Biochimica et Biophysica Acta 1854 (2015) 127–137

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbapap

The *E. coli* thioredoxin folding mechanism: The key role of the C-terminal helix

Diego S. Vazquez [a], Ignacio E. Sánchez [b,*], Ana Garrote [a], Mauricio P. Sica [c,*], Javier Santos [a,*]

[a] Instituto de Química y Fisicoquímica Biológicas "Prof. Alejandro C. Paladini" (IQUIFIB), Universidad de Buenos Aires, Junín 956, 1113AAD Buenos Aires, Argentina
[b] Laboratorio de Fisiología de Proteínas, Departamento de Química Biológica, Facultad de Ciencias Exactas y Naturales and IQUIBICEN-CONICET, Universidad de Buenos Aires, C1428EGA Buenos Aires, Argentina
[c] CONICET e Instituto de Energía y Desarrollo Sustentable (IEDS), Centro Atómico Bariloche, Comisión Nacional de Energía Atómica (CNEA), Av. E. Bustillo km 9.500, San Carlos de Bariloche, Río Negro, Argentina

ARTICLE INFO

Article history:
Received 16 August 2014
Received in revised form 23 October 2014
Accepted 10 November 2014
Available online 22 November 2014

Keywords:
Molecular dynamics
Protein stability
Transition state ensemble
Intermediate state
Folding kinetics

ABSTRACT

In this work, the unfolding mechanism of oxidized *Escherichia coli* thioredoxin (EcTRX) was investigated experimentally and computationally. We characterized seven point mutants distributed along the C-terminal α-helix (CTH) and the preceding loop. The mutations destabilized the protein against global unfolding while leaving the native structure unchanged. Global analysis of the unfolding kinetics of all variants revealed a linear unfolding route with a high-energy on-pathway intermediate state flanked by two transition state ensembles *TSE*1 and *TSE*2. The experiments show that CTH is mainly unfolded in *TSE*1 and the intermediate and becomes structured in *TSE*2. Structure-based molecular dynamics are in agreement with these experiments and provide protein-wide structural information on transient states. In our model, EcTRX folding starts with structure formation in the β-sheet, while the protein helices coalesce later. As a whole, our results indicate that the CTH is a critical module in the folding process, restraining a heterogeneous intermediate ensemble into a biologically active native state and providing the native protein with thermodynamic and kinetic stability.

1. Introduction

Protein folding involves the coordination of a myriad of interaction events. Intra- and intermolecular contacts establish a cooperative and plastic interaction network that determines protein topology, stability, dynamics, and biological activity.

Thioredoxin (TRX) is a key protein in the archaea, bacteria, and eukaryote domains. TRX, TRX reductase and the co-enzyme NADPH are components of the TRX system that controls the global protein dithiol/disulfide balance in the cell [1,2] and are involved in a large number of biochemical processes [3]. The role of TRX as an oxidoreductase enzyme is crucial in the peroxide detoxification process, as a partner of thiol peroxidases [1]. TRX, the physiological electron donor of methionine sulfoxide reductases, acts in the reversible reduction of the oxidized methionine as well [4]. Along with the ribonucleotide reductase enzyme, both TRX and glutaredoxin are involved in the reduction of ribonucleotides to desoxiribonucleotides [3]. TRX also acts by controlling the redox state of transcription factors [5,6].

Escherichia coli TRX (EcTRX) has been extensively used in protein engineering to investigate the bases of protein stability, folding and protein function [7–17]. Full-length EcTRX is a small (108 residues) globular and monomeric protein with a characteristic α/β topology, and its structure has been already investigated by both NMR [18–21] and X-ray crystallography [22] (Fig. 1). The structural features and the knowledge gained by several research groups make EcTRX a very interesting model to try to obtain a complete picture of the protein folding process. The active site of EcTRX is characterized by the presence of the –WCGPC– catalytic motif, which is characteristic of the TRX family, and the reversible oxidation of these Cys residues forms an –S–S– bond.

Both chemical [23] and temperature [24–26] equilibrium unfolding experiments indicate that EcTRX, a thermodynamically stable protein (~8 kcal mol^{-1}, in the case of the oxidized form), behaves as a two-state folder. The EcTRX folding kinetics was also studied [27–29]. Folding traces of oxidized EcTRX show high complexity, and at least five multiple phases are necessary to describe the process (followed by Trp fluorescence). Among them, a burst phase is observed, characterized by a time constant smaller than the observation dead time (2.5 ms), which, more likely, corresponds to the initial hydrophobic collapse [30]. In addition, the slowest phase arises from the *trans* to *cis* isomerization

Abbreviations: CD, circular dichroism; CTH, C-terminal α-helix; DLS, dynamic light scattering; ESI-MS, electrospray ionization mass spectrometry; GdmCl, guanidinium chloride; HB, hydrogen bond; I^* and I^*-sim, high energy intermediate state inferred from experiments and simulations, respectively; MALS, multiangle light scattering; MDS, molecular dynamics simulation; *N*, the native state; NMR, nuclear magnetic resonance; R_S, the hydrodynamic radius; SASA, solvent accessible surface area; SEC, size exclusion chromatography; SDS-PAGE, sodium dodecyl sulfate polyacrylamide gel electrophoresis; *TSE*, the transition state ensemble; TRX, thioredoxin; EcTRX, TRX from *E. coli*; *U*, the unfolded state

* Corresponding authors. Tel.: +54 11 49648290x108; fax: +54 11 49625457x118.
E-mail addresses: nachoquique@gmail.com (I.E. Sánchez), mp.sica@gmail.com (M.P. Sica), jsantos@qb.ffyb.uba.ar (J. Santos).

http://dx.doi.org/10.1016/j.bbapap.2014.11.004
1570-9639/

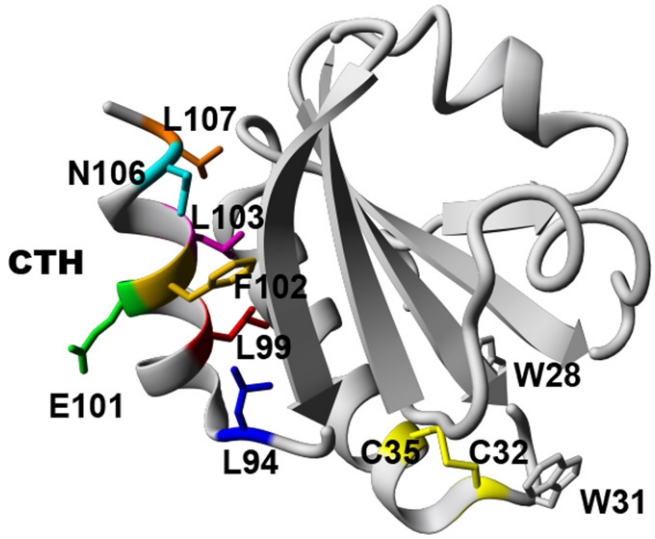

Fig. 1. Ribbon diagram of EcTRX. Residues that were mutated in this work are shown in sticks. The disulfide bridge (residues 31–35) is in yellow. In addition, the side chains of W28 and W31 are shown in gray. The model was generated with YASARA view software [84] and the PDB ID: 2TRX.

of Pro76. On the other hand, the unfolding of EcTRX may be described by a single exponential. The rate of unfolding for wild-type EcTRX in water, *i.e.*, extrapolated to zero urea concentration, was estimated as 9.5×10^{-8} s^{-1}, which yields a half-life time in water of around four months [31]. This fact supports the idea that EcTRX is also a kinetically stable protein [31].

In addition, there is invaluable thermodynamic and kinetic information for an extended list of EcTRX point mutants [13,14,26,31–33]. Godoy-Ruiz and coworkers have found, based on results from urea-induced folded/unfolded experiments, that there is a large fraction of residues that occupy unstructured regions in the EcTRX *TSE*, yielding a high energy barrier, presumably as the result of the evolution towards a highly kinetically stable conformation [31]. Remarkably, Hamid Wani and coworkers [34] showed that the unfolding of EcTRX includes an intermediate state when the reaction is performed at pH 3.0. They suggest that at pH 7.0, the energy of this intermediate state is too high for it to be detected. Thus, how the transition state ensembles (*TSE*) are organized seems to be an open question.

Experiments between different couples of EcTRX fragments point to the plasticity of the core of this protein [35–40] and to the relevance of C-terminal α-helix (CTH) on the consolidation of the native state [41, 42]. Experiments involving fragment TRX1–93 showed that the absence of the CTH (residues 94–108) results in the stabilization of a premolten globule-like state, with only a residual secondary structure and without signatures of a persistent tertiary structure. In particular, complementation between fragment TRX1–93 and peptide TRX94–108 indicated that interactions mediated by apolar residues of CTH strongly contribute to the secondary structure propensity of this secondary structure element [43] and simultaneously stabilize the tertiary structure of EcTRX native-like complex [44]. Interestingly, it was suggested that the alteration of the C-terminal region of the protein may considerably affect folding kinetics [30,45], thus indicating that this region plays a key role, not only contributing to the native state stability, but also acting on the stabilization of *TSE*.

As previously described [41,43], helix α5 consists of a set of spatially aligned buried apolar residues L99, F102, L103 and L107 that establishes contact among them (apparently stabilizing the α-helical structure of the last 15 residues of the protein, as predicted by the AGADIR algorithm [46–49], Table S1) and also with residues of helix α3 and residues from the β-sheet (tertiary interactions). In particular, L99 is highly conserved and the apolar residues Ile, Leu and Val are very frequently found in position 103 along the TRX family (L103 in the *E. coli* sequence), suggestive of the key role these residues exhibit in the EcTRX structure (a sequence logo for the CTH is shown in Fig. S1). In addition, L94, a less conserved residue, is located in the connector between CTH and the last strand, establishing contacts with L99 that could also be relevant in contributing to the native stability of EcTRX. The side chain of residue N106 establishes an intra-helical hydrogen bond (HB) with the carbonyl oxygen of F102 and one tertiary hydrogen bond between the carbonyl oxygen of N106 and the amine group of K82. Finally, E101 is close in space to the charged groups of residues K96, K100, and D104 from the CTH. To evaluate the role in the unfolding mechanism of some interactions established by key residues of the CTH (helix α5) in the native state of EcTRX, a series of point mutants was designed (Fig. 1). We chose to mutate L94, L99, F102, L103, N106 and L107 to alanine, and E101 to glycine. We expected that these mutations would destabilize the native state of EcTRX relative to the unfolded state because of the truncation of the side chains forming stabilizing interactions. Mutation E101G should further destabilize the native state because of the higher entropic cost of restricting the conformation of a glycine to the α-helical region.

In this work we investigated experimentally and computationally the role of CTH in the EcTRX unfolding mechanism. Based on the observation and analysis of non-linear rate profiles (kinks or rollovers) in the unfolding arm of EcTRX *chevron plot* [50,51], equilibrium unfolding experiments and structure-based simulation results, we suggest that the EcTRX folding/unfolding mechanism includes an on-pathway high-energy intermediate state. In addition, we characterized the involvement of CTH in the energetics of the transition-state ensembles. Finally, based on molecular dynamics simulation (MDS) results, we described the folding landscape and the conformational ensembles across the folding process.

