Mutational studies on resurrected ancestral proteins reveal conservation of site-specific amino acid preferences throughout evolutionary history

Valeria A. Risso[1,a], Fadia Manssour-Triedo[1,a], Asunción Delgado-Delgado[1], Rocio Arco[1], Alicia Barroso-delJesus[2], Alvaro Ingles-Prieto[1,3], Raquel Godoy-Ruiz[4], Jose A. Gavira[5], Eric A. Gaucher[6], Beatriz Ibarra-Molero[1], and Jose M. Sanchez-Ruiz[1,*]

[1]Departamento de Quimica Fisica, Facultad de Ciencias, Universidad de Granada, 18071-Granada, Spain.

[2]Unidad de Genómica, Instituto de Parasitología y Biomedicina López-Neyra, IPBLN-CSIC, PTS Granada, Avda. del Conocimiento s/n, Armilla, 18016 Granada, Spain.

[3]Present address: IST Austria, Am Campus 1, 3400 Klosterneuburg, Austria.

[4]Department of Biochemistry and Molecular Biology, University of Maryland School of Medicine, Baltimore, Maryland 21201, United States.

[5]Laboratorio de Estudios Cristalográficos, Instituto Andaluz de Ciencias de la Tierra (Consejo Superior de Investigaciones Científicas – Universidad de Granada), Avenida de las Palmeras 4, 18100-Armilla, Granada, Spain.

[6]School of Biology, School of Chemistry and Biochemistry, and Parker H. Petit Institute for Bioengineering and Biosciences, Georgia Institute of Technology, Atlanta, Georgia 30332, United States.

[a]V.A.R. and F.M.-T. contributed equally to this work.

[*]**Corresponding author:** Email: sanchezr@ugr.es

1. Determination of the stability for variants of the laboratory resurrection corresponding to the thioredoxin of the last bacterial common ancestor (variants of LBCA thioredoxin)

1.1. General description of the studied variants of LBCA thioredoxin

We show below an alignment of the sequences of the extant and laboratory-resurrected ancestral proteins studied in this work, with labels on the positions we targeted for mutation:

```
                 **   **  *  *      * *            *    **   *
E. coli    SDKIIHLTDDSFDTDVLKADGAILVDFWAEWCGPCKMIAPILDEIADEYQGKLTVAKL    58
LBCA       -MSVIEINDENFEEEVLKSDKPVLVDFWAPWCGPCRMIAPIIEELAEEYEGKVKFAKV    57
           1.......10........20........30........40........50........

            **          *  *         **    *          *  *
E. coli    NIDQNPGTAPKYGIRGIPTLLLFKNGEVAATKVGALSKGQLKEFLDANLA   108
LBCA       NVDENPETAAKYGIMSIPTLLLFKNGEVVDKLVGARPKEALKERIEKHL-   106
           60........70........80........90.......100........
```

We aimed to reproduce in the ancestral LBCA thioredoxin background mutations we previously studied (refs. Godoy-Ruiz et al. 2004, 2005 and 2006 in the main text) in the extant *E. coli* thioredoxin background. As described in detail in the main text, in some cases the mutation had to be carried out in the reverse direction in the ancestral background. All the variants studied, together with their relation with the mutations targeted, are described in supplementary table S1.

We used the following procedure to estimate of the degree of evolutionary conservation of the local environments for the mutated positions. For the X-ray crystal structures of both, *E. coli* thioredoxin and LBCA thioredoxin (PDB codes 2H6X and 4BA7, respectively) we determined the amino acid residues present in the neighborhood of each mutated position. We considered a residue to be in the neighborhood of a given position if at least one of its atoms is within 6 Å of one side-chain atom of the residue present at the position. A simple metric of local environment conservation for each position (see Figure 2C in the main text) was derived as the fraction of the of neighbor residues in LBCA thioredoxin that coincide with the neighbor residues in *E. coli* thioredoxin.

1.2 Determination of thermal stability by differential scanning calorimetry (DSC)

Thermal stabilities of the thioredoxin variants (supplementary table S1) were measured with a VP-Capillary DSC [Microcal, GE Healthcare]. Protein solutions for DSC experiments were prepared by exhaustive dialysis against 50 mM Hepes buffer, pH 7 and the buffer from the last dialysis step was used to fill the reference cell in the calorimetric experiments. Protein concentration was determined spectrophotometrically using extinction coefficient values determined from the sequence. Given the high denaturation temperature of all variants of LBCA thioredoxin (around 110 ºC), most DSC experiments were performed up to a temperature of 140 ºC in order to measure reliably the post-transition baselines. However, the first scans in experiments performed to assess the calorimetric reversibility of the thermal denaturation process were terminated at 120 ºC (a temperature corresponding approximately to the end of the calorimetric transitions) before cooling

and proceeding to perform a reheating run. All experiments were performed under an overpressure of 4 atm to prevent boiling at temperatures above 100 ºC. Such an overpressure increases the boiling point of water by many degrees but has a negligible effect on a denaturation process in a condensed phase. Most DSC experiments were performed at a scan-rate of 90 degrees per hour, although a scan rate of 200 degrees per hour was used in experiments aimed at determining reversibility. Also, studies into the scan-rate dependence of the calorimetric transitions employed three different scan-rates within the 45-180 degrees per hour range. A typical calorimetric run involved several buffer-buffer baselines to ensure proper equilibration of the calorimeter followed by runs with several protein variants with intervening buffer-buffer baselines. Since the main purpose of these experiments was the determination of mutation effects with respect to the "wild-type" LBCA-thioredoxin, a scan with this protein was included every 4-5 scans with mutant variants to ascertain instrumental stability. Representative examples of the different calorimetric experiments performed are collected in supplementary figure S1.

