HHS Public Access

Author manuscript

Nat Struct Mol Biol. Author manuscript; available in PMC 2011 November 01.

Published in final edited form as:

Nat Struct Mol Biol. 2011 May ; 18(5): 592–596. doi:10.1038/nsmb.2020.

Single-molecule paleoenzymology probes the chemistry of resurrected enzymes

Raul Perez-Jimenez[1,*], **Alvaro Inglés-Prieto**[2], **Zi-Ming Zhao**[3], **Inmaculada Sanchez-Romero**[2], **Jorge Alegre-Cebollada**[1], **Pallav Kosuri**[1,4], **Sergi Garcia-Manyes**[1], **T. Joseph Kappock**[5], **Masaru Tanokura**[6], **Arne Holmgren**[7], **Jose M. Sanchez-Ruiz**[2], **Eric A. Gaucher**[3,8,9], and **Julio M. Fernandez**[1,*]

[1] Department of Biological Sciences, Columbia University, New York, New York, USA

[2] Facultad de Ciencias, Dpto. de Química-Física, Universidad de Granada, Granada, Spain

[3] Georgia Institute of Technology, School of Biology, Atlanta, Georgia, USA

[4] Department of Biochemistry and Molecular Biophysics, Columbia University, New York, New York, USA

[5] Department of Biochemistry, Purdue University, West Lafayette, Indiana 47907-2063

[6] Department of Applied Biological Chemistry, Graduate School of Agricultural and Life Sciences, The University of Tokyo, Tokyo

[7] Division of Biochemistry, Department of Medical Biochemistry and Biophysics, Karolinska Institutet, SE-171 77, Stockholm, Sweden

[8] Georgia Institute of Technology, School of Chemistry, Atlanta, Georgia, USA

[9] Georgia Institute of Technology, Parker H. Petit Institute for Bioengineering and Bioscience, Atlanta, Georgia, USA

Abstract

A journey back in time is possible at the molecular level by reconstructing proteins from extinct organisms. Here we report the reconstruction, based on sequence predicted by phylogenetic analysis, of seven Precambrian thioredoxin enzymes (Trx), dating back between ~1.4 and ~4 billion years (Gyr). The reconstructed enzymes are up to 32° C more stable than modern enzymes and the oldest show significantly higher activity than extant ones at pH 5. We probed their mechanisms of reduction using single-molecule force spectroscopy. From the force-dependency of the rate of reduction of an engineered substrate, we conclude that ancient Trxs utilize chemical

*Correspondence and requests for materials should be addressed to R.P.-J. (raulpjc@biology.columbia.edu) or J.M.F (jfernandez@columbia.edu).

Author contribution. R.P.-J., J.M.S.-R., E.A.G. and J.M.F. designed the research; Z.-M.Z. and E. A.G. conducted the phylogenetic analysis and sequence reconstruction; A.I.-P and J.M.S.-R. expressed and purified the ancestral enzymes and conducted the calorimetric measurements and analysis; T. J. K. provided *A. aceti* Trx; M.T. provided *S. tokodaii* Trx; A. H. provided Human TRX; R.P.-J., I.S.-R., J.A.-C., P.K. and S.G.-M. performed AFM experiments; R.P.-J. and I.S.-R. conducted AFM data analysis; R.P.-J., E.A.G. and J.M.F wrote the paper; all authors participated in revising the manuscript.

mechanisms of reduction similar to those of modern enzymes. While Trx enzymes have maintained their reductase chemistry unchanged, they have adapted over a 4 Gyr time span to the changes in temperature and ocean acidity that characterize the evolution of the global environment from ancient to modern Earth.

Experimental paleogenetics and paleobiochemistry provide an opportunity to investigate in the laboratory the molecular history of modern organisms[1–4]. The study of resurrected proteins can also reveal valuable information regarding the adaptation of extinct forms of life to climatic, ecological and physiological alterations[5–8]. Despite numerous experimental examples of protein resurrection in the laboratory[1], the majority of the resurrected proteins have reconstructed enzymes thought to exist a few million years (Myr)[1–2,9]. Consequently, many hypotheses about ancient life remain untested, especially for time periods associated with dramatic changes in biological systems[10]. Such time traveling is largely limited by the ambiguity in the historical models used for ancestral sequence inference. For instance, uncertainties in databases, sequence alignments, failures in evolutionary theories and uncertainty in the construction of phylogenetic trees are common sources of ambiguity[1]. Nonetheless, several approaches are commonly used to overcome these limitations and some examples of reconstructed proteins from the Precambrian (4.5–0.5 Gyr ago) have been reported[1,11]. One example is the resurrection of an Elongation Factor (EF) gene family from ancient bacteria 0.5 to 3.5 Gyr old [9,12]. The ancestral EFs demonstrated that ancient bacteria lived in a hot environment and also revealed a correlation between the thermostability of ancient life and the temperature of ancient oceans as inferred from geological records.

These reconstruction and resurrection studies have paved the way to formulate interesting questions about ancient organisms and the biomolecules supporting these ancient life forms. For instance, little is known about how the chemistry of primitive enzymes arose and how the environmental conditions affected the evolution of their chemistry[13]. These questions cannot be directly answered by examining fossil records. In an effort to understand the evolution of enzymatic reactions we have reconstructed and tested thioredoxin enzymes (Trx) from extinct organisms that lived in the Precambrian.

Thioredoxins belong to a broad family of oxidoreductase enzymes ubiquitous in all living organisms[14]. The archetypical active site (CXXC) and the Trx fold are well conserved throughout evolution, indicating that Trx enzymes were likely present in primitive forms of life. In previous work, we have shown that by using single molecule force-clamp spectroscopy the chemical mechanisms of disulfide reduction by Trx enzymes can be examined in detail at the sub-Ångström scale[15–16]. Hence, the combination of single-molecule force spectroscopy and the resurrection of ancestral proteins may reveal novel insights into the reductase activity of these sulfur-based enzymes.

RESULTS

Reconstruction of ancestral Trx enzymes

Owing to the large number of extant Trxs sequences available, we have constructed a highly articulated phylogenetic tree encompassing over 200 diverse Trx sequences from the three

domains of life (Supplementary Fig. 1 and Supplementary List for GI numbers). From this tree we sampled several biologically relevant nodes for sequence reconstruction and laboratory resurrection. Divergence dates estimates were applied to nodes in the tree assuming the root of the tree lies between bacteria and the common ancestor of archaea/eukaryotes[17].

