ISSN 1479-2621
Plant Genetic Resources: Characterization and Utilization (2011) **9**(2); 347–351
doi:10.1017/S1479262111000244

Protein disulphide isomerase family in bread wheat (*Triticum aestivum* L.): protein structure and expression analysis

A. R. Paolacci[1], M. Ciaffi[1]*, A. P. Dhanapal[2], O. A. Tanzarella[1], E. Porceddu[1] and E. d'Aloisio[2]

[1]*Dipartimento di Agrobiologia e Agrochimica, Università della Tuscia, Viterbo, Italy and*
[2]*Scuola Superiore Sant'Anna, Pisa, Italy*

Abstract

The deduced amino-acid sequences of 17 protein disulphide isomerase (PDI) and PDI-like cDNAs of wheat assigned to nine homoeologous groups were searched for conserved motives by comparison with characterized sequences in different protein databases. The wheat protein sequences encoded by genes of different homoelogous groups showed a high level of structural similarity with the corresponding protein sequences of other species clustering into the same phylogenetic group. The proteins of five groups (I–V) share two thioredoxin-like active domains and show structural similarities with the corresponding proteins of higher eukaryotes, whereas those of the remaining three groups (VI–VIII) contain a single thioredoxin-like active domain. The expression analysis of the nine non-homoeologous wheat genes, which was carried out by quantitative RT-PCR in developing caryopses and in seedlings subjected to temperature stresses, showed their constitutive although highly variable transcription rate. The comprehensive structural and transcriptional characterization of the *PDI* and *PDI*-like genes of wheat performed in this study represents a basis for future functional characterization of the PDI gene family in the hexaploid context of bread wheat.

Keywords: expression analysis; protein disulphide isomerase; protein structure; temperature stress; *Triticum aestivum*

Introduction

The protein disulphide isomerase (PDI) family includes several genes whose products are responsible for diversified metabolic functions, including secretory protein folding, chaperone activity and redox signalling. PDI and PDI-like proteins differ for number and position of their thioredoxin-like active (type a) and inactive (type b) domains, for presence/absence of other domains and of the Lysine. Aspartic acid, Glutamic acid, Leucine (KDEL) retention signal in the endoplasmic reticulum (ER). The a-type domains are homoeologous to that of thioredoxin and contain the catalytic site (CXXC), whereas the b-type domains lack the active tetrapeptide, but have a secondary structure similar to that of thioredoxin.

Phylogenetic analysis resolved the PDI family into eight groups (Houston *et al.*, 2005), five with two TRX-like active domains (I–V) and three with only one (VI–VIII). The purpose of this research was to characterize the genes of the PDI family in wheat and to compare the structure of their deduced amino-acid sequences and their expression with those of homoeologous genes of other plant species. Former studies in wheat had been limited to the genes encoding the typical PDI (Ciaffi *et al.*, 2006), which is of special interest for its potential involvement in determining the technological properties of flour.

* Corresponding author. E-mail: ciaffi@unitus.it

Material and methods

Developing caryopses from 5 to 38 d after anthesis at 5–6 d intervals (seven samplings) were collected from 20 bread wheat plants (*Triticum aestivum* cv. Chinese Spring). For temperature stress experiments, 20-d-old CS seedlings were exposed to either 33 or 4°C for 48 h; each treatment included two biological replications of 20 plants. Shoots were harvested just before (control), after 24 h and after 48 h of temperature stress, frozen in liquid nitrogen and stored at −80°C.

The deduced protein sequences were analysed by searching for conserved motifs in CDD, InterPro and SMART databases; their subcellular locations were predicted by Target P1.1 and ChloroP 1.1, the presence of the signal peptide was confirmed by Signal P3.0 and the transmembrane regions were determined by TMHMM version 2.0; protein identity was estimated using DNAMAN software. Quantitative RT-PCR (qRT-PCR) analyses and data normalization were performed according to Paolacci *et al.* (2009). Two biological replicates, resulting from two different RNA extractions and qRT-PCR reactions, were used in quantification analysis; moreover, three technical replicates were analysed for each biological replicate.

