Three-Dimensional Solution Structure of the Reduced Form of *Escherichia coli* Thioredoxin Determined by Nuclear Magnetic Resonance Spectroscopy†

H. Jane Dyson,*,‡ Garry P. Gippert,‡ David A. Case,‡ Arne Holmgren,§ and Peter E. Wright*,‡

Department of Molecular Biology, Research Institute of Scripps Clinic, La Jolla, California 92037, and Department of Physiological Chemistry, Karolinska Institutet, S-10401 Stockholm, Sweden

Received November 14, 1989; Revised Manuscript Received January 10, 1990

ABSTRACT: The three-dimensional solution structure of reduced (dithiol) thioredoxin from *Escherichia coli* has been determined with distance and dihedral angle constraints obtained from ^{1}H NMR spectroscopy. Reduced thioredoxin has a well-defined global fold consisting of a central five-strand β-sheet and three long helices. The β-strands are packed in the sheet in the order $\beta_1\beta_3\beta_2\beta_4\beta_5$, with β_1, β_3, and β_2 parallel and β_2, β_4, and β_5 arranged in an antiparallel fashion. Two of the helices connect strands of the β-sheet: α_1 between β_1 and β_2 and α_2 between β_2 and β_3. Strands β_4 and β_5 are connected by a short loop that contains a β-bulge. Strands β_3 and β_4 are connected by a long loop that contains a series of turn-like or 3_{10} helical structures. The active site Cys-Gly-Pro-Cys sequence forms a protruding loop between strand β_2 and helix α_2. The structure is very similar overall to that of oxidized (disulfide) thioredoxin obtained from X-ray crystal structure analysis but differs in the local conformation of the active site loop. The distance between the sulfurs of Cys 32 and Cys 35 increases from 2.05 Å in the disulfide bridge to 6.8 ± 0.6 Å in the dithiol of reduced thioredoxin, as a result of a rotation of the side chain of Cys 35 and a significant change in the position of Pro 34. This conformational change has important implications for the mechanism of thioredoxin as a protein disulfide oxidoreductase.

Thioredoxin is a ubiquitous, multifunctional protein, with the conserved active site residues Cys-Gly-Pro-Cys forming a disulfide bridge in the oxidized form (thioredoxin-S_2) (Holmgren, 1985; Gleason & Holmgren, 1988). Thioredoxin-S_2 is reduced by NADPH and the flavoprotein thioredoxin reductase to form thioredoxin-$(SH)_2$, which contains a dithiol (Holmgren & Reichard, 1967). Thioredoxin-$(SH)_2$ is a powerful protein disulfide reductase, and thioredoxin catalyzes dithiol–disulfide exchange reactions (Holmgren, 1985). Functional properties demonstrate significant differences between the two oxidation states of thioredoxin. One striking example is that only thioredoxin-$(SH)_2$ binds to gene 5 protein of phage T7 to generate a highly active DNA polymerase (Mark & Richardson, 1976; Huber et al., 1986); thioredoxin-S_2 is completely inactive (Adler & Modrich, 1983). The differences observed between the two forms of the protein have not yet been characterized structurally. Early work (Stryer et al., 1967; Holmgren, 1972; Holmgren & Roberts, 1976) indicated that the conformational change upon reduction of thioredoxin-S_2 was small and probably localized to the active site. Recent NMR experiments show only limited changes in the chemical shifts of protons in the two oxidation states (Dyson et al., 1988, 1989).

Thioredoxin from *Escherichia coli* contains 108 residues (M_r 11 700), and the oxidized form has been crystallized as a Cu^{2+} complex (Holmgren & Söderberg, 1970). The three-dimensional structure has been determined by X-ray crystallography to 2.8-Å resolution (Holmgren et al., 1975) and has been recently refined to 1.7 Å (Katti et al., 1990). The most striking feature of thioredoxin-S_2 is that the active site disulfide is located on a protrusion of the molecule at the end of a β-strand (β_2) and followed by a long α-helix (Holmgren et al., 1975). Since reduced thioredoxin has so far proved impossible to crystallize, its detailed three-dimensional structure is still unknown. Complete resonance assignments have recently been made for the ^{1}H NMR[1] spectrum of reduced and oxidized *E. coli* thioredoxin (Dyson et al., 1989). Assignments for the oxidized protein have also been reported by LeMaster and Richards (1988). An extensive series of distance constraints have been derived for reduced thioredoxin and have been utilized in distance geometry and restrained MD calculations to derive three-dimensional structures. In the present paper we report the results of these calculations and describe the solution structure of reduced thioredoxin together with the conformational changes accompanying reduction of the active site disulfide bridge of thioredoxin-S_2.

Experimental Procedures

NMR Measurements. Reduced thioredoxin was prepared and used for NMR spectroscopy as described previously (Dyson et al., 1989). Samples for NMR spectroscopy were ~4 mM in 0.1 M potassium phosphate buffer (pH 5.7) in either ^2H_2O or 90% ^1H_2O/10% ^2H_2O. Spectra were recorded at 25 and 35 °C on a Bruker AM500 spectrometer. All spectra were acquired in the phase-sensitive mode, with time-proportional phase incrementation (Drobny et al., 1979, Marion & Wüthrich, 1983). 2QF-COSY (Rance et al., 1983) and NOESY (Jeener et al., 1979; Kumar et al., 1981) spectra with mixing times of 30, 50, 80, 120, and 150 ms were recorded by standard methods. The quality of the NOESY spectra has been illustrated previously, and complete resonance assignments have been reported (Dyson et al., 1989).

† This research was supported by Grant GM36643 from the National Institutes of Health and Grant 13x-3529 from the Swedish Medical Research Council and by the Karolinska Institutet.

‡ Research Institute of Scripps Clinic.

§ Karolinska Institutet.

[1] Abbreviations: NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; NOESY, two-dimensional nuclear Overhauser effect spectroscopy; 2QF-COSY, double quantum filtered two-dimensional correlated spectroscopy; MD, molecular dynamics; rmsd, root mean square deviation.

0006-2960/90/0429-4129$02.50/0

Assignment of NOE Cross Peaks. Assignment of cross peaks in the NOESY spectrum to specific proton pairs was carried out in several stages. First, cross peaks were assigned on the basis of the ^{1}H chemical shifts at 35 °C (Dyson et al., 1989). Where ambiguities occur as a result of resonance overlap, some assignments were made on the basis of multiplet structure and cross-peak shape and width. NOESY spectra recorded at 25 °C were helpful in resolving ambiguities involving amide protons. By use of these procedures, a large number of NOESY cross peaks could be assigned unambiguously, giving distance constraints for input to the initial structure calculations.