We resurrected Trx enzymes belonging to the last bacterial common ancestor (LBCA in Fig. 1), the last archaeal common ancestor (LACA) and the archaeal/eukaryotic common ancestor (AECA) (Fig. 1). These organisms are thought to have inhabited Earth 4.2-3.5 Gyr ago (Fig. 1a) after diverging from the last universal common ancestor (LUCA)[7],[17]. We also selected a node corresponding to the last eukaryotic common ancestor (LECA) that lived in the Proterozoic, ~1.60 Gyr ago. Two other internal nodes in the bacterial lineages were selected; the last common ancestor of cyanobacterial and deinococcus/thermus groups (LPBCA, representing the origins of photosynthetic bacteria) which existed ~2.50 Gyr ago and represents the origin of photosynthetic bacteria, and the last common ancestor of γ-proteobacteria, ~1.61 Gyr old (LGPCA). Finally, we also chose the last common ancestor of animals and fungi (LAFCA) that lived ~1.37 Gyr ago (Fig. 1a). While it would be interesting to observe the chemical transition between LUCA and the last common ancestors of Bacteria and Archaea/Eukaryotes, the reconstruction of Trx from LUCA cannot be inferred without making far-reaching assumptions about ancient life. For instance, the last common ancestor of bacteria represents an ancient population while the last common ancestor of archaea/eukaryotes represents another ancient population. It is practical to infer that LUCA lies somewhere along the branch connecting these two ancient populations. We cannot, however, identify the position of LUCA (i.e., the ancestral node) without an out-group. We have reason to believe LUCA existed on this branch (based on ancient gene duplication events that gave rise to paralogs still present in all modern organisms) but we do not know where on the branch LUCA existed. As such, this prevents us from inferring the sequence of LUCA (*pace* mid-point rooting).

The sequences of the ancestral Trx enzymes were reconstructed using statistical methods based on maximum likelihood[3],[9]. For a given node in the tree, we calculated posterior probability values for all 20 amino acids considering each site of the inferred sequence. These values represent the probability that a certain residue occupied a specific position in the sequence at a particular point in the phylogeny. The posterior probabilities were calculated on the basis of an amino acid replacement matrix[18]. The most probabilistic ancestral sequence (M-PAS) at a specific node was then reconstructed by assigning to each site the residue with the highest posterior probability. We determined the posterior probability distribution of the inferred amino acids across 106 sites for the selected sequences (Fig. 1b). The M-PASs of interest are summarized in Supplementary Fig. 2. The genes encoding these sequences were synthesized and the proteins were expressed and purified from *E. coli* cells.

Thermal stability of ancient Trx enzymes

As a first step toward investigating the physico-chemical properties of these resurrected enzymes, we used differential scanning calorimetry (DSC) to measure their thermal

stabilities. The denaturation temperature (T_m) can provide an idea about the temperature range in which the proteins are operative. A plot of the T_m of the resurrected enzymes against geological time is shown in Fig. 1c. It is striking that the three oldest Trxs enzymes from LBCA, AECA and LACA, show a similar T_m of ~113° C. As observed in Fig. 1c (inset), LBCA Trx maintains a highly populated native state up to ~105° C, where the thermal transition begins. By contrast, we determined a T_m for modern *E. coli* and human Trxs of 88.8 and 93.3° C, respectively. The ΔT_m between the oldest and modern Trx is ~25° C, a similar value to that determined for bacterial EF[12], which corroborates the hypothesis of the thermophilic nature of LBCA, AECA and LACA[7]. From the data in Fig. 1c we determined a paleotemperature trend yielding a decrease in the T_m of 5.8 ±1.8 K/Gyr.

It must be noted that while the thermodynamic denaturation temperatures determined for the ancestral Trxs follow a similar cooling trend than that of ancient oceans, the actual values are about 50° C higher than the ocean temperatures inferred from maximum $\delta^{18}O$ (see ref. [12]). An interesting possibility is that Trx evolution operates primarily on *kinetic* stability and this could be reflected in thermodynamic stability[19]. However, other than the obvious loss of function upon denaturation, the particular way in which the value of T_m is related to Trx enzyme fitness is still unknown.

Force-dependent chemical kinetics of disulfide reduction

It is also of great interest to examine the chemical mechanisms of disulfide bond reduction utilized by the resurrected enzymes. Given the ancient origin of the reconstructed Trx enzymes, with some of them predating the buildup of atmospheric oxygen, it seemed logical to assume that their catalytic chemistry might be closer to that of simple sulfur based molecules which utilize a straightforward collision-driven substitution nucleophilic bimolecular (S_N2) mechanism of reduction[20]. By contrast, Trx enzymes are known to utilize a complex mixture of chemical mechanisms including a critical substrate binding and rearrangement reaction that accounts for the vast increase in the efficiency of Trx over the simpler sulfur compounds that were available in early geochemistry[15–16].

Recently, we have developed a single molecule force-spectroscopy based assay that measures the effect of applying a well-controlled force to a disulfide bonded substrate, on its rate of reduction by a nucleophile. This assay readily distinguishes the simple S_N2 chemistry of nucleophiles such as hydroxide, glutathione and L-cysteine, from the more complex reduction chemistry of the Trx enzymes[15–16,21–24], making it the ideal assay to probe the chemistry of the resurrected enzymes. Fig. 2 demonstrates our approach. As a substrate, we use an engineered polyprotein made of eight repeats of the I27 immunoglobulin-like protein modified by mutating amino acids 32nd and 75th to cysteine $(I27_{G32C–A75C})_8$. The cysteines oxidize spontaneously, forming disulfide bonds that are hidden within each folded I27 domain in the chain. We pick up and stretch single polyproteins using an AFM, in solutions containing the desired nucleophile. In a typical experiment, we first apply a constant force to the polyprotein (175–185 pN, 0.2–0.3 s) which rapidly unfolds the $I27_{G32C–A75C}$ modules up to the disulfide bond. The unfolding events result in a stepwise increase in the length of the polyprotein where each module contributes with ~11 nm in length (Fig. 2a, Supplementary Fig., 3). After unfolding, every disulfide bond is exposed to the solvent. If

active Trx enzymes are present in the solution, single reduction events of ~14 nm per module can be readily observed (Fig. 2a, b; Supplementary Fig. 3, 4). Remarkably, all the ancestral enzymes that we resurrected were able to trigger staircases of reduction events (Fig 2b and Suppementary Fig. 3, 4) indicating that they were all active. In order to measure the reduction rate, we sum 15–80 reduction staircases similar to the one shown in Fig. 2b, and fit the resulting average with a single exponential. We repeat this procedure for different pulling forces (Fig. 2c). The resulting set of data measures the force-dependency of the rate of reduction of the disulfide bond (Fig. 2d).

