Consolidation of the Thioredoxin Fold by Peptide Recognition: Interaction between *E. coli* Thioredoxin Fragments 1−93 and 94−108[†]

Javier Santos,[‡,§] Cristina Marino-Buslje,[‡] Claudia Kleinman,[||] 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, Department of Science and Technology, University of Quilmes, Roque Sáenz Peña 352, B1876XD Bernal, Argentina, and Canadian Institute for Advanced Research, Département de Biochimie, Université de Montréal, Montréal, Québec, Canada

Received December 21, 2006; Revised Manuscript Received March 1, 2007

ABSTRACT: *Escherichia coli* thioredoxin (TRX) catalyzes redox reactions via the reversible oxidation of the conserved active center WCGPC. TRX is a monomeric α/*β* protein with a fold characterized by a central *β*-sheet surrounded by α-helical elements. The interaction of the C-terminal α-helix (helix 5) of TRX against the remainder of the protein involves the close packing of hydrophobic surfaces, opening the possibility of studying a fine-tuned molecular recognition phenomenon. To evaluate the relevance of this interaction on the folding mechanism of TRX, we characterize TRX1−93, a truncated variant of TRX devoid of the last stretch of 15 amino acid residues that includes helix 5. TRX1−93 may possibly represent a molecular form where the folding process becomes interrupted, giving rise to a structure exhibiting the features of a molten globule state. This was assessed by circular dichroism, intrinsic fluorescence, binding of the probe ANS, size-exclusion chromatography, limited proteolysis, and calorimetry. Remarkably, fragment TRX1−93 interacts with peptide TRX94−108 (K_D ∼2−12 *μ*M), bringing forth the restoration of native-like signatures and enzymic function. This represents a molecular event of reciprocal structure selection where both partners gain order, thus leading to long-range consequences on conformation. In this context, the binding of the C-terminal helix could signify a late event in the consolidation of the overall TRX fold.

Understanding how the amino acid sequence encodes the three-dimensional (3D) protein structure remains one of the main challenges of biophysics, and much effort has been devoted to devising suitable experimental systems to address this problem. In this regard, the association of peptides leading to native proteins focuses folding studies on the interaction of complementary protein surfaces. Several truncated proteins are known to act as templates for the binding and folding of complementary peptides (*1*−*7*). For instance, a structured but inactive N-terminally truncated ribonuclease (the S protein) binds the cognate S peptide to regenerate a folded and active ribonuclease (RNase S) that differs from the wild type by having one hydrolyzed peptide bond (*8*). Productive association, in a mutually induced binding/folding process, can also take place between protein fragments that would remain unfolded if in isolation. Related issues are the specificity of the complementary interactions and the tolerance of the adjusting surfaces to mutations (*5*, *7*, *9*−*11*).

Thioredoxins (TRXs[1]) are proteins of about 100 residues with a distinctive α/*β* topology. They participate in diverse redox reactions via the reversible oxidation of the conserved active center WCGPC. In addition, *E. coli* TRX is essential for phage T7 DNA replication as a subunit of the T7 DNA polymerase complex and also for the assembly of the filamentous phages f1 and M13 (*12*). *E. coli* TRX has been thoroughly characterized by X-ray diffraction and by NMR methods (*13*−*16*). It has also been used extensively as a model in protein engineering to understand the basis of protein stability (*17*−*22*).

The equilibrium unfolding of *E. coli* TRX is well described by a two-state model (*23*, *24*). In addition, the unfolding kinetics also appears to be a two-state process, without an appreciable accumulation of intermediates (*23*). On the contrary, the refolding kinetics is more complex, where multiple phases become apparent (*23*, *25*). This process would occur via a burst-phase intermediate with an elevated content of *β*-structure that folds to the native state through

[†] This research was supported by grants from the University of Buenos Aires (UBA), Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Agencia Nacional de Promoción Científica y Tecnológica (ANPCyT), National University of Quilmes (UNQ), and Comisión Nacional de Programas de Investigación Sanitaria (CONAPRIS).

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

[‡] University of Buenos Aires.

[§] University of Quilmes.

[||] Université de Montréal.

[1] Abbreviations: ANS, 8-anilino-1-naphthalene sulfonic acid; ASA, accessible surface area; *fos*, the fraction of occupied sites; CD, circular dichroism; Di-FTC-insulin, di-fluoresceinthiocarbamyl-insulin; DSS, the homobifunctional cross-linker disuccinimidyl suberate; MSA, multiple sequence alignment; PAGE, polyacrylamide gel electrophoresis; RP-HPLC; reversed-phase high-performance liquid chromatography; SEC, size-exclusion chromatography; TRX, thioredoxin; TRX1-93, the recombinant fragment spanning sequence 1−93 of full-length TRX; TRX94-108, the synthetic peptide comprising residues 94−108 of TRX plus a C-terminal tyrosine tag. (The numbering scheme used in the entry pdb 2trx is also used throughout the article.)

10.1021/bi6026264 CCC: $37.00
Published on Web 04/07/2007

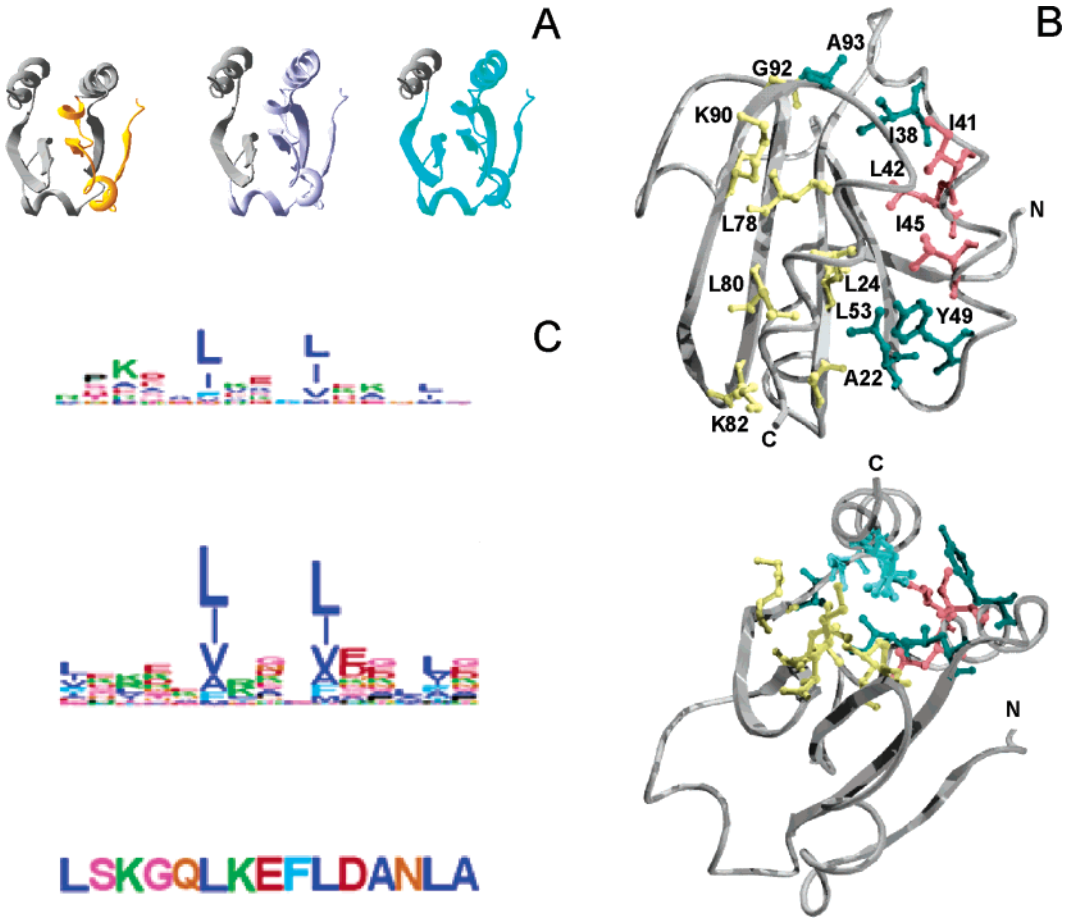

FIGURE 1: (A) Complementation systems based on TRX: ribbon representation of complexes involving interacting fragments that give rise to native-like structures. From left to right: TRX1−37/TRX38−108, TRX1−73/TRX74−108, and TRX1−93/TRX94−108, on the basis of the known structure of full-length TRX1−108 (2trx). (B) Area of interaction between α-helix 5 and the LI surface of TRX. The top figure shows a ribbon diagram with side chains of amino acids in close contact (atom-to-atom distance <4.5 Å) with L94, L99, L103, and L107 (the L row, not shown). The bottom figure shows the above model rotated 90° along the *x*-axis, where the L row is shown in light blue. Yellow, red, and blue indicate the location of relevant side chains in the *β*-, α-, and turn regions. N and C represent the N- and C-termini, respectively. (C) Sequence logos of TRX94−108, generated from a multiple sequence alignment of 283 unique amino acid sequences (top) and from 1,000 amino acid sequences compatible with the TRX1−108 structure, as obtained by Gibbs sampling using a pairwise contact plus the solvent-accessibility potential (center). At the bottom, the sequence of TRX94−108 is shown as a reference.

multiple routes (*26*, *27*). Finally, hydrogen-exchange experiments on the native state suggest the subsistence of *β*-structure in a high-energy intermediate preceding the globally unfolded state or more likely in the globally unfolded state itself (*24*).

