Engineering functional thermostable proteins using ancestral sequence reconstruction

Received for publication, January 31, 2022, and in revised form, August 23, 2022 Published, Papers in Press, August 27, 2022,
https://doi.org/10.1016/j.jbc.2022.102435

Raine E. S. Thomson[1], **Saskya E. Carrera-Pacheco**[2], **and Elizabeth M. J. Gillam**[1,*]

From the [1]*School of Chemistry and Molecular Biosciences, The University of Queensland, Brisbane, Australia;* [2]*Centro de Investigación Biomédica (CENBIO), Facultad de Ciencias de la Salud Eugenio Espejo, Universidad UTE, Quito, Ecuador*

Edited by F. Peter Guengerich

Natural proteins are often only slightly more stable in the native state than the denatured state, and an increase in environmental temperature can easily shift the balance toward unfolding. Therefore, the engineering of proteins to improve protein stability is an area of intensive research. Thermostable proteins are required to withstand industrial process conditions, for increased shelf-life of protein therapeutics, for developing robust 'biobricks' for synthetic biology applications, and for research purposes (*e.g.*, structure determination). In addition, thermostability buffers the often destabilizing effects of mutations introduced to improve other properties. Rational design approaches to engineering thermostability require structural information, but even with advanced computational methods, it is challenging to predict or parameterize all the relevant structural factors with sufficient precision to anticipate the results of a given mutation. Directed evolution is an alternative when structures are unavailable but requires extensive screening of mutant libraries. Recently, however, bioinspired approaches based on phylogenetic analyses have shown great promise. Leveraging the rapid expansion in sequence data and bioinformatic tools, ancestral sequence reconstruction can generate highly stable folds for novel applications in industrial chemistry, medicine, and synthetic biology. This review provides an overview of the factors important for successful inference of thermostable proteins by ancestral sequence reconstruction and what it can reveal about the determinants of stability in proteins.

Native protein structures are complex three-dimensional arrangements of functional groups, which have evolved to carry out discrete biological functions that almost always depend on the maintenance of specific spatial relationships. However, native protein structures typically represent a metastable balance between conformational flexibility and stability that can be disturbed by environmental factors such as heat, organic solvents, chaotropic agents, and pH (1). Both enthalpic and entropic factors determine how a linear polymer of amino acid residues folds reproducibly into a specific structure, including intramolecular interactions between different structural elements and the degree of solvation of polar and hydrophobic regions of the structure (2). Any change to the sequence of a protein can affect these factors and therefore alter the ability of a polypeptide chain to fold into a functional structure.

Nature has explored only a small proportion of the available sequence space, so there is much scope to engineer novel proteins with useful properties. However, to be useful for industrial applications, most novel proteins must fold easily into stable domains (3, 4) (an exception being intrinsically disordered proteins), and so, an understanding of factors that underpin stable structures is essential for effective protein design. Studies have shown that more robust protein scaffolds are better able to accept potentially destabilizing mutations that confer novel activities or properties (5, 6). Indeed the robustness of different folds is a key factor behind the power law describing the extent to which different folds have been exploited in evolution: inherently stable folds are observed more commonly (7).

Enzymes represent a particular case where evolution has produced versatile and specific catalysts that can lower the activation energy of chemical reactions. Just as in nature, in industry, enzymes have the potential to improve the efficiency and sustainability of many chemical processes. Increasing the operational temperature of chemical reactions improves yield and reduces waste by enhancing reaction rates, improving reagent solubility and reducing microbial contamination; however, most native enzymes have limited stability even under their normal physiological conditions and are rapidly denatured at elevated temperatures. Since the biocatalyst (*i.e.*, the enzyme or a cell containing it) is often the most expensive part of a biocatalytic process, to be commercially competitive against chemocatalysis, the enzymes used need to have long operational lifetimes (3). While enzymes from thermophilic organisms are one option, it is rarely possible to find an enzyme in a thermophile with the catalytic profile of interest. Therefore, the operational stability of 'mesophilic' enzymes usually needs to be extended, and intensive efforts over the last ~40 years have been put toward engineering enzymes to be more thermostable.

Industrial biocatalysis is not the only motivation for stabilizing proteins however. Thermostable enzymes have also found wide application in basic research, for example, the PCR is only possible due to the use of thermostable polymerases,

* For correspondence: Elizabeth M. J. Gillam, e.gillam@uq.edu.au.

originally sourced from thermophiles, which enable the iterative replication and amplification of specific DNA templates. The development of numerous protein therapeutics has provided added impetus for engineering other types of protein for thermostability. Thermostable proteins have a longer shelf life and can be used in a wider range of therapeutic contexts than less stable proteins. More recently, the emergence of synthetic biology has expanded the use of independently folding and stable protein domains as biobricks in bioinspired devices. Yet another motivation for stabilizing proteins by engineering is that more stable homologs of proteins are often needed for structural and mechanistic studies, since they are typically more easily expressed and purified, and stand up better to biophysical characterization.

While analysis of proteins from thermophiles has provided valuable information on factors that can stabilize particular protein folds, not all proteins of interest have thermophilic homologs and it has become clear that the success of stabilization strategies is often dependent on the structural context. Good structural data are important for most rational and computational approaches to enhancing protein thermostability. However, structures are not always available, and the alternative, 'blind' approach of directed evolution usually requires intensive characterization of large libraries of mutants. Fortuitously, another source of inspiration from nature has emerged in recent years, namely the resurrection of thermostable ancestral enzymes, which alongside consensus approaches, leverages the huge expansion in available sequence from genome sequencing projects. This review will briefly summarize traditional approaches to engineering proteins for thermostability, then explore the use of ancestral sequence reconstruction (ASR) as an alternative strategy for engineering and elucidation of the determinants of thermostability.

Conventional approaches to the engineering of thermostability

The free energy difference between the folded and unfolded states of a protein is only ∼5 to 15 kcal/mol (1) and often only a few interactions are needed to stabilize a protein. However, determining the appropriate changes to make, without unwanted effects on protein function, has been an ongoing challenge. Figure 1 compares the alternative approaches to engineering thermostability in terms of information required, typical screening effort required, and the extent of sequence space that can be sampled.

Rational and computational design

Rational design methods have been used most commonly and have involved designing in improved hydrophobic core packing, salt bridges, and disulfide bonds. Alternatively constraining the most flexible regions of proteins by shortening loops, replacing glycine, and introducing proline residues has been useful. Critically, all of these approaches rely on having structural information of the protein of interest and involve some hypothesis as to the basis to the putative stabilization effect.

The success of different rational approaches will vary with the structural context presented by an individual protein and be affected by the complex landscape of epistatic interactions. Recently, rational design has been facilitated by numerous computational tools (reviewed by (6, 8–18), which have achieved notable successes (*e.g.*, (19) where the thermostability of an alcohol dehydrogenase was increased to ∼94 °C). Many computational tools rely on machine learning and extensive databases for training data. However, in such cases, the quality of the data available determines the accuracy of such tools and the available data are biased toward particular types of mutation (10). Current computational tools have difficulty modeling small but often critical alterations in stability (2, 20). Expansion and standardization of the information available from databases, plus high throughput approaches that can afford comprehensive data obtained under comparable conditions, such as deep mutational scanning and analysis of combinatorial data (21), may facilitate better predictions by augmenting training data. Artificial intelligence approaches, such as AlphaFold (22), should also make the prediction and design of protein stability more robust and are likely to lead to another step change. Importantly, AlphaFold predicts structures that can then be used as inputs for other methods that require them, such as PROSS (23).

Directed evolution

Directed evolution emerged in the 1990s as a useful 'blind' or 'brute force' technique for stabilization of proteins that was independent of any prior hypothesis concerning the mechanism of stabilization. It mimics the process of natural selection by using iterative rounds of genetic diversification (such as random mutagenesis or recombination of related sequences) combined with phenotypic screening and selection for high thermal stability and other required properties. In the absence of structural information on which to base hypotheses, random mutagenesis can be used to find residues that determine stability, which can then be targeted by saturation mutagenesis.

Directed evolution approaches employing entirely random methods for sequence diversification require large screening efforts to detect useful mutants (Fig. 1). While it is possible to, for example, assess activity at a stringent temperature in high throughput fashion, more detailed analysis of melting temperatures (T_ms) or temperatures at which half the population of proteins remains intact or active (T_{50} values) is resource intensive. Therefore, strategies that focus directed evolution efforts on smaller, more fertile areas of sequence space have been sought. The focus has been on identifying flexible regions to target (*e.g.*, by iterative saturation mutagenesis combined with B-factor analysis (24)) or using structure-guided computational approaches (25, 26).

Any given random mutation is more likely to be deleterious or neutral than beneficial (27), which places a limit on the number of random point mutations (typically one to two, maximum) that can be introduced per sequence, per iteration. Therefore, only a relatively small area of sequence space

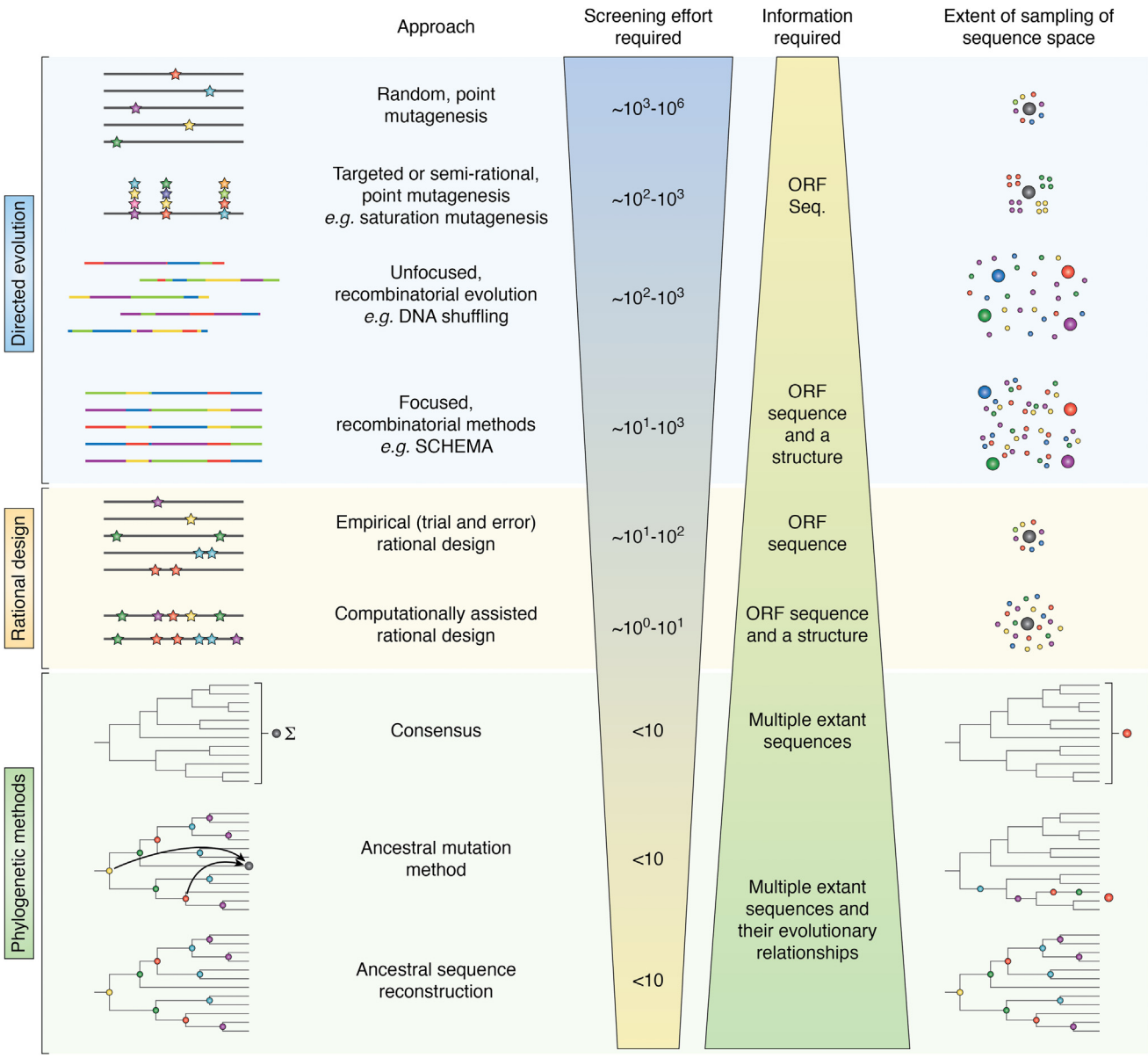

Figure 1. Comparison of approaches to engineering thermostability, in terms of typical screening effort (library size) and information required, and the extent of sequence space that can be sampled. Approaches are grouped broadly into directed evolution, rational (including computer-aided) design, and phylogenetic methods (*i.e.*, evolutionary methods that rely on data mining of sequences in natural evolutionary trees as opposed to directed evolution experiments). Note that there is some overlap between approaches (*e.g.*, site saturation mutagenesis can be used for rational design as well as directed evolution strategies; computational methods can be used to augment directed evolution and phylogenetic approaches), and different methods are often combined.

