Higher temperatures worsen the effects of mutations on protein stability

Dimitrios - Georgios Kontopoulos^{1,2*}, Ilias Patmanidis³, Timothy G. Barraclough^{2,4}, Samraat Pawar²

*For correspondence:

dgkontopoulos@gmail.com

¹ Science and Solutions for a Changing Planet DTP, Imperial College London, London, UK;
² Department of Life Sciences, Imperial College London, Silwood Park, Ascot, Berkshire, UK;
³ Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Groningen, the Netherlands; ⁴ Department of Zoology, University of Oxford, Oxford, Oxfordshire, UK

Abstract Understanding whether and how temperature increases alter the effects of mutations on protein stability is crucial for understanding the limits to thermal adaptation by organisms. Currently, it is generally assumed that the stability effects of mutations are independent of temperature. Yet, mutations should become increasingly destabilizing as temperature rises due to the increase in the energy of atoms. Here, by performing an extensive computational analysis on the essential enzyme adenylate kinase in prokaryotes, we show, for the first time, that mutations become more destabilizing with temperature both across and within species. Consistent with these findings, we find that substitution rates of prokaryotes decrease nonlinearly with temperature. Our results suggest that life on Earth likely originated in a moderately thermophilic and thermally fluctuating environment, and indicate that global warming should decrease the per-generation rate of molecular evolution of prokaryotes.

24 Introduction

25 Genetic mutations are the primary source of the extensive phenotypic diversity that can be observed
 26 today, providing the raw material upon which selection can act. Thus, understanding the factors that
 27 influence the effects of mutations, as well as the rate at which they occur, is necessary for predicting
 28 the impacts of environmental change on organisms. From a thermodynamic standpoint, a given
 29 nonsynonymous mutation (i.e., a mutation that alters the amino acid sequence of a protein) can be
 30 classified as beneficial or detrimental based on its effect on the stability of the protein (**Zeldovich**
 31 **et al., 2007b; Chen and Shakhnovich, 2009; Quan et al., 2016**). Assuming a folding model with two
 32 states (unfolded and folded), we can define protein stability as the difference in Gibbs free energy
 33 ($\Delta G = G_{\text{unfolded}} - G_{\text{folded}}$). Nonsynonymous mutations alter ΔG by changing the free energy of the
 34 folded state. Therefore, the impact of a specific nonsynonymous mutation on protein stability
 35 can be calculated as $\Delta\Delta G = \Delta G_{\text{mutant}} - \Delta G_{\text{wild}}$. Under this notation, mutations with positive $\Delta\Delta G$
 36 values are beneficial (stabilizing) and those with negative values are detrimental (destabilizing).
 37 Most mutations are destabilizing, with the average $\Delta\Delta G$ close to -1 kcal/mol (**Zeldovich et al., 2007b;**
 38 **Tokuriki et al., 2008; Tokuriki and Tawfik, 2009**). Furthermore, an analysis of 2,188 point mutation
 39 experiments has shown that $\Delta\Delta G$ values are almost independent of ΔG_{wild} (**Chen and Shakhnovich,**
 40 **2009**), i.e., nonsynonymous mutations affect highly stable and less stable proteins quite similarly.
 41 For context, a mutation that leads to a ΔG_{mutant} value of 0 kcal/mol will result in half of the protein
 42 population being in the unfolded state (**Zeldovich et al., 2007b; Chen and Shakhnovich, 2009**). This
 43 is expected to strongly impair the activity of the protein and can lead to the death of the organism,
 44 especially if the protein is essential for survival (**Zeldovich et al., 2007b; Chen and Shakhnovich,**
 45 **2009**).

46 The impact of temperature on ΔG is well-recognised and, for proteins that follow the two-
 47 state folding/unfolding model, the relationship between ΔG and temperature can be described
 48 by the Gibbs-Helmholtz equation (**Privalov 1990**; pink solid line in Figure 1; see also section I in
 49 Appendix 1). In contrast, there is no theoretical basis for the influence of temperature on the
 50 average $\Delta\Delta G$, possibly because it is extremely challenging to experimentally quantify the effects of
 51 all possible nonsynonymous mutations for a given protein at multiple temperatures. Temperature
 52 rises have often been reported to make mutations increasingly deleterious (temperature-sensitive
 53 mutations; **Guo et al. 1996; Dieterle et al. 2010; Ndo et al. 2018**), but this is typically hypothesised

Figure 1. The two hypotheses tested in this study. The pink solid line represents the protein stability (ΔG) as a function of temperature, whereas the lengths of the purple arrows stand for the average effects of nonsynonymous mutations ($\Delta\Delta G$). Panel A represents the conventional assumption that mutational effects are independent of temperature, whereas in panel B, mutations become increasingly destabilizing as temperature increases. The dashed line indicates the position on the vertical axis of the critical threshold of $\Delta G = 0$ kcal/mol, where the protein population is equally divided into the folded and unfolded states. The two hypotheses make different predictions for the relationship between per-generation substitution rate and temperature. In panel C, the substitution rate has a similar shape to that of the protein stability curve. In contrast, in panel D, the substitution rate is nearly invariant at low temperatures (where mutations have a negligible effect on protein stability), but decreases nonlinearly at higher temperatures.

54 to be exclusively driven by the decline in protein stability at high temperatures (*Puurttinen et al.*
55 *2016; Berger et al. 2020*; Figure 1A). Under this hypothesis, $\Delta\Delta G$ is assumed to be independent of
56 temperature and constant (the “temperature-invariant $\Delta\Delta G$ ” hypothesis). Thus, mutations should
57 become deleterious also at low temperatures where protein stability can be similarly low (*Puurttinen*
58 *et al., 2016*).

59 In this study we introduce a new, alternative hypothesis (Figure 1B), that $\Delta\Delta G$ is not constant but
60 becomes more negative with temperature (the “temperature-dependent $\Delta\Delta G$ ” hypothesis). Such a
61 relationship is expected from molecular biophysics. Consider a protein structure from a hyperther-
62 mophilic organism which will have evolved under strong selection against heat denaturation, with
63 this being reflected in its amino acid composition. In other words, the amino acid frequencies of a
64 hyperthermophilic protein tend to differ from those of its psychrophilic or mesophilic orthologs

(see e.g., *Nakashima et al. 2003; Zeldovich et al. 2007a*). At temperatures typically experienced by a hyperthermophile, a nonsynonymous mutation would (on average) strongly destabilize a key conformation of the protein, due to the high internal (kinetic and potential) energy of the atoms. In contrast, if the same mutation occurred at a lower temperature, atoms would have lower energy and the mutation would be better tolerated. Thermal shifts in the opposite direction (e.g., if a mesophilic protein is exposed to high temperatures) would have the opposite result (i.e., the mutation would become more detrimental). If this hypothesis holds, then there should be a negative interspecific relationship between environmental temperature and the average $\Delta\Delta G$ value—estimated at a temperature close to those experienced by each species. A similar pattern would be expected between temperature and the average mutational effect within species.

Another difference between the two hypotheses lies in their predictions regarding the effects of temperature on the rate of molecular evolution (Figure 1C,D). Near the high temperature extreme, both hypotheses agree that mutations should bring the protein's stability close to the critical threshold of 0 kcal/mol. Thus, thermophiles should have lower rates of sequence substitution, due to i) a higher fraction of deleterious mutations and potentially ii) the evolution of a low mutation rate through optimisation of DNA proofreading and repair mechanisms (*Kimura, 1987*). In cold environments, however, the temperature-invariant $\Delta\Delta G$ hypothesis would also predict a low substitution rate (*Puurtinen et al., 2016*), as the resulting protein stability would be again close to the critical threshold (Figure 1A). In contrast, a low substitution rate in psychrophiles is not expected under the temperature-dependent $\Delta\Delta G$ hypothesis (Figure 1B).

To test the two hypotheses, we used the essential enzyme adenylate kinase (ADK). This enzyme plays an important role in cellular energy homeostasis as it catalyses the biochemical reaction $\text{MgATP} + \text{AMP} \rightleftharpoons \text{MgADP} + \text{ADP}$, regulating adenine nucleotide levels in the cell. Studies in bacteria have shown that the stability of ADK is directly linked to fitness, suggesting the presence of strong selection (*Couñago and Shamoo, 2005; Couñago et al., 2006*). We first collated a dataset of 70 ADKs of bacteria and archaea, spanning the entire temperature spectrum of life on Earth (from psychrophiles to hyperthermophiles; Appendix 1—table 1), and examined their evolutionary patterns. We then performed molecular dynamics simulations to obtain realistic conformations of each ADK at a temperature close to the natural environment of each species, and inferred the effects of all possible mutations (between 3,382 and 4,256 mutations per ADK depending on sequence length) on each conformation *in silico*. Lastly, we estimated the effects of temperature on

i) the median $\Delta\Delta G$ across and within species, and on ii) substitution rate (based on 44 genes).

Results

Dataset of ADK structures

We initially queried the Protein Data Bank (PDB) for experimentally determined ADK structures of bacteria and archaea. As this search yielded results from only 18 unique species (as of February 2018), we also performed homology modelling (*Brändén and Tooze, 1999*) to obtain a larger and more phylogenetically diverse dataset. Briefly, homology modelling allows for the inference of the 3D structure of a protein sequence (query sequence) based on similar protein sequences whose structures are experimentally determined (template sequences). To this end, we manually selected 70 bacterial and archaeal ADK sequences from the UniProt database (*The UniProt Consortium 2017*; Appendix 1—table 1), covering a broad range in terms of taxonomy (16 phyla; see Appendix 1—figure 4) and thermal preferences, and for which reliable homology models could be inferred using the Phyre2 web server (*Kelley et al. 2015*; see the Methods section). Regarding the thermal environment, in particular, we divided ADKs into four groups: psychrophiles (adapted to low temperatures and with an optimum temperature for growth $\leq 20^\circ\text{C}$), mesophiles (adapted to moderate temperatures and with an optimum temperature for growth between 20 and 40°C), thermophiles (adapted to high temperatures and with an optimum temperature for growth between 40 and 80°C), and hyperthermophiles (adapted to very high temperatures and with an optimum temperature for growth $> 80^\circ\text{C}$). We tried to avoid ADKs from species which did not conform to these criteria (e.g., psychrotolerant species that are capable of growing at low temperatures but grow optimally at temperatures far above 20°C).

3D structure and evolution of prokaryotic ADKs

The ADK structure consists of three domains (Figure 2): the large and rigid CORE domain, the LID domain which binds ATP, and the NMPbind domain which binds AMP (*Li et al., 2015*). The length of the LID domain varies across species within our dataset, with most bacterial ADKs having a long LID domain and most archaeal ADKs a short one. Furthermore, archaeal ADKs with short LID domains have two additional β -strands in their CORE domain. These β -strands allow such ADKs to form homotrimers, whereas ADKs that lack them are monomeric (*Vonrhein et al., 1998*). The major conformational states that ADKs adopt are two: the open (inactive) conformation (Figure

Figure 2. Open and closed conformations of the ADK of *Streptococcus pneumoniae* (PDB accession codes 4NTZ and 4NU0 respectively). The main structural domains are explicitly shown with different colors in panels A and C. Panels B and D are two-dimensional maps of the surface of each conformation, colored according to the charge of exposed amino acids (positively charged in blue, negatively charged in red). The open conformation (A,B) is energetically favourable in the absence of substrate. Upon substrate binding, the free energy of the closed conformation (C,D) becomes lower than that of the open conformation, and thus the former is preferred. A and C were rendered with UCSF Chimera v. 1.12 (Pettersen et al., 2004), whereas B and D with Structuprint v. 1.001 (Kontopoulos et al., 2016).

2A,B) and the closed (active) conformation (Figure 2C,D). As the protein transitions between these two states, other intermediate conformations may be observed. In the absence of ligands, ADKs tend to shift between open and closed-like conformations, whereas ligand binding induces the closed conformation (Kovermann et al., 2015; Li et al., 2015; Kovermann et al., 2017).

