



OPEN

Structure of the substrate-binding b' domain of the Protein disulfide isomerase-like protein of the testis

Sara Bastos-Aristizabal, Guennadi Kozlov & Kalle Gehring

Department of Biochemistry, Groupe de recherche axé sur la structure des protéines, McGill University, 3649 Promenade Sir William Osler, Montréal, Québec, Canada H3G 0B1.

SUBJECT AREAS:
CHAPERONES
X-RAY CRYSTALLOGRAPHY
SOLUTION-STATE NMR
ENDOPLASMIC RETICULUMReceived
16 December 2013Accepted
10 March 2014Published
25 March 2014Correspondence and
requests for materials
should be addressed to
K.G. (kalle.gehring@
mcgill.ca)

Protein Disulfide Isomerase-Like protein of the Testis (PDILT) is a testis-specific member of the PDI family. PDILT displays similar domain architecture to PDIA1, the founding member of this protein family, but lacks catalytic cysteines needed for oxidoreduction reactions. This suggests special importance of chaperone activity of PDILT, but how it recognizes misfolded protein substrates is unknown. Here, we report the high-resolution crystal structure of the b' domain of human PDILT. The structure reveals a conserved hydrophobic pocket, which is likely a principal substrate-binding site in PDILT. In the crystal, this pocket is occupied by side chains of tyrosine and tryptophan residues from another PDILT molecule, suggesting a preference for binding exposed aromatic residues in protein substrates. The lack of interaction of the b' domain with the P-domains of calreticulin-3 and calmeglin hints at a novel way of interaction between testis-specific lectin chaperones and PDILT. Further studies of this recently discovered PDI member would help to understand the important role that PDILT plays in the differentiation and maturation of spermatozooids.

Protein disulfide isomerase-like protein of the testis (PDILT) is a member of the Protein Disulfide Isomerase (PDI) family, a family responsible for catalyzing formation of native disulfide bonds in the endoplasmic reticulum (ER). This multifunctional and diverse family consists of approximately 20 proteins, each containing at least one thioredoxin-like domain¹. PDIA1, the founding member and the most studied protein in the PDI family, displays the **abb'xa'** domain architecture, where **a** and **a'** are catalytic domains, **b** and **b'** are non-catalytic, and **x** denotes a conserved linker (Fig. 1a). PDILT is a 67-kDa protein that resides exclusively in the ER of male germ cells. Only discovered a few years ago, PDILT is distinguished from other PDI members for its tissue specific distribution and the absence of typical catalytic motifs². Full-length PDILT shares 27% of sequence identity with PDIA1, and the **a'** domain is the most strongly conserved between these proteins with 48% identity². However, PDILT differs from PDIA1 and other PDIs in several intriguing aspects.

The most striking difference between PDIA1 and PDILT lies in the **a** and **a'** domains. Instead of the classical redox-active CXXC motif, PDILT lacks the catalytic cysteines needed for oxidoreduction and instead possesses SKQS and SKKC motifs in the **a** and **a'** domains, respectively (Fig. 1a). The SXXC motif, also present in TMX2, is highly conserved in PDILT among different species. This motif is also found strictly conserved among bacterial thioredoxin proteins, but its functional implications remain unknown³. The absence of catalytic cysteines suggests that PDILT primarily functions as a chaperone, one of the principal functions of PDIA1. PDIA1 directly binds to unfolded proteins via its **b'** domain⁴.

PDILT does not display oxidoreductase activity *in vitro* and interacts with the Δ -somatostatin peptide and the BPTI protein in non-native conformation, so it is likely that PDILT operates as a folding assistant for spermatid-specific proteins⁵. PDILT is highly glycosylated, is the only PDI-family member under developmental control⁶⁻⁸ and contains a particularly long C-terminal extension of 90 residues with no known homology to other proteins².

Recent studies have yielded new hints about the role that PDILT plays in the testis. Interactions of PDILT with binding partners, such as lectin chaperones calreticulin-3 (CRT3) and calmeglin (CMG), also specifically expressed in testis, suggest that it plays a role in specific complexes that catalyze the folding of spermatogenesis-specific proteins in the late stages of maturation and differentiation of the spermatids^{5,9}. CMG and CRT3, along with PDILT, are only expressed in the most mature germ cells and their absence causes infertility⁹. Additionally, PDILT has been reported to form a chaperone complex with ADAM3 (a disintegrin and metallo-peptidase domain 3). This protein is important in fertility and requires the assistance of at least seven other genes