Conspicuously, *E. coli* TRX has offered a plastic framework for complementation assays. At least two combinations of fragments are productive: TRX1−37/TRX38−108 and TRX1−73/TRX74−108 (Figure 1, panel A) (*4*, *6*, *7*, *28*). Furthermore, although fragments TRX1−37, TRX38−73, and TRX74−108 behave as disordered monomers when isolated in solution, native-like TRX interactions arise from combinations of these fragments in the ternary mixture or in a binary mixture containing only TRX1−37 and TRX74−108-fragments (*7*). ^{1}H-^{15}N HSQC NMR experiments have indicated that the initial step in the folding of TRX involves the interaction between the *β*2 and *β*4 strands (*6*, *7*). In agreement with these studies, hydrogen-exchange (HX) results support the notion of a highly stable *β*-sheet, comprising strands *β*2, *β*3, and *β*4, forming the core of the protein and contributing to the residual structure persisting in the unfolded state (*24*).

In this context, the packing of the C-terminal α-helix of TRX (helix 5) against the remainder of the protein offers the possibility of probing a series of interactions that might play key roles in stabilizing the native fold, a fact that makes it a particularly interesting topic for study. Among other residues, L24, L42, L53, L80, I38, I41, and I45, belonging to the *β*-sheet floor (strands 2, 3, and 4) and neighboring helix 2 (herein referred to as the LI surface), lie in close contact (atom-to-atom distance <4.5 Å) to the four leucine residues belonging to the C-terminal α-helix: L94, L99, L103, and L107 (herein called the L row). The latter arrangement conforms a cluster of hydrophobic interactions on one side of helix 5 (Figure 1, panel B). When I38, I41, or I45 are separately replaced by valine residues, changes in the parameters of stability are observed. The authors speculate that the optimization of the environment of these three residues is the reflection of an evolutionary process (*21*). Most likely, these changes will bring about the occurrence of packing defects in the environment of the hydrophobic core of the protein, thus compromising its stability. Nevertheless, Hellinga et al. demonstrated that positions L42 and L78 in the hydrophobic core of *E. coli*

TRX are highly tolerant to substitutions, even by charged residues (*29*). However, the least stable mutant (L78D) in their set still shows a $\Delta G°$ of 8.7 kcal/mol, highlighting the extreme stability of this protein. In this context, the extended thermodynamic range for accommodating mutations makes it difficult to evaluate the informational content at particular positions along the amino acid sequence.

In the work presented herein, we characterize TRX1−93, a TRX fragment comprising the first 93 residues of the full protein, thus devoid of the last stretch of 15 amino acid residues comprising helix 5. Indeed, we were able to demonstrate that this fragment interacts with synthetic TRX94−108. Even though these two fragments do not show native-like properties in isolation, they interact with each other very efficiently to yield a native-like complex, thus suggesting that binding can select the correct conformation from an ensemble of partially folded states. This system opens the possibility of dissecting key interactions leading to the consolidation of the native TRX fold.

MATERIALS AND METHODS

General Details. Protein and peptide purity were evaluated by SDS−PAGE (*30*) and RP-HPLC. Molecular graphics were prepared using Swiss PDB Viewer 3.7 (Glaxo Wellcome Experimental Research). Accessible surface area (ASA) was calculated using Surface Racer (*31*). In protein concentration measurements, 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 1,280 M^{-1} cm^{-1}. Limited trypsinolysis was performed in 25 mM TrisHCl, 200 mM NaCl, and 1 mM EDTA at pH 7.0 (buffer A) with the addition of 1 mM DTT at 25 °C and a weight ratio of protease to TRX of 1:100. Proteolytic fragments were separated by SDS−PAGE (16%) and stained by colloidal Coomassie Blue (*32*) or silver staining. Protein mass was analyzed by HPLC-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 on line 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 (solvent A: 2% acetic acid and 2% acetonitrile in water; solvent B: 2% acetic acid and 96% acetonitrile in water). The ESI-MS spectra range was 300−2000 amu, and data were deconvoluted with the program Xcalibur provided with the instrument.

Protein and Peptide Preparation. The cDNAs for TRX and TRX1−93 were prepared by PCR mutagenesis using Platinum Pfx DNA Polymerase (Invitrogen, CA), appropriate primers (XbaITRX tttccctctagaaataattttgtttaactttaagaaggagatatacatatg, NdeITRX tttaagaaggagatatacatatgtctgataaaattattcatctg, TRX1−108 gcatggatccttaggccaggttagcgtcgaggaa, and TRX1−93 gcatggatccttatgcacccactttggttgccgc (purchased from QIAGEN)), and pTRXfus (Invitrogen, CA) as the template. PCR products were purified by PureLink (Invitrogen, CA), cut with restriction enzymes, and ligated into XbaI/BamHI of pET9a (Novagen), generating pETRX1−108 and pETRX1−93. Transformed *E. coli* BL21 (DE3) cells were grown in Luria−Bertani medium at 37 °C to A_{600nm} ∼1.0. Finally, TRX1−108 and the fragment TRX1−93 were overexpressed for 3 h after induction with 1 mM IPTG.

For the purification of TRX1−108, cells were washed with 20% sucrose, 2.5 mM EDTA, 20 mM TrisHCl at pH 7.4, centrifuged at 5,000−7,000 rpm, and subjected to osmotic shock by incubation in cold water. The isolated fluid (∼500 mL) was mixed with 50 mL of 100 mM TrisHCl at pH 7.4 and 4 °C, and loaded onto a DE52 Sepharose column equilibrated with 10 mM TrisHCl at pH 7.4. Elution was performed with increasing concentrations of NaCl up to 1.0 M. Fractions containing TRX (evaluated by SDS−PAGE and UV absorption) were loaded onto a second column, a preparative Sephadex G-100 chromatography (SEC, 93 cm × 2.7 cm) column, previously equilibrated with 20 mM TrisHCl at pH 7.0 and 200 mM NaCl. This step completes the purification from other proteins of higher molecular weight and from DNA fragments. Then, pure TRX fractions were pooled and extensively dialyzed against distilled water. Finally, the protein was lyophilized. Similar TRX purification protocols were described (*28*, *33*).

Taking advantage of its pronounced tendency to aggregate during *in vivo* protein expression, fragment TRX1−93 was purified from inclusion bodies (*34*). These were dissolved in 6.0 M urea, 10 mM DTT, and 5 mM glycine at pH 7.0 for 20 min. Then, sodium acetate (pH 4.5) was added up to a 10 mM concentration. This solution was centrifuged at 10,000 rpm for 30 min before loading it onto a fast-flow SP-Sepharose column to remove insoluble particles. Protein elution was performed with increasing concentrations of NaCl 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. The protein was eluted in a linear gradient from 40 to 60% aqueous acetonitrile (0.05% TFA). Fragment TRX1−93 typically elutes at ∼50% acetonitrile. Finally, fractions containing >90% pure TRX1−93, as judged by analytical HPLC, were lyophilized. Peptide TRX94−108, LSKGQLKEFLDANLAY, was synthesized and purified up to the desalting step by GenScript Corp. (NJ). Afterward, the peptide was purified by an HPLC system (Rainin Dynamax, NY) equipped with a reverse-phase C18 semipreparative column (Vydac) equilibrated in 0.05% aqueous TFA. TRX94−108 was eluted in a linear gradient from 30 to 45% aqueous acetonitrile and 0.05% TFA (the peptide typically elutes at ∼40% acetonitrile). Fractions containing >90% pure peptide were pooled and lyophilized. The sequence of the peptide was corroborated by ESI mass spectrometry.

Physicochemical Characterization

Size-Exclusion Chromatography. Changes in hydrodynamic volumes were monitored by chromatography on an SEC-FPLC system (Pharmacia Biotech, Sweden) equipped with a 280 nm detector and a Superdex S-200 HR 10/30 column (Pharmacia Biotech, Sweden) equilibrated at room temperature in buffer A. The choice of this chromatographic resin allowed us to confidently test for the presence of soluble aggregates in the samples. The flow rate was 0.4 or 0.5 mL/min, and the injection volume was 200 μL. Samples were centrifuged at 13,000 rpm before loading onto the column that was previously calibrated with appropriate molecular weight markers (*35*).

