Prediction and experimental testing of *Bacillus acidocaldarius* thioredoxin stability[1]

Emilia PEDONE*, Raffaele CANNIO†, Michele SAVIANO‡, Mosè ROSSI*§ and Simonetta BARTOLUCCI*[2]

*Dipartimento di Chimica Organica e Biologica, Università degli Studi di Napoli 'Federico II', via Mezzocannone 16, 80134 Naples, Italy, †Istituto delle Scienze dell'Alimentazione, via Roma (ang. via Rubbilli), 83100 Avellino, Italy, ‡Centro di Studio di Biocristallografia, Consiglio Nazionale delle Ricerche, Dipartimento di Chimica, Università degli Studi di Napoli 'Federico II', via Mezzocannone 4, 80134 Naples, Italy, and §Istituto di Biochimica delle Proteine ed Enzimologia, Consiglio Nazionale delle Ricerche, via Marconi 10, 80125 Naples, Italy

In order to investigate further the determinants of protein stability, four mutants of thioredoxin from *Bacillus acidocaldarius* were designed: K18G, R82E, K18G/R82E, and D102X, in which the last four amino acids were deleted. The mutants were constructed on the basis of molecular dynamic studies and the prediction of the structure of thioredoxin from *B. acidocaldarius*, performed by a comparative molecular modelling technique using *Escherichia coli* thioredoxin as the reference protein. The mutants obtained by PCR strategy were expressed in *E. coli* and then characterized. CD spectroscopy, spectrofluorimetry and thermodynamic comparative studies permitted comparison of the relative physicochemical behaviour of the four proteins with that of the wild-type protein. As predicted for the molecular dynamic analysis at 500 K *in vacuo*, the wild-type structure was more stable than that of the mutants; in fact the T_m of the four proteins showed a decrease of about 15 °C for the double and the truncated mutants, and a decrease of about 12 °C for the single mutants. A difference in the resistance of the proteins to denaturants such as guanidine HCl and urea was revealed; the wild-type protein always proved to be the most resistant. The results obtained show the importance of hydrogen bonds and ion pairs in determining protein stability and confirm that simulation methods are able to direct protein engineering in site-directed mutagenesis.

Keywords: circular dichroism, molecular dynamics, protein engineering, thermostability.

INTRODUCTION

Greater knowledge of the molecular mechanisms that allow proteins from thermophilic organisms to withstand high temperatures is of potential significance for a better understanding of the forces contributing to protein stability, and for the design of stable proteins of biotechnological importance. Different strategies of stabilization have been proposed on the basis of comparison of thermophile homologues with their mesophilic counterparts. According to several sources, the substitution of amino acids preferentially occurring in mesophiles by amino acids preferentially occurring in thermophiles [1] may be a promising way to shift the point of denaturation towards higher temperatures. Research has also focused on the incorporation of additional favourable interactions into the protein structure available, thereby presumably decreasing the flexibility of the structure [2]. The engineering of disulphide bridges [3], electrostatic interactions [4], hydrophobic interactions [5], or the introduction of amino acids with a high α-helix propensity or an abundance of ion pairs at the protein surface [6] have all been used to increase the stability of proteins or peptides. The role of ionic networks in the thermal stability of proteins was functionally demonstrated by site-directed mutagenesis, in which specific salt bridges in α-helices were disrupted [7].

On the basis of this knowledge thioredoxin was chosen as a model system to investigate the structural basis of the stability of proteins. Thioredoxin is a small (12 kDa) heat-stable protein found in all living cells from Archaea representatives to humans [8]. The active site of thioredoxin is conserved with two redox-active-cysteine residues in the sequence Trp-Cys-Gly-Pro-Cys. The structure of both reduced and oxidized *Escherichia coli* thioredoxin has been elucidated by NMR [9], and the three-dimensional structure of the oxidized protein resolved by X-ray crystallography [10]. Reduced and oxidized thioredoxin proved to have identical secondary structures and tertiary folds consisting of a central five-stranded β-sheet with two parallel and two antiparallel junctions. This central sheet is surrounded by α-helices. The active-site disulphide (Cys^{32}-Cys^{35}) is located in a short loop at the N-terminal of the α2-helix, after the second strand of the β-sheet. Recently, the physicochemical features of the thioredoxin from *Bacillus acidocaldarius* (BacTrx), its primary structure and the high-level expression in *E. coli* of the coding gene have been reported [11]. Investigation of the structural stability by CD, differential scanning calorimetry and nanogravimetry showed that it is endowed with a higher conformational heat capacity when compared with the thioredoxin from *E. coli*.

Previous molecular dynamic studies on *E. coli* thioredoxin and BacTrx, carried out at different temperatures *in vacuo* and in water solution, showed that protein stability could be predicted using these techniques and provided some details on the molecular basis of thermostability in proteins. The molecular dynamics were not inconsistent with the experimental findings, that the replacement of residues in BacTrx stabilized the protein (with respect to *E. coli* thioredoxin) and that one possible cause of this stabilization was reduced backbone flexibility. In addition, it was possible to correlate the increase in protein stabilization with the conformational and structural changes caused by single amino

Abbreviations used: BacTrx, thioredoxin from *Bacillus acidocaldarius*; GuHCl, guanidine hydrochloride.

[1] This paper is dedicated to the memory of our friend and colleague Giacomino Randazzo.

[2] To whom correspondence should be addressed (e-mail bartoluc@unina.it).

acid replacements [12]. Based on these results, four mutants of BacTrx were designed to allow further investigation of protein stability and simulation methods were used for protein engineering and/or site-directed mutagenesis strategies.

Simulation studies carried out previously [12] highlighted that the amino acid residues Lys-18 and Arg-82 were especially interesting because Lys-18 was hydrogen bonded to Asp-48 and Gln-105, and Arg-82 was hydrogen bonded to Gln-15, Gly-16 and Asp-17. In the present study, site-directed mutagenesis was applied in BacTrx to replace these two key residues Lys-18 and Arg-82 with a glycine and glutamate residue respectively present in *E. coli* thioredoxin, thus removing the electrostatic interaction. A double mutant (K18G/R82E) and a deletion mutant, in which four amino acids at the C-terminal tail from D102 were deleted (D102X), were also obtained. This tail played an important role in protecting the protein core and moreover it was found to be linked to many of the salt bridges. The four mutants were expressed in *E. coli* at high levels, the proteins were purified and the physicochemical behaviour was compared with the wild-type by means of molecular dynamics simulations, CD and spectrofluorimetry.

MATERIALS AND METHODS

Materials

Thioredoxin from *E. coli* was purchased from IMCO (Stockholm, Sweden). Bovine insulin was from Sigma. Molecular-mass standards for SDS/PAGE were obtained from Gibco-BRL. Nile Red was obtained from Aldrich Chimica. pTrc99A and pUC18 *Sma*I/BAP plasmids were from Pharmacia. pGEM5Zf(−) was from Promega. Radioactive materials were obtained from Amersham International. Deoxynucleotides and restriction and modification enzymes were purchased from Boehringer. All materials used for gene amplification were supplied by Stratagene Cloning Systems. All synthetic oligonucleotides were purchased from PRIMM (Milan, Italy). All aqueous solutions were made using water purified by a Milli-Q water system (Millipore). All chemicals used were of the highest grade available.

