On the Potential Origins of the High Stability of Reconstructed Ancestral Proteins

Devin L. Trudeau,[1] Miriam Kaltenbach,[1] and Dan S. Tawfik*[,1]
[1]Department of Biomolecular Sciences, Weizmann Institute of Science, Rehovot, Israel
***Corresponding author:** E-mail: dan.tawfik@weizmann.ac.il.
Associate editor: Helen Piontkivska

Abstract

Ancestral reconstruction provides instrumental insights regarding the biochemical and biophysical characteristics of past proteins. A striking observation relates to the remarkably high thermostability of reconstructed ancestors. The latter has been linked to high environmental temperatures in the Precambrian era, the era relating to most reconstructed proteins. We found that inferred ancestors of the serum paraoxonase (PON) enzyme family, including the mammalian ancestor, exhibit dramatically increased thermostabilities compared with the extant, human enzyme (up to 30 °C higher melting temperature). However, the environmental temperature at the time of emergence of mammals is presumed to be similar to the present one. Additionally, the mammalian PON ancestor has superior folding properties (kinetic stability)—unlike the extant mammalian PONs, it expresses in *E. coli* in a soluble and functional form, and at a high yield. We discuss two potential origins of this unexpectedly high stability. First, ancestral stability may be overestimated by a "consensus effect," whereby replacing amino acids that are rare in contemporary sequences with the amino acid most common in the family increases protein stability. Comparison to other reconstructed ancestors indicates that the consensus effect may bias some but not all reconstructions. Second, we note that high stability may relate to factors other than high environmental temperature such as oxidative stress or high radiation levels. Foremost, intrinsic factors such as high rates of genetic mutations and/or of transcriptional and translational errors, and less efficient protein quality control systems, may underlie the high kinetic and thermodynamic stability of past proteins.

***Key words:* ancestral inference, protein stability, protein evolution, serum paraoxonase (PON), thermostability, consensus effect.**

Downloaded from https://academic.oup.com/mbe/article/33/10/2633/2925581 by guest on 20 August 2022

Article

Introduction

Given a large enough set of sequences of contemporary, evolutionarily related proteins, various models can be used to infer a phylogenetic tree and the amino acid sequences of the corresponding ancestral nodes (Pauling and Zuckerkandl 1963). Reconstructions of ancestral proteins have provided fascinating insights regarding the biophysical and biochemical properties of past proteins, and regarding the environmental and evolutionary forces that shaped them. For example, recent studies of ancestral proteins have shed light on the evolution of ligand binding (Carroll et al. 2008; Wilson et al. 2015) and enzyme specificity (Bar-Rogovsky et al. 2013; Smith et al. 2013; Boucher et al. 2014; Howard et al. 2014). Specifically, in relation to this paper, resurrected ancestral proteins were found to exhibit extraordinarily high stability, in particular high thermostability. The latter trait has been correlated with a high temperature of the environment at their presumed era of existence (Gaucher et al. 2008; Hobbs et al. 2012; Akanuma et al. 2013; Risso et al. 2013, 2014, 2015; Hart et al. 2014).

A major limitation of ancestral reconstruction regards the statistical nature of the prediction. Uncertainty is especially high in the prediction of amino acids in sites of intermediate to high divergence, typically in surface residues that are not involved in function. In such positions, more than one amino acid may be predicted, often with similar probabilities. The prediction is also inevitably biased by the "consensus effect"—by default, if at a given position, a certain amino acid is present in the majority of contemporary sequences, it would be assigned as the ancestral state. Highly conserved sites such as residues that mediate function, very slowly evolving core residues, and positions in which exchanges are correlated with other positions (covariance, or coevolution [Sullivan et al. 2012; Ashenberg et al. 2013; Pollock and Goldstein 2014; Risso et al. 2015]) are less biased by the consensus effect. However, the consensus amino acid is inevitably the most likely ancestral state in relatively diverged sites, especially those where deviations from the consensus amino acid are sporadic in relation to the occurrence of the deviating sequences along the tree (hereafter called "potentially biased positions"). It is widely known that replacing a rare amino acid at a given site with the most common amino acid in the alignment of homologous sequences (the consensus amino acid) often increases the thermodynamic and/or kinetic stability of proteins (Pey et al. 2008). Indeed, consensus mutations are routinely used to stabilize proteins (Steipe et al. 1994; Lehmann et al. 2002; Kiss et al. 2009; Komor et al. 2012; Sullivan et al. 2012; Goldsmith and Tawfik 2013). Therefore,

© The Author 2016. Published by Oxford University Press on behalf of the Society for Molecular Biology and Evolution.
All rights reserved. For permissions, please e-mail: journals.permissions@oup.com

a bias toward the consensus in the inferred ancestor may also lead to a bias in thermostability. In many cases, the historical ancestor may indeed have contained the consensus amino acid, possibly even at the majority of potentially biased positions. However, there is no reason to assume that, by default, any given historical ancestor had the consensus amino acid at all the potentially biased positions.

We have encountered the issue of hyperstability of ancestral proteins in the course of characterizing the evolutionary history of an enzyme family dubbed serum paraoxonases, or PONs (Bar-Rogovsky et al. 2013, 2015). PONs comprise a family of lactonases that also promiscuously hydrolyse phosphotriesters such as paraoxon. PONs are found mostly in mammals and also in other animals, but not in single-cell eukaryotes. More recently, distantly related PON-like genes were identified in bacteria and assigned as lactonases that hydrolyze homoserine lactones serving as quorum sensing signals. We previously studied the evolution of substrate specificity in the mammalian PON lineage, and to this end, we reconstructed both the mammalian and vertebrate PON ancestors (Bar-Rogovsky et al. 2013). Although the prediction of the mammalian ancestor was found to be robust, incorporation of alternatively predicted amino acids was found to induce significant variations in the profile of enzymatic activities of the vertebrate ancestor (Bar-Rogovsky et al. 2015). The latter was therefore not included in the study of the evolutionary history of PONs' substrate specificity (Bar-Rogovsky et al. 2015), but in the context of this study, its stability has been examined. This paper stems from the unexpected observation that, as reported here, the mammalian PON ancestor's melting temperature (T_m) was ~30 °C higher than that of human PON1, even though the temperature at its time of emergence is not thought to have been higher than at present, and further, mammals regulate their internal body temperature. The vertebrate ancestor also exhibited high stability, but not as high as the mammalian one (T_m value ~20 °C higher than human PON1). We thus describe various biophysical properties that relate to the hyperstability of the PON ancestors, and discuss their possible origins.

