PROTEINS: Structure, Function, and Genetics 11:13–28 (1991)

Structural and Functional Relations Among Thioredoxins of Different Species

Hans Eklund,[1] Florence K. Gleason,[2] and Arne Holmgren[3]

[1]Department of Molecular Biology, Swedish University of Agricultural Sciences, Biomedical Center, S-751 24 Uppsala, Sweden, [2]Department of Plant Biology, University of Minnesota, St. Paul, Minnesota 55108, and [3]Department of Physiological Chemistry, Karolinska Institute, S-104 01 Stockholm, Sweden

ABSTRACT **Three-dimensional models have been constructed of homologous thioredoxins and protein disulfide isomerases based on the high resolution x-ray crystallographic structure of the oxidized form of *Escherichia coli* thioredoxin. The thioredoxins, from archebacteria to humans, have 27–69% sequence identity to *E. coli* thioredoxin. The models indicate that all the proteins have similar three-dimensional structures despite the large variation in amino acid sequences. As expected, residues in the active site region of thioredoxins are highly conserved. These include Asp-26, Ala-29, Trp-31, Cys-32, Gly-33, Pro-34, Cys-35, Asp-61, Pro-76, and Gly-92. Similar residues occur in most protein disulfide isomerase sequences. Most of these residues form the surface around the active site that appears to facilitate interactions with other enzymes.**

Other structurally important residues are also conserved. A proline at position 40 causes a kink in the alpha-2 helix and thus provides the proper position of the active site residues at the amino end of this helix. Pro-76 is important in maintaining the native structure of the molecule. In addition, residues forming the internal contact surfaces between the secondary structural elements are generally unchanged such as Phe-12, Val-25, and Phe-27.

Key words: thiol-dependent redox proteins, redox-active disulfide, sequence homology, three-dimensional structure, molecular modeling, protein domains

INTRODUCTION

Thioredoxin is a small protein (M_r approximately 12,000) found in all living cells from archebacteria to humans. The active site of thioredoxin is conserved and has two redox-active cysteine residues in the sequence: —Trp-Cys-Gly-Pro-Cys—. The oxidized form (thioredoxin-S_2) contains a disulfide bridge that is reduced to a dithiol by NADPH and the flavoprotein, thioredoxin reductase. The reduced form (thioredoxin-$(SH)_2$) is a powerful protein disulfide oxido-reductase. Depending on the chemical environment, thioredoxin can catalyze reduction of a protein disulfide bond or oxidation of a dithiol to a disulfide (for reviews, see references 1, 2, and 3).

Thioredoxin can also serve as a reducing agent for ribonucleotide reductase from *Escherichia coli* and other organisms.[4] Thioredoxin has a similar function in sulfate reduction[5,6] and methionine sulfoxide reduction.[7] In addition, *E. coli* thioredoxin-$(SH)_2$ is an essential subunit of phage T7 DNA polymerase[8] and is also required for the assembly of the filamentous viruses f1 and M13.[9,10] There is a large and growing list of functional roles for thioredoxin in different biological systems involving regulation of protein activity by thiol redox control. An example is the modulation of enzymes in the Calvin cycle by light-promoted reduction of thioredoxin.[11] Specific thioredoxins can then reduce enzymes like fructose-1,6-bisphosphatase (thioredoxin f) and malate dehydrogenase (thioredoxin m).

Thioredoxin has been studied most extensively in *E. coli*. The three-dimensional structure of thioredoxin-S_2 has been determined by x-ray crystallography to 2.8 Å[12] and recently refined to 1.68 Å resolution.[13] Thioredoxin is a highly structured molecule with 90% of its residues involved in secondary structural elements. As illustrated in Figure 1, the molecule has a central core of five strands of β-sheet enclosed by four α-helices. The active site disulfide, formed by residues 32 and 35, is located at the amino end of α2 in a short segment that is separated from the rest of the helix by a kink due to Pro-40. The 14-membered disulfide ring composed of residues Cys_{32}-Gly-Pro-Cys_{35} is located on the surface of the protein but on the shielded side of the α-helix. Cys-32 is exposed to solvent, whereas Cys-35 is recessed and interacts with residues in other parts of the molecule such as Pro-76. The two residues between the cysteines, Gly-33 and Pro-34, have no protruding

Received September 4, 1990; revision accepted January 14, 1991.

Address reprint requests to Dr. Hans Eklund, Department of Molecular Biology, Swedish University of Agricultural Sciences, Biomedical Center, S-751 24 Uppsala, Sweden.

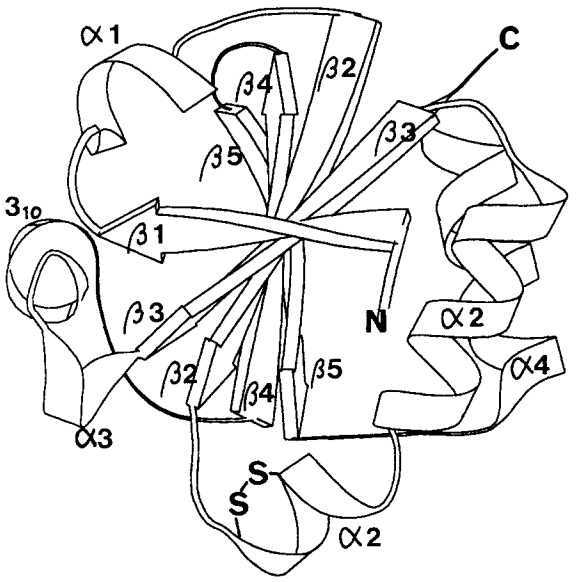


Fig. 1. Schematic drawing of *E. coli* thioredoxin based on the x-ray coordinates.[13,14] Five beta strands (β1-β5), shown as elongated arrows, form the internal core of the protein. Four alpha helices and a short stretch of 3_{10} helix surround the central β-sheet. The disulfide bridge (S—S) is located after β2 and forms the N-terminal portion of α2. Amino (N) and carboxyl (C) ends of the protein are indicated.

side-chains and the surface around the active site is flat. One part of the area around the active site is hydrophobic and has been suggested to be a main interaction site for other proteins.[15,16]

A number of recent reports on thioredoxin have expanded our knowledge about this protein. New structural information based on crystallography[13] and NMR[17,18] is available. In addition to the 10 prokaryotic sequences, several primary structures of mammalian thioredoxins have been determined (see Fig. 2). Mutant proteins have been isolated[9,40,41] or made by site-directed mutagenesis[42–45] and characterized.

In addition to the small thioredoxin protein, a number of mammalian proteins of high molecular weight have been found to contain thioredoxin domains. For example, protein disulfide isomerase (PDI) contains two regions with internal homology that exhibit sequence similarity to thioredoxin.[36] PDI is a substrate for thioredoxin reductase, indicating that the domains are folded and recognized as thioredoxins.[46] Similar structures have been reported in other protein disulfide isomerases[47–49] including the β subunit of proline hydroxylase,[37] and phosphoinositol phospholipase C.[38] Recently a developmentally regulated gene of *Trypanosoma brucei* has been shown to encode a homolog of PDI.[39]

Based on model building, we have summarized and reviewed structural features of thioredoxin and identified essential residues that are conserved for structural or functional reasons. We suggest that these similarities and differences serve as a guide for further investigation of thioredoxin function by site-directed mutagenesis.

METHODS

Primary structures of thioredoxins of different species, determined either by protein sequencing or derived from cDNA sequences, were taken from the literature and are presented in Figure 2. For protein disulfide isomerase and related proteins, only the thioredoxin domains are shown. The sequences were initially aligned with the *E. coli* structure because of similarity mainly in the conserved active site region. Unlike the straightforward sequence alignment in prokaryotic thioredoxins, the alignment of the avian and different mammalian proteins with *E. coli* thioredoxin is more complex. The sequences of the N-terminal third of these proteins before the active site are one residue shorter than in *E. coli* thioredoxin. If these are positioned without gaps, the alignment gives rise to poor three-dimensional structures with internal charged residues. The most plausible tertiary structure is obtained if a deletion of one residue is made in the bend between the β-1 and α-1 elements, and an insertion of one residue is made in the bend between α-1 and β-2 elements. This alignment yields eight identities between the mammalian sequences and *E. coli* thioredoxin and, in addition, places hydrophobic residues on the internal side of the first helix (residues 12 and 16).

The region between α-2 and β-4 (residues 50–75), exhibits little similarity between the eukaryotic and bacterial thioredoxin sequences and is one residue shorter in the eukaryotic proteins. From structural considerations, we have chosen to make this deletion in the bend immediately after the helix α-2, which places hydrophobic residues at internal positions in strand β-3. The last of the deletions must be made between residues 93 and 96 where the mammalian thioredoxins are generally one residue shorter.

Models of the thioredoxins listed in Figure 2 have been constructed from the refined *E. coli* thioredoxin structure using computer graphics. The residues that differ from the *E. coli* sequence have been replaced using FRODO[50,51] and modeled as closely as possible to the corresponding residues in the *E. coli* thioredoxin structure. Structurally favorable rotamers were chosen.[52] Coordinates for *E. coli* thioredoxin to 1.68 Å have been deposited in the Protein Data Bank (Brookhaven National Laboratory, Upton, NY).[13] (Coordinates of the thioredoxin models are available from H. Eklund on request.)

RESULTS AND DISCUSSION

Sequence Alignment

By using the sequence similarity in the active site region, a preliminary alignment of various thiore-

TABLE I. Homology Between Thioredoxins From Different Species (In Percent Identical Residues)*

SPECIES	*E. coli* (108)	*C. nephridii* (105)	*C. nephridii* C2 (107)	*Anabaena* 7119 (106)	*A. nidulans* (107)	*Anabaena* 7120 (110)	*R. sphaeroides* (105)	*Rs. rubrum* (104)	Chlorobium (108)	Chromatium (107)	Spinach m (114)	Spinach f (113)	Chicken (104)	Rabbit (104)	Mouse (105)	Calf thymus (104)	Human (104)
E. coli	100	50	48	48	47	43	47	56	44	67	46	30	27	27	26	27	27
C. nephridii		100	51	50	52	29	56	52	45	52	50	30	30	32	30	32	30
C. nephridii C2			100	44	43	35	50	40	43	50	41	30	29	33	31	35	34
Anabaena 7119				100	84	39	52	51	51	52	57	30	26	29	26	29	30
A. nidulans R2					100	42	50	50	47	52	54	31	27	28	27	30	31
Anabaena 7120						100	35	37	30	39	32	23	21	21	21	22	22
R. sphaeroides							100	54	46	52	43	31	31	32	30	32	31
Rs. rubrum								100	40	52	43	31	31	28	26	28	30
Chlorobium									100	44	41	29	28	29	30	30	30
Chromatium										100	46	31	26	25	24	25	26
Spinach m											100	29	25	27	26	26	26
Spinach f												100	29	33	32	33	35
Chicken													100	68	69	69	68
Rabbit														100	79	88	87
Mouse															100	84	83
Calf thymus																100	90
Human																	100

