ISSN 1479-2621
Plant Genetic Resources: Characterization and Utilization (2011) **9**(2); 338–341
doi:10.1017/S147926211100013X

Protein disulphide isomerase promoter sequence analysis of *Triticum urartu*, *Aegilops speltoides* and *Aegilops tauschii*

Arun Prabhu Dhanapal[1], Mario Ciaffi[2], Enrico Porceddu[2]* and Elisa d'Aloisio[1]

[1]*International Doctoral Programme on Agrobiodiversity, Scuola Superiore Sant'Anna – Plant Genetic Resources, Centro Ricerche ENEA Cassacia, Rome 00123, Italy and*
[2]*Dipartimento di Agrobiologia e Agrochimica, Università della Tuscia, Viterbo, Italy*

Abstract

Protein disulphide isomerase (PDI) catalyses the formation, reduction and isomerization of disulphide bonds in the newly synthesized secretory proteins. Plant PDIs have 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. In hexaploid wheat cultivar (AABBDD) Chinese Spring (CS), the genomic, complementary DNA and promoter sequences of the three homoeologous gene encoding PDI had been isolated and characterized in a previous study revealing high levels of sequence conservation. In this study, we report the isolation and sequencing of a ~700 bp region, comprising ~600 bp of the putative promoter region and 88 bp of the first exon of the typical PDI gene, in five accessions each from *Triticum urartu* (AA), *Aegilops speltoides* (BB) and *Aegilops tauschii* (DD). Sequence analysis indicated large variation among sequences belonging to the different genomes, while 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 regulatory motifs conferring endosperm-specific expression.

Keywords: promoter; protein disulphide isomerase; regulatory elements; variability and wild species

Introduction

Protein disulphide isomerase (PDI) catalyses the formation and isomerization of disulphide bonds in nascent proteins targeted to the endoplasmic reticulum. Different studies on its expression and intracellular localization in wheat and maize (Shimoni *et al*., 1995; Li and Larkins, 1996), as well as the characterization of the rice *esp2* mutant that lacks PDI expression (Takemoto *et al*., 2002), suggest that PDI is involved in regulating the folding of seed storage proteins and their deposition into the protein bodies of the endosperm. Wheat storage proteins are synthesized in the developing endosperm cells and targeted to their endoplasmic reticulum lumen, wherein they are folded and connected by intermolecular disulphide bonds to form large aggregates (Shewry and Tatham, 1997), whose molecular weight is directly related to higher dough elasticity and better technological quality. Even though wheat storage proteins have been the object of a wide range of studies both at chemical and at genetic levels (reviewed by Shewry *et al*. (2003, 2009)), knowledge of factors affecting their folding and deposition is still extremely limited. The genomic and complementary DNA sequences of three PDI homoeologous genes located in the chromosomes 4A, 4B and 4D of bread wheat cv. Chinese Spring (CS) and their putative

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

promoters had been cloned and sequenced previously (Ciaffi *et al.*, 2006). The three genes showed a very high sequence conservation in the coding region and their exon/intron structures, which consisted of ten exons. The putative promoter sequences of the three genes shared some regulatory motifs involved in endosperm-specific expression; consistently with this observation, expression analysis of the three homoeologous PDI genes showed the constitutive presence of their mRNAs in several wheat tissues, but the expression level was much higher in developing caryopses. This study reports the isolation and characterization of a ~700 bp region, comprising ~600 bp of the putative promoter sequence and 88 bp of the first exon of the *PDI* gene from wild diploid relatives of wheat, *Triticum urartu* (AA), *Aegilops speltoides* (BB) and *Aegilops tauschii* (DD).

Materials and methods

A total of 15 accessions, five each from *T. urartu* (AA), *A. speltoides* (BB) and *A. tauschii* (DD), were used in this study. A total of 120 plants (8 plants/accession) were grown and analysed. Flag leaves were collected at heading stage from plants grown in green house (January–June 2008), immediately frozen in liquid nitrogen and kept at −80°C until use. About 200 mg of leaf tissue were ground in liquid nitrogen, and genomic DNA was extracted using Sigma Gen Elute Plant Genomic DNA Kit (G2N-350; Sigma Aldrich, St. Louis, MO, USA). F2 (5′-ACTCCAAATTTGGAACGGG-3′) and R1 (5′-GGTG-AGCACTGCCTCGGG-3′) primers were chosen to amplify the genomic DNA, and amplification products were directly sequenced. Sequences were analysed in Chromas version 2.3 (http://technelysium.com.au/chromas.html) to identify any unresolved bases and subjected to visual inspection. Multiple sequence alignment was performed with Clustal X software version 2.011 (Larkin *et al.*, 2007), International Union of Biochemistry as DNA weight matrix with default parameters. Haplotype diversity, polymorphic sites and evolutionary relationship were calculated using DnaSP v5 software (Librado and Rozas, 2009). Sequences were then searched for regulatory elements in PlantCARE (http://sphinx.rug.ac.be:8080/PlantCARE/), a database of plant promoters, and PLACE databases (http://www.dna.affrc.go.jp/htdocs/PLACE/; Higo *et al.*, 1999).

