Nuclear magnetic resonance characterization of the N-terminal thioredoxin-like domain of protein disulfide isomerase

JOHAN KEMMINK,¹ NIGEL J. DARBY,¹ KLAAS DIJKSTRA,² RUUD M. SCHEEK,²
AND THOMAS E. CREIGHTON¹

¹European Molecular Biology Laboratory, Meyerhofstrasse 1, D-69012 Heidelberg, Germany

²The Groningen Biomolecular Science and Biotechnology Institute, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands

(RECEIVED August 22, 1995; ACCEPTED October 4, 1995)

Abstract

A genetically engineered protein consisting of the 120 residues at the N-terminus of human protein disulfide isomerase (PDI) has been characterized by ¹H, ¹³C, and ¹⁵N NMR methods. The sequence of this protein is 35% identical to *Escherichia coli* thioredoxin, and it has been found also to have similar patterns of secondary structure and β -sheet topology. The results confirm that PDI is a modular, multidomain protein. The last 20 residues of the N-terminal domain of PDI are some of those that are similar to part of the estrogen receptor, yet they appear to be an intrinsic part of the thioredoxin fold. This observation makes it unlikely that any of the segments of PDI with similarities to the estrogen receptor comprise individual domains.

Keywords: domain structure; estrogen receptor; NMR; protein disulfide isomerase; thioredoxin; triple resonance

Protein disulfide isomerase (PDI) is a major protein of the endoplasmic reticulum of eukaryotes that is involved in disulfide bond formation in proteins translocated into the endoplasmic reticulum (reviewed by Noiva & Lennarz [1992] and Freedman [1992]). Genetic analysis has shown that it is essential for yeast viability (Farquhar et al., 1991; Scherens et al., 1991). It also has other functions in the assembly of oligomeric proteins. It is the β subunit of prolyl-4-hydroxylase, which has an $\alpha_2\beta_2$ quaternary structure (Pihlajaniemi et al., 1987) and is one subunit of the heterodimeric triglyceride transfer protein complex (Wetterau et al., 1990, 1991).

Although its 3D structure has not been determined, the amino acid sequence of PDI (Edman et al., 1985) provides insights into

its structure and reaction mechanisms. Two segments of the sequence, designated *a* and *a'*, are homologous to the important redox protein thioredoxin. One of this protein's functions is to reduce various cellular constituents, in a reaction mechanism that cycles the active site -Cys-Gly-Pro-Cys- between the dithiol and disulfide forms (Holmgren, 1985). The corresponding sequence in both the *a* and *a'* segments of PDI is -Cys-Gly-His-Cys-. The involvement of these residues in catalyzing the incorporation and rearrangement of disulfide bonds in proteins has been confirmed (Krause et al., 1991; Lyles & Gilbert, 1991a, 1991b, 1994; Freedman, 1992; Lundström et al., 1992; Vuori et al., 1992; Darby & Creighton, 1995).

In addition to the homologous *a* and *a'* segments, PDI has two other segments that are homologous to each other but that are not detectably homologous to any other protein; they are designated *b* and *b'* and comprise, respectively, residues 154–225 and 257–333 of mature human PDI. In addition, there is a C-terminal segment (designated *c*) of 21 residues that is predominantly acidic in nature. The order in the primary structure of these segments is *a-b-b'-a'-c* (Edman et al., 1985).

More recently, homology of PDI with the estrogen-binding domain of the estrogen receptor has been detected, and some activities of PDI have been shown to be inhibited by estrogens (Tsibris et al., 1989). PDI has also been found to be involved in complexes of the estrogen receptor with its DNA (Landel et al., 1994). Residues 101–144 of mature PDI are similar to residues 350–392 of the estrogen receptor and correspond to exon

Reprint requests to: Thomas E. Creighton, European Molecular Biology Laboratory, Meyerhofstrasse 1, D-69012 Heidelberg, Germany; e-mail: creighton@embl-heidelberg.de.

Abbreviations: 2D, two-dimensional; 3D, three-dimensional; HSQC, heteronuclear single-quantum coherence spectroscopy; NOE, nuclear Overhauser effect; NMR, nuclear magnetic resonance; NOESY, 2D NOE spectroscopy; TOCSY, total correlation spectroscopy; ¹⁵N-NOESY-HSQC, 3D ¹H NOESY ¹H-¹⁵N HSQC; ¹⁵N-TOCSY-HSQC, 3D ¹H TOCSY ¹H-¹⁵N HSQC; HNCA, HN(CO)CA, 3D ¹H-¹⁵N-¹³C α correlation spectroscopy; HNCO, HN(CA)CO, 3D ¹H-¹⁵N-¹³C α correlation spectroscopy; COCAH, 3D ¹³C α -¹H α correlation spectroscopy; (CO)N(CO)CAH, 3D ¹⁵N-¹³C α -¹H α correlation spectroscopy; TPPI, time proportional phase incrementation; PDI, protein disulfide isomerase; PFG, pulsed-field gradient; TSP, trimethylsilylpropionic acid.

