Trypanothione Reductase of *Trypanosoma congolense*: Gene Isolation, Primary Sequence Determination, and Comparison to Glutathione Reductase†

Spencer L. Shames,‡ Bruce E. Kimmel,§ Oliver P. Peoples,∥ Nina Agabian,§ and Christopher T. Walsh*,‡

Departments of Chemistry, Biology, and Applied Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, and Intercampus Program in Molecular Parasitology, University of California at San Francisco/Berkeley, San Francisco, California 94143

Received October 30, 1987; Revised Manuscript Received January 22, 1988

ABSTRACT: The gene encoding trypanothione reductase, the redox disulfide-containing flavoenzyme that is unique to the parasitic trypanosomatids (Shames et al., 1986), has been isolated from the cattle pathogen *Trypanosoma congolense*. Library screening was carried out with inosine-containing oligonucleotide probes encoding sequences determined from two active site peptides isolated from the purified *Crithidia fasciculata* enzyme. The nucleotide sequence of the gene was determined according to the dideoxy chain termination method of Sanger. The structural gene is 1476 nucleotides long and encodes 492 amino acids. We have identified the active site peptide containing the redox-active disulfide, a peptide corresponding to the histidine-467 region of human erythrocyte glutathione reductase, as well as the flavin binding domain that is highly conserved in all disulfide-containing flavoprotein reductase enzymes. Alignment of five tryptic peptides (80 residues) isolated from the *C. fasciculata* trypanothione reductase with the primary sequence of the *T. congolense* enzyme showed 88% homology with 76% identity. Additionally, a sequence comparison of the glutathione reductase from *Escherichia coli* or human erythrocytes to *T. congolense* trypanothione reductase reveals >50% homology. A search for the amino acid residues in the primary sequence of trypanothione reductase functionally active in binding/catalysis in human erythrocyte glutathione reductase shows that only the two arginine residues (Arg-37 and Arg-347), shown by X-ray crystallographic data to hydrogen bond to the GS_1 glutathione glycyl carboxylate, are absent.

The trypanosomatid parasites are causative agents of a variety of serious diseases in both humans and domestic animals. These include African sleeping sickness (*Trypanosoma gambiense* and *Trypanosoma rhodesiense*) and Chagas disease (*Trypanosoma cruzi*) in humans and nagana (*Trypanosoma congolense* and *Trypanosoma brucei*) in cattle. Numerous clinical and research observations indicate that these parasites may be selectively susceptible to oxidant stress by the reduced oxygen metabolites superoxide, hydrogen peroxide, and hydroxyl radical. These later toxic species may be generated, for example, by one-electron redox cycling of such trypanocidal agents as nifurtimox and hematoporphyrin (Fairlamb, 1982; Schirmer et al., 1987).

Glutathione, as the major thiol in mammalian cells, plays an integral role in a variety of cellular functions including the protection against oxidative stress as carried out by the glutathione peroxidase/glutathione reductase enzyme couple. Whereas all trypanosomatids show superoxide dismutase activity, no iron-dependent peroxidase has been discovered, thus suggesting that a glutathione-dependent peroxidase/reductase couple might play a crucial role in parasite survival in the host blood stream. The recent discovery that trypanosomatids contain very little glutathione and no classical glutathione reductase but instead contain the 24-membered macrocyclic disulfide N^1,N^8-bis(glutathionyl)spermidine, given the trivial name trypanothione (Fairlamb et al., 1985), led us to purify and characterize the specific trypanothione reductase from the insect trypanosomatid *Crithidia fasciculata* (Shames et al., 1986). We established the structural and mechanistic similarity of this new disulfide-containing flavoprotein to human erythrocyte glutathione reductase but found that the host and parasite enzymes showed mutually exclusive binding and catalysis of their respective disulfide-containing substrates. Most recently, Krauth-Siegel et al. (1987) demonstrated strong homology between the trypanothione reductase from the human pathogen *T. cruzi* and the enzyme from *C. fasciculata*. In similar fashion to the crithidial enzyme, the *T. cruzi* reductase also showed strong specificity toward the binding of its physiological disulfide-containing substrate, trypanothione.

The mutually exclusive substrate binding seen with parasite trypanothione and host glutathione reductase enzymes suggests this key antioxidant enzyme to be an attractive target for the design of selective antiparasitic agents. Although a high-resolution X-ray crystal structure may be essential to understand fully the functional basis for the substrate differentiation observed, present efforts are severely hampered by the lack of enzyme availability. Thus, to gain insights into the structural differences between host glutathione and parasite trypanothione reductase enzymes and to provide a means for obtaining larger yields of trypanothione reductase, we have undertaken the cloning and sequencing of the gene that encodes trypanothione reductase in the cattle pathogen *T. congolense*. Additionally, the provision of a primary sequence has yielded the necessary basis for an analytical comparison of the primary structures of host and parasite disulfide reductase enzymes.

Materials and Methods

Materials. β-Cyanoethyl diisopropyl phosphoramidites were from Pharmacia. M13mp18, M13mp19, deoxyadenosine

† This work was supported in part by NIH grant GM 21643 and the McCarther Foundation. S.L.S. is a recipient of a National Institutes of Health postdoctoral fellowship. N.A. is a Burrough-Wellcome Scholar of Molecular Parasitology.

* Address correspondence to this author at the Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115.

‡ Departments of Chemistry and Biology.

§ Intercampus Program in Molecular Parasitology.

∥ Department of Applied Biology.