Additional NOESY cross peaks were assigned on the basis of distances calculated from structures generated by distance geometry and restrained MD. All possible distances between the various proton pairs which might contribute to an ambiguous cross peak were calculated from the list of chemical shifts by a program written by G. Gippert. Stringent criteria were used to assign ambiguous NOESY cross peaks on the basis of distance calculations: the only assignments used in further structure calculations were those for which the average calculated distance was no more than 3 Å greater than the upper-bound distance estimated from the experimental NOE intensity and for which the distance between all other possible proton pairs was more than 8 Å greater than the estimated upper bound. This procedure requires that proton assignments be complete, to ensure that all possibilities have been considered.

Evaluation of Distance Constraints. NOE cross-peak volumes were determined by use of software of Dr. Dennis Hare. The classification of an NOE as strong, medium, weak, or very weak was made from visual inspection of NOE buildup curves, as well as from a consideration of the absolute magnitudes of the volume integrals. Constraints were quoted as a range of permitted distance between an upper and a lower bound. Upper bounds were determined by empirical calibration versus known distances in regular secondary structure (Billeter et al., 1982; Williamson et al., 1985), with corrections for uniform motional averaging (Braun et al., 1981) and for methyl groups (Wagner et al., 1987). The resulting upper bounds were 2.5, 3.0, 3.5, and 5.0 Å for backbone–backbone NOEs and 3.0, 4.0, and 5.0 Å for NOEs involving side-chain protons. The upper bounds for side-chain protons are highly conservative and allow for the possibility of side-chain motion. In all cases, the lower bounds were taken to be the sum of the van der Waals radii.

Other Structural Constraints. Backbone ϕ dihedral angle constraints were obtained from $^3J_{HN\alpha}$ coupling constants calculated from the peak-to-peak separation in cross peaks in 2QF-COSY spectra zero filled to 16K real points for greater digital resolution in ω_2. For $^3J_{HN\alpha} < 5.5$ Hz, the dihedral angle constraint was $-90° < \phi < -40°$ (distance geometry) and $-90° < \phi < +40°$ (MD), and for $^3J_{HN\alpha} > 10$Hz, the constraint was $-160° < \phi < -80°$. The broadness of many NH resonances caused difficulties in measurement of small coupling constants: the apparent peak-to-peak separation in the COSY spectrum approaches a lower-limiting value related to the line width, and the cross peak is severely attenuated (Neuhaus et al., 1985). In many cases, upper limits on coupling constants could be estimated by computer simulation based on the width of NOESY cross peaks at $\omega_2 = \omega_{NH}$ and the observed attenuation of the corresponding NH–C$^\alpha$H COSY cross peak. (Full details of this method will be published elsewhere.) $^3J_{\alpha\beta}$ coupling constants for the active site Cys 32 and Cys 35 residues were obtained by computer simulation of the C$^\alpha$H–C$^\beta$H cross peaks

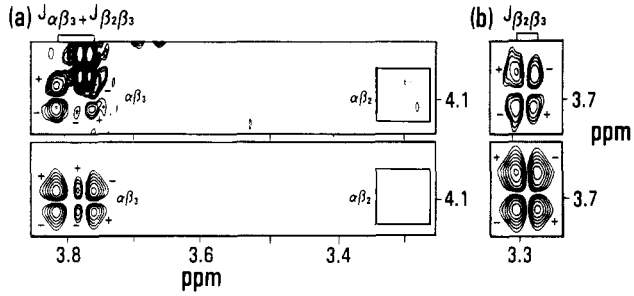

FIGURE 1: Regions of a 2QF-COSY spectrum of reduced thioredoxin showing the (a) C$^\alpha$H–C$^{\beta_3}$H and (b) C$^{\beta_2}$H–C$^{\beta_3}$H cross peaks of Cys 35 together with the cross peaks simulated for this spin system by the computer program SPINNER written by Dr. M. Rance. Note that in the experimental spectrum (a) there is a strong cross peak from Thr 14 superimposed on the upper part of the C$^\alpha$H–C$^{\beta_3}$H cross peak. Boxes in (a) indicate the position of the expected C$^\alpha$H–C$^{\beta_2}$H cross peak, which is not observed due to the small size of the $^3J_{\alpha\beta_2}$ coupling constant relative to the resonance line width. This is faithfully reproduced in the simulated spectrum. The digital resolution in both experimental and calculated spectra is 0.35 Hz/point in ω_2 and 11 Hz/point in ω_1. The spectrum was simulated with a line width of 12 Hz and the following parameters: C$^\alpha$H, 4.09 ppm; C$^{\beta_2}$H, 3.29 ppm; C$^{\beta_3}$H, 3.79 ppm; $^3J_{\alpha\beta_3}$, 12.5 Hz; $^3J_{\beta\beta}$, −15 Hz; $^3J_{\alpha\beta_2}$, 3.5 Hz. The $^3J_{\alpha\beta_3}$ and $^3J_{\beta\beta}$ coupling constraints were extracted directly from the experimental spectrum, while that for $^3J_{\alpha\beta_2}$ was estimated. The pattern of coupling constants indicates that the Cys 35 side-chain conformation is close to the g^2t^3 ($\chi_1 = 180°$) or t^2g^3 ($\chi_1 = -60°$) rotameric states. The rotamers were discriminated, and stereospecific assignments for the C$^\beta$H protons were made on the basis of the relative intensity of the NOE cross peaks between the NH and the C$^\beta$H protons of Cys 35. The stereospecific assignments are consistent with observed NOEs between the C$^\beta$H protons and other protons in the active site region.

in the 2QF-COSY spectrum (Figure 1). For both residues there is one large (~12 Hz) and one small (~4 Hz) $^3J_{\alpha\beta}$ coupling constant. The pattern of intraresidue NOEs (C$^\beta$H–NH, C$^\beta$H–C$^\alpha$H) observed allows stereospecific assignments of the C$^\beta$H proton resonances to be made and χ_1 dihedral angle constraints to be imposed (Billeter & Wüthrich, 1984; Arseniev et al., 1988). In the final cycle of restrained MD refinement, χ_1 for Cys 32 and Cys 35 was constrained to 180 ± 60°. For distance geometry calculations (but not for MD refinement) 17 hydrogen-bond constraints (NH···O, 1.8–2.0 Å; N···O, 2.7–3.0 Å) were imposed on the basis of slowly exchanging amide protons (Dyson et al., 1989).