3 of the PDI gene (Tsibris et al., 1989). Likewise, residues 163–211 of PDI appear to be related to residues 304–349 of the estrogen receptor. For these reasons, PDI has been concluded to be constructed of six domains, in the order *a-e-b-b'-a'-c*, where *e* is the first segment homologous to the estrogen receptor (Freedman et al., 1994). However, there are difficulties with this conclusion. Why does the second PDI segment homologous to the estrogen receptor clearly comprise part of the putative *b* domain? The first segment could also be considered part of the *a* domain. Why are the orders in the primary structures of the two similar segments in PDI and in the estrogen receptor reversed in the two proteins?

The functions of the segments of sequence *e*, *b*, *b'*, and *c* are currently unknown, but the primary sequence of PDI, the patterns of exon/intron boundaries, and some limited proteolysis studies all indicate that it is a modular multidomain protein (Freedman et al., 1994). Recent studies on the PDI-*a* domain prepared in *Escherichia coli* have demonstrated that, like homologous proteins such as glutaredoxin, thioredoxin, DsbA, and DsbC, the active site consists of one highly reactive cysteine residue with an anomalously low pK_a and a second inaccessible cysteine residue (unpubl. obs.). The PDI-*a* domain catalyzes a subset of PDI activities (Darby & Creighton, 1995); it is a good oxidant, as a result of the instability of the active site disulfide bond. This contrasts with thioredoxin, which has a very stable disulfide bond (Holmgren, 1985). The origins of the high reactivity and the instability of the PDI-*a* disulfide bond compared to that of thioredoxin is unclear. Here, the structure of the *a* domain that is homologous to thioredoxin is characterized by heteronuclear multidimensional NMR methods. This protein also contains part of the putative *e* domain, and the results are relevant to the domain structure of PDI.

Results

The PDI-*a* construct characterized here consists of residues Asp 1–Ala 120 of the human PDI sequence, plus an initiating methionine residue. These residues were chosen to include the

entire segment clearly homologous to *E. coli* thioredoxin, plus all adjoining polar residues. The last 20 residues of this construct are similar to the sequence of the estrogen receptor and are normally considered to be part of the putative *e* domain (Tsibris et al., 1989; Freedman et al., 1994). The protein was prepared in ^{15}N - and $^{15}\text{N}/^{13}\text{C}$ -labeled forms so as to utilize the most up-to-date NMR methods. Its two cysteine residues were in the disulfide form.

The PDI-*a* domain clearly displayed dispersion of the chemical shifts characteristic of a folded protein, as can be judged from ^1H - ^{15}N HSQC spectrum shown in Figure 1. A total of 118 peaks could be identified in this spectrum. Twelve peaks could be assigned to side-chain ^1H - ^{15}N correlations: W35 (N ϵ 1H), W111 (N ϵ 1H), N16 (N δ H₂), N90 (N δ H₂), N110 (N δ H₂), Q75 (N ϵ H₂), and Q76 (N ϵ H₂). The other 106 peaks (two peaks are double) represent 108 backbone NH correlations. The expected number of backbone correlations is 114, so six ^1H - ^{15}N backbone correlations were not found in the ^1H - ^{15}N HSQC spectrum (see above).

Backbone assignments

HNCA, HN(CO)CA, HNCO, and HN(CA)CO spectra were used to accomplish the sequential assignment of the backbone resonances. First, the correlations present in the HNCA experiment ($^1\text{H}(i)$ - $^{15}\text{N}(i)$ - $^{13}\text{C}\alpha(i)$ and $^1\text{H}(i)$ - $^{15}\text{N}(i)$ - $^{13}\text{C}\alpha(i-1)$) and the HN(CO)CA experiment ($^1\text{H}(i)$ - $^{15}\text{N}(i)$ - $^{13}\text{C}\alpha(i-1)$) were used to search for the residues on the N-terminal side of residue *i*. In most of the cases the $^{13}\text{C}\alpha(i-1)$ frequencies were not unique, resulting in a number of possible connecting residues. In almost all cases, these ambiguities could be solved by inspection of the correlations present in the HNCO ($^1\text{H}(i)$ - $^{15}\text{N}(i)$ - $^{13}\text{CO}(i-1)$) and HN(CA)CO ($^1\text{H}(i)$ - $^{15}\text{N}(i)$ - $^{13}\text{CO}(i)$) and $^1\text{H}(i)$ - $^{15}\text{N}(i)$ - $^{13}\text{CO}(i-1)$) experiments. Identification of glycine residues by their characteristic $^{13}\text{C}\alpha$ chemical shift and of unique spin-systems in the ^{15}N -TOCSY-HSQC spectrum enabled us to solve the remaining ambiguities and to position assigned stretches of residues in the amino-acid sequence of the protein. The six expected backbone ^1H - ^{15}N correlations not

Fig. 1. ^1H - ^{15}N HSQC spectrum of ^{15}N -enriched PDI-*a* domain recorded at pH 5.1, 300 K, and 600 MHz (^1H). Spectral widths were 8,196.7 Hz (^1H) and 2,403.8 Hz (^{15}N). Residue-specific assignments of backbone ^1H and ^{15}N frequencies are indicated, as well as the side-chain $^1\text{H}\epsilon 1$ and $^{15}\text{N}\epsilon 1$ frequencies of the two tryptophan residues. The horizontal lines connect the side-chain NH_2 frequencies of asparagine and glutamine residues.

found in the 2D ^1H - ^{15}N HSQC spectrum were also absent in the experiments described here.

