J. Mol. Biol. (1991) **220**, 209–216

Determination of the Positions of Bound Water Molecules in the Solution Structure of Reduced Human Thioredoxin by Heteronuclear Three-dimensional Nuclear Magnetic Resonance Spectroscopy

Julie D. Forman-Kay[1], Angela M. Gronenborn[1]†, Paul T. Wingfield[2] and G. Marius Clore[1]†

[1]*Laboratory of Chemical Physics*
Building 2, National Institutes of Diabetes and Digestive and Kidney Diseases
National Institutes of Health, Bethesda, MD 20892, U.S.A.

[2]*Protein Expression Laboratory*
Building 6B, Office of the Director
National Institutes of Health, Bethesda, MD 20892, U.S.A.

(Received 21 March 1991; accepted 15 April 1991)

The presence of bound water molecules in the solution structure of reduced human thioredoxin has been investigated using three-dimensional ^{1}H rotating frame Overhauser ^{1}H–^{15}N multiple quantum coherence spectroscopy. It is demonstrated that the backbone amide protons of Lys21, Lys39, Lys82, Gly83 and Asn102, as well as the side-chain amide group of Asn102, are in close proximity to bound water molecules. Examination of the high-resolution solution structure of reduced human thioredoxin reveals that these results are best accounted for by four bound water molecules. Subsequent simulated annealing calculations carried out on the basis of interproton distance and hydrogen bonding restraints to the bound water molecules, supplemented by the original set of experimental restraints used in the calculation of the three-dimensional structure of human thioredoxin, permit a more precise localization of the bound water positions. Potential hydrogen bonds to these water molecules are described and a comparison is made to corresponding bound water molecules in the crystal structure of oxidized *Escherichia coli* thioredoxin.

Keywords: bound water; human thioredoxin; solution structure; 3D n.m.r.

Thioredoxin is a ubiquitous protein involved in a variety of oxidation–reduction reactions in widely differing species ranging from bacteriophages and prokaryotes to plant, avians and mammals (Holmgren, 1989). Thioredoxins are small stable proteins of molecular mass approximately 12 kDa with an active site comprising the amino acid sequence Cys-Gly-Pro-Cys, which forms a closed 14-membered ring in the oxidized species. In a series of recent papers we have presented the ^{1}H and ^{15}N nuclear magnetic resonance (n.m.r.‡) assignments (Formań-Kay *et al.*, 1989, 1990) of reduced human thioredoxin (105 residues) and the determination of its high-resolution three-dimensional structure in solution (Forman-Kay *et al.*, 1991). In this communication we extend the previous structural analysis to the determination and localization of bound water molecules using three-dimensional ^{1}H rotating frame Overhauser ^{1}H–^{15}N multiple quantum coherence (^{1}H–^{15}N ROESY-HMQC) spectroscopy. We show that four bound water molecules participate in stabilizing hydrogen bonds and

† Authors to whom all correspondence should be addressed.

‡ Abbreviations used: n.m.r., nuclear magnetic resonance; ROESY, rotating frame nuclear Overhauser spectroscopy; HMQC, heteronuclear multiple quantum coherence; ^{1}H–^{15}N ROESY-HMQC, 3-dimensional ^{1}H rotating frame Overhauser ^{1}H–^{15}N multiple quantum coherence spectroscopy; ROE, rotating frame nuclear Overhauser effect; NOE, laboratory frame nuclear Overhauser effect; NOESY, laboratory frame nuclear Overhauser effect spectroscopy; 3D, 3-dimensional; SA, simulated annealing; r.m.s. root-mean-square.

