High-Resolution Three-Dimensional Structure of Reduced Recombinant Human Thioredoxin in Solution†

Julie D. Forman-Kay,‡,§ G. Marius Clore,*,‡ Paul T. Wingfield,‖ and Angela M. Gronenborn*,‡

Laboratory of Chemical Physics, Building 2, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892, Molecular Biophysics and Biochemistry Department, Yale University, New Haven, Connecticut 06511, and Protein Expression Laboratory, Building 6B, Office of the Director, National Institutes of Health, Bethesda, Maryland 20892

Received July 23, 1990; Revised Manuscript Received October 12, 1990

ABSTRACT: The solution structure of recombinant human thioredoxin (105 residues) has been determined by nuclear magnetic resonance (NMR) spectroscopy combined with hybrid distance geometry–dynamical simulated annealing calculations. Approximate interproton distance restraints were derived from nuclear Overhauser effect (NOE) measurements. In addition, a large number of stereospecific assignments for β-methylene protons and torsion angle restraints for ϕ, ψ, and χ_1 were obtained by using a conformational grid search on the basis of the intraresidue and sequential NOE data in conjunction with $^3J_{HN\alpha}$ and $^3J_{\alpha\beta}$ coupling constants. The structure calculations were based on 1983 approximate interproton distance restraints, 52 hydrogen-bonding restraints for 26 hydrogen bonds, and 98 ϕ, 71 ψ, and 72 χ_1 torsion angle restraints. The 33 final simulated annealing structures obtained had an average atomic rms distribution of the individual structures about the mean coordinate positions of 0.40 ± 0.06 Å for the backbone atoms and 0.78 ± 0.05 Å for all atoms. The solution structure of human thioredoxin consists of a five-stranded β-sheet surrounded by four α-helices, with an active site protrusion containing the two redox-active cysteines. The overall structure is similar to the crystal and NMR structures of oxidized [Katti, S. K., LeMaster, D. M., & Eklund, H. (1990) *J. Mol. Biol. 212*, 167–184] and reduced [Dyson, J. H., Gippert, G. P., Case, D. A., Holmgren, A., & Wright, P. (1990) *Biochemistry 29*, 4129–4136] *Escherichia coli* thioredoxin, respectively, despite the moderate 25% amino acid sequence homology. Several differences, however, can be noted. The human α_1 helix is a full turn longer than the corresponding helix in *E. coli* thioredoxin and is characterized by a more regular helical geometry. The helix labeled α_3 in human thioredoxin has its counterpart in the 3_{10} helix of the *E. coli* protein and is also longer in the human protein. In contrast to these structural differences, the conformation of the active site loop in both proteins is very similar, reflecting the perfect sequence identity for a stretch of eight amino acid residues around the redox-active cysteines.

Thioredoxin is a ubiquituous dithiol oxidoreductase found in many organisms and involved in numerous biochemical processes. Although the *Escherichia coli* protein is the most widely studied, thioredoxins from other bacteria, bacteriophages, plants, and mammals have also been isolated and sequenced (Holmgren, 1985, 1989). Thioredoxin functions as a hydrogen donor for ribonucleotide reductase in the biosynthesis of deoxyribonucleotides (Laurent et al., 1964; Engstrom et al., 1974), constitutes a structural unit in the phage T7 DNA polymerase complex (Mark & Richardson, 1976), plays an important role in plant chloroplasts in the expression of photosynthetic enzymes (Buchanan, 1980), and acts as a general reductant for disulfides in proteins, including insulin, oxytocin, and fibrinogen; it is therefore potentially involved in the correct folding of proteins with mixed disulfides (Holmgren, 1985). Human thioredoxin has recently been implicated in immune function on the basis of its high abundance in immunocompetent cell lines, such as EBV[1]-transformed B-cells (Rimski et al., 1986; Wakasugi et al., 1987) and HTLV-1-transformed T-cells (Tagaya et al., 1989), suggesting a possible involvement of thioredoxin-mediated dithiol reduction in the induction of interleukin 2 receptor expression (Tagaya et al., 1987).

E. coli thioredoxin has been frequently used as a model system in studies on protein structure and folding since it is a small one-domain protein with no structurally required disulfide bonds and possesses extreme thermal stability (Wilson et al., 1986; Kelley et al., 1986; Kelley & Richards, 1987; Hiraoki et al., 1988; Langsetmo et al., 1989; Lin & Kim, 1989). X-ray structures of the oxidized form of *E. coli* (Holmgren et al., 1975) and phage T4 (Söderberg et al., 1978) thioredoxins at 2.8 Å have been reported, and a refined crystal structure of *E. coli* thioredoxin at 1.68-Å resolution has recently become available (Katti et al., 1990). No crystals of the reduced form have been obtained to date. The structure of reduced *E. coli* thioredoxin, however, has been determined in solution by NMR spectroscopy (Dyson et al., 1990).

† This work was supported by the Intramural AIDS Targeted Antiviral Program of the Office of the Director, NIH (G.M.C. and A.M.G.). J.D.F.-K. acknowledges a graduate fellowship from the Molecular Biophysics and Biochemistry Department of Yale University and support from NIH Grant GM-22778 (to F. M. Richards, Yale University).

‡ Laboratory of Chemical Physics, National Institutes of Health.
§ Yale University.
‖ Protein Expression Laboratory, National Institutes of Health.

[1] Abbreviations: *E. coli*, *Escherichia coli*, DNA, deoxyribonucleic acid; ADF, adult T-cell leukemia derived factor; EBV, Epstein–Barr virus; HTLV-1, human T-cell leukemia virus I; NMR, nuclear magnetic resonance; DTT, dithiothreitol; NOE, nuclear Overhauser effect; NOESY, two-dimensional NOE spectroscopy; COSY, two-dimensional correlated spectroscopy; PE-COSY, primitive exclusive correlated spectroscopy; HMQC, heteronuclear multiple-quantum coherence spectroscopy; HMQC-NOESY, relayed ^1H–^{15}N HMQC-NOESY spectroscopy; HMQC-J, purged HMQC; rms, root mean square; SA, simulated annealing.

Human thioredoxin exhibits only a limited amount of sequence identity (25%) with the *E. coli* protein, which has, in part, prompted this detailed structural study of the human enzyme, the most extensive stretch of sequence identity being confined to eight amino acids in the active site. Additional impetus for these studies was provided by the potential immunological role played by human thioredoxin. The protein has been identified and isolated as adult T-cell leukemia derived factor (ADF) in the search for interleukins capable of activating immune cells (Tagaya et al., 1989; Wollman et al., 1988).

Recently, the complete ^1H and ^{15}N resonance assignments of human thioredoxin were presented, as well as the delineation of elements of regular secondary structure (Forman-Kay et al., 1989, 1990). In this paper, the focus is extended to the determination of the high-resolution structure of reduced human thioredoxin using two-dimensional NMR spectroscopy and hybrid distance geometry–simulated annealing calculations. A comparison of the resulting structure with those of oxidized and reduced *E. coli* thioredoxin is also included.

Experimental Procedures

Sample Preparation. Human thioredoxin was expressed and purified from an *E. coli* strain containing a temperature-sensitive repressor and a plasmid with the thioredoxin gene under the control of the λ P_l promoter and the phage Mu *ner* gene ribosome-binding site, as described by Wollman et al. (1988). During the course of our studies we noted, subsequent to nucleotide sequencing of the gene (A. M. Gronenborn and S. Stahl, unpublished data), a discrepancy in the amino acid sequence with respect to that originally reported by Wollman et al. (1988). We found that the sequence contains a Lys at position 39 instead of an Asn. This difference, however, has no direct bearing on either the previous (Forman-Kay et al., (1989) or present NMR studies. The aliphatic chemical shifts for Lys-39 of the 80% N-Met/20% N-Val species are $C^\alpha H$ 4.39 ppm, $C^\beta H$ 2.21 and 2.06 ppm, $C\gamma H$ 1.32 ppm, $C^\delta H$ 1.73 ppm, and $C^\epsilon H$ 2.94 ppm. The corresponding chemical shifts in the 70% N-Val/30% N-Met species are $C^\alpha H$ 4.33 ppm, $C^\beta H$ 2.17 and 2.04 ppm, $C^\gamma H$ 1.33 ppm, $C^\delta H$ 1.73 ppm, and $C^\epsilon H$ 2.94 ppm. Bacteria were grown either in full medium or in minimal medium with $^{15}NH_4Cl$ as the sole nitrogen source. Two species of the protein differing in the presence or absence of N-terminal methionine were produced due to differential posttranslational processing (Forman-Kay et al., 1989). N-Terminal amino acid sequencing determined the ratio of the two species to be 80% N-Met/20% N-Val for protein produced from bacteria grown in full medium and 30% N-Met/70% N-Val for that produced from bacteria grown in minimal medium (Forman-Kay et al., 1989, 1990).

For NMR samples, 2 mM protein was reduced in excess dithiothreitol (DTT), dialyzed against 100–150 mM phosphate buffer containing less than 0.2 mM DTT at pH 7.0 or pH 5.5, lyophilized, redissolved in argon-purged 99.996% D_2O or 90% H_2O/10% D_2O, sealed in tubes with airtight rubber septa, and blanketed with argon for 20 min (Forman-Kay et al., 1989). Although human thioredoxin has five cysteine residues, two in the active site (Cys-32 and Cys-35) and three elsewhere in the sequence (Cys-62, Cys-69, and Cys-73), no disulfide bonds are present, with the exception of that between the active site cysteines in the oxidized form of the protein only. Thus, in the reduced form of the protein studied here, all cysteines exist as free sulfhydryls.

NMR Spectroscopy. All experiments were recorded on a Bruker AM600 spectrometer equipped with digital phase shifters and a "reverse"-mode proton probe. Quadrature in the indirectly detected dimension was achieved throughout by using the time-proportional incrementation method (Redfield & Kuntz, 1975; Marion & Wüthrich, 1983). Examples of the quality of the majority of the spectra recorded have been published previously (Forman-Kay et al., 1989, 1990). Interproton distance restraints were derived from ^1H–^1H NOESY (Jeener et al., 1979; Macura et al., 1981) and relayed ^1H–^{15}N HMQC-NOESY (Gronenborn et al., 1989a,b) spectra. NOESY spectra were recorded at 15, 25, and 40 °C at pH 5.5 and pH 7.0 in H_2O and D_2O at mixing times of 50, 150, and 200 ms. For the 50-ms NOESY experiment, a 5% random variation of the mixing time was used to remove zero-quantum coherence. Typically, 800–1024 t_1 increments of 2K data points were collected for the NOESY spectra and zero-filled in the F_1 dimension to give a final digital resolution of ~8 Hz in both dimensions. The relayed ^1H–^{15}N HMQC-NOESY experiment was recorded in H_2O at 40 °C and pH 5.5; 512 t_1 increments of 2K data points were collected. $^3J_{\alpha\beta}$ coupling constants were measured from a 35° mixing pulse PE-COSY (Mueller, 1987; Marion & Bax, 1988b) experiment recorded in D_2O at 40 °C and pH 5.5. For this experiment, 1024 t_1 increments of 4K data points were collected and zero-filled to 8 Hz/point in F_1 and 0.5 Hz/point in F_2. $^3J_{HN\alpha}$ coupling constants were derived from a version of an HMQC experiment with a purge ^1H 90°$_y$ immediately prior to the acquisition period, known as HMQC-COSY (Gronenborn et al., 1989a) or HMQC-J (Kay & Bax, 1990). This experiment was recorded in H_2O with 1248 t_1 increments of 2K data points to give an acquisition time of 250 ms in the t_1 dimension. Zero filling in F_1 gave a final digital resolution of 0.5 Hz/point in the ^{15}N dimension.

In the case of experiments carried out in H_2O, water suppression was achieved either by coherent saturation or by using an on-resonance jump–return pulse sequence in place of the last 90° pulse (Plateau & Guéron, 1982). Radiation damping was minimized by setting the preparation pulse 45° out of phase with the evolution and detection pulses (Driscoll et al., 1989). Spectra obtained with the jump–return water supression sequence were processed by zeroing the first point, by linearly base-line correcting the free induction decay prior to Fourier transformation in the F_2 dimension, and by linearly base-line correcting the F_2 cross sections before Fourier transformation in the F_1 dimension (Driscoll et al., 1989). For all experiments, base-line distortions were minimized by optimizing the receiver phase and the delay time between the detection pulse and the acquisition period so that the zero- and first-order phase correction parameters for the F_2 Fourier transform were 90° and 180°, respectively (Marion & Bax, 1988a).