Results

Ancestral PON sequences were inferred from a phylogenetic tree generated with PhyML using the maximum-likelihood-based algorithm FastML (Pupko et al. 2000) (fig. 1A). The most-probable ancestor of mammalian PONs was ancN9* (previously called "N9*"—for clarity, all ancestors are denoted here as "anc"). Another version of this ancestor, dubbed ancN9, was inferred from an alignment and tree constructed without lizard or zebra finch PONs that were not available at the time of reconstruction. AncN9 and ancN9* differ at 11 positions, yet have essentially identical activities and substrate specificities (Bar-Rogovsky et al. 2013). Further, evaluation of the phenotypic variability in a library of sequences carrying alternative predictions at ambiguously predicted sites indicated that as far as the enzymatic specificity is concerned, the mammalian ancestor's phenotype is robustly inferred (Bar-Rogovsky et al. 2015).

The most probable ancestor of vertebrate PONs was dubbed ancN6. The activity and substrate specificity of the latter remained ambiguous, because our library screen indicated significant variations in specificity upon introduction of the alternatively predicted amino acids at sites where the predication was ambiguous (Bar-Rogovsky et al. 2015). Nonetheless, the thermodynamic stability of ancN6 was examined and is described here.

Extant PON1 proteins do not express in *Escherichia coli* in a soluble and functional form (Aharoni et al. 2004; Gaidukov et al. 2009). Our reference is therefore a recombinant mammalian PON1 (dubbed reG3C9) that was directly evolved for soluble expression in *E. coli* by screening a DNA shuffling library of mammalian PON genes (Aharoni et al. 2004). As previously determined, variant reG3C9 is considerably more stable than serum-purified human PON1 (Gaidukov et al. 2009). Herein, we used the very same conditions applied before to compare the stabilities of ancN9, ancN9* and ancN6 to reG3C9, thus providing the basis for comparison of the ancestral PONs to human PON1.

An alignment of the three inferred N6, N9 and N9* ancestors, the recombinant PON1-G3C9 variant, and human PON1 is provided in figure 1B, also listing the consensus amino acid at each position.

Thermal Stability of PON Ancestors

We expressed reG3C9, ancN9, ancN9* and ancN6 in *E. coli* and purified them as described (Gaidukov et al. 2009; Bar-Rogovsky et al. 2015). Due to the dissociation of a structural calcium anion that is essential for their core packing, PONs do not fold reversibly, neither upon chemical denaturation nor after heat denaturation (Kuo and La Du 1995). We could therefore test their thermostability by assaying residual activity at room temperature after incubation at a range of temperatures. Variant reG3C9 exhibited an apparent melting temperature (the temperature at which the residual activity is half of the initial activity, or T_m) of 55.6 ± 0.1 °C (fig. 2) which is lower than the previously reported 60.4 °C (Gaidukov et al. 2009). Although purification and stability measurements were carried out as described in this previous study (see Materials and Methods), one possible difference is that the presently used reG3C9 had a 6- instead of 8-Histidine C-terminal tag. Other differences may relate to the detergent concentration that tends to vary due to adsorption to storage tubes. Accordingly, all PON variants studied here were expressed, purified, and characterized in parallel to minimize experimental variation. The ancestral variants exhibited a distinctly higher T_m than reG3C9: ancN9; 68.9 ± 1.1 °C, ancN9*, 77.1 ± 0.5 °C, and ancN6, 63.4 ± 0.4 °C (table 1). For comparison, under conditions where reG3C9 exhibited a Tm of 60.4 °C, human PON1 exhibited a Tm of 47.2 °C (Gaidukov et al. 2009). Thus, by the most conservative estimate, the ancestral PONs exhibit T_m values that are up to 30 °C higher than human PON1 (i.e., without normalizing human PON1's Tm to the T_m value observed for reG3C9 in the current measurement; table 1).

Downloaded from https://academic.oup.com/mbe/article/33/10/2633/2925582 by guest on 20 August 2022

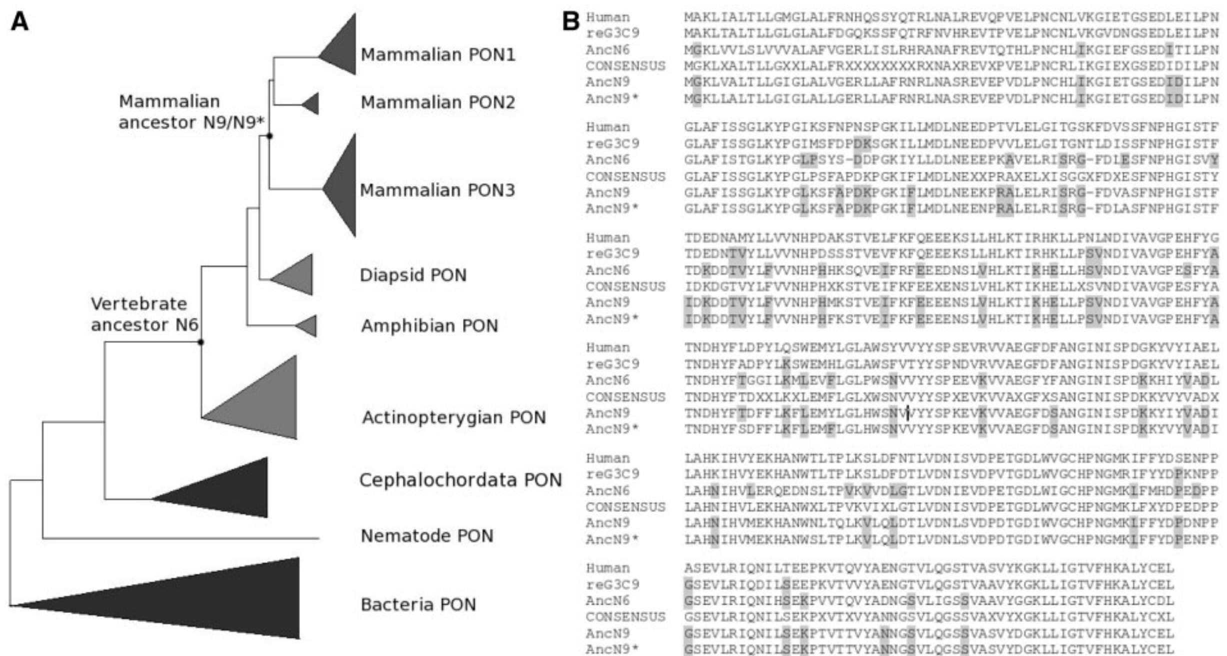

Fig. 1. (*A*) Schematic phylogenetic tree of PON sequences. This scheme is based on the actual phylogentic tree presented in Bar-Rogovsky et al. (2013). Noted are the reconstructed mammalian PON ancestor N9 (and the alternate ancestor N9* inferred from a phylogeny that includes the zebrafinch and lizard PON proteins) and the reconstructed vertebrate PON ancestor N6. (*B*) Alignment of the three ancestral PON variants discussed here, human PON1, and the recombinant PON1 variant G3C9. The consensus PON sequence is shown in bold, with X representing residues that could not be defined with a 30% cut-off. Residues that accommodated the consensus in the ancestral PONs relative to Human PON are highlighted in gray. A complete alignment is found in supplementary material, Supplementary Material online.