2. Materials and methods

2.1. Subcloning, expression and purification of EcTRX mutants

EcTRX mutants were obtained by PCR-based site-directed mutagenesis. Expression and purification were carried out as previously described in Santos et al. [41,44]. Briefly, transformed *E. coli* BL21 (DE3) cells were grown in a Luria–Bertani medium at 37 °C to $OD_{600nm} = 0.9–1.0$. Finally, overexpression of each EcTRX variant (3 h, 37 °C, 250 rpm) was induced using 1.0 mM IPTG. For the protein purification, cells were disrupted in 20 mM Tris–HCl, pH 7.0, and centrifuged at 6000 rpm. The isolated fluid was loaded onto a DE52 Sepharose column equilibrated with 20 mM Tris–HCl, pH 7.0. Elution was performed with increasing concentrations of NaCl up to 1.0 M. Fractions containing EcTRX (evaluated by SDS-PAGE and UV absorption) were loaded onto a preparative Sephadex G-100 chromatography (SEC, 93 cm × 2.7 cm) column, previously equilibrated with 10 mM Tris–HCl and 100 mM NaCl, pH 7.0. Then, pure EcTRX fractions were pooled and extensively dialyzed against distilled water. Finally, the protein was lyophilized and preserved at −20 °C.

2.2. Mass spectrometry

Protein samples were analyzed by RP-HPLC–MS using a 1.0 mm × 30 mm Vydac C8 column, operating at 40 μL min^{-1}, connected to a Surveyor HPLC System on-line with an LCQ Duo (ESI ion trap) mass spectrometer (Thermo Fisher, San José, CA, USA). Samples were eluted using a 15 min gradient from 10% to 100% solvent B (solvent A: 2% acetic acid, 2% ACN; solvent B: 2% acetic acid, 96% ACN). Protein characterization was performed by full scan 300–2000 amu and deconvoluted by the XCalibur software.

2.3. Trp fluorescence spectroscopy

Fluorescence spectra were recorded in the range of 305–500 nm using an excitation wavelength of 295 nm and a 1.0 cm path length quartz cell with stirring. The spectral slit-width was set to 8 nm for both monochromators. After acquisition, each spectrum was corrected by blank subtraction. Buffer was 25 mM sodium phosphate, pH 7.0 and spectra were acquired at 25 °C.

2.4. Circular dichroism spectroscopy

Spectra were acquired using a Jasco 810 spectropolarimeter (Jasco Corporation, Japan) equipped with a peltier for temperature control. Protein samples were prepared in 25 mM sodium phosphate, pH 7.0 to a final concentration of 10 and 25 μM for far-UV and near-UV, respectively. Circular dichroism (CD) spectra were recorded in the range of 190–260 nm (using a 0.1 cm path-length cell) and 240–340 nm (using 0.5 cm or 1 cm path-length cells). Data acquisition was carried out at 25 °C and at least five spectra were acquired at a speed scan of 20 nm min^{-1} and a time constant of 1 s, and averaged. Finally, a scan of buffer was properly smoothed and subtracted from the corresponding average spectrum.

2.5. Fourth-derivative UV-absorption spectroscopy

At least 10 scans were recorded at intervals of 0.1 nm in the range of 240–340 nm and a scan speed of 40 nm min^{-1} using a Jasco V-550 UV–visible spectrophotometer (Jasco Corporation, Japan) equipped with peltier equilibrated at 25 °C. Protein samples (10.0 μM final concentration) were prepared in 25 mM sodium phosphate, pH 7.0. A scan corresponding to the buffer was acquired in the same conditions, smoothed and subtracted to the averaged spectrum of each protein sample. An *ad-hoc* Excel® spreadsheet was used to transform the raw numerical data to the corresponding fourth-derivative spectra.

2.6. Equilibrium unfolding followed by fluorescence measurements

Lyophilized proteins were dissolved in 25 mM sodium phosphate, pH 7.0 at room temperature (10 μM final concentration) and incubated overnight in increased concentrations of GdmCl in a range of 0.0 to 5.0 M. Unfolding was followed by Trp fluorescence intensity as described above. Excitation and emission were 295 and 354 nm, respectively. A two-state model was fitted to experimental data according to Santoro and Bolen [52] by the nonlinear least squares method to define the difference in free energy of unfolding ($\Delta G^0_{NU\ H2O}$) and the denaturant concentration in which the unfolded fraction of molecules is 0.5 (C_m). Global fitting for all variants, including the wild-type EcTRX, was performed using a single m_{NU} value (the dependence of the difference in free energy on denaturant concentration) using Eq. (1),

$$S = \frac{\left(S_{0,N} + m_N[GdmCl]\right) + \left(S_{0,U} + m_U[GdmCl]\right)e^{\left(\frac{\Delta G^0_{NU} + m_{NU}[GdmCl]}{RT}\right)}}{1 + e^{\left(\frac{\Delta G^0_{NU} + m_{NU}[GdmCl]}{RT}\right)}} \quad (1)$$

where $S_{0,N}$ and $S_{0,U}$ are the intrinsic fluorescence signals for the native and unfolded states, respectively; m_N and m_U are the slopes of the pre- and post-transition regions, respectively, assuming a linear dependence of S_N and S_U with denaturant concentration.

Thermal denaturation of EcTRX variants was monitored by changes in the fluorescence signal of SYPRO orange dye by heating the holder from 4 to 95 °C at a rate of 1 °C min^{-1}. The experiment was performed in a real-time PCR system (Bio-Rad). Excitation and emission ranges were 470–500 and 540–700 nm, respectively. Protein concentration was 0.16 mg mL^{-1} and the buffer was 20 mM sodium phosphate, pH 7.0.

It is worthy of note that the EcTRX I^* state described in this paper (see below) is a high-energy intermediate. This determines that the I^* state does not significantly populate in solution, in equilibrium experiments. In this case, the folding transition can be described using a two-state model; the only observable states under these conditions are the native and the unfolded states. Thermodynamics can be applied and free energy of unfolding calculated, without any assumption concerning the microscopic kinetic constants.

2.7. Fast chemical unfolding kinetics

Unfolding kinetics was recorded using a *Spectra Kinetic Monochromator* stopped-flow device (Applied Photophysics, Leatherhead, UK) equipped with a fluorescence detection system and a thermal bath. The excitation wavelength was 280 nm and emission was recorded through a 320 nm filter. Unfolding kinetics were initiated by a 10-fold dilution of protein samples (10 μM final concentration) into increasing GdmCl concentrations up to 7.8 M. The buffer was 25 mM sodium phosphate, pH 7.0. All measurements were carried out at 25 °C. In all cases, the concentration of denaturant solutions was verified by refractometry ($\Delta N_{8.007M} = 0.132$).

2.8. Analysis of kinetic data

A simple exponential equation was fitted to the unfolding time-traces performed at each denaturant concentration (Eq. (2)):

$$F_{(t)} = A \times e^{(-k_{obs}t)} + F_\infty \quad (2)$$

where A is the difference between the initial and final signal; k_{obs} is the observed rate constant and F_∞ is the fluorescence signal when time tends to infinity (equilibrium). In addition, the residuals of each fitting were plotted and compared with those obtained using more complex models like two or three exponentials plus a drift of the signal.

The unfolding kinetic data (k_{obs}) for EcTRX variants were globally fitted to a three-state model including an obligatory high-energy intermediate (Eqs. (3)–(11), [51,53]),

$$U \underset{k_{IU}}{\overset{k_{UI}}{\rightleftharpoons}} I^* \underset{k_{NI}}{\overset{k_{IN}}{\rightleftharpoons}} N. \quad (3)$$

The apparent rate constants are the solutions to Eq. (4):

$$k_{obs} = \frac{-B \pm \sqrt{B^2 - 4C}}{2} \quad (4)$$

where

$$B = -(k_{UI} + k_{IU} + k_{IN} + k_{NI}) \quad (5)$$

$$C = k_{UI}(k_{IN} + k_{NI}) + k_{IU}k_{NI}. \quad (6)$$

The value that each coefficient k_i takes varies with the denaturant concentration as follows:

$$k_{i,GdmCl} = k_i e^{m_i[GdmCl]}. \quad (7)$$

In addition, at low denaturant concentration (*TSE*1 limit), $k_{UI} \ll k_{IN}$ and given that the crossing of *TSE*2 can be treated as a pre-equilibrium, we can approximate k_f and k_u as:

$$k_f(TSE1) = k_{UI} \quad (8)$$

$$k_u(TSE1) = \frac{k_{NI}}{k_{IN}} k_{IU} \quad (9)$$

and at high denaturant concentration *TSE*2 represents the higher barrier (*TSE*2 limit), thus the corresponding equations are:

$$k_f(TSE2) = \frac{k_{UI}}{k_{IU}} k_{IN} \tag{10}$$

$$k_u(TSE2) = k_{NI}. \tag{11}$$

Eqs. (8)–(11) were used to plot lines in Fig. 4 corresponding to the *TSE*2 and *TSE*1 limits.

All activation free energies were assumed to vary linearly with GdmCl concentration, with a proportionality constant *m*. Since the intermediate is not populated, only the ratio of the rate constants for the steps leading from the intermediate (k_{IN}/k_{IU}) can be obtained from the data fitting. For the same reason, we cannot calculate values for m_{IN} and m_{IU}; in spite of this, we can fit the value for (m_{IN}–m_{IU}). For practical reasons, we did not fit the ratio directly but kept k_{IN} and its denaturant dependence fixed at a value that will not make this step rate-limiting under any conditions ($k_{IN} = 10^5$ s^{-1}, $m_{IN} = 0$ kcal mol^{-1} M^1) [51]. Individual φ (phi) values, which compare the effect of amino acid replacements introduced by site-directed mutagenesis, were calculated according to Eq. (12):

$$\varphi = \frac{\Delta G^{0,TSE_i^\dagger}_{f,\ wt} - \Delta G^{0,TSE_i^\dagger}_{f,\ mut}}{\Delta G^{0}_{Eq,\ wt} - \Delta G^{0}_{Eq,\ mut}} \tag{12}$$

where $\Delta G^{0,TSE_i^\dagger}_{f,wt}$ and $\Delta G^{0,TSE_i^\dagger}_{f,mut}$ are the Gibbs terms applied to the activation free energies of folding corresponding to wild-type and mutant EcTRX, respectively, whereas $\Delta\Delta G^0_{eq}$ is $\Delta G^0_{Eq,wt} - \Delta G^0_{Eq,mut}$, the difference between equilibrium free energy differences of folding for wild-type and mutant variant, respectively.