Protocol for Structure Calculations. Structures were calculated a priori with the distance geometry program DISGEO (Havel & Wüthrich, 1984). The resulting structures were refined by a combination of energy minimization and restrained MD. The refinement procedure used the AMBER all-atom force field (Weiner et al., 1986), modified as described elsewhere (Lee et al., 1989), to compute the intrinsic strain energy with a half-parabola penalty function added for proton–proton distances that violate the NMR upper-bound constraints. Dihedral angles falling outside the prescribed limits were penalized by a function of the form $K[1 - \cos(\phi - \phi_0)]$, where ϕ_0 is the end point of the "allowed" range and $K = 32$ kcal/mol.

The distance geometry structures were energy minimized in this combined potential and then heated to 1200 K during 1 ps by a temperature-regulated MD algorithm (Berendsen et al., 1984) with a temperature relaxation time τ of 0.6 ps. This was followed by 4 ps of equilibration at 1200 K with τ = 0.2 ps and a 6-ps cooling period, with a target temperature of 0 K and τ = 2 ps. A final cycle of 1000 energy minimization steps produced the refined structures. During the initial minimization and the first 3 ps of the dynamics, the force constant for distance violations was slowly increased from 1 to 32 kcal/(mol·Å^2); this final value was kept constant for the

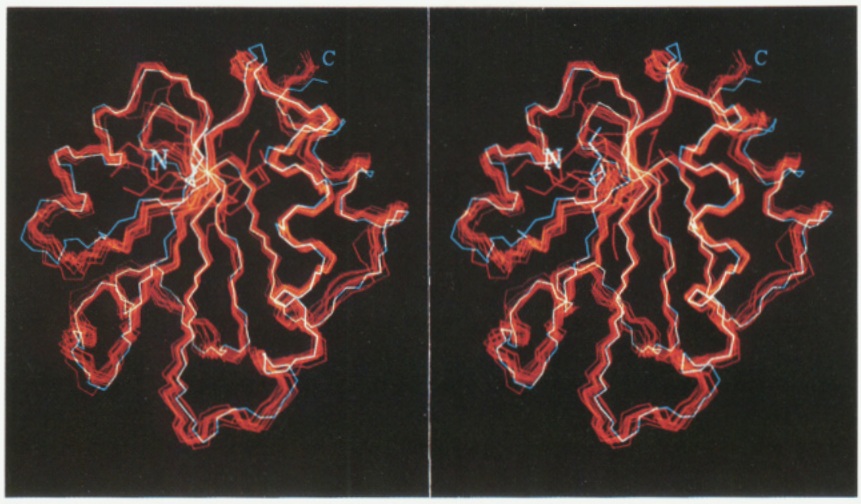

FIGURE 2: Stereoview of superposition of 12 refined solution structures of reduced *E. coli* thioredoxin (red), together with the refined X-ray crystal structure of oxidized thioredoxin (Katti et al., 1990) (blue). Backbone C^{α}, C, and N atoms are shown. The structures are superimposed for minimum rms deviation between backbone heavy atoms from residues 3 to 108. The active site is located on the projecting loop at the bottom of the structure.

cooling and minimization steps. Each refinement required about 70 min of cpu time on a Cray XMP/SE.

With this protocol, a number of preliminary calculations were carried out. Early DISGEO structures were expanded, with poor packing of the long central helix against the β-sheet, due in part to a lack of NOE constraints between helix and sheet and in part to loss of constraint integrity due to pseudoatom corrections (Wüthrich et al., 1983). Local geometry in the two long helices consisted of mixed left and right helical turns in an irregular pattern. The resulting structures violated many of the NOEs [d_{NN}, $d_{\alpha N}(i,i+3)$, and $d_{\alpha\beta}(i,i+3)$] diagnostic of helix; only turns with right-handed helical geometry satisfied all of these NMR constraints. A mixed pseudoatom/all-atom representation was used in DISGEO to tighten constraints to Gly $C^{\alpha}H$, Ile $C^{\gamma}H$, and all $C^{\beta}H$ protons. The resulting structures were better packed, although helices were still not completely regular and there were some incorrect cis peptide bonds. Restrained MD, with tight dihedral angle constraints on the helical residues, was then used to regularize the helices and resulted in structures suitable for distance calculations to assign more long-range NOE cross peaks. The constraint list was further extended to include NOEs assigned by careful examination of NOESY spectra of reduced thioredoxin acquired at different temperatures in 90% ^1H_2O/10% ^2H_2O solutions and in ^2H_2O.

With these additional constraints in hand, the calculations were begun anew to obtain the final structures reported here. A new family of 28 distance geometry structures was generated with the full constraint list but without the intraresidue NH–$C^{\alpha}H$ NOE constraints, since we found empirically that helices embedded more reliably without these constraints. The helical constraints that had been used in the preliminary calculations were no longer necessary and were omitted from all subsequent calculations. The 17 structures with the lowest residual constraint violations were subjected to five cycles of restrained MD refinement, during which further checks of constraints against the spectra were made. Several more NOE cross peaks were assigned on the basis of these structures. The final MD refinement, performed with a 10-ps cooling period with τ = 4 ps, included χ_1 constraints and stereospecific NOEs to the $C^{\beta}H$ protons of Cys 32 and Cys 35. This final refinement used a total of 853 interresidue and 391 intraresidue distance constraints and 43 dihedral angle constraints (Table I). The coordinates of the final structures, together with a complete list of constraints and additional descriptive statistics about the structures, have been deposited with the Protein Data Bank.

Table I: NMR Constraints Used in Final Structure Calculations and Refinements

Interresidue NOEs					
		type[a]			
bound (Å)	no.	SEQ	MBB	LBB	LNG
2.5	38	38	0	0	0
3.0	93	56	1	5	31
3.5	97	44	27	16	10
4.0	123	29	2	0	92
5.0	502	83	59	38	322
total	853	250	89	59	455

Intraresidue NOEs (Intraresidue NH–$C^{\alpha}H$ NOEs Excluded)	
bound (Å)	no.
3.0	43
4.0	168
5.0	180
total	391

Dihedral Angle Constraints	
$^3J_{HN\alpha} < 5.5$ Hz	17
$^3J_{HN\alpha} > 10$ Hz	24
χ_1 constraints	2

[a] Terminology of Wüthrich (1986): SEQ, sequential backbone–backbone or backbone–β $(i,i+1)$; MBB, medium-range backbone–backbone or backbone–β $(i,i+2$ to $i,i+4)$; LBB, long-range backbone–backbone or backbone–β $(|j - i| \geq 5)$; LNG, all other interresidue NOEs.