Molecular dynamic simulations

All calculations and graphical analyses were run on a Silicon Graphics Indigo2 workstation. The INSIGHT/DISCOVER package (Biosym Technologies) was used to perform energy minimization and molecular dynamic simulations *in vacuo* at 500 K and pH 7.0, using the AMBER force field [13].

The molecular dynamic simulations on the four mutants were carried out *in vacuo* at 500 K. In all simulations, the side-chains of arginine and lysine residues were positively charged, whereas the side chains of glutamate and aspartate residues were negatively charged. The starting structures used in the simulations were those obtained previously for the wild-type [12], and necessary changes to, or deletions of, the residues were made. The dihedral angles of the side chains were then modelled by choosing favourable rotamers. Computational conditions were chosen to avoid any boundary effects [14].

In all simulations, which were performed with a time step of 1.0 fs, the proteins were equilibrated for 80 ps. After this first step, a further 80 ps of simulations with no rescaling were carried out, as energy conservation had been observed and the average temperature had remained essentially constant around the target values. The time span of the runs was optimized, checking that the prolongation of the simulations did not influence the results. Co-ordinates and velocities for the four simulations were transferred on to a disk every 10 steps during the last 80 ps of the simulations and were used for the statistical analysis.

Construction of mutated thioredoxins

The mutations were introduced into *B. acidocaldarius* thioredoxin coding gene (*BacTrx*) by PCR-mediated mutagenesis; the template was the expression vector pTrx described previously [11].

Lys-18 and Arg-82 were replaced with glycine and glutamate residues respectively; both of these substitutions were introduced in the double mutant. The C-terminal truncated protein was obtained by deleting the last four amino acids of BacTrx [11].

The amplification was performed according to Saiki [15] for 35 cycles at an annealing temperature of 55 °C (Cetus Cycler Temp; Perkin-Elmer) using *Pfu* polymerase (Stratagene) for R82E and *Taq* polymerase for all of the other amplifications.

The K18G mutant was obtained by amplifying the BacTrx coding sequence with the 5′ mutagenic primer: 5′GACTGATGCTAATTTTCAGCAGGCGATTCAAGGCGATgggCCC3′, which allowed the insertion of the point mutation (lower case letters) and the 3′ primer BacTrx C-terminal oligonucleotide: 5′GATTACTGTAATACATCTGCTAACTGTGCCTCCAAC3′.

The wild-type *Hin*dII fragment of the *BacTrx* gene was substituted into the plasmid pUC18BacTrx [11] with the corresponding mutagenized amplicon. The reconstituted coding sequence was then inserted into the pTrc99A plasmid, between the *Nco*I and *Pst*I sites, producing the expression vector pTrxK18G.

R82E was constructed by amplifying the *BacTrx* gene with the 5′ mutagenic primer: CgagCCTGTGAAACAGCTGATCGGG, which allowed the insertion of the point mutation (lower case letters) and the 3′ primer BacTrx C-terminal oligonucleotide.

The amplified DNA fragment and the *Nco*I/*Hae*III region from the pTrx plasmid were cloned into pGEM5Zf(−), which had been digested previously with *Nco*I–*Eco*RV. The reconstituted coding sequence was then inserted into the pTrc99A plasmid, between the *Nco*I and *Sal*I sites, producing the expression vector pTrxR82E.

The *Eco*RV–*Xho*I gene fragment containing the K18G mutation was inserted in the R82E mutant, replacing the corresponding region of the wild-type, in order to generate the double mutant K18G/R82E.

The D102X mutant for the last four amino acids was obtained by amplifying the pTrx plasmid with the 5′ primer N-terminal oligonucleotide: 5′GCCCATGGCTACGATGACGTTGACTGATG3′ and the mutagenic 3′ primer: 5′TACttaTGCTAACTGTGCCTCCAAC3′, which inserted a stop codon (lower case letters), allowing earlier termination of translation. The reconstituted coding sequence inserted into the pTrc99A plasmid between the *Nco*I and *Sal*I sites, producing the expression vector pTrxD102X.

The insertion of the correct mutations into the genes was confirmed by DNA sequencing with the Sanger dideoxynucleotide termination method, using the Sequenase version 2.0 kit (United States Biochemicals–Amersham).

Expression and purification of mutant thioredoxins

E. coli RB791-competent cells [16] were transformed with the newly constructed plasmids and protein expression was performed following the procedure described previously [11]. Cell pellets from 1 litre cultures were re-suspended in 20 ml of 10 mM Tris/HCl, pH 8.4 and crude extracts were prepared by disrupting

the cells with 20 min pulses at 20 kHz (Sonicator Ultrasonic liquid processor; Heat System Ultrasonics Inc.) and ultracentrifugation at 160000 ***g*** for 30 min.

The purification of recombinant wild-type protein was carried out as described previously [11], and the mutant proteins were prepared with slight modifications. The crude extracts were subjected to heat treatment at temperatures adapted to the different mutants and centrifuged at 5000 ***g*** at 4 °C for 15 min. The supernatants were concentrated five-fold by ultrafiltration in an Amicon cell (membrane cut-off 1000 Da) and reheated under the same conditions. The crude extracts were extensively dialysed against 20 mM Tris/HCl, pH 8.4 (buffer A) and then applied, in two separate runs, to a 2.6-cm × 60-cm column (HiLoad Superdex 75; Pharmacia) connected to an FPLC system (Pharmacia) and eluted with 20 mM Tris/HCl, pH 8.4/0.2 M NaCl (buffer B) at a flow rate of 2 ml/min. The active fractions were pooled and concentrated. The Superdex pool was loaded on to a 0.5-cm × 5-cm column (Mono Q HR; Pharmacia) connected to an FPLC system (Pharmacia) equilibrated with buffer A. Bound proteins were eluted with a linear gradient of 0–0.4 M NaCl in buffer A at a flow rate of 0.5 ml/min.

Mutant thioredoxins generated by PCR-mediated mutagenesis were expressed in *E. coli* Rb791, as previously described for the wild-type BacTrx [11]. The optimized overexpression of mutant K18G and K18G/R82E was obtained by exposing the cells to 1 mM isopropyl β-thiogalactoside at a cell density of $D = 1$. The ideal cell densities for the expression of the R82E and D102X mutants were $D = 0.5$ and $D = 1.5$ respectively. In order to optimize the production of the recombinant proteins, transformed cells were exposed to the inducer for time scales ranging from 2–24 h and the maximum level of expression was obtained after 20 h of induction.

The crude extract of *E. coli* was subjected to two thermal precipitation steps at 60 °C for 20 min for K18G, R82E and K18G/R82E and at 50 °C for D102X to remove almost 30 % of the mesophilic host proteins. After gel-filtration and anion-exchange chromatography, only one peak was observed with reverse-phase HPLC and only one protein band was revealed by SDS/PAGE.

Protein concentration

Protein concentration was determined by the method of Smith et al. [17] using BSA as the standard. In order to confirm protein concentrations, N-terminal sequences were performed with the Edman degradation using a pulsed liquid-phase sequencer (model 477A; Applied Biosystems) equipped with a 120A analyser for the on-line detection of phenylthiohydantoin amino acids. Absorbance spectra in the range 200–300 nm were also monitored.