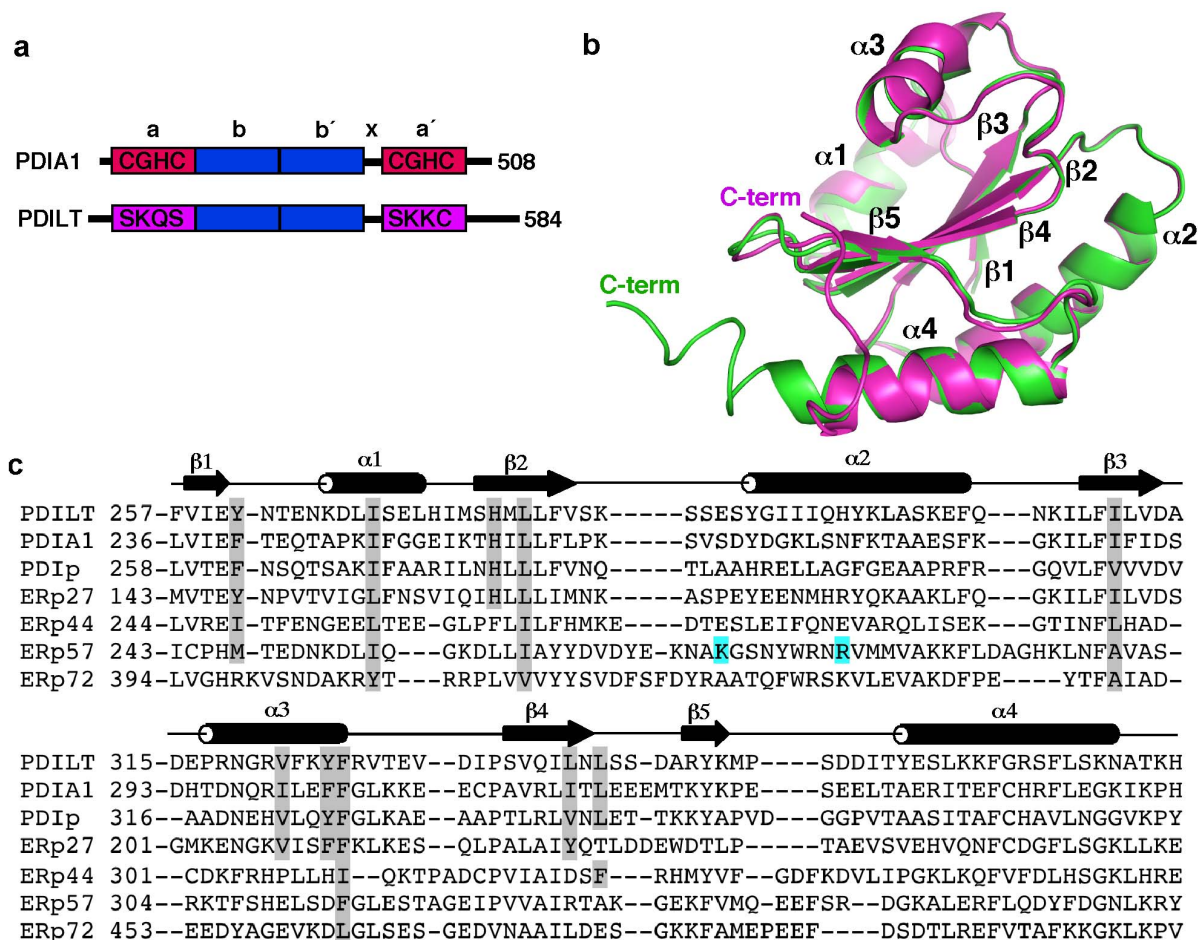


Figure 1 | Crystal structure of the **b'** domain of PDILT. (a) Domain organization of human PDILT and PDIA1. Catalytic motifs are shown in **a** and **a'** thioredoxin-like domains. (b) Structure of PDILT **b'** domain displays a typical thioredoxin-fold with a central five-stranded β -sheet ($\beta 1$ – $\beta 5$) and four flanking α -helices ($\alpha 1$ – $\alpha 4$). An overlay of chains A (green) and C (magenta) shows divergence in the C-terminal region (C-term) of the construct. (c) Structure-based sequence alignment of the **b'** domains of several members of PDI family. Amino acids forming the hydrophobic pocket in the PDILT structure and conserved between aligned proteins are highlighted in gray. Residues of ERp57 involved in the interaction with calreticulin (CRT) and calnexin (CNX) are highlighted in cyan. Secondary structure corresponds to the PDILT structure.

in order to be delivered properly onto the sperm surface^{9–11}. This suggests that PDILT plays a role in the folding pathway of certain and specific spermatogenesis-proteins and enzymes required for the proper functioning of the sperm. PDILT possesses the **b'** domain, which is associated with binding of substrates and protein partners in the PDI family members¹², but the molecular details of PDILT interactions with protein substrates remain elusive.