The aforementioned structural differences among ADKs in our dataset (e.g., the length of the LID domain) may systematically affect the median $\Delta\Delta G$ value, representing distinct evolutionary strategies against deleterious mutations. Thus, to better understand how variation in prokaryotic ADK structure emerges, we added an extra (outgroup) sequence (UniProt:Q9HKM7) to our dataset of 70 ADKs and reconstructed their evolutionary tree. The new sequence belongs to the archaeum *Thermoplasma acidophilum*, is described as a “putative adenylate kinase”, and is much shorter than any other ADK in our dataset. The 70 ADKs were found to form two major phylogenetic clusters

(shown in purple and mustard colors; Figure 3A). The purple cluster contains structures from bacteria and archaea with both short and long LID domains, but without the two extra β -strands, causing these ADKs to form monomers (Figure 3B, subpanels I, II, and III). In contrast, the mustard cluster contains archaeal ADKs that can form homotrimers (Figure 3B, subpanel IV). Both clusters contain ADKs from two major archaeal clades: the Euryarchaeota and the TACK superphylum. This suggests that the two clusters are the outcome of a gene duplication that took place in the archaeal lineage, before the split of Euryarchaeota and TACK. The large distance that separates the two clusters, the divergence in secondary and quaternary structure between the two clusters, as well as systematic differences in their site-specific rates of evolution (inset of Figure 3A) are all consistent with this hypothesis. It is worth emphasising here that the activity of the ADK protein is directly linked to organismal fitness (Couñago and Shamoo, 2005; Couñago et al., 2006). Thus, the significant structural differences that can be observed between the two ADK types would be highly unlikely to occur without the presence of two gene copies. One copy would remain functional, whereas the other would be able to evolve under relaxed selection pressure. Given that no species in our dataset expressed both ADK types (according to the UniProt database), either gene copy may have become a pseudogene or may have been removed through genome streamlining (Ranea et al., 2005; Giovannoni et al., 2014). The hypothesis of a single gene duplication, followed by multiple gene losses is consistent with the phylogeny of the 70 species in our dataset, which we reconstructed from 44 genes (Figure 3C; see Methods and section IV in Appendix 1). Nevertheless, it is also possible that the distribution of the two ADK types across species may be partly the outcome of horizontal gene transfer events. In any case, we henceforth use the terms “ancestral-type” to refer to ADKs from the purple cluster and “archaeal-type” to refer to those from the mustard cluster.

158 **Mutations become more destabilizing with temperature across and within species**

159 Next, we examined if the median effect of all possible mutations on each ADK varies with environmental temperature across species. To this end, we conducted molecular dynamics simulations (200 ns long, with 10 independent replicates) for each ADK at a temperature close to those experienced by each species: 6.85°C for psychrophiles, 26.85°C for mesophiles, 56.85°C for thermophiles, and 81.85°C for hyperthermophiles. These simulations enabled us to a) obtain reliable conformations for each ADK at its biologically meaningful temperature, and to b) remove any minor structural biases arising from homology modelling using multiple templates. Sampled ADK conformations

Figure 3. Evolution of ADKs of bacteria and archaea. A: Prokaryotic ADKs are broadly divided into two phylogenetic clusters, possibly as a result of an early gene duplication event. The ADKs of the purple (ancestral-type) cluster are monomeric, whereas those of the mustard (archaeal-type) cluster are homotrimeric. The archaeal sequence that belongs to neither cluster is the outgroup sequence. Inset: Distributions of TIGER scores (a proxy for site-specific evolutionary rate; *Cummins and McInerney 2011*) for the ADKs of the two phylogenetic clusters, which we aligned separately. Boxplot edges represent the first and third quartiles, whereas the solid line stands for the median. Whiskers extend up to the most remote data point within 1.5 interquartile ranges from each boxplot edge. The difference between the two TIGER score distributions is statistically significant according to the two-sided two-sample Kolmogorov-Smirnov test (*Corder and Foreman, 2014*). Their low median values highlight that both ADK types (but predominantly those that are monomeric) exhibit very high rates of evolution. This explains the poor statistical support (lack of black or grey circles) of most tree nodes. B: Four representative ADK structures from bacteria (subpanels I and II) and archaea (subpanels III and IV). The main differences among them (i.e., LID length and presence/lack of two additional β -strands) are highlighted in blue and red. Structures I-III are part of the purple phylogenetic cluster, whereas IV is part of the mustard cluster. C: The proposed hypothesis regarding the emergence of the two ADK types (shown as purple and mustard circles). The arrow indicates the position of the hypothesised gene duplication event. Ancestral species (tree nodes) that presumably had both ADK types are shown with both circles.

were grouped into clusters (from 2 to 20), and an average (representative) structure was calculated for each conformational cluster (see Methods and Appendix 1—figure 7 for an example). We then submitted each representative structure to the STRUM web service (*Quan et al., 2016*), a machine learning tool that has been trained and tested on experimentally-determined $\Delta\Delta G$ values from the ProTherm database (*Kumar et al., 2006*). STRUM can provide reasonably accurate $\Delta\Delta G$ estimates for all possible single point mutations of a given protein, based on sequence and structure information. After obtaining estimates of the median $\Delta\Delta G$ value per ADK conformation, we calculated the weighted median mutational effect across all conformations ($\widetilde{\Delta\Delta G}$), with weights equal to the proportion of each conformational cluster across simulations.

We checked the $\widetilde{\Delta\Delta G}$ values for two possible biases: the type of the conformation (open versus closed) and the quality of the initial homology model. For the former, we used the R package bio3d (v. 2.3-4; *Grant et al. 2006*) to calculate the radius of gyration of each conformation, which is a measure of protein structure compactness. For the latter, we used the proportion of amino acids in disallowed regions (see the Methods section) as a metric. No systematic association could be detected between $\widetilde{\Delta\Delta G}$ and these two factors (Appendix 1—figure 8).

To investigate if $\widetilde{\Delta\Delta G}$ varies with temperature, we fitted a number of alternative regression models with the MCMCglmm R package (v. 2.26; *Hadfield 2010*; see the Methods section). We found that $\widetilde{\Delta\Delta G}$ becomes linearly more negative (mutations more destabilizing) with temperature (Figure 4A, Appendix 1—table 4). The two ADK types share a common slope for temperature but have different intercepts, with archaeal-type ADKs being more robust to nonsynonymous mutations than ancestral-type ADKs. A systematic effect of LID length on $\widetilde{\Delta\Delta G}$ could not be detected, and controlling for phylogeny did not improve the fit either (Appendix 1—table 4).

Further support for a non-constant $\widetilde{\Delta\Delta G}$ was obtained through the estimation of its phylogenetic heritability, i.e., the extent to which closely related species (which often tend to live in similar thermal environments) have more similar $\widetilde{\Delta\Delta G}$ values than randomly selected species. A complete absence of phylogenetic heritability would be expected if the average effect of mutations was effectively constant across temperatures and species. In contrast, for both ADK types, $\widetilde{\Delta\Delta G}$ exhibited intermediate phylogenetic heritability. This indicates that it evolves along the phylogeny but not in a purely random and gradual (Brownian motion) manner (Figure 4B-C), likely reflecting the rapid evolution of the ADK sequence (see the inset of Figure 3A).

We additionally checked if the temperature-dependent $\Delta\Delta G$ hypothesis also holds within species.

Figure 4. $\Delta\Delta G$ varies systematically with temperature and is phylogenetically heritable. A: Across species, nonsynonymous mutations become on average more destabilizing with temperature. Each data point represents the weighted median effect of all possible mutations for a single ADK at a temperature close to those experienced by the species. Mutations are less detrimental to archaeal-type ADKs, suggesting a potential evolutionary benefit for species whose genomes code for this ADK type. All coefficients of the model had 95% Highest Posterior Density (HPD) intervals that did not include zero. B-C: For both ancestral (panel B) and archaeal (panel C) ADK types, $\Delta\Delta G$ evolves along the species phylogeny, reflecting the evolution of species' thermal strategies (e.g., adaptation to psychrophilic temperatures). Nevertheless, the phylogenetic heritability is lower than that expected by a random walk in the parameter space (Brownian motion), possibly in part due to episodes of convergent evolution or interspecific recombination of the ADK gene (Feil et al., 1996). $\Delta\Delta G$ values of internal nodes were estimated using the stable model (Elliot and Mooers, 2014), which can accommodate patterns of both gradual and rapid trait evolution. Panels B and C were rendered using the phytools R package v. 0.6-60 (Revell, 2012).

For this, we selected five ancestral-type ADKs (for consistency reasons) from our dataset: i) from *Aquifex aeolicus* (a hyperthermophilic bacterium), ii) from *Thermus thermophilus* (a thermophilic bacterium), iii) from *Mycobacterium tuberculosis* (a mesophilic bacterium), iv) from *Sporosarcina globispora* (a psychrophilic bacterium), and v) from *Methanosarcina acetivorans* (a mesophilic archaeum). We also performed ancestral sequence reconstruction and homology modelling to infer the ADK of the last universal common ancestor (LUCA; section VIII in Appendix 1). An analysis of the resulting sequence showed that the enzyme's optimum temperature for catalysis (T_{opt}) is estimated at 65.23°C (95% prediction interval = [54.79, 75.67]). Given that T_{opt} values of prokaryotic ADKs tend to fall close to the temperature of maximum growth rate (Appendix 1—figure 12), this result suggests that the LUCA likely experienced weakly thermophilic conditions.

We then performed molecular dynamics simulations for each of these 6 ADKs at 8 different temperatures, from 6.85°C to 94.35°C. This allowed us to infer the effects of mutations on each ADK across a wide temperature range, including temperatures in which the species would not be able to survive. As previously, representative structures were submitted to STRUM to obtain estimates of their median $\Delta\Delta G$. Finally, we used weighted least squares regression to fit six mathematical equations (Appendix 1—figure 9) to the relationship between median $\Delta\Delta G$ and temperature of each species. The best-fitting equation was determined according to the small sample size-corrected Akaike Information Criterion (Sugiura 1978; AICc). For all ADKs, median $\Delta\Delta G$ becomes more negative with temperature (Figure 5), as expected. It is also worth noting that the shape of the relationship between median $\Delta\Delta G$ and temperature varies to an extent across ADKs of different species (see also Appendix 1—figure 10).

Substitution rates of prokaryotes decrease with temperature

Finally, we tested the predictions of the two hypotheses of Figure 1 for the relationship between substitution rate and temperature across the 70 species in our ADK dataset. It is worth noting that this relationship may be confounded by other (potentially temperature-dependent) factors such as cell volume or generation time (see sections X and XI in Appendix 1). Therefore, to control for any effects of these two factors on substitution rate, we used the following equation (see section XI in Appendix 1):

$$\ln(K \cdot t_{\text{gen}}) = \ln(\gamma) + \alpha \cdot \ln(V) + g(T). \quad (1)$$

Here, K is the substitution rate per time, t_{gen} the generation time, T temperature, V the cell volume,

Figure 5. The relationship of median mutational effects versus temperature for the ancestral-type ADKs of five extant species and of the Last Universal Common Ancestor. Each data point corresponds to a different representative conformation, whereas the color depth stands for the proportion of protein snapshots per conformation at a given temperature. The vertical line shows a temperature near those typically experienced by each species (i.e., the temperatures shown in Figure 4A). The observed relationships vary both in their position along the vertical axis and in their sensitivity to increasing temperature.

226 γ a normalization constant (independent of cell volume and temperature), α a dimensionless expo-
 227 nent, whereas $g(T)$ is a temperature dependence function. To accommodate the predictions of both
 228 hypotheses of this study (Figure 1), we specified $g(T)$ as either a logarithmic or a quadratic equation.
 229 Given that we had previously reconstructed and calibrated a 44-gene phylogeny of the species in
 230 our dataset (see Methods and section IV in Appendix 1), we had already obtained estimates of the
 231 substitution rate for each branch of the phylogeny. Thus, we used the median K estimate of the
 232 terminal branch corresponding to each of the 70 species. For temperature (T), we assigned to each
 233 species the temperature value for its group (i.e., 6.85°C for psychrophiles, 26.85°C for mesophiles,
 234 56.85°C for thermophiles, and 81.85°C for hyperthermophiles). Regarding generation time (t_{gen}) es-
 235 timates, we analysed a large dataset of prokaryotic growth rates (Smith et al., 2019) and obtained a
 236 general equation that links generation time with environmental temperature (section X in Appendix
 237 1). Lastly, cell volume estimates were calculated from linear cell dimension measurements which
 238 we collected from the literature for all but one species (section IX in Appendix 1).