0022-2836/91/140209–08 $03.00/0

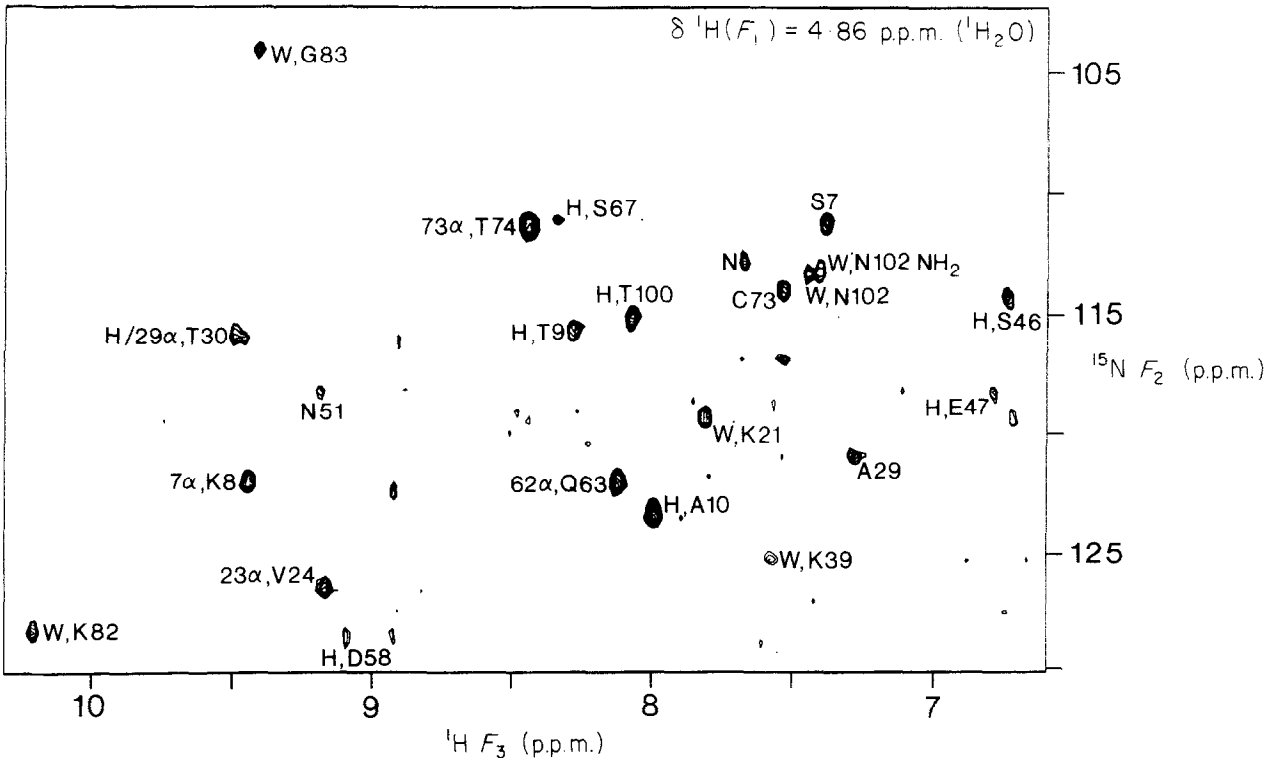

Figure 1. The $NH(F_3)$–$^{15}N(F_2)$ plane at the $^1H(F_1)$ frequency of the water resonance (4·86 p.p.m. (parts per million)) of the 600 MHz 3D ^1H–^{15}N ROESY-HMQC spectrum (mixing time 33 ms) of reduced recombinant human thioredoxin at 16 °C. Only negative cross-peaks are plotted; positive cross-peaks that arise from chemical exchange are not shown. Peaks in the spectrum are labeled i, j, where i refers either to a water proton (W), a protein hydroxyl or water proton (H), or an aliphatic proton, marked with the residue number and atom type, and j refers to the residue name and number of the amide ^{15}N and NH atoms. Intraresidue $C^\alpha H(i)$-$NH(i)$ ROE cross-peaks are labeled only with residue name and number. The peak marked N is from a side-chain amide NH_2 which could not be assigned, while other unlabeled peaks are derived from noise ridges or artifacts in the F_1 or F_2 dimension. The spectrum was recorded on a Bruker AM600 spectrometer as described by Clore *et al.* (1990). The sample contained ~2 mM-reduced human thioredoxin, 150 mM-sodium phosphate (pH 5·4), 0·02 mM-dithiothreitol, and 90% H_2O/10% ^2H_2O, and was maintained in a reduced state as described by Forman-Kay *et al.* (1989). The spectral width in the $^1H(F_1)$, $^{15}N(F_2)$ and $^1H(F_3)$ dimensions were 10·96 p.p.m. 28·36 p.p.m. and 14·62 p.p.m., respectively. The ^1H carrier position was set at the water frequency of 4·86 p.p.m. and the ^{15}N carrier was placed at 116·45 p.p.m. The spectrum was recorded with 128 complex × 32 complex × 512 real points in t_1, t_2 and t_3, respectively, and the total measuring time was ~45 h. After zero-filling in all dimensions, the absorptive part of the final processed data matrix comprised 256 $(F_1) \times 64$ $(F_2) \times 512$ (F_3) points. The spectrum was processed on a Sun Sparc workstation using in-house routines (Kay *et al.*, 1989) for the Fourier transform in the F_2 dimension, together with the commercially available software package NMR2 (New Methods Research, Inc. Syracuse, NY) for processing the F_1–F_3 planes.

constitute an integral part of the solution structure of human thioredoxin.

The basis for the detection of bound water lies in the rotating frame Overhauser effect (ROE), which permits one to distinguish through-space (<4 Å; 1 Å = 0·1 nm) nuclear Overhauser interactions from chemical exchange, as these two processes give rise to cross-peaks of opposite sign (Bothner-By *et al.*, 1984; Bax & Davis, 1985). In a two-dimensional ROESY experiment, ROE cross-peaks between protein protons and water are located along a single trace of the spectrum, parallel to the F_2 axis at the F_1 frequency of the H_2O resonance. Because of extensive resonance overlap such an experiment is only suitable for small proteins (Otting & Wüthrich, 1989). For larger proteins, it is therefore necessary to record a 3D ^1H–^{15}N ROESY-HMQC spectrum in which the ^1H–^1H interactions are separated out according to the chemical shift of the ^{15}N atoms directly attached to the destination protons, namely backbone and side-chain amide protons (Clore *et al.*, 1990). All interactions with water can then be observed in a single $NH(F_3)$–$^{15}N(F_2)$ cross-section through the 3D spectrum at the $^1H(F_1)$ frequency of the water resonance.

