm path-length pressurized cell using a J-720 spetropolarimeter (Jasco) equipped with a programmable temperature controller. For each measurement, temperature was typically increased at a rate of 1.0 °C/min from 40–110 °C and protein denaturation was monitored by CD spectral changes of the proteinaceous solution at 222 nm. The midpoint unfolding temperature, T_m, was calculated from two-state denaturation curves normalized to their native- and denatured-state baselines, measurements made for each protein at pH 6.0 and pH 7.6.

Enzyme activity assays were performed in a 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 DTT, 5.0 mM ADP, and 5.0 mM GTP. Increases in ATP, produced by the NDK-catalyzed transfer of phosphate from GTP to ADP, were monitored with the Kinase-Glo Luminescence Kinase Assay Kit (Promega), and a Perkin-Elmer Wallac 1420 Victor2 microplate reader (Winpact Scientific) was used to measure the luminescence. Measurements were typically taken in triplicate at 10 °C intervals from 40 to 90 °C.

ACKNOWLEDGMENTS. We thank M. Bessho, R. Furukawa, M. Harada, A. Nagano, and T. Sasamoto for discussion and assistance with both phylogenetic and experimental procedures. The analytical work reported was carried out by A.K.G. during two extended visits to the laboratories of A.Y., a collaboration arranged by J.W.S. and A.Y. and funded by the Center for the Study of Evolution and Origin of Life at University of California, Los Angeles and the Tokyo University of Pharmacy and Life Sciences.

