


Research Article

# A meta-analysis of the activity, stability, and mutational characteristics of temperature-adapted enzymes

 Stewart Gault<sup>1</sup>, Peter M. Higgins<sup>1,2</sup>, Charles S. Cockell<sup>1</sup> and Kaitlyn Gillies<sup>1</sup>

<sup>1</sup>UK Centre for Astrobiology, SUPA School of Physics and Astronomy, University of Edinburgh, James Clerk Maxwell Building, Peter Guthrie Tait Road, Edinburgh EH9 3FD, U.K.;

<sup>2</sup>Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, U.K.

**Correspondence:** Stewart Gault (s.a.gault@sms.ed.ac.uk)



Understanding the characteristics that define temperature-adapted enzymes has been a major goal of extremophile enzymology in recent decades. In the present study, we explore these characteristics by comparing psychrophilic, mesophilic, and thermophilic enzymes. Through a meta-analysis of existing data, we show that psychrophilic enzymes exhibit a significantly larger gap ( $T_g$ ) between their optimum and melting temperatures compared with mesophilic and thermophilic enzymes. These results suggest that  $T_g$  may be a useful indicator as to whether an enzyme is psychrophilic or not and that models of psychrophilic enzyme catalysis need to account for this gap. Additionally, by using predictive protein stability software, HoTMuSiC and PoPMuSiC, we show that the deleterious nature of amino acid substitutions to protein stability increases from psychrophiles to thermophiles. How this ultimately affects the mutational tolerance and evolutionary rate of temperature adapted organisms is currently unknown.

## Introduction

Extremophiles on Earth have become adapted to both high and low 'extreme' environmental temperatures. In the process of evolving to survive in such environments, they have had to adapt their biomolecular machinery to function at extreme environmental temperatures [1–3]. As enzymes are the major facilitators of biological reactions, they represent an important window through which temperature adaptation of organisms can be understood. Temperature-adapted enzymes exhibit adaptations to both their activity and stability. Psychrophilic enzymes, those adapted to low temperature environments, exhibit greater activity at low temperatures compared with mesophilic and thermophilic enzymes [4]. Thermophilic enzymes on the other hand are adapted to be both active and stable at extremely high environmental temperatures, even upwards of 100°C [5–7]. These adaptations are achieved through specific changes to an enzyme's amino acid composition [8–13], secondary structure [14], and the number and type of intramolecular bonds present in the enzyme [15–19]. In this study, thermophilic (from environments of ~55–60°C) and hyperthermophilic (from environments >80°C) enzymes are grouped together.

Many studies of temperature-adapted enzymes focus on what may be considered the main physical characteristics of an enzyme: its optimum temperature ( $T_{opt}$ ) and its melting temperature ( $T_m$ ). Unsurprisingly, it is observed that psychrophilic enzymes exhibit a lower  $T_{opt}$  and  $T_m$  than their mesophilic and thermophilic homologues. However, it was also observed that some psychrophilic enzymes exhibited a  $T_{opt}$  that was far from their  $T_m$  [20]. Here, we term this difference between  $T_{opt}$  and  $T_m$  as an enzyme's 'temperature gap' ( $T_g$ ). It was initially suggested that this gap was due to the active site of psychrophilic enzymes being more thermolabile than the rest of the protein in order to have sufficient flexibility to achieve catalysis at low environmental temperatures [4,20,21]. However, alternative hypotheses have been

Received: 04 February 2021

Revised: 29 March 2021

Accepted: 19 April 2021

Accepted Manuscript online:  
19 April 2021

Version of Record published:  
30 April 2021

proposed to account for this, such as the equilibrium model [22], macromolecular rate theory [23–25], and the loss of temperature-sensitive enzyme–substrate interactions [26]. However, as most studies focus on one type of enzyme across a small sample of species, it is difficult to understand how representative this phenomenon is across many enzyme types. Therefore, the first aim of the present study is to determine whether a large  $T_g$  can be characterised as a general feature of psychrophilic enzymes and to what extent we also see this phenomenon in mesophilic and thermophilic enzymes.

Another suggested characteristic of temperature-adapted organisms is that thermophiles exhibit particularly low mutational tolerance [27,28]. It has been suggested that the high temperatures of a thermophile’s environment make it particularly constrained by temperature-sensitive mutations. However, it has also been suggested that microbial communities actually evolve faster in extreme environments [29], seemingly in contrast with the predictions made by Drake [27]. This raises the question as to whether mutations themselves have a greater effect on thermophilic enzyme stability, or do thermophiles simply live closer to their proteome’s thermodynamic edge of stability than do mesophiles or psychrophiles? Thus, the second aim of the present study was to determine whether protein mutation software, PoPMuSiC [30] and HoTMuSiC [31], predicts a difference in effect to an enzyme’s Gibbs free energy of folding ( $\Delta\Delta G_f$ ) or melting temperature ( $\Delta T_m$ ) upon mutation among psychrophiles, mesophiles, and thermophiles.

In the present study, it is shown through meta-analysis that the  $T_{opt}$  and  $T_m$  of an enzyme increases from psychrophiles to thermophiles, as is expected. It is also shown that, while most enzymes exhibit a  $T_g$ , the  $T_g$  of psychrophilic enzymes is significantly larger than that of both mesophilic and thermophilic enzymes and in certain cases  $T_g$  provides the best indication of whether an enzyme is psychrophilic or not. Additionally we show that the average amino acid substitution is more deleterious to thermophilic enzyme stability compared with psychrophilic enzymes, with a general increase in the deleterious effect from psychrophiles through to thermophiles. Owing to the small absolute predicted differences between the stability parameters for the temperature-adapted enzymes, it is unknown how this would affect the mutational tolerance of thermophiles compared with mesophiles and psychrophiles over evolutionary timescales.