The $NH(F_3)$–$^{15}N(F_2)$ plane at the F_1 ^1H_2O frequency of the 3D ^1H–^{15}N ROESY-HMQC spectrum of reduced human thioredoxin recorded at 16 °C is shown in Figure 1. To facilitate the distinction of ROE effects involving bound water from those involving $C^\alpha H$ protons with chemical shifts close to or degenerate with the water resonance, the experiment was also recorded at 25 °C. The signal-to-noise ratio for each experiment was optimized by choosing a mixing time approximately equal to the average spin-locked relaxation time, $T_{1\rho}$, of the

amide protons (~33 ms at 16°C and ~42 ms at 25°C) (Bothner-By *et al.*, 1984; Bax & Davis, 1985). About 30 negative cross-peaks, representing potential ROE interactions between protein amide and H_2O protons, were identified in this cross-section of the two spectra. Over half of these corresponded to ROEs between NH and $C^{\alpha}H$ protons, while most of the remaining peaks could be accounted for by ROEs between NH and hydroxyl protons on the basis of close proximity to a Ser or Thr side-chain hydroxyl in the solution structure of the protein (Forman-Kay *et al.*, 1991). In the final analysis, only six cross-peaks could be attributed unambiguously to ROE interactions with bound water. These involved the NH protons of Lys21, Lys39, Lys82, Gly83 and Asn102, and the side-chain NH_2 group of Asn102. None of these amide protons were within 3·5 Å of a rapidly exchanging side-chain proton, effectively ruling out the possibility of indirect magnetization transfer to water. Due to the close spatial proximity in the human thioredoxin structure of the NH protons of Lys82 and Gly83, and similarly the NH and NH_2 protons of Asn102, the ROEs with water could be accounted for by four bound water molecules (referred to hereinafter as W1 to W4).

To further characterize the positions of these bound water molecules, a series of distance restraints suitable for simulated annealing calculations (Nilges *et al.*, 1988*a*,*b*,*c*) was derived. Since ROE intensities provide good quantitative distance information (Bax *et al.*, 1986; Clore *et al.*, 1990; Bauer *et al.*, 1990), experimental interproton distance restraints between the NH and water protons were obtained by comparison with intensities of two strong $C^{\alpha}H(i)$–$NH(i+1)$ sequential ROEs observed at the water frequency between Cys73 and Thr74 and between Asn93 and Lys94, and their corresponding interproton distances found in the previously calculated structures (2·2 to 2·3 Å). The calibration also took into account the fact that a peak in the ^{1}H–^{15}N ROESY-HMQC spectrum arising from a protein–protein interaction appears to be about twice as intense as one reflecting the same interproton distance from protein to water, due to the substantially longer T_1 value of water relative to that of protein protons (Clore *et al.*, 1990). The resulting upper interproton distance limits were 2·7 Å from W1 protons to Lys21 NH and from W3 to Lys82 NH and Gly83 NH, 2·8 Å from W4 to Asn102 NH and NH_2, and 3·0 Å from W2 to Lys39 NH. The lower distance bounds in all cases were set to 1·8 Å, the sum of the van der Waals' radii of two protons.

In general, the ROE experiment can only detect interactions involving bound water molecules with a lifetime of greater than about one nanosecond, which implies that their positions are stabilized by hydrogen bonds (Clore *et al.*, 1990). Consequently, in addition to the interproton distance restraints to the water, hydrogen bonding restraints were derived from the water protons to all main-chain or side-chain carbonyl atoms within ~6 Å of the co-ordinates of the amide to which an ROE was observed and whose orientation might allow favorable hydrogen bonding interactions (Clore *et al.*, 1991). As the actual identity of the hydrogen bonding partners is unknown, each hydrogen bonding restraint was described as an $\langle r^{-6} \rangle^{-1/6}$ average over all possibilities (i.e. to carbonyl *a* or carbonyl *b* or ... or carbonyl *n*). Because this quantity is heavily weighted towards the shortest distance(s) within each group, it permits the appropriate hydrogen bonding partner(s) to be selected automatically during the course of the simulated annealing calculations (Clore *et al.*, 1985). To ensure suitable geometry for the hydrogen bonding interactions, two distance restraints were used for each hydrogen bond, one from the donor heavy atom to the acceptor (2·4 to 3·5 Å), and one from the proton to the acceptor (1·4 to 2·5 Å) (Baker & Hubbard, 1984). Only hydrogen bond restraints with the water protons acting as donors to protein carbonyl atoms were used initially since these interactions are found to be present in nearly all protein crystal structures containing water, while hydrogen bonds from protein amide atoms to water oxygen atoms are less frequent (Baker & Hubbard, 1984).