Results and discussion

Analysis of the 120 sequences revealed that there was no variability among plants within accessions, thus a total of 15 sequences one from each accession were further analysed and compared with the three homoeologous sequences from CS (GPDI-4A, AJ868102; GPDIA-4B, AJ868103; GPDI-4D, AJ868104; Ciaffi *et al.*, 2006). The alignment of the five sequences derived from the accessions of *T. urartu* revealed the presence of two haplotypes: one common to the accessions from Lebanon, Turkey, Jordan and Syria and the other characteristic of the accession from Iran; in total, two variable sites were identified (Fig. 1). Both cases involved transversions (G–T and C–G). Haplotype diversity (Hd) was 0.400. This high level of conservation of *PDI* gene promoter in the proximal region of *T. urartu* may reflect its important function or the effect of a bottleneck. The alignment of the five sequences from *A. speltoides* revealed a higher number of polymorphic sites and the presence of five haplotypes, each characteristic of one of the five analysed accessions. Out of the 16 identified single nucleotide polymorphisms (SNPs), whose positions are reported in Fig. 1, ten were transitions (A–G and C–T) and six were transversions (A–C, A–T, G–C and G–T); moreover, one Indel at position −7 bp from the ATG was also identified. Hd was 1.000. The number of polymorphic sites found in *A. speltoides* was thus higher than in *T. urartu*, consistently with its partial outcrossing nature. Finally, the alignment of

AA genome (*T. urartu*)

Accessions	SNP/haplotype position
[	44]
[	45]
[	18]
IG45475-1_Lebanon	GC
IG115817-2_Jordan	GC
IG44831-3_Syria	GC
IG45108-4_Turkey	GC
IG45477-5_Iran	TG

BB genome (*A. speltoides*)

Accessions	SNP/haplotype position	
[	112333333	334555]
[	1356011244	792699]
[	9832038245	677335]
IG48766-1_Lebanon	GTGATTTTAA	CCATGC
IG46811-2_Turkey	....CCC...	TTCC..
IG46812-3_Turkey	.CCGCC.AG.	TTCC.T
IG46593-4_Syria	T...CCC...	TTCCC.
IG46597-5_Syria	....CCC..T	TTCC..

DD genome(*A. tauschii*)

Accessions	SNP/haplotype position
[	11]
[	89]
[	11]
AE525-1_Iran	AG
AE526-2_Iran	..
AE527-3_Iran	..
AE541-4_Iran	CT
AE1068-5_Syria	CT

Fig. 1. SNP positions identified in the ~700 bp amplicon comprising the partial PDI promoter sequences and partial first exon from *T. urartu* (AA), *A. speltoides* (BB) and *A. tauschii* (DD). Dots denote consensus sequences.

https://doi.org/10.1017/S1479262111000013X Published online by Cambridge University Press

Table 1. Regulatory motifs identified in the ~700 bp amplicon comprising the partial protein disulphide isomerase promoter sequences and partial first exon from *T. urartu* (AA), *A. speltoides* (BB) and *A. tauschii* (DD)[a]

	T. urartu (AA)		*A. speltoides* (BB)		*A. tauschii* (DD)	
Regulatory motifs	Consensus sequence	Position from ATG	Consensus sequence	Position from ATG	Consensus sequence	Position from ATG
TATA box	TATTAAA	−79	TATTAAA	−83[b], −79	TATTAAA	−79
PROLAMINE box	TGAAAAGT	−226	CGAAAAGT	−231[b], −227	CGAAAAGT	NA
AACA	AACA	−323, −547, −588	AACA	−243[b], −270[b], −296[b], −548[b], −568[b], −247, −274, −300, −552, −572	AACA	−223, −242, −269, −324, −547, −589
Skn-1-like	ACGAC	−480	ACGAC	−485[b], −481	ATGAC	−480
GCN-4 4-like	CGACTCA, CATGTCA	−356, −492	CGAGTCA, CATGTCA	−362[b], −358, −497[b], −493	CGTGTCA	−492
ACGT- box	GTACGTG	−442	GTACGTG	−447[b], −443	GTACGTG	−442
CACGTG	CACGTG	−426	CACGTG	−431[b], −427	CACGTG	−426

NA, not applicable.
[a] Their consensus sequence and position are also reported. [b] The position of motifs for IG48766-1.