RESULTS

The best 12 of the final refined structures are in excellent agreement with the NMR data, and none contain distance constraint violations greater than 0.32 Å (Table II). The structures display only small deviations from idealized covalent geometry and have good nonbonded contacts. Only glycine residues have positive average values of ϕ, with the exception of Asn 83, which has a positive ϕ in all of the final NMR structures (average $\phi = +46.8 \pm 2.4°$ in the 12 best structures). The backbone conformations of reduced thioredoxin are shown superimposed in Figure 2. Overall, the conformation of the polypeptide backbone is well-defined except at the relatively unconstrained N-terminus (residues 1–3). The average backbone rms deviation from the mean is 0.56 Å for

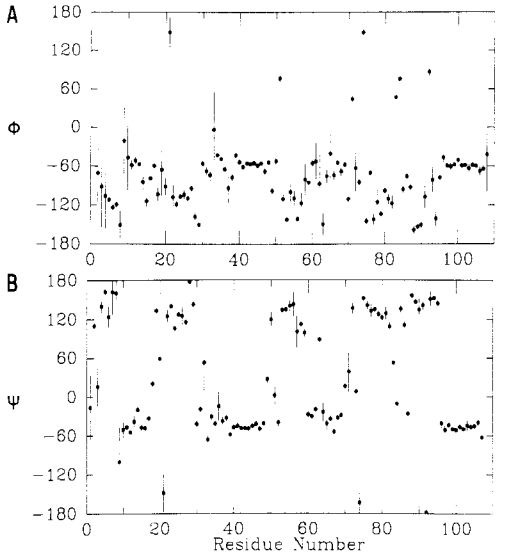

FIGURE 3: (A) ϕ and (B) ψ dihedral angles plotted for each amino acid residue. The average value for the 12 final MD-refined structures is represented by a dot, and the vertical lines represent the standard deviations of these values.

residues 3–108 (0.77 Å for all residues) and 1.38 Å for side-chain heavy atoms (Table III). The conformations of most interior side chains are very well-defined (average rms deviation from the mean of 0.96 Å). Many surface side chains are relatively unconstrained by NOEs, and this is reflected in their higher average rms deviations (1.72 Å).

Description of the Structure. The backbone conformations shown in Figure 2 contain all of the elements of secondary structure and supersecondary structure deduced previously for reduced thioredoxin on the basis of observed patterns of NOE connectivities (Dyson et al., 1989). The protein consists of a central five-strand β-sheet and three long helices, arranged in the order $\beta\alpha\beta\alpha\beta\beta\beta\alpha$. The various elements of secondary structure were identified from an analysis of average backbone dihedral angles (Figure 3) and hydrogen-bonding patterns (Table IV). For this purpose, a hydrogen bond is deemed to be formed if the O–H distance is less than 2.5 Å and the O–H–N angle is greater than 135°. We emphasize that no hydrogen-bond constraints were included in restrained MD refinement. The hydrogen-bond constraints that were used in the distance geometry calculations are indicated in Table IV; note that the overwhelming majority of the hydrogen bonds found in the final structures were not specified as constraints at any stage of the calculations. Removing all hydrogen-bond constraints at the MD stage allows exploration of alternate conformations. This flexibility may be important at the ends of sheets and helices; indeed, as indicated in Table IV, one of the original hydrogen-bond constraints (connecting strands β_2 and β_3) is missing from 9 of the 12 MD-refined structures.

Table II: Summary of Residual Constraint Violations

range (Å)	av no. of distance constraint violations: MD[a]	DG[b]	X-ray[c]
0.1–0.2	13	50	21
0.2–0.3	1.4	30	19
0.3–0.4	0.1[d]	17	17
0.4–0.5		13	15
>0.5		18	63
max violation (Å)	0.32	1.8	5.7
range (deg)	**av no. of dihedral angle constraint violations: MD[a]**	**DG[b]**	**X-ray[c]**
0–5	2.5	2.4	0
5–10	0.3	2.0	0
>10	0	1.0	2[e]

Restrained Molecular Dynamics Structure	
av AMBER energy (kcal)	−1334.5
AMBER energy range (kcal)	−1301.0 to −1366.6
av total energy (kcal)	−1315.3
total energy range (kcal)	−1280.7 to −1349.7
rms deviations of bond lengths from ideality (Å)	0.005
rms deviations of bond angles from ideality (deg)	1.8

[a] 12 best-restrained molecular dynamics structures. [b] 21 distance geometry structures. [c] Violations of NMR constraints in X-ray structure of oxidized thioredoxin (Katti et al., 1990). [d] Only one violation greater than 0.3 Å. [e] Includes violation of χ_1 dihedral angle for Cys 35. The observed coupling constants for Cys 35 in reduced thioredoxin are $^3J_{\alpha\beta_2} \sim 3.5$ Hz and $^3J_{\alpha\beta_3} = 12.5$ Hz; those expected from the X-ray structures of oxidized thioredoxin are $^3J_{\alpha\beta_2} = 12.5$ Hz and $^3J_{\alpha\beta_3} = 3.5$ Hz.

Table III: Average rms Deviation from Average Structure for Distance Geometry and Restrained Molecular Dynamics Structures[a]

residues (atoms)	DG[b]	MD[c]
1–108 (backbone)	1.164	0.768
3–108 (backbone)	1.138	0.564
3–108 (side chain)	1.990	1.380
3–108 (all heavy)	1.606	1.045
helical residues (backbone)[d]	1.026	0.433
sheet residues (backbone)[e]	0.926	0.352
internal residues (side-chain heavy atoms)[f]	1.709	0.962
external residues (side-chain heavy atoms)[f]	2.157	1.715

[a] Only heavy atoms are considered. [b] 21 distance geometry structures. [c] 12 best-restrained molecular dynamics structures. [d] Helical residues: 11–18, 34–49, and 95–108. [e] β-Sheet residues: 4–7, 22–29, 53–59, and 76–92. [f] Internal residues have a side-chain surface area of <25 Å^2.

The central twisted β-sheet forms the core of the molecule (Figure 4). Strand β_1, as defined by ϕ and ψ angles, runs from Ile 4 to Thr 8, but only residues 4–7 appear to form part of the β-sheet according to hydrogen-bonding patterns (Table

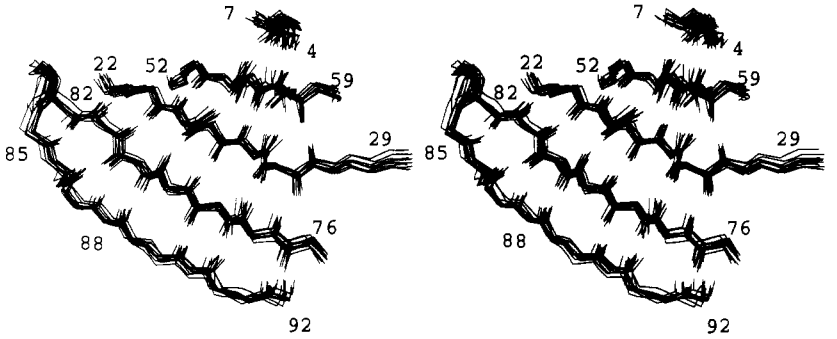

FIGURE 4: Superposition of 12 refined solution structures of reduced thioredoxin showing the central β-sheet. Carbonyl oxygens and amide protons involved in hydrogen bonds are shown as well as backbone C^α, C, and N atoms.