around the starting protein can be explored by point mutagenesis, due to the likelihood that deleterious mutations will accrue (Fig. 1). However, directed evolution approaches based on recombination of naturally occurring sequences can sample a larger volume of sequence space. Such libraries are enriched in functional mutants since, almost always, the residue introduced at a given position is found naturally, that is, has been 'vetted' by evolution in at least one of the parents (not eliminated by purifying selection). However, regions of homologous proteins that have diverged in different evolutionary branches and acquired different epistatic relationships with other structural elements in a protein fold, can be incompatible when fragments of homologs are recombined, leading to loss of stabilizing interactions or introduction of steric clashes or electrostatic repulsion.

Computational approaches have been applied to improve directed evolution strategies, just as for rational design. In particular, structure-based approaches have been used to increase the average structural integrity of mutant libraries created by recombinatorial evolution. Chief amongst these approaches is SCHEMA, which uses the sequences of homologous proteins and a representative structure to estimate optimal positions for recombination to minimize the disruption of interactions that stabilize the protein fold (28, 29). In an extension of this approach, Gaussian processes, a Bayesian learning technique that was trained on 242 measurements of

individual cytochrome P450 chimeras generated by a SCHEMA approach, was used to model the stability landscape of chimeric cytochrome P450 libraries and allowed the identification of a mutant that showed a further 5.3 °C increase in T_{50} (26). Importantly, mutants identified by these 'augmented' recombination approaches differ from the starting points in dozens to hundreds of positions, meaning they would not be readily identified by conventional rational or random point mutagenesis methods. In that respect, they are analogous to extensive fold optimization approaches enabled by Rosetta and other recent computational approaches to protein (re)design. The difference is that mutation and selection is used as the 'algorithm', leveraging evolutionarily proven folds found in nature as templates.

The consensus approach

Over the last ∼25 years, alternative approaches to leveraging the information implicit in natural evolutionary pathways for engineering thermostability have emerged, namely the consensus approach and ASR. Assuming the function of a protein confers a growth advantage on the organism, natural selection will tend to select for stabilizing residues and against residues that destabilize the structure. Thus, consensus residues are at least unlikely to be frankly destabilizing, unless they confer a selection advantage that is independent of, and greater than, the destabilizing effect. Therefore, consensus residues that are at least marginally stabilizing would tend to dominate a position over long-term evolution (30, 31). Many studies have taken advantage of this approach to improve stability by introducing 'consensus' residues at one or more positions in a protein of interest (*e.g.*, (30, 32–38); Figure 1). One advantage of this strategy is that it only requires a set of homologous sequences. However, the inference of which residues represent the 'consensus' can be heavily biased by imbalances in the amount of sequence information available for certain organisms relative to others. Consequently, it can be hard to dissociate stochastic or historical effects from the true consensus residues at a given position.

ASR

ASR has frequently yielded ancestor proteins that are more thermostable than their extant counterparts, providing some support for the hypothesis that primordial organisms were thermophilic. The earliest example was the inference of an ancestral sequence of 3-isopropylmalate dehydrogenase (IPMDH) from the last universal common ancestor (39). Seven ancestral residues introduced into an extant IPMDH found in an extreme thermophile, *Sulfolobus* strain 7, increased the thermostability of the extant form, supporting the idea that last universal common ancestor was a thermophile. Multiple complete elongation factor Tu (EF Tu) proteins from Precambrian (>∼500 million years ago; Ma) bacteria (40) were inferred by ASR using a phylogeny consisting of forms found in mesophilic, thermophilic, and hyperthermophilic bacteria. The EF Tu inferred at the node representing the most recent common ancestor of mesophilic bacteria was resurrected and found to have an optimal substrate-binding temperature of ∼55 °C compared to ∼37 °C for an extant EF Tu from mesophilic bacteria. The most basal ancestor of all lineages showed a comparable optimal temperature for substrate binding (∼65 °C) to extant forms from thermophiles. Analysis of seven intermediate ancestors revealed a trend of progressively increased thermostability going back in time from 0.5 to 3.5 billion years ago (Ga); the T_m values of the youngest ancestors were ∼44 to 48 °C compared to ∼65 to 74 °C for the oldest ancestors (41).

This foundational work was followed by similar studies in which sets of ancestral proteins of various evolutionary ages were resurrected and assessed for their thermostability as a means of assessing the experimental support for the existence of a thermophilic universal common ancestor (42–44), understanding the evolution of thermophily (45–47), and exploring the properties of ancestral proteins (48, 49). These studies have covered a broad range of protein families, phylogenetic taxa (bacteria and eukarya; including plants, animals, and fungi), and evolutionary ages (from a few hundred thousand years up to four billion years old). The prevailing observation has been that the mesostable proteins in existence today evolved from more thermostable forms. Enhancements in stability of ancestral proteins over directly related extant forms have ranged from a few degrees to more than 40 °C (Fig. 2; Table 1). However, it is also clear from recent bacterial phylogenies (45, 50) that thermophily may also have developed *de novo* in specific lineages of microorganisms that have evolved to fill niches in high temperature environments.

The use of ASR as an engineering technique

The observation that ancestral proteins were frequently more thermostable than their extant descendants (39, 41, 42, 44, 45, 51–53) inspired the use of ASR as a tool for engineering thermostable proteins, particularly for industrial biocatalysis. The earliest studies used the ancestral mutation method (Fig. 1), where a subset of residues from the inferred ancestor was introduced into an existing protein to alter its properties (51, 54). The T_{50} of an IPMDH from *Thermus thermophilus* was enhanced by 3.2 to 5.5 °C by introducing four residues from the IPMDH inferred in the common ancestor of Bacteria and Archaea (51, 55). Likewise, the T_m of β-amylase from *Bacillus circulans* was enhanced by 0.2 to 3.2 °C by introducing seven residues from the most ancient bacterial β-amylase (54).

This ancestral mutation method was further developed by Cole and Gaucher in the evolution-guided engineering approach, Reconstructing Evolutionary Adaptive Paths (REAP) (56), whereby phylogenetic and sequence analysis was used to identify amino acid substitutions in evolutionary branches that are likely to alter or enhance protein properties. REAP has also been promoted as a method for identifying ancestral residues that enhance thermostability (57); however, to date, there are no explicit examples of its implementation for this purpose to our knowledge.

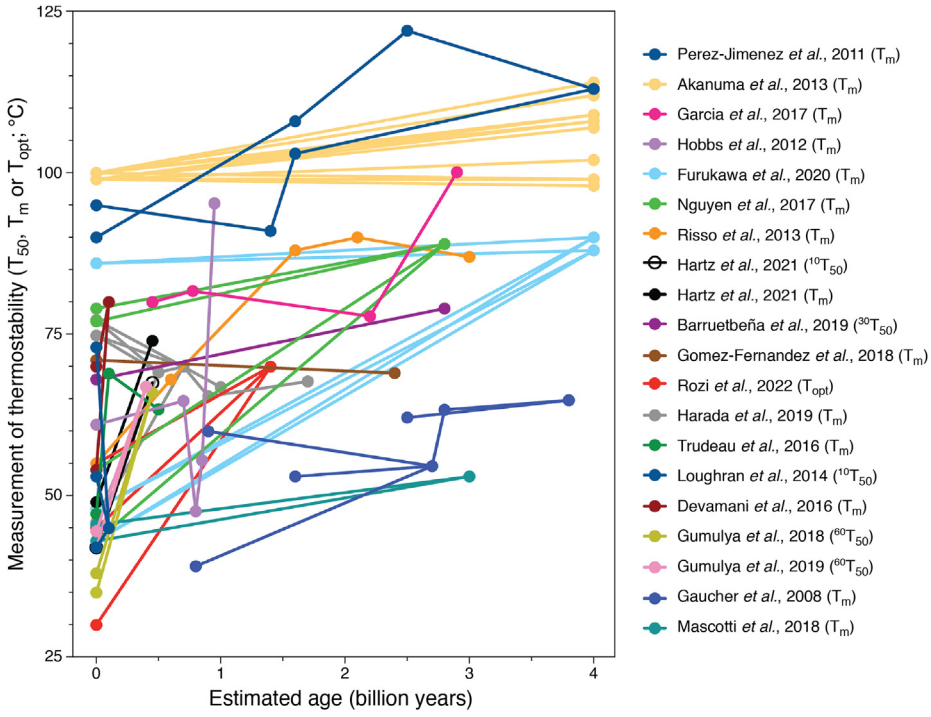

Figure 2. Changes in experimentally determined thermostability (T_{50} or T_m) *versus* estimated evolutionary age observed in resurrected ancestors compared to their related extant forms. An overall trend is seen toward greater thermostability in older ancestors but the magnitude of the effect differs markedly between different proteins and with the overall stability of the extant form. The data used in this analysis are from the studies listed in Table 1; only those that proposed an estimated age for respective ancestors are shown here. Different colors represent individual studies and for each phylogeny, directly related lineages are connected by solid lines. Sources in the order shown in the figure are: (41–43, 45, 47–49, 63, 64, 97, 109, 115–122).

More recent studies, facilitated by inexpensive gene synthesis, have resurrected complete ancestral forms for developing stabilized proteins. Some reports have focused on a specific application, such as using an ancestral coagulation factor VIII in infusion therapy to treat hemophilia (58), an ancestral phenylalanine/tyrosine ammonia-lyase for supplementary treatment of hereditary tyrosinemia (59) or an ancestral spiroviolene synthase to more easily obtain a crystal structure to elucidate structure and mechanism of the extant form (60). However, other studies have been undertaken on protein families generally appreciated for their potential as biocatalysts in chemical and pharmaceutical industries, bioremediation, biomass decomposition, biosensing, and cell imaging, amongst others (61–64). As with rational and directed evolution approaches, ASR has been incorporated into computational design algorithms (23).

The ASR process and factors influencing success

Despite the general success of ASR in generating stabilized proteins, there is a degree of uncertainty in all resurrected ancestors and errors in the inference may result in failure to express a folded, functional, or thermostable protein (65). There is no way to verify that a resurrected protein is historically accurate, but for the purposes of engineering a thermostable variant, this is less important than for studies of protein evolution. Nevertheless, the way in which the reconstruction is done can affect the sequence and consequently inferences drawn regarding the characteristics of the resultant proteins, such as stability or specific activity (65, 66). Studies on simulated datasets have revealed factors that lead to inconsistent or atypical results (67–69). Uncertainties in each of the individual inputs (*i.e.*, the multiple sequence alignment (MSA), tree and evolutionary model) compound in the inference process, with relatively greater impacts seen in the inference of ancestors from highly diverse groups of sequences, that is, where there is more ambiguity in the alignment and positions of insertions of deletions (indels).