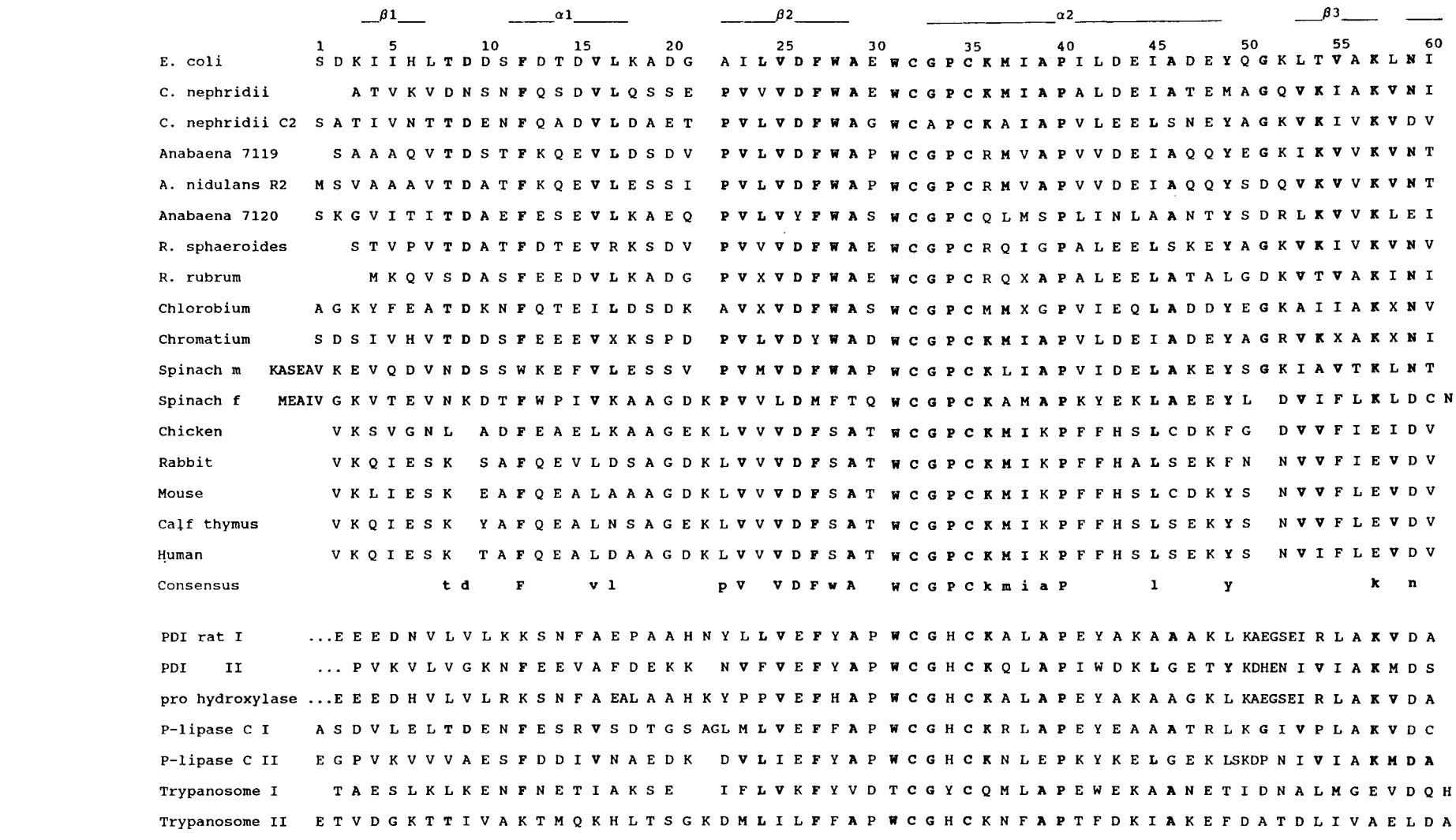
*Sequence alignments were made from the active site disulfide as shown in Figure 2. Only positive matches are scored. No penalties were assigned to gaps or loops. Percentages were calculated based on the number of residues (in parentheses) in the thioredoxin listed in the top line.

doxins can be made as seen in Figure 2. The chain length is approximately the same for the thioredoxins of all species, but there are differences of a few residues at the amino and carboxyl ends. About half of the polypeptide chains are shorter (1–3 residues) than that of *E. coli* thioredoxin at the amino end. Only the thioredoxins m and f from spinach chloroplasts have longer N-terminal sequences, perhaps retaining a portion of a signal peptide required for transport into the chloroplast.[30] Most sequences are one residue shorter than the *E. coli* thioredoxin sequence at the carboxyl end. The primary structure of spinach thioredoxin m was reported to be nine residues shorter than that of *E. coli* thioredoxin. However, in a revised structure based on the gene sequence, it has the same length.[29]

Some species have two or more thioredoxins with major differences in the primary structure.[24,30] An extreme example is spinach chloroplast thioredoxins m and f, which exhibit only 29% sequence identity (Table I). As noted previously by Schürmann and coworkers,[53] this may represent an example of gene duplication and evolutionary drift toward specificity for activation of fructose-1,6-bisphophatase by thioredoxin f and interaction of thioredoxin m with other proteins. The different mechanisms for reducing thioredoxin, either NADPH and thioredoxin reductase or ferredoxin and ferredoxin-thioredoxin reductase, are another example of functional divergence.[2]

Approximately 50% amino acid identity occurs among prokaryotic thioredoxins (Table I). However, there are some notable exceptions. The cyanobacterial thioredoxins from *Anabaena* sp. 7119 and *Anacystis nidulans* R2 exhibit 84% homology, reflecting the close phylogenetic relationship of the organisms. Most residues in the active site region are conserved. Models of both proteins show that residue substitutions occur primarily on the exterior of the protein and in an area distant from the disulfide bridge. In contrast, the unusual *Anabaena* sp strain 7120 thioredoxin has only 39% amino acid identity to the protein from *Anabaena* 7119. Since two thioredoxins occur in *Anabaena* 7119 and 7120, the situation in filamentous cyanobacteria may be analogous to that in chloroplasts where two dissimilar thioredoxins have evolved for specialized purposes.[54] Thioredoxins from presumably related purple photosynthetic bacteria such as *Rhodobacter sphaeroides* and *Rhodospirillum rubrum* show little additional homology with each other or any of the other thioredoxins from photosynthetic organisms. Perhaps phylogenetic schemes based on various physiological criteria are not very accurate or some lines have diverged and evolved more rapidly than others. In any case, it seems that evolutionary comparisons of thioredoxin sequences must be interpreted with caution and provide little information on evolution of species.

Both the low frequency of gaps in the alignment of


β1
α1
β2
α2
β3
1 5 10 15 20 25 30 35 40 45 50 55 60
E. coli SDKIIHLTDDSFDTDVLKADG AILVDFWAEWCGPCKMIAPILDEIADEYQGKLTVAKLNI
C. nephridii ATVKVDNSNFQSDVLQSSE PVVVDFWAEWCGPCKMIAPALDEIATEMAGQVKIAKVNI
C. nephridii C2 SATIVNTTDENFQADVLDAET PVLVDFWAGWCAPCKAIAPVLEELSNEYAGKVKIVKVDV
Anabaena 7119 SAAAQVTDSTFKQEVLDSDV PVLVDFWAPWCGPCRMVAPVVDEIAQQYEGKIKVVKVNT
A. nidulans R2 MSVAAAVTDATFKQEVLESSI PVLVDFWAPWCGPCRMVAPVVDEIAQQYSDQVKVVKVNT
Anabaena 7120 SKGVITITDAEFESEVLKAEQ PVLVYFWASWCGPCQLMSPLINLAANTYSDRLKVVKLEI
R. sphaeroides STVPVTDATFDTEVRKSDV PVVVDFWAEWCGPCRQIGPALEELSKEYAGKVKIVKVNV
R. rubrum MKQVSDASFEEDVLKADG PVXVDFWAEWCGPCRQXAPALEELATALGDKVTVAKINI
Chlorobium AGKYFEATDKNFQTEILDSDK AVXVDFWASWCGPCMMXGPVIEQLADDYEGKAIIAKXNV
Chromatium SDSIVHVTDDSFEEEVXKSPD PVLVDYWADWCGPCKMIAPVLDEIADEYAGRVKXAKXNI
Spinach m KASEAV KEVQDVNDSSWKEFVLESSV PVMVDFWAPWCGPCKLIAPVIDELAKEYSGKIAVTKLNT
Spinach f MEAIV GKVTEVNKDTFWPIVKAAGDKPVVLDMFTQWCGPCKAMAPKYEKLAEEYL DVIFLKLDCN
Chicken VKSVGNL ADFEAELKAAGEKLVVVDFSATWCGPCKMIKPFFHSLCDKFG DVVFIEIDV
Rabbit VKQIESK SAFQEVLDSAGDKLVVVDFSATWCGPCKMIKPFFHALSEKFN NVVFIEVDV
Mouse VKLIESK EAFQEALAAAGDKLVVVDFSATWCGPCKMIKPFFHSLCDKYS NVVFLEVDV
Calf thymus VKQIESK YAFQEALNSAGEKLVVVDFSATWCGPCKMIKPFFHSLSEKYS NVVFLEVDV
Human VKQIESK TAFQEALDAAGDKLVVVDFSATWCGPCKMIKPFFHSLSEKYS NVIFLEVDV
Consensus td F vl pV VDFwA WCGPCkmiaP l y k n
PDI rat I ...EEEDNVLVLKKSNFAEPAAHNYLLVEFYAPWCGHCKALAPEYAKAAAKLKAEGSEIRLAKVDA
PDI II ... PVKVLVGKNFEEVAFDEKK NVFVEFYAPWCGHCKQLAPIWDKLGETYKDHENIVIAKMDS
pro hydroxylase ...EEEDHVLVLRKSNFAEALAAHKYPPVEFHAPWCGHCKALAPEYAKAAGKLKAEGSEIRLAKVDA
P-lipase C I ASDVLELTDENFESRVSDTGSAGLMLVEFFAPWCGHCKRLAPEYEAAATRLKGIVPLAKVDC
P-lipase C II EGPVKVVVAESFDDIVNAEDK DVLIEFYAPWCGHCKNLEPKYKELGEKLSKDPNIVIAKMDA
Trypanosome I TAESLKLKENFNETIAKSE IFLVKFYVDTCGYCQMLAPEWEKAANETIDNALMGEVDQH
Trypanosome II ETVDGKTTIVAKTMQKHLTSGKDMLILFFAPWCGHCKNFAPTFDKIAKEFDATDLIVAELDA

