Wheat Protein Disulfide Isomerase (PDI) Promoter Sequence Analysis in *Triticum aestivum* cv Chinese Spring and its Wild Relatives

Arun Prabhu Dhanapal1,2,*, Enrico Porceddu1,3

1Plant Genetic Resources, Scuola Superiore Sant’Anna, Piazza Martiri della Libertà, Pisa, Italy
2Present addressee: Division of Plant Sciences, University of Missouri, Columbia, MO, USA
3Dipartimento di Agrobiologia e Agrochimica, Universita della Tuscia, Viterbo, Italy

*Corresponding author: a.dhanapal@sssup.it

Received September 13, 2013; Revised September 26, 2013; Accepted September 29, 2013

**Abstract** Protein disulphide isomerase (PDI) is an oxidoreductase enzyme abundant in the endoplasmic reticulum (ER). Plant PDIs has been shown to be involved in the folding and deposition of seed storage proteins, which makes this enzyme particularly interesting in wheat, as flour quality is strongly affected by composition and structure of seed storage proteins. Promoter sequences of three homoeologous genes encoding typical PDI, located on chromosome group four of bread wheat, and PDI promoter sequence analysis of *Triticum urartu*, *Aegilops speltoides* and *Aegilops tauschii* had also been reported previously. In this study, we report the isolation, cloning and sequencing of a ~1450 bp region, comprising ~1350 bp of the putative promoter region and 88 bp of the first exon of the typical PDI gene, from *Triticum urartu* (AA), *Aegilops speltoides* (BB) and *Aegilops tauschii* (DD). Sequence analysis indicated close similarity was found within each species and with the corresponding homoeologous PDI sequences of *Triticum aestivum* cv. CS (AABBDD) resulting in an overall high conservation of the sequence in proximal region then distal region of promoter conferring endosperm-specific expression.

**Keywords:** protein disulphide isomerase, cloning, promoter, regulatory elements, wheat wild relatives


1. **Introduction**

Wheat is adapted to temperate regions of the world and was one of the first crops to be domesticated. Spread over all continents, it is today one of the most important food source for human beings. Bread wheat (*Triticum aestivum*) belongs to the tribe *Triticeae*, and has three genomes A, B and D, each organised in seven homoeologous chromosome groups. The diploid progenitors of the three genomes A, B, and D have been identified in *Triticum urartu*, *Aegilops speltoides* and *Aegilops tauschii*, although the progenitor of B genome is still matter of debate [1]. Throughout their evolutionary history, multiple polyploidization events occurred between species of the *Triticum* and *Aegilops* genera and human manipulation of wild species led to the domestication of different cultivated lineages [2,3,4,5].

Wheat grains are consumed under many different forms, such as flour for leavened, flat and steamed breads, biscuits, cookies, cakes, breakfast cereal, pasta, noodles, couscous and for fermentation to make beer, alcohol, vodka and biofuel. It is composed by starch, proteins and other compounds, which accumulate in significant quantities during its development, in particular grain storage proteins, are responsible for the quality of the end product. The majority of them are prolamins, which account for > 90 % of the total protein content in the wheat grain. Different quality of grain storage proteins produced by the same genotype under different environmental condition has stimulated research on protein folding and assembling.

Secretory protein folding and disulfide bond formation takes place within the Endoplasmic reticulum (ER lumen), but the precise mechanisms involved, and the role of other proteins such as molecular chaperones, are not fully understood [6,7]. PDI plays an important role in assisting protein folding and assembly, catalyzing thiol-disulfide oxidation, reduction and isomerisation, this latter occurring directly by intramolecular disulfide rearrangement or through cycles of reduction and oxidation [8,9,10]. During the maturation of the secretory proteins, disulfide bonds cross-linking specific cysteines are added to stabilize a protein or to join covalently different polypeptides. These bonds are crucial for the stability of the final protein structure, thus mispairing of cysteine residues can prevent proteins from attaining their native conformation and lead to misfolding [11].

