

Cloning and sequencing of a human thioredoxin reductase

Pamela Y. Gasdaska, John R. Gasdaska, Shawn Cochran, Garth Powis*

Arizona Cancer Center, 1515 N. Campbell Avenue, Tucson, AZ 85724, USA

Received 21 August 1995

Abstract The DNA sequence encoding human placental thioredoxin reductase has been determined. Of the 3826 base pairs sequenced, 1650 base pairs were in an open reading frame encoding a mature protein with 495 amino acids and a calculated molecular mass of 54,171. Sequence analysis showed strong similarity to glutathione reductases and other NADPH-dependent reductases. Human thioredoxin reductase contains the redox-active cysteines in the putative FAD binding domain and has a dimer interface domain not previously seen with prokaryote and lower eukaryote thioredoxin reductases.

Key words: Thioredoxin reductase; Human placenta

1. Introduction

Regulation of the intracellular redox environment is a feature of all cells. A redox enzyme that plays an important role in cell proliferation is thioredoxin reductase (EC 1.6.4.5). It is a member of the pyridine nucleotide-disulfide oxidoreductase family that includes glutathione reductase, lipoamide dehydrogenase, mercuric ion reductase and NADH peroxidase [1]. Members of the family are homodimeric proteins, each subunit of which has a redox active disulfide site and a tightly, but non-covalently bound FAD group that mediates the transfer of reducing equivalents from NADPH to a disulfide bond of the enzyme, and then to the disulfide bond of the substrate.

Physiological substrates for thioredoxin reductase are the redox proteins thioredoxin [2] and protein disulfide isomerase [3]. There may be other substrates but they remain to be identified [4]. Thioredoxin reductase has diverse functions in the cell. Through thioredoxin it provides reducing equivalents for ribonucleotide reductase, the first unique step in DNA synthesis [5], for methionine sulfoxide reductase [6], and for vitamin K epoxide reductase [7]. Thioredoxin also catalyzes protein folding [3] and exerts specific redox control of some transcription factors regulated in this way include NF- κ B [8,9], TFIIIC [10], BZLF1 [11], and the glucocorticoid receptor [12]. The transcription factor AP-1 (Fos/Jun heterodimer) is subject to redox control by the nuclear redox factor Ref-1 which, in turn, is reduced by thioredoxin [13]. Thioredoxin has also been found to stimulate the growth of a variety of normal and cancer cell lines in culture [14–17]. The redox activity of thioredoxin is essential for its cell growth stimulating activity [16,17] and thioredoxin may be reduced by a thioredoxin reductase on the surface of cells [17].

Because of its role in cell proliferation, mammalian thiore-

doxin reductase is a potential target for the development of drugs to control abnormal cell proliferation. It is known that some antitumor quinone drugs [18], nitrosoureas [19], and the cell differentiating agent 13-*cis*-retinoic acid [20] are mechanism-based inhibitors of mammalian thioredoxin reductase which may contribute to their activity.

E. coli thioredoxin reductase has been cloned and sequenced [21,22] and its biochemical and physical properties extensively studied [23,24]. Eukaryotic thioredoxin reductases have so far been only cloned from *Penicillium chrysogenum* [25], *Saccharomyces cerevisiae* [26], and *Arabidopsis thaliana* [27] and they show 44–50% sequence identity to the bacterial enzyme. We now report the cloning and sequencing of a putative thioredoxin reductase from human placenta.