Here, we report the high-resolution crystal structure of the non-catalytic **b'** domain of human PDILT determined at 2.0 Å. The structure identifies a hydrophobic pocket that is conserved in a number of PDI family members and suggests a preference for binding aromatic residues in protein substrates.

Results

Structure of the **b' domain.** In order to better understand how PDILT interacts with its substrates, we obtained crystals of the **b'** domain of human PDILT (residues 258–386) that diffracted to 2.0 Å using synchrotron radiation. The three-dimensional structure was solved by molecular replacement using **b'** domain of human PDIA1 (PDB accession number 3BJ5), which shares 32% identity with PDILT **b'** domain sequence, as a search model. A summary of data collection and refinement statistics is given in Table 1.

The structure of the PDILT **b'** domain displays a thioredoxin-like fold comprising a five-stranded β -sheet with the $\beta 1\beta 3\beta 2\beta 4\beta 5$

topology, flanked by four helices (Fig. 1b). All strands are parallel, except strand $\beta 4$, which is antiparallel. There are four molecules (chains A–D) in the asymmetric unit, which are very similar to each other, as all chains are superimposed using backbone $C\alpha$ atoms with an RMSD oscillating from 0.1 to 0.8 Å. However, there is a feature that divides them into two groups according to the position of the x-linker (residues 369–386 of the construct). Chains A and B have the x-linker pointing away from the thioredoxin domain, while the x-linker folds towards the β -sheet in chains C and D (Fig. 1b). The x-linker residues Q375–P381 were not observed in the electron density map of chains C and D likely due to its intrinsic mobility.

The conserved hydrophobic pocket in PDILT is primed for binding hydrophobic residues.

Analysis of the structure reveals a hydrophobic pocket between helices $\alpha 1$ and $\alpha 3$, the location of the principal substrate-binding pocket in PDIA1 (Fig. 2). Interestingly, in the crystal this pocket is occupied by two aromatic residues (Y383 and W384) from the x-linker of another PDILT molecule (Fig. 2a and Supplemental Fig. S1). These residues most likely mimic the hydrophobic residues exposed in misfolded substrates. Specifically, the side chain of Y383 makes hydrophobic interactions with the side chains of V322, F326, L279 and L339 of the **b'** domain, while the side chain of W384 interacts with the side chains of L341 and I269 of the **b'** domain of PDILT (Fig. 2b). Also, polar interactions help stabilize the binding. In particular, the hydroxyl group of Y383 forms a hydrogen



Table 1 | Data collection and refinement statistics

Data Collection	
Crystal	PDILT b' domain
X-ray wavelength (Å)	0.977
Space group	P2 ₁
Unit cell dimensions (Å)	
a, b, c (Å)	60.98, 119.74, 33.29
α, β, γ (°)	90.00, 90.90, 90.00
Resolution range (Å)	50.0–2.00 (2.03–2.00) ¹
Unique reflections	30080
Redundancy	3.2 (2.8)
Data completeness (%)	99.2 (96.4)
R _{merge} (%)	0.043 (0.211)
I/σ(I)	24.7 (4.2)
Refinement statistics	
Resolution range (Å)	60.97–2.00
R _{work} /R _{free}	0.195/0.238
No. of non-H atoms	
Protein	4082
Water	218
Formate	12
Average B-factors (Å ²)	
Protein	15.3
Water	32.1
Formate	24.1
r.m.s. deviations	
Bond lengths (Å)	0.007
Bond angles (°)	0.97
Ramachandran plot statistics (%) ²	
Most favored regions	91.4
Additional allowed regions	8.6
Disallowed regions	0

¹Highest resolution shell is shown in parentheses.
²Ramachandran statistics were computed using PROCHECK³⁴.

bond with the imidazole ring of H277. A hydrogen bond is also observed between the backbone carbonyl group of W384 and the side chain of K266 in the α1-helix of the b' domain. Compared with W384, Y383 goes deeper into the hydrophobic pocket and likely plays a more important role in the binding.