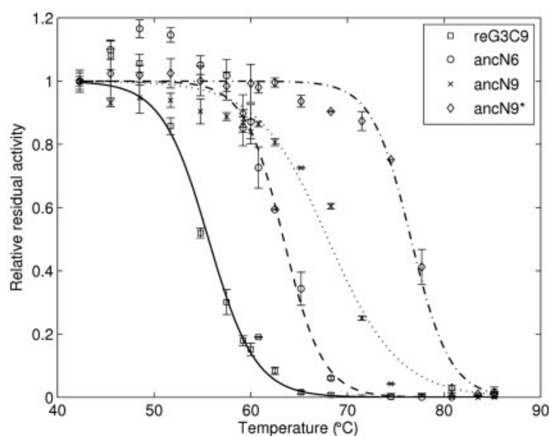

Fig. 2. Resistance of PON variants to thermal denaturation. The studied PON variants were incubated at different temperature for 30 min and their residual enzymatic activity was then measured at room temperature. Relative residual activity was derived from the initial rates of arylester hydrolysis and plotted as a fraction of the activity of the enzyme incubated at ambient temperature. The fit is based on a 2-parameter sigmoidal curve (see Materials and Methods).

Resistance to Calcium Chelation

We next tested the stability of the different PON variants to denaturation due to chelation of their structural calcium ion. PONs have two calcium ions—one ion resides at the bottom of the active site and is essential for catalysis (the "catalytic calcium"); the other is buried and is essential for maintaining the enzyme's structural integrity (the "structural calcium"). Although the catalytic calcium can be reversibly removed, removal of the structural calcium results in irreversible unfolding (Gaidukov 2009). Chelation using ethylenediaminetetraacetic acid (EDTA) causes human PON1 to lose over 80% of its activity within 15 min (Gaidukov et al. 2009). As shown in figure 3, following 2 h of incubation under the same conditions, reG3C9 lost >50% of its activity, and by 24 h, 80% of the activity was lost. In contrast, the ancestral variants ancN9, ancN9*, and ancN6 exhibited less than 20% decrease in activity over 96 h of incubation. Thus, the ancestral PONs exhibit extraordinarily high configurational stability regardless of whether the trigger for unfolding is high temperature or chelation of the structural calcium.

Resistance to Ionizing Radiation

As outlined in the introduction, we suggest that factors other than, or in addition to, high environmental temperature could have driven the hyperstability of past proteins. One example is ionizing radiation, and other sources of oxidizing species, that result in the chemical inactivation of proteins. Indeed, exposure to short-wavelength ultraviolet radiation (UV-C) is known to damage proteins, and proteomes seem to adapt to environments of high radiation (Krisko and Radman 2010). In this context, we tested the response of the studied PON variants to UV-C radiation. We irradiated PON variants with UV-C light (254 nm, 400–450 $\mu W/cm^2$) and measured their residual activity. As shown in figure 4, the variants vary in their ability to withstand radiation: The half-life of reG3C9 is approximately 2 h, whereas the half-life

Downloaded from https://academic.oup.com/mbe/article/33/10/2633/2925582 by guest on 20 August 2022

Table 1. Stability and Consensus Parameters of PON Variants

	Thermostability (T_m)	Chelation Stability (half-life)	UV-C Stability (half-life)	Consensus Identity (30%, 50%)[a]
Human PON1	47.2 °C[a]	<15 min[b]	Not measured	81%, 90%
reG3C9	55.6 ± 0.1 °C	20 min	2 h	80%, 85%
ancN6	63.4 ± 0.4 °C	>96 h	1 h	85%, 93%
ancN9	68.9 ± 1.1 °C	>96 h	4 h	92%, 97%
ancN9*	77.1 ± 0.5 °C	>96 h	>4 h	93%, 98%

[a]Consensus identity was determined by generating the consensus sequence for all PON sequences included in the tree schematically shown in figure 1, and comparing it to the amino acid sequence of each variant (with cut-offs of 30% and 50%, respectively).
[b]From Gaidukov et al. (2009). Although the same assay conditions were used, the previous measurements seem to have yielded higher T_m values as indicated by reG3C9 that exhibited a T_m of 60.4 °C (Gaidukov et al. 2009)—that is, 5 °C higher than observed here.

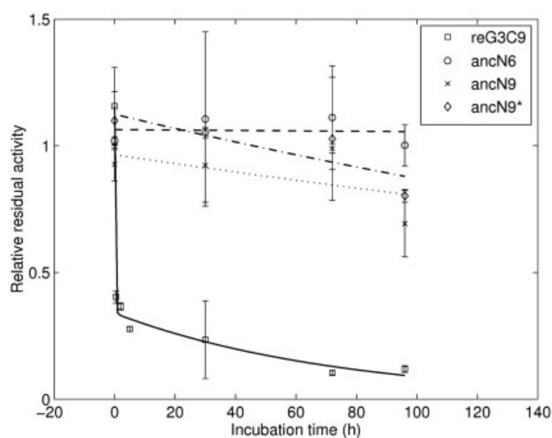

Fig. 3. Resistance of PON variants to calcium chelation. PON variants were incubated with the chelator EDTA (4 mM, at 37 °C and in the presence of 8 mM β-mercaptoethanol). Residual activity at various time points was determined by initial rates of dihydrocoumarin hydrolysis and plotted as a fraction of activity relative to enzyme incubated in activity buffer with no chelator. The fit line for reG3C9 corresponds to a two-phase exponential decay (see Materials and Methods).

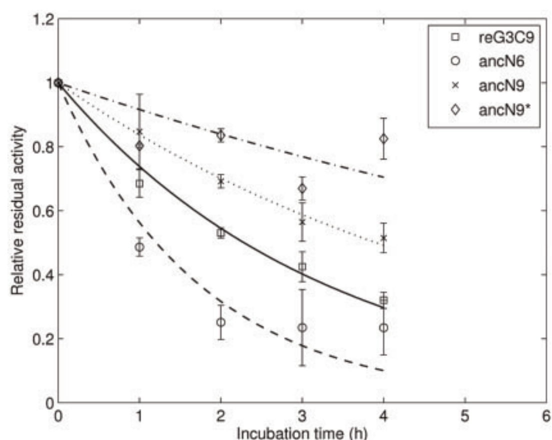

Fig. 4. The resistance of PON variants to UV-C radiation. PON variants were incubated under 254 nm UV light for up to 4 h. Residual activity at various time points was determined by initial rates of arylester hydrolysis and plotted as a fraction of activity of the untreated enzyme. The fit represents an exponential decay.

of ancN6, ancN9, and ancN9* is 1, 4, and ≥8 h, respectively. The two most thermostable PON variants, ancN9 and ancN9*, have half-lives that are over twice as long as reG3C9.