Results and discussion

Structural characteristics and domain organization of the deduced amino-acid sequences of three homoeologous cDNAs coding for the typical PDI (Ciaffi *et al.*, 2006) and of 14 PDI-like cDNAs, of nine different homoeologous groups are described in Table 1 and Supplementary Fig. S1 (available online only at http://journals.cambridge.org). Except TaPDIL8-1, all proteins contained a putative signal peptide for translocation into the ER. TaPDIL2-1 and TaPDIL3-1 showed a multidomain organization similar to the typical PDI, with the addition of an N-terminal domain (*c*) rich of acidic residues and a putative calcium-binding site. TaPDIL4-1 had a *D* domain consisting of a C-terminal α-helical of about 100 amino acids with unknown function and a potential ER-translocation signal. Like all proteins of the fourth group, it lacked the ER-retention signal; consequently, it might be targeted to a different subcellular location or be retained as part of an heteromeric complex; however, in *D. discoideum*, the C-terminal part of the *D* domain is responsible for the ER retention of the PDI-D protein (Monnat *et al.*, 2000). The proteins of the seventh group have a transmembrane segment which, even without the KDEL signal, could retain the protein in the ER by anchoring it to the membrane; in man, *Drosophila* and *C. elegans*, all proteins with similar structures are involved in developmental control (Clissold and Bicknell, 2003). TaPIDL8-1 has two transmembrane segments; the C-terminal segment is part of the DUF1692 domain, found in several proteins. ERGIC-32 is one of them (Breuza *et al.*, 2004), which is localised in the ER–Golgi intermediate compartment (ERGIC), the first anterograde/retrograde sorting station in the mammalian secretory pathway. Protein sorting seems only one of the ERGIC functions. Other presumed functions are less clear; it appears involved in protein quality control and in the retrotranslocation to the cytosol of permanently misfolded proteins. Further studies will be necessary to understand whether or not in plants the PDI-like proteins belonging to the VIII phylogenetic group are localized in a compartment similar to the ERGIC of mammalian and to investigate their function.

The functional features of the proteins encoded by the PDI gene family of wheat can be predicted on the basis of their domain structure and of the presence/absence of three major determinants whose role has been identified in humans (Ellgaard and Ruddock, 2005). Besides the active site tetrapeptide -CXHC-, typically -CGHC-, there are two additional prominent determinants for the activity of the PDI-family proteins. A conserved arginine modulates the pK_a of the active-site cysteines and is involved in the timing, allowing a single catalyst to act as isomerase and oxidase and facilitating the release of non-productive folding substrates. Moreover, the catalytic cycle for oxidation or reduction requires numerous proton transfer reactions; in the thioredoxins, a charged glutamic acid–lysine pair, located under the active site, is the proton acceptor. Table 1 reports the active-site tetrapeptide and the positions of conserved charge pair sequence and arginine residues determined on the basis of multiple alignments of the a-type domains of wheat PDI-like proteins and human typical PDI (Supplementary Fig. S2, available online only at http://journals.cambridge.org).

The transcription level of the nine genes of the wheat PDI family was investigated in a series of seed developmental stages (Fig. 1(a–f)) from cellularisation to physiological maturity, showing a considerable variation between the genes. The transcripts of some PDI genes containing two active thioredoxin domains, such as *TaPDIL1-1*, *TaPDIL2-1*, *TaPDIL4-1* and *TaPDIL5-1*, increased dramatically during the early stages of seed development (5–10 d post anthesis; DPA), then declined steadily during the major part of the grain filling period (15–27 DPA). Consequently, their temporal expression preceded that of seed storage protein genes and reached a maximum before the more intense synthesis of gluten proteins. For *TaPDIL1-1*, *TaPDIL4-1* and *TaPDIL5-1*, it is noteworthy that a second peak of transcription was detected by qRT-PCR between 27 and 32 DPA, when the deposition of seed storage reserves decreased and seeds