Fluorescence Measurements. Steady-state fluorescence measurements were performed in an Aminco Bowman Series

2 spectrofluorometer operating in the ratio mode and equipped with a thermostated cell holder connected to a circulating water bath set at 20 °C. A 1 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 310−450 nm. The spectral slit-widths were set to 4 nm for both monochromators. For ANS-binding experiments, fragment TRX1−93 was refolded from 6 M urea by dialysis at 4 °C in buffer A with the addition of 1 mM DTT in the absence of peptide TRX94−108. Then, fragment TRX1−93 (10 μM) was incubated with peptide TRX94−108 (20 μM) for 30 min at room temperature, a time sufficient to recover a stable native-like CD signal. A control sample of isolated TRX1−93 was treated under the same conditions. The dye ANS (8-anilino-1-naphthalene sulfonic acid) was added to both samples at a final concentration of 50 μM. An additional control sample consisted of ANS (50 μM) added to a TRX1−108 solution (10 μM). The excitation wavelength was 350 nm, and the emission spectra were collected between 400 and 600 nm. The bandwidth used was 4.0 nm for both excitation and emission. For estimating the concentration of the dye, a value of the extinction coefficient of 4,950 M^{-1} cm^{-1} at 350 nm was used (*36*). A simple hyperbolic equation was enough to fit the binding data and estimate the dissociation constant (parameter K_D) (*37*).

CD Spectroscopy. Measurements were carried out with a Jasco J-810 spectropolarimeter calibrated with (+) 10-camphor sulfonic acid. Far and near-UV CD spectra were collected in 20 mM TrisHCl buffer, 100 mM NaF, 1.0 mM DTT, and 0.1 mM EDTA at pH 7.0 (buffer B). 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 $dmol^{-1}$. The protein concentration ranged from 7 to 52 μM.

The 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 on the Y/W environment and probes the consolidation of the tertiary structure that ensues the formation of the complex. In this case, TRX1−93 (28 μM) was incubated at room temperature for 1 h with peptide TRX94−108 at various concentrations ranging from 0 to 87 μM. Because no time dependence of the signal is observed under these conditions, the equilibrium of the complex formation reaction is assumed to be complete.

Extrapolation of the observed signal at a large excess of peptide TRX94−108 allowed us to estimate the ellipticity of the assembled complex. At a given amount of added peptide TRX94−108, the fraction of occupied sites (*fos*) was calculated as the ratio between the measured signal and that of the complex, and the concentration of free peptide TRX94−108 ($[L_{free}]$) as the difference between the total ($[L_{total}]$) and bound (($[L_{bound}]$) forms of the ligand. Then, the dissociation constant (K_D) is determined from eq 1.

$$fos = \frac{1}{\frac{K_D}{[L_{free}]} + 1} \quad (1)$$

Thermal unfolding experiments were carried out in 20 mM sodium phosphate, 100 mM NaF, 1.0 mM DTT, and 0.1 mM EDTA at pH 7.0. This buffer was chosen for its transparency and low dependence on temperature. Unfolding transitions as a function of temperature were monitored by the CD signal at 220 nm. The protein concentration was 7.0 μM, and a 1.0 cm cell was used. Temperature was varied from 0 to 95 °C, at a 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 (*38*)

$$\Delta G_{NU} = -RT \ln\left(\frac{f_U}{f_N}\right) = \Delta H_{T_m} + \Delta C_P(T - T_m) - T\left(\left(\frac{\Delta H_{T_m}}{T_m}\right) + \Delta C_P \ln\left(\frac{T}{T_m}\right)\right) \quad (2)$$

$$S = f_N(S_{0,N} + l_N T) + f_U(S_{0,U} + l_U T) \quad (3)$$

where f_U and f_N are the unfolded and folded fractions at equilibrium ($f_U + f_N = 1$), respectively; T_m is the temperature at which f_U is equal to 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; and l_N and l_U are the slopes of the pre and post-transition regions, respectively, assuming a linear dependence of S_N and S_U with temperature.

Radioiodination of Peptide TRX94−108. Peptide TRX94−108 (100 μL, 4.3 mg/mL in water) was mixed with 1.0 M sodium phosphate at pH 7.0 (12 μL), [^{125}I]Na (4 μL, approximately 0.5 μCi), and chloramine-T (58 μL, 3 mg/mL). This mixture was incubated for 3 min at room temperature. Immediately thereafter, the reaction was quenched with sodium metabisulfite (2 μL, 10 mg/mL), and the solution was loaded onto a Sephadex G-25 (Superfine grade) column (0.5 cm × 38 cm) previously equilibrated in buffer A. The column was eluted with the same buffer, and the radioactivity in the fractions was quantified by scintillation counting with a gamma detector (1272 CliniGamma, LKB, Finland). In addition, the UV spectrum of the main fraction containing [^{125}I]-TRX94−108 was recorded, (This spectrum is identical to that of the unlabeled peptide, confirming that under these experimental conditions, there is no oxidizing damage of the peptide as a result of the chloramine-T reaction.) A solution of [^{125}I]-TRX94−108 (0.26 mM, 50,-880 cpm/μL, 500 μL) was used for the TRX1−93 binding experiments.

Isothermal Titration Calorimetry (ITC). The ITC experiments were performed on a Nano ITC III 5300 (Calorimetry Sciences Corp.) at 25 °C. The unfolded fragment TRX1−93 was dialyzed extensively against buffer A and 1 mM DTT at pH 7.0 and 4 °C, and the lyophilized peptide TRX94−108 was dissolved into the same dialysis buffer. Fragment TRX1−93 at a concentration of 81 μM was placed into the cell and titrated with peptide TRX94−108 (1.03 mM solution, loaded into a 250 μL syringe). Intervals of 60 min were set between 10 μL injections to ensure the completion of the folding reaction.

Fluorometric Assay of TRX Activity. Bovine pancreas insulin was chemically modified with fluorescein isothiocyanate and purified according to Heuck and Wolosiuk (*39*) to yield the fluorogenic substrate di-fluoresceinthiocarbamyl-insulin (Di-FTC-insulin). The catalytic activity of TRX as

protein disulfide reductase was assayed as follows. TRX variants were dissolved in buffer A with the addition of 1 mM DTT. 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.

Bioinformatic Analyses. PSI-BLAST (*40*) was used to collect 283 non-redundant amino acid sequences sharing the TRX fold from SwissProt: five iterations were run and an *e*-value cutoff of 0.001 was set. Multiple sequence alignment (MSA) was built using MUSCLE (*41*). Sequence conservation scores were calculated with the ConSurf server (*42*, *43*) using the JTT amino acid substitution model and a Bayesian method for score calculation.

In order to visualize the central aspects of the MSA results, a sequence logo was constructed, computing the frequency of each of the 20 amino acids *a* at each position *i*, yielding a vector $q_i(a)$ of the site-specific profiles. Each frequency is computed as the number of occurrences of amino acid *a*, at position *i* (N_i (a)) divided by the number of sequences in the alignment (N). To ponder gaps in the alignment, the following correction was introduced:

$$q_i(a) = \frac{N_i(a)}{N} + \frac{N_i(\text{gaps})}{20} \quad (4)$$

In the graphic representation, the total height h_i at each position is proportional to the Shannon information as follows:

$$h_i = \sum_a q_i(a) \ln q_i(a) \quad (5)$$

However, a total of 1,000 theoretical sequences compatible with the TRX1−108 structure were generated by Gibbs sampling using a statistical potential optimized for the inverse folding problem (*44*). The potential consists of a pairwise contact term supplemented with a solvent accessibility term. The results were represented as sequence logos in the same fashion as that for the MSA analysis.

RESULTS

Fragment TRX1−93

Molecular Design. An examination of the TRX structure led us to pay particular attention to the cluster depicted in Figure 1 (panel B). This region comprises both α- and β-segments that should organize appropriately to form a hydrophobic concave surface onto which the amphipathic C-terminal helix is tightly docked. Most likely, the correct packing of this crucial region, which includes several long-distance tertiary interactions, will consolidate the overall fold of the protein. Our efforts were then directed toward excising the polypeptide chain of TRX so as to isolate helix 5 from the remainder of the protein. If successful, this approach would provide a means for dissecting the interaction network to evaluate its importance for determining and stabilizing the final fold.

To that end, we prepared the recombinant fragment TRX1−93. Under the assumption of a native-like fold, this truncated form would partially expose the surface, otherwise hidden, belonging to β-strands 2, 3, and 4, and helix 2. A more detailed analysis of the balance of polar/apolar accessible surface reveals that although TRX1−93 exposes a less polar area (2,381 Å^2) than the full protein (2,645 Å^2), this indeed occurs at the expense of a reduction in total area: 5,467 and 5,767 Å^2, for the former and the latter, respectively. In other words, the truncated form actually exposes a similar amount of hydrophobic area (3,086 Å^2, 56% of total area) compared to that exposed by the native protein (3,121 Å^2, 54% of total area).