1.3 Assessing the protein concentration dependence of the calorimetric transitions for the variants of LBCA thioredoxin

A protein concentration dependence of the calorimetric transition for the thermal denaturation of the extant thioredoxin from *E. coli* was reported many years ago by Ladbury et al. (1993) and attributed to dimerization processes of both, the native and the unfolded states. The protein concentrations at which these processes appear relevant are higher than the concentration used, both in this work and in our previous mutational studies on *E. coli* thioredoxin (refs. Godoy-Ruiz et al. 2004, 2005 and 2006 in the main text). Nevertheless, LBCA thioredoxin has low sequence identity with *E. coli* thioredoxin and it cannot be ruled out a priori that association processes occur at lower concentrations in LBCA thioredoxin. Therefore, for all the 21 variants studied in this work we performed at least two experiments (in some cases three) at different concentrations within the 0.2-1 mg/mL range. No significant concentration effect was found for any of the variants (supplementary table S2). Panel B in supplementary figure S1 shows a pictorial representation of this result. For each experiment, the difference between the T_m value and the mean value of the variant ($\langle T_m \rangle$) is calculated. These T_m-$\langle T_m \rangle$ differences are plotted versus protein concentration. It is apparent that the differences are always small and cluster around zero. We conclude that protein association processes do not significantly affect the experimentally determined calorimetric transitions.

1.4 Assessment the calorimetric reversibility of the thermal denaturation of the variants of LBCA thioredoxin

Experiments aimed at determining calorimetric reversibility were carried out for the 21 variants studied in this work at a scanning rate of 200 degrees per hour (see representative examples in panel C of supplementary figure S1). Calorimetric reversibility was estimated as the ratio of the excess heat capacity values (measured with respect to the extrapolation of the pre-transition baseline) for the first and reheating scans. In all cases, the scans were terminated at 120 ºC, roughly a temperature corresponding to the end of the calorimetric transitions, to avoid distortions caused by fast irreversible processes at higher temperatures. In all cases partial, but significant reversibility ratios typically within the 50%-80% range were obtained (see supplementary table

S3). Most reversibility experiments were carried out at a low concentration of about 0.2 mg/mL. For LBCA thioredoxin, however, additional experiments at higher concentration were carried out. No significant protein concentration effect on the degree of reversibility was detected, thus suggesting that aggregation is not the main process responsible for partial irreversibility. One possibility is that partial irreversibility for the proteins studied here is linked to chemical alterations of the unfolded state, which are known to be fast at high temperatures (Zale and Klibanov 1983). However, as it will be made clear by the results and analyses described below, the specific process responsible for partial irreversibility is not relevant for the calculation of mutation effects on the stability of LBCA thioredoxin.

1.5 Scan rate effect on the calorimetric transitions

The significant levels of calorimetric reversibility found (supplementary table S3) suggest that DSC transitions for the thermal denaturation of variants of LBCA thioredoxin reflect to a good degree of approximation an equilibrium unfolding process, although their high temperature tail may likely be distorted by the kinetics of the process responsible for partial irreversibility. In order to sustain this interpretation, we have determined scan rate effect on the transitions [a well-known metric of kinetic distortion (Sanchez-Ruiz 1992, 2010, 2011)] for several selected variants. For each variant, experiments were performed at three different scan rates: 45, 90 and 180 degrees per hour. The results (see representative examples in panel A of supplementary figure S1) do confirm the above interpretation, as the scan-rate effect is confined to the high temperature tail (note the different slopes) and does not significantly affect the denaturation temperature values.

1.6 Assessing the two-state character of the thermal unfolding of variants of LBCA thioredoxin

In order to assess the two-state character of the unfolding of LBCA thioredoxin variants, we have fitted the experimental DSC profiles on the basis of the following equation (Thórólfsson et al. 2002)

$$C_P = C_P(N) + \frac{\Delta H \Delta H^{vH}}{RT^2}\frac{K}{(1+K)^2} \qquad \text{(S1)}$$

where $C_P(N)$ is the pre-transition (i.e. "native") baseline, ΔH is the unfolding enthalpy and ΔH^{vH} is the so-called van`t Hoff enthalpy that describes the temperature-dependence of the unfolding equilibrium constant:

$$K = exp\left\{-\frac{\Delta H^{vH}}{R}\left(\frac{1}{T}-\frac{1}{T_m}\right)\right\} \qquad \text{(S2)}$$

where T_m is the denaturation temperature (temperature at which K=1).