Typical PDI/Classical PDI is the most prominent member of a family of related proteins (PDI-like) characterised by one, two or three thioredoxin-like active sites.
domains [12]. It has been cloned and sequenced in many plant species, such as alfalfa [13], barley [14], maize [15], castor bean [16] and soybean [17,18,19,20], common and durum wheat [21,22]. A detailed knowledge of the complexity and diversity of genes encoding PDI and PDI-like proteins in *A. thaliana*, wheat and other plant species was described by [23,24,25].

Wheat genes coding for typical PDI in bread wheat have been located in homoeologous chromosome group four [26]. Analyses performed on 23 species of *Triticum* and *Aegilops* [27], indicated that PDI restriction fragments were highly conserved within each species and confirmed that plant PDI is encoded either by one or few copy sequences, respectively in diploid and polyploid species. The nucleotide sequences of the three genes located respectively on genomes A, B, and D (designated as *GPDI*-4A, *GPDI*-4B and *GPDI*-4D) were 3561bp, 3527bp and 3466 bp long. The comparison of typical PDI gene sequences of wheat, rice and *Arabidopsis* showed a significant conservation of the exon/intron structure across species [28]. More detailed study on the complexity and diversity of genes encoding PDI and PDI-like proteins in wheat and other plant species also showed a significant conservation of the exon/intron structure [23,24,29].

The expression analysis of the typical PDI homoeologous genes located on chromosomes 4A, 4B and 4D of bread wheat cv Chinese Spring (CS) [28] showed that the PDI transcripts, although constitutively present at a low-level in all the analyzed tissues, are equally abundant in the developing caryopses, but are differentially expressed in spikelets, roots and leaves [21,28]. The PDI-4A transcription was higher in spikelets that of PDI-4B were higher in roots while the PDI-4D transcripts were more abundant in leaves. The transcription levels of the three genes were higher in the early stage of seed development (6-14 DAA) and decreased during middle to late stage of (18-34 DAA) grain filling [28]. Within the upstream putative promoter region of the three homoeologous genes cloned from bread wheat cv CS, respectively 1352 bp for *PromPDI*-4A, 1370 bp for *PromPDI*-4B and 1292 bp for *PromPDI*-4D long, several cis-acting elements involved in endosperm specific expression were detected, consistently with the higher PDI expression detected in the kernels. The variability and evolutionary relationship in a ~700 bp region, comprising ~600 bp of the 5′ upstream putative promoter region and 88 bp of the first exon of the typical PDI gene from the diploid species *Triticum uratu* (AA) *Aegilops speltoides* (BB) and *Aegilops tauschii* (DD) have been reported [30,31,32]. This paper reports the cloning and characterization of ~1350 bp of the 5′ upstream putative promoter region and 88 bp of the first exon of the typical PDI gene from the diploid species *Triticum uratu* (AA) *Aegilops speltoides* (BB) and *Aegilops tauschii* (DD).

### 2. Materials and Methods

#### 2.1. Plant Material

A total of three accessions *Triticum uratu* (IG 44831, AA), *Aegilops speltoides* (IG 46812, BB) and *Aegilops tauschii* (AE 1068, DD) were used in this study. Single plant per accession was grown and used for DNA extraction. Flag leaves were collected at heading stage from plants grown in greenhouse (January- June 2008), immediately frozen in liquid nitrogen and kept at –80 °C until use. About 200 mg of leaf tissue was ground in liquid nitrogen and genomic DNA was extracted using Sigma Gen Elute Plant Genomic DNA Kit (G2N-350, Sigma Aldrich, St. Louis, Mo.).

#### 2.2. Primers Design and PCR Amplification

Primers were designed on the basis of the known homoeologous PDI promoter and gene sequences isolated from bread wheat cv Chinese Spring [28], using DNAMAN 4.15 program. Putative promoters of typical PDI were amplified by using the following primer pairs *PDIAPDF1*-PDIAPDR1 and *PDIAPDF2*-PDIAPDR2 for A Genome, *PDIAPDF3*-PDIAPDR3 and *PDIAPDF2*-PDIAPDR4 for B genome and *PDIAPDF3*-PDIAPDR1 and *PDIAPDF2*-PDIAPDR6 for D genome (Table 1), cloned and sequenced in each accession.