Experimental Restraints. The interproton distance restraints were derived from NOESY spectra on the basis of the ^1H and ^{15}N resonance assignments obtained previously with a variety of homonuclear and heteronuclear experiments (Forman-Kay et al., 1989, 1990). For proteins larger than 10 kDa, ambiguities in the assignment of NOE cross peaks arising from substantial chemical shift overlap become increasingly frequent. This problem was resolved primarily by extensive comparisons between NOESY and HMQC-NOESY spectra recorded under a variety of conditions, by making use of data obtained at different temperatures, pH values, and mixing times, and by use of samples with varying degrees of isotopic enrichment and N-terminal processing. For the purposes of deriving distance restraints, NOESY spectra were recorded at 15, 25, and 40 °C in both D_2O and H_2O, with the sample at pH values of 5.5 and 7.0, and with mixing times

of 50, 100, and 200 ms. The pattern and relative intensities of the NOEs observed in the spectra recorded at the three temperatures and two pH values were essentially the same, clearly indicating that no noticeable conformational changes occur over this range of experimental conditions. The small shifts in resonance frequency, however, greatly facilitated the assignment of NOEs in crowded regions of the spectrum. In addition, the NH/aromatic region of spectra recorded in D_2O immediately after lyophilization from H_2O was analyzed to assign cross peaks associated with the more slowly exchanging NH protons.

Attempts were also undertaken to identify NOEs arising from those resonances that are duplicated in the spectra due to the N-terminal heterogeneity (Forman-Kay et al., 1989). By comparison of spectra recorded on the 80% N-Met/20% N-Val sample to those obtained with the 30% N-Met/70% N-Val sample, a duplication of resonances for residues 2–4, 23–25, 40–59, and 101–102 and the aromatic ring protons of residue 77 was revealed. The pattern and relative intensities of the NOE connectivities for both species of thioredoxin, with and without the N-terminal methionine, are very similar. Consequently, the structures of the two forms were considered to be virtually identical, and data from both the 80% N-Met/20% N-Val sample and from the 30% N-Met/70% N-Val sample were utilized. Also, since no assignment of resonances for atoms in the N-terminal methionine could be made, no NOE restraints on the position of this first amino acid were imposed.

All interproton distance restraints were classified into one of four ranges, 1.8–2.7, 1.8–3.3, 1.8–5.0, and 3.0–6.0 Å, corresponding to strong, medium, weak, and very weak NOEs, respectively. Intraresidue and sequential NOE intensities involving NH, $C^{\alpha}H$, and $C^{\beta}H$ protons were derived from a 50-ms NOESY spectrum in order to minimize effects of spin diffusion, while intensities for other NOE cross peaks were obtained from spectra recorded with 100–150 ms mixing times. Upper limits for distances involving methyl protons and nonstereospecifically assigned β-methylene protons were corrected appropriately for center averaging (Wüthrich et al., 1983), and an additional 0.5 Å was added to the upper limits for distances involving methyl protons to account for the higher apparent intensity of methyl resonances (Wagner et al., 1987; Clore et al., 1987).

Stereospecific assignments of β-methylene protons and ϕ, ψ, and χ_1 torsion angle restraints were obtained on the basis of the $^3J_{HN\alpha}$ and $^3J_{\alpha\beta}$ coupling constants, and approximate distance restraints were derived from intraresidue and sequential NOEs involving the NH, $C^{\alpha}H$, and $C^{\beta}H$ protons by means of a conformational grid search with the program STEREOSEARCH (Nilges et al., 1990), as described previously (Kraulis et al., 1989). Two databases were employed in the search procedure, a systematic one with idealized geometry in which the ϕ, ψ, and χ_1 angles of the central residue of a tripeptide fragment were varied in 10° steps and a crystal one comprising tripeptide segments from 34 well-refined X-ray structures. The $^3J_{HN\alpha}$ coupling constants were obtained from an HMQC-J experiment on ^{15}N-labeled protein and corrected for finite line-width and dispersive line-shape contributions as described previously (Forman-Kay et al., 1990), while the $^3J_{\alpha\beta}$ couplings were measured as passive couplings in a PE-COSY spectrum. The minimum ranges employed for ϕ, ψ, and χ_1 torsion angle restraints were ±30°, ±50°, and ±20° (Kraulis et al., 1989).

Stereospecific assignments were obtained initially in this manner for 17 of the 53 nondegenerate β-methylene proton pairs in human thioredoxin. In cases where both $^3J_{\alpha\beta}$ coupling constants were approximately 6–7 Hz, indicative of motional averaging, conformational heterogeneity about the C^{α}–C^{β} bond was assumed (i.e., a mixture of rotamer populations), and no stereospecific assignments were derived. This was the case for 12 of the nondegenerate β-methylene proton pairs in thioredoxin. In addition, stereospecific assignments for the methyl groups of 8 of the 11 Val residues were obtained on the basis of the relative intensity of the NH–$C^{\gamma}H$ and $C^{\alpha}H$–$C^{\gamma}H$ NOEs according to the criteria laid out by Zuiderweg et al. (1985).

All peptide bonds were restrained to be planar and trans, with the exception of that between Thr-74 and Pro-75, which was restrained to be planar and cis on the basis of the observation of a strong $C^{\alpha}H(i)$–$C^{\alpha}H(i + 1)$ NOE between Thr-74 and Pro-75 (Forman-Kay et al., 1989). A cis peptide bond between Ile-75 and Pro-76 has also been reported in the X-ray structure of oxidized *E. coli* thioredoxin (Katti et al., 1990) as well as in the NMR solution structure of the reduced *E. coli* protein (Dyson et al., 1990).

Several restraints for backbone hydrogen bonds were also included in the calculations. The original set of hydrogen-bond restraints included only those found in regions of regular secondary structure (Forman-Kay et al., 1989) that could be assigned unambiguously on the basis of a qualitative interpretation of the NOE and amide exchange data according to the criteria layed out by Wagner et al. (1987). Additional hydrogen bonds to account for slowly exchanging amide protons were added subsequently on the basis of the initial structure calculations (see below). Two distance restraints were used for each hydrogen bond, one between the hydrogen and the acceptor atom of 1.5–2.3 Å and one between the donor heavy atom and the acceptor atom of 2.4–3.3 Å.

Structure Calculations. Three-dimensional structures were computed from the experimental restraints by using the hybrid distance geometry–dynamical simulated annealing method of Nilges et al. (1988a) with the programs DISGEO (Havel & Wüthrich, 1984; Havel, 1986) and XPLOR (Brünger et al., 1987; Brünger, 1988). The latter is derived from the program CHARMM (Brooks et al., 1983) and has been specially adapted for dynamical simulated annealing calculations with NMR restraints (Clore et al., 1985, 1986). All calculations were carried out on Micro VAX III or Stellar GS1000 computers. The resulting structures were visualized on an Evans and Sutherland PS390 graphics system with the programs FRODO (Jones, 1978), XPLOR interfaced with some of the function networks of FRODO, and O (Jones, 1990). Analysis of secondary and supersecondary structure was aided by the program DEFINE-STRUCTURE (Richards & Kundrot, 1988).

The hybrid protocol involves two stages. In the first stage, a set of substructures, comprising about a third of the total number of atoms and including the N, C, C^{α}, $C^{\alpha}H$, C^{β}, and nonterminal C^{γ} and C^{δ} atoms, as well as pseudoatoms for the aromatic rings, is embedded from n-dimensional distance space into Cartesian coordinate space (Havel et al., 1983) without checking triangle inequalities. These substructures are very crude but have approximately the correct overall fold. The remaining atoms are then added with the side chains placed in an extended conformation, and the resulting structures are subjected to the protocol of dynamical simulated annealing described by Nilges et al. (1988a) with a few minor modifications (Clore et al., 1990a; Omichinski et al., 1990). The procedure involves solving Newton's equations of motion subject to a simplified target function that comprises quadratic harmonic terms for bonds, angles, and improper torsion angles

(i.e., planes and chirality restraints), a quartic van der Waals repulsion term, and square-well quadratic potential terms for interproton distance and torsion angle restraints (Nilges et al., 1988a–c). The dynamical simulated annealing protocol employed is designed to overcome large energy barriers along the path toward the global minimum region of the target function and to sample efficiently and comprehensively the conformational space consistent with the experimental restraints. Indeed, during the early stages of the calculations chains can readily pass through one another to attain the correct polypeptide fold. It should also be noted that *no* electrostatic, hydrogen-bonding, or 6–12 Lennard-Jones van der Waals terms are used in the calculations.

Calculational Strategy. In order to make maximal use of the experimental data, an iterative strategy to the structure calculation was employed by carrying out a series of successive calculations with more and more restraints incorporated at each successive stage (Kraulis et al., 1990; Clore et al., 1990a; Omichinski et al., 1990). Thus, analysis of the initial low-resolution structures and successively higher resolution structures obtained after subsequent rounds of simulated annealing permitted many ambiguities in the assignment of NOE cross peaks arising from chemical shift degeneracy to be resolved, thereby allowing the extraction of further approximate interproton distance restraints. Similarly, this approach permitted additional stereospecific assignments to be made, the identification of more hydrogen bonds associated with slowly exchanging amide protons, and the narrowing of torsion angle restraints. The latter was possible when preliminary structures had dihedral angles that fell within only one range of a set of multiple ranges of angles derived from the conformational database search with the program STEREOSEARCH (Nilges et al., 1990).

By this means 15 additional stereospecifically assigned β-methylene proton pairs, as well as stereospecific assignments for all 5 glycine α-methylenes, the 3 other valine γ-methyl groups, and 4 of the 6 leucine δ-methyl groups, were obtained. The glycine α-methylene stereo assignment was based on results from the conformational grid search program using a modified systematic database with glycine as the central residue of the tripeptide fragment (Nilges et al., 1990).

In some cases where the proton resonances were degenerate, but the set of structures had a well-defined conformation in the vicinity of that residue, it was possible to assign the NOE restraints to only one of a pair of protons, even though both protons have the same chemical shift. In fast-flipping Phe and Tyr side chains, the $C^{\delta}H$ and $C^{\epsilon}H$ protons on the two sides of the ring exhibit chemical shift averaging. For well-defined aromatic residues, the NOE distance restraints to the $C^{\delta}H$ and $C^{\epsilon}H$ protons can often be assigned to only one side of the ring, thereby discriminating between the $C^{\delta 1}H$ and $C^{\delta 2}H$ or the $C^{\epsilon 1}H$ and $C^{\epsilon 2}H$ protons. In preliminary structures of thioredoxin, all nine Phe rings and the one Tyr ring had well-defined conformations, with the χ_2 side-chain torsion angle lying in a range of about $\pm 30°$, so that NOE distance restraints could be separated for the $C^{\delta}H$ and $C^{\epsilon}H$ protons on either side of the ring. Likewise, for degenerate β-methylene protons, some of the NOE restraints involving the β-protons could be confined to only one proton of the pair on the basis of the relative distances of these β-protons to other protons nearby in the structure. This restriction of the distance restraints was only made when the structures clearly demonstrated that a well-defined conformation of this side chain existed and only one of the pair of protons was close enough to give rise to a direct NOE connectivity. This was the case for 7 of the 15 degenerate β-methylenes in human thioredoxin. It should be noted that spin diffusion does not complicate this procedure. In the case of the aromatic rings, cross relaxation between the two $C^{\delta}H$ protons and between the two $C^{\epsilon}H$ protons is insignificant, as the distance between these proton pairs is ~ 5 Å. For a β-methylene group, on the other hand, cross relaxation between the two β-methylene protons is highly effective. Thus, in the case of a proton X in close proximity to, say, the $C^{\beta 3}H$ proton but 1–2 Å further away from the $C^{\beta 2}H$ proton, both a direct X–$C^{\beta 3}H$ NOE and a weaker indirect X–$C^{\beta 2}H$ NOE will be observed. However, the sum of these two NOEs, and hence the intensity of the corresponding NOESY cross peak involving these degenerate β-methylene protons, will essentially be unaffected by the spin diffusion process, as the buildup of the indirect NOESY cross peak is accompanied by a concomitant loss of intensity in the direct one. Thus, the restriction of a given NOE restraint to one particular $C^{\beta}H$ proton of a β-methylene pair with degenerate chemical shifts, in cases where the side-chain conformation is well-defined and where the cross-relaxation pathway to one of the $C^{\beta}H$ protons is clearly direct, does not introduce any errors or artifacts and is compatible with a conservative interpretation of the NOE data. It should also be stressed that this procedure does not lead to any change in the resulting conformations but simply increases the precision and accuracy of the structure determination.