For fitting purposes, a linear pre-transition baseline was assumed [i.e., $C_P(N)=a+bT$ where a and b are fitting parameters] while the temperature dependence of ΔH and ΔH^{vH} was taken as negligible within the narrow temperature range of the calorimetric transitions. That is, ΔH and ΔH^{vH} are taken as temperature-independent for each given fit and the values obtained are assigned to the obtained denaturation temperature value. Overall, five fitting parameters are involved in the analysis based on equations S1 and S2: a, b, T_m, ΔH and ΔH^{vH}.

The model embodied in equations S1 and S2 immediately leads to the established calorimetric two-state test (Privalov and Khechinashvili 1974). That is, for a two-state unfolding process, the "calorimetric" (ΔH) and van`t Hoff (ΔH^{vH}) enthalpies should show a good agreement, while disagreement between the two values does not mean, of course, that there are two different enthalpy change values for the same process, but that there are significant deviations from two-state behavior (i.e., significant population of intermediate states). It is also to be noted that equation S1 does not include a description of the post-transition baseline. The reason is that, as discussed above, the high-temperature tails of the experimentally determined calorimetric transitions are distorted by the kinetics of irreversible protein denaturation. For fitting purposes, therefore, they are truncated at temperature a few degrees above the temperature corresponding to the maximum heat capacity.

Equation S1 (together with equation S2) was used to fit calorimetric profiles for all the LBCA thioredoxin variants studied in this work. The non-linear, least-squares fittings were carried out using a program written by us in the MLAB environment (Civilized Software Inc.). Fits were excellent (see representative examples in panel D of supplementary figure S1) and the values derived for the energetically relevant parameters (T_m, ΔH and ΔH^{vH}) are collected in supplementary table S2. In all cases, there is good agreement between the calorimetric and van't Hoff enthalpy and the van`t Hoff to calorimetric enthalpy ratios, $r= \Delta H^{vH}/ \Delta H$ are reasonably close to unity indicating low-population of intermediate states and acceptable adherence to two-state unfolding.

1.7 The calculation of mutational effects on unfolding free energy changes

Overall, the results reported here support a simple Lumry-Eyring mechanism (Sanchez-Ruiz 1992, 2010; Lumry and Eyring 1954) for the thermal denaturation of LBCA thioredoxin variants under the conditions studied:

$$N \leftrightarrow U \rightarrow F \qquad \text{(S3)}$$

in which the reversible unfolding step conforms acceptably to the two-state model (N↔U) and does not include association processes or significantly populated intermediates, while the kinetically-determined step leading to irreversibly denatured protein (U→F) distorts the calorimetric profiles only at the high-temperature tail of the transitions.

Certainly, kinetic distortions at the higher temperatures within the transition ranges prevent us from deriving values of the unfolding heat capacity change (ΔC_P) from the DSC profiles. Furthermore, an average ΔC_P value for all the variants studied cannot be estimated in this case from the slope of a plot of calorimetric enthalpy versus denaturation temperature. The reason is that the narrow range of the experimental T_m values, together with the experimental scatter of the calorimetric enthalpy values (on the order of a few tens of kJ/mol: see supplementary table S2), prevent us from observing the small temperature dependence of ΔH linked to the ΔC_P value. The lack of a reliable ΔC_P value prevents us from using the integrated Gibbs-Helmoltz equation to calculate the mutational effects on unfolding free energy ($\Delta\Delta G$s). Nevertheless, since the mutational effects on denaturation temperature (ΔT_m values) are always small (i.e., they can be

considered as small perturbations), ΔΔGs can be calculated in this case from the ΔT_ms using the equation derived by John Schellman (ref. Schellman 1987 in the main text),

$$\Delta\Delta G = \frac{\Delta H^0}{T_m^0}\Delta T_m \qquad \text{(S4)}$$

where ΔH^0 and $T_m{}^0$ are the unfolding enthalpy and the denaturation temperature for the "unperturbed" system (i.e. LBCA thioredoxin). Note that, unlike the integrated Gibbs-Helmholtz equation, Schellman equation does not require a ΔC_P value. Values of ΔΔG shown in figure 3 in the main text and collected in supplementary table S2 have been calculated using equation S4. Note, however, the same conclusions are reached using the T_m values as a metric of protein stability and describing the mutation effects on stability on the basis of mutation ΔT_m values. To clearly show this point, figure 3 in the main text includes a ΔT_m scale. It is also worth noting that stability-related purifying selection in the thioredoxin system very likely operates on the basis of kinetic stability (related to the rate of irreversible denaturation *in vivo*) and that ΔΔG and ΔT_m can be viewed as suitable metrics of the mutational effects on kinetic stability (for details, see ref. Godoy-Ruiz 2006 in the main text).