## Methods

### Dataset construction

Two datasets were created for the present study. Dataset 1 contains the  $T_{opt}$  and  $T_m$  data for homologous temperature-adapted enzymes from psychrophiles, mesophiles, and thermophiles which were included following a literature search of published data. Dataset 1 also contains the calculated  $T_g$ .  $T_g$  is defined here as the temperature gap between an enzyme’s  $T_m$ , and its  $T_{opt}$  and is calculated from the following equation:

$$T_g = T_m - T_{opt}$$

Dataset 2 contains the Protein Data Bank (PDB) IDs of homologous temperature-adapted enzymes from psychrophiles, mesophiles, and thermophiles which were found following a literature search or from searching through the PDB itself.

Each dataset had certain criteria which had to be met before data were entered into the dataset. For dataset 1, only wildtype enzymes were included. This meant that variants generated through random/targeted mutagenesis were excluded. This means that the data obtained for the studied enzymes result from their natural evolutionary history, whereas generated variants may have contained alterations which are not represented or permissible in the natural environment and as such may have affected the results. For an individual enzyme, the  $T_m$  and  $T_{opt}$  values were only taken from separate publications if it was clear that both studies were using the same enzyme from the same source organism. Reports of  $T_{50}$  values were excluded as they primarily reflect the kinetic stability of an enzyme rather than the global stability which is inferred from  $T_m$  measurements. Reports in which an enzyme’s  $T_m$  was lower than its  $T_{opt}$  were excluded. Such reports were rare. Furthermore an enzyme was only included in dataset 1 if both  $T_m$  and  $T_{opt}$  could be obtained, otherwise  $T_g$  could not be calculated. This has bearing for the thermophilic results as there were instances of thermophilic enzymes exhibiting high  $T_{opt}$  values, however the  $T_m$  values were experimentally unobtainable in the respective studies. These restrictions on data mean that the results presented here may represent a lower estimate of the mean  $T_m$ ,  $T_{opt}$ , and  $T_g$  of thermophilic enzymes.

Dataset 2 had similar entry requirements, such as only natural enzymes were included, and generated variants were excluded. As the mutational software used in the present study is structure based, a PDB ID was required for entry into dataset 2. Enzymes were taken as psychrophilic, mesophilic, and thermophilic based on how the source literature characterised them.

## Predicting the effect of mutations to protein stability

In order to predict the effect of mutations on the stability of temperature-adapted enzymes, two pieces of software were used, HoTMuSiC and PoPMuSiC [30,31] (available at <https://soft.dezyme.com/>). Both pieces of software require a PDB ID as input. HoTMuSiC predicts the effect of a mutation to a protein's melting temperature ( $\Delta T_m$ ), therefore a positive value is interpreted as stabilising and a negative value is destabilising. PoPMuSiC predicts the effect of a mutation to a protein's  $\Delta\Delta G_f$  and so a negative value is stabilising, and a positive value is destabilising. For data analysis, the mean effect of mutations to the respective proteins was recorded. Together the two pieces of software report on different, but complementary parts of a protein's temperature stability curve.

## Statistics

Statistical analysis was performed on GraphPad Prism. The results were analysed for statistically significant differences using one-way ANOVAs followed by post-hoc Tukey's multiple comparisons tests. If the group variances were found to be significantly different using a Bartlett's test, then a Welch's ANOVA was employed instead, followed by post-hoc Dunnett's T3 multiple comparisons tests. This was implemented for the  $T_g$  and  $\Delta T_m$  data. The ANOVA results and post-hoc test results are provided as supplementary information.

## Results

### Enzyme activity and stability

The first hypothesis tested in the present study is to what extent can  $T_{opt}$ ,  $T_m$ , and  $T_g$  be described as defining characteristics of temperature-adapted enzymes. Figure 1 shows the  $T_{opt}$  (A),  $T_m$  (B), and  $T_g$  (C) of enzymes from temperature-adapted organisms. The results displayed in Figure 1A show that the  $T_{opt}$  of an enzyme increases with increasing environmental temperatures and that the  $T_{opt}$  values were significantly different in pairwise comparisons ( $P$ -values, psychrophile-mesophile =  $4.2 \times 10^{-9}$ , psychrophile-thermophile =  $5 \times 10^{-10}$ , mesophile-thermophile =  $1.4 \times 10^{-9}$ ). The mean  $T_{opt}$  values ( $\pm$ SEM) for psychrophilic, mesophilic, and thermophilic enzymes are 32.97 ( $\pm$ 2.16), 55.03 ( $\pm$ 2.52), and 78.03 ( $\pm$ 2.25) $^{\circ}$ C respectively.

Similarly, Figure 1B shows that the  $T_m$  of an enzyme increases from psychrophiles to thermophiles and that  $T_m$  values were significantly different in pairwise comparisons ( $P$ -values, psychrophile-mesophile = 0.004, psychrophile-thermophile =  $5 \times 10^{-10}$ , mesophile-thermophile =  $5 \times 10^{-10}$ ). The mean  $T_m$  values for psychrophilic, mesophilic, and thermophilic enzymes are 55.02 ( $\pm$ 2.25), 62.37 ( $\pm$ 2.02), and 86.77 ( $\pm$ 2.38) $^{\circ}$ C respectively.