0006-2960/88/0427-5014$01.50/0

Table I: Oligonucleotide Probes Utilized for Trypanosomatid Gene Screening

Redox-Active Disulfide-Containing N-Terminal Active Site Peptide										
peptide	Gly	Gly	Thr	Cys	Val	Asn	Val	Gly	Cys	Val
probe	GGI	GGI	ACI	TGI	GTI	AAI	GTI	GGI	TGI	GT
Histidine-467 C-Terminal Active Site Peptide										
peptide	Asp	Phe	Tyr	Asn	Thr	Ile	Gly	Val	His	Pro	Thr
probe	GAI	TTI	TAI	AAI	ACI	ATI	GGI	GTI	CAI	CCI	AC

5′-(α-[^{35}S]thiotriphosphate), deoxy- and dideoxynucleoside triphosphates, the −40 universal sequencing primer, and the Klenow fragment of DNA polymerase I were from Amersham. Restriction enzymes and T4 DNA ligase were obtained from either IBI or New England Biolabs. Calf intestinal alkaline phosphatase was from Boehringer-Mannheim. All other reagents were of the highest quality available from Sigma and Aldrich.

Peptide Isolation and Sequence Determination. Tryptic digestion of *C. fasciculata* trypanothione reductase followed by high-performance liquid chromatography (HPLC) purification of the resulting tryptic peptides was carried out as described for the redox-active disulfide-containing active site peptide (Shames et al., 1986). Four peptides, chosen randomly, were sequenced by William Lane at the Harvard Microchemistry Facility (Cambridge, MA) on an Applied Biosystems 470A protein sequencer equipped with an on-line 120A phenylthiohydantoin (PTH) analyzer.

Oligonucleotide Probes. Synthetic oligonucleotide probes were synthesized on a Pharmacia gene assembler according to the manufacturer's directions. Following detritylation of the final residue, the oligomers were deprotected fully and removed from the solid support by treatment with 28% ammonium hydroxide at 55 °C for 14 h. The oligonucleotides were purified by preparative gel electrophoresis on a 20% acrylamide gel. The purified oligonucleotides were excised and extracted into 3 mL of 0.1 M TEAB buffer (triethylammonium bicarbonate, pH 7.5) over 12 h at 37 °C. Desalting and final isolation were carried out on a Sep-pack C18 cartridge (Waters Associates) that was equilibrated in 25 mM TEAB. After the crude oligonucleotide mixture was applied, the resin was washed with 25 mM TEAB and the oligonucleotide eluted with 35 mM TEAB containing 50% v/v acetonitrile. The fractions (0.9 mL) containing the oligonucleotide were evaporated on a Savant Instruments Speed Vac, and the purified oligonucleotide was dissolved in 0.5 mL of water.

The oligonucleotides were phosphorylated on the 5′-ends with [γ-^{32}P]ATP and T4 polynucleotide kinase. The labeled probes were purified by preparative electrophoresis on a 20% acrylamide gel and visualized by autoradiography. Following extraction of the oligonucleotides into 0.5 mL of 0.5 M ammonium acetate buffer containing 1 mM ethylenediaminetetraacetic acid (EDTA) for 24 h at 37 °C, the acrylamide fragments were removed by filtration through a 0.2 μm Nalgene syringe filter. The oligonucleotide probes had specific activities of $\sim 1 \times 10^8$ cpm μg^{-1} and were used without further purification.

T. congolense Genomic DNA Library Preparation. A size-selected *T. congolense* genomic DNA library was prepared in the pUC8 vector. *T congolense* genomic DNA (10 μg) was digested to completion with *Bam*HI under conditions recommended by the manufacturer. The digested DNA was size-selected for fragments between 2.5 and 4.0 kilobases (kb) by preparative electrophoresis on a 0.8% agarose gel. The selected DNA was isolated by electroelution into dialysis tubing followed by ethanol precipitation. The isolated fragments were ligated into the pUC8 vector that had been digested with *Bam*HI and treated with calf intestine alkaline phosphatase. The ligation reaction was transformed into *E. coli* strain HB101 (BRL), and the transformants were selected by growth on ampicillin-containing agar plates. The library obtained by this method contained ca. 4000 recombinants.

Oligonucleotide Screening of T. congolense pUC8 Library. Two thousand recombinants were screened with the ^{32}P-labeled His-467 probe. Colonies were lifted onto nitrocellulose membranes, lysed, and screened as described by Maniatis et al. (1982). The two positives obtained were subsequently rescreened by the same methods.

DNA Sequence Determination and Analysis. DNA sequencing was carried out according to the dideoxy chain termination method of Sanger et al. (1977). DNA fragments obtained by restriction endonuclease digestion of the pCGTR plasmid were purified by preparative agarose gel electrophoresis and subsequently isolated by electroelution on an IBI electroelution apparatus followed by ethanol precipitation. The isolated fragments were ligated into the corresponding restriction sites of M13 vectors mp18 or mp19 that had been pretreated with calf intestine alkaline phosphatase. G/C-rich regions were resequenced using 7-deaza-GTP in place of GTP in the sequencing reaction mixtures. Computer-assisted sequence analysis was carried out with the Staden programs (Staden & McLachlan, 1982; Staden, 1982a) as modified by the MIT Seq Group. The extent of linear alignment (Dayhoff et al., 1983) and diagonal conservative correlation (Staden, 1982b) of the trypanothione reductase to both the *E. coli* and human erythrocyte glutathione reductase protein sequences was computed on the basis of mutation data scoring matrices.