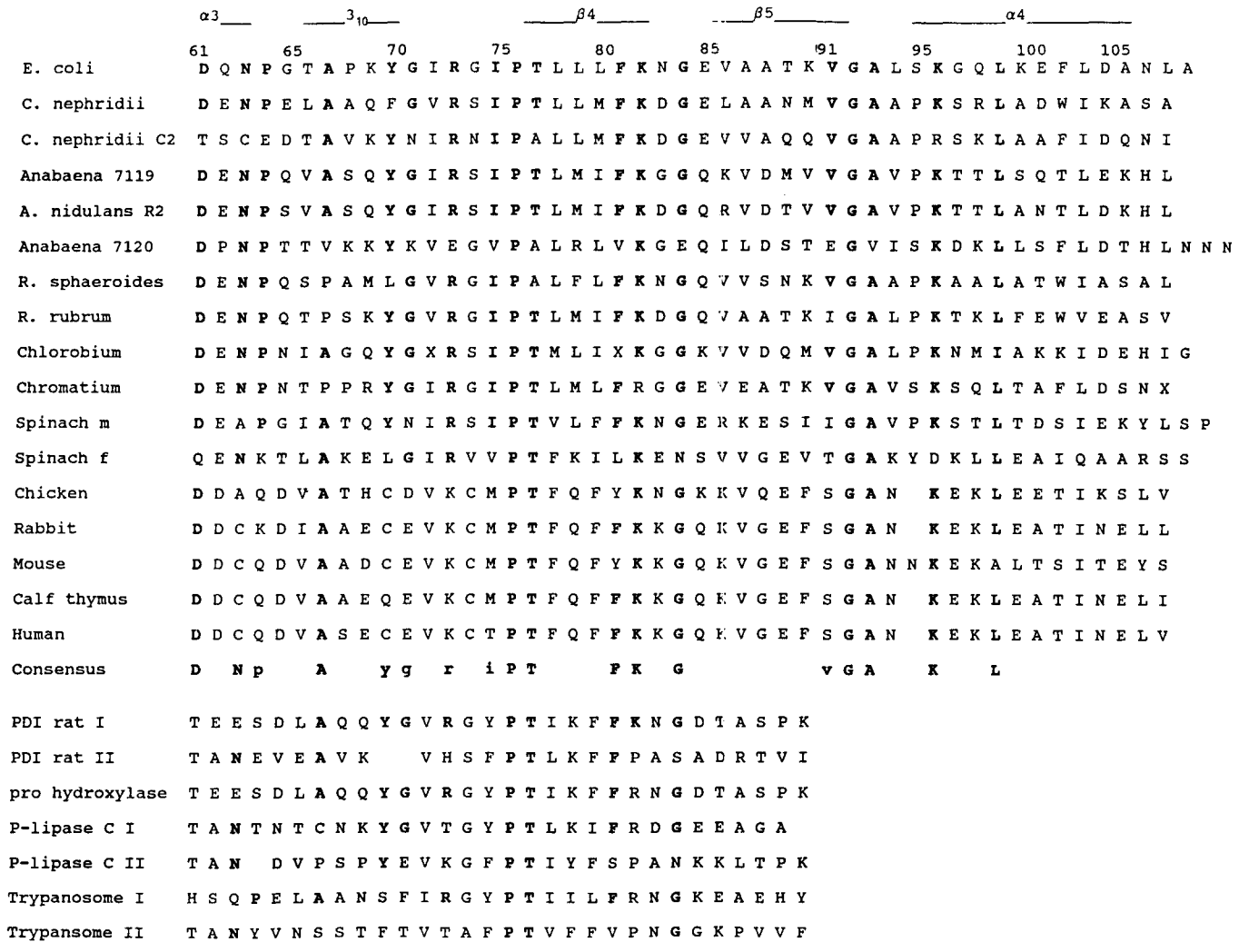


Fig. 2. Primary structures of thioredoxins and related protein disulfide isomerases. The amino acid sequences are aligned from the redox active cysteine residues. Numbers refer to the *E. coli* sequence. Structural elements (sheet or helix) are indicated across the top. Thioredoxin sequence data were taken from the following sources: *E. coli*,[19] *Corynebacterium nephridii*,[20] *Corynebacterium nephridii* thioredoxin C2,[21] *Anabaena* sp. 7119,[22] *Anacystis nidulans* R2,[23] *Anabaena* sp. 7120,[24] *Rhodobacter sphaeroides*,[25] *Rhodospirillum rubrum*,[26] *Chlorobium thiosulfatophilum*,[27] *Chromatium vinosum*,[28] spinach chloroplast thioredoxin m,[29] spinach chloroplast thioredoxin f,[30] chicken,[31] rabbit,[32] mouse,[33] calf thymus,[34] human.[33,35] Partial sequences are shown for the protein disulfide isomerases, below the concensus line. These sequences were obtained from: rat I and II PDI domains,[36] proline hydroxylase, β-subunit,[37] phospholipase I and II,[38] trypanosome isomerase.[39] Standard one letter abbreviations are used for the amino acids. X denotes either leu or ile. Rationale used for placement of gaps and loops is given in the Methods section.

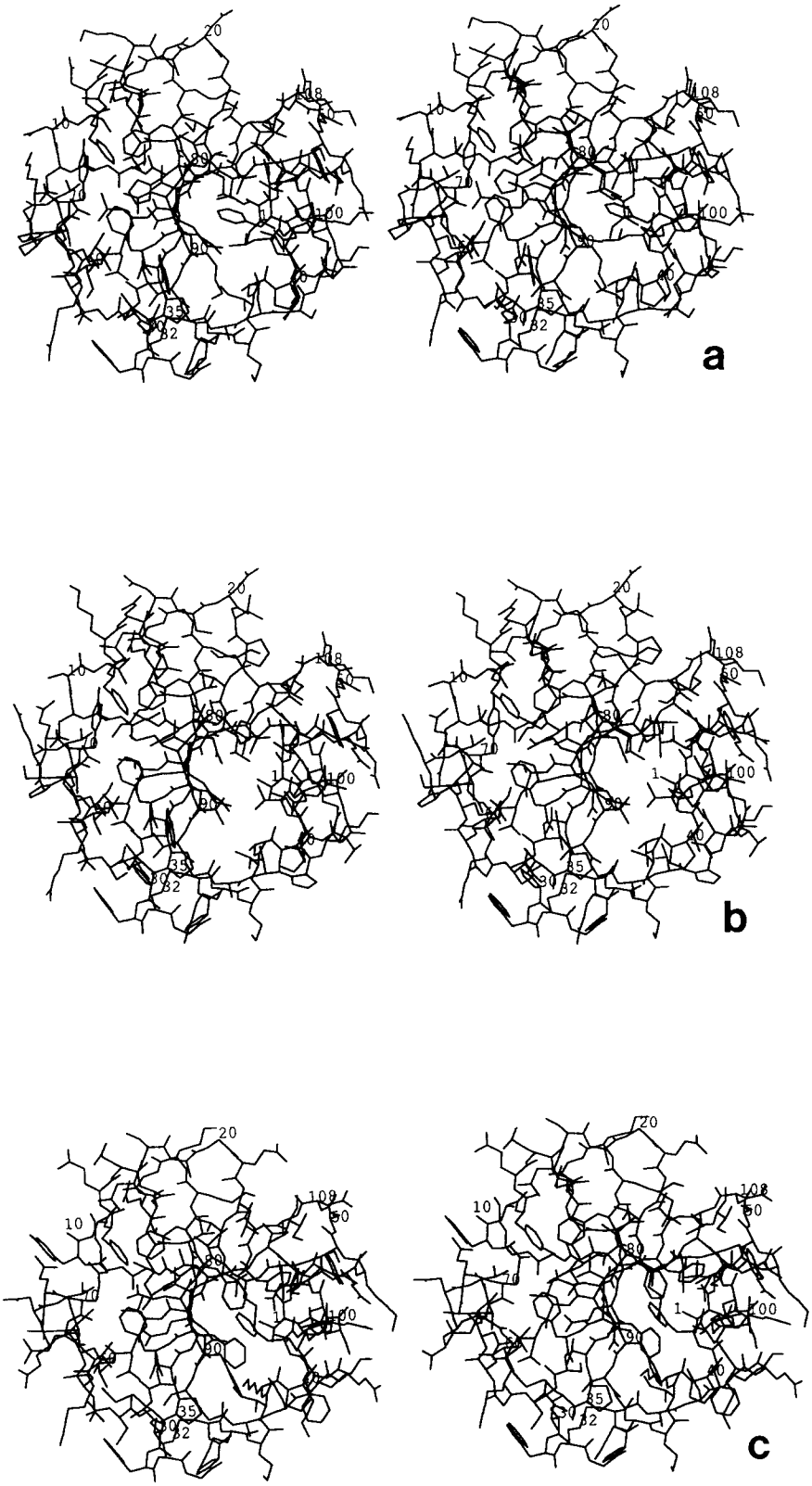


Fig. 3 a–c. Legend appears on page 19.

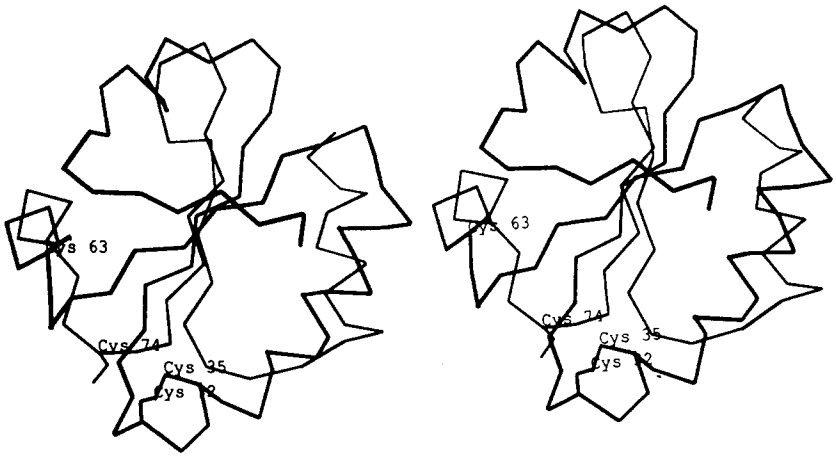


Fig. 4. Stereo drawing of a model of the backbone structure of calf thymus thioredoxin. Only the positions of the cysteine residues are indicated. Unlike bacterial thioredoxins, the mammalian proteins have additional cysteine residues at positions 63 and 74. As seen in the model, these are situated on opposite "sides" of the molecule and cannot form a disulfide bond. The model is supported by structural information from NMR, which also indicates that the molecules are more likely to form intermolecular disulfide bonds under conditions of high protein concentration.[18]

the thioredoxin family and the fact that these gaps are only one residue long are unusual. Most protein families usually have a common structural core but exhibit large variations in loop length.[55] Well-known examples are the serine protease family,[56] the cytochrome c,[57] and the alcohol dehydrogenase families.[58] The glutaredoxins, which catalyze glutathione-disulfide oxidoreductions, also exhibit differences in chain length with loops containing variable numbers of residues.[59–62] The glutaredoxin family is essentially unrelated to the thioredoxin family in primary structure, but may have a similar tertiary structure.[16] It includes one "thioredoxin" from bacteriophage T4, which in its mechanistic[63] and chemical properties is a glutaredoxin[16,62] and is not discussed in this work.