These data illustrate the notion that higher stability may relate to factors other than, or in addition to, high environmental temperature. Higher thermostability could, for example, endow higher resistance to chemical modifications induced by radiation, presumably by the same mechanism by which higher stability endows higher tolerance to mutations (see Discussion). However, thermostability does not strictly correlate with radiation stability (Chen et al. 2012): ancN6 is significantly more thermostable than reG3C9, but its UV half-life is reduced by about 50%. Among other factors, this deviation may relate to radiation damages inflicted on residues near or within the active site rather than on the protein scaffold, thus inactivating the enzymatic activity despite a hyperstable scaffold.

The Ancestral PONs Exhibit Unusually High Folding Efficiency

In addition to the unusually high stability of the folded, native states (as indicated by resistance to thermal denaturation and calcium chelation), the ancestral PONs also exhibited remarkably high folding efficiency. The latter is indicated by the high levels of soluble, functional enzymes obtained upon overexpression of the ancestral PONs in *E. coli* (supplementary fig. S1, Supplementary Material online). Expression of human or rabbit PON1, on the other hand, resulted in an inactive enzyme in the insoluble pellet fraction, with no soluble, functional enzyme observed (Aharoni et al. 2004). Rabbit and mouse PON3s yielded very low levels of soluble and functional enzyme, yet only when fused to thioredoxin at their N-termini. Fusion-free forms of PON1 and PON3 (i.e., with the original membrane-anchoring N-terminus) that functionally express in *E. coli* could only be obtained after directed evolution, yielding the variant reG3C9 used in this work (Aharoni et al. 2004), and other recombinant variants including PON3 variants described elsewhere (Khersonsky et al. 2009). In the case of PON2, none of the extant mammalian PONs tested expressed in a soluble functional form, and directed evolution failed to yield variants able to fold properly in *E. coli*, even when fused to thioredoxin. Moreover, wild-type PONs are glycosylated, and deglycosylation of human PON1 led to loss of activity (Brushia et al. 2001). But foremost, being

Downloaded from https://academic.oup.com/mbe/article/33/10/2633/2925582 by guest on 20 August 2022

comprised of beta-sheets (PONs have a six-bladed beta-propeller structure), PONs are generally prone to misfolding and aggregation. However, the inferred mammalian ancestor, even in a nonglycosylated form, exhibits high folding efficiency, relative not only to glycosylated human PON1 but also to a recombinant PON1 that was directly evolved toward higher foldability and stability.

Consensus Effects in Ancestral Reconstruction

To assess the degree to which the consensus effect may contribute to the observed hyperstability of the ancestral PONs, we calculated the identity of individual variants to the consensus sequence obtained from the same multiple sequence alignment used to infer the ancestors (fig. 1).

As shown in table 1, when using a cut-off of 30% to define the consensus amino acid, human PON1 exhibits 81% identity to the consensus (257 of 315 positions with a defined consensus contain the consensus amino acid). The recombinant G3C9 is 80% identical to the consensus (253/315 positions). In contrast, mammalian ancestors N9 and N9* have 92% (293/315) and 93% identity to the consensus, respectively, and N6 shows 85% identity. Applying a more stringent cut-off of 50% resulted in a similar trend—that is, the ancestral PONs are much closer to the consensus than human PONs, or for that matter, any another extant mammalian sequence (human PON = 214/239 = 90%, ancN6 = 223/239 = 93%, ancN9 = 233/239 = 97%, ancN9* = 234/239 = 98%). It therefore appears that, at least in the case of PONs, higher configurational stability correlates with higher similarity to the consensus ($R^2 = 0.9$; table 1). This effect is partly due to the fact that the mammalian clade is the most populated in the phylogeny, and the mammalian PON sequences are also far less diverged than the vertebrate ones. The correlation between consensus and thermostability is obviously limited—notably, the recombinant variant G3C9 is significantly more stable than human PON1, but G3C9 is as close to the consensus as is human PON1. Nonetheless, introduction of consensus mutations led to increased stability of a recombinant PON3 (Khersonsky et al. 2009), as well as of many other proteins (Steipe et al. 1994; Lehmann et al. 2002; Kiss et al. 2009; Komor et al. 2012; Sullivan et al. 2012). Further, introduction of the consensus amino acid may increase not only the stability of the native state (thermostability) but also foldability or kinetic stability (Godoy-Ruiz et al. 2006; Pey et al. 2008).

The consensus bias may also depend on the inference method. Specifically, maximum likelihood was found to be more biased toward the consensus than Bayesian methods (Williams et al. 2006). However, in our case, inference of the mammalian ancestor by the Bayesian maximum a posteriori method (Ronquist 2004) gave a nearly identical sequence to the experimentally constructed maximum-likelihood sequence (95.5%, identity) and the Bayesian ancestor had 1.3% higher identity to the consensus PON (supplementary table S1, Supplementary Material online). Random sampling of amino acids from the marginal posterior probabilities is also likely to minimize the consensus bias (Williams et al. 2006). Indeed, a similar strategy was applied for testing the robustness of prediction of the enzymatic properties (Bar-Rogovsky et al. 2015). Accordingly, we randomly sampled 100 mammalian ancestors from the posterior distribution. The posterior distribution samples were less biased than the Bayesian maximum a posteriori ancestor (the average identity to the consensus being ~1.2% lower). However, the identity of this random sample to the consensus sequence was essentially the same as for the maximum-likelihood ancestor (supplementary table S1, Supplementary Material online). It therefore appears that in the case of the mammalian PON ancestor, different inference methods and random sampling do not significantly reduce the consensus bias. Nonetheless, these strategies are essential in minimizing biases with respect to protein stability (Williams et al. 2006) and other traits (Bar-Rogovsky et al. 2015). Moreover, as indicated above, the correlation between thermostability and identity to the consensus is not a simple, linear one. However, we only examined the latter (identity to consensus) but did not test the former (i.e., experimentally determined thermostability of 100 randomly sampled variants).

Discussion

Vertebrates originated approximately 500 Ma (Maloof et al. 2010) and placental mammals approximately 65–100 Ma (Murphy et al. 2001; Luo 2007) when the average temperatures are predicted to have been similar to the current one (Robert and Chaussidon 2006). Mammals are also unlikely to have emerged in a high temperature niche, and body temperature regulation (homeothermy) makes their proteins even less sensitive to environmental temperatures. Therefore, the hyperstability of ancestral PONs, and especially the 30 °C higher T_m of the mammalian ancestor, is unlikely to relate to adaptation toward high environmental temperature. Two questions therefore arise: 1) Is the high stability of the inferred mammalian ancestor a property of the actual historical ancestor; or is it, at least in part, an artifact of the inference from extant sequences? 2) Assuming that the actual ancestor did exhibit a high degree of stability, what might have been the evolutionary forces driving this hyperstability?