Results

Gene Isolation. The sequence of two active site peptides,[1] which had been isolated by tryptic digestion of *C. fasciculata* trypanothione reductase, was used to prepare oligonucleotide probes for the initial screening of trypanosomatid genomic DNA (Table I). Because only limited information concerning codon usage in trypanosomatids is available, the probes were synthesized with inosine in each wobble position to alleviate problems with high background often found with highly degenerate probes (Takahashi et al., 1985; Ohtsuka et al., 1985). Initially, the genomic DNA of several pathogenic trypanosomatids including *T. brucei*, *T. gambiense*, *T. cruzi*, *T. equiperdeium*, and *T. congolense* was screened by Southern hybridization. Under the conditions employed (see Materials and Methods), both *Bam*HI and *Hin*dIII digestions of *T. congolense* DNA showed the strongest hybridization to the oligonucleotide probes with the His-467 probe yielding the lowest background. This latter probe showed highly specific hybridization and revealed only a single *Bam*HI fragment of 3.2 kb at high stringency. Thus, a library of size-selected *Bam*HI-digested *T. congolense* genomic DNA of 2.5–4.0 kb

[1] The isolation and sequencing of the redox-active disulfide-containing active site peptide were reported in Shames et al. (1986). Identification of the histidine-containing peptide as an active site peptide was based on the strong sequence homology to the region surrounding histidine-467 in human erythrocyte glutathione reductase.

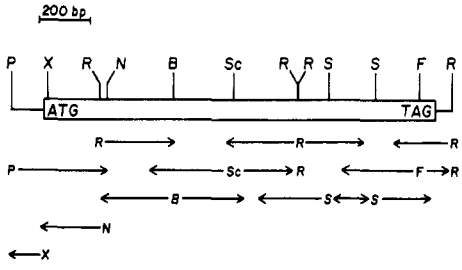

FIGURE 1: Strategy utilized to sequence the trypanothione reductase gene. The restriction sites from which sequence data were obtained are shown. The extent of sequence determined from each site is indicated by the arrows. The restriction sites presented are as follows: P, *Pst*I; X, *Xmn*1; R, *Rsa*1; N, *Nae*1; B, *Bal*1; Sc, *Sac*1; S, *Sau*3A1; F, *Fsp*1. The boxed portion represents the trypanothione reductase structural gene.

was constructed in the pUC8 vector and the library subsequently screened with the His-467 probe as described under Materials and Methods. Two positive clones containing 3.2-kb *Bam*HI inserts were identified and named pCGTR1 and pCGTR2. Restriction analysis showed pCGTR1 and pCGTR2 to contain the same DNA fragment ligated in opposite orientations (data not shown). Because of a preferred insert orientation, all subsequent work was carried out with plasmid pCGTR2.

To verify the presence of the reductase gene on the *Bam*HI insert in pCGTR2, the clone was digested with *Rsa*1 and hybridized with the His-467 probe; a single 580 base pair (bp) *Rsa*1 fragment was found to show strong hybridization. This fragment was subcloned into the *Sma*I site of M13mp19 and sequenced. Translation of the resulting sequence revealed the C-terminus of a glutathione reductase like protein with a C-terminal active site peptide identical in homology with that found for the *C. fasciculata* enzyme. The presence of the trypanothione reductase gene in toto was determined by restriction analysis of the pCGTR2 *Bam*HI insert. The restriction map revealed 1.6 kb of DNA on the 5′-side of the 580-bp *Rsa*1 C-terminal fragment. Thus, given the size of the *C. fasciculata* enzyme of ca. 500 amino acids and the fact that no introns have been found in any trypanosomatid structural genes to date, the *Bam*HI insert of pCGTR2 should encompass the entire gene.

DNA Sequence of Trypanothione Reductase Gene. The complete DNA sequence of the trypanothione reductase gene was obtained according to the M13/Sanger dideoxy chain termination method. To locate the N-terminal region, the DNA sequence obtained via bidirectional sequencing from the unique *Pst*I site was translated in all six reading frames and examined for the presence of the redox-active disulfide-containing active site peptide. This peptide was expected to lie near the N-terminus given the apparent homology of trypanothione reductase to both *E. coli* and human erythrocyte glutathione reductase enzymes. Once the active site was located, the complete nucleotide sequence was determined by the strategy shown in Figure 1.

Amino Acid Sequence and Composition of T. congolense Trypanothione Reductase. Because the N-terminal of *C. fasciulata* trypanothione reductase was blocked and could not be sequenced (Shames et al., 1986), we relied on the N-terminal redox-active disulfide-containing active site peptide to serve as a reading frame reference. With this strategy, a single open reading frame beginning with an ATG start codon 152 bp from the *Pst*I site was found. The open reading frame thus consists of 1476 nucleotides, which encodes 492 amino acid residues (Figure 2). From the amino acid composition, a subunit molecular weight of 53 443 was determined. This is

Table II: Amino Acid Composition Comparison of *T. congolense* Trypanothione Reductase and Human Erythrocyte Glutathione Reductase

amino acid	trypanothione reductase	glutathione reductase[a]	amino acid	trypanothione reductase	glutathione reductase[a]
Ala	38	42	Leu	36	34
Arg	16	17	Lys	34	34
Asn	21	17	Met	13	15
Asp	23	21	Phe	21	14
Cys	8	10	Pro	23	24
Gln	12	11	Ser	30	31
Glu	23	29	Thr	34	31
Gly	46	43	Trp	5	3
His	10	16	Tyr	14	13
Ile	35	29	Val	43	44

[a] Data from Krauth-Siegel et al. (1982).

in excellent agreement with the value of 53 800 determined experimentally for the *C. fasciculata* enzyme. A comparison of the amino acid compositions of *T. congolense* trypanothione reductase and human erythrocyte glutathione reductase is shown in Table II.