the sequences from the five analysed accessions of *A. tauschii* (DD) revealed the presence of only two haplotypes, one common to three out of the four analysed accessions from Iran and one shared by accession number AE541-4 from Iran and the accession from Syria (Fig. 1). In total, only two variable sites were identified (Fig. 1), both of which were transversions (G-T and C-G). Hd was 0.600. The presence of fewer polymorphic sites is probably partially due to self-fertilizing nature of the species (Dvorak *et al.*, 1998) and partially to the lower level of overall diversity present. The 15 sequences were further compared with their counterparts described previously in bread wheat cv. CS (Ciaffi *et al.*, 2006), and the percentage of nucleotide identity is reported (Supplementary Table S1, available online only at http://journals.cambridge.org) Furthermore, the evolutionary relationship between the 18 sequences, five sequences each from *T. urartu*, *A. speltoides*, *A. tauschii* and the three homoeologous sequences from the bread wheat cv. CS, was investigated (data not shown). The sequences from *A. speltoides* (BB) closely clustered together with that located on chromosome 4B of CS, those from *A. tauschii* (DD) with that on 4D and those from *T. urartu* (AA) with that on 4A.

All sequences were further searched for regulatory elements in two databases, PlantCARE and PLACE. Several regulatory elements, among which many involved in endosperm-specific expression, were identified. The TATA box, a *cis*-regulatory element found in the promoter region of many eukaryotic genes and considered as the core promoter sequence, was identified at position −79 from the ATG (Table 1). The combination of four motifs, namely GCN-4 motif, prolamine box, AACA and ACGT box, seems to be a minimal requirement for endosperm-specific expression (Wu *et al.*, 1998, 2000). All four motifs were identified in the putative PDI promoter sequences analysed. Furthermore, the Skn-1-like motif, an additional motif that together with ACGT and the prolamine box modulates the level of endosperm-specific expression, was also identified. The presence of these regulatory motifs is consistent with the higher PDI expression found in developing caryopses with respect to other tissues (Ciaffi *et al.*, 2001). The regulatory elements are identified, and their consensus sequence and position are reported (Table 1). The present study revealed a high level of conservation of regulatory elements conferring endosperm-specific expression between sequences from CS and those isolated in this study. Further studies should involve large number of accessions as well the characterization of the proximal and distal ends of the typical PDI promoter in diploid and tetraploid progenitor from diverse geographic origin of the world.

https://doi.org/10.1017/S147926211100013X Published online by Cambridge University Press

References

Ciaffi M, Paolacci AR, Dominici L, Tanzarella OA and Porceddu E (2001) Molecular characterization of gene sequences coding for protein disulfide isomerase (PDI) in durum wheat (*Triticum turgidum* ssp. *durum*). *Gene* 265: 147–156.

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.

Dvorak J, Luo MC, Yang ZL and Zhang HB (1998) The structure of the *A. tauschii* genepool and the evolution of hexaploid wheat. *Theoretical and Applied Genetics* 97: 657–670.

Higo K, Ugawa Y, Iwamoto M and Higo H (1999) PLACE: a database of plant *cis*-acting regulatory DNA elements. *Nucleic Acids Research* 26: 358–359.

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ and Higgins DG (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* 23: 2947–2948.

Li CP and Larkins BA (1996) Expression of protein disulfide isomerase is elevated in the endosperm of the maize *floury*-2 mutant. *Plant and Molecular Biology* 30: 873–882.

Librado P and Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25: 1451–1452.

Shewry PR and Tatham AS (1997) Disulphide bonds in wheat gluten proteins. *Journal of Cereal Science* 25: 207–227.

Shewry PR, Halford NG, Tatham AS, Popineau Y, Lafiandra D and Belton PS (2003) The high molecular weight subunits of wheat glutenin and their role in determining wheat processing properties. *Advances in Food and Nutrition Research* 45: 219–302.

Shewry PR, Underwood C, Wan Y, Lovegrove A, Bhandari D, Toole G, Mills ENC, Denyer K and Mitchell RAC (2009) Storage product synthesis and accumulation in developing grains of wheat. *Journal of Cereal Science* 50: 106–112.

Shimoni Y, Zhu X, Levanoy H, Segal G and Galili G (1995) Purification, characterization, and intracellular localization of glycosylated protein disulfide isomerase from wheat grains. *Plant Physiology* 108: 327–335.

Takemoto Y, Coughlan SJ, Okita TW, Hikaru S, Masahiro O and Tohihiro K (2002) The rice mutant *esp2* greatly accumulates the glutenin precursor and deletes the protein disulfide isomerase. *Plant Physiology* 128: 1212–1222.

Wu CY, Suzuki A, Washida H and Takaiwa F (1998) The GCN4 motif in a rice glutelin gene is essential for endosperm-specific gene expression and is activated by Opaque-2 in transgenic rice plants. *Plant Journal* 14: 673–683.

Wu CY, Washida H, Onodera Y, Harada K and Takaiwa F (2000) Quantitative nature of the Prolamin-box, ACGT and AACA motifs in a rice glutelin gene promoter: minimal *cis*-element requirements for endosperm-specific gene expression. *Plant Journal* 23: 415–421.