It is important to emphasize that since *no* electrostatic and hydrogen-bonding terms are included in the target function for the simulated annealing calculations, the hydrogen bonds that emerge from the structures are not artifacts of the calculations but are the direct result of the experimental restraints. In this manner, 32 of the 33 slowly exchanging NHs were assigned unambiguously to hydrogen bonds. No restraints were used in helical segments in cases where the $C^{\alpha}H$–NH (i, $i + 4$) NOEs were not observed. In total, 15 hydrogen bonds across the core of the β-sheet and 5 within helices were restrained initially, and 3 more at the edges of the β-sheet, 2 across the β-bulge, and 1 in the 91 (O)–93 (NH) irregular turn were added later, based on preliminary structures.

The final structure calculations were based on 1983 approximate interproton distance restraints comprising 478 sequential ($|i - j| = 1$), 353 short-range ($1 < |i - j| \leq 5$), and 463 long-range ($|i - j| > 5$) interresidue restraints; 689 intraresidue restraints; 52 distance restraints for 26 backbone hydrogen bonds; and 241 torsion angle restraints made up of 98 ϕ, 71 ψ, and 72 χ_1 restraints. The distribution of the experimental distance restraints over the structure is illustrated in Figure 1, where the NOE-derived interproton distance restraints and the hydrogen-bonding distance restraints are mapped onto a smoothed backbone ribbon representation of the 3D structure of human thioredoxin. The intraresidue, sequential, and short-range restraints are shown in Figure 1A, while the long-range restraints are illustrated in Figure 1B. The former are distributed evenly along the entire structure, while the latter tend to be concentrated in the core of the protein.

A total of 75 simulated annealing structures converged from 98 different initial distance geometry substructures. Although, in global terms, they all satisfied the experimental restraints, displayed very small deviations from idealized geometry, and had good nonbonded contacts as evidenced by a low value of the van der Waals repulsion term, as well as negative values for the calculated Lennard-Jones van der Waals energy, about 55% of the structures exhibited some structural problems on a local level. Usually this could be traced to a backbone mirror

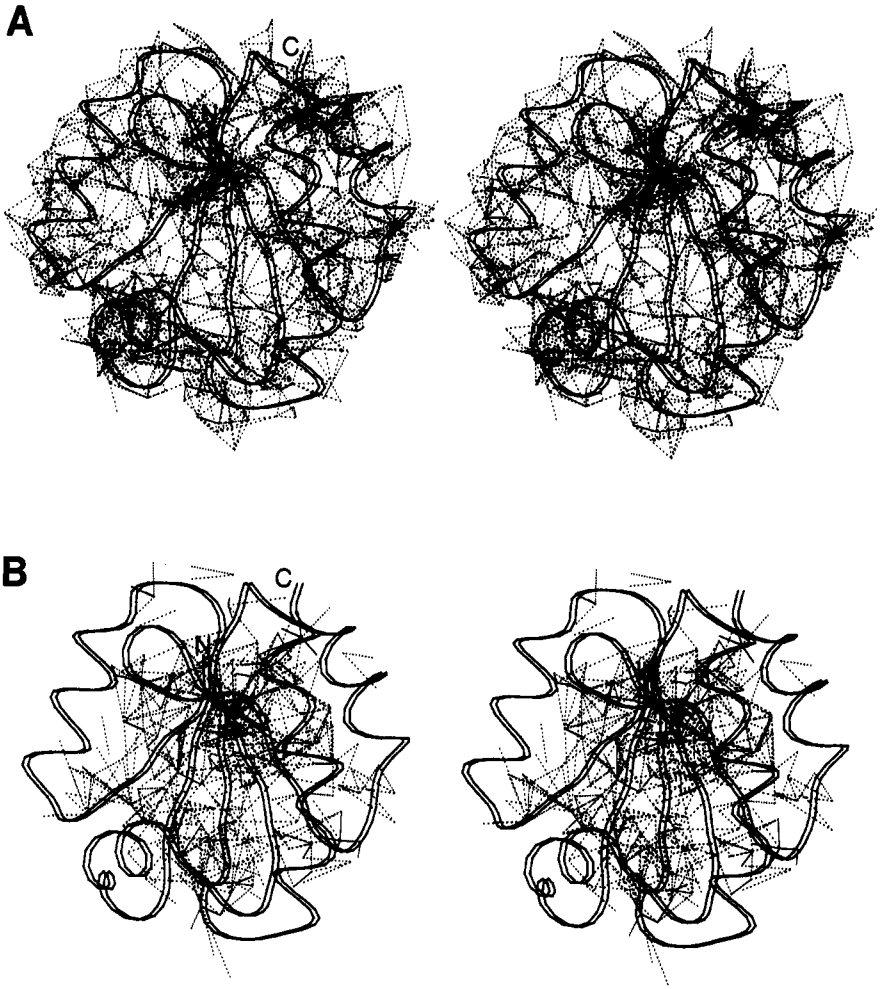

FIGURE 1: Stereoview of the experimental NOE interproton distance restraints and hydrogen-bond distance restraints superimposed on a smoothed backbone ribbon drawing of the restrained minimized mean structure, $(\overline{SA})_r$, of reduced human thioredoxin. (A) Intraresidue, sequential ($|i - j| = 1$), and short-range ($1 < |i - j| \leq 5$) restraints; (B) long-range ($|i - j| > 5$) restraints.

image, with the ψ of one residue and the ϕ of the following residue flipped about 180° from the rest of the structures. At times, this structural problem involved an atypical side-chain conformation at one position. These local structural deviations often occurred at a different location in each of the anomalous structures. For example, an alternate χ_1 rotamer population at a particular residue would be observed in only one of the 75 structures, while another alternate rotamer conformation would be observed for a different residue in another structure. In order to ascertain whether such alternate conformations were consistent with the experimental observations, all interproton distance contacts less than 5 Å that involved these residues were examined. In those cases where an interproton distance contact of less than 2.3–3 Å was predicted from the structure but for which no corresponding NOE could be observed, the particular structure involved was excluded from further analysis. It is also worth noting that these excluded structures also displayed a larger number of small violations and systematically higher rms deviations with respect to the experimental and geometric restraints, as well as systematically higher values of the various terms in the target function. Of the 75 simulated annealing structures calculated, 42 were excluded in this manner. The total number of structures in the final set was thus 33, and these are considered in the structural analysis presented in the Results and Discussion sections. It should be noted, however, that the mean coordinate positions and the overall atomic rms distributions about these positions are not affected in any significant manner by the inclusion or exclusion of the 42 discarded structures.

Results

Structural Statistics. The 33 final SA structures of reduced human thioredoxin exhibit an atomic rms distribution about the mean coordinate positions of 0.40 ± 0.06 Å for the backbone atoms and 0.78 ± 0.05 Å for all atoms, excluding the N-terminal methionine, which is ill-defined. In the case of all the SA structures, the NMR restraints are satisfied within the experimental errors with no interproton distance and torsion angle violations greater than 0.4 Å and 4°, respectively, and the deviations from idealized covalent geometry are very small (Table I). In addition, the Lennard-Jones van der Waals energy (which is not used in the target function for the simulated annealing calculations) is large and negative, with an average value of -387 ± 8 kcal·mol^{-1}, indicating that there are no bad nonbonded contacts.

The overall structure, and especially the core of the molecule, is very well defined, as the statistics presented in Table I demonstrate. The atomic and angular rms distributions for the 33 SA structures, as well as the variation in surface accessibility, are shown in Figure 2 as a function of residue number. Comparison of the secondary structure, indicated at the bottom of Figure 2A, with the rms deviations in atomic positions demonstrates that the regular secondary structure elements are extremely well defined, while the loops between them have a much greater variation in conformation. A noticeable exception to this general observation can be found in

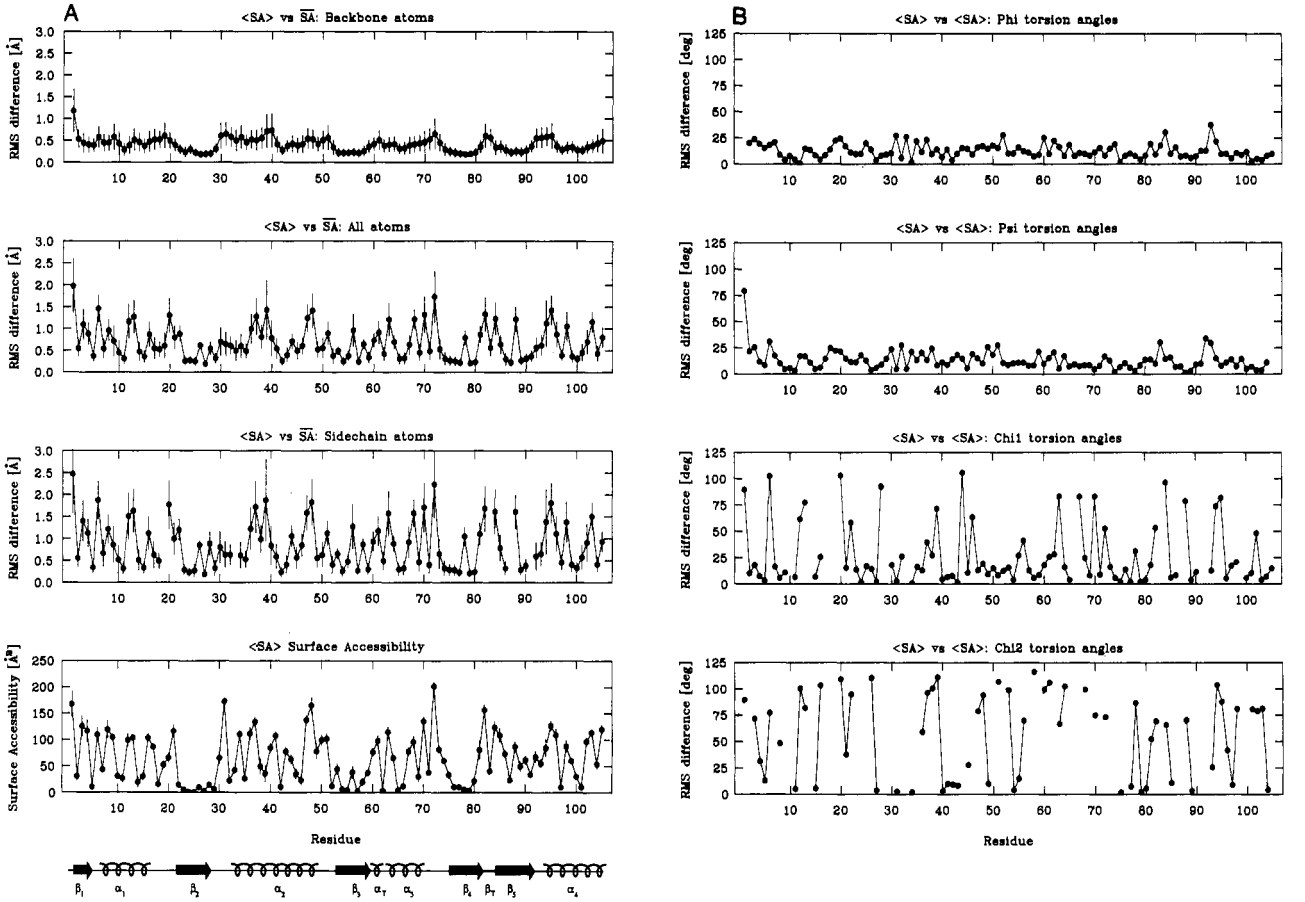

FIGURE 2: (A) Atomic rms distribution of the 33 individual SA structures about the mean structure, $\overline{SA}$, for backbone atoms (N, C^{α}, C, and O), all atoms, and side-chain atoms as a function of residue number, together with the mean surface accessibility of each residue. The mean values are represented as solid circles, and the error bars indicate the standard deviations in these values. The bottom of the figure presents a schematic diagram of the secondary structure of human thioredoxin, with β-strands shown as arrows, helices shown as helical coils, and the thick bar labeled β_T representing the β-hairpin turn. (B) Mean pairwise rms differences of the backbone ϕ and ψ and side-chain χ_1 and χ_2 torsion angles for the 33 SA structures as a function of residue number.

helix α_2, where the central proline, in addition to introducing a bend in the helix, leads to a larger degree of variability.