Expression and Characterization of TRX Variants. TRX1−108 and TRX1−93 were efficiently expressed in *E. coli* (up to 60 mg/L of culture). Mass spectrometry confirmed the expected chemical structures: 11,673.3 ± 2.0 Da for full-length oxidized TRX and 10,045.7 ± 0.5 Da for the reduced form of the fragment TRX1−93. Remarkably, the fragment was mainly localized in inclusion bodies (data not shown). This *in vivo* behavior was the first indication that TRX1−93 does not fold as efficiently as the parent TRX. Refolding by dialysis at 4 °C for 16 h allowed us to recover up to an 81 μM concentration of soluble TRX1−93.

Circular Dichroism. We used CD spectroscopy to address the question of whether the TRX1−93 fragment is able to efficiently acquire the native TRX fold in isolation. Clearly, the spectral features of soluble TRX1−93 are at variance with those of native TRX (Figure 2). Only residual secondary structure is evident from the analysis of the far-UV region, and no defined asymmetric environment for the aromatic residues becomes evident from the greatly diminished signal in the near-UV region (Figure 2, panel B). In addition, the CD spectra of the fragment show no protein concentration dependence in the 7−51 μM range (data not shown). Only after incubation in 6.0 M GdmCl, TRX1−93 becomes completely unfolded at pH 7.0 and room temperature, as judged by the obliteration of signals attributed to ordered secondary structure in the far-UV CD spectra (Figure 2, panel A). Taking into consideration all of the evidence as a whole, under nondenaturing conditions, TRX1−93 includes the features of a molten globule state (*45*, *46*).

Intrinsic Fluorescence Emission. Additional evidence for the remaining structure in TRX1−93 came from tryptophan fluorescence analysis. A comparison of the emission spectra (taken upon excitation at 295 nm) suggests that tryptophan residues in TRX1−93 lie in a slightly less quenched environment than they do in the full-length protein (Figure 3, panel A).

Significantly, the spectrum of TRX1−93 is incompatible with an unfolded conformation. Moreover, in this fragment, tryptophan residues might share nearly the same solvent accessibility (λ_{max} = 343 nm) as that of the native state (λ_{max} = 344 nm). This is a strong indication that at least the part of the molecule comprising the environment of these fluorophores preserves native-like structure, preventing full solvation of the polypeptide chain, as is indeed observed for the unfolded state (λ_{max} = 352 nm in 4 M GdmCl). Importantly, for the aggregation state of TRX1−93 (see below), the emission maximum of this fragment does not

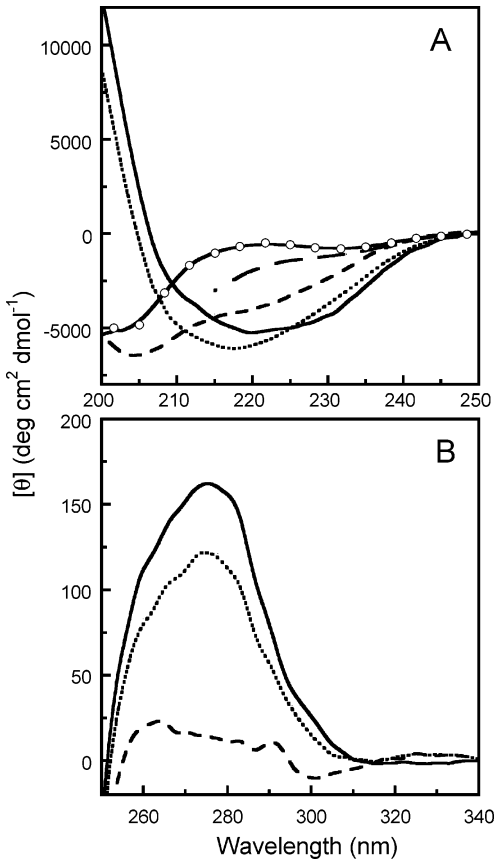

FIGURE 2: Circular dichroism spectra of TRX fragments. The (A) far-UV and (B) near-UV spectra are shown: TRX1−108 (—), TRX1−93 (---), TRX1−93 in 6 M GdmCl (— —), TRX94−108 (−○−), and TRX1−93/TRX94−108 (·····). All measurements were carried out at 20 °C in buffer B (see Materials and Methods). The concentration of TRX1−108 and TRX1−93 was 50 μM, and that of peptide TRX94−108 was 100 μM.

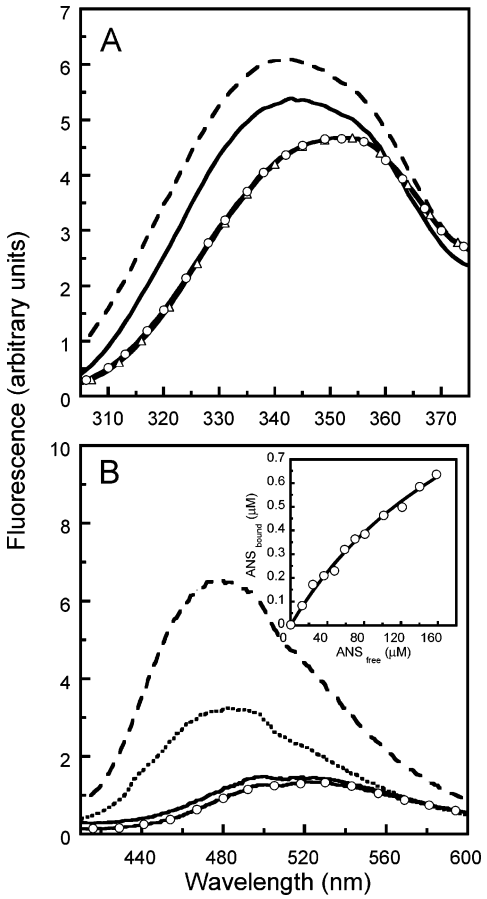

FIGURE 3: Fluorescence emission spectra of TRX variants. The intrinsic fluorescence emission is shown in panel A, where the symbols represent the following: TRX1−93 (---) or TRX1−108 (—) dissolved in buffer A and TRX1−93 (−○−) or TRX1−108 (−△−) in 4.0 M GdmCl. The excitation wavelength was set at 295 nm. Panel B shows the emission spectra of free ANS (−○−), TRX1−93 (- - -), TRX1−108 (—), and the complex TRX1−93/TRX94−108 (···). The excitation wavelength here was set at 350 nm. The inset shows the binding of ANS to TRX1−93 (1.0 μM) at 20 °C in buffer A with the addition of DTT (1 mM). The solid line corresponds to the optimal curve fitting to the data by nonlinear regression, after which the dissociation constant (K_D) was derived.

change with protein concentration in the range 4−70 μM (data not shown).

Size-Exclusion Chromatography. To directly assess whether soluble TRX1−93 adopts a compact conformation, fragment TRX1−93 was analyzed by SEC-FPLC (Figure 4). At the lowest concentration assayed (16 μM), this species shows a Stokes radius (R_S = 22.0 ± 0.5 Å) larger than that of the full-length protein (R_S = 18.0 ± 0.7 Å). However, a clear dependence on protein concentration was observed for the elution volume of the main fraction (Figure 4, inset), a fact consistent with a change in the aggregation state of TRX1−93. Indeed, R_S changes to 26.0 ± 0.5 Å at the highest concentration assayed (81 μM). This evidence is consistent with the existence of a rapid equilibrium (in the time scale of the filtration experiment) between the somewhat expanded monomeric and dimeric species with similar CD and fluorescence signals (see above). The aggregation process does not appear to progress toward higher aggregation states, as judged by the invariance of the relative proportion of the eluting species.

The cross-linking of TRX1−93 with disuccinimidyl suberate (DSS) followed by separation by SDS−PAGE reveals that only a monomer and a dimer are significantly populated at equilibrium. This result confirms that a switch toward the latter occurs upon increasing the protein concentration in the same range as that assayed before (data not shown).

ANS Binding. The molten globule character of TRX1−93 was further evidenced by the occurrence of hydrophobic patches on the surface of this fragment. Indeed, upon the binding of ANS to TRX1−93, the fluorescence intensity increases dramatically, and λ_{max} shifts from 500 to 480 nm (Figure 3, panel B). Note that at the low protein concentration needed for this experiment (1 μM), the fragment exists entirely as the monomeric species. Furthermore, the measured dissociation constant for the ligand ($K_{D\ ANS}$ = 240 μM) lies within the range of values usually attributed to the binding of ANS to molten globule-like states (Figure 3, the inset in panel B,) (*47*). Most significantly, the addition of TRX94−108 to TRX1−93 actually *inhibits* ANS binding, as revealed by the large decrease in fluorescence observed (∼50%) after adding a 2-fold molar excess of the former relative to the latter (Figure 3, panel B). One should note that ANS does not bind to native TRX. Most likely, this result would suggest the refolding of TRX1−93 into a native-like conformation, presumably as a consequence of the association of peptide TRX94−108 to the former. In this case, one would expect a substantial reduction of the solvent-accessible hydrophobic area. To explore the possibility of complex formation in

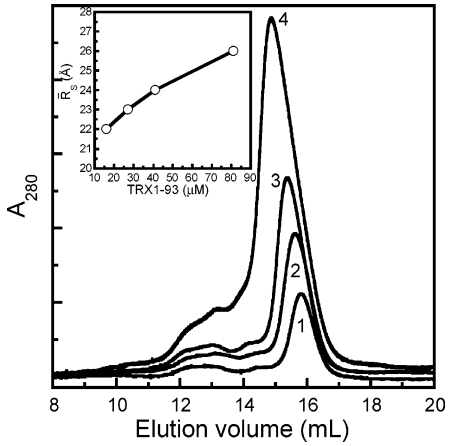

FIGURE 4: Size-exclusion chromatography of fragment TRX1−93. Protein (100−200 μL) at different concentrations, 16 (1), 27 (2), 41 (3), and 81 μM (4), was sampled onto a Superdex S-200 column equilibrated in buffer A at room temperature. Elution with the same buffer was carried out at a flow rate of 0.4 mL/min. The column was previously calibrated with appropriate molecular weight markers. The inset shows the dependence of the average hydrodynamic radius ($\bar{R}_S$) *of fragment TRX1−93 with protein concentration.*

further detail, the interaction of TRX1−93 with TRX94−108 was studied next.