Assay and analysis of thioredoxin

Thioredoxin activity was assessed according to the method of Holmgren [18] by following the dithiothreitol-dependent reduction of bovine insulin disulphides. The purity of the proteins was estimated by SDS/PAGE [15 % (w/v) gels] [19] and the proteins were revealed by the silver staining procedure of Rabilloud et al. [20]. In addition, the proteins were analysed by non-denaturing electrophoresis [15 % (w/v) polyacrylamide slab gel] and visualized by staining with Coomassie Brilliant Blue.

Molecular mass

The molecular mass of the proteins was estimated using electrospray mass spectra recorded on a Bio-Q triple quadrupole instrument (Micromass). Samples were dissolved in 1 % (v/v) acetic acid/50 % (v/v) acetonitrile and injected into the ion source at a flow rate of 10 ml/min using a Phoenix syringe pump. Spectra were collected and elaborated by using the MASSLYNX software provided by the manufacturer. Calibration of the mass spectrometer was performed with horse heart myoglobin (16951.5 Da).

CD studies

CD spectra were recorded on a JASCO J-710 spectropolarimeter equipped with a Peltier thermostatic cell holder (model PTC-343; JASCO). Samples were at a concentration of 7×10^{-6} M in 10 mM sodium phosphate pH 7.0; a 1-mm path length cell was used. Molar ellipticity per residue (θ), expressed in deg. $cm^2 \cdot dmol^{-1}$, was calculated from the formula: $\theta = MW\ \Theta / 100\ cnl$, where Θ is the ellipticity observed (mdeg), MW the molecular mass, c the protein concentration ($mol \cdot dm^{-3}$), l the path length of the cuvette (cm) and n the number of the amino acid residues.

Denaturation/renaturation studies

Thermal denaturation of each protein was monitored by following the change in dichroism at 223 nm and recording spectra at 185–260 nm every 10 °C from 25 °C to 98 °C. The sample was heated at the rate of 1 °C/min. Similar spectra were observed after a prolonged incubation at the same temperature.

Protein unfolding by guanidine hydrochloride (GuHCl) and urea, at concentrations that avoided aggregation, was monitored by CD at 223 nm over 210–260 nm in 0–8 M GuHCl or 0–9 M urea at 25 °C. In order to monitor unfolding equilibrium transitions, samples were incubated at 25 °C for 1 h, 4 h and 24 h. Structure recovery of the proteins denatured in 8 M GuHCl or 9 M urea was checked after 20-fold dilution in 10 mM sodium phosphate, pH 7.0. The spectra were corrected by subtracting a buffer blank with the same GuHCl or urea concentrations. Predictions of the secondary structure percentages from the CD spectra were calculated by using the CD spectra of proteins with a known secondary structure as the reference, in combination with singular value decomposition and a variable selection procedure [21]. The method was modified using *E. coli* thioredoxin as an additional reference spectrum. The following recommended selection criteria were maintained: (a) the sum of secondary structures was between 0.9 and 1.10, (b) the secondary structures were greater than -0.05, and (c) the fit of the reconstructed CD lay within the noise level of experimental data, i.e. the root square of error was less than 0.22. The data obtained from the secondary structure analysis of the spectra recorded at 98 °C (maximum temperature achievable) for K18G, R82E, K18G/R82E and D102X were analysed by the standard linear decomposition method using Chang's CD reference spectra [22].

Thermodynamic studies

The transition curve at any stage of the denaturation is characterized by the fraction (f) of the protein in the denatured state as:

$$f = \theta_{obs} - \theta_{nat} / \theta_{den} - \theta_{nat}$$

where θ_{nat} and θ_{den} represent values of molar ellipticity characteristic of the native (N) and the denatured (D) state, respectively. The equilibrium constant K for the reaction $N \leftrightarrow D$ can be calculated by the equation:

$$K = f/1 - f = \theta_{nat} - \theta_{obs} / \theta_{obs} - \theta_{den}$$

The Gibbs free energy change from the native to the denatured state is given by $\Delta G^\circ = -RT\ln K$ at a different fixed temperature value. ΔG° was plotted against the temperature (25–98 °C), GuHCl (0–8 M) or urea (0–9 M) molarity for each protein.

The temperature of the midpoints of the transitions, T_m, for each protein was determined from the curve fitting of the θ_{obs} at 223 nm against the temperature plots. The concentrations of GuHCl or urea at the midpoints of the transitions were determined from the curve fitting of the θ_{obs} at 223 nm against the molarity of GuHCl or urea plots.

The results were processed using the Curve Fit Program and INPLOT 4 Program for IBM computers.

The enthalpy change of unfolding at T_m, ΔH_m, and the entropy change of unfolding at T_m, ΔS_m, were calculated by van't Hoff analysis. The difference, $\Delta\Delta G$, between the free-change of unfolding of the mutant and the wild-type proteins calculated at the T_m of the wild-type protein was estimated by the relationship given by Becktel and Schellman [23], $\Delta\Delta G = \Delta T_m \Delta S_m$ (wild-type), where ΔT_m is the change in T_m of the mutant protein relative to the wild-type protein, and ΔS_m (wild-type) is the entropy change of the wild-type protein at T_m. The Gibbs free energy change from the native to the denatured state, $\Delta\Delta G^\circ$, in the temperature range from 30 °C to 90 °C was calculated for each protein.

Nile Red binding

The binding study with Nile Red was conducted in a thermostatically controlled spectrofluorimeter (Perkin-Elmer LS-50B); samples were excited at 550 nm and the emission spectra 580–730 nm were recorded. The four mutant and the wild-type proteins (0.25 μg/μl) were dissolved in 10 mM sodium phosphate, pH 7.0, and incubated at a pre-determined temperature. At the same time, further protein samples were dissolved in either 10 mM sodium phosphate, pH 7.0/1 M GuHCl for 1 h or in 10 mM sodium phosphate, pH 7.0/2 M urea for 24 h before being incubated at a pre-determined temperature. Nile Red (1 μM) was added to the sample and the mixture was allowed to equilibrate for 2 min at the incubation temperature before spectra were recorded.

RESULTS

Choice of mutations

In a previous comparative study of the conformational and structural behaviour between the *E. coli* thioredoxin and BacTrx using computational techniques [12], certain features concerning the molecular basis of thermostability for this class of proteins were noted. The computational analysis revealed that the replacement of some residues in the BacTrx sequence could reduce the backbone flexibility and stabilize it when compared with *E. coli* thioredoxin. One of the most important factors leading to increased thermostability of BacTrx was the formation of new hydrogen bond and salt-bridge interactions between the side chains of the substituted residues. Some of these interactions also gave rise to hydrogen-bond networks that linked different secondary elements and increased the unfolding resistance. In fact the two NH of Arg-82 are H-bonded to the CO Gln-15, to the CO Gly-16 and to Oδ Asp-17 groups, while the two Lys-18 Nζ are H-bonded to the Oδ Asp-48 and Oϵ Gln-105. In addition, the analysis of the molecular dynamic simulations *in vacuo* at 300 K and 500 K, and in solution at 300 K, revealed that the helix in the C-terminal section played an important role in protecting the protein core.