In our previous work we showed that the chemical mechanisms of disulfide reduction can be distinguished by their sensitivity to the force applied to the substrate[16]. We demonstrated that simple thiol reducing agents showed a force-dependency where the rate always increased exponentially with the applied force[21–22]. By contrast, modern Trx enzymes showed a negative force dependency in the range of 30–200 pN[16]. This mechanism is consistent with a Michaelis-Menten binding reaction followed by a force-inhibited reorientation of the substrate disulfide bond, necessary for an S_N2 reaction to occur[15]. In a second mechanism, the rate of reduction increases exponentially at forces above 200 pN. This mechanism can be described by a simple S_N2 reaction and is found only in Trx enzymes of bacterial origin. Present in all Trxs enzymes there is a force-independent mechanism of reduction that can be ascribed to a single electron transfer reaction[16].

Surprisingly, the same three reduction mechanisms can be observed in the ancient enzymes with similar patterns to those found in extant Trxs (Fig. 2d, 3). Indeed, we can easily fit the force-dependency of the reduction rate measured from the resurrected enzymes using the three-state kinetic model that was previously used with modern Trxs[15–16] (see methods section and Supplementary Table 1). One might expect that Trx enzymes from primitive forms of life should have less-developed chemical mechanisms. For instance, one of the main factors controlling the chemistry of Trx catalysis is the geometry of the binding groove. In the case of modern bacterial-origin Trxs, the binding groove is less pronounced than in eukaryotic Trxs[16]. This structural difference is responsible for the different chemical behavior observed in eukaryotic *versus* bacterial Trxs. It seems reasonable that ancient enzymes would have had a less-structured groove making their chemistry more similar to that of simple reducing agents like L-cysteine or TCEP[22]. However, the chemistry of Trx enzymes seems to have been established very early in evolution, about 4 Gyr ago, in the same manner that it is observed today. This observation suggests that the step from simple reducing compounds to well-structured and functional enzymes occurred early in molecular evolution[10].

Nonetheless, several features of the catalytic mechanisms of some ancestral Trxs are intriguing. For instance, it is remarkable the high activity observed for AECA and LACA Trxs when the substrate is pulled at forces below 200 pN (Fig 2d and 3b). From the fitting of the reduction rate *versus* force data to the three-state kinetic model, an extrapolation to zero force yields rate constants of 30×10^5 M^{-1} s^{-1} for AECA Trx and 29×10^5 M^{-1} s^{-1} for LACA Trx. Interestingly, the extrapolation to zero force in the rest of ancestral Trxs predicts rate constants ranging from 3.7×10^5 M^{-1} s^{-1} to 6.6×10^5 M^{-1} s^{-1} (Supplementary Fig. 5). Although these latter values are within the same range to those found in extant Trx enzymes

using AFM (Fig. 3 and ref [16]) and also using bulk experiments[25], we can observe a trend in the reconstructed enzymes showing higher reduction rates at forces below 200 pN (Fig. 3). We speculate that this trend may be related to their substrate specificity. Ancient enzymes may be less substrate specific than modern ones, therefore being more efficient with generic substrates such as those used here. Another interesting feature is the small upward slope observed at low forces for LBCA Trx with a maximum at ~100 pN (Fig. 3a). Although structural information would be needed to fully address this point, it seems possible that the binding between substrate and enzyme is not optimum at zero force. A better conformation appears to be achieved by applying force to the substrate.

We also measured the activity of the ancestral enzymes using the conventional insulin assay (Supplementary Fig. 6). We found than the values of insulin precipitation rates obtained with this assay are similar than previously determined for *E. coli* Trx[26–27].

Activity of ancestral Trxs under acidic conditions

LBCA, AECA and LACA thrived in an anoxygenic environment likely rich in sulfur compounds and CO_2 whereas LPBCA, LECA, LGPCA and LAFCA lived in an global oxygenic environment[10] (Fig. 1a). The high level of CO_2 in the Hadean was partly responsible for the proposed low pH of the ancient oceans (~5.5)[28–29]. Therefore, following the hypothesis that early life lived in seawater, the natural habitat in which LBCA, AECA and LACA lived was likely to have been acidic in addition to hot. This is especially important given that the reactivity of Trx enzymes is due, in part, to the low pK_a value of the reactive Cys: ~6.5 *vs.* 8.0 for free cysteine[14]. This low pK_a value is a consequence of complex electrostatic interactions between several residues that stabilize the deprotonated form of the reactive cysteine[30]. Thus, Trx activity is highly sensitive to pH and modern enzymes would not work well at low pH because the catalytic thiol would be protonated and inactive. To examine these considerations we compared the reactivity of LACA, AECA and LBCA enzymes with the extant human and *E. coli* Trx enzymes at pH 5. The remarkable ability of the resurrected enzymes to operate in low pH environments is apparent from these experiments. We measured the force dependency of reduction for AECA, LACA and LBCA at pH 5, over the 50–150 pN force range (Fig. 4a). For AECA Trx, an extrapolation to zero force gives a reduction rate constant of 19×10^5 M^{-1} s^{-1} (Fig. 4a, solid line); similarly for LACA we estimated a rate constant of 6.2×10^5 M^{-1} s^{-1} whereas for LBCA Trx the reduction rates observed at pH 5 are strikingly similar to those measured at pH 7.2 (Fig. 4a). These are very high values similar to those measured for some modern Trx enzymes at neutral pH (ref [16]). We also compared the rate of reduction measured at 100 pN for LBCA, LACA and AECA with the rates of some modern Trx enzymes also measured at pH 5 (Fig. 4b).