Collection, curation, and alignment of sequences

Three major components are required for an ASR (Fig. 3): an MSA of available extant sequences, a phylogenetic tree showing their relationship to each other, and an evolutionary (substitution) model. Firstly, all available extant protein sequences descended from the ancestor of interest are aligned, along with sequences from an evolutionarily related outgroup. The sequence collection is perhaps the one most important factor influencing the quality of the inference, as there is the risk of including erroneous extant sequences or misaligning them (67, 69). A basic principle of any computational method applies: trash input leads to trash output. Including as wide a set of extant sequences as possible will strengthen the probability that the reconstruction will be accurate (70, 71). It is essential that the extant sequences are error free, yet sequence databases are rife with sequences containing transcription errors and miscalled exons, introns,

Table 1
Changes in thermostability of resurrected ancestors inferred in ASR studies compared to their extant counterparts

Protein	Taxon	Measure of stability	Ancestor stability (°C)	Descendant[a] stability (°C)	Δ stability (°C)	Estimated age[b]	Reference
Elongation factor Tu	All bacteria	Optimal binding	65	38–65	+0 to +17	Precambrian	(40)
	Mesophilic bacteria		55	38	+17	Precambrian	
Elongation factor Tu	All bacteria	T_m	65–73	40–64	+1 to +33	3.8 Ga	(41)
	Mitochondria and bacteria: Proteobacteria, cyanobacteria, *Thermus*, *Deinococcus*, Chloroflexi, chloroplast		63	40–64	−1 to +24	2.8 Ga	
	Proteobacteria, mitochondria		55–58	39–58	+2 to +18	2.7 Ga	
	Bacteria (Firmicutes)		60–62	46–48	+12 to +16	2.5 Ga	
	Mitochondria		51–53	–	–	1.6 Ga	
	Bacteria (α-Proteobacteria)		44–50	–	–	0.9 Ga	
	Bacteria (γ-Proteobacteria)		40	–	–	0.8 Ga	
Thioredoxin	All bacteria	T_m	113	89	+24	4.2 Ga	(48)
	All archaea, eukaryotes		113	91–122	−9 to +20	4.1 Ga	
	All archaea		113	122	−9	4 Ga	
	Bacteria: Cyanobacteria, Aquificae, *Deinococcus*, Chloroflexi, chloroplast.		122	1-	−1	2.5 Ga	
	Bacteria (γ-Proteobacteria)		108	89	+19	1.6 Ga	
	All eukaryotes		103	91–93	+10 to +12	1.6 Ga	
	Animals/fungi		91	93	−2	1.4 Ga	
3-isopropylmalate dehydrogenase	Bacteria (*Bacillus*)	T_m	65.3	47.6–64.7	+0.6 to +18	0.95 Ga	(45)
	Bacteria (*Bacillus*)		55.5	47.6–64.7	+0.6 to +18	0.85 Ga	
	Bacteria (*Bacillus*)		47.6	61–64.7	−17 to −13	0.8 Ga	
	Bacteria (*Bacillus*)		64.7	61	+3.7	0.7 Ga	
	All bacteria	T_m	88–90	43–86	+2 to +47	4 Ga	(115)
Nucleoside diphosphate kinase	All bacteria	T_m	98–109	99	−1 to +10	4 Ga	(42)
	All archaea		99–113	100	−1 to +13	4 Ga	
	Cyanobacteria	T_m	100	67–93	+7 to +33	2.9 Ga	(43)
	Nostocales		78	-	–	2.2 Ga	
	Viridiplantae		81–83	59–74	+7 to +24	0.775 Ga	
	Embryophyta		64–80	-	–	0.45 Ga	
	Cyanobacteria	T_m	68	46–75	−7 to +22	1.7 Ga	(119)
	Cyanobacteria		67	46–75	−8 to +21	1.0 Ga	
	Cyanobacteria		65	46–70	−5 to +19	0.9 Ga	
	Cyanobacteria		70	46–70	0 to +24	0.7 Ga	
	Cyanobacteria		70	46	+24	0.6 Ga	
	Cyanobacteria		69	46	+23	0.5 Ga	
β-lactamase	Bacteria (Gram +ve & -ve)	T_m	87	51–65	+22 to +36	3 Ga	(49)
	Bacteria (Gram +ve)		85–90	55–59	+26 to +35	2.1 Ga	
	Bacteria (γ-Proteobacteria)		88	55–59	+29 to +33	1.6 Ga	
	Bacteria (Enterobacteria)		68	55–59	+9 to +13	0.6 Ga	
Ribonuclease H1	Bacteria (α/β/γ/δ-Proteobacteria, *Thermus*, *Deinococcus*)	T_m	77	51–89	−12 to +26	–	(46)
	Bacteria (*Thermus*, *Deinococcus*)		77	89	−12	–	
	Bacteria (*Thermus*)		83	89	−6	–	
	Bacteria (α/β/γ/δ-Proteobacteria)		70	51–68	+2 to +19	–	
	Bacteria (γ-Proteobacteria)		68	51–68	+0 to +17	–	
	Bacteria (γ-Proteobacteria)		67	51–68	−1 to +16	–	
	Bacteria (Enterobacteria)		68	51–68	+0 to +17	-	
Hydroxynitrile lyase	Plants (Tracheophytes)	T_m	80	54–70	+10 to +26	<0.1 Ga	(121)
F420-dependent dehydrogenase	Bacteria/Archaea	T_m	53	43–46	+7 to +10	>3 Ga	(122)
Peroxidase	Plants	T_{50}	45	42–73	−28 to −3	0.11 Ga	(120)
Periplasmic binding protein	All Bacteria	T_m	∼75	52–80	−5 to +23	-	(123)
			∼80	–			
Serum paraoxonase	Vertebrates	T_m	63	47	+16	0.5 Ga	(97)
	Mammals		69	47	+22	0.1 Ga	
Haloalkane dehalogenase (and luciferase)	Bacteria/fungi	T_m	74	50–76	−2 to +24	–	(62)
	Bacteria/fungi		71	50–76	−5 to +21	–	
	Bacteria		73	54–76	−3 to +19	–	
	Bacteria		76	54–75	−1 to +22	–	
	Bacteria		75	54–59	+16 to +21	–	

Table 1—*Continued*

Protein	Taxon	Measure of stability	Ancestor stability (°C)	Descendant[a] stability (°C)	Δ stability (°C)	Estimated age[b]	Reference
	Cnidarians/Echinoderms	T_m	71	64	+7	–	(124)
Generic/Ligninolytic peroxidase	Fungi (Polyporales)	T_{50}	67	54–69	−2 to +13	0.15 Ga	(125)
	Fungi (Polyporales)		62	54–69	−7 to +10	–	
	Fungi (Polyporales)		69	54–58	+11 to +15	–	
	Fungi (Polyporales)		58	54	+4	–	
Adenylate kinase	Bacteria (Firmicutes)	T_m	89	48–88	+1 to +41	2.6–3 Ga	(47)
	Bacteria (Aerobic Firmicutes)		87	48–77	+10 to +39	–	
	Bacteria (Aerobic Firmicutes		77	48–76	+1 to +29	–	
	Bacteria (*Bacilli*)		76	48–76	0 to +28	–	
	Bacteria (*Bacilli*)		73	48–76	+3 to +25	–	
	Bacteria (*Bacilli*)		66	54–76	+10 to +12	–	
	Bacteria (*Bacilli*)		80	76	+4	–	
	Bacteria (*Bacilli*)		73	54	+19	–	
Diterpene cyclase	Bacteria (*Streptomyces*)	T_m	64	57–71	−7 to +7	–	(101)
	Bacteria (*Streptomyces*)		71	56–57	+14 to +15	–	
	Bacteria (*Streptomyces*)		56	57	−1	-	
Chalcone isomerase (CHI)/CHI-like	Land plants	T_m	82	40–80	+2 to +42	–	(126)
	Land plants		80	–	-	–	
	Land plants		88	50	+38	–	
Rubisco	Bacteria (Proteobacteria, Cyanobacteria,Firmicutes)	T_m	–	71	–	3.2	(117)
	Proteobacteria		69	71	−2	2.4	
	β/γ proteobacteria		–	71	–	1.9	
L-threonine 3-dehydrogenase	Bacteria (β–Proteobacteria, Cytophagia, Sphingobacteria, Flavobacteria)	T_m	56	50	+6	–	(100)
Ketol-acid reductoisomerase	Bacteria (Proteobacteria, Bacteroidetes, Verrucomicrobia, Fibrobacteres, Spirochetes)	T_m	59	43	+6	–	(63)
Cytochrome P450 (CYP3 Family)	Vertebrates	$^{60}T_{50}$	66	35–38	+28 to +31	0.45 Ga	(63)
Cytochrome P450 (CYP2D Subfamily)	Tetrapods	$^{60}T_{50}$	67	42–45	+22 to +25	0.4 Ga	(64)
Cytochrome P450, CYP11A Subfamily	Vertebrates	T_m	74	49	+25	0.4 Ga	(109)
		$^{10}T_{50}$	67.5	42	+25.5	–	
	Mammals	T_m	49	49	+0		
		$^{10}T_{50}$	45	42	+3		
Triosephosphate isomerase	Opisthokonta	T_m	66	59–66	+0 to +7	–	(127)
	All animals		66	54	+12	–	
	Vertebrates		54	–	–	–	
	Fungi		66	59–66	+0 to +7	–	
	Fungi		66	59	+7	–	
Endoglucanase	Bacteria (Firmicutes)	$^{30}T_{50}$	79	65–85	−6 to +14	2.8 Ga	(116)
L-arginine oxidase	Bacteria (γ-Proteobacteria)	$^{10}T_{50}$	92	65–81	+11 to +27	–	(128)
			81	65–74	+7 to +16	–	
			74	65	+9	–	
L-amino acid oxidase	–	$^{10}T_{50}$	40	63–64	−23 to +24	–	(65, 129)
	–		∼64	63	+1	–	
	–		∼63	–	–	–	
Fatty acid photo-decarboxylase	Algae	T_m	31–36/44–49.4	14–24/35.5–36.5	+7 to +22 +7.5 to +13.9	–	(130)
Geranylgeranylglyceryl phosphate synthase	Crenarchaeota, Thaumarchaeota, Euryarchaeota	T_m	>95–108	78–126	−31 to −18	–	(131)
	Bacteroidetes		>95	58–105	−10 to +37	–	
	Crenarchaeota, Thaumarchaeota, Euryarchaeota		>95–113	78–126	−31 to +35	–	
	Thaumarchaeota, Euryarchaeota		88–89	78–126	−38 to +11	–	
	Crenarchaeota		>95	–	–	–	
	Euryarchaeota		>95–118	>95–126	−31 to +23	–	
	Thaumarchaeota		81–85	78–80	+1 to +7	–	
	Bacteroidetes		58	>95–105	−37 to −47	–	
	Euryarchaeota		>95–103	>95–126	-31 to +8	–	
Lipase	Gram-negative bacteria	T_m	72	–	–	1.4 Ga	(118)
		T_{opt}	70	30–55	+15 to +40		

[a] Descendant stability refers to both extant forms and younger ancestors that are direct descendants of the corresponding form listed under Ancestor stability. Studies are arranged in order of the date of the first study on the protein concerned.

