

MIT Open Access Articles

Reconstructed Ancestral Enzymes Impose a Fitness Cost upon Modern Bacteria Despite Exhibiting Favourable Biochemical Properties

The MIT Faculty has made this article openly available. **Please share** how this access benefits you. Your story matters.

Citation: Hobbs, Joanne K. et al. "Reconstructed Ancestral Enzymes Impose a Fitness Cost upon Modern Bacteria Despite Exhibiting Favourable Biochemical Properties." *Journal of Molecular Evolution* 81.3-4 (2015): 110-120.

As Published: <http://dx.doi.org/10.1007/s00239-015-9697-5>

Publisher: Springer US

Persistent URL: <http://hdl.handle.net/1721.1/105120>

Version: Author's final manuscript: final author's manuscript post peer review, without publisher's formatting or copy editing

Terms of Use: Article is made available in accordance with the publisher's policy and may be subject to US copyright law. Please refer to the publisher's site for terms of use.



Reconstructed Ancestral Enzymes Impose a Fitness Cost upon Modern Bacteria Despite Exhibiting Favourable Biochemical Properties

Joanne K. Hobbs^{1,2} · Erica J. Prentice¹ · Mathieu Groussin^{3,4} · Vickery L. Arcus¹

Received: 26 June 2015 / Accepted: 3 September 2015 / Published online: 9 September 2015
© Springer Science+Business Media New York 2015

Abstract Ancestral sequence reconstruction has been widely used to study historical enzyme evolution, both from biochemical and cellular perspectives. Two properties of reconstructed ancestral proteins/enzymes are commonly reported—high thermostability and high catalytic activity—compared with their contemporaries. Increased protein stability is associated with lower aggregation rates, higher soluble protein abundance and a greater capacity to evolve, and therefore, these proteins could be considered “superior” to their contemporary counterparts. In this study, we investigate the relationship between the favourable in vitro biochemical properties of reconstructed ancestral enzymes and the organismal fitness they confer in vivo. We have previously reconstructed several ancestors of the enzyme LeuB, which is essential for leucine

biosynthesis. Our initial fitness experiments revealed that overexpression of ANC4, a reconstructed LeuB that exhibits high stability and activity, was only able to partially rescue the growth of a $\Delta leuB$ strain, and that a strain complemented with this enzyme was outcompeted by strains carrying one of its descendants. When we expanded our study to include five reconstructed LeuBs and one contemporary, we found that neither in vitro protein stability nor the catalytic rate was correlated with fitness. Instead, fitness showed a strong, negative correlation with estimated evolutionary age (based on phylogenetic relationships). Our findings suggest that, for reconstructed ancestral enzymes, superior in vitro properties do not translate into organismal fitness in vivo. The molecular basis of the relationship between fitness and the inferred age of ancestral LeuB enzymes is unknown, but may be related to the reconstruction process. We also hypothesise that the ancestral enzymes may be incompatible with the other, contemporary enzymes of the metabolic network.

Electronic supplementary material The online version of this article (doi:10.1007/s00239-015-9697-5) contains supplementary material, which is available to authorized users.

✉ Joanne K. Hobbs
jhobbs@uvic.ca

✉ Vickery L. Arcus
varcus@waikato.ac.nz

¹ School of Science, Faculty of Science and Engineering, University of Waikato, Private Bag 3105, Hamilton 3240, New Zealand

² Present Address: Department of Biochemistry and Microbiology, University of Victoria, PO Box 3055, STN CSC, Victoria, BC V8W 3P6, Canada

³ Laboratoire de Biometrie et Biologie Evolutive, Université de Lyon, Université Lyon 1, CNRS, UMR5558, Villeurbanne, France

⁴ Present Address: Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA

Keywords Ancestral sequence reconstruction · Protein resurrection · Enzyme evolution · Fitness · Stability · 3-Isopropylmalate dehydrogenase

Introduction

The evolution of protein structure, function and stability are intrinsically linked to organismal fitness, yet protein evolution is often not considered in the context of whole organisms (DePristo et al. 2005). Therefore, in recent years, there has been a drive to integrate the fields of protein biophysics and molecular evolution in order to better understand protein evolution (Harms and Thornton 2013; Liberles et al. 2012; Serohijos and Shakhnovich

2014; Soskine and Tawfik 2010). One technique that has proved particularly useful in this pursuit is ancestral sequence reconstruction (ASR). ASR uses a combination of phylogenetics, evolutionary theory, synthetic biology and protein biochemistry to infer the sequences of ancestral proteins and then characterise them in the laboratory. It has provided otherwise unobtainable insight into many evolutionary questions, such as ligand specificity in steroid hormone receptors (Bridgham et al. 2006, 2009; Eick et al. 2012), spectral tuning in visual pigments (Chang et al. 2002; Chinen et al. 2005; Yokoyama et al. 2008) and substrate specificity amongst fungal α -glucosidases (Voordeckers et al. 2012).

Since the first ancestral reconstructions by Malcolm et al. (1990) and Stackhouse et al. (1990), there have now been more than 40 published ASR studies and two common properties exhibited by inferred ancestral proteins/enzymes have emerged: (i) high stability—both thermostability (usually measured as the temperature midpoint of thermal denaturation, T_m , or optimum temperature for activity, T_{opt}) and kinetic stability (measured as the free energy for unfolding, ΔG_{N-U}^\ddagger)—(Akanuma et al. 2013; Butzin et al. 2013; Gaucher et al. 2003, 2008; Groussin et al. 2015; Hobbs et al. 2012; Miyazaki et al. 2001; Perez-Jimenez et al. 2011; Risso et al. 2013; Watanabe et al. 2006a; Watanabe and Yamagishi 2006b) k_{cat} and (ii) high catalytic activity and/or catalytic efficiency (usually expressed as k_{cat} and k_{cat}/K_M , respectively) (Stackhouse et al. 1990; Akanuma et al. 2013; Butzin et al. 2013; Groussin et al. 2015; Hobbs et al. 2012; Perez-Jimenez et al. 2011; Risso et al. 2013; Watanabe and Yamagishi 2006b; Jermann et al. 1995). These increases in stability and catalytic activity compared with contemporary proteins/enzymes can be dramatic, for example Risso et al. (2013) reported that their inferred ancestral β -lactamases were more stable than their contemporary descendants by as much as 35 °C, and Perez-Jimenez et al. (2011) found that their reconstructed ancestral thioredoxins displayed catalytic rate constants up to 30-fold higher than modern thioredoxins at pH 5, as well as being up to 32 °C more stable. We have previously used ASR to reconstruct presumed ancestral sequences of the core metabolic enzyme LeuB (3-isopropylmalate dehydrogenase, IPMDH, EC 1.1.1.85) from the *Bacillus* genus (Hobbs et al. 2012) and the Firmicutes phylum (Groussin et al. 2015), and also observed enhancements in stability and activity. In our *Bacillus* study, three out of the four reconstructed LeuBs were thermophilic (represented by high T_{opt} and T_m values) and exhibited k_{cat} values ≥ 2 -fold higher than that exhibited by the contemporary *B. subtilis* LeuB enzyme (Hobbs et al. 2012). The reconstructed enzyme ANC4 (which sits at the base of the *Bacillus* tree; Fig. 1), in particular, was biochemically impressive in that it exhibited not only a >6 -fold

increase in k_{cat} and >4 -fold increase in k_{cat}/K_M compared with LeuB from the contemporary thermophile *B. caldovelox* (BCVX). In addition, it exhibited exceptionally high kinetic stability (i.e. very slow to unfold in the presence of a denaturant). More recently, we reconstructed two variants of LeuB from the last common ancestor of the Firmicutes phylum (Fig. 1; Groussin et al. 2015). The two variants were inferred using either a species tree-aware gene phylogeny (LeuB_{S-aw}) or a species tree-unaware gene phylogeny (LeuB_{S-unaw}). Both of these enzymes were highly thermophilic (T_{opt} values ≥ 9 °C higher than that of the BCVX enzyme), exhibited k_{cat} values ≥ 3 -fold higher than the BCVX enzyme and one of the enzymes possessed a ΔG_{N-U}^\ddagger value equal to that of ANC4 (Groussin et al. 2015).