The structural alignment with the b' domain of PDIA1 using Dali database¹³ gave a Z-score of 17.6 and an RMSD of 1.4 Å over 113 residues, showing that the structures are very similar (Fig. 2c). The major difference is in the orientation of the helix α1, which creates a larger and deeper hydrophobic pocket in PDILT. Also, the overall charge surrounding the hydrophobic pocket is negative for PDIA1, while this area is very positively charged in PDILT. Additionally, K266, S270, L339 and L341 in the substrate-binding pocket of PDILT are not conserved between PDI family members (Fig. 1c). These differences may be the key determinants for PDILT specificity towards *in vivo* substrates. The size of the hydrophobic pocket in the PDILT structure is more similar to that in the C-terminal domain of ERp27 in the polyethylene glycol-bound conformation with a Z-score of 18.6 and RMSD of 1.7 Å over 111 residues (Fig. 2d)¹⁴. Overall, the structure-based sequence alignment shows a high conservation of residues forming the hydrophobic pocket in PDILT, PDIA1, PDIp and ERp27, explaining why these PDIs are able to directly interact with unfolded proteins (Fig. 1c). Interestingly, H277 of PDILT is positioned at a base of the hydrophobic pocket and is strictly conserved for these PDIs. Our structure demonstrates that the side chain of this histidine forms a hydrogen bond with the hydroxyl group of interacting tyrosine residue and could be employed similarly by other PDIs. The corresponding histidine H256 of PDIA1 was predicted to hydrogen bond with 3-hydroxyl group of 17β-estradiol¹⁵.

Small amphipathic peptides Δ-somatostatin (AGCKNFFWK-TFTSC) and mastoparan (INLKALAALAKKIL) have been used in several studies of the PDI family proteins in order to mimic unfolded substrates. Interactions of these peptides with PDIA1 bb' domains

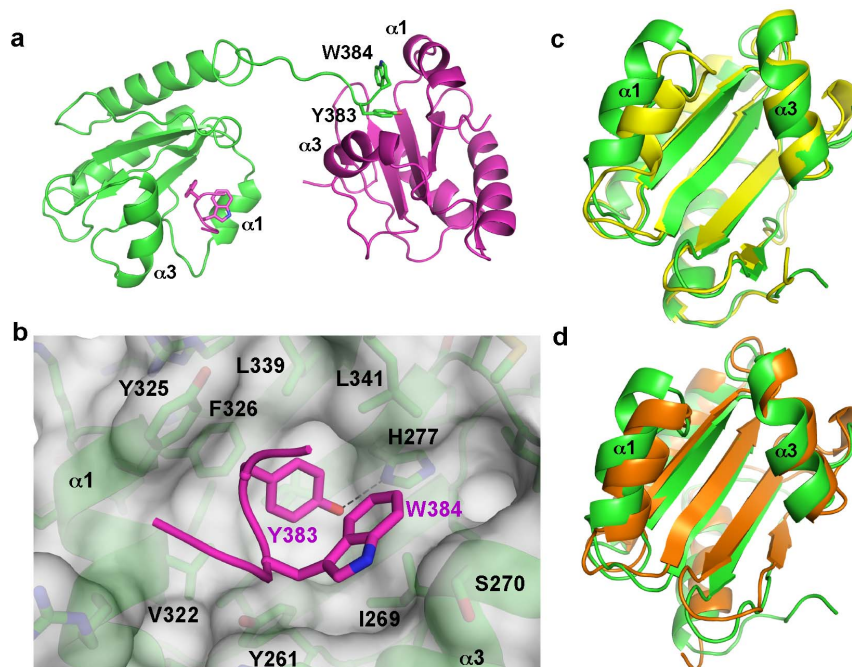


Figure 2 | A large hydrophobic pocket on the b' domain of PDILT is able to interact with aromatic side chains. (a) The residues Y383 and W384 from the C-terminal tail of the crystallized fragment bind to a hydrophobic pocket located between helices α1 and α3 of an adjacent PDILT molecule. (b) The base of the large pocket is lined with hydrophobic residues with the exception of H277, which forms a hydrogen bond with the hydroxyl group of Y383. (c) Overlay of the crystal structures of b' domain from PDILT (green) and PDIA1 (yellow; PDB accession number 3BJ5) reveals differing orientations of helix α1, resulting in a larger substrate-binding pocket in PDILT. (d) Overlay of the crystal structures of b' domain from PDILT (green) and the C-terminal domain of ERp27 (orange; PDB accession number 4F9Z).



have been characterized previously^{4,16,17}. Interestingly, Δ -somatostatin binds to the hydrophobic pocket of PDIA1 with a higher affinity compared to mastoparan (K_d of 35 μ M and 130 μ M, respectively)¹⁶. Since PDIA1 and PDILT display similar substrate-binding pockets, NMR titrations of ¹⁵N-labeled PDILT b' domain with mastoparan are expected to show significant shifts of the NMR signals from PDILT residues of the hydrophobic pocket, similar to those observed for PDIA1 bb' domains. Surprisingly, mastoparan did not interact with the PDILT b' domain, since NMR spectra remained unchanged after the addition of up to 1 : 10 protein/peptide molar ratio (Supplemental Fig. S2). Interestingly, the binding of somatostatin to PDILT was previously reported⁵ and could be due to the presence of aromatic residues that are absent in mastoparan.

