The Formation of Diselenide Bridges in Proteins by Incorporation of Selenocysteine Residues: Biosynthesis and Characterization of $(Se)_2$-Thioredoxin†

Sabine Müller,‡ Hans Senn,§ Bernard Gsell,§ Walter Vetter,§ Christian Baron,‡ and August Böck*,‡

Lehrstuhl für Mikrobiologie der Universität München, Maria-Ward-Strasse 1a, D-80638 München, Germany, and Department of Pharmaceutical Research, New Technologies, F. Hoffmann-La Roche Ltd., CH-4002 Basel, Switzerland

Received September 14, 1993®

ABSTRACT: A system was devised which allows the efficient substitution of cysteine residues in a protein by selenocysteine. It involves overexpression of the respective gene with the aid of the T7 promotor/polymerase system in a cysteine auxotrophic strain. The induction of the T7 polymerase formation was performed in cysteine-supplemented medium followed by wash-out of the cysteine and production of the desired gene product in the presence of selenocysteine. The system was applied to substitute the two cysteine residues in *Escherichia coli* thioredoxin. Analysis of the purified gene product by electrospray mass spectrometry and HPLC revealed that both cysteine residues were replaced in approximately 75–80% of the protein, only one cysteine residue was substituted in about 5–10%, and no substitution had taken place in 12–17% of the protein. The occurrence of diselenide, seleno–sulfur, and disulfide bridges in the purified gene product was revealed by ES/MS and chemical modification studies. The diselenide bridge represents an entity in protein structures which has hitherto not been described. The redox property of the selenocysteine variant of thioredoxin [$(Se)_2$-thioredoxin] was found to be substantially different from that of thioredoxin. Only the latter could be reduced under native conditions in the presence of an excess of β-mercaptoethanol. The oxidized $(Se)_2$-thioredoxin was then separated from the selectively reduced and carboxymethylated protein by anion-exchange chromatography. The purity of the isolated $(Se)_2$-thioredoxin was at least 92%.

The replacement of methionine by selenomethionine in a protein has been reported for a long time (Cowie & Cohen, 1957; Huber & Criddle, 1967; Frank et al., 1985), and it is presently used in biophysical research for the determination of protein structures in the crystalline state by X-ray multiwavelength anomalous diffraction methods (Hendrickson et al., 1990; Hendrickson, 1991). This allows the potential solution of the protein three-dimensional structure using only one crystal. Selenomethionine-containing proteins have also found application for structural and enzyme mechanistic studies in solution by high-field NMR spectroscopy (Mullen et al., 1986; Gettins & Wardlaw, 1991; Gettins & Crew, 1991; House et al., 1992, 1993).

The system employed for substitution of methionine by its selenium analogue involves the growth of a methionine auxotrophic strain of an organism in the presence of selenomethionine under the condition of overexpression of the desired protein from a cloned gene. Selenomethionine seems to be well tolerated by the organism. No significant discrimination between the sulfur and the selenium variant was experienced in these experiments, and the gene product appeared to fully maintain its function.

Selenomethionine was the predominant (or the only) selenated amino acid in a protein when the organism was grown in the presence of selenite in the medium. Hartmanis and Stadtman (1982) and Sliwkowski and Stadtman (1985) cultivated *Clostridium kluyveri* in the presence of low selenite concentrations (0.5–1 μM) and demonstrated the efficient replacement of methionine residues in the enzymes thiolase and β-hydroxybutyryl-CoA dehydrogenase. Although these enzymes contain a rather high proportion of cysteine residues, no substitution by selenocysteine was detected. A possible reason for preferential incorporation of selenomethionine could be that free selenocysteine is not a substrate for the cysteinyl-tRNA ligase from this organism but is readily converted into selenocystathionine and thereby channeled into the methionine pathway. Significant replacement of cysteine by selenocysteine was, however, found by Kramer and Ames (1988) in a *selD* mutant of *Salmonella typhimurium* (previously designated *selA1*) and was later also shown to occur in an *Escherichia coli selD* strain (Stadtman et al., 1989). The product of *selD* has a function in the pathway for the specific insertion of selenocysteine into proteins; it catalyzes the formation of an activated selenium compound (Ehrenreich et al., 1992) identified as phosphoroselenoate (Veres et al., 1992). The lesion in *selD* may result in the buildup of a high intracellular concentration of selenide, which again could result in an increased formation of selenocysteine and its charging to $tRNA^{Cys}$ by cysteinyl-tRNA ligase. The only artificial selenoprotein containing a selenocysteine residue which has been prepared and characterized so far is selenosubtilisin (Wu & Hilbert, 1989; Syed at al., 1993). However, it was obtained from subtilisin by chemical modification of a serine residue, the side-chain hydroxyl group of which was substituted by hydrogen selenide.

In this paper we report on the development of a system which allows the efficient biosynthetic substitution of cysteine residues by selenocysteine. As a model protein for selenocysteine incorporation, we have chosen thioredoxin from *E. coli*. The reason was that the 3D structure of this protein is known with high resolution in the crystal and in solution (Holmgren et al., 1975; Dyson et al., 1988, 1990; Katti et al.,

† This work was supported by grants from the Bundesministerium für Forschung und Technologie (Genzentrum München) and the Fonds der Chemischen Industrie.

* Corresponding author.
‡ Universität München.
§ F. Hoffmann-La Roche Ltd.

® Abstract published in *Advance ACS Abstracts*, February 15, 1994.

0006-2960/94/0433-3404$04.50/0

Table 1: Strains and Plasmids

strain or plasmid	genotype	ref
strains		
MC4100	F⁻ *araD139* Δ(*argF-lac*) U169 *ptsF25 deoC1 relA1 flbB5301 rpsL 150* λ–	Casabadan & Cohen, 1979
JM39/2	F⁺ *cysE51*	Denk, 1987
BL21 (DE3)	F⁻ *hsdS gal* λ DE3	Studier & Moffatt, 1986
BL21 *cys*	*BL21 (DE3) selB::kan cysE51*	this work
plasmids		
pUC118-*trxA*	Ap^r *trxA*	Krause & Holmgren, 1991
pT7-5	Ap^r T7 *ϕ10*	Tabor & Richardson, 1985
pSM1	pT7-5-*trxA*	this work

1990). This will allow a detailed structural comparison of thioredoxin with its selenocysteine-containing variant [designated further on as $(Se)_2$-thioredoxin]. *E. coli* thioredoxin contains two cysteine residues, Cys32 and Cys35, which are engaged in a disulfide bridge in the oxidized protein. Replacement of the two active site cysteine residues which undergo disulfide bridging during catalysis can provide interesting functional information on the reduction properties of the species containing selenocysteine. Finally, the folding pathway of many proteins includes folding intermediates with different disulfide bonding patterns (Creighton, 1988). Substitution of selected cysteine residues by selenocysteine may identify critical steps of the folding event.

EXPERIMENTAL PROCEDURES

Materials. Pure thioredoxin from *E. coli* and rifampicin were obtained from Serva Feinbiochemica (Heidelberg, Germany); dithiothreitol was from Bio-Rad Laboratories (Hercules, CA, USA); $NaBH_4$, β-mercaptoethanol, and ultrapure guanidine hydrochloride were from Fluka AG (Buchs, Switzerland); DL-selenocystine and ultrapure iodoacetic acid, sealed under argon, were from Sigma (St. Louis, MO, USA); acetonitrile was from Biosolve LTD (Amsterdam); trifluoroacetic acid was from Baker Inc. (Phillipsburg, NJ, USA); and HPLC-grade water was from BDH (Poole, England).

