Structural Selection of a Native Fold by Peptide Recognition. Insights into the Thioredoxin Folding Mechanism†

Javier Santos,[‡,§] Mauricio P. Sica,[§] Cristina Marino Buslje,[‡] Ana M. Garrote,[‡,§] Mario R. Ermácora,[§] and José M. Delfino*[,‡]

Department of Biological Chemistry and Institute of Biochemistry and Biophysics (IQUIFIB), School of Pharmacy and Biochemistry, University of Buenos Aires, Junín 956, C1113AAD, Buenos Aires, Argentina, and Department of Science and Technology, University of Quilmes, Roque Sáenz Peña 352, B1876XD, Bernal, Argentina

Received October 21, 2008; Revised Manuscript Received December 9, 2008

ABSTRACT: Thioredoxins (TRXs) are monomeric α/β proteins with a fold characterized by a central twisted β-sheet surrounded by α-helical elements. The interaction of the C-terminal α-helix 5 of TRX against the remainder of the protein involves a close packing of hydrophobic surfaces, offering the opportunity of studying a fine-tuned molecular recognition phenomenon with long-range consequences on the acquisition of tertiary structure. In this work, we focus on the significance of interactions involving residues L94, L99, E101, F102, L103 and L107 on the formation of the noncovalent complex between reduced TRX1−93 and TRX94−108. The conformational status of the system was assessed experimentally by circular dichroism, intrinsic fluorescence emission and enzymic activity; and theoretically by molecular dynamics simulations (MDS). Alterations in tertiary structure of the complexes, resulting as a consequence of site specific mutation, were also examined. To distinguish the effect of alanine scanning mutagenesis on secondary structure stability, the intrinsic helix-forming ability of the mutant peptides was monitored experimentally by far-UV CD spectroscopy upon the addition of 2,2,2-trifluoroethanol, and also theoretically by Monte Carlo conformational search and MDS. This evidence suggests a key role of residues L99, F102 and L103 on the stabilization of the secondary structure of α-helix 5, and on the acquisition of tertiary structure upon complex formation. We hypothesize that the transition between a partially folded and a native-like conformation of reduced TRX1−93 would fundamentally depend on the consolidation of a cooperative tertiary unit based on the interaction between α-helix 3 and α-helix 5.

Selective interactions among protein modules are vital in biological processes, such as protein folding or molecular recognition phenomena. The insightful view of protein folding as a hierarchical process involving the assembly of interacting units to attain the global fold highlights the relationship between folding and binding (*1, 2*). This idea harmonizes with the hypothesis that proteins are constructed from a small catalogue of identifiable parts that fit together in a restricted number of ways. In this view, the wide-ranging tendency of the backbone to form secondary structure elements could predetermine the overall shape of the fold, by restricting the general topology (*3, 4*). In this fashion, the interaction between complementary surfaces of these small parts through specific side-chain/side-chain interactions would constitute a very important issue in the process of acquiring and fixing a unique tertiary structure from a large but limited collection of potential backbone conformations (*5*).

Thioredoxins (TRX) are proteins of about 100 residues with a distinctive α/β topology: a motif containing a central twisted β-sheet flanked by helices (*6*−*9*). These proteins participate in diverse redox reactions via the reversible oxidation of the conserved active center WCGPC. This activity is crucial to regulate the intracellular redox environment, cellular growth, cell cycle and resistance against oxidative stress. In addition, TRX has been involved in control of mammalian cell apoptosis and in carcinogenesis (*10*−*13*). The TRX fold also emerges as a dimerization domain in which the redox active site is not present. Such is the case of the eukaryotic protein chaperone erp29 (*14, 15*). Additional examples of the widespread involvement of TRX in biological processes come from the well-studied *Escherichia coli* TRX. As a subunit of the T7 phage DNA polymerase complex, it is essential for phage DNA replication and also for the assembly of the filamentous phages f1 and M13 (*16*). Furthermore, TRX is a fundamental key protein for the activity of TRX-dependent thiol peroxidases (Tpxs), that represent a group of antioxidant enzymes widely distributed in pathogenic bacteria (*Mycobacteria*) and protozoal parasites (*Toxoplasma*, *Trypanosoma*, *Leishmania*, and

† This research has been supported by grants from the University of Buenos Aires (UBA), the Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), the Agencia Nacional de Promoción Científica y Tecnológica (ANPCyT), and the National University of Quilmes (UNQ).

* To whom correspondence should be addressed. E-mail: delfino@qb.ffyb.uba.ar. Phone: 54 11 4964 8291, extension 116. Fax: 54 11 4962 5457.

‡ Department of Biological Chemistry and Institute of Biochemistry and Biophysics (IQUIFIB), School of Pharmacy and Biochemistry, University of Buenos Aires, Junín 956, C1113AAD, Buenos Aires, Argentina.

§ Department of Science and Technology, University of Quilmes, Roque Sáenz Peña 352, B1876XD, Bernal, Argentina.

10.1021/bi801969w CCC: $40.75 © 2009 American Chemical Society
Published on Web 01/02/2009

Plasmodia) (*17, 18*). On the other hand, TRXs have been extensively used to study protein folding and protein stability (*19*−*25*). Recently, fragments of *E. coli* TRX have emerged as model systems of natively disordered proteins (*26*).

The existence of families of natively unfolded proteins in nature cannot be circumvented. From this point of view, the plasticity of backbone and/or complementing surfaces may be a fine-tuned property. In this regard, it could be important to characterize the molecular basis of the stabilization and interaction of protein modules. Thus, reconstituting systems—like complementing fragments in dynamic association/dissociation equilibrium—show great promise as tools for recognizing minimum structural requirements for binding and folding. This kind of system consists of two or more protein fragments that associate through noncovalent interactions to yield a complex with native-like properties. Other particularly interesting systems are molecular chimeras constructed with sections of different proteins sharing low sequence identity, but similar tertiary structure. In most of these cases, it has proven difficult to study the effect of the change introduced due to the extensive surface area involved in the binding/folding event. The high-resolution structure determination of these molecules is fascinating and has helped to understand the basis of protein stability. An excellent example is the human-*E. coli* chimera of TRX (*27, 28*).

From the point of view of plasticity, *E. coli* TRX seems to be an exceptionally malleable system. In addition to the native-like, chimerical protein mentioned above, pairs of fragments TRX1−37/TRX38−108, TRX1−73/TRX74−108, and TRX1−93/TRX94−108 are endowed with the capability of associating and reconstituting native structures with biological TRX activity (*26, 29*−*31*). In the latter case, interaction between the disordered peptide TRX94−108 in isolation (the C-terminal α-helix 5 in the native structure) and fragment TRX1−93 embodies a molecular recognition event attached to reciprocal structure selection and structure acquisition (*31*). Interestingly, the association between these fragments is circumscribed to a few interactions and a small occluded area. We have previously shown that packing of α-helix 5 against the remainder of the protein plays a key role in stabilizing the native fold of TRX (*31*). This fact highlights the role played by the so-called *L row*: four leucine residues clustered along one side of α-helix 5. Pairs of leucine side-chains spaced at positions i and i+4 are known to stabilize alanine-based peptide helices (*32, 33*). This situation represents the most frequently observed class of pairwise side-chain interaction found in α-helices and requires that the χ_1 rotamers of residues i and i+4 be *trans* and *gauche*+, respectively (*33, 34*). In this work, we analyze by alanine scanning mutagenesis the significance of interactions involving L94, L99, F102, L103 and L107 residues where three out of four leucine residues are positioned at i, i+4. In this regard, we wonder whether a particular set of group interactions could encompass a key role in the acquisition of secondary structure of the peptide and native tertiary structure of the complex. In support of this possibility, bioinformatics analysis showed that L99 and L103 are highly conserved along evolution (*31*). Together with previous results, which identify residues F102 and L107 as contributing to protein stability and folding (*35, 36*), the results from this analysis suggest that the C-end of TRX is an important module for protein cooperativity. Our broad goal is to evaluate the informational content of peptide TRX94−108 in regard to efficiently achieving the folded state of TRX.

MATERIALS AND METHODS

General Details. Protein and peptide purity were evaluated by SDS−PAGE (*37*) and RP-HPLC, respectively. Molecular graphics were prepared using PyMOL (http://pymol.sourceforge.net) and VMD (*38*). Accessible surface area (ASA[1]) was calculated using MOLMOL software (*39*). For protein concentration measurement, the theoretical extinction coefficient of unfolded TRX1−108 was used for fragment TRX1−93 (ε = 14,060 M^{-1} cm^{-1} at 280 nm). For peptide TRX94−108, the extinction coefficient at 280 nm was calculated as 1,490 M^{-1} cm^{-1} (*40*). Protein mass was measured by HPLC/ESI-MS using a 1.0 mm × 30 mm Vydac C8 column, eluted at a flow rate of 40 μL/min, in a Surveyor HPLC system connected online with an LCQ Duo (ESI ion trap) mass spectrometer (Thermo Finnigan, San José, CA). Proteins were eluted using a 15 min linear gradient from 10 to 100% solvent B in solvent A (solvent A, 2% acetic acid, 2% acetonitrile in water; solvent B, 2% acetic acid, 96% acetonitrile in water). ESI-MS spectra range was 300−2000 amu, and data were deconvoluted with the program Xcalibur provided with the instrument.

Protein and Peptide Preparation. A complete description of the preparation of fragment TRX1−93 is in Santos et al. (*31*). Briefly, taking advantage of its pronounced tendency to aggregate during protein expression in *E. coli*, fragment TRX1−93 was purified from inclusion bodies. These were dissolved in 6.0 M urea, 10 mM DTT, 5 mM glycine, pH 7.0 for 20 min. Then, sodium acetate was added up to a 10 mM final concentration (pH 4.5). This solution was centrifuged at 10,000 rpm for 30 min before loading it into a fast-flow SP-Sepharose column to remove insoluble particles. Protein elution was performed with a NaCl gradient up to 1.0 M. Fractions containing TRX1−93 were pooled and injected into an HPLC system (Rainin Dynamax, NY) equipped with a reverse phase C4 semipreparative column (Vydac) equilibrated in 0.05% aqueous TFA. Protein was eluted with a linear gradient from 40 to 60% aqueous acetonitrile (0.05% TFA). Fragment TRX1−93 typically elutes at ∼50% acetonitrile. Finally, fractions containing ≥95.0% pure TRX1−93 (as judged by analytical RP-HPLC) were lyophilized. Peptide TRX94−108, "LSKGQLKEFLDANLAY", was synthesized and partially purified up to the desalting step by GenScript Corp. The purification was completed by HPLC (Rainin Dynamax, NY) using a reverse phase C18 semipreparative column equilibrated in 0.05% aqueous TFA. TRX94−108 was eluted with a linear gradient

[1] Abbreviations: ASA, accessible surface area; CD, circular dichroism; di-FTC-insulin, difluoresceinthiocarbamyl-insulin; DTT, dithiothreitol; ESI-MS, electrospray ionization mass spectrometry; MC, Monte Carlo analysis; MDS, molecular dynamics simulations; PAGE, polyacrylamide gel electrophoresis; rmsd, root-mean-square deviation; RP-HPLC, reversed-phase high-performance liquid chromatography; SDS, sodium dodecyl sulfate; SEC, size-exclusion chromatography; TFE, 2,2,2-trifluoroethanol; TRX, thioredoxin; TRX1−93, the recombinant fragment spanning the sequence 1−93 of full-length TRX; TRX94−108, the synthetic peptide comprising residues 94 to 108 of TRX or mutants thereof, plus a C-terminal tyrosine tag (the numbering scheme used in entry pdb 2TRX is used throughout the paper).

from 30 to 45% aqueous acetonitrile, 0.05% TFA; the peptide typically elutes at ~40% acetonitrile. Fractions containing ≥98.0% pure peptide were pooled and lyophilized. The same procedure was carried out in the case of mutant peptides TRX94−108: L94A, L99A, F102A, L103A, L107A and E101G. The covalent structure of these peptides was verified by ESI mass spectrometry.