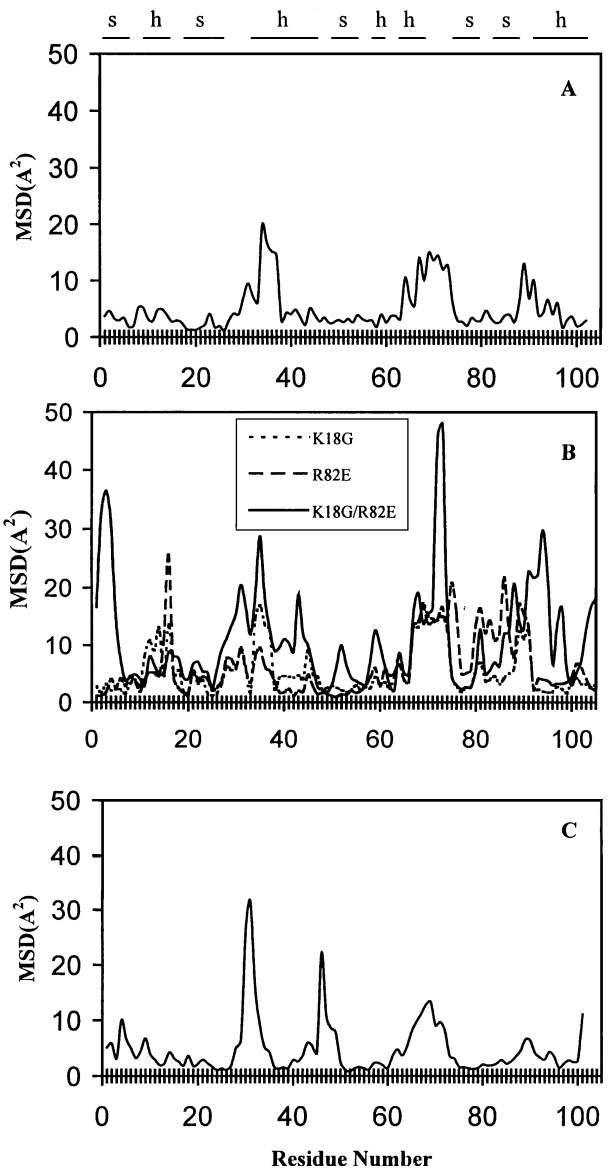

Figure 1 Results of molecular dynamic studies on BacTrx and mutated proteins

Maximum mean-square displacement (MSD) of each residue with respect to the minimized form as a function of the residue number. Molecular dynamic simulations *in vacuo* at 500 K of (**A**) BacTrx; (**B**) K18G, R82E and K18G/R82E; and (**C**) D102X. The secondary structure elements of thioredoxin are shown above the border of the top panel: h, helix; s, strand.

Using these data, four mutants were designed in which Lys-18 was replaced by glycine (K18G), present in *E. coli* thioredoxin, and Arg-82 was substituted by the corresponding glutamate present in *E. coli* thioredoxin (R82E). In the double mutant (K18G/R82E), both residues were replaced by the corresponding glycine and glutamate present in *E. coli* thioredoxin. In the D102X mutant, the last four amino acids at the C-terminal tail from D102 were deleted. The first three mutants were designed to verify the influence of the hydrogen bonds in the behaviour of BacTrx because Gly-21 in *E. coli* thioredoxin lacks hydrogen bonds, whereas Glu-85 is hydrogen bonded only to Lys-82; the D102X mutant was useful for clarifying the influence of the C-terminal tail in protein unfolding.

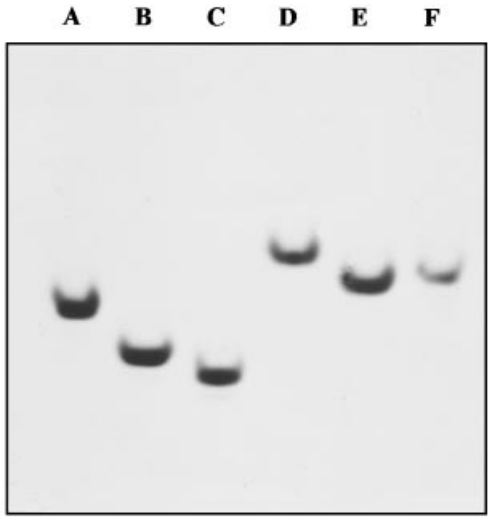

Figure 2 Native PAGE showing relative mobility of BacTrx and its mutant derivatives

Lanes (A) K18G, (B) R82E, (C) K18G/R82E, (D) D102X, (E) BacTrx and (F) *E. coli* thioredoxin.

Molecular dynamic analysis

Molecular dynamic techniques were used to analyse the dynamic behaviour of the four BacTrx mutants. In particular, four molecular dynamic simulations were carried out *in vacuo* at 500 K to obtain information on the thermal stability of the mutants and to compare the results with those of the wild-type protein.

Figure 1 illustrates the maximum-mean-square displacement of each residue with respect to the minimized form as a function of the residue number for the wild type and for the four mutants at 500 K. Comparative analysis of average structures at 500 K *in vacuo* suggests that the BacTrx structure is more stable than that of the mutants. In fact, the Lys-18 Nζ acts as H-bond donor with the Oδ1 Asp-48 and Oϵ Gln-105, and the NH of the Arg-82 side chain acts as H-bond donor with the CO Gln-15, the CO Gly-16 and the Oδ1 Asp-17. It is worth noting that these substitutions break the hydrogen-bond interactions which linked different secondary structure elements and bring about a structural instability.

The analysis of mean-square displacement for the proteins shown in Figure 1 revealed an increase in mobility for the active site in K18G/R82E and D102X. For K18G a mobility increase was observed only in the α1 helix region; for R82E, the α1, β4 and β5 secondary structures were similarly affected. Not surprisingly, the K18G/R82E presented a high degree of flexibility in all protein regions, with maxima in the N- and C-terminal tails and in the β4 region, which constitutes a portion of the protein core. In addition, the increase in flexibility of the Ile-72 residue in the loop region between the 3_{10} and β_4 elements can be explained in terms of destabilization of the β-sheet core.

An accurate inspection of the simulations of D102X carried out *in vacuo* at 500 K shows that the tail deletion dramatically increased protein mobility in the α2–β3 region, with high values for residues Ala-36 and His-46 of the α2 helix. This behaviour can be ascribed to the loss of the α4 helix, which protects the hydrophobic core, and a decrease in the hydrogen-bond interactions involving Ala-36 and His-46 because of the structural consequence of the partial unfolding of the α2 helix.