Subsequently, the $\text{H}\alpha$ protons were assigned using the COCAH and (CO)N(CO)CAH experiments. The correlations that can be found in these two experiments are of the type $^{13}\text{CO}(i)$ - $^{13}\text{C}\alpha(i)$ - $^1\text{H}\alpha(i)$ and $^{15}\text{N}(i+1)$ - $^{13}\text{C}\alpha(i)$ - $^1\text{H}\alpha(i)$, respectively. The frequencies of $^{13}\text{CO}(i)$, $^{13}\text{C}\alpha(i)$ and $^{15}\text{N}(i+1)$ were found in the backbone experiments described above making unambiguous assignment of the $^1\text{H}\alpha(i)$ proton possible in almost all cases. The $^1\text{H}\alpha$ protons of a number of the 12 residues (including 6 proline residues) that showed no peaks in the ^1H - ^{15}N HSQC spectrum could be assigned by using the sequential information present in the HNCO and HN(CO)CA experiment in combination with the COCAH. The $^1\text{H}\alpha$ frequencies of five out of the six proline residues were found this way, as well as the $^1\text{H}\alpha$ frequencies of H8, W35, H38, and A101. More difficult were the assignments of D7, P34, and G37, because they were preceded by a residue also not having a peak in the ^1H - ^{15}N HSQC spectrum. G37 was tentatively assigned to two unassigned peaks in the COCAH spectrum with the same $^{13}\text{C}\alpha$ chemical shift, which is typical of glycine residues (48.0 ppm). D7 and P34 were tentatively assigned on the basis of the COCAH spectrum and observed NOE connectivities. The backbone resonance assignments are listed in Table 1 of the supplementary material in the Electronic Appendix.

Secondary structure

Characterization of the secondary structure of the protein relied on chemical shift data and NOEs. Qualitative information about the type of secondary structure present in the sequence can be obtained by determination of the "secondary" chemical shifts (Spera & Bax, 1991; Wishart et al., 1991). Positive deviations of the carbon shifts occur in helices, whereas negative deviations are found in extended structures; the opposite holds for the $^1\text{H}\alpha$ chemical shifts. The secondary shifts of the $^1\text{H}\alpha$, $^{13}\text{C}\alpha$, and ^{13}CO are shown in Figure 2.

Sequential and medium-range NOEs also provide information about the secondary structure (Wüthrich, 1986): Relatively strong $\text{H}\alpha(i)$ - $\text{NH}(i+1)$ NOEs in combination with relatively weak $\text{NH}(i)$ - $\text{NH}(i+1)$ NOEs and no $\text{H}\alpha(i)$ - $\text{NH}(i+3)$ NOEs are characteristic of extended structures, whereas relatively weak $\text{H}\alpha(i)$ - $\text{NH}(i+1)$ NOEs in combination with relatively strong $\text{NH}(i)$ - $\text{NH}(i+1)$ NOEs and the presence of $\text{H}\alpha(i)$ - $\text{NH}(i+3)$ NOEs are characteristic of helical conformations. An overview of the observed backbone NOEs is shown in Figure 3. As indicated, four helical regions could be discerned, which match approximately the corresponding helices in *E. coli* thioredoxin. Definition of the β -sheet of the protein is further supported by the observation of interstrand NOEs in the ^{15}N -NOESY-HSQC and homonuclear NOESY experiments. These are shown in Figure 4. The topology of the sheet, consisting of three strands in a parallel arrangement and three strands in an antiparallel arrangement, is the same as in *E. coli* thioredoxin (Dyson et al., 1989). The last strand (K97-Y99) in the PDI- α domain is, however, considerably shorter than in *E. coli* thioredoxin, and it is also positioned differently in the sequence.

Cis proline

A common feature observed in the structures of proteins belonging to the thioredoxin family (e.g., thioredoxin [Katti et al.,

Fig. 2. Secondary shifts for the backbone $^1\text{H}\alpha$ ($\Delta\text{H}\alpha$), $^{13}\text{C}\alpha$ ($\Delta\text{C}\alpha$), and ^{13}CO (ΔCO) resonances of the PDI- α domain. The $^1\text{H}\alpha$ and $^{13}\text{C}\alpha$ secondary chemical shifts were obtained by subtracting the "random coil" shifts given by Wishart et al. (1991). The glycine $^1\text{H}\alpha$ secondary shift was obtained by averaging the two $^1\text{H}\alpha$ shifts before subtracting the "random coil" shift. The ^{13}CO secondary shifts were obtained by subtracting the "random coil" shifts given by Howarth and Lilley (1978). Vertical lines delineate the limits of the elements of secondary structure determined here and presented in Figure 3.

1990], glutaredoxin [Sodano et al., 1991; Xia et al., 1992], DsbA [Martin et al., 1993]) is a conserved *cis*-proline residue situated between the third α -helix and the fourth β -strand. That a *cis*-proline peptide bond was also present at the corresponding position in PDI- α (P83) was apparent from an observed NOE connectivity between the $^1\text{H}\alpha$ protons of Y82 and P83, which is typical of a *cis* isomer. No NOE could be found between the $^1\text{H}\alpha$ of Y82 and the $^1\text{H}\delta$ protons of P83, which would be characteristic of a *trans* peptide bond.