As is evident from Figure 2A, the high definition of the structures is not confined solely to the backbone. More than 70% of the side chains exhibit atomic rms deviations with respect to the mean of less than 1 Å. Of the less well defined side chains, 10 residues (Lys-3, Gln-12, Glu-13, Lys-39, Lys-48, Gln-78, Lys-81, Lys-82, Glu-95, and Glu-103) display multiple χ_1 side-chain conformations as ascertained on the basis of values of 6–8 Hz for both of the $^3J_{\alpha\beta}$ coupling constants. Twelve additional side chains (Gln-4, Glu-6, Asp-20, Leu-22, Met-37, Ser-44, Glu-47, Glu-56, Gln-63, Glu-68, Glu-88, and Lys-94) had degenerate or near degenerate β-methylene proton chemical shifts. Almost all of the ill-defined side chains exhibit a surface accessibility in the restrained minimized mean structure, $(\overline{SA})r$, of greater than 40% of the accessible surface area of the residue in a tripeptide Gly-X-Gly (Chothia, 1976), indicative of a surface location. Many of these residues are charged and include 10 of the 12 Lys residues, 9 of the 10 Glu residues, and 3 of the 7 Asp residues in the protein. The well-defined side chains exhibit less than 40% surface accessibility, indicating their buried nature.

Figure 2B displays the angular rms differences between the 33 SA structures for backbone ϕ, ψ and side-chain χ_1, χ_2 torsion angles. The overall angular rms deviation for ϕ and ψ torsion angles is only 10° and 11°, respectively. All backbone torsion angles of the non-glycine residues also fall in the generally allowed regions of the Ramachandran ϕ, ψ plots for all 33 structures (Figure 3), with the exception of Glu-70 (with

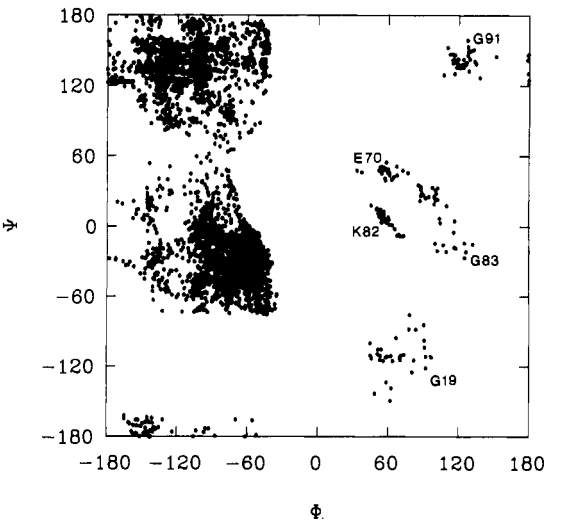

FIGURE 3: Ramachandran ϕ, ψ plot for the 33 SA structures.

ϕ = 57°) and Lys-82 (with ϕ = 56°), which lie in the left-handed α_L-helix region. The latter region is rarely occupied by residues other than Gly or Asn. Glu-70 is involved in a turn responsible for reversing the direction of the polypeptide chain at the end of an α-helix, while Lys-82 is located at the i+1 position of a type I' β-hairpin turn from Lys-81 to Gln-84.

Description of the Overall Structure. The overall structure of human thioredoxin consists of a five-stranded β-sheet surrounded by four α-helices and a helical turn. Figure 4

Table I: Structural Statistics and Atomic RMS Differences[a]

(A) Structural Statistics

	⟨SA⟩	(SA)r
RMS deviations from exptl distance restraints (Å)[b]		
all (2035)	0.025 ± 0.001	0.028
sequential ($\|i-j\| = 1$) (478)	0.032 ± 0.002	0.034
short range ($1 < \|i-j\| \leq 5$) (353)	0.019 ± 0.002	0.020
long range ($\|i-j\| > 5$) (463)	0.022 ± 0.002	0.030
intraresidue (689)	0.024 ± 0.002	0.026
H-bond (52)[c]	0.024 ± 0.002	0.027
RMS deviations from exptl dihedral restraints (deg) (241)[b]	0.231 ± 0.034	0.156
F_{NOE} (kcal·mol^{-1})[d]	40 ± 2	47
F_{tor} (kcal·mol^{-1})[d]	0.80 ± 0.24	0.36
F_{repel} (kcal·mol^{-1})[d]	19 ± 2	39
$E_{L\text{-}J}$ (kcal·mol^{-1})[e]	−387 ± 8	−368
Deviations from idealized covalent geometry		
bonds (Å) (1644)	0.004 ± 0	0.005
angles (deg) (2980)	1.870 ± 0.002	1.879
impropers[f] (deg) (637)	0.506 ± 0.006	0.516

(B) Atomic RMS Differences (Excluding Residue 1)

	atomic rms differences (Å)	
	backbone atoms	all atoms
⟨SA⟩ vs $\overline{SA}$	0.40 ± 0.06	0.78 ± 0.05
(SA)r vs $\overline{SA}$	0.13	0.37
⟨SA⟩ vs (SA)r	0.42 ± 0.05	0.87 ± 0.06
internal residues only[g]		
⟨SA⟩ vs $\overline{SA}$	0.34 ± 0.04	0.50 ± 0.07
(SA)r vs $\overline{SA}$	0.12	0.24
⟨SA⟩ vs (SA)r	0.36 ± 0.04	0.55 ± 0.08

[a] The notation is as follows: ⟨SA⟩ are the 33 final simulated annealing structures; $\overline{SA}$ is the mean structure obtained by averaging the coordinates of the 33 individual structures, best fitted to each other (excluding residue 1); (SA)r is the restrained minimized mean structure obtained by restrained minimization of the mean structure, $\overline{SA}$. The number of terms for the various restraints is given in parentheses. [b] None of the structures exhibited distance violations greater than 0.4 Å or dihedral angle violations greater than 4°. [c] For each hydrogen bond there are two restraints: $r_{NH\text{-}O}$, 1.7–2.3 Å, and $r_{N\text{-}O}$, 2.4–3.3 Å; thus 52 distance restraints were used for 26 restrained hydrogen bonds. [d] The values of the square-well NOE and torsion angle potentials [cf. eqs 2 and 3 in Clore et al. (1986b)] are calculated with force constants of 50 kcal·mol^{-1}·Å^{-2} and 200 kcal·mol^{-1}·rad^{-2}, respectively. The value of the quartic van der Waals repulsion term [cf. eq 5 in Nilges et al. (1988a)] is calculated with a force constant of 4 kcal·mol^{-1}·Å^{-4} with the hard-sphere van der Waals radii set to 0.8 times the standard values used in the CHARMM empirical energy function (Brooks et al., 1983). [e] $E_{L\text{-}J}$ is the Lennard-Jones van der Waals energy calculated with the CHARMM (Brooks et al. 1983) empirical energy function. It is not included in the target function used during the simulated annealing. [f] The improper torsion angles are used to maintain planarity and chirality. All peptide bonds are restrained as planar and trans with the exception of that between Thr-74 and Pro-75, which is restrained to be planar and cis (see Experimental Procedures). [g] Internal residues are defined as those with less than 40% of the surface of the residue in a tripeptide Gly-X-Gly (Chothia, 1976) accessible in the restrained minimized mean structure, (SA)r, of human thioredoxin.

shows the superposition of the backbone (N, C^{α}, and C) atoms for the 33 SA structures, in two different stereoviews. The approximate dimensions of the protein are 30 Å × 30 Å × 20 Å, excluding side chains. About 90% of the residues in human thioredoxin are involved in regular secondary structure or turns involving hydrogen bonds. The second and last α-helices are almost parallel to each other and lie on one face of the twisted sheet, while the other helices fall on the other side of the sheet. This packing arrangement is evident from Figure 4B, which is a view along the edge of the sheet with the helices above and below the twisted plane of the sheet. The active site forms a protruding loop from the end of the second β-strand at Ala-29 through the start of the second helix and can be seen in the lower right-hand corner of Figure 4A.

Hydrogen Bonding. A hydrogen bond is considered to be present if it is found in a substantial number of the individual structures. The definition used for identification of a hydrogen bond is that the distance between the acceptor and donor heavy atoms is less than 3.4 Å and the angle between them is greater than 110°. In the individual structures, 71 ± 3 hydrogen bonds were found, with 72 present in the restrained minimized average structure. Backbone–backbone hydrogen bonds in all structures were located in the β-sheet and α-helices. The mean value of the N–O backbone hydrogen-bond distances in the restrained minimized mean structure, (SA)r, is 3.02 ± 0.22 Å in the β-sheet and 2.97 ± 0.25 in the α-helices. In addition, a number of side-chain to backbone hydrogen bonds were found to occur in more than 50% of these structures. These involved hydrogen bonds between the $N^{\delta 1}H$ imidazole proton of His-43 and its own backbone carbonyl oxygen atom, the side-chain carboxyl group of Asn-64 and the backbone amide protons of either or both Asp-64 and Val-65, and the side-chain carboxyl group of Asp-58 and the amide protons of either or both Thr-30 and Asp-60.

β-Sheet. The core of human thioredoxin consists of a five-stranded β-sheet. The individual strands comprise residues from Val-2 to Ile-5 (β_1), Leu-22 to Ser-28 (β_2), Ile-53 to Val-59 (β_3), Pro-75 to Lys-81 (β_4), and Gln-84 to Gly-91 (β_5). The sheet has the characteristic right-handed twist, defined in terms of the hydrogen-bonding direction or the peptide planes viewed along a strand (Chothia et al., 1977), with the outermost strands at an angle of about 100° to each other. Both parallel (p) and antiparallel (a) arrangements of the strands are present and display the pattern $\beta_1(p)\beta_3(p)\beta_2(a)\beta_4(a)\beta_5$, which is illustrated in Figure 5A. The β-hairpin around Lys-81 to Gln-84 and the β-bulge interrupting β_5 at Val-86 to Gly-87 are also shown in Figure 5B. All the hydrogen bonds within the β-sheet, the β-hairpin turn, and the β-bulge are present in 100% of the structures, with the exception of the three hydrogen bonds located in β_1 between the amide groups of Lys-3, Ile-5, and Glu-56 and the carbonyl oxygen atoms of Phe-54, Glu-56, and Lys-3, respectively, which are observed in 50%–90% of all structures. A hydrogen bond between Val-59 (NH) and Ser-28 (O) indicates that Val-59 is a part of β_3, although the dihedral angles of this residue are not characteristic of the β-structure. Indeed, Val-59 is located at the end of the β-strand where the polypeptide chain forms a helical turn comprising residues Asp-58 to Cys-62 with a 5–1 hydrogen bond between Asp-58 (O) and Cys-62 (NH). The β-bulge (Richardson, 1981) in strand β5 is formed by Val-86 and Gly-87 hydrogen bonding to Phe-79 on the opposing strand.

Helices. All four helices in the structure of thioredoxin lie on the periphery of the β-sheet region. The helical structure comprising Asp-58 to Cys-62 is more aptly described as a helical turn rather than a helix. Helix 1 extends from Ser-7 to Ala-17 (α_1), helix 2 from Gly-33 to Tyr-49 (α_2), helix 3 from Cys-62 to Glu-70 (α_3), and helix 4 from Lys-94 to Val-105 (α_4). The numbering of helices is slightly different from that used by Holmgren et al. (1975) and Katti et al. (1990) for the structure of *E. coli* oxidized thioredoxin. The human α_3 helix corresponds to the *E. coli* 3_{10} helix, and the helical turn in human thioredoxin has its equivalent in α_3 of the *E. coli* structure. Illustrations of all four helices and the helical turn, indicating their hydrogen-bonding pattern, are presented in Figure 6.