239 We then fitted a series of alternative regression models with MCMCglmm and performed model
 240 selection to identify the most appropriate model (Appendix 1—table 6). Through this process, we

Figure 6. The generation time-corrected substitution rate decreases with temperature. This nonlinear monotonic relationship is consistent with our finding that mutational effects are temperature-dependent, becoming increasingly deleterious as temperature increases. Temperature explains 84% of variance on its own, whereas an additional 10% of variance is explained by a phylogenetic random effect on the intercept.

241 detected a strong negative association between generation time-corrected substitution rates and
 242 temperature (Figure 6). In contrast, a systematic effect of cell volume on substitution rate was not
 243 found, as the exponent of cell volume was statistically indistinguishable from zero.

244 Discussion

245 Nonsynonymous mutations have hitherto been considered to become more deleterious with
 246 temperature exclusively because of the decline in protein stability (ΔG) beyond a critical temperature
 247 (Figure 1A). Our study shows that besides temperature-driven changes in protein stability, mutations
 248 themselves ($\Delta\Delta G$) become on average increasingly destabilizing as temperature increases (Figure
 249 1B). This aggravation of the effects of mutations on protein stability is consistent with the increased
 250 energy of atoms at high temperatures.

251 In particular, our interspecific analysis of the weighted median effects of all possible mutations
 252 ($\widetilde{\Delta\Delta G}$) across two types of prokaryotic ADKs revealed that $\widetilde{\Delta\Delta G}$ decreases linearly with environmen-
 253 tal temperature and evolves across species (Figure 4), despite the relatively limited phylogenetic
 254 signal present in the alignment of ADK sequences (Figure 3A). Furthermore, ADK-specific relation-
 255 ships between median $\Delta\Delta G$ and temperature were negative (as expected) and varied predominantly
 256 in their elevation along the median $\Delta\Delta G$ axis (vertical axis of Figure 5 subpanels, Appendix 1—figure
 257 10), and to a lesser extent in their temperature sensitivity. Variation in temperature sensitivity is

likely to reflect the conditions of the local environment, as thermal generalist species or strains (with low temperature sensitivity) should be selected for at environments with large intergenerational temperature fluctuations (*Gilchrist, 1995*). In particular, given that the LUCA had a relatively low temperature sensitivity, we would expect it to have emerged in a strongly fluctuating thermal environment. Indeed, currently proposed environments for the origin of life (e.g., alkaline hydrothermal vents, subaerial hot springs, and submarine hydrothermal sediments; *Mulkidjanian et al. 2012; Sojo et al. 2016; Westall et al. 2018; Camprubí et al. 2019*) are characterised by strong gradients in temperature, pH, and other physicochemical variables. Such environments are also consistent (under certain assumptions) with the LUCA being adapted to a moderate thermophilic (but not hyperthermophilic) lifestyle, as suggested by the relatively high robustness to mutations of its ADK, its estimated T_{opt} for catalysis (65.23°C; Appendix 1—figure 12), as well as some previous studies (*Di Giulio, 2003; Gaucher et al., 2003, 2008; Catchpole and Forterre, 2019*).

Consistent with all the above results, we detected a monotonic nonlinear decrease of the per-generation substitution rate (averaged across 44 genes) with temperature (Figure 6). According to the neutral theory of molecular evolution (*Kimura, 1987*), such a pattern could occur if i) the proportion of deleterious mutations increases with temperature (as we show in this study) and/or if ii) mutation rates decrease with temperature. An investigation of the latter cannot be done without new laboratory experiments, given that almost all available prokaryotic mutation rate measurements are from mesophiles (Appendix 1—figure 16). A third factor that may partly drive the decline in substitution rates with temperature is the effective population size. More precisely, in species with low effective population size, weakly deleterious mutations may not be removed by selection, yielding an elevated substitution rate (*Woolfit, 2009*). Thus, if the effective population size increases with temperature across prokaryotes, this would partly explain the high substitution rates of psychrophiles. Once again, currently existing datasets are inadequate for addressing this hypothesis.

To the best of our knowledge, such systematic inter- and intraspecific relationships between the median mutational effect and temperature have never been reported before in the literature. This is likely explained by the complete lack of experimentally-determined datasets of $\Delta\Delta G$ (averaged across all possible mutations) for a given protein at multiple temperatures. The generality of these relationships could be investigated by performing similar analyses for other kinases (or other proteins in general) in future studies. If, for example, the slope of the interspecific relationship

varies very little across different proteins, it could suggest the presence of a hard biophysical constraint on thermal adaptation. Similarly, determining the temperature sensitivity of the median $\Delta\Delta G$ value for many ADKs across thermal gradients could shed more light on the likely effects of the local environment.

Our finding of a decline in substitution rates with temperature appears to contradict the results of two previous studies by Gillooly and co-workers (Gillooly *et al.*, 2005, 2007). These studies found that across animals, nucleotide and amino acid substitution rates increase with temperature and decrease with body mass, consistent with a dependence of mutation rate on metabolic rate. The discrepancy between their results and ours probably arises from two main differences in study design. First, Gillooly and co-workers assumed that generation time has the same temperature and body size dependence as mass-specific metabolic rate. Second, because they only examined the substitution rates of animals, their datasets did not cover the entire temperature range, but mainly mesophilic temperatures.

Overall, the present study sheds new light on the mechanism by which nonsynonymous mutations become increasingly deleterious as temperature increases. Our results are consistent with and help explain the monotonic decrease in the per-generation substitution rates of prokaryotes with temperature, across a 75-degree temperature range. These findings further our understanding of the mechanisms underlying thermal adaptation in prokaryotes, from the origin of life to the present day, improving—among others—our ability to predict the potential of populations and species to develop evolutionary responses to ongoing global change.

Methods

Homology modelling

We performed homology modelling using the intensive mode of the Phyre2 server which can produce a homology model based on multiple templates available from the PDB. The choice of the most appropriate template(s) is made by taking into account i) the confidence of the alignment between the query and template sequences, and ii) the coverage of the query sequence by the template(s). Any amino acids of the query sequence that are not covered by the template(s) are modelled using an *ab initio* approach. As the accuracy of *ab initio* modelling is usually lower, we rejected any homology models for which more than four amino acids had to be modelled *ab initio*.

We also ensured that the sequence identity between each query sequence and its chosen

template(s) was sufficiently high. For this, we used Rost's empirical equation (**Rost, 1999**):

$$p_{\text{identical}} = 480 \cdot L^{-0.32} \cdot (1 + \exp(-L/1000)). \quad (2)$$

$p_{\text{identical}}$ is the percentage of sequence identity and L the length of the alignment. Plotting this equation for different values of L gives a line that partitions the parameter space into a "safe zone" for homology modelling (i.e., where the resulting structural models are expected to be reliable) and a "twilight zone" (Appendix 1—figure 2).

Finally, we performed a Ramachandran plot (**Brändén and Tooze, 1999**) analysis to assess the quality of the resulting homology models. The two axes of the Ramachandran plot correspond to the backbone dihedral angles ϕ and ψ of amino acids of a given protein structure. Based on a large number of experimentally determined structures, the parameter space defined by the two axes has been empirically partitioned into the "most favourable regions", "additional allowed regions", "generously allowed regions", and "disallowed regions". Thus, the amino acids of a high quality model are expected to be mostly found in the most favoured and additional allowed regions, with none or very few amino acids in disallowed regions. We generated Ramachandran plots for each ADK structure with PROCHECK (**Laskowski et al., 1993**) and examined the proportions of amino acids in each region of the plot (Appendix 1—figure 3).

Phylogenetic analyses

ADK tree inference

ADK sequences were aligned with MAFFT (v7.123b; **Katoh and Standley 2013**) using the L-INS-i algorithm. We then ran ModelTest-NG (**Darriba et al., 2020**) to identify the most appropriate substitution model for tree reconstruction, according to AICc. The best-fitting model was WAG, with Γ -distributed rate variation across alignment sites (**Gu et al., 1995**). We next performed 300 maximum likelihood tree searches with IQ-TREE (v. 1.6.7; **Nguyen et al. 2015**) and selected the tree with the highest log-likelihood. For this, we used the default options with the exception of the Nearest Neighbor Interchange tree search, whose thorough mode was activated (flag "-allnni"). The statistical support for the nodes of the best tree was estimated by executing 1,000 nonparametric bootstrap replicates.

345 Species tree reconstruction

346 To obtain a robust species tree, we collected further molecular sequences of multiple highly
347 phylogenetically informative genes from the species in the study, where possible (Appendix 1—
348 table 2). In particular, our dataset for species tree inference comprised i) amino acid sequences from
349 the two ADK types (which were treated as two distinct genes), ii) 16S and 23S rRNA sequences from
350 the SILVA rRNA database (*Quast et al., 2012*), and iii) amino acid sequences of the 40 protein-coding
351 “PhyEco” marker genes (*Wu et al., 2013*). More precisely, the PhyEco genes are universally present
352 across bacteria and archaea, have been shown to lead to the inference of robust species trees, and
353 exhibit low copy number variation. The study that introduced these 40 genes (*Wu et al., 2013*) had
354 provided FASTA files of bacterial and archaeal amino acid sequences for each PhyEco gene. We
355 filtered these files to extract sequences that belonged to the species in our study, and performed
356 blastp (*Altschul et al., 1997*) searches against the UniProt database to identify sequences from
357 missing species. This process was followed by manual curation to eliminate any false positive hits.

358 Sequences were aligned with MAFFT as previously described, followed by the removal of phylo-
359 genetically uninformative homoplastic sites in each alignment, as identified with Noisy (v. 1.5.12;
360 *Dress et al. 2008*). We then ran PartitionFinder 2 (*Lanfear et al., 2017*) to determine the optimal
361 partitioning scheme for nucleotide and amino acid sequences. In other words, PartitionFinder 2
362 identifies groups of genes or proteins (partitions) with similar rates or patterns of evolution, as
363 well as the most appropriate substitution model for them. That model can then be fitted to the
364 entire partition, instead of separately to each gene or protein within the partition. We used the
365 greedy search algorithm (*Lanfear et al., 2012*) of PartitionFinder 2, AICc as the criterion, and RAXML
366 (*Stamatakis, 2014*) for substitution model fitting. The models that were considered were the GTR
367 (for nucleotide sequences) and the WAG, MTREV, DAYHOFF, VT, BLOSUM62, CPREV, RTREV, and
368 MTMAM (for amino acid sequences). The resulting partitioning scheme was used to reconstruct
369 the species phylogeny with IQ-TREE. 20 tree searches were executed, all of which led to the same
370 topology (Appendix 1—figure 4). As previously, the statistical support was estimated via 1,000
371 bootstrap replicates.

372 For time-calibration of the phylogeny, we used BEAST (v. 2.5.2; *Bouckaert et al. 2014*) and a
373 lognormal relaxed clock (*Drummond et al., 2006*), keeping the topology fixed to that inferred by
374 IQ-TREE. We used the same partitioning scheme as before, but replaced any substitution models

that were not available in BEAST 2 with the general WAG model. To obtain a starting tree with ultrametric branch lengths, we used the *chronos* function in the *ape* R package (v. 5.2; *Paradis and Schliep 2019*), which is based on penalised likelihood (*Paradis, 2013*). Two independent BEAST 2 runs were executed for 450 million generations, with posterior samples being obtained every thousand generations, after discarding samples from the first 112.5 million generations as burn-in. To ensure that the two runs converged to a common posterior distribution, we calculated the effective sample size and the potential scale reduction factor for each model parameter (*Gelman and Rubin 1992; Brooks and Gelman 1998*; Appendix 1—figure 5). Our convergence criteria were i) an effective sample size of at least 200 and ii) a potential scale reduction factor value below 1.1. The same convergence diagnostics were performed for the estimated substitution rate of each tree branch (Appendix 1—figure 6). We summarised the posterior distribution of trees by calculating the median age estimate for each node. As we had not fixed the age for any node, the branch lengths of the resulting tree were in units of substitutions per site. To convert them to units of time, we set the root to a depth of 3,700 mya, according to current estimates of the date of the origin of life on Earth (*Pearce et al., 2018*). Substitution rates were also divided by 3,700.

Molecular dynamics simulations

Simulation protocol

Molecular dynamics simulations were performed using the Desmond software and its graphical interface, Maestro (Version 10.6.013, Release 2016-2). Our protocol was based on Desmond's molecular dynamics simulation tutorial, distributed with the software suite.