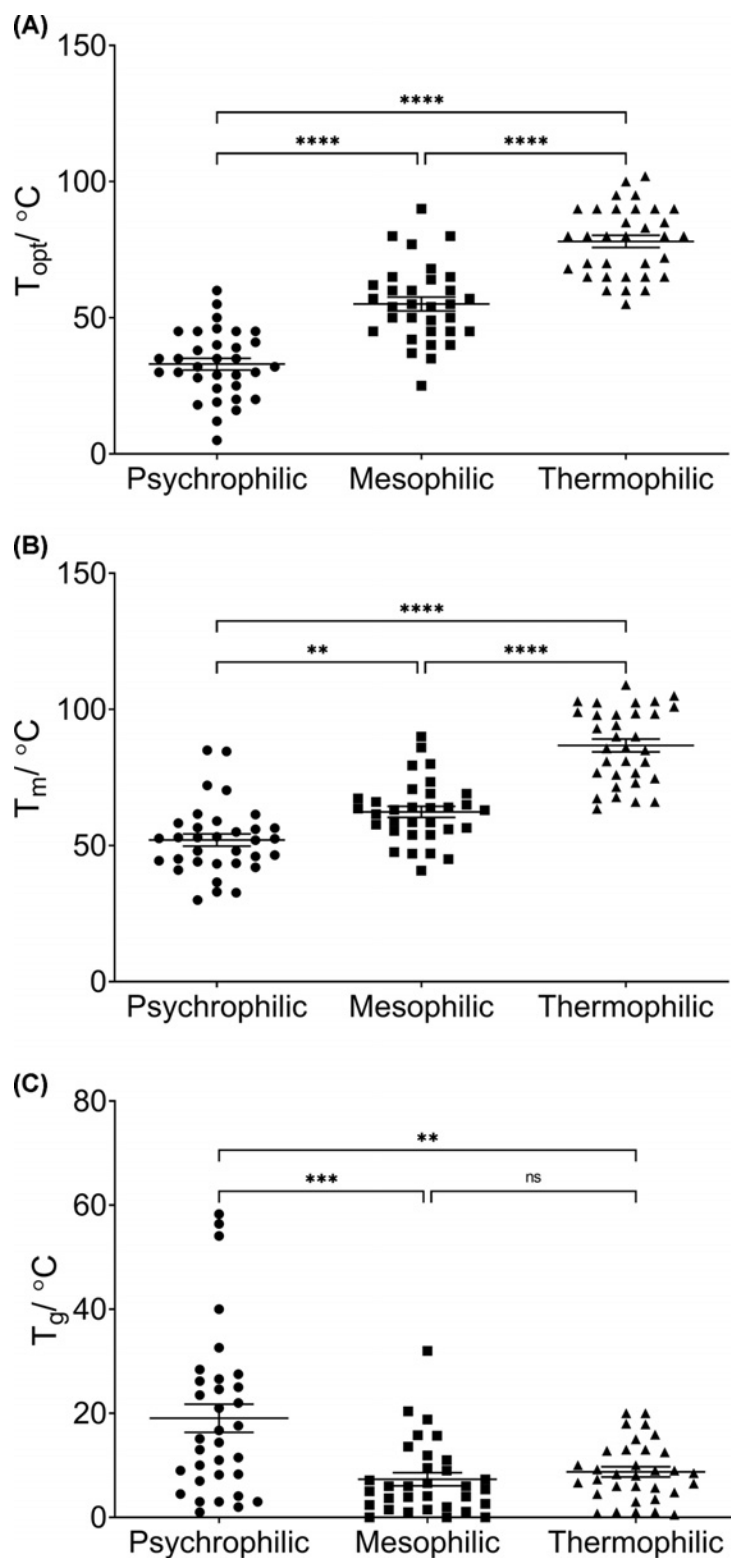
The statistically significant difference between the means of both  $T_{opt}$  and  $T_m$  for all three groups of organisms suggests that, on average,  $T_{opt}$  and  $T_m$  can be described as defining characteristics of an enzyme from organisms adapted to different temperature conditions. Namely, that psychrophiles exhibit the lowest  $T_{opt}$  and  $T_m$  as they inhabit the lowest temperature environments, while the opposite is true for the thermophiles with the mesophiles exhibiting intermediate values.

Figure 1C shows that while all enzymes exhibited a  $T_g$ , it is only statistically significantly different when comparing psychrophilic enzymes to mesophilic or thermophilic enzymes ( $P$ -values = 0.000896 and 0.00276 respectively). There is no statistical difference between the  $T_g$  of mesophilic enzymes and thermophilic enzymes ( $P$ -value = 0.765462). The mean  $T_g$  for psychrophiles is 19.05 ( $\pm$ 2.71) $^{\circ}$ C whereas for mesophiles and thermophiles it is 7.34 ( $\pm$ 1.26) and 8.74 ( $\pm$ 0.99) $^{\circ}$ C, respectively. So while most enzymes exhibit a  $T_g$ , it is significantly greater in psychrophilic enzymes. These results suggest that a large  $T_g$  may be considered as an indicative characteristic of psychrophilic enzymes in general, analogous to their canonical characteristics of a lower  $T_{opt}$  and  $T_m$ .