Due to spontaneous precipitation of insulin at pH below 6, we used DTNB as a substrate for disulfide reduction to further verify the ability of the oldest enzymes to work at pH 5 (Fig. 4c). This analysis of reconstructed enzymes indicates that that ancient Trx enzymes were well adapted to function under acidic conditions and that Trx enzymes were able to maintain similar reduction rate constants as they evolved into more alkaline environments. A feature of the thioredoxin family of enzymes is that many of them are secreted to the extracellular

environment where most disulfide-bonded proteins are found[31–32]. From this perspective, thioredoxin enzymes are perhaps one of the few enzymes where a correlate can be established between their pH sensitivity and the environmental conditions found outside cells[31–32]. It is interesting to compare the acid tolerance of the resurrected enzymes with enzymes from extant extremophiles. For example, Trx from *Sulfolobus tokodaii* (thermophilic archaea[33]), with a melting temperature of 122° C (Supplementary Fig. 7), is active at pH 7 (0.12×10^5 M^{-1} s^{-1} at 50 pN), but does not show detectable activity at pH 5 (Fig. 4b), which is not surprising given that *Sulfolobus* regulates its cytosolic pH[34]. By contrast, Trx from *Acetobacter aceti* (acidophilic bacteria[35] that grows at pH 4) is active at pH 5 (0.6×10^5 M^{-1} s^{-1} at 100 pN), reflecting its acidic cytosol[35–36].

DISCUSSION

Our results show that the chemical mechanisms observed in modern Trx enzymes were already present in reconstructed Trxs predicted to be present in Precambrian organism. Our data support the view that ancestral Trx enzymes from LBCA, AECA and LACA that lived in the mid-to-late Hadean were highly resistant to temperature and active in relatively acidic conditions. Our results support the view that in early life Trx enzymes were present in hot environments and these environments have progressively cooled from 4 to 0.5 Gyr ago[10,12,37–38]. However, it is also possible that a much cooler early Earth was populated by psychrophiles, mesophiles and thermophiles where the latter could have been the only survivors of cataclysmic events (e.g., the late heavy bombardment or global glaciations on Early Earth[10,39]). Our findings also support the idea that many important biochemical pathways in the modern biosphere were already established by 3.5 Gyr ago[10]. For instance, metabolism is one of the most conserved cellular processes. Important pathways like energy production, sugar degradation, cofactor biosynthesis or amino acid processing are highly conserved from bacteria to human and were likely present in LUCA[40]. Evolution clearly operates at multiple levels of biological organization; however, enzymatic mechanisms accompanying adaptive changes seem to be highly conserved. The ability of enzymes to maintain specific chemical reactivity and mechanisms in disparate environments is necessary for the diversification of life. As we show here, this ability is exemplified by Trx enzymes and it could also be the case of universal proteins, e.g., ubiquitin, RNase, ATPase or other metabolic enzymes that have been maintained in nearly all organisms throughout the history of life. We propose that the experimental resurrection of ancestors of these universal proteins together with the sensitivity of single-molecule techniques can be a powerful tool towards understanding the origin and evolution of life on Earth.

METHODS

Phylogenetic analysis and ancestrtal sequence reconstruction

A total of 203 thioredoxin sequences from the three domains of life were retrieved from GenBank. A list of GI number can be found in the supplementary information (Supplementary List). Sequences were aligned using MUSCLE[41] and further corrected manually. The multiple sequence alignment is available upon request. The phylogenetic analysis was carried out by the minimum evolution distance criterion with 1000 bootstrap

replicates using PAUP* 4.0 beta (Sinauer Associates, Inc.). Ancestral sequences were reconstructed using PAML version 3.14 and incorporated the gamma distribution for variable replacement rates across sites[42]. For each site of the inferred sequences, posterior probabilities were calculated for all 20 amino acids. The amino acid residue with the highest posterior probability was then assigned at each site.

Protein Expression and Purification

Genes enconding the ancestral Trxs enzymes were synthesized and codon-optimized for expression in *E. coli* cells (Epoch Biolabs, TX). The genes were cloned into pQE80L vector (Qiagen) and transformed in *E. coli* BL21 (DE3) cells (Invitrogen). Cells were incubated overnight in LB medium at 37 °C and protein expression was induced with 1mM IPTG. Cell pellets were sonicated and the His 6-tagged proteins were loaded onto *His GraviTrap* affinity column (GE Healthcare). The purified protein was verified by SDS-PAGE. The proteins were then loaded into PD-10 desalting column (GE Healthcare) and finally dialysed against 50 mM HEPES, pH 7.0 buffer. The preparation of $(I27_{G32C-A75C})_8$ was carried out as described previously[15],[43].

DSC experiments

Thermal stabilities of ancestral and modern Trx enzymes were measured with a VP-Capillary DSC (MicroCal). Protein solutions were dialyzed into a buffer of 50 mM HEPES, pH 7. The scan speed was set to 1.5 K/min. Several buffer-buffer baselines were first obtained for proper equilibration of the calorimeter. Concentrations were 0.3–0.7 mg/mL and were determined spectrophotometrically at 280 nm using theoretical extinction coefficients and molecular weights. The experimental traces were analyzed following the two-state thermodynamic model[44].

AFM experiments

The atomic force microscope used is a custom-made design previousy described[45]. Cantilever model MLCT of silicon nitride were used (Veeco). We calibrate the cantilever using the equipartition theorem[46] giving rise to a typical spring constant of 0.02 N/m. The AFM works in the force-clamp mode with length resolution of 0.5 nm. The feedback response can reach 5 ms. The buffer used in the experiment is 10 mM HEPES, pH 7.2, 150 mM NaCl, 1 mM EDTA, 2mM NADPH. We add Trx enzymes to a desire concentration. The buffer also contains Trx reductase 50nM to keep Trx enzymes in their reduced state. *E. coli* Trx reductase works well with bacterial-origin Trx enzymes whereas eukaryotic Trx reductase works with Archaea/Eukaryote Trx enzymes; however, similar results are obtained when using DTE 200 μM to keep Trx enzymes reduced, thus demonstrating that modern Trx reductases maintain fully reduced ancestral Trx enzymes. For the experiments at pH 5 we used 20 mM sodium acetate buffer and 200 μM DTE.

To perform the experiment we deposited 3–6 μl of substrate ~0.1 mg/mL on a gold-covered coverslide. A drop of ~100 μl containing the Trx solution was then added. The force-clamp protocol consists of three pulses of force. In the first pulse the cantilever tip was pressed against the surface at 800 pN for 2s. In the second pulse the attached $(I27_{G32C-A75C})_8$ is stretched at 175–185 pN for 0.2–0.3s. The third pulse is the test force where the reduction

events are captured. This pulse is applied at different forces 30–500 pN time enough to capture all the possible reduction events.