2. Materials and methods

Human thioredoxin reductase was purified to homogeneity from human placenta as we have previously described [28] and amino terminal and internal amino acid sequences of the tryptic digest obtained (W.M. Keck Foundation, Biotechnology Resource Laboratory, New Haven, CT) as follows: *N-terminal sequence*, Gly-Pro-Glu-Asp-Leu-Pro-Lys-Ser/Lys-Tyr; *internal sequences*, Phe-Leu-Ile-Ala-Thr-Gly-Glu-Arg-Pro, and Val-Val-Gly-Phe-His-Val-Leu-Gly-Pro-Asn-Ala-Gly-Glu-Val-Thr-Gln-Gly-Phe-Ala-Ala-Ala-Leu-Lys. The N-terminal sequence was confirmed by a separate analysis (Dr. Ronald Niece, University of Wisconsin, Madison, WI). A degenerate oligonucleotide, GTN GTN GGN TTY CAY GTN CTN GGN CCN AAY GCN GGN GAR GTN ACN CAR GGN TTY GC (N = A/G/C/T, Y = C/T and R = A/G) (TR5, Bio-Synthesis Labs, Lewisville, TX), was used to screen a λ gtII human placenta 5'-stretch cDNA library (Clontech, Palo Alto, CA). Approximately 2×10^5 plaques were hybridized in 0.5 M sodium phosphate, pH 7.2, with 10 mM EDTA, 7% sodium dodecyl sulfate and 1% bovine serum albumin at 50°C for 72 h with 4 ng/ml TR5 end-labelled with [γ - 32 P]dATP using T4 polynucleotide kinase according to manufacturers instructions (Boehringer-Mannheim, Indianapolis, IN). λ DNA was purified using Prep-Eze columns (5 Prime to 3 Prime Inc., Boulder, CO) and amplified by PCR (25 cycles, 94°C for 1 min, 60°C for 1 min and 72°C for 1 min; 1 cycle of 72°C for 7 min) using λ gtII insert screening amplimers (Clontech). cDNA fragments were purified by Qiaex (Qiagen, Chatsworth, CA), digested with *Eco*RI and cloned into the *Eco*RI site of Bluescript (Stratagene, La Jolla, CA). cDNA was manually sequenced on both strands using Sequenase Version 2.0 T7 DNA polymerase (US Biochemicals, Cleveland, OH). Sequence information was compiled and analyzed using the algorithms available through GCG (Genetics Computer Group Inc., Madison, WI). DNA and protein databases were searched using the computer programs FASTA and BLASTP performed at the NCBI (National Center for Biotechnology Information) using the BLAST Network Service (GCG). Gapped sequence alignments and identity/similarity comparisons were made using the computer programs PILEUP and GAP (GCG).

The measurement of thioredoxin reductase activity was by a modification of the method of Holmgren [29] that used the thioredoxin-dependent reduction of insulin with 5,5'-dithiobis-2-nitrobenzoic acid as the post-reaction redox chromophore, as previously described [18]. Antibodies were raised in rabbits using the synthetic peptide Val-Val-Gly-Phe-His-Val-Leu-Gly-Pro-Asn-Ala-Gly-Glu-Val-Thr-Gln-Gly-Phe-Ala-Ala (Macromolecular Structure Facility, University of Arizona, Tucson, AZ) derived from protein sequencing of the native en-

*Corresponding author. Fax: (1) (520) 626-4848.

zyme. Western blots were visualized using affinity purified goat anti-rabbit IgG (H+L) alkaline phosphatase conjugate (BioRad).

2.1. Expression in *E. coli*

The QIAexpressionist system (Qiagen Inc., Chatsworth, CA) was used to add 6 histidine residues (6 × His) to the N-terminal end of the expressed protein. A PCR product was generated using cDNA fragment 30B as a template and the oligonucleotide primers (Bio-Synthesis Labs) GCGGATCCGATGACGATGACAAAGGCCCTGAAGATCTTCCCAAG (which includes a *Bam*HI site, an enterokinase cleavage site and thioredoxin reductase sequence from base 446 to 466) and GCGTCGACCTACCACATGGGCTTGAGAC (which includes a *Sal*I site and thioredoxin reductase sequence from 2215 to 2196). The PCR product was ligated into the pQE-30 vector (Qiagen) using *Bam*HI and *Sal*I restriction sites. *E. coli* strain M15[pRep4] was transformed with the thioredoxin reductase/pQE-30 construct. The 6xHis tagged protein was expressed after induction with isopropyl- β -D-thiogalactoside (IPTG) and purified according to the manufacturer's protocol. The histidine tags were removed with enterokinase according to the manufacturer's instructions (Boehringer-Mannheim, Indianapolis, IN). Expression of thioredoxin reductase as a fusion protein with glutathione S-transferase was accomplished using the GST Gene Fusion System (Pharmacia Biotech, Uppsala, Sweden). A PCR product was generated using clone 30B as a template and oligonucleotide primers (Bio-Synthesis Labs) GCGAATTCCTTATCAGGAGGGCAGACTTC (which includes an *Eco*RI site and thioredoxin reductase sequence from base 405 to 425) and GTCGGCCGCTACCACATGGGCTTGAGAC (which includes a *Not*I site and thioredoxin reductase sequence from base 2196 to 2215). The PCR product was ligated into the pGEX-4T-2 vector (Pharmacia Biotech) using *Eco*RI and *Not*I restriction sites. *E. coli* strain JM105 was transformed with the thioredoxin reductase/pGEX-4T-2 construct. The fusion product was expressed after induction with IPTG. Glutathione S-transferase was removed by treatment with thrombin and purified according to the manufacturer's instructions (Pharmacia Biotech).

