Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Purification, crystallization and preliminary crystallographic studies on the N-terminal fragment of human protein disulfide isomerase

Xin-Quan Wang, Lu-Lu Gui, Yong Dai, Chih-Chen Wang, Wen-Rui Chang and Dong-Cai Liang*

National Laboratory of Biomacromolecules, Institute of Biophysics, Academia Sinica, Beijing 100101, People's Republic of China

Correspondence e-mail: dcliang@sun5.ibp.ac.cn

Received 6 May 1999
Accepted 14 September 1999

A fragment of human protein disulfide isomerase composed of the thioredoxin-like *a* and *b* domains (*ab*) has been expressed in *Escherichia coli* as a fusion protein with glutathione-S-transferase and purified after thrombin cleavage. Two forms of *ab* crystal were obtained with polyethylene glycol as precipitant and different additives at pH 7.5. The space group of form I is $P4_12_12$ or $P4_32_12$, with unit-cell dimensions a = 81.5, c = 259.7 Å. The space group of form II is $P4_122$ or $P4_322$, with unit-cell dimensions a = 82.7, c = 86.5 Å.

1. Introduction

Protein disulfide isomerase (PDI, E.C. 5.3.4.1), a multifunctional protein in the lumen of endoplasmic reticulum and one of the two foldases, catalyzes the formation and rearrangement of disulfide bonds (Freedman *et al.*, 1994). Experimental evidence also shows that PDI has chaperone activity (Cai *et al.*, 1994; Puig & Gilbert, 1994; Yao *et al.*, 1997). In addition, PDI is a component of prolyl-4-hydroxylase (Pihlajaniemi *et al.*, 1987) and the microsomal triglyceride transfer protein (Wetterau *et al.*, 1990) and is essential for their structural stability and biological function (John *et al.*, 1993; Wetterau *et al.*, 1991).

According to the cDNA-deduced primary structure (Edman *et al.*, 1985) and results from protein-engineering studies (Darby & Creighton, 1995; Darby *et al.*, 1996), PDI has been deduced to be composed of four domains in the order –*a*–*b*–*b*′–*a*′– and a stretch of acidic residues, labelled *c*, at the C-terminus (Edman *et al.*, 1985). Domains *a* and *a*′ are homologous to each other (47% identity) and also to thioredoxin (Edman *et al.*, 1985). Both the *a* and *a*′ domains contain the active-site motif -Cys-Gly-His-Cys-, which participates in thiol–disulfide exchange during the catalysis process (Hawkins & Freedman, 1991; Lundstrom & Holmgren, 1993). There is 28% sequence identity between the *b* and *b*′ domains (Edman *et al.*, 1985), which do not have the active-site motif -Cys-Gly-His-Cys-. The functions of *b* and *b*′ are not clearly known. The *a* and *a*′ domains have been individually expressed in *Escherichia coli* as stable and soluble proteins which have 14 and 9% of the activity of the intact PDI molecule in catalyzing the oxidative formation of disulfide bonds (Darby & Creighton, 1995). Neither domain has substantial activity in catalyzing disulfide-bond isomerization. Addition of *b* and *b*′ domains to the isolated *a* domain gradually increases the catalytic efficiency of PDI, especially in the catalysis of disulfide-bond rearrangements in folded substrates (Darby *et al.*, 1998). It has been reported that different domains of PDI contribute to the substrate-binding site and that the acidic C-terminal extension is not critical for the enzyme-subunit function (Klappa *et al.*, 1998; Koivunen *et al.*, 1999). These results show that other structural elements, such as *b* and *b*′ domains, are necessary for the full spectrum of PDI activities.

Although structures of individual *a* and *b* domains of human PDI in solution have been determined by NMR techniques (Kemmink *et al.*, 1996, 1997), the crystal structures of human PDI and its multi-domain constructs have not yet been reported. In this paper, we report the purification and crystallization of the N-terminal fragment of human PDI consisting of *a* and *b* domains.

2. Experimental procedure

2.1. Construction of the expression plasmid

The plasmid pBR322-PDI, which contains the full length of the human PDI cDNA, was a generous gift from Professor K. Kivirikko, University of Oulu, Finland. The primer I (5′-CG**GGATCC**GACGCCCCCGAGGAG-3′) was designed to hybridize with the first 15 nucleotides at the 5′-terminus of the human PDI cDNA sequence and contains a *Bam*HI site (in bold) just before the sequence. The primer II (5′-CG**GAATTC**ACTCTTGGG-CAAG-3′) with a *Eco*RI site (in bold) hybridizes with the sequence between base pairs 729 and 741 of the human PDI cDNA. The double-stranded DNA fragment coding for the sequence Asp1–Ser247 of human PDI was generated by the polymerase chain reaction with the above two primers and the pBR322-PDI plasmid as a template. The PCR

© 1999 International Union of Crystallography
Printed in Denmark – all rights reserved

product was then ligated into expression vector pGEX4T1 (Pharmacia) digested with *Bam*HI and *Eco*RI in-frame with the glutathione S-transferase (GST) fusion codons to construct the expression plasmid pGEX4T1-ab. The foreign DNA sequences cloned into the expression plasmid were verified by nucleotide sequencing.