To prepare each ADK for the simulations, we first capped the N- and C-termini and added missing hydrogens. To remove close contacts or clashes between atoms, we relaxed the structure into a local energy minimum, in vacuum (i.e., in the absence of solvent), using the OPLS 2005 force field (*Banks et al., 2005*). The minimized protein was solvated in a periodic orthorhombic simulation box with Simple Point Charge water molecules (*Berendsen et al., 1981*). Na⁺ and Cl⁻ ions were added to neutralize the charge of the protein as appropriate, and to generate a NaCl salt concentration of 0.15 M. The resulting system was relaxed using Desmond's standard protocol.

We next performed ten independent simulations (with randomised initial atomic velocities) per ADK. These were performed under the NPT ensemble, i.e., with temperature and pressure being maintained constant using the Nosé-Hoover chain thermostat (*Martyna et al., 1992*) and

the Martyna-Tobias-Klein barostat (*Martyna et al., 1994*) respectively. We varied the temperature of the simulations by the thermal group: 280 K (6.85°C) for psychrophiles, 300 K (26.85°C) for mesophiles, 330 K (56.85°C) for thermophiles, and 355 K (81.85°C) for hyperthermophiles. The pressure was kept constant at the default value of 1.01325 bar (the standard atmospheric pressure) using isotropic coupling and a relaxation time of 2 ps. The cutoff radius for van der Waals and short range electrostatic interactions was set to 9 Å. The integration time step was left at the default value of 2 fs, with snapshots of the system being recorded every 500 ps until each simulation was 200 ns long.

Trajectory analysis

We analyzed the ten replicate molecular dynamics trajectories per ADK by first calculating the root-mean-square deviation (RMSD) matrix across all recorded snapshots (conformations) of the protein. The RMSD between two conformations of the same protein, **a** and **b**, is defined as

$$\text{RMSD}(\mathbf{a}, \mathbf{b}) = \sqrt{\frac{\sum_{i=1}^{N_{\text{atoms}}} [(\mathbf{a}_{ix} - \mathbf{b}_{ix})^2 + (\mathbf{a}_{iy} - \mathbf{b}_{iy})^2 + (\mathbf{a}_{iz} - \mathbf{b}_{iz})^2]}{N_{\text{atoms}}}}, \quad (3)$$

where x, y, and z are the axes of the three-dimensional space. As the entire protein moves in space during the course of a molecular dynamics simulation, an objective comparison of two conformations requires finding the superposition between them that minimizes the RMSD.

It is worth stressing that ADK structures contain a large number of amino acids that participate in neither helices nor β -strands. Such amino acids are usually highly flexible and their motions are largely random. Therefore, to reduce the influence of highly flexible amino acids on the identification of protein conformations, we excluded such amino acids from the calculation of the RMSD matrix. For this, we converted the trajectories from the Desmond format to DCD format with the VMD program (*Humphrey et al., 1996*). VMD was also used to generate a Protein Structure File. These two files were then analysed with Carma (*Glykos, 2006*). Carma can automate the secondary structure analysis of molecular dynamics trajectories by calling the STRIDE secondary structure assignment tool (*Frishman and Argos, 1995*) and analysing its output. This allowed us to identify the amino acids of each ADK that participated in neither helices nor β -strands for at least 80% of the combined duration of the ten replicate trajectories. Such amino acids were excluded from the superposition and the RMSD calculation.

The resulting RMSD matrix was clustered using the Partitioning Around Medoids algorithm

(*Reynolds et al., 2006*), as implemented in the cluster R package (*Maechler et al., 2017*). The only requirement of this algorithm is the number of clusters to detect, whereas the optimal number of clusters can be determined by comparing the mean silhouette score (*Rousseeuw, 1987*) of different clusterings. In our case, we set the initial number of clusters to all possible values between 2 and 20, and selected the clustering with the highest mean silhouette score. Finally, we used Carma to obtain the average (representative) conformation for each cluster.

The relationship between median mutational effects and temperature across species

To check for a systematic effect of temperature on $\widetilde{\Delta\Delta G}$ values of different species, we fitted a series of models with MCMCglmm. The response variable was $\widetilde{\Delta\Delta G}$, whereas possible predictors (other than an intercept) included: i) the temperature at which molecular dynamics simulations were performed, ii) the ADK type (ancestral or archaeal), and iii) the length of the LID. All possible combinations of these predictors were specified in different models, including an empty model that only contained an intercept. A phylogenetic variant of each model was also fitted, using the species phylogeny that we inferred. In that case, species identity was treated as a random effect on the intercept. An estimate of uncertainty for each $\widetilde{\Delta\Delta G}$ value was obtained through bootstrapping and was integrated into the analysis. We used the default prior for the fixed effects, and a relatively uninformative inverse- Γ prior for the random effect and residual variances. For each model, two replicate MCMC chains were executed for three million generations, with samples being obtained every thousand generations. We discarded samples from the first three hundred thousand generations as burn-in and ensured that the two chains per model had sufficiently converged on identical posterior distributions by examining the effective sample size for each parameter and the potential scale reduction factor, as described earlier. For identifying the most appropriate model, we first rejected models that had at least one coefficient whose 95% Highest Posterior Density (HPD) interval included zero. From the remaining models, we chose the model with the lowest Deviance Information Criterion (*Spiegelhalter et al. 2002*; DIC) value.

We also estimated the phylogenetic heritability of $\widetilde{\Delta\Delta G}$ to understand if closely related species have more similar values than species chosen at random. This involved fitting a phylogenetic regression of $\widetilde{\Delta\Delta G}$ using MCMCglmm with only an intercept as a predictor. This model was fitted separately to each ADK type by executing two replicate chains for ten million generations, with samples being obtained every thousand generations after the first million generations. The

463 phylogenetic heritabilities of $\widetilde{\Delta\Delta G}$ for the two ADK types were then estimated as the ratio of the
464 variance explained by the random effect (species identity corrected for phylogeny) over the sum of
465 the random effect variance and the residual variance.

466 To visualise the evolution of $\widetilde{\Delta\Delta G}$ for each ADK type on the species phylogeny (Figure 4B,C), we
467 estimated ancestral states by fitting the stable model of trait evolution (*Elliot and Mooers, 2014*). For
468 this, four MCMC chains were executed for a hundred million generations, sampling every hundred
469 generations after a burn-in phase of 25 million generations. Convergence of the four chains was
470 examined as previously described.

471 **Estimation of the relationship between substitution rate and temperature**

472 We fitted variants of Equation (1) with MCMCglmm, with and without a phylogenetic correction on
473 the intercept by executing two MCMC chains per model for a million generations, with posterior
474 samples being recorded every thousand generations after the first hundred thousand. Because cell
475 volume estimates were available as ranges for most species, we fitted models that included cell
476 volume 20 times. Each time, the volume of each species was sampled from a uniform distribution
477 with bounds that corresponded to the minimum and maximum volume calculations.

478 **Acknowledgments**

479 We thank Hira Tanvir for providing some of the cell volume data that were used in this study,
480 and Mark E. Ritchie, Andrew Hirst, Guy Woodward, Vassiliki Koufopanou, Bhavin S. Khatri, Vincent
481 Savolainen, and the Michael J. E. Sternberg Lab for useful discussions and comments. We are also
482 particularly grateful to the developers and maintainers of the STRUM web service, as a major com-
483 ponent of this study depended on it. Finally, we thank Imperial College London's High Performance
484 Computing service (doi:10.14469/hpc/2232) for access to the necessary computational resources
485 for performing the molecular dynamics simulations.

486 **Competing interests**

487 The authors declare that no competing interests exist.

Additional files

Data availability

The data and source code that support the conclusions of this study are available at <https://doi.org/10.6084/m9.figshare.12635837.v1> and https://github.com/dgkontopoulos/Kontopoulos_et_al_mutations_vs_temperature_2020, respectively.

References

- Altschul SF**, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research*. 1997; 25(17):3389–3402.
- Angilletta MJ**. Thermal adaptation: a theoretical and empirical synthesis. Oxford University Press; 2009.
- Ashkenazy H**, Penn O, Doron-Faigenboim A, Cohen O, Cannarozzi G, Zomer O, Pupko T. FastML: a web server for probabilistic reconstruction of ancestral sequences. *Nucleic Acids Research*. 2012; 40(W1):W580–W584.
- Banks JL**, Beard HS, Cao Y, Cho AE, Damm W, Farid R, Felts AK, Halgren TA, Mainz DT, Maple JR, et al. Integrated modeling program, applied chemical theory (IMPACT). *Journal of Computational Chemistry*. 2005; 26(16):1752–1780.
- Berendsen HJ**, Postma JP, van Gunsteren WF, Hermans J. Interaction models for water in relation to protein hydration. In: *Intermolecular forces* Springer; 1981.p. 331–342.
- Berger D**, Stångberg J, Baur J, Walters RJ, Elevated temperature increases genome-wide selection on de novo mutations; 2020. <https://doi.org/10.1101/268011>.
- Bouckaert R**, Heled J, Kühnert D, Vaughan T, Wu CH, Xie D, Suchard MA, Rambaut A, Drummond AJ. BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Computational Biology*. 2014; 10(4):e1003537.
- Brändén CI**, Tooze J. Introduction to Protein Structure. Garland Science; 1999.
- Brooks SP**, Gelman A. General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics*. 1998; 7(4):434–455.
- Campubí E**, De Leeuw J, House C, Raulin F, Russell M, Spang A, Tirumalai M, Westall F. The Emergence of Life. *Space Science Reviews*. 2019; 215(8):56.
- Catchpole RJ**, Forterre P. The evolution of reverse gyrase suggests a nonhyperthermophilic Last Universal Common Ancestor. *Molecular Biology and Evolution*. 2019; 36(12):2737–2747.
- Chen P**, Shakhnovich EI. Lethal mutagenesis in viruses and bacteria. *Genetics*. 2009; 183(2):639–650.
- Corder GW**, Foreman DI. Nonparametric statistics: a step-by-step approach. Second ed. John Wiley & Sons; 2014.

518 **Couñago R**, Chen S, Shamoo Y. In vivo molecular evolution reveals biophysical origins of organismal fitness.
519 *Molecular Cell*. 2006; 22(4):441–449.

520 **Couñago R**, Shamoo Y. Gene replacement of adenylate kinase in the gram-positive thermophile *Geobacillus*
521 *stearothermophilus* disrupts adenine nucleotide homeostasis and reduces cell viability. *Extremophiles*. 2005;
522 9(2):135–144.

523 **Cummins CA**, McInerney JO. A method for inferring the rate of evolution of homologous characters that
524 can potentially improve phylogenetic inference, resolve deep divergence and correct systematic biases.
525 *Systematic Biology*. 2011; 60(6):833–844.

526 **Darriba D**, Posada D, Kozlov AM, Stamatakis A, Morel B, Flouri T. ModelTest-NG: a new and scalable tool for the
527 selection of DNA and protein evolutionary models. *Molecular Biology and Evolution*. 2020; 37(1):291–294.

528 **DeLong JP**, Okie JG, Moses ME, Sibly RM, Brown JH. Shifts in metabolic scaling, production, and efficiency across
529 major evolutionary transitions of life. *Proceedings of the National Academy of Sciences*. 2010; 107(29):12941–
530 12945.

531 **Di Giulio M**. The universal ancestor was a thermophile or a hyperthermophile: tests and further evidence.
532 *Journal of Theoretical Biology*. 2003; 221(3):425–436.

533 **Dieterle MG**, Wiest AE, Plamann M, McCluskey K. Characterization of the temperature-sensitive mutations *un-7*
534 and *png-1* in *Neurospora crassa*. *PLoS One*. 2010; 5(5):e10703.

535 **Dress AWM**, Flamm C, Fritzsche G, Grünwald S, Kruspe M, Prohaska SJ, Stadler PF. Noisy: identification of
536 problematic columns in multiple sequence alignments. *Algorithms for Molecular Biology*. 2008; 3(1):7.

537 **Drummond AJ**, Ho SYW, Phillips MJ, Rambaut A. Relaxed phylogenetics and dating with confidence. *PLoS*
538 *Biology*. 2006; 4(5):e88.

539 **Elliot MG**, Mooers AØ. Inferring ancestral states without assuming neutrality or gradualism using a stable
540 model of continuous character evolution. *BMC Evolutionary Biology*. 2014; 14(1):226.