The apparent biochemical superiority of the inferred ancestral enzymes over their contemporary counterparts begs the question: if these enzymes do indeed approximate the ancestral state and are so “good” biochemically, why didn’t their seemingly favourable biochemical properties prevail during evolution? Many studies suggest that high protein stability should be evolutionarily advantageous. Firstly, a significant proportion of cellular ATP is used in protein synthesis therefore long protein half-lives should conserve energy (Cox and Cook 2007). Secondly, it is known that less stable proteins populate unfolded states more frequently, and therefore have higher aggregation rates, than more stable proteins (DePristo et al. 2005). For example, Bershtein et al. (2012) found that the T_m values of dihydrofolate reductase mutants were positively correlated with soluble protein abundance at 30 °C. In turn, soluble protein abundance was positively correlated with organism fitness. Tomala et al. (2014) also showed that overexpression of destabilised mutants of two yeast proteins was associated with decreased fitness. Thirdly, increased protein stability has been associated with mutational robustness, or a greater capacity to sample sequence space in the course of evolution (Bloom et al. 2005, 2006). For example, both Bloom et al. (2006) and Studer et al. (2014) observed that only previously stabilised protein variants could tolerate mutations that confer enhanced or novel function. However, there are potential disadvantages to increased stability. Protein degradation rates are largely determined by a protein’s stability (Parsell and Sauer 1989) and the increased resistance to proteases exhibited by stable proteins can make them difficult to regulate (DePristo et al. 2005). In addition, increased protein stability is often accompanied by a trade-off in terms of catalytic activity (Somero 1995), although this appears not to be the case for many inferred ancestral enzymes. From an evolutionary perspective, assuming that the increased thermostability of ancestral proteins is not the result of a bias in the in silico reconstruction method (Williams et al.

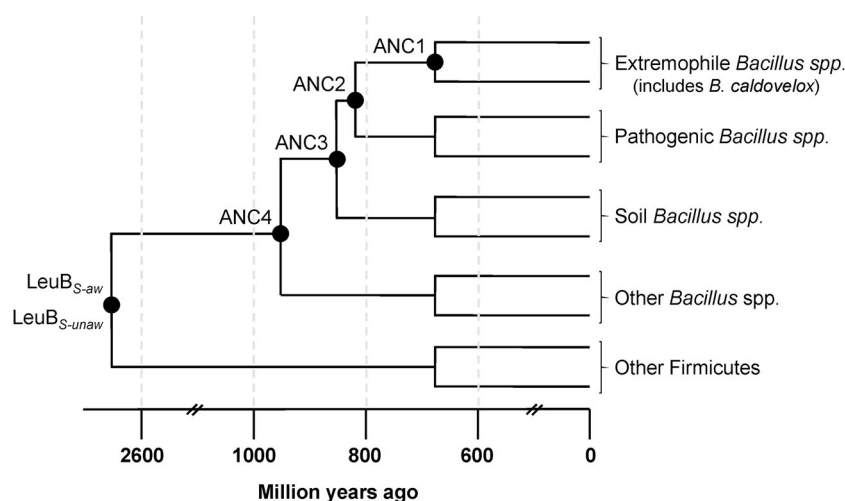


Fig. 1 Schematic representation of the phylogenetic relationships between previously characterised contemporary and ancestral LeuB enzymes. The schematic cladogram shown is a combination of the *Bacillus* phylogeny from (Hobbs et al. 2012) and the Firmicutes

phylogeny from (Groussin et al. 2015). The evolutionary time scale was determined previously (Hobbs et al. 2012) and is based on calibration points taken from Battistuzzi et al. (2004). Pathogenic and soil *Bacillus* clade names are taken from Alcaraz et al. (2010)

2006; see “Result and Discussion” section), the marginal stability of contemporary proteins may imply that an unknown selection pressure has caused the stability of proteins to decrease during evolution, or simply that stability is a neutral trait (Bloom et al. 2007; Taverna and Goldstein 2002; Zeldovich et al. 2007). For example, as the majority of mutations are destabilising rather than stabilising, it has been argued that proteins only retain high stability if it is evolutionarily beneficial (Taverna and Goldstein 2002). Whilst increased thermostability and activity appear to be common features of reconstructed ancestral proteins/enzymes, this is not always the case. For example, Hart et al. (2014) recently reconstructed ancestors of ribonuclease H1 (RNH) from two lineages of bacteria, one mesophilic and one thermophilic, and found that the ancestors exhibited intermediate thermostabilities between those of contemporary RNHs from *Escherichia coli* and *Thermus thermophilus*. Furthermore, in our previous study of LeuBs from the *Bacillus* genus (Hobbs et al. 2012), one of our inferred ancestral LeuBs (ANC2) exhibited a psychrophilic/mesophilic T_{opt} , low $\Delta G_{N-U}^{\ddagger}$ and only moderate k_{cat} .

Reconstructed ancestral proteins provide a unique opportunity to study the biochemical, biophysical and organismal aspects of protein evolution, and the relationships between them. To date, only a small number of ASR studies have included in vivo evaluation of inferred ancestral proteins and these have related to functionality rather than fitness (e.g. transcriptional response to different ligands [Bridgham et al. 2006, 2009, 2010; Eick et al. 2012], ability to rescue different phenotypes (Finnigan et al. 2012), ability to confer resistance to β -lactams (Risso et al. 2013)]. In this

study, we utilise our previously reconstructed ancestral LeuB enzymes to investigate the organismal fitness conferred by these inferred ancestral enzymes in relation to their favourable biochemical properties.

Materials and Methods

Reconstruction of Ancestral LeuBs

The ancestral inference and reconstruction of ANC1, ANC2, ANC3 and ANC4 (Hobbs et al. 2012) and LeuB_{S-aw} and LeuB_{S-unaw} (Groussin et al. 2015) have been reported previously. In brief, the sequences of ANC1, ANC2, ANC3 and ANC4 were inferred under the maximum likelihood criterion using a phylogeny of *Bacillus* species based on LeuB protein sequences and the LG substitution model implemented in PAML (Yang 2007). The sequences of LeuB_{S-aw} and LeuB_{S-unaw} were inferred using the site-heterogeneous EX_EHO mixture model implemented in bppAncestor (Dutheil and Boussau 2008) with either a species tree-aware phylogeny (LeuB_{S-aw}), which included both LeuB and ribosomal sequence information, or a species tree-unaware phylogeny (LeuB_{S-unaw}) based on LeuB sequences alone.

Bacterial Strains, Plasmids and Culture Conditions

The *E. coli* $\Delta leuB$ strain and its parent K12 BW25113 were obtained from the Keio collection (Baba et al. 2006). Genes encoding for LeuB from *B. caldovelox* and the ancestral LeuBs were codon optimised for expression in *E. coli*,

chemically synthesised by Genart (Life Technologies) and cloned into pPROEX HTb as reported previously (Groussin et al. 2015; Hobbs et al. 2012). All strains were routinely cultured in Luria–Bertani (LB) broth containing ampicillin and/or kanamycin at 100 or 50 µg/ml, respectively. When required, IPTG was added to a final concentration of 1 mM. For additional k_{cat} determinations, LeuB proteins were expressed, purified and assayed as previously described (Hobbs et al. 2012).