Multiangle Laser Light Scattering Measurements. Peptide samples were subjected to size exclusion chromatography (HPLC-SEC) on a Pharmacia Superdex 200 (10/30) column. Eluted peaks were monitored at 690 nm with a light-scattering detector (miniDAWN, Wyatt Technology Corp, Santa Barbara, CA) and at 280 nm with a UV detector (Rainin Dynamax, NY). The molar mass of the protein sample was determined from Zimm plots with Astra software (Wyatt Technology Corp, Santa Barbara, CA). A differential index of refraction (d*n*/d*c*) of 0.186 mL/g for the analyzed proteins was assumed (*41*).

CD Spectroscopy. Measurements were carried out with a Jasco J-810 spectropolarimeter calibrated with (+) 10-camphor sulfonic acid. Near-UV CD spectra were collected in 20 mM TrisHCl buffer, 100 mM NaF, 1.0 mM DTT, 0.1 mM EDTA, pH 7.0. Protein and peptide concentrations were 30.0 μM and 150 μM, respectively. Far-UV CD spectra were collected in 10 mM sodium phosphate, 100 mM NaF, pH 7.0. Cells of 1.0 and 0.1 cm were used for near- and far-UV measurements, respectively. Data were acquired at a scan speed of 20 nm/min and at least five scans were averaged. Finally, blank (buffer) scans were subtracted from the spectra and values of ellipticity were generally expressed in units of deg cm^2 dmol^{-1}. TFE-peptide mixtures were prepared by adding aliquots of a peptide stock solution (1.0 mM) into aqueous TFE (400 μL of 0−80%, v/v). Final peptide concentration was 100 μM.

Binding of peptide TRX94−108 to fragment TRX1−93 was followed by the evolution of the CD signal at 280 nm. The band centered at this wavelength reports the tyrosine/tryptophan environment and probes the consolidation of tertiary structure that follows the formation of the complex.

Unfolding transitions as a function of temperature were also monitored by the CD signal at 280 nm. Protein and peptide concentrations were 30.0 μM and 150 μM, respectively, and a 1.0 cm cell was used. Experiments were performed in 20 mM sodium phosphate, 100 mM NaCl, 1.0 mM DTT, 0.1 mM EDTA, pH 7.0. Temperature was varied in the range 0−95 °C at a scan rate of 2 °C min^{-1}, and the melting curve was sampled at 0.2 min intervals. The following equations were fitted to the data (*42*):

$$\Delta G_{NU} = -RT \ln\left(\frac{f_U}{f_N}\right) = \Delta H_{T_m} + \Delta C_P(T - T_m) - T\left(\left(\frac{\Delta H_{T_m}}{T_m}\right) + \Delta C_P \ln\left(\frac{T}{T_m}\right)\right)$$

$$S = f_N(S_{0,N} + l_N T) + f_U(S_{0,U} + l_U T)$$

where f_U and f_N are the unfolded and folded fractions at equilibrium, respectively; T_m is the temperature at which $f_U = f_N$; S is the observed CD signal; $S_{0,N}$ and $S_{0,U}$ are the intrinsic CD signals for the native and unfolded states, respectively; l_N and l_U are the slopes of the pre and post transitions, respectively, assuming a linear dependence of S_N and $S_U$with temperature.

Fluorescence Measurements. Steady-state fluorescence measurements were performed in an Aminco Bowman Series II spectrofluorometer operating in the ratio mode and equipped with a thermostatted cell holder connected to a circulating water bath set at 25 °C. A 0.3 or 1.0 cm path cuvette sealed with a Teflon cap was used. When the intrinsic fluorescence of proteins was measured, the excitation wavelength was 295 nm, and emission data were collected in the range 305−450 nm. The spectral slit-widths were set to 8 nm for both monochromators. Fragment and peptide concentrations were 30.0 μM and 150 μM, respectively. The experiments were carried out in 20 mM TrisHCl buffer, 100 mM NaCl, 1.0 mM DTT, 0.1 mM EDTA, pH 7.0.

Fluorometric Assay of TRX Activity. Bovine pancreas insulin was chemically modified with fluorescein isothiocyanate and purified according to Heuck and Wolosiuk (*43*) to yield the fluorogenic substrate difluoresceinthiocarbamylinsulin (di-FTC-insulin). The catalytic activity of TRX as protein disulfide reductase was assayed as follows. Fragment and peptide TRX were dissolved in buffer (20 mM TrisHCl buffer, 100 mM NaCl, pH 7.0) with the addition of 1 mM DTT, and their concentrations were 30.0 μM and 150 μM, respectively. The reaction was started by the addition of di-FTC-insulin to this solution (0.1 μM final concentration), and the intensity of the fluorescent signal (F) at 519 nm (excitation wavelength is 495 nm) was continuously recorded at room temperature as a function of time. Normalized intensities were expressed as the ratio $((F - F_o)/F_o)$, where F_o is the value of the fluorescent intensity of the protein substrate. Finally, the specific catalytic activity was estimated through the initial slope of the kinetic trace.

Molecular Dynamics Simulations. Molecular dynamics simulations (MDS) were carried out with GROMACS 3.3 (*44*) and GROMOS 53a6 force field (*45*) to investigate the conformational landscape of TRX94−108 and their mutants, in solution or complexed with TRX1−93. In all cases, the initial structures were generated from the coordinates of chain A in the crystallographic structure 2TRX. The structure of each peptide or complex was embedded in a dodecahedral periodic cell (90 or 114 nm^3 respectively) and solvated with SPC (simple point charge) water molecules (*46*). The size of this periodic cell avoids contacts between mirror images, even if peptide TRX94−108 is completely unfolded. Sodium and chloride ions were added to neutralize charges in the peptides and to simulate a 150 mM salt concentration. Conditioning consisted of one thousand steps of energy minimization and subsequent linear heating (from 0 to 300 K in 10 ps) by position-restrained MDS. The conditioned system was the starting point for the simulation.

The system was simulated as an isobaric−isothermal ensemble at 300 K and 1 bar, using weak temperature and pressure coupling (*47*) (0.1 and 1.0 ps^{-1}, respectively). All hydrogen atoms were considered explicitly. The masses of the hydroxyl hydrogen atoms were increased 4-fold at the expense of the oxygen atoms masses. The positions of the remaining hydrogen atoms were calculated as virtual-interaction sites (*48*). All protein covalent bonds were constrained by LINCS (*49*). This setup allowed us to use a 4 fs time-step, according to the work of Feenstra and co-workers (*50*). Cut-off distances for neighbor searching, Lennard-Jones, and Coulombic interactions were 9, 9 and 14 Å, respectively. Long-range interactions were computed

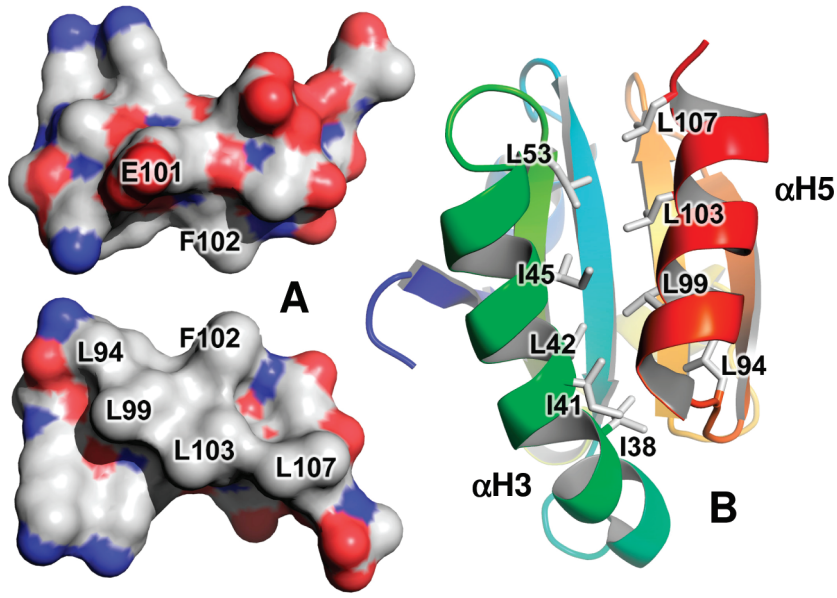

FIGURE 1: (A) Van der Waals surface model of peptide TRX94−108 colored by atom type (the rear view results from a 180° rotation of the top view around a horizontal axis), and (B) a ribbon diagram model of TRX (2TRX)—decorated with the side-chains of key residues as sticks—showing the interaction between α-helix 3 and α-helix 5.

according to the generalized reaction field method proposed by Tironi et al. (*51*). Cluster analysis was carried out using the method of Daura et al. (*52*).

Conformational Search of Peptides. Molecular modeling experiments employed MacroModel 8.6 equipped with Maestro 6.5 graphical interface installed on a Linux Rocks Cluster System 4.2.1. Molecular models of the core pentapeptides (LKEFL and mutants) were initially constructed into extended conformations (phi = 180°, psi = 180°) using the model building facility implemented in Maestro (*53*). In all cases, N- and C-termini were acetylated and N-methyl amidated, respectively. Coordinates for these structures were used as input for MacroModel/BatchMin, the calculation module of this program. The conformational search of peptides was carried out with an optimized Monte Carlo method (MC) (*54*). This procedure allows free rotation around all phi/psi dihedral angles. The energy was minimized using the OPLS 2001 force field (*55*) and the Polak−Ribière conjugate gradient, with a final gradient of 0.01 kcal $Å^{-1}$ mol^{-1} as the criterion for convergence. Aqueous solution conditions were simulated using the continuum dielectric water solvent model (GB/SA) as implemented in MacroModel. Extended cutoff distances were defined at 8 Å for van der Waals, 50 Å for electrostatics and 4 Å for H-bonds. After each MC step the conformer was partially minimized (500 iterations), and the subset of resulting structures was further minimized until convergence. In this process, duplicate and high energy structures were discarded. In the end, all nonenantiomeric conformers within a 50 kJ mol^{-1} window above the global minimum were analyzed. All backbone atoms were taken into account for comparisons. In all cases, each run included 50,000 MC steps, which sufficed to sample the conformational ensemble.