Table 1. Characteristics of the wheat PDI and PDI-like proteins

Name	Length	Mw	pI	N-glycosilation sites (putative)	Domain composition[a]	Active site sequence	Conserved charge pair sequence[b]	Conserved Arg[b]
TaPDIL1-1a	515	56.59	4.99	1:N283	a-b-b′-a′	CGHC, CGHC	E62-K96, E406-K439	R136, R475
TaPDIL1-1b	512	56.44	5.03	1:N283	a-b-b′-a′	CGHC, CGHC	E62-K96, E406-K439	R136, R475
TaPDIL1-1c	515	56.63	4.96	1:N283	a-b-b′-a′	CGHC, CGHC	E62-K96, E406-K439	R136, R475
TaPDIL2-1	588	63.80	4.61	2:N109,N212	c-a-b-b′-a′	CGHC, CGHC	E123-K157, E464-K497	R193, R535
TaPDIL3-1	541	59.61	4.95	1:N150	c-a-b-b′-a′	CERS, CVDC	*L90*-K124, E429-R462	R160, *P492*
TaPDIL4-1a	367	40.27	6.17	0	a°-a-D	CGHC, CGHC	E54-K87, E173-K211	R125, R244
TaPDIL4-1b	367	40.26	6.17	0	a°-a-D	CGHC, CGHC	E54-K87, E173-K211	R125, R244
TaPDIL5-1a	440	47.15	5.12	1:N170	a°-a-b	CGHC, CGHC	E51-K89, E188-K226	R119, R257
TaPDIL5-1b	440	47.22	5.36	1:N170	a°-a-b	CGHC, CGHC	E51-K89, E188-K226	R119, R257
TaPDIL6-1a	151	16.99	4.96	0	a	CKHC	*Q56-S95*	R126
TaPDIL6-1b	149	16.65	5.30	0	a	CKHC	*Q54-S93*	R124
TaPDIL7-1a	413	46.30	4.91	1:N275	a-b-b′-t	CGHC	D56-K90	R126
TaPDIL7-1b	417	46.62	4.91	2:N176,N279	a-b-b′-t	CGHC	D60-K94	R130
TaPDIL7-1c	413	46.32	4.87	2:N172,N275	a-b-b′-t	CGHC	D56-K90	R126
TaPDIL7-2a	418	46.40	5.12	0	a-b-b′-t	CGHC	D64-K98	R134
TaPDIL7-2b	418	46.34	5.03	1:N384	a-b-b′-t	CGHC	D64-K98	R134
TaPDIL8-1	485	54.41	6.90	ND	t-a-t	CYWS	N164-K203	R249

[a] a, Active site containing the thioredoxin-like domain; b, inactive thioredoxin-like domain (superscript prime and degree symbols are included to distinguish multiple a and b domains on the basis of their position and not of sequence homology); c, acidic segment; D, Erp29c domain; t, transmembrane domain. [b] Positions of conserved charge pair sequence and arginine residues important for the catalytic activity of different proteins of the human PDI family were determined on the basis of multiple alignments of the a-type domains of wheat PDI-like proteins and human classical PDI (Accession number P07237) (Supplementary Fig. S2, available online only at http://journals.cambridge.org). ND, not determined because TaPDIL8-1 lacks a putative N-terminal signal peptide. Proteins without signal peptides are unlikely to be exposed to N-glycosilation machinery and thus may not be glycosylated *in vivo* even though they contain potential motifs.

started desiccating. Seemingly, the proteins encoded by these genes may play an important role in later stages of seed development, when there is a dramatic increase of large gluten polymers stabilized by the formation and/or the rearrangement of inter-chain disulphide bonds.