Consensus Biases in Ancestral Inference

Our observation that the most stable ancestral PON variants are also the closest to the consensus sequence suggests that the observed hyperstability relates, at least in part, to the consensus effect. The specific combination of consensus and other sequence exchanges that give rise to the hyperstability of the inferred PON ancestors remains unknown. Higher stability may relate to improved core packing, as previously observed with PON1 (Roodveldt et al. 2005), reduced mobility of loops and/or to a range of surface effects including changes in surface charges (Socha and Tokuriki 2013). However, the ancestral variants show no global changes in charge (pI) compared with extant PONs. Indeed, predicting which sequence exchanges result in higher stability is challenging—overall trends for large data sets can be reliably predicted but not the effects of individual mutations (Potapov et al. 2009; Khan and Vihinen 2010). Regardless of the physicochemical origins of stabilization, the beneficial effects of

Downloaded from https://academic.oup.com/mbe/article/33/10/2633/2925582 by guest on 20 August 2022

individual consensus mutations and combinations thereof, on both the stability of the native state and smoother folding, have been widely demonstrated (e.g., see Steipe et al. 1994; Godoy-Ruiz et al. 2006; Komor et al. 2012; Sullivan et al. 2012; Mesa-Torres et al. 2014). However, the origins of the consensus effect are still under investigation (Godoy-Ruiz et al. 2006; Bloom and Glassman 2009) in particular because of the high degree of epistasis that underlies the divergence of protein sequences (Ashenberg et al. 2013; Pollock and Goldstein 2014; Risso et al. 2015). In principle, there are two possible scenarios: 1) The consensus as derived from the extant family sequences originates from the historical ancestor. Because of genetic drift, extant family members diverged from the historical ancestor, with each lineage deviating at a different subset of positions, and with a concomitant loss of stability over time. Because the deviations are sporadic, the historical, ancestral state can still be readily identified as the family consensus. 2) The consensus reflects an optimal state that is arbitrarily realized in some positions and not others. Within a given protein, a large set of positions drifts relatively rapidly, going toward and away from consensus such that, overall, a certain stability threshold is maintained. In this case, the inferred ancestral state, where many positions are in the consensus state, does not reflect the historical ancestor. The latter, like any other extant sequence, had a considerable number of positions in the off-consensus state. This scenario, if applied across the board implies that the stability of the historical ancestor was not fundamentally different from extant family members. After all, today's genes are tomorrow's ancestors.

A crucial question, therefore, is whether and how ancestral proteins differ from consensus proteins. Protein variants consisting of the consensus amino acids at all positions have been shown to be hyperstable yet often exhibit severely compromised activity (Lehmann et al. 2002; Kiss et al. 2009). Further, a comparison of consensus β-lactamases with the inferred ancestors indicated that although all ancestral variants had increased stability (T_m increases ranging from 20–40 °C), one of these consensus variants did not express, another had similar activity to the extant TEM-1 β-lactamase and only the third consensus β-lactamase had increased thermostability (>20 °C) (Risso et al. 2013). This result suggests that the observed properties of reconstructed ancestors are fundamentally different from those of artificially constructed consensus variants (Risso et al. 2013). We have not characterized the consensus PON, but the outcome is highly likely to be similar to the β-lactamase and other consensus proteins.

So while we agree with the assertion that consensus variants differ from ancestors, in our view, in itself, this result is not an indication that the consensus effect does not bias ancestral inference. This is because the concept of "consensus variants," that is, variants with all position in the consensus state, and even more so the inference of such variants, is highly problematic.

For example, in our analysis, we chose 30% and 50% cut-offs when deciding how common an amino acid needed to be. These cut-offs resulted in excluding only 40 positions from the analysis for the 30% cut-off, or 116 positions for the 50% cut-off. What is the optimal, let alone the correct cut-off, is entirely unknown. Accordingly, Risso et al. (2013) did not apply any frequency cut-off. Further, which amino acids should occupy the remaining, nonconsensus positions: Those of the extant reference sequence, or maybe the ancestral ones? Finally, which sequences should be included in the alignment for the purpose of identifying the consensus? For these reasons, an experimental characterization of a consensus PON (or any other protein family) is a futile exercise, as there is no consistent, reproducible way of deriving the consensus sequence. Ancestral inference relates to a phylogenetic tree and is therefore less sensitive to sampling biases, whereas the consensus sequence is a direct output of the sequence alignment and is therefore highly sensitive to sampling. Overall, that "consensus variants" fail to exhibit biochemical function and/or hyperstability, as do ancestors, is not an indication of the lack of consensus bias in ancestral reconstruction.

The correlation between consensus and thermostability is not universal to all lineages. A recent study determined the thermostability and folding properties of a series of ancestral bacterial ribonuclease H1s (RNHs) along the divergence of the extant mesophilic *E. coli* RNH and the thermophilic *Thermus thermophilus* RNH (Hart et al. 2014). In fact, the most thermostable variant along this trajectory was the extant *T. thermophilus* RNH, and not the common ancestor. Further, *T. thermophilus* RNH also is the least similar to the consensus (supplementary fig. S1, Supplementary Material online). Thus, in contrast to PONs, and many other cases, where reconstruction indicated a highly stable ancestor and a decrease in stability in essentially all diverging lineages, this study indicated divergence toward higher stability in the thermophilic clade and a decrease in stability along the mesophilic one. Thus, the degree to which the consensus effect biases individual reconstructions varies, and at present, this degree cannot be quantified.

The consensus bias may also depend on the sample of available extant sequences. We lack, for example, sequences of the platypus and of marsupial PONs, and these may critically impact the inferred ancestral sequences. As discussed above, the applied inference method also has an effect (Williams et al. 2006). Indeed, even with a 50% cut-off for consensus identity, the probability that the consensus state is correctly predicted at all sites is vanishingly small (0.5^{355} for a 355 amino acid protein like PON). Increased certainty through larger sequence alignments, and/or testing multiple ancestor/consensus variants may reduce but not totally alleviate these uncertainties.

Finally, several observations support the notion that reconstructed ancestors do reflect the hyperstability of the actual historical ancestors. Precambrian ancestors of very different proteins showed hyperstability (Gaucher et al. 2008; Hobbs et al. 2012; Akanuma et al. 2013; Risso et al. 2013, 2014, 2015). The observed increases in T_m values are often >20 °C, and some cases >30 °C, whereas the increases achieved by laboratory engineering and directed evolution, including by introducing consensus mutations, are typically much lower. For example, the reconstructed Precambrian β-lactamases

Downloaded from https://academic.oup.com/mbe/article/33/10/2633/2925582 by guest on 20 August 2022

exhibit approximately 35 °C higher T_m compared with the commonly studied descendant TEM-1 (Risso et al. 2013). Directed evolution of the latter for higher stability has been pursued (Hecky and Müller 2005; Kather et al. 2008), including by selecting for optimal combinations of consensus mutations (Dellus-Gur et al. 2013). However, even after several rounds of directed evolution, the observed increases in T_m are in the range of 14–18 °C.