Torsion angles and interatomic distances were measured automatically with the command GEOM on the set of minimized structures. These values were further formatted using Linux scripts and processed with a macro implemented in Visual BASIC on a Microsoft Excel spreadsheet that assigns secondary structures classes to each amino acid position (available from the authors), according to refs *56, 57*. To achieve a realistic view of the relative abundance of conformers in equilibrium at 25 °C present in the population, each structure was correspondingly weighted by the Boltzmann factor calculated on the energy difference with respect to the global minimum. The final plotted values represent running averages of the relative abundance of secondary structure classes or interatomic distances.

Application of the FoldX Algorithm. A 3D model of the complex (TRX94−108/ TRX1−93) was prepared as for MDS and subjected to an optimization procedure using the repair function of FoldX prior to calculating the contribution to the total free energy per residue with SequenceDetail (*58, 59*).

RESULTS

Preparation of the Complementary Moieties. One of the outstanding properties of α-helix 5 in the native state of full-length TRX is its evident amphiphilicity. In the nonpolar face, residues L94, L99, F102, L103 and L107 establish a network of intra helical contacts as well as interactions with the rest of the protein (Figure 1). In the opposite side, charged residues K96, K100, E101 and D104 are solvent exposed. In this context, it is not unreasonable to assume that perturbations in the apolar network could bear a direct consequence on the consolidation of the TRX fold.

To test this hypothesis, a series of synthetic mutant peptides were prepared where each apolar residue of α-helix 5 was replaced by alanine. In addition, E101 was replaced by glycine to test the possibility of locally reducing the helical propensity of the peptide without affecting the apolar interface. The synthetic peptides were purified by HPLC to homogeneity (>98%), and their hydrodynamic behavior was assayed by size exclusion chromatography and light scattering. An important feature for further analysis is that all peptides exhibit monomeric behavior in aqueous solution, even at the highest concentration assayed (100 μM). For this purpose, each peptide (1 mM) was injected into a Superdex

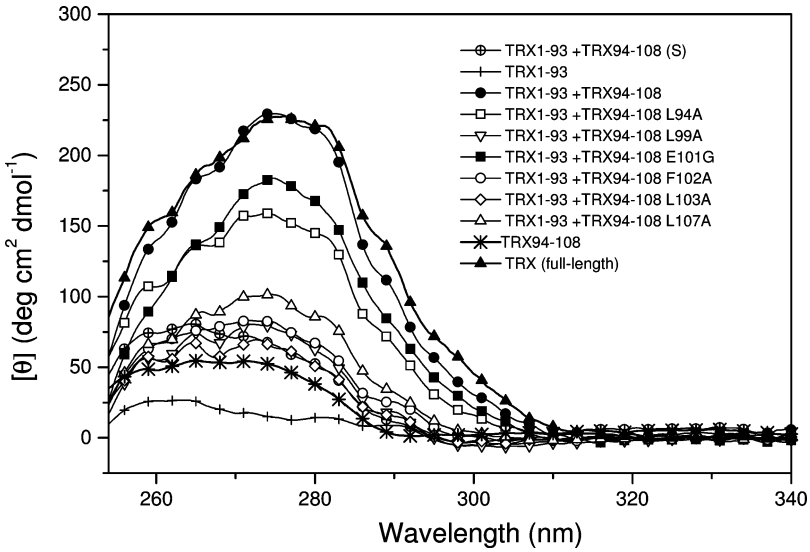

FIGURE 2: Binding of peptide TRX94−108 and mutants to fragment TRX1−93 as followed by near-UV circular dichroism. Fragment TRX1−93, incubated with wild-type peptide TRX94−108 (●), or mutants: L94A (□), L99A (▽), F102A (○), L103A (◇), E101G (■) or L107A (Δ). The spectra of the full-length TRX (▲), isolated fragment TRX1−93 (+), isolated wild-type peptide TRX94−108 (*) and the sum (S) of spectra corresponding to isolated fragment TRX1−93 and peptide TRX94−108 (⊕) are shown as control samples.

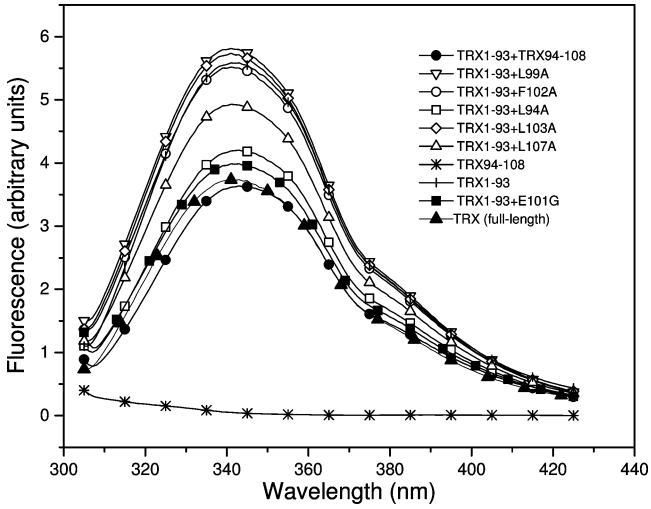

FIGURE 3: Fluorescence emission spectra of TRX1−93 added to peptides. Isolated TRX1−93 (|), incubated with wild type peptide TRX94−108 (●), or mutants: L94A (□), L99A (▽), F102A (○), L103A (◇), E101G (■) or L107A (Δ) mutants. The fluorescence spectrum of isolated peptide TRX94−108 (*) and full-length TRX (▲) are shown as control samples.

S-200 column and the main peak detected by light-scattering yields values of 1944 ± 200 (wild type), 1832 ± 293 (L99A), 1851 ± 222 (L103A), 2127 ± 400 (L107A), 2080 ± 300 (L94A), 1854 ± 87 (E101G) and 1859 ± 150 Da (F102A).

The complementary part of the complex, fragment TRX1−93, was efficiently expressed in *E. coli*, with a yield of 40 mg/L of culture and purified to >95% homogeneity. The molecular masses of the fragment and mutant peptides determined by ESI-MS differed by less than 1 Da from those theoretically calculated.

Exploring the Structure of Complexes. After overnight dialysis from 6.0 to 0.005 M urea, 1.0 mM DTT, at 4 °C, the reduced form of fragment TRX1−93 adopts a partially folded state sensitive to limited proteolysis and lacking characteristic spectroscopic signatures of the native state (Figures 2 and 3). This isolated fragment is catalytically inactive (Figure 4).

Adding wild-type peptide TRX94−108 to the reduced form of fragment TRX1−93 induces the formation of a complex in which aromatic residues become structured in an asymmetric environment (as attested by near-UV CD) and causes a significant quenching of tryptophan fluorescence (Figures 2 and 3, see also ref *31*), giving rise to spectral signatures identical in shape and magnitude to those shown by the full-length protein (see also ref *60*). Contrary to this behavior, mutation of residues L99 or L103 to alanine brings about a complete loss in the ability of the peptide to interact with fragment TRX1−93, thus preventing the acquisition of the full-fledged tertiary structure. Even a very large excess of peptide (5:1 on a molar basis) is not sufficient to fix the tertiary structure, as evidenced by the absence of both the near-UV CD signal and tryptophan fluorescence quenching (Figures 2 and 3, respectively). Accordingly, only residual redox activity is observed for these mixtures, as measured by the insulin reduction assay (Figure 4).

Interestingly, the mutations L94A and E101G preserve most of the ability to bind to the fragment TRX1−93, as demonstrated by both spectroscopy and enzymic activity. Remarkably, the redox activity observed for the L94A complex is significantly enhanced (nearly 3-fold) with respect to the wild type complex (the latter exhibits about 10% the activity of the full-length protein, see ref *31*). This could result from direct effects on the active site, given its proximity to the mutated side chains.

On the other hand, peptide L107A shows a weak capacity for forming a structured complex, as judged by near-UV CD spectroscopy (Figure 2). Nevertheless, the significant quenching of tryptophan fluorescence (Figure 3) and the substantial enzymic activity observed (Figure 4) suggest complex formation.

Temperature unfolding experiments followed by near-UV CD at 280 nm for wild-type, E101G and L94A complexes yield cooperative transitions from which thermodynamic parameters could be derived: (a) T_m = 49.7, 42.6 and 38.0 °C and (b) $\Delta G°_{NU(25°C)}$ = 3.23, 2.69 and 1.35 kcal mol^{-1} for wild type, E101G and L94A, respectively. For the sake of

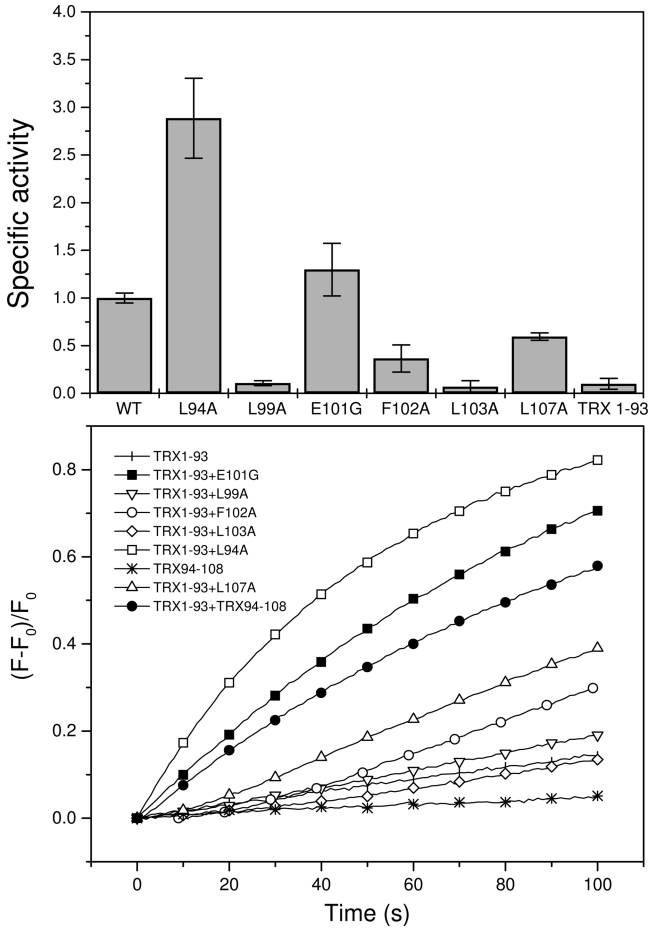

FIGURE 4: TRX1−93/TRX94−108-catalyzed reduction of insulin. The time course of the protein disulfide reduction exhibited by mixtures of TRX1−93 and peptide TRX94−108 or mutants assayed against di-FTC insulin (lower panel): Isolated TRX1−93 (|), TRX1−93 incubated with wild type peptide (●), or mutants L94A (□), L99A (▽), F102A (○), L103A (◇), E101G (■) or L107A (Δ). A sample containing isolated peptide TRX94−108 is showed as a control (*). The upper panel represents the specific activity of each sample expressed relative to the wild-type complex. The average and standard deviation of data corresponding to two independent experiments, each run in triplicate, are plotted as bars with errors.

comparison, data for full-length TRX was also included (T_m = 76.4 °C; $\Delta G°_{NU(25°C)}$ = 7.55 kcal mol^{-1}: Figure 5). The stability and cooperative unfolding along with the conservation of significant catalytic activity strongly support a native-like nature of these complexes. Unfortunately, the small signal-to-noise ratio observed for L107A precluded a reliable assessment of the unfolding transition (data not shown).