### Effect of mutations

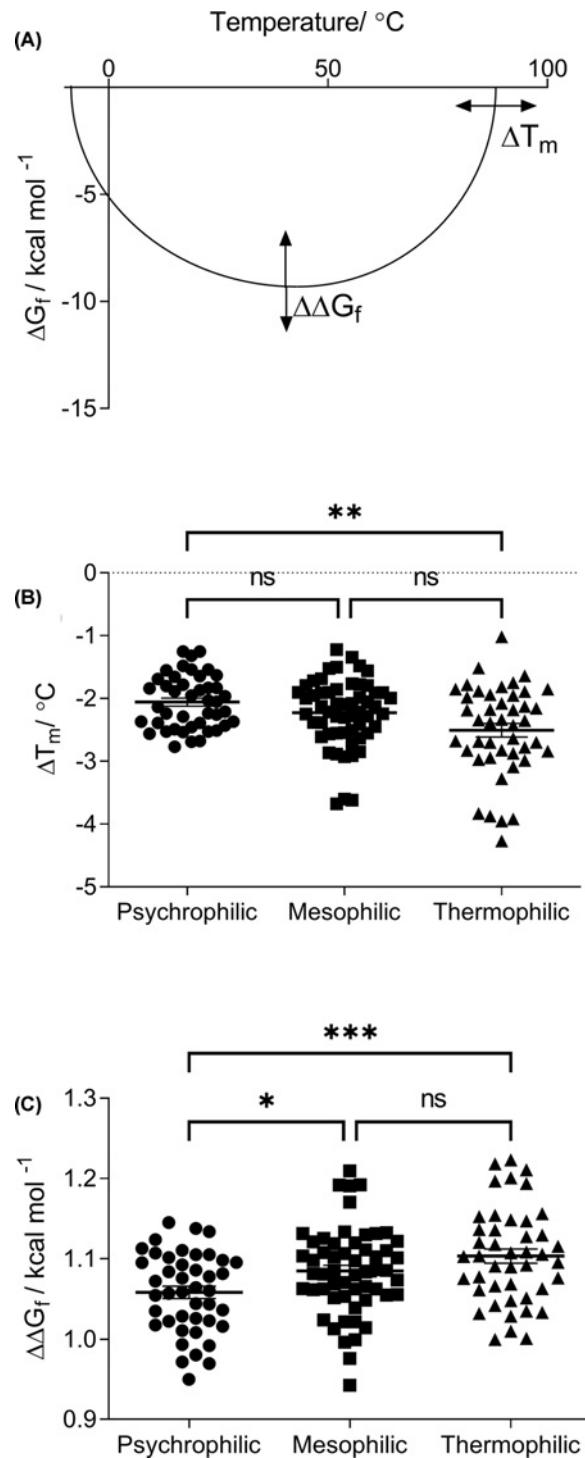
The second hypothesis tested in the present study was that there was a difference in the effect of a mutation (specifically amino acid substitutions) to an enzyme's  $\Delta\Delta G_f$  or melting temperature ( $\Delta T_m$ ) among psychrophiles, mesophiles, and thermophiles. Figure 2A shows a representative protein stability curve which could be produced with results from differential scanning calorimetry with the Gibbs free energy of folding on the y-axis and temperature on the x-axis. A protein's stability curve shows a region of peak stability where  $\Delta G_f$  is most negative, and also exhibits two melting points where the curve intersects the x-axis. On Figure 2A, the horizontal and vertical arrows represent the changes to a protein's melting temperature and Gibbs free energy of folding predicted by HoTMuSiC and PoPMuSiC respectively.

Figure 2B shows the predicted  $\Delta T_m$  to enzymes from psychrophiles, mesophiles, and thermophiles upon mutation by HoTMuSiC. The average  $\Delta T_m$  was  $-2.06$ ,  $-2.23$ , and  $-2.51$  $^{\circ}$ C for psychrophiles, mesophiles, and thermophiles



**Figure 1. The activity and stability parameters of temperature-adapted enzymes**

Panel (A) represents the optimum temperature for enzyme activity ( $T_{opt}$ ), while (B) shows the melting temperatures ( $T_m$ ) of the individual enzymes. Panel (C) shows the temperature gap between  $T_{opt}$  and  $T_m$ , denoted as  $T_g$ . The individual data points for psychrophiles are represented by circles, mesophiles by squares, and the thermophiles by triangles. All data points are plotted with the mean  $\pm$  the SEM. \* represent the statistical significance results from Tukey's multiple comparisons tests for panels (A,B) and Dunnett's T3 multiple comparisons tests for panel (C) (\*\* =  $P < 0.01$ , \*\*\* =  $P < 0.001$ , \*\*\*\* =  $P < 0.0001$ , ns = not significant).



**Figure 2. The effects of mutations to temperature-adapted enzymes**

Panel (A) shows a representative protein stability curve expressed as its Gibbs free energy of folding ( $\Delta G_f$ ) across temperature. The stability curve exhibits two melting points where it crosses the x-axis, and a peak of stability where the curve has its most negative y value. Horizontal and vertical arrows represent the changes to protein stability predicted by HoTMuSiC and PoPMuSiC respectively ( $\Delta T_m$  and  $\Delta \Delta G_f$ ). Panel (B) shows the  $\Delta T_m$  predicted by HoTMuSiC to enzymes from psychrophiles, mesophiles, and thermophiles as a result of single amino acid mutations. Panel (C) shows the  $\Delta \Delta G_f$  predicted by PoPMuSiC to enzymes from psychrophiles, mesophiles, and thermophiles. All data points are plotted with the mean  $\pm$  the SEM. \* represent the statistical significance results from Dunnett's T3 multiple comparisons tests for panel (B) and Tukey's multiple comparisons tests for panel (C) (\* =  $P < 0.05$ , \*\* =  $P < 0.01$ , \*\*\* =  $P < 0.001$ , ns = not significant).



respectively. The results suggest that the average reduction in the melting temperature of an enzyme upon mutation increases from psychrophiles to thermophiles, which agrees with previous literature [32]. The only Dunnett's T3 multiple comparisons test to produce a statistically significant result was between the psychrophiles and the thermophiles ( $P$ -value = 0.0019). The difference among the three categories is small in terms of absolute numbers, but as percentages they suggest that the average mutation to a thermophilic enzyme is ~10–25% more destabilising than those to their mesophilic and psychrophilic counterparts. To what extent such differences would have an effect over evolutionary timescales is currently unknown.

Figure 2C shows the predicted  $\Delta\Delta G_f$  to enzymes from psychrophiles, mesophiles and thermophiles upon mutation by PoPMuSiC. The average  $\Delta\Delta G_f$  was 1.058, 1.085, and 1.103 kcal.mol<sup>-1</sup> for psychrophiles, mesophiles, and thermophiles respectively. Similar to the  $\Delta T_m$  results, the average  $\Delta\Delta G_f$  upon mutation increases from psychrophiles through to the thermophiles. Post-hoc Tukey's multiple comparisons tests showed statistically significant differences between psychrophiles-mesophiles and psychrophiles-thermophiles with  $P$ -values of 0.0451 and 0.000459, respectively.

From these results it is demonstrated that the average mutation to an enzyme not only lowers the melting temperature, but also reduces the thermodynamic stability, thus constricting the global folded phase space. Furthermore, it is shown that mutations are more deleterious to thermophilic enzymes than they are to mesophilic or psychrophilic enzymes.