3. Results

3.1. Cloning and sequence analysis

Hybridization of a λ gtII human placenta cDNA library with the degenerate oligonucleotide yielded 4 positive plaques, 22, 30A, 30B and 30C. cDNA inserts from 22 and 30C were found by sequence analysis to be identical 1.6 kb fragments. The other two cDNA inserts, 30A and 30B, with sizes of 2.4 kb and 3.7 kb, overlapped. Clone 30A contained the reported sequence from base 1450 to 3826, and 30B contained the sequence from base 1 to 3695. Fig. 1 is a consensus of the results from 4 separate sequencing experiments. This shows that human thioredoxin reductase cDNA has 3826 bases with the longest open reading frame starting with the ATG start codon at base 284. An additional ATG start codon in the same reading frame is located at position 440. Base pair 440 was selected as the translational start-site depicted in Fig. 1 on the basis of its proximity to the N-terminal sequence of the mature protein. There is also a large AT-rich, 3'-untranslated region (3'-UTR) which includes a consensus polyadenylation sequence (AATAAA) immediately preceding the poly(A) tail. The predicted amino acid sequence of thioredoxin reductase contains the N-terminal amino acid sequence, minus the initiating methionine and the first aspartate, as well as the 2 internal amino acid sequences identified from the purified enzyme. The predicted amino acid sequence of thioredoxin reductase gives a protein with 495 amino acids and a molecular weight of 54,171 kDa which is slightly smaller than the experimentally derived molecular weights previously described [28]. Examina-

Table 1
Comparison of human thioredoxin reductase with other proteins

Protein	Identity (%)	Similarity (%)	High score	Reference
Glutathione reductases				
<i>Caenorhabditis elegans</i> (probable)	44	64	322	[35]
<i>E. coli</i>	40	60	139	[36]
Human	35	57	178	[37,38]
<i>Pisum sativum</i>	38	63	148	[39]
<i>Saccharomyces cerevisiae</i>	37	59	138	[40]
<i>Burkholderia cepacia</i>	38	62	145	[41]
<i>Mus musculus</i>	36	58	177	[37,42]
<i>Haemophilus influenzae</i>	40	30	191	[43]
<i>Streptococcus thermophilus</i>	41	64	148	[44]
<i>Glycine max</i>	38	61	242	[45]
<i>Spinacia oleracea</i>	35	54	235	[46]
Thioredoxin reductases				
<i>E. coli</i>	24	53	65	[22]
<i>Streptomyces clavuligerus</i>	23	48	46	[47]
<i>Arabidopsis thaliana</i>	22	46	61	[27]
<i>Eubacterium acidaminophilum</i>	31	51	77	[48]
<i>Penicillium chrysogenum</i>	26	48	ND	[25]
<i>Saccharomyces cerevisiae</i>	23	48	ND	[26]
Other				
Human dihydrolipoamide dehydrogenase	30	53	106	[49]
<i>Trypanosoma cruzi</i> trypanothione reductase	36	56	138	[50]
<i>Pseudomonas aeruginosa</i> mercuric reductase	29	52	89	[51]

Percentages are derived from the computer program GAP using the algorithm of Needleman and Wunsch [34]. High score refers to the maximal segment pair score as determined by the computer program BLASTP using the basic local alignment search tool of Altschul et al. [33]. ND represents scores not determined.

tion of the sequence showed no consensus N-linked glycosylation sites (N-X-S, N-X-T).

The cDNA and deduced amino acid sequence was used to search sequence databases. Several of the most significant sequence similarities are given in Table 1. Human thioredoxin