In the case of the F102A mutant, although a low but significant enzymic activity occurs, neither quenching of tryptophan fluorescence nor near-UV CD changes become apparent, suggesting that the tendency to complement of this mutant is so weak that only the additional stabilization by the substrate drives the formation of a native-like complex.

All the results support the idea that mutations L99A, F102A and L103A produce the largest destabilization of the complex.

Helical Tendency of Peptide TRX94−108 and Its Mutants. On average, the isolated peptide TRX94−108 is unstructured in aqueous solution. Nevertheless, 2,2,2-trifluoroethanol (TFE), a solvent that preferentially stabilizes structures rich

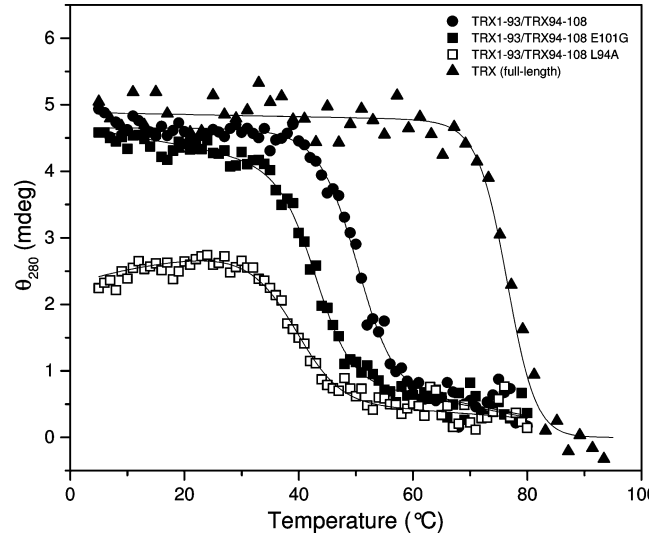

FIGURE 5: Thermal unfolding of TRX1−93/TRX94−108 complexes. The near-UV CD signal at 280 nm is monitored as a function of temperature: wild type (●), L94A (□) and E101G (■) complexes. Thermal unfolding of the reduced form of the full-length TRX (▲) is shown as control samples. Solid lines represent the nonlinear regression fitting of curves to the data (see Materials and Methods).

in intramolecular H bonds (*61, 62*), promotes an α-helical conformation for this peptide. This fact opens the way to testing the intrinsic propensity of this peptide and mutants to form helical structures. The temperature-induced unfolding of the TFE-stabilized helix monitored by far-UV CD shows cooperativity (Figure 6, panel A), with a $T_m \sim 40$ °C, at 25% (V/V) TFE. Increasing the ionic strength by the addition of NaCl or NaF brings about a small increment in the far-UV CD signal of peptide TRX94−108 (data not shown).

TRX94−108 mutant peptides display spectra compatible with random-coil structures in aqueous solution, in a similar fashion to the wild-type peptide. In all cases, addition of TFE induces CD bands consistent with the appearance of α-helical conformers, i.e. negative bands at 208 and 222 nm and positive ellipticity in the 192−193 nm range (Figure 6, panel B, inset). Interestingly, mutant peptides L99A and L103A show a much lesser tendency (∼30% relative to wild-type) to form helical structure, as it becomes evident in 20% TFE. Thus, the overall helix trend results: wild-type > L107A > L94A > F102A ∼ E101G > L99A ∼ L103A (Figure 6, panel B). This attenuated helix propensity might indicate the weakening in local stabilizing interactions, such as hydrophobic side-chain contacts. The above trend agrees reasonably well with the predictions by AGADIR (*63, 64*): wild-type > F102A > L94A ∼ L107A > L99A ∼ L103A > E101G. In view of the amphipaticity of the helix, a detergent like SDS would also have the capability of stabilizing this structure. Indeed, SDS at submicellar concentration (2.0 mM) promotes helical structure formation, as attested by far-UV CD. In agreement with the TFE results, more SDS is required for inducing similar helix content in the case of L99A and L103A with respect to wild-type (Ernesto A. Román, Pablo Rosi, José M. Delfino, F. Luis González-Flecha, and Javier Santos, unpublished results).

Nevertheless, the fact that L107A, despite showing more helical propensity than L94A and E101G, interacts less efficiently with the cognate fragment TRX1−93, suggests

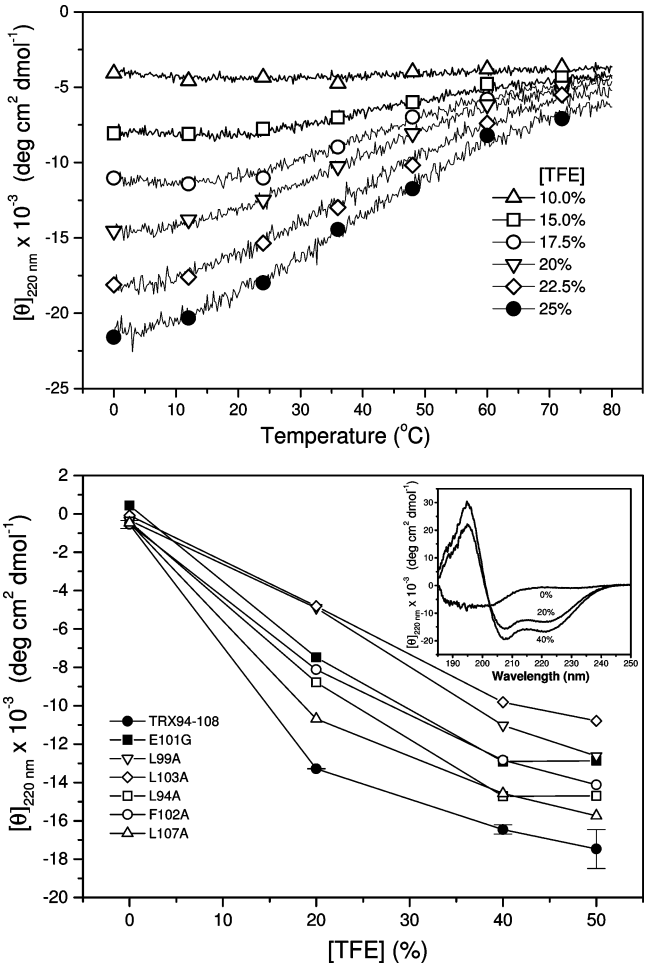

FIGURE 6: Intrinsic helix-forming tendency of peptide TRX94−108 and mutants as monitored by far-UV CD spectroscopy. (A, upper panel) Thermal unfolding behavior of wild type peptide at different TFE concentrations, followed by the molar ellipticity at 220 nm. (B, lower panel) Induction of secondary structure by TFE: Wild-type (●), and mutant peptides L94A (□), L99A (▽), F102A (○), L103A (◇), and L107A (Δ) were incubated for 1 h at 25 °C with TFE, the final TFE concentration of each sample is indicated. The far-UV CD spectra were recorded, and the molar ellipticity at 220 nm is plotted. The inset shows the CD spectra corresponding to the wild-type peptide dissolved in 0, 20 or 40% aqueous TFE solutions.

that additional factors such as the occurrence of particular groups able to establish specific tertiary contacts might come into play for an efficient structure consolidation of the complex.

Monte Carlo Conformational Search. Positions L99 and L103 appear to be critical for helix integrity (Figure 6). Therefore we undertook a thorough conformational analysis of the core region bound between these positions. After a full search of the phi−psi space a clear tendency favoring helix structure becomes evident for all peptides except for that where the central E101 is replaced by glycine. In the latter, the added conformational backbone freedom—rather than the loss of the negative charge of the side-chain (as attested by the behavior of the E101Q mutant, not shown)—favors alternative compact H-bonded structures. In general, α-helical structures prevail, although $\alpha/3_{10}$ and α/π helices including bifurcated H bonds at the acceptor carbonyl oxygen are also represented (Figure 7, panel A). Significantly, helical structures with the lowest energies superimpose very well

Table 1

	rmsd values (Å)					
conformation[b]	LKEFL[a]	AKEFL[a]	LKEFA[a]	LKGFL[a]	LKEAL[a]	LKQFL[a]
1	0.66[c]	0.17	0.66	1.41	0.67	0.65
2	1.03	1.47	0.17	1.82	0.20	0.65
3	0.84	0.70	0.81	1.20	1.68	1.70
4	0.21	1.15	1.66	1.37	1.76	0.16
5	1.67	0.80	1.68	1.42	0.86	1.82

[a] Peptides were initially constructed into extended conformations (phi = 180°, psi = 180°), and in all cases, N- and C-termini were acetylated and N-methyl amidated, respectively. [b] The five conformations of lowest energy after the Monte Carlo search and energy minimization. [c] For the rmsd calculation, all backbone atoms were considered and the TRX structure (2TRX) was used as a reference.

with the corresponding segment in TRX (see rmsd of backbone positions in Table 1). In a similar fashion, the calculated helical conformations place the terminal leucine residues in close contact (Figure 7, panel B), thus minimizing the exposure of hydrophobic area to the aqueous solvent and giving rise to amphipaticity. Different leucine rotamer combinations are allowed within a 10−15 kJ mol^{-1} window above the global minimum, permitting close packing in all cases (data not shown).

Molecular Dynamics Simulations. To better understand the effect of the mutations on the helix propensity of peptides, we performed MDS (Figure 8). Four runs of 20 ns per peptide were carried out, starting from the native conformation. The helical structure of the peptides was lost during the simulations and no refolding event was ever observed, which is consistent with the absence of structure in these peptides as shown in the CD experiment described above. Interestingly, in wild type and mutants with leucine residues at positions 99 and 103, the central helical structure remains even after unfolding has started at the C-terminus. Taken together, this evidence suggests that the peptides are committed to the helical state mainly by the local side-chain interactions of central residues 99 and 103. Expectedly, in the mutant E101G, the centrally located glycine residue allows alternative conformations, and in the case of mutant F102A in one of the simulations a β-hairpin was formed.