Strains, Plasmids and Growth Conditions. The strains of *E. coli* and the plasmids used in this work are listed in Table 1. The medium employed in all cultivations except in the overexpression experiments was a modification of M9 minimal medium (Miller, 1972). It contained either $MgSO_4$ (2 mM) or L-cysteine at the indicated concentration as sulfur source. It was supplemented with 1% glycerol (w/v) as carbon source, 100 μg/mL of 18 amino acids (without cysteine or methionine), 5 μg/mL thiamine, and 1 mL of a trace element solution containing 400 μM H_3BO_3, 30 μM $CoCl_2$, 1 μM $CuSO_4$, 80 μM $MnCl_2$, 10 μM $ZnSO_4$, and 10 μM $FeCl_3$. In the expression experiments a modification of the minimal medium described by Senn et al. (1987) was employed. Glucose as carbon source was replaced by glycerol, and $FeSO_4$ and $MgSO_4$ were replaced by equimolar amounts of $FeCl_2$ and magnesium acetate. The medium also was fortified with 100 μg/mL of L-leucine.

The strains were grown aerobically in gyrotatory water bath shakers in Erlenmeyer flasks filled to one-tenth of their volume with medium. Larger volumes of cultures were obtained upon growth in a Biostat M fermenter (Braun Diessel Biotechnology, Melsungen, Germany) with a working volume of 1.2 L. The agitation rate was 700–800 rpm, and the aeration rate was 1 volume of air per volume of medium per minute. The temperature was 37 °C throughout.

Recombinant DNA Techniques. Standard recombinant DNA methods were carried out with established procedures (Maniatis et al., 1982). *E. coli* strain BL 21 was transformed either with the RbCl method (Hanahan, 1985) or by electroporation. Phage P1-mediated transductions were performed according to Miller (1972).

Isolation of Bulk tRNA and Aminoacylation with L-Cysteine. Bulk tRNA was isolated from *E. coli* MC4100 by phenol extraction and fractionated by ethanol and 2-propanol precipitation (Zubay, 1962). Cysteinyl-tRNA synthetase assays were conducted in 250-μL reaction mixtures containing 50 mM Tris buffer, pH 7.4, 10 mM $MgCl_2$, 30 mM NH_4Cl, 2 mM ATP, 2 mM dithiothreitol (DTT), 1 pmol of purified cysteinyl-tRNA synthetase (kindly donated by Dr. Y. Hou), and varying concentrations of L-[^{35}S]cysteine (20 mCi/mmol). All solutions had been made oxygen-free by consecutive degassing and gassing with N_2 before use. Reactions were started by the addition of 0.5 μM tRNA and allowed to take place at room temperature in the anaerobic hood (Coy Laboratory Products Inc., Ann Arbor, MI) under an atmosphere of 95% N_2 and 5% H_2. Samples (45 μL) were taken at 0 time and after 1, 2, 3, and 5 min and transferred onto Whatman 3 MM filter disks, which were immediately immersed into ice-cold 10% trichloroacetic acid. They were washed in 5% trichloroacetic acid containing 5 mg of L-cysteine/mL and 80% ethanol (–20 °C) and dried, and their radioactivity was determined by liquid scintillation counting. When DL-selenocysteine was to be added, it was reduced in the presence of 18 mM KBH_4 for 3 h prior to addition.

Design of the trxA Expression System. A *cysE* derivative of *E. coli* strain BL21 was constructed by introducing the *cysE51* allele from strain JM39/2 via phage P1 transduction. To facilitate transduction, a kanamycin resistance determinant located in the closely linked *selB* gene was first introduced into strain JM39/2. The resulting strain was then used as donor for transducing *cysE* together with *selB::kan* into BL21.

The gene coding for thioredoxin from *E. coli* (*trxA*) was set under the control of the T7 promoter of plasmid pT7-5. To this end, the 0.48-kb *Hin*dIII–*Eco*RI fragment carrying *trxA* was isolated from plasmid pUC118-*trxA* and ligated into the multicloning site of plasmid pT7-5, yielding plasmid pSM1. pSM1 was transformed into strain BL21*cysE51*.

For overexpressing *trxA* in the presence of selenocystine an overnight culture of BL21*cysE51*/pSM1 grown in LB medium (Miller, 1972) was used to inoculate 1.2 L of M9 expression medium containing ampicillin, kanamycin, and cysteine hydrochloride (50 μg/mL) to an OD_{600} of 0.1. When the culture had reached an optical density of about 1 (after about 4.5 h), IPTG was added to a concentration of 1 mM and after a further 10 min chloramphenicol to a concentration of 10 μg/mL. Five minutes after the addition of the antibiotic the culture was transferred into prechilled centrifuge beakers (GS3, DuPont) and the cells were sedimented at 6000 rpm for 5 min. They were washed twice in cold saline (1.2 L each) and resuspended in the production medium (Senn et al., 1987)

containing 400 μg/mL rifampicin and 600 μM DL-selenocystine. Incubation was continued for further 2.5 h. The cells were then harvested (10 min at 6000 rpm) and washed twice in an equal volume of buffer T (50 mM Tris buffer, pH 7.5, containing 1 mM EDTA).

Purification of $(Se)_2$-Thioredoxin. The purification of the thioredoxin variants was adopted from a procedure published by Krause and Holmgren (1991). In short, the cells were suspended in buffer T (1 mL/g cell wet weight) containing 1 mM phenylmethanesulfonyl fluoride and broken by passage through a French press cell at 16 000 psi. Cell debris was removed by centrifugation at 30000*g* for 30 min. The resulting S30 extract was then further clarified by centrifugation at 120000*g* for 2 h. The supernatant was mixed with 0.1 volume of a 7% streptomycin sulfate solution, kept at 0 °C for 30 min, and freed from the precipitated nucleic acids by centrifugation for 30 min at 30000*g*. The supernatant was then dialyzed twice against 2 L of buffer T.

The dialysate was applied to a DEAE-Sepharose ion-exchange column (30-mL column volume) which had been equilibrated in buffer T. The column was washed with 30 mL of buffer T and developed with a linear gradient (500 mL) reaching from 0 to 300 mM NaCl in the same buffer, at a flow rate of 30 mL/h. The elution of the thioredoxin species was monitored by SDS–polyacrylamide gel electrophoresis (Laemmli, 1970) and staining of the gels with silver nitrate. Fractions containing the protein were pooled and brought to 85% saturation of ammonium sulfate. The precipitate was collected by centrifugation and dissolved in 5 mL of buffer T. The protein solution was then applied to a Sephadex G-50 column (300-mL volume) equilibrated with buffer T. The column was developed with the same buffer at a flow rate of 20 mL/h. Thioredoxin-containing fractions were pooled and lyophilized.

The protein concentration of crude extracts was determined according to Whitaker and Granum (1980). The concentration of the purified protein was determined by calculating the molar extinction coefficient at 280 nm from its known molecular weight and content of aromatic amino acids.