2.2. Expression, purification and activity assay

Protein production was carried out in *E. coli* strain BL21 containing the pGEX4T1-ab plasmid. Cells grown overnight at 310 K in LB medium containing 50 µg ml^{-1} ampicillin were diluted 50-fold and grown at 301 K to an absorbance of 0.6 at 600 nm. Gene expression was induced by addition of IPTG to a final concentration of 0.1 m*M* followed by additional growth for 4 h.

The cell pellet from 1 l culture was suspended in 50 ml PBS (140 m*M* NaCl, 2.7 m*M* KCl, 10.1 m*M* Na_2HPO_4, 1.8 m*M* KH_2PO_4, pH 7.3) and lysed by sonication at 277 K. Triton X-100 was added to a final concentration of 1%(*v*/*v*) with gentle stirring for 30 min at room temperature and cellular debris was removed by centrifugation. The supernatant was loaded onto a glutathione Sepharose 4B column (Pharmacia, 10 ml) equilibrated with PBS containing 1%(*v*/*v*) Triton X-100. After extensive washing with 100 ml PBS, the bound fusion protein on the matrix was eluted with 10 m*M* reduced glutathione in 50 m*M* Tris–HCl (pH 8.0). Following the removal of glutathione by extensive dialysis against 50 m*M* Tris–HCl (pH 8.5) and cleavage of the fusion protein at 277 K for 48 h with thrombin, the eluate was applied to the glutathione Sepharose 4B column again and the flow-through containing *ab* but devoid of GST was applied to an FPLC Superdex 75HR gel-filtration column (Pharmacia) and eluted with 50 m*M* phosphate buffer (pH 7.0) containing 0.15 *M* NaCl. Fractions of *ab* identified with 12% sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) were pooled, thoroughly dialyzed against 50 m*M* NH_4HCO_3, lyophilized and stored at 253 K.

The disulfide isomerase and thiol-protein oxidoreductase (TPOR) activities were assayed using the methods of Lambert & Freedman (1983).

Figure 1
Photograph of a form I crystal of maximum dimensions 0.3 × 0.05 × 0.05 mm.

Figure 2
Photograph of a form II crystal of dimensions 0.6 × 0.1 × 0.1 mm.

2.3. Crystallization

Crystallization experiments were performed using the hanging-drop vapour-diffusion method. Crystallization conditions were screened by sparse-matrix sampling (Jancarik & Kim, 1991). Thin needle-shaped crystals were obtained with polyethylene glycol as precipitant. With the optimized crystallization conditions, form I crystals were obtained by mixing 3 µl protein solution (15 mg ml^{-1}, 25 m*M* HEPES pH 7.5) with 3 µl reservoir solution containing 20–23%(*w*/*v*) PEG 5000 MME, 200 m*M* NH_4Ac, 100 m*M* HEPES buffer (pH 7.5) and 5%(*v*/*v*) glycerol. The crystals grew to dimensions 0.3 × 0.05 × 0.05 mm after two weeks at 293 K (Fig. 1). When 5%(*v*/*v*) glycerol was replaced by 5%(*v*/*v*) 1,6-hexanediol, form II crystals with dimensions 0.6 × 0.1 × 0.1 mm were obtained by mixing 4 µl protein solution (15 mg ml^{-1}, 25 m*M* HEPES, pH 7.5) with 2 µl reservoir solution (Fig. 2).

2.4. X-ray data collection

Diffraction data were collected from a form I crystal with a Weissenberg camera (Sakabe, 1991) installed on beamline 16A (λ = 1.00 Å) at the Photon Factory (Tsukuba, Japan). The data set includes 36 IP frames with a 2.5° oscillation angle per frame. Diffraction data from a form II crystal were recorded over 280 1° oscillation images at room temperature using a MAR Research image-plate system. Data were processed with the *HKL* suite (Otwinowski & Minor, 1997).

3. Results

The purified *ab* fragment is composed of 261 residues, including residues 1–247 of the PDI molecule (491 residues) with an additional Gly and Ser at the N-terminus and Glu, Phe, Pro, Gly, Arg, Leu, Glu, Arg, Pro, His, Arg and Asp at the C-terminus. The extra two residues at the N-terminus are produced by thrombin cleavage in the GST–*ab*, while the extra 12 residues at the C-terminus are the result of using the stop codon in the expression vector. Fragment *ab* has 17.3 ± 0.5% (n = 4) of the isomerase activity and 12.9 ± 1.6% (n = 4) of the TPOR activity of the intact PDI molecule.