To validate that the gene isolated in fact encoded a trypanothione reductase, the primary sequence determined was screened for homology to the five tryptic peptides isolated from the *C. fasciculata* enzyme. As shown in Figure 3, there is strong homology between the two trypanosomatid enzymes with 76% identity over a total of 80 amino acid residues. When the conservative changes indicated by dashed lines in Figure 3 are also examined, an overall homology of 88% is calculated. The numbers above key residues indicate their position in the *T. congolense* enzyme.

Primary Sequence Comparison to E. coli and Human Erythrocyte Glutathione Reductase Enzymes. Linear alignment of the primary sequence of *T. congolense* trypanothione reductase and *E. coli* and human erythrocyte glutathione reductase enzymes as determined by the method of Dayhoff et al. (1983) is shown in Figure 4. The alignment shows identity between trypanothione reductase and *E. coli* and human erythrocyte glutathione reductase enzymes of 38% and 41%, respectively. However, when conservative changes are also scored in the alignment as carried out by the Staden programs, greater than 50% homology is revealed in each case (Figure 5). It is striking that optimal alignment of the primary sequence of trypanothione reductase with both bacterial and human glutathione reductase enzymes reveals a C-terminal extension of 20 amino acid residues.

DISCUSSION

Our initial mechanistic and physical studies on *C. fasiculata* trypanothione reductase suggested that this enzyme was a catalyst built along the same lines as human erythrocyte glutathione reductase (Shames et al., 1986). Isolation of the gene that encodes trypanothione reductase from the related trypanosomatid *T. congolense* has not only highlighted further the similarities between these two biological catalysts but has now revealed some of the important structural features that distinguish host and parasite enzymes.