Structural Characterization of Seleno-Labeled Thioredoxins. The identification and structural characterization of the isolated proteins and their reduced and carboxymethylated forms were carried out using isoelectric focusing gel electrophoresis (IEF), reverse-phase HPLC, and electrospray mass spectrometry (ES/MS).

Lyophilized and salt-free samples were directly submitted to ES/MS analysis using a PE/SCIEX API III mass spectrometer (Whitehouse et al., 1985; Fenn et al., 1990). Protein samples (5 μL) at concentrations ranging from 0.2 to 2.0 μg in 0.1 N aqueous acetic acid containing 50% acetonitrile were injected into a flow of acetonitrile/water (1/1) and sprayed from a nozzle at 5 kV with pneumatic assistance into the ion source at a flow rate of 30 μL/min (Bruins et al., 1987). In a typical run the spectra were scanned from m/z 700 to m/z 1700 at 2.5 s/scan in steps of 0.2 Da. All scans were averaged. The accuracy of the m/z measurements was better than ±0.2 Da in the averaged scans. Mass-scale calibration was performed by means of ions from polypropylene glycol. All molecular masses shown are reconstructed from the differently charged ions of the m/z spectra, using the PE/SCIEX program.

HPLC was performed using a Hewlett-Packard (HP) 1050TI system, equipped with a diode array UV/vis detector (1040 II Serie) monitoring at wavelengths of 218 and 280 nm. The analytical columns C4-Dynamax 300A 4.6 × 300 mm (Rainin) and Hytach peptide column 4.6 × 100 mm (HP) were used. The columns were thermostated in a column heating chamber at 35 and 40 °C, respectively. Samples were injected on the column by the HP autoinjector 1050 equipped with a 100-μL loop. The HP Chemstation was used for signal integration and data handling. All the eluents were filtered (0.2 μm) and degassed on line with a HP1050 vacuum degassing system, and the flow rate was set to 0.5 mL (Hytach column) and 1.0 mL (Dynamax column). The mobile phase was 0.1% TFA in water (A) and 0.75% TFA in acetonitrile (B). The following time-dependent linear gradients were applied: Dynamax column (35 °C): from 0 to 6 min, 0% B; at 10 min, 5% B; at 25 min, 50% B; at 35 min, 60% B. Hytach column (40 °C): from 0 to 5 min, 0% B; at 25 min, 50% B; at 35 min, 70% B.

Isoelectric focusing gel electrophoresis (IEF) was performed on agarose IEF gels, pH range 3–10 (Iso-gel from FMC, Bioproducts, Rockland, ME, USA) using the multiphore II horizontal electrophoresis system (Pharmacia) in the constant-power mode (25 W) (Wu et al., 1989). For the determination of the isoelectric point (p*I*), a set of calibration proteins with p*I* values ranging from 4.75 to 10.60 (from BDH, Poole, England) were used. The gels were stained with Coomassie blue R250 and destained in a mixed solvent (water/methanol/acetic acid, 6/3/1).

Chemical Modifications. The single disulfide bridge of *E. coli* thioredoxin (obtained from Serva Feinbiochemica) was reduced under native conditions with β-mercaptoethanol (MSH) and alkylated with iodoacetic acid at pH 8.3 to form reduced monocarboxymethylated thioredoxin (Kallis & Holmgren, 1980). The dicarboxymethylated form of the protein was obtained when the alkylation reaction was carried out in the presence of 5 M guanidine hydrochloride. Under these conditions the sulfhydryl groups from Cys32 and Cys35 are both alkylated (Kallis & Holmgren, 1980). Alternatively, the dialkylated form of thioredoxin and $(Se)_2$-thioredoxin was also obtained by reduction of the protein with an excess of dithiothreitol (DDT) or sodium borohydride followed by alkylation under native and denaturing conditions (Kelley et al., 1987). The protein was separated from excess reagent and desalted either by gel filtration (Sephadex G-25) or by dialysis. The alkylated thioredoxin preparations were checked by HPLC, IEF, and ES/MS for identity and purity.

The purified protein isolate expressed from recombinant *E. coli* grown on a medium containing selenocystine was reduced and alkylated according to two different protocols:

(1) One milligram of protein was dissolved in 1 mL of 0.1 M acetic acid containing 45% acetonitrile (solution pH 3.3). The following reagents were added at the time points indicated: 20 μL of 0.4% aqueous β-mercaptoethanol (0 min); 20 μL of 1 M iodoacetic acid (30 min), reaction pH 4.4; 30 μL of 3 M Tris buffer, pH 12, and 1.35 mL of water (32 min), reaction pH 6.1; 100 μL of 3 M Tris buffer, pH 8.3 (34 min), reaction pH 8.1. The alkylation reaction was allowed to proceed for one additional hour and was stopped by adding 5 μL of neat β-mercaptoethanol.

(2) One milligram of protein was dissolved in 2.35 mL of water; 20 μL of 1M iodoacetic acid was added together with 100 μL of 3 M Tris buffer, pH 8.3 (reaction pH 8.0); after 5 min 5 μL of neat β-mercaptoethanol or a corresponding molar equivalent of dithiothreitol (DTT) was added, and after an additional 15 min the protein was separated from the reagents by two dialysis steps (50 mM NH_4HCO_3, pH 7.8, and water) or by gel filtration (Sephadex G-25).

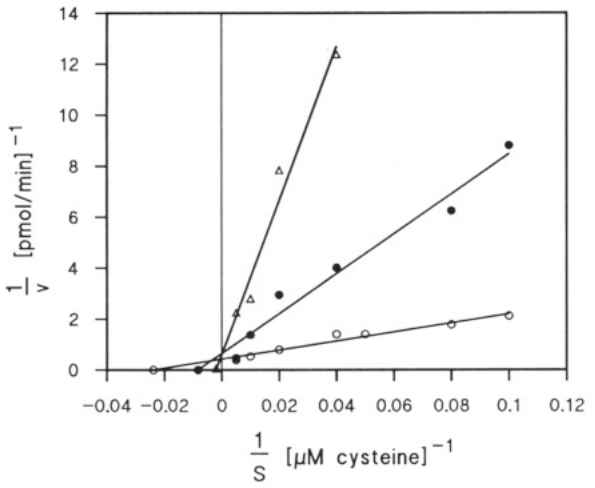

FIGURE 1: Lineweaver–Burk plot of initial velocities of the cysteinyl-tRNA ligase reaction in the absence (○) and in the presence of 100 μM (●) and 400 μM (△) L-selenocysteine.

All the reaction steps were carried out at room temperature in the dark and under argon atmosphere.

The reduced and alkylated protein samples were analyzed by HPLC, ES/MS, and IEF gel electrophoresis.

Separation of $(Se)_2$-Thioredoxin from Thioredoxin. Oxidized $(Se)_2$-thioredoxin was separated from thioredoxin by selective reduction of the thioredoxin with β-mercaptoethanol and alkylation with iodoacetate followed by anion-exchange chromatography on DEAE-Sepharose. In a typical experiment, 24 mg of protein of the thioredoxin/$(Se)_2$-thioredoxin mixture was subjected to selective alkylation according to method 2 (see above). The reaction mixture was then dialyzed twice against 2 L each of buffer T and then applied to a DEAE-Sepharose column (20-mL volume) equilibrated in buffer T. The column was first washed with 35 mL of buffer T and then developed with 200 mL of a NaCl gradient reaching from 0 to 100 mM. Samples of each fraction collected were analyzed by electrophoresis in 15% polyacrylamide gels under nondenaturing conditions followed by staining with silver nitrate.