The traces are collected and analyzed using custom-written software in Igor Pro 6.03 (Wavemetrics). The traces containing the reduction events at each force were summated, normalized and fitted with a single exponential. From the fitting we can obtain a time constant, τ, and thus the reduction rate at a given force ($r = 1/\tau$). Bootstrapping method was used to obtain the error of the reduction rates. The bootstrapping was run 1000 times for each reduction rate obtaining a distribution from where the s.e.m. can be calculated.

Thioredoxin bulk enzymatic measurements

Bulk-solvent oxidoreductase activity for ancestral thioredoxins was determined using the insulin precipitation assay as described elsewhere [26–27,43]. In order to further verify the activity of ancestral Trxs enzymes at acidic pH, we used DTNB (5,5′-dithiobis-(2-nitrobenzoic acid)) as substrate at pH 5. In this assay, Trxs enzymes were preactivated by incubation with 1 mM DTT. The reaction was initiated by adding active Trx to a final concentration of 4 μM to the cuvette containing 1mM DTNB in 20 mM sodium acetate buffer, pH 5. Change in absorbance at 412 nm due to the formation of TNB was followed during 1 min. Activity was determined from the slope $d\Delta A_{412}/dt$. A control experiment lacking Trx was registered and subtracted as baseline.

AFM data analysis

The data were fitted following a three-state kinetic model previously described[15–16]. In this model three different chemical mechanisms are taken into account. The rate constants used in the kinetic model are:

$$k_{01}=\alpha_0[\mathrm{Trx}]$$
$$k_{12}=\beta_0\exp\left(F\Delta x_{12}/k_BT\right)+\lambda_0$$
$$k_{02}=\gamma_0[\mathrm{Trx}]\exp\left(F\Delta x_{02}/k_BT\right)+\lambda_0$$
$$k_{10}=\delta_0$$

Rate constants k_{01} and k_{02} depend on Trx concentration in a linear manner. k_{12} and k_{02} exponentially depend on force. The kinetic model is solved using matrix analysis and parameters α_0, β_0, Δx_{12}, γ_0, Δx_{02}, λ and δ_0 can be obtained for each ancestral enzyme. The optimal kinetic parameters are calculated by numerical optimization using the downhill simplex method [47] (Supplementary Table 1). An extensive explanation of the different chemical mechanisms can be found in reference [16] where the origin of this chemical diversity was explained on the basis of the structural features of the binding groove.

Igor Pro software was used for data analysis. All figures were generated using Igor Pro and Adobe Illustrator CS4.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by grant from NIH to J.M.F.; grant from Spanish Ministry of Science and Innovation to J.M.S.-R.; NASA Exobiology grant to E.A.G; R.P.-J. and S.G.-M. thank Fundación Ibercaja and Fundación Caja Madrid for financial support. J.A.-C. thanks Fundación Alfonso Martín Escudero. We thank Dr. Beatriz Ibarra-Molero (University of Granada) for assistance with bulk enzymatic assays.

References

1. Benner SA, Sassi SO, Gaucher EA. Molecular paleoscience: systems biology from the past. Adv Enzymol Relat Areas Mol Biol. 2007; 75:1–132. xi. [PubMed: 17124866]
2. Thornton JW. Resurrecting ancient genes: experimental analysis of extinct molecules. Nat Rev Genet. 2004; 5:366–75. [PubMed: 15143319]
3. Liberles, DA. Ancestral sequence reconstruction. Vol. xii. Oxford University Press; Oxford; New York: 2007. p. 252
4. Hall BG. Simple and accurate estimation of ancestral protein sequences. Proc Natl Acad Sci U S A. 2006; 103:5431–6. [PubMed: 16567642]
5. Thornton JW, Need E, Crews D. Resurrecting the ancestral steroid receptor: ancient origin of estrogen signaling. Science. 2003; 301:1714–7. [PubMed: 14500980]
6. Thomson JM, et al. Resurrecting ancestral alcohol dehydrogenases from yeast. Nat Genet. 2005; 37:630–5. [PubMed: 15864308]
7. Boussau B, Blanquart S, Necsulea A, Lartillot N, Gouy M. Parallel adaptations to high temperatures in the Archaean eon. Nature. 2008; 456:942–5. [PubMed: 19037246]
8. Chang BS, Jonsson K, Kazmi MA, Donoghue MJ, Sakmar TP. Recreating a functional ancestral archosaur visual pigment. Mol Biol Evol. 2002; 19:1483–9. [PubMed: 12200476]
9. Gaucher EA, Thomson JM, Burgan MF, Benner SA. Inferring the palaeoenvironment of ancient bacteria on the basis of resurrected proteins. Nature. 2003; 425:285–8. [PubMed: 13679914]
10. Nisbet EG, Sleep NH. The habitat and nature of early life. Nature. 2001; 409:1083–91. [PubMed: 11234022]
11. Pollock, DD.; Chang, BSW. Ancestral sequence reconstruction. Liberles, DA., editor. Oxford University Press; Oxford; New York: 2007. p. 85-94.
12. Gaucher EA, Govindarajan S, Ganesh OK. Palaeotemperature trend for Precambrian life inferred from resurrected proteins. Nature. 2008; 451:704–7. [PubMed: 18256669]
13. Zalatan JG, Herschlag D. The far reaches of enzymology. Nat Chem Biol. 2009; 5:516–20. [PubMed: 19620986]
14. Holmgren A. Thioredoxin. Annu Rev Biochem. 1985; 54:237–71. [PubMed: 3896121]
15. Wiita AP, et al. Probing the chemistry of thioredoxin catalysis with force. Nature. 2007; 450:124–7. [PubMed: 17972886]
16. Perez-Jimenez R, et al. Diversity of chemical mechanisms in thioredoxin catalysis revealed by single-molecule force spectroscopy. Nat Struct Mol Biol. 16:890–6. 209. [PubMed: 19597482]
17. Hedges, SB.; Kumar, S. The Timetree of life. Oxford University Press; Oxford: 2009. p. xxip. 551
18. Yang Z, Kumar S, Nei M. A new method of inference of ancestral nucleotide and amino acid sequences. Genetics. 1995; 141:1641–50. [PubMed: 8601501]
19. Godoy-Ruiz R, et al. Natural selection for kinetic stability is a likely origin of correlations between mutational effects on protein energetics and frequencies of amino acid occurrences in sequence alignments. J Mol Biol. 2006; 362:966–78. [PubMed: 16935299]
20. Kice, JL. Nucleophilic Substitution at Different Oxidation States of Sulfur. In: Edwards, JO., editor. Progress in Inorganic Chemistry. 2007. p. 147-206.
21. Wiita AP, Ainavarapu SR, Huang HH, Fernandez JM. Force-dependent chemical kinetics of disulfide bond reduction observed with single-molecule techniques. Proc Natl Acad Sci U S A. 2006; 103:7222–7. [PubMed: 16645035]