Overall, we surmise that, at present, one cannot exclude the possibility that the hyperstability observed in certain reconstructed ancestors, such as the mammalian PON ancestor is, at least in part, an inference artifact. The error ranges in relation to the inference of stability might also be larger than for other properties such as affinity or catalytic efficiency, and conversely, in some cases the stability of the historical ancestor may be underestimated. In the future, statistical methods that address this bias may become available (see also Williams et al. 2006), thus alleviating the caveats associated with the inference of stability of ancestral proteins.

Possible Origins of Ancestral Hyperstability

As for the second question posed above, we would like to note that protein stability certainly relates to environmental temperature, but not only to temperature. For example, the reconstructed ancestral PONs exhibited a markedly high stability toward calcium chelation (fig. 3). Similarly, the mammalian PON ancestors were found to be more resistant to UV irradiation (fig. 4). The correlation between thermostability and radiation tolerance is not strict, most likely because radiation simultaneously affects proteins in a variety of different ways (Gray and Winkler 2015). In any case, one would not wish to draw any conclusions from these properties of the ancestral PONs about the calcium concentration in the serum of the mammalian ancestor, nor about the atmospheric levels of short-wavelength UV radiation. Rather, these properties exemplify the notion that selection toward higher stability may be driven by an entire range of physical and chemical challenges.

Another key challenge known to be buffered by higher stability is genetic mutations (Bloom et al. 2006; Tokuriki and Tawfik 2009). However, proteins of current organisms known to have exceedingly high rates of genetic mutations such as RNA viruses seem to exhibit ordinary, or even low protein stability (Tokuriki et al. 2009). A much more likely driving force for hyperstability are phenotypic mutations—transcriptional and/or translational errors. Translational mutations are, on average, $>10^5$ more frequent than genetic mutations, and occur at very high rates (up to 10^{-1} per codon) in some organisms (Ribas de Pouplana et al. 2014). Proteins were found to evolve toward higher stability in response to higher rates of these phenotypic mutations, primarily by acquiring stabilizing, consensus mutations (Goldsmith and Tawfik 2009; Bratulic et al. 2015). Phenotypic mutations also pave the road to the divergence of new proteins by genetic mutations that follow at much later stages (Whitehead et al. 2008; Yanagida et al. 2015). The hyperstability of some ancestral proteins may therefore relate, among other factors, to historically low fidelity of translation and/or transcription. Along the same veins, the absence or low efficiency of components of the proteome quality control machinery (chaperones, etc.) may also have driven proteins toward higher stability (Risso et al. 2013). Indeed, as discussed above, the mammalian PON ancestor also shows much higher folding efficiency than extant PONs.

To conclude, we surmise that in the case of the ancestral PONs, and possibly in other cases, one cannot exclude the possibility that the actual, historical ancestor(s) exhibited significantly lower stability than the reconstructed ones. At present, however, there might be ways to minimize various biases leading to high stability such as repeated sampling (Williams et al. 2006), but we are unaware of a way by which inference biases such as the consensus bias can be systematically quantified per given phylogeny and ancestor. We therefore advocate for more caution when interpreting the biophysical properties of reconstructed ancestors, as well as for further research into potential driving forces other than growth temperature that shape the stability of proteins.

Materials and Methods

Expression and Purification of PON Variants

PON variants were expressed and purified as described (Gaidukov et al. 2009). Briefly, His-tagged PON variants were cloned into a pET32a vector and expressed from *E. coli* BL21 (DE3). Following ammonium sulfate (55% w/v) precipitation, PON variants were purified on a Nickel-NTA resin (GE Healthcare), eluting with imidazole at 150 mM. Purified enzymes were dialyzed and stored in activity buffer (50 mM Tris pH 8.0, 50 mM NaCl, 1 mM $CaCl_2$, 0.1% tergitol). Protein concentrations were determined using the BCA assay. Purified protein yields from 500 ml cultures were, on average, about 5 mg.

Thermostability Assay

Aliquots of PON variants (20 µl of 1 µM enzyme) were incubated at a range of temperatures for 30 min. Samples were diluted to 200 µl in activity buffer with 2 mM of substrate, 4-nitrophenyl acetate. Absorbance at 405 nm was measured over 10 min in a PowerWave HT Microplate Spectrophotometer (Bio-Tek, Winooski, VT) and initial rates of 4-nitrophenyl acetate hydrolysis were determined. The temperature at which the residual activity was half of the maximum activity (T_m) was determined by plotting residual activities against the temperature using Matlab (Mathworks, Natick, MA) and fitting 2-parameter sigmoidal curves:

$$y = 1/\left[1 + (t/c)^b\right],$$

where y is the relative residual activity, t is the temperature, c is the T_m, and b is a scaling factor. Reported values relate to averages from at least two independent measurements.

Calcium Chelation Assay

Inactivation buffer (500 µl of 4 mM EDTA, 8 mM β-mercaptoethanol, 50 mM Tris–HCl, pH 8.0) was added to 500 µl of 2 µM enzyme in TBS plus 1 mM $CaCl_2$, and samples were incubated at 37 °C. Aliquots of 5 µl were removed at specified

Downloaded from https://academic.oup.com/mbe/article/33/10/2633/2925582 by guest on 20 August 2022

time points and diluted in 200 μl of activity buffer with 2 mM substrate, dihydrocoumarin. Absorbance at 270 nm was measured over 10 min in order to determine initial rates of dihydrocoumarin hydrolysis. The activity of the chelated samples was compared with unchelated samples (identically treated except for incubation in TBS + 1mM $CaCl_2$) to determine the residual activity. As previously observed (Gaidukov et al. 2009), PON1's loss of activity upon calcium chelation was found to follow a 2-phase exponential decay. Thus, residual activity as a function of incubation time was fitted to:

$$y = a * \exp(-k_{fast} * t) + b * \exp(-k_{slow} * t),$$

where y is the relative, residual activity, t is time, a and b are scaling factors, and k_{fast} and k_{slow} are rate constants for the fast and slow phases, respectively. The ancestral PONs exhibited very slow inactivation rates, and were thus fitted for a single-exponential decay.

Reported values were averaged from two independent measurements.

UV-C Radiation Assay

Aliquots of 20 μl of PON variants in activity buffer were incubated at 4 °C under 254 nm UV light (Spectroline USA handheld lamp, 400–450 μW/cm^2) for up to 4 h. Residual activity was determined with 4-nitrophenyl acetate as above. The irradiation time at which the residual activity was half of the maximum activity was determined by plotting relative residual activities against radiation time and fitting to a single exponential decay.

Reported values were averaged from two independent measurements.

Phylogenetic Reconstruction and Consensus Identity Determination

The same multiple sequence alignment used to infer the ancestral PONs by FastML (Bar-Rogovsky et al. 2013) as used to create a consensus amino acid sequence. The most common amino acid at each position was found and the resulting sequence was aligned to the human PON1, rePON1-G3C9, ancN9, ancN9*, and ancN6 PON variants. If the most common amino acid was present in <30% of the sequences, no consensus was defined and this position was excluded from the analysis. The "consensus identity" was then calculated as the number of positions at which the ancestral amino acid matches the consensus divided by the total number of positions where a consensus amino acid could be defined. A more rigorous cut-off of 50% was also tested (239 out of 355 total sites with a defined consensus).