Discussion

The results presented here demonstrate that the PDI- α domain of residues 1–120 of human PDI adopts a stable folded conformation very similar to that of *E. coli* thioredoxin. Very similar patterns of secondary structure were found. In addition, a proline residue that is preceded by a *cis* peptide bond in *E. coli* thioredoxin and other members of this family was also found to have a *cis* peptide bond. Very similar structures would be expected with two proteins having homologous primary structures that are 35% identical. Freedman et al. (1988) were able to build a

Fig. 3. Schematic representation of the observed sequential $\text{NH}(i)\text{-NH}(i+1)$ and $\text{H}\alpha(i)\text{-NH}(i+1)$ NOEs and the observed medium range $\text{H}\alpha(i)\text{-NH}(i+3)$ NOEs. Sequential NOEs are classified into strong, medium, or weak as indicated by the bar size. Open bars in the case of the sequential NOEs and dashed lines in the case of medium range NOEs indicate ambiguity caused by spectral overlap. Results from the amide exchange measurements are indicated in the top line. An open circle indicates that the NH proton was only observed in the $^1\text{H}\text{-}^{15}\text{N}$ spectrum recorded directly after addition of D_2O . Hatched circles indicate that there was ambiguity caused by spectral overlap in the HSQC spectrum. Elements of secondary structure deduced for PDI-*a* are indicated in the bottom line. Limits of helix α_3 are those indicated by further structural data not presented here. The absence of one NOE indicates that helix α_4 is interrupted near its middle.

realistic model of PDI-*a* based on the *E. coli* thioredoxin structure.

The folded structure of PDI-*a* confirms that PDI has a modular structure, because PDI-*a* has only 120 of the 491 residues of human PDI. Moreover, PDI-*a* unfolds reversibly at urea concentrations where similar unfolding transitions occur in intact PDI (unpubl. obs.), suggesting that the PDI-*a* domain has similar stability in isolation and when part of intact PDI. It also has some of the functional activities of PDI (Darby & Creighton, 1995).

Fig. 4. Representation of the topology of the five-stranded β -sheet. Solid arrows indicate unambiguous NOE contacts, and dashed arrows are uncertain due to spectral overlap.

Most of the 120 residues of PDI-*a* were part of the stable folded structure; only residues 1–3 and 117–120 at the termini appeared not to be fixed in the folded state. The stable structure included residues 101–116, which have been assigned to the *e* domain (Tsibris et al., 1989; Freedman et al., 1994). These residues appear to be part of the normal thioredoxin fold; they are detectably homologous to *E. coli* thioredoxin, and they adopted a largely helical conformation, as do the corresponding residues in thioredoxin. If residues 101–120 were part of another domain, they would be expected to be primarily unfolded in the absence of the remainder of that domain. Furthermore, searches in the protein sequence databases reveal many other sequences that are more homologous to the putative PDI-*e* domain than the estrogen receptor. These observations, plus the fact that the other segment of PDI said to be homologous to the estrogen receptor occurs in the *b* domain, cast doubt on the general significance of these amino acid similarities and on whether they have any structural or functional implications.

PDI-*a* and thioredoxin adopt very similar structures and both have the -Cys-Gly-Xxx-Cys- active site (Xxx = His or Pro, respectively) residues at the N-terminus of homologous α -helices. Consequently, both have the first cysteine residue sulfur atom exposed and reactive, whereas the second is buried, which has been confirmed chemically for PDI-*a* (unpubl. obs.). In contrast to these similarities, these two proteins differ dramatically in their functional properties. Thioredoxin is a potent reductant, whereas PDI is an oxidant, and their disulfide bonds differ in stability by a factor of 104 (unpubl. obs.). That of PDI-*a* is very much less stable and more similar to that of the related bacterial protein DsbA (Zapun et al., 1993; Wunderlich & Glockshuber, 1993). The instability of the disulfide bond of DsbA has been attributed to electrostatic interactions that stabilize the thiolate anion of the exposed cysteine residue, thereby stabilizing the reduced state (Nelson & Creighton, 1994). This electrostatic interaction could involve the exposed NH groups in the first turn

of the α -helix (Hol, 1985; Kortemme & Creighton, 1995). It may be significant, therefore, that these NH groups could not be observed in PDI-*a*. They are presumably undergoing rapid hydrogen exchange with the solvent, perhaps as a result of having altered effective pK_a values due to the electrostatic interactions. Such effects would be expected to be much stronger in PDI-*a* than in *E. coli* thioredoxin, which has a cysteine with a less perturbed pK_a , and the corresponding NH groups were observed by NMR in *E. coli* thioredoxin (Dyson et al., 1989). The PDI-*a* domain should be an excellent subject for investigating such effects and their functional relevance for the diverse properties of these two proteins and other members of the thioredoxin family.

Materials and methods

Materials

The cloned gene for human PDI was very kindly provided as cDNA clone S-138 by K. Kivirikko (University of Oulu, Finland). The PDI-*a* domain was produced as described elsewhere (Darby & Creighton, 1995).

Sample preparation

Construction of a pET 12a-based vector for expression of PDI-*a* has been described previously. This plasmid was transformed into BL21 (DE3) cells that also carried the plasmids pLysS, to control leakthrough expression, and pUBS520, encoding the rare *E. coli* t-RNA^{AGA/AGG} (Schenk et al., 1995), which slightly improves protein expression.