The proportionality constants α_{GdmCl} to quantify the energetic sensitivity of each *TSE* to changes in denaturant concentration relative to the ground states were calculated according to Eq. (13):

$$\alpha_{GdmCl} = \frac{m_f}{m_{NU}} = \left(1 - \frac{m_u}{m_{NU}}\right). \tag{13}$$

α_{GdmCl} is interpreted as the relative amount of accessible surface area buried in the *TSE*, m_{NU} is the dependence of the difference in free energy of equilibrium unfolding on denaturant concentration, m_u and m_f are defined as follows:

$$m_{f,u} = \frac{\partial \Delta G^{0,TSE_i^\dagger}_{f,\ u}}{\partial [GdmCl]}. \tag{14}$$

2.9. Oxidoreductase activity assay

TRX enzymatic activity was measured as described in Santos et al. [41]. Briefly, lyophilized protein (10 μM final concentration) was dissolved in 25 mM sodium phosphate and 1 mM DTT, pH 7.0 at room temperature and incubated for 1 h to reduce the disulfide bridge to free thiols. The reaction was initiated by the addition of DiFTC-insulin (0.1 μM final concentration) [54,55] and the fluorescence signal was recorded. Excitation and emission wavelengths were 320 and 520 nm, respectively. The bandwidth used was 8.0 nm for both excitation and emission. The apparent straight line curves (first seconds) were fitted by linear regression and the percentage ratio between the slope of each mutant and the wild-type protein was considered as the specific activity in this work.

2.10. Hydrodynamic behavior

Size-exclusion chromatography (SEC) was performed in a FPLC system equipped with UV detector (Jasco Corporation, Japan), multiangle light scattering (MALS) and dynamic light scattering (DLS) modules (Wyatt Technology). SEC experiments were carried out using a Superose 12 (Pharmacia Biotech, Sweden) equilibrated in 25 mM Tris–HCl and 100 mM NaCl, pH 7.0 at room temperature. The flow rate was 0.4 mL min^{-1} and the injection volume was 100 μL. Protein samples were prepared in the same buffer and centrifuged (16,000 rpm, 4 °C) before injection. Data processing was performed using ASTRA software [56].

2.11. Computer simulations

All-atom Gõ-like MDS with structure-based force fields [57] were carried out with the GROMACS package [58] and the empiric gromos-G43a1 force field using periodic boundary conditions on the *in vacuo* model [59]. Topologies for the EcTRX structure (PDB ID: 2TRX, [22]) were built applying the shadow contact map method using a cutoff of 1.5 nm [60]. Basically, these force fields only consider attractive interaction between pairs of heavy atoms in close proximity to the native structure and avoid the inclusion of non-direct interactions. A total of 35 simulations were performed, each one at a constant reduced temperature, spanning a range of values between 60 and 140 (60, 62, 63, 64, 66, 69, 72, 74, 75, 76, 78, 80, 82, 85, 86, 90, 92, 98, 99, 101, 102, 103, 104, 105, 106, 107, 109, 112, 114, 116, 124, 127, 131, 135 and 140 °C). This protocol ensures the whole set of simulations spans all possible scenarios for the protein, from completely folded to completely unfolded. On the other hand, the temperature intervals ensure a good superimposition between the histograms obtained from each simulation, which is necessary for a reliable WHAM analysis. To keep the temperature constant along each simulation, a Berendsen thermostat [61] was used to suppress fluctuations of the kinetic energy of the system with a relaxation constant of 1.0 ps^{-1}. Each simulation comprised 20 million steps with a 0.5 fs time step, and the coordinates were taken at every 0.5 ps. Protein covalent bonds were constrained using LINCS [62]. Bidimensional energy maps were obtained with the weighted-histogram analysis method (WHAM) [63].

The solvent-accessible surface area was calculated with a POPS online server [64] using a water probe with a 1.4 Å radius and the PDB ID: 2TRX. Intramolecular interactions were calculated with UCSF Chimera software [65] using default parameters. Only chain A (PDB ID: 2TRX) was considered. Helicity of CTH was calculated with the AGADIR online server [46–49].

3. Results

3.1. Conformation of EcTRX variants

All EcTRX variants were soluble when over-expressed in *E. coli* and purified as monomers (as judged by light scattering MALS and DLS, Table 1) in the oxidized form (disulfide bond between C32 and C35), as judged by the absence of free thiols in a protein solution (not shown). The masses of the EcTRX variants evaluated by ES-MS differ in less than 2.0 Da from the expected masses deduced from protein sequences (Table 1). Far- and near-UV CD spectra (Fig. 2A and B), UV absorption spectra and fluorescence λ_{MAX} values (Fig. 2C and D and Table 1) of the variants show native-like signatures. All mutants show more than 70% of the specific enzymic activity of EcTRX (Table 1). Additionally, we were able to obtain crystals for variants L94A, E101G, N106A and L107A (Diego S. Vazquez and Javier Santos, unpublished results, Table 1). Taken together, these results suggest that these mutations in the CTH do not produce a major disruption in neither the secondary structure nor the tertiary packing of the native state of EcTRX.

Table 1
Spectroscopic characterization of the native state of EcTRX variants.

Variant	Hydrodynamic radii[a] (Å)	Mass[b] (Da)	λ_{MAX} native-state[c]	Enzymatic activity[d] (%)	Crystallization[e]
Wild-type	18.0	11,674.1 (11,675.4)	343.6	100	Yes
L94A	18.0	11,629.7 (11,631.4)	344.3	80	Yes
L99A	17.4	11,632.1 (11,631.4)	344.2	105	No
E101G	17.7	11,602.1 (11,603.3)	344.2	175	Yes
F102A	16.2	11,595.9 (11,599.3)	342.8	80	No
L103A	18.3	11,632.1 (11,631.4)	344.1	70	No
N106A	17.9	11,630.0 (11,628.4)	344.0	108	Yes
L107A	16.8	11,630.0 (11,631.4)	344.1	95	Yes

[a] The oligomeric state of each EcTRX variant was determined by multiangle light scattering (MALS) and the hydrodynamic radii determined by light scattering (DLS) measurements. The error was lower than 5%.
[b] The molecular mass measured by ESI-MS and the expected mass according to the sequence between parentheses are both shown.
[c] The maximal wavelength emission (λ_{MAX}) in fluorescence spectra are in nm. For all variants in the unfolded state, on average, $\lambda_{MAX} = 348.4$ nm.
[d] Specific enzymatic activity using DiFTC-insulin as substrate [54,55].
[e] EcTRX variants were subjected to a crystallization trial using a standard battery of conditions. Experiments were carried out at different temperatures (4, 16 and 22 °C). In particular, sodium acetate was screened in the range of pH 4–6 and ethanol, 2-methyl 2,4-pentanediol (MPD) were tested as precipitants (Diego S. Vazquez and Javier Santos, unpublished results).

3.2. Thermodynamic stability of EcTRX variants

To quantify the effect of each mutation on stability, equilibrium GdmCl-induced unfolding experiments were performed (Fig. 3). Unfolding was followed by Trp fluorescence and data were fitted to a two-state model ($N \leftrightarrow U$). The reversibility of the process was corroborated by dilution of high [GdmCl] samples. The m_{NU} values for the EcTRX variants were similar to each other and to the value obtained for the wild-type enzyme, suggesting that unfolding involves similar changes in an accessible surface area (ΔASA_{NU}). We performed a global fit to the data for all variants using a common m_{NU} value. The results are shown in Table 2. As judged by the differences in free energy, each EcTRX variant has an unfolded fraction $\leq 10^{-4}$ in the absence of GdmCl (Figs. 3 and S2). The change in stability upon mutation varied between a minimal destabilization for N106 ($\Delta\Delta G^0_{NU} = 0.3$ kcal mol^{-1}), and a large destabilization for L103A ($\Delta\Delta G^0_{NU} = 3.9$ kcal mol^{-1}, Table 2).

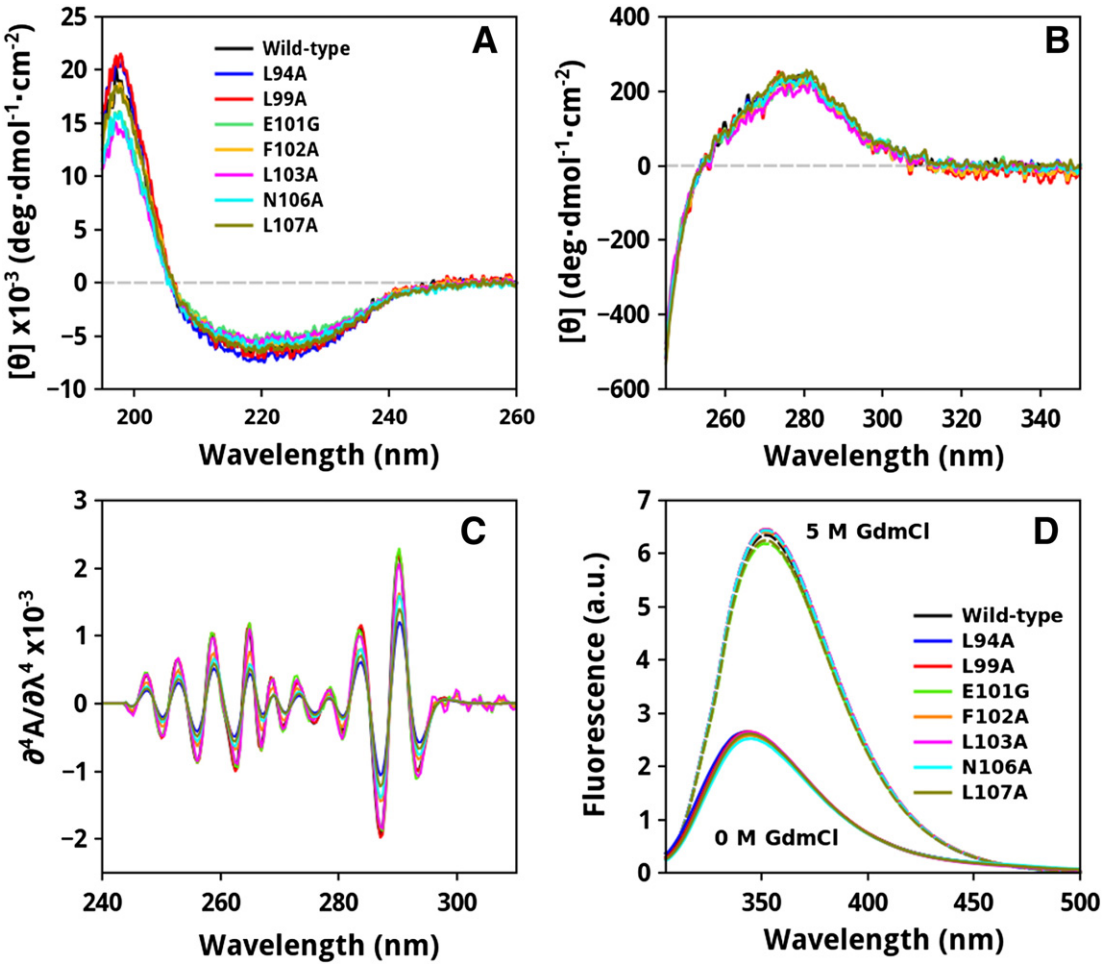

Fig. 2. CD spectra of wild-type EcTRX and mutants. Far- (A) and near-UV CD (B) spectra corresponding to the wild-type (black) L94A (blue), L99A (red), E101G (green), F102A (orange), L103A (magenta), N106A (cyan) and L107A (brown) are shown. The buffer was 25 mM sodium phosphate, pH 7.0 and spectra were acquired at 25 °C. Fourth derivative UV-absorption (C) and Trp fluorescence spectra (D) of EcTRX variants corresponding to wild-type (black) L94A (blue), L99A (red), E101G (green), F102A (orange), L103A (magenta), N106A (cyan) and L107A (brown) are shown. The buffer was 25 mM sodium phosphate, pH 7.0. For Trp fluorescence, the excitation wavelength was 295 nm. To unfold the proteins, 5.0 M GdmCl was added (dashed lines). In all cases, samples were thermostatized at 25 °C.