(1)	gaattcgggtggagctcctgaaggagggcctgagctcttctcattctcctcaaatctcttgtaagctctgctgctgggtg	(75)	
(76)	aaaccagacaaagccgcagccagggatggagcagcgcggggagcgcctcgcggcgagcagcagcagcattcgc	(150)	
(151)	cctgggtcagcagctgctgctcctggagagggatatttaagcgtgctcgcagcagcagcagcagcagcattc	(225)	
(226)	aaaccagacaaagctcctgctgctgctgctcctcagccgcctcagcagcagcagcagcagcagcagcagc	(300)	
(301)	ctggcgcctgaaggagaaactctcggatctggcgcggagaaacagcagctgctgctgctgagcagcagcagc	(375)	
(376)	gataggcgcctcagctggctcgaaccttgagagcttatcaggaagcagcagcagcagcagcagcagcagc	(450)	
(1)	G P E D L P K S Y D Y D L I I I G G G	(19)	
(446)	GCC CTT GAA GAT CTT CCG AAG TCC TAT GAC TAT GGC AAT CTT AAT GGA GGT GGC	(502)	
(20)	S G G L A A A K E A Q A Q Y G K K V M V	(38)	
(503)	TCA GGA GGT CTG GCT CTT AAG GAG GCA GGC CAA TAT GGC AAG AAG GTG ATG GTC	(559)	
(39)	L D F V T P T F L G T R M G L G G T C	(57)	
(560)	CTG GAC TTT GTC ACT CCC ACC CTT CTT GGA ACT GAG TGG GGT CTT GGA GGA ACA TGT	(616)	
(58)	V N V G C I P C K L M H Q A L L G Q	(76)	
(617)	GTG AAT GTG GGT TGC ATA CTT AAA AAA CTT ATG CAT CAA GCA GCT TTT TTA GGA CAA	(763)	
(77)	A L Q D S R N Y G W K V E E T V K H D	(95)	
(674)	GCC CTG CAA GAC TCT GGA AAT TAT GGA TGG AAA GTC GAG GAG ACA GTT AAG CAT GAT	(730)	
(96)	W D R M I E A A V Q N H I I G S L N W G Y	(114)	
(731)	TGG GAC AGA ATG ATA GAA GCT GTA CAA CAT CAC ATT GGC TCT TTT AAT TGG GGC TAC	(787)	
(115)	R V A L R E K K K V T V E N A Y G Q F I	(133)	
(788)	CGA GTA GCT CTG CGG GAG AAA AAA GTC GTC TAT GAG AAT GCT TAT GGG CAA TTT ATT	(844)	
(134)	G P H R I K A A T N H K K K E K I Y S A	(152)	
(845)	GCT CTT CAC AGG ATT AAG GCA ACA TAA AAT AAA GGC AAA GAA AAA ATT TAT TCA GCA	(901)	
(153)	E S F L I A T G E R P R Y L G I P G D	(171)	
(902)	GAG AGT TTT CTT ACT GCC ACT GGT GAA GCA CCA CTA TAT TTT GGC ATC CTT GGT GAC	(958)	
(172)	K E Y C I S S D F T S L Y C P G K A G	(190)	
(959)	AAA GAA TAC TGC ATC AGC AGT GAT CTT TTT TCC TTT GCT TAC TGC CGG GGT AAG	(1015)	
(191)	T L V V G A S Y V A L E C A G F L A G	(209)	
(1016)	ACC CTT GTT GTT GGA GCA TCC TAT GTC TTT GAG GCT GCT GGA TTT CTT CTT GGT	(1072)	
(210)	I G L G V T V H V R S I L R L G F D Q	(228)	
(1073)	ATT GGT TTA GGC GTC ACT GTT ATG GTT AGG TCC ATT CTT CTT AGA GGA TTT GAC CAG	(1229)	
(229)	D M A N K I G E H M E E N G I K F I R	(247)	
(1130)	GAC ATC GCC AAC AAA ATT GGT GAA CAC ATG GAA GAA CAC GGC ATC AAG TTT ATA AGA	(1186)	
(248)	Q F V P I K V E Q I E A G T P G R L R	(266)	
(1187)	CAG TFC GTA CCA ATT AAA GTT GAA CAA ATT GAA GCA GGC ACA CCA GGC CTA AGA	(1243)	
(267)	V V A Q S T N S E E I E I E G E Y N T A	(285)	
(1244)	GTA GTA GCT CAG TCC ACC AAT AGT GAG GAA ATT GAA GGA GAA TAT AAT ACG GTG	(1300)	
(286)	M L A I G R D A C T C T R K I I G L E A T V G	(304)	
(1301)	ATG CTG GCA ATA GGA AGA GAT GCT TCC ACA GAA AAA ATT GGC TTA GAA ACC GTA GGG	(1357)	
(305)	V K I N E K T G K I P V T D E E Q T N	(323)	
(1358)	GTG AAG ATA AAT AAG ACT GGA AAA ATA CTT CTT ACA GAT GAA GAA CAG AAT AAT	(1414)	
(324)	V P Y I Y A I G D I L E D K V E L T P	(342)	
(1415)	GTG CTT TAC TAT TAT GCT ACT GGC GAT ATA TTT GAG GAT AAG GTG GAG CTT ACC ACA	(1471)	
(343)	V A I Q A G R L L A A Q R L Y A G S T V	(361)	
(1472)	GTT GCA ATC CAG GCA GGA AGA TTT CTG GCT CAG AGG CTC TAT GCA GGT TCC ACT GTC	(1528)	
(362)	K C D Y E N V P T T V F T P L E Y G A	(380)	
(1529)	AAG TGT GAC TAT GAA AAT GTT CCA ACC ACT GTA TTT ACT CTT TTT GAA TAT GGT GCT	(1585)	
(381)	C G G L S E E K A G T V E K P F G E E N I E V	(399)	
(1586)	TGT GGT TCT TCT GAG GAG GAA GCT GTG GAG AAT TTT GGC GAA GAA AAT AAT GAG GTT	(1642)	
(400)	Y H S Y F W P L L A T T G A T T I P S R D N N K	(418)	
(1643)	TAC CAT AGT TAT TTT TGG CCA TTT GAA TGG AGT ATT CCG TCA GAA GAT AAC AAC AAA	(1699)	
(419)	C Y A K I I I C N T K D N E R V V G F H	(437)	
(1700)	TGT TAT GCA AAA ATA ATC TGT AAT ACT AAA GAC AAT GAA CTT GTT GTG GGC TTT CAC	(1756)	
(438)	V L G P N A G E V T Q G F A A A L K X C	(456)	
(1757)	GTA CTG GGT CCA AAT GCT GGA GAA GTT ACA CAA GGC TTT GCA GCT GCG CTT AAA Tgt	(1813)	
(457)	G L T K K K O L D S T I G I H P V C A E	(475)	
(1814)	GGA CTT ACC AAA AAG CAG CTT GAC AGC ACA ATT GGA ATC CAC CTT GTC TGT GCA GAG	(1870)	
(476)	V F T T L S V T K R S G A S I L Q A G	(494)	
(1871)	GTA TTC ACA ATA TTT TCT GTG ACC AAG GGT TCT GGG GCA AGC ATC CTC CAG GCT GGC	(1927)	
(495)	C end	(495)	
(1928)	TGC TGA gtttaagcccaagctgtgagctgtgtgcaaacagcagcagcagcagcagcagcagcagcagc	(2000)	
(2001)	caagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc	(2075)	
(2076)	cctggatctcttggatagagctgtggaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc	(2150)	
(2151)	tgacatttggcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc	(2225)	
(2226)	agcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc	(2300)	
(2301)	tgacatttggcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc	(2375)	
(2376)	agcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc	(2450)	
(2451)	atgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc	(2525)	
(2526)	tcttctctcatatctcccaaacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc	(2600)	
(2601)	tgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc	(2675)	
(2676)	tatagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc	(2750)	
(2751)	tctgctcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt	(2825)	
(2826)	aaagcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt	(2900)	
(2901)	caatcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc	(2975)	
(2976)	taactatcaatagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc	(3050)	
(3051)	tgacatttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt	(3125)	
(3126)	actgcatcattataaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc	(3200)	
(3201)	gctgctcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt	(3275)	
(3276)	aaagcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt	(3350)	
(3351)	gccacatcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc	(3425)	
(3426)	tcacgctctctctcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt	(3500)	
(3501)	atcctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc	(3575)	
(3576)	ggcgttctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctct	(3650)	
(3651)	gaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc	(3725)	
(3726)	catgcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt	(3800)	
(3801)	aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa		