The *T. congolense* trypanothione reductase subunit is 492 amino acid residues compared to 478 for erythrocyte glutathione reductase and 450 for the *E. coli* enzyme. Alignment of *T. congolense* trypanothione reductase and human erythrocyte glutathione reductase places the N-terminal methionine with residue 17 of the host enzyme. However, the *E. coli* reductase, which lacks the first 17 residues compared to the erythrocyte enzyme, shows very close N-terminal alignment

```
      10                  30                  50                  70                  90                  110
CTGCAGCGTTTCTTTTTCTTTTACTTTAGTATTTTGGTATTATCTCCATCATTCCTATTTTAATTTTTTTAACACATATTCTTTTTCTTTGCACACACACACACACACAAGAAGGCCCT

      130                 150                 170                 190                 210                 230
                              M  S  K  A  F  D  L  V  I  I  G  A  G  S  G  G  L  E  A  G  W  N  A  A  T  L  Y  K  K  R
TTCCAAGATTTTTCGTTTCAATCGCTTTTCTATGTCGAAGGCGTTCGACCTCGTTATCATTGGCGCTGGTTCAGGAGGACTGGAAGCTGGTTGGAACGCCGCTACACTATACAAGAAACG

      250                 270                 290                 310                 330                 350
  V  A  V  V  D  V  Q  T  V  H  G  P  P  F  F  A  A  L  G  G  T  C  V  N  V  G  C  V  P  K  K  L  M  V  T  G  A  Q  Y  M
TGTTGCTGTGGTTGACGTACAAACCGTCCATGGACCACCTTTTTTCGCTGCACTTGGCGGCACGTGTGTTAACGTGGGTTGCGTGCCGAAGAAGCTCATGGTCACCGGGGCACAGTACAT

      370                 390                 410                 430                 450                 470
  D  Q  L  R  E  S  A  G  F  G  W  E  F  D  A  S  T  I  K  A  N  W  K  T  L  I  A  A  K  N  A  A  V  L  D  I  N  K  S  Y
GGATCAACTACGCGAGTCTGCCGGCTTCGGCTGGGAATTTGATGCCTCAACAATTAAAGCAAACTGGAAAACGTTGATAGCAGCCAAGAATGCGGCTGTATTGGACATAAACAAAAGCTA

      490                 510                 530                 550                 570                 590
  E  D  M  F  K  D  T  E  G  L  E  F  F  L  G  W  G  A  L  E  Q  K  N  V  V  T  V  R  E  G  A  D  P  K  S  K  V  K  E  R
TGAAGACATGTTCAAAGATACAGAGGGTCTTGAGTTCTTTTTGGGTTGGGGAGCATTGGAGCAAAAAAATGTAGTCACCGTTCGCGAAGGTGCAGACCCGAAAAGTAAAGTGAAGGAACG

      610                 630                 650                 670                 690                 710
  L  Q  A  E  H  I  I  I  A  T  G  S  W  P  Q  M  L  K  I  P  G  I  E  H  C  I  S  S  N  E  A  F  Y  L  E  E  P  P  R  R
CCTTCAGGCGGAGCACATCATCATAGCGACCGGCTCATGGCCACAGATGCTTAAGATTCCCGGCATTGAGCACTGCATAAGCAGTAATGAAGCGTTCTACCTGGAAGAACCTCCCCGTCG

      730                 750                 770                 790                 810                 830
  V  L  T  V  G  G  G  F  I  S  V  E  F  A  G  I  F  N  A  Y  K  P  V  G  G  K  V  T  L  C  Y  R  N  N  P  I  L  R  G  F
TGTGTTGACCGTTGGCGGTGGTTTTATTTCCGTTGAGTTTGCTGGTATTTTTAATGCATACAAACCCGTGGGTGGCAAGGTAACACTATGCTACCGCAACAATCCCATTCTTCGTGGATT

      850                 870                 890                 910                 930                 950
  D  Y  T  L  R  Q  E  L  T  K  Q  L  V  A  N  G  I  D  I  M  T  N  E  N  P  S  K  I  E  L  N  P  D  G  S  K  H  V  T  F
TGACTACACGCTTCGGCAAGAGCTCACGAAACAGCTCGTTGCTAATGGTATTGACATCATGACAAATGAAAACCCCTCCAAGATTGAACTCAATCCTGATGGGTCTAAACATGTCACCTT

      970                 990                 1010                1030                1050                1070
  E  S  G  K  T  L  D  V  D  V  V  M  M  A  I  G  R  L  P  R  T  G  Y  L  Q  L  Q  T  V  G  V  N  L  T  D  K  G  A  I  Q
TGAAAGCGGTAAGACATTGGATGTGGACGTTGTCATGATGGCAATTGGTCGTCTTCCCCGCACTGGATACTTGCAGTTGCAGACTGTCGGAGTAAACCTCACTGACAAGGGTGCTATACA

      1090                1110                1130                1150                1170                1190
  V  D  E  F  S  R  T  N  V  P  N  I  Y  A  I  G  D  V  T  G  R  I  M  L  T  P  V  A  I  N  E  G  A  S  V  V  D  T  I  F
AGTGGACGAGTTTTCTCGTACCAACGTACCCAACATCTACGCAATTGGGGATGTAACGGGTCGTATTATGCTGACTCCGGTGGCTATTAATGAGGGAGCCAGTGTCGTAGATACCATTTT

      1210                1230                1250                1270                1290                1310
  G  S  K  P  R  K  T  D  H  T  R  V  A  S  A  V  F  S  I  P  P  I  G  T  C  G  L  T  E  E  E  A  A  K  S  F  E  K  V  A
TGGTAGCAAGCCGCGGAAAACTGATCACACTCGTGTTGCCAGTGCCGTATTCTCCATTCCCCCAATTGGAACATGTGGTCTCACTGAAGAAGAGGCGGCAAAGAGTTTTGAAAAGGTTGC

      1330                1350                1370                1390                1410                1430
  V  Y  L  S  C  F  T  P  L  M  H  N  I  S  G  S  K  Y  K  K  F  V  A  K  I  I  T  D  H  G  D  G  T  V  V  G  V  H  L  L
GGTTTACCTTTCATGTTTTACCCCACTAATGCACAATATCAGTGGGTCAAAATACAAAAAGTTTGTGGCGAAGATTATCACAGATCACGGTGATGGGACAGTGGTTGGTGTGCACCTTCT

      1450                1470                1490                1510                1530                1550
  G  D  S  S  P  E  I  I  Q  A  V  G  I  C  M  K  L  N  A  K  I  S  D  F  Y  N  T  I  G  V  H  P  T  S  A  E  E  L  C  S
TGGGGACAGTTCACCAGAAATCATCCAAGCTGTTGGTATCTGCATGAAACTTAATGCAAAGATAAGTGACTTTTACAACACTATTGGTGTGCACCCGACAAGCGCAGAAGAGTTATGCTC

      1570                1590                1610                1630                1650                1670
  M  R  T  P  S  H  Y  Y  I  K  G  E  K  M  E  T  L  P  D  S  S  L  *
CATGCGCACACCCTCTCACTACTATATAAAAGGGGAGAAGATGGAAACACTACCGGACTCGAGCCTCTAGACGAAAGACAGCAGCGACGTGGCGATACCCAGTGGACTTTCGGCAAATAT
```

FIGURE 2: DNA sequence and translation of the trypanothione reductase gene. The coding region starts at position 152 and is terminated at position 1628. The regions to which the oligonucleotide probes were targeted are underlined.

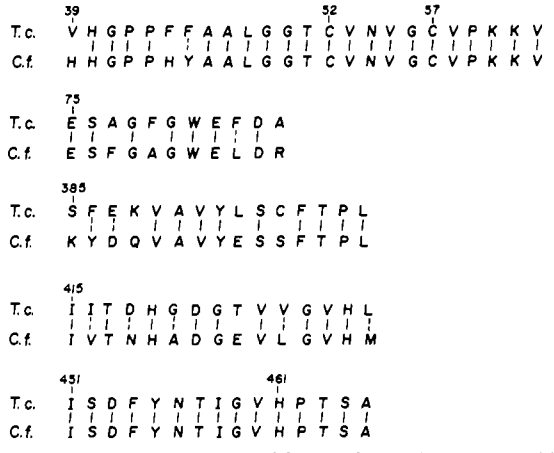

FIGURE 3: Sequence comparison of five *C. fasciculata* trypanothione reductase tryptic peptides to the primary sequence of *T. congolense* trypanothione reductase. Identities are indicated by solid lines, whereas conservative changes are indicated by dashed lines. The position of each peptide in the primary sequence of the *T. congolense* enzyme is given.

with the trypanosomatid trypanothione reductase at the start. The optimal alignment of all three enzymes suggests a 20 amino acid C-terminal extension for the parasite reductase. Two significant gaps were inserted in both the host and *E. coli* enzymes to optimize the alignment shown in Figure 4. Between residues 53 and 54 of the host glutathione reductase the trypanothione reductase sequence contains the nonapeptide VHGPPFFAA, whereas between residues 133 and 134 the decapeptide QKNVVTVREG insertion is found. Interestingly, both of these insertions occur at bends that follow β-sheet structures in the glutathione reductase (Karplus & Schulz, 1987). The strong sequence homology observed upon comparison of the redox-active disulfide-containing active site peptide from crithidial and *T. congolense* trypanothione reductase enzymes, as shown in Figure 3, suggests that the nonapeptide insertion -HGPP-*X*AA (where *X* is either F or Y) may be characteristic of all trypanothione reductase enzymes.