Structure computations were carried out in an iterative manner with the program XPLOR (Brünger *et al.*, 1986; Brünger, 1988), using a minor modification of the simulated annealing protocol described by Nilges *et al.* (1988*a*). The target function comprises harmonic potential terms for covalent geometry (i.e. bonds, angles, planes and chirality), a simple quartic van der Waals' repulsion term to describe the non-bonded contacts, and square-well potential terms for the experimental interproton distance, torsion angle and hydrogen bonding restraints. Initially, the force constants for the repulsive van der Waals' term and experimental distance and torsion angle restraints terms were set to low values of 0·001 kcal mol^{-1} $Å^{-4}$, 0·5 kcal mol^{-1} $Å^{-2}$ and 0·5 kcal mol^{-1} $Å^{-2}$, respectively (1 cal = 4·184 J), so that the bound water molecules and protein atoms could move freely. These force constants were increased over 50 cycles of 75 femtoseconds of annealing at 1000 K to 0·25 kcal mol^{-1} $Å^{-4}$, 50 kcal mol^{-1} $Å^{-2}$ and 200 kcal mol^{-1} $Å^{-2}$, respectively. This was followed by slow cooling to 300 K over 25 cycles of 50 femtoseconds of annealing with the repulsive van der Waals' force constant set to 4·0 kcal mol^{-1} $Å^{-4}$ and, finally 500 steps of restrained Powell minimization.

The starting co-ordinates for simulated annealing were obtained by adding the four water molecules to the previously determined set of 33 simulated annealing (SA) structures (Forman-Kay *et al.*, 1991) at positions close to the co-ordinates of the amide protons to which ROEs were observed and of the possible carbonyl oxygen hydrogen bond acceptors. In the first series of calculations, the set of 2276 n.m.r.-derived restraints used to determine the original solution structure of human thioredoxin (Forman-Kay *et al.*, 1991) was supplemented by the additional restraints to the four bound water mole-

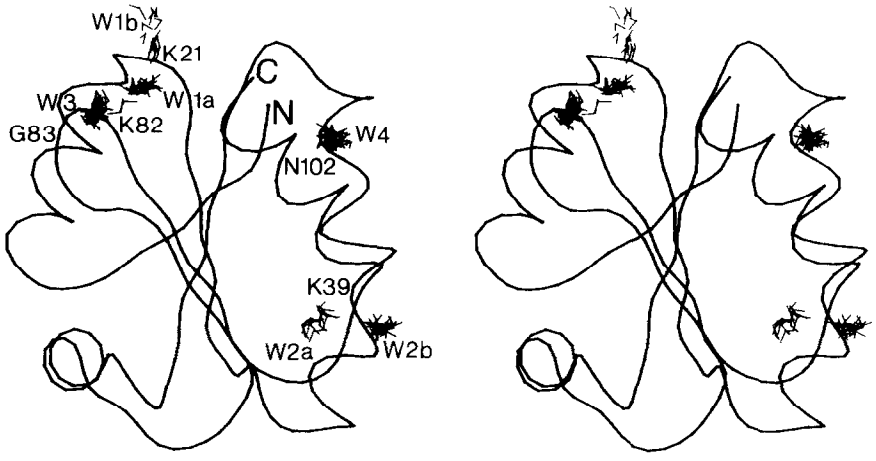

Figure 2. Stereo view showing a superposition of the bound water molecules calculated for all 33 simulated annealing structures on a smoothed backbone ribbon drawing of the restrained minimized mean structure, (SA)*r*, of reduced human thioredoxin. The amide protons of the residues to which ROEs from water protons are observed have been labeled, and the water molecules are indicated by W1 to W4. The restrained minimized mean structure (SA)*r* is obtained by restrained minimization of the mean structure SA using the same target function employed in the simulated annealing calculations. The mean structure SA is derived by averaging the co-ordinates of all 33 simulated annealing structures best fitted to each other.

cules described above. In some of the resulting structures, W1, W2 and W4 were found to be close (<2·5 Å) to the NH protons of Asp20, Ile38 and Glu103, respectively, although no corresponding ROEs for these interactions could be observed. Therefore, in the second set of calculations, repulsive distance restraints were added, imposing lower bounds of 2·8 Å between the water molecules and these three NH protons. In the third and final set of calculations, more potential hydrogen bonding partners with bound water were added on the basis of proximity and appropriate orientation, as well as additional hydrogen bonding restraints with water oxygen acting as an acceptor for protein protons. In particular, these involved the water oxygen atom of W3 acting as an acceptor for the NH protons of Lys82 and Gly83, and the water oxygen atom of W4 acting as an acceptor for the NH and NH_2 protons of Asn102 and the side-chain $O^{\eta}H$ proton of Tyr49. These restraints were also treated as $\langle r^{-6} \rangle^{-1/6}$ averages over all possible proton donors (H_a or H_b or . . . H_n), adding two more modeled hydrogen bonds.

The positions of the four bound water molecules in the 33 SA structures superimposed on an overall smoothed backbone atom representation of the restrained minimized mean structure, (SA)*r*, are shown in Figure 2, while Figure 3 provides more detailed stereo views illustrating the potential stabilizing interactions. In the case of W1 and W2 an exact determination of the location of the water could not be made, since two different possible positions consistent with the experimental data are observed in the collection of SA structures. This probably reflects the smaller number of restraints involving these two water molecules. The atomic r.m.s. deviation of the water oxygen atoms from their mean co-ordinate positions within each group is as follows: W1a, 1·6(±1·1) Å; W1b, 1·6(±0·7) Å; W2a, 0·7(±0·3) Å; W2b, 0·4(±0·2) Å; W3, 0·8(±0·4) Å; W4, 0·6(±0·3) Å.