Interaction between Fragment TRX1−93 and Peptide TRX94−108

Circular Dichroism. When fragment TRX1−93 is incubated with peptide TRX94−108, large changes in the far- and near-UV CD regions are indeed observed (Figure 2). Particularly, the evolution of the spectrum of TRX1−93 in the near-UV CD region as peptide TRX94−108 is added might be indicative of the reconstitution of the native spectrum (TRX1−108). In fact, if a 2:1 molar excess of peptide TRX94−108 is present in the mixture, ~75% of a characteristic signal at 280 nm is recovered. The resulting spectrum closely resembles that observed for the native state, showing identical details of fine structure (Figure 2, panel B). Correspondingly, the far-UV CD spectrum of the same mixture acquires a shape similar to that of the spectrum of TRX1−108 (Figure 2, panel A). This statement results from the comparison of the spectra of native TRX1−108 with those of the mixture, after subtracting in the latter the contribution of the remaining amount of peptide 94−108 in its free form. (Notice that this peptide exhibits a typical random coil spectrum.)

Size-Exclusion Chromatography. To further explore the extent of the compactness of the complex, a mixture of fragment TRX1−93 and TRX94−108 (in a molar excess of 2:1 of the latter) was analyzed by SEC-FPLC (Figure 5, panel A), in a fashion similar to that of the experiments already described (Figure 4). By comparison, isolated fragments TRX1−93 and TRX94−108 and the full-length protein TRX1−108 were also tested. At the lowest concentration assayed (17 μM), the elution profile of fragment TRX1−93 is consistent with the monomer−dimer equilibrium, where the expanded monomer clearly prevails (R_S = 22.0 ± 0.5 Å). Remarkably, the TRX1−93/TRX94−108 mixture exhibits the appearance of a new peak centered at a position (R_S = 19.0 ± 0.5 Å) compatible with that of the full-length

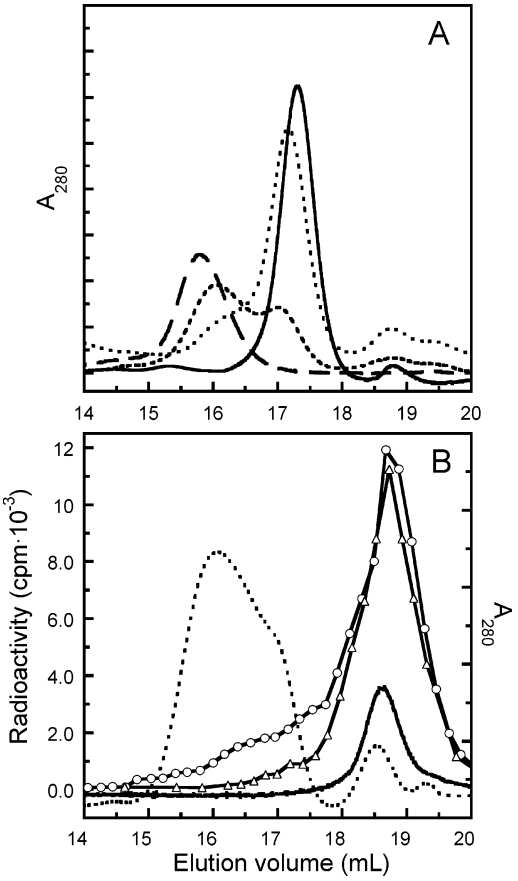

FIGURE 5: Size-exclusion chromatography of the complex TRX1−93/TRX94−108. Chromatographic conditions were identical to those described in Figure 4. (A) The average hydrodynamic radius ($\bar{R}_S$) was measured for isolated TRX1−93 or the complex formed after incubation for 30 min with TRX94−108: TRX1−108 (40 μM, (—)), TRX1−93 (17 μM, (− −)), and TRX1−93/TRX94−108 (17 μM/34 μM, (- - -) or 50 μM/100 μM, (•••)). (B) A similar SEC-FPLC experiment carried out with [^{125}I]-TRX94−108 is shown. Here, both the absorbance at 280 nm (free TRX94−108, (—); complex TRX1−93/TRX94−108, (---)) and the eluted radioactivity (free TRX94−108, (−△−); complex TRX1−93/TRX94−108, (−○−)) were monitored along the run.

protein, a fact indicative of a transition to a native-like complex. Predictably, in a more concentrated sample (50 μM), the peak attributed to the complex increases dramatically at the expense of the TRX1−93 peak. In addition, the presence of large molecular aggregates does not rise upon increasing protein concentration (not shown in the graph).

A size-exclusion experiment with radiolabeled TRX94−108 was also carried out to ascertain the presence of this peptide in the peak assigned to the complex (Figure 5, panel B). The free TRX94−108 species elutes at a position consistent with its low molecular weight and can be detected both by the absorbance of the Tyr residue and because it is the main radioactivity peak in the sample. However, upon incubating a mixture of radiolabeled TRX94−108 with TRX1−93 prepared as described above, the radioactivity profile denotes the appearance of a shoulder eluting precisely at the position of the complex.

Affinity of the Association. To study the peptide−fragment interaction, we measured the dissociation constant (K_D) of the TRX1−93/TRX94−108 complex. To this end, CD in the near-UV region seems most appropriate to monitor peptide binding. The signal at 280 nm serves to monitor the

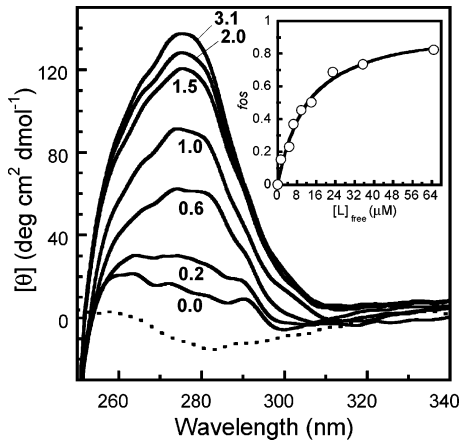

FIGURE 6: Binding of peptide TRX94−108 to fragment TRX1−93 as followed by near-UV circular dichroism. Fragment TRX1−93 was titrated with peptide TRX94−108 (concentration from 0 to 104 μM) in the same buffer as that described in Figure 2. After 1 h of mixing, spectra were recorded. The concentration of fragment TRX1−93 was 35 μM. Numbers (from 0.0 to 3.1) indicate the peptide-to-fragment molar ratio, and a blank of the buffer is shown as a dotted line. The inset shows the evolution of the signal at 280 nm as a function of free peptide TRX94−108. The fraction of occupied sites (*fos*) and the dissociation constant were derived after nonlinear regression fitting of the curve to the data according to eq 1 (see Materials and Methods).

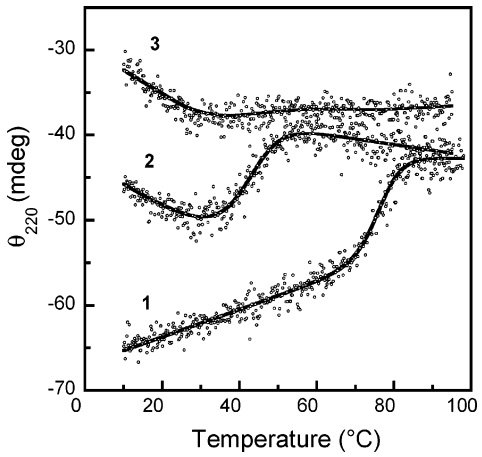

FIGURE 7: Thermal unfolding behavior of isolated fragment TRX1−93 (3) and the complex TRX1−93/TRX94−108 (2). For comparison, full-length TRX1−108 (1) is shown side by side. The concentration of TRX1−93 or TRX1−108 was 7 μM, and that of TRX94−108 was 14 μM. The measurements were carried out in 20 mM sodium phosphate, 100 mM NaF, 0.1 mM EDTA, and 1 mM DTT at pH 7.0. The solid lines represent the nonlinear regression fitting of curves to the data, according to eqs 2 and 3 (see Materials and Methods).

acquisition of tertiary structure (Figure 6) and poses an advantage over the spectrum in the far-UV region because no contribution of the free peptide TRX94−108 exists between 250 and 310 nm (data not shown). In addition, TRX1−93 shows a somewhat poor signal in this spectral range that increases dramatically upon complex formation. To a solution of TRX1−93 (28 μM), different amounts of peptide TRX94−108 were added up to a 3-fold molar excess ratio. After an appropriate transformation of the spectral data, a rectangular hyperbola could be fitted (the inset in Figure 6), from which a value of $K_D = 12$ μM was calculated after nonlinear regression. At the highest molar excess of peptide TRX94−108 assayed, almost 85% of the expected signal for native TRX was achieved.