Aside from the overall similarity in amino acid composition between host and parasite enzymes, extensive homologies

```
          1       10        20        30        40        50        60
          :        :        :         :         :         :         :
GR:       ACRQEPQPQGPPPAAGAVASYDYLVIGGGSGGLASARRAAEL-GARAAVVESHK------
TR:       ---------------MSKAFDLVIIGAGSGGLEAGWNAATLYKKRVAVVDVQTVHGPPF
gor:      ---------------MTKHYDYIAIGGGSGGIASINRAA-MYGQKCALIEAKE------
Common:                          D    G GSGG      AA       A

         61       70        80        90        100       110       120
          :        :        :         :         :         :         :
GR:       ---LGGTCVNVGCVPKKVMWNTAVHSEFM---HD-HADYG--FPSCEGKFNWRVIKEKRD
TR:       FAALGGTCVNVGCVPKKLM---VTGAQYMDQLRE-SAGFGWEFDASTIKANWKTLIAAKN
gor:      ---LGGTCVNVGCVPKKVMWH---AAQIREAIHMYGPDYG--FDTTINKFNWETLIASRT
Common:      LGGTCVNVGCVPKK M                    G  F     K NW

        121       130       140       150       160       170       180
          :        :        :         :         :         :         :
GR:       AYVSRLNAIYQNNLTKSH-IEIIRGHAAFT----------SDPKPTIEVSGKKYTAPHIL
TR:       AAVLDINKSYEDMFKDTEGLEFFLGWGALEQKNVVTVTEGADPKSKVKER---LQAEHII
gor:      AYIDRIHTSYENVL-GKNNVDVIKGFARF-----------VDAK-TLEVNGETITADHIL
Common:   A        Y             G                D K           A HI

        181       190       200       210       220       230       240
          :        :        :         :         :         :         :
GR:       IATGGMPSTPHESQIPGASLGITS-DGFFQLEELPGRSVIVGAGYIAVEMAGILSA---L
TR:       IATGSWP---QMLKIPGIEHCISSNEAFY-LEEPPRRVLTVGGGFISVEFAGIFNAYKPV
gor:      IATGGRPSHP---DIPGVEYGIDS-DGFFALPALPERVAVVGAGYIAVELAGVING---L
Common:   IATG  P       IPG    I S   F  L   P R   VG G I VE AG

        241       250       260       270       280       290       300
          :        :        :         :         :         :         :
GR:       GSKTSLMIRHDKVLRSFDSMISTNCTEELENAGVEVLKFSQVKEVKKTLSGLEVSMVTAV
TR:       GGKVTLCYRNNPILRGFDYTLTQELTKQLVANGIDIM--TNENPSKIELNPDGSKHVTFE
gor:      GAKTHLFVRKHAPLRSFDPMISETLVEVMNAEGPQLH--TNAIPKAVVKNTDGSLTLELE
Common:   G K  L  R    LR FD              G

        301       310       320       330       340       350       360
          :        :        :         :         :         :         :
GR:       PGRLPVMTMIPDVDCLLWAIGRVPNTKDLSLNKLGIQTDDKGHIIVDEFQNTNVKGIYAV
TR:       SGK----TL--DVDVVMMAIGRLPRTGYLQLQTVGVNLTDKGAIQVDEFSRTNVPNIYAI
gor:      DGR----SE--TVDCLIWAIGREPANDNINLEAAGVKTNEKGYIVVDKYQNTNIEGIYAV
Common:    G          VD    AIGR P      L   G     KG I VD    TN   IYA

        361       370       380       390       400       410       420
          :        :        :         :         :         :         :
GR:       GDVCGKALLTPVAIAAGRKLAHRLFEYKEDSKLDYNNIPTVVFSHPPIGTVGLTEDEAIH
TR:       GDVTGRIMLTPVAINEGASVVDTIFGSKP-RKTDHTRVASAVFSIPPIGTCGLTEEEAAK
gor:      GDNTGAVELTPVAVAAGRRLSERLFNNKPDEHLDYSNIPTVVFSHPPIGTVGLTEPQARE
Common:   GD  G   LTPVA   G       F  K     D       VFS PFIGT GLTE  A

        421       430       440       450       460       470       480
          :        :        :         :         :         :         :
GR:       KYGIENVKTYSTSFTPMYHAVTKRK-TKCVMKMVCANKEEKVVGIHMQGLGCDEMLQGFA
TR:       SF--EKVAVYLSCFTPLMHNISGSKYKKFVAKIITDHGDGTVVGVHLLGDSSPEIIQAVG
gor:      QYGDDQVKVYKSSFTAMYTAVTTHR-QPCRMKLVCVGSEEKIVGIHGIGFGMDEMLQGFA
Common:         V  Y   FT               K          VG H  G    E  Q

        481       490       500       510       520       530
          :        :        :         :         :         :
GR:       VAVKMGATKADFDNTVAIHPTSSEELVTLR------------------
TR:       ICMKLNAKISDFYNTIGVHPTSAEELCSMRTPSHYYIKGEKMETLPDSSL
gor:      VALKMGATKKDGDNTVAIHPTAAEEFVTMR------------------
Common:      K A   D  NT   HPT  EE    R
```

FIGURE 4: Linear alignment of the protein sequences of *T. congolense* trypanothione reductase and human erythrocyte and *E. coli* glutathione reductase enzymes. The sequence data of both human erythrocyte (Krauth-Siegel et al., 1982) and *E. coli* (Greer & Perham, 1986) glutathione reductase were obtained from the NBRF data bank. Identities were located by the method of Dayhoff (1983) and are shown below all three sequences.