541 **Engqvist MK**. Correlating enzyme annotations with a large set of microbial growth temperatures reveals
542 metabolic adaptations to growth at diverse temperatures. *BMC Microbiology*. 2018; 18(1):177.

543 **Feil E**, Zhou J, Smith JM, Spratt BG. A comparison of the nucleotide sequences of the *adk* and *recA* genes of
544 pathogenic and commensal *Neisseria* species: Evidence for extensive interspecies recombination within *adk*.
545 *Journal of Molecular Evolution*. 1996; 43(6):631–640.

546 **Frishman D**, Argos P. Knowledge-based protein secondary structure assignment. *Proteins*. 1995; 23(4):566–579.

547 **García-Carreras B**, Sal S, Padfield D, Kontopoulos DG, Bestion E, Schaum CE, Yvon-Durocher G, Pawar S. Role
548 of carbon allocation efficiency in the temperature dependence of autotroph growth rates. *Proceedings of the*
549 *National Academy of Sciences*. 2018; 115(31):E7361–E7368.

550 **Gaucher EA**, Govindarajan S, Ganesh OK. Palaeotemperature trend for Precambrian life inferred from resur-
551 rected proteins. *Nature*. 2008; 451(7179):704–707.

552 **Gaucher EA**, Thomson JM, Burgan MF, Benner SA. Inferring the palaeoenvironment of ancient bacteria on the
553 basis of resurrected proteins. *Nature*. 2003; 425(6955):285–288.

554 **Gelman A**, Rubin DB. Inference from iterative simulation using multiple sequences. *Statistical Science*. 1992;
555 7(4):457–472.

556 **Gibson B**, Eyre-Walker A. Investigating evolutionary rate variation in bacteria. *Journal of Molecular Evolution*.
557 2019; 87(9-10):317–326.

558 **Gibson B**, Wilson DJ, Feil E, Eyre-Walker A. The distribution of bacterial doubling times in the wild. *Proceedings*
559 *of the Royal Society B: Biological Sciences*. 2018; 285(1880):20180789.

560 **Gilchrist GW**. Specialists and generalists in changing environments. I. Fitness landscapes of thermal sensitivity.
561 *The American Naturalist*. 1995; 146(2):252–270.

562 **Gillooly JF**, Allen AP, West GB, Brown JH. The rate of DNA evolution: effects of body size and temperature on the
563 molecular clock. *Proceedings of the National Academy of Sciences*. 2005; 102(1):140–145.

564 **Gillooly JF**, McCoy MW, Allen AP. Effects of metabolic rate on protein evolution. *Biology Letters*. 2007; 3(6):655–
565 659.

566 **Giovannoni SJ**, Thrash JC, Temperton B. Implications of streamlining theory for microbial ecology. *The ISME*
567 *Journal*. 2014; 8(8):1553–1565.

568 **Glykos NM**. Carma: a molecular dynamics analysis program. *Journal of Computational Chemistry*. 2006;
569 27(14):1765–1768.

570 **Grant BJ**, Rodrigues APC, ElSawy KM, McCammon JA, Caves LSD. Bio3d: An R package for the comparative
571 analysis of protein structures. *Bioinformatics*. 2006; 22(21):2695–2696.

572 **Gu X**, Fu YX, Li WH. Maximum likelihood estimation of the heterogeneity of substitution rate among nucleotide
573 sites. *Molecular Biology and Evolution*. 1995; 12(4):546–557.

574 **Guo Q**, Penman M, Trigatti BL, Krieger M. A single point mutation in ϵ -COP results in temperature-sensitive,
575 lethal defects in membrane transport in a Chinese hamster ovary cell mutant. *Journal of Biological Chemistry*.
576 1996; 271(19):11191–11196.

577 **Hadfield JD**. MCMC Methods for Multi-Response Generalized Linear Mixed Models: The MCMCglmm R Package.
578 *Journal of Statistical Software*. 2010; 33(2):1–22.

579 **Humphrey W**, Dalke A, Schulten K. VMD – Visual Molecular Dynamics. *Journal of Molecular Graphics*. 1996;
580 14:33–38.

581 **Katoh K**, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance
582 and usability. *Molecular Biology and Evolution*. 2013; 30(4):772–780.

583 **Kelley LA**, Mezulis S, Yates CM, Wass MN, Sternberg MJ. The Phyre2 web portal for protein modeling, prediction
584 and analysis. *Nature Protocols*. 2015; 10(6):845–858.

585 **Kimura M**. Molecular evolutionary clock and the neutral theory. *Journal of Molecular Evolution*. 1987; 26(1-
586 2):24–33.

587 **Kontopoulos DG**, van Seville E, Lange M, Yvon-Durocher G, Barraclough TG, Pawar S. Phytoplankton thermal
588 responses adapt in the absence of hard thermodynamic constraints. *Evolution*. 2020; 74(4):775–790.

589 **Kontopoulos DG**, Smith TP, Barraclough TG, Pawar S. Adaptive evolution shapes the present-
590 day distribution of the thermal sensitivity of population growth rate. *bioRxiv*. 2020;
591 <https://www.biorxiv.org/content/10.1101/712885v3>.

592 **Kontopoulos DG**, Vlachakis D, Tsiliki G, Kossida S. Structuprint: a scalable and extensible tool for two-
593 dimensional representation of protein surfaces. *BMC Structural Biology*. 2016; 16(1):4.

594 **Kovermann M**, Ådén J, Grundström C, Sauer-Eriksson AE, Sauer UH, Wolf-Watz M. Structural basis for catalyti-
595 cally restrictive dynamics of a high-energy enzyme state. *Nature Communications*. 2015; 6:7644.

596 **Kovermann M**, Grundström C, Sauer-Eriksson AE, Sauer UH, Wolf-Watz M. Structural basis for ligand binding to
597 an enzyme by a conformational selection pathway. *Proceedings of the National Academy of Sciences*. 2017;
598 114(24):6298–6303.

599 **Kumar MS**, Bava KA, Gromiha MM, Prabakaran P, Kitajima K, Uedaira H, Sarai A. ProTherm and ProNIT:
600 thermodynamic databases for proteins and protein–nucleic acid interactions. *Nucleic Acids Research*. 2006;
601 34(suppl_1):D204–D206.

602 **Lanfear R**, Calcott B, Ho SYW, Guindon S. PartitionFinder: combined selection of partitioning schemes and
603 substitution models for phylogenetic analyses. *Molecular Biology and Evolution*. 2012; 29(6):1695–1701.

604 **Lanfear R**, Frandsen PB, Wright AM, Senfeld T, Calcott B. PartitionFinder 2: new methods for selecting partitioned
605 models of evolution for molecular and morphological phylogenetic analyses. *Molecular Biology and Evolution*.
606 2017; 34(3):772–773.

607 **Lanfear R**, Thomas JA, Welch JJ, Brey T, Bromham L. Metabolic rate does not calibrate the molecular clock.
608 *Proceedings of the National Academy of Sciences*. 2007; 104(39):15388–15393.

609 **Laskowski RA**, MacArthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical
610 quality of protein structures. *Journal of Applied Crystallography*. 1993; 26(2):283–291.

611 **Li D**, Liu MS, Ji B. Mapping the dynamics landscape of conformational transitions in enzyme: the adenylate
612 kinase case. *Biophysical Journal*. 2015; 109(3):647–660.

613 **Li G**, Rabe KS, Nielsen J, Engqvist MK. Machine learning applied to predicting microorganism growth temperatures
614 and enzyme catalytic optima. *ACS Synthetic Biology*. 2019; 8(6):1411–1420.

615 **Long H**, Miller SF, Williams E, Lynch M. Specificity of the DNA mismatch repair system (MMR) and mutagenesis
616 bias in bacteria. *Molecular Biology and Evolution*. 2018; 35(10):2414–2421.

617 **Lynch M**, Ackerman MS, Gout JF, Long H, Sung W, Thomas WK, Foster PL. Genetic drift, selection and the
618 evolution of the mutation rate. *Nature Reviews Genetics*. 2016; 17(11):704.

619 **Maechler M**, Rousseeuw P, Struyf A, Hubert M, Hornik K. cluster: Cluster Analysis Basics and Extensions; 2017,
620 R package version 2.0.6.

621 **Martin TL**, Huey RB. Why “suboptimal” is optimal: Jensen’s inequality and ectotherm thermal preferences. *The*
622 *American Naturalist*. 2008; 171(3):E102–E118.

623 **Martyna GJ**, Klein ML, Tuckerman M. Nosé–Hoover chains: the canonical ensemble via continuous dynamics.
624 *The Journal of Chemical Physics*. 1992; 97(4):2635–2643.

625 **Martyna GJ**, Tobias DJ, Klein ML. Constant pressure molecular dynamics algorithms. *The Journal of Chemical*
626 *Physics*. 1994; 101(5):4177–4189.

627 **Mulkidjanian AY**, Bychkov AY, Dibrova DV, Galperin MY, Koonin EV. Origin of first cells at terrestrial, anoxic
628 geothermal fields. *Proceedings of the National Academy of Sciences*. 2012; 109(14):E821–E830.

629 **Nakagawa S**, Schielzeth H. A general and simple method for obtaining R^2 from generalized linear mixed-effects
630 models. *Methods in Ecology and Evolution*. 2013; 4(2):133–142.

631 **Nakashima H**, Fukuchi S, Nishikawa K. Compositional changes in RNA, DNA and proteins for bacterial adaptation
632 to higher and lower temperatures. *Journal of Biochemistry*. 2003; 133(4):507–513.

633 **Ndo C**, Poumachu Y, Metitsi D, Awono-Ambene HP, Tchuinkam T, Gilles JLR, Bourtzis K. Isolation and characteri-
634 zation of a temperature-sensitive lethal strain of *Anopheles arabiensis* for SIT-based application. *Parasites &*
635 *Vectors*. 2018; 11(2):659.

636 **Nguyen LT**, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for
637 estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution*. 2015; 32(1):268–274.

638 **Paradis E**. Molecular dating of phylogenies by likelihood methods: a comparison of models and a new
639 information criterion. *Molecular Phylogenetics and Evolution*. 2013; 67(2):436–444.

640 **Paradis E**, Schliep K. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R.
641 *Bioinformatics*. 2019; 35(3):526–528.

642 **Pearce BKD**, Tupper AS, Pudritz RE, Higgs PG. Constraining the time interval for the origin of life on Earth.
643 *Astrobiology*. 2018; 18(3):343–364.

- 644 **Pettersen EF**, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. UCSF Chimera—a visualiza-
645 tion system for exploratory research and analysis. *Journal of Computational Chemistry*. 2004; 25(13):1605–
646 1612.
- 647 **Privalov PL**. Cold denaturation of proteins. *Critical Reviews in Biochemistry and Molecular Biology*. 1990;
648 25(4):281–306.
- 649 **Puurtinen M**, Elo M, Jalasvuori M, Kahilainen A, Ketola T, Kotiaho JS, Mönkkönen M, Pentikäinen OT. Temperature-
650 dependent mutational robustness can explain faster molecular evolution at warm temperatures, affecting
651 speciation rate and global patterns of species diversity. *Ecography*. 2016; 39(11):1025–1033.
- 652 **Quan L**, Lv Q, Zhang Y. STRUM: structure-based prediction of protein stability changes upon single-point
653 mutation. *Bioinformatics*. 2016; 32(19):2936–2946.
- 654 **Quast C**, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO. The SILVA ribosomal RNA
655 gene database project: improved data processing and web-based tools. *Nucleic Acids Research*. 2012;
656 41(D1):D590–D596.
- 657 **Ranea JAG**, Grant A, Thornton JM, Orengo CA. Microeconomic principles explain an optimal genome size in
658 bacteria. *Trends in Genetics*. 2005; 21(1):21–25.
- 659 **Revell LJ**. phytools: an R package for phylogenetic comparative biology (and other things). *Methods in Ecology*
660 *and Evolution*. 2012; 3(2):217–223.
- 661 **Reynolds AP**, Richards G, de la Iglesia B, Rayward-Smith VJ. Clustering rules: a comparison of partitioning and
662 hierarchical clustering algorithms. *Journal of Mathematical Modelling and Algorithms*. 2006; 5(4):475–504.
- 663 **Robinson JL**, Pyzyna B, Atrasz RG, Henderson CA, Morrill KL, Burd AM, DeSoucy E, Fogleman RE, Naylor JB, Steele
664 SM, Elliott DR, Leyva KJ, Shand RF. Growth kinetics of extremely halophilic *Archaea* (family *Halobacteriaceae*) as
665 revealed by Arrhenius plots. *Journal of Bacteriology*. 2005; 187(3):923–929.
- 666 **Rost B**. Twilight zone of protein sequence alignments. *Protein Engineering*. 1999; 12(2):85–94.
- 667 **Rousseeuw PJ**. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of*
668 *Computational and Applied Mathematics*. 1987; 20:53–65.
- 669 **Savage VM**, Gillooly JF, Brown JH, West GB, Charnov EL. Effects of body size and temperature on population
670 growth. *The American Naturalist*. 2004; 163(3):429–441.
- 671 **Schoolfield RM**, Sharpe PJH, Magnuson CE. Non-linear regression of biological temperature-dependent rate
672 models based on absolute reaction-rate theory. *Journal of Theoretical Biology*. 1981; 88(4):719–731.
- 673 **Smith TP**, Thomas TJH, García-Carreras B, Sal S, Yvon-Durocher G, Bell T, Pawar S. Community-level respiration
674 of prokaryotic microbes may rise with global warming. *Nature Communications*. 2019; 10(5124):1–11.