22. Koti Ainavarapu SR, Wiita AP, Dougan L, Uggerud E, Fernandez JM. Single-molecule force spectroscopy measurements of bond elongation during a bimolecular reaction. J Am Chem Soc. 2008; 130:6479–87. [PubMed: 18433129]
23. Garcia-Manyes S, Liang J, Szoszkiewicz R, Kuo TL, Fernandez JM. Force-activated reactivity switch in a bimolecular chemical reaction. Nature Chemistry. 2009; 1:236–242.
24. Liang J, Fernandez JM. Mechanochemistry: One Bond at a Time. ACS Nano. 2009
25. Holmgren A. Reduction of disulfides by thioredoxin. Exceptional reactivity of insulin and suggested functions of thioredoxin in mechanism of hormone action. J Biol Chem. 1979; 254:9113–9. [PubMed: 39074]
26. Suarez M, et al. Using multi-objective computational design to extend protein promiscuity. Biophys Chem. 2010; 147:13–9. [PubMed: 20034725]
27. Holmgren A. Thioredoxin catalyzes the reduction of insulin disulfides by dithiothreitol and dihydrolipoamide. J Biol Chem. 1979; 254:9627–32. [PubMed: 385588]
28. Walker JCG. Possible Limits on the Composition of the Archean Ocean. Nature. 1983; 302:518–520.
29. Russell MJ, Hall AJ. The emergence of life from iron monosulphide bubbles at a submarine hydrothermal redox and pH front. J Geol Soc Lond. 1997; 154:377–402.
30. Dyson HJ, et al. Effects of buried charged groups on cysteine thiol ionization and reactivity in Escherichia coli thioredoxin: structural and functional characterization of mutants of Asp 26 and Lys 57. Biochemistry. 1997; 36:2622–36. [PubMed: 9054569]
31. Xu SZ, et al. TRPC channel activation by extracellular thioredoxin. Nature. 2008; 451:69–2. [PubMed: 18172497]
32. Windle HJ, Fox A, Ni Eidhin D, Kelleher D. The thioredoxin system of Helicobacter pylori. J Biol Chem. 2000; 275:5081–9. [PubMed: 10671551]
33. Ming H, et al. Crystal structure of thioredoxin domain of ST2123 from thermophilic archaea Sulfolobus tokodaii strain7. Proteins. 2007; 69:204–8. [PubMed: 17557330]
34. Baker-Austin C, Dopson M. Life in acid: pH homeostasis in acidophiles. Trends Microbiol. 2007; 15:165–71. [PubMed: 17331729]
35. Starks CM, Francois JA, MacArthur KM, Heard BZ, Kappock TJ. Atomic-resolution crystal structure of thioredoxin from the acidophilic bacterium Acetobacter aceti. Protein Sci. 2007; 16:92–8. [PubMed: 17192591]
36. Menzel U, Gottschalk G. The Internal-Ph of Acetobacterium-Wieringae and Acetobacter-Aceti during Growth and Production of Acetic-Acid. Archives of Microbiology. 1985; 143:47–51.
37. Knauth LP, Lowe DR. High Archean climatic temperature inferred from oxygen isotope geochemistry of cherts in the 3.5 Ga Swaziland Supergroup, South Africa. Geol Soc Am Bull. 2003; 115:566–580.
38. Schulte M. The Emergence of Life on Earth. Oceanography. 2007; 20:42–49.
39. Gogarten-Boekels M, Hilario E, Gogarten JP. The effects of heavy meteorite bombardment on the early evolution--the emergence of the three domains of life. Orig Life Evol Biosph. 1995; 25:251–64. [PubMed: 7708385]
40. Peregrin-Alvarez JM, Tsoka S, Ouzounis CA. The phylogenetic extent of metabolic enzymes and pathways. Genome Res. 2003; 13:422–7. [PubMed: 12618373]
41. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 2004; 32:1792–7. [PubMed: 15034147]
42. Yang ZH. PAML: a program package for phylogenetic analysis by maximum likelihood. Comput Appl Biosci. 1997; 13:555–556. [PubMed: 9367129]
43. Perez-Jimenez R, et al. Force-clamp spectroscopy detects residue co-evolution in enzme catalysis. J Biol Chem. 2008; 283:27121–9. [PubMed: 18687682]
44. Ibarra-Molero B, Loladze VV, Makhatadze GI, Sanchez-Ruiz JM. Thermal versus guanidine-induced unfolding of ubiquitin. An analysis in terms of the contributions from charge-charge interactions to protein stability. Biochemistry. 1999; 38:8138–49. [PubMed: 10387059]
45. Fernandez JM, Li H. Force-clamp spectroscopy monitors the folding trajectory of a single protein. Science. 2004; 303:1674–8. [PubMed: 15017000]

46. Florin EL, et al. Sensing Specific Molecular-Interactions with the Atomic-Force Microscope. Biosensors & Bioelectronics. 1995; 10:895–901.

47. Nelder JA, Mead R. A Simplex-Method for Function Minimization. Computer Journal. 1965; 7:308–313.

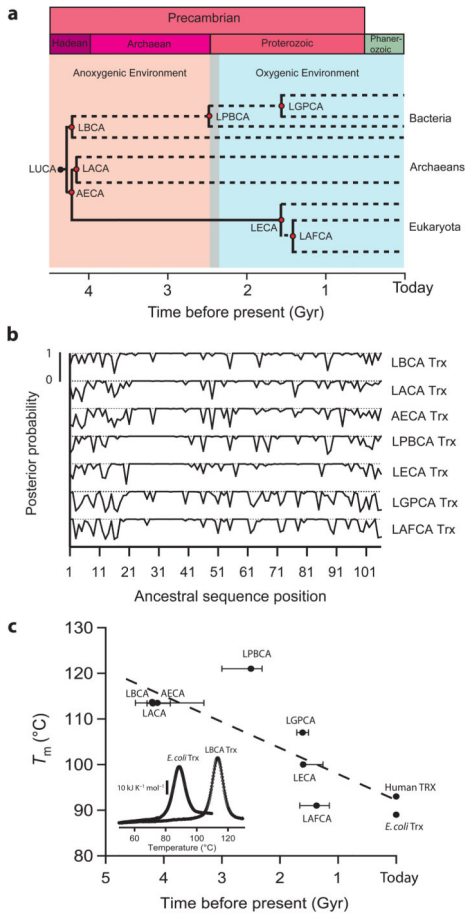

Fig. 1. Phylogenetic analysis of Trx enzymes and ancestral sequence reconstruction