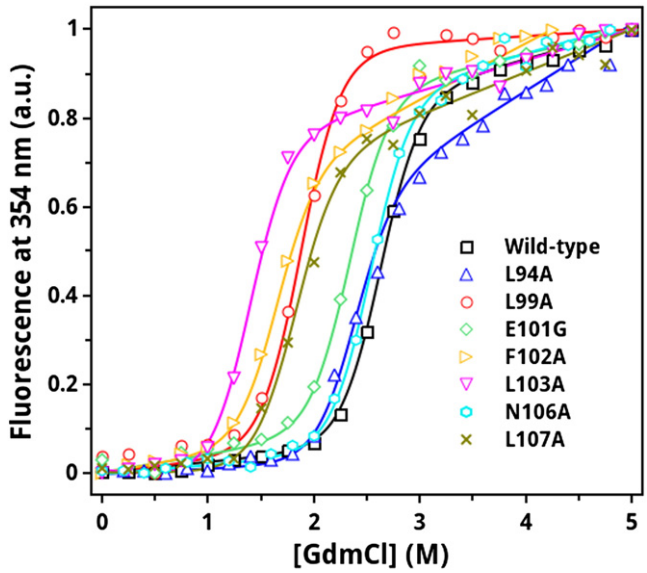

Fig. 3. Equilibrium unfolding experiments followed by Trp fluorescence. The buffer was 25 mM sodium phosphate, pH 7.0. The lines represent the best fit of a two-state model ($N \leftrightarrow U$) to the experimental data (symbols). Wild-type (□), L94A (Δ), L99A (○), E101G (◊), F102A (►), L103A (▼), N106A (●) and L107A (×). The excitation and emission wavelengths were 295 and 354 nm, respectively. In all cases, samples were thermostatized at 25 °C.

Thermal unfolding experiments followed by SYPRO-orange dye fluorescence were also performed. The changes in the T_m range between a small decrease of 1.8 ± 1.2° for the N106A variant and a large decrease of 17.3 ± 0.8° for the L103A variant. The T_m values show a good correlation with the results from the chemical unfolding experiments (Fig. S3). The crystallization experiments failed for the L99A, F102A and L103A variants, also in agreement with the chemical unfolding data. We concluded that the designed mutations noticeably altered the stability of EcTRX against global unfolding and, therefore, are useful to probe the folding mechanism of the protein.

3.3. Unfolding kinetics of EcTRX variants

We studied the unfolding kinetics of wild-type EcTRX and its variants using stopped-flow fluorescence in the presence of variable concentrations of GdmCl (excitation at 280 nm, emission collected above 320 nm) (Figs. 4 and S4).

The unfolding kinetics of five variants, including the wild-type protein, L99A, E101G, N106A and L107A were well described by mono- and biexponential curves at high and low GdmCl concentrations. The slowest phase (λ_2) accounts for <10% of the observed amplitude, and was not observed for variants L94A, F102A, and L103A. λ_2 did not vary neither with GdmCl concentration nor with protein mutation. Both the GdmCl-independence and the magnitude of λ_2 are compatible with a phase limited by the *cis/trans* isomerization [66–68] of one or several of the Pro residues of EcTRX, which in the native state has four *trans* and one *cis* X-Pro peptide bonds (Fig. S5 and Table S2). In the following, we will analyze only the fastest, isomerization independent phase λ_1.

The fluorescence signal at the end of the kinetics parallels the equilibrium curve for all variants, indicating that unfolding was followed until equilibrium (Fig. 4A, inset). The fluorescence signal at the beginning of the recorded kinetics follows the extrapolation of the native baseline for all variants, indicating that there are no unfolding reactions taking place in the 3 ms mixing time of the instrument (Fig. 4A, inset).

The logarithm of the observed rate constants for EcTRX unfolding shows a biphasic behavior for all variants (Figs. 4, panels A, B and C, and S5). The ln (k_{obs}) varies linearly with the GdmCl concentration between the denaturation midpoint and approximately 5 M GdmCl. A rollover is observed at higher GdmCl concentrations, and for some variants a second linear region is observed at the highest GdmCl concentrations. This behavior is compatible with a three-state linear mechanism for EcTRX unfolding, with an on-pathway high energy intermediate I^* flanked by two consecutive transition state ensembles *TSE*1 (between U and I^*) and *TSE*2 (between I^* and N) $U \leftrightarrow I^* \leftrightarrow N$ (Fig. 5A). At low denaturant concentrations, the unfolding rate constant is limited by *TSE*1, and at high denaturant concentrations, the unfolding rate constant is limited by *TSE*2 (Figs. 4A and 5A) [51]. We fitted such a three-state model to the data for all variants [50], using global *m*-values for the dependence of the free energies of activation with denaturant concentration. The results of the fit are shown in Table 3. Given that GdmCl efficiently screens electrostatic interactions, it might produce a differential effect on the stability of each *TSE*. However, the effect of GdmCl on the screening of electrostatic interactions is expected in the range of 0–1.5 M GdmCl [69], whereas the rollover is observed at significantly higher concentrations of denaturant (higher than 4.5–5 M). In addition, as GdmCl is a stronger denaturant than urea, the use of the former enables us to study a larger range of denaturant concentrations than in urea.

We can use the data in Table 3 to characterize the molecular interactions present in the transition state ensembles for the unfolding of EcTRX. First, we considered *m*-values as a proxy for changes in solvent exposure upon unfolding [70]. From the equilibrium *m*-value for global unfolding and the kinetic *m*-values (Table 3), we calculated the α_{GdmCl} values for *TSE*1 and *TSE*2 ($\alpha_{GdmCl} = 1 - m_u / m_{Eq}$, Fig. 5C). α_{GdmCl} describes the interaction of the different states along the folding coordinate with the denaturant GdmCl. α_{GdmCl} can range between 0 (unfolded-like) and 1 (native-like) and usually assumes values between 0.6 and 1 [71]. The α_{GdmCl} values for *TSE*1 and *TSE*2 are 0.40 ± 0.05 and 0.85 ± 0.05. We interpreted that *TSE*2 is considerably more native-like than *TSE*1.

Second, we used the unfolding dynamics of the designed EcTRX variants to extract information about the energetics of the CTH in *TSE*1 and *TSE*2. Fig. 5B shows the structure-induced rate-equilibrium free energy relationships [71] that compare the effect of mutations on the stability of *TSE*1 and *TSE*2 relative to the unfolded state (kinetic experiments) with the effect these mutations have on the stability of the native state relative to the unfolded state (equilibrium experiments) [71]. Taking into account the magnitude of the errors of the individual φ-values (Table 3), and given the small $\Delta\Delta G^0_{NU}$ for mutant N106A, we decided to carry out the analysis considering the CTH structural element as a whole, even when this approach means the loss of resolution at the residue level. The slopes in the plot (global φ-values for the CTH, Fig. 5B) can essentially assume values between 0 (unfolded-like energetics) and 1 (native-like energetics). All variants behave as a group for both *TSE*1 and *TSE*2, indicating that the CTH behaves as a single cooperative unit along the EcTRX unfolding pathway (Table 3). The slopes are

Table 2
GdmCl and temperature-induced unfolding parameters for EcTRX variants.

Variant	$\Delta G^0_{NU\ H2O}$ (kcal/mol)	C_m (M)	T_m (°C)
Wild-type	8.1 ± 0.3	2.6 ± 0.2	87.7 ± 0.7
L94A	7.4 ± 0.3	2.4 ± 0.2	83.3 ± 0.6
L99A	5.8 ± 0.2	1.9 ± 0.1	76.1 ± 0.7
E101G	7.2 ± 0.3	2.3 ± 0.2	81.6 ± 1.5
F102A	5.1 ± 0.3	1.7 ± 0.2	77.6 ± 1.0
L103A	4.2 ± 0.2	1.4 ± 0.1	70.4 ± 0.4
N106A	7.8 ± 0.3	2.5 ± 0.2	85.9 ± 1.0
L107A	5.8 ± 0.3	1.9 ± 0.2	78.5 ± 1.6

The parameters were calculated by non-linear least square global fitting of the data shown in Fig. 3 using a global m_{NU} for all variants, as described in Materials and methods. $m_{NU} = 3.09 \pm 0.12$ kcal mol^{-1} M^{-1}. The units of C_m and free energy differences are M and kcal mol^{-1}, respectively. GdmCl equilibrium unfolding experiments were performed at 25 °C. Temperature unfolding was followed by SYPRO-orange dye fluorescence increments.

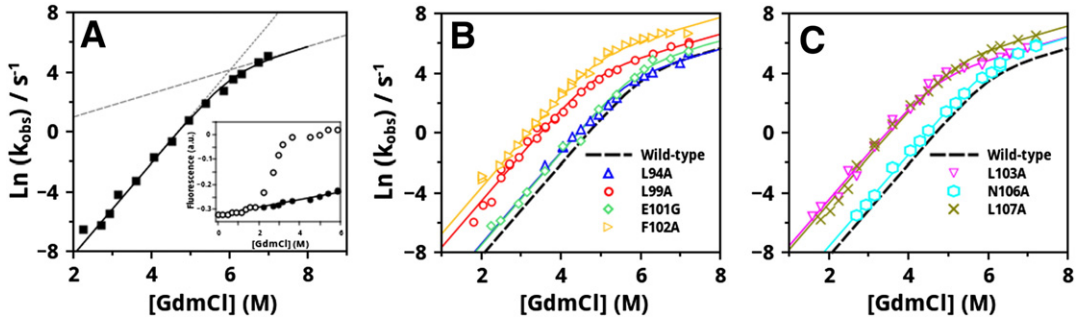

Fig. 4. Unfolding kinetics of EcTRX variants. (A) Unfolding branch of wild-type EcTRX is shown. The black line represents the fitting of a three-state model to the data as described in Materials and methods. The dashed and dotted lines correspond to the *TSE*2 and *TSE*1 limits. Inset: the initial (●) and the final (○) fluorescence signals observed for the unfolding reactions of wild-type are shown. Denaturant dependence of ln (k_{obs}) for EcTRX mutants, (B) L94A (Δ), L99A (○), E101G (◊) and F102A (►) and (C) for L103A (▼), N106A (⬢) and L107A (×). The fitting corresponding to the wild-type EcTRX is shown in dashed lines.

0.29 ± 0.14 for *TSE*1 and 0.82 ± 0.11 for *TSE*2, indicating that the CTH is unfolded-like in *TSE*1 and native-like in *TSE*2.