The atomic r.m.s. shift between the mean co-ordinate positions before and after the inclusion of water is 0·19 Å for backbone atoms and 0·31 Å for all atoms. This is well within the atomic r.m.s. distribution of the original set of 33 structures about their mean co-ordinate positions, indicating that the structures were not perturbed in any significant way by the addition of the water. In addition, the atomic r.m.s. distribution of the current bound water structures about the mean co-ordinate positions is 0·40(±0·05) Å for backbone and 0·80(±0·05) Å for all atoms, which is virtually identical to the original r.m.s. deviations (Forman-Kay *et al.*, 1991). An analysis of the atomic r.m.s. shift of the co-ordinates of the two average structures as a function of residue number parallels the r.m.s. deviation among the set of original SA structures, showing that no local changes are induced by the water. Finally, the average atomic r.m.s. shift between the initial and final co-ordinates for the 33 structures is 0·60(±0·08) Å for backbone and 1·15(±0·08) Å for all atoms. These values confirm the statistical expectation that the average atomic r.m.s. difference between all possible pairwise combinations of n structures is $[2n/(n-1)]^{1/2}$ times the atomic r.m.s. distribution about the mean co-ordinate positions (Clore & Gronenborn, 1989), indicating that during the course of the simulated annealing calculations the atomic co-ordinates can move freely from their starting positions to any position within the confines of the conformational space consistent with the experimental data.

The r.m.s. deviations of the calculated structures from the experimental restraints and from idealized covalent geometry were essentially the same as those found for the original solution structures

