An Elongated Form of T4 Glutaredoxin with Four Extra Residues†

Matti Nikkola,*,‡ Åke Engström,§ Markku Saarinen,‡ Margareta Ingelman,‡ Thorleif Joelson,‡ and Hans Eklund‡

Department of Molecular Biology, Swedish University of Agricultural Sciences, and Department of Immunology, Uppsala University, Uppsala Biomedical Center, Uppsala, Sweden

Received December 29, 1992; Revised Manuscript Received April 23, 1993

ABSTRACT: Two different forms of T4 glutaredoxin (thioredoxin) arising from the same gene on a multicopy plasmid in an *Escherichia coli* expression system have been isolated and characterized. Up to one-fourth of purified T4 glutaredoxin has an extension of four amino acids in the carboxy terminus, with the sequence aspartate, arginine, isoleucine, lysine. This four-residue extension may be caused by a translational +1 frameshift at the UGA terminator codon.

T4 glutaredoxin, previously called T4 thioredoxin, is a small globular redox protein, found in cells of *Escherichia coli* infected by the bacteriophage T4. The protein was identified as a thioredoxin on the basis of its activity with the thioredoxin reductase of *E. coli* and the ribonucleotide reductase encoded by phage T4 (Berglund, 1969). The primary structure of T4 glutaredoxin (Figure 1) with the typical thioredoxin motif of two cysteines was established by sequencing T4 glutaredoxin isolated from phage-infected cells (Sjöberg & Holmgren, 1972).

The *nrdC* gene coding for T4 glutaredoxin has been cloned, sequenced, and expressed successfully in systems employing phage λ p_L expression vectors (LeMaster, 1986; Joelson *et al.*, 1990). A number of mutant glutaredoxins have been constructed with site-directed mutagenesis, and their activities with different enzymes and substrates have been assayed (Joelson *et al.*, 1990; Nikkola *et al.*, 1991, 1993).

The three-dimensional structure of T4 glutaredoxin was established using crystals of glutaredoxin purified from phage-infected cells (Sjöberg & Söderberg, 1976), and the structure was solved at 2.8-Å resolution (Söderberg *et al.*, 1978). The refinement of the X-ray structure of T4 glutaredoxin at 1.45 Å has been completed with crystals of a mutant glutaredoxin (V15G;Y16P) diffracting far beyond the crystals of wild-type glutaredoxin (Eklund *et al.*, 1992).

In this paper, we describe a newly identified secondary form of T4 glutaredoxin, arising from the same *nrdC* gene, but with four extra residues following the last residue of the previously described amino acid sequence, and the purification and separation of the two forms.

MATERIALS AND METHODS

Purification of Glutaredoxins. The purification, cloning, and mutagenesis of T4 glutaredoxins have been described (Berglund & Sjöberg, 1970; Sjöberg, 1972; Joelson *et al.*, 1990; Nikkola *et al.*, 1991).

```
          10   (VY)  20          30          40          50
MFKVYGYDSN IHKCGPCDNA KRLLTVKKQP FEFININPEK GVFDDEKIAE

          60          70          80
LLTKLGRDTQ IGLTMPQVFA PDGSHIGGFD QLREYFK
                    70          80          90
               X: PQVFA PDGSHIGGFD QLREYFKDRI K
```

FIGURE 1: Primary structure of T4 glutaredoxin V15G;Y16P. The active-site sequence of wild-type T4 glutaredoxin is shown in parentheses. X denotes the amino acid sequence of the cyanogen bromide fragment X of the longer form of glutaredoxin showing the four extra residues at the carboxy terminus.