Molecular Modeling. The molecular modeling of $(Se)_2$-thioredoxin was done with the Roche in-house program MOLOC (Müller et al., 1988). The two cysteine residues in the X-ray structure of thioredoxin (Katti et al., 1990) were replaced by selenocysteine. The resulting structure was energy minimized to optimize the contacts. The force field parameters for the disulfide and the carbon–sulfur bond were reparametrized to reflect the increased bond length and van der Waals radius of selenium as compared to that of sulfur (Emsley, 1992).

RESULTS

Charging of $tRNA^{Cys}$ with Selenocysteine. An important condition for the replacement of cysteine residues by selenocysteine is that the cysteine-specific tRNA is aminoacylated with selenocysteine by cysteinyl-tRNA ligase. Since labeled selenocysteine was not available, it was determined whether the analogue is able to compete with cysteine in the aminoacylation reaction. Figure 1 gives the results of the kinetic analysis of the cysteinyl-tRNA ligase reaction in the absence and presence of selenocysteine. The inhibition was strictly competitive; a K_i value (45 μM) was obtained, being in the same range of magnitude as the K_m for the cognate substrate L-cysteine (42 μM). The results provide evidence that under

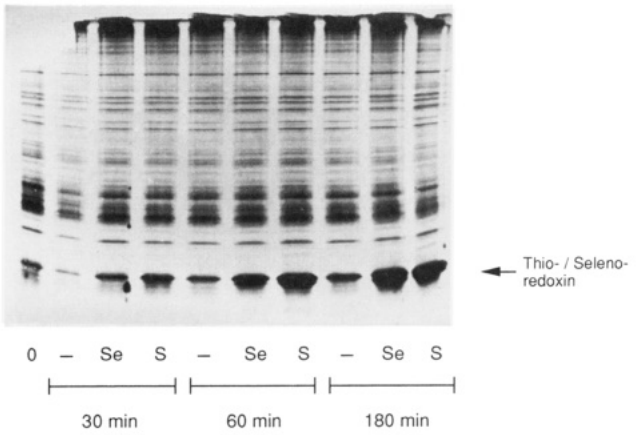

FIGURE 2: Thioredoxin/$(Se)_2$-thioredoxin formation by strain BL21*cysE51*/pSM1. Cells were induced for the formation of T7 RNA polymerase by the addition of 1 mM IPTG for 10 min and then deprived of cysteine. Incubation was continued for the indicated times either in the absence of any supplement (–), in the presence of 600 μM DL-selenocystine (Se), or in the presence of 400 μM cystine (S). Lane 0 is from cells harvested before induction. An SDS gel electropherogram stained with silver nitrate is shown.

in vivo conditions there may be efficient charging of $tRNA^{Cys}$ with selenocysteine provided that the cells are deprived of L-cysteine.

Design and Optimization of the Expression System. Since selenocysteine is a compound highly toxic for *E. coli* (our unpublished results), an expression system had to be chosen which does not rely on growth of the host strain. The system described by Studier and Moffat (1986) was employed, therefore, which involves a strain, BL21, with a chromosomally encoded gene for the phage T7 RNA polymerase and a plasmid carrying the gene to be expressed under the control of a phage T7 promoter. A cysteine auxotrophic derivative of BL21 was constructed to facilitate cysteine deprivation. It was transformed with plasmid pSM1 which harbors the *trxA* gene coding for the *E. coli* thioredoxin under the control of the phage T7 promoter.

The rationale for the efficient incorporation of selenocysteine into thioredoxin was to grow strain BL21*cysE51* in the presence of cysteine and then to induce the formation of the T7 RNA polymerase, followed by washing out the cysteine and supplementing the medium with selenocysteine. A number of physiological parameters and medium conditions had to be optimized to reach efficient formation of thioredoxin containing selenocysteine. These were in particular the following: (1) the supplementation of the medium with L-cysteine hydrochloride at 50 μg/mL which was just sufficient to allow growth to the density of the culture desired for induction of expression but was low enough to achieve easy wash-out of the rest not consumed by the bacteria; (2) induction of expression of the T7 RNA polymerase gene by 1 mM IPTG when the culture reached an OD_{600} of 1 followed by a 10-min expression phase; (3) addition of chloramphenicol at 10 μg/mL, chilling the culture, washing the cells extensively with saline, and transferring them into the production medium containing 600 μM DL-selenocystine.

Figure 2 shows the effect of some of these parameters and of the length of the production phase on the formation of thioredoxin/$(Se)_2$-thioredoxin. Without induction only a small amount of gene product was formed. Addition of selenocysteine stimulated the synthesis of gene product albeit at a rate significantly lower than that taking place in the presence of cysteine. A production time of 2.5 h was chosen in all further experiments. Gene product made in the absence

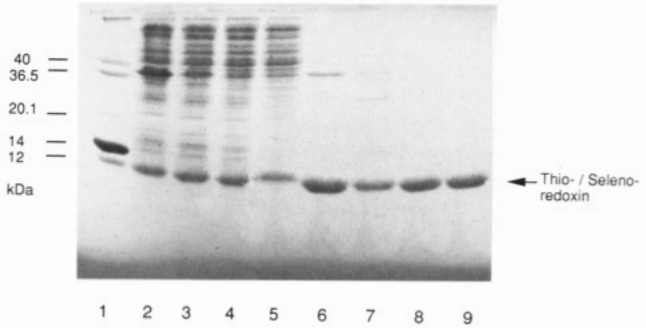

FIGURE 3: Purification of thioredoxin/$(Se)_2$-thioredoxin as followed by SDS gel electrophoresis and staining with silver nitrate. Lanes: 1, size standard; 2, broken cell; 3, S30 extract; 4, S100 extract; 5, supernatant of streptomycin precipitation; 6, DEAE pool I (side fractions); 7, DEAE pool II (main fractions); 8, Sephadex pool I; and 9, Sephadex pool II.

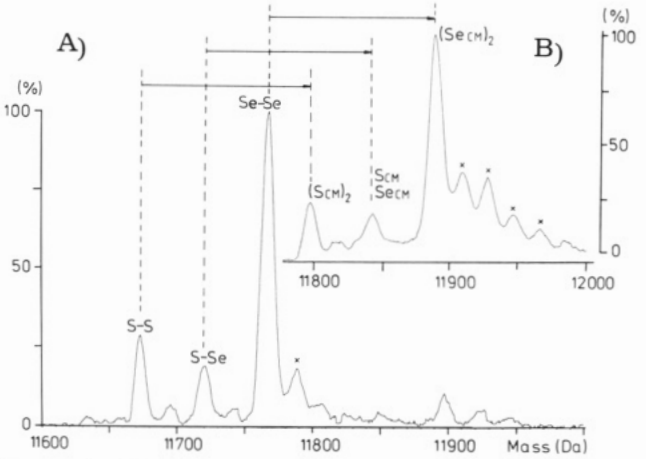

FIGURE 4: Electrospray mass spectrometric analysis of (A) the purified protein isolate expressed from recombinant *E. coli* grown on a medium containing selenocystine and (B) its reduced (denaturing conditions, $NaBH_4$) and carboxymethylated (CM) form. The expected mass shift for the incorporation of two carboxymethyl groups would be 118 Da. The measured masses (±2 Da) of the three main peaks in (A) are 11 673.8, 11 721.4, and 11 767.8. These peaks correspond to the theoretical molecular masses of the oxidized form of thioredoxin (S-S) (11 674.4), of the mixed (SeS)-thioredoxin (11 721.3), and of $(Se)_2$-thioredoxin (11 768.2), respectively. The corresponding peaks in the ES/MS spectrum of (B), $(SeCM)_2$, (SCM, SeCM), and $(SeCM)_2$, are all shifted by 119 ± 1 Da to higher masses as is expected for the dicarboxymethylated form of these proteins. The peaks labeled with "×" represent the mono and multiple sodium salt forms ($n$$Na^+$) of the main component.

of cysteine or selenocysteine possibly is due to the supply of cysteine originating from protein turnover.