PCR reaction mixture included 10 ng/µl genomic DNA, 0.20 mM NTPs 5ul, 0.05 units/µl of Taq 0.50 µl (go-taq, Promega), 0.40µM primer 2.0 µl (0.20µM per each), 5X buffer 10 µl and 25.5 µl of ddH2O used per reaction (final volume of 50 µl). PCR condition included initial denaturation step at 95°C for three min, then 32 cycles of (95°C for one min, 59°C for 35 sec, 72°C for one min) and final elongation at 72°C for amplification of the sequences from the A, B and D Genome.

#### 2.3. Cloning and Sequencing

Genome specific primer pair (Table 1) was used to amplify, clone and sequence the DNA extracted from each one of the three wild relatives. PCR products of expected size were excised from the gel, purified using the High Pure Purification kit (Roche) according to manufacturer’s instructions, and cloned into the pGEM-T easy plasmid vector (Promega). Plasmids were transformed by heat shock into *Escherichia coli* strain DH5α. Bacteria were plated onto LB medium containing ampicillin, X-Gal and IPTG, and recombinant plasmids were identified by blue/white screening. For each primer combination two independent PCR reactions were performed and a total of six clones were sequenced. Plasmid DNA for sequencing reaction was prepared from three ml overnight cultures using a plasmid miniprep kit (Qiagen). Sequencing was performed on both strands by the ABI PRISM 377 DNA sequencer (PE Applied Biosystem) using an ABI Prism Dye Terminator sequencing kit (PE Applied Biosystem) and sequenced both with vector and sequence specific primers.

#### 2.4. Data Analysis

Sequences were analysed in Chromas version 2.3 (http://technelysium.com.au/chromas.html) to identify any unresolved bases and subjected to visual inspection. Analysis of three new sequences were done along with the three homoeologous gene and promoter sequences from CS (GPDI-4A, AJ868102; GPDIA-4B, AJ868103; *GPDI*-4D, AJ868104; PromPDI4A, AJ868108; PromPDI4B, AJ868109; PromPDI4D, AJ868110; [28]). Sequences
were then searched for regulatory elements in PlantCARE (http://sphinx.rug.ac.be:8080/PlantCARE/), a database of plant promoters and (http://www.dna.affrc.go.jp/htdocs/PLACE/ [33]) PLACE database. The six sequences were multiple aligned with Clustal X software version 2.011 [34], IUB as DNA weight matrix with default parameters. A phylogenetic tree was constructed on obtained data by using the neighbour-joining (NJ) method [35], using MEGA version4 with neighbour joining option and 1000 bootstrap [36]. All three cloned sequences have been submitted in EMBL Nucleotide Sequence Database with sequence ID HE588130, HE588131 and HE588132 for Triticum uratu (IG 44831), Aegilops speltoides (IG 46812) and Aegilops tauschii (AE 1068) respectively.

3. Results

3.1. PDI Promoter Sequence

The PDI promoter sequences from one plant each of Triticum uratu (IG 44831), Aegilops speltoides (IG 46812) and Aegilops tauschii (AE 1068) were amplified using the following primer pairs (Table 1) corresponding respectively to sequences comprising putative PDI promoter and part of the first exons of the PDI gene belonging to the A, B and D genome was visualized by gel electrophoresis. Two independent PCR reactions were performed for each of the two primer pair/DNA combinations used and in all cases a specific amplification product of the expected electrophoretic mobility was obtained and cloned. For every each cloned amplification product three clones were randomly chosen and sequenced, for a total of 6 clones for each primer pair/DNA combination. The six DNA sequences deriving from the same primer pair/DNA combination resulted 100% identical.

The length of the cloned region upstream of the translation start codon was 1334 bp for Triticum uratu (AA, accession IG 44831), 1332 bp for Aegilops speltoides (DD, accession AE 1068); the three sequences showed 92.38 % identity in the promoter region (Figure 2). The specificity and uniqueness of their respective amplification products was confirmed through cloning all fragments and sequencing.