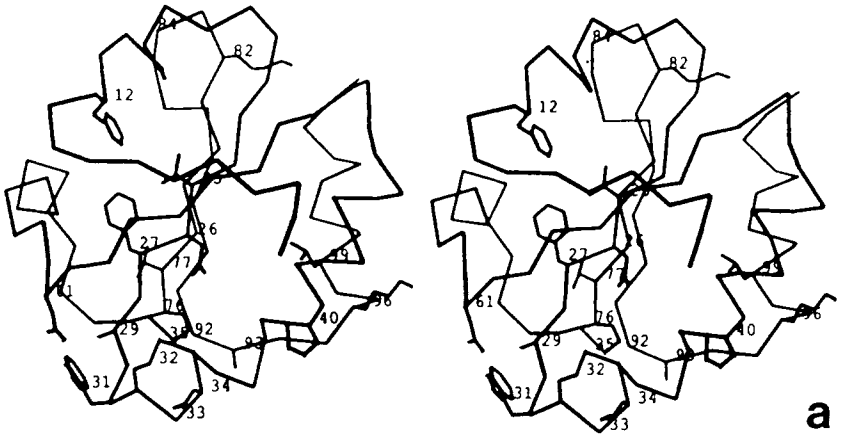
In our comparisons among the thioredoxin family, we have attempted to obtain good three-dimensional structures based on known crystallographic parameters. We predict that thioredoxins and the thioredoxin domains of the protein disulfide isomerase family all have tertiary structures similar to the *E. coli* protein.

Three-Dimensional Structures of the Thioredoxins

Although the amino acid sequences of the thioredoxins from different species show only 26–67% identity to that of the *E. coli* thioredoxin, the derived three-dimensional structures are all similar. The secondary structural elements are retained in all thioredoxins except for the spinach thioredoxins m and f where the extra residues at the amino terminal extend beyond the frame of the thioredoxin secondary structure. Many of the residues that are important for maintaining the tertiary structure and function are conserved (see Table II). The same internal hydrophobic cores are retained, and residue substitutions in these cores are generally conservative. As expected, a variety of substitutions are found at the molecular surface, but these should affect only the local structure of the molecule.

Most of the molecule can be considered to be within the core of the protein as defined by Chothia and Lesk.[55] The deviation of atomic positions between the molecules should then follow the normal deviation related to sequence similarities. For molecules with an amino acid sequence identity of 50%, the root mean square deviation for their Cα atoms is approximately 1.5 Å; for molecules with a sequence identity of 25%, the corresponding figure is 2.5 Å. These root mean square deviations can be regarded as an estimation of the errors in the present models since they have main chain conformations almost identical to that of *E. coli* thioredoxin.

Fig. 3. **(a)** Stereo drawing based on the crystallographic coordinates for oxidized *E. coli* thioredoxin. The disulfide bridge is near the bottom of the drawing and the location of the cysteines is indicated at 32 and 35. Residues Asp20, Lys 100, and the carboxyl terminal, residue 108, are labeled as additional reference points. **(b)** Stereo drawing of the model of thioredoxin from *Anabaena* 7119. The orientation of the model is the same as the structure in 3a. Residues are numbered as in *E. coli* thioredoxin. The model shows the same secondary structural elements as seen above. Two changes in the active site region of the cyanobacterial protein can be noted. Proline 30 replaces the glutamic acid seen in the *E. coli* protein. An arginine residue at 36, in place of lysine, protrudes farther over the surface in the active site region. **(c)** Stereo drawing of the model of thioredoxin from calf thymus. The orientation of the model is the same as the structure in 3a. Residues are numbered as in *E. coli* thioredoxin. Although the mammalian thioredoxin exhibits less than 30% amino acid identity to the bacterial protein, the same secondary structural elements are obvious. Note the conserved cluster of aromatic amino acids in the core of the protein in both models, which corresponds to the same cluster seen in the crystal structure of *E. coli* thioredoxin.



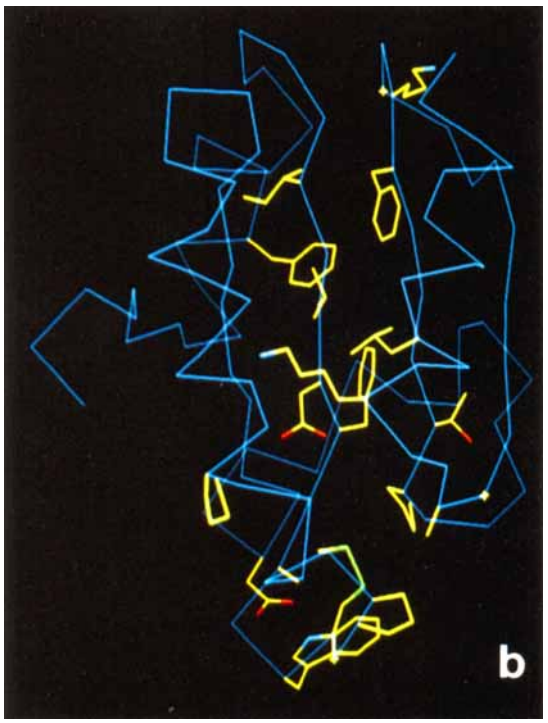


Fig. 5. **(a)** Stereo drawing of *E. coli* thioredoxin based on x-ray data. The orientation is the same as in previous diagrams. Side chains of residues that are conserved in most thioredoxins and related proteins are shown. In addition to residues in the active site region, other conserved residues are in the interior of the protein and maintain the structural integrity of the N and C terminal folding units of the molecule. **(b)** Diagram of the backbone structure of *E. coli* thioredoxin. The alpha carbon backbone is shown in blue. Side chains of conserved residues in thioredoxins from various species are shown in yellow. The active site disulfide is at the bottom. The diagram clearly shows that conserved residues in the interior of the molecule occur between the two folding domains.

The three-dimensional model of *Anabaena* 7119 thioredoxin is shown in Figure 3b as an example of the large group with approximately 50% homology to *E. coli* thioredoxin, shown in Figure 3a. The overall structures are very similar, but a few obvious changes can be noted. Residue substitutions occur close to the active site in *Anabaena* thioredoxin, which are also observed in other thioredoxins such as the m type from spinach chloroplasts (see Fig. 2). One of these is a proline at position 30 in place of glutamic acid and a second, an arginine for lysine substitution, at position 36. These occur on the surface near the active site but opposite the side of the suggested protein interaction area[16] and apparently have little effect on redox activity. *Anabaena* thioredoxin interacts with *E. coli* thioredoxin reductase and ribonucleotide reductase with slightly higher K_m values. Insertion of the *Anabaena* thioredoxin gene in a mutant of *E. coli* lacking thioredoxin restores the wild-type phenotype except for support of T7 phage replication.[64] The substitution of the more rigid proline in position 30 does not alter the main chain conformation. However, substitution of glycine for serine at position 74 does alter the backbone structure. This change, however, is easily accommodated since it is on the surface. Its position in the model was based on the conformation of similar turns found in other proteins.[65] The serine side chain in this position in the *Anabaena, Anacystis,* and spinach m thioredoxins occurs in the suggested enzyme interaction area and may facilitate functions unique to photosynthetic organisms.

Mutagenesis of *E. coli* thioredoxin has shown that substitution of Gly-74 by an Asp results in a protein that is unable to support phage replication. In addition, the mutated protein is not a substrate for thioredoxin reductase[42] (see Table III). About half of the prokaryotic thioredoxins have a glycine in this position; the other half generally have a serine. Glycine-serine mutations seem to be relatively common among related proteins.[66] Since the *Anabaena* thioredoxin also does not support T7 replication but *R. sphaeroides* does,[76] this residue may be critical for interaction with the viral DNA polymerase. However, since the cyanobacterial thioredoxin is still a good substrate for *E. coli* thioredoxin reduc-

TABLE II. Highly Conserved Residues in Thioredoxins of Different Species

Residue	Number*	Possible function
Asp-9	10	Structural, H bonded to main chain N atoms of residues 65 and 66 at the amino end of the 3_{10} helix.
Phe-12	16	Structural, forms the internal hydrophobic section of helix α-1
Val-16	11	As above, often substituted with Leu.
Leu-17	9	As above, but more variable.
Pro-22	10	Structural, forms a bend; is Ala in *E. coli* and Leu in eukaryotes.
Val-23	16	Structural cluster. These three residues form the hydrophobic surface of β-2 in the N-terminal which interacts with β-4 in the C-terminal portion of the molecule.
Leu-24	6	Leu-24 is often substituted by Val.
Val-25	16	
Asp-26	16	Functional, partially buried carboxyl group 6.5 Å from sulfur of Cys-32 and 3.7 Å from Lys-57; charge stabilization in active site.
Phe-27	15	Structural, part of a hydrophobic cluster with residues 16, 25, 81, and 70.
Trp-28	11	Is slightly exposed on one edge and may account for fluorescence properties of prokaryotic thioredoxins.
Ala-29	16	Structural, in van der Waals contact with Trp-31. A larger residue would shift the position of the side chain of Trp-31.
Trp-31	17	Forms part of the active site surface; is firmly positioned by H-bonds to the side chain of Asp-61, and by van der Waals contact to Ala-29. Required for protein interactions.
Cys-32	17	Active site cysteine.
Gly-33	16	Maintains conformation of active site and influences redox potential.
Pro-34	17	As above.
Cys-35	17	Active site cysteine.
Lys-36	11	Possible stabilization of thiolate in active site; is often substituted by Arg. Important in protein interactions.
Met-37	11	Surface residue that may be involved in protein interactions.
Ala-39	9	Structural, part of an internal hydrophobic core.
Pro-40	17	Structural, makes a kink in α-2; maintains conformation of disulfide.
Tyr-49	13	Structural stabilization, H-bonded to Asp-104.
Lys-57	12	Protein interactions; is close to the active site and could influence the pK_a of active site cysteines.
Asn-59	9	Structural, H-bonded to main chain N of Asp-61; maintains correct position of Ala-29, Trp-31, and Asp-61 cluster. Is often replaced by Asp, which can also form this H-bond.
Asp-61	16	Structural, clustered with Ala-29 and Trp-31.
Asn-63	10	
Pro-64	10	Structural, 60% of residues in this type of structure in proteins are Pro.
Ala-67	13	Structural, is on internal side of 3_{10} helix.
Arg-73	11	Very obvious protrusion from surface, but no obvious function; may be important in protein interactions; is Lys in animal thioredoxins.
Ile-75	10	Structural, important in packing that requires hydrophobic residue at this position; is met in animal thioredoxins.
Pro-76	17	Structural, must be in *cis*-configuration; in van der Waals contact with disulfide bridge.
Thr-77	14	Structural, H-bonded to main chain atoms of residues 74 and 77.
Lys-82	16	Structural, charge stabilization at carboxyl end of α-4; one Arg substitution.
Gly-84	15	Structural, 70% of residues at equivalent positions in this type of turn are glycines.
Val-91	8	With Gly-92 and Ala-93, forms surface for protein interactions; is serine in animal thioredoxins.
Gly-92	17	Protein interaction area.
Ala-93	16	Protein interactions.
Lys-96	15	Protein interactions; one Arg substitution; Asp substitution in spinach thioredoxin f.