Purification of $(Se)_2$-Thioredoxin. For purification of $(Se)_2$-thioredoxin cells from a 1.2-L fermenter culture were used as starting material; an established procedure consisting of ultracentrifugation, streptomycin sulfate precipitation, DEAE-Sepharose chromatography, and Sephadex gel filtration was followed (Krause & Holmgren, 1991). The path of purification is displayed by the SDS gel of Figure 3. An essentially homogeneous preparation was obtained. The yield was 14.4 mg of gene product from 105 mg of protein of the crude cell-free extract.

Structural Characterization of the Purified Protein Isolate. The seleno-labeled thioredoxin isolate was characterized by electrospray mass spectrometry (ES/MS), HPLC, and isoelectric focusing gel electrophoresis (IEF). The ES/MS spectrum of the purified protein sample (Figure 4A) reveals the presence of three differently composed components. The mass value of the main component (11 767.8 Da) completely agrees with that predicted for oxidized $(Se)_2$-thioredoxin in which the two cysteine residues of thioredoxin, Cys32 and Cys35, had been replaced by selenocysteines. The minor protein components were identified as thioredoxin (11 673.8 Da) and the "mixed" (SeS)-thioredoxin (11721.4 Da), respectively. From the relative peak intensities in Figure 4A, it was estimated that the protein sample is composed of approximately 73% $(Se)_2$-thioredoxin, 10% (SeS)-thioredoxin, and 17% thioredoxin.

The molecular structures of the three components were further verified by reduction of the protein isolate with an excess of either β-mercaptoethanol, dithiothreitol, or $NaBH_4$, followed by exhaustive alkylation of the cysteine and selenocysteine residues with iodoacetic acid under denaturing conditions. The dicarboxymethylated form of the three different redoxin components were thus obtained (Figure 4B). From the relative peak intensities in the MS spectrum of Figure 4B, an initial sample composition of approximately 80% $(Se)_2$-thioredoxin, 8% (SeS)-thioredoxin, and 12% thioredoxin was estimated.

A unique property of thioredoxin is its Na^+ binding capacity (Figure 4). A fully Na^+-free population of the protein could not be observed by extensive dialysis or by desalting on a Sephadex G-25 column. Also multiple K^+ and Rb^+ ions were bound by the protein if these salts are added to the thioredoxin solution (ES/MS results, not shown). This affinity of thioredoxin for alkali ions has not been reported before and is surprising.

The thioredoxin component of the protein isolate could further be separated from $(Se)_2$-thioredoxin by reverse-phase HPLC (Figure 5A). However, the minor (SeS)-thioredoxin component was not resolved in the chromatogram. It probably has the same retention time as $(Se)_2$-thioredoxin (see below). When the protein isolate from two independent fermentations was analyzed, practically identical HPLC and ES/MS results were obtained. A thioredoxin content of 15–16% was determined by HPLC for both isolates. This shows that the biochemical method described for the incorporation of selenocysteines into proteins works very reproducibly.

In the IEF gel electrophoresis experiment, the three different components of the protein sample did not separate and formed one single band with a pI value of 4.7 (Figure 6A). The isoelectric point of the reduced and dicarboxymethylated sample, however, is more acidic (pI = 4.3) due to the additional negative charges carried by the carboxymethyl groups (Figure 6C).

Selective Reduction of Thioredoxin and Separation from Oxidized $(Se)_2$-Thioredoxin. An intriguing observation was made when the time course of the reduction of the protein isolate with β-mercaptoethanol was followed by ES/MS under native conditions (Figure 7). A selective reduction of the thioredoxin component occurred within a few minutes after the addition of the reducing agent as is demonstrated by the appearance of a new molecular component in the spectrum at 11 751.8 Da. The mass value of this component corresponds exactly with that expected for the mixed disulfide adduct of reduced thioredoxin with one molecule of β-mercaptoethanol.

The $(Se)_2$-thioredoxin component was not reduced under these experimental conditions (Figure 7), indicating that its redox potential is substantially lower than that of thioredoxin. The third component of the protein sample, the mixed (SeS)-thioredoxin, was only slightly affected by β-mercaptoethanol. This is concluded from the weak intensity increase observed at the mass position of the corresponding protein–β-mercaptoethanol adduct (11 797.8 Da). This peak did not further

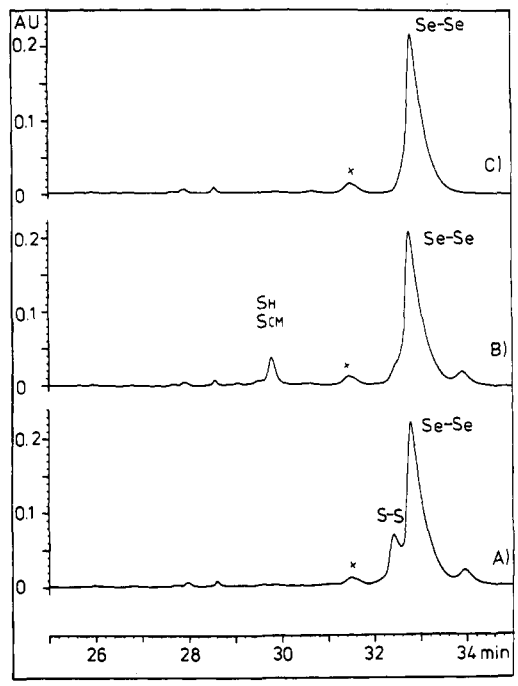

FIGURE 5: HPLC analysis of (A) the purified protein isolate expressed from recombinant *E. coli* grown on a medium containing selenocystine; (B) the same protein isolate, but with the thioredoxin component (S-S) selectively reduced and monocarboxymethylated (SH, SCM); (C) the $(Se)_2$-thioredoxin (Se-Se) component that was isolated from (A) by anion-exchange (DEAE) chromatography (see text). The HPLC separation was on a C4-Dynamax 300A, 4.6 × 300 mm column at 35 °C; the elution system consisted of 0.1% TFA in water (solvent A) and 0.75% TFA in acetonitrile (solvent B). The proteins were eluted with a gradient optimized for the separation of the (S-S) and (Se-Se) components (see Experimental Procedures). The retention times of pure oxidized thioredoxin (Serva) and its reduced and monocarboxymethylated form (HPLC not shown) are identical with the (S-S) and (SH, SCM) component of traces A and B, respectively. The minor peak marked with "×" also occurs in the HPLC of pure thioredoxin.