2. Determination of the effect of the K/L exchange at position 90 on the stability extant and ancestral thioredoxins.

We used differential scanning calorimetry to characterize the thermal denaturation of the extant *E. coli* thioredoxin, the laboratory resurrections of the tiorredoxins corresponding to the LGPCA, LPBCA and LBCA nodes (figure 2 in the main text), the variants of *E. coli* and LGPCA thioredoxin with the K90L mutation and the variants of the LPBCA and LBCA thioredoxins. Experimental details of the calorimetric experiments are substantially identical to those described in section 1.2. Calorimetric transitions showed a significant degree of reversibility (within the 50%-90% range: see supplementary table S3) and both scan-rate and concentration effects on the transitions were found to be small. The overall picture appears therefore, similar to that we have described in section 1.7 for the variants of LBCA thioredoxin that include I↔V and E↔D exchanges; i.e., to a substantial extent the calorimetric transitions reflect an equilibrium unfolding process, but may be distorted at the high-temperature tail by the kinetics of an irreversible denaturation process. Actually, the model described by equations S1 and S2 provided visually acceptable fits to the DSC transitions truncated at temperature a few degrees above the temperature corresponding to the maximum heat capacity. However, we found that, in some cases, the discrepancies between the calorimetric and van't Hoff enthalpies were somewhat larger than those collected in supplementary table S2 for variants of LBCA thioredoxin with I↔V and E↔D mutations. This may be due to the presence of significantly populated intermediate states or to larger kinetic distortions of the DSC transitions. Therefore, we deemed convenient in this case to use the denaturation temperature value as a simple metric of stability and describe the mutational effects on stability on the basis of the ΔT_m values (see panel B in Figure 4 of the main text). It must be noted that, in any case, the use of Schellman equation (equation S4) is not advisable because of the large mutational changes in T_m involved (up to about 15 degrees). Note also that, precisely because of these large mutational effects, the evolutionary conservation of the K over L

preference (panel B in Figure 4 of the main text) is strongly supported (despite the fact that, in this case, T_m is a rather empirical metric of stability).

3. Determination of the effect of the M/T exchange at position 182 on the stability extant and ancestral β-lactamases.

We used differential scanning calorimetry to characterize the thermal denaturation of the extant TEM-1 *E. coli* β-lactamase, the extant *B. licheniformis* β-lactamase, the laboratory resurrections of the β-lactamases of the GPBCA, GNCA and PNCA nodes (see Figure 6 in the main text), the variant of the TEM-1 β-lactamase with the M182T mutation and the variants of the *B. licheniformis*, GPBCA, GNCA and PNCA β-lactamases with the T182M mutation. Experimental details of the calorimetric experiments are substantially identical to those described in section 1.2, except that 25 mM Hepes buffer pH 7 was used (instead of the 50 mM Hepes buffer pH 7 used in the thioredoxin experiments). Only small protein concentration effects on denaturation temperature values were found (see panel C in supplementary figure S4) and very small scan rate effects are observed for the laboratory-resurrected ancestral lactamases, as we have previously reported. In fact, the DSC transitions for the resurrected Precambrian lactamases are well described by a two-state equilibrium model, as we have previously reported (see ref. Risso et al. 2013 in the main text). However, significant scan rate effects are detected for the extant TEM-1 and *B. licheniformis* β-lactamases, although transitions obtained at the higher scan rates conformed acceptably to the equilibrium two-state model. Furthermore, scan-rate effects were essentially identical for the wild-type and the mutant variants (see panel D in supplementary figure S4). That is, whatever kinetic distortions are responsible for the scan-rate effect on the extant lactamases transitions, they do not significantly affect the mutational ΔT_m values. Therefore, we deem acceptable to use Schellman equation (equation S4) to calculate $\Delta\Delta G$ from ΔT_m for both, the extant and ancestral prion backgrounds (see panel B in Figure 6 of the main text).

We also studied the effect of the T182M mutation on the ancestral ENCA β-lactamase background (see panel A in Figure 6 of the main text). The T182M was clearly destabilizing in this background, as expected from the conserved T over M preference. However, the T182M variant of ENCA β-lactamase showed a two-transition calorimetric profile, suggesting decreased cooperativity. Since the calculation of the $\Delta\Delta G$ value is not straightforward in this case, we did not include ENCA β-lactamase in the data reported in Figure 6 of the main text.

4. Competition experiments performed to determine the effect on organismal fitness of the K/L exchange at position 90 in thioredoxin.

4.1 Setting up the system to study fitness effects of thioredoxin variants

We have used the thioredoxin deficient *E. coli* strain FA41 to study the effect on fitness of thioredoxin variants. When compared to its parental strain DHB4, FA41 shows a defect in growth, which is readily apparent even in optimum growth conditions (see panel A in supplementary figure S5). As shown below, complementation of FA41 with thioredoxin by the system used here restores the ability of this strain to grow at normal rate.

FA41 has been adapted for expression of thioredoxin variants expressed under the control of a T7 promoter in pET derivative plasmids (Novagen). Expression of target genes in these vectors requires the presence of the specific RNA polymerase from the T7 phage. The gene for T7 RNA polymerase was introduced in FA41 by lysogenization with λDE3 (Novagen), a modified λD69 phage bearing it under the control of the lacUV5 promoter.