Analyses by qRT-PCR of six seedling samples consisting of two temperature treatments (4 and 33°C) each for 24 and 48 h and their controls (Fig. 1(g) and (h)) indicated that the nine genes were differentially regulated by cold and heat stresses. The expression of *TaPDIL4-1*and *TaPDIL8-1* was not affected by temperature treatments, whereas that of *TaPDIL2-1* was significantly down-regulated by both cold and heat stresses and that of *TaPDIL7-2* was about twice lower in shoots exposed to low temperatures (Fig. 1(g)). A significant induction was observed for at least one of the two temperature treatments for the remaining five *PDI*-like genes (Fig. 1(h)). The most significant variation in transcription rate induced by temperature treatments was observed for *TaPDIL1-1* (typical PDI) and *TaPDIL5-1*.

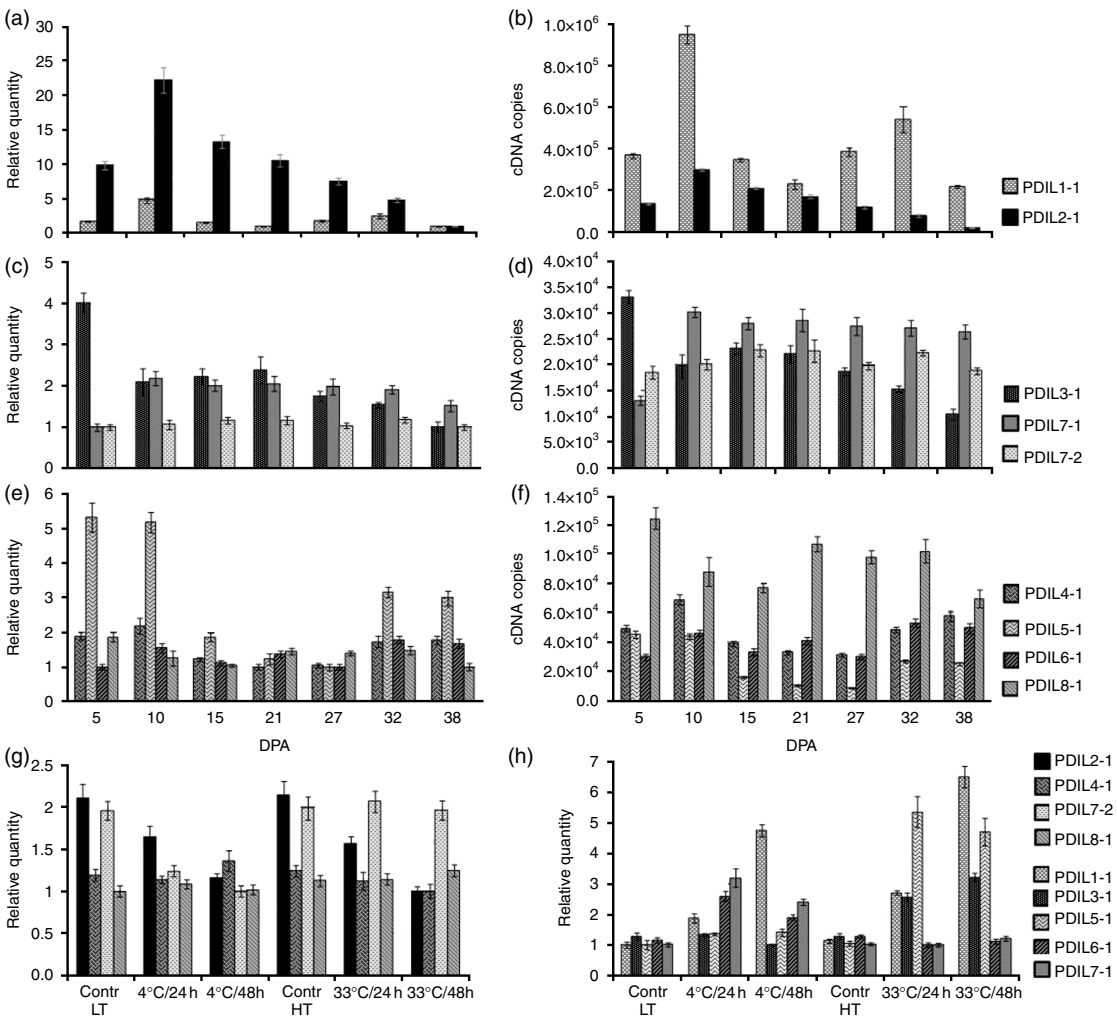

Fig. 1. Expression analysis of PDI and PDI-like genes in developing wheat caryopses and in wheat seedlings exposed to high- (HT) and low-temperature (LT) treatments. Relative (a, c and e) and absolute (b, d and f) quantification of the expression level in developing caryopses collected between 5 and 38 DPA of nine PDI and PDI-like genes. Relative (g and h) quantification of the expression level in six samples consisting of seedlings exposed to two temperature treatments (4 and 33°C) for 24 and 48 h and their controls of nine PDI and PDI-like genes. The 14 cDNA pools (two biological replicates, seven developing caryopsis samples; a–f) and the 12 cDNA pools (two biological replicates, six seedling samples; g and h) were tested in triplicate and normalized using the geometric average of the relative expression of the two reference genes encoding cell division control protein and ADP-ribosylation factor. Relative expression levels of the nine genes were referred to that of a calibrator set to the value 1, which was represented by the developmental stage (a, c and e) or by the temperature treatments (g and h) with the lowest expression. Absolute expression levels of the nine genes were expressed as number of cDNA copies/(g of reverse transcribed total RNA (b, d and f). Normalized values of relative and absolute expressions of the nine genes are given as average ± SD. Contr, control samples grown at 18°C.