Given that I^* is a high-energy state under these experimental conditions, and consequently, it does not populate, we did not gain many direct structural information about the conformational I^* ensemble, neither from the unfolding kinetic nor from equilibrium unfolding experiments. Nevertheless, partial compaction of the protein and the absence of a stabilized C-terminal α-helix in I^* suggested by experiments liken I^* to the molten globule state in which conformation is less rigid and it undergoes structural fluctuations that are larger in amplitude than the native state, including local unfolding describing unlocked states with increased conformational entropy.

3.4. Computer simulations of EcTRX folding

Since I^* does not populate in our unfolding experiments to a significant degree, we were unable to obtain information about the effect of mutations on I^* stability. To answer this question, we performed structure-based molecular dynamics simulations of wild-type EcTRX. We modeled folding on a perfectly funneled energy landscape using a force field that only considers attractive interactions between pairs of atoms in contact in the native structure. Several simulations were run at different temperatures with the protein in its oxidized form (see Materials and methods). Energy landscapes were obtained by WHAM analysis, with the number of native interactions *Q* as the reaction coordinate.

The energy landscape shows three basins corresponding to a native-like state, an unfolded state and a high-energy intermediate I^*-sim, separated by two transition state ensembles (*TSE*1-sim and *TSE*2-sim) (Fig. 5D). This reaction route is in agreement with our modeling of the experimental results (Fig. 5A). It is worthy of note that non-native contacts are not considered by the model implemented herein. This fact defines a funneled energy landscape, without frustration at the level of the interactions, in which only native contacts will promote

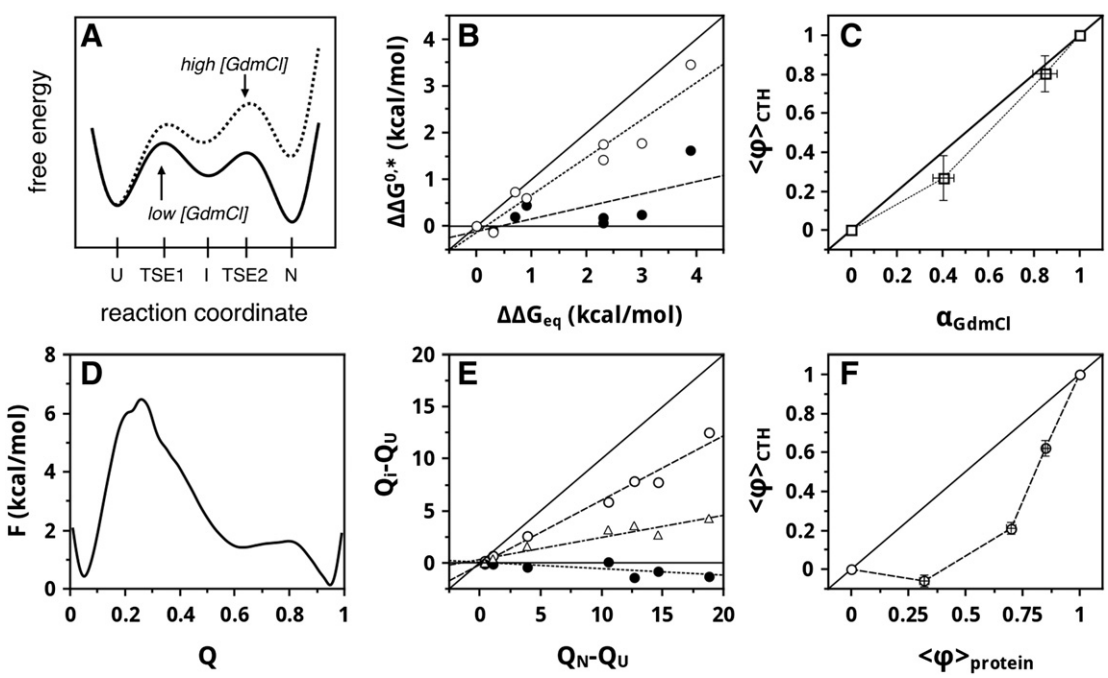

Fig. 5. Structure-induced Leffler Plot for CTH of EcTRX. (A) Schema of the reaction coordinate proposed in this paper, where *U*, I^*, *N*, are the unfolded, the high-energy intermediate and the native states, respectively. *TSE*1 and *TSE*2 are the transition state ensembles between *U* and I^*, and between I^* and *N*, respectively. (B) Relation between kinetics and equilibrium free energies. Dotted and dashed lines represent linear fits to the data and the slope of each curve corresponds to the average φ value (<φ>) of *TSE*1 and *TSE*2, respectively. (C) Relations between <φ> and α_{GdmCl} for each transition state. (D) The energy landscape for TRX obtained by WHAM analysis of MDS. The reaction coordinate (*Q*) is the fraction of native interactions between atoms. (E) The difference between the average number of contacts formed by residues L94, L99, E101, F102, L103, N106 or L107 in each ensemble (*TSE*1 (●), I^* (Δ) and *TSE*2 (○)) and the average number of contacts formed by each residue in the unfolded state (Q_i–Q_U) are plotted against (Q_N–Q_U), the difference between the number of contacts in the native and unfolded ensembles. (F) The average phi value for the residues mentioned above ($<\varphi>_{CTH}$, the slope of the linear dependences in E) corresponding to each ensemble is plotted against the <φ> for the whole protein.

Table 3
Folding kinetics of EcTRX variants.

Variant	k_{IU} (s^{-1})	k_{NI} (s^{-1})	$\Delta\Delta G_{TSE1}$ (kcal/mol)	$\Delta\Delta G_{TSE2}$ (kcal/mol)	φ_{TSE1}	φ_{TSE2}
Wild-type	0.10 ± 0.05	0.55 ± 0.31	0.00 ± 0.76	0.00 ± 0.63	–	–
L94A	0.24 ± 0.12	0.51 ± 0.27	0.20 ± 0.75	0.74 ± 0.62	n.d.	n.d.
L99A	1.58 ± 0.72	1.40 ± 0.69	0.08 ± 0.7	1.75 ± 0.57	0.04 ± 0.30	0.76 ± 0.28
E101G	0.13 ± 0.06	0.90 ± 0.49	0.45 ± 0.76	0.61 ± 0.63	n.d.	n.d.
F102A	1.26 ± 0.56	4.31 ± 2.08	0.25 ± 0.73	1.78 ± 0.61	0.08 ± 0.24	0.59 ± 0.22
L103A	2.05 ± 0.93	1.18 ± 0.57	1.62 ± 0.69	3.45 ± 0.57	0.42 ± 0.18	0.88 ± 0.17
N106A	0.09 ± 0.04	1.13 ± 0.62	−0.08 ± 0.76	−0.13 ± 0.63	n.d.	n.d.
L107A	0.78 ± 0.36	2.40 ± 1.20	0.18 ± 0.74	1.43 ± 0.61	0.08 ± 0.32	0.62 ± 0.29

The parameters were calculated by non-linear least square global fitting of the data shown in Fig. 4 to a linear three-state model for irreversible unfolding with an on-pathway high-energy intermediate using a global m_{IU} and m_{NI} for all variants, as described in Materials and methods [53]. The fit yields values $m_{IU} = 2.30 \pm 0.06$ M^{-1} and $m_{NI} = 0.79 \pm 0.08$ M^{-1}. Since the intermediate does not populate significantly during unfolding, its free energy is not accessible and only the difference in free energy between *TSE*1 and *TSE*2 can be determined. Thus, k_{IN} and m_{IN} were fixed to values of 10^5 s^{-1} and 0 M^{-1} [51]. Only φ-values for which the error propagated from the data is 0.32 or lower are reported.

the folding reaction. However, barriers and intermediates along the folding pathway accounting for the presence of topological frustration can be observed. In other words, even though the intermediate state is stabilized by interactions present in the native state, it may have a non-native topology. For instance, during the folding process, certain partially folded structures cannot access the fully folded state because it has a set of native contacts which do not allow the formation of the remainder as a result of geometrical constraints [72].

*TSE*1-sim and *TSE*2-sim feature about 30% and 80% of native contacts (Fig. 5D), again in qualitative agreement with our experimental results (Fig. 5A). The basin for I^*-sim presents about 60% of native contacts.

Next, we evaluated the behavior of the CTH along the simulated folding/unfolding route and compared it with the experiment. Subsequently, we took the difference in the number of contacts that a mutated residue makes in *N* and *U* (Q_N–Q_U) as a proxy for the effect of mutations on the stability of the native state, relative to the unfolded state $\Delta\Delta G^0_{NU}$ [73]. Similarly, we took the difference in the number of contacts that a mutated residue makes in a transition state ensemble and *U* (Q_{TSE}–Q_U) as a proxy for the effect of mutations on the stability of *TSE*1 and *TSE*2 relative to the unfolded state. Fig. 5E shows the structure-induced rate-equilibrium free energy relationships [71] corresponding to the simulation for the CTH residues probed in the experiment. All perturbations behave as a group for both *TSE*1 and *TSE*2, indicating that the CTH behaves as a single cooperative unit along the EcTRX folding simulation. The slopes are −0.06 ± 0.03 for *TSE*1, and 0.62 ± 0.04 for *TSE*2 (Fig. 5F), indicating that the CTH is unfolded-like in *TSE*1 and native-like in *TSE*2 and in agreement with the experiment (Fig. 5B). The slopes for the CTH in the experiment are higher than in the simulation. This discrepancy may be due to transient non-native interactions that are not included in our native-centric model [74].

We performed the same calculations for all EcTRX residues in order to evaluate the global degree of folding along the simulated route. We obtained an average φ-value of 0.32 ± 0.02 for *TSE*1 and 0.85 ± 0.01 for *TSE*2 (Fig. 5F), which are in agreement with the assessment of solvent exposure given by the α_{GdmCl} values of 0.40 ± 0.05 for *TSE*1 and 0.85 ± 0.05 for *TSE*2 (Fig. 5C). In both the experiment and simulations, structure formation in the CTH lags behind the averaged structure formation in the protein as a whole.

We interpreted that our simplified model for the folding of EcTRX reproduces the experimental results. In the next section, we will analyze the simulations to gather information that was not accessible in our experiments.

3.5. TRX folding on an energy landscape

The fractional number of native contacts per residue (Q_i) at different values of the reaction coordinate *Q* is shown in Fig. 6A (bottom). In *TSE*1-sim ($Q_{protein} = 0.27$), the first four strands of the β-sheet are partially folded and the rest of the protein, including the CTH, is mainly unfolded. On the other hand, the ensemble I^*-sim ($Q_{protein} = 0.6$) is more structured in the β-sheet and, additionally, helix α1 is folded. Helix α3 and the CTH show a low degree of folding in I^*-sim. All regions that show some degree of structure in I^*-sim are more structured in *TSE*2-sim ($Q_{protein} = 0.8$). Moreover, helix α3 and the CTH are almost folded, whereas the fifth β-strand is partially folded in this ensemble. Thus, the less structured ensemble I^*-sim exhibits a native-like structure only in the first four strands of the β-sheet. Residues in the strands β2–4 present the highest values of contacts per residue in the native structure (Fig. 6A), suggesting that this might be one of the reasons for the formation of the β-sheet in this ensemble.