Lysogenization of FA41 was performed following manufacturer recommendations (λDE3 Lysogenization Kit, Novagen). This system is designed so that growing colonies are expected to be stable λDE3 lysogens. In effect, since λDE3 lacks the infection, it cannot integrate or be excised from the chromosome. Therefore, lysogenization of λDE3 requires co-infection with a Helper Phage providing it. In addition, a Selection Phage is also included in the infection event; this phage cannot infect λDE3 lysogens and kills non-λDE3 cells, ensuring that only lysogens survive. Specifically, the following procedure was carried out: one independent colony of the thioredoxin minus strain FA41 was grown at 37ºC in LB medium supplemented with 10mM $MgSO_4$ and 0.2% maltose (to induce expression of the lambda receptor). At OD_{600} of 0.5, 1µl of the culture was co-infected with 10^8 pfu of λDE3, 10^8 pfu of Helper Phage and 10^8 pfu of Selection Phage, and the suspension was incubated for 20 min at 37ºC to allow phage adsorption. Mixtures were plated on LB and incubated overnight at 37ºC. Growing colonies were putative λDE3 lysogens. A number of these candidates was tested for resistance to infection by Selection Phage, suggesting that they were in fact λDE3 lysogens. One of these colonies was randomly selected and submitted to two tests for further verification of the presence of T7 RNA polymerase in the FA41 chromosome.

The first test was carried out by analyzing the inducibility of thioredoxin gene expression under the control of the T7 promoter. For that, λDE3 lysogenized FA41 was made competent and transformed with pET30a(+)::trxA. The transformed clone was incubated in the presence of 1 mM IPTG (Isopropyl ß-D-1-thiogalactopyranoside) and the corresponding cell extract was compared with that of non-induced cells in SDS PAGE analyses. A protein band with the expected size of thioredoxin was obtained under inducing conditions. This result indicated that the FA41 selected clone was able to overexpress thioredoxin, and therefore contained the T7 RNA polymerase.

Further verification of the presence of T7 RNA polymerase in the FA41 chromosome was derived from an *in vivo* test. It should be first mentioned that lysogenization with λDE3 had no effect on growth of FA41, which, as stated above, is deficient in respect to the wild type strain DHB4 (Fig S5). Thioredoxin in pET30a(+)::trxA complemented the growth deficiency of the thioredoxin deficient strain. In fact, basal expression of T7 RNA polymerase from the lacUV5 promoter in λDE3 lysogens occurs even in the absence of an inducer (Studier and Moffatt 1986). Therefore, we expected that under non inducing conditions, basal expression of thioredoxin occurs in cultures of FA41 λDE3 (pET30a(+)::trxA). In typical growth curve experiments, this strain shows similar growth as the wild type strain DHB4 (panel A in supplementary figure S5). Therefore, complementation of the FA41 with thioredoxin in this system restores the growth defect of this strain.

4.2 Estimation of relative proportions of wt and K90L thioredoxins in competition assays

In the competition experiments carried out in this work, the relative amounts of the two different bacterial strains in a mixed culture have been calculated using Sanger DNA sequencing technique and the QSV Analyzer software (see ref. Carr et al. 2009 in the main text). The quantification method exploits the simple principle that the proportion of two sequence variants in a mixture will determine the relative fluorescence heights of the base peak that represent each variant in a Sanger sequence electropherogram.

In our case, both bacterial strains (wt vs K90L) differ in a 2-nt mutation within the transgene coding for thioredoxin protein. Thus, by Sanger analysis of the plasmid DNA region containing this mutation, quantification of fluorescence peak height at this variable position can be used to infer the relative amounts of the two concomitant strains (see panel B in supplementary figure S5). However, prior to quantification, two data normalization steps are required. First, in order to correct for the sequence context-dependent incorporation of dideoxinucleotides (maximum height for the two variants are not expected to be equal), problem data is referenced to control data obtained from two standard electropherograms containing each only one of the sequence variants. Secondly, differences in the amount of the template DNA used in each sequencing reaction are normalized by comparing the variant nucleotide peak height to that of other invariant nucleotides within the same trace. All these data transformations are implemented in a computer algorithm developed by Bonthron and co-workers (ref. Carr et al. 2009 in the main text) for the QSV Analyzer software.

Our particular experimental design had an additional hurdle to be overcome. For accurate peak height analysis, there should be no other nearby sequence variation, since this may affect the peak height at the position under study. The wt sequence differs from the mutated sequence (K90L) in two consecutive nucleotides AA→TT. Since a 3´ variation is expected to affect to lesser extent, we chose the most 5´ variable position to be the one analyzed. During the calibration assays with known strain ratios we still detected a slightly but consistent over-estimation of the wt strain (AA sequence). We also observed a very good linear correlation with expected values (R = 0,99, Rsqr = 0,98, see panel C in supplementary figure S5), thus enabling us to use a factor for easily correct the (rather small) experimental deviation. In order to calculate the correction factor, a set of control samples with known strain ratios were included in each experiment and they were analyzed in parallel. The mean wt over-estimation value (expressed as a percentage) was calculated from control samples (in general ranging from 4% to 5%) and subtracted to the experimental data obtained from problem samples.

5. Supplementary references

Ladbury JE, Wynn R, Hellinga HW, Sturtevant JM. 1993. Stability of oxidized *Escherichia coli* thioredoxin and its dependence on protonation of the aspartic residue in the 26 position. *Biochemistry* 32:7526-7530. .

Lumry R, Eyring E. 1954. Conformational changes of proteins. *J Phys Chem* 58:110-120.

Privalov PL, Khechinashvili NN. 1974. A Thermodynamic approach to the problem of stabilization of globular protein structure: a calorimetric study. *J Mol Biol* 86:665-684.

Sanchez-Ruiz JM. 1992. Theoretical analysis of Lumry-Eyring in differential scanning calorimetry. *Biophys J* 6:921-935.