In addition to the peptide simulations, two runs of 60 ns molecular dynamics were performed with the isolated reduced fragment TRX1−93 and with each TRX1−93/TRX94−108 complex. The TRX1−93 fragment exhibited high rmsd values throughout the simulations (>5 Å, Figure 9, panel A) supporting the view whereby interactions with peptide TRX94−108 constitute a major stabilizing factor of the TRX fold. As per the complex simulations, after a few nanoseconds, wild-type, L94A, F102A and L107A complexes reached a plateau of rmsd (2.0−3.5 Å) (Figure 9, panel A), and the overall native structure of the complexes was preserved. The only significant perturbations, likely caused by the lack of continuity of the backbone after residue 93, were localized at loops 29−32, 60−63 and 71−77, which, in the crystallographic structure, interact with the loop 92−95. In the simulations with the mutant E101G, the complex was preserved but the TRX94−108 moiety underwent partial unfolding (Figure 9, panel D).

Finally, throughout the simulation of the mutant complex L99A, the Cα-rmsd increases monotonically, suggesting that the complex has a marked tendency to unfold. Interestingly, during the unfolding of the L99A complex, helix 94−108

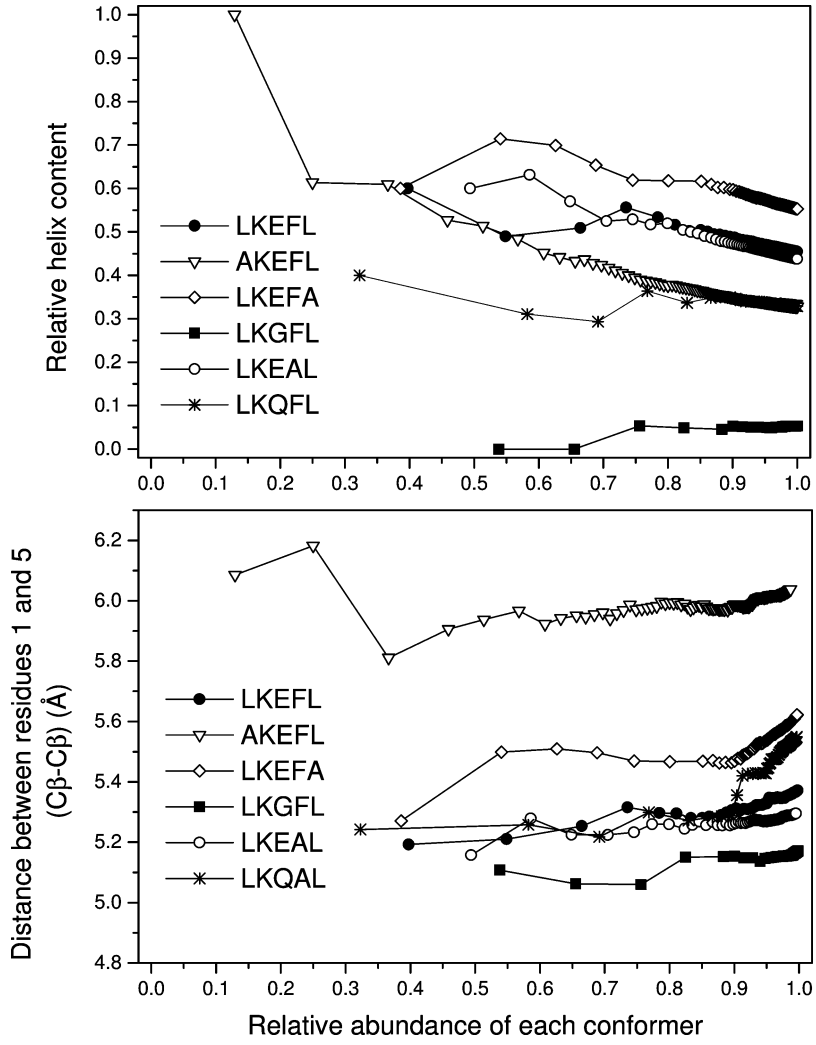

FIGURE 7: The MC conformational search of the pentapeptide core of α-helix 5 in TRX. Plotted values represent: running averages of the relative abundance of (A, upper panel) helical conformation or (B, lower panel) interatomic distances between β carbon atoms of residues 1 and 5, present in the population of the wild-type LKEFL and mutant peptides. Each structure was correspondingly weighted by the Boltzmann factor calculated on the energy difference with respect to the global minimum.

unwinds concertedly with α-helix 3 (residues 38−53; Figure 9, panels B and C). Although all mutations perturb to some degree α-helix 3, this effect is particularly strong in the L99A variant. This is also in accordance with cluster analysis of the interstructure rmsd matrix for each simulation, where similar structures with rmsd < 0.2 Å are grouped in the same cluster. For the wild-type, L94A, L107A and F102A complexes, 4−5 clusters were calculated, while for E101G and L103A and L99A complexes, 12, 12, and 27 clusters were respectively found (Figure 9, panel C). This result highlights the importance of the hydrophobic core involving central leucine residues 99 and 103 and side-chains of I38, L41 and I46 in the stabilization of both helices, thus giving rise to the overall consolidation of the tertiary TRX structure (see also Figure 1, panel B).

It is worth noting that the structure of the active site, including the environment of residues W31 and W28, is highly dependent on the interaction between the above two helices. Therefore, their disruption is expected to affect enzyme activity and spectroscopic properties. In addition, disorder of the hydrophobic core between these helices could be the main reason underlying the inability of fragment TRX1−93 to achieve a native-like state, as shown by the experiments and the simulations.

DISCUSSION

A fundamental input in the process of tertiary structure acquisition of a polypeptide chain is the selection of the native structure from a large but limited set of allowed conformations. Peptide TRX94−108 (corresponding to α-helix 5 in TRX structure) is not stable by itself as a helix in aqueous solution, although its helical propensity can be enhanced by TFE (Figure 6) and it adopts a helical conformation in the complex with the complementary fragment TRX1−93. Thus, the folding process involves both local and long-range interactions that are cooperatively coordinated as the TRX structure consolidates.

In the binding−folding process, L99 and L103 emerge as critical residues. These amino acids occupy a central position in the α-helix, possess minimal solvent accessible surface area (ASA) and exhibit reduced mobility (Figure 10, panels A, D and E). Furthermore, these residues are moderately conserved throughout TRX evolution (Figure 10, panel B), and whenever mutations take place at these positions, the hydrophobic profile is conserved. In fact, the appearance of the three most frequent amino acid residues at position 99 is 46% for L, 21% for I, and 16% for F; and at position 103, 34% for L, 32% for I, and 15% for V (see Supporting

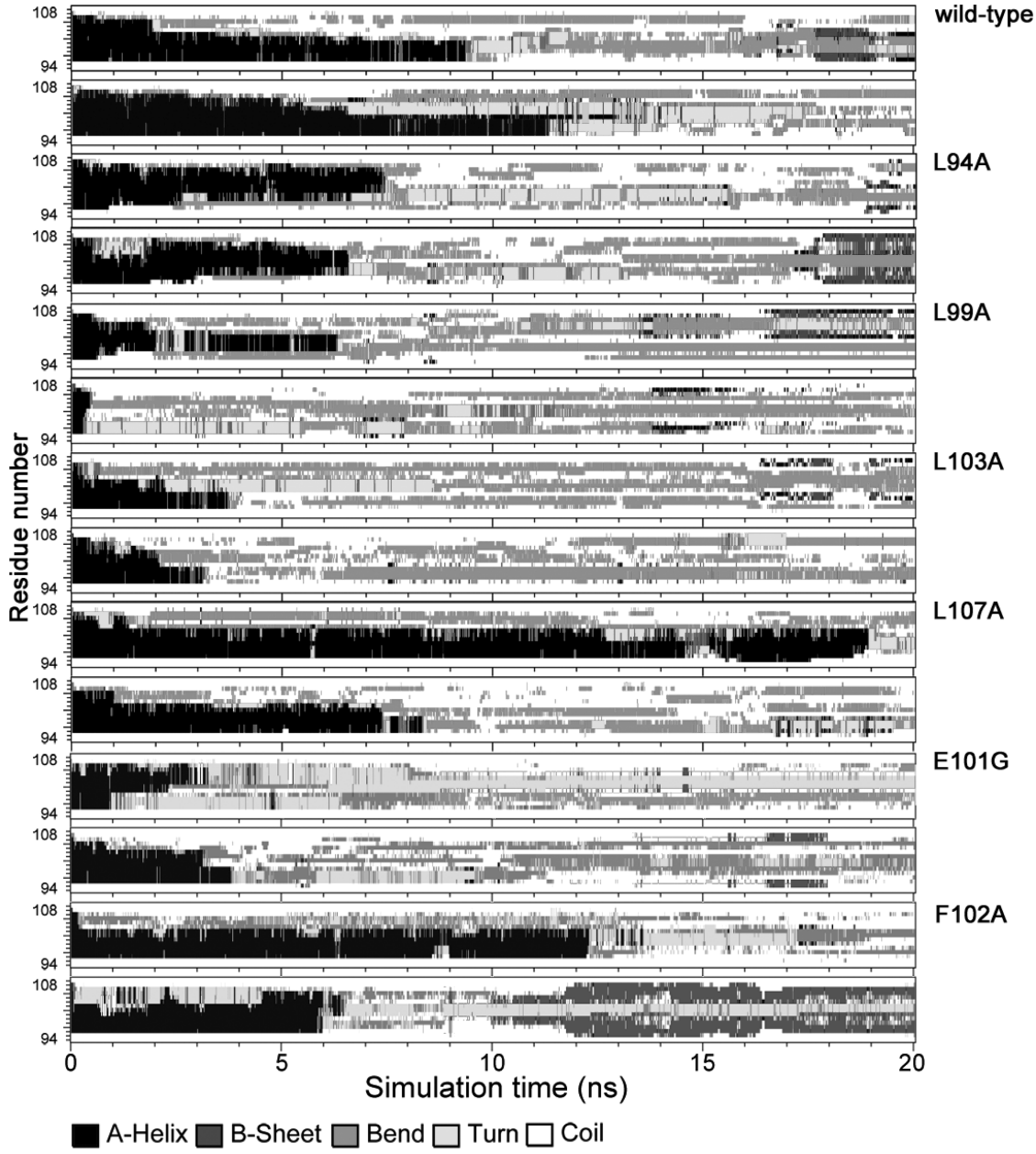

FIGURE 8: Helix propensity of isolated peptides TRX94−108 studied by molecular dynamics simulations (MDS). For each peptide, the results of two 20 ns runs are plotted. The different types of secondary structure were calculated for each residue with the DSSP program according to Kabsch and Sander criteria (*67*) and are indicated in shades of gray at the bottom of the figure.