3.2. Similarity of Sequences with CS

A high degree of conservation was detected in the putative promoter sequences of the entire three wild genomes cloned. Differences were due to both nucleotide substitutions and short insertions/deletions. The promoter sequence of Triticum uratu (AA, accession IG 44831) showed 99.40 % identity with one deletion and nucleotide substitution. Aegilops tauschii (DD, accession AE 1068) promoter sequence showed 99.04 % identity with few insertion and nucleotide substitution. Whereas promoter sequence of Aegilops speltoides (BB, accession IG 46812) showed only 89.59 % identity with large number of deletions and nucleotide substitution.

3.3. Phylogenetic Analysis

The evolutionary relationship between six sequences, of the putative promoter was studied by phylogeny reconstruction. These included one cloned sequence from Triticum uratu, Aegilops speltoides, Aegilops tauschii for a total of 3 new sequences and the three homoeologous sequences from CS previously reported [28]. The phylogenetic tree was constructed using three different methods namely the neighbour-joining (NJ) method, Minimum Evolution (ME) and Maximum Parsimony (MP) methods. As the results of the three methods were similar, only NJ tree is presented here (Figure 1).

![Figure 1. Phylogenetic tree of PDI gene promoter sequence of Triticum uratu (TU AA), Aegilops speltoides (AS BB) and Aegilops taushcii (TT DD) with three homoeologous promoter sequences from CS](image)

<table>
<thead>
<tr>
<th>PCR target for Cloning</th>
<th>Primer</th>
<th>Forward</th>
<th>Reverse</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA GENOME</td>
<td>PDIAPDF1</td>
<td>5’-TTCCAGCTTGAGATGAGGC-3’</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PDIAPDF2</td>
<td>5’-ACTCAAATTTGGAACGGG-3’</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PDIAPDR1</td>
<td>5’-TTGCCGGTTCAGGACTAGG-3’</td>
<td>5’-GGTGAGCACCTCCTCGGG-3’</td>
</tr>
<tr>
<td></td>
<td>PDIAPDR2</td>
<td>5’-TTGCCGGTTCAGGACTAGG-3’</td>
<td>5’-GGTGAGCACCTCCTCGGG-3’</td>
</tr>
<tr>
<td></td>
<td>PDIAPDF3</td>
<td>5’-TTGCCGGTTCAGGACTAGG-3’</td>
<td>5’-GGTGAGCACCTCCTCGGG-3’</td>
</tr>
<tr>
<td></td>
<td>PDIAPDF2</td>
<td>5’-TTGCCGGTTCAGGACTAGG-3’</td>
<td>5’-GGTGAGCACCTCCTCGGG-3’</td>
</tr>
<tr>
<td>BB GENOME</td>
<td>PDIAPDF3</td>
<td>5’-TTGCCGGTTCAGGACTAGG-3’</td>
<td>5’-GGTGAGCACCTCCTCGGG-3’</td>
</tr>
<tr>
<td></td>
<td>PDIAPDF2</td>
<td>5’-TTGCCGGTTCAGGACTAGG-3’</td>
<td>5’-GGTGAGCACCTCCTCGGG-3’</td>
</tr>
<tr>
<td>DD GENOME</td>
<td>PDIAPDR1</td>
<td>5’-TTGCCGGTTCAGGACTAGG-3’</td>
<td>5’-GGTGAGCACCTCCTCGGG-3’</td>
</tr>
<tr>
<td></td>
<td>PDIAPDR2</td>
<td>5’-TTGCCGGTTCAGGACTAGG-3’</td>
<td>5’-GGTGAGCACCTCCTCGGG-3’</td>
</tr>
<tr>
<td></td>
<td>PDIAPDR4</td>
<td>5’-TTGCCGGTTCAGGACTAGG-3’</td>
<td>5’-GGTGAGCACCTCCTCGGG-3’</td>
</tr>
<tr>
<td></td>
<td>PDIAPDR6</td>
<td>5’-TTGCCGGTTCAGGACTAGG-3’</td>
<td>5’-GGTGAGCACCTCCTCGGG-3’</td>
</tr>
</tbody>
</table>
Figure 2. Multiple sequence alignment of PDI promoter sequences from *Triticum uratu* (TU_AA), *Aegilops speltoides* (AS_BB) and *Aegilops tauschii* (TT_DD)
<table>
<thead>
<tr>
<th>Motif</th>
<th>Function</th>
<th>Sequence</th>
<th>A Genome</th>
<th>B Genome</th>
<th>D Genome</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Triticum aestivum (4A)</td>
<td>Triticum urartu (AA)</td>
<td>Triticum aestivum (4B)</td>
<td>Aegilops speltoides (BB)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>+ Strand</td>
<td>- Strand</td>
<td>+ Strand</td>
<td>- Strand</td>
</tr>
<tr>
<td>Skn-1</td>
<td>Cis-acting regulatory element for endosperm expression</td>
<td>ACGAC      ACGAC      ATGAC</td>
<td>-480</td>
<td>-480</td>
<td>-50</td>
</tr>
<tr>
<td></td>
<td></td>
<td>+ Strand (number)</td>
<td>- Strand (number)</td>
<td>+ Strand (number)</td>
<td>- Strand (number)</td>
</tr>
<tr>
<td>AACA</td>
<td>Cis-acting element conserved in rice glutelin genes and involved in</td>
<td>AACC</td>
<td>1</td>
<td>8</td>
<td>5</td>
</tr>
<tr>
<td>CAAT box</td>
<td>Cis-acting element common in promoter and enhancer regions</td>
<td>12</td>
<td>12</td>
<td>4</td>
<td>6</td>
</tr>
</tbody>
</table>