E. coli expression strain N4830 (Gottesman *et al.*, 1980) with the genotype [F- *su*-, *his*-, *ilv*-, gal+, Δ8(λ*cI857*, Δ*BAM*, Δ*HI*] harboring the expression plasmid pMG561 (Heaphy *et al.*, 1987) with the cloned *nrdC* gene was grown in 2-L batch cultures in rich medium (2×YT) with vigorous shaking at +34 °C until the optical density A_{640} reached 0.5–1.0, or overnight. The phage λ promoter in the expression plasmid was derepressed by shifting the temperature to 40–42 °C. The cultures were shaken at the higher temperature for 2–3 h, after which the cultures were rapidly cooled down, and the cells were harvested by centrifugation. The yield of cells was typically 5–6 g/L of culture.

The cells were washed once and the cell pellets stored at −80 °C. The cells were lysed by grinding the frozen cells with alumina. The lysis buffer was 50 mM Tris-HCl, pH 8.2, and 2 mM EDTA. Cell debris was removed by centrifugation. Nucleic acids were precipitated by addition of streptomycin sulfate to a concentration of 1%, followed by centrifugation. The supernatant was cut with ammonium sulfate at 40 and 85% saturation, at +4 °C, and the pellet from the latter cut was dissolved in 50 mM Tris-HCl, pH 8.2, and 2 mM EDTA and dialyzed against the same buffer.

The proteins from the previous step were applied onto a DEAE CL-6B (Pharmacia) ion-exchange column and eluted with a 0–0.2 M NaCl gradient in 50 mM Tris-HCl, pH 8.2, and 2 mM EDTA at +4 °C. The fractions were analyzed on 20% homogeneous SDS–polyacrylamide gels (High Density PhastGel, Pharmacia), where the two different forms could be detected well separated in the same lane. The fractions that contained glutaredoxin were pooled and concentrated, and dialyzed against 50 mM MES (4-morpholineethanesulfonic acid), pH 6.5, and 2 mM EDTA. The sample was then applied onto an S-Sepharose Fast Flow or a Mono-S column (both from Pharmacia) at room temperature. The different forms of glutaredoxin were eluted from the column in two separate peaks using a 0–0.2 M NaCl gradient in 50 mM MES, pH 6.5, and 2 mM EDTA. The fractions

† This study was supported by grants from the Finnish Academy, the Swedish Natural Science Research Council, the Swedish National Board for Technical Developments, and the Swedish Research Council for Engineering Sciences (to A.E.).

* Correspondence should be addressed to this author at the Department of Molecular Biology, Swedish University of Agricultural Sciences, Biomedical Center, Box 590, S-75124, Uppsala, Sweden. Telephone: +46-18-174550. Telefax: +46-18-536971.

‡ Department of Molecular Biology, Swedish University of Agricultural Sciences.

§ Department of Immunology, Uppsala University.

0006-2960/93/0432-7133$04.00/0

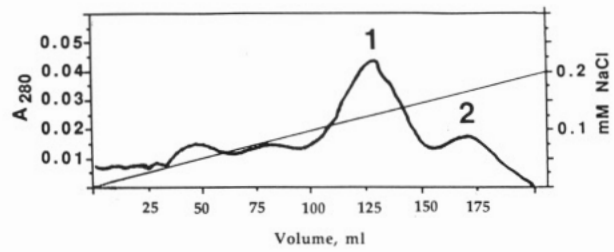

FIGURE 2: Separation of 1.5 mg (total) of the two forms of glutaredoxin by ion-exchange chromatography with S-Sepharose Fast Flow in a Pharmacia 16/50 column with a flow rate of 10 mL/min. The first peak corresponds to the shorter form and the second to the longer form of T4 glutaredoxin.

containing glutaredoxin were first concentrated in an Amicon cell using YM5 membranes, and further in Centricon 3 centrifugation cartridges. The protein concentrations were estimated using extinction coefficients determined earlier (Joelson *et al.*, 1990).

Analysis of Primary Structure Differences. Two hundred micrograms of T4 glutaredoxin of the two separate peaks was cleaved with cyanogen bromide in 70 μL of 70% formic acid. Digestion with *Staphylococcus aureus* V8 protease was done in 0.1 M NH_4HCO_3 with an enzyme to substrate ratio of 1 to 25 for 3 h at 37 °C.