Information). Interestingly, the occurrence of F at position 99 would be admissible, in view of the fact that a hydrophobic, rigid and bulky side chain here would lie at the boundary of the largest cavity known to exist in TRX (∼ 45 Å^3). Suggestively, the L99F replacement would only result in a marginally higher calculated energy, as predicted by FoldX (data not shown). On the other hand, the AGADIR program predicts this same pair of leucine residues as structural determinants putatively conferring helical character to peptide TRX94−108. Specifically, residues L99 and L103 appear to be involved in (i) intrahelical stabilization and (ii) surface interaction between fragment TRX1−93 and peptide TRX94−108.

Most probably, mutations by alanine of leucine residues of α-helix 5 would create cavities. For mutations L94A and L107A, the void space would occur on the protein surface, because these residues are partially solvent exposed in TRX (exhibiting 8.1 and 16.7% accessibility, respectively: Figure 10, panel A). By contrast, for the mutants L99A and L103A, in which the replaced leucine residues are totally buried (0.7 and 0% accessibility, respectively: Figure 10, panel A) the putative cavities would be buried in the core region (a pre-existing cavity increases by 155 Å^3 in the former case, and a new one -amounting to 158 Å^3 appears in the latter, both calculated on the 2TRX structure). To investigate this matter further, the corresponding full-length mutants were prepared to study the effect of side-chain suppression at these positions on the overall TRX structure and how this perturbation could presumably be compensated by rearrangements in the tertiary structure. In agreement with the results presented herein, full-length mutants TRX L99A and L103A are significantly destabilized ($\Delta\Delta G$ = 3.0 and 3.7 kcal mol^{-1}, respectively, with respect to wild type TRX; Ana Garrote, Bruno Manta, José M. Delfino, and Javier Santos, unpublished results). In light of these facts, in the context of the complex TRX1−93/TRX94−108, mutations L99A and L103A could bring about

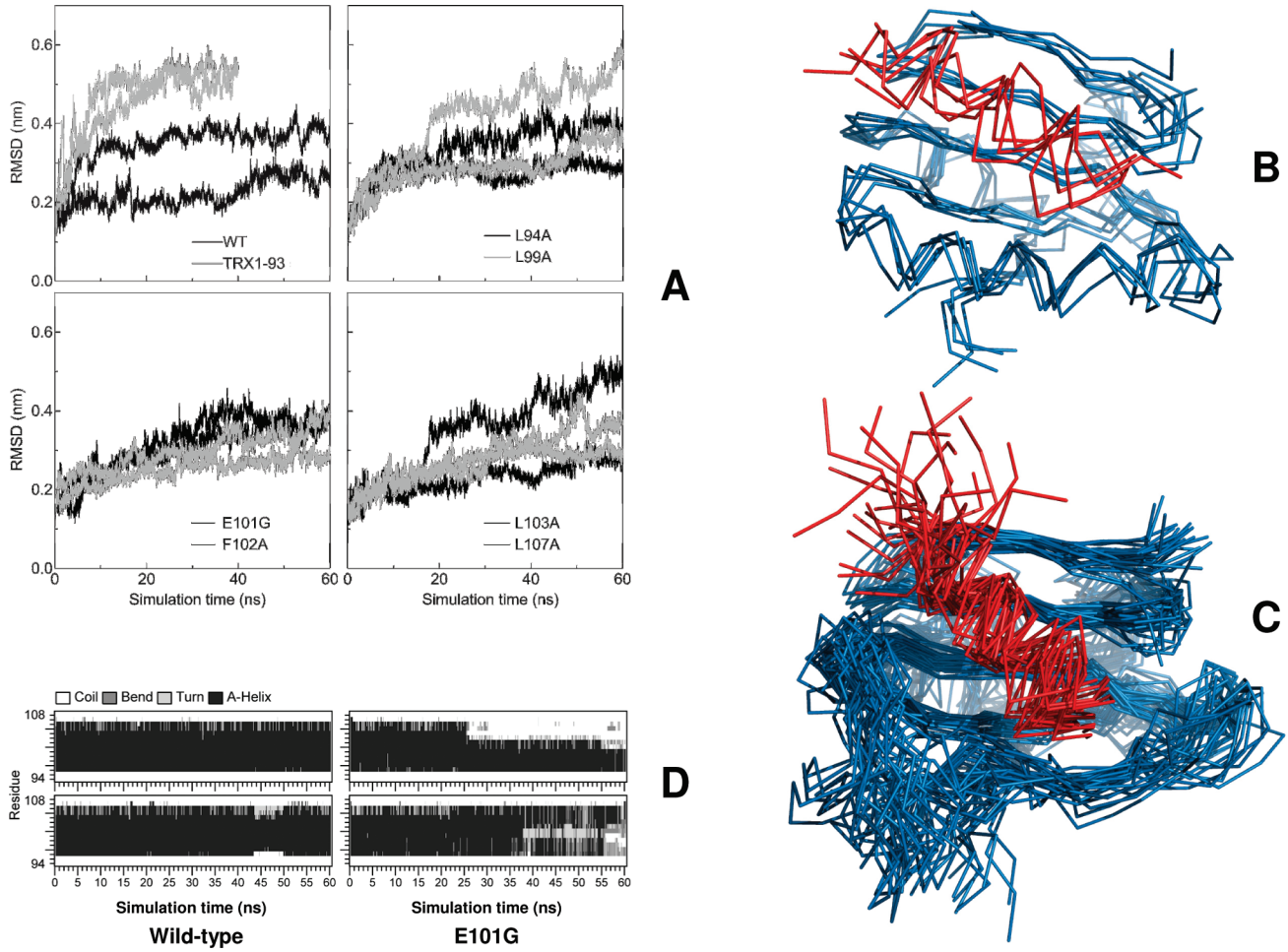

FIGURE 9: Molecular dynamics simulations (MDS) of the TRX1−93/TRX94−108 complexes. Two MDS runs of 60 ns each were carried out for each complex, and two additional runs for isolated fragment TRX1−93 in its reduced form were carried out as a control sample. (A) Plot of the time-course of the heavy atom rmsd (in nm) with respect to the crystallographic structure. Representative clusters of superimposed structures along the simulation of wild-type (B) and L99A (C) complexes. (D) A secondary structure analysis of wild-type peptide TRX94−108 or mutant E101G complexed to TRX1−93 along two 60 ns runs are plotted (see legend to Figure 8).

a similar destabilization, decreasing significantly the population of native-like conformations, as observed in experiments (Figure 2 and Figure 3). On the other hand, the estimations arising from the application of the FoldX algorithm (*58*) also suggest a key contribution of residues L99A and L103A to the stabilization of the complex (Figure 10, panel C).

In addition to the role of these leucine residues, F102 could add to complex stability as well. The phenyl side-chain lies in close contact with residues A88, K90, Q98, and N106, in the center of a superficial net, including both aliphatic and amino−aromatic interactions (*6*). Accordingly, the F102A mutant complex is marginally stable and exhibits residual enzymic activity (Figure 4). Furthermore, mutation F102A significantly reduces the α-helix propensity of peptide TRX94−108 in solution, suggesting that this residue would be implicated in both long-range and local interactions. Consistently, the FoldX algorithm also predicts F102 as a key residue for complex stabilization (Figure 10, panel C).

By contrast to central leucine mutations, the L94A mutant complex appears much less compromised in its stability with respect to the wild-type complex ($\Delta\Delta G$ = 1.9 kcal mol^{-1}: Figure 5), a fact consistent with its less prominent role for helix stabilization or tertiary structure consolidation as predicted from the MDS analysis, the larger extent of solvent exposure of this residue, and its less strict conservation.

In the case of the complex involving the E101G mutation—as expected from predictions from MDS, the MC conformational search and AGADIR—the experimental results indicate a significant destabilization of the helical structure, as a consequence of the expected added conformational freedom. Nevertheless, results show that this mutation does not hamper complex formation, probably because of the outward orientation of this residue.

Interestingly, the sequence of the central part of α-helix 5—that contains the conserved residues L99 and L103 (peptide sequence LKEFL)—is present in other folds (1JIH, 1P1I, 1QGH, 1U6G, 1WG8, 1YDW, 2C2A, and 2QP2), where in all cases this peptide is either embedded in a longer helix or forms by itself a short helix turn. Consistent with this observation, the MC conformational search of peptide LKEFL shows that first, the minimum energy structure adopts a helical conformation with the leucine side-chains in close contact, and second, these structures appear to be vastly represented in the conformational space (Figure 7).

In particular, we paid special attention to 1JIH, the yeast DNA polymerase ETA. Even though the thumb domain of this protein bears no structural homology to TRX, it contains a nine-residue α-helical element (α-helix 20: GQLKEFLDA, Figure 11) identical in sequence and structurally very similar to α-helix 5 in TRX (backbone rmsd = 0.22 Å). Subse-

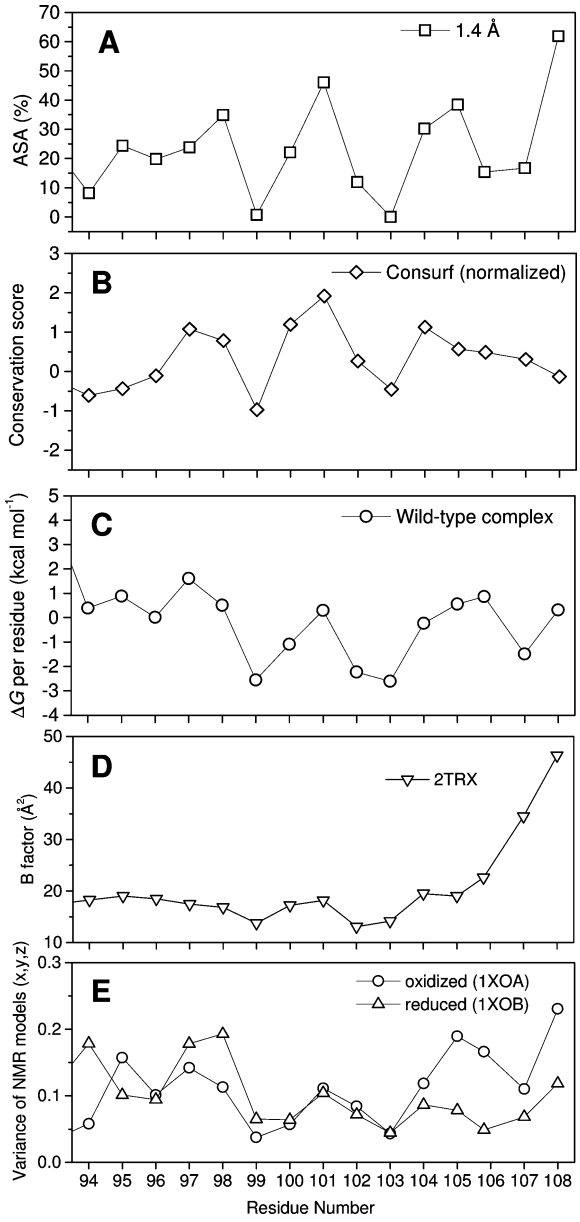

FIGURE 10: Structural analyses of peptide TRX94−108: (A) solvent accessible surface area (ASA) per residue as % of the total surface of each residue, calculated from 2TRX with the MOLMOL software; (B) the conservation scores (normalized, where high values represent low conservation) calculated from 456 homologue sequences with the ConSurf web server; (C) the predicted contribution of each residue to the stability of the wild-type complex (where negative ΔG values represent higher stability) using the FoldX web server; (D) *B* factors from the 2TRX crystallographic structure; (E) variance of Cα positions calculated for oxidized and reduced NMR models of TRX.