Sanchez-Ruiz JM. 2010. Protein kinetic stability. *Biophys Chem* 148:1-15.

Sanchez-Ruiz JM. 2011. Probing free-energy surfaces with differential scanning calorimetry. *Annu Rev Phys Chem* 62:231-255.

Studier FW, Moffatt BA. 1986. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J Mol Biol* 189: 113-130.

Thórólfsson M, Ibarra-Molero B, Fojan P, Petersen SB, Sanchez-Ruiz JM, Martinez A. 2002. L-phenylalanine binding and domain organization in human phenylalanine hydroxylase: a differential scanning calorimetry study. *Biochemistry* 41:7573-7585.

Zale SE, Klibanov AM. 1983. On the role of reversible denaturation (unfolding) in the irreversible thermal inactivation of enzymes. *Biotechnol Bioeng* 25:2221-2230.

6. Legends to the supplementary figures

Fig. S1 Thermal denaturation of LBCA thioredoxin and variants of LBCA thioredoxin by differential scanning calorimetry (DSC). A) Scan-rate dependence of the DSC transitions for the thermal denaturation of LBCA thioredoxin and some representative examples of variants of LBCA thioredoxin (I75V, D61E and E48D). The heat capacity profiles shown have been corrected for the instrumental baseline and shifted in the y-axis for the sake of clarity. Note that there is no significant scan-rate effect on the temperatures of the maximum of the calorimetric transitions. There is, however, some effect on the high temperature baselines, indicating that irreversible denaturation processes distort the calorimetric profiles at higher temperatures. B) Lack of significant protein concentration effect on the denaturation temperature values for LBCA thioredoxin variants. For each experiment, the difference between the T_m value and the variant mean value is calculated. The span of the y-axis approximately matches the experimental range of T_m values for all the LBCA thioredoxin variants. C) Representative examples of experiments performed to determine calorimetric reversibility of variants of LBCA thioredoxin. The first scans are shown with closed circles and the reheating scans with open circles. D) Representative examples of the fits of the model embodied in equations S2 and S3 to the DSC profiles for the thermal denaturation of LBCA thioredoxin variants. Data have been corrected for the instrumental baseline but not for water displacement in the sample cell (essentially a constant shift that does not have any influence in the fit). For this reason, apparent heat capacities are negative in some temperature ranges. E) Representative DSC profiles for the thermal denaturation of LBCA thioredoxin variants. Profiles shown have been corrected for the instrumental baseline and shifted in the y-axis for the purpose of clarity. They are shown in the order they were performed. To

ascertain instrumental stability, the first and last experiments correspond to "wild type" LBCA thioredoxin.

Fig. S2 Phylogenetic tree used for ancestral sequence reconstruction of thioredoxin enzymes (see ref. Perez-Jimenez et al. 2011 in the main text) color-coded according to the residue present at the indicated positions. Small circles highlight the evolutionary trajectory from the last bacterial common ancestor (LBCA) to *E. coli*. Many nodes along that trajectory can be dated on the basis of published estimates (ref. Hedges and Kumar 2009 in the main text). For some intermediate nodes, an interpolation in a plot age versus sequence identity was used. The node ages were used to construct the plot in Figure 7 of the main text. Note, however, that, in Figure 7, a single amino letter is used when the same amino acid is present two nodes of very similar ages (this accounts for the fact that there are more nodes in the trees shown here than there are amino acid letters in Figure 7).

Figure S3 Effect of the K/L exchange at position 90 on the stability of the extant *E. coli* thioredoxin and several laboratory resurrections corresponding to Precambrian nodes in the evolution of thioredoxins (see panel A in Figure 2 of the main text). Note that, in all cases, the exchange is stabilizing in the K→L direction, regardless of the amino acid actually present in the "wild-type" background. The transitions shown have been corrected for the chemical baseline.

Figure S4 Evolutionary history of the amino acid present at position 182 in β-lactamases and effect of the M/T exchange at that position in the stability of extant and laboratory-resurrected ancestral lactamases. A) Phylogenetic tree used for ancestral sequence reconstruction of β-lactamases (see ref. Risso et al. 2013 in the main text) color-coded according to the residue present at position 182. Circles highlight the evolutionary trajectory from the PNCA ancestor to *E. coli*. B) Effect of the M/T exchange of the stability of resurrected ancestral β-lactamase corresponding to the GNCA node (compare with panel C of Figure 6 of the main text). Note that, in all cases, the exchange is stabilizing in the M→T direction. C) Lack of significant protein concentration effect of the denaturation temperature values for variants of extant and ancestral β-lactamases. D) Effect of scan rate on the denaturation temperature values for variants of extant and ancestral β-lactamases. Note that, while some scan-rate dependence is apparent in some cases, mutational effects on denaturation temperature are nearly independent of scan rate (compare open and closed data points with the same color).

Figure S5 Experimental of the competition experiments at determining the effect of K/L exchange at position 90 on organismal fitness. A) Representative growth curves of DHB4 (wild type), FA41 (thioredoxin minus), FA41 λDE3 and FA41 λDE3 (pET30a(+)::trxA). Overnight cultures were diluted 1/200 in fresh LB medium, and OD600 was monitored with time. B) Representative electropherograms of samples from pure cultures, (1) wt strain, (2) K90L strain and (3) a mixed culture. The analyzed position is marked by a vertical red line. C) Linear correlation between the experimental values obtained during calibration assays using mixed cultures of known ratios and their respective expected values. Data is expressed as wt strain ratio.