For preparation of ¹⁵N-labeled PDI-*a*, 5 L of minimal media containing 1 g/L ¹⁵NH₄Cl supplemented with minerals and co-factors (Hoffman & Spicer, 1991), plus carbanicillin (50 μ g/mL), kanamycin (50 μ g/mL), and chloramphenicol (40 μ g/mL), was inoculated from an overnight culture and grown at 37 °C until the absorbance at 600 nm reached 0.4–0.5. Isopropyl- β -D-thiogalactopyranoside was added to a concentration of 0.5 mM and growth allowed to proceed for a further 6 h, when cells were harvested. The protein was isolated from the soluble cell extract by sequential ammonium sulfate precipitation, phenyl sepharose chromatography, and anion exchange chromatography; it was converted into the oxidized form as described previously.

¹³C/¹⁵N-double labeled PDI-*a* was prepared in the same way as the ¹⁵N-labeled sample, but the ¹⁵NH₄Cl was omitted and a double labeled sugar mix (J. Ashurst et al., in prep.) was added to a final concentration of 4 g/L.

NMR spectroscopy

General

All NMR spectra were recorded at 300 K on two samples: uniformly ¹⁵N-labeled protein and uniformly ¹³C/¹⁵N-labeled protein both dissolved at 2 mM in 93% H₂O/7% D₂O (v/v), pH 5.1. Spectra were recorded on three different spectrometers: a Bruker AMX500 (EMBL), a Bruker AMX600 (EMBL), and a Varian Unity 500 (GBB). All spectrometers were equipped with triple-resonance probe heads and 3-channel interfaces. The Bruker AMX600 and the Varian Unity 500 were equipped with PFG

accessories. Off-resonance excitation, when appropriate, was accomplished by the application of phase-modulated pulses with an amplitude profile corresponding to the center lobe of a sinc function (Patt, 1992). In addition, the carrier frequency of one of the channels on the Varian Unity 500 spectrometer could be switched during the experiment without the loss of phase coherence.

Quadrature detection in the indirect dimensions was accomplished in all experiments by TPPI (Marion & Wüthrich, 1983). The ¹H carrier frequency was always positioned at the water resonance. Carrier frequencies of ¹⁵N and ¹³C are given in ppm relative to liquid NH₃ (Live et al., 1984) and TSP (Bax & Subramanian, 1986), respectively.

During acquisition, heteronuclear decoupling was accomplished by application of GARP (Shaka et al., 1985) or broadband WALTZ-16 (Shaka et al., 1983) decoupling schemes.

The spectral widths in the proton dimensions were in all cases 6,849.3 Hz for the experiments at 500 MHz and 8,192.6 Hz for the experiments at 600 MHz.

¹H-¹H homonuclear NMR

Conventional NOESY (Jeener et al., 1979) and TOCSY (Braunschweiler & Ernst, 1983) experiments were recorded on the uniformly ¹⁵N-labeled sample using the Bruker AMX600 spectrometer. In the NOESY experiments, mixing times of 100 ms and 200 ms were used. In the TOCSY experiment, coherence mixing was accomplished by applying a modified MLEV16 mixing sequence (Kadkhodaei et al., 1993) of 60 ms duration. Suppression of the ¹H₂O signal was accomplished using the WATERGATE solvent suppression scheme (Piotto et al., 1992). Maximum values for t_1 and t_2 acquisition were 31.2 ms and 124.9 ms.

¹H-¹⁵N double resonance NMR

¹H-¹⁵N double resonance NMR spectra were recorded on the uniformly ¹⁵N-labeled sample using the Bruker AMX600 spectrometer. The ¹H-¹⁵N HSQC (Bodenhausen & Ruben, 1980) spectrum (Fig. 1) was recorded using PFGs for the suppression of the intense H₂O signal and artifacts as described by Bax and Pochapsky (1992). The spectral width in ω_1 (¹⁵N) was 2,403.8 Hz, and the maximum values for t_1 and t_2 acquisition were 53.2 and 124.9 ms, respectively. The ¹⁵N carrier was placed at 118.4 ppm.

¹⁵N-NOESY-HSQC and ¹⁵N-TOCSY-HSQC experiments (Marion et al., 1989; Zuiderweg & Fesik, 1989; Fesik & Zuiderweg, 1990) were recorded on the uniformly ¹⁵N-labeled sample using the Bruker AMX-600 spectrometer. Suppression of the ¹H₂O signal was accomplished using the WATERGATE solvent suppression scheme (Piotto et al., 1992). The spectral width in ω_2 (¹⁵N) was 2,403.8 Hz. The maximum values of t_1 , t_2 , and t_3 were 24.4, 15.8, and 62.5 ms. The ¹⁵N carrier was placed at 118.4 ppm.

Two ¹⁵N-NOESY-HSQC data sets were recorded, using mixing times of 150 and 200 ms. In the ¹⁵N-TOCSY-HSQC experiment, coherence transfer was accomplished using a modified MLEV16 mixing sequence (Kadkhodaei et al., 1993) of 60-ms duration.