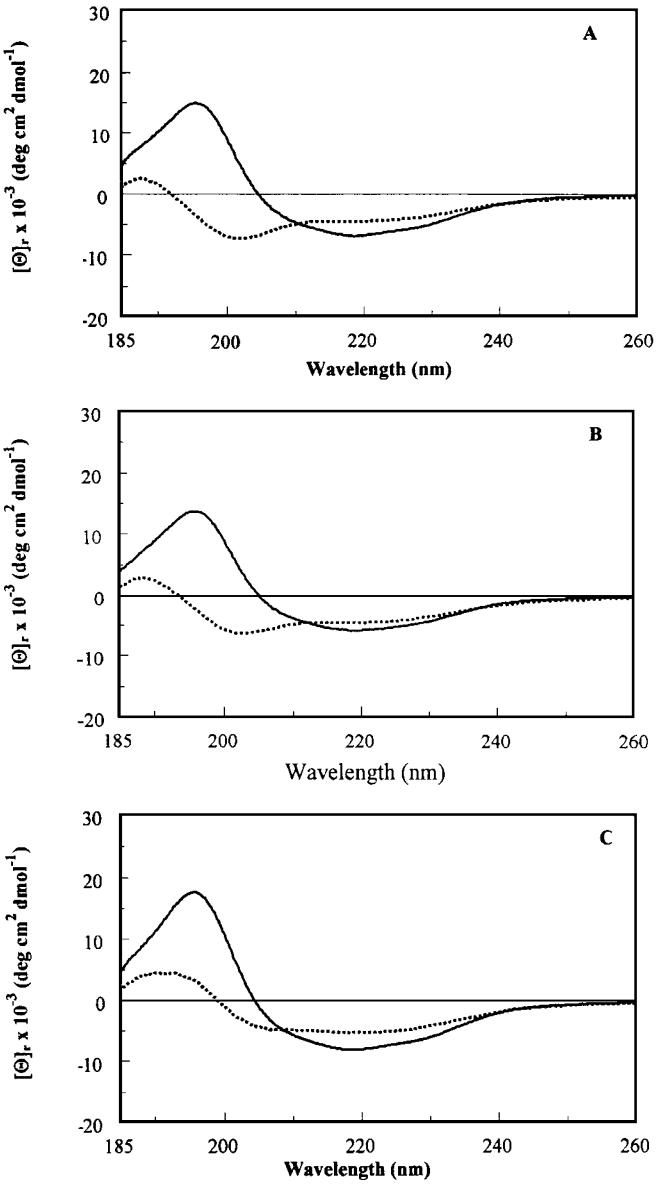

Figure 3 Far-UV CD spectra of BacTrx, K18G/R82E and D102X at 25 °C and 98 °C

(**A**) K18G/R82E, (**B**) D102X and (**C**) BacTrx at 25 °C (unbroken line) and 98 °C (dotted line). Samples were at a concentration of 7×10^{-6} M in 10 mM sodium phosphate, pH 7.0. The ellipticity ([Θ]) is given as the molar ellipticity/residue.

Production and characterization of wild-type and mutant thioredoxins

The protein yields from 1 litre of cultures were 15–20 mg for K18G, R82E and K18G/R82E and about half as much for D102X. Analysis of native PAGE (Figure 2), revealed only one band with different electrophoretic mobility for the mutants when compared with the BacTrx and *E. coli* thioredoxin according to the ionic nature of the mutations introduced.

The molecular mass of the proteins was analysed by electrospray MS. The observed mass of BacTrx was 11573.2$\pm$0.5 Da compared with the calculated mass of 11576 Da for the reduced form. The measured masses of K18G, R82E, K18G/R82E and D102X were 11502.2$\pm$0.5 Da, 11546.0$\pm$0.6 Da, 11475.0$\pm$

Table 1 Predicted secondary structure of BacTrx and mutated thioredoxin proteins at (a) 25 °C or (b) 98 °C

(a)

	Secondary structure (%)				
Protein ...	BacTrx	K18G	R82E	K18G/R82E	D102X
α-Helix	33.4	34.0	42.0	20.1	17.4
β-Sheet or extended	39.3	26.6	31.7	49.6	49.8
Disordered structure	27.3	39.4	26.3	30.3	32.8

(b)

	Secondary structure (%)				
Protein ...	BacTrx	K18G	R82E	K18G/R82E	D102X
α-Helix	13.8	9.4	10.3	7.8	10.0
β-Sheet or extended	53.1	35.4	33.7	35.3	35.0
Disordered structure	33.1	55.2	56.0	56.9	55.0

0.4 Da and 11118.7 $\pm$ 0.4 Da respectively, and thus in good agreement with the introduced mutation. It was therefore possible to conclude that the mutant thioredoxins contained the desired substitutions.

All thioredoxins obtained proved to be active using the Holmgren assay [18], thus demonstrating that the substitutions were not disruptive. The most dramatic effect was with D102X, which showed a 30% decrease in insulin reductase activity.

The redox state of the various thioredoxins was checked by spectrofluorimetry; all the proteins proved to be in an oxidized form [24].

Denaturation of wild-type and mutant thioredoxins

In order to compare the stabilities of the mutant and wild-type proteins against thermal denaturation, CD spectra in the 185–260 nm region were recorded and the magnitude of the CD band at 223 nm was followed at different temperatures. While near-UV CD spectra reflect alterations in tertiary structure, ellipticity at 223 nm is usually correlated with the α-helical content of a protein. Positive increases in the rotation values at 223 nm are assumed to be a measure of the decrease in the α-helical content of a polypeptide [25]. The reversibility of the transition was then checked by lowering the temperature.

Figure 3 (A–C) show the CD spectra in the far-UV region at 25 °C and 98 °C for the double mutant, D102X and BacTrx respectively. The data presented in Figure 3 and the calculated secondary structure (Table 1) show that, when the temperature is increased, all the mutants undergo significant structural variations, including an increase in disordered structures. However, the spectra of BacTrx revealed greater structural preservation up to 98 °C. The predicted secondary structure of BacTrx and of the two single mutants at 25 °C indicated that the single mutations scarcely affected the overall secondary structure. On the contrary, as predicted, the double mutant K18G/R82E and D102X showed a decrease in α-helical content and a consequent increase in extended and disordered structures. This was a result of the replacement of two important amino acids in the double mutant K18G/R82E and the deletion of the C-terminal tail, which is fundamental in the stabilization of the structure of BacTrx, in D102X. After denaturation, caused by increasing the temperature from 25 °C to 98 °C, the proteins were still able to catalyse the reduction of bovine insulin disulphides, with approx. 70% residual activity.

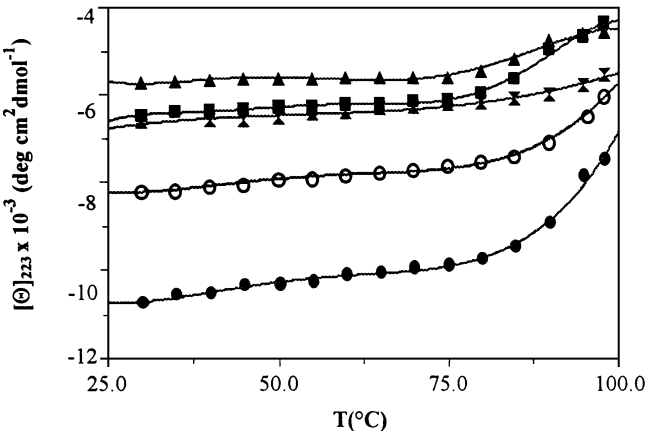

Figure 4 Melting curves for BacTrx and mutated proteins

Temperature dependency of ellipticity ($[\Theta]_{223}$) for BacTrx (⊠), K18G (●), R82E (○), K18G/R82E (■) and D102X (▲).