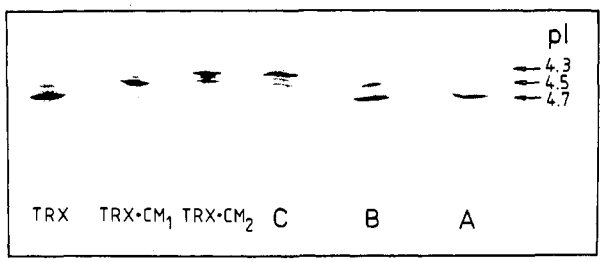

FIGURE 6: Isoelectric focusing gel electrophoresis (IEF) of (A) the purified protein isolate composed of $(Se)_2$-thioredoxin (Se-Se), (SeS)-thioredoxin (Se-S), and thioredoxin (S-S); (B) the selectively reduced and monocarboxymethylated form of (A), where only the (S-S) component is reduced; (C) the fully reduced and dicarboxymethylated form of (A). For comparison, the IEF gel of pure oxidized thioredoxin (TRX) and its reduced monocarboxymethylated ($TRX{\cdot}CM_1$) and dicarboxymethylated forms ($TRX{\cdot}CM_2$) are also shown. The $TRX{\cdot}CM_2$ preparation contains also some monoalkylated form $TRX{\cdot}CM_1$ as was also observed in the ES/MS spectrum (not shown). The isoelectric points (p*I*) measured for the three main IEF bands are indicated on the right.

increase after 60-min reaction time, indicating that only part (approximately 50%) of the mixed seleno–sulfur protein could be reduced.

The selectively reduced protein sample was then monocarboxymethylated with iodoacetic acid under weakly alkaline conditions (Kallis & Holmgren, 1980). The reduction and alkylation conditions had been optimized in order to yield fully monoalkylated thioredoxin and, possibly, also (SeS)-

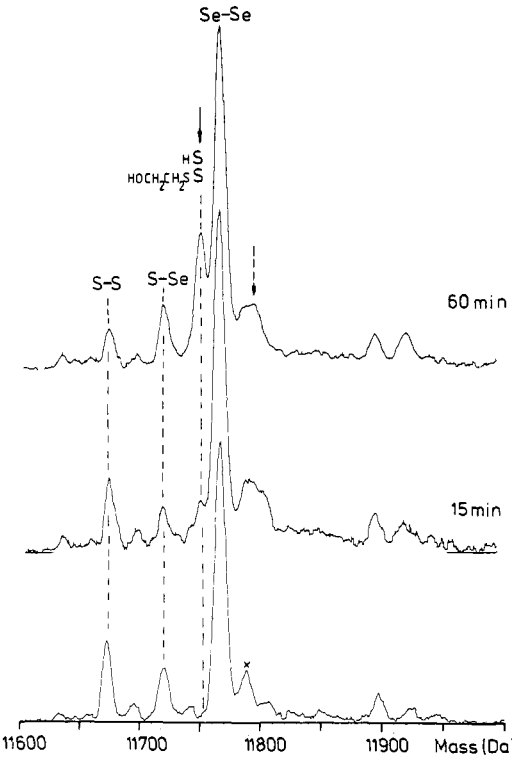

FIGURE 7: Electrospray mass spectrometric analysis of the reduction of the purified protein isolate (mixture of S-S, S-Se, and Se-Se; see Figure 5A) with β-mercaptoethanol (MSH). The protein was dissolved in 0.1 N aqueous acetic acid containing 50% acetonitrile, and the ES/MS spectra were measured before and after adding 1% β-mercaptoethanol. Only the thioredoxin component (S-S) was substantially reduced under these conditions. The new peak (arrow) which appears in the mass spectrum at 11 751.8 ± 2 Da (at t = 15 and 60 min) corresponds to the mixed disulfide adduct (SH, $SSCH_2CH_2OH$) that is formed between one of the two cysteine residues of thioredoxin (S-S) and MSH (theoretical mass 11 751.5). The beginning accumulation of a corresponding adduct between (SeS)-thioredoxin and β-mercaptoethanol can be inferred from the (unresolved) peak at 11 797.8 Da (dotted arrow). The peak labeled with "×" represents the sodium salt form of $(Se)_2$-thioredoxin.

thioredoxin without affecting oxidized $(Se)_2$-thioredoxin (see Experimental Procedures). Aliquots of the reaction mixture were analyzed by HPLC, IEF, and ES/MS. The HPLC result indicates the disappearance of the thioredoxin component and the formation of the monoalkylated species eluting at lower retention times (Figure 5B). In the IEF gel, the monoalkylated thioredoxin component appeared as a separated band at pI = 4.5, above the band of oxidized $(Se)_2$-thioredoxin (Figure 6, lane B). The (SeS)-thioredoxin component was not separated by HPLC, but, as inferred from analysis of the reaction mixture by ES/MS, had greatly diminished (<5%). This indicates that also a partial reduction and alkylation of (SeS)-thioredoxin by iodoacetic acid had occurred under the experimental conditions applied.

The oxidized $(Se)_2$-thioredoxin could be separated from the monocarboxymethylated minor components of the reaction mixture by anion-exchange chromatography (DEAE-Sepharose Fast Flow) (Figure 8A). The major peak contains oxidized $(Se)_2$-thioredoxin whereas the follow-up peak harbors the carboxymethylated thioredoxin (see Figure 8B). The isolated $(Se)_2$-thioredoxin exhibited a single symmetric peak in the HPLC with a retention time of 32.8 min (Figure 5C). From the relative peak intensities, the protein was determined to be at least 92% pure. The molecular mass as determined by ES/MS corresponded exactly to the expected mass of oxidized $(Se)_2$-thioredoxin (11 767.8 Da). The remaining (SeS)-

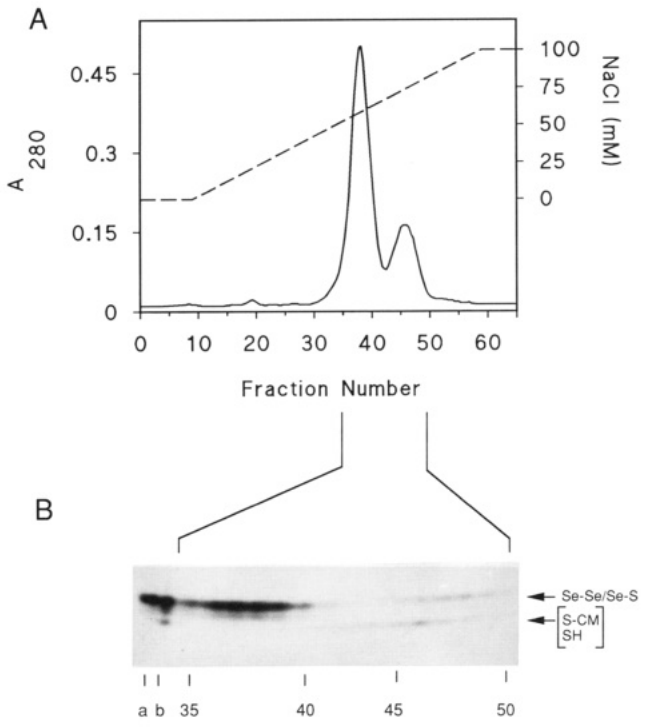

FIGURE 8: Purification of oxidized $(Se)_2$-thioredoxin after selective alkylation of the thioredoxin component. Part A of the figure gives the elution pattern as followed by A_{280} measurements. Part B presents the electropherogram of selected fractions in 15% polyacrylamide gels. Lane a gives the separation pattern of the mixture (3 μg of protein) before alkylation; lane b, that of 2.5 μg of the mixture after alkylation. The following lanes contain 6-μL samples of fractions 35–52 of the column eluate. Total volume of each fraction was 4 mL.

thioredoxin and thioredoxin impurities were estimated from the ES/MS spectrum to be <5% and <1%, respectively.