¹H-¹³C-¹⁵N triple resonance NMR

Constant time versions of the HNCA, HN(CO)CA, and HNCO experiments (Grzesiek & Bax, 1992) were recorded on

the uniformly $^{13}\text{C}/^{15}\text{N}$ -labeled protein sample using the Bruker AMX500 spectrometer. Water suppression was accomplished by presaturation during the relaxation delay. The ^{15}N carrier was placed at 119.9 ppm, the $\text{C}\alpha$ carrier at 56.2 ppm, and the CO carrier at 177.1 ppm. Spectral widths of 2,000.0 Hz in ω_1 (^{15}N) and 5,000.0 Hz ($^{13}\text{C}\alpha$) or 2,500.0 Hz (^{13}CO) in ω_2 were used. Maximum values for t_1 (^{15}N) and t_3 (^1H) were 16.0 and 74.8 ms, respectively, in all 3 experiments. The maximum values for t_2 (^{13}C) were 10.4 for the HNCA, 15.6 ms for the HN(CO)CA, and 25.6 ms for the HNCO.

The HN(CA)CO (Clubb et al., 1992), the COCAH (Dijkstra et al., 1994), and the (CO)N(CO)CAH (K. Dijkstra et al., in prep.) experiments were performed on the Varian UNITY 500 spectrometer. Water suppression was accomplished using PFGs as purge pulses (Bax & Pochapsky, 1992). Spectral widths were 2,000.0 Hz in ω_1 (^{15}N) and 2,500.0 Hz in ω_2 (^{13}CO) in the case of the HN(CA)CO experiment; 5,000.0 Hz in ω_1 ($^{13}\text{C}\alpha$) and 2,500.0 Hz in ω_2 (^{13}CO) in the case of the COCAH experiment; 5,000.0 Hz in ω_1 ($^{13}\text{C}\alpha$) and 2,500.0 Hz in ω_2 (^{15}N) in the case of the (CO)N(CO)CAH experiment. The ^{15}N carrier was positioned at 116.7 ppm, the ^{13}CO carrier at 177.5 ppm, and the $^{13}\text{C}\alpha$ carrier at 58.2 ppm. Maximum values for t_1 (^{15}N), t_2 (^{13}CO), and t_3 (^1H) acquisition in the case of the HN(CA)CO experiment were: 16.0, 24.0, and 74.8 ms, respectively; for t_1 ($^{13}\text{C}\alpha$), t_2 (^{13}CO), and t_3 (^1H) in the case of the COCAH experiment: 7.4, 11.8, and 74.8 ms; for t_1 ($^{13}\text{C}\alpha$), t_2 (^{15}N), and t_3 (^1H) in the case of the (CO)N(CO)CAH experiment: 6.4, 12.8, and 74.8 ms.

Hydrogen exchange experiments

Fast-exchanging amide protons were identified by dissolving the ^{15}N -labeled protein freeze-dried from H_2O into 100% D_2O . A series of ^1H - ^{15}N HSQC experiments were recorded on the Bruker AMX500 spectrometer (^{15}N carrier at 119.9 ppm; 2,000 Hz spectral width). Maximum values for t_1 (^{15}N) and t_2 (^1H) acquisition were 64.0 and 149.5 ms.

Data processing and analysis

Data were processed on a Silicon Graphics Iris Indigo using the program Snarf v0.8.9 (Frans van Hoesel, University of Groningen). Linear prediction was used to extend the number of time domain data points in t_1 and t_2 in the case of the 3D experiments. Before Fourier transformation, the resulting time domain data points were multiplied by a suitable window function (Lorentz-Gauss or sine-bell). After Fourier transformation in three dimensions, the following data sets were obtained ($\omega_3 \times \omega_2 \times \omega_1$): $1,024 \times 128 \times 128$ (^1H , ^{15}N , $^{13}\text{C}\alpha/^{13}\text{CO}$) real points for the HNCA, HN(CO)CA, HNCO, and HN(CA)CO experiments; $512 \times 512 \times 128$ (^1H , ^1H , ^{15}N) real points for the ^{15}N -NOESY-HSQC and ^{15}N -TOCSY-HSQC experiments; $1,024 \times 128 \times 128$ (^1H , $^{13}\text{C}\alpha$, $^{13}\text{CO}/^{15}\text{N}$) real points for the COCAH and (CO)N(CO)CAH experiments.

The NOESY, TOCSY, and ^1H - ^{15}N HSQC experiments were processed to give data sets of $1,024 \times 1,024$ real points.

Supplementary material in Electronic Appendix

The ^1H , ^{13}C , and ^{15}N chemical shifts of the backbone atoms of the PDI- α domain are given in Table 1.

Acknowledgments

We thank K. Kivirikko for generously supplying the cloned gene for human PDI and Elke Penka for excellent technical assistance.