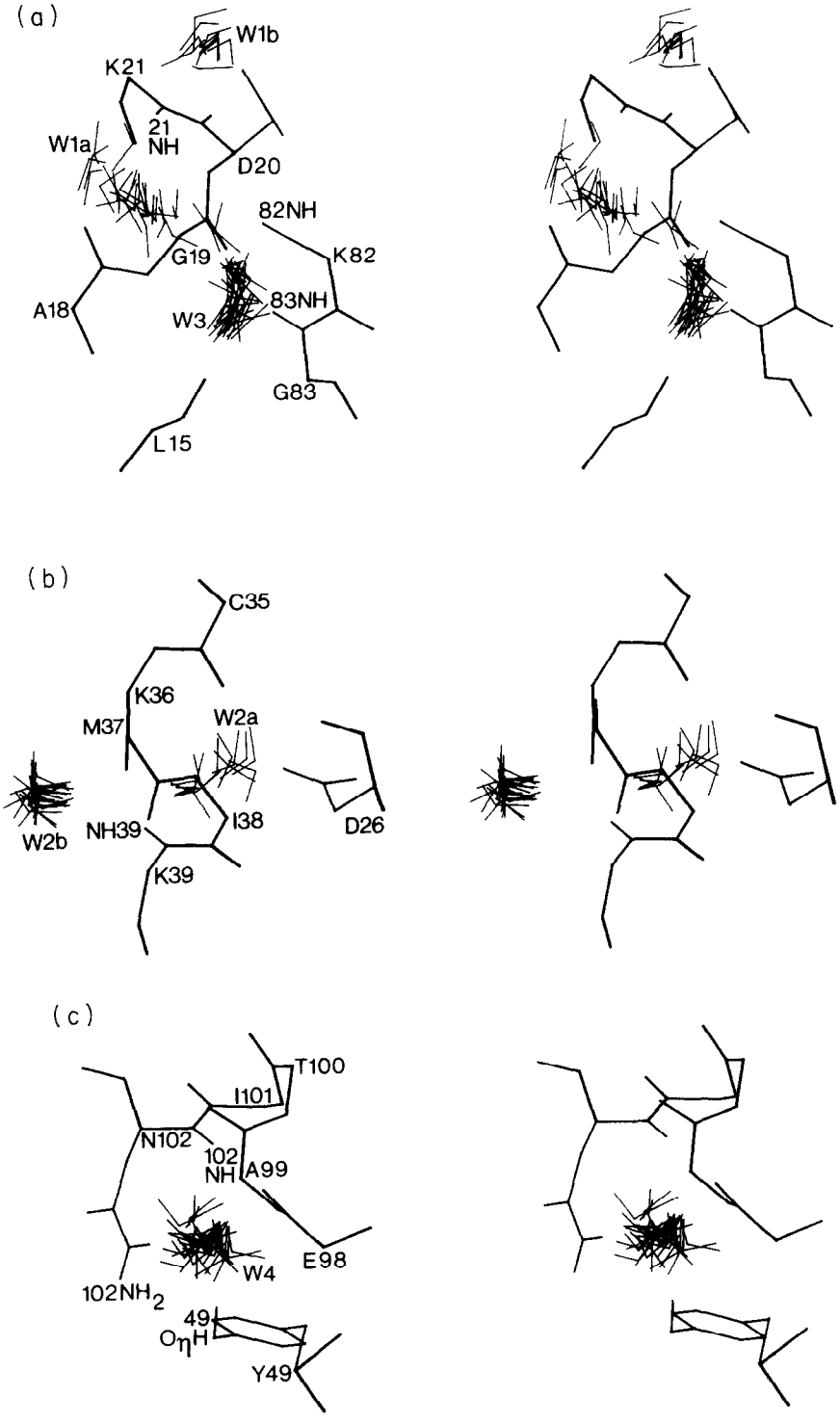

Figure 3. Stereo views showing details of the interactions involving the bound water molecules of all 33 simulated annealing structures superimposed on the restrained minimized mean structure (SA)*r*. (a) Waters W1 and W3 and residues 15, 18 to 21 and 82 to 83, with the side-chain of Asp20; (b) W2 and residues 26 and 35 to 39, with the side-chain of Asp26; and (c) W4 and residues 49 and 98 to 102, with side-chains of Tyr49 and Asn102. The backbone atoms as well as potential protein hydrogen bonding partners to the water are illustrated. Residue numbers are indicated at the C^{α} position, amide protons involved in ROEs to water are also labeled, and the backbone (N, C^{α}, C′ and O) atoms are drawn with heavy lines.

(Forman-Kay *et al.*, 1991), with the addition of a 0·010(±0·014) Å r.m.s. deviation from the nine experimental ROE-derived water-amide distance restraints (6 attractive and 3 repulsive), and 0·016(±0·007) Å r.m.s. deviation from the 12 hydrogen bonding restraints for the six modeled interactions to four groups of potential carbonyl acceptors and two groups of possible amide proton donors. No distance violations greater than 0·3 Å were seen for any of these restraints. The values of

the various pseudo-energy terms for the experimental restraints, covalent geometry and non-bonded contacts, as well as the calculated 6–12 Lennard–Jones van der Waals' energy, were also virtually identical to those found in the previously calculated structures lacking water (Forman-Kay *et al.*, 1991).

These results correlate with previous analyses of water structure from crystallographic studies, which show bound water involved in bridging hydrogen bonds located near bends and distortions of α-helices, in turns and between strands of β-sheets (Blundell *et al.*, 1983; Edsall & MacKenzie, 1983; Baker & Hubbard, 1984). Thus, in the structure of reduced human thioredoxin, W1 is close to an irregular loop, W2 to a proline-induced kink in α-helix α_2, W3 to a type I′ hairpin β-turn between strands β_4 and β_5, and W4 to the C-terminal end of helix α_4. Each water molecule is involved in multiple hydrogen bonding interactions. Not all of the possibilities, however, which were given in the hydrogen bond restraints list, are actually present in the structures, indicating that the resulting hydrogen bonds are significant. Hydrogen bonds are considered to be present if the distance between the acceptor and donor heavy atoms is less than 3·5 Å and the angle between the acceptor, hydrogen and donor heavy atom is greater than 90° in at least 20% of the calculated structures. All four water molecules act as hydrogen bond donors to one or more protein carbonyl groups. W1 is involved in interactions with the backbone carbonyl atoms of Ala18, Gly19, Asp20 or Lys21 in a loop between α_1 and β_2. The W1b group displays interactions with only the backbone and side-chain carbonyl atoms of Asp20. W2 interacts with the backbone carbonyls of residues one helical turn away from Lys39 to whose NH it shows an ROE, with W2a and W2b showing a preference for interactions with Cys35 and Lys36, respectively. The backbone carbonyl atoms of Leu15 and Gly19 hydrogen bond to W3. Finally, W4 hydrogen bonds to the carbonyl atom of Glu98 and the side-chain hydroxyl oxygen of Tyr49. The oxygen atoms of W3 and W4 also act as hydrogen bond acceptors for the amide protons to which they show ROEs, namely, the backbone NH protons of Lys82 and Gly83 in the case of W3 and the backbone NH and side-chain NH_2 protons of Asn102 in the case of W4. In addition, the oxygen atom of W4 appears to accept a hydrogen bond from the side-chain O$^{\eta}$H atom of Tyr49.

Bound water molecules that participate in hydrogen bonds with protein atoms constitute an integral part of the protein and are often referred to as "internal" waters, as opposed to "external" hydration layer solvent molecules. All four bound waters of human thioredoxin, in at least one of the possible positions (W1a, W2a, W3 and W4), are buried within the protein, with small average accessible surface areas for the oxygen atom ranging from 0 to 6 Å^2. The average accessible surface areas for W1b and W2b, however, are slightly greater than 25 Å^2. In general, internal waters in protein structures are observed to participate in three to four hydrogen bonds per solvent molecule (Baker & Hubbard, 1984). In human thioredoxin, the average number of hydrogen bonding interactions per structure is 2 for W1a, 1 to 1·5 for W2a and W2b, ~2·5 for W3 and >3 for W4. The most external water group W1b, however, has significantly fewer interactions than W1a, averaging only about one per structure. The bound water molecules which have both donor and acceptor characteristics are involved in bridging hydrogen bonds. Thus, W3 bridges hydrogen bonds from the NH protons of Lys82 and Gly83 to the carbonyl atoms of Leu15 and Gly19, while W4 bridges hydrogen bonds from the backbone and side-chain amide protons of Asn102 and the side-chain hydroxyl proton of Tyr49 to the carbonyl atom of Glu98.