Separation of fragments was performed by reversed-phase chromatography on a RP18 column, 2.1 × 30 mm (Brownlee Labs), and eluted with a gradient of 1% per minute of acetonitrile in water with a flow rate of 0.2 mL/min. Both solvents contained 0.1% trifluoroacetic acid. The eluate was monitored at 214 nm, and fractions were collected manually.

Plasma desorption mass spectrometry (Sundqvist & MacFarlane, 1985) was done with a BIOION 20 instrument, and amino acid sequence analysis was done with a 477A amino acid sequencer (both instruments from Applied Biosystems).

RESULTS

The final purification of the mutant T4 glutaredoxin V15G;Y16P by ion-exchange chromatography produced two fractions containing glutaredoxin (Figure 2). The proteins in the two fractions (peaks 1 and 2) had a slightly different mobility upon SDS gel electrophoresis (Figure 3).

Mass spectrometric analysis of the two fractions revealed that the detected m/z (9934) of the protein in peak 1 was close to the calculated m/z of the mutant glutaredoxin V15G;Y16P of 9921. The detected m/z of 10 450 for the protein in peak 2 is significantly higher than the value expected for glutaredoxin. Amino acid sequence analysis of the two proteins showed an identical sequence for the five amino-terminal residues analyzed.

The two proteins were treated with CNBr, and the two digests were again analyzed by plasma desorption mass spectrometry. Interpretation of the resulting mass analysis data (Table I) shows that the expected m/z for all cyanogen bromide fragments of glutaredoxin could be found in the analysis of the digest of the protein in peak 1. However, in the analysis of the cyanogen bromide digest of the protein in peak 2, the expected m/z of fragment 66–87 was missing, and an m/z of 3019 appeared. The peptide corresponding to the m/z of 3019 was isolated by reverse-phase chromatography and analyzed by mass spectrometry after digestion with *Staphylococcus aureus* protease. The resulting spectra showed the presence of the glutaredoxin *S. aureus* protease fragment 66–84 (m/z of 2071) and another fragment (fragment Y in Table I) with an m/z of 971 instead of the expected 457 for fragment 85–87.

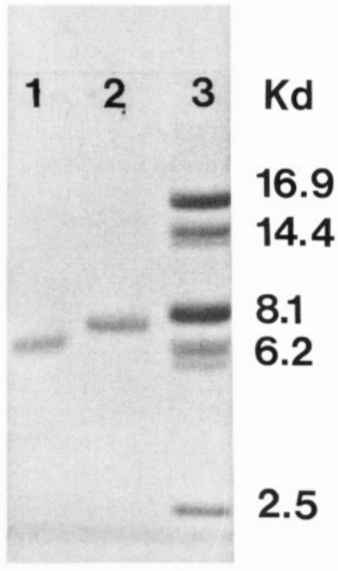

FIGURE 3: SDS–polyacrylamide gel electrophoresis of the protein peaks from Figure 2. Lane 1, shorter form of glutaredoxin (peak 1); lane 2, longer form of glutaredoxin (peak 2); lane 3, molecular mass markers.

Table I: Mass Spectrometric Analysis of Peaks 1 and 2 of Glutaredoxin and of Fragments Obtained by Cyanogen Bromide and *S. aureus* Protease Digestion[a]

	glutaredoxin calcd m/z	peak 1 detected m/z	peak 2 detected m/z
intact protein	9921	9934	10450
CNBr 1–37 Met(1) = sulfoxide	4316	4316	4310
CNBr 38–65	3071	3071	3066
CNBr 66–87	2509	2509	nd
fragment X			3019
S. aureus protease fragments of X			
66–84	2071		2071
85–87	457		nd
fragment Y			971

[a] Numbering of amino acids as in Figure 1.

```
Gln-Leu-Arg-Glu-Tyr-Phe-Lys-End
CAA-TTG-CGG-GAA-TAC-TTT-AAA-TGA-TAG-AAT-TAA-ATG-AAC
                            ::
                           -GAT-AGA-ATT-AAA-TGA-
Gln-Leu-Arg-Glu-Tyr-Phe-Lys--Asp-Arg-Ile-Lys-End
```

FIGURE 4: Nucleotide sequence around the C-terminus of the *nrdC* gene of bacteriophage T4 coding for T4 glutaredoxin (LeMaster, 1986) with the corresponding carboxy-terminal amino acid sequences of the two forms of T4 glutaredoxin.