quently, we speculate that in the context of the polymerase this helix should likely be stabilized by tertiary interactions, as is the case of α-helix 5 of TRX that contacts α-helix 3. Indeed, the thumb domain of the polymerase includes a second α-helix in close contact (α-helix 21), but contrary to TRX, here this element is oriented in an opposite direction (see Figure 11). More intriguing is that α-helix 21 in DNA polymerase ETA shares "retro similarity" to α-helix 3 of TRX (Figure 11). Perhaps this allows us to postulate a way of helix packing leading to stabilization. This concept might result useful in protein engineering, but more importantly, it might permit to infer the existence of a natural structural

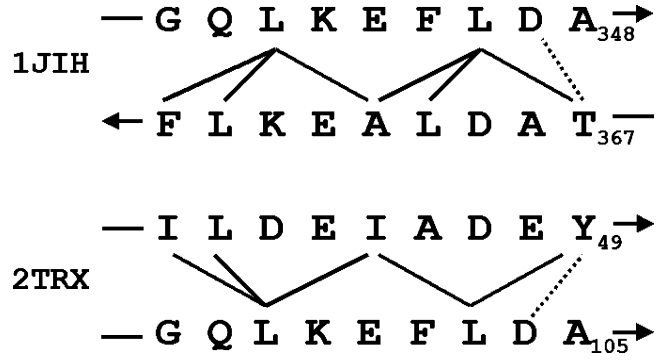

FIGURE 11: The amino acid sequence GQLKEFLDA constitutes a helical element in both TRX (2TRX) and the yeast DNA polymerase ETA (1JIH). In the former, α-helix 5 is presumably stabilized by tertiary interactions involving amino acid residues belonging to α-helix 3. In 1JIH, α-helix 20 lies in close contact to α-helix 21. Putative stabilizing contacts for this element are shown. Solid lines (—) identify apolar residues that are in close contact (at a distance ≤4.5Å), and amino acid residues H-bonded through their side-chains are indicated with dotted lines (•••). Arrows indicate the C-terminal end of α-helices.

motif for folding of an unstable amphipathic helix. It might be possible that mechanisms like this could underlie the origin of cooperativity in autonomous folding units. Biophysical experiments will be designed to evaluate the role of these α-helix/α-helix interactions in the case of TRX as well as in the case of the short 1JIH α-helical thumb domain.

It is a fact that isolated fragment TRX1−93, in its *reduced* form, is unable to adopt a proper tertiary fold. Nevertheless, we wonder whether this fragment can ever *efficiently* follow a path to a native-like topology (other than interacting with the cognate peptide TRX94−108). In this regard, an unexpected observation on the structure of *M. tuberculosis* thioredoxin adds substance to this matter. Here, the crystal packing causes that a five residue C-terminal stretch belonging to helix 5 fits into a groove adjacent to the active site of a neighboring thioredoxin molecule. In this case, the C-terminal part of helix 5 is somewhat unstructured and authors suggest that crystal packing forces would contribute to the stabilization of this unfavorable conformation (*65*). This would indicate that the TRX fold could in principle tolerate the presence of partially unfolded α-helix 5. Furthermore, it is important to point out that the *oxidized* form of TRX1−93 can fold into a conformation exhibiting ∼50% of the near-UV CD signal characteristic of the native state of TRX (Javier Santos, unpublished results). It is well-known that the S−S link stabilizes full-length TRX in 2.4 kcal mol^{-1} (by comparison, the reported free energies of unfolding are 6.2 and 8.6 kcal mol^{-1} for the reduced and the oxidized forms, respectively (*66*)). In the case of *M. tuberculosis* thioredoxin structure, the frustration may be compensated by the extra stability conferred by favorable crystal packing interactions, whereas in the case of the oxidized form of fragment TRX1−93, partial stabilization of the native structure (or destabilization of the unfolded state) might result as a consequence of disulfide formation at the close pair C32/C35. More experiments should be carried out to determine if fragment TRX1−93 could be properly engineered to minimize frustration. We speculate that the energetic landscape of fragment TRX1−93 might be represented by a flat funnel where a molten globule conformation populates the dominating basin. This funnel would deepen when, for

instance, specific interactions, such as those satisfied by peptide TRX94−108, contribute with added stability to the consolidation of the native state. Importantly, our first round of NMR experiments ([^{15}N,^{1}H]-HSQC) also points to the notion that structuring occurs upon the addition of peptide TRX94−108 to the TRX1−93 fragment, increasing dramatically the spread of the NH peaks, thus suggesting progress of the folding reaction toward the native-like state of TRX (Javier Santos, Andrés Binolfi, Claudio Fernández, and José M. Delfino, unpublished results). This evidence taken together, one might state that sequence TRX1−93 would include the signature of the native structure, but the efficiency of the folding process would be highly dependent on the action of peptide TRX94−108. In other words, a kinetic trapping of a partially folded form represented by fragment TRX1−93 could be circumvented by reducing the energetic barrier that separates it from the native state through the addition of peptide TRX94−108.

In summary, in the reduced form fragment TRX1−93 would represent a partially folded state. The picture that emerges from MDS points to a mobile TRX scaffold where α-helix 3 is disordered. This prediction agrees with spectroscopic and functional data presented herein, indicating that fragment TRX1−93 would present significant conformational alterations in the active site, a region that is immediately connected to α-helix 3. Specific interactions—like those involving L99, F102 and L103—between peptide TRX94−108 and partially folded TRX1−93 play a pivotal role to stabilize the native-like tertiary structure of the complex.

ACKNOWLEDGMENT

The authors thank Drs. Mariano Grasselli for his gift of di-FTC-insulin, Bruno Manta for his helpful comments and suggestions, and Mariano C. González-Lebrero for prolific discussion and valuable insights in the conformational search of the peptides.

SUPPORTING INFORMATION AVAILABLE

The occurrence of amino acid residues at positions 99 and 103 of thioredoxins was evaluated after examining a multiple sequence alignment (MSA) constructed with the sequence of TRX as seed of a PSI-BLAST search. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Tsai, C. J., and Nussinov, R. (2001) Transient, highly populated, building blocks folding model. *Cell Biochem. Biophys. 34*, 209–235.
2. Haspel, N., Tsai, C. J., Wolfson, H., and Nussinov, R. (2003) Reducing the computational complexity of protein folding via fragment folding and assembly. *Protein Sci. 12*, 1177–1187.
3. Gong, H., Fleming, P. J., and Rose, G. D. (2005) Building native protein conformation from highly approximate backbone torsion angles. *Proc. Natl. Acad. Sci. U.S.A. 102*, 16227–16232.
4. Fleming, P. J., Gong, H., and Rose, G. D. (2006) Secondary structure determines protein topology. *Protein Sci. 15*, 1829–1834.
5. Rose, G. D., Fleming, P. J., Banavar, J. R., and Maritan, A. (2006) A backbone-based theory of protein folding. *Proc. Natl. Acad. Sci. U.S.A. 103*, 16623–16633.
6. Katti, S. K., LeMaster, D. M., and Eklund, H. (1990) Crystal structure of thioredoxin from Escherichia coli at 1.68 Å resolution. *J. Mol. Biol. 212*, 167–184.
7. Dyson, H. J., Gippert, G. P., Case, D. A., Holmgren, A., and Wright, P. E. (1990) Three-dimensional solution structure of the reduced form of Escherichia coli thioredoxin determined by nuclear magnetic resonance spectroscopy. *Biochemistry 29*, 4129–4136.
8. Chandrasekhar, K., Krause, G., Holmgren, A., and Dyson, H. J. (1991) Assignment of the 15N NMR spectra of reduced and oxidized Escherichia coli thioredoxin. *FEBS Lett. 284*, 178–183.
9. Jeng, M. F., Campbell, A. P., Begley, T., Holmgren, A., Case, D. A., Wright, P. E., and Dyson, H. J. (1994) High-resolution solution structures of oxidized and reduced Escherichia coli thioredoxin. *Structure 2*, 853–868.
10. Arner, E. S., and Holmgren, A. (2006) The thioredoxin system in cancer. *Semin. Cancer Biol. 16*, 420–426.
11. Kakolyris, S., Giatromanolaki, A., Koukourakis, M., Powis, G., Souglakos, J., Sivridis, E., Georgoulias, V., Gatter, K. C., and Harris, A. L. (2001) Thioredoxin expression is associated with lymph node status and prognosis in early operable non-small cell lung cancer. *Clin. Cancer Res. 7*, 3087–3091.
12. Powis, G., and Montfort, W. R. (2001) Properties and biological activities of thioredoxins. *Annu. Rev. Biophys. Biomol. Struct. 30*, 421–455.
13. Pekkari, K., and Holmgren, A. (2004) Truncated thioredoxin: physiological functions and mechanism. *Antioxid. Redox Signaling 6*, 53–61.
14. Mkrtchian, S., and Sandalova, T. (2006) ERp29, an unusual redox-inactive member of the thioredoxin family. *Antioxid. Redox Signaling 8*, 325–337.
15. Liepinsh, E., Baryshev, M., Sharipo, A., Ingelman-Sundberg, M., Otting, G., and Mkrtchian, S. (2001) Thioredoxin fold as homodimerization module in the putative chaperone ERp29: NMR structures of the domains and experimental model of the 51 kDa dimer. *Structure 9*, 457–471.
16. Gleason, F. K., and Holmgren, A. (1988) Thioredoxin and related proteins in procaryotes. *FEMS Microbiol. Rev. 4*, 271–297.
17. Jaeger, T. (2007) Peroxiredoxin systems in mycobacteria. *Subcell. Biochem. 44*, 207–217.
18. Deponte, M., Rahlfs, S., and Becker, K. (2007) Peroxiredoxin systems of protozoal parasites. *Subcell. Biochem. 44*, 219–229.
19. Godoy-Ruiz, R., Perez-Jimenez, R., Ibarra-Molero, B., and Sanchez-Ruiz, J. M. (2005) A stability pattern of protein hydrophobic mutations that reflects evolutionary structural optimization. *Biophys. J. 89*, 3320–3331.
20. Hellinga, H. W., Wynn, R., and Richards, F. M. (1992) The hydrophobic core of Escherichia coli thioredoxin shows a high tolerance to nonconservative single amino acid substitutions. *Biochemistry 31*, 11203–11209.
21. Jeng, M. F., and Dyson, H. J. (1995) Comparison of the hydrogen-exchange behavior of reduced and oxidized Escherichia coli thioredoxin. *Biochemistry 34*, 611–619.
22. Kelley, R. F., and Stellwagen, E. (1984) Conformational transitions of thioredoxin in guanidine hydrochloride. *Biochemistry 23*, 5095–5102.
23. Ladbury, J. E., Kishore, N., Hellinga, H. W., Wynn, R., and Sturtevant, J. M. (1994) Thermodynamic effects of reduction of the active-site disulfide of Escherichia coli thioredoxin explored by differential scanning calorimetry. *Biochemistry 33*, 3688–3692.
24. Maier, C. S., Schimerlik, M. I., and Deinzer, M. L. (1999) Thermal denaturation of Escherichia coli thioredoxin studied by hydrogen/deuterium exchange and electrospray ionization mass spectrometry: monitoring a two-state protein unfolding transition. *Biochemistry 38*, 1136–1143.
25. Perez-Jimenez, R., Godoy-Ruiz, R., Ibarra-Molero, B., and Sanchez-Ruiz, J. M. (2005) The effect of charge-introduction mutations on E. coli thioredoxin stability. *Biophys. Chem. 115*, 105–107.
26. Georgescu, R. E., Garcia-Mira, M. M., Tasayco, M. L., and Sanchez-Ruiz, J. M. (2001) 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*, 1477–1485.
27. Dangi, B., Dobrodumov, A. V., Louis, J. M., and Gronenborn, A. M. (2002) Solution structure and dynamics of the human-Escherichia coli thioredoxin chimera: insights into thermodynamic stability. *Biochemistry 41*, 9376–9388.
28. Louis, J. M., Georgescu, R. E., Tasayco, M. L., Tcherkasskaya, O., and Gronenborn, A. M. (2001) Probing the structure and stability of a hybrid protein: the human-E. coli thioredoxin chimera. *Biochemistry 40*, 11184–11192.
29. Ghoshal, A. K., Swaminathan, C. P., Thomas, C. J., Surolia, A., and Varadarajan, R. (1999) Thermodynamic and kinetic analysis of the Escherichia coli thioredoxin-C′ fragment complementation system. *Biochem. J. 339*, 721–727.