ERp57-CNX recognition mechanism is not conserved in PDILT. Calnexin (CNX) and calreticulin (CRT), which are CMG and CRT3 homologs, use their negatively charged arm-like P-domains to bind to a large positively charged patch in the b' domain of ERp57, another member of the PDI family^{18–20}. In order to determine whether PDILT interacts with CMG and CRT3 in a similar manner, we performed NMR titrations of the ¹⁵N-labeled P-domains of CMG or CRT3 with unlabeled b' domain of PDILT. Addition of the b' domain did not result in any chemical shift changes, demonstrating no interaction (Supplemental Fig. S3). To explain this result, we compared the binding site for CNX and CRT in ERp57 b' domain with the PDILT structure. Both domains could be superimposed using their backbone atoms of 111 residues (out of 129 residues in PDILT b' domain) with an RMSD of 2.6 Å and a Z score of 13.4 using Dali database (Fig. 3). In ERp57, the CNX-binding surface is centered on helix $\alpha 2$ and is characterized by a pronounced positive charge (Fig. 3a). Interestingly, PDILT does not display a positively charged patch in the corresponding surface. This is consistent with the fact that PDILT and ERp57 b' -domains share only 11% identity, and the amino acids implicated in the ERp57/CNX interaction are not conserved in PDILT (Fig. 1c).

Discussion

Despite many years of studies, molecular details of substrate recognition by protein disulfide isomerases remain elusive. In particular, structural characterization of substrate-binding sites provides important information about this process. Here, we determined a high-resolution structure of the substrate-binding domain of PDILT, which is also the first structure of any PDILT domain. Importantly, the structure reveals a large hydrophobic pocket that seems to be primed for binding aromatic residues. In PDIA1, this hydrophobic pocket is a binding site for protein substrates and stabilizes hydrophobic residues exposed in unfolded regions of protein substrates^{16,17,21–23}. The higher affinity of the b' domain for aromatic tryptophan and tyrosine residues was previously observed for the other tissue-specific PDI, PDIp. Using a seven amino acid peptide phage-display library, Ruddock and co-workers showed that tyrosine or tryptophan residues within a folding polypeptide trigger the recognition and binding to PDIp²⁴. It is possible that PDILT uses similar preference for aromatic residues reflecting the selectivity for spermatogenesis-specific proteins. However, further studies are required to determine how PDILT recognizes its substrates *in vivo*.

PDILT interacts *in vivo* with testis-specific lectin-type chaperones, calnexin (CMG) and calreticulin 3 (CRT3)^{5,10}. Our results suggest that the mechanism of binding of PDILT to CMG and CRT3 is different from the one that occurs in ERp57/CNX, and other domains are necessary to mediate binding. For instance, the C-terminal tail of ERp57 is an additional site of interaction with CNX¹⁹. While the C-terminus alone does not interact with CNX, the basic stretch (residues ⁴⁹⁴KPKKKK⁵⁰⁰K of ERp57) is needed for high-affinity binding. Interestingly, the C-terminal tail of PDILT also contains a positively charged region (residues KKKTSEEEVVVVAKPK-

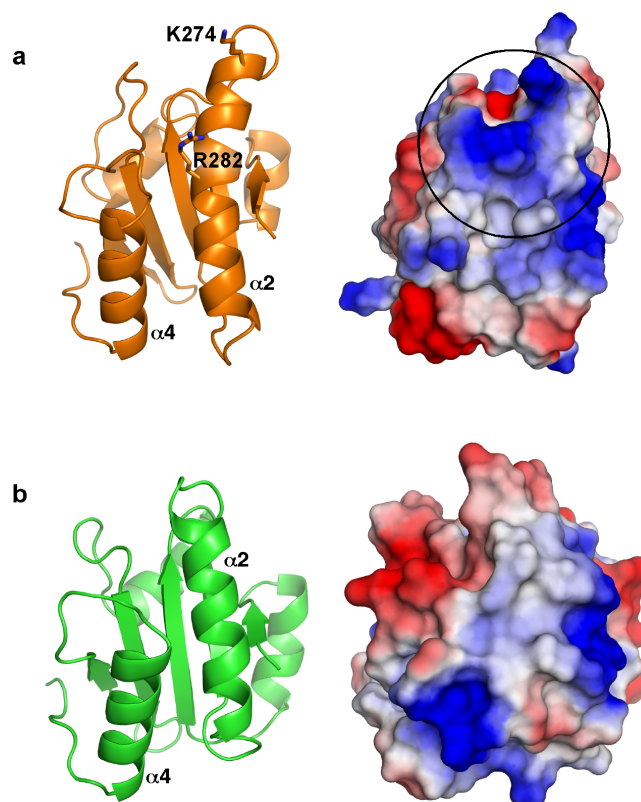


Figure 3 | Structural comparison between PDILT and ERp57 b' domains. Cartoon representations and electrostatic surfaces for ERp57 (a) and PDILT (b) are displayed in the same orientation. The residues K274 and R282 of ERp57 b' domain essential for binding to CNX and CRT are shown as stick representations. The positively charged CNX-binding site (circled) is not conserved in PDILT.