Thermodynamic Stability of the Complex. Because cooperativity is a feature of native proteins, we decided to evaluate this property in the complex TRX1−93/TRX94−108 by temperature unfolding followed by circular dichroism. The melting transition of TRX variants at pH 7.0 was monitored by the loss of secondary structure upon heating the samples at constant rate (Figure 7). Under our experimental conditions, more than 90% of the signal was recovered after reversing the temperature ramp (data not shown). As can be appreciated, the complex exhibits a clear cooperative behavior, although the transition is centered at a much lower midpoint temperature than that of TRX1−108: 40 vs 76 °C, respectively. Undoubtedly, the complex is significantly less stable than the wild-type protein: $\Delta G_{NU} = 1.3 \pm 0.5$ and 5.2 ± 0.4 kcal mol^{-1} at 25 °C, respectively. Nevertheless, it is remarkable that the temperature range for the transition of the complex is relatively narrow and comparable to that of the full-length protein, pointing to similar values of ΔC_P (1.6 and 1.2 kcal K^{-1} mol^{-1}), a fact that would suggest a comparable extent of solvent exposure of hydrophobic residues occurring upon unfolding. However, the CD signal of the isolated TRX1−93 fragment does not change appreciably between 25 and 95 °C. This result suggests that TRX1−93 does not exhibit a temperature-induced cooperative transition and that the residual structure in the conformational ensemble might remain even at a temperature as high as 95 °C. (Notice, however, that the CD signal at 220 nm for TRX1−93 is sensitive to 6 M GdmCl (cf. Figure 2, panel A).)

Isothermal Titration Calorimetry. To further investigate the thermodynamics of the association between TRX94−108 and TRX1−93, the heat changes that ensue the formation of the complex were measured (Figure 8). A long interval (60 min) was set between each injection of peptide TRX94−108 because it was previously observed that the association process is slow ($k_A \sim 2.9 \cdot 10^{-3}$ s^{-1}, equivalent to $t_{1/2} = 4$ min, as determined from the evolution of the CD signal at 220 nm upon adding a 2-fold molar excess of TRX94−108 over TRX1−93; data not shown). The reaction is exothermic (Figure 8, panel A), and the dissociation constant ($K_D = 2$ μM), estimated after a nonlinear fitting of a binding model considering a single site to the data (Figure 8, panel B), does not significantly differ from the K_D determined by CD (see above). The calculated binding enthalpy is 13.7 kcal mol^{-1}, and the stoichiometry determined for the titration is 0.7. This somewhat low value might result from the fact that in the course of the experiment, close to 25% of the protein fragment aggregates, as estimated by SEC-FPLC on a sample submitted to the ITC experiment (data not shown). This arises from the high protein concentration required (81 μM) and the long incubation time (18 h at 25 °C).

Limited Proteolysis. Resistance to proteolysis is another distinctive attribute of native TRX. Fragment TRX1−93 is highly sensitive to trypsin (Figure 9), a feature that likely results from its molten globule structure. On the contrary, the partial resistance to trypsinolysis exhibited by the TRX1−93/TRX94−108 complex would point out a consolidation of its structure giving rise to reduced conformational flexibility, a fact probably indicative of dynamics closer to that of the native state.

Catalytic Activity of TRX1−93 and the Complex TRX1−93/TRX94−108. Function is the paramount property of the

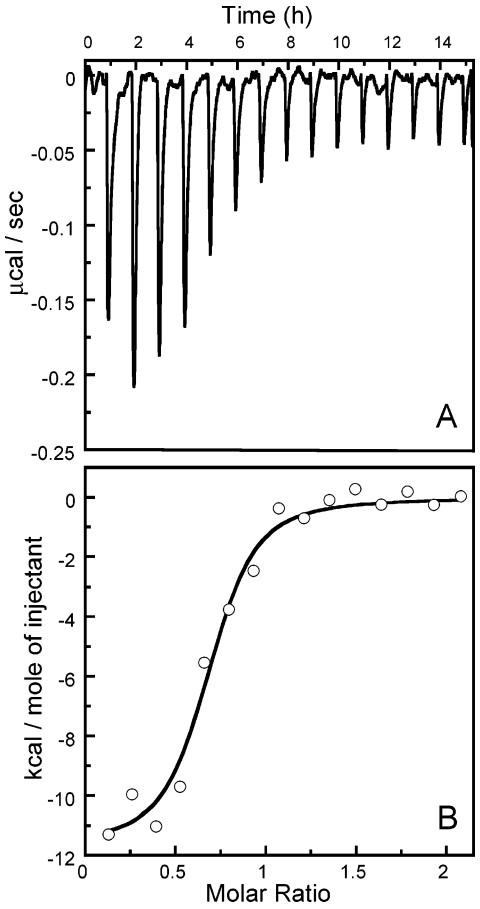

FIGURE 8: Isothermal calorimetric titration of fragment TRX1−93 with peptide TRX94−108. Peptide TRX94−108 was injected into the sample cell (1 mL volume) containing fragment TRX1−93 in buffer A with the addition of 1 mM DTT, at 25 °C. The sample cell was stirred at 400 rpm. (A) Rate of the heat evolved after each injection. The peaks represent the difference between the heat evolved in the sample cell and that in the reference cell containing water. The heat of dilution of peptide TRX94−108 has not been subtracted. (B) Solid line corresponding to the nonlinear regression fitting of a simple binding model to the data after subtraction of the heat of dilution.

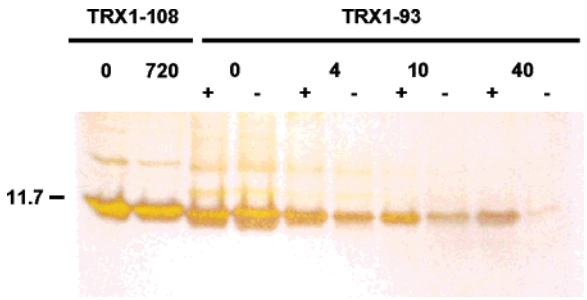

FIGURE 9: Relative stability to proteolysis of fragment TRX1−93 and complex TRX1−93/TRX94−108 as analyzed by SDS−PAGE. The incubation times (in min) are indicated at the top of each lane, and symbols (+) and (−) indicate the presence or absence of peptide TRX94−108 in the reaction, respectively. As a control sample, TRX1−108 was incubated for the time indicated in the presence of an identical concentration of trypsin. The molecular mass, 11.7 kDa, corresponds to that of full-length TRX1−108.

native state. To ascertain whether the title variants are catalytically competent, these were compared side by side with full-length TRX by employing a highly sensitive fluorometric assay of protein disulfide reductase activity (Figure 10). TRX1−93 exhibits only marginal activity (0.7%

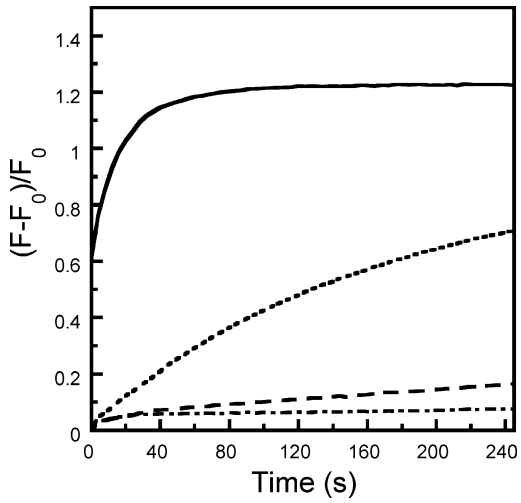

FIGURE 10: TRX-catalyzed reduction of insulin. The time course of the protein disulfide reduction exhibited by TRX variants assayed against di-FTC-insulin: TRX1−108 (—), TRX1−93 (---), TRX94−108 (— · —), and TRX1−93/TRX94−108 (···). All measurements were carried out at 25 °C in buffer A and 1 mM DTT (see Materials and Methods). The concentration of TRX1−108 was 17 μM, that of TRX1−93 (free or in the complex) was 40 μM, and that of TRX94−108 (free or in complex) was 120 μM.

of the specific activity shown by full-length TRX), but the catalytic efficiency dramatically increases in the case of the complex TRX1−93/TRX94−108. (The specific activity of the complex is 7.1%; note that isolated peptide TRX94−108 is not active.) At the longest time assayed (250 s), the complex is able to reduce close to 30% as much protein substrate as full-length TRX, following a time course of catalysis not distinguishable from that of the latter.