Table IV: Backbone–Backbone Hydrogen Bonds Formed in at Least 7 of the 12 Best-Restrained Molecular Dynamics Structures of Reduced Thioredoxin

donor	acceptor	frequency	donor	acceptor	frequency
5 Ile N	55 Val O	12	56 Ala N	24 Leu O	11
12 Phe N	8 Thr O	11	57 Lys N	5 Ile O	12
13 Asp N	9 Asp O	10	58 Leu N	26 Asp O	7[a]
14 Thr N	10 Asp O	10	60 Ile N	28 Trp O	12[a]
16 Val N	12 Phe O	11	62 Gln N	59 Asn O	11
17 Leu N	12 Phe O	11	63 Asn N	60 Ile O	7
19 Ala N	16 Val O	11	67 Ala N	64 Pro O	8
21 Gly N	19 Ala O	11	70 Tyr N	66 Thr O	7
23 Ile N	81 Phe O	12[a]	71 Gly N	68 Pro O	12
24 Leu N	54 Thr O	12[a]	72 Ile N	67 Ala O	11
25 Val N	79 Leu O	12[a]	77 Thr N	27 Phe O	9
26 Asp N	56 Ala O	12[a]	78 Leu N	90 Lys O	12
27 Phe N	77 Thr O	12[a]	79 Leu N	25 Val O	12[a]
28 Trp N	58 Leu O	12[a]	80 Leu N	88 Ala O	12[a]
32 Cys N	29 Ala O	11	81 Phe N	23 Ile O	12[a]
33 Gly N	31 Trp O	7	82 Lys N	85 Glu O	12
35 Cys N	32 Cys O	9	83 Asn N	21 Gly O	10
36 Lys N	33 Gly O	11	85 Glu N	82 Lys O	12
37 Met N	34 Pro O	11	87 Ala N	80 Leu O	12
38 Ile N	34 Pro O	9	88 Ala N	80 Leu O	8
39 Ala N	35 Cys O	12	90 Lys N	78 Leu O	12
42 Leu N	38 Ile O	12	92 Gly N	76 Pro O	12
43 Asp N	39 Ala O	12	99 Leu N	95 Ser O	11
44 Glu N	40 Pro O	12	100 Lys N	96 Lys O	12
45 Ile N	41 Ile O	12[a]	101 Glu N	97 Gly O	10[a]
46 Ala N	42 Leu O	12[a]	102 Phe N	98 Gln O	12
47 Asp N	43 Asp O	12	103 Leu N	99 Leu O	12[a]
48 Glu N	44 Glu O	12	104 Asp N	100 Lys O	12[a]
49 Tyr N	45 Ile O	12	105 Ala N	101 Glu O	12
50 Gln N	47 Asp O	12	106 Asn N	102 Phe O	12
52 Lys N	49 Tyr O	12	107 Leu N	103 Leu O	11
53 Leu N	49 Tyr O	12	108 Ala N	104 Asp O	11

[a] Indicates hydrogen-bond constraints used in distance geometry calculations, based on observation of slow NH proton exchange. No hydrogen-bond constraints were used in the restrained MD calculations. An additional hydrogen-bond constraint involving Thr 54 HN and Ala 22 CO was imposed in the distance geometry calculations, but this hydrogen bond is retained in only three of the restrained dynamics structures.

IV). Strand β_2 comprises residues Ala 22 to Ala 29, all of which form part of the β-sheet. Strand β_3 consists of residues Leu 53 to Asn 59. While the ϕ and ψ dihedral angles of Ile 60 ($-56 \pm 4°$ and $-28 \pm 6°$, respectively) show that it is not part of β_3, its amide proton is involved in intrasheet hydrogen bonding to the Trp 28 CO in all structures (Table IV). The β-sheet is completed by strands β_4 (Pro 76 to Phe 81) and β_5 (Ala 88 to Val 91), which form an antiparallel β-bulge hairpin structure. The strands are arranged in the order $\beta_1\beta_3\beta_2\beta_4\beta_5$: strands β_1, β_3, and β_2 are parallel; strands β_2, β_4, and β_5 are antiparallel.

Reduced thioredoxin contains three well-defined helices which are amphipathic and pack against the central β-sheet. The helix between residues 11–18 is irregular: the ϕ and ψ angles for residues 12, 13, and 17 have low rms deviations and are approximately helical, but ϕ for residues 14–16 differ significantly from standard helical values. CO_i–NH_{i+4} hydrogen bonds are found for Asp 10–Thr 14 (10 structures) and Phe 12–Val 16 (11 structures) (Table IV). The Phe 12 CO is also hydrogen bonded to Leu 17 NH. The helix is terminated by one loop of 3_{10} helix, according to dihedral angles and a CO_i–NH_{i+3} hydrogen bond between Val 16 and Ala 19 in 11 structures. Residues Pro 34 to Tyr 49 form a well-defined helix with a distortion at Pro 40. The mean ϕ and ψ angles ($-59.3°$ and $-42.5°$, respectively) between Pro 34 and Glu 48 are characteristic of α-helix, and with the exception of Lys 36–Pro 40 and Met 37–Ile 41, CO_i–NH_{i+4} hydrogen bonds are formed from residues 35–49 in all of the NMR structures. The helix is terminated by a turn of 3_{10} helix (47–50 hydrogen bond in all structures). Another well-defined α-helix extends from Ser 95 to Ala 108. The ϕ and ψ dihedral angles are tightly constrained and are close to standard helical values, and an α-helical pattern of CO_i–NH_{i+4} hydrogen bonds occurs throughout the helix in all structures.

In addition, the structures appear helical in the region from residues 64–70. The predominant hydrogen-bonding pattern is suggestive of 3_{10} helix, although this is not clear-cut. Residues 68–71 appear to form a type I turn-like structure, with a Pro 68–Gly 71 hydrogen bond in all cases. Ala 67 CO is hydrogen-bonded to Ile 72 NH in 11 of the structures.