GPPVQKKKPK); therefore, it is possible that the C-terminus of PDILT in combination with other domains is required for binding to CMG and CRT3. More studies are needed to determine the mechanism underlying these interactions. Noteworthy, ERp57 is not able to bind to its substrates directly and requires CRT or CNX for substrate recognition²⁵, while the PDILT b' domain is more PDIA1-like and should be able to directly bind the non-native substrates through the hydrophobic pocket.

The specificity of substrate binding by protein disulfide isomerases is still poorly understood. PDIA1 acts on very broad range of substrates^{26,27}, while PDILT is expected to interact with spermatogenesis-specific protein substrates unique to this cell type. This structure is another step towards understanding of how substrate binding is mediated by PDI family members and PDILT in particular. The study of ER protein folding in the testis will lead to better understanding of pathways associated with male infertility.

Methods

Protein expression and purification. The PDILT b' domain (V258–Q386) was amplified by PCR and sub-cloned into pGEX-6P-1 (Amersham Pharmacia) using BamHI and XhoI restriction sites. Two modules of the CMG P-domain comprising the repeats 3 and 4 (residues 300–369) were cloned in pET15b (Novagen) containing N-terminal His-tag. For CRT3, the full-length P-domain (residues 198–291, 3-modules) was also cloned into pET15b using NdeI and BamHI restriction sites. Clones were confirmed by sequencing.

Protein was expressed in *E. coli* BL21 (DE3) cells. Expression of PDILT b' domain with a N-terminal GST-tag were accomplished by growing cells in LB medium at 37°C and shaking until the optical density (OD) at 600 nm reached 0.8, then protein production was induced by adding isopropyl-thiogalactopyranoside (IPTG) to a final concentration of 1 mM. Cells were incubated with shaking at 30°C for another 4 hours. The GST fusion protein was purified by affinity chromatography on glutathione-Sepharose resin (GE Healthcare). Once eluted from the resin, the tag was



removed by cleavage with PreScission protease leaving a N-terminal GPLGS sequence. The resulting PDILT b' domain protein with a molecular weight of 15.5 kDa was additionally purified using size-exclusion (Superdex-75) chromatography using 20 mM HEPES pH 7.5, 0.15 M NaCl buffer, and then concentrated using a Centricon tube (Millipore) with the 3-kDa cut-off. His-tagged proteins were purified by affinity chromatography on Ni-NTA (nickel-nitrilotriacetic acid) resin (QIAGEN). Eluted proteins were then cleaved with thrombin protease to remove the His-tag, leaving a N-terminal GSHM sequence. Proteins were further purified as described above.

For NMR experiments, the ^{15}N -labeled proteins were expressed by growing in M9 minimal medium with ^{15}N -ammonium chloride as the unique source of nitrogen. Expression and purification conditions were the same as for unlabeled proteins.

Protein crystallization. Crystallization conditions for PDILT b' domain were identified using commercial PACT suite (QIAGEN). The best PDILT b' domain crystals were obtained by equilibrating a 0.6 μl drop of the protein solution (7.2 mg/ml) in 20 mM HEPES pH 7.5, 150 mM NaCl, mixed with 0.6 μl of reservoir solution containing 0.2 M sodium formate, 0.1 M Bicine pH 9.0, 18% PEG3350, 0.1 M cobaltous chloride and suspended over 1 ml of reservoir solution. Crystals grew in 4 days at 22°C. For data collection, crystals were cryoprotected by addition of 20% (v/v) of ethylene glycerol and flash cooled in a N_2 cold stream. Crystals of PDILT b' domain belong to the primitive monoclinic system, space group P2₁, with four protein molecules per asymmetric unit corresponding to a solvent content of 36.9%.

Structure solution and refinement. Diffraction data from native PDILT b' domain crystals were collected using a single wavelength of 0.977 Å regime with an ADSC Quantum-210 CCD detector (Area Detector Systems Corp.) on beamline A1 at the Cornell High-Energy Synchrotron Source (CHESS). Data processing and scaling were obtained with HKL-2000²⁸. The structure of human PDIA1 b' x domain (residues 230–368, PDB accession code 3BJ5) was used as the search model for molecular replacement. The initial model obtained with Phaser²⁹ was then completed manually with the program Coot³⁰. The resulting model was improved by several cycles of refinement, using the program REFMAC³¹ and model refitting, followed by TLS (translation-libration-screw) refinement³². Structure figures were made with PyMOL (<http://www.pymol.org>).