DISCUSSION

It is universally accepted that proteins contain in their sequence all the necessary information to reach the native state. Nevertheless, despite intense research during the past four decades, the folding code remains elusive. Furthermore, the problem becomes more intricate in view of the fact that different sequences give rise to similar structures. In this context, useful clues to decipher the code may be extracted from the study of peptide complementation as a model of folding.

In particular, in this article we describe and characterize a novel system of this type, constituted by fragments 1−93 and 94−108 of *E. coli* TRX. In comparison with other complementing fragments of TRX (cf. Figure 1, panel A), an important aspect of the present system is that it allows one to place the focus on the informational content of a more restricted region of the TRX structure. In other words, this would enable us to address key conformational features in a clear-cut fashion.

The bulk of spectroscopic evidence presented herein supports the notion that fragment TRX1−93 in isolation adopts a non-native structure: this polypeptide folds into an expanded molten globule, with residual tertiary interactions. Upon binding peptide TRX94−108, fragment TRX1−93 consolidates a native-like fold, minimally exposing the hydrophobic surface area to the aqueous solvent.

In this regard, one should bear in mind that isolated peptide TRX94−108 does not exhibit defined secondary structure

in solution, as ascertained by far-UV CD spectra. Nevertheless, in all likelihood it would acquire helical structure upon complex formation. Fragment interaction confers to fragment TRX1−93 less susceptibility to proteolysis, pointing to a dynamics of the complex more closely resembling that of the native state, that is, the polypeptide backbone would not be as flexible as in the case of the isolated fragment, a result that agrees well with the emergence of rigid environments for aromatic residues, as can be judged by near-UV CD spectra. All evidence indicates that TRX1−93 efficiently refolds when it meets the cognate peptide TRX94−108, establishing an interaction network that physically represents the reciprocal exchange of information between the interacting moieties. Proof of this is the ability of the complex to perform catalysis, the hallmark of the native state.

Paradoxically, it is worth noting that Jacobs et al. conclude, on the basis of a distance constraint model, that a main-chain break between residues 92 and 94 would produce two stable and non-interacting fragments (*48*). The first part of this prediction finds a generally good agreement with our finding that fragment TRX1−93 does not adopt a fully disordered state but rather folds into a molten globule state. However, at variance with their interpretation, we believe that embracing self-stability does not necessarily imply the lack of an ability to interact with the missing moiety. On the contrary, stabilization of a native-like conformation would provide a scaffold favorable to peptide recognition. In turn, binding of the cognate peptide would provide further structural stabilization. From this viewpoint, herein we present experimental evidence supporting the consolidation of a stable structure for the complex TRX1−93/TRX94−108.

In regard to complex formation, it is worth discussing thermodynamic stability in connection with the preservation of structure. From an experimental perspective, an evaluation of the thermal stability of TRX peptide complexes can provide insights into this matter. Despite the lower stability observed for TRX1−93/TRX94−108 (T_m ~40 °C; K_D = 2−12 μM at 25 °C) compared to that of the wild-type protein (T_m ~76 °C), this fact does not pose an obstacle to presumably attaining the same 3D structure, as is also the case for other complexes such as TRX1−73/TRX74−108 (T_m ~60 °C, as estimated by DSC; K_D = 49 nM) (*28*, *49*) or TRX1−37/TRX38−108 (K_D = 4−6 μM) (*4*, *49*).

From a structural viewpoint, one should discuss the consequence of preserving matching hydrophobic surfaces at the contact region between fragments TRX1−93 and TRX94−108. Indeed, screening experiments point to the critical importance in the early steps of the folding process played by the formation of a hydrophobic cluster comprising the L row (belonging to TRX94−108) closing onto the LI surface (in TRX1−93, see Introduction). In particular, the change L107P in helix 5 produces a significant decrease in the folding rate of TRX, bringing about an increase in protein export (*50*). In this case, the authors conclude that the proline at that position could partially destabilize the interaction with the LI surface in the folded state. Independently, Chiu et al. established the significance of residue F102 for the structural integrity of TRX (*51*). In this case, pulse proteolysis was used to show that mutant F102S unfolds at a lower urea concentration than wild-type TRX. Moreover, they demonstrated that the decreased stability shown by mutant F102S affects the affinity of binding to protein gp5, a step of crucial importance where helix 5 plays a key role in forming the processive T7 DNA polymerase complex.

Our own evidence and that of others taken together allow us to put forward the prominent position played by the C-terminal helix as an essential element of TRX structure. Interestingly, three out of four leucine residues in the L row (L99, L103, and L107) are conserved throughout TRX evolution, and whenever mutations occur at L99 and L103, a hydrophobic profile is conserved (Figure 1, panel C). The above statement is supported by the following sequence analysis. By using the 1 to 9 scoring scale by Landau et al. (*43*), the conservation scores are 7 and 6 for positions L99 and L103, respectively, suggesting their importance for establishing the hydrophobic core. As for L107, this residue is not conserved in the MSA because many sequences (174 out of 283) are shorter than 107 residues. However, when it is longer than that, a hydrophobic amino acid is always found at this position: I, L, or V in 90% of the sequences. By contrast, as is evident in the sequence profile, conservation of L94 does not seem relevant for the protein family. Noticeably, L94 does not strictly belong to helix 5, as one can conclude after inspection of the TRX structure.

To further assess the significance of the L row for attaining the TRX fold, we applied a newly developed protein design algorithm (*44*). This procedure allowed us to generate sequences compatible with the TRX structure. By using a pairwise contact potential and a solvent accessibility criterion, a sequence profile very similar to the MSA of natural sequences was obtained (Figure 1, panel C). Consistently, L99 and L103 emerge from this analysis with enhanced importance. Strikingly, the AGADIR server (*52*, *53*) predicts this same pair of leucine residues as key structural determinants conferring helical behavior to peptide TRX94−108 in solution (data not shown). Hypothetically, if the conformational ensemble of this peptide had a small but significant population in helical form, then this could have accelerated the association/folding reaction of the complex. In the future, in order to assess the informational content of this peptide more thoroughly, computational as well as biophysical experiments will be designed to evaluate its intrinsic or induced conformational trend.

Furthermore, the results presented in this article may shed light on the mechanism of TRX folding. On the basis of NMR studies of complementing fragments TRX1−37, TRX38−73, and TRX74−108, Tasayco et al. postulated that the zippering of two rather hydrophobic regions, those corresponding to strands β2 and β4 in the native state and located at the bottom of the LI surface (Figure 1, panel B), constitutes the first operational step leading to folded TRX (*7*). In line with this, we hypothesize that the binding of helix 5 to the LI surface could represent a late event in the consolidation of the TRX fold. In this regard, unfolding (or the absence) of helix 5 could permit the population of a partially folded state of TRX under native conditions. As mentioned in the Introduction, the existence of partially folded states for TRX has been well documented. In addition, evidence from temperature unfolding experiments point to the existence of a collapsed unfolded state for TRX, with the persistence of residual secondary structure. This was presumed from the presence of average chiral structure, as detected by far-UV CD, and the simultaneous lack of

protection against amide proton exchange (*54*). However, under non-physiological conditions, the application of Fourier-transform IR spectroscopy revealed the presence of conformations with properties of molten globule-like states, that is, characterized by well-defined secondary structure in the absence of rigid tertiary interactions (*55*).

In summary, TRX1−93 could eventually represent a new paradigm useful for interpreting differences among non-native states. Finally, identification of short sequence signatures with a significant influence on the acquisition of 3D structure such as those present in TRX94−108 might enable one to pinpoint the existence of interactions essential for molecular recognition. In turn, this might lead to clues regarding the origin of the cooperativity distinctive of the folded state.

CONCLUSIONS

The effect of excising helix 5 (included in TRX94−108) from the remainder of TRX causes the resulting fragment (TRX1−93) to halt progress toward the native state. Instead, this variant remains in a molten globule conformation. Remarkably, complementation of fragment TRX1−93 with the cognate peptide TRX94−108 brings forth the restoration of native-like conformation and function. This system underscores key interactions leading to the consolidation of the native TRX fold. We believe that results presented herein point to the stabilizing role played by helix 5. We conjecture that the primary interaction brought about by the apposition of the so-called L row onto the LI surface would illustrate a molecular event of mutual selection, with long-range consequences on conformation, and possibly represent a late event in the consolidation of the overall TRX fold.

ACKNOWLEDGMENT

We thank Dr. Mauricio P. Sica for his valuable suggestions and comments on our experiments, Dr. Mariano Grasselli for his gift of di-FTC-insulin, and Ms. Valeria A. Risso for her assistance with the operation of the nanocalorimeter.