## Discussion

In this meta-analysis, we have collated and presented data which further expand our understanding of the defining characteristics of temperature-adapted enzymes. It was shown that the  $T_{opt}$  and  $T_m$  of enzymes increased with increasing environmental temperatures. In contrast, it was shown that an enzyme's  $T_g$ , the gap between the optimum and melting temperature of an enzyme, is significantly larger in psychrophiles, and is in fact a defining characteristic of psychrophilic enzymes that could allow for the prediction of enzymatic psychrophilicity. Additionally it was shown that the average amino acid mutation is predicted to be more destabilising to thermophilic enzymes than it is to mesophilic or psychrophilic enzymes.

Our data allow for several important observations. There is a considerable overlap in the  $T_m$  values for psychrophiles and mesophiles, suggesting that increased psychrophilic enzyme activity at lower temperatures has not necessarily come at a cost to overall protein stability. This suggests that global protein stability is not a major constraint on psychrophilic enzyme adaptation and evolution. Conversely, thermophilic enzyme stability is more clearly an adaptive feature as seen from the larger difference between the thermophilic and mesophilic  $T_m$  means.

The results show that not all psychrophilic enzymes necessarily have psychrophilic characteristics. This is perhaps best exemplified by one of the enzymes included in our dataset, the thermostable psychrophilic glutathione reductase from an Arctic *Sphingomonas* with a  $T_{opt}$  and  $T_m$  of 60 and 84.6°C respectively [33], values typically associated with thermophilic enzymes. In this case, the large  $T_g$  value of 24.6°C is the best predictive indicator that this enzyme came from a psychrophilic organism. Additionally, few psychrophilic enzymes exhibit  $T_{opt}$  values which would be considered similar to the expected environmental temperature of a psychrophile.

It should also be noted that the thermophile  $T_m$  and  $T_{opt}$  (and consequently  $T_g$ ) values represent a lower estimate of their true population. This is due to exclusion of studies which did not report both the  $T_m$  and  $T_{opt}$ . This largely results from the limitations of circular dichroism apparatus and differential scanning calorimeters used in such studies, which prevent the measurement of high  $T_m$  values. It raises a question as to whether there are thermophilic enzymes which are so thermostable that they resist melting until their carbon backbone begins to physically dissociate. The sample size of thermophilic enzymes was further reduced due to the propensity to report  $T_{50}$  measurements in the literature. This is understandable due to the considerable biotechnological interest in thermophilic enzymes [1], where their kinetic stability at elevated temperatures is of more interest than the temperature at which global unfolding occurs.

These data also raise the question of the correlation between enzyme type and the size of  $T_g$ . Evidence for a correlation was seen with the luciferase enzymes included in our dataset. They exhibit high  $T_g$  values in both psychrophiles and mesophiles. Our dataset contained four luciferase enzymes. The three psychrophilic luciferase  $T_g$  values were 56.4, 58.3, and 54.1°C with a mesophilic firefly luciferase exhibiting a  $T_g$  of 15.8°C. Of additional interest is the observation that all three psychrophilic luciferases were more thermostable than the mesophilic firefly luciferase, by as much as 31°C.

While our results show that a large  $T_g$  is a defining characteristic of psychrophilic enzymes, they cannot elucidate the precise source of this phenomenon. We can however discuss the consequences of each hypothesis with regards to our analysis. Multiple explanations have been proposed to explain this observation such as, active

site unfolding [4,20,21], an equilibrium model [22], macromolecular rate theory [23–25], and the loss of specific temperature-sensitive enzyme–substrate interactions [26]. The initial explanation that the active site of  $\alpha$ -amylase from the psychrophile *Pseudoalteromonas haloplanktis* is particularly thermolabile [21] possesses strong explanatory power and fits with observations that increased active site flexibility and dynamics are key to achieving greater enzymatic activity at low environmental temperatures [34–37]. Within the framework of this hypothesis, our results would suggest that, as a population, psychrophilic enzymes possess significantly more thermolabile active sites than do mesophiles or thermophiles compared with the stability of the whole enzymes. An equilibrium model interpretation of the data would suggest that psychrophilic enzymes reach the equilibrium temperature ( $T_{eq}$ ), the point at which half the enzyme is active, much before they reach their  $T_m$ . This would suggest that the ratio of active to inactive enzyme forms ( $E_{act}/E_{inact}$ ) is particularly temperature sensitive in psychrophiles and therefore results in a larger  $T_g$ . The loss of temperature-sensitive enzyme–substrate interactions proposed by Sočan et al. [26] is largely a molecular level interpretation of the equilibrium model as they propose a ‘dead-end model’ where an inactive enzyme forms with increasing temperature. This would suggest that the interactions between substrates and psychrophilic enzymes is significantly weaker than those of mesophilic and thermophilic enzymes and therefore is the source of the large  $T_g$  in psychrophilic enzymes. Macromolecular rate theory would predict that the change in heat capacity of activation ( $\Delta C_p^\ddagger$ ) is significantly lower in psychrophilic enzymes compared with mesophilic and thermophilic enzymes. This would cause a larger  $T_g$  in psychrophiles due to the increasing curvature of the temperature-dependent activity profile as  $\Delta C_p^\ddagger$  is lowered. No single hypothesis may explain the  $T_g$  phenomenon and diverse hypotheses may be applicable to different enzymes. It will require precise measurements on the molecular level to determine the true origin of  $T_g$  for each enzyme.

The lower  $T_g$  values for mesophilic and thermophilic enzymes may be useful for validating ancestrally reconstructed enzymes. Ancestral reconstruction tends to produce more thermostable enzymes [38,39], however there is a concern that this may be an artifact due to biases in the reconstruction process [40]. Therefore based on our meta-analysis, if these ancestral enzymes were indeed more thermophilic, then one should not expect to find that  $T_g$  increases significantly when constructing an ancestral enzyme from the modern day mesophilic form.