Residues 29–32 in the active site form a well-defined type I β-turn with a CO_i-NH_{i+3} hydrogen bond in 11 of the structures. A type II β-turn is formed in all structures by residues 49–52, connecting helix 2 and strand β_3. Tyr 49 CO has a bifurcated hydrogen bond to the NH's of Lys 52 and Leu 53 in all 12 structures. Residues 59–62 form a hydrogen-bonded type III turn. Residues 74–77 form a type VIb β-turn with no CO_i–NH_{i+3} hydrogen bond. A hydrogen bond between the $O^\gamma H$ of Thr 77 (the resonance of which is observed in H_2O solution) and the CO of Gly 74 is found in 11 of the structures. Residues 8–12 clearly form a reverse turn, but the relatively high backbone rms deviations (due mainly to resonance overlap between Asp 9, Asp 10, and Ser 11) in this region preclude determination of its type. Residues 83–87 form a β-bulge loop connecting strands β_4 and β_5. Four interstrand backbone NH–CO hydrogen bonds are present in all structures (92–76, 78–90, 90–78, and 80–88). A further hydrogen bond (88–80) is present in eight of the structures, indicating the start of the β-bulge loop. Leu 80 CO is also hydrogen bonded to Ala 87 NH in all structures. The CO and NH groups of Lys 82 and Glu 85 are hydrogen bonded across the end of the loop in a type I′ β-turn, leaving Val 86 and Ala 87 as a bulge opposite Phe 81. This structure is similar to the "wide β-bulge" of Richardson (1981).

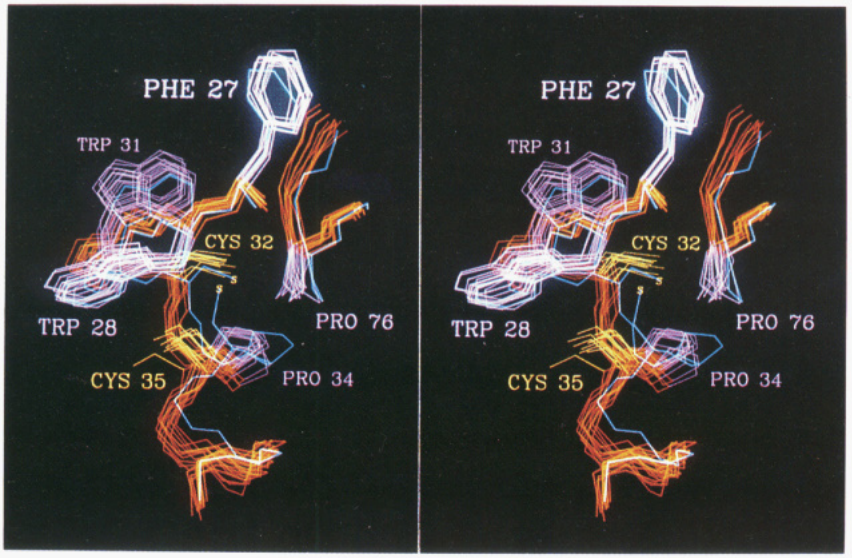

FIGURE 5: Comparison of the active site region (residues 27–37) of the 12 refined solution structures obtained for reduced thioredoxin with that of the X-ray crystal structure of oxidized thioredoxin (Katti et al., 1990). Backbone C^{α}, C, and N atoms of reduced thioredoxin are shown in red, with Phe, Trp, and Pro side chains in purple and Cys side chains in yellow. Both backbone and side-chain heavy atoms are shown in blue for oxidized thioredoxin.

Active Site Conformation. The active site of reduced thioredoxin contains the two sulfhydryl groups at positions 32 and 35 and is located on a protruding loop connecting strand β_2 and α_2 (Figure 5). The β-structure extends to Ala 29 in all structures, followed by a hydrogen-bonded type I turn between Ala 29 and Cys 32. Residues Pro 34 to Tyr 49 form the α_2-helix. The beginning of the helix appears to be wound more tightly than is customary for α-helix, with hydrogen bonds between Cys 32 and Cys 35 (9 structures) and Gly 33 and Lys 36 (11 structures) and bifurcated hydrogen bonds between Pro 34 and Met 37 (11 structures) and Pro 34 and Ile 38 (9 structures). The three structures that lack the Cys 32–Cys 35 hydrogen bond also have values for ψ (Cys 32) and ϕ (Gly 33) that are out of family, and the distance between the S^{γ} atoms is on average 1 Å longer in these structures than in the other nine.

Discussion

Thioredoxin is larger than other proteins for which NMR structures have been reported. Difficulties were encountered in assignment of NOE cross peaks because of the broadness of resonances and cross-peak overlap. To obtain high-resolution structures, it proved necessary to perform the calculations iteratively, beginning with a limited constraint set derived from cross peaks that could be unambiguously assigned and using the calculated structures to make additional NOE cross-peak assignments at each stage. Once a sufficient number of constraints had been identified in the spectra, the final round of structural refinement proceeded smoothly, using protocols that have been applied to many other proteins.

It is evident from Figure 6 and Table III that some regions of the structure are better constrained by the NMR data than others and that less well-constrained areas of the molecule also exhibit the greatest divergence between structures. Loops and turns tend to be less well-constrained, probably due to their position on the surface. In some external loops, a small number or even a single long-range NOE served to define the local conformation, indicating the necessity for extremely careful and conservative NOE cross-peak assignments if significant errors in NMR structures at this level of definition are not to result.

The structure of thioredoxin–$(SH)_2$ generated from NMR

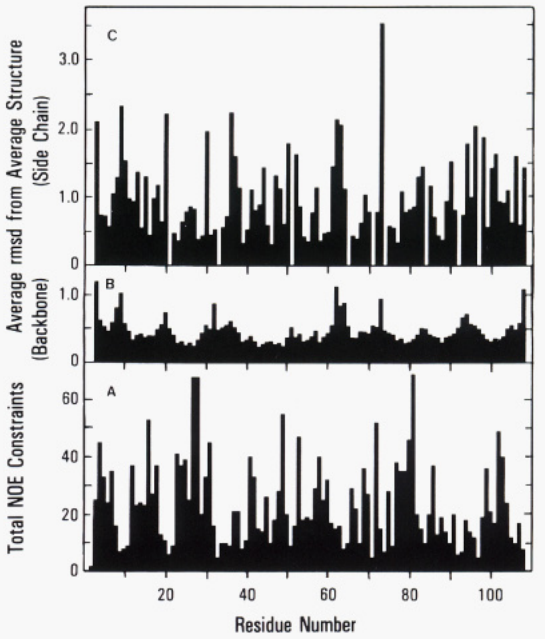

FIGURE 6: (A) Schematic diagram showing number of NOE connectivities (inter- and intraresidue) used in the final calculations versus residue number. (B) Average rms deviation from the mean for backbone heavy atoms as a function of residue number for 12 structures, superimposed between residues 3 and 108. (C) Average rms deviation from the mean for side-chain heavy atoms as a function of residue number for 12 structures, superimposed between residues 3 and 108.