Cloning of *leuB* Under Control of *leu* Promoter

Genes encoding for contemporary and ancestral LeuBs were cloned into pUC19 under the control of the native *E. coli* *leu* operon promoter using overlap extension PCR. The *leu* operon promoter sequence (including RBS) as identified by Haughn et al. (1986) was amplified from *E. coli* K12 BW25113 genomic DNA using the primers *leu* promoter fwd and *leu* promoter rev. The *leuB* genes were amplified from the pPROEX HTb expression constructs using forward primers that contained a 14 bp overlap with the 3' end of the *leu* promoter PCR fragment. The promoter and *leuB* fragments were then joined together by PCR using the primers *leu* promoter fwd and pPROEX *leuB* rev. Following gel extraction and purification, the promoter-*leuB* inserts were ligated into pUC19 between the *Eco*RI and *Hind*III sites and transformed into *E. coli* DH5 α . Constructs were confirmed by sequencing, transformed into the Δ *leuB* strain and plated on LB agar. All primer sequences can be found in Online Resource 1.

Rescue Experiments

Control and test strains were tested and scored for their ability to grow on minimal agar at decreasing cell densities as previously described (Finnigan et al. 2012). Following overnight incubation in LB broth, cells were pelleted by centrifugation and washed twice with M9 broth before being diluted and spotted (20 µl) onto M9 minimal agar containing 10 g/l glucose as the sole carbon source. For Δ *leuB* (pPROEX) strains, 1 mM IPTG and ampicillin were also included in the agar. Plates were incubated at 37 °C and growth judged after 24 h. All rescue experiments were performed in duplicate.

Growth Rate Determinations

Growth rate determinations were performed in Erlenmeyer flasks in a shaking incubator at 37 °C and 200 rpm in triplicate. For growth curves in rich medium, flasks were inoculated 1/100 with overnight starter cultures grown in LB broth and OD_{600nm} readings taken every 15 min for

~4 h using a ThermoSpectronic Helios spectrophotometer. For growth curves in M9 minimal medium, LB starter cultures were used to inoculate M9 starter cultures which, after overnight incubation, were then used to inoculate growth curve flasks. OD_{600nm} readings of minimal medium cultures were taken every ~45 min for 13 h. Growth rate constants during exponential growth were calculated using the programme GrowthRates (Hall et al. 2014). The growth of Δ *leuB* (pUC19-ANC4) in minimal medium was also monitored over a 24-h period in a FLUOStar Optima microplate reader (BMG Labtech). M9 starter cultures were used to inoculate (1/50) 200 µl minimal medium in a round-bottomed 96-well plate in triplicate. The plate was loosely sealed with clingfilm and incubated at 37 °C, with OD_{600nm} readings taken every 30 min preceded by 10 s of shaking. Statistical analysis of growth rate data was performed in GraphPad Prism 6. Calculation of the Spearman rank-order correlation coefficient and correction of *P* values for multiple testing was performed using R version 3.1.0.

Competition Assays

The relative fitness of Δ *leuB* (pUC19-ANC4) was compared with that of Δ *leuB* (pUC19-BCVX) and Δ *leuB* (pUC19-ANC1) using a standard 24-h pairwise competition assay (Lenski et al. 1991). Two strains were used to inoculate a single 5 ml aliquot of M9 minimal medium and incubated at 37 °C with shaking at 200 rpm for 24 h. Following incubation, appropriate dilutions of the culture were plated onto LB agar containing kanamycin and ampicillin and the total number of cells counted. As there was no morphological way to distinguish between colonies of the two strains, a diagnostic restriction digest was employed following PCR amplification of *leuB*. For each pairing, 16 colonies were picked and used as the template in a PCR with the primers, *leuB* universal fwd and *leuB* universal rev (Online Resource 1). Following confirmation of successful *leuB* gene amplification (1116 bp for BCVX, 1110 bp for ANC1 and ANC4), the remaining PCR products were digested with *Nsp*I at 37 °C for 2 h, and the banding patterns were visualised by gel electrophoresis (the ANC4 gene contains no *Nsp*I cut sites, whereas BCVX and ANC1 both contain a single cut site resulting in two fragments of 768/762 and 348 bp). Known colonies of BCVX, ANC1 and ANC4 were used as controls, and no clones were sequenced. The proportion of colonies of each strain as determined by the diagnostic restriction digest was applied to the total number of colonies, and the relative fitness of Δ *leuB* (pUC19-ANC4) was calculated as previously described (Lenski et al. 1991). All pairings were performed in triplicate.

Table 1 In vitro biochemical properties of contemporary and ancestral LeuBs

Enzyme	$K_M^{(IPM)}$ (mM)	$K_M^{(NAD)}$ (mM)	K_{CAT} (S^{-1})	T_{opt} ($^{\circ}C$)	$\Delta G_{N-U}^{\ddagger}$ ($kJ\ mol^{-1}$)
BCVX ^a	1.1	0.8	53.8	69	100.7
ANC1 ^a	1.3	0.5	141.8	73	100.9
ANC2 ^a	1.0	0.9	41.7	49	91.1
ANC3 ^a	2.7	1.0	102.3	60	95.6
ANC4 ^a	1.7	1.0	362.2	70	110.8
LeuB _{S-aw} ^b	1.6	6.5	181.2	85	110.9
LeuB _{S-unaw} ^b	6.8	5.5	441.2	78	91.4

K_M is a measure of the affinity of an enzyme for its substrate; a lower value indicates higher affinity. k_{cat} indicates the turnover number, or catalytic rate, of an enzyme and varies with temperature. T_{opt} is the temperature at which an enzyme exhibits its highest k_{cat} . $\Delta G_{N-U}^{\ddagger}$ indicates the free energy of unfolding for an enzyme and is a measure of kinetic stability; an enzyme with a high $\Delta G_{N-U}^{\ddagger}$ value is slow to unfold. K_M and k_{cat} values shown were determined at the respective T_{opt} of each enzyme

^a Data were taken from Hobbs et al. (2012)

^b Data were taken from Groussin et al. (2015)

Results and Discussion

Overexpression of ANC4 Only Partially Rescues $\Delta leuB$ Phenotype

In our previous reconstruction of LeuB enzymes from the *Bacillus* genus, we identified ANC4 as an exceptional enzyme in terms of its high catalytic rate and kinetic stability (Hobbs et al. 2012). Two of its descendants, BCVX and ANC1 (Fig. 1), are similar to ANC4 in terms of thermoactivity, but less kinetically stable and exhibit lower k_{cat} values (Table 1). Therefore, we decided to compare the fitness costs conferred by these three enzymes in vivo. LeuB catalyses the third step in the leucine biosynthesis pathway, the oxidative decarboxylation of 3-isopropylmalate (IPM) to 2-isopropyl-3-oxosuccinate (with reduction of NAD^+). In the absence of an external source of leucine, *leuB* is an essential gene, and its deletion results in leucine auxotrophy. In an initial rescue experiment, we used our IPTG-inducible protein overexpression constructs for BCVX, ANC1 and ANC4 to transform a $\Delta leuB$ strain of *E. coli* and then tested these strains for their ability to grow on minimal medium at the decreasing cell densities (Fig. 2). The auxotrophy of the $\Delta leuB$ strain was confirmed by its inability to grow on minimal medium. The $\Delta leuB$ strains expressing BCVX and ANC1 both grew well on minimal medium and at all cell densities, whereas the strain expressing ANC4 grew poorly even at the highest densities. This implies that expression of ANC4 is associated with a fitness cost. We know from our previous work that all three enzymes are expressed solubly in *E. coli* at 37 °C and in an active form (Hobbs et al. 2012). Furthermore, all three *leuB* genes were codon optimised for expression in *E. coli* during gene synthesis (Hobbs et al. 2012). However, we have not previously tested ANC4 for