The mutational data presented here are in strong agreement with the well-established observation that mutations are on average destabilising. The  $\Delta T_m$  values reported here are less destabilising than those presented in previous literature [32] which ranged from approx.  $-1.3$  to  $-5^\circ\text{C}$ , with thermophilic proteins predicted to experience more destabilising mutations. This may be due to the focus on enzymes in the present study, which may produce more stabilising mutations than the average non-enzymatic protein. This could be explained by the fact that the active site of an enzyme generally contributes little to stability, therefore mutating it tends to introduce stabilising interactions [41–44] or have more neutral effects. Our data does however point towards an increasing trend in this deleterious nature with increasing environmental temperatures. Therefore, studies regarding the trajectories and timescales of enzyme evolution may require varied weighting of mutational effects depending on the thermophilicity of the enzymes in question.

The observation that mutations are more deleterious to thermophilic enzymes agrees with the hypothesis put forward by Drake [27]. If there is a tight coupling between a thermophile’s environmental temperature and its enzymes’ temperature stabilities, then a difference in  $\Delta T_m$  of  $0.5^\circ\text{C}$  may be sufficient to make the average mutation particularly potent against thermophile survivability. So while thermophilic proteins may be more tolerant to mutations at  $\sim 30^\circ\text{C}$  compared with their mesophilic counterparts [45], the coupling of environmental temperature and  $T_m$  would produce the phenomenon of lower mutational tolerance *in situ*. In contrast with Drake [27], Li et al. [29] have reported that microbial communities evolve faster in extreme environments. Drake reported that the  $d_N/d_S$  (the non-synonymous/synonymous mutation ratio) for thermophiles was lower for thermophiles compared with mesophiles, 0.09 versus 0.14 respectively, suggesting thermophiles tolerate less mutation. However, Li et al. report that communities of thermophiles from hot springs have a higher  $d_N/d_S$  than communities from the surface ocean, freshwater or soil ( $d_N/d_S$  values of 0.126, 0.061, 0.087, and 0.087 respectively). Li et al. also reported higher relative evolutionary rates (rERs) for thermophilic communities compared to freshwater and soil communities. It is hard to directly compare the two studies though, as Drake [27] considered other mutation types such as chain terminations and indel mutations. On the other hand, Drake [27] examined two species of thermophiles, so it is difficult to extrapolate those results to all thermophiles, whereas Li et al. have reported data at the community level, making their work potentially more representative of thermophiles as a class of organism. The experimental determination of whether psychrophiles and mesophiles can tolerate higher mutational loads than thermophiles, while critical for answering this question, is limited by the long time-course required to culture and grow such organisms.

## Conclusion

The aim of the present study was to further explore the characteristics of temperature-adapted enzymes. It was shown, in strong agreement with theory, that the  $T_{opt}$  and  $T_m$  increases with an organism's environmental temperature. It was also shown that a large  $T_g$  is a defining characteristic of psychrophilic enzymes and in certain cases is a better predictor of psychrophilicity than either  $T_{opt}$  or  $T_m$ . The average effect of an amino acid mutation to temperature-adapted enzymes was also explored. It was found that the average  $\Delta T_m$  and  $\Delta \Delta G_f$  becomes more deleterious, with increasing environmental temperature. The difference in deleterious effect was small and the effect of this over evolutionary timescales is unknown.

## Data Availability

The source data for all results are provided as Supplementary Data. Dataset 1 contains the  $T_m$ ,  $T_{opt}$  and  $T_g$  values for each enzyme, their source organism, and their literature source. Dataset 2 contains the PDB IDs of all enzymes used in the mutation results with their average  $\Delta T_m$  and  $\Delta \Delta G_f$ . Both datasets contain a summary table. The ANOVA results and post-hoc test results for each analysis are also provided as a supplementary data file.

## Competing Interests

The authors declare that there are no competing interests associated with the manuscript.

## Funding

This work was supported by the EPSRC (to Stewart Gault); the Nuffield Foundation (to Kaitlyn Gillies); and the STFC [grant number ST/R000875/1 (to Peter M. Higgins and Charles S. Cockell)].

## Author Contribution

Stewart Gault conceived the study and created the datasets. Stewart Gault, Peter M. Higgins, Charles S. Cockell and Kaitlyn Gillies contributed to data analysis, manuscript drafting and editing.

## Abbreviations

PDB, Protein Data Bank;  $T_g$ , temperature gap;  $T_m$ , melting temperature;  $T_{opt}$ , optimum temperature;  $\Delta \Delta G_f$ , Gibbs free energy of folding.