Table 2: Main regulatory motifs found in the 5' upstream of typical PDI gene sequence in *Triticum aestivum*, *Triticum urartu*, *Aegilops speltoides* and *Aegilops tauschii* in both positive and negative strands from ATG.
3.4. Conserved Regulatory Motifs

The search for regulatory motifs in the promoter sequences of *Triticum uratu* (IG 44831), *Aegilops speltoides* (IG 46812) and *Aegilops tauschii* (AE 1068) upstream the coding region in the database of plant promoters (PlantCARE) and PLACE database [http://www.dna.affrc.go.jp/htdocs/PLACE/; [33]] detected a TATA box located at ~79 nt from the start codon, and a number of different cis-acting regulatory elements (Table 2) including several motifs (AACA, prolamin box, GCN4, Skn-1) involved in the regulation of endosperm specific genes [41,42,44,44].

Several CAAT-boxes were present in both (+) strand and (-) strands of A, B and D genome sequences controlling endosperm and tissue specific expression. AACA motif was present in (+) strand of all genomes, eight in A genome, eleven in B genome and 10 in D genome, but in (-) strand, only five in A, seven in B and six in D genome were identified. A prolamine-box was discovered on (+) strands at position -226 bp in *Triticum uratu* and at position -227 bp in *Aegilops speltoides*. No prolamine-box was detected in the target region of *Aegilops tauschii*. The prolamine-box was present in an identical position in the corresponding promoter sequences of the PDI gene in CS located on homoeologous chromosomes 4A and 4B of CS as described in previous study [28]. Skn-1 like element were found at position -480 bp, -481bp and -480 bp on (+) strand respectively in the A, B and D chromosomes. In the homoelogous sequences of CS, [28] identified them at identical position in the corresponding promoter sequences from CS. Similar result was also obtained from series of partial promoter region from *Triticum uratu*, *Aegilops speltoides* and *Aegilops tauschii* reported previously [32,49,50].