Supplementary Material

Supplementary figures S1 and S2 and table S1 are available at *Molecular Biology and Evolution* online (http://www.mbe.oxfordjournals.org/).

Acknowledgments

The authors thank Moshe Goldsmith for helpful discussions. They also thank Susan Marquese and Joe Thornton for providing them with the RNH alignments and trees, and Jose Manuel Sanches-Ruiz for his critical comments and counter-arguments regarding the consensus bias effect described in the manuscript, which they have tried to express to the best of our ability. This work was supported by the Defense Threat Reduction Agency through grant [HDTRA1-11-C-0026]. D.L.T. is supported by a postdoctoral fellowship from the Alternative Energy Research Initiative at the Weizmann Institute of Science.

References

Aharoni A, Gaidukov L, Yagur S, Toker L, Silman I, Tawfik DS. 2004. Directed evolution of mammalian paraoxonases PON1 and PON3 for bacterial expression and catalytic specialization. *Proc Natl Acad Sci U S A*. 101:482–487.

Akanuma S, Nakajima Y, Yokobori S, Kimura M, Nemoto N, Mase T, Miyazono K, Tanokura M, Yamagishi A. 2013. Experimental evidence for the thermophilicity of ancestral life. *Proc Natl Acad Sci U S A*. 110:11067–11072.

Ashenberg O, Gong LI, Bloom JD. 2013. Mutational effects on stability are largely conserved during protein evolution. *Proc Natl Acad Sci U S A*. 110:21071–21076.

Bar-Rogovsky H, Hugenmatter A, Tawfik DS. 2013. The evolutionary origins of detoxifying enzymes: the mammalian serum paraoxonases (PONs) relate to bacterial homoserine lactonases. *J Biol Chem*. 288:23914–23927.

Bar-Rogovsky H, Stern A, Penn O, Kobl I, Pupko T, Tawfik DS. 2015. Assessing the prediction fidelity of ancestral reconstruction by a library approach. *Protein Eng Des Sel*. 28:507–518.

Bloom JD, Glassman MJ. 2009. Inferring stabilizing mutations from protein phylogenies: application to influenza hemagglutinin. *PLoS Comput Biol*. 5:e1000349.

Bloom JD, Labthavikul ST, Otey CR, Arnold FH. 2006. Protein stability promotes evolvability. *Proc Natl Acad Sci U S A*. 103:5869–5874.

Boucher JI, Jacobowitz JR, Beckett BC, Classen S, Theobald DL. 2014. An atomic-resolution view of neofunctionalization in the evolution of apicomplexan lactate dehydrogenases. *ELife* doi: 10.7554/eLife.02304

Bratulic S, Gerber F, Wagner 2015. Mistranslation drives the evolution of robustness in TEM-1 β-lactamase. *Proc Natl Acad Sci U S A*. 112:12758–12763.

Brushia RJ, Forte TM, Oda MN, La Du B, Bielicki JK. 2001. Baculovirus-mediated expression and purification of human serum paraoxonase 1A. *J Lipid Res*. 42:951–958.

Carroll SM, Bridgham JT, Thornton JW. 2008. Evolution of hormone signaling in elasmobranchs by exploitation of promiscuous receptors. *Mol Biol Evol*. 25:2643–2652.

Chen CD, Lin CH, Chuankhayan P, Huang YC, Hsieh YC, Huang TF, Guan HH, Liu MY, Chang WC, Chen CJ. 2012. Crystal structures of complexes of the branched-chain aminotransferase from Deinococcus radiodurans with α-ketoisocaproate and L-glutamate suggest the radiation resistance of this enzyme for catalysis. *J Bacteriol*. 194:6206–6216.

Dellus-Gur E, Toth-Petroczy A, Elias M, Tawfik DS. 2013. What makes a protein fold amenable to functional innovation? Fold polarity and stability trade-offs. *J Mol Biol*. 425:2609–2621.

Gaidukov L, Bar D, Yacobson S, Naftali E, Kaufman O, Tabakman R, Tawfik DS, Levy-Nissenbaum E. 2009. In vivo administration of BL-3050: highly stable engineered PON1-HDL complexes. *BMC Clin Pharmacol*. 9:18.

Gaucher EA, Govindarajan S, Ganesh OK. 2008. Palaeotemperature trend for Precambrian life inferred from resurrected proteins. *Nature* 451:704–707.

Godoy-Ruiz R, Ariza F, Rodriguez-Larrea D, Perez-Jimenez R, Ibarra-Molero B, Sanchez-Ruiz JM. 2006. Natural selection for kinetic stability is a likely origin of correlations between mutational effects on protein energetics and frequencies of amino acid occurrences in sequence alignments. *J Mol Biol*. 362:966–978.

Downloaded from https://academic.oup.com/mbe/article/33/10/2633/2925582 by guest on 20 August 2022

Goldsmith M, Tawfik DS. 2009. Potential role of phenotypic mutations in the evolution of protein expression and stability. *Proc Natl Acad Sci U S A*. 106:6197–6202.

Goldsmith M, Tawfik DS. 2013. Enzyme engineering by targeted libraries. *Methods Enzymol*. 523:257–283.

Gray HB, Winkler JR. 2015. Hole hopping through tyrosine/tryptophan chains protects proteins from oxidative damage. *Proc Natl Acad Sci U S A*. 112:10920–10925.

Hart KM, Harms MJ, Schmidt BH, Elya C, Thornton JW, Marqusee S. 2014. Thermodynamic system drift in protein evolution. *PLoS Biol*. 12:e1001994.

Hecky J, Müller KM. 2005. Structural perturbation and compensation by directed evolution at physiological temperature leads to thermostabilization of beta-lactamase. *Biochemistry* 44:12640–12654.

Hobbs JK, Shepherd C, Saul DJ, Demetras NJ, Haaning S, Monk CR, Daniel RM, Arcus VL. 2012. On the origin and evolution of thermophily: reconstruction of functional precambrian enzymes from ancestors of Bacillus. *Mol Biol Evol*. 29:825–835.

Howard CJ, Hanson-Smith V, Kennedy KJ, Miller CJ, Lou HJ, Johnson AD, Turk BE, Holt LJ. 2014. Ancestral resurrection reveals evolutionary mechanisms of kinase plasticity. *ELife* doi: 10.7554/eLife.04126.

Kather I, Jakob RP, Dobbek H, Schmid FX. 2008. Increased folding stability of TEM-1 beta-lactamase by in vitro selection. *J Mol Biol*. 383:238–251.

Khan S, Vihinen M. 2010. Performance of protein stability predictors. *Hum Mutat*. 31:675–684.

Kuo CL, La Du BN. 1995. Comparison of purified human and rabbit serum paraoxonases. *Drug Metab Dispos*. 23:935–944.