The CTH is unfolded in *U* and *TSE*1-sim (Fig. 6B) and is partially folded in the high-energy intermediate (the average φ-sim for I^*-sim state for residues studied in this work is 0.23 ± 0.03) and *TSE*2-sim (Figs. 6C and S6). In the latter ensemble, the structure in the CTH is evidenced at residues L94, K96, L99, K100, L103 and L107, which are also the CTH residues with the highest values of contacts per residue in the native structure (Fig. 6A).

These contacts mainly involve other CTH residues and, to a lesser degree, residues in helix α3 and strands β2–5. We have also determined the propensity of the CTH sequence segment to be in a helical conformation as a function of $Q_{protein}$ using DSSP (Fig. 6D) [75]. The helical propensity of the CTH sequence segment is low in *U*-sim, *TSE*1-sim, *I*-sim and *TSE*2-sim. Our MD results suggest that two conformations coexist in *N*-sim. Both present most native contacts but in one of them the CTH residues are not in a helical conformation. More experiments will be carried out to explore the significance of this result and the existence of this heterogeneity in the native ensemble.

4. Discussion

The folding dynamics of EcTRX under different conditions have been interpreted in terms of a two-state process in both equilibrium and unfolding kinetics [31], a two-state process in equilibrium and complex refolding kinetics with populated intermediates [76], a three-state process in equilibrium and complex unfolding kinetics with populated intermediates [34] and a two-state process in equilibrium and three-state unfolding kinetics with a high-energy intermediate (this work). We will now discuss whether these apparent differences reflect the behavior of a single underlying mechanism under different experimental conditions.

Sanchez-Ruiz and coworkers reported a single *TSE* with an alpha-value of 0.47 and unfolded-like side chain energetics [31]. This is partially compatible with our description of *TSE*1-sim, which has an alpha-value of 0.4 and an average φ-value of 0.4 for the residues mutated in [31] (data not shown). Udgaonkar and coworkers detected a populated equilibrium intermediate for EcTRX at pH 3, which they proposed is higher in free energy than both *U* and *N* at pH 7 [34]. Their work also describes two populated kinetic unfolding intermediates. We speculated that the high-energy intermediate in our model for EcTRX unfolding may be one of the intermediates reported by [34] or a mixture thereof. Chafotte

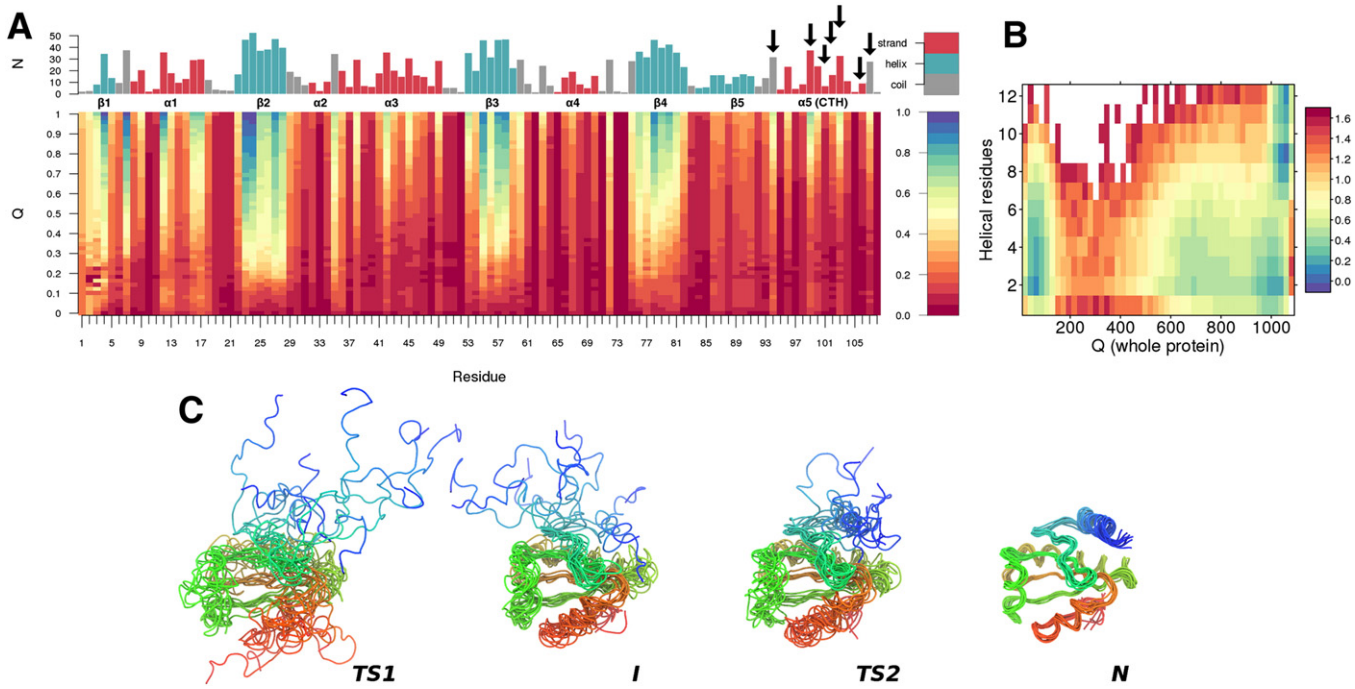

Fig. 6. Molecular dynamics characterization of the EcTRX folding mechanism. (A) Top panel: number of contacts per residue in the native state. The bar colors indicate the type of secondary structure. Black arrows indicate the residues that were mutated in this work. Bottom panel: fractional number of native contacts per residue (Q_i) with respect to Q of the whole protein. The color scale indicates the value of Q_i. (B) Unfolding of the CTH in the I^* state described by a bidimensional energy landscape using the number of residues in an alpha-helical conformation in the stretch 94–108 as the second coordinate. The color scale corresponds to the energy scale. (C) Representative conformations of the ensembles corresponding to *TSE*1, I^*, *TSE*2 and *N* states visited during folding dynamics. The structures were selected according to the number of native contacts corresponding to the transition states of the energy landscape shown in Fig. 5D.

and coworkers reported complex refolding kinetics, which are related to the isomerization of Xaa-Pro peptide bonds. The unfolding kinetics described in [31,34] and in this work started from native protein and cannot be directly compared with refolding kinetics starting from a mixture of fast- and slow-folding unfolded state isomers. However, Chafotte and coworkers have reported a rapidly formed intermediate that is unrelated to peptide bond isomerization [76]. Perhaps the high-energy intermediate that causes the unfolding rollover in our experiments is the same species as the one found by Chafotte and coworkers in refolding experiments. Additional computational and experimental work may further converge on a unified description of the energy landscape for EcTRX folding.

Previous work also led to proposals for the structure of (the) folding intermediate(s). The burst-phase intermediate in [76] has a predominantly β-structure according to CD spectroscopy. Furthermore, hydrogen-exchange experiments of native EcTRX in the EX2 regime [77] suggest that the amide hydrogens belonging to the β-sheet have higher local stabilities than amide hydrogens belonging to the α-helices. These data fit well with the structure we have proposed for the high-energy unfolding intermediate (Fig. 6). The pH 3 equilibrium intermediate described by Udgaonkar and coworkers [34] is prone to aggregation, which may be a disadvantageous property *in vivo*. In this context, EcTRX domains might have evolved towards a high kinetic stability [31] and folding kinetics in which aggregation-prone intermediates do not populate [78,79].

Previous work pointed at the CTH as an essential element of the EcTRX fold. When the CTH is truncated, the rest of the protein adopts a pre-molten globule state devoid of stable tertiary structure and enzymatic activity [41,44]. The crucial role of the CTH is supported by the thermodynamic analysis of point mutants presented in this work. Mutation of CTH wild-type residues to alanine leads in most cases to a significant reduction in stability (Table 2). These results obtained for oxidized EcTRX are in agreement with the behavior of reduced L107P, L107R and L107A EcTRX variants, which present reduced stability and impaired folding [30]. These results, together with our studies, indicate that in both forms of EcTRX (reduced and oxidized) the CTH has a key role in the structure acquisition process.

Mutations lying on the apolar face of CTH may affect both the secondary structure propensity of the CTH and the consolidation of the EcTRX core by perturbing intrahelical contacts and tertiary interactions. The decrease in stability observed for our EcTRX variants correlates well with both a decrease of helical propensity of CTH peptides in TFE solutions [44] and with incomplete consolidation of the tertiary structure when the mutant CTH peptides are combined with the fragment TRX1–93 [44]. Thus, the mutations analyzed in this work probably lead to native state destabilization by decreasing both secondary structure propensity and tertiary hydrophobic interactions. It is noteworthy that the CTH sequence adopts an amphipathic helical structure in the presence of both its intramolecular interaction surface comprising strands β4, β5 and helices α2 and α3, and of flexible apolar surfaces such as a C18 HPLC matrix or the C12 tail of SDS [43]. Thus, it seems improbable that the CTH folds in the absence of an apolar surface provided by the remaining protein. In agreement with that, the folding of the CTH in our simulations takes place only after the partial folding of its intramolecular interaction surface (Fig. 6).

We observed in our simulations two CTH conformations in the native basin (Fig. 6B). Both present tertiary contacts but only one of them displays significant helicity. This suggests that the lowest free energy route for CTH docking-and-folding involves a non-helical intermediate, reminiscent of simulation results for other binding-and-folding reactions [80]. In this context, as fragment TRX1–93 does not exhibit significant enzymatic activity, one might ask whether the native-like conformations carrying the C-terminal residues in non-helical conformation are inactive. If they were inactive, the activity of the enzyme could be modulated by protein–protein or protein–ligand interactions that alter the CTH conformation.

In the EcTRX folding reaction, apolar residues of CTH (L99, F102, L103 and L107) seem to interact late and in a cooperative way, locally

and with the rest of the protein, stabilizing the secondary structure of CTH and the tertiary structure of the TRX, promoting final consolidation of the protein. As judged by the X-ray structure, in addition to the apolar contacts, polar residues of CTH establish side-chain/side-chain interactions K96–E44 (3.18 Å), K96–E48 (2.54 Å), K100–E48 (3.29 Å), and D104–Y49 (2.88 Å) and side-chain/backbone interactions N106–K82 (2.85 Å). Interestingly, interaction between K100–E48 and contacts between L99–I45 and L103–Y49 (all tertiary contacts) are identified as stabilization centers of the EcTRX structure using the SCide program [81–83], suggesting their involvement in cooperative long-range contacts. Thus, the last stretch of the protein (CTH) would act as an apolar/polar tertiary structure zipper. In this context, it is reasonable to propose a folding mechanism in which apolar/polar interactions take place with dehydration of the apolar interfaces and a concomitant adjustment of CTH with a subsequent stabilization of its helical conformation in a cooperative fashion.

On the other hand, the stability of the isolated helix α5 (CTH in helical conformation, in the absence of non-covalent interactions with the rest of the protein) is so low that, in principle, a scenario where the helix is formed previously to interact with the rest of the protein, and exposing a considerable apolar area, is far less likely than the one described above (Fig. S7).