[b] Age estimates are only included where explicitly stated in the source.

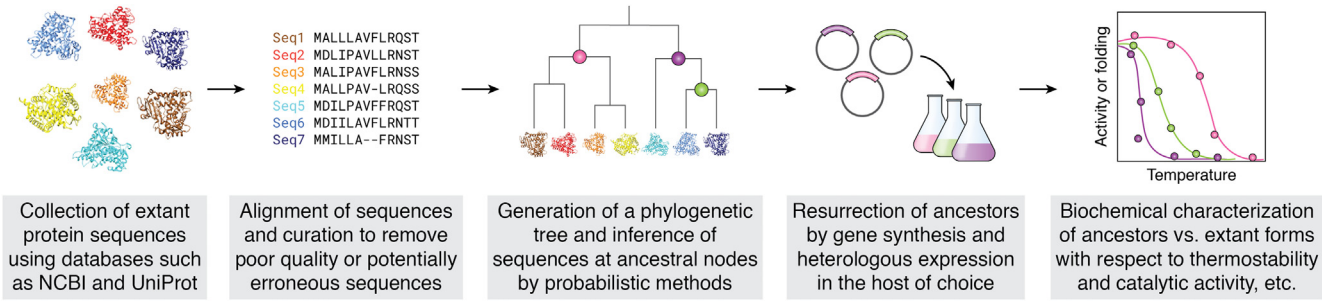

Figure 3. Outline of the ASR process. Extant protein sequences collected from sequence databases are iteratively aligned and curated to remove poor quality or potentially erroneous data then used to generate a phylogenetic tree. The tree, alignment, and an amino acid substitution model are used as inputs for ancestral inference using probabilistic methods. Ancestors from points of interest in the evolutionary tree are then reverse translated and the corresponding ORFs synthesized and expressed in a heterologous host, for example, *E. coli*. The resurrected ancestors can then be characterized for various biochemical properties or used as templates for further protein engineering. ASR, ancestral sequence reconstruction.

insertions, deletions, and frameshifts, many of which probably result from the imperfect interpretation of start, stop, and splice sites in the corresponding nucleotide sequences from genome sequencing studies. Since the speed of sequencing has vastly outpaced other biochemical approaches over the last decade, experimental verification is performed on only relatively few of the available sequences, and partial or erroneous sequences are frequently loaded into public sequence databases. Occasionally, the source for a DNA sequence is also misattributed, which can cause problems at the subsequent stage of tree generation. This means it is important to manually curate the MSA to remove sequences likely to contain errors, while maintaining the broadest possible coverage of the extant sequence space (Fig. 4). In our experience, this process is the most labor-intensive part of undertaking an ASR, especially with very large sequence alignments, yet is rarely mentioned in descriptions of the approach and often not given sufficient consideration in computational tools that incorporate ASR in strategies for the design of thermostable proteins.

Various alignment tools are available, of which maximum likelihood (ML) methods are generally preferred. However, handling of indels is a point of difference (reviewed in (71)); problems with interpreting indels have been shown to cause artefactual lengthening of ancestors (69). Standard two-dimensional arrays of aligned residues involve implicit judgments (or guesses!), as to which residues in an alignment are homologous, whereas in practice, such relationships are far from obvious in highly variable regions of an alignment or between distantly related proteins. Methods that represent relationships between residues in sets of proteins as partial order alignment graphs that do not require decisions to be made as to the position of deletions and insertions are under development (70, 71) and should minimize the confounding effects of subjective decisions about alignments on ancestral inferences (69).

Inference of the phylogenetic tree

The second requirement is a phylogenetic tree explaining the evolutionary relationships between these sequences (Fig. 3). This can be taken from the literature if a well-corroborated gene tree is available that includes all extant branches for which sequences are available; however, it is more commonly inferred from the MSA using a statistical (*e.g.*, ML) approach. It is important that the tree is as accurate as possible and bootstrapping is used to evaluate the topology, as the inference at each ancestral node is dependent on its position in the tree, the lineages that it gave rise to, and their order of evolution. Even small differences in the MSA can affect the relative position of different branches. Importantly, the gene tree for the protein under reconstruction is not necessarily the same as the accepted species tree due to factors such as incomplete lineage sorting (72). Nonetheless, in the absence of significant horizontal gene transfer, there is usually general agreement in the overall topology and inferences have been shown to be relatively robust to difference in phylogenetic trees (68). Similarly, the third input, the choice of evolutionary model (discussed further below) has been found to have less impact on ASR accuracy than the alignment (73).

Ancestral inference methods

Once the MSA and phylogenetic tree have been refined, a method of statistical inference is applied, which uses the information in the MSA and phylogenetic tree and an evolutionary model to predict the ancestral state at all internal nodes of the tree (Fig. 3). There are three inference methods that have typically been used for ASR studies, maximum parsimony (MP; (74)), ML (75) and Bayesian inference (BI; (76)). There is no definitively correct inference method, and ASR tools are continually being developed to increase accuracy. No single ASR tool has been preferred in the literature; however, ML methods are used most commonly.

MP methods

MP methods were the first to be developed and are based on the principle of parsimony, that is, the simplest explanation of an event or observation is the preferred explanation. In the context of ASR, MP infers ancestral states that minimize the total number of character changes required to give rise to the sequences observed at the tips of the phylogenetic tree.

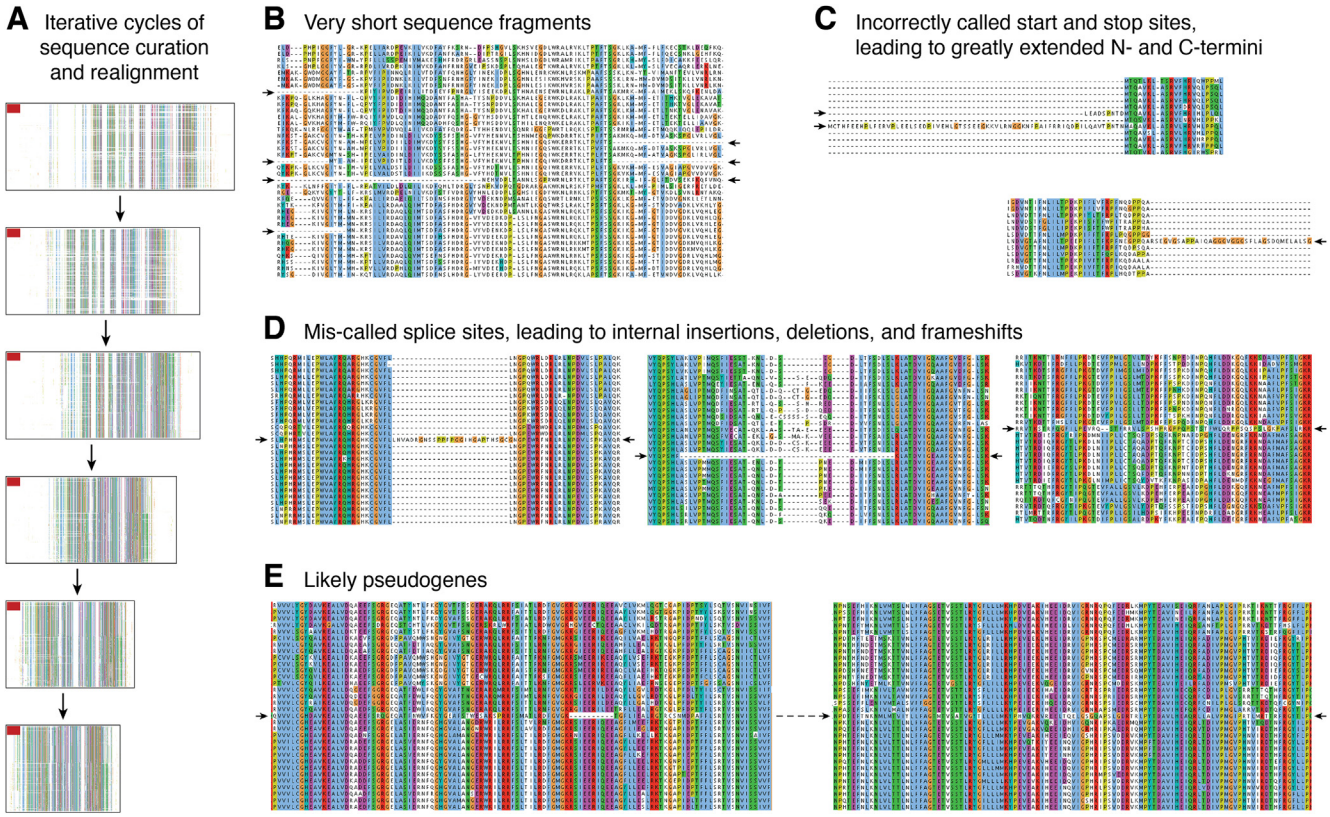

Figure 4. Examples of sequence curation required for ASR. *A*, representative changes in the overall alignment during sequence curation. The *red* rectangle at the *top left* of each image shows an equivalent area of the alignment. The overall number of sequences decreases during curation as sequences with likely artefacts (insertions, deletions, and frameshifts) resulting from miscalling of start, stop, and splice sites are removed. Removal of such sequences, especially those containing insertion artefacts, improves the ability to align the remaining sequences such that the overall alignment length decreases markedly. *B–E*, *arrows* indicate sequences with likely artefacts. *B*, very short sequence fragments are typically removed since they may not encode a functional protein, whereas sequences that lack a small proportion of the overall coding sequence at the N or C termini can be retained without disrupting the ASR. *C*, incorrectly called start and stop sites lead to massively extended sequences, which appear as clear outliers in sequence alignments. If these sequences are retained in the alignment used for the ASR, the inferred ancestors will have similar artefactual extensions, so extensions are typically pruned to the consensus start and stop sites. *D*, artefactual insertions, deletions, and frameshifts appear as sequences with marked differences to phylogenetic near-neighbors over an extended area of the alignment. Such artefacts are readily visible in highly conserved regions but may not be apparent in regions of higher variability or in alignments with highly diverse sequences. Biochemical expertise can also be used to interpret the likelihood of these sequences being correct, that is, from what is known about the structure, a prediction be made as to whether the fold would tolerate such a disruption to the typical sequence. *E*, likely pseudogenes are evident from a pattern of numerous, possibly minor deviations from the sequence of phylogenetic near-neighbors distributed across the ORF. ASR, ancestral sequence reconstruction.

While efficient, the parsimony method has several shortcomings (77). First, at positions that have changed more than once across the tree, there are often several equally parsimonious ancestral states and there is no way of selecting which is most likely to be correct. This becomes more problematic as the degree of diversity between the terminal extant sequences increases (78) and therefore only ancestors of extant sequences that are well conserved can be unambiguously reconstructed, making this method unsuitable for highly diverse protein groups. A second criticism is that MP oversimplifies evolution and does not consider amino acid substitution biases. An MP algorithm ranks all evolutionary changes as equally probable when, in reality, some mutations, for example, conservative amino acid changes, are more likely than others (79). Thirdly, MP methods assume that the same amount of evolutionary time has passed along every branch of the tree, ignoring branch length, and therefore preferentially choosing evolutionary paths in which one mutation occurred along a short branch rather than alternatives where multiple changes have occurred along a long branch (80).