The purified $(Se)_2$-thioredoxin could finally be reduced under native conditions when the reducing agent in the reduction protocol, β-mercaptoethanol, was replaced by an equimolar amount of dithiothreitol. The ES/MS analysis of the purified reaction product revealed the formation of the reduced and dicarboxymethylated form of $(Se)_2$-thioredoxin as it was previously reported for DTT reduction of thioredoxin under native conditions (Kelley et al., 1987). Also the HPLC and IEF results are in agreement with the ES/MS result.

DISCUSSION

The system described above allows the efficient although not complete replacement of cysteine residues in a protein by selenocysteine. In several experiments substitution ratios between 73% and 80% could be achieved. The residual incorporation of cysteine may result from the replenishment of the free pool of cysteine by protein turnover. The fact that the formation of other cellular proteins is blocked under the overexpression conditions employed (presence of rifampicin) might contribute to the buildup of this pool. We have, however, not conducted any systematic studies on the substitution ratios under conditions where the synthesis of cellular proteins continues.

The ratios of the different kinds of gene products formed is intriguing. On the basis that it is solely the stoichiometry of $tRNA^{Cys}$ molecules charged with selenocysteine relative to those charged with cysteine which determines the ratio of incorporation, one should expect that the amount of the Se-S and S-Se forms is higher than that of the thioredoxin (S-S form). The opposite, however, is the case. Trivial explanations would be that the "mixed" form is less stable or preferentially lost during purification. A purely theoretical possibility, however, is that there is some discrimination at the ribosome against insertion of selenocysteine, which would result in an unproportionately high thioredoxin ratio.

$(Se)_2$-thioredoxin is the only protein known so far with a diselenide bridge . It exhibits significant redox activity differences when compared to native thioredoxin. The intrinsically lower redox potential of the inorganic diselenide/selenol couple compared to that of the disulfide/thiol pair (Emsley, 1992) is maintained if not accentuated by the protein environment. The observation that in the native state of the proteins only thioredoxin but not $(Se)_2$-thioredoxin was reduced by an excess of β-mercaptoethanol implies a substantially lower redox potential for $(Se)_2$-thioredoxin than for thioredoxin. The reduced state of $(Se)_2$-thioredoxin, however, could be generated under native conditions by using the stronger reducing agent DTT. Therefore, the redox potential of $(Se)_2$-thioredoxin lies between the redox potentials of β-mercaptoethanol (–0.22 V) and thioredoxin (–0.23 to –0.26 V) on one side and dithiothreitol (–0.323 V) on the other side (Holmgren, 1968; Berglund & Sjöberg, 1970; Szajewski & Whitesides, 1980).

Under denaturing conditions (5 M guanidine hydrochloride), however, both proteins were equally affected by the reducing agent. The diselenide bridge thus seems to be efficiently stabilized and shielded by the protein environment. Therefore, it must be expected that, under physiological conditions, the redox properties of $(Se)_2$-thioredoxin will differ from those of thioredoxin; e.g., it might not be reduced by the NADPH-dependent thioredoxin reductase. Indeed, preliminary experiments showed that $(Se)_2$-thioredoxin cannot be reduced by thioredoxin reductase from *E. coli* with $NADPH_2$ as electron donator (M. Meyer and J. R. Andreesen, unpublished experiments).

The separation of the fully substituted protein from the thiol-containing forms may be essential in many cases for the use of the protein in biochemical or biophysical studies. Therefore, $(Se)_2$-thioredoxin was separated from thioredoxin by making use of their different redox properties, which allowed the selective reduction and carboxymethylation of thioredoxin. The modified protein could then be separated from $(Se)_2$-thioredoxin due to the additional negative charge carried by the carboxymethylated thiol group. The (SeS)-thioredoxin, which is the second minor component in the protein isolate, was not completely removed from $(Se)_2$-thioredoxin by this procedure, indicating that the Se-S bridge was only partially reduced by β-mercaptoethanol. The (SeS)-thioredoxin is composed of two species with either Cys32 or Cys35 replaced by selenocysteine. It may well be that the two seleno–sulfur bridges, SeCys32/Cys35 and Cys32/SeCys35, respectively, have different redox potentials in two isomeric proteins and that only one of them was selectively reduced by the β-mercaptoethanol. The observation that approximately half of the seleno–sulfur form of the protein could be separated from $(Se)_2$-thioredoxin does indeed support this view.

There may be additional separation problems in the case of proteins possessing more cysteine residues than thioredoxin. Possibilities to overcome these difficulties could reside in the reversible modification of either the selenol or thiol function and in the subsequent separation of the modified protein from the unmodified one and removal of the substituent. It is known that the selenol group is reactive at acidic pH where the thiol group does not react.

The method for the replacement of cysteine residues by its seleno analogue should find application as a new labeling tool

for biomolecules to support biostructural research. It can be used for the determination of protein structures, in analogy to selenomethionine-labeled proteins (Hendrickson et al., 1990; Hendrickson, 1991), by X-ray multiwavelength anomalous diffraction methods. Molecular modeling of $(Se)_2$-thioredoxin using the refined X-ray coordinates of thioredoxin together with the geometrical parameters for the selenium atom showed that no conformational differences between the two proteins are to be expected. A detailed conformational characterization of $(Se)_2$-thioredoxin by ^1H NMR is presently underway in our laboratory.

The uniform replacement of cysteine residues by selenocysteine could also potentially open new possibilities for assignment of unknown disulfide connectivities in protein structures by heteronuclear $^1H/^{77}Se$-correlated NMR spectroscopy of the homologous selenium-labeled protein. An unknown disulfide pattern in the natural protein could thus be inferred by homology. A prerequisite for such an analysis, however, would be the formation of identical diselenide and disulfide connectivities, respectively, in the two homologous proteins. This can be expected if the ^1H NMR spectra of the two proteins are identical. The existing biochemical problem of disulfide bridge mapping in cysteine-rich proteins by biochemical methods (Creighton, 1989; Klaus et al., 1993) could thus be resolved. The actual implementation of the technique, however, may need enrichment of the NMR-active ^{77}Se nucleus (spin $^1/_2$) above its level of natural abundance of 7.6%. No separation of the heterogeneous seleno–sulfur protein, however, seems necessary, as the spin system of the two diselenide-linked selenocysteine residues and its nearest neighbors in the sequence can be recognized unequivocally also in a heterogeneous protein sample (Wüthrich, 1986).