As a whole, experiments and simulations indicate that the CTH provides thermodynamic and kinetic stability to the native conformation of EcTRX by a combination of local and tertiary interactions. The folding of the CTH in the context of the whole protein can be partly understood from its behavior as an isolated peptide. In turn, the role of intermediates in folding the whole protein could be studied by mutation of the CTH.

Acknowledgments

This work was supported by the *Agencia Nacional de Promoción Científica y Tecnológica* (ANPCyT), PICT2010-2595, the *Consejo Nacional de Investigaciones Científicas y Técnicas* (CONICET), PIP11220110100723 and the *Universidad de Buenos Aires* (UBACyT), 2014-2017. We are grateful to Dr. Gonzalo de Prat Gay and his group for providing us with the stopped flow equipment.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.bbapap.2014.11.004.

References

[1] J. Lu, A. Holmgren, The thioredoxin antioxidant system, Free Radic. Biol. Med. 66 (2014) 75–87.
[2] E.J. Stewart, F. Aslund, J. Beckwith, Disulfide bond formation in the *Escherichia coli* cytoplasm: an *in vivo* role reversal for the thioredoxins, EMBO J. 17 (1998) 5543–5550.
[3] G. Powis, W.R. Montfort, Properties and biological activities of thioredoxins, Annu. Rev. Pharmacol. Toxicol. 41 (2001) 261–295.
[4] A. Vlamis-Gardikas, The multiple functions of the thiol-based electron flow pathways of *Escherichia coli*: eternal concepts revisited, Biochim. Biophys. Acta 1780 (2008) 1170–1200.
[5] Y. Kabe, K. Ando, S. Hirao, M. Yoshida, H. Handa, Redox regulation of NF-kappaB activation: distinct redox regulation between the cytoplasm and the nucleus, Antioxid. Redox Signal. 7 (2005) 395–403.
[6] K. Hirota, M. Murata, Y. Sachi, H. Nakamura, J. Takeuchi, K. Mori, J. Yodoi, Distinct roles of thioredoxin in the cytoplasm and in the nucleus. A two-step mechanism of redox regulation of transcription factor NF-kappaB, J. Biol. Chem. 274 (1999) 27891–27897.
[7] V. Dillet, H.J. Dyson, D. Bashford, Calculations of electrostatic interactions and pKas in the active site of *Escherichia coli* thioredoxin, Biochemistry 37 (1998) 10298–10306.
[8] M.Y. Kim, C.S. Maier, D.J. Reed, M.L. Deinzer, Site-specific amide hydrogen/deuterium exchange in *E. coli* thioredoxins measured by electrospray ionization mass spectrometry, J. Am. Chem. Soc. 123 (2001) 9860–9866.
[9] M.Y. Kim, C.S. Maier, D.J. Reed, P.S. Ho, M.L. Deinzer, Intramolecular interactions in chemically modified *Escherichia coli* thioredoxin monitored by hydrogen/deuterium exchange and electrospray ionization mass spectrometry, Biochemistry 40 (2001) 14413–14421.
[10] J.M. Louis, R.E. Georgescu, M.L. Tasayco, O. Tcherkasskaya, A.M. Gronenborn, Probing the structure and stability of a hybrid protein: the human–*E. coli* thioredoxin chimera, Biochemistry 40 (2001) 11184–11192.
[11] R. Perez-Jimenez, R. Godoy-Ruiz, B. Ibarra-Molero, J.M. Sanchez-Ruiz, The effect of charge-introduction mutations on *E. coli* thioredoxin stability, Biophys. Chem. 115 (2005) 105–107.
[12] R. Perez-Jimenez, J. Li, P. Kosuri, I. Sanchez-Romero, A.P. Wiita, D. Rodriguez-Larrea, A. Chueca, A. Holmgren, A. Miranda-Vizuete, K. Becker, S.H. Cho, J. Beckwith, E. Gelhaye, J.P. Jacquot, E.A. Gaucher, J.M. Sanchez-Ruiz, B.J. Berne, J.M. Fernandez, Diversity of chemical mechanisms in thioredoxin catalysis revealed by single-molecule force spectroscopy, Nat. Struct. Mol. Biol. 16 (2009) 890–896.
[13] A.L. Pey, D. Rodriguez-Larrea, J.A. Gavira, B. Garcia-Moreno, J.M. Sanchez-Ruiz, Modulation of buried ionizable groups in proteins with engineered surface charge, J. Am. Chem. Soc. 132 (2010) 1218–1219.
[14] M. Suarez, P. Tortosa, M.M. Garcia-Mira, D. Rodriguez-Larrea, R. Godoy-Ruiz, B. Ibarra-Molero, J.M. Sanchez-Ruiz, A. Jaramillo, Using multi-objective computational design to extend protein promiscuity, Biophys. Chem. 147 (2010) 13–19.
[15] R. Perez-Jimenez, A. Ingles-Prieto, Z.M. Zhao, I. Sanchez-Romero, J. Alegre-Cebollada, P. Kosuri, S. Garcia-Manyes, T.J. Kappock, M. Tanokura, A. Holmgren, J.M. Sanchez-Ruiz, E.A. Gaucher, J.M. Fernandez, Single-molecule paleoenzymology probes the chemistry of resurrected enzymes, Nat. Struct. Mol. Biol. 18 (2011) 592–596.
[16] H. Garcia-Seisdedos, B. Ibarra-Molero, J.M. Sanchez-Ruiz, Probing the mutational interplay between primary and promiscuous protein functions: a computational–experimental approach, PLoS Comput. Biol. 8 (2012) e1002558.
[17] H. Garcia-Seisdedos, B. Ibarra-Molero, J.M. Sanchez-Ruiz, How many ionizable groups can sit on a protein hydrophobic core? Proteins 80 (2012) 1–7.
[18] H.J. Dyson, G.P. Gippert, D.A. Case, A. Holmgren, P.E. Wright, Three-dimensional solution structure of the reduced form of *Escherichia coli* thioredoxin determined by nuclear magnetic resonance spectroscopy, Biochemistry 29 (1990) 4129–4136.
[19] K. Chandrasekhar, G. Krause, A. Holmgren, H.J. Dyson, Assignment of the ^{15}N NMR spectra of reduced and oxidized *Escherichia coli* thioredoxin, FEBS Lett. 284 (1991) 178–183.
[20] M.J. Stone, K. Chandrasekhar, A. Holmgren, P.E. Wright, H.J. Dyson, Comparison of backbone and tryptophan side-chain dynamics of reduced and oxidized *Escherichia coli* thioredoxin using ^{15}N NMR relaxation measurements, Biochemistry 32 (1993) 426–435.
[21] K. Chandrasekhar, A.P. Campbell, M.F. Jeng, A. Holmgren, H.J. Dyson, Effect of disulfide bridge formation on the NMR spectrum of a protein: studies on oxidized and reduced *Escherichia coli* thioredoxin, J. Biomol. NMR 4 (1994) 411–432.
[22] S.K. Katti, D.M. LeMaster, H. Eklund, Crystal structure of thioredoxin from *Escherichia coli* at 1.68 Å resolution, J. Mol. Biol. 212 (1990) 167–184.
[23] A. Chakrabarti, S. Srivastava, C.P. Swaminathan, A. Surolia, R. Varadarajan, Thermodynamics of replacing an alpha-helical Pro residue in the P40S mutant of *Escherichia coli* thioredoxin, Protein Sci. 8 (1999) 2455–2459.
[24] J.E. Ladbury, N. Kishore, H.W. Hellinga, R. Wynn, J.M. Sturtevant, Thermodynamic effects of reduction of the active-site disulfide of *Escherichia coli* thioredoxin explored by differential scanning calorimetry, Biochemistry 33 (1994) 3688–3692.
[25] R.E. Georgescu, M.M. Garcia-Mira, M.L. Tasayco, J.M. Sanchez-Ruiz, Heat capacity analysis of oxidized *Escherichia coli* thioredoxin fragments (1–73, 74–108) and their noncovalent complex. Evidence for the burial of apolar surface in protein unfolded states, Eur. J. Biochem. 268 (2001) 1477–1485.
[26] R. Godoy-Ruiz, R. Perez-Jimenez, B. Ibarra-Molero, J.M. Sanchez-Ruiz, Relation between protein stability, evolution and structure, as probed by carboxylic acid mutations, J. Mol. Biol. 336 (2004) 313–318.
[27] R.F. Kelley, J. Wilson, C. Bryant, E. Stellwagen, Effects of guanidine hydrochloride on the refolding kinetics of denatured thioredoxin, Biochemistry 25 (1986) 728–732.
[28] J. Wilson, R.F. Kelley, W. Shalongo, D. Lowery, E. Stellwagen, Equilibrium and kinetic measurements of the conformational transition of thioredoxin in urea, Biochemistry 25 (1986) 7560–7566.
[29] R.F. Kelley, W. Shalongo, M.V. Jagannadham, E. Stellwagen, Equilibrium and kinetic measurements of the conformational transition of reduced thioredoxin, Biochemistry 26 (1987) 1406–1411.
[30] D. Huber, A. Chaffotte, M. Eser, A.G. Planson, J. Beckwith, Amino acid residues important for folding of thioredoxin are revealed only by study of the physiologically relevant reduced form of the protein, Biochemistry 49 (2010) 8922–8928.
[31] R. Godoy-Ruiz, F. Ariza, D. Rodriguez-Larrea, R. Perez-Jimenez, B. Ibarra-Molero, J.M. Sanchez-Ruiz, Natural selection for kinetic stability is a likely origin of correlations between mutational effects on protein energetics and frequencies of amino acid occurrences in sequence alignments, J. Mol. Biol. 362 (2006) 966–978.
[32] D. Rodriguez-Larrea, R. Perez-Jimenez, I. Sanchez-Romero, A. Delgado-Delgado, J.M. Fernandez, J.M. Sanchez-Ruiz, Role of conservative mutations in protein multi-property adaptation, Biochem. J. 429 (2010) 243–249.
[33] R. Godoy-Ruiz, R. Perez-Jimenez, B. Ibarra-Molero, J.M. Sanchez-Ruiz, A stability pattern of protein hydrophobic mutations that reflects evolutionary structural optimization, Biophys. J. 89 (2005) 3320–3331.
[34] A. Hamid Wani, J.B. Udgaonkar, HX–ESI-MS and optical studies of the unfolding of thioredoxin indicate stabilization of a partially unfolded, aggregation-competent intermediate at low pH, Biochemistry 45 (2006) 11226–11238.
[35] M.L. Tasayco, K. Chao, NMR study of the reconstitution of the beta-sheet of thioredoxin by fragment complementation, Proteins 22 (1995) 41–44.
[36] R.E. Georgescu, E.H. Braswell, D. Zhu, M.L. Tasayco, Energetics of assembling an artificial heterodimer with an alpha/beta motif: cleaved *versus* uncleaved *Escherichia coli* thioredoxin, Biochemistry 38 (1999) 13355–13366.