Fig. 1. Nucleotide sequence and deduced amino acid sequence of human thioredoxin reductase. Nucleotide residues are numbered from (1) to (3826) while amino acids are numbered from [–2] to [495] with the ATG triplet encoding the proposed initiating methionine residue numbered [–2]. Residues underlined correspond to sequences confirmed by the peptide sequencing and N-terminal sequencing of the native enzyme isolated from human placenta. The calculated molecular mass of the mature enzyme is 54,171.

		bFl	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+</
--	--	-----	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	-----

Fig. 2. Progressive gapped alignment of related sequences: Hsglu = human glutathione reductase [37,38], Ecglu = *E. coli* glutathione reductase [36], Hstrx = human thioredoxin reductase, Ceglu = *C. elegans* probable glutathione reductase [35], Ectrx = *E. coli* thioredoxin reductase [22], Pctrx = *P. chrysogenum* thioredoxin reductase [25]. Identical residues are indicated by dots (•); + indicates residues conserved among Hsglu, Ecglu, Hstrx and Ceglu, with # indicating their active-site cysteines. % indicates active site cysteines of Ectrx and Pctrx. Underlined residues are probable FAD binding regions ADP (12–42; ADP) and Flavin (322–332; flavin) of Hstrx on the basis of similarity to human glutathione reductase. ► indicates domain boundaries of Hstrx as determined by sequence comparison with Hsglu with F1 and F2 representing the FAD domain, N the NADPH domain and I the interface domain.

reductase is most similar to a deduced protein sequence derived from genomic sequencing of *Caenorhabditis elegans*, identified as a probable glutathione reductase (Fig. 2). Sequence identity to a number of eukaryotic and prokaryotic glutathione reductases ranges from 35% to 44%, and for thioredoxin reductases 23% to 31%. Identity to other pyridine nucleotide-disulfide oxidoreductases was 29% to 36%. Genetic distance measurements between all sequences listed in Table 1 favor a pairwise alignment between human thioredoxin reductase and *C. elegans* glutathione reductase, indicating that the *C. elegans* sequence is likely to be a thioredoxin reductase.