constraints is similar overall to the corresponding structure obtained by X-ray crystallography for thioredoxin–S_2 (Holmgren et al., 1975; Katti et al., 1990) (Figure 2). However, some local differences are observed between the crystal structure of oxidized thioredoxin and the solution structures of the reduced protein, as evidenced by the large number of NOE constraints violated by the X-ray structure (Table II). Many of the violations greater than 1 Å involve the side chain of His 6 (violations up to 5.7 Å) and other residues of the β_1-strand. These differences probably arise from crystal packing and effects of coordination of Cu^{2+} ions at Ser 1, Asp 2, and Asp 10 (Katti et al., 1990). We note that the NOEs involving His 6 that are grossly violated by the

X-ray structure are also observed in the NOESY spectra of oxidized thioredoxin, suggesting differences between crystal and solution structures rather than between oxidized and reduced protein.[2] These issues will be addressed in detail once NMR structure determination for oxidized thioredoxin has been completed. Several other distance constraints for reduced thioredoxin that are violated by the X-ray structure of the oxidized protein are between protons in the active site and neighboring residues 58–60. Additional NOEs between protons of Cys 35 and Pro 76, predicted by the X-ray structure, are observed in the NOESY spectrum of oxidized thioredoxin but are absent from the spectra of the reduced protein.

The active site of thioredoxin is located on a surface loop and forms part of a hydrophobic surface that is probably involved in molecular recognition (Eklund et al., 1984). The 12 final NMR structures obtained for reduced thioredoxin are well-defined in the active site region and exhibit only subtle conformational differences from the X-ray structure of the oxidized protein (Figure 5). There is a small shift of the backbone between residues 33 and 36, and the side chains of Pro 34 and Pro 76 appear to shift slightly between oxidation states. The side chain of Cys 32 adopts a similar conformation in reduced and oxidized thioredoxin, but the Cys 35 side chain in the reduced state is rotated. This conformational difference is clearly evident in the χ_1 angle for Cys 35, which changes from $\chi_1 = -62°$ in the X-ray structure of the oxidized protein to $\chi_1 = 176 \pm 4°$ in the NMR structures of reduced thioredoxin. The distance between the S^γ atoms is increased from 2.05 Å in the disulfide bridge of the oxidized protein to a mean distance of 6.8 ± 0.6 Å in the NMR structures of reduced thioredoxin. (In a subset of structures that excludes three out-of-family structures, the mean S···S distance is 6.5 ± 0.1 Å.)

The rotation of the side chain of Cys 35 in reduced relative to oxidized thioredoxin brings its SH group away from the hydrophobic molecular surface formed by Gly 33, Pro 34, Ile 75, Pro 76, Gly 92, and Ala 93 previously implicated in binding interaction with other proteins (Eklund et al., 1984). The position of the SH group of Cys 32 remains essentially unchanged in the structures of reduced and oxidized thioredoxin. Chemical modification of thioredoxin-$(SH)_2$ with iodoacetic acid has shown that only Cys 32 reacts at pH 7.5 (Kallis & Holmgren, 1980). Furthermore, from inactivation rate versus pH, the pK_a value of the Cys 32 thiol group has been estimated to be 6.7. This suggests that Cys 32 provides the thiolate anion that is the nucleophile in thioredoxin-catalyzed protein disulfide reduction (Kallis & Holmgren, 1980). A mechanism has been proposed involving a transient mixed disulfide involving Cys 32 and one of the substrate half-cysteine residues (Holmgren, 1985; Kallis & Holmgren, 1980). The reduction cycle would be completed by attack of the thiolate anion of Cys 35 on the mixed disulfide to give thioredoxin-S_2 and a dithiol in the protein substrate. The elements of this proposed mechanism require further analysis using site-directed mutagenesis, enzyme kinetics, and structural studies by NMR.

At the present resolution of the NMR structures it is evident that, with the exception of the slight conformational changes in the active site region described above, the structures of reduced and oxidized thioredoxin are highly similar. In this context it is of considerable interest that the reduced protein is significantly less stable [10 °C in denaturation temperature (Hiraoki et al., 1988)] than the oxidized form. The additional stability of the oxidized thioredoxin cannot be attributed simply to the presence of the disulfide bridge: since only two residues separate the cysteine residues, only minimal thermal stabilization is to be expected (Pace et al., 1988). The explanation for the substantial difference in stability of the two oxidation states must be sought elsewhere, perhaps in differences in hydrogen bonding or other interactions in the active site region. Work is in progress in this laboratory to refine further the solution structures of reduced thioredoxin and to calculate and refine structures of the oxidized protein in order to better understand the relationship between molecular structure, stability, and catalytic activity.

[2] After submission of this paper, Dr. Hans Eklund provided additional X-ray coordinates for oxidized thioredoxin showing two alternate conformations for the side chain of His 6. One of these conformations is very similar to that determined for His 6 in reduced thioredoxin in solution; the other gives rise to the large constraint violations noted in the text. There is no indication that the latter conformation of His 6 is populated in solution; characteristic NOEs to Thr 8 predicted from the X-ray structure are absent from the NOESY spectra of both oxidized and reduced thioredoxin.

Acknowledgments

We thank Linda Tennant and Pia Lundman for expert technical assistance, Dr. Hans Eklund for making available coordinates of the refined X-ray crystal structure of thioredoxin-S_2 before publication, and Jerri Columpus and Sonia Reed for assistance with the manuscript.

References

Adler, S., & Modrich, P. (1983) *J. Biol. Chem. 258*, 6956–6962.

Arseniev, A., Schultze, P., Wörgötter, E., Braun, W., Wagner, G., Vasák, M., Kägi, J. H. R., & Wüthrich, K. (1988) *J. Mol. Biol. 201*, 637–657.

Berendsen, H. J. C., Postma, J. P. M., Van Gusteren, W. F., DiNola, A., & Haak, J. R. (1984) *J. Chem. Phys. 81*, 3684–3690.

Billeter, M., & Wüthrich, K. (1984) *J. Mol. Biol. 180*, 741–751.

Billeter, M., Braun, W., & Wüthrich, K. (1982) *J. Mol. Biol. 155*, 321–346.

Braun, W., Bosch, C., Brown, L. R., Go, N., & Wüthrich, K. (1981) *Biochim. Biophys. Acta 667*, 377–396.