https://doi.org/10.1017/S1479262111000244 Published online by Cambridge University Press

References

Breuza L, Halbeisen R, Jeno P, Otte S, Barlowe C, Hong W and Hauri HP (2004) Proteomics of endoplasmic reticulum–Golgi intermediate compartment (ERGIC) membranes from brefeldin A-treated HepG2 cells identifies ERGIC-32, a new cycling protein that interacts with human Erv46. *The Journal of Biological Chemistry* 279: 47242–47253.

Ciaffi M, Paolacci AR, D'Aloisio E, Tanzarella OA and Porceddu E (2006) Cloning and characterization of wheat PDI (protein disulfide isomerase) homoeologous genes and promoter sequences. *Gene* 366: 209–218.

Clissold PM and Bicknell R (2003) The thioredoxin-like fold: hidden domains in protein disulfide isomerases and other chaperone proteins. *Bioessays* 25: 603–611.

Ellgaard L and Ruddock LW (2005) The human protein disulphide isomerase family: substrate interactions and functional properties. *European Molecular Biology Organization Reports* 6: 28–32.

Houston NL, Fan C, Xiang QY, Schulze JM, Jung R and Boston RS (2005) Phylogenetic analyses identify 10 classes of the protein disulfide isomerase family in plants, including single-domain protein disulfide isomerase-related proteins. *Plant Physiology* 137: 762–778.

Monnat J, Neuhaus EM, Pop MS, Ferrari DM, Kramer B and Soldati T (2000) Identification of a novel saturable endoplasmic reticulum localization mechanism mediated by the C-terminus of a *Dictyostelium* protein disulfide isomerase. *Molecular Biology of the Cell* 11: 3469–3484.

Paolacci AR, Tanzarella OA, Porceddu E and Ciaffi M (2009) Identification and validation of reference genes for quantitative RT-PCR normalization in wheat. *BMC Molecular Biology* 10: 11.

https://doi.org/10.1017/S1479262111000244 Published online by Cambridge University Press