NMR titrations. All NMR experiments were performed at 25°C on a Bruker DRX 600 MHz spectrometer. NMR samples were prepared with 90% of NMR buffer (20 mM HEPES pH 7.5 and 150 mM NaCl) and 10% of D₂O. NMR titrations were carried out by acquiring ^1H - ^{15}N heteronuclear single quantum correlation (HSQC) spectra of 0.2 mM of ^{15}N -labeled protein. Subsequent spectra were taken after the addition of an unlabeled ligand. ^{15}N -labeled CMG P-domain (residues 300–369) was titrated with unlabeled PDILT b' domain to the final ratio of labeled to unlabeled protein of 1 : 2. Purified mastoparan (INLKALAALAKKIL) peptide was added to the ^{15}N -labeled PDILT b' domain to a final protein/peptide ratio of 1 : 10. Also, ^{15}N -labeled PDILT b' domain was titrated with unlabeled CRT3 P-domain (residues 198–291) to a ratio of labeled/unlabeled protein of 1 : 2. Spectra were processed by NMRPipe³³ and analyzed by NMRview.

- Kozlov, G., Maattanen, P., Thomas, D. Y. & Gehring, K. A structural overview of the PDI family of proteins. *FEBS J.* **277**, 3924–36 (2010).
- van Lith, M., Hartigan, N., Hatch, J. & Benham, A. M. PDILT, a divergent testis-specific protein disulfide isomerase with a non-classical SXXC motif that engages in disulfide-dependent interactions in the endoplasmic reticulum. *J. Biol. Chem.* **280**, 1376–83 (2005).
- Fomenko, D. E. & Gladyshev, V. N. Identity and functions of CxxC-derived motifs. *Biochemistry* **42**, 11214–25 (2003).
- Klappa, P., Ruddock, L. W., Darby, N. J. & Freedman, R. B. The b' domain provides the principal peptide-binding site of protein disulfide isomerase but all domains contribute to binding of misfolded proteins. *EMBO J.* **17**, 927–935 (1998).
- van Lith, M. *et al.* A developmentally regulated chaperone complex for the endoplasmic reticulum of male haploid germ cells. *Mol. Biol. Cell* **18**, 2795–804 (2007).
- Guo, X. *et al.* Proteomic analysis of proteins involved in spermiogenesis in mouse. *J. Proteome Res.* **9**, 1246–56 (2010).
- Meistrich, M. L., Mohapatra, B., Shirley, C. R. & Zhao, M. Roles of transition nuclear proteins in spermiogenesis. *Chromosoma* **111**, 483–8 (2003).
- Yu, Y. E. *et al.* Abnormal spermatogenesis and reduced fertility in transition nuclear protein 1-deficient mice. *Proc. Natl. Acad. Sci. U S A* **97**, 4683–8 (2000).
- Ikawa, M. *et al.* Calsperin is a testis-specific chaperone required for sperm fertility. *J. Biol. Chem.* **286**, 5639–46 (2011).
- Tokuhiro, K., Ikawa, M., Benham, A. M. & Okabe, M. Protein disulfide isomerase homolog PDILT is required for quality control of sperm membrane protein ADAM3 and male fertility [corrected]. *Proc. Natl. Acad. Sci. U S A* **109**, 3850–5 (2012).