## References

- Vieille, C. and Zeikus, G.J. (2001) Hyperthermophilic enzymes: sources, uses, and molecular mechanisms for thermostability. *Microbiol. Mol. Biol. Rev.* **65**, 1–43, <https://doi.org/10.1128/MMBR.65.1.1-43.2001>
- Sterner, R. and Liebl, W. (2001) Thermophilic adaptation of proteins. *Crit. Rev. Biochem. Mol. Biol.* **36**, 39–106, <https://doi.org/10.1080/20014091074174>
- Feller, G. and Gerday, C. (2003) Psychrophilic enzymes: hot topics in cold adaptation. *Nat. Rev. Microbiol.* **1**, 200–208
- Feller, G. (2013) Psychrophilic enzymes: from folding to function and biotechnology. *Scientifica* **2013**, 512840, <https://doi.org/10.1155/2013/512840>
- Kamondi, S., Szilágyi, A., Barna, L. and Závodszy, P. (2008) Engineering the thermostability of a TIM-barrel enzyme by rational family shuffling. *Biochem. Biophys. Res. Commun.* **374**, 725–730, <https://doi.org/10.1016/j.bbrc.2008.07.095>
- Dong, G., Vieille, C. and Zeikus, J.G. (1997) Cloning, sequencing, and expression of the gene encoding amylopullulanase from *Pyrococcus furiosus* and biochemical characterization of the recombinant enzyme. *Appl. Environ. Microbiol.* **63**, 3577–3584, <https://doi.org/10.1128/AEM.63.9.3577-3584.1997>
- Brown, S.H. and Kelly, R.M. (1993) Characterization of amylolytic enzymes, having both alpha-1,4 and alpha-1,6 hydrolytic activity, from the thermophilic archaea *Pyrococcus furiosus* and *Thermococcus litoralis*. *Appl. Environ. Microbiol.* **59**, 2614–2621, <https://doi.org/10.1128/AEM.59.8.2614-2621.1993>
- Panja, A.S., Bandyopadhyay, B. and Maiti, S. (2015) Protein thermostability is owing to their preferences to non-polar smaller volume amino acids, variations in residual physico-chemical properties and more salt-bridges. *PLoS ONE* **10**, e0131495–e0131495, <https://doi.org/10.1371/journal.pone.0131495>
- Zhou, X.X., Wang, Y.B., Pan, Y.J. and Li, W.F. (2008) Differences in amino acids composition and coupling patterns between mesophilic and thermophilic proteins. *Amino Acids* **34**, 25–33, <https://doi.org/10.1007/s00726-007-0589-x>
- Metpally, R.P.R. and Reddy, B.V.B. (2009) Comparative proteome analysis of psychrophilic versus mesophilic bacterial species: Insights into the molecular basis of cold adaptation of proteins. *BMC Genomics* **10**, 11, <https://doi.org/10.1186/1471-2164-10-11>
- Kahlke, T. and Thorvaldsen, S. (2012) Molecular characterization of cold adaptation of membrane proteins in the vibronaceae core-genome. *PLoS ONE* **7**, 1–9, <https://doi.org/10.1371/journal.pone.0051761>
- Panja, A.S., Maiti, S. and Bandyopadhyay, B. (2020) Protein stability governed by its structural plasticity is inferred by physicochemical factors and salt bridges. *Sci. Rep.* **10**, 1822, <https://doi.org/10.1038/s41598-020-58825-7>