ML methods

With the recognition that MP approaches oversimplified evolution, phylogenetic methods based on likelihood estimation were developed and have been most commonly used for ASR to date. ML accounts for the fact that not all mutation events are equally likely to occur (81) by incorporating the use of an amino acid substitution rate matrix that describes the probability of different mutations based on a hypothetical evolutionary model. ML evaluates the probability of every possible ancestral state at every residue based on the probability that all the residues found at that site at the tips of the tree would have evolved given this ancestral amino acid state, the phylogeny, and the evolutionary model. The inferred ancestral sequence is that which maximizes the likelihood at all positions. The more widely used amino acid substitution models are the Dayhoff (82), Whelan, and Goldman (WAG; (83)), Le and Gascuel (LG; (84)), and Jones-Taylor-Thornton (JTT; (85)). All are empirical models developed from different databases of protein sequences and commonly implemented in ASR tools. It is not possible *a priori* to

determine the most suitable model to use; typically, a number of models will be tested in parallel and the one that best fits the protein family of interest (typically as evaluated from the gene tree obtained) will be chosen. While these empirical models remain popular, many more sophisticated substitution models have been developed that take into account constraints on protein folding and epistatic interactions and which have been found to be more accurate (86). However, these models are not well established yet in ASR studies because they are more computationally intensive and not as easily incorporated into the currently used phylogenetic frameworks.

ML ancestors can be inferred in a "marginal" or "joint" manner (75). In a joint reconstruction, the most likely ancestral state at all internal nodes is inferred, whereas a marginal reconstruction infers the most likely sequence at a single node. By focusing on a single node, marginal reconstruction is more efficient and tends to be more commonly used in ASR studies; however, it does not necessarily give the globally optimal sequence and can only be considered as an approximation to the joint reconstruction (87).

One shortcoming of ML is that it does not account for uncertainty in the reconstruction. It assumes that the phylogenetic tree and evolutionary model are accurate, which is often not true, particularly for highly divergent proteins, and this can lead to errors in the inference.

BI

Like ML, BI is a probabilistic method; however, BI incorporates uncertainty into the reconstruction (76). Rather than providing a single best estimate for an internal node, BI provides the posterior probability of the ancestral state. There are two methods of BI, the more simplistic empirical method (75) and the complex hierarchical method (76). Empirical BI is computationally similar to ML, but rather than calculating the most likely ML character state based on their respective probability distributions, the probability distributions are reported directly. In an empirical Bayesian approach, the posterior probability distribution is calculated based on a single phylogenetic tree and evolutionary model and does not account for uncertainty in these parameters. Therefore, empirical BI still faces the issue of inference errors due to inaccurate assumptions. The more complex hierarchical Bayesian approach incorporates uncertainty about the phylogeny and evolutionary model into the reconstruction. This method calculates the posterior probability of the ancestral state by averaging its probability over all possible trees and models of evolution, weighted by how likely these trees and models are, given the observed data (76). While the hierarchical Bayesian approach is superior in its ability to incorporate uncertainty, it is computationally intensive and realistically limited to analyzing relatively small numbers of sequences.

ML was shown to be sufficiently robust to phylogenetic uncertainty that there was no significant benefit from using BI (68). In addition, a study assessing the accuracy of MP, ML, and BI methods found that ML was the most accurate with an average of ∼94% of all sites correctly inferred, compared to ∼92% using BI and ∼90% using MP (88).

The reliability of outcomes from ASR

ASR can only infer the most probable ancestor based on the inputs provided. Therefore, in many studies, multiple ancestral sequences have been inferred by different methods to assess how robust the observed ancestral properties (*e.g.*, thermostability) are to the alternative inputs and the use of different algorithms (analogous to sampling the experimental error by performing replicates in a typical biochemical experiment). Alternative tree topologies (40–42, 68, 89), evolutionary/amino acid substitution models (41, 90), and methods of statistical inference (45, 46, 63) have been assessed in parallel. Marginal ML approaches provide the posterior probability of each residue at each position in an ancestor, so multiple, plausible ancestors by choosing alternative residues in ambiguous positions of the ancestral sequence (91)—the so-called 'ancestral cloud' approach (41, 42, 49, 63, 92, 93)—or by resurrecting an 'alt-all' ancestor that has the least likely, yet still plausible, alternative residue at all ambiguous sites (94). In all such studies, the measured thermostability of the alternative ancestors has proven to be remarkably robust to methodological differences.

Early studies using the ancestral mutation method also indirectly addressed the idea that thermostability seen in ancestral proteins was the result of chance. When ancestral residues were introduced into extant IPMDH (39), isocitrate dehydrogenase (44) and glycyl tRNA synthetase (52) the proportion of mutants that showed improved thermostability were 5/7, 4/5, and 6/8, respectively. The likelihood of increasing the thermostability of a protein using random mutagenesis is substantially lower with one of the most successful reports being that of esterase, where only 1/3 mutants were found to have improved stability (95, 96), consistent with most random mutations being likely to be deleterious than advantageous with respect to any property (27).

Ancestral reconstructions versus *consensus approaches*

It has been proposed that there is an inherent systematic bias in the statistical inference methods used in ASR that results in overestimation of protein stability. Evidence for this was first proposed in a seminal study that assessed the accuracy of MP, ML, and BI using computational population evolution simulations (88). ML was found to overestimate stability by ∼1.5 kcal/mol compared to 0.4 and 0.05 kcal/mol, using MP and BI, respectively. It was proposed that the stabilizing bias of ML and MP was due to the tendency of these methods to infer consensus residues as the most likely ancestral residues (88, 97).

Various studies have compared the thermostability of consensus variants *versus* ancestral proteins (42, 56, 98–101) and in most cases have found the ancestral form to be superior. Cole and Gaucher (56) generated a consensus EF Tu protein along with the ancestral form from the Last Bacterial Common Ancestor (LBCA) with which it shared 76% sequence

identity. While the consensus variant showed a ∼20 °C increase in T_m over an extant EF Tu from *Escherichia coli* (60 °C *versus* 39 °C), the LBCA ancestor showed a higher T_m (73 °C). Likewise, a consensus nucleoside diphosphate kinase (T_m of 84 °C) showed lower thermostability than 14 ancestors obtained using several inference methods and tree topologies, for the nodes representing the LBCA and Last Archeal Common Ancestor (T_m values of 99–114 °C; 42). Four ancient bacterial β-lactamases (0.5, 1, 1.5, and 2 billion years old; 49) and three consensus β-lactamase sequences were generated (99). Two of these consensus sequences were inferred from the same set of sequences used initially (49), one of which did not express and the other showed a T_m of 60 °C *versus* 88 °C for the ancestral form. The third consensus was calculated from a broader set of sequences than used to infer the ancestors and showed a T_m of 79 °C, which was lower than three of the four resurrected ancestors (49). Other studies that assessed the effect of consensus mutations on β-lactamases were only able to reach a maximum T_m of 61.5 °C (31) and 66.2 °C (32) compared to ∼90 °C achieved by ASR (99).

It is reasonable to expect that many residues in a consensus mutant are actually ancestral; since, as explained by Tawfik *et al.* (97), a good proportion of consensus residues may have originated in the ancestor, with the sequences in successive lineages changing through genetic drift. The stochastic nature of genetic drift means that the ancestral state can still be identified as the consensus (97). However, it is highly unlikely that any given ancestor would show the consensus amino acid at all the potentially biased positions.

Experimental characterization of consensus proteins has revealed that they often display severely compromised activity, are completely inactive, or are not expressed at all (33, 99, 101, 102), issues that are rarely reported for ancestral sequences. These deficiencies may be due to the consensus approach failing to account for epistatic interactions, that is, combining incompatible residues that arise in different lineages as a result of divergent evolution. In addition, unless some type of weighting is applied, consensus sequences are biased toward the clades or species that have received intensive attention in sequencing projects and therefore make up a larger proportion of available sequence information.

Unlike the methods of ML and MP, BI is not thought to be biased toward consensus residues and may only slightly overestimate stability (88). Observing high thermostability in ancient Bayesian-inferred ancestors would be strong evidence that thermostability is not an artifact of the inference method; however BI is less commonly used overall than ML for ASR studies. Various studies have compared ancestors inferred using ML and BI to assess whether there is a difference, yet no consistent bias has been seen. The sequences of the ML and BI versions of two IPMDH (LeuB) ancestors differed by ∼7% to 10%, that is, 25 to 36 amino acids (45); the optimal temperature for activity (T_{opt}) of the ML inferred ancestors was 46 °C and 70 °C *versus* 64 °C and 68 °C, respectively, for the corresponding BI versions. The T_m values of ML ancestors for three adenylate kinases were +4, +0.3, and -6 °C different from the corresponding BI ancestors (47). Importantly, in addition to assessing the differences in thermostability, studies have compared the theoretical accuracies of the two methods and have shown ML approaches to be just as accurate and in some cases more accurate than BI, despite their potential to overestimate stability (68, 88, 103).

Ultimately, even if ML does bias toward thermostability, an extra ∼1.5 kcal/mol (88) can be equated to a ∼6 °C change in T_m (based on experiments measuring the effect of point mutations on the thermal stability of bacteriophage T4 lysozyme that indicate a change of ∼4 °C in the folding temperature for every kcal/mol change in ΔΔG (104)). This increase is well below what has been observed in many ASR studies that have shown enhancements in stability of ∼30 to 35 °C in resurrected ancestors (Table 1).

What can be learned from the types of interactions underpinning ancestral thermostability?

As ancestral forms often have dozens, if not hundreds, of residue changes from an extant form or between any two nodes in a given tree, it is difficult to identify those changes that are responsible for conferring stability, as noted in many studies (45, 51, 97, 105, 106). Some studies have proposed stabilizing mechanisms based on sequence information alone; however, it is uncommon that strong correlations are observed between any particular biochemical property and thermostability (45). Increased hydrophobicity has also been found to correlate with higher T_m in ancestral EF Tu proteins (107). Other proposed stabilizing mechanisms include improved core packing, reduced mobility of loops, and changes in surface charges, all similar to observations from studies improving the thermostability of extant forms (97).

Understanding the context of a residue in the protein tertiary structure is important in predicting its stabilizing effect, and this has been facilitated by crystallization of some ancestral proteins (42, 47, 105, 106, 108). A comparison of the crystal structures of three adenylate kinase ancestors with two extant forms revealed several unique salt bridges that disappear sequentially in forms with decreasing thermostability but are found in more thermostable extant forms (47). The crystal structure of an ancestral nucleoside diphosphate kinase revealed a reduction in its nonpolar accessible surface and increased numbers of intersubunit ion pairs and hydrogen bonds (42). Molecular dynamics simulations performed using the crystal structures of ancestral, consensus, and extant EF Tu proteins revealed stabilizing networks of ionic and hydrophobic interactions and a greater average buried area in more thermostable forms. However, even with a crystal structure, it is not always straightforward to identify stabilizing interactions. Despite the availability of crystal structures of seven Precambrian thioredoxins that were up to 24 °C more stable than extant forms, no significant differences were identified between the extant and ancestral forms, in terms of polar or apolar solvent-accessible surface areas, the number of hydrogen bonds or salt bridges, or surface charge distributions (105).

One way to confirm the stabilizing residues of interest is to test their effect experimentally. Studies using the ancestral

mutation method have assessed the stabilizing effect of one or a few ancestral residues and revealed hydrophobic packing, hydrogen bonding, and the formation of ion pairs (39, 44, 47, 51, 52, 54) to be important. In other mutagenesis studies, residues in the ancestral proteins have been altered to abolish stability as a way of identifying important interactions (47, 63, 109).

Thus, analysis of ancestral structures has provided complementary information to studies on proteins from thermophiles and underscored the importance of a variety of stabilizing interactions. Importantly, however, ASR has allowed identification of thermostable homologs of proteins for which thermostable variants are either not available or phylogenetically remote (*e.g.*, for the very diverse cytochrome P450 family), simply on the basis of abundant sequence information. This should make it more straightforward to determine the causative changes above the background noise of neutral drift. In our experience, a correlation is generally seen between evolutionary age and thermostability (110), so choosing ancestors that are more or less closely related should enable specific changes to be identified, by restricting the number of differences between the forms in question. However, to date, insufficient studies have been performed for the full benefit of this approach to be realized.