Table S1 Positions in LBCA thioredoxin targeted for mutation in this work. The *E. coli* thioredoxin numbering is used. The positions labeled in red correspond to the positions studied in opposite mutation directions in the extant and the ancestral backgrounds and positions labeled in blue correspond to position studied the same direction.

Position	Amino acid present in *E. Coli* thioredoxin	Amino acid present in LBCA thioredoxin	Mutations performed in LBCA thioredoxin	Mutations targeted ("*E. Coli* direction")
4	I	V	V4I	I → V
5	I	I	I5V	I → V
9	D	D	D9E	D → E
10	D	E	E10D	D → E
13	D	E	E13D	D → E
16	V	V	V16I	V → I
23	I	V	V23I	I → V
25	V	V	V25I	V → I
38	I	I	I38V	I → V
43	D	E	E43D	D → E
44	E	E	E44D	E → D
47	D	E	E47D	D → E
48	E	E	E48D	E → D
60	I	V	V60I	I → V
61	D	D	D61E	D → E
72	I	I	I72V	I → V
75	I	I	I75V	I → V
85	E	E	E85D	E → D
86	V	V	V86I	V → I
91	V	V	V91I	V → I
101	E	E	E101D	E → D

Table S2 Effect of mutations on the unfolding energetics of variants of LBCA thioredoxin. All experiments analyzed to determine the data in this table were performed at a scan rate of 90 degrees/hour with calorimetric scans terminated at 140 ºC. Mutation effects on denaturation temperature [ΔTm=Tm(variant)-Tm("wild type" LBCA thioredoxin)] are calculated using for the "wild type" LBCA thioredoxin a value of 386.66 K (113.51 ºC) calculated as the average from the values obtained in several experiments (standard deviation 0.1 degree). Calorimetric and van´t Hoff enthalpies, as well as the corresponding van' Hoff to calorimetric enthalpy ratio, were derived from the fitting of a pseudo-two-state model (equations S1 and S2) to the calorimetric profiles. Mutation effects on unfolding free energy (ΔΔG values) were calculated from the corresponding ΔTm values using Schellman equation (ref. Shellman 1987 in the main text) with an unfolding enthalpy value for the "wild type" LBCA thioredoxin of 536 kJ/mol (standard deviation 40 kJ/mol). The effects given here correspond to the mutations as they were performed on the ancestral LBCA thioredoxin background and also (data shown between brackets) to the effects in the "*E. coli* direction".

Mutation on LBCA (mutations in the "*E. coli* direction")	Conc. (mg/mL)	ΔTm (degrees)	ΔH (kJ/mol)	ΔH^{vH} (kJ/mol)	r ($\Delta H^{vH}/\Delta H$)	ΔΔG (kJ/mol)
V4I (I4V)	0.29	0.28 (-0.28)	496	482	0.97	0.39 (-0.39)
V4I (I4V)	0.58	0.32 (-0.32)	484	496	1.03	0.44 (-0.44)
I5V	0.48	-0.09	456	528	1.16	-0.13
I5V	0.76	-0.17	499	566	1.13	-0.24
D9E	0.32	-1.17	541	543	1.00	-1.62
D9E	0.64	-1.08	507	564	1.11	-1.50
E10D (D10E)	0.46	-0.13 (0.13)	524	599	1.14	-0.18 (0.18)
E10D (D10E)	0.91	0.09 (-0.09)	537	585	1.09	0.13 (-0.13)
E13D (D13E)	0.34	-1.61 (1.61)	407	377	0.93	-2.23 (2.23)
E13D (D13E)	0.50	-1.16 (1.16)	468	357	0.76	-1.61 (1.61)
E13D (D13E)	0.88	-0.99 (0.99)	517	378	0.73	-1.37 (1.37)
V16I	0.22	-2.19	440	489	1.11	-3.04
V16I	0.44	-2.29	454	468	1.03	-3.17
V16I	0.84	-2.45	497	573	1.15	-3.40
V23I (I23V)	0.34	-4.08 (4.08)	506	546	1.08	-5.66 (5.66)
V23I (I23V)	0.60	-3.99 (3.99)	448	561	1.25	-5.53 (5.53)