*Number of sequences in Figure 2 for which the residue is conserved. The total number of sequences is 17.

tase, substitution by serine obviously does not cause a significant alteration in the protein structure, as corroborated by the model. The mammalian and avian thioredoxins have Cys in the corresponding position. This type of substitution is relatively rare and may be relevant to species specific function.

A model of calf thymus thioredoxin is shown in Figure 3c as an example of the mammalian thioredoxins. Despite the large number of amino acid differences from the bacterial protein, the tertiary structure is essentially conserved. The prediction from this model is in agreement with secondary

TABLE III. Effects of Mutations in *E. coli* Thioredoxin on Structure-Function Relationships

Mutation	Effect of replacement	Reference
D26A	Increased stability	74
C32S	No redox activity	42,43
	Active in T7 and filamentous phage replication	
C35S	No redox activity	42,43
	Active in T7 and filamentous phage replication	
C32S; C35S	No redox activity	42,43
	Active in T7 and filamentous phage replication	
P34S	No effect on redox activity or stability	45,71
	Active in T7 and filamentous phage replication	42,43
P34H	Altered redox potential	68
K36E	Higher K_m with interacting enzymes	45
	Ionic strength effect	75
	Active in T7 and filamentous phage replication	45
G74D	Not active in T7 or filamentous phage replication	42,43
P76A	Does not show slow phase on refolding	71
	Not active in filamentous phage assembly	42
G92S	Temperature sensitive in filamentous phage assembly	42
	Slight support for T7 replication	43
	Reduced activity with thioredoxin reductase	42
G92D	Reduced activity with thioredoxin reductase	40
	Not active in T7 or filamentous phage replication	42,43

structure determination of the human thioredoxin which was recently done by NMR.[18] The extra cysteine residues in avian and mammalian thioredoxins have been suggested to form intramolecular disulfide bridges.[34,67] However, these residues are not in proximity (Fig. 4) and would not normally form intramolecular disulfide bridges without a major change in the tertiary structure. The NMR data show that these cysteines easily form intermolecular disulfides in oxidized human thioredoxin.[18]

Mammalian Proteins Containing Thioredoxin Domains

Edman and coworkers[36] reported that the sequence of rat protein disulfide isomerase (PDI) contained two regions with identity to *E. coli* thioredoxin. The protein disulfide isomerase group of sequences has been listed below the consensus line in Figure 2. The similarity to thioredoxin is striking and most of the essential residues are present in this extended family. The three-dimensional structure of the thioredoxin domains of PDIs should also be very similar to that of thioredoxin. More gaps must be placed for optimal alignment and there is a long insertion near position 50 (*E. coli* thioredoxin numbering). This represents a longer loop between α-2 and the 3_{10} helices. We have made no attempt to model the PDIs outside the thioredoxin domains.

The overall conservation of tertiary structure suggests that protein disulfide isomerases have a mechanism of action that is similar to thioredoxin. This has been confirmed by the discovery that PDI is a substrate for mammalian thioredoxin reductase and has thioredoxinlike activity in reduction of insulin.[46] The relative activities of PDI and thioredoxin are different, which is consistent with an altered redox potential that may be due to substitution of Pro-34 with a histidine residue.[68]

CONSERVED RESIDUES

Hydrophobic core

E. coli thioredoxin can be cleaved into two fragments with cyanogen bromide at the sole methionine at position 37. Although neither fragment shows any enzymatic activity, the two peptides will associate to reform a thioredoxinlike molecule in the absence of a covalent bond between them. The peptides can only be separated under denaturing conditions in high ionic strength buffers, which implies reassociation via an hydrophobic interaction.[69] In like manner, specific tryptic cleavage at Arg-73 generates two fragments that also reassociate.[70] The thioredoxin molecule can be visualized as consisting of two large folding units, one βαβαβ and one ββα unit (see Fig. 1). The cleavage at Arg-73 divides the molecule roughly into these folding units. In the thioredoxin structure, interaction between β-2 and β-4 strands is mainly responsible for this association and is vital for maintaining the active site conformation in the native protein. The importance of this interaction to the general stability of the thioredoxin structure is further substantiated by a comparison of amino acid sequences. A large number of the residues in the interface between the folding units are conserved (Table II, Fig. 5). In the first folding unit, residues Pro-22, Val-23, Leu-24, Val-25, Asp-26, Phe-27, Trp-28, Ala-29, and Tyr-49 are at this interface, whereas Ile-75, Pro-76, Thr-77, Gly-84, and Lys-96 in the second folding unit are at the interface. Although some amino acid substitu-

tions do occur in these regions, they are relatively conservative changes and can be readily accommodated without large alterations in the overall structure. Production of chimeric proteins between *E. coli* and *Anabaena* 7119 thioredoxins illustrates the importance of this interaction for activity. Two hybrid proteins were made by fusion of the bacterial genes at the active site. The *Anabaena-E. coli* hybrid protein is both as stable and enzymatically active as wild-type *E. coli* thioredoxin. In contrast, the *E. coli-Anabaena* hybrid thioredoxin is unstable and a relatively poor disulfide reductase. Amino acid substitutions in the β-4 strand of *Anabaena* thioredoxin apparently lead to poor interaction with the amino-terminal folding unit of the *E. coli* protein.[44]

Protein-Protein Interactions

A previous comparison of two thioredoxins, glutaredoxin and T4 thioredoxin, suggested functionally conserved regions in these proteins.[15,16] All have a hydrophobic area close to the accessible Cys-32 of the active site, which was suggested to be important for interactions with thioredoxin reductase and ribonucleotide reductase. This area is still conserved in the larger data base now under consideration. All thioredoxins and PDI domains have a proline in the position equivalent to Pro-76. This proline is in the less usual *cis* configuration. Pro-76 is located directly behind the disulfide bridge and in van der Waals contact with it (Fig. 5; see also Fig. 8). It may serve to stabilize the active site. Site-directed mutation of this proline to alanine eliminates the slow phase observed in thioredoxin folding from the denatured state (Table III).[71] The mutant protein also does not support filamentous phage assembly.[42] Further analysis of mutant thioredoxins suggests that the *cis* configuration facilitates redox interaction with other proteins.[62]

Gly-92 is also invariant in the interaction area. All thioredoxins listed in Figure 2 have this residue. Gly-92 may be important for the proper kink between the last strand and the α-4 helix, which would allow an interacting protein such as thioredoxin reductase to dock properly with the active site. Different *E. coli* thioredoxin mutants in which Gly-92 is changed to a Ser or an Asp have remarkably different properties from the wild type protein. G92S thioredoxin is able to support phage replication at low temperature (30°C) or when present on a high copy number plasmid.[42] It is also a substrate for thioredoxin reductase with approximately half the activity of the wild-type protein. In contrast, G92D thioredoxin will not support viral replication or assembly under any conditions.[40] It is only about 20% as active as wild-type thioredoxin with thioredoxin reductase. When G92D is used as a substrate for the reductase, both the K_m (sevenfold higher) and the V_{max} (one-third that of wild-type thioredoxin) are altered.[40–43] An acidic side-chain in this position disturbs interaction between thioredoxin and other proteins, whereas the more conservative substitution of a serine yields a partially active protein. These results are summarized in Table III.

The positive charge closest to the active site disulfide bridge is Lys-36 in the *E. coli* thioredoxin. This residue is often Lys or Arg in the thioredoxins of other species. It has been suggested that the positive charge of this residue can influence the pK_a of Cys-32 and stabilize the thiolate anion at physiological pH.[72] In a mutant thioredoxin in which this lysine was changed to a glutamic acid, protein-protein interactions were primarily affected. Thioredoxin K36E shows a significantly reduced rate of complex formation with thioredoxin reductase, which can be partially overcome by increasing the ionic strength.[75] However, thioredoxin K36E is a good reducing agent at pH 7. The data imply that other interactions, such as the helix dipole or Lys-57,[77] are of primary importance in maintaining the thiolate anion at neutral pH and that the positive charge at position 36 serves mainly to facilitate thioredoxin-protein interactions.[45]

A positive charge is also present in thioredoxins and glutaredoxins at approximately the same position on the shielded side of the active site disulfide.[16] All thioredoxins in this investigation have Lys-57 conserved except for the animal thioredoxins. However, in these proteins there is a lysine at position 39 (see Fig. 6). Lys-39 is equivalent in space to a corresponding Lys residue in T4 thioredoxin. Thus, although the positively charged residue is not strictly conserved in the primary structure, the position is conserved in the three-dimensional structure. The role of this residue may be important for protein interactions, or it may influence the pK_a and reactivity of the cysteines.

Another highly conserved residue is Asp-26 in the β-2 strand. It is located at a position where one would expect a hydrophobic residue since it is in the central strand of the core of the pleated sheet. Buried carboxyl side chains do occur in proteins but are generally hydrogen bonded.[13] Asp-26 is not completely buried and one of its oxygens is slightly exposed at the surface of the protein only 4.8 Å from the charge of the conserved Lys-57 (Lys-39 in the animal thioredoxins)[13] (see Fig. 6). There is a narrow channel from the carboxyl group to the surface of the molecule as shown in Figure 7. It has been suggested that Asp-26 interacts with a positively charged residue in the reductases.[12] In the crystallographic investigation of oxidized thioredoxin done at pH 4.5, Asp-26 may be protonated. Since it is only 6.5 Å from the sulfur atom of Cys-35, it could also influence the pK_a of this residue. In the structure of reduced thioredoxin, determined by two-dimensional NMR, the Asp-26 carboxyl group is close to both the sulfhydryl group of Cys-35 and the ε-amino group of Lys-57.[77] A mutant thioredoxin in which