- Yamaguchi, R., Fujihara, Y., Ikawa, M. & Okabe, M. Mice expressing aberrant sperm-specific protein PMIS2 produce normal-looking but fertilization-incompetent spermatozoa. *Mol. Biol. Cell* **23**, 2671–9 (2012).
- Maattanen, P., Kozlov, G., Gehring, K. & Thomas, D. Y. ERp57 and PDI: multifunctional protein disulfide isomerases with similar domain architectures but differing substrate-partner associations. *Biochem. Cell Biol.* **84**, 881–9 (2006).
- Holm, L. & Rosenstrom, P. Dali server: conservation mapping in 3D. *Nucleic Acids Res.* **38**, W545–9 (2010).
- Kober, F. X. *et al.* The crystal structure of the protein-disulfide isomerase family member ERp27 provides insights into its substrate binding capabilities. *J. Biol. Chem.* **288**, 2029–39 (2013).
- Fu, X. M., Wang, P. & Zhu, B. T. Characterization of the estradiol-binding site structure of human protein disulfide isomerase (PDI). *PLoS One* **6**, e27185 (2011).
- Denisov, A. Y. *et al.* Solution structure of the bb' domains of human protein disulfide isomerase. *FEBS J.* **276**, 1440–9 (2009).
- Byrne, L. J. *et al.* Mapping of the ligand-binding site on the b' domain of human PDI: interaction with peptide ligands and the x-linker region. *Biochem. J.* **423**, 209–17 (2009).
- Kozlov, G. *et al.* Crystal structure of the bb' domains of the protein disulfide isomerase ERp57. *Structure* **14**, 1331–9 (2006).
- Pollock, S. *et al.* Specific interaction of ERp57 and calnexin determined by NMR spectroscopy and an ER two-hybrid system. *EMBO J.* **23**, 1020–9 (2004).
- Frickel, E. M. *et al.* TROSY-NMR reveals interaction between ERp57 and the tip of the calreticulin P-domain. *Proc. Natl. Acad. Sci. U S A* **99**, 1954–9 (2002).
- Gruber, C. W., Cemazar, M., Heras, B., Martin, J. L. & Craik, D. J. Protein disulfide isomerase: the structure of oxidative folding. *Trends Biochem. Sci.* **31**, 455–64 (2006).
- Freedman, R. B., Klappa, P. & Ruddock, L. W. Protein disulfide isomerases exploit synergy between catalytic and specific binding domains. *EMBO Rep.* **3**, 136–40 (2002).
- Klappa, P., Hawkins, H. C. & Freedman, R. B. Interactions between protein disulfide isomerase and peptides. *Eur. J. Biochem.* **248**, 37–42 (1997).
- Ruddock, L. W., Freedman, R. B. & Klappa, P. Specificity in substrate binding by protein folding catalysts: tyrosine and tryptophan residues are the recognition motifs for the binding of peptides to the pancreas-specific protein disulfide isomerase PDip. *Protein Sci.* **9**, 758–64 (2000).
- Rutkevich, L. A. & Williams, D. B. Participation of lectin chaperones and thiol oxidoreductases in protein folding within the endoplasmic reticulum. *Curr. Opin. Cell Biol.* **23**, 157–66 (2011).
- Hatahet, F. & Ruddock, L. W. Substrate recognition by the protein disulfide isomerases. *FEBS J.* **274**, 5223–34 (2007).
- Hatahet, F. & Ruddock, L. W. Protein disulfide isomerase: a critical evaluation of its function in disulfide bond formation. *Antioxid. Redox Signal.* **11**, 2807–50 (2009).
- Otwinowski, Z. M. W. Processing of X-ray Diffraction Data Collected in Oscillation Mode. *Methods Enzym.* **276**, 307–326 (1997).
- McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
- Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–32 (2004).
- Murshudov, G. N., Vagin, A. A., Lebedev, A., Wilson, K. S. & Dodson, E. J. Efficient anisotropic refinement of macromolecular structures using FFT. *Acta Crystallogr. D Biol. Crystallogr.* **55**, 247–55 (1999).
- Winn, M. D., Murshudov, G. N. & Papiz, M. Z. Macromolecular TLS refinement in REFMAC at moderate resolutions. *Methods Enzym.* **374**, 300–21 (2003).
- Delaglio, F. *et al.* NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277–93 (1995).
- Laskowski, R. A. M., M. W., Moss, D. S. & Thornton, J. M. PROCHECK - a program to check the stereochemical quality of protein structures. *J. App. Cryst.* **26**, 283–291 (1993).

Acknowledgments

Data acquisition at the Macromolecular Diffraction (MacCHESS) facility of the Cornell High Energy Synchrotron Source (CHESS) was supported by the National Science Foundation award DMR 0225180 and the National Institutes of Health award RR-01646. This study was funded by an NSERC Discovery grant RGPIN-238873 to K.G.

Author contributions

S.B.A. and G.K. designed the experiments; S.B.A. performed the experiments; S.B.A. and G.K. analyzed the data; S.B.A., G.K. and K.G. wrote the main manuscript; G.K. and S.B.A. prepared figures. All authors reviewed the manuscript.

Additional information

Database: Coordinates have been deposited into the Protein Data Bank database under the accession number 4NWY

Supplementary information accompanies this paper at <http://www.nature.com/scientificreports>

Competing financial interests: The authors declare no competing financial interests.



How to cite this article: Bastos-Aristizabal, S., Kozlov, G. & Gehring, K. Structure of the substrate-binding b' domain of the Protein disulfide isomerase-like protein of the testis. *Sci. Rep.* **4**, 4464; DOI:10.1038/srep04464 (2014).



This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported license. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-sa/3.0>