Drobny, G., Pines, A., Sinton, S., Weitekamp, D., & Wemmer, D. (1979) *Symp. Faraday Soc. 13*, 49–55.

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

Dyson, H. J., Holmgren, A., & Wright, P. E. (1989) *Biochemistry 28*, 7074–7087.

Eklund, H., Cambillau, C., Sjöberg, B.-M., Holmgren, A., Jörnvall, H., Höög, J.-O., & Brändén, C.-I. (1984) *EMBO J. 3*, 1443–1449.

Gleason, F. K., & Holmgren, A. (1988) *FEMS Microbiol. Rev. 54*, 271–298.

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

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

Holmgren, A. (1972) *J. Biol. Chem. 247*, 1992–1998.

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

Holmgren, A., & Reichard, P. (1967) *Eur. J. Biochem. 2*, 187–196.

Holmgren, A., & Söderberg, B.-O. (1970) *J. Mol. Biol. 54*, 387–390.

Holmgren, A., & Roberts, G. (1976) *FEBS Lett. 71*, 261–265.

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

Huber, H. E., Russel, M., Model, P., & Richardson, C. C. (1986) *J. Biol. Chem. 261*, 15006–15012.

Jeener, J., Meier, B. H., Bachmann, P., & Ernst, R. R. (1977) *J. Chem. Phys. 71*, 4546-4553.
Kallis, G. B., & Holmgren, A. (1980) *J. Biol. Chem. 255*, 10261-10265.
Katti, S., LeMaster, D. M., & Eklund, H. (1990) *J. Mol. Biol.* (in press).
Kumar, A., Wagner, G., Ernst, R. R., & Wüthrich, K. (1981) *J. Am. Chem. Soc. 103*, 3654-3658.
Lee, M. S., Gippert, G., Soman, K. Y., Case, D. A., & Wright, P. E. (1989) *Science 245*, 635-637.
LeMaster, D. M., & Richards, F. M. (1988) *Biochemistry 27*, 142-150.
Marion, D. I., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967-974.
Mark, D. F., & Richardson, C. C. (1976) *Proc. Natl. Acad. Sci. U.S.A. 73*, 780-784.
Neuhaus, D., Wagner, G., Vasák, M., Kägi, J. H. R., & Wüthrich, K. (1985) *Eur. J. Biochem. 151*, 257-273.
Pace, C. N., Grimsley, G. R., Thomson, J. A., & Barnett, B. J. (1988) *J. Biol. Chem. 263*, 11820-11825.
Rance, M., Sørensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 117*, 479-485.
Richardson, J. S. (1981) *Adv. Protein Chem. 34*, 167-339.
Stryer, L., Holmgren, A., & Reichard, P. (1967) *Biochemistry 6*, 1016-1020.
Wagner, G., Braun, W., Havel, T. F., Schaumann, T., Go, N., & Wüthrich, K. (1987) *J. Mol. Biol. 196*, 611-641.
Weiner, S. J., Kollman, P. A., Nguyen, D. T., & Case, D. A. (1986) *J. Computat. Chem. 7*, 230-252.
Williamson, M. P., Havel, T. F., & Wüthrich, K. (1985) *J. Mol. Biol. 182*, 295-315.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.
Wüthrich, K., Billeter, M., & Braun, W. (1983) *J. Mol. Biol. 169*, 949-961.

Mitochondrial Energy-Linked Nicotinamide Nucleotide Transhydrogenase: Effect of Substrates on the Sensitivity of the Enzyme to Trypsin and Identification of Tryptic Cleavage Sites†

Mutsuo Yamaguchi, Sadao Wakabayashi,‡ and Youssef Hatefi*

Division of Biochemistry, Department of Molecular and Experimental Medicine, Research Institute of Scripps Clinic, La Jolla, California 92037

Received November 27, 1989; Revised Manuscript Received January 8, 1990

ABSTRACT: The mitochondrial nicotinamide nucleotide transhydrogenase catalyzes hydride ion transfer between NAD(H) and NADP(H) in a reaction that is coupled to proton translocation across the inner mitochondrial membrane. The enzyme (1043 residues) is composed of an N-terminal hydrophilic segment (~400 residues long) which binds NAD(H), a C-terminal hydrophilic segment (~200 residues long) which binds NADP(H), and a central hydrophobic segment (~400 residues long) which appears to form about 14 membrane-intercalating clusters of ~20 residues each. Substrate modulation of transhydrogenase conformation appears to be intimately associated with its mechanism of proton translocation. Using trypsin as a probe of enzyme conformation change, we have shown that NADPH (and to a much lesser extent NADP) binding alters transhydrogenase conformation, resulting in increased susceptibility of several bonds to tryptic hydrolysis. NADH and NAD had little or no effect, and the NADPH concentration for half-maximal enhancement of trypsin sensitivity of transhydrogenase activity (35 μM) was close to the K_m of the enzyme for NADPH. The NADPH-promoted trypsin cleavage sites were located 200-400 residues distant from the NADP(H) binding domain near the C-terminus. For example, NADPH binding greatly increased the trypsin sensitivity of the K_{410}-T_{411} bond, which is separated from the NADP(H) binding domain by the 400-residue-long membrane-intercalating segment. It also enhanced the tryptic cleavage of the R_{602}-L_{603} bond, which is located within the central hydrophobic segment. These results, which suggest a protein conformation change as a result of NADPH binding, have been discussed in relation to the mechanism of proton translocation by the transhydrogenase.

The mitochondrial nicotinamide nucleotide transhydrogenase catalyzes the stereospecific transfer of a hydride ion between the 4A position of NAD(H) and the 4B position of NADP(H) in a reaction that is coupled to proton translocation across the inner membrane with a H^+/H^- stoichiometry close to unity (see eq 1; H_c^+ and H_m^+ are protons on the cytosolic and the matrix side of the inner membrane, respectively) (Fisher & Earle, 1982).

$$NADH + NADP + H_c^+ \rightleftarrows NAD + NADPH + H_m^+ \quad (1)$$

The bovine enzyme has been sequenced from cDNA clones (Yamaguchi et al., 1988a), and the sequence of its signal peptide (43 residues) has been determined from the sequence of the mRNA (Yamaguchi et al., 1988b). The mature transhydrogenase has 1043 amino acids and a molecular

† Supported by U.S. Public Health Service Grant GM 24887.
* To whom correspondence should be addressed.
‡ Present address: Department of Biology, Faculty of Science, Osaka University, Toyonaka, Osaka 560, Japan.

0006-2960/90/0429-4136$02.50/0