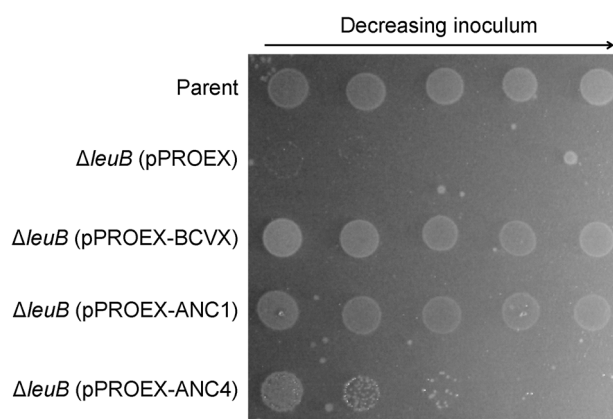


Fig. 2 Rescue of *E. coli* $\Delta leuB$ by overexpression of contemporary and ancestral LeuBs. Five strains were spotted onto M9 minimal agar containing 1 mM IPTG at different culture dilutions: Keio collection parental strain (carrying pPROEX HTb to enable it to grow on ampicillin), $\Delta leuB$ carrying pPROEX HTb empty vector, $\Delta leuB$ expressing LeuB from *B. caldovelox* (BCVX), $\Delta leuB$ expressing the ancestral LeuB ANC1 and $\Delta leuB$ expressing the ancestral LeuB ANC4. Growth was judged following overnight incubation at 37 °C

activity at a temperature as low as 37 °C, and we postulated that the cause of the apparent fitness cost may be weak activity at 37 °C as a result of its increased stability. To test this hypothesis, we purified the three enzymes and determined their k_{cat} values at 37 °C. The trend in k_{cat} between the three enzymes at 37 °C was the same as at T_{opt} , i.e. ANC4 exhibited the highest k_{cat} ($7.9\ s^{-1}$), with ANC1 being the next highest ($3.4\ s^{-1}$) and BCVX the lowest ($1.8\ s^{-1}$). In addition, all three enzymes exhibited higher affinities for IPM at 37 °C than at T_{opt} (0.12, 0.34 and 0.37 mM for BCVX, ANC1 and ANC4, respectively). Therefore, ANC4 exhibits sufficient activity at 37 °C to theoretically rescue the $\Delta leuB$ strain.

ANC4 is Evolutionarily Less Fit than BCVX and ANC1 in the Absence of Leucine

In our initial $\Delta leuB$ rescue experiment (above), we made use of the protein overexpression constructs we had previously generated (expression of LeuB under the control of the strong IPTG-inducible *trc* promoter) (Hobbs et al. 2012). However, in order to quantitatively compare the fitness effects conferred by BCVX, ANC1 and ANC4, we needed to remove the potentially confounding factors of overexpression and artificial induction. In *E. coli*, *leuB* is located within the *leuABCD* operon, which is controlled by a single promoter upstream of *leuA* (Haughn et al. 1986). Therefore, we cloned the genes encoding BCVX, ANC1 and ANC4 into pUC19 under the control of the native *leu* operon promoter and then transformed the $\Delta leuB$ strain with these constructs (we initially attempted to use allelic replacement to place our LeuBs on the chromosome of the thermophile *Bacillus stearothermophilus*; however, our attempts were unsuccessful due to the low transformation efficiency of this bacterium). The growth of these strains in minimal medium was followed by regular OD_{600nm} measurements over a ~13-h period (Fig. 3a). Where possible, the data were then used to calculate the growth rate constant. The fitness cost associated with expression of ANC4 in the previous rescue experiment was confirmed as the $\Delta leuB$ (pUC19-ANC4) strain failed to grow detectably over the course of the experiment. The strains complemented with BCVX and ANC1 both grew well and reached similar cell densities. Whilst $\Delta leuB$ (pUC19-BCVX) exhibited a slightly higher growth rate constant than $\Delta leuB$ (pUC19-ANC1), it was not statistically significant (unpaired *t* test, $P = 0.9469$). We also determined the growth rates of the three strains in rich medium to determine

whether simply the presence of the gene encoding for ANC4 conferred a fitness cost. In LB broth, the growth rate constant of $\Delta leuB$ (pUC19-ANC4) did not differ significantly from that of $\Delta leuB$ (pUC19-BCVX) or $\Delta leuB$ (pUC19-ANC1) (one-way ANOVA with Dunnett’s multiple comparisons test, $P > 0.05$).

Despite the lack of detectable growth of $\Delta leuB$ (pUC19-ANC4) in minimal medium over 13 h, ANC4 does not appear to be completely inactive in this system as $\Delta leuB$ (pUC19-ANC4) showed some growth on minimal agar following overnight incubation and monitoring of OD_{600nm} over 24 h in a plate reader indicated slow, although inconsistent, growth (increase of ~ 0.1 OD_{600nm} in 24 h). As we were unable to use the growth rate of $\Delta leuB$ (pUC19-ANC4) to compare its fitness with $\Delta leuB$ (pUC19-BCVX) and $\Delta leuB$ (pUC19-ANC1), we determined the relative fitness (*W*) of $\Delta leuB$ (pUC19-ANC4) using pairwise competition assays. Following pairings in triplicate, the relative fitness of $\Delta leuB$ (pUC19-ANC4) were found to be 0.18 ± 0.18 and 0.36 ± 0.18 (\pm SEM) compared with $\Delta leuB$ (pUC19-BCVX) and $\Delta leuB$ (pUC19-ANC1), respectively. This highlights the evolutionary disadvantage that ANC4 confers upon a contemporary organism in the absence of leucine.

In Vivo Fitness is not Correlated with Stability or Catalytic Rate

Despite its high catalytic activity and stability, our growth rate and competition assay data indicate that ANC4 is less fit than BCVX and ANC1 in the context of a modern microorganism. Therefore, we hypothesised that these seemingly “improved” biochemical and biophysical properties may actually be the cause of ANC4’s fitness cost. To test this hypothesis, we expanded our growth rate

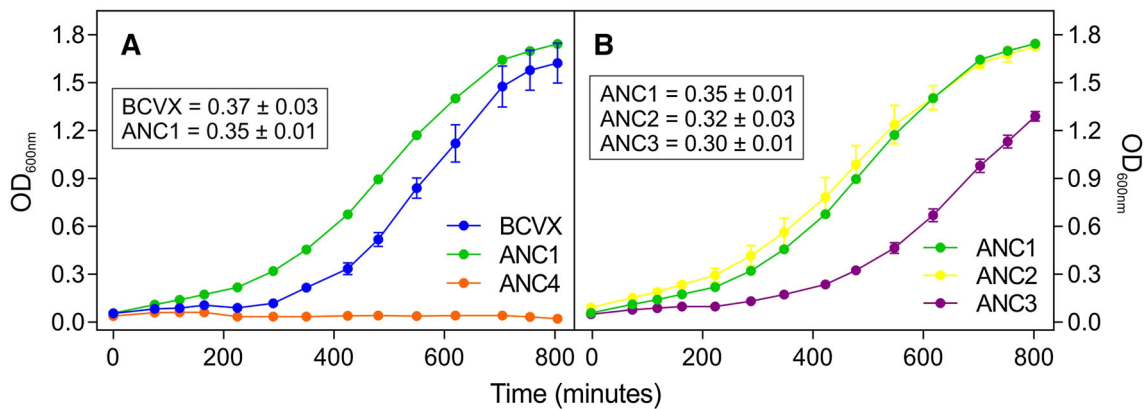


Fig. 3 Growth curves of $\Delta leuB$ complemented with different LeuBs under the control of the *leu* promoter. Growth rate determinations of *E. coli* $\Delta leuB$ complemented with a contemporary or ancestral LeuB expressed under the control of the *leu* promoter were performed in triplicate (error bars represent the SEM). **a** Comparison of growth