3.2. Protein expression

Expression of the cDNA in *E. coli* as either a glutathione S-transferase fusion protein or with a 6 × His tag yielded a protein product that co-migrated on SDS-PAGE and was immunologically indistinguishable from the native protein (Fig. 3). However, the purified protein had no thioredoxin reductase activity and spectral analysis showed no absorption at 450 nm characteristic of bound FAD.

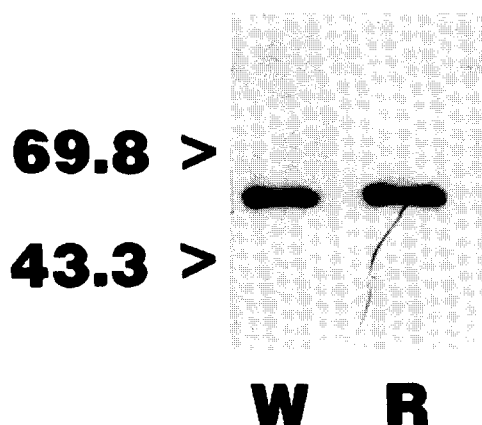


Fig. 3. Western blot of thioredoxin reductase run on an 8% SDS-PAGE gel. W represents pure thioredoxin reductase isolated from human placenta, and RX represents thioredoxin reductase expressed in *E. coli* as a fusion protein with glutathione *S*-transferase (GST) after removal of GST.

4. Discussion

We report the cloning of cDNA for human thioredoxin reductase from a human placenta library. This represents the first sequence available for thioredoxin reductase cDNA from a higher-order eukaryote. The deduced amino acid sequence clearly identifies the protein as a member of the pyridine-nucleotide-disulfide reductase family with a strong resemblance, up to 44% identity, to reported glutathione reductases. Similarity to reported thioredoxin reductases from prokaryotes and lower-order eukaryotes (plants and fungi) is lower, but still significant (Table 1).

It was expected that mammalian thioredoxin reductases, with an estimated subunit molecular weight of 58 kDa for the bovine enzyme [2] and 65 kDa for the human enzyme [28], would have a different domain structure than the smaller 35 kDa bacterial thioredoxin reductases. While structural details can only be estimated from sequence information, comparisons to proteins where structural information exists can be predic-

tive. Extensive structural information derived from X-ray crystal data is available for human glutathione reductase [30,31] and *E. coli* thioredoxin reductase [32]. It is, therefore, possible to compare their sequences with the putative human thioredoxin reductase we have cloned and to infer a structure of the reported sequence. A gapped sequence alignment [33] with selected thioredoxin reductases and glutathione reductases (Fig. 2) delineates the regions of similarity with human thioredoxin reductase.

The domain structure of the putative human thioredoxin reductase is proposed to be similar to glutathione reductase based on sequence comparison (Fig. 4). The features that distinguish *E. coli* thioredoxin reductase from human glutathione reductase are mirrored in the comparison between *E. coli* thioredoxin reductase and human thioredoxin reductase. Most of the size difference between human thioredoxin reductase and the smaller bacterial enzyme can be attributed to the presence of a dimer interface domain. The redox active cysteines of human thioredoxin reductase are located in the FAD domain with a 4-amino acid bridge between cysteines. The active site disulfide of *E. coli* thioredoxin reductase, on the other hand, is part of the NADPH domain with only a 2-amino acid bridge between cysteines. It is reasonable to surmise that the 3-dimensional structure will be very similar to glutathione reductase.

To further study the enzyme, we attempted to express the active enzyme in *E. coli*. While we were able to detect expression of the protein by Western blotting utilizing 2 distinct expression systems, we were unable to detect thioredoxin reductase activity in either case. The protein also had no glutathione reductase activity. The expressed protein co-migrated with the native enzyme on SDS-PAGE as expected but lacked the 450 nm absorption maxima that would be expected if the enzyme contained FAD. The inability to bind FAD may be related to incorrect folding of the protein by *E. coli* expression system since all the structural sequence components for effective FAD binding are present. Further work will pursue expression of the enzyme in other eukaryotic systems in order to provide definitive evidence that the cDNA we have cloned is indeed thioredoxin reductase, and to further study the biochemistry of the enzyme.