(a) Schematic phylogenetic tree showing the geological time in which different extinct organisms lived, i.e., last bacterial common ancestor (LBCA); last archaeal common ancestor (LACA); archaea/eukaryota common ancestor (AECA) and last eukaryotic common ancestor (LECA). Other internal nodes are: the last common ancestor of cyanobacterial and deinococcus/thermus groups (LPBCA), the last common ancestor of γ-proteobacteria (LGPCA), and the last common ancestor of animals and fungi (LAFCA). The dashed lines represent further bifurcations. Divergence times are compiled from multiple sources and are summarized in the Timetree of Life[17]. The figure indicated the global environment; although both aerobic and anaerobic organisms are found in modern environments. **(b)** Posterior probability distribution of the inferred amino acids across 106 sites for the interested internal nodes. The inferred amino acid at each site for the interested internal node is the residue with the highest posterior probability. **(c)** Denaturation temperatures (T_m) *vs.* geological time for ancestral Trx enzymes. Modern *E. coli* and Human Trx enzymes are also indicated. The dashed line represents a lineal fit to the data. The inset shows experimental DSC thermograms for *E. coli* Trx and LBCA Trx. The instrumental uncertainty of DSC measurements is < 0.5 °C.

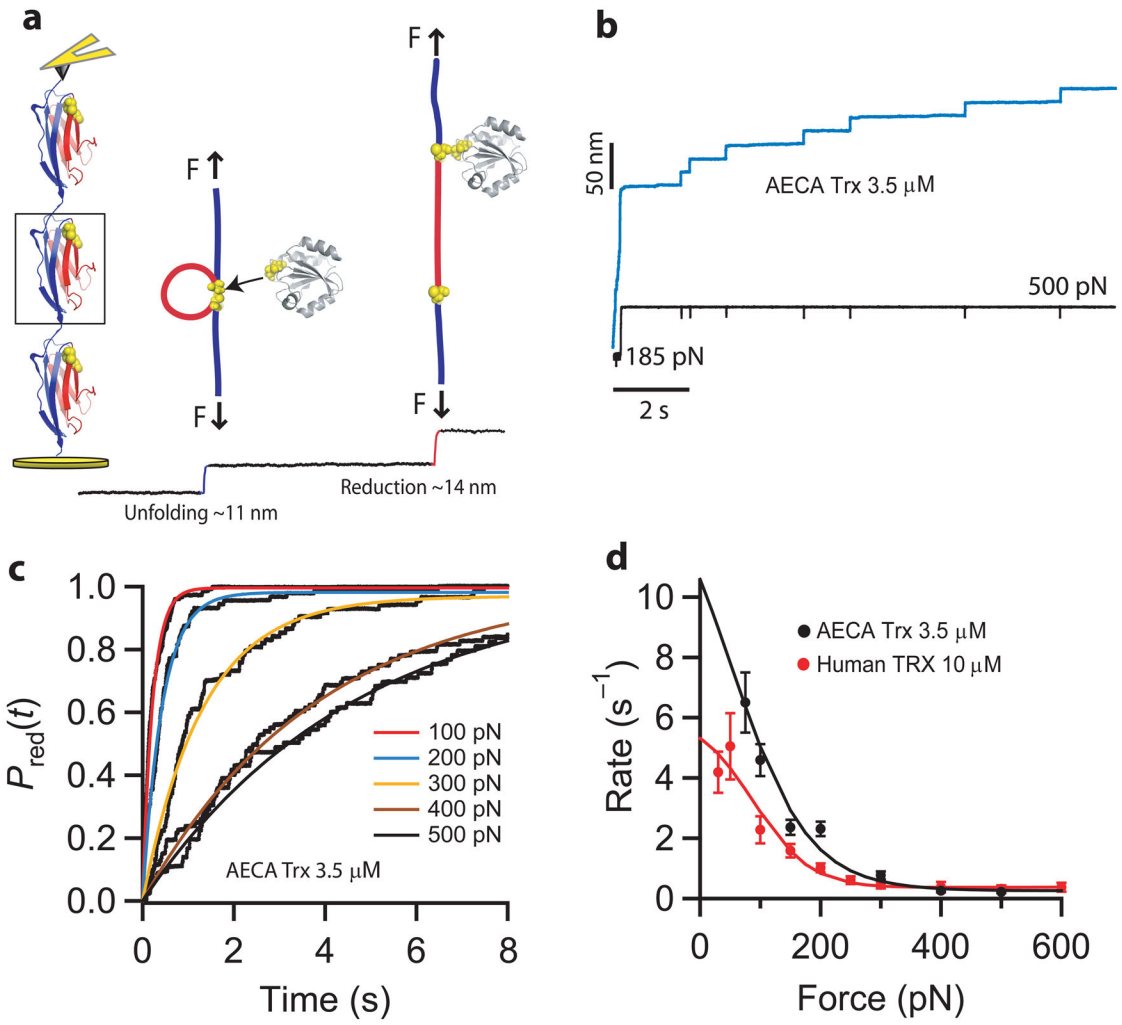

Fig. 2. Single-molecule disulfide reduction assay

(a) Schematic representation of the singe-molecule disulfide reduction assay. A first pulse of force rapidly unfolds the $I27_{G32C–A75C}$ domains (11 nm step). When the disulfide bond is exposed to the solvent a single Trx molecule can reduce it (14 nm step) **(b)** Experimental force-clamp trace showing single disulfide reductions of a $(I27_{G32C–A75C})_8$ polyprotein. The unfolding pulse was set at 185 pN for 0.2 s and the test-pulse force at 500 pN. **(c)** Probability of reduction ($P_{red}(t)$) resulted from summing and normalizing the reduction test pulse at different forces for AECA Trx (3.5 μM). **(d)** Force-dependency of disulfide reduction by AECA Trx; human Trx is also shown for comparison. Both Trx enzymes show a similar pattern: a negative force-dependency of the reduction rate, from 30–200 pN, consistent with a Michalis-Menten mechanisms and a force-independent mechanism, from 200 pN and up, described by an electron transfer reaction[16]. Notice the higher activity for AECA Trx (3.5 μM for AECA Trx *vs.* 10 μM for human TRX). The lines represent fittings to the kinetic model (see methods). The error bars are given by the s.e.m. obtained using the bootstrap method.