curves for $\Delta leuB$ (pUC19-BCVX), $\Delta leuB$ (pUC19-ANC1) and $\Delta leuB$ (pUC19-ANC4). **b** Comparison of growth curves for $\Delta leuB$ (pUC19-ANC1), $\Delta leuB$ (pUC19-ANC2) and $\Delta leuB$ (pUC19-ANC3). Inset boxes contain the growth rate constant for each strain in per hour (\pm SEM)

study to include other ancestral LeuBs we have previously studied in vitro: ANC2 and ANC3 (Hobbs et al. 2012), and LeuB_{S-aw} and LeuB_{S-unaw} (Groussin et al. 2015; Fig. 1). We were unable to include the two additional contemporary LeuBs we have previously characterised in vitro (BPSYC and BSUB) as these enzymes do not fold correctly when expressed at 37 °C (Hobbs et al. 2012). In vitro, ANC2 and ANC3 are similar to BCVX and ANC1 in that they exhibit moderate k_{cat} values; however, they are considerably less kinetically stable ($\Delta G_{\text{N-U}}^{\ddagger}$, Table 1). LeuB_{S-aw} and LeuB_{S-unaw} are also interesting enzymes to study in terms of fitness as LeuB_{S-unaw} has a k_{cat} even higher than that of ANC4 but low kinetic stability, and LeuB_{S-aw} has a moderate k_{cat} value similar to that of ANC1 but is as kinetically stable as ANC4. We have previously analysed the amino acid differences between our ancestral LeuBs and found no sequence- or structure-based rationale for their differing in vitro biochemical/biophysical properties (Groussin et al. 2015; Hobbs et al. 2012).

All four additional ancestral *leuB* genes were cloned under the control of the *leu* promoter and transformed into the $\Delta leuB$ strain. In minimal medium, $\Delta leuB$ (pUC19-ANC2) grew similarly to $\Delta leuB$ (pUC19-ANC1) but with a slightly slower growth rate (Fig. 3b). $\Delta leuB$ (pUC19-ANC3) grew markedly, although not statistically significantly, slower than $\Delta leuB$ (pUC19-ANC1) and $\Delta leuB$ (pUC19-ANC2) (one-way ANOVA with Dunnett's multiple comparisons test, $P > 0.05$), and did not reach as high a cell density after 13 h. $\Delta leuB$ (pUC19-LeuB_{S-aw}) and $\Delta leuB$ (pUC19-LeuB_{S-unaw}) failed to grow detectably during the course of the experiment but did show some growth on minimal agar (data not shown). These results do not support our hypothesis that the high k_{cat} and/or $\Delta G_{\text{N-U}}^{\ddagger}$ of ANC4 are the cause of its negative effect on fitness in vivo as $\Delta leuB$ (pUC19-ANC3) exhibited a noticeably lower growth rate than $\Delta leuB$ (pUC19-BCVX) and $\Delta leuB$ (pUC19-ANC1) despite having a similar k_{cat} and a lower $\Delta G_{\text{N-U}}^{\ddagger}$ (although it does possess a twofold higher K_{M} for IPM). Furthermore, both $\Delta leuB$ (pUC19-LeuB_{S-aw}) and $\Delta leuB$ (pUC19-LeuB_{S-unaw}) displayed poor fitness in vivo even though these enzymes each only exhibit one “improved” biochemical/biophysical property (i.e. high k_{cat} or high $\Delta G_{\text{N-U}}^{\ddagger}$).

In order to perform a full correlation analysis between in vivo fitness and in vitro biochemical/biophysical properties of LeuB enzymes, we needed a fitness score for all the enzymes, including those for which we were unable to obtain growth rates. As such, we decided to use growth on minimal agar as a semi-quantitative measure of fitness. Given our previous concerns regarding the reconstruction accuracy of LeuB_{S-unaw} (Groussin et al. 2015), we decided

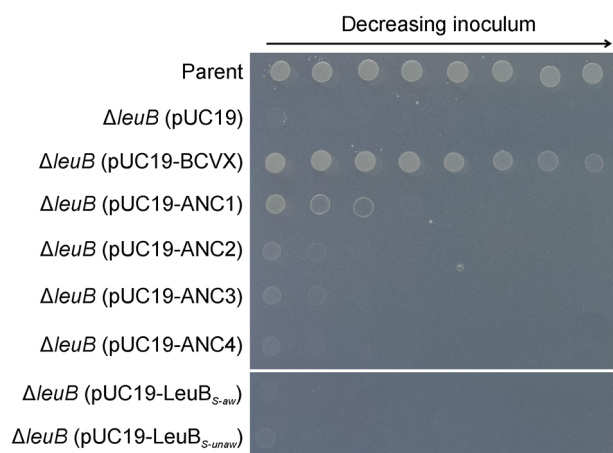


Fig. 4 Relative fitness of $\Delta leuB$ complemented with different LeuBs under the control of the *leu* promoter. Parental, control and complemented strains were spotted onto M9 minimal agar at a range of culture dilutions. A fitness score between 1 and 8 was assigned to each strain according to the number of spots visible following 24-h incubation at 37 °C

to exclude this enzyme from the correlation analysis. The remaining six complemented strains (plus control strains) were spotted onto minimal agar at a range of culture dilutions and the fitness score taken as the number of spots showing visible growth after 24 h. The fitness scores for $\Delta leuB$ complemented with BCVX, ANC1, ANC2, ANC3, ANC4 and LeuB_{S-aw} were 8, 4, 2, 2, 1 and 1, respectively (Fig. 4). Plotting these fitness scores against each of the biochemical and biophysical parameters shown in Table 1 suggested nonlinearity, and monotonicity for several of the parameters, and therefore, correlation analysis was performed using the nonparametric Spearman rank-order correlation coefficient (Fig. 5). Despite the previously reported associations between increased protein stability and lower aggregation rates, higher soluble abundance and a greater capacity to evolve (DePristo et al. 2005; Bershtein et al. 2012; Tomala et al. 2014; Bloom et al. 2006), we found no statistically significant correlation, either positive or negative, between the measures of LeuB stability (T_{opt} and $\Delta G_{\text{N-U}}^{\ddagger}$) and organism fitness (Fig. 5). There is also no statistically significant correlation between k_{cat} and fitness, which is consistent with the lack of correlation reported by Bershtein et al. (2012). We must acknowledge here that, whilst our data do not indicate that any of the biochemical/biophysical parameters display a correlation with fitness that is statistically significant, we are working with a relatively small dataset so we cannot exclude the possibility that correlations do exist.