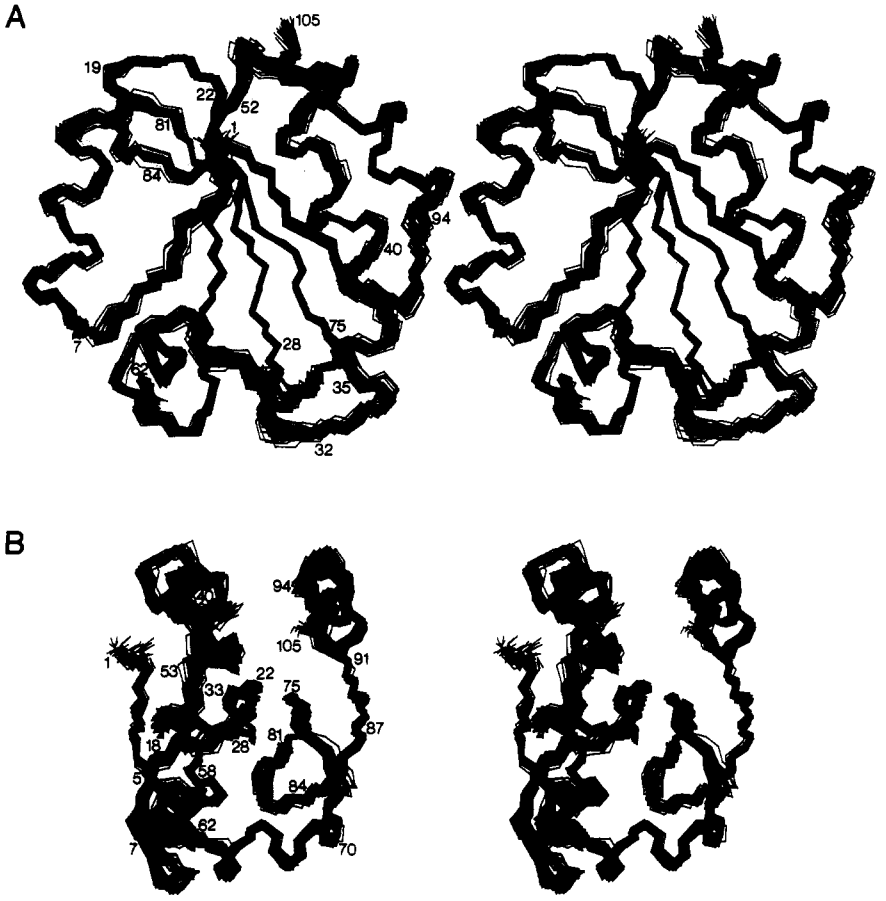

FIGURE 4: Two stereoviews showing the best-fit superpositions of the backbone (N, C^{α}, and C) atoms of the 33 SA structures of reduced human thioredoxin. (A) View toward the five strands of the twisted β-sheet with the surrounding helices. (B) View down the edge of the sheet, illustrating helices α_2 and α_4 interacting in a parallel orientation on one side and the other helices located on the opposite side.

Helix α_1 is very well defined and regular with 5–1 hydrogen bonds occurring all the way along the helix in over 75% of all structures, except near the end where the hydrogen bond between Leu-15 (NH) and Phe-11 (O) is not present in most structures. At this point the helix becomes more tightly wound and 3_{10}-like with bifurcated 5–1 and 4–1 hydrogen bonds. Helix α_2, on the other hand, is distorted and exhibits no continuous hydrogen-bonding pattern. The N-terminal residues of the α_2 helix display bifurcated 5–1 and 4–1 hydrogen bonding, and many of the hydrogen bonds along the length of the rest of this helix are 4–1. Well-defined helical hydrogen bonds are found between His-43 (NH) and Lys-39 (O) and at the C-terminal end of α_2. There is a bend of approximately 30° in the middle of the α_2 helix at Pro-40, rendering the entire helix curved. The other proline in this helix, Pro-34, is accommodated more easily into the helix due to its N-terminal position. The helical turn at Asp-58 to Cys-62 includes one 5–1 hydrogen bond, present in 100% of the structures, and immediately precedes α_3, with an angle of ~90° between the two helices. The dihedral backbone angles in this region are not characteristic of an α-helix, since a backbone hydrogen bond is formed between Val-59 (NH) and the backbone carbonyl oxygen of Ser-28 in the adjacent β-strand. Thus Val-59 and Asp-58 can be regarded as belonging to both the β_3 strand and the helical turn. Helix α_3 is well-defined and displays regular helical hydrogen-bonding patterns in well over 75% of the structures, and helix α_4 exhibits a slight bend at the beginning with a 4–1 hydrogen bond in 75% of the structures that is followed by regular hydrogen-bonding patterns in 100% of the structures for the remainder of α_4. The C-terminal end of the polypeptide chain appears to tighten into a 3_{10} helix with a few bifurcated 5–1 and 4–1 interactions.

β-Turns. Four β-turns are found in the human thioredoxin structure, some of which are distorted to varying degrees. The first turn is an irregular type I reverse turn located at Ser-7 to Ala-10, which exhibits bifurcated 4–1 and 5–1 hydrogen bonds between Ser-7 (O) and Ala-10 (NH) and between Ser-7 (O) and Phe-11 (NH). In addition, a stabilizing interaction between the O^{γ} atom of Ser-7 and the NH of Ala-10 is present in a few of the structures. The turn from Tyr-49 to Val-52 between the end of α_2 and the beginning of β_3 is also an irregular turn with a possible hydrogen bond between the NH of Val-52 and the carbonyl oxygen of Tyr-49. This turn is illustrated in Figure 7. A type VIb reverse turn occurs at Cys-73 to Thr-76, involving the cis peptide bond at Pro-75, following a region with extended conformation from Glu-70 to Cys-73 and leading into β_4. Two potential stabilizing interactions can also be observed between either Cys-73 (O) or Pro-75 (O) and the $O^{\gamma}H$ proton of Thr-76. The β-hairpin turn from Lys-81 to Gln-84 is of the type I′ variety, with Gly-83 in the common $i+2$ position for glycine residues and hydrogen bonds between the NH of Lys-81 and the carbonyl oxygen of Gln-84, as well as between the NH of Gln-84 and the carbonyl oxygen of Lys-81 (Figure 5B). An unusual turn located between β_5 and α_4 comprising Gly-91 to Lys-94 can also be observed, exhibiting a hydrogen bond between Gly-91 (O) and Asn-93 (NH). A conformational search of high-resolution X-ray structures (Jones & Thirup, 1986), however, revealed that 20 tetrapeptide fragments in 17 different X-ray structures could be superimposed on residues 91–94 of human thioredoxin with backbone atomic rms differences ranging from 0.045 to 0.13 Å.

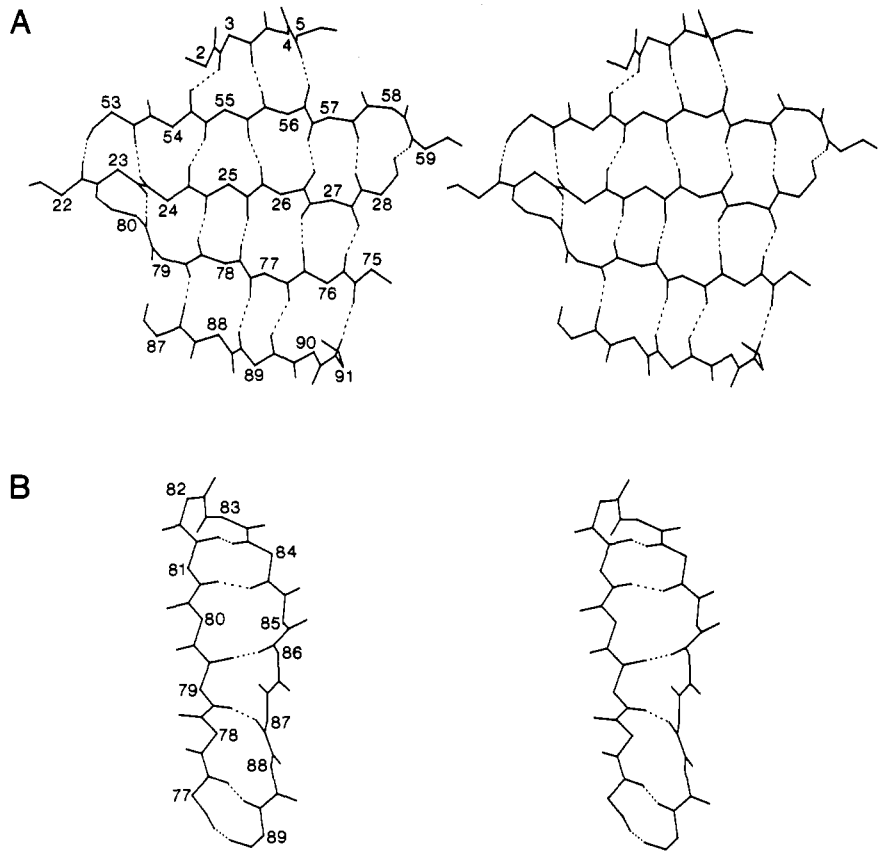

FIGURE 5: Stereoviews of the β-sheet conformation and hydrogen-bonding pattern seen in the restrained minimized mean structure, $(\overline{SA})r$, of human thioredoxin. (A) Major portion of the sheet including β_1, β_2, β_3, and parts of β_4 and β_5. (B) The β-hairpin turn and β-bulge with portions of β_4 and β_5. All backbone atoms (N, NH, C^{α}, C, and O) are shown, and the hydrogen bonds are shown by dashed lines. Residues are labeled near the C^{α} position.

Active Site. Thioredoxin has an unusual active site located in a protruding loop on the exterior of the protein rather than in a cleft, which is more commonly the case. This is necessitated by its functional role in the reduction of other protein substrates. The active site loop comprises the stretch of residues from Ala-29 to Lys-36, beginning after β_2 and extending into the first turn of α_2. Figure 8 illustrates the conformation of residues in the active site of the human thioredoxin structure around the highly conserved region Cys-32–Gly-33–Pro-34–Cys-35. The present structure contains the two cysteines in their reduced, sulfhydryl form, with no disulfide bridge between them. They are located across the helical loop formed by the first few residues of α_2, which are more tightly wound than in a regular helix, including bifurcated 5–1 and 4–1 hydrogen bonds from the amide protons of Lys-36 and Met-37 to the backbone carbonyls of Gly-33 and Pro-34, respectively (see Figure 6B). The sulfhydryl groups are not pointing exactly toward each other but would only require a slight change from the observed value of $-60 \pm 12°$ for the χ_1 torsion angle of Cys-35 to allow the formation of an oxidized disulfide bridge. The χ_1 angle of Cys-32 is $-179 \pm 19°$, and the average distance between the S^{γ} atoms of Cys-32 and Cys-35 is 3.1 $\pm$ 0.2 Å. The residues involved in the active site do not interact extensively with residues in distant regions of the protein, leading to fewer long-range NOE connectivities and fewer packing constraints on the structure than for most of the protein molecule. This necessarily produces a reduction in precision of the conformation in this region, which is evident from Figure 4A where the active site region is located in the lower right-hand corner of the structure.

Packing of Secondary Structural Elements. The core of thioredoxin is a β-sheet that is flanked on either side by the four long helices and the helical turn. These helices are amphipathic and interact with the predominantly hydrophobic sheet along their hydrophobic sides, leaving polar and charged side chains exposed on the surface. Exceptions to this rule include three acidic side chains, Asp-26, Glu-56, and Asp-58, that are involved in stabilizing interactions outlined in the Discussion section.