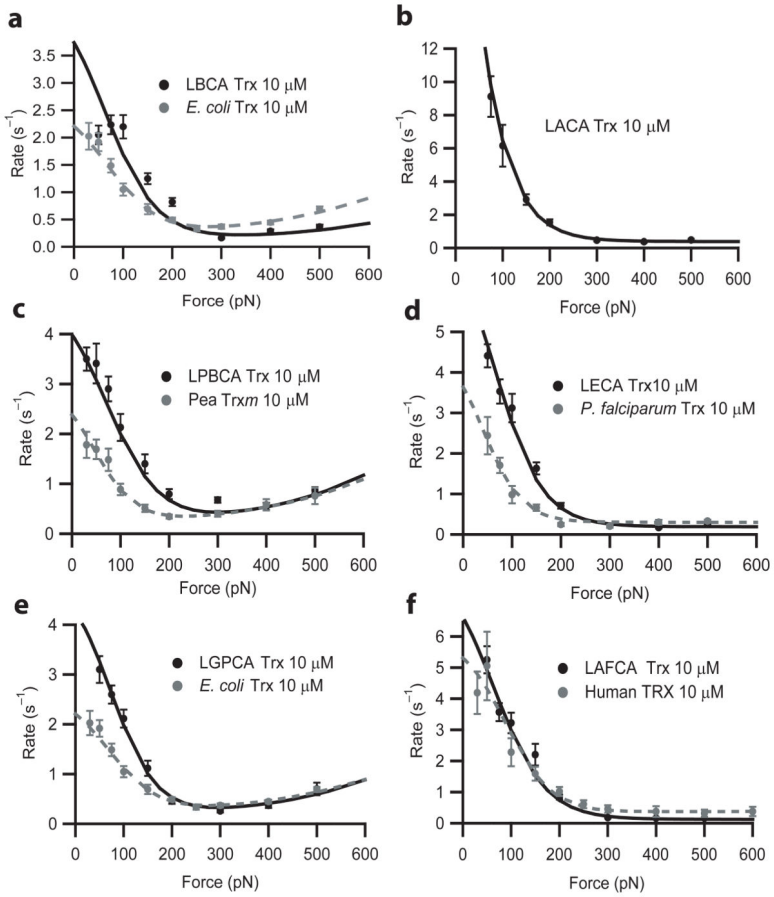

Fig. 3. Force-dependence of disulfide reduction by ancestral Trx enzymes

The reduction rate at a given force is obtained by summing, averaging and fitting to a single exponential numerous traces (15–80) like the one shown in Fig. 2b. The solid lines are fitting to the kinetic model. The grey circles and dashed lines represent the rate vs. force dependence for modern Trxs: **(c)** Pea Trx*m* from chloroplast, **(d)** *P. falciparum* Trx, **(a,e)** *E. coli* Trx and **(f)** Human Trx (all extracted from ref.[16]). These modern Trxs are descendants of the ancestral Trxs in the same plot. The error bars are given by the s.e.m. obtained using the bootstrap method.

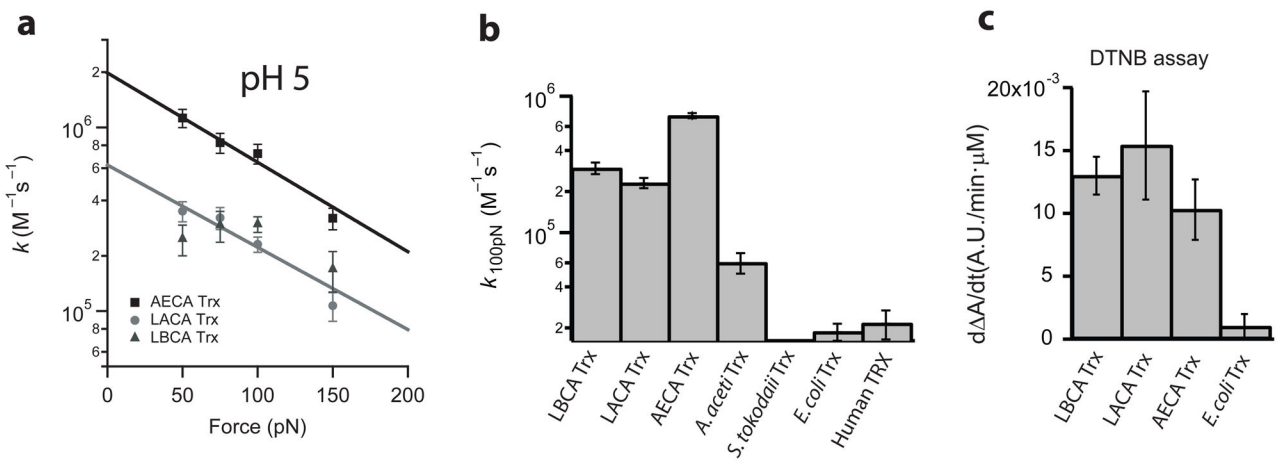

Fig. 4. Rate constants of disulfide bond reduction at pH 5

(a) A high activity for AECA (squares) and LACA (circles) Trxs can be observed at pH 5 when the substrate is pulled at low forces (50–150 pN). LBCA Trx (triangles) shows similar activity than that at pH 7.2 also with a similar trend (Fig. 3a). The solid lines are exponential fit to the experimental data. **(b)** Rate constants for disulfide reduction by ancestral Trx enzymes are considerably higher than those measured for modern *E. coli* and human Trx. While thioredoxin from the acidophile *Acetobater aceti* shows activity at pH 5, enzymes from the thermophilic *Sulfolobus tokodaii* do not show a detectable rate of reduction at the same pH. All experiments were conducted at a pulling force of 100 pN. The error bars are given by the s.e.m. obtained using the bootstrap method. **(c)** Activity of ancestral Trxs and modern *E. coli* Trx measured using DTND as substrate at pH 5 and determined by monitoring spectrophotometrically the formation of TNB at 412 nm. The error bars represent s.d. from three different measurements.