Khersonsky O, Rosenblat M, Toker L, Yacobson S, Hugenmatter A, Silman I, Sussman JL, Aviram M, Tawfik DS. 2009. Directed evolution of serum paraoxonase PON3 by family shuffling and ancestor/consensus mutagenesis, and its biochemical characterization. *Biochemistry* 48:6644–6654.

Kiss C, Temirov J, Chasteen L, Waldo GS, Bradbury ARM. 2009. Directed evolution of an extremely stable fluorescent protein. *Protein Eng*. 22:313–323.

Komor RS, Romero PA, Xie CB, Arnold FH. 2012. Highly thermostable fungal cellobiohydrolase I (Cel7A) engineered using predictive methods. *Protein Eng*. 25:827–833.

Krisko A, Radman M. 2010. Protein damage and death by radiation in Escherichia coli and Deinococcus radiodurans. *Proc Natl Acad Sci U S A*. 107:14373–14377.

Lehmann M, Loch C, Middendorf A, Studer D, Lassen SF, Pasamontes L, van Loon AP, Wyss M. 2002. The consensus concept for thermostability engineering of proteins: further proof of concept. *Protein Eng*. 15:403–411.

Luo Z. 2007. Transformation and diversification in early mammal evolution. *Nature* 450:1011–1019.

Maloof AC, Porter SM, Moore JL, Dudas FO, Bowring SA, Higgins JA, Fike DA, Eddy MP. 2010. The earliest Cambrian record of animals and ocean geochemical change. *Geol Soc Am Bull*. 122:1731–1774.

Mesa-Torres N, Yunta C, Fabelo-Rosa I, Gonzalez-Rubio JM, Sánchez-Ruiz JM, Salido E, Albert A, Pey AL. 2014. The consensus-based approach for gene/enzyme replacement therapies and crystallization strategies: the case of human alanine-glyoxylate aminotransferase. *Biochem J*. 462:453–463.

Murphy WJ, Eizirik E, O'Brien SJ, Madsen O, Scally M, Douady CJ, Teeling E, Ryder OA, Stanhope MJ, de Jong WW, Springer MS. 2001. Resolution of the early placental mammal radiation using Bayesian phylogenetics. *Science* 294:2348–2351.

Pauling L, Zuckerlandl E. 1963. Chemical paleogenetics: molecular restoration studies of extinct forms of life. *Acta Chem Scand*. 17(1 Suppl):9–16

Pey AL, Rodriguez-Larrea D, Bomke S, Dammers S, Godoy-Ruiz R, Garcia-Mira MM, Sanchez-Ruiz JM. 2008. Engineering proteins with tunable thermodynamic and kinetic stabilities. *Proteins* 71:165–174.

Pollock DD, Goldstein RA. 2014. Strong evidence for protein epistasis, weak evidence against it. *Proc Natl Acad Sci U S A*. 111:E1450.

Potapov V,Cohen M, Schreiber G. 2009. Assessing computational methods for predicting protein stability upon mutation: good on average but not in the details. *Protein Eng Des Sel*. 22:553–560.

Pupko T, Pe'er I, Shamir R, Graur D. 2000. A fast algorithm for joint reconstruction of ancestral amino acid sequences. *Mol Biol Evol*. 17:890–896.

Ribas de Pouplana L, Santos MA, Zhu JH, Farabaugh PJ, Javid B. 2014. Protein mistranslation: friend or foe? *Trends Biochem Sci*. 39:355–362.

Risso VA, Gavira JA, Gaucher EA, Sanchez-Ruiz JM. 2014. Phenotypic comparisons of consensus variants versus laboratory resurrections of Precambrian proteins. *Proteins* 82:887–896.

Risso VA, Gavira JA, Mejia-Carmona DF, Gaucher EA, Sanchez-Ruiz JM. 2013. Hyperstability and substrate promiscuity in laboratory resurrections of Precambrian β-lactamases. *J Am Chem Soc*. 135:2899–2902.

Risso VA, Manssour-Triedo F, Delgado-Delgado A, Arco R, Barroso-delJesus A, Ingles-Prieto A, Godoy-Ruiz R, Gavira JA, Gaucher EA, Ibarra-Molero B, et al. 2015. Mutational studies on resurrected ancestral proteins reveal conservation of site-specific amino acid preferences throughout evolutionary history. *Mol Biol Evol*. 32:440–455.

Robert F, Chaussidon M. 2006. A palaeotemperature curve for the Precambrian oceans based on silicon isotopes in cherts. *Nature* 443:969–972.

Ronquist F. 2004. Bayesian inference of character evolution. *Trends Ecol Evol*. 19:475–481.

Roodveldt C, Aharoni A, Tawfik DS. 2005. Directed evolution of proteins for heterologous expression and stability. *Curr Opin Struct Biol*. 15:50–56.

Smith SD, Wang S, Rausher MD. 2013. Functional evolution of an anthocyanin pathway enzyme during a flower color transition. *Mol Biol Evol*. 30:602–612.

Socha RD, Tokuriki N. 2013. Modulating protein stability—directed evolution strategies for improved protein function. *FEBS J*. 280:5582–5595.

Steipe B, Schiller B, Pluckthun A, Steinbacher S. 1994. Sequence statistics reliably predict stabilizing mutations in a protein domain. *J Mol Biol*. 240:188–192.

Sullivan BJ, Nguyen T, Durani V, Mathur D, Rojas S, Thomas M, Syu T, Magliery TJ. 2012. Stabilizing proteins from sequence statistics: the interplay of conservation and correlation in triosephosphate isomerase stability. *J Mol Biol*. 420:384–399.

Tokuriki N, Oldfield CJ, Uversky VN, Berezovsky IN, Tawfik DS. 2009. Do viral proteins possess unique biophysical features? *Trends Biochem Sci*. 34:53–59.

Tokuriki N, Tawfik DS. 2009. Stability effects of mutations and protein evolvability. *Curr Opin Struct Biol*. 19:596–604.

Whitehead DJ, Wilke CO, Vernazobres D, Bornberg-Bauer E. 2008. The look-ahead effect of phenotypic mutations. *Biol Direct*. 3:18.

Williams PD, Pollocks DD, Blackburne BP, Goldstein RA. 2006. Assessing the accuracy of ancestral protein reconstruction methods. *PLoS Comput Biol*. 2:e69.

Wilson C, Agafonov RV, Hoemberger M, Kutter S, Zorba A, Halpin J, Buosi V, Otten R, Waterman D, Theobald DL, et al. 2015. Kinase dynamics. Using ancient protein kinases to unravel a modern cancer drug's mechanism. *Science* 347:882–886.

Yanagida H, Gispan A, Kadouri N, Rozen S, Sharon M, Barkai N, Tawfik DS. 2015. The Evolutionary Potential of Phenotypic Mutations. *PLoS Genet*. 11:e1005445.

Downloaded from https://academic.oup.com/mbe/article/33/10/2633/2925582 by guest on 20 August 2022