Helices α_1, α_2, and α_4 are oriented approximately parallel to the direction of the sheet (see Figure 4A), with helices α_2 and α_4 on one side of the sheet (see Figure 4B) and the long axis of the second half of the bent α_2 helix at an angle of about 30–40° to the long axis of α_4. These two helices show many of the characteristics of coiled-coil helix interactions (McLachlan & Stewart, 1975). In particular, the first and fourth residues in the seven-residue repeats are hydrophobic amino acids involved in helix–helix interactions, while the fifth and seventh residues are polar or charged. This interaction extends from Ile-38 to Tyr-49 of α_2 and from Lys-94 to Ile-101 of α_4, with numerous NOE connectivities observed between these stretches of polypeptide chain. The axial separation between helices α_2 and α_4 is approximately 8 Å. Helices α_1, α_3, and the helical turn lie on the opposite side of the sheet, are almost at right angles, and are not involved in many direct interactions with each other.

Side Chains. Thioredoxin contains 29 charged amino acids, comprising 10 Glu, 7 Asp, and 12 Lys residues, giving the protein a net negative charge. Six acidic residues, Asp-20, Glu-56, Glu-68, Glu-88, Glu-98, and Glu-95, are involved in potential salt bridges that were observed in the NMR structures. The carboxylate group of Asp-20 is in close proximity to the $N^{\zeta}H_3^+$ groups of Lys-81 or Lys-82; Glu-56 can interact with Lys-36 or Lys-39, Glu-68 with Lys-8, and Glu-88 with

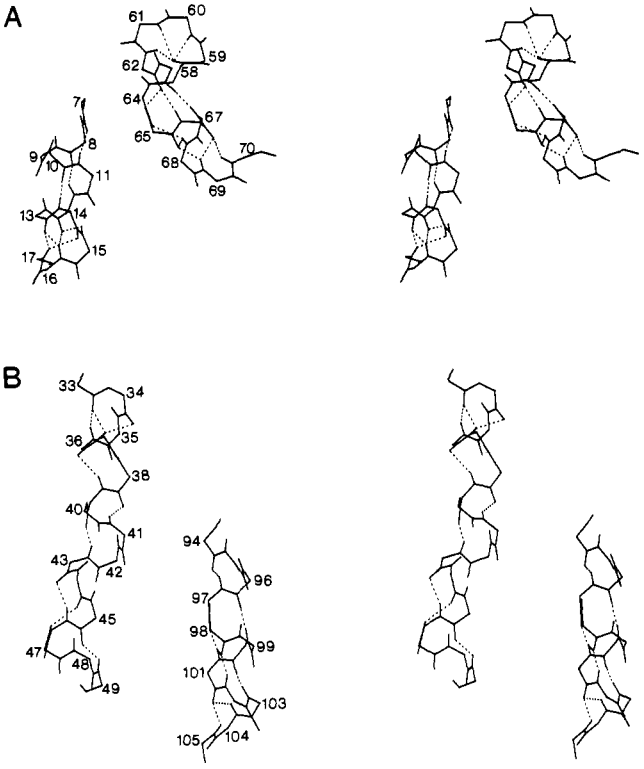

FIGURE 6: Stereoviews of the backbone (N, NH, C^{α}, C, and O) conformation of the helices in the restrained minimized mean structure, $(\overline{SA})r$, of human thioredoxin illustrating the hydrogen-bonding pattern. α_1, the short helix, and α_3 are shown in (A), while α_2 and α_4 are displayed in (B). Note the distortions in α_2 and the bend induced at the position of Pro-40. The hydrogen bonds are shown by dashed lines. Residues are labeled near the C^{α} position.

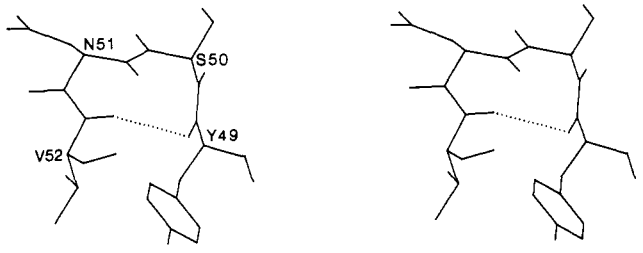

FIGURE 7: Stereoview of the reverse turn from Tyr-49 to Val-52 in the restrained minimized mean structure, $(\overline{SA})r$, of human thioredoxin. Residues are labeled near the C^{α} position.

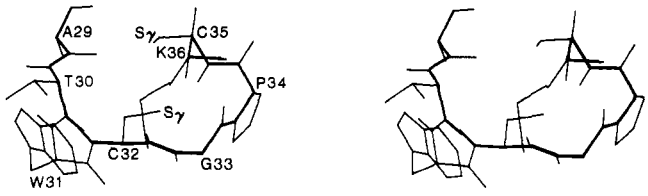

FIGURE 8: Stereoview of the active site loop, from Ala-29 to Lys-36 in the restrained minimized mean structure, $(\overline{SA})r$, of human thioredoxin. The backbone (N, NH, C^{α}, C, and O) atoms and the side-chain heavy atoms are shown, and the path of the backbone N, C^{α}, and C atoms is highlighted by a thicker line. Residues are labeled near the C^{α} position. Note the proximity of the S^{γ} atoms of the two cysteine residues, Cys-32 and Cys-35.

Lys-85. Although not seen in all structures, Glu-95 and Glu-98 can potentially interact with Lys-94 and Lys-48, respectively.

Most of the aromatic residues, including nine phenylalanines, one tyrosine, one tryptophan, and one histidine, are located in the protein core. Trp-31 is solvent exposed, being part of the protruding loop region of the active site, and Phe-41 and His-43 are located on the exposed surface of α_2. As noted previously, the histidine side chain is held in position by an intraresidue hydrogen bond between the $N^{\delta 1}H$ of the imidazole ring and the backbone carbonyl oxygen. Figure 9 illustrates two views of the protein showing residues within the core and some residues on the surface of the human thioredoxin structure, demonstrating that the interior side chains are as well-defined as the backbone atoms. Figure 9A includes all of the aromatic residues except Trp-31. The internal aromatic side chains interact with each other as well as with other large hydrophobic residues such as Val, Leu, Ile, and Thr, shown in Figure 9B, forming a tightly packed hydrophobic core. Most of the internal aromatic rings are packed in an approximately mutually orthogonal arrangement comprising two groups: Phe-11, Phe-27, and Phe-80 on the one hand and Phe-42, Tyr-49, Phe-54, Phe-77, Phe-79, and Phe-80 on the other.

Discussion

Overall Structure Quality. The determination of the solution structure of reduced human thioredoxin presented here demonstrates that it is possible to obtain high-resolution structures of proteins larger than 10 kDa by employing current NMR technology. The general strategy used relies on the exploitation of spectra obtained under a large number of different experimental conditions combined with an iterative approach to resolve ambiguities and to obtain as many stereospecific assignments as possible. Heteronuclear techniques have been employed to obtain accurate $^3J_{HN\alpha}$ coupling constants that in conjunction with $^3J_{\alpha\beta}$ coupling constant information and approximate intraresidue and sequential distance restraints allow one to obtain stereospecific assignments of a large number of β-methylene protons and γ-methyls of valines, together with an extensive set of ϕ, ψ, and χ_1 torsion angle restraints, by means of conformational grid searches. Frequently, further stereospecific assignments of specific NOEs can be made after inspection of an earlier set of structures. This was possible in the present structure determination for the α-methylene protons of glycine residues and a large number of aromatic side-chain $C^{\delta}H$ and $C^{\epsilon}H$ protons, as well as some degenerate β-methylene protons. The primary goal of all of these approaches is to increase the number of experimental restraints used in the structure calculations, since the precision and accuracy of the final structure is directly correlated to the number of restraints (Gronenborn & Clore, 1990; Clore & Gronenborn, 1991). Altogether, 1983 interproton distance and 241 torsion angle restraints were used in the structure determination, resulting in an average of 18.9 interproton distance and 2.3 torsion angle restraints per residue, in addition to restraints for the 26 hydrogen bonds. An additional calculation of the average number of NOEs affecting each residue's conformation can be carried out that treats intraresidue and interresidue NOEs separately in order to take into account the fact that interresidue NOEs affect two residues, while intraresidue NOEs affect only a single residue. This yields an average of 31.2 NOEs affecting the conformation of each residue.

This large number of experimental restraints resulted in a well-defined high-resolution structure for human thioredoxin with an atomic rms distribution about the mean coordinate positions of 0.40 Å for the backbone atoms and 0.78 Å for all atoms. As evidenced from Table I, which lists the structural statistics, all structures satisfy the experimental restraints within the experimental errors, and no distance violations >0.4 Å were observed. In addition, good covalent geometry is observed for bonds, angles, planes, and chiral centers, and there are no bad nonbonded contacts.

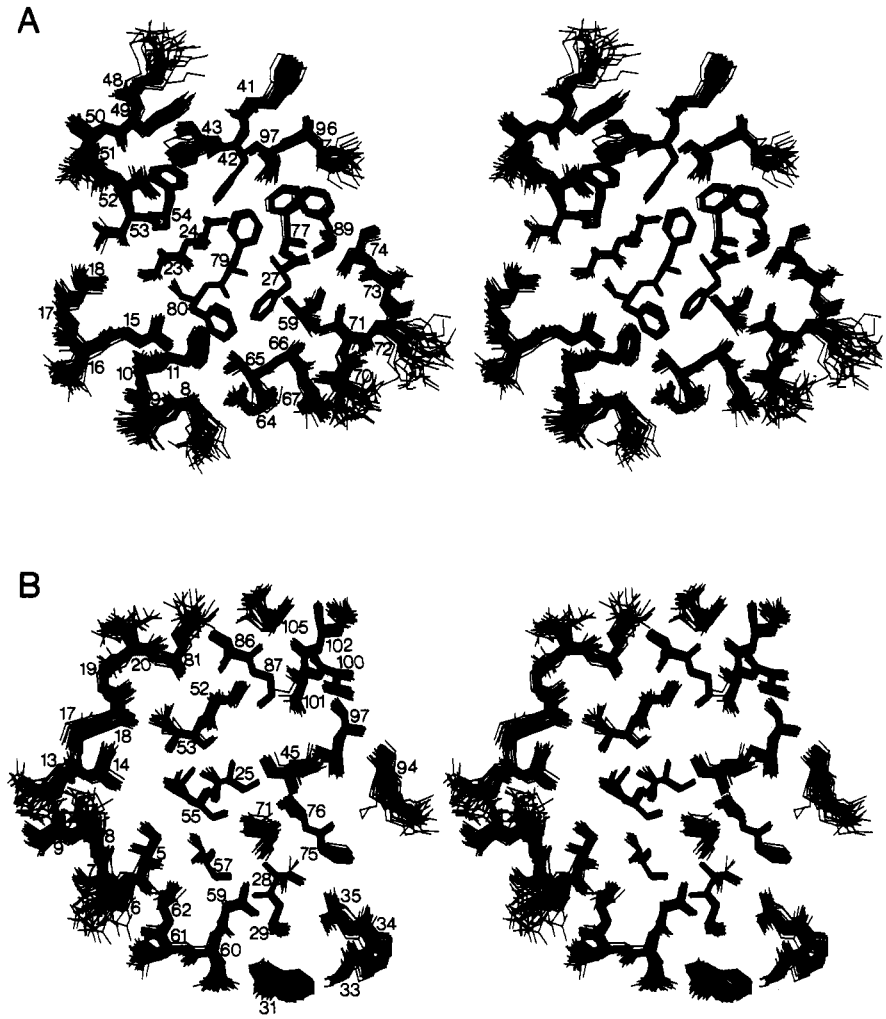

FIGURE 9: Two stereoviews showing best-fit superpositions of all atoms (excluding protons) of the 33 SA structures showing selected residues from the core and surface of human thioredoxin. (A) Aromatic residues, mostly located within the hydrophobic core and packed in a mutually orthogonal arrangement, including Phe-11, Phe-27, Phe-41, Phe-42, His-43, Phe-54, Tyr-49, Phe-77, Phe-79, Phe-80, and Phe-89, with additional residues Lys-8, Thr-9, Ala-10, Leu-15, Asp-16, Ala-17, Ala-18, Val-23, Val-24, Lys-48, Ser-50, Asn-51, Val-52, Ile-53, Val-59, Asp-64, Val-65, Ala-66, Ser-67, Glu-70, Val-71, Lys-72, Cys-73, Thr-74, Lys-96, and Leu-97. (B) Residues Ile-5, Glu-6, Ser-7, Lys-8, Thr-9, Glu-13, Ala-14, Ala-17, Ala-18, Gly-19, Asp-20, Val-25, Ser-28, Ala-29, Trp-31, Gly-33, Pro-34, Cys-35, Leu-45, Val-52, Ile-53, Leu-55, Val-57, Val-59, Asp-60, Asp-61, Cys-62, Val-71, Pro-75, Thr-76, Lys-81, Val-86, Gly-87, Lys-94, Leu-97, Thr-100, Ile-101, Asn-102, and Val-105.