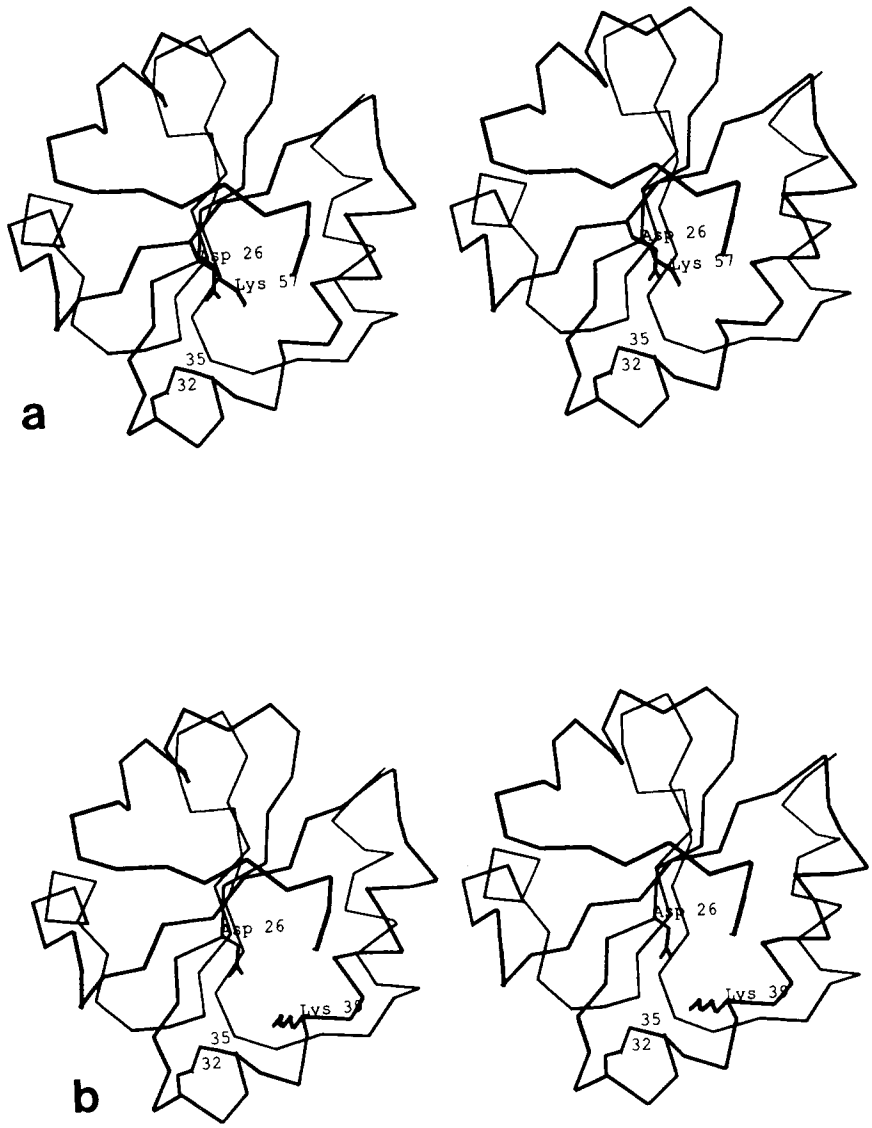


Fig. 6. **(a)** Stereo drawing of *E. coli* thioredoxin based on the x-ray data. Side chains of Asp-26 and Lys-57 are shown. The carboxyl group of Asp-26 is clearly buried in the interior of the molecule but is in proximity to Lys-57. **(b)** Stereo drawing of a model of calf thymus thioredoxin showing positions of Asp-26 and Lys-39. Asp-26 is conserved in most thioredoxins. However, in avian and mammalian thioredoxins, Lys-57 is replaced by a glutamic acid residue.

Asp-26 has been replaced by Ala[74] functions better as a reducing agent for ribonucleotide reductase at pH 8.2 than at pH 7.6 (Gleason, unpublished observations). This indicates that Asp-26 does influence the redox properties of the active site cysteines.

Trp-31 and Ala-29 are invariant. The Trp side chain covers an important part of the active site surface and may facilitate the optimum interactions with enzymes such as ribonucleotide reductase and thioredoxin reductase. Ala-29 is covered by the Trp-31 side chain and the Ala methyl group is in van der Waals contact with the ring. Its function is probably to give the Trp side chain the proper orientation. The side chain of Trp-31 is rigidly stabilized at the surface of the active site by hydrogen bonding (Figs. 5, 8). The aromatic ring of Trp-31 is H-bonded to both side chain oxygen atoms of Asp-61. This latter residue is highly conserved, but is substituted by a threonine in *Corynebacterium* C-2 thioredoxin and the PDIs. The position of Asp-61 is in turn stabilized by hydrogen bonds from the side chain oxygen of Asn-59 to N61 and N62. The corresponding residues in other thioredoxins and PDIs are Asn, Asp, or Glu, which can also form these hydrogen bonds. Other H-bonded interactions also stabilize the active site region. The last hydrogen bonded residue of strand β-3 is Trp-28 in prokaryotic thioredoxins, which binds to N60 through its main chain carbonyl oxy-

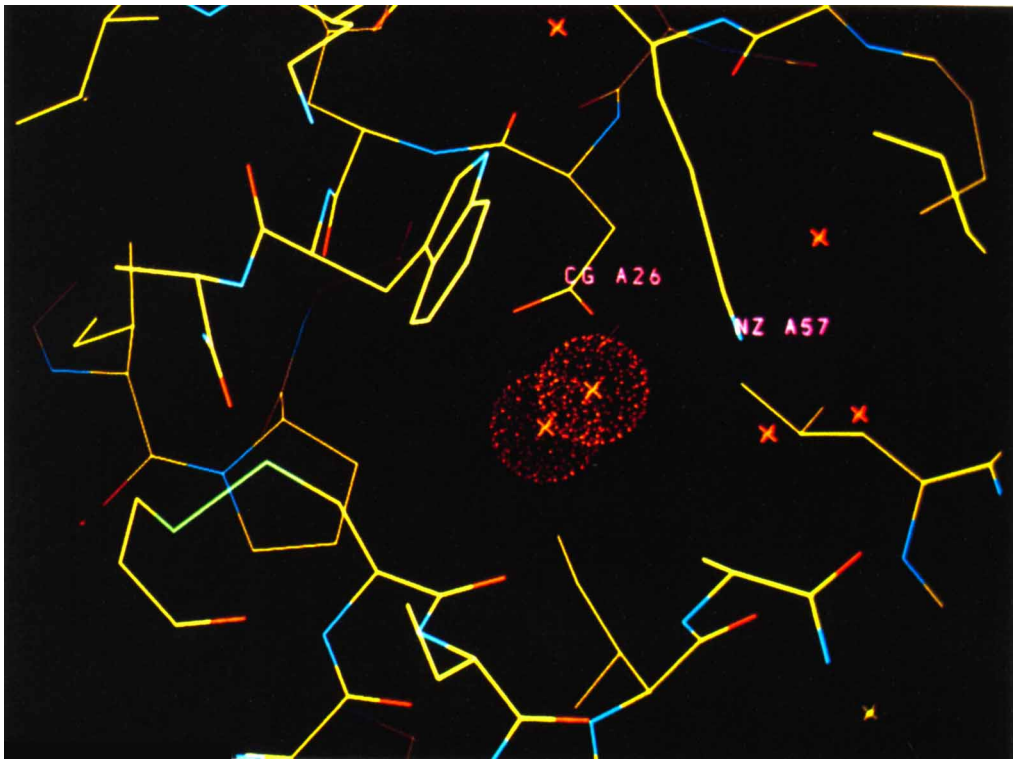


Fig. 7. Diagram showing the position of the buried Asp-26 in *E. coli* thioredoxin. A narrow channel is seen connecting the carboxyl group to the exterior. In the crystal structure, this is occupied by two water molecules.

gen. The bend is stabilized by a normal 1–4 main chain hydrogen bond.

Prolines and glycines are common in or close to bends in proteins. Pro-22, Pro-64, Gly-84, and Gly-92 are highly conserved residues in thioredoxins and fulfill this function. In addition, some of the residues close to bends participate in H-bonding to main chain residues and thus stabilize the structure, as Asn-59 mentioned above. Asp-9 connects the end of the first strand of the structure to main chain atoms of the second folding half of the molecule. This Asp is also highly conserved.

The Active Site

The active site of thioredoxin-S_2 is a 14-membered disulfide ring that undergoes reduction to a dithiol. The structure of *E. coli* thioredoxin-$(SH)_2$ has recently been studied by two dimensional ^{1}H-NMR techniques[17] and a three-dimensional structure has been constructed by distance geometry and molecular dynamics methods.[77] The secondary structure and global fold are very similar to that of thioredoxin-S_2. Except for changes in the disulfide ring, notably a rotation of the side chain of Cys-35, structural alterations are minor and localized to the hydrophobic surface area.[77]

The active site is located in the amino end of α-2 on the surface of the molecule. This part of the helix is separated from the rest of the structure by a kink caused by Pro-40. As a result, the active site helix protrudes from the rest of the molecule but the sulfur atoms of the S—S bridge are located on the inside of the helix and are covered by it (Figs. 8, 9). The residues between the cysteines form the exposed side of the first part of α-2 and have the sequence Gly-Pro in thioredoxins. There is only one exception, a Gly to Ala substitution in *C. nephridii* C2 thioredoxin. However, the change is a relatively minor one that is easily accommodated and has been reported to have little impact on the redox function of this thioredoxin.[21] Although studies of model peptides show considerable strain in small disulfide loops,[73] the conformation found for the disulfide bridge residues in the *E. coli* thioredoxin structure show normal torsion angles.[13] A 14-membered disulfide ring is also found in all glutaredoxins, the active site of thioredoxin reductase, and the thioredoxin domains of PDIs. The residues between the cysteines vary, reflecting different substrate specificities. All PDIs have the sequence, Cys-Gly-His-Cys. The sequence of glutaredoxins is Cys-Pro-Tyr-Cys and presumably accommodates the binding of reduced glutathione by this protein. The unusual thioredoxin from T4 phage has properties of both thioredoxin and glutaredoxin,

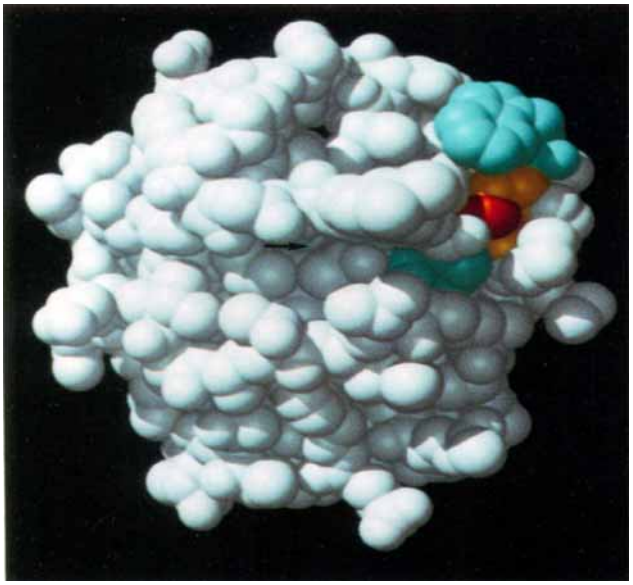


Fig. 8. Space-filling model of *E. coli* thioredoxin based on the x-ray coordinates. Cysteine residues at the active site are colored yellow. The exposed sulfur of Cys-32 is shown in red. In contrast, the sulfur of Cys-35 is not exposed and cannot be seen in the model. Trp-31, which partially covers the disulfide, is in blue. Pro-76, which interacts with the active site disulfide, is also in blue and is partially exposed behind the sulfur of Cys-32. The narrow channel leading to the buried Asp-26 is indicated by the arrow.

interacting with both glutathione and *E. coli* thioredoxin reductase.[63] It has the active site sequence, Cys-Val-Tyr-Cys. Stepwise mutation of the Val and Tyr residues to Gly and Pro produces viral thioredoxins with properties more like those of the *E. coli* protein and diminished binding of glutathione.[62] Alteration of Pro-34 to Ser in *E. coli* thioredoxin by site directed mutagenesis produced a protein with folding properties and stability like those of wild-type thioredoxin,[71] thus confirming that the residues between the cysteines have little effect on overall tertiary structure. The P34S thioredoxin also restored the wild-type phenotype in *E. coli* thioredoxin-deficient mutants, including support of phage replication.[42,45] Only minor changes were found in kinetic constants in interactions with thioredoxin reductase and ribonucleotide reductase. This mutation appears to have little effect on redox activity or stability (Table III). A proline to serine mutation is relatively common in protein families.[66] The fact that Pro-34 is so highly conserved in all thioredoxins implies additional, unexplored functions for thioredoxin.