References

- Bax A, Pochapsky SS. 1992. Optimized recording of heteronuclear multi-dimensional NMR spectra using pulse field gradients. *J Magn Reson* 99:638–643.
- Bax A, Subramanian S. 1986. Sensitivity-enhanced two-dimensional heteronuclear shift correlation NMR spectroscopy. *J Magn Reson* 67:565–569.
- Bodenhausen G, Ruben DJ. 1980. Natural abundance nitrogen 15 NMR by enhanced heteronuclear spectroscopy. *Chem Phys Lett* 69:185–189.
- Braunschweiler L, Ernst RR. 1983. Coherence transfer by isotropic mixing: Application to proton correlation spectroscopy. *J Magn Reson* 53:521–528.
- Clubb RT, Thanabal V, Wagner G. 1992. A constant-time three-dimensional triple-resonance pulse scheme to correlate intraresidue $^1\text{H}^{\text{N}}$, ^{15}N and $^{13}\text{C}'$ chemical shifts in ^{15}N - ^{13}C -labeled proteins. *J Magn Reson* 97:213–217.
- Darby NJ, Creighton TE. 1995. Functional properties of the individual thioredoxin-like domains of protein disulfide isomerase. *Biochemistry* 34:11725–11735.
- Dijkstra K, Kroon GJA, van Nuland NAJ, Scheek RM. 1994. The COCAH experiment to correlate intra-residue carbonyl, $\text{C}\alpha$ and $\text{H}\alpha$ resonances in proteins. *J Magn Reson* 97:213–217.
- Dyson HJ, Holmgren A, Wright PE. 1989. Assignment of the proton NMR spectrum of reduced and oxidized thioredoxin: Sequence-specific assignments, secondary structure and global fold. *Biochemistry* 28:7074–7087.
- Edman JC, Ellis L, Blacher RW, Roth RA, Rutter WJ. 1985. Sequence of protein disulfide isomerase and implications of its relationship to thioredoxin. *Nature* 317:267–270.
- Farquhar R, Honey N, Murant SJ, Bossier P, Schultz L, Montgomery D, Ellis RW, Freedman RB, Tuite MF. 1991. Protein disulfide isomerase is essential for viability in *Saccharomyces cerevisiae*. *Gene* 108:81–89.
- Fesik SW, Zuiderweg ERP. 1990. Heteronuclear three-dimensional NMR spectroscopy of isotopically labeled macromolecules. *Q Rev Biophys* 23:97–131.
- Freedman RB. 1992. Protein folding in the cell. In: Creighton TE, ed. *Protein folding*. New York: W.H. Freeman & Co. pp 455–539.
- Freedman RB, Hawkins HC, Murant SJ, Reid L. 1988. Protein disulfide isomerase: A homologue of thioredoxin implicated in the biosynthesis of secretory proteins. *Biochem Soc Trans* 16:96–99.
- Freedman RB, Hirst TR, Tuite MF. 1994. Protein disulfide isomerase: Building bridges in protein folding. *Trends Biochem Sci* 19:331–336.
- Grzesiek S, Bax A. 1992. Improved 3D triple-resonance NMR techniques applied to a 31 kDa protein. *J Magn Reson* 96:432–440.
- Hoffman DW, Spicer LD. 1991. Isotopic labeling of specific amino acid types as an aid to NMR spectrum assignment of the methionine repressor protein. In: Villafranca JJ, ed. *Techniques in protein chemistry II*. San Diego: Academic Press. pp 409–419.
- Hol WGJ. 1985. The role of the α -helix dipole in protein function and structure. *Prog Biophys Mol Biol* 45:149–195.
- Holmgren A. 1985. Thioredoxin. *Annu Rev Biochem* 54:237–271.
- Howarth OW, Lilley DMJ. 1978. Carbon-13-NMR of peptides and proteins. *Progr NMR Spectrosc* 12:1–40.
- Jeener J, Meier BH, Bachman P, Ernst RR. 1979. Investigation of exchange processes by two-dimensional NMR spectroscopy. *J Chem Phys* 71:4546–4553.
- Kadkhodaei M, Hwang TL, Tang J, Shaka AJ. 1993. A simple windowless mixing sequence to suppress cross relaxation in TOCSY experiments. *J Magn Reson A* 105:104–107.
- Katti SK, Lemaster DM, Eklund H. 1990. Crystal structure of thioredoxin from *Escherichia coli* at 1.68 Å resolution. *J Mol Biol* 212:167–184.
- Kortemme T, Creighton TE. 1995. Ionisation of cysteine residues at the termini of model α -helical peptides. Relevance to unusual thiol pK_a values in proteins of the thioredoxin family. *J Mol Biol*. Forthcoming.
- Krause G, Lundström J, Lopez Barea J, Pueyo de la Cuesta C, Holmgren A. 1991. Mimicking the active site of protein disulfide isomerase by substitution of proline 34 in *Escherichia coli* thioredoxin. *J Biol Chem* 266:9494–9500.
- Landel CC, Kushner PJ, Greene GL. 1994. The interaction of human estrogen receptor with DNA is modulated by receptor-associated proteins. *Mol Endocrinol* 8:1407–1419.
- Live DH, Davis DG, Agosta WC, Cowburn D. 1984. Long range hydrogen