30. Chaffotte, A. F., Li, J. H., Georgescu, R. E., Goldberg, M. E., and Tasayco, M. L. (1997) Recognition between disordered states: kinetics of the self-assembly of thioredoxin fragments. *Biochemistry 36*, 16040–16048.
31. Santos, J., Marino-Buslje, C., Kleinman, C., Ermacora, M. R., and Delfino, J. M. (2007) Consolidation of the thioredoxin fold by peptide recognition: interaction between E. coli thioredoxin fragments 1−93 and 94−108. *Biochemistry 46*, 5148–5159.
32. Baldwin, R. L., and Rose, G. D. (1999) Is protein folding hierarchic? I. Local structure and peptide folding. *Trends Biochem. Sci. 24*, 26–33.
33. Luo, P., and Baldwin, R. L. (2002) Origin of the different strengths of the (i,i+4) and (i,i+3) leucine pair interactions in helices. *Biophys. Chem. 96*, 103–108.
34. Klingler, T. M., and Brutlag, D. L. (1994) Discovering structural correlations in alpha-helices. *Protein Sci. 3*, 1847–1857.
35. Chiu, J., Tillett, D., and March, P. E. (2006) Mutation of Phe102 to Ser in the carboxyl terminal helix of Escherichia coli thioredoxin affects the stability and processivity of T7 DNA polymerase. *Proteins 64*, 477–485.
36. Huber, D., Cha, M. I., Debarbieux, L., Planson, A. G., Cruz, N., Lopez, G., Tasayco, M. L., Chaffotte, A., and Beckwith, J. (2005) A selection for mutants that interfere with folding of Escherichia coli thioredoxin-1 in vivo. *Proc. Natl. Acad. Sci. U.S.A. 102*, 18872–18877.
37. Schägger, H., and von Jagow, G. (1987) Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal. Biochem. 166*, 368–379.
38. Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: visual molecular dynamics. *J. Mol. Graphics 14*, 33–38, 27−28.
39. Koradi, R., Billeter, M., and Wuthrich, K. (1996) MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graphics 14*, 51–5.
40. Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S. e., Wilkins, M. R., Appel, R. D., and Bairoch, A. (2005) Protein Identification and Analysis Tools on the ExPASy Server, pp 571−607.
41. Wen, J., Arakawa, T., and Philo, J. S. (1996) Size-Exclusion Chromatography with On-Line Light-Scattering, Absorbance, and Refractive Index Detectors for Studying Proteins and Their Interactions. *Anal. Biochem. 240*, 155–166.
42. Fersht, A. (1999) *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding*; Freeman and Company: New York.
43. Heuck, A. P., and Wolosiuk, R. A. (1997) Di-fluoresceinthiocarbamyl-insulin: A Fluorescent substrate for the Assay of Protein Disulfide Oxidoreductase Activity. *Anal. Biochem. 248*, 94–101.
44. Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A. E., and Berendsen, H. J. (2005) GROMACS: fast, flexible, and free. *J. Comput. Chem. 26*, 1701–1718.
45. Oostenbrink, C., Villa, A., Mark, A. E., and van Gunsteren, W. F. (2004) A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6. *J. Comput. Chem. 25*, 1656–1676.
46. Berendsen, H. J. C., Grigera, J. R., and Straatsma, T. P. (1987) The missing term in effective pair potentials. *J. Phys. Chem. 91*, 6269–6271.
47. Berendsen, H. J. C., Postma, J. P. M., Gunsteren, W. F. v., DiNola, A., and Haak, J. R. (1984) Molecular dynamics with coupling to an external bath. *J. Chem. Phys. 81*, 3684–3690.
48. Berendsen, H. J. C., and van Gunsteren, W. F. (1984) Molecular dynamics simulations: Techniques and approaches, in *Molecular Liquids-Dynamics and Interactions* (Orville-Thomas, W. J., Barnes, A. J., and Yarwood, J., Eds.) pp 475−500, Reidel, Dordrecht, Florence (Italy).
49. Hess, B., Bekker, H., Berendsen, H. J. C., and Fraaije, J. G. E. M. (1997) LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem. 18*, 1463–1472.
50. Feenstra, K. A., Hess, B., and Berendsen, H. J. C. (1999) Improving efficiency of large time-scale molecular dynamics simulations of hydrogen-rich systems. *J. Comput. Chem. 20*, 786–798.
51. Tironi, I. G., Sperb, R., Smith, P. E., and Gunsteren, W. F. v. (1995) A generalized reaction field method for molecular dynamics simulations. *J. Chem. Phys. 102*, 5451–5459.
52. Daura, X., Gademann, K., Jaun, B., Seebach, D., van Gunsteren, W. F., and Mark, A. E. (1999) Peptide Folding: When Simulation Meets Experiment. *Angew. Chem., Int. Ed. 38*, 236–240.
53. Mohamadi, F., Richards, M. G. J., Guida, W. C., Liskamp, R., Lipton, M., Craig, C., George, C., Thomas, H., and Still, W. C. (1990) MacroModel−an integrated software system for modeling organic and bioorganic molecules using molecular mechanics. *J. Comput. Chem. 11*, 440–467.
54. Chang, G., Guida, W. C., and Still, W. C. (1989) An internal-coordinate Monte Carlo method for searching conformational space. *J. Am. Chem. Soc. 111*, 4379–4386.
55. Jorgensen, W. L., Maxwell, D. S., and Tirado-Rives, J. (1996) Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J. Am. Chem. Soc. 118*, 11225–11236.
56. Evans, J. S., Chan, S. I., and Goddard-Iii, W. A. (1995) Prediction of polyelectrolyte polypeptide structures using Monte Carlo conformational search methods with implicit solvation modeling. *Protein Sci. 4*, 2019–2031.
57. Evans, J. S., Mathiowetz, A. M., Chan, S. I., and Goddard-Iii, W. A. (1995) De novo prediction of polypeptide conformations using dihedral probability grid Monte Carlo methodology. *Protein Sci. 4*, 1203–1216.
58. Schymkowitz, J., Borg, J., Stricher, F., Nys, R., Rousseau, F., and Serrano, L. (2005) The FoldX web server: an online force field. *Nucleic Acids Res. 33*, 382–388.
59. Guerois, R., Nielsen, J. E., and Serrano, L. (2002) Predicting changes in the stability of proteins and protein complexes: a study of more than 1000 mutations. *J. Mol. Biol. 5*, 369–387.
60. Krause, G., and Holmgren, A. (1991) Substitution of the conserved tryptophan 31 in Escherichia coli thioredoxin by site-directed mutagenesis and structure-function analysis. *J. Biol. Chem. 266*, 4056–4066.
61. Luo, P., and Baldwin, R. L. (1997) Mechanism of helix induction by trifluoroethanol: a framework for extrapolating the helix-forming properties of peptides from trifluoroethanol/water mixtures back to water. *Biochemistry 36*, 8413–8421.
62. Roccatano, D., Colombo, G., Fioroni, M., and Mark, A. E. (2002) Mechanism by which 2,2,2-trifluoroethanol/water mixtures stabilize secondary-structure formation in peptides: a molecular dynamics study. *Proc. Natl. Acad. Sci. U.S.A. 99*, 12179–12184.
63. Fisinger, S., Serrano, L., and Lacroix, E. (2001) Computational estimation of specific side chain interaction energies in alpha helices. *Protein Sci. 10*, 809–818.
64. Lacroix, E., Viguera, A. R., and Serrano, L. (1998) Elucidating the folding problem of alpha-helices: local motifs, long-range electrostatics, ionic-strength dependence and prediction of NMR parameters. *J. Mol. Biol. 284*, 173–191.
65. Hall, G., Shah, M., McEwan, P. A., Laughton, C., Stevens, M., Westwell, A., and Emsley, J. (2006) Structure of Mycobacterium tuberculosis thioredoxin C. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 62*, 1453–1457.
66. Kelley, R. F., Shalongo, W., Jagannadham, M. V., and Stellwagen, E. (1987) Equilibrium and kinetic measurements of the conformational transition of reduced thioredoxin. *Biochemistry 26*, 1406–1411.
67. Kabsch, W., and Sander, C. (1983) Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers 22*, 2577–2637.

BI801969W