ACKNOWLEDGMENTS

This work was supported by grants from the Swedish Natural Science Research Council and Swedish Medical Research Council. F.K.G. acknowl-

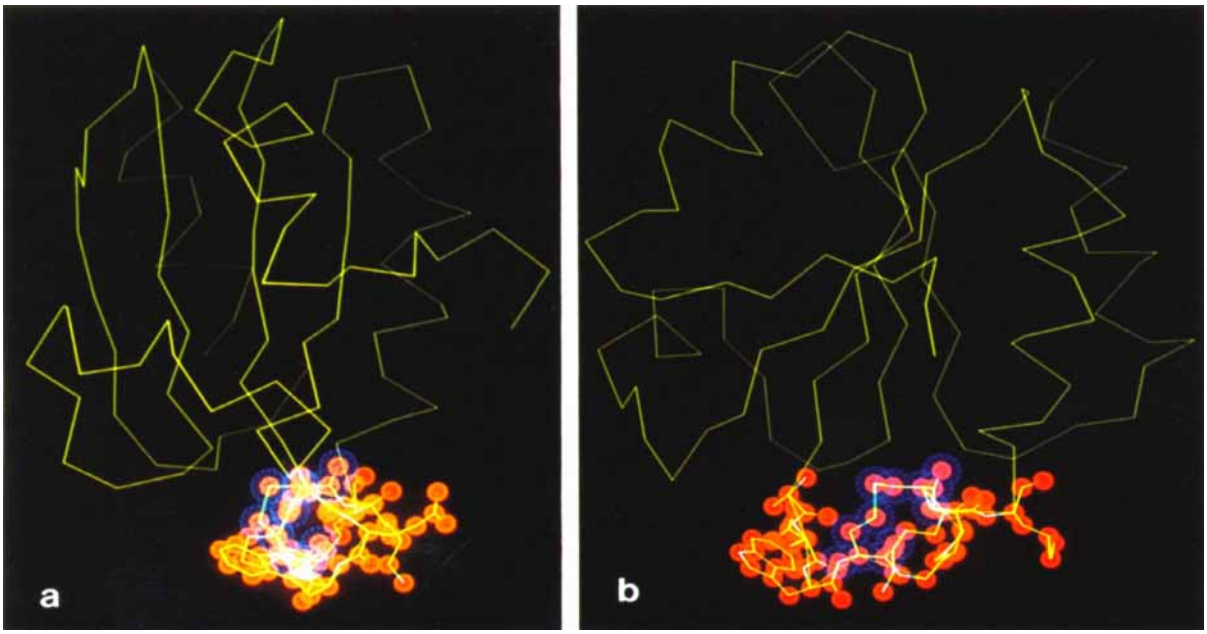


Fig. 9. **(a)** Drawing of the alpha carbon backbone of *E. coli* thioredoxin. The active site residues are highlighted by yellow spheres. The active site cystine is further enhanced in blue. The model clearly shows that the active site is exposed at the surface of the molecule but that the disulfide itself is partially covered by other residues in the active site region. **(b)** Same as above but with the molecule oriented as in Figure 1.

edges support from the U.S. National Science Foundation, DCB-8817071, and the Swedish Medical Research Council visiting scientist fellowship, 13V-8469.

REFERENCES

1. Holmgren, A. Thioredoxin. Annu. Rev. Biochem. 54:237–271, 1985.
2. Gleason, F.K., Holmgren, A. Thioredoxin and related proteins in prokaryotes. FEMS Microbiol. Rev. 54:271–298, 1988.
3. Holmgren, A. Thioredoxin and glutaredoxin systems. J. Biol. Chem. 264:13963–13966, 1989.
4. Thelander, L., Reichard, P. Reduction of ribonucleotides. Annu. Rev. Biochem. 48:133–158, 1979.
5. Black, S., Harte, E.M., Hudson, B., Wartofsky, L. A specific enzymatic reduction of L(−)methionine sulfoxide and a related non-specific reduction of disulfides. J. Biol. Chem. 235:2916–2910, 1960.
6. Wilson, L.G., Asahi, T., Bandurski, R.S. Yeast sulfate-reducing systems. I. Reduction of sulfate to sulfite. J. Biol. Chem. 236:1822–1829, 1961.
7. Porque, P.G., Baldesten, A., Reichard, P. The involvement of the thioredoxin system in the reduction of methionine sulfoxide and sulfate. J. Biol. Chem. 245:2371–2374, 1970.
8. Mark, D.F., Richardson, C.C. *Escherichia coli* thioredoxin: A subunit of bacteriophage T7 DNA polymerase. Proc. Natl. Acad. Sci. U.S.A. 73:780–784, 1976.
9. Russel, M., Model, P. Thioredoxin is required for filamentous phage assembly. Proc. Natl. Acad. Sci. U.S.A. 82:29–33, 1985.
10. Lim, C-J., Haller, B., Fuchs, J.A. Thioredoxin is the bacterial protein encoded by *fip* that is required for filamentous bacteriophage f1 assembly. J. Bacteriol. 161:799–802, 1985.
11. Crawford, N.A., Yee, B.C., Droux, M., Carlson, D.E., Buchanan, B.B. Ferredoxin/thioredoxin system. Meth. Enzymol. 167:415–427, 1988.
12. Holmgren, A., Söderberg, B.-O., Eklund, H., Bränden, C.-I. Three-dimensional structure of *Escherichia coli* thioredoxin-S_2 to 2.8 Å resolution. Proc. Natl. Acad. Sci. U.S.A. 72:2305–2309, 1975.
13. Katti, S.K., LeMaster, D.M., Eklund, H. Crystal structure of thioredoxin from *Escherichia coli* at 1.68 Å resolution. J. Mol. Biol. 212:167–184, 1990.
14. Priestle, J.P. RIBBON: A stereo cartoon drawing program for proteins. J. Appl. Cryst. 21:572–576, 1988.
15. Bränden, C.-I., Eklund, H., Söderberg, B.-O. Structural and functional similarities in thioredoxin from *Escehrichia coli* and bacteriophage T4. In: "Functions of Glutathione." Larsson, A., Orrenius, S., Holmgren, A., Mannervik, B. (eds.). New York: Raven Press, 1983: 223–230.
16. Eklund, H., Cambillau, C., Sjöberg, B.-M., Holmgren, A., Jörnvall, H., Höög, J.-O., Bränden, C.-I. Conformational and functional similarities between glutaredoxin and thioredoxins. EMBO J. 3:1410–1417, 1984.
17. Dyson, H.J., Holmgren, A., Wright, P.E. Assignment of the proton NMR spectrum of reduced and oxidized thioredoxin: Sequence-specific assignments, secondary structure, and global fold. Biochemistry 28:7074–7087, 1989.
18. Forman-Kay, J.D., Clore, G.M., Driscoll, P.C., Wingfield, P., Richards, F.M., Gronenborn, A.M. A proton nuclear magnetic resonance assignment and secondary structure determination of recombinant human thioredoxin. Biochemistry 28:7088–7097, 1989.
19. Höög, J.-O., vonBahr-Lindström, H., Josephson, S., Wallace, B., Kushner, S.R., Jörnvall, H., Holmgren, A. Nucleotide sequence of the thioredoxin gene from *Escherichia coli*. Biosci. Rep. 4:917–923, 1985.
20. Meng, M., Hogenkamp, H.P.C. Purification, characterization and amino acid sequence of thioredoxin from *Corynebacterium nephridii*. J. Biol. Chem. 262:12114–12119, 1981.
21. McFarlan, S., Hogenkamp, H.P.C., Eccleston, E.D., Howard, J.B., Fuchs, J.A. Purification, characterization and revised amino acid sequence of a second thioredoxin from *Corynebacterium nephridii*. Eur. J. Biochem. 179: 389–398, 1989.
22. Gleason, F.K., Whittaker, M.M., Holmgren, A., Jörnvall, H. The primary structure of thioredoxin from the filamentous cyanobacterium *Anabaena* sp. 7119. J. Biol. Chem. 260:9567–9573, 1985.
23. Muller, E.G.D., Buchanan, B.B. Thioredoxin is essential for photosynthetic growth. J. Biol. Chem. 264:4008–4014, 1989.
24. Alam, J., Curtis, S.E., Gleason, F.K., Gerami-Nejad, M., Fuchs, J.A. Isolation, sequence and expression in *Escherichia coli* of an unusual thioredoxin gene from the cyanobacterium, *Anabaena* sp. strain PCC 7120. J. Bacteriol. 171:162–171, 1989.
25. Clement-Metral, J.D., Holmgren, A., Cambillau, C., Jörnvall, H., Eklund, H., Thomas, D., Lederer, F. Amino acid sequence-determination and three-dimensional modelling of thioredoxin from the photosynthetic bacterium, *Rhodobacter sphaeroides* Y. Eur. J. Biochem. 172:413–419, 1988.
26. Johnson, T.C., Yee, B.C., Carlson, D.E., Buchanan, B.B., Johnson, R.S., Mathews, W.R., Biemann, K. Thioredoxin from *Rhodospirillum rubrum:* Primary structure and relation to thioredoxins from other photosynthetic bacteria. J. Bacteriol. 170:2406–2408, 1988.
27. Mathews, W.R., Johnson, R.S., Cornwell, K.L., Johnson, T.C., Buchanan, B.B., Biemann, K. Mass spectrometrically derived amino acid sequence of thioredoxin from *Chlorobium,* an evolutionarily prominent photosynthetic bacterium. J. Biol. Chem. 262:7537–7545, 1987.
28. Johnson, R.S., Biemann, K. The primary structure of thioredoxin from *Chromatium vinosum* determined by high performance mass spectrometry. Biochemistry 26: 1209–1214, 1987.
29. Wedel, N., Clausmeyer, S., Gardet-Salvi, L., Herrmann, R.G., Schürmann, P. Nucleotide sequence of cDNAs encoding the entire precursor polypeptide for thioredoxin m from spinach chloroplasts. Plant Mol. Biol. (in press), 1990.
30. Kamo, M., Tsugita, A., Wiessner, C., Wedel, N., Barling, D., Herrmann, R.G., Aguilar, F., Gardet-Salvi, L., Schürmann, P. Primary structure of spinach chloroplast thioredoxin f. Eur. J. Biochem. 182:315–322, 1989.
31. Jones, S.W., Luk, K.-C. Isolation of a chicken thioredoxin cDNA clone. J. Biol. Chem. 263:9607–9611, 1988.
32. Johnson, R.S., Mathews, W.R., Biemann, K., Hopper, S. Amino acid sequence of thioredoxin isolated from rabbit bone marrow determined by tandem mass spectrometry. J. Biol. Chem. 263:9589–9597, 1988.
33. Tagaya, Y., Maeda, Y., Mitsui, A., Kondo, N., Matsui, H., Hamuro, J., Brown, N., Arai, K.-i., Yokota, T., Wakasugi, H., Yodoi, J. ATL-derived factor (ADF), an IL-2 receptor/Tac inducer homologous to thioredoxin; possible involvement of dithiol-reduction in the IL-2 receptor induction. EMBO J. 8:757–764, 1989.
34. Holmgren, A., Palmberg, C., Jörnvall, H., Hernberg, C. Bovine thioredoxin and thioredoxin reductase: Amino acid sequence of calf thymus thioredoxin and reactions of oxidized and reduced forms with thioredoxin reductase. Eur. J. Biochem. manuscript.
35. Wollman, E.E., d'Auriol, L., Rimsky, L., Shaw, A., Jacquot, J.-P., Wingfield, P., Graber, P., Dessarps, F., Robin, P., Galibert, F., Bertoglio, J., Fradelizi, D. Cloning and expression of a cDNA for human thioredoxin. J. Biol. Chem. 263:15506–15512, 1988.
36. Edman, J.C., Ellis, L., Blacher, R.W., Roth, R.A., Rutter, W.J. Sequence of protein disulfide isomerase and implications of its relationship to thioredoxin. Nature 317:267–270, 1985.
37. Pihlajaniemi, T., Helaakoski, T., Tasanen, K., Myllylä, R., Huhatala, M.L., Koivu, J., Kivirikko, K.I. Molecular cloning of the β-subunit of human prolyl-4-hydroxylase. This subunit and protein disulfide isomerase are products of the same gene. EMBO J. 6:643–649, 1987.
38. Bennett, C.F., Balcarck, J.M., Varrichio, A., Crooke, S.T. Molecular cloning and complete amino acid sequence of form I phosphoinositide-specific phospholipase C. Nature 334:268–270, 1988.
39. Hsu, M.P., Muhich, M.L., Boothroyd, J.C. A developmentally regulated gene of trypanosomes encodes a homologue of rat protein-disulfide isomerase and phosphoinositol phosphorylase C. Biochemistry 28:6440–6446, 1989.
40. Holmgren, A., Kallis, G.-B., Nordström, B. A mutant