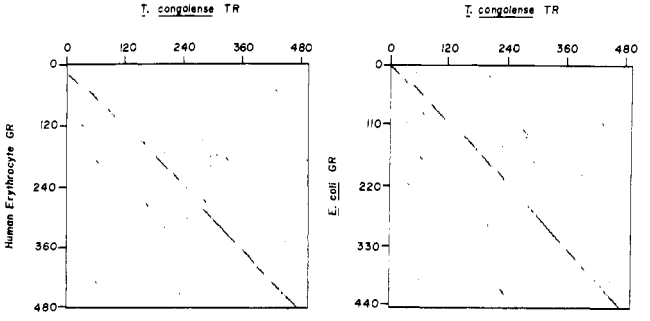

FIGURE 5: Diagonal dot matrix comparison of the protein sequence of *T. congolense* trypanothione reductase with that of both human erythrocyte and *E. coli* glutathione reductase enzymes. The conservative homology between the polypeptide chains was calculated according to a mutation scoring matrix with a window of 15 amino acids at 0.2% probability in each case.

between glutathione and trypanothione reductase enzymes occur in three regions. At the glycine-rich binding site for the ADP moiety of flavin adenine dinucleotide (FAD), residues 22–33 of glutathione reductase and residues 6–17 of trypanothione reductase (and *E. coli* glutathione reductase), the

Table III: Comparison of Catalytically Important Residues in Human Erythrocyte Glutathione Reductase and *T. congolense* Trypanothione Reductase Primary Sequences

GSSG reductase	function from X-ray analysis	S–S reductase
$GGGSGGL_{33}$	FAD binding site	$GAGSGGL_{17}$
Cys-58	redox-active disulfide	Cys-52
Cys-63	redox-active disulfide	Cys-57
Lys-57	binds carboxylate of GS_1 γ-Glu moiety	Lys-61
Tyr-114	stacks between GS moieties in GSSG	Tyr-110
His-467	active site base	His-461
Glu-472	H bonds to active site histidine	Glu-466
Arg-37	binds carboxylate of GS_1 glycine via H bonds	Trp-21[a]
Arg-347	binds carboxylate of GS_1 glycine via H bonds	Ala-343[a]

[a] No positive side chain for hydrogen bonding to GS_1 glycyl carboxylate.

βαβ-fold orientation (Wierenga et al., 1986) is preserved. Following the nonapeptide insertion amino proximal to the active site, there is a 14-residue identity around the canonical catalytic redox-active disulfide. In the C-terminal third of each protein sequence there is again highly extensive homology, including the upper dimer interface domain that corresponds to residues 439–469 of glutathione reductase (Karplus & Schulz, 1987). Not unexpectedly, there is also homology in the domain known to bind nicotinamide adenine dinucleotide phosphate (NADP) on glutathione reductase; the crucial Arg-218 and Arg-224 (Schirmer & Schulz, 1987) are found at positions 222 and 228 in trypanothione reductase. Lastly, key tripeptide residues at domain interfaces, $ATG_{155-157}$ between the FAD and reduced NADP (NADPH) domains and $ALG_{288-290}$ between the NADPH and central domains (Thieme et al., 1981), are identically conserved in trypanothione reductase at positions 159–161 and 284–286, respectively. Taken together, these data suggest that the overall domain folding patterns of host and parasite enzymes will be highly similar.

A central question that arises from initial evaluation of the primary sequence comparison is why trypanothione reductase from *C. fasciculata* (Shames et al., 1986), *T. cruzi* (Krauth-Siegel et al., 1986), and presumptively *T. congolense* reduces oxidized trypanothione but shows very poor binding affinity for the host cognative substrate oxidized glutathione. Collected in Table III are some residues in human erythrocyte glutathione reductase identified by both mechanistic and crystallographic analysis to be important residues for binding or catalysis. These residues include the active site disulfide constructed by Cys-58 and Cys-63, the Tyr-114 that stacks between the glycine moieties of bound glutathione, His-467 that is the active site base, and Glu-472 that hydrogen bonds to His-467 to orient the imidazole ring. As shown in Table III, all these residues have apparent counterparts in the *T. congolense* encoded trypanothione reductase sequence. A major difference occurs, however, in the two arginine residues at positions 37 and 347 of glutathione reductase. These basic residues lie close enough to bound glutathione in the crystals to allow hydrogen bonding with the glycyl carboxylate of the GS_1 moiety; this latter moiety engages in mixed disulfide linkage to Cys-58 in reductive catalytic turnover (Emil Pai and Georg Schulz, personal communication). These two arginine residues are missing from the *T. congolense* trypanothione reductase sequence substituted with Trp-21 and Ala-343 in the corresponding positions, respectively. Thus, these two amino acid replacements, taken together with the three peptide

insertions discussed previously, may contribute significantly to the very weak binding of oxidized glutathione to trypanothione reductase, a feature that may provide a basis for subsequent structural investigations.