Table 2 GuHCl and urea denaturation values calculated from sigmoidal curves of BacTrx and mutated thioredoxin proteins

Melting temperatures (T_m) in the presence of *1 M GuHCl or †2 M urea were calculated from sigmoidal curves of BacTrx and mutated thioredoxin proteins.

Protein	Urea (M)	GuHCl (M)	T_m (°C)*	T_m (°C)†
BacTrx	6.7	2.9	84.5	88.2
K18G	5.2	2.6	78.1	78.8
R82E	5.5	2.7	81.5	81.4
K18G/ R82E	4.6	2.4	76.3	73.1
D102X	4.0	2.0	65.2	70.6

The denaturation curves of the proteins, obtained by plotting the ellipticity at 223 nm against temperature, are shown in Figure 4. All proteins exhibited reversible denaturation and activity was completely restored when tested using the Holmgren assay [18] by following the dithiothreitol-dependent reduction of bovine

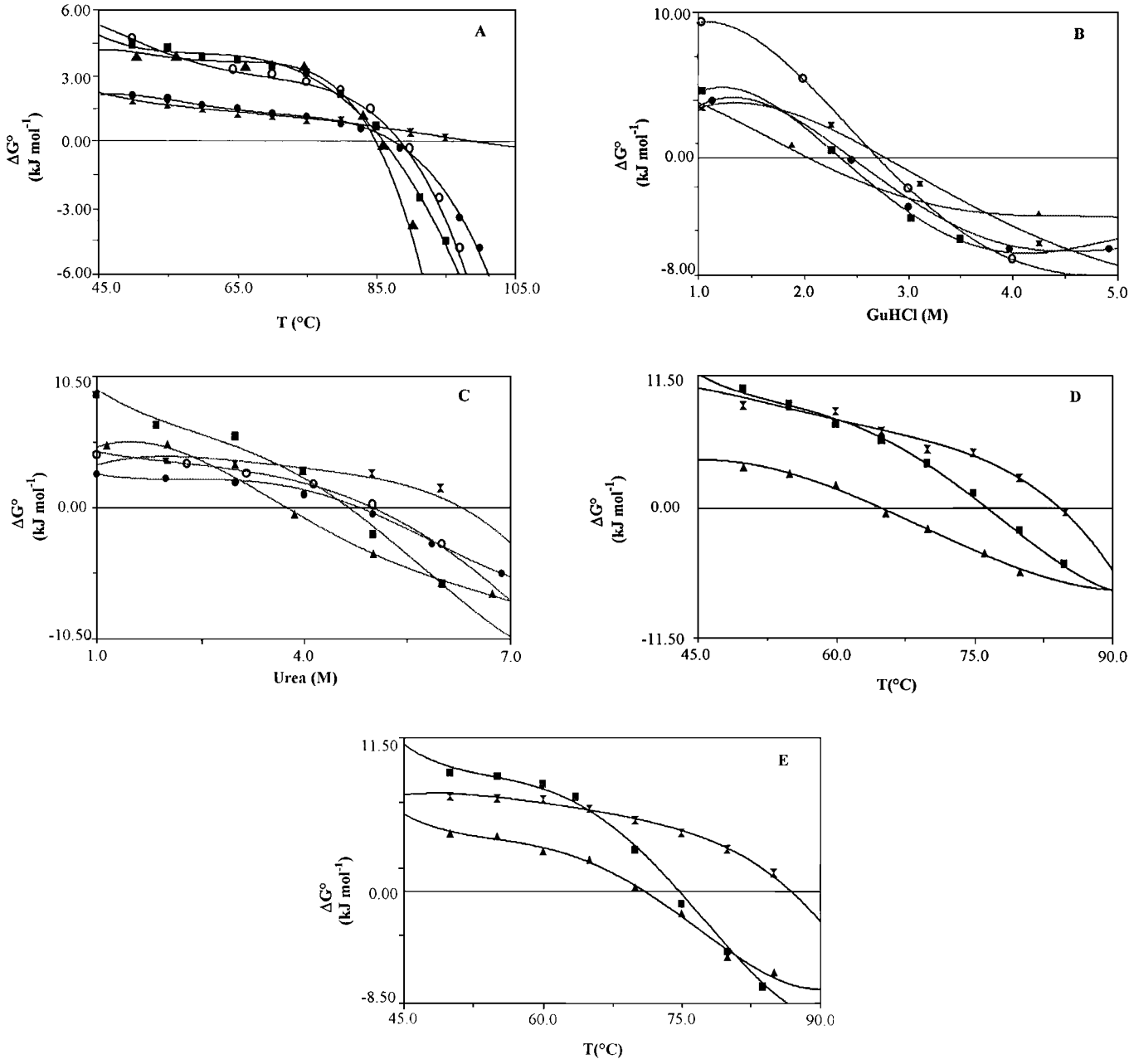

Figure 5 ΔG° as a function of temperature, GuHCl and urea

(**A**) ΔG° with temperature; (**B**) ΔG° with GuHCl concentration; (**C**) ΔG° with urea concentration; (**D**) ΔG° with temperature in the presence of 1 M GuHCl; (**E**) ΔG° with temperature in the presence of 2 M urea. BacTrx (⨂), K18G (●), R82E (○), K18G/R82E (■) and D102X (▲). ΔG° values were calculated from the results obtained from CD spectra.

insulin disulphides. The double mutant and the truncated mutant were significantly less stable, with a T_m of 88.7 °C and 86.1 °C respectively compared with a T_m of 102.6 °C for the wild-type protein (the BacTrx value was extrapolated from the sigmoidal curve because there was no observable transition in the denaturation curve); the T_m value calculated was in agreement with the value obtained by differential scanning calorimetry [11]. Similarly K18G and R82E, at 98 °C, did not reach the complete plateau of the sigmoidal curve and an extrapolated T_m value of 90.6 °C was calculated for both mutant proteins. It is interesting to note that not all of the mutations proved disruptive and, in fact, the T_m of each of the mutants was found to be more stable [26] than that of *E. coli* thioredoxin, thus demonstrating that, even after mutation, BacTrx is still endowed with higher conformational heat stability.

The stability of the proteins in GuHCl (0–8 M) and urea (0–9 M) was analysed by CD at a constant temperature (25 °C) after incubation in the presence of GuHCl for 1 h or in the presence of urea for 24 h. The minimum times necessary to complete the unfolding equilibrium transitions and the optimal conditions for complete reversible denaturation have been determined previously. The proteins revealed different resistances to denaturants and the wild-type always proved the most resistant (Table 2).

Table 3 Parameters characterizing the thermal denaturation of BacTrx and its mutants

[a] The melting temperature (T_m) is the temperature at the midpoint of the thermal denaturation transition shown in Figure 4. [b] The difference in the melting temperature between BacTrx and mutant proteins, ΔT_m, was calculated as T_m (mutant) $- T_m$ (BacTrx). [c] The difference between the free-energy change of unfolding of BacTrx and the mutated proteins at the T_m of BacTrx, $\Delta\Delta G$, was calculated by the relationship given by Becktel and Schellman [23], $\Delta\Delta G = \Delta T_m \Delta S_m$ (BacTrx), as described in the Materials and methods section; negative values indicate less stable proteins than BacTrx.