Interpretation of the nucleotide sequence coding for glutaredoxin shows that a deletion of a single base in the normally used stop codon would create a protein with a mass 512.8 higher than the expected mass for glutaredoxin (Figure 4). The observed differences in the expected mass in the mass spectrometric analyses of the intact protein in peak 2 and the cyanogen bromide and *S. aureus* protease derived peptides (X and Y in Table I) are 529, 510, and 514, respectively.

Amino acid sequence analysis of the cyanogen bromide fragment X, with an m/z of 3019, showed the amino acid sequence corresponding to the peptide starting at position 66 in glutaredoxin, but with an extension of four amino acids at the carboxy terminus with the sequence aspartic acid, arginine, isoleucine, and lysine.

All the results shown here originate from the two glutaredoxin fractions of the mutant glutaredoxin V15G;Y16P. However, we have found the longer form in many other

glutaredoxin preparations and have purified some of the longer forms of the glutaredoxins. The recoveries of total T4 glutaredoxin from the preparations varied from less than a milligram to 10 mg/L of culture. The longer form constitutes up to one-fourth of the total glutaredoxin in the final stages of purification.

DISCUSSION

Translational frameshifts have been found in a number of organisms [for a review, see Atkins *et al.* (1990)]. Such frameshifts are now thought to be an important mechanism of gene regulation, even in bacteriophages like T4, which served as the model organism for the elucidation of triplet reading in translation (Barnett *et al.*, 1967; Crick & Brenner, 1967).

The *nrdC* gene coding for T4 glutaredoxin has two adjacent stop codons after the last protein coding codon and a third stop codon in the vicinity (Figure 4). The stop codons are of different types (UGA, UAG, and UAA), and ought to form a very powerful way of halting any further translation. The observed protein sequence could be caused by a translational frameshift in the stop codon UGA by a four-base translocation of the mRNA in the ribosome. If the +1 frameshift occurs, skipping over the U in the stop codon, the observed protein sequence will be formed. Lysine is the final amino acid of both forms of glutaredoxin. Moreover, the corresponding nucleotide sequences are identical with the sequence UU-AAA-UGA. It is puzzling why we do not see the same phenomenon at the second UGA.

The biological relevance of the observed frameshift in the expression of the T4 glutaredoxin gene is unknown. Glutaredoxin from *E. coli* has been reported having two forms differing in size (Björnberg & Holmgren, 1991; Sandberg *et al.*, 1991). In the glutaredoxin coded by the bacterial genome, the difference is caused by a highly charged N-terminal extension of five residues. The sequence is Met-Arg-Arg-Glu-Ile, and it arises from a start codon upstream from the previously reported start codon. The activity of the extended species was similar to the previously reported shorter form, both in the glutathione binding assay and in the reduction of the ribonucleotide reductase according to Björnberg and Holmgren (1991). Sandberg *et al.* (1991), however, report that the elongated form of the *E. coli* glutaredoxin is more active than the shorter form and possibly associated with membrane.

T4 glutaredoxin was first described in cells of *E. coli* infected by the bacteriophage T4 (Berglund, 1969). Moreover, the protein was first crystallized using material isolated from phage-infected cells. The structural studies on the mutant glutaredoxins have been limited to the one described here, the glutaredoxin V15G;Y16P, as all mutant glutaredoxins have failed to crystallize with the exception of one single preparation of the V15G;Y16P mutant glutaredoxin.