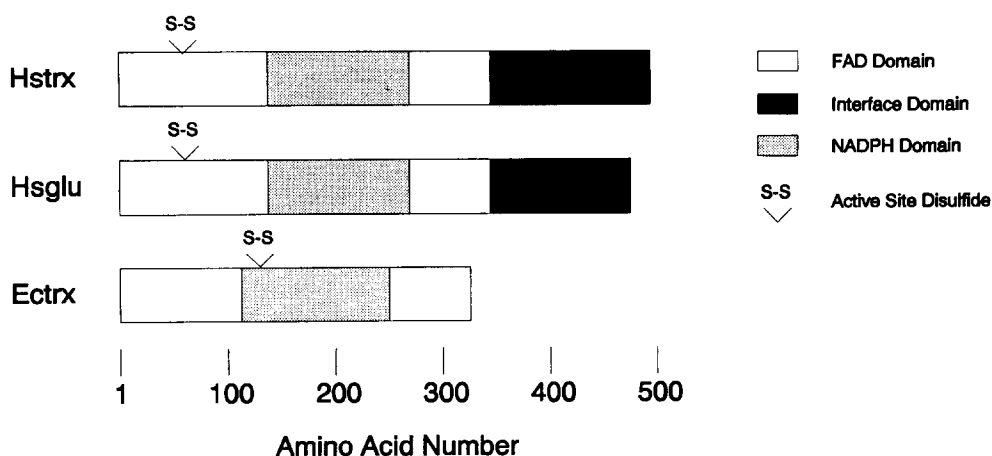


Fig. 4. Proposed domain structure of human thioredoxin reductase (Hstrx) compared to human glutathione reductase (Hsglu) and *E. coli* thioredoxin reductase (Ectrx).

Acknowledgments: Supported by NIH Grant CA48725.