Recently a refined crystal structure of oxidized *Escherichia coli* thioredoxin has been determined at 1·65 Å resolution (Katti *et al.*, 1990). This bacterial thioredoxin has ~25% sequence identity to the human enzyme and a very similar overall conformation (Forman-Kay *et al.*, 1991). In the *E. coli* thioredoxin crystal there are two molecules within the unit cell surrounded by 140 water molecules. It is therefore of interest to ascertain whether any of these water molecules correspond to the four found in the solution structure of human thioredoxin. To this end the mean structure (SA) of human thioredoxin was best fitted to each of the two *E. coli* thioredoxin molecules, referred to as A and B, using the program O (Jones & Kjeldgaard, 1990), and the correspondence of water molecules was made by visual inspection on an Evans & Sutherland PS390 graphics system and by analysis of hydrogen bonding interactions. Since the structural homology is weakest in the loop region between α_1 and β_2 and the precision of the localization of W1 is only at the 1·5 Å level, it is difficult to determine exact correspondences for this water molecule. A number of possibilities exist: in particular, W89A and W25A from molecule A and W133B and W40B from molecule B exhibit similar hydrogen bonding interactions to W1. More convincing matches are observed for the other bound water molecules. W3A and W80B correspond to W2a, and W47A, W28A and W24B display similar hydrogen bonds to W3. While there is no water in molecule A of the *E. coli* structure that corresponds to W4 in the human structure, the oxygen of W110B from molecule B corresponds exactly to the position of W4. Thus, W110B hydrogen bonds to the NH of Asp104 and the O$^{\eta}$H of Tyr49, which are equivalent to Asn102 and Tyr49 of the human protein, respectively. There appears to be no bound water molecule in either of the two *E. coli* molecules that is equivalent to W2b.

It is interesting to note that W25A/W40B and W3A/W80B, which correspond to W1 and W2a, respectively, are two of the 25 pairs of water molecules related by non-crystallographic symmetry in the *E. coli* structure with 0·22 Å and 0·24 Å deviations, respectively (Katti *et al.*, 1990). Further,

those water molecules in the *E. coli* structure that match W2a and W3 are among the 20 solvent molecules with the lowest temperature factors ($\leq$ 20 Å^2) and occupancies of close or equal to unity ($\geq$0·99). Water molecules corresponding to W1 and W4, on the other hand, have somewhat higher temperature factors (25 to 44 Å^2) and lower occupancies (0·73 to 1·00). The presence of corresponding water molecules in the human and *E. coli* thioredoxin structures reinforces the notion that tightly bound water detected by n.m.r. constitutes an integral component of the structure which may be conserved in homologous proteins.

A number of speculations regarding the role of water in stabilizing structural features of reduced human thioredoxin may be made on the basis of the results presented in this paper. Although thioredoxin has a large amount of secondary structure, comprising both β-sheet and α-helices, and many of the backbone amide and carbonyl atoms are involved in hydrogen bonding within these structural elements, the turns and loop regions leave numerous unsatisfied hydrogen bonds. The presence of two bound water molecules which participate in bridging hydrogen bonds in or adjacent to the loop between α_1 and β_2 (W1) and between this loop and the type I' β-turn connecting β4 and β5 (W3), could add significant stabilization to the conformation in this region. In addition, distortions of α-helices also leave hydrogen bonds unsatisfied. The α_2 helix of thioredoxin, with a proline in the center at position 40, has a highly irregular structure with very few $i, i+4$ hydrogen bonds. The bound water, W2, located at a bend in the helix, could stabilize as well as facilitate the induction of curvature by means of a bifurcated hydrogen bond from the protein carbonyl atom to both the solvent and backbone amide atom (Blundell *et al.*, 1983). Of the two possible locations identified for W2, the W2a position seems more likely owing to its more internal location relative to W2b, and the fact that its position is closely related to that of a bound water molecule in the *E. coli* structure. It is also interesting to note that W2a is reasonably close to the side-chain carboxyl group of Asp26 (Fig. 3(b)), one of the three buried acidic amino acid residues in the protein, and may possibly be involved in a stabilizing interaction analogous to that observed in the *E. coli* enzyme. The bound water W4 is located in a surface crevice on the C-terminal helix α_4, possibly serving as a stabilizing link between helices α_4 and α_2 at this location.

In conclusion, the positions of four internal bound water molecules have been determined in the solution structure of reduced human thioredoxin. These water molecules are involved in stabilizing specific structural features by interacting with backbone and side-chain atoms which would otherwise have unsatisfied hydrogen bonds. As these stabilizing elements within the protein must have a lifetime greater than about one nanosecond to permit their observation by n.m.r. (Clore *et al.*, 1990), these four bound water molecules constitute a significant component of the solution structure of the protein. Knowledge of the location of these bound water molecules may increase understanding of the role of solvent in the structure of this particular protein, as well as of the general role of water in protein function and conformational stability.

This work was supported by the AIDS Directed Anti-Viral Program of the Office of the Director of the National Institutes of Health (G.M.C., A.M.G. and P.T.W.).

References

Baker, E. N. & Hubbard, R. E. (1984). Hydrogen bonding in globular proteins. *Prog. Biophys. Mol. Biol.* **44**, 97–179.

Bauer, C. J., Frenkiel, T. A & Lane, A. M. (1990). A comparison of the ROESY and NOESY experiments for large molecules with applications to nucleic acids. *J. Magn. Reson.* **87**, 144–152.