Protein	ΔH_m (kJ · mol^{-1})	ΔS_m (kJ · mol^{-1} · K^{-1})	T_m (°C)[a]	ΔT_m (°C)[b]	$\Delta\Delta G_m$ (kJ · mol^{-1})[c]	$\Delta\Delta G°$ (30 °–90 °C) (kJ · mol^{-1})
BacTrx	288.6	0.785	102.6			2.4
K18G	276.0	0.76	90.6	−12	−9.42	8.3
R82E	301.6	0.83	90.6	−12	−9.57	8.22
K18G/R82E	196.2	0.547	88.7	−13.9	−10.9	9.51
D102X	339.2	0.95	86.1	−16.5	−12.9	9.63

The heat stability of each protein was determined by the trend of the melting curves (T_m) obtained from the change in the far-UV CD at 223 nm at various temperatures and in the presence of 1 M GuHCl or 2 M urea (concentrations of denaturant at which proteins are fully active) (see Figure 5). The calculated T_m are shown in Table 2. A plateau was reached in all the curves. The wild-type was always the most stable protein, with a T_m of 84.5 °C in 1 M GuHCl or 88.2 °C in 2 M urea. These values are approx. 20 °C higher than the T_m obtained for D102X, 65.2 °C and 70.6 °C respectively.

Thermodynamic evaluations

Since the ΔH_m values for the wild-type protein obtained with van't Hoff analysis or by differential scanning calorimetry [11] were similar (69 kcal/mol), a two-state transition for unfolding can be assumed. The Gibbs free energy from the native to the denatured state ($\Delta G°$) was plotted against temperature (Figure 5A), GuHCl concentration (Figure 5B) or urea concentration (Figure 5C). In addition, data were also obtained by plotting $\Delta G°$ against temperature in the presence of 1 M GuHCl (Figure 5D), or temperature in the presence of 2 M urea (Figure 5E). Figure 5 confirms higher thermostability and resistance to denaturants for the wild-type protein when compared with the mutant thioredoxins and, in fact, it was observed that the $\Delta G° = 0$ for the wild-type protein corresponded to higher values in all the conditions examined. The $\Delta G°$ plotted against GuHCl concentration shows no great difference among the proteins.

The stability of each mutant at the T_m of the wild-type was estimated using the relation $\Delta\Delta G = \Delta T_m \Delta S_m$ (see the Materials and methods section and [23]). The values thus obtained are summarized in Table 3. It is worth noting that by introducing a single mutation, a great decrease in thermostability was obtained with a $\Delta\Delta G_m$ of -9.42 kJ · mol^{-1} for K18G and -9.57 kJ · mol^{-1} for R82E. The different values of $\Delta\Delta G°$, observed in the transition from 30 °C to 90 °C, confirm the differences in thermostability among the proteins; 2.4 kJ · mol^{-1} were needed by the wild-type protein for the transition, whereas 9.63 kJ · mol^{-1} were necessary for the same transition in the less stable D102X.

Structural investigations by Nile Red binding

Nile Red is a useful compound for the detection of major or minor hydrophobic areas on protein surfaces [27], as the dye fluorescence is greatly enhanced upon binding to hydrophobic surfaces. At room temperature, BacTrx and the mutant proteins interact weakly with the dye thus, as expected, proving very similar in surface hydrophobicity. Fluorescence spectra were recorded in conditions considered critical by CD experiments:

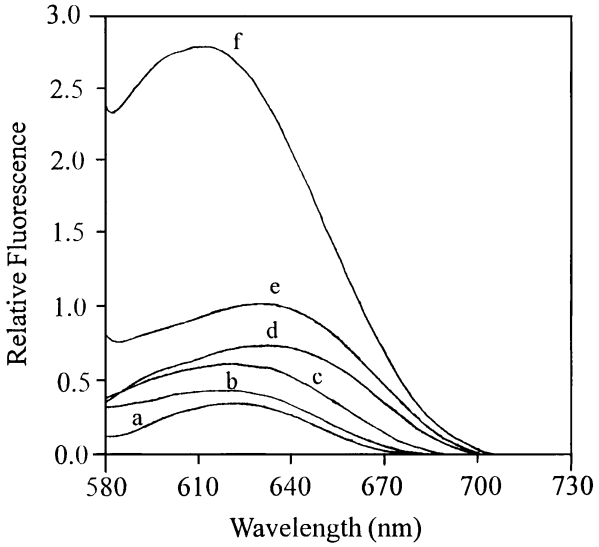

Figure 6 Fluorescence of Nile Red in the presence of BacTrx proteins

The relative fluorescence of Nile Red was measured after incubation of the proteins at 85 °C for 1 h. Emission spectra are shown for (a) buffer, (b) BacTrx, (c) R82E, (d) K18G, (e) K18G/R82E and (f) D102X. In each case the concentration of the protein was 0.25 μg/μl and that of Nile Red was 1 μM.

after incubation for 1 h at 85 °C (T_m of D102X) (Figure 6), or 1 h in the presence of 1 M GuHCl at 25 °C and then for 30 min at 65 °C (T_m of D102X), or 24 h in the presence of 2 M urea at 25 °C and then for 45 min 70 °C (T_m of D102X). A consistent increase in emission intensity and a blue shift to 608 nm as a result of binding of Nile Red was observed for D102X. A similar trend was observed in all the denaturation conditions examined, except for the wild-type protein where no marked binding of Nile Red was observed in any of the conditions tested.

DISCUSSION

In the present study, structure-based modelling was used to predict the protein engineering of residues which are crucial for the stability of *B. acidocaldarius* thioredoxin. An analysis of the three-dimensional model constructed for BacTrx showed a shortening of loop regions and an increase in the number of ion pairs and hydrogen bonds, referring to the well-known three-dimensional structure of *E. coli* thioredoxin [9,10]. On the basis of molecular dynamic simulations, we substituted the Lys-18 and

Arg-82 residues, which seem to be involved in stabilizing hydrogen bonds and ion pairs, with the corresponding residues present in the less stable *E. coli* thioredoxin, . Both residues were substituted in the double mutant K18G/R82E. The structural role of the C-terminal helix, which is anchored by many salt bridges to the rest of the *B. acidocaldarius* core (Lys-85–Asp-102, Lys-18–Gln-105, Lys-85–Gln-99, Gln-99–Gln-91, Gln-105–Asp-48), was also investigated by deleting four amino acids from the residue D102. The effects of the substitutions on the stability of BacTrx were monitored by a wide array of modelling, spectroscopic and thermodynamic tests. The predictions were in total agreement with the experimental data. The results obtained by comparing the mutants and the wild-type protein clearly confirmed the important role of the Lys-18 and Arg-82 residues, and the C-terminal α-helix region in stabilizing the whole structure of BacTrx. The four thioredoxin mutants proved to be less resistant to denaturants and to temperature changes, with a shift in T_m of 12–16 °C with respect to the wild-type protein. This trend was confirmed by the thermodynamic studies, where a negative $\Delta\Delta G_m$ compared with the wild-type was observed [28]. When the two single mutants were considered, K18G was found to be the less stable, perhaps because of a different structural organization, already suggested in simulation studies and secondary structure predictions (Figure 1, Table 1). *E. coli* thioredoxin, used as a control [11], was always shown to be the least thermoresistant when compared with the mutant proteins, except for the D102X protein where a more dramatic change had been introduced into the structure.