We have previously estimated the evolutionary ages of our inferred LeuBs (Fig. 1; Groussin et al. 2015; Hobbs et al. 2012) by comparing their phylogenetic positions to a published timescale of prokaryotic evolution (Battistuzzi

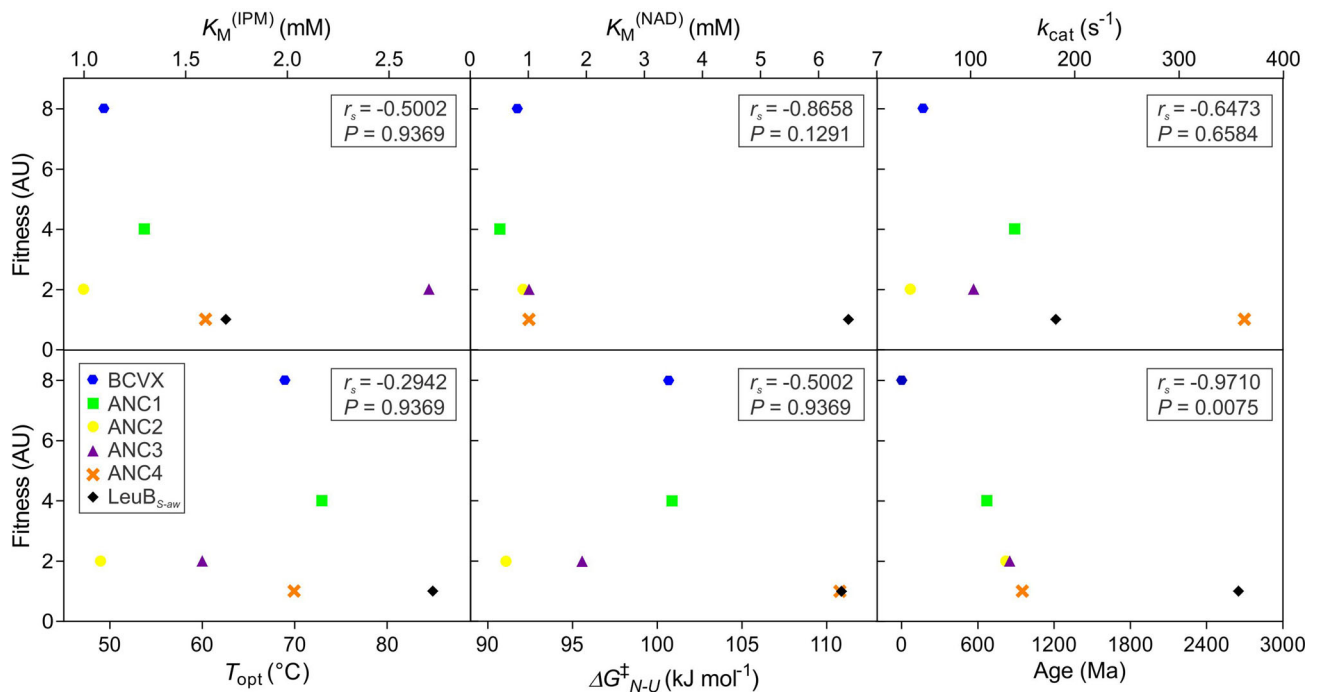


Fig. 5 Correlation analysis of biochemical, biophysical and evolutionary LeuB parameters with organism fitness. Scatterplots show the biochemical and biophysical data from Table 1, in addition to the age of the LeuB enzymes, plotted against the relative fitness scores determined from Fig. 4. Points are colour coded according to the different enzymes (see legend). Inset boxes contain the Spearman

rank order correlation coefficient (r_s) and two-tailed P value for each correlation. The P -values shown have been corrected for multiple testing using Holm's method; the P value for the correlation between evolutionary age and fitness is the same when corrected using the Bonferroni method

et al. 2004). We use these “ages” here simply as a measure of the distances of our inferred ancestors from the leaves. Interestingly, when we included these ages in our correlation analysis we found there to be a strong, negative and statistically significant correlation between estimated evolutionary age and fitness (Fig. 5; $r_s = -0.9710$, $P = 0.0075$). Of course, we must acknowledge that there are a number of assumptions/caveats associated with this correlation. Firstly, we must assume that the inferred sequences are in some way representative of the ancestral state; inferred sequences simply represent the “most likely” sequences given the reconstructed phylogeny and inference parameters. Secondly, the evolutionary ages inferred here are *broad estimates*; for example, the 95 % confidence interval for the LeuB_{S-aw} node is 2367–3013 Ma, as reported by Battistuzzi et al. (2004). Thirdly, the accuracy of ancestral inference decreases with increasing distance from the leaves, therefore the observed correlation between evolutionary age and fitness may actually be the result of decreased ancestral inference accuracy. Notwithstanding these caveats, this correlation remains interesting and certainly worthy of further investigation.

Williams et al. (2006) have previously reported that maximum likelihood (ML) inference methods, such as those used in the inference of LeuB, may have a tendency

to gradually overestimate the stability of ancestral proteins along the phylogenetic gene tree. ML is known to be biased in its estimation of ancestral characters. The first bias concerns the use of the state with maximum posterior probability for a given node at a given site. This approach has a tendency to assign the ancestral states that are the most frequently observed at a given site, progressively excluding observed states with low frequencies when reconstructing ancestral states upwards along the phylogenetic tree (Yang 2006). These less frequent amino acids tend to decrease the stability of the protein with respect to highly frequent states. Therefore, one could speculate that the molecular basis of the correlation between evolutionary age and fitness is over-stabilisation. We do not, however, believe this to be the case as our measures of stability and fitness are not well correlated (Fig. 5). A second known bias with ML methods is the tendency to incorporate into ancestral sequences the states that have the highest equilibrium frequency in the employed substitution model (Yang 2006). Thus, it has been suggested that this bias may result in a preponderance of hydrophobic residues in ancestral sequences, as hydrophobic residues tend to have high equilibrium frequencies in substitution models (Gaucher et al. 2008; Williams et al. 2006). This, in turn, could affect the properties of an ancestral protein. To address this

point, we compared the amino acid compositions of our ancestral LeuBs. We found no propensity towards hydrophobic residues within our ancestral LeuBs, and no statistically significant correlation between protein pI or the proportion of hydrophobic, polar, charged, basic or acidic residues and estimated evolutionary age (Online Resource 2). We cannot, of course, exclude the possibility that our ancestral sequences contain biases or errors that are undetectable from sequence analysis but still impactful. Similarly, we analysed our LeuB sequences for their aggregation potential (using TANGO; Fernandez-Escamilla et al. 2004) and, again, found no trend with inferred evolutionary age. All of our LeuBs are predicted to have a low average percentage aggregation per residue (<2 %) and, in fact, BCVX contains the greatest number of aggregation-prone residues and aggregation-prone stretches. Our experimental data also support this prediction that our LeuBs are not aggregated in vitro as we have been able to crystallise ANC1 (Prentice 2013) and ANC4 (Hobbs et al. 2012), and obtain differential scanning calorimetry curves for ANC1, ANC2, ANC3 and ANC4 (Hobbs et al. 2012). In addition, we have no evidence to suggest that the observed lack of fitness of our oldest ancestral LeuBs is due to toxicity in *E. coli* as they are all expressed to a high level during recombinant protein production (see Supplementary Fig. S2 in Hobbs et al. 2012 for an example).