675 **Sojo V**, Herschy B, Whicher A, Camprubi E, Lane N. The origin of life in alkaline hydrothermal vents. *Astrobiology*.
676 2016; 16(2):181–197.

677 **Spiegelhalter DJ**, Best NG, Carlin BP, van der Linde A. Bayesian measures of model complexity and fit. *Journal*
678 *of the Royal Statistical Society: Series B (Statistical Methodology)*. 2002; 64(4):583–639.

679 **Stamatakis A**. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies.
680 *Bioinformatics*. 2014; 30(9):1312–1313.

681 **Sugiura N**. Further analysis of the data by Akaike's information criterion and the finite corrections. *Communica-*
682 *tions in Statistics - Theory and Methods*. 1978; 7(1):13–26.

683 **The UniProt Consortium**. UniProt: the universal protein knowledgebase. *Nucleic Acids Research*. 2017;
684 45(D1):D158–D169.

685 **Tokuriki N**, Stricher F, Serrano L, Tawfik DS. How protein stability and new functions trade off. *PLoS Computa-*
686 *tional Biology*. 2008; 4(2):e1000002.

687 **Tokuriki N**, Tawfik DS. Stability effects of mutations and protein evolvability. *Current Opinion in Structural*
688 *Biology*. 2009; 19(5):596–604.

689 **Vonrhein C**, Bönisch H, Schäfer G, Schulz GE. The structure of a trimeric archaeal adenylate kinase. *Journal of*
690 *Molecular Biology*. 1998; 282(1):167–179.

691 **Westall F**, Hickman-Lewis K, Hinman N, Gautret P, Campbell K, Bréhéret JG, Foucher F, Hubert A, Sorieul S, Dass
692 AV, Kee TP, Georgelin T, Brack A. A hydrothermal-sedimentary context for the origin of life. *Astrobiology*. 2018;
693 18(3):259–293.

694 **Woolfit M**. Effective population size and the rate and pattern of nucleotide substitutions. *Biology Letters*. 2009;
695 5(3):417–420.

696 **Wu D**, Jospin G, Eisen JA. Systematic identification of gene families for use as “markers” for phylogenetic and
697 phylogeny-driven ecological studies of bacteria and archaea and their major subgroups. *PLoS One*. 2013;
698 8(10):e77033.

699 **Zeldovich KB**, Berezhovsky IN, Shakhnovich EI. Protein and DNA sequence determinants of thermophilic
700 adaptation. *PLoS Computational Biology*. 2007; 3(1):e5.

701 **Zeldovich KB**, Chen P, Shakhnovich EI. Protein stability imposes limits on organism complexity and speed of
702 molecular evolution. *Proceedings of the National Academy of Sciences*. 2007; 104(41):16152–16157.

703 Appendix 1

I. The Gibbs-Helmholtz equation

If all environmental variables other than temperature (e.g., pH, salt concentration) are kept constant, the relationship between protein stability (ΔG) and temperature (T) for proteins unfolding according to the two-state model can be described by the Gibbs-Helmholtz equation (Privalov, 1990):

$$\Delta G(T) = \Delta H_m \cdot \left[1 - \frac{T}{T_m} \right] - \Delta C_p \cdot \left[T_m - T + T \cdot \ln \left(\frac{T}{T_m} \right) \right] \quad (1)$$

Here, T_m (the “melting temperature”, in K units) is the temperature at which $\Delta G = 0$ due to heat denaturation. At T_m , the protein population is equally distributed between the two states (folded and unfolded). ΔH_m (kcal · mol⁻¹) is the difference in enthalpy between the folded and unfolded states at T_m . ΔC_p (kcal · mol⁻¹ · K⁻¹) is the difference in heat capacity between the two states, which is assumed to be independent of temperature.

Appendix 1—figure 1. An example of a protein stability curve according to the Gibbs-Helmholtz equation. The temperature is shown in °C for clarity but in the equation it is in K units. Note that ΔH_m , which mainly controls the maximum height of the stability curve, is normalised at $T = T_m$.

II. List of analysed prokaryotic ADKs

Appendix 1—table 1. Prokaryotic ADKs that were included in this study.

Species	UniProt accession ID	Thermal group
<i>Acidothermus cellulolyticus</i>	A0LRP1	thermophile
<i>Anoxybacillus flavithermus</i>	B7GJ88	thermophile

<i>Aquifex aeolicus</i>	O66490	hyperthermophile
<i>Bacillus subtilis</i>	P16304	mesophile
<i>Bacillus thuringiensis</i>	A0R8K1	mesophile
<i>Bordetella petrii</i>	A9IQ39	mesophile
<i>Burkholderia pseudomallei</i>	Q3JVB1	mesophile
<i>Cenarchaeum symbiosum</i>	A0RUE3	psychrophile
<i>Chlamydomonas reinhardtii</i>	Q822Z1	mesophile
<i>Clostridium perfringens</i>	Q0TMR7	mesophile
<i>Clostridium thermocellum</i>	A3DJJ3	thermophile
<i>Colwellia psychrerythraea</i>	Q47XA8	psychrophile
<i>Corynebacterium urealyticum</i>	B1VEX6	mesophile
<i>Cupriavidus pinatubonensis</i>	Q475G1	mesophile
<i>Dehalococcoides mccartyi</i>	Q3Z960	mesophile
<i>Desulfovibrio gigas</i>	C7U112	mesophile
<i>Devosia psychrophila</i>	A0A0F5PRB6	psychrophile
<i>Escherichia coli</i>	P69441	mesophile
<i>Flavobacterium psychrophilum</i>	A6GZB9	psychrophile
<i>Francisella philomiragia</i>	B0TXR2	mesophile
<i>Francisella tularensis</i>	Q5NFR4	mesophile
<i>Gemmatimonas aurantiaca</i>	C1A6S6	mesophile
<i>Geobacillus stearothermophilus</i>	P27142	thermophile
<i>Geobacter daltonii</i>	B9M6F7	mesophile
<i>Geoglobus ahangari</i>	A0A0F7IDB9	hyperthermophile
<i>Halarchaeum acidiphilum</i>	U3AED7	mesophile
<i>Korarchaeum cryptofilum</i>	B1L775	hyperthermophile
<i>Lactobacillus gasseri</i>	Q046A5	mesophile
<i>Lutispora thermophila</i>	A0A1M6DWR4	thermophile
<i>Mesoplasma florum</i>	Q6F1X3	mesophile
<i>Metallosphaera sedula</i>	A4YCZ0	thermophile
<i>Methanococcus aeolicus</i>	A6UW35	mesophile
<i>Methanococcus maripaludis</i>	A4FXE7	mesophile
<i>Methanococcus voltae</i>	P43411	mesophile
<i>Methanopyrus kandleri</i>	Q8TZB0	hyperthermophile
<i>Methanosarcina acetivorans</i>	Q8TRS3	mesophile
<i>Methanothermobacter thermautotrophicus</i>	O26135	thermophile
<i>Methanothermococcus thermolithotrophicus</i>	P43410	thermophile

<i>Methanotorris igneus</i>	P43408	hyperthermophile
<i>Methylobacillus flagellatus</i>	Q1GZB7	mesophile
<i>Moorella thermoacetica</i>	Q2RFR8	thermophile
<i>Mycobacterium tuberculosis</i>	P9WKF5	mesophile
<i>Mycobacterium ulcerans</i>	A0PM98	mesophile
<i>Natranaerobius thermophilus</i>	B2A4G2	thermophile
<i>Nitrosococcus oceani</i>	Q3J7A3	mesophile
<i>Octadecabacter arcticus</i>	M9RLZ6	psychrophile
<i>Photobacterium profundum</i>	Q6LTE1	psychrophile
<i>Photorhabdus luminescens</i>	Q7N0P5	mesophile
<i>Polaromonas naphthalenivorans</i>	A1VVK2	mesophile
<i>Prochlorococcus marinus</i>	A8G742	mesophile
<i>Pseudomonas fluorescens</i>	C3KEC6	mesophile
<i>Psychromonas ingrahamii</i>	A1STI3	psychrophile
<i>Rhodferax ferrireducens</i>	Q21TN6	mesophile
<i>Rhodospirillum centenum</i>	B6IRS7	mesophile
<i>Rickettsia akari</i>	A8GPC9	mesophile
<i>Sporosarcina globispora</i>	P84139	psychrophile
<i>Streptococcus pneumoniae</i>	B8ZKQ1	mesophile
<i>Sulfolobus acidocaldarius</i>	P35028	thermophile
<i>Sulfolobus solfataricus</i>	Q9UX83	hyperthermophile
<i>Thermoanaerobacter pseudethanolicus</i>	B0KCM0	thermophile
<i>Thermococcus sibiricus</i>	C6A2Q1	hyperthermophile
<i>Thermomonas hydrothermalis</i>	A0A1M5AE38	thermophile
<i>Thermoplasma acidophilum</i>	Q9HIT1	thermophile
<i>Thermosipho africanus</i>	B7IHW7	thermophile
<i>Thermotoga maritima</i>	Q9X1I8	hyperthermophile
<i>Thermus thermophilus</i>	Q5SHQ9	thermophile
<i>Vibrio cholerae</i>	Q9KTB7	mesophile
<i>Vibrio vulnificus</i>	Q7MMR5	mesophile
<i>Wolbachia pipientis</i>	Q73HA6	mesophile
<i>Xylella fastidiosa</i>	B2I757	mesophile

III. Homology modelling of prokaryotic ADKs

Appendix 1—figure 2. Percentage of sequence identity versus alignment length for all template sequences of this study. Note that many of our homology models were built using multiple templates and, therefore, multiple data points can correspond to the same homology model. Data points at 100% sequence identity indicate ADKs whose structures were available from the PDB, and for which homology modelling was done only to add any missing amino acids or not at all.

Appendix 1—figure 3. Distribution of amino acids across the regions of the Ramachandran plot for each ADK structure of our dataset. Boxplot edges represent the first and third quartiles, whereas the solid line stands for the median. Whiskers extend up to the most remote data point within 1.5 interquartile ranges from each boxplot edge. Given that all structures had >94% of their amino acids in the most favoured and additional allowed regions, and at worst just over 3% of amino acids in disallowed regions, their quality can be considered sufficiently high.

IV. Species tree reconstruction

Appendix 1—table 2 lists the genes from which we reconstructed the species tree. The best partitioning scheme is shown in Appendix 1—table 3. It is worth noting that the two ADK types were assigned to two different partitions and models. The resulting species phylogeny is shown in Appendix 1—figure 4. Finally, Appendix 1—figures 5 and 6 show convergence diagnostics for the time-calibration of the phylogeny with BEAST 2.

Appendix 1—table 2. Molecular sequences used for species tree inference.