Perspectives for future use of ASR to engineer thermostability

Proteins are the fundamental agents that achieve chemistry on biological timescales, transmit and receive signals at the molecular level, and serve as structural modules from which many cellular structures are built up in nature. Therefore, they are also the principal feedstock and inspiration for efforts to (re) design biological catalysts, protein therapeutics, metabolic pathways, signal transduction relays, biosensors, synthetic gene circuits, and other novel bioinspired 'devices' for chemical, biotechnological, and synthetic biology applications. However, a two-step approach is often needed to protein engineering, the first being to make the protein more thermostable in order to buffer the potentially destabilizing effects of mutations needed for subsequent optimization of function (12).

It is ironic that such forward looking fields should gain inspiration from ASR, an approach that looks back in time through the evolutionary record. Natural sequence diversity is a rich resource of functional structures, but using ASR to explore the evolutionary history of protein sequence, structure, and function markedly extends the toolbox available for protein engineers. Ancestral proteins represent additional diversity that is enriched in functional and robust proteins; the 'extinct' intermediates in the evolutionary record must vastly outnumber the collection of forms extant today. Importantly, neither structural data nor extensive screening of mutant libraries is required for ASR, only extensive sequence information (that is increasingly available from genome sequencing efforts) combined with bioinformatic tools for interpretation of protein evolution. Much is to be gained by integrating across multiple approaches to protein engineering, for example, recombining stable structural modules from ancestral proteins (111) and applying advanced methods for computational design to ancestors used as robust scaffolds for designing new proteins (12).

While there is no ideal method for inferring the most probable ancestor and ASR is highly dependent on the quality of the sequence information and alignment used, structure-aware approaches (reviewed recently in (112)) and tailoring of evolutionary models show great promise for improving confidence in the inferences obtained. Improvements in machine learning (113) and particularly *ab initio* structure prediction (*e.g.*, AlphaFold), should accelerate the improvement of ASR approaches, the engineering of proteins using ancestral templates, and the interpretation of information gained from studying ancestral proteins. Indeed, stable ancestors offer opportunities for obtaining insights into the structure and function of poorly characterized protein families that may not be feasible, or are at least much more challenging to achieve, with extant proteins.

One particularly exciting prospect is to rerun evolution *in vitro* from robust ancestors and apply different, artificial selection pressures, both to optimize the properties of proteins to match the needs of industrial or medical applications, but also to reveal how such properties develop in the absence of confounding, pleiotropic influences that constrain evolution *in vivo*. Combining such experiments with machine learning (113), molecular dynamics simulations, and advanced biophysical methods for structure determination, including the analysis of (un)folding pathways (114), should provide insights into how changes in sequence, structure, and conformation affect function. Such studies promise to inform our understanding of natural proteins but also efforts to reshape them to provide clever, bioinspired solutions to global challenges across fields as diverse as medicine, industrial chemistry, agriculture, and environmental management, indeed any area in which proteins can serve useful purposes.

Acknowledgments—This work was supported by Australian Research Council Discovery Project Grants DP160100865 and DP200102837, and by AstraZeneca Innovative Medicines and Early Development, Cardiovascular, Renal and Metabolism, Gothenburg. We thank Dr Gabriel Foley for helpful comments on preliminary drafts of this review and Anthony Bengochea for providing sequence alignments as examples for Fig. 4. The authors acknowledge the usefulness of consulting a table in the PhD thesis submitted by Kurt Harris (132) in directing them to existing literature which was re-analysed in Table 1. The authors apologize for being unable, due to space constraints, to comprehensively reference the many studies done to engineer thermostability over the last ~40 years; rather, typical examples of each approach have been cited where appropriate.

Author contributions—E. M. J. G. conceptualization; R. E. S. T. and S. E. C. P. investigation; R. E. S. T., S. E. C. P., and E. M. J. G. writing–original draft; R. E. S. T., S. E. C. P., and E. M. J. G. writing–review & editing; R. E. S. T., S. E. C. P., and E. M. J. G. visualization; E. M. J. G. supervision.

Conflict of interest—The authors are engaged in directed evolution efforts to produce thermostable cytochrome P450 enzymes for biocatalysis and synthetic biology applications, some of which have

been licensed for application in pharmaceutical and fine chemical production under the tradename "CYPerior." The authors declare that they have no conflicts of interest with the contents of this article.

Abbreviations—The abbreviations used are: ASR, ancestral sequence reconstruction; BI, Bayesian inference; IPMDH, 3-isopropylmalate dehydrogenase; LBCA, Last Bacterial Common Ancestor; ML, maximum likelihood; MP, maximum parsimony; MSA, multiple sequence alignment.

References

1. Pace, C. N. (1990) Conformational stability of globular proteins. *Trends Biochem. Sci.* **15**, 14–17
2. Baker, D. (2019) What has *de novo* protein design taught us about protein folding and biophysics? *Protein Sci.* **28**, 678–683
3. Bommarius, A. S., and Paye, M. F. (2013) Stabilizing biocatalysts. *Chem. Soc. Rev.* **42**, 6534–6565
4. Burton, S. G., Cowan, D. A., and Woodley, J. M. (2002) The search for the ideal biocatalyst. *Nat. Biotech.* **20**, 37–45
5. Tokuriki, N., and Tawfik, D. S. (2009) Stability effects of mutations and protein evolvability. *Curr. Opin. Struct. Biol.* **19**, 596–604
6. Socha, R. D., and Tokuriki, N. (2013) Modulating protein stability - directed evolution strategies for improved protein function. *FEBS J.* **280**, 5582–5595
7. Magner, A., Szpankowski, W., and Kihara, D. (2015) On the origin of protein superfamilies and superfolds. *Sci. Rep.* **5**, 8166
8. Dombkowski, A. A., Sultana, K. Z., and Craig, D. B. (2014) Protein disulfide engineering. *FEBS Lett.* **588**, 206–212
9. Pongsupasa, V., Anuwan, P., Maenpuen, S., and Wongnate, T. (2022) Rational-design engineering to improve enzyme thermostability. In: Magnani, F., C, M., F, P., eds. *Enzyme Engineering: Methods and Protocols*, Humana Press, New York, NY: 159–178
10. Modarres, H. P., Mofrad, M. R., and Sanati-Nezhad, A. (2016) Protein thermostability engineering. *RSC Adv.* **6**, 115252–115270
11. Ó'Fágáin, C. (2011) Engineering protein stability. In: Walls, D., Loughran, S. T., eds. *Protein Chromatography: Methods and Protocols*, Humana Press, New York, NY: 103–136
12. Weinstein, J., Khersonsky, O., and Fleishman, S. J. (2020) Practically useful protein-design methods combining phylogenetic and atomistic calculations. *Curr. Opin. Struct. Biol.* **63**, 58–64
13. Eijsink, V. G. H., Bjork, A., Gaseidnes, S., Sirevag, R., Synstad, B., van den Burg, B., *et al.* (2004) Rational engineering of enzyme stability. *J. Biotechnol.* **113**, 105–120
14. Eijsink, V. G. H., Gaseidnes, S., Borchert, T. V., and van den Burg, B. (2005) Directed evolution of enzyme stability. *Biomol. Eng.* **22**, 21–30
15. Steipe, B. (2004) Consensus-based engineering of protein stability: from intrabodies to thermostable enzymes. *Protein Eng.* **388**, 176–186
16. Razvi, A., and Scholtz, J. M. (2006) Lessons in stability from thermophilic proteins. *Protein Sci.* **15**, 1569–1578
17. Wijma, H. J., Floor, R. J., and Janssen, D. B. (2013) Structure- and sequence-analysis inspired engineering of proteins for enhanced thermostability. *Curr. Opin. Struct. Biol.* **23**, 588–594
18. Sun, Z. T., Liu, Q., Qu, G., Feng, Y., and Reetz, M. T. (2019) Utility of B-factors in protein science: interpreting rigidity, flexibility, and internal motion and engineering thermostability. *Chem. Rev.* **119**, 1626–1665
19. Aalbers, F. S., Fürst, M. J. L. J., Rovida, S., Trajkovic, M., Gómez Castellanos, J. R., Bartsch, S., *et al.* (2020) Approaching boiling point stability of an alcohol dehydrogenase through computationally-guided enzyme engineering. *eLife* **9**, e54639
20. Huang, P., Chu, S. K. S., Frizzo, H. N., Connolly, M. P., Caster, R. W., and Siegel, J. B. (2020) Evaluating protein engineering thermostability prediction tools using an independently generated dataset. *ACS Omega* **5**, 6487–6493
21. Nisthal, A., Wang, C. Y., Ary, M. L., and Mayo, S. L. (2019) Protein stability engineering insights revealed by domain-wide comprehensive mutagenesis. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 16367–16377
22. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., *et al.* (2021) Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589
23. Goldenzweig, A., Goldsmith, M., Hill, S. E., Gertman, O., Laurino, P., Ashani, Y., *et al.* (2016) Automated structure- and sequence-based design of proteins for high bacterial expression and stability. *Mol. Cell* **63**, 337–346
24. Reetz, M. T., Carballeira, J. D., and Vogel, A. (2006) Iterative saturation mutagenesis on the basis of B factors as a strategy for increasing protein thermostability. *Angew. Chem. - Int. Ed.* **45**, 7745–7751
25. Romero, P. A., and Arnold, F. H. (2012) Random field model reveals structure of the protein recombinational landscape. *PLoS Comput. Biol.* **8**, e1002713
26. Romero, P. A., Krause, A., and Arnold, F. H. (2013) Navigating the protein fitness landscape with Gaussian processes. *Proc. Natl. Acad. Sci. U. S. A.* **110**, E193–E201
27. Guo, H. H., Choe, J., and Loeb, L. A. (2004) Protein tolerance to random amino acid change. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 9205–9210
28. Voigt, C. A., Martinez, C., Wang, Z. G., Mayo, S. L., and Arnold, F. H. (2002) Protein building blocks preserved by recombination. *Nat. Struct. Biol.* **9**, 553–558
29. Otey, C. R., Silberg, J. J., Voigt, C. A., Endelman, J. B., Bandara, G., and Arnold, F. H. (2004) Functional evolution and structural conservation in chimeric cytochromes P450: calibrating a structure-guided approach. *Chem. Biol.* **11**, 309–318
30. Steipe, B., Schiller, B., Plückthun, A., and Steinbacher, S. (1994) Sequence statistics reliably predict stabilizing mutations in a protein domain. *J. Mol. Biol.* **240**, 188–192
31. Bershtein, S., Goldin, K., and Tawfik, D. S. (2008) Intense neutral drifts yield robust and evolvable consensus proteins. *J. Mol. Biol.* **379**, 1029–1044
32. Amin, N., Liu, A. D., Ramer, S., Aehle, W., Meijer, D., Metin, M., *et al.* (2004) Construction of stabilized proteins by combinatorial consensus mutagenesis. *Protein Eng. Des. Sel.* **17**, 787–793
33. Lehmann, M., Loch, C., Middendorf, A., Studer, D., Lassen, S. F., Pasamontes, L., *et al.* (2002) The consensus concept for thermostability engineering of proteins: further proof of concept. *Protein Eng.* **15**, 403–411
34. Lehmann, M., Pasamontes, L., Lassen, S. F., and Wyss, M. (2000) The consensus concept for thermostability engineering of proteins. *Biochim. Biophys. Acta* **1543**, 408–415
35. Kohl, A., Binz, H. K., Forrer, P., Stumpp, M. T., Plückthun, A., and Grütter, M. G. (2003) Designed to be stable: crystal structure of a consensus ankyrin repeat protein. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 1700–1705
36. Sullivan, B. J., Durani, V., and Magliery, T. J. (2011) Triosephosphate Isomerase by consensus design: dramatic differences in physical properties and activity of related variants. *J. Mol. Biol.* **413**, 195–208
37. Rath, A., and Davidson, A. R. (2000) The design of a hyperstable mutant of the Abp1p SH3 domain by sequence alignment analysis. *Protein Sci.* **9**, 2457–2469
38. Di Nardo, A. A., Larson, S. M., and Davidson, A. R. (2003) The relationship between conservation, thermodynamic stability, and function in the SH3 domain hydrophobic core. *J. Mol. Biol.* **333**, 641–655
39. Miyazaki, J., Nakaya, S., Suzuki, T., Tamakoshi, M., Oshima, T., and Yamagishi, A. (2001) Ancestral residues stabilizing 3-isopropylmalate dehydrogenase of an extreme thermophile: experimental evidence supporting the thermophilic common ancestor hypothesis. *J. Biochem.* **129**, 777–782
40. Gaucher, E. A., Thomson, J. M., Burgan, M. F., and Benner, S. A. (2003) Inferring the palaeoenvironment of ancient bacteria on the basis of resurrected proteins. *Nature* **425**, 285–288
41. Gaucher, E. A., Ganesh, O. K., and Govindarajan, S. (2008) Palaeotemperature trend for Precambrian life inferred from resurrected proteins. *Nature* **451**, 704–707
42. Akanuma, S., Nakajima, Y., Yokobori, S., Kimura, M., Nemoto, N., Mase, T., *et al.* (2013) Experimental evidence for the thermophilicity of ancestral life. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 11067–11072