Our biochemical data and sequence analysis have failed to identify a reconstruction bias behind, or molecular basis for, the fitness cost associated with our oldest inferred ancestral LeuBs. Therefore, we hypothesise that ancestral enzymes of increasing evolutionary age may be imposing a fitness cost on modern bacteria due to an increasing degree of uncoupling between the ancestral LeuB and the other enzymes of the leucine biosynthetic pathway. Biosynthesis of leucine begins with an intermediate from the valine biosynthetic pathway that is processed by LeuA (IPM synthase), LeuB, LeuCD (IPM isomerase) and an aminotransferase into leucine (Vartak et al. 1991). In *E. coli* and *B. subtilis*, this pathway is regulated via two mechanisms: feedback inhibition of LeuA by leucine, and leucine-mediated transcriptional attenuation of the *leuABCD* operon (Freundlich et al. 1962; Ward and Zahler 1973). As the expression of LeuA, LeuB and LeuCD are controlled together, we can assume that the functionalities of these enzymes have evolved in unison. Therefore, the insertion of an ancestral LeuB into a modern leucine biosynthesis pathway has the potential to cause downstream effects, such as increased or decreased expression of the other enzymes, build-up of pathway intermediates and/or reduced pathway efficiency. Furthermore, the product of leucine biosynthesis is involved in the regulation of valine and isoleucine biosynthesis (Freundlich et al. 1962), and LeuB is involved

in a number of protein–protein interactions according to the IntAct database (Orchard et al. 2014), including one with the large subunit of acetolactate synthase (*ilvI*) from the isoleucine biosynthesis pathway, so the consequences of a less than optimal leucine biosynthetic pathway could be far reaching for metabolism in general. This hypothesis is speculative, but warrants further research.

Conclusions

Reconstructed ancestral enzymes are frequently reported to be more stable and more catalytically active than their contemporary descendants (e.g. Groussin et al. 2015; Hobbs et al. 2012; Perez-Jimenez et al. 2011; Risso et al. 2013), but the effect of these enzymes on in vivo organismal fitness has not been addressed. Here, using our previously reconstructed LeuB enzymes, we have shown that older-inferred ancestral enzymes confer a fitness cost to modern bacteria, and that there is a strong negative correlation between the estimated evolutionary age of an inferred enzyme and the fitness it confers. Conversely, the two in vitro biochemical/biophysical properties that are commonly associated with ancestral enzymes do not appear to be correlated with fitness. The data presented in this study provide a working hypothesis as to why ancestral enzymes are *ancestral* and not found in contemporary organisms—because they do not confer a high level of fitness and are outcompeted by their contemporary counterparts—but there are a number of outstanding questions. Firstly, the data presented here are for a single candidate enzyme, and it is unknown whether the same correlation between estimated evolutionary age and organism fitness would be observed with other reconstructed ancestral enzymes. Secondly, we do not know the molecular basis of the fitness cost associated with the inferred ancestral enzymes in this study. With regard to both of these points, we need to again acknowledge that ancestral inference accuracy decreases with the increasing evolutionary age, and therefore, the underlying cause of the observed unfitness may be the low accuracy of ancestral inference. If this is the case, we may expect to observe the same correlation between estimated evolutionary age and fitness for other reconstructed ancestral enzymes. However, the accuracy of inference is not only dependent on the distance from the leaves but also on protein-specific factors, such as branch lengths and the robustness of the protein fold to mutation (Tokiriki and Tawfik 2009). Even if inference accuracy is behind this trend, it still does not provide any insight into the molecular basis of the fitness effects. It is possible that we may never uncover the molecular basis of the trend in evolutionary age and fitness we have observed; for example, the molecular basis of the fitness cost associated with

each inferred enzyme may be different or multifaceted (e.g. being dependent on a combination of one or more of the biochemical properties we have studied), or the cause of the in vivo fitness cost may not be observable or relevant in vitro. However, there are a number of experiments which could be used to test our hypothesis regarding the uncoupling of LeuB from the rest of the leucine biosynthesis pathway. A recent study by Bershtein et al. (2013) suggests that overexpression of the chaperone complex GroEL/ES or deletion of the Lon protease can rescue the slow growth of *E. coli* strains carrying destabilised mutant proteins by affecting the balance between protein production, folding and degradation. Although our ancestral LeuBs are not destabilised in comparison with BCVX, this could be an interesting experiment to perform. The potentially most revealing experiment, however, would be to place one of our unfit ancestral LeuBs on the chromosome, forward-evolve the strain until its fitness improves and then map the resulting mutations (ideally, this experiment would be performed in *Bacillus*). The effect of mutations that map to LeuB itself could then be studied in vitro, but we suspect that mutations would also be found further afield in the rest of the *leuABCD* operon and its regulatory regions. Compensatory mutations in other genes may also occur, including those that encode for proteins that interact with LeuB, which in turn would provide valuable insight into the evolution of amino acid biosynthesis, wider metabolic networks and gene redundancy. These experiments are ongoing in our laboratory.

In conclusion, we have shown that superior in vitro biochemical properties exhibited by reconstructed ancestral enzymes do not translate into greater in vivo organismal fitness. Additional experiments are required to determine the molecular basis behind this phenomenon.

Acknowledgments We would like to thank Dr Jonathan Abell for his assistance with statistical analysis. This work was supported by a research grant from the Marsden Fund of New Zealand awarded to VLA and a University of Waikato Research Trust Contestable Fund grant awarded to JKH. EJP was supported by a number of Masters scholarships from the University of Waikato. MG was supported by the French Agence Nationale de la Recherche (ANR) and is a contributor to the Ancestrome project (ANR-10-BINF-01-01).

Compliance with Ethical Standards

Conflict of interest The authors declare that they have no conflict of interest.

References

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 USA* 110:11067–11072

- Alcaraz LD, Moreno-Hagelsieb G, Eguarte LE, Souza V, Herrera-Estrella L, Olmedo G (2010) Understanding the evolutionary relationships and major traits of *Bacillus* through comparative genomics. *BMC Genom* 11:332–348
- Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL, Mori H (2006) Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol*. doi:10.1038/msb4100050
- Battistuzzi FU, Feijao A, Hedges SB (2004) A genomic timescale of prokaryote evolution: insights into the origin of methanogenesis, phototrophy, and the colonization of land. *BMC Evol Biol* 4:44–57
- Bershtein S, Mu W, Shakhnovich EI (2012) Soluble oligomerization provides a beneficial fitness effect on destabilizing mutations. *Proc Natl Acad Sci USA* 109:4857–4862
- Bershtein S, Mu W, Serohijos AWR, Zhou J, Shakhnovich EI (2013) Protein quality control acts on folding intermediates to shape the effects of mutations on organismal fitness. *Mol Cell* 49:133–144
- Bloom JD, Silberg JJ, Wilke CO, Drummond DA, Adami C, Arnold FH (2005) Thermodynamic prediction of protein neutrality. *Proc Natl Acad Sci USA* 102:606–611
- Bloom JD, Labthavikul ST, Otey CR, Arnold FH (2006) Protein stability promotes evolvability. *Proc Natl Acad Sci USA* 103:5869–5874
- Bloom JD, Raval A, Wilke CO (2007) Thermodynamics of neutral protein evolution. *Genetics* 175:255–266
- Bridgham JT, Carroll SM, Thornton JW (2006) Evolution of hormone-receptor complexity by molecular exploitation. *Science* 312:97–101
- Bridgham JT, Ortlund EA, Thornton JW (2009) An epistatic ratchet constrains the direction of glucocorticoid receptor evolution. *Nature* 461:515–519
- Bridgham JT, Eick GN, Larroux C, Deshpande K, Harms MJ, Gauthier MEA, Ortlund EA, Degnan BM, Thornton JW (2010) Protein evolution by molecular tinkering: diversification of the nuclear receptor superfamily from a ligand-dependent ancestor. *PLoS Biol* 8:e1000497
- Butzin NC, Lapiere P, Green AG, Swithers KS, Gogarten JP, Noll KM (2013) Reconstructed ancestral myo-inositol-3-phosphate synthases indicate that ancestors of the Thermococcales and *Thermotoga* species were more thermophilic than their descendants. *PLoS ONE* 8:e84300
- Chang BS, Jonsson K, Kazmi MA, Donoghue MJ, Sakmar TP (2002) Recreating a functional ancestral archosaur visual pigment. *Mol Biol Evol* 19:1483–1489
- Chinen A, Matsumoto Y, Kawamura S (2005) Reconstitution of ancestral green visual pigments of zebrafish and molecular mechanism of their spectral differentiation. *Mol Biol Evol* 22:1001–1010
- Cox RA, Cook GM (2007) Growth regulation in the mycobacterial cell. *Curr Mol Med* 7:231–245
- DePristo MA, Weinreich DM, Hartl DL (2005) Missense meanderings in sequence space: a biophysical view of protein evolution. *Nature* 6:678–687
- Dutheil J, Boussau B (2008) Non-homogeneous models of sequence evolution in the Bio ++ suite of libraries and programs. *BMC Evol Biol* 8:255
- Eick GN, Colucci JK, Harms MJ, Ortlund EA, Thornton JW (2012) Evolution of minimal specificity and promiscuity in steroid hormone receptors. *PLoS Genet* 8:e1003072
- Fernandez-Escamilla AM, Rousseau F, Schymkowitz J, Serrano L (2004) Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins. *Nat Biotechnol* 22:1302–1306
- Finnigan GC, Hanson-Smith V, Stevens TH, Thornton JW (2012) Evolution of increased complexity in a molecular machine. *Nature* 481:360–364