References

- [1] Williams Jr., C.H., Prongay, A.J., Lennon, B.W. and Kuriyan, J. (1991) in: *Flavins and Flavoproteins* (Cutri, B., Zannetti, G. and Ronchi, S. eds.) pp. 497–504, Walter de Gruyter and Co., Berlin.
- [2] Luthman, M. and Holmgren, A. (1982) *Biochemistry* 21, 6628–6633.
- [3] Lundström, J. and Holmgren, A. (1990) *J. Biol. Chem.* 265, 9114–9120.
- [4] Derman, A.I., Prinz, W.A., Belin, D. and Beckwith, J. (1993) *Science* 262, 1744–1747.
- [5] Laurent, T.C., Moore, E.C. and Reichard, P. (1964) *J. Biol. Chem.* 239, 3436–3444.
- [6] Holmgren, A. (1985) *Annu. Rev. Biochem.* 54, 237–271.
- [7] Silverman, R.B. and Nandi, D.L. (1988) *Biochem. Biophys. Res. Commun.* 155, 1248–1254.
- [8] Sorachi, K.-I., Sugie, K., Maekawa, N., Takami, M., Kawabe, T., Kumagai, S., Imura, H. and Yodoi, J. (1992) *Immunobiology* 185, 193–206.
- [9] Hayashi, T., Ueno, Y. and Okamoto, T. (1993) *J. Biol. Chem.* 268, 11380–11388.
- [10] Cromlish, J.A. and Roeder, R.G. (1989) *J. Biol. Chem.* 264, 18100–18109.
- [11] Bannister, A.J., Cook, A. and Kouzarides, T. (1991) *Oncogene* 6, 1243–1250.
- [12] Grippo, J.F., Tienrungraj, W., Dahmer, M.K., Housley, P.R. and Pratt, W.B. (1983) *J. Biol. Chem.* 258, 13658–13664.
- [13] Abate, C., Patel, L., Rauscher III, F.J. and Curran, T. (1990) *Science* 249, 1157–1161.
- [14] Wakasugi, N., Tagaya, Y., Wakasugi, A., Mitsui, M., Maeda, M., Yodoi, J. and Tursz, T. (1990) *Proc. Natl. Acad. Sci. USA* 87, 8282–8286.
- [15] Yodoi, J. and Tursz, T. (1991) *Adv. Cancer Res.* 57, 381–411.
- [16] Oblong, J.E., Berggren, M., Gasdaska, P.Y. and Powis, G. (1994) *J. Biol. Chem.* 269, 11714–11720.
- [17] Gasdaska, J.R., Berggren, M. and Powis, G. (1995) *Cell Growth Differ.* (in press).
- [18] Mau, B.-L. and Powis, G. (1992) *Biochem. Pharmacol.* 43, 1621–1626.
- [19] Schallreuter, K.U., Gleason, F.K. and Wood, J.M. (1990) *Biochim. Biophys. Acta* 1054, 14–20.
- [20] Schallreuter, K.U. and Wood, J.M. (1989) *Biochem. Biophys. Res. Commun.* 160, 573–579.
- [21] Russel, M. and Model, P. (1985) *J. Bacteriol.* 163, 238–242.
- [22] Russel, M. and Model, P. (1988) *J. Biol. Chem.* 263, 9015–9019.
- [23] Kuriyan, J., Krishna, T.S.R., Wong, L., Guenther, B., Pahler, A., Williams, C.H. and Model, P. (1991) *Nature* 352, 172–174.
- [24] Prongay, A.J., Engelke, D.R. and Williams, C.H. Jr. (1989) *J. Biol. Chem.* 264, 2656–2664.
- [25] Cohen, G., Argaman, A., Schreiber, R., Mislovati, M. and Aharonowitz, Y. (1994) *J. Bacteriol.* 176, 973–984.
- [26] Chae, H., Chung, S. and Rhee, S. (1994) *GenBank Accession No.* U10274.
- [27] Jacquot, J.-P., Rivera-Madrid, R., Marinho, P., Kollarova, M., Le Marechal, P., Miginiac-Maslow, M. and Meyer, Y. (1994) *J. Mol. Biol.* 235, 1357–1363.
- [28] Oblong, J.E., Gasdaska, P.Y., Sherrill, K. and Powis, G. (1993) *Biochemistry* 32, 7271–7277.
- [29] Holmgren, A. (1977) *J. Biol. Chem.* 252, 4600–4606.
- [30] Karplus, P.A. and Schulz, G.E. (1989) *J. Mol. Biol.* 210, 163–180.
- [31] Karplus, P.A., Pai, E.F. and Schulz, G.E. (1989) *Eur. J. Biochem.* 178, 693–703.
- [32] Nagasawa, H., Kataoka, H., Isogai, A., Tamura, S., Suzuki, A., Ishizaki, H., Mizoguchi, A. and Fujiwara, Y. (1984) *Science* 226, 1344–1345.
- [33] Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) *J. Mol. Biol.* 215, 403–410.
- [34] Needleman, S.B. and Wunsch, C.D. (1970) *J. Mol. Biol.* 48, 443–453.
- [35] Sulston, J., Du, Z., Thomas, K., Wilson, R., Hillier, L., Staden, R., Halloran, N., Green, P., Thierry-Mieg, J., Qiu, L., Dear, S., Coulson, A., Craxton, M., Durbin, R., Berks, M., Metzstein, M., Hawkins, T., Ainscough, R. and Waterston, R. (1992) *Nature* 356, 37–41.
- [36] Gree, S. and Perham, R.N. (1986) *Biochemistry* 25, 2736–2742.
- [37] Tutic, M., Lu, X., Schirmer, R.H. and Werner, D. (1990) *Eur. J. Biochem.* 188, 523–528.
- [38] Krauth-Siegel, R.L., Blatterspiel, R., Saleh, M., Schiltz, E., Schirmer, R.H. and Untucht-Grau, R. (1982) *Eur. J. Biochem.* 121, 259–267.
- [39] Creissen, G., Edwards, E.A., Enard, C., Wellburn, A. and Mullineaux, P. (1992) *Plant J.* 2, 129–131.
- [40] Collinson, L.P. and Dawes, I.W. (1994) *GenBank Accession No.* L35342.
- [41] Daubaras, D.L., Hershberger, C., Kitano, K. and Chakrabarty, A. (1995) *GenBank Accession No.* U19883.
- [42] Werner, D. (1993) *GenBank Accession No.* X76341.
- [43] Barcak, G.J. and Heimer, S.R. (1995) *GenBank Accession No.* U20229.
- [44] Pebay, M., Holl, A., Simonet, J. and Decaris, B. (1994) *GenBank Accession No.* Z29494.
- [45] Tang, X. and Webb, M.A. (1994) *Plant Physiol.* 104, 1081–1082.
- [46] Aona, M., Fujiyama, K., Sano, T. and Tanaka, K. (1994) *GenBank Accession No.* D37870.
- [47] Cohen, G., Yanko, M., Mislovati, M., Argaman, A., Schreiber, R., Av-Gay, Y. and Aharonowitz, Y. (1993) *J. Bacteriol.* 175, 5159–5167.
- [48] Luebbbers, M. and Andreesen, J.R. (1993) *Eur. J. Biochem.* 217, 791–798.
- [49] Otulakowski, G. and Robinson, B.H. (1987) *J. Biol. Chem.* 262, 17313–17318.
- [50] Sullivan, F.X. and Walsh, C.T. (1991) *Mol. Biochem. Parasitol.* 44, 145148–148.
- [51] Brown, N.L., Ford, S.J., Pridmore, R.D. and Fritzinger, D.C. (1983) *Biochemistry* 22, 4089–4095.