REFERENCES

1. Dunn, I. S., and Jennings, P. A. (1990) Protein modification from mutational analysis of an autologous peptide fragment, *Protein Eng. 4*, 205−213.
2. Musi, V., Spolaore, B., Picotti, P., Zambonin, M., De Filippis, V., and Fontana, A. (2004) Nicked apomyoglobin: a noncovalent complex of two polypeptide fragments comprising the entire protein chain, *Biochemistry 43*, 6230−6240.
3. Smith, V. F., and Matthews, C. R. (2001) Testing the role of chain connectivity on the stability and structure of dihydrofolate reductase from *E. coli*: fragment complementation and circular permutation reveal stable, alternatively folded forms, *Protein Sci. 10*, 116−128.
4. 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.
5. Ojennus, D. D., Fleissner, M. R., and Wuttke, D. S. (2001) Reconstitution of a native-like SH2 domain from disordered peptide fragments examined by multidimensional heteronuclear NMR, *Protein Sci. 10*, 2162−2175.
6. Tasayco, M. L., and Chao, K. (1995) NMR study of the reconstitution of the beta-sheet of thioredoxin by fragment complementation, *Proteins 22*, 41−44.
7. Tasayco, M. L., Fuchs, J., Yang, X. M., Dyalram, D., and Georgescu, R. E. (2000) Interaction between two discontiguous chain segments from the beta-sheet of *Escherichia coli* thioredoxin suggests an initiation site for folding, *Biochemistry 39*, 10613−10618.
8. Richards, F. M., and Vithayathil, P. J. (1959) The preparation of subtilisn-modified ribonuclease and the separation of the peptide and protein components, *J. Biol. Chem. 234*, 1459−1465.
9. Neira, J. L., Davis, B., Ladurner, A. G., Buckle, A. M., Gay G de, P., and Fersht, A. R. (1996) Towards the complete structural characterization of a protein folding pathway: the structures of the denatured, transition and native states for the association/folding of two complementary fragments of cleaved chymotrypsin inhibitor 2. Direct evidence for a nucleation-condensation mechanism, *Folding Des. 1*, 189−208.
10. 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.
11. 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.
12. Gleason, F. K., and Holmgren, A. (1988) Thioredoxin and related proteins in prokaryotes, *FEMS Microbiol. Rev. 4*, 271−297.
13. 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.
14. Hendrickson, W. A., Horton, J. R., and LeMaster, D. M. (1990) Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD): a vehicle for direct determination of three-dimensional structure, *EMBO J. 9*, 1665−1672.
15. 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.
16. Dyson, H. J., Jeng, M. F., Model, P., and Holmgren, A. (1994) Characterization by 1H NMR of a C32S, C35S double mutant of *Escherichia coli* thioredoxin confirms its resemblance to the reduced wild-type protein, *FEBS Lett. 339*, 11−17.
17. 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.
18. Bolon, D. N., and Mayo, S. L. (2001) Polar residues in the protein core of *Escherichia coli* thioredoxin are important for fold specificity, *Biochemistry 40*, 10047−10053.
19. Mancusso, R., Cruz, E., Cataldi, M., Mendoza, C., Fuchs, J., Wang, H., Yang, X., and Tasayco, M. L. (2004) Reversal of negative charges on the surface of *Escherichia coli* thioredoxin: pockets versus protrusions, *Biochemistry 43*, 3835−3843.
20. 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.
21. 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.
22. Chakraborty, K., Thakurela, S., Prajapati, R. S., Indu, S., Ali, P. S., Ramakrishnan, C., and Varadarajan, R. (2005) Protein stabilization by introduction of cross-strand disulfides, *Biochemistry 44*, 14638−14646.
23. Kelley, R. F., and Stellwagen, E. (1984) Conformational transitions of thioredoxin in guanidine hydrochloride, *Biochemistry 23*, 5095−5102.
24. Bhutani, N., and Udgaonkar, J. B. (2003) Folding subdomains of thioredoxin characterized by native-state hydrogen exchange, *Protein Sci. 12*, 1719−1731.
25. Kelley, R. F., Wilson, J., Bryant, C., and Stellwagen, E. (1986) Effects of guanidine hydrochloride on the refolding kinetics of denatured thioredoxin, *Biochemistry 25*, 728−732.
26. Georgescu, R. E., Li, J. H., Goldberg, M. E., Tasayco, M. L., and Chaffotte, A. F. (1998) 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*, 10286−10297.
27. Bhutani, N., and Udgaonkar, J. B. (2001) GroEL channels the folding of thioredoxin along one kinetic route, *J. Mol. Biol. 314*, 1167−1179.

28. 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.
29. 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 substitution, *Biochemistry 31*, 11203−11209.
30. 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.
31. Tsodikov, O. V., Record, M. T., Jr., and Sergeev, Y. V. (2002) Novel computer program for fast exact calculation of accessible and molecular surface areas and average surface curvature, *J. Comput. Chem. 23*, 600−609.
32. Curto, L. M., Caramelo, J. J., and Delfino, J. M. (2005) Δ98Δ, a functional all-β-sheet abridged form of intestinal fatty acid binding protein, *Biochemistry 44*, 13847−13857.
33. Langsetmo, K., Fuchs, J., and Woodward, C. (1989) *Escherichia coli* thioredoxin folds into two compact forms of different stability to urea denaturation, *Biochemistry 28*, 3211−3220.
34. Santos, J., Gebhard, L. G., Risso, V. A., Ferreyra, R. G., Rossi, J. P., and Ermacora, M. R. (2004) Folding of an abridged beta-lactamase, *Biochemistry 43*, 1715−1723.
35. Uversky, V. N. (1993) Use of fast protein size-exclusion liquid chromatography to study the unfolding of proteins which denature through the molten globule, *Biochemistry 32*, 13288−13298.
36. Daniel, E., and Weber, G. (1966) Cooperative effects in binding by bovine serum albumin. I. The binding of 1-anilino-8-naphthalenesulfonate. Fluorimetric titrations, *Biochemistry 5*, 1893−1900.
37. Arighi, C. N., Rossi, J. P., and Delfino, J. M. (2003) Temperature-induced conformational switch in intestinal fatty acid binding protein (IFABP) revealing an alternative mode for ligand binding, *Biochemistry 42*, 7539−7551.
38. Fersht, A. (1999) *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding*, Freeman, New York.
39. 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.
40. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, *Nucleic Acids Res. 25*, 3389−3402.
41. Edgar, R. C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput, *Nucleic Acids Res. 32*, 1792−1797.
42. Glaser, F., Pupko, T., Paz, I., Bell, R. E., Bechor-Shental, D., Martz, E., and Ben-Tal, N. (2003) ConSurf: identification of functional regions in proteins by surface-mapping of phylogenetic information, *Bioinformatics 19*, 163−164.
43. Landau, M., Mayrose, I., Rosenberg, Y., Glaser, F., Martz, E., Pupko, T., and Ben-Tal, N. (2005) ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures, *Nucleic Acids Res. 33*, 299−302.
44. Kleinman, C. L., Rodrigue, N., Bonnard, C., Philippe, H., and Lartillot, N. (2006) A maximum likelihood framework for protein design, *BMC Bioinformatics 7*, 326.
45. Ptitsyn, O. B., and Uversky, V. N. (1994) The molten globule is a third thermodynamical state of protein molecules, *FEBS Lett. 341*, 15−18.
46. Vassilenko, K. S., and Uversky, V. N. (2002) Native-like secondary structure of molten globules, *Biochim. Biophys. Acta 1594*, 168−177.
47. Shi, L., Palleros, D. R., and Fink, A. L. (1994) Protein conformational changes induced by 1,1′-bis(4-anilino-5-naphthalenesulfonic acid): preferential binding to the molten globule of DnaK, *Biochemistry 33*, 7536−7546.
48. Jacobs, D. J., Livesay, D. R., Hules, J., and Tasayco, M. L. (2006) Elucidating quantitative stability/flexibility relationships within thioredoxin and its fragments using a distance constraint model, *J. Mol. Biol. 358*, 882−904.
49. Yang, X. M., Georgescu, R. E., Li, J. H., Yu, W. F., Haierhan, and Tasayco, M. L. (1999) Recognition between disordered polypeptide chains from cleavage of an alpha/beta domain: self- versus non-self-association, *Pac. Symp. Biocomput. '96* 590−600.
50. 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.
51. 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.
52. 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.
53. Fisinger, S., Serrano, L., and Lacroix, E. (2001) Computational estimation of specific side chain interaction energies in alpha helices, *Protein Sci. 10*, 809−818.
54. 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.
55. Pedone, E., Bartolucci, S., Rossi, M., Pierfederici, F. M., Scire, A., Cacciamani, T., and Tanfani, F. (2003) Structural and thermal stability analysis of *Escherichia coli* and *Alicyclobacillus acidocaldarius* thioredoxin revealed a molten globule-like state in thermal denaturation pathway of the proteins: an infrared spectroscopic study, *Biochem. J. 373*, 875−883.

BI6026264