Form I crystals belong to the tetragonal system with unit-cell dimensions a = 81.5, c = 259.7 Å. Based on the systematic absences observed, the space group was determined to be $P4_12_12$ or $P4_32_12$. Given the molecular weight of 28 kDa, the unit cell has a V_m of 3.8 Å^3 Da^{-1} (Matthews, 1968) for two *ab* molecules per asymmetric unit, a V_m of 2.6 Å^3 Da^{-1} for three *ab* molecules per asymmetric unit or a V_m of 1.9 Å^3 Da^{-1} for four *ab* molecules per asymmetric unit. The final data set [$I > 2\sigma(I)$] has a merging R factor for symmetry-related reflections of 0.09 at 3.5 Å resolution. The completeness of the data (7695 unique reflections) is 73% for the resolution range 30–3.5 Å and 30% in the 3.63–3.50 Å shell.

Form II crystals also belong to the tetragonal system, but the unit-cell dimensions are a = 82.7, c = 86.5 Å. Systematic absences $l \neq 4n$ were observed along the 00l direction and there were no systematic absences $h = 2n + 1$ along the h00 direction. Inspection of the Laue symmetry of the reciprocal space showed that the space group was $P4_122$ or $P4_322$. There is one *ab* molecule in the asymmetric unit, with a V_m of 2.6 Å^3 Da^{-1}, corresponding to a solvent content of 53%. The final data set, consisting of 3840 unique reflections with $I > 2\sigma(I)$, is 67% complete in the 30–3.0 Å resolution range and 30% in the resolution shell 3.11–3.00 Å. The merging R factor of symmetry-related reflections of this data set is 0.12.

Compared with the form I crystals, form II crystals are more suitable for structure determination, having one molecule in the asymmetric unit. Attempts to solve the structure by the molecular-replacement method with the solution structure of

domain *a* as a search model have been unsuccessful. The structure determination of the *ab* fragment by MIR method is in progress.

References

Cai, H., Wang, C. C. & Tsou, C. L. (1994). *J. Biol. Chem.* **269**, 24550–24552.

Darby, N. J. & Creighton, T. E. (1995). *Biochemistry*, **34**, 11725–11735.

Darby, N. J., Kemmink, J. & Creighton, T. E. (1996). *Biochemistry*, **35**, 10517–10528.

Darby, N. J., Penka, E. & Vincentelli, R. (1998). *J. Mol. Biol.* **276**, 239–247.

Edman, W. J., Ellis, L., Blacher, R. W., Roth, R. A. & Rutter, W. J. (1985). *Nature (London)*, **317**, 267–270.

Freedman, R. B., Hirst, T. R. & Tuite, M. F. (1994). *Trends Biochem. Sci.* **19**, 331–336.

Hawkins, H. C. & Freedman, R. B. (1991). *Biochem. J.* **275**, 335–339.

Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.

John, D. C. A., Grant, M. E. & Bulleid, N. J. (1993). *EMBO J.* **12**, 1587–1595.

Kemmink, J., Darby, N. J., Dijstra, K., Nilges, M. & Creighton, T. E. (1996). *Biochemistry*, **35**, 7684–7691.

Kemmink, J., Darby, N. J., Dijstra, K., Nilges, M. & Creighton, T. E. (1997). *Curr. Biol.* **7**, 239–245.

Klappa, P., Ruddock, L. W., Darby, N. J. & Freedman, R. B. (1998). *EMBO J.* **17**, 927–935.

Koivunen, P., Pirneskoski, A., Karvonen, P., Ljung, J., Hilaakoski, T., Notbohm, H. & Kivirikko, K. I. (1999). *EMBO J.* **18**, 65–74.

Lambert, N. & Freedman, R. B. (1983). *Biochem. J.* **213**, 225–234.

Lundstrom, J. & Holmgren, A. (1993). *Biochemistry*, **32**, 6649–6655.

Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **273**, 307–326.

Pihlajaniemi, T., Helaakoshi, T., Tasanen, K., Mgllyla, R., Huhtala, M. L., Koivu, J. & Kivirriko, K. I. (1987). *EMBO J.* **6**, 643–649.

Puig, A. & Gilbert, H. (1994). *J. Biol. Chem.* **269**, 7764–7771.

Sakabe, N. (1991). *Nucl. Instrum. Methods Phys. Res. A*, **303**, 448–463.

Wetterau, J. R., Combs, K. A., McLean, L. R., Spinner, S. N. & Aggerbeck, L. W. (1991). *Biochemistry*, **30**, 9728–9735.

Wetterau, J. R., Combs, K. A., Spinner, S. N. & Joiner, B. J. (1990). *J. Biol. Chem.* **265**, 9800–9807.

Yao, Y., Zhou, Y. C. & Wang, C. C. (1997). *EMBO J.* **16**, 651–658.