- bond mediated effects in peptides: ^{15}N NMR study of gramicidin S in water and organic solvents. *J Am Chem Soc* 106:1939–1941.
- Lundström J, Krause G, Holmgren A. 1992. A Pro to His mutation in the active site of thioredoxin increases its disulfide isomerase activity 10-fold. *J Biol Chem* 267:9047–9052.
- Lyles MM, Gilbert HF. 1991a. Catalysis of the oxidative folding of ribonuclease A by protein disulfide isomerase: Dependence of the rate on the composition of the redox buffer. *Biochemistry* 30:613–619.
- Lyles MM, Gilbert HF. 1991b. Catalysis of the oxidative folding of ribonuclease A by protein disulfide isomerase: Pre-steady-state kinetics and the utilization of the oxidizing equivalents of the isomerase. *Biochemistry* 30:619–625.
- Lyles MM, Gilbert HF. 1994. Mutations in the thioredoxin sites of protein disulfide isomerase reveal functional nonequivalence of N- and C-terminal domains. *J Biol Chem* 269:30946–30952.
- Marion D, Driscoll PC, Kay LE, Wingfield PT, Bax A, Gronenborn AM, Clore GM. 1989. Overcoming the overlap problem in assignment of ^1H -NMR spectra of larger proteins by use of three-dimensional heteronuclear ^1H - ^{15}N Hartmann-Hahn multiple quantum coherence and nuclear-Overhauser multiple quantum coherence spectroscopy: Application to interleukin 1b. *Biochemistry* 28:6150–6156.
- Marion D, Wüthrich K. 1983. Application of phase sensitive two-dimensional spectroscopy (COSY) for measurements of ^1H - ^1H spin-spin coupling constants in proteins. *Biochem Biophys Res Commun* 113:967–974.
- Martin JL, Bardwell JCA, Kuriyan J. 1993. Crystal structure of the DsbA protein required for disulphide bond formation *in vivo*. *Nature* 365:464–468.
- Nelson JW, Creighton TE. 1994. Reactivity and ionization of the active site cysteine residues of DsbA, a protein required for disulfide bond formation *in vivo*. *Biochemistry* 33:5974–5983.
- Noiva R, Lennarz WJ. 1992. Protein disulfide isomerase. A multifunctional protein resident in the lumen of the endoplasmic reticulum. *J Biol Chem* 267:3553–3556.
- Patt SL. 1992. Single- and multiple-frequency-shifted laminar pulses. *J Magn Reson* 96:94–102.
- Pihlajaniemi T, Helaakoski T, Tasanen K, Myllylä R, Huhtala ML, Koivu J, Kivirikko KI. 1987. Molecular cloning of the β -subunit of human prolyl 4-hydroxylase. This subunit and protein disulphide isomerase are products of the same gene. *EMBO J* 6:643–649.
- Piotto M, Saudek V, Sklenar V. 1992. Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *J Biomol NMR* 2: 661–665.
- Schenk PM, Baumann S, Mattes R, Steinbiss HH. 1995. Improved high-level expression system for eukaryotic genes in *Escherichia coli* using T7 RNA polymerase and rare $^{A^{\text{t}}}\text{tRNAs}$. *Biotechniques* 19:196–200.
- Scherens B, Dubois E, Messenguy F. 1991. Determination of the sequence of the yeast YCL313 gene localized on chromosome III. Homology with the protein disulfide isomerase (PDI gene product) of other organisms. *Yeast* 7:185–193.
- Shaka AJ, Barker PB, Freeman R. 1985. Computer optimized decoupling scheme for wideband applications and low level operation. *J Magn Reson* 64:547–552.
- Shaka AJ, Keeler J, Freeman R. 1983. Evaluation of a new broadband decoupling sequence: WALTZ-16. *J Magn Reson* 53:313–340.
- Sodano P, Xia TH, Bushweller JH, Björnberg O, Holmgren A, Billeter M, Wüthrich K. 1991. Sequence-specific ^1H NMR assignments and determination of the three-dimensional structure of reduced *E. coli* glutaredoxin. *J Mol Biol* 221:1311–1324.
- Spera S, Bax A. 1991. Empirical correlation between protein backbone conformation and $\text{C}\alpha$ and $\text{C}\beta$ ^{13}C nuclear magnetic resonance chemical shifts. *J Am Chem Soc* 113:5490–5492.
- Tsibris JCM, Hunt LT, Ballejo G, Barker WC, Toney LJ, Spellacy WN. 1989. Selective inhibition of protein disulfide isomerase by estrogens. *J Biol Chem* 264:13967–13970.
- Vuori K, Myllylä R, Pihlajaniemi T, Kivirikko KI. 1992. Expression and site-directed mutagenesis of human protein disulfide isomerase in *Escherichia coli*. *J Biol Chem* 267:7211–7214.
- Wetterau JR, Aggerbeck LP, Laplaud PM, McLean LR. 1991. Structural properties of the microsomal triglyceride-transfer protein complex. *Biochemistry* 30:4406–4412.
- Wetterau JR, Combs KA, Spinner SN, Joiner BJ. 1990. Protein disulfide isomerase is a component of the microsomal triglyceride transfer protein complex. *J Biol Chem* 265:9800–9807.
- Wishart DS, Sykes BD, Richards FM. 1991. Relationship between nuclear magnetic resonance chemical shift and protein secondary structure. *J Mol Biol* 222:311–333.
- Wunderlich M, Glockshuber R. 1993. Redox properties of protein disulfide isomerase (DsbA) from *Escherichia coli*. *Protein Sci* 2:717–726.
- Wüthrich K. 1986. *NMR of proteins and nucleic acids*. New York: Wiley.
- Xia TH, Bushweller JH, Sodano P, Billeter M, Björnberg O, Holmgren A, Wüthrich K. 1992. NMR structure of oxidized *Escherichia coli* glutaredoxin: Comparison with reduced *E. coli* glutaredoxin and functionally related proteins. *Protein Sci* 1:310–321.
- Zapun A, Bardwell JCA, Creighton TE. 1993. The reactive and destabilizing disulfide bond of DsbA, a protein required for protein disulfide bond formation *in vivo*. *Biochemistry* 32:5083–5092.
- Zuiderweg ERP, Fesik SW. 1989. Heteronuclear three-dimensional NMR spectroscopy of the inflammatory protein C5a. *Biochemistry* 28:2387–2391.