[37] A.K. Ghoshal, C.P. Swaminathan, C.J. Thomas, A. Surolia, R. Varadarajan, Thermodynamic and kinetic analysis of the *Escherichia coli* thioredoxin–C′ fragment complementation system, Biochem. J. 339 (Pt 3) (1999) 721–727.
[38] M.L. Tasayco, J. Fuchs, X.M. Yang, D. Dyalram, R.E. Georgescu, Interaction between two discontiguous chain segments from the beta-sheet of *Escherichia coli* thioredoxin suggests an initiation site for folding, Biochemistry 39 (2000) 10613–10618.
[39] W.F. Yu, C.S. Tung, H. Wang, M.L. Tasayco, NMR analysis of cleaved *Escherichia coli* thioredoxin (1–73/74–108) and its P76A variant: *cis/trans* peptide isomerization, Protein Sci. 9 (2000) 20–28.
[40] J. Yang, S. Paramasivam, D. Marulanda, M. Cataldi, M.L. Tasayco, T. Polenova, Magic angle spinning NMR spectroscopy of thioredoxin reassemblies, Magn. Reson. Chem. 45 (Suppl. 1) (2007) S73–S83.
[41] J. Santos, C. Marino-Buslje, C. Kleinman, M.R. Ermacora, J.M. Delfino, Consolidation of the thioredoxin fold by peptide recognition: interaction between *E. coli* thioredoxin fragments 1–93 and 94–108, Biochemistry 46 (2007) 5148–5159.
[42] A. Binolfi, C.O. Fernandez, M.P. Sica, J.M. Delfino, J. Santos, Recognition between a short unstructured peptide and a partially folded fragment leads to the thioredoxin fold sharing native-like dynamics, Proteins 80 (2012) 1448–1464.
[43] E.A. Roman, P. Rosi, M.C. Gonzalez Lebrero, R. Wuilloud, F.L. Gonzalez Flecha, J.M. Delfino, J. Santos, Gain of local structure in an amphipathic peptide does not require a specific tertiary framework, Proteins 78 (2010) 2757–2768.
[44] J. Santos, M.P. Sica, C.M. Buslje, A.M. Garrote, M.R. Ermacora, J.M. Delfino, Structural selection of a native fold by peptide recognition. Insights into the thioredoxin folding mechanism, Biochemistry 48 (2009) 595–607.
[45] D. Huber, M.I. Cha, L. Debarbieux, A.G. Planson, N. Cruz, G. Lopez, M.L. Tasayco, A. Chaffotte, J. Beckwith, A selection for mutants that interfere with folding of *Escherichia coli* thioredoxin-1 *in vivo*, Proc. Natl. Acad. Sci. U. S. A. 102 (2005) 18872–18877.
[46] E. Lacroix, A.R. Viguera, L. Serrano, Elucidating the folding problem of alpha-helices: local motifs, long-range electrostatics, ionic-strength dependence and prediction of NMR parameters, J. Mol. Biol. 284 (1998) 173–191.
[47] V. Munoz, L. Serrano, Elucidating the folding problem of helical peptides using empirical parameters. II. Helix macrodipole effects and rational modification of the helical content of natural peptides, J. Mol. Biol. 245 (1995) 275–296.
[48] V. Munoz, L. Serrano, Elucidating the folding problem of helical peptides using empirical parameters. III. Temperature and pH dependence, J. Mol. Biol. 245 (1995) 297–308.
[49] V. Munoz, L. Serrano, Development of the multiple sequence approximation within the AGADIR model of alpha-helix formation: comparison with Zimm–Bragg and Lifson–Roig formalisms, Biopolymers 41 (1997) 495–509.
[50] A. Bachmann, T. Kiefhaber, Apparent two-state tendamistat folding is a sequential process along a defined route, J. Mol. Biol. 306 (2001) 375–386.
[51] I.E. Sanchez, T. Kiefhaber, Evidence for sequential barriers and obligatory intermediates in apparent two-state protein folding, J. Mol. Biol. 325 (2003) 367–376.
[52] M.M. Santoro, D.W. Bolen, A test of the linear extrapolation of unfolding free energy changes over an extended denaturant concentration range, Biochemistry 31 (1992) 4901–4907.
[53] M. Dellarole, I.E. Sanchez, E. Freire, G. de Prat-Gay, Increased stability and DNA site discrimination of "single chain" variants of the dimeric beta-barrel DNA binding domain of the human papillomavirus E2 transcriptional regulator, Biochemistry 46 (2007) 12441–12450.
[54] A.P. Heuck, R.A. Wolosiuk, Di-fluoresceinthiocarbamyl-insulin: a fluorescent substrate for the assay of protein disulfide oxidoreductase activity, Anal. Biochem. 248 (1997) 94–101.
[55] A.P. Heuck, R.A. Wolosiuk, Fluoresceinthiocarbamyl-insulin: a potential analytical tool for the assay of disulfide bond reduction, J. Biochem. Biophys. Methods 34 (1997) 213–225.
[56] A. Oliva, M. Llabres, J.B. Farina, Estimation of uncertainty in size-exclusion chromatography with a double detection system (light-scattering and refractive index), Talanta 78 (2009) 781–789.
[57] P.C. Whitford, J.K. Noel, S. Gosavi, A. Schug, K.Y. Sanbonmatsu, J.N. Onuchic, An all-atom structure-based potential for proteins: bridging minimal models with all-atom empirical forcefields, Proteins 75 (2009) 430–441.
[58] E. Lindahl, B. Hess, D. van der Spoel, GROMACS 3.0: a package for molecular simulation and trajectory analysis, Mol. Model. Annu. 7 (2001) 306–317.
[59] W.R.P. Scott, P.H. Hünenberger, I.G. Tironi, A.E. Mark, S.R. Billeter, J. Fennen, A.E. Torda, T. Huber, P. Krüger, W.F. van Gunsteren, The GROMOS biomolecular simulation program package, J. Phys. Chem. A 103 (1999) 3596–3607.
[60] J.K. Noel, P.C. Whitford, K.Y. Sanbonmatsu, J.N. Onuchic, SMOG@ctbp: simplified deployment of structure-based models in GROMACS, Nucleic Acids Res. 38 (2010) W657–W661.
[61] H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren, A. DiNola, J.R. Haak, Molecular dynamics with coupling to an external bath, J. Chem. Phys. 81 (1984) 3684–3690.
[62] B. Hess, H. Bekker, H.J.C. Berendsen, J.G.E.M. Fraaije, LINCS: a linear constraint solver for molecular simulations, J. Comput. Chem. 18 (1997) 1463–1472.
[63] S. Kumar, J.M. Rosenberg, D. Bouzida, R.H. Swendsen, P.A. Kollman, THE weighted histogram analysis method for free-energy calculations on biomolecules. I. The method, J. Comput. Chem. 13 (1992) 1011–1021.
[64] L. Cavallo, J. Kleinjung, F. Fraternali, POPS: a fast algorithm for solvent accessible surface areas at atomic and residue level, Nucleic Acids Res. 31 (2003) 3364–3366.
[65] E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, T.E. Ferrin, UCSF Chimera—a visualization system for exploratory research and analysis, J. Comput. Chem. 25 (2004) 1605–1612.
[66] T. Kiefhaber, H.H. Kohler, F.X. Schmid, Kinetic coupling between protein folding and prolyl isomerization. I. Theoretical models, J. Mol. Biol. 224 (1992) 217–229.
[67] T. Kiefhaber, F.X. Schmid, Kinetic coupling between protein folding and prolyl isomerization. II. Folding of ribonuclease A and ribonuclease T1, J. Mol. Biol. 224 (1992) 231–240.
[68] U. Reimer, G. Scherer, M. Drewello, S. Kruber, M. Schutkowski, G. Fischer, Side-chain effects on peptidyl–prolyl *cis/trans* isomerisation, J. Mol. Biol. 279 (1998) 449–460.
[69] R. Perez-Jimenez, R. Godoy-Ruiz, B. Ibarra-Molero, J.M. Sanchez-Ruiz, The efficiency of different salts to screen charge interactions in proteins: a Hofmeister effect? Biophys. J. 86 (2004) 2414–2429.
[70] J.K. Myers, C.N. Pace, J.M. Scholtz, Denaturant m values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding, Protein Sci. 4 (1995) 2138–2148.
[71] T. Kiefhaber, I.E. Sánchez, A. Bachmann, Characterization of protein folding barriers with rate equilibrium free energy relationships, Protein Folding Handbook, Wiley-VCH Verlag GmbH, 2008, pp. 411–444.
[72] S. Gosavi, L.L. Chavez, P.A. Jennings, J.N. Onuchic, Topological frustration and the folding of interleukin-1 beta, J. Mol. Biol. 357 (2006) 986–996.
[73] H. Nymeyer, N.D. Socci, J.N. Onuchic, Landscape approaches for determining the ensemble of folding transition states: success and failure hinge on the degree of frustration, Proc. Natl. Acad. Sci. U. S. A. 97 (2000) 634–639.
[74] P. Das, S. Matysiak, C. Clementi, Balancing energy and entropy: a minimalist model for the characterization of protein folding landscapes, Proc. Natl. Acad. Sci. U. S. A. 102 (2005) 10141–10146.
[75] W. Kabsch, C. Sander, Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features, Biopolymers 22 (1983) 2577–2637.
[76] R.E. Georgescu, J.H. Li, M.E. Goldberg, M.L. Tasayco, A.F. Chaffotte, Proline isomerization-independent accumulation of an early intermediate and heterogeneity of the folding pathways of a mixed alpha/beta protein, *Escherichia coli* thioredoxin, Biochemistry 37 (1998) 10286–10297.
[77] N. Bhutani, J.B. Udgaonkar, Folding subdomains of thioredoxin characterized by native-state hydrogen exchange, Protein Sci. 12 (2003) 1719–1731.
[78] S. Gianni, Y. Ivarsson, P. Jemth, M. Brunori, C. Travaglini-Allocatelli, Identification and characterization of protein folding intermediates, Biophys. Chem. 128 (2007) 105–113.
[79] T.R. Jahn, M.J. Parker, S.W. Homans, S.E. Radford, Amyloid formation under physiological conditions proceeds *via* a native-like folding intermediate, Nat. Struct. Mol. Biol. 13 (2006) 195–201.
[80] A.G. Turjanski, J.S. Gutkind, R.B. Best, G. Hummer, Binding-induced folding of a natively unstructured transcription factor, PLoS Comput. Biol. 4 (2008) e1000060.
[81] Z. Dosztanyi, C. Magyar, G. Tusnady, I. Simon, SCide: identification of stabilization centers in proteins, Bioinformatics 19 (2003) 899–900.
[82] Z. Dosztányi, I. Simon, Stabilization centers in various proteins, Theor. Chem. Acc. 101 (1999) 27–32.
[83] E. Tudos, A. Fiser, A. Simon, Z. Dosztanyi, M. Fuxreiter, C. Magyar, I. Simon, Noncovalent cross-links in context with other structural and functional elements of proteins, J. Chem. Inf. Comput. Sci. 44 (2004) 347–351.
[84] J. Van Durme, J. Delgado, F. Stricher, L. Serrano, J. Schymkowitz, F. Rousseau, A graphical interface for the FoldX forcefield, Bioinformatics 27 (2011) 1711–1712.