As pointed out previously, the resonances of a number of residues exhibit chemical shift duplication arising from the presence or absence of the N-terminal methionine. These involve residues 2–4, 23–25, 40–59, and 101–102 and the aromatic ring of Phe-77 (Forman-Kay et al., 1989, 1990). With the structure of human thioredoxin in hand, it is of interest to establish the spatial relationship between these residues and the site of origin for their different chemical shifts at the N-terminus. The C^{α}–C^{α} distances between residue 2 and residues 23–25 range from 9.2 to 10.4 Å, residues within the stretch 40–59 are up to 11.5–16.3 Å away, and residues 101–102 are at a distance of 15.2 and 18.3 Å. Given that the pattern and relative intensities of the NOEs involving the duplicated resonances are the same within experimental error, this demonstrates quite dramatically that chemical shifts are extremely sensitive to very small structural and/or electronic changes and that effects can be transmitted across very large distances. All these residues are located in connected secondary structure elements. Thus, residues 2–4 in strand β_1 are hydrogen bonded to residues in strand β_3 (residues 53–59), which in turn are hydrogen bonded to residues 23–25 in the next β-strand (β_2) (see Figure 5A). The aromatic ring of Phe-77 overlays residues 24–25. Residues 40–49 are located in helix α_2, which precedes β_3 and in turn interacts with helix α_4 (Figure 6B), which contains residues 101–102. Thus any effects can be efficiently transmitted via a concerted action involving all of these residues.

Comparison with the X-ray Crystal Structure of Oxidized and the NMR Solution Structure of Reduced E. coli Thioredoxin. It is interesting to compare the present solution structure of reduced human thioredoxin with that of the other known thioredoxins, in order to investigate the relationship between structural elements and primary sequence. Three other thioredoxin structures have been reported: the X-ray structure of T4 thioredoxin (Söderberg et al., 1978), the X-ray structure of oxidized *E. coli* thioredoxin (Holmgren et al., 1975; Katti et al., 1990), and the NMR solution structure of reduced *E. coli* thioredoxin (Dyson et al., 1990). Since the primary sequence and tertiary structure of the *E. coli* enzyme are closest to those of the human enzyme, with T4 thioredoxin being more distant, the comparison has been restricted to *E. coli* thioredoxin and primarily to the refined X-ray structure, since the coordinates of the NMR structure are not yet available from the Brookhaven Protein Data Bank.

The X-ray structure of oxidized *E. coli* thioredoxin contains two asymmetrically positioned protein molecules in the crystal unit cell (Katti et al., 1990), which differ by 0.66 Å in C^{α} positions. The NMR structure of reduced human thioredoxin

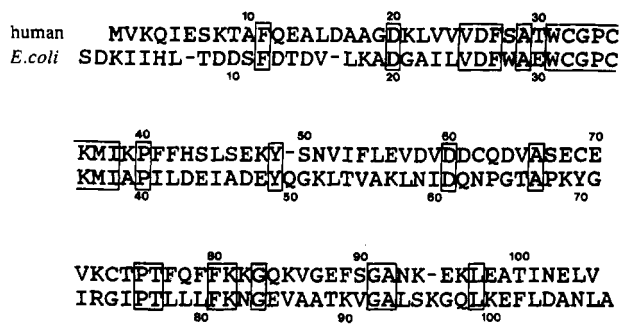

FIGURE 10: Sequence alignment of human and *E. coli* thioredoxin based on a structural superposition of the two proteins. The C^{α} atomic rms difference for this alignment, excluding residues 6 and 16–19, is ~1.5 Å.

was superimposed on each of the two molecules in the X-ray structure by using the program O (Jones, 1990), yielding an rms deviation for the C^{α} positions of 1.49 and 1.54 Å for molecules A and B, respectively. Almost all of the residues could be superimposed, comprising 100 of human thioredoxin's 105 residues and excluding only Glu-6, Asp-16, Ala-17, Ala-18, and Gly-19. The sequence alignment between human and *E. coli* thioredoxin based on the result of the structural superposition is shown in Figure 10. In this alignment there are two single amino acid deletions in the *E. coli* sequence and two in the human sequence. This high degree of structural homology contrasts markedly with the rather moderate degree of amino acid sequence identity, which is only 25%. When conservative amino acid changes are taken into account, however, the sequence similarity can increase within the range of 30%–55%, depending on the definition of "conservative", more in line with the structural homology of the two proteins. If only internal residues with surface accessibilities of less than 50 $Å^2$ are considered, comprising 43 of the 105 residues of human thioredoxin, the sequence identity is 30%, and the sequence similarity is more than twice that value when conservative changes are allowed.

Although the overall tertiary structures superimpose rather well, distinct differences between the structures can be observed. Human thioredoxin is shorter, lacking residues corresponding to the first two and the last amino acids of *E. coli* thioredoxin and containing single amino acid deletions in the loops connecting α_2 with β_3 and β_5 with α_4. There are two primary regions of significant structural difference located in the first helix, α_1, and in the helix identified as α_3 in human thioredoxin and as the 3_{10} helix in the *E. coli* structure. The human α_1 helix (Ser-7 to Ala-17) is a full turn longer than the corresponding helix in *E. coli* thioredoxin (residues Ser-11 to Leu-17) and has a more regular helical geometry. The C-terminal end of the helix and the loop extending from the helix to the second β-strand include the only significant stretch of residues of human thioredoxin that did not superimpose well on *E. coli* thioredoxin. The helix labeled α_3 in human thioredoxin (Cys-62 to Glu-70) is also longer than the corrersponding 3_{10} helix of the *E. coli* protein (Thr-66 to Ile-70). This is most likely due to the presence of a proline in *E. coli* thioredoxin at position 64 that prevents helix formation earlier in the sequence. In addition, helix α_3 in *E. coli* thioredoxin contains a proline at position 68 in the center of the helix that introduces the irregular hydrogen-bonding patterns that lead to its designation as a 3_{10} helix. The complete absence of prolines in the human sequence for this stretch of amino acids allows a regular α-helical geometry to propagate.

The two amino acids, Glu-70 (ϕ = 57°) and Lys-82 (ϕ = 56°), that exhibit ϕ angles in the unusual left-handed α_L-helical region of the Ramachandran plot are both located in turns, and it is interesting to note that the corresponding residues, Gly-71 and Asn-83, in the X-ray structure of oxidized *E. coli* thioredoxin have similar ϕ angles of ~70° and ~50°, respectively (Katti et al., 1990).

The tightly packed hydrophobic core formed by Phe-11, Val-25, Phe-27, Val-65, Val-71, and Phe-80 (Figure 9) is closely similar to the core of the *E. coli* structure, both in terms of sequence and conformation. Indeed, only a few conservative changes in amino acid sequence can be noted here, namely, Leu-15/Val-16, Val-65/Thr-66, and Val-71/Ile-72. One of the aromatic residues present in the *E. coli* sequence, Tyr-70, which lies at the exterior of the hydrophobic core, is replaced in the human enzyme by a cysteine (Cys-69). Interestingly, the $S^{\gamma}H$ proton of Cys-69 is within hydrogen-bonding distance of the $O^{\epsilon 1}$ atom of Gln-78.

The three internal charged amino acids, Asp-26, Glu-56, and Asp-58, with surface accessibilities of less than 40%, are all involved in stabilizing interactions. The carboxylate group of Asn-26 is in close proximity to the $O^{\gamma}H$ proton of Ser-28, allowing for potential hydrogen-bond formation. This constrasts with the situation in the *E. coli* structure, where Asp-26 pairs with the side chain of Lys-57 via a water molecule. An equivalent interaction cannot occur in the human protein since Lys-57 in *E. coli* is replaced by Glu-56 in human thioredoxin. The carboxylate group of Glu-56 is involved in a potential electrostatic interaction with the side-chain $N^{\zeta}H_3^+$ group of Lys-36 or Lys-39, while the carboxylate group of Asp-58 lies within hydrogen-bonding distance of the backbone amides of residues Thr-30, Asp-60, and Asp-61.

A comparison of the conformation of the active site for human and *E. coli* thioredoxin reveals that they are very similar indeed. The eight completely conserved residues (Trp-31 to Ile-38), encompassing the two reactive site cysteines, can be superimposed with an rms difference of 0.68 Å for the C^{α} atoms. This may seem surprising given that the structure of human thioredoxin has been determined in the reduced state and contains free sulfydryls, whereas the *E. coli* structure is of the oxidized protein and contains a disulfide bridge between Cys-32 and Cys-35. The χ_1 torsion angles for Cys-32 and Cys-35 are $-179 \pm 19°$ and $-60 \pm 12°$, respectively, in reduced human thioredoxin and 166° and −62° in oxidized *E. coli* thioredoxin. In the reduced *E. coli* thioredoxin NMR structure the χ_1 angle of Cys-35 is in a different rotamer conformation with a value of $176 \pm 4°$. The side chain of Cys-35 is rotated away from the hydrophobic surface, resulting in a Cys-32(S^{γ}) to Cys-35(S^{γ}) distance of 6.8 ± 0.6 Å (Dyson et al., 1990). The equivalent distance for the reduced human structure is 3.1 ± 0.2 Å, somewhat larger than that for the disulfide-bonded oxidized X-ray structure of 2.05 Å (Katti et al., 1990) but significantly smaller than that found in reduced *E. coli* thioredoxin. In order to draw any firm conclusions about the significance of the different active site cysteine conformations and to be able to relate this structural information to the redox function of thioredoxin, it will be necessary to study either the human or the *E. coli* protein in different redox states under comparable experimental conditions.

A second interesting observation concerning the active site involves a potential stabilizing interaction in which the Cys-32 S^{γ} atom accepts a hydrogen bond from the backbone amide of Cys-35. [In the restrained minimized mean structure, $\overline{(SA)}r$, the distances between the S^{γ} atom of Cys-35 and the NH proton and backbone nitrogen atom of Cys-35 are ~2.6 Å and ~3.5 Å, respectively, and the N (Cys-35)–NH (Cys-35)–S^{γ} (Cys-32) angle is ~154°.] The chemical shift of the

Cys-35 NH exhibits a substantial pH dependence, with resonances at 8.34 ppm at pH 7.0 and 7.59 ppm at pH 5.5. This probably reflects the different ionization states of the Cys-32 sulfhydryl. Furthermore, this hydrogen-bonding interaction predicts a low pK for the Cys-32 sulfhydryl, which would be in agreement with the finding of a pK of 6.7 in *E. coli* thioredoxin (Kallis & Holmgren, 1980). This is consistent with the Cys-32 side chain being protonated at pH 5.5 but deprotonated and negatively charged at pH 7.0. As a result, further deshielding of the neighboring Cys-35 amide proton would occur at higher pH values with an increase in the strength of the hydrogen bond, producing a concomitant downfield shift of the Cys-35 NH resonance.

Concluding Remarks. The availability of the high-resolution NMR structure of human thioredoxin presented in this paper opens the door for a whole range of further studies. In particular, investigations of backbone and side-chain dynamics by heteronuclear relaxation measurements (Clore et al., 1990b,c), hydrogen-exchange kinetics (Englander & Kallenbach, 1983; Wagner, 1983), and reversible folding–denaturation studies (Udgaonkar & Baldwin, 1988; Roder et al., 1988; Baum et al., 1989) can be interpreted in detail in the light of the three-dimensional structure. In addition, the rational application of site-directed mutagenesis can now be used to systematically probe structure–function relationships (Leatherbarrow & Fersht, 1986) as well as factors involved in stabilizing and destabilizing the folded state (Matouschek et al., 1989). Such studies are currently underway in our laboratory.

ACKNOWLEDGMENTS

We thank Dr. F. M. Richards for useful discussions.