The crystal packing of wild-type and mutant glutaredoxin V15G;Y16P does not allow for any extra residues after the final Lys-87. The carboxy termini of the proteins are well-defined in difference maps, and there is no electron density which could correspond to the extra four residues. Moreover, in the V15G;Y16P mutant crystal structure, the carboxy terminus forms hydrogen bonds to the amino-terminal part of the last α-helix of another molecule (Eklund *et al.*, 1992). In the crystal structure of wild-type T4 glutaredoxin, the carboxy terminus of molecule A is positioned at 5 Å from residues Val-26 and Arg-22 in molecule B. Also, in molecule B of this structure, the carboxy terminus is close to Lys-40 in molecule A. The failed crystallization experiments with mutant T4 glutaredoxins could in part have been caused by this newly discovered heterogeneity of our glutaredoxin preparations.

LeMaster (1986) reported that the purified T4 glutaredoxin appeared as a single band in SDS gel electrophoresis. The expression system used was almost identical to ours, with the gene coding for glutaredoxin coupled to the thermoinducible phage λ p_L promoter. The host strain was M5219 with the temperature-sensitive repressor C_{I857}. Nevertheless, the difference we report here is small enough to go undetected in standard denaturing gels.

ACKNOWLEDGMENT

We thank Jim Mattson for excellent technical assistance in the purification of mutant proteins.

REFERENCES

Atkins, J. F., Weiss, R. B., & Gesteland, R. F. (1990) *Cell 62*, 413–423.

Barnett, L., Brenner, S., Crick, F. H. C., Schulman, R. G., & Watts-Tobin, R. J. (1967) *Philos. Trans. R. Soc. London, B 252*, 487–560.

Berglund, O. (1969) *J. Biol. Chem. 244*, 6306–6308.

Berglund, O., & Sjöberg, B.-M. (1970) *J. Biol. Chem. 245*, 6030–6035.

Björnberg, O., & Holmgren, A. (1991) *Protein Expression Purif. 2*, 287–295.

Crick, F. H. C., & Brenner, S. (1967) *J. Mol. Biol. 26*, 361–363.

Eklund, H., Ingelman, M., Söderberg, B.-O., Uhlin, T., Nordlund, P., Nikkola, M., Sonnerstam, U., Joelson, T., & Petratos, K. (1992) *J. Mol. Biol. 288*, 596–618.

Gottesman, M. E., Adhya, S., & Das, A. (1980) *J. Mol. Biol. 140*, 57–75.

Heaphy, S., Singh, M., & Gait, M. J. (1987) *Biochemistry 26*, 1688–1696.

Joelson, T., Sjöberg, B.-M., & Eklund, H. (1990) *J. Biol. Chem. 265*, 3183–3188.

LeMaster, D. (1986) *J. Virol. 59*, 759–760.

Nikkola, M., Gleason, F. K., Saarinen, M., Joelson, T., Björnberg, O., & Eklund, H. (1991) *J. Biol. Chem. 266*, 16105–16112.

Nikkola, M., Gleason, F. K., & Eklund, H. (1993) *J. Biol. Chem. 268*, 3845–3849.

Sandberg, V. A., Kren, B., Fuchs, J. A., & Woodward, C. (1991) *Biochemistry 30*, 5475–5484.

Sjöberg, B.-M. (1972) *J. Biol. Chem. 247*, 8058–8062.

Sjöberg, B.-M., & Holmgren, A. (1972) *J. Biol. Chem. 247*, 8063–8068.

Sjöberg, B.-M., & Söderberg, B.-O. (1976) *J. Mol. Biol. 100*, 415–419.

Söderberg, B.-O., Sjöberg, B.-M., Sonnerstam, U., & Brändén, C.-I. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 5827–5830.

Sundqvist, B., & MacFarlane, R. D. (1985) *Mass. Spectrom. Rev. 4*, 421–460.