1. Schopf JW (1993) Microfossils of the Early Archean Apex chert: New evidence of the antiquity of life. *Science* 260:640–646.
2. Wacey D, et al. (2006) The ~3.4 billion-year-old Strelley Pool Sandstone: A new window into early life on Earth. *Int J Astrobiol* 5:333–342.
3. Schopf JW (2013) Response by Bill Schopf and brief remarks about fundamental unsolved problems in paleontology. *J Paleontol* 87:526–528.
4. Veizer J, Prokoph A (2015) Temperatures and oxygen isotopic composition of Phanerozoic oceans. *Earth Sci Rev* 146:92–104.
5. Knauth LP, Lowe DR (1978) Oxygen isotope geochemistry of cherts from the Onverwacht Group (3.4 billion years), Transvaal, South Africa, with implications for secular variations in the isotopic composition of cherts. *Earth Planet Sci Lett* 41:209–222.
6. Knauth LP, Lowe DR (2003) High Archean climatic temperature inferred from oxygen isotope geochemistry of cherts in the 3.5 Ga Swaziland Supergroup, South Africa. *Geol Soc Am Bull* 115:566–580.
7. Robert F, Chaussidon M (2006) A palaeotemperature curve for the Precambrian oceans based on silicon isotopes in cherts. *Nature* 443:969–972.
8. Perry EC (1967) The oxygen isotope chemistry of ancient cherts. *Earth Planet Sci Lett* 3: 62–66.
9. Kasting JF, et al. (2006) Paleoclimates, ocean depth, and the oxygen isotopic composition of seawater. *Earth Planet Sci Lett* 252:82–93.
10. Hren MT, Tice MM, Chamberlain CP (2009) Oxygen and hydrogen isotope evidence for a temperate climate 3.42 billion years ago. *Nature* 462:205–208.
11. Degens ET, Epstein S (1962) Relationship between O^{18}/O^{16} ratios in coexisting carbonates, cherts, and diatomites. *Am Assoc Pet Geol Bull* 46:534–542.
12. Weis D, Wasserburg GJ (1987) Rb-Sr and Sm-Nd isotope geochemistry and chronology of cherts from the Onverwacht Group (3.5 AE), South Africa. *Geochim Cosmochim Acta* 51:973–984.
13. de Wit MJ, Furnes H (2016) 3.5-Ga hydrothermal fields and diamictites in the Barberton Greenstone Belt-Paleoarchean crust in cold environments. *Sci Adv* 2:e1500368.
14. Chakrabarti R, Knoll AH, Jacobsen SB, Fischer WW (2012) Si isotope variability in Proterozoic cherts. *Geochim Cosmochim Acta* 91:187–201.
15. Pauling L, Zuckerlandl E (1963) Chemical paleogenetics: Molecular "restoration studies" of extinct forms of life. *Acta Chem Scand* 17:S9–S16.
16. Woese CR (1987) Bacterial evolution. *Microbiol Rev* 51:221–271.
17. Galtier N, Tourasse N, Gouy M (1999) A nonhyperthermophilic common ancestor to extant life forms. *Science* 283:220–221.
18. Boussau B, Blanquart S, Necsulea A, Lartillot N, Gouy M (2008) Parallel adaptations to high temperatures in the Archaean eon. *Nature* 456:942–945.
19. Gaucher EA, Thomson JM, Burgan MF, Benner SA (2003) Inferring the palaeoenvironment of ancient bacteria on the basis of resurrected proteins. *Nature* 425:285–288.
20. Gaucher EA, Govindarajan S, Ganesh OK (2008) Palaeotemperature trend for Precambrian life inferred from resurrected proteins. *Nature* 451:704–707.
21. Hobbs JK, et al. (2012) On the origin and evolution of thermophily: Reconstruction of functional precambrian enzymes from ancestors of Bacillus. *Mol Biol Evol* 29:825–835.
22. Akanuma S, et al. (2013) Experimental evidence for the thermophilicity of ancestral life. *Proc Natl Acad Sci USA* 110:11067–11072.
23. Akanuma S, Yokobori S, Nakajima Y, Bessho M, Yamagishi A (2015) Robustness of predictions of extremely thermally stable proteins in ancient organisms. *Evolution* 69:2954–2962.
24. Farquhar J, Bao H, Thiemens M (2000) Atmospheric influence of Earth's earliest sulfur cycle. *Science* 289:756–759.
25. Anbar AD, et al. (2007) A whiff of oxygen before the great oxidation event? *Science* 317:1903–1906.
26. Schopf JW (2012) The fossil record of cyanobacteria. *Ecology of Cyanobacteria II: Their Diversity in Space and Time*, ed Whitton BA (Springer, Durham, UK), pp 15–36.
27. Golubic S, Sergeev VN, Knoll AH (1995) Mesoproterozoic Archaeoellipsoides: akinetes of heterocystous cyanobacteria. *Lethaia* 28:285–298.
28. Tomitani A, Knoll AH, Cavanaugh CM, Ohno T (2006) The evolutionary diversification of cyanobacteria: Molecular-phylogenetic and paleontological perspectives. *Proc Natl Acad Sci USA* 103:5442–5447.
29. Holland HD, Beukes NJ (1990) A paleoweathering profile from Griqualand West, South Africa: Evidence for a dramatic rise in atmospheric oxygen between 2.2 and 1.9 bybp. *Am J Sci* 290-A:1–34.
30. Butterfield NJ (2009) Modes of pre-Ediacaran multicellularity. *Precambrian Res* 173: 201–211.
31. Wellman CH, Gray J (2000) The microfossil record of early land plants. *Philos Trans R Soc Lond B Biol Sci* 355:717–731, discussion 731–732.
32. Ruhfel BR, Gitzendanner MA, Soltis PS, Soltis DE, Burleigh JG (2014) From algae to angiosperms-inferring the phylogeny of green plants (Viridiplantae) from 360 plastid genomes. *BMC Evol Biol* 14:23.
33. Hanson-Smith V, Kolaczkowski B, Thornton JW (2010) Robustness of ancestral sequence reconstruction to phylogenetic uncertainty. *Mol Biol Evol* 27:1988–1999.
34. Eick GN, Bridgham JT, Anderson DP, Harms MJ, Thornton JW (2016) Robustness of reconstructed ancestral protein functions to statistical uncertainty. *Mol Biol Evol* 34: 247–261.
35. Yang Z (1997) PAML: A program package for phylogenetic analysis by maximum likelihood. *Comput Appl Biosci* 13:555–556.
36. Brock TD (1967) Life at high temperatures. Evolutionary, ecological, and biochemical significance of organisms living in hot springs is discussed. *Science* 158:1012–1019.
37. Hayes JM (1983) Geochemical evidence bearing on the origin of aerobiosis: A speculative hypothesis. *Earth's Earliest Biosphere, its Origins and Evolution*, ed Schopf JW (Princeton Univ Press, Princeton), pp 291–301.
38. Golubic S, Lee S-J (1999) Early cyanobacterial fossil record: Preservation, palaeoenvironments and identification. *Eur J Phycol* 34:339–348.
39. Hofmann HJ (1976) Precambrian microflora, Belcher Islands, Canada: Significance and systematics. *J Paleontol* 50:1040–1073.
40. Klein C, Beukes NJ, Schopf JW (1987) Filamentous microfossils in the early Proterozoic Transvaal Supergroup: Their morphology, significance, and paleoenvironmental setting. *Precambrian Res* 36:81–94.
41. Katoh K, Misawa K, Kuma K, Miyata T (2002) MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res* 30:3059–3066.
42. Quast C, et al. (2013) The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res* 41:D590–D596.
43. Stamatakis A (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30:1312–1313.
44. Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* 52:696–704.
45. Darriba D, Taboada GL, Doallo R, Posada D (2012) jModelTest 2: More models, new heuristics and parallel computing. *Nat Methods* 9:772.