With the *T. congolense* trypanothione reductase structural gene cloned and sequenced, efforts to express and overproduce the enzyme in *E. coli* are under way. Large quantities of the enzyme will be necessary for both mechanistic studies and the testing of inhibitors that may select between parasite trypanothione reductase and host glutathione reductase. Furthermore, since preliminary crystals of the *T. cruzi* trypanothione reductase have been reported recently (Krauth-Siegel et al., 1987) but enzyme availability was severely limited, use of the *T. congolense* gene as a probe to facilitate *T. cruzi* gene cloning, sequencing, and enzyme overproduction is a current strategy in these laboratories.

Registry No. Trypanothione reductase, 102210-35-5; DNA (trypanothione reductase gene), 114691-51-9; trypanothione reductase (protein moiety reduced), 114718-82-0; glutathione reductase, 9001-48-3.

References

Dayhoff, M. O., Barker, W. C., & Hunt, L. T. (1983) *Methods Enzymol. 91*, 524–545.

Fairlamb, A. H. (1982) *Trends Biochem. Sci. (Pers. Ed.) 7*, 249–253.

Fairlamb, A. H., Blackburn, P., Ulrich, P., Chait, B. T., & Cerami, A. (1985) *Science (Washington, D.C.) 227*, 1485–1487.

Greer, S., & Perham, R. N. (1986) *Biochemistry 25*, 2736–2742.

Karplus, P. A., & Schulz, G. E. (1987) *J. Mol. Biol. 195*, 701–729.

Krauth-Siegel, R. L., Blatterspiel, R., Saleh, M., Schiltz, E., Schirmer, R. H., & Untucht-Grau, R. (1982) *Eur. J. Biochem. 121*, 259–267.

Krauth-Siegel, R. L., Enders, B., Henderson, G. B., Fairlamb, A. H., & Schirmer, R. H. (1987) *Eur. J. Biochem. 164*, 123–128.

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

Ohtsuka, E., Matsuki, S., Ikehara, M., Takahashi, Y., & Matsubara, K. (1985) *J. Biol. Chem. 260*, 2605–2608.

Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463–5467.

Schirmer, R. H., & Schulz, G. E. (1987) in *Pyridine Nucleotide Coenzymes: Chemical, Biochemical, and Medical Aspects* (Dolphin, D., Poulson, R., & Avramovic, O., Eds.) Vol. 2B, pp 333–379, Wiley, New York.

Schirmer, R. H., Schollhammer, T., Eisenbrand, G., & Krauth-Seigel, R. L. (1987) *Free Radical Res. Commun. 3*, 3–12.

Shames, S. L., Fairlamb, A. H., Cerami, A., & Walsh, C. T. (1986) *Biochemistry 25*, 3519–3526.

Staden, R. (1982a) *Nucleic Acids Res. 10*, 4731–4751.

Staden, R. (1982b) *Nucleic Acids Res. 10*, 2951–2961.

Staden, R., & McLachlan, A. D. (1982) *Nucleic Acids Res. 10*, 141–156.

Takahashi, Y., Kato, K., Hayashizaki, Y., Wakabayashi, T., Ohtsuka, E., Matsuki, S., Ikehara, M., & Matsubara, K. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 1931–1935.

Thieme, R., Pai, E. F., Schirmer, R. H., & Schulz, G. E. (1981) *J. Mol. Biol. 152*, 763–782.

Wierenga, R. K., Terpstra, P., & Wim, G. J. H. (1986) *J. Mol. Biol. 187*, 101–107.

Estradiol-Stimulated Nuclear Ribonucleoprotein Transport in the Rat Uterus: A Molecular Basis†

Raghava Varman Thampan

School of Life Sciences, University of Hyderabad, Hyderabad 500 134, India

Received July 15, 1987; Revised Manuscript Received January 7, 1988

ABSTRACT: The present investigation probes the intranuclear molecular changes that serve to link the nuclear binding of estradiol with the hormone-stimulated ribonucleoprotein (RNP) transport in the rat uterus. Within 2 min of in vitro exposure of isolated uterine nuclei to 10 nM 17β-estradiol a Mg^{2+}-dependent nuclear ATPase becomes activated and reaches its peak activity. This is immediately followed by a phase of ATP resynthesis. This newly synthesized ATP serves as the substrate for the nuclear protein kinases. Cyclic AMP inhibits this ATP resynthesis and, as a consequence, prevents the estradiol-stimulated nuclear protein kinase activity and the exit of the RNP–estradiol complex from the nuclei. cGMP is stimulatory to the estradiol-mediated nuclear ribonucleoprotein transport.

17β-Estradiol (E_2) has a direct stimulatory influence upon ribonucleoprotein (RNP) transport in the rabbit and rat uteri (Vazquez-Nin et al., 1978, 1979; Thampan, 1985). Exposure of uterine nuclei either in vivo or in vitro to physiological concentrations of estradiol results in hormone binding to nuclear RNP and an immediate release of the hormone–RNP complex from the nuclei. It has been found that under in vivo conditions hormone withdrawal from the system results in retention of the RNP within the nuclei. The hormone binding to the nuclear RNP brings about the release of the RNP from the nuclei. The estradiol–RNP complexes thus released are

† This investigation was supported by a Re-entry Grant Award (Project No. 82051) from the World Health Organization.

0006-2960/88/0427-5019$01.50/0