ID	Description	Type
16S	Small subunit rRNA	Nucleotide
23S	Large subunit rRNA	Nucleotide
ADK (ancestral type)	Adenylate kinase	Amino acid
ADK (archaeal type)	Adenylate kinase	Amino acid
PhyEco_BA_01	Ribosomal protein S2	Amino acid
PhyEco_BA_02	Ribosomal protein S10	Amino acid
PhyEco_BA_03	Ribosomal protein L1	Amino acid
PhyEco_BA_04	Translation elongation factor EF-2	Amino acid
PhyEco_BA_05	Translation initiation factor IF-2	Amino acid
PhyEco_BA_06	Metalloendopeptidase	Amino acid
PhyEco_BA_07	Ribosomal protein L22	Amino acid
PhyEco_BA_08	Ffh signal recognition particle protein	Amino acid
PhyEco_BA_09	Ribosomal protein L4/L1e	Amino acid
PhyEco_BA_10	Ribosomal protein L2	Amino acid
PhyEco_BA_11	Ribosomal protein S9	Amino acid
PhyEco_BA_12	Ribosomal protein L3	Amino acid
PhyEco_BA_13	Phenylalanyl-tRNA synthetase β subunit	Amino acid
PhyEco_BA_14	Ribosomal protein L14b/L23e	Amino acid
PhyEco_BA_15	Ribosomal protein S5	Amino acid

PhyEco_BA_16	Ribosomal protein S19	Amino acid
PhyEco_BA_17	Ribosomal protein S7	Amino acid
PhyEco_BA_18	Ribosomal protein L16/L10E	Amino acid
PhyEco_BA_19	Ribosomal protein S13	Amino acid
PhyEco_BA_20	Phenylalanyl-tRNA synthetase α subunit	Amino acid
PhyEco_BA_21	Ribosomal protein L15	Amino acid
PhyEco_BA_22	Ribosomal protein L25/L23	Amino acid
PhyEco_BA_23	Ribosomal protein L6	Amino acid
PhyEco_BA_24	Ribosomal protein L11	Amino acid
PhyEco_BA_25	Ribosomal protein L5	Amino acid
PhyEco_BA_26	Ribosomal protein S12/S23	Amino acid
PhyEco_BA_27	Ribosomal protein L29	Amino acid
PhyEco_BA_28	Ribosomal protein S3	Amino acid
PhyEco_BA_29	Ribosomal protein S11	Amino acid
PhyEco_BA_30	Ribosomal protein L10	Amino acid
PhyEco_BA_31	Ribosomal protein S8	Amino acid
PhyEco_BA_32	tRNA pseudouridine synthase B	Amino acid
PhyEco_BA_33	Ribosomal protein L18P/L5E	Amino acid
PhyEco_BA_34	Ribosomal protein S15P/S13e	Amino acid
PhyEco_BA_35	Porphobilinogen deaminase	Amino acid
PhyEco_BA_36	Ribosomal protein S17	Amino acid
PhyEco_BA_37	Ribosomal protein L13	Amino acid
PhyEco_BA_38	Phosphoribosylformylglycinamide cyclo-ligase	Amino acid
PhyEco_BA_39	Ribonuclease HII	Amino acid
PhyEco_BA_40	Ribosomal protein L24	Amino acid

Appendix 1—table 3. Optimal partitioning scheme according to AICc.

Partition #	Gene ID	Model
1	16S	GTR + Γ + I
2	23S	GTR + Γ + I
3	ADK (ancestral type)	WAG + Γ + I
4	ADK (archaeal type)	CPREV + Γ
5	PhyEco_BA_01	WAG + Γ + I
	PhyEco_BA_15	
	PhyEco_BA_22	
	PhyEco_BA_23	
	PhyEco_BA_28	
6	PhyEco_BA_02	RTREV + Γ + I
	PhyEco_BA_14	
7	PhyEco_BA_03	WAG + Γ + I
	PhyEco_BA_17	
	PhyEco_BA_18	
	PhyEco_BA_25	
	PhyEco_BA_31	
	PhyEco_BA_37	
8	PhyEco_BA_04	WAG + Γ + I
	PhyEco_BA_16	
9	PhyEco_BA_05	WAG + Γ + I
10	PhyEco_BA_06	WAG + Γ + I
	PhyEco_BA_20	
11	PhyEco_BA_07	VT + Γ + I

12	PhyEco_BA_08	WAG + Γ + I
	PhyEco_BA_09	
	PhyEco_BA_12	
13	PhyEco_BA_33	WAG + Γ + I
	PhyEco_BA_40	
14	PhyEco_BA_10	RTREV + Γ + I
	PhyEco_BA_11	
15	PhyEco_BA_38	WAG + Γ + I
16	PhyEco_BA_13	WAG + Γ + I
17	PhyEco_BA_19	RTREV + Γ + I
	PhyEco_BA_21	
18	PhyEco_BA_36	VT + Γ + I
	PhyEco_BA_24	
19	PhyEco_BA_34	WAG + Γ + I
20	PhyEco_BA_26	WAG + Γ
21	PhyEco_BA_27	VT + Γ + I
22	PhyEco_BA_29	RTREV + Γ
23	PhyEco_BA_30	WAG + Γ + I
24	PhyEco_BA_32	VT + Γ + I
25	PhyEco_BA_35	WAG + Γ + I
26	PhyEco_BA_39	BLOSUM62 + Γ + I

Appendix 1—figure 4. The maximum likelihood phylogeny of the species in our dataset of ADKs, as inferred with IQ-TREE. Each phylum is shown in a different colour. Additionally, the phyla that belong to the TACK superphylum are explicitly marked. In contrast to the ADK gene tree, most nodes of the species tree have sufficiently high statistical support (black and grey circles).

Appendix 1—figure 5. Effective sample size and potential scale reduction factor values for each parameter of the evolutionary model fitted with BEAST 2.

Appendix 1—figure 6. Effective sample size and potential scale reduction factor values for each branch-specific substitution rate, as estimated with BEAST 2.

V. Analysis of molecular dynamics trajectories

Appendix 1—figure 7. Illustration of the procedure for obtaining representative ADK conformations, based on the molecular dynamics trajectories of *Francisella philomiragia*. A: The RMSD matrix among all snapshots is calculated after excluding highly flexible amino acids. B: Conformations are grouped into major structural clusters, visualised here using multidimensional scaling. Note that the reported goodness of fit refers to the ability of two dimensions to properly reflect the RMSD values among data points, and is not a measure of clustering quality. C: From each cluster, an average structure is produced. The flexibility of amino acids for each conformation is shown on a colour scale that ranges from dark orange (highly flexible) to white to dark blue (least flexible). The percentages show the proportion of protein snapshots that belong to each conformation.

VI. Investigation of biases in the estimates of mutational effects

Appendix 1—figure 8. $\widetilde{\Delta\Delta G}$ estimates were not found to be influenced by either the radius of gyration (panel A) or the quality of the homology model (panel B).

VII. The inter- and intraspecific relationships between median mutational effects and temperature

Appendix 1—table 4 lists the models that were fitted to characterise the relationship between $\widetilde{\Delta\Delta G}$ and temperature across species. Models with at least one coefficient whose 95% HPD interval included zero are not shown.

Appendix 1—table 4. Candidate models for the interspecific relationship between $\widetilde{\Delta\Delta G}$ and temperature. The model with the lowest DIC is highlighted.

Model	Phylogenetic correction	Mean DIC
$\widetilde{\Delta\Delta G} \sim \text{Intercept}$	✓	-27.1691
$\widetilde{\Delta\Delta G} \sim \text{Intercept} + \text{Temperature}$	✓	-32.4251
$\widetilde{\Delta\Delta G} \sim \text{Intercept} + \text{ADK type}$	✓	-30.8497
$\widetilde{\Delta\Delta G} \sim \text{Intercept} + \text{Temperature} + \text{ADK type}$	✓	-35.3310
$\widetilde{\Delta\Delta G} \sim \text{Intercept}$	✗	-28.4569
$\widetilde{\Delta\Delta G} \sim \text{Intercept} + \text{LID length}$	✗	-33.9572
$\widetilde{\Delta\Delta G} \sim \text{Intercept} + \text{ADK type}$	✗	-53.9365
$\widetilde{\Delta\Delta G} \sim \text{Intercept} + \text{Temperature} + \text{ADK type}$	✗	-63.9542

To estimate the species-specific relationships between median $\Delta\Delta G$ and temperature, we

fitted six mathematical equations to the data (Appendix 1—figure 9). The most appropriate shape for each species (Figure 5) was determined using AICc.

Appendix 1—figure 9. The six mathematical models that we fitted to the dataset of median $\Delta\Delta G$ versus temperature for each ADK. These models capture a wide range of shapes, from a linear decrease to nonlinear changes or even a constant value across temperatures.

Due to the variation in the mathematical form of the fitted equations, it was not possible to directly compare the relationships on the basis of model coefficients. For an objective comparison, we instead executed PCA on a matrix of predicted median $\Delta\Delta G$ values (for each ADK) at temperatures from 6.85°C to 94.35°C with a step of 0.05°C. This showed that species-specific relationships can be adequately described using two principal components, on which we also projected the phylogeny (Appendix 1—figure 10). The first component captures 95.4% of the variance and stands for the average robustness to mutations across all temperatures. The second principal component represents the temperature sensitivity of the median $\Delta\Delta G$. The values of internal nodes (except for the node corresponding to LUCA) were estimated using the stable model of trait evolution (*Elliot and Mooers, 2014*). More precisely, we executed four MCMC chains for 1.5 billion generations, sampling every 2,000 generations after the first 375 million generations.

Appendix 1—figure 10. Projection of the relationships between median $\Delta\Delta G$ and temperature to the first two principal components. Most of the variance (PC1 axis) corresponds to shifts in mutational robustness (i.e., changes along the vertical axis of plots of Figure 5). A much smaller amount of variance (PC2 axis) is due to variation in the temperature sensitivity of median $\Delta\Delta G$ among ADKs of different species.

It is worth noting that the ADK of the LUCA had the lowest temperature sensitivity across the six species, suggesting a likely evolutionary trend towards increasing temperature sensitivity. However, given that i) the ADK gene is under strong selection and that ii) it evolves rapidly and in a convergent manner across species (Figure 3A), such a trend in temperature sensitivity could be an artefact of the ancestral reconstruction of the LUCA sequence. To understand whether this trend is real, future studies could perform a similar analysis for genes that are characterised by non-convergent and less rapid evolution.

VIII. The ADK of the Last Universal Common Ancestor

VIII.I Estimated sequence and 3D structure

To reconstruct the ADK sequence of the LUCA, we submitted to the FASTML server (*Ashkenazy et al., 2012*) the alignment of ancestral-type ADKs, the species tree, and the most appropriate model of sequence substitution (WAG; Appendix 1—table 3). The resulting sequence is as follows:

```

10      20      30      40      50      60      70      80
MNIILLGPPGAGKGTQAKRLAEKLGIPHISTGDI FREAVKKGTELGKKAKEYMDRGELVPDEVVIGIVKERLSQPDCKN
GFILDGFPRTVKQAEALDEMLDEKGKKLDAVINFEVPDEVVIKRISGRRICPECGRVYHIIYNPPKVDGVCDVCGGELY
QREDDKEETVRKRLEVYREQTEPLIDYYKKKGILKNIDGTGSIEEVYKQILSVLGKSKK

```

Its homology model that we obtained via Phyre2 is shown in Appendix 1—figure 11.

Appendix 1—figure 11. The inferred ADK structure of the LUCA. Helices are shown in orange, whereas β -strands are in blue. The LID domain is long, similarly to that of most ancestral-type ADKs.

VIII.II Estimation of the temperature of optimal catalysis (T_{opt})

In microbes, the temperature of optimal catalysis of a given enzyme (T_{opt}) has been found to correlate with the temperature of maximum growth rate (T_{pk} ; *Engqvist 2018; Li et al. 2019*). The strength of this correlation varies across enzymes, possibly because the activity of certain enzymes (e.g., ADKs; *Couñago and Shamoo 2005; Couñago et al. 2006*) is more tightly linked to organismal fitness than others. Thus, estimating T_{opt} for the ADK of the LUCA should provide a rough understanding of the environmental conditions in which LUCA emerged.

Li et al. (2019) showed that the T_{opt} of an enzyme can be relatively accurately predicted using a random forest model. The inputs of the model are i) amino acid frequencies for the enzyme of choice and ii) the organism's growth rate T_{pk} . The model can also be trained without T_{pk} data, but at a cost of slightly worse performance. To this end, we used the code and training dataset of *Li et al. (2019)* to build a random forest model that would predict T_{opt} from amino acid frequencies alone. The resulting model achieved an R^2 of 0.9193 (RMSE = 5.328) against its training dataset. We next used this model to infer the likely T_{opt} value of the ADK sequence of the LUCA (Appendix 1—figure 12).