Former studies in wheat had been restricted to the characterization of genes encoding the typical PDI and their promoter, and the cloning and characterization of complete set of genes encoding PDI and PDI like proteins in bread wheat (*Triticum aestivum* cv Chinese Spring) and the comparison of their sequence, structure and expression with homologous genes from other plant species, which is of special interest for its involvement in determining bread making qualities and technological properties of flour. The interest of extending the study to promoters of wild species is due to high conservation of gene, due to their relevant metabolic functions, as well as to the interesting expression pattern found in previous work. Present study confirmed the conservation of important motifs conferring the endosperm specific expression with evolutionary relationship between wild relatives and bread wheat (*Triticum aestivum* cv Chinese Spring). An intensive analysis carried out on the transcription levels of nine PDI and PDI-like genes in a set of 29 samples of wheat, including tissues, developmental stages and temperature stresses, showed their constitutive although very variable expression. Highly diversified expression rates and expression patterns were evidenced not only by genes belonging to different phylogenetic groups, but also in close paralog genes. This variable expression pattern may be due to differing regulatory elements and their numbers. The very high expression of the gene TaPDIL1-1, encoding the typical PDI, in the developing caryopses, is consistent with its hypothesised role in the folding, aggregation and deposition of seed storage proteins. This function of the typical PDI has been demonstrated experimentally in maize, rice and soybean [23,24,45,46,47,48,51,52].

Future studies should involve characterization of proximal and distal end of the Typical PDI promoter in diploid and tetraploid progenitor from diverse geographic origin of world to determine to determine the similarity and identity. Functional analysis of Chinese spring PDI promoter (full and partial sequence) driving the GUS gene and selective silencing of the PDI and PDI-like genes in wheat plants. The characterization of the regulatory motifs through the expression studies of the progressive deletions of their promoters, as well as the expression analysis of the PDI gene from accession of wild and cultivated wheat
is currently under way. This analysis will help in elucidating the function of some regulatory elements controlling the spatial and temporal specific expression of the PDI and PDI–like genes.

5. Conclusions

The very high expression of the gene TaPDIL1-1, encoding the typical PDI, in the developing caryopses, which is consistent with its hypothesised role in the folding, aggregation and deposition of seed storage proteins was confirmed. This function of the typical PDI has been demonstrated experimentally in maize, rice and soybean. Future studies should involve characterization of proximal and distal end of the Typical PDI promoter in diploid and tetraploid progenitor from diverse geographic origin of world to determine to determine the similarity and identity. Functional analysis of Chinese spring PDI promoter (full and partial sequence) driving the GUS gene and identity. Functional analysis of Chinese spring PDI promoter (full and partial sequence) driving the GUS gene and identity. Functional analysis of Chinese spring PDI promoter (full and partial sequence) driving the GUS gene and identity. Functional analysis of Chinese spring PDI promoter (full and partial sequence) driving the GUS gene and identity. Functional analysis of Chinese spring PDI promoter (full and partial sequence) driving the GUS gene and identity. Functional analysis of Chinese spring PDI promoter (full and partial sequence) driving the GUS gene and identity. Functional analysis of Chinese spring PDI promoter (full and partial sequence) driving the GUS gene and identity. Functional analysis of Chinese spring PDI promoter (full and partial sequence) driving the GUS gene and identity. Functional analysis of Chinese spring PDI promoter (full and partial sequence) driving the GUS gene and identity.

Acknowledgement

I thank Shalini Narasimhan for helping me in editing grammatical errors and finalizing the manuscript. My Ph.D programme scholarship entitled “International Doctoral Scholarship for Agrobiodiversity” from Scuola Superiore Sant’ Anna Piazza Martiri della Libertà, 33 - 56127 Pisa, Italy from February 2007 to February 2010 is highly acknowledged.

Conflict of Interest

The authors declare that they have no conflict of interest.

References


[50] Dhanapal, A.P., d’Aloisio, E., Ciaffi, M., Porceddu, E., “Conserved regulatory motifs identified in Protein Disulfide Isomerase Promoter sequence analysis in wheat wild relatives.” *Proceedings of the International Workshop on Plant Genetic Resources For Food And Agriculture-General aspects and research opportunities*, 5-6 November 2009b, Rome, Italy.