43. Garcia, A. K., Schopf, J. W., Yokobori, S.-i., Akanuma, S., and Yamagishi, A. (2017) Reconstructed ancestral enzymes suggest long-term cooling of Earth's photic zone since the Archean. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 4619
44. Iwabata, H., Watanabe, K., Ohkuri, T., Yokobori, S.-i., and Yamagishi, A. (2005) Thermostability of ancestral mutants of *Caldococcus noboribetus* isocitrate dehydrogenase. *FEMS Microbiol. Lett.* **243**, 393–398
45. Hobbs, J. K., Shepherd, C., Saul, D. J., Demetras, N. J., Haaning, S., Monk, C. R., *et al.* (2012) On the origin and evolution of thermophily: reconstruction of functional Precambrian enzymes from ancestors of Bacillus. *Mol. Biol. Evol.* **29**, 825–835
46. Hart, K. M., Harms, M. J., Schmidt, B. H., Elya, C., Thornton, J. W., and Marqusee, S. (2014) Thermodynamic system drift in protein evolution. *PLoS. Biol.* **12**, e1001994
47. Nguyen, V., Wilson, C., Hoemberger, M., Stiller, J. B., Agafonov, R. V., Kutter, S., *et al.* (2017) Evolutionary drivers of thermoadaptation in enzyme catalysis. *Science* **355**, 289–293
48. Perez-Jimenez, R., Inglés-Prieto, A., Zhao, Z., Sanchez-Romero, I., Alegre-Cebollada, J., Kosuri, P., *et al.* (2011) Single-molecule paleoenzymology probes the chemistry of resurrected enzymes. *Nat. Struct. Mol. Biol.* **18**, 592–596
49. Risso, V. A., Gavira, J. A., Mejia-Carmona, D. F., Gaucher, E. A., and Sanchez-Ruiz, J. M. (2013) Hyperstability and substrate promiscuity in laboratory resurrections of Precambrian β-lactamases. *J. Am. Chem. Soc.* **135**, 2899–2902
50. Coleman, G. A., Davin, A. A., Mahendrarajah, T. A., Szantho, L. L., Spang, A., Hugenholtz, P., *et al.* (2021) A rooted phylogeny resolves early bacterial evolution. *Science* **372**, eabe0511
51. Watanabe, K., Ohkuri, T., Yokobori, S., and Yamagishi, A. (2006) Designing thermostable proteins: ancestral mutants of 3-isopropylmalate dehydrogenase designed by using a phylogenetic tree. *J. Mol. Biol.* **355**, 664–674
52. Shimizu, H., Yokobori, S. I., Ohkuri, T., Yokogawa, T., Nishikawa, K., and Yamagishi, A. (2007) Extremely thermophilic translation system in the common ancestor commonote: ancestral mutants of glycyl-tRNA synthetase from the extreme thermophile Thermus thermophilus. *J. Mol. Biol.* **369**, 1060–1069
53. Boussau, B., Blanquart, S., Necsulea, A., Lartillot, N., and Gouy, M. (2008) Parallel adaptations to high temperatures in the Archaean eon. *Nature* **456**, 942–U974
54. Yamashiro, K., Yokobori, S.-I., Koikeda, S., and Yamagishi, A. (2010) Improvement of *Bacillus circulans* β-amylase activity attained using the ancestral mutation method. *Protein Eng. Des. Sel.* **23**, 519–528
55. Watanabe, K., and Yamagishi, A. (2006) The effects of multiple ancestral residues on the Thermus thermophilus 3-isopropylmalate dehydrogenase. *FEBS Lett.* **580**, 3867–3871
56. Cole, M. F., and Gaucher, E. A. (2011) Exploiting models of molecular evolution to efficiently direct protein engineering. *J. Mol. Evol.* **72**, 193–203
57. Cole, M. F., Cox, V. E., Gratton, K. L., and Gaucher, E. A. (2013) Reconstructing evolutionary adaptive paths for protein engineering. *Met. Mol. Biol.* **978**, 115–125
58. Zakas, P. M., Brown, H. C., Knight, K., Meeks, S. L., Spencer, H. T., Gaucher, E. A., *et al.* (2016) Enhancing the pharmaceutical properties of protein drugs by ancestral sequence reconstruction. *Nat. Biotech.* **35**, 35–37
59. Hendrikse, N. M., Larsson, A. H., Gelius, S. S., Kuprin, S., Nordling, E., and Syren, P.-O. (2020) Exploring the therapeutic potential of modern and ancestral phenylalanine/tyrosine ammonia-lyases as supplementary treatment of hereditary tyrosinemia. *Sci. Rep.* **10**, 1315
60. Schriever, K., Saenz-Mendez, P., Rudraraju, R. S., Hendrikse, N. M., Hudson, E. P., Biundo, A., *et al.* (2021) Engineering of ancestors as a tool to elucidate structure, mechanism, and specificity of extant terpene cyclase. *J. Am. Chem. Soc.* **143**, 3794–3807
61. Wilding, M., Peat, T. S., Kalyaanamoorthy, S., Newman, J., Scott, C., and Jermiin, L. S. (2017) Reverse engineering: transaminase biocatalyst development using ancestral sequence reconstruction. *Green. Chem.* **19**, 5375–5380
62. Babkova, P., Sebestova, E., Brezovsky, J., Chaloupkova, R., and Damborsky, J. (2017) Ancestral haloalkane dehalogenases show robustness and unique substrate specificity. *ChemBioChem* **18**, 1448–1456
63. Gumulya, Y., Baek, J.-M., Wun, S.-J., Thomson, R. E. S., Harris, K. L., Hunter, D. J. B., *et al.* (2018) Engineering highly functional thermostable proteins using ancestral sequence reconstruction. *Nat. Catal.* **1**, 878–888
64. Gumulya, Y., Huang, W., D'Cunha, S. A., Richards, K. E., Thomson, R. E. S., Hunter, D. J. B., *et al.* (2019) Engineering thermostable CYP2D enzymes for biocatalysis using combinatorial libraries of ancestors for directed evolution (CLADE). *ChemCatChem* **11**, 841–850
65. Nakano, S., Kozuka, K., Minamino, Y., Karasuda, H., Hasebe, F., and Ito, S. (2020) Ancestral L-amino acid oxidases for deracemization and stereoinversion of amino acids. *Comm. Chem.* **3**, 181
66. Park, Y., Patton, J. E. J., Hochberg, G. K. A., and Thornton, J. W. (2020) Comment on "Ancient origins of allosteric activation in a Ser-Thr kinase". *Science* **370**, eabc8301
67. Aadland, K., and Kolaczkowski, B. (2020) Alignment-integrated reconstruction of ancestral sequences improves accuracy. *Genome Biol. Evol.* **12**, 1549–1565
68. Hanson-Smith, V., Kolaczkowski, B., and Thornton, J. W. (2010) Robustness of ancestral sequence reconstruction to phylogenetic uncertainty. *Mol. Biol. Evol.* **27**, 1988–1999
69. Vialle, R. A., Tamuri, A. U., and Goldman, N. (2018) Alignment modulates ancestral sequence reconstruction accuracy. *Mol. Biol. Evol.* **35**, 1783–1797
70. [preprint] Foley, G., Mora, A., Ross, C. M., Bottoms, S., Sützl, L., Lamprecht, M. L., *et al.* (2020) Identifying and engineering ancient variants of enzymes using graphical representation of ancestral sequence predictions (GRASP). *bioRxiv*. https://doi.org/10.1101/2019.12.30.891457
71. Ross, C. M., Gabriel, F., Bodén, M., and Gillam, E. M. J. (2021) Using the evolutionary history of proteins to engineer insertion-deletion mutants from robust, ancestral templates using Graphical Representation of Ancestral Sequence Predictions (GRASP). In: Magnani, F., Marabelli, C., Paradisi, F., eds. *Enzyme Engineering: Methods and Protocols*, Humana Press, NY, NY. pp in press
72. Mendes, F. K., and Hahn, M. W. (2016) Gene tree discordance causes apparent substitution rate variation. *Syst. Biol.* **65**, 711–721
73. Abadi, S., Azouri, D., Pupko, T., and Mayrose, I. (2019) Model selection may not be a mandatory step for phylogeny reconstruction. *Nat. Commun.* **10**, 934
74. Fitch, W. M. (1971) Toward defining the course of evolution: minimum change for a specific tree topology. *Syst. Zool.* **20**, 406–416
75. Yang, Z., Kumar, S., and Nei, M. (1995) A new method of inference of ancestral nucleotide and amino acid sequences. *Genetics* **141**, 1641–1650
76. Huelsenbeck, J. P., and Bollback, J. P. (2001) Empirical and hierarchical Bayesian estimation of ancestral states. *Syst. Biol.* **50**, 351–366
77. Joy, J. B., Liang, R. H., McCloskey, R. M., Nguyen, T., and Poon, A. F. Y. (2016) Ancestral reconstruction. *PLoS Comput. Biol.* **12**, e1004763
78. Zhang, J. Z., and Nei, M. (1997) Accuracies of ancestral amino acid sequences inferred by the parsimony, likelihood, and distance methods. *J. Mol. Evol.* **44**, S139–S146
79. Li, W.-H. (1997) *Molecular Evolution*, Sinauer Associates, Sunderland, MA
80. Cunningham, C. W., Omland, K. E., and Oakley, T. H. (1998) Reconstructing ancestral character states: a critical reappraisal. *Trends Ecol. Evol.* **13**, 361–366
81. Cai, W., Pei, J. M., and Grishin, N. V. (2004) Reconstruction of ancestral protein sequences and its applications. *BMC Evol. Biol.* **4**, 3
82. Dayhoff, R. M., Schwartz, R. V., and Orcut, B. C. (1972) A model of evolutionary change in proteins. In: Dayhoff, M. O., ed. *Atlas of Protein Sequence and Structure*, National Biomedical Research Foundation, WA, DC: 89–99
83. Whelan, S., and Goldman, N. (2001) A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Mol. Biol. Evol.* **18**, 691–699