- Freundlich M, Burns RO, Umbarger HE (1962) Control of isoleucine, valine, and leucine biosynthesis. I. Multivalent repression. *Proc Natl Acad Sci USA* 48:1804–1808
- Gaucher EA, Thomson JM, Burgan MF, Benner SA (2003) Inferring the palaeoenvironment of ancient bacteria on the basis of resurrected proteins. *Nature* 425:285–288
- Gaucher EA, Govindarajan S, Ganesh OK (2008) Palaeotemperature trend for Precambrian life inferred from resurrected proteins. *Nature* 451:704–707
- Groussin M, Hobbs JK, Szöllösi GJ, Gribaldo S, Arcus VL, 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
- Hall BG, Acar H, Nandipati A, Barlow M (2014) Growth rates made easy. *Mol Biol Evol* 31:232–238
- Harms MJ, Thornton JW (2013) Evolutionary biochemistry: revealing the historical and physical causes of protein properties. *Nat Rev Genet* 14:559–571
- Hart KM, Harms MJ, Schmidt BH, Elya C, Thornton JW, Marqusee S (2014) Thermodynamic system drift in protein evolution. *PLoS Biol* 12:e1001994
- Haughn GW, Wessler SR, Gemmill RM, Calvo JM (1986) High A + T content conserved in DNA sequences upstream of *leuABCD* in *Escherichia coli* and *Salmonella typhimurium*. *J Bacteriol* 166:1113–1117
- 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
- Jermann TM, Opitz JG, Stackhouse J, Benner SA (1995) Reconstructing the evolutionary history of the artiodactyl ribonuclease superfamily. *Nature* 374:57–59
- Lenski RE, Rose MR, Simpson SC, Tadler SC (1991) Long-term experimental evolution in *Escherichia coli*: I. Adaptation and divergence during 2,000 generations. *Am Nat* 138:1315–1341
- Liberles DA, Teichmann SA, Bahar I et al (2012) The interface of protein structure, protein biophysics, and molecular evolution. *Protein Sci* 21:769–785
- Malcolm BA, Wilson KP, Matthews BW, Kirsch JF, Wilson AC (1990) Ancestral lysozymes reconstructed, neutrality tested, and thermostability linked to hydrocarbon packing. *Nature* 345:86–89
- Miyazaki J, Nakaya S, Suzuki T, Tamakoshi M, Oshima T, 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
- Orchard S, Ammari M, Aranda B et al (2014) The MIntAct project—IntAct as a common curation platform for 11 molecular interaction databases. *Nucleic Acids Res* 42:D358–D363
- Parsell DA, Sauer RT (1989) The structural stability of a protein is an important determinant of its proteolytic susceptibility in *Escherichia coli*. *J Biol Chem* 264:7590–7595
- Perez-Jimenez R, Ingles-Prieto A, Zhao ZM, Sanchez-Romero I, Alegre-Cebollada J, Kosuri P, Garcia-Manyes S, Kappock TJ, Tanokura M, Holmgren A, Sanchez-Ruiz JM, Gaucher EA, Fernandez JM (2011) Single-molecule paleoenzymology probes the chemistry of resurrected enzymes. *Nat Struct Mol Biol* 18:592–596
- Prentice EJ (2013) Characterisation of enzyme evolution through ancestral enzyme reconstruction. Dissertation, University of Waikato
- 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
- Serohijaw AWR, Shakhnovich EI (2014) Merging molecular mechanism and evolution: theory and computation at the interface of biophysics and evolutionary population genetics. *Curr Opin Struct Biol* 26:84–91
- Somero GN (1995) Proteins and temperature. *Annu Rev Physiol* 57:43–68
- Soskine M, Tawfik DS (2010) Mutational effects and the evolution of new protein functions. *Nat Rev Genet* 11:572–582
- Stackhouse J, Presnell SR, McGeehan GM, Nambiar KP, Benner SA (1990) The ribonuclease from an extinct bovid ruminant. *FEBS Lett* 262:104–106
- Studer RA, Christin PA, Williams MA, Orengo CA (2014) Stability-activity tradeoffs constrain the adaptive evolution of RubisCO. *Proc Natl Acad Sci USA* 111:2223–2228
- Taverna DM, Goldstein RA (2002) Why are proteins marginally stable? *Proteins* 46:105–109
- Tokiriki N, Tawfik DS (2009) Stability effects of mutations and protein evolvability. *Curr Opin Struct Biol* 19:596–604
- Tomala K, Pogoda E, Jakubowska A, Korona R (2014) Fitness costs of minimal sequence alterations causing protein instability and toxicity. *Mol Biol Evol* 31:703–707
- Vartak NB, Liu L, Wang BM, Berg CM (1991) A functional *leuABCD* operon is required for leucine synthesis by the tyrosine-repressible transaminase in *Escherichia coli* K-12. *J Bacteriol* 173:3864–3871
- Voordeckers K, Brown CA, Vanneste K, van der Zande E, Voet A, Maere S, Verstrepen KJ (2012) Reconstruction of ancestral metabolic enzymes reveals molecular mechanisms underlying evolutionary innovation through gene duplication. *PLoS Biol* 10:e1001446
- Ward JB Jr, Zahler SA (1973) Regulation of leucine biosynthesis in *Bacillus subtilis*. *J Bacteriol* 116:727–735
- Watanabe K, Yamagishi A (2006) The effects of multiple ancestral residues on the *Thermus thermophilus* 3-isopropylmalate dehydrogenase. *FEBS Lett* 580:3867–3871
- Watanabe K, Ohkuri T, Yokobori S, Yamagishi A (2006) Designing thermostable proteins: ancestral mutants of 3-isopropylmalate dehydrogenase designed by using a phylogenetic tree. *J Mol Biol* 355:664–674
- Williams PD, Pollock DD, Blackburne BP, Golstein RA (2006) Assessing the accuracy of ancestral protein reconstruction methods. *PLoS Comput Biol* 2:e69
- Yang Z (2006) Computational molecular evolution. Oxford University Press, Oxford
- Yang Z (2007) PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol* 24:1586–1591
- Yokoyama S, Yang H, Starmer WT (2008) Molecular basis of spectral tuning in the red- and green-sensitive (M/LWS) pigments in vertebrates. *Genetics* 179:2037–2043
- Zeldovich KB, Chen P, Shakhnovich EI (2007) Protein stability imposes limits on organism complexity and speed of molecular evolution. *Proc Natl Acad Sci USA* 104:16152–16157