Appendix 1—figure 12. T_{opt} values of prokaryotic ADKs against the temperature of maximum growth rate (T_{pk}). The dashed line is the one-to-one line, the estimated T_{opt} of the LUCA is shown in blue, whereas the 95% prediction interval is shown in grey. Even though the characteristics of the thermal environment of the LUCA remain under debate, the close relationship between T_{opt} and T_{pk} in four other species suggests that LUCA was likely a moderate thermophile. The data points were taken from *Couñago and Shamoo (2005)* and *Li et al. (2019)*.

IX. Estimation of cell volume from cell dimension measurements

We estimated the cell volume (V) of spherical microbes (Appendix 1—figure 13A) as

$$V = \frac{\pi}{6} \cdot \text{diameter}^3.$$

For microbes whose shape was reported as similar to a disc, a filament, or a cylinder (Appendix 1—figure 13B), we used the following equation:

$$V = \frac{\pi}{4} \cdot \text{width}^2 \cdot \text{length}.$$

Last, for rod-shaped microbes (Appendix 1—figure 13C), we estimated their cell volume as the sum of the volumes of a cylinder and two half-spheres:

$$V = \frac{\pi}{6} \cdot \text{diameter}^3 + \frac{\pi}{4} \cdot \text{width}^2 \cdot (\text{length} - \text{diameter}).$$

Appendix 1—figure 13. The three-dimensional geometrical shapes based on which we estimated microbial cell volumes in this study. Note that in panel C, the diameter of the two half-spheres is equal to the width of the cell.

X. Estimation of prokaryotic generation times as a function of temperature and/or cell volume

X.I Temperature

The relationship between population growth rate and temperature in ectotherms has the shape of a unimodal curve (the thermal performance curve; *Angilletta 2009*; Appendix 1—figure 14A). As previously mentioned, the temperature at which the curve peaks is defined as T_{pk} , whereas the population growth rate at T_{pk} is B_{pk} . T_{pk} tends to be slightly above the mean environmental temperature (*Martin and Huey, 2008; Angilletta, 2009*). Calculating the inverse of B_{pk} gives an estimate of the minimum generation time (t_{gen}) of a microbial species or strain.

As *Smith et al. (2019)* have shown, in prokaryotes, B_{pk} tends to increase with T_{pk} until $\approx 45^{\circ}\text{C}$, after which it levels off. To quantify this relationship in a continuous manner, we fitted a few alternative models to their dataset. More precisely, we specified $\ln(B_{pk})$ as the response variable, whereas the predictor was T_{pk} , either using a natural logarithm transformation or a square root transformation. We note that we did not use a Boltzmann-Arrhenius term, even if the Metabolic Theory of Ecology framework also predicts an exponential decrease of t_{gen}

with temperature (*Savage et al., 2004*), due to the implicit assumption that the trait being studied (i.e., t_{gen}) is directly governed by the influence of temperature on enzyme kinetics. Such an assumption has been previously proven invalid (e.g., see *García-Carreras et al. 2018*, *Kontopoulos et al. 2020a*, and *Kontopoulos et al. 2020b*). Species identity was accounted for, as a random effect on the intercept. For each model, we fitted both a non-phylogenetic and a phylogenetically-corrected variant, in the latter case using the relative chronogram of *Kontopoulos et al. (2020b)*. All models were fitted with MCMCglmm, by running two independent chains per model for five million generations, with parameter samples being obtained from the posterior distribution every thousand generations. Samples from the first 500,000 generations were treated as burn-in and were discarded. Convergence was examined as previously described. The best-fitting model was determined according to DIC. To examine the proportion of variance explained by the fixed effects alone ($\text{Var}_{\text{fixed}}$), by the random effect ($\text{Var}_{\text{random}}$), or left unexplained ($\text{Var}_{\text{resid}}$), we calculated the marginal and conditional R^2 values (*Nakagawa and Schielzeth, 2013*):

$$R_m^2 = \frac{\text{Var}_{\text{fixed}}}{\text{Var}_{\text{fixed}} + \text{Var}_{\text{random}} + \text{Var}_{\text{resid}}}, \quad (2)$$

$$R_c^2 = \frac{\text{Var}_{\text{fixed}} + \text{Var}_{\text{random}}}{\text{Var}_{\text{fixed}} + \text{Var}_{\text{random}} + \text{Var}_{\text{resid}}}. \quad (3)$$

The results are shown in Appendix 1—table 5 and Appendix 1—figure 14.

Appendix 1—table 5. Alternative models for the relationship of $\ln(B_{\text{pk}})$ with T_{pk} across prokaryotes. The model with the lowest DIC is highlighted.

Model	Phylogenetic correction	Mean DIC
$\ln(B_{\text{pk}}) \sim \text{Intercept} + \ln(T_{\text{pk}})$	✓	494.3650
$\ln(B_{\text{pk}}) \sim \text{Intercept} + \sqrt{T_{\text{pk}}}$	✓	496.6180
$\ln(B_{\text{pk}}) \sim \text{Intercept} + \ln(T_{\text{pk}})$	✗	473.9934
$\ln(B_{\text{pk}}) \sim \text{Intercept} + \sqrt{T_{\text{pk}}}$	✗	475.6145

Appendix 1—figure 14. The relationships of maximum population growth rate (B_{pk}) and minimum generation time (t_{gen}) with T_{pk} . A: The thermal performance curve of population growth rate of the archaeum *Natrinema pellirubrum* (Robinson et al., 2005). Gold circles stand for the experimental measurements, whereas the black line is the fit of the four-parameter variant of the four-parameter Sharpe-Schoolfield model (Schoolfield et al., 1981). T_{pk} is the temperature at which growth rate takes its maximum value, B_{pk} . The left vertical axis is in units of s^{-1} , whereas the right vertical axis is in myr^{-1} . B: In prokaryotes, B_{pk} increases with the natural logarithm of T_{pk} . This explains 13% of variance, whereas species identity leads to 92% of the variance being explained. The equation that is shown has B_{pk} in units of s^{-1} and T_{pk} in $^{\circ}\text{C}$. C: t_{gen} decreases as a power law of T_{pk} .

X.II Cell volume

We further examined whether cell volume also influences the minimum generation times of prokaryotes. To this end, we estimated cell volumes for as many species in the dataset of *Smith et al. (2019)* as possible. We then fitted models with MCMCglmm that had $\ln(B_{pk})$ as the response and $\ln(T_{pk})$ and/or $\ln(\text{cell volume})$ as fixed effects. Species identity was again specified as a random effect on the intercept, both with and without a phylogenetic correction. Given that the cell volume of most species was available as ranges and not as a single value, we incorporated the uncertainty in our models by fitting them 30 times. For each of the thirty runs, we set each species' volume by sampling from a uniform distribution with bounds equal to the minimum and maximum cell volume of the species. Two independent MCMC chains were run for each model, for two million generations. Samples from the posterior were obtained every thousand generations after the first two hundred thousand generations and were examined for sufficient convergence. Model selection was done according to DIC, averaged across the thirty runs.

Models with phylogenetic correction had higher DIC values than their non-phylogenetic equivalents. Across the latter, the 95% HPD interval of the coefficient of $\ln(\text{cell volume})$ almost always included zero. This was the case both when $\ln(\text{cell volume})$ was the only fixed effect (Appendix 1—figure 15A) and when it was accompanied by $\ln(T_{pk})$ (Appendix 1—figure 15B). This shows that cell volume does not systematically influence B_{pk} across prokaryotes.

Appendix 1—figure 15. Investigation of a possible effect of cell volume on B_{pk} and, by extension, on generation time. Circles show the mean of the posterior distribution of the exponent of cell volume, whereas whiskers indicate the 95% HPD interval. The estimates in panel A were obtained through a regression of $\ln(B_{pk})$ against the natural logarithm of cell volume. In contrast, in panel B, the regression also contained the natural logarithm of T_{pk} as another fixed effect.

XI. The relationship between substitution rate and temperature

According to the neutral theory of molecular evolution (*Kimura, 1987*), the rate of molecular substitution per time (K) can be predicted by the generation time (t_{gen}), the mutation rate per generation (U), and the proportion of selectively neutral mutations (f_0):

$$K = \frac{U \cdot f_0}{t_{gen}}. \quad (4)$$

This equation makes the assumption that selectively beneficial mutations are extremely

rare and can be ignored, i.e., $1 - f_0$ stands for the proportion of deleterious mutations. Temperature influences t_{gen} (Appendix 1—figure 14), but also potentially the numerator of Appendix 1—equation (4), given that mutations become more detrimental with temperature as we show in this study. Thus, f_0 and possibly U should decrease with temperature. In contrast, *Gillooly et al. (2005, 2007)* have argued that, similarly to mass-specific metabolic rate, U should increase with temperature and associate (positively in prokaryotes, negatively in protists and metazoans; *DeLong et al. 2010*) with body size (or cell volume). Such a pattern would be consistent with the “metabolic rate hypothesis” which posits that mutation rate depends on metabolic rate, due to the production of oxidative free radicals as a byproduct of metabolism. Thus, to estimate the potential effects of cell volume (V) and temperature (T) on the numerator of Appendix 1—equation (4), the equation can be written as

$$K = \frac{\gamma \cdot V^\alpha \cdot e^{g(T)}}{t_{\text{gen}}} \quad (5)$$

Here, γ is a species-specific normalisation constant that is independent of cell volume and temperature, α is a dimensionless exponent, whereas $g(T)$ a temperature dependence function for which we used either a logarithmic ($g(T) = \zeta \cdot \ln(T)$) or a quadratic equation ($g(T) = \eta \cdot T + \theta \cdot T^2$). It is worth noting that we did not model temperature as a Boltzmann-Arrhenius function (*Gillooly et al., 2007*) because the latter implies that substitution rates are directly governed by the effects of temperature on the activity of underlying enzymes. This assumption has been previously shown to be invalid for metazoans (*Lanfear et al., 2007*).

Appendix 1—equation (5) can be simplified by removing T and/or V if these are found to have no systematic effect on K . To understand if and how temperature drives substitution rates, we multiplied both sides of Appendix 1—equation (5) with t_{gen} and took the natural logarithm of both sides:

$$\ln(K \cdot t_{\text{gen}}) = \ln(\gamma) + \alpha \cdot \ln(V) + g(T). \quad (6)$$

The mean DIC values of alternative fitted models are shown in Appendix 1—table 6. Models which included the natural logarithm of cell volume as a fixed effect are not shown, as the coefficient of cell volume always had a 95% HPD interval that included zero.

Appendix 1—table 6. Candidate models for the estimation of generation time-corrected substitution rates. The model with the lowest DIC is highlighted.

Model	Phylogenetic correction	Mean DIC
$\ln(K \cdot t_{\text{gen}}) \sim \text{Intercept}$	✓	161.3176
$\ln(K \cdot t_{\text{gen}}) \sim \text{Intercept} + \ln(\text{Temperature})$	✓	29.3025
$\ln(K \cdot t_{\text{gen}}) \sim \text{Intercept} + \text{Temperature} + \text{Temperature}^2$	✓	15.2959
$\ln(K \cdot t_{\text{gen}}) \sim \text{Intercept}$	✗	188.1174
$\ln(K \cdot t_{\text{gen}}) \sim \text{Intercept} + \ln(\text{Temperature})$	✗	24.9966
$\ln(K \cdot t_{\text{gen}}) \sim \text{Intercept} + \text{Temperature} + \text{Temperature}^2$	✗	27.1478

XII. Literature data on the effects of temperature on mutation rate

Appendix 1—figure 16. Mutation rate measurements for bacteria adapted to different thermal environments. Each data point corresponds to a different species, whereas colors represent the experimental approach. Mutation accumulation experiments (currently the most accurate method; *Gibson et al. 2018*) are shown in orange. Boxplot edges represent the first and third quartiles, whereas the solid line stands for the median. Whiskers extend up to the most remote data point within 1.5 interquartile ranges from each boxplot edge. Most measurements are for species adapted to mesophilic environments and, therefore, no conclusion on the effect of temperature can be drawn. Data points were obtained from *Lynch et al. (2016)*; *Gibson et al. (2018)*; *Long et al. (2018)*; *Gibson and Eyre-Walker (2019)*.