SUPPLEMENTARY MATERIAL AVAILABLE

The coordinates of the 33 final simulated annealing structures of reduced human thioredoxins, as well as those of the restrained minimized mean structure, have been deposited in the Brookhaven Protein Data Bank, along with the complete list of experimental restraints used in the calculations. (The PDB code numbers are 4TRX, 3TRX, and R3TXRMR, respectively.)

REFERENCES

Baum, J., Dobson, C. M., Evans, P. A., & Hanley, C. (1989) *Biochemistry 28*, 7–13.

Brünger, A. T. (1988) *XPLOR Manual*, Yale University, New Haven, CT.

Brünger, A. T., Clore, G. M., Gronenborn, A. M., & Karplus, M. (1987) *Protein Eng. 1*, 399–406.

Chothia, C. (1976) *J. Mol. Biol. 105*, 1–14.

Chothia, C., Levitt, M., & Richardson, D. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 4130–4134.

Clore, G. M., & Gronenborn, A. M. (1991) *Annu. Rev. Biophys. Biophys. Chem. 20*, 29–63.

Clore, G. M., Gronenborn, A., M., Brünger, A. T., & Karplus, M. (1985) *J. Mol. Biol. 186*, 435–455.

Clore, G. M., Brünger, A. T., Karplus, M., & Gronenborn, A. M. (1986a) *J. Mol. Biol. 191*, 523–551.

Clore, G. M., Nilges, M., Sukumaran, D. K., Brünger, A. T., Karplus, M., & Gronenborn, A. M. (1986b) *EMBO J. 5*, 2729–2735.

Clore, G. M., Gronenborn, A. M., Nilges, M., & Ryan, C. (1987) *Biochemistry 26*, 8012–8023.

Clore, G. M., Appella, E., Yamada, M., Matsushima, K., & Gronenborn, A. M. (1990a) *Biochemistry 29*, 1689–1696.

Clore, G. M., Szabo, A., Bax, A., Kay, L. E., Driscoll, P. C., & Gronenborn, A. M. (1990b) *J. Am. Chem. Soc. 112*, 4989–4991.

Clore, G. M., Driscoll, P. C., Wingfield, P. T., & Gronenborn, A. M. (1990c) *Biochemistry 29*, 7387–7401.

Driscoll, P. C., Clore, G. M., Berêss, L., & Gronenborn, A. M. (1989) *Biochemistry 28*, 2178–2187.

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

Englander, S. W., & Kallenbach, N. R. (1983) *Q. Rev. Biophys. 16*, 521–655.

Engstrom, N. E., Holmgren, A., Larsson, A., & Sonderhall, S. (1974) *J. Biol. Chem. 249*, 205–210.

Forman-Kay, J. D., Clore, G. M., Driscoll, P. C., Wingfield, P. T., Richards, F. M., & Gronenborn, A. M. (1989) *Biochemistry 28*, 7088–7097.

Forman-Kay, J. D., Gronenborn, A. M., Kay, L. E., Wingfield, P. T., & Clore, G. M. (1990) *Biochemistry 29*, 1566–1572.

Gronenborn, A. M., & Clore, G. M. (1990) *Anal. Chem. 62*, 2–15.

Gronenborn, A. M., Bax, A., Wingfield, P. T., & Clore, G. M. (1989a) *FEBS Lett. 243*, 93–98.

Gronenborn, A. M., Wingfield, P. T., & Clore, G. M. (1989b) *Biochemistry 28*, 5081–5089.

Havel, T. F. (1986) *DISGEO, Quantum Chemistry Program Exchange No. 507*, Indiana University, Bloomington, IN.

Havel, T. F., & Wüthrich, K. (1984) *Bull. Math. Biol. 46*, 673–698.

Havel, T. F., Kuntz, I. D., & Crippen, G. M. (1983) *Bull. Math. Biol. 45*, 665–720.

Hiraoki, T., Brown, S., Stevenson, K., & Vogel, H. (1988) *Biochemistry 27*, 5000–5008.

Holmgren, A. (1985) *Annu. Rev. Biochem. 54*, 237–271.

Holmgren, A. (1989) *J. Biol. Chem. 264*, 13963–13966.

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

Jeener, J., Meier, B. H., Bachmann, P., & Ernst, R. R. (1979) *J. Chem. Phys. 71*, 4546–4553.

Jones, T. A. (1978) *J. Appl. Crystallogr. 11*, 268–272.

Jones, T. A. (1990) *O Manual*, University of Uppsala, Uppsala, Sweden.

Jones, T. A., & Thirup, S. (1986) *EMBO J. 5*, 819–822.

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

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

Kay, L. E., & Bax, A. (1990) *J. Magn. Reson. 86*, 10–126.

Kelley, R., & Richards, F. M. (1987) *Biochemistry 26*, 6765–6774.

Kelley, R., Wilson, J., Bryant, C., & Stellwagen, E. (1986) *Biochemistry 25*, 728–732.

Kraulis, P. J., Clore, G. M., Nilges, M., Jones, T. A., Pettersson, G., Knowles, J., & Gronenborn, A. M. (1989) *Biochemistry 28*, 7241–7257.

Langsetmo, K., Fuchs, J., & Woodward, C. (1989) *Biochemistry 28*, 3211–3220.

Laurent, T. C., Moore, E. C., & Reichard, P. (1964) *J. Biol. Chem. 239*, 3426–3444.

Leatherbarrow, R. J., & Fersht, A. R. (1986) *Protein Eng. 1*, 7–16.

Lin, T.-Y., Kim, P. S. (1989) *Biochemistry 28*, 5282–5287.

Macura, C., Huang, Y., Suter, D., & Ernst, R. R. (1981) *J. Magn. Reson. 43*, 259–281.

Marion, D., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967-974.
Marion, D., & Bax, A. (1988a) *J. Magn. Reson. 79*, 352-356.
Marion, D., & Bax, A. (1988b) *J. Magn. Reson. 80*, 528-533.
Mark, D. F., & Richardson, C. C. (1976) *Proc. Natl. Acad. Sci. U.S.A. 73*, 780-784.
Matouschek, A., Kellis, J. T., Jr., Serrano, L., & Fersht, A. (1989) *Nature 340*, 122-126.
McLachlan, A., & Stewart, M. (1975) *J. Mol. Biol. 98*, 293.
Mueller, L. (1987) *J. Magn. Reson. 72*, 191-196.
Nilges, M., Clore, G. M., & Gronenborn, A. M. (1988a) *FEBS Lett. 229*, 317-324.
Nilges, M., Gronenborn, A. M., Brünger, A. T., & Clore, G. M. (1988b) *Protein Eng. 2*, 27-38.
Nilges, M., Clore, G. M., & Gronenborn, A. M. (1988c) *FEBS Lett. 239*, 129-136.
Nilges, M., Clore, G. M., & Gronenborn, A. M. (1990) *Biopolymers 29*, 813-822.
Omichinski, J., Clore, G. M., Apella, E., Sakaguchi, K., & Gronenborn, A. M. (1990) *Biochemistry 29*, 9324-9334.
Plateau, P., & Guéron, M. (1982) *J. Am. Chem. Soc. 104*, 7310-7311.
Redfield, A. G., & Kunz, S. D. (1975) *J. Magn. Reson. 19*, 250-254.
Richards, F. M., & Kundrot, C. E. (1988) *Proteins 3*, 71-84.
Richardson, J. (1981) *Adv. Protein Chem. 34*, 167-339.
Rimski, L., Wakasuki, H., Ferrara, P., Robin, P., Capdevielle, J., Tursz, T., Fradelizi, D., & Bertoglio, J. (1986) *J. Immunol. 136*, 3304-3310.
Roder, H., Elöve, G. A., & Englander, W. (1988) *Nature 335*, 700-704.
Söderberg, B.-O., Sjöberg, B.-M., Sonnerstam, V., & Brändén, C.-I. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 5827-5830.
Tagaya, Y., Taniguchi, Y., Naramura, M., Okada, M., Suzuki, N., Kanamori, H., Nikaido, T., Honjo, T., & Yodoi, J. (1987) *Immunol. Lett. 15*, 221-228.
Tagaya, Y., Maeda, Y., Mitsui, A., Kondo, N., Matsui, H., Hamuro, J., Brown, N., Arai, K., Yokota, T., Wakasugi, H., & Yodoi, J. (1989) *EMBO J. 8*, 757-764.
Udgaonkar, J. B., & Baldwin, R. L. (1988) *Nature 335*, 694-699.
Wagner, G. (1983) *Q. Rev. Biophys. 16*, 1-57.
Wagner, G., Braun, W., Havel, T. F., Schaumann, T., Go, N., & Wüthrich, K. (1987) *J. Mol. Biol. 196*, 611-639.
Wakasuki, H., Rimski, L., Mahe, Y., Kamel, A. M., Fradelizi, D., Tursz, T., & Bertoglio, J. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 804-808.
Wilson, J., Kelley, R., Shalongo, W., Lowery, D., & Stellwagen, E. (1986) *Biochemistry 25*, 7560-7566.
Wollman, E. E., d'Auriol, L., Rimsky, L., Shaw, A., Jacquot, J.-P., Wingfield, P. T., Graber, P., Dessarps, F., Robin, P., Galibert, F., Bertoglio, J., & Fradelizi, D. (1988) *J. Biol. Chem. 263*, 15506-15512.
Wüthrich, K., Billeter, M., & Braun, W. (1983) *J. Mol. Biol. 169*, 949-961.
Zuiderweg, E. R. P., Boelens, R., & Kaptein, R. (1985) *Biopolymers 24*, 601-611.

Folding of Staphylococcal Nuclease A Studied by Equilibrium and Kinetic Circular Dichroism Spectra†

Tatsuro Sugawara,‡ Kunihiro Kuwajima,* and Shintaro Sugai§
Department of Polymer Science, Faculty of Science, Hokkaido University, Kita-ku, Sapporo, Hokkaido 060, Japan
Received July 2, 1990; Revised Manuscript Received December 5, 1990

ABSTRACT: The urea-induced unfolding of staphylococcal nuclease A has been studied by circular dichroism both at equilibrium and by the kinetics of unfolding and refolding (pH 7.0 and 4.5 °C), as a function of Ca^{2+} and thymidine 3′,5′-diphosphate (pdTp) concentration. The results are as follows. (1) The unfolding transition is shifted to higher concentrations of urea by Ca^{2+} and pdTp, and the presence of both ligands further stabilizes the protein. (2) In the first stage of kinetic refolding, the peptide ellipticity changes rapidly within the dead time of stopped-flow measurement (15 ms), indicating accumulation of a transient intermediate. This intermediate is remarkably less stable than those of other globular proteins previously studied. (3) Dependence of the folding and unfolding rate constants on urea concentration indicates that the critical activated state of folding ("transition state") has considerable structural organization. The transition state does not, however, have the capacity to bind Ca^{2+} and pdTp, as indicated by the effects of these ligands on the unfolding rate constant. (4) There are at least four different phases in the refolding kinetics in native conditions below 1 M urea. In the absence of pdTp, there are two phases in unfolding, while in the presence of pdTp the unfolding kinetics show a single phase. Some characteristics of the transient intermediate and of the transition state for folding are discussed.

Elucidation of the mechanism of protein folding remains one of the major challenges in biochemistry. Two important approaches to this problem are (1) characterization of transient intermediates formed early in refolding and (2) examination of the kinetics of folding and unfolding to investigate the nature of the transition state for folding (Kim & Baldwin, 1982, 1990). Kinetic circular dichroism (CD)[1] measurements in the

† This work was supported in part by Grants-in-Aid for Scientific Research from the Ministry of Education, Science and Culture of Japan (61420050 and 01580258) and also from the Kurata Foundation.
* Author to whom correspondence should be addressed.
‡ Present address: Institute of Biological Science, Mitsui Pharmaceuticals Inc., Mobara, Chiba 2974, Japan.
§ Present address: Institute of Life Science, Soka University, Tangicho, Hachioji, Tokyo 192, Japan.

[1] Abbreviations: CD, circular dichroism; EGTA, [ethylenebis(oxyethylenenitrilo)]tetraacetic acid; NMR, nuclear magnetic resonance; pdTp, thymidine 3′,5′-diphosphate; SNase, staphylococcal nuclease; UV, ultraviolet.

0006-2960/91/0430-2698$02.50/0