Bax, A. & Davis, D. G. (1985). Practical aspects of two-dimensional transverse NOE spectroscopy. *J. Magn. Reson.* **63**, 207–213.

Bax, A., Sklenar, V. & Summers, M. F. (1986). Direct identification of relayed nuclear Overhauser effects. *J. Magn. Reson.* **70**, 327–331.

Blundell, T., Barlow, D., Borkakoti, N. & Thornton, J. (1983). Solvent-induced distortions and the curvature of α-helices. *Nature (London)*, **306**, 281–283.

Bothner-By, A. A., Stephens, R. L., Lee, J. T., Warren, C. D. & Jeanloz, R. W. (1984). Structure determination of a tetrasaccharide: transient nuclear Overhauser effect in the rotating frame *J. Amer. Chem. Soc.* **106**, 811–813.

Brünger, A. T. (1988). *X-PLOR Manual*, Yale University, New Haven, CT.

Brünger, A. T., Clore, G. M., Gronenborn, A. M. & Karplus, M. (1986). Three-dimensional structures of proteins determined by molecular dynamics with interproton distance restraints: application to crambin. *Proc. Nat. Acad. Sci., U.S.A.* **83**, 3801–3805.

Clore, G. M. & Gronenborn, A. M. (1989). Determination of three-dimensional structures of proteins and nucleic acid in solution by nuclear magnetic resonance spectroscopy. *CRC Crit. Rev. Biochem. Mol. Biol.* **24**, 479–564.

Clore, G. M., Gronenborn, A. M., Brünger, A. T. & Karplus, M. (1985). Solution conformation of a heptadecapeptide comprising the DNA binding helix F of the cyclic AMP receptor protein of *Escherichia coli*: combined use of ^{1}H nuclear magnetic resonance and restrained molecular dynamics *J. Mol. Biol.* **186**, 435–455.

Clore, G. M., Bax, A., Wingfield, P. T. & Gronenborn, A. M. (1990). Identification and localization of bound internal water in the solution structure of interleukin 1β by heteronuclear three-dimensional ^{1}H rotating-frame Overhauser ^{15}N–^{1}H multiple quantum coherence NMR spectroscopy. *Biochemistry*, **29**, 5671–5676.

Clore, G. M., Wingfield, P. T. & Gronenborn, A. M. (1991). High-resolution three-dimensional structure of interleukin-1β in solution by three and four-dimensional nuclear magnetic resonance spectroscopy. *Biochemistry*, **30**, 2315–2323.

Edsall, J. T. & McKenzie, H. A. (1983). Water and proteins II: The location and dynamics of water in

protein systems and its relation to their stability and properties. *Advan. Biophys.* **16**, 53–183.

Forman-Kay, J. D., Clore, G. M., Driscoll, P. C., Wingfield, P. T., Richards, F. M. & Gronenborn, A. M. (1989). A proton nuclear magnetic resonance assignment and secondary structure determination of recombinant human thioredoxin. *Biochemistry*, **28**, 7088–7097.

Forman-Kay, J. D., Gronenborn, A. M., Kay, L. E., Wingfield, P. T. & Clore, G. M. (1990). Studies on the solution conformation of human thioredoxin using heteronuclear ^{15}N–^{1}H nuclear magnetic resonance spectroscopy. *Biochemistry*, **29**, 1566–1572.

Forman-Kay, J. D., Clore, G. M., Wingfield, P. T. & Gronenborn, A. M. (1991). The high resolution three-dimensional structure of reduced recombinant human thioredoxin in solution. *Biochemistry*, **30**, 2685–2698.

Holmgren, A. (1989). Thioredoxin and glutaredoxin systems. *J. Biol. Chem.* **264**, 13963–13966.

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

Katti, S. K., LeMaster, D. M. & Eklund, H. (1990). Crystal structure of thioredoxin from *Escherichia coli* at 1·68 Å resolution. *J. Mol. Biol.* **212**, 167–184.

Kay, L. E., Marion, D. & Bax, A. (1989). Practical aspects of 3D heteronuclear NMR of proteins. *J. Magn. Reson.* **84**, 72–84.

Nilges, M., Clore, G. M. & Gronenborn, A. M. (1988*a*). Determination of three-dimensional structures of proteins from interproton distance data by hybrid distance geometry-simulated annealing calculations. *FEBS Letters*, **299**, 317–324.

Nilges, M., Gronenborn, A. M., Brünger, A. T. & Clore, G. M. (1988*b*). Determination of three-dimensional structures of proteins by simulated annealing with interproton distance restraints: application to crambin, potato carboxypeptidase inhibitor and barley serine proteinase inhibitor. *Protein Eng.* **2**, 27–38.

Nilges, M., Clore, G. M. & Gronenborn, A. M. (1988*c*). Determination of three-dimensional structures of proteins from interproton distance data by dynamical simulated annealing from a random array of atoms. *FEBS Letters*, **239**, 129–136.

Otting, G. & Wüthrich, K. (1989). Studies of protein hydration in aqueous solution by direct NMR observation of individual protein-bound water molecules. *J. Amer. Chem. Soc.* **111**, 1871–1875.

Edited by P. E. Wright