The replacement of cysteine residues by selenocysteine may further be of interest for the study of protein folding pathways as the kinetic and thermodynamic properties of the selenol group are distinctly different from those of the thiol group. It is therefore a priori unknown whether the seleno analogue of a protein does still fold properly if disulfide-bonded intermediates have to be formed in the specific folding pathway of the protein (Goldenberg, 1992; Creighton et al., 1993). Early disulfide-bonded intermediates may be trapped efficiently as seleno analogues, or, alternatively, the folding to the native state may occur under different reducing and pH conditions.

ACKNOWLEDGMENT

We thank Dr. A. Holmgren for the donation of plasmid pUC118-*trxA* and Dr. Y. Hou for the gift of purified cysteinyl-tRNA ligase. Dr. Daniel Bur (Hoffmann-La Roche) is gratefully acknowledged for his help in the molecular modeling of $(Se)_2$-thioredoxin. We are grateful to M. Meyer and J. R. Andreesen for their unpublished information on $(Se)_2$-thioredoxin as a substrate of thioredoxin reductase.

REFERENCES

Berglund, O., & Sjöberg, B. M. (1970) *J. Biol. Chem. 267*, 22440–22445.

Bruins, A. P., Lovey, T. R., & Henion, J. D. (1987) *Anal. Chem. 59*, 2642–2646.

Casadaban, M. J., & Cohen, S. N. (1979) *Proc. Natl. Acad. Sci. U.S.A. 67*, 4530–4533.

Cowie, D. B., & Cohen, G. N. (1957) *Biochim. Biophys. Acta 26*, 252–261.

Creighton, T. E. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 5082–5086.

Creighton, T. E. (1989) in *Protein structure, a practical approach* (Creighton, T. E., Ed.) pp 155–167, IRL Press Ensyham, Oxford, England.

Creighton, T. E., Bagley, C. J., Cooper, L., Darby, N. J., Freedman, R. B., Kemmink, J., & Sheikh, A. (1993) *J. Mol. Biol. 232*, 1176–1196.

Denk, D. (1987) Dissertation, Ludwig-Maximilians-Universität, München.

Dyson, H. J., Holmgren, A., & Wright, P. E. (1988) *FEBS Lett. 228*, 254–258.

Dyson, H. J., Gippert, G. P., Case, D. A., Holmgren, A., & Wright, P. E. (1990) *Biochemistry 29*, 4129–4136.

Ehrenreich, A., Forchhammer, K., Tormay, P., Veprek, B., & Böck, A. (1992) *Eur. J. Biochem. 206*, 767–773.

Emsley, J. (1992) *The Elements*, Clarendon Press, Oxford.

Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F., & Whithouse, C. M. (1990) *Mass Spectrom. Rev. 9*, 37–70.

Frank, P., Licht, A., Tullius, T. D., Hodgson, K. O., & Pecht, I. (1985) *J. Biol. Chem. 260*, 5518–5525.

Gettins, P., & Crews B. C. (1991) *J. Biol. Chem. 266*, 4804–4809.

Gettins, P., & Wardlaw, S. (1991) *J. Biol. Chem. 226*, 3422–3426.

Goldenberg, D. P. (1992) *Trends Biochem. Sci. 17*, 257–261.

Hanahan, D. (1985) in *DNA Cloning* (Glover, D. M., Ed.) Vol. 1, pp 109–135, IRL Press, Oxford.

Hartmanis, M. G. N., & Stadtman, T. C. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 4912–4916.

Hendrickson, W. A. (1991) *Science 254*, 51–58.

Hendrickson, W. A., Horton, J. R., & LeMaster, D. M. (1990) *EMBO J. 9*, 1665–1672.

Holmgren, A. (1968) *J. Biochem. 6*, 475–484.

Holmgren, A., Söderberg, B.-O., Eklund, H., & Brändén, C.-I. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72*, 2305–2309.

House, K. L., Dunlap, R. B., Odom, J. D., Wu, Z.-P., & Hilvert, D. (1992) *J. Am. Chem. Soc. 114*, 8573–8579.

House, K. L., Garber, A. R., Dunlap, R. B., Odom, J. D., & Hilvert D. (1993) *Biochemistry 32*, 3468–3473.

Huber, R. E., & Criddle, R. S. (1967) *Arch. Biochem. Biophys. 122*, 164–173.

Kallis, G. B., & Holmgren, A. (1980) *J. Biol. Chem. 255*, 10261–10265.

Katti, S. K., LeMaster, D. M., & Eklund, H. (1990) *J. Mol. Biol. 212*, 167–184.

Kelley, R. F., Shalango, W., Jagannadham, M. V., & Stellwagen, E. (1987) *Biochemistry 26*, 1406–1411.

Klaus, W., Broger, C., Gerber, P., & Senn, H. (1993) *J. Mol. Biol. 232*, 897–906.

Kramer, G. F., & Ames, B. N. (1988) *J. Bacteriol. 170*, 736–743.

Krause, G., & Holmgren, A. (1991) *J. Biol. Chem. 266*, 4056–4066.

Laemmli, U. K. (1970) *Nature 227*, 680–685.

Maniatis, T., Fritsch, E. F., & Sambrook, J. (1982) *Molecular Cloning: a Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Miller, J. H. (1972) *Experiments in Molecular Genetics*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Mullen, G. P., Dunlap R. B., & Odom, J. D. (1986) *Biochemistry 25*, 5625–5632.

Müller, K., Ammann, H. J., Doran, D. M., Gerber, P. R., Gubernator, K., & Schrepfer, G. (1988) *Com. Soc. Chim. Belg. 97*, 655–667.

Odom, J. D. (1983) *Structure and Bonding 54*, 1–26.

Senn, H., Eugster, A., Otting, G., Suter, K., & Wüthrich, K. (1987) *Eur. Biophys. J. 14*, 301–306.

Sliwkowski, M. X., & Stadtman, T. C. (1985) *J. Biol. Chem. 260*, 3140–3144.

Stadtman, T. C., Davis, J. N., Zehelein, E., & Böck, A. (1989) *BioFactors 2*, 35–44.

Studier, F. W., & Moffatt, B. A. (1986) *J. Mol. Biol. 189*, 113–130.
Syed, R., Wu, Z.-P., Hogle, J. M., & Hilvert, D. (1993) *Biochemistry 32*, 6157–6164.
Szajewski, R. P., & Whitesides, G. M. (1980) *J. Am. Chem. Soc. 102*, 2011–2026.
Tabor, S., & Richardson, C. C. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 1074–1078.
Veres, Z., Tsai, L., Scholz, T. D., Politino, M., Balaban, R. S., & Stadtman, T. C. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 2975–2979.
Whitaker, J. R., & Granum, P. E. (1980) *Anal. Biochem. 109*, 156–159.
Whitehouse, C. M., Dreyer, R. N., Yamashita, M., & Fenn, J. B. (1985) *Anal. Chem. 57*, 675–679.
Wu, J. T., Pieper, R. K., Wu, L. H., & Peters, J. L. (1989) *Clin. Chem. 35*, 778–782.
Wu, Z.-P., & Hilvert, D. (1989) *J. Am. Chem. Soc. 111*, 4513–4514.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.
Zubay, G. (1962) *J. Mol. Biol. 4*, 347–356.