Binding studies with Nile Red [29] tested the response of BacTrx and the mutant proteins to denaturing agents. In fact, using similar experimental denaturing conditions, a significant increase in the amount of bound dye was detected only for the D102X mutant and to a lesser extent for K18G/R82E, because of exposure of hydrophobic surfaces found only in the mutant proteins.

From a thermodynamic point of view, it is worth noting that the difference in the free energy associated with the transition from 30 °C to 90 °C ($\Delta\Delta G$ °) was lower for the wild-type protein compared with the two mutants K18G/R82E and D102X. This is in complete accordance with the knowledge that thermophilic proteins have a very compact structure and the structure of the two constructed mutants was more flexible.

Interestingly, comparison of the dependency of ΔG° on GuHCl concentration revealed no great difference among the proteins, possibly because GuHCl particularly affects salt bridges and thus eliminates the contribution of these to the stability of the proteins examined.

According to the simulation data, these findings confirm that networks of ion pairs and hydrogen bonds in specific settings may contribute significantly to the thermal stability and denaturant resistance of proteins [30].

It has already been shown that structurally mutant thioredoxins exhibit perturbed thermodynamic properties [31] and changes in heat capacity [32], whereas their biological activity is not affected. The authors concluded that in thioredoxins the link between stability and function is not as strong as that found in other systems. According to the data in the present study, the mutant K18G, R82E and K18G/R82E proteins proved to be fully active with the Holmgren assay. The molecular dynamic simulations confirmed that the reduced mobility of the catalytic site was preserved. The effect of the C-terminal deletion was more dramatic, and the D102X protein proved not only to be less stable but also less active in the Holmgren assay. Simulation data on this mutant show greater mobility of the catalytic site when compared with the other mutant and wild-type proteins.

These results show that simulation procedures, molecular biology and biochemical techniques can be used together to direct protein engineering and/or site-directed mutagenesis.

This work was supported by grants from the Ministero dell'Università e della Ricerca Scientifica e Tecnologica (40%), the BIOTECH Programme of the European Union 'Extremophiles as Cell Factories', contract n ° BIO4-CT96-0488, and by a C. N. R. Target Project on Biotechnology C.N.R. Programma Biotecnologie Industriali e Metodologie Innovative.

REFERENCES

1 Menendez-arias, L. and Argos, P. (1989) J. Mol. Biol. **206**, 397–406
2 Jaenicke, R. (1996) FASEB J. **10**, 84–92
3 Takagi, H., Takahashi, T., Momose, H., Inouye, M., Maeda, Y., Matsuzawa, H. and Takahisa, O. (1990) J. Biol. Chem. **265**, 6874–6878
4 Serrano, L., Horowitz, A., Avron, B., Bycroft, M. and Fersht, A. R. (1990) Biochemistry **29**, 9343–9352
5 Biro, J., Fabry, S., Dietmaier, W., Bogedain, C. and Hensel, R. (1990) FEBS Lett. **275**, 4441–4444
6 Wallon, G., Kryger, G., Lovett, S. T., Oshima, T., Ringe, D. and Petsko, G. A. (1997) J. Mol. Biol. **266**, 1016–1031
7 De Prat, G., Johnson, C. M. and Fersht, A. R. (1994) Protein Eng. **7**, 103–108
8 Eklund, H., Gleason, K. and Holmgren, A. (1991) Proteins: Struct. Funct. Genet. **11**, 13–28
9 Jeng, M. F., Campbell, A. P., Begley, T., Holmgren, A., Case, D. A., Wright, P. E. and Dyson, H. J. (1994) Structure **2**, 853–868
10 Katti, S. K., LeMaster, D. M. and Eklund, H. (1990) J. Mol. Biol. **212**, 167–184
11 Bartolucci, S., Guagliardi, A., Pedone, E., de Pascale, D., Cannio, R., Camardella, L., Rossi, M., Nicastro, G., de Chiara, C., Facci, P. et al. (1997) Biochem. J. **328**, 277–285
12 Pedone, E., Bartolucci, S., Rossi, M. and Saviano, M. (1998) J. Biomol. Struct. Dyn. **16**, 437–446
13 Weiner, S. F., Kollman, P. A., Nguyen, D. T. and Case, D. A. (1986) J. Comp. Chem. **7**, 230–252
14 Saviano, M., Aida, M. and Corongiu, G. (1991) Biopolymers **31**, 1017–1024
15 Saiki, R. K. (1990) PCR Protocols. A Guide to Methods and Applications. Academic Press Inc., San Diego
16 Brent, R. and Ptashne, M. (1981) Proc. Natl. Acad. Sci. U.S.A. **78**, 4204–4208
17 Smith, P. K., Krohn, R. I., Hermanson, G. T., Mallia, A. K., Gartner, F. H., Provenzano, M. D., Fujimoto, E. K., Goeke, N. M., Olson, B. J. and Klenk, D. C. (1985) Anal. Biochem. **150**, 76–85
18 Holmgren, A. (1979) J .Biol. Chem. **254**, 9627–9632
19 Laemmli, U.K. (1970) Nature (London) **227**, 680–685
20 Rabilloud, T., Vaillard, L., Gilly, C. and Lawrence, J. J. (1994) Cell. Mol. Biol. **40**, 57–75
21 Johnson, W. C. (1988) Annu. Rev. Biophys. Biophys. Chem. **17**, 145–166
22 Chang, C. T., Wu, C. S. C. and Yang, J. T. (1978) Anal. Biochem. **91**, 13–31
23 Becktel, W. J. and Schellman, J. A. (1987) Biopolymers **26**, 1859–1877
24 Holmgren, A. (1972) J. Biol. Chem. **247**, 1992–1998
25 Johnson, W. C. (1990) Proteins: Struct. Funct. Genet. **7**, 205–214
26 Ladbury, J. E., Wynn, R., Hellinga, H. W. and Sturtevant, J. E. (1993) Biochemistry **32**, 7526–7530
27 Sackett, D. L. and Wolff, J. (1987) Anal. Biochem. **167**, 228–234
28 Kawamura, S., Kakuta, Y., Tanaka, I., Hikichi, K., Kuhara, S., Yamasaki, N. and Kimura, M. (1996) Biochemistry **35**, 1195–1200
29 Kotik, M. and Zuber, H. (1993) Eur. J. Biochem. **211**, 267–280
30 Pappenberg, G., Schurig, H. and Jaenicke, R. (1997) J. Mol. Biol. **274**, 676–683
31 Hellinga, H. W., Wynn, R. and Richards, F. M. (1992) Biochemistry **31**, 11203–11209
32 Ladbury, J. E., Wynn, R., Hellinga, H. W. and Sturtevant, J. M. (1995) Biochemistry **34**, 2148–2152

Received 23 July 1998/23 November 1998; accepted 2 February 1999