- 13 Gianese, G., Argos, P. and Pascarella, S. (2001) Structural adaptation of enzymes to low temperatures. *Protein Eng. Des. Sel.* **14**, 141–148, <https://doi.org/10.1093/protein/14.3.141>
- 14 Kumar, S., Tsai, C.J. and Nussinov, R. (2000) Factors enhancing protein thermostability. *Protein Eng. Des. Sel.* **13**, 179–191, <https://doi.org/10.1093/protein/13.3.179>
- 15 Szilágyi, A. and Závodszy, P. (2000) Structural differences between mesophilic, moderately thermophilic and extremely thermophilic protein subunits: results of a comprehensive survey. *Structure* **8**, 493–504, [https://doi.org/10.1016/S0969-2126\(00\)00133-7](https://doi.org/10.1016/S0969-2126(00)00133-7)
- 16 Tanner, J.J., Hecht, R.M. and Krause, K.L. (1996) Determinants of enzyme thermostability observed in the molecular structure of *Thermus aquaticus* d-Glyceraldehyde-3-phosphate dehydrogenase at 2.5 Å resolution. *Biochemistry* **35**, 2597–2609, <https://doi.org/10.1021/bi951988q>
- 17 Goldstein, R.A. (2007) Amino-acid interactions in psychrophiles, mesophiles, thermophiles, and hyperthermophiles: insights from the quasi-chemical approximation. *Protein Sci.* **16**, 1887–1895, <https://doi.org/10.1110/ps.072947007>
- 18 Bae, E. and Phillips, G.N.J. (2004) Structures and analysis of highly homologous psychrophilic, mesophilic, and thermophilic adenylate kinases. *J. Biol. Chem.* **279**, 28202–28208, <https://doi.org/10.1074/jbc.M401865200>
- 19 Pucci, F. and Rooman, M. (2017) Physical and molecular bases of protein thermal stability and cold adaptation. *Curr. Opin. Struct. Biol.* **42**, 117–128, <https://doi.org/10.1016/j.sbi.2016.12.007>
- 20 D'Amico, S., Marx, J.C., Gerday, C. and Feller, G. (2003) Activity-stability relationships in extremophilic enzymes. *J. Biol. Chem.* **278**, 7891–7896, <https://doi.org/10.1074/jbc.M212508200>
- 21 Siddiqui, K.S., Feller, G., D'Amico, S., Gerday, C., Giaquinto, L. and Cavicchioli, R. (2005) The active site is the least stable structure in the unfolding pathway of a multidomain cold-adapted  $\alpha$ -Amylase. *J. Bacteriol.* **187**, 6197–6205, <https://doi.org/10.1128/JB.187.17.6197-6205.2005>
- 22 Daniel, R.M., Danson, M.J., Eisenthal, R., Lee, C.K. and Peterson, M.E. (2008) The effect of temperature on enzyme activity: new insights and their implications. *Extremophiles* **12**, 51–59, <https://doi.org/10.1007/s00792-007-0089-7>
- 23 Hobbs, J.K., Jiao, W., Easter, A.D., Parker, E.J., Schipper, L.A. and Arcus, V.L. (2013) Change in heat capacity for enzyme catalysis determines temperature dependence of enzyme catalyzed rates. *ACS Chem. Biol.* **8**, 2388–2393, <https://doi.org/10.1021/cb4005029>
- 24 van der Kamp, M.W., Prentice, E.J., Kraakman, K.L., Connolly, M., Mulholland, A.J. and Arcus, V.L. (2018) Dynamical origins of heat capacity changes in enzyme-catalysed reactions. *Nat. Commun.* **9**, 1177, <https://doi.org/10.1038/s41467-018-03597-y>
- 25 Arcus, V.L., Prentice, E.J., Hobbs, J.K., Mulholland, A.J., Van der Kamp, M.W., Pudney, C.R. et al. (2016) On the temperature dependence of enzyme-catalyzed rates. *Biochemistry* **55**, 1681–1688, <https://doi.org/10.1021/acs.biochem.5b01094>
- 26 Sočan, J., Purg, M. and Åqvist, J. (2020) Computer simulations explain the anomalous temperature optimum in a cold-adapted enzyme. *Nat. Commun.* **11**, 2644, <https://doi.org/10.1038/s41467-020-16341-2>
- 27 Drake, J.W. (2009) Avoiding dangerous missense: thermophiles display especially low mutation rates. *PLoS Genet.* **5**, e1000520, <https://doi.org/10.1371/journal.pgen.1000520>
- 28 Zeldovich, K.B., Chen, P. and Shakhnovich, E.I. (2007) Protein stability imposes limits on organism complexity and speed of molecular evolution. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 16152–16157, <https://doi.org/10.1073/pnas.0705366104>
- 29 Li, S.J., Hua, Z.S., Huang, L.N., Li, J., Shi, S.H., Chen, L.X. et al. (2014) Microbial communities evolve faster in extreme environments. *Sci. Rep.* **4**, 6205, <https://doi.org/10.1038/srep06205>
- 30 Dehouck, Y., Kwasigroch, J.M., Gilis, D. and Rooman, M. (2011) PoPMuSiC 2.1: a web server for the estimation of protein stability changes upon mutation and sequence optimality. *BMC Bioinformatics* **12**, 151, <https://doi.org/10.1186/1471-2105-12-151>
- 31 Pucci, F., Bourgeas, R. and Rooman, M. (2016) Predicting protein thermal stability changes upon point mutations using statistical potentials: introducing HoTMuSiC. *Sci. Rep.* **6**, 23257, <https://doi.org/10.1038/srep23257>
- 32 Pucci, F. and Rooman, M. (2016) Improved insights into protein thermal stability: from the molecular to the structurome scale. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **374**, 20160141
- 33 VuThi, H., Jang, S.H. and Lee, C. (2019) Cloning and characterization of a thermostable glutathione reductase from a psychrophilic Arctic bacterium *Sphingomonas* sp. *FEMS Microbiol. Lett.* **366**, <https://doi.org/10.1093/femsle/fnz218>
- 34 Fields, P.A. and Somero, G.N. (1998) Hot spots in cold adaptation: localized increases in conformational flexibility in lactate dehydrogenase A4 orthologs of Antarctic notothenioid fishes. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 11476–11481, <https://doi.org/10.1073/pnas.95.19.11476>
- 35 Chiuri, R., Maiorano, G., Rizzello, A., del Mercato, L.L., Cingolani, R., Rinaldi, R. et al. (2009) Exploring local flexibility/rigidity in psychrophilic and mesophilic carbonic anhydrases. *Biophys. J.* **96**, 1586–1596, <https://doi.org/10.1016/j.bpj.2008.11.017>
- 36 Fedøy, A.E., Yang, N., Martinez, A., Leiros, H.K.S. and Steen, I.H. (2007) Structural and functional properties of isocitrate dehydrogenase from the psychrophilic bacterium *Desulfotalea psychrophila* reveal a cold-active enzyme with an unusual high thermal stability. *J. Mol. Biol.* **372**, 130–149, <https://doi.org/10.1016/j.jmb.2007.06.040>
- 37 Saavedra, H.G., Wrabl, J.O., Anderson, J.A., Li, J. and Hilser, V.J. (2018) Dynamic allostery can drive cold adaptation in enzymes. *Nature* **558**, 324–328, <https://doi.org/10.1038/s41586-018-0183-2>
- 38 Wheeler, L.C., Lim, S.A., Marqusee, S. and Harms, M.J. (2016) The thermostability and specificity of ancient proteins. *Curr. Opin. Struct. Biol.* **38**, 37–43, <https://doi.org/10.1016/j.sbi.2016.05.015>
- 39 Gaucher, E.A., Govindarajan, S. and Ganesh, O.K. (2008) Palaeotemperature trend for Precambrian life inferred from resurrected proteins. *Nature* **451**, 704–707, <https://doi.org/10.1038/nature06510>
- 40 Williams, P.D., Pollock, D.D., Blackburne, B.P. and Goldstein, R.A. (2006) Assessing the accuracy of ancestral protein reconstruction methods. *PLoS Comput. Biol.* **2**, e69, <https://doi.org/10.1371/journal.pcbi.0020069>
- 41 Meiering, E.M., Serrano, L. and Fersht, A.R. (1992) Effect of active site residues in barnase on activity and stability. *J. Mol. Biol.* **225**, 585–589, [https://doi.org/10.1016/0022-2836\(92\)90387-Y](https://doi.org/10.1016/0022-2836(92)90387-Y)

- 42 Shoichet, B.K., Baase, W.A., Kuroki, R. and Matthews, B.W. (1995) A relationship between protein stability and protein function. *Proc. Natl. Acad. Sci. U.S.A.* **92**, 452–456, <https://doi.org/10.1073/pnas.92.2.452>
- 43 Beadle, B.M. and Shoichet, B.K. (2002) Structural bases of stability-function tradeoffs in enzymes. *J. Mol. Biol.* **321**, 285–296, [https://doi.org/10.1016/S0022-2836\(02\)00599-5](https://doi.org/10.1016/S0022-2836(02)00599-5)
- 44 Xie, Y., An, J., Yang, G., Wu, G., Zhang, Y., Cui, L. et al. (2014) Enhanced enzyme kinetic stability by increasing rigidity within the active site. *J. Biol. Chem.* **289**, 7994–8006, <https://doi.org/10.1074/jbc.M113.536045>
- 45 Bloom, J.D., Labthavikul, S.T., Otey, C.R. and Arnold, F.H. (2006) Protein stability promotes evolvability. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 5869–5874, <https://doi.org/10.1073/pnas.0510098103>