84. Le, S. Q., and Gascuel, O. (2008) An improved general amino acid replacement matrix. *Mol. Biol. Evol.* **25**, 1307–1320
85. Jones, D. T., Taylor, W. R., and Thornton, J. M. (1992) The rapid generation of mutation data matrices from protein sequences. *Comput. Appl. Biosci.* **8**, 275–282
86. Arenas, M. (2015) Trends in substitution models of molecular evolution. *Front. Genet.* **6**, 319
87. Pupko, T., Pe'er, I., Shamir, R., and Graur, D. (2000) A fast algorithm for joint reconstruction of ancestral amino acid sequences. *Mol. Biol. Evol.* **17**, 890–896
88. 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
89. Groussin, M., Hobbs, J. K., Szollosi, G. J., Gribaldo, S., Arcus, V. L., and Gouy, M. (2015) Toward more accurate ancestral protein genotype-phenotype reconstructions with the use of species tree-aware gene trees. *Mol. Biol. Evol.* **32**, 13–22
90. Akanuma, S., Yokobori, S. I., Nakajima, Y., Bessho, M., and Yamagishi, A. (2015) Robustness of predictions of extremely thermally stable proteins in ancient organisms. *Evolution* **69**, 2954–2962
91. Oliva, A., Pulicani, S., Lefort, V., Bréhélin, L., Gascuel, O., and Guindon, S. (2019) Accounting for ambiguity in ancestral sequence reconstruction. *Bioinformatics* **35**, 4290–4297
92. Bar-Rogovsky, H., Stern, A., Penn, O., Kobl, I., Pupko, T., and Tawfik, D. S. (2015) Assessing the prediction fidelity of ancestral reconstruction by a library approach. *Protein Eng. Des. Sel.* **28**, 507–518
93. Clifton, B. E., Kaczmarski, J. A., Carr, P. D., Gerth, M. L., Tokuriki, N., and Jackson, C. J. (2018) Evolution of cyclohexadienyl dehydratase from an ancestral solute-binding protein. *Nat. Chem. Biol.* **14**, 542–547
94. Eick, G. N., Bridgham, J. T., Anderson, D. P., Harms, M. J., and Thornton, J. W. (2016) Robustness of reconstructed ancestral protein functions to statistical uncertainty. *Mol. Biol. Evol.* **34**, 247–261
95. Kuchner, O., and Arnold, F. H. (1997) Directed evolution of enzyme catalysts. *Trends Biotech.* **15**, 523–530
96. Giver, L., Gershenson, A., Freskgard, P.-O., and Arnold, F. H. (1998) Directed evolution of a thermostable esterase. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 12809–12813
97. Trudeau, D. L., Kaltenbach, M., and Tawfik, D. S. (2016) On the potential origins of the high stability of reconstructed ancestral proteins. *Mol. Biol. Evol.* **33**, 2633–2641
98. Akanuma, S., Iwami, S., Yokoi, T., Nakamura, N., Watanabe, H., Yokobori, S.-i., *et al.* (2011) Phylogeny-based design of a B-subunit of DNA gyrase and its ATPase domain using a small set of homologous amino acid sequences. *J. Mol. Biol.* **412**, 212–225
99. Risso, V. A., Gavira, J. A., Gaucher, E. A., and Sanchez-Ruiz, J. M. (2014) Phenotypic comparisons of consensus variants *versus* laboratory resurrections of Precambrian proteins. *Proteins Struct. Funct. Bioinf.* **82**, 887–896
100. Nakano, S., Motoyama, T., Miyashita, Y., Ishizuka, Y., Matsuo, N., Tokiwa, H., *et al.* (2018) Benchmark analysis of native and artificial NAD^+-dependent enzymes generated by a sequence-based design method with or without phylogenetic data. *Biochemistry* **57**, 3722–3732
101. Hendrikse, N. M., Charpentier, G., Nordling, E., and Syrén, P. O. (2018) Ancestral diterpene cyclases show increased thermostability and substrate acceptance. *FEBS J.* **285**, 4660–4673
102. Kiss, C., Temirov, J., Chasteen, L., Waldo, G. S., and Bradbury, A. R. M. (2009) Directed evolution of an extremely stable fluorescent protein. *Protein Eng. Des. Sel.* **22**, 313–323
103. Hall, B. G. (2006) Simple and accurate estimation of ancestral protein sequences. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 5431–5436
104. Jaenicke, R. (1999) Stability and folding of domain proteins. *Prog. Biophys. Mol. Biol.* **71**, 155–241
105. Ingles-Prieto, A., Ibarra-Molero, B., Delgado-Delgado, A., Perez-Jimenez, R., Fernandez, J. M., Gaucher, E. A., *et al.* (2013) Conservation of protein structure over four billion years. *Structure* **21**, 1690–1697
106. Okafor, C. D., Pathak, M. C., Fagan, C. E., Bauer, N. C., Cole, M. F., Gaucher, E. A., *et al.* (2018) Structural and dynamics comparison of thermostability in ancient, modern, and consensus elongation factor Tus. *Structure* **26**, 118–129.e3
107. Gromiha, M. M., Pathak, M. C., Saraboji, K., Ortlund, E. A., and Gaucher, E. A. (2013) Hydrophobic environment is a key factor for the stability of thermophilic proteins. *Proteins: Struct. Funct. Bioinf.* **81**, 715–721
108. Bart, A. G., Harris, K. L., Gillam, E. M. J., and Scott, E. E. (2020) Structure of an ancestral mammalian family 1B1 cytochrome P450 with increased thermostability. *J. Biol. Chem.* **295**, 5640–5653
109. Hartz, P., Strohmaier, S. J., EL-Gayar, B. M., Abdulmughni, A., Hutter, M. C., Hannemann, F., *et al.* (2021) Resurrection and characterization of ancestral CYP11A1 enzymes. *FEBS J.* **288**, 6510–6527
110. Harris, K. L., Thomson, R. E. S., Gumulya, Y., Foley, G., Carrera-Pacheco, S. E., Syed, P., *et al.* (2022) Ancestral sequence reconstruction of a cytochrome P450 family involved in chemical defence reveals the functional evolution of a promiscuous, xenobiotic-metabolizing enzyme in vertebrates. *Mol. Biol. Evol.* **39**, msac116
111. Khersonsky, O., and Fleishman, S. J. (2016) Why reinvent the wheel? Building new proteins based on ready-made parts. *Protein Sci.* **25**, 1179–1187
112. Spence, M. A., Kaczmarski, J. A., Saunders, J. W., and Jackson, C. J. (2021) Ancestral sequence reconstruction for protein engineers. *Curr. Opin. Struct. Biol.* **69**, 131–141
113. Mazurenko, S., Prokop, Z., and Damborsky, J. (2020) Machine learning in enzyme engineering. *ACS Catal.* **10**, 1210–1223
114. Avadhani, V. S., Mondal, S., and Banerjee, S. (2022) Mapping protein structural evolution upon unfolding. *Biochemistry* **61**, 303–309
115. Furukawa, R., Toma, W., Yamazaki, K., and Akanuma, S. (2020) Ancestral sequence reconstruction produces thermally stable enzymes with mesophilic enzyme-like catalytic properties. *Sci. Rep.* **10**, e15493
116. Barruetabeña, N., Alonso-Lerma, B., Galera-Prat, A., Joudeh, N., Barandiaran, L., Aldazabal, L., *et al.* (2019) Resurrection of efficient Precambrian endoglucanases for lignocellulosic biomass hydrolysis. *Comm. Chem.* **2**, 1–13
117. Gomez-Fernandez, B. J., Garcia-Ruiz, E., Martin-Diaz, J., Gomez de Santos, P., Santos-Moriano, P., Plou, F. J., *et al.* (2018) Directed *-in vitro-* evolution of Precambrian and extant Rubiscos. *Sci. Rep.* **8**, 5532
118. Rozi, M. F. A. M., Rahman, R. N. Z. R. A., Leow, A. T. C., and Ali, M. S. M. (2022) Ancestral sequence reconstruction of ancient lipase from family I.3 bacterial lipolytic enzymes. *Mol. Phylogenet. Evol.* **168**, 107381
119. Harada, M., Nagano, A., Yagi, S., Furukawa, R., Yokobori, S.-i., and Yamagishi, A. (2019) Planktonic adaptive evolution to the sea surface temperature in the Neoproterozoic inferred from ancestral NDK of marine cyanobacteria. *Earth Planet. Sci. Lett.* **522**, 98–106
120. Loughran, N. B., O'Connell, M. J., O'Connor, B., and Ó'Fágáin, C. (2014) Stability properties of an ancient plant peroxidase. *Biochimie* **104**, 156–159
121. Devamani, T., Rauwerdink, A. M., Lunzer, M., Jones, B. J., Mooney, J. L., Tan, M. A. O., *et al.* (2016) Catalytic promiscuity of ancestral esterases and hydroxynitrile lyases. *J. Am. Chem. Soc.* **138**, 1046–1056
122. Mascotti, M. L., Kumar, H., Nguyen, Q.-T., Ayub, M. J., and Fraaije, M. W. (2018) Reconstructing the evolutionary history of F420-dependent dehydrogenases. *Sci. Rep.* **8**, e17571
123. Whitfield, J. H., Zhang, W. H., Herde, M. K., Clifton, B. E., Radziejewski, J., Janovjak, H., *et al.* (2015) Construction of a robust and sensitive arginine biosensor through ancestral protein reconstruction. *Protein Sci.* **24**, 1412–1422
124. Chaloupkova, R., Liskova, V., Toul, M., Markova, K., Sebestova, E., Hernychova, L., *et al.* (2019) Light-emitting dehalogenases: reconstruction of multifunctional biocatalysts. *ACS Catal.* **9**, 4810–4823
125. Ayuso-Fernández, I., Martínez, A. T., and Ruiz-Dueñas, F. J. (2017) Experimental recreation of the evolution of lignin-degrading enzymes from the Jurassic to date. *Biotechnol. Biofuels* **10**, 67
126. Kaltenbach, M., Burke, J. R., Dindo, M., Pabis, A., Munsberg, F. S., Rabin, A., *et al.* (2018) Evolution of chalcone isomerase from a non-catalytic ancestor. *Nat. Chem. Biol.* **14**, 548–555
127. Schulte-Sasse, M., Pardo-Ávila, F., Pulido-Mayoral, N. O., Vázquez-Lobo, A., Costas, M., García-Hernández, E., *et al.* (2019) Structural,

thermodynamic and catalytic characterization of an ancestral triosephosphate isomerase reveal early evolutionary coupling between monomer association and function. *FEBS J.* **286**, 882–900

128. Nakano, S., Niwa, M., Asano, Y., and Ito, S. (2019) Following the evolutionary track of a highly specific L-arginine oxidase by reconstruction and biochemical analysis of ancestral and native enzymes. *Appl. Environ. Microbiol.* **85**, e00459–e00478
129. Nakano, S., Minamino, Y., Hasebe, F., and Ito, S. (2019) Deracemization and stereoinversion to aromatic d-amino acid derivatives with ancestral l-amino acid oxidase. *ACS Catal.* **9**, 10152–10158
130. Sun, Y., Calderini, E., and Kourist, R. (2021) A reconstructed common ancestor of the fatty acid photo-decarboxylase clade shows photo-decarboxylation activity and increased thermostability. *ChemBioChem* **22**, 1833–1840
131. Kropp, C., Straub, K., Linde, M., and Babinger, P. (2021) Hexamerization and thermostability emerged very early during geranylgeranylglyceryl phosphate synthase evolution. *Protein Sci.* **30**, 583–596
132. Harris, Kurt L. (2020) Ancestral reconstruction of cytochrome P450 family 1, 4 and cytochrome P450 reductase: insights into evolution and applications in biocatalysis *Ph.D. Thesis*. The University of Queensland, Brisbane, Australia: 54–56