thioredoxin from *Escherichia coli* tsnC7007 that is nonfunctional as a subunit of phage T7 DNA polymerase. J. Biol. Chem. 256:3118–3124, 1981.
41. Russel, M., Model, P. Replacement of the *fip* gene of *Escherichia coli* by an inactive gene cloned on a plasmid. J. Bacteriol. 159:1034–1039, 1984.
42. Russel, M., Model, P. The role of thioredoxin in filamentous phage assembly. Construction, isolation and characterization of mutant thioredoxins. J. Biol. Chem. 261: 14997–15005, 1986.
43. Huber, H.E., Russel, M., Model, P., Richardson, C.C. Interaction of mutant thioredoxins of *Escherichia coli* with the gene 5 protein of phage T7. J. Biol. Chem. 261:15006–15012, 1986.
44. Lim, C.-J., Gleason, F.K., Jacobson, B.A., Fuchs, J.A. Characterization of *Escherichia coli-Anabaena* sp. hybrid thioredoxins. Biochemistry 27:1401–1408, 1988.
45. Gleason, F.K., Lim, C-J., Gerami-Nejad, M., Fuchs, J.A. Characterization of *Escherichia coli* thioredoxins with altered active site residues. Biochemistry 29:3701–3709, 1990.
46. Lundström, J., Holmgren, A. Protein disulfide-isomerase is a substrate for thioredoxin reductase and has thioredoxin-like activity. J. Biol. Chem. 265:9114–9120, 1990.
47. Freedman, R.B., Hawkins, H.C., Murant, S.J., Reid, L. Protein disulphide-isomerase: A homologue of thioredoxin implicated in the biosynthesis of secretory proteins. Biochem. Soc. Trans. 16:96–99, 1988.
48. Yamauchi, K., Yamamoto, T., Nayashi, H., Koya, S., Takikawa, H., Toyoshima, K., Horiuchi, R. Sequence of membrane-associated thyroid hormone binding protein from bovine liver: Its identity with protein disulfide isomerase. Biochem. Biophys. Res. Commun. 146:1485–1492, 1987.
49. Cheng, S.-y., Gong, Q.-h., Parkinson, C., Robinson, E.A., Appella, E., Merlino, G.T., Pastan, I. The nucleotide sequence of a human cellular thyroid hormone binding protein present in endoplasmic reticulum. J. Biol. Chem. 262: 11221–11227, 1987.
50. Jones, T.A. A graphics model building and refinement system for macromolecules. J. Appl. Crystallogr. 11:268–272, 1978.
51. Jones, T.A. Interactive computer graphics: FRODO. Meth. Enzymol. 115:157–171, 1985.
52. Ponder, J.W., Richards, F.M. Tertiary templates for proteins. J. Mol. Biol. 193:775–791, 1987.
53. Tsugita, A., Maeda, K., Schürmann, P. Spinach chloroplast thioredoxins in evolutionary drift. Biochem. Biophys. Res. Commun. 115:1–7, 1983.
54. Gleason, F.K. Function of two dissimilar thioredoxins in the cyanobacterium, *Anabaena* sp. 7120. Current Research in Photosynthesis, IV:16.175–16.178, 1990.
55. Chothia, C., Lesk, A.M. The relation between the divergence of sequence and structure in proteins. EMBO J. 5: 823–826, 1986.
56. Delbaere, L.T.J., Hutcheon, W.L.B., James, M.N.G., Thiessen, W.E. Tertiary structural differences between microbial serine proteases and pancreatic serine enzymes. Nature 257:758–763, 1975.
57. Salemme, F.R. Structure and function of cytochromes c. Annu. Rev. Biochem. 46:299–329, 1977.
58. Eklund, H., Bränden, C.-I. Alcohol dehydrogenase. In: "Biological macromolecules and assemblies. Vol 3: Active sites of enzymes." Jurnak, F.A., McPherson, A. (eds.). New York: John Wiley & Sons, 1987.
59. Höög, J.-O., Jörnvall, H., Holmgren, A., Carlquist, M., Persson, M. The primary structure of *Escherichia coli* glutaredoxin. Eur. J. Biochem. 136:223–232, 1983.
60. Klintrot, I-M., Höög, J.-O., Jörnvall, H., Holmgren, A., Luthman, M. The primary structure of calf thymus glutaredoxin. Eur. J. Biochem. 144:417–423, 1984.
61. Papayannopoulos, I.A., Gan, Z.-R., Wells, W.W., Biemann, K. A revised sequence of calf thymus glutaredoxin. Biochem. Biophys. Res. Commun. 159:1448–1454, 1989.
62. Nikkola, M., Gleason, F.K., Saarinen, M., Eklund, H., manuscript in preparation, 1990.
63. Holmgren, A. Glutathione-dependent enzyme reactions of the phage T4 ribonucleotide reductase system. J. Biol. Chem. 253:7424–7430, 1978.
64. Lim, C.-J., Gleason, F.K., Fuchs, J.A. Cloning, expression, and characterization of the *Anabaena* thioredoxin gene in *Escherichia coli*. J. Bacteriol. 168:1258–1264, 1986.
65. Jones, T.A., Thirup, S. Using known substructures in protein model building and crystallography. EMBO J. 5:819–822, 1986.
66. Creighton, T.E. "Proteins." New York: W.H. Freeman, 1984:11.
67. Holmgren, A. Bovine thioredoxin system. Purification of thioredoxin reductase from calf liver and thymus and studies of its function in disulfide reduction. J. Biol. Chem. 252:4600–4606, 1977.
68. Krause, G., Lundström, J., Lopez-Barea, J., Pueyo de La Cuesta, C., Holmgren, A. Mimicking the active site of protein disulfide isomerase by substitution of proline-34 in *Escherichia coli* thioredoxin. J. Biol. Chem., in press, 1991.
69. Holmgren, A., Slaby, I. Thioredoxin-C′: Mechanism of noncovalent complementation and reactions of the refolded complex and the active site containing fragment with thioredoxin reductase. Biochemistry 18:5591–5599, 1979.
70. Slaby, I., Holmgren, A., Structure and enzymatic functions of thioredoxin refolded by complementation of two tryptic peptide fragments. Biochemistry 18:5584–5591, 1979.
71. Kelley, R.F., Richards, F.M. Replacement of proline-76 with alanine eliminates the slowest kinetic phase in thioredoxin folding. Biochemistry 26:6765–6774, 1987.
72. Kallis, G.-B., Holmgren, A. Differential reactivity of the functional sulfhydryl groups of cysteine-32 and cysteine-35 present in the reduced form of thioredoxin from *Escherichia coli*. J. Biol. Chem. 255:10261–10265, 1980.
73. Hardy, P.M., Ridge, B., Rydon, H.N., dos S.P. Serrao, F.O. Polypeptides. Part XV. The synthesis and oxidation of some L-cysteinyl polyglycyl-L-cysteines. J. Chem. Soc. 1722–1731, 1971.
74. Langsetmo, K., Sung, Y.-C., Fuchs, J., Woodward, C. *E. coli* thioredoxin stability is greatly enhanced by substitution of aspartic acid 26 by alanine. In: "Current Research in Protein Chemistry." Villafranca, J. (ed). New York Academic Press, 1990: 449–456.
75. Navarro, J.A., Gleason, F.K., Cusanovich, M.A., Fuchs, J.A., Meyer, T.E., Tollin, G. Kinetics of electron transfer from thioredoxin reductase to thioredoxin, Biochemistry, 30:2192–2195, 1991.
76. Clement-Metral, J.D. Cloning, nucleotide sequence and expression of *Rhodobacter sphaeroides* y thioredoxin gene. Current Research in Photosynthesis, III:12.521–12.524, 1990.
77. Dyson, H.J., Gippert, G.P., Case, D.A., Holmgren, A., Wright, P.E. Three-dimensional solution structure of the reduced form of *Escherichia coli* thioredoxin determined by nuclear magnetic resonance spectroscopy. Biochemistry 29:4129–4136, 1990.