V23I (I23V)	0.68	-3.89 (3.89)	500	538	1.08	-5.39 (5.39)
V25I	0.35	-3.14	495	556	1.12	-4.35
V25I	0.70	-3.86	409	326	0.80	-5.35
I38V	0.33	-1.52	518	581	1.12	-2.11
I38V	0.65	-1.51	525	577	1.10	-2.09
E43D (D43E)	0.20	-1.36 (1.36)	488	547	1.12	-1.88 (1.88)
E43D (D43E)	0.40	-1.60 (1.60)	446	527	1.18	-2.22 (2.22)
E43D (D43E)	0.80	-1.46 (1.46)	464	520	1.12	-2.02 (2.02)
E44D	0.38	-2.57	395	321	0.81	-3.56
E44D	0.76	-2.73	384	415	1.08	-3.78
E47D (D47E)	0.44	-1.94 (1.94)	485	602	1.24	-2.69 (2.69)
E47D (D47E)	0.88	-1.78 (1.78)	524	584	1.12	-2.47 (2.47)
E48D	0.33	-4.16	454	576	1.27	-5.77
E48D	0.63	-3.99	514	567	1.10	-5.53
E48D	0.66	-3.99	505	550	1.09	-5.53
V60I (I60V)	0.31	-0.53 (0.53)	650	600	0.92	-0.74 (0.74)
V60I (I60V)	0.82	-0.27 (0.27)	504	570	1.13	-0.37 (0.37)
D61E	0.43	0.46	565	606	1.07	0.64
D61E	0.63	0.55	537	622	1.16	0.76
D61E	0.87	0.55	528	605	1.15	0.76
I72V	0.34	0.07	577	565	0.98	0.10
I72V	0.68	0.11	551	585	1.06	0.15
I75V	0.44	0.95	381	317	0.83	1.32
I75V	0.62	0.97	517	624	1.21	1.35
I75V	0.88	1.28	381	368	0.97	1.77
E85D	0.42	-1.57	410	362	0.88	-2.18
E85D	0.97	-2.11	527	581	1.10	-2.93
V86I	0.27	-0.48	557	564	1.01	-0.67
V86I	0.54	-1.1	393	358	0.91	-1.53
V91I	0.34	-0.43	471	526	1.12	-0.60
V91I	0.69	-0.61	412	516	1.25	-0.85
E101D	0.30	-1.82	524	477	0.91	-2.52
E101D	0.60	-1.84	493	481	0.98	-2.55

Table S3 Degree of calorimetric reversibility for the thermal denaturation of thioredoxin variants. Experiments aimed at assessing reversibility were performed at a scan rate of 200 K/min and with a first scan terminated at 120 ºC, except when otherwise noted.

Variant	Concentration (mg/mL)	Percentage of reversibility
LBCA trx	0.21	70
LBCA trx	0.21	73
LBCA trx	0.21	86
LBCA trx	0.67	79
LBCA trx	0.67	76
LBCA trx V4I	0.23	68
LBCA trx I5V	0.21	68
LBCA trx D9E	0.25	49
LBCA trx E10D	0.30	74
LBCA trx E13D	0.23	75
LBCA trx V16I	0.22	68
LBCA trx V23I	0.27	60
LBCA trx V25I	0.23	60
LBCA trx I38V	0.26	73
LBCA trx E43D	0.22	69
LBCA trx E44D	0.25	63
LBCA trx E47D	0.29	64
LBCA trx E48D	0.26	45
LBCA trx V60I	0.27	67
LBCA trx D61E	0.17	73
LBCA trx I72V	0.27	74
LBCA trx I75V	0.22	71
LBCA trx E85D	0.19	79
LBCA trx V86I	0.18	60
LBCA trx V91I	0.23	67
LBCA trx E101D	0.23	57
E. coli trx	0.17	88[1]
E. coli trx K90L	0.25	90[1]
LGPCA trx	0.18	66[2]
LGPCA trx K90L	0.16	69[2]
LPBCA Trx	0.25	50[3]
LPBCA Trx L90K	0.22	44[3]
LBCA trx	0.20	55[3]
LBCA trx L90K	0.25	70[2]

[1] first scan terminated at 105 ºC; [2] first scan terminated at 125 ºC and [3] first scan terminated at 130 ºC.

Table S4 Effect of mutations involving E↔D exchanges at α-helix positions on the stability of LBCA thioredoxin (this work) and *E. coli* thioredoxin (refs. Godoy-Ruiz et al. 2004, 2005 and 2006 in the main text).

Mutation	helix	ΔΔG (kJ/mol) on LBCA thioredoxin	ΔΔG (kJ/mol) on *E. coli* thioredoxin
D13E	α1	1.6	1.0
D43E	α2	2.0	0.8
D44E	α2	3.7	4.1
D47E	α2	2.6	1.6
D48E	α2	5.6	5.3
D101E	α4	2.5	2.6

Table S5. Data collection and refinement statistics (values in parentheses are for the highest-resolution shell).

	LPBCA L89K
PDB identifier	**4ulx**
Data collection	
Beam line	BM30A (ESRF)
Space Group	*P 212121*
Cell dimensions	
a, b, c (Å)	35.17, 38.88, 61.04
ASU	1
Resolution (Å) *	32.79 - 2.35 (2.434 - 2.35)
R_{merge} *	0.1701 (1.069)
I/σ_I *	11.14 (2.06)
Completeness (%)*	99.95 (100.00)
Unique reflections*	3761 (366)
Multiplicity*	6.8 (7.0)
CC(1/2)	0.996 (0.707)
Refinement	
Resolution (Å)	32.79 - 2.35
R_{work}/R_{free} (%)	19.92 / 24.37
No. atoms	894
Protein	877
Ligands	1
Water	16
Average B-factors ($Å^2$)	48.4
R.m.s deviations	
Bond lengths (Å)	0.003
Bond angles (^{0})	0.97
Ramachandran (%)	
Favored	97
Outliers	0