Australian Journal of Crop Science AJCS
AJCS 8(1):44-54 (2014) ISSN:1835-2707

Isolation and characterization of Dehydration-Responsive Element-Binding Factor 2 (DREB2) from Indian wheat (*Triticum aestivum* L.) cultivars

Bharati Pandey[1, 2], Pradeep Sharma[1]*, Manoj Saini[1], Dev Mani Pandey[2], Indu Sharma[1]

[1]Plant Biotechnology, Directorate of Wheat Research, Karnal-132001, India
[2]Department of Biotechnology, Birla Institute of Technology, Mesra, Ranchi- 835 215, India

***Correspondence author: neprads@yahoo.com**

Abstract

Plants have developed a sophisticated stress response system to cope with and adapt to different types of abiotic stress imposed by the frequenlty adverse environmets. It is essential to understand plants stress response mechanisms in the attempt to improve crop yield under stressful condition. Dehydration responsive element binding (DREB) proteins are important transcription factors known to regulate diverse processes of plant development and stress responses. In this study, *DREB2* gene was amplified using specific primer from a diverse set of bread wheat cultivars. PCR products were purified using a gel extraction kit and sequenced. Obtained sequences were aligned and submitted to the NCBI database with HQ804651-55, FC556845-53 accession numbers. Alignment of corresponding amino acid sequences showed conservation of AP2 domain in Indian wheat genotypes. Transcripts level of the genes were shown to be most abundant (17%) in leaves, followed by flower (11%) and the least abundant in stem and seed (8%). To understand the structure function relationships, tertiary structure of DREB2 protein from wheat were built by homology modelling based on the crystal structure of GCC-box binding domain of *Arabidopsis thaliana*. Predicted models were refined by energy minimization and further validated by Procheck and root mean square deviation (RSMD) value. This is the first documentation of the 3-D model of DREB2 in wheat which is extremely reliable and stable. More similarities were found between AP2/EREBP protein of *A. thaliana* and *T. aestivum* by conducting protein docking with the DNA containing GCC-box. A protein was found to interact through their β-sheet, with the major DNA groove by hydrogen and hydrophobic bond, which provides structural stability to the molecule. Data generated in this study will be useful for conducting genomic research to determine the precise role of the *AP2/ERF* gene under water stress responses with the ultimate goal of improving crops.

Keywords: AP2 domain, *DREB*, Docking, Expression analysis, Transcription factor, Wheat.
Abbreviations: TF_Transcription factor, AP2_APETALA2 factor, DRE_Dehydration responsive element, ERF_Ethylene responsive element binding factor, DREB_Dehydration responsive element binding factor, EST_Expressed sequence tag, RMSD_Root mean square deviation, Mw_Molecular weight, DDD_Digital differential display, EREBP_Ethylene responsive element binding protein.

Introduction

Abiotic stresses are serious threats to agriculture and results in environmental deterioration. Among the abiotic stress factors that have shaped and continue shaping plant evolution, water availability is the most important. Plant growth and productivity are affected by various abiotic stresses, such as extreme temperatures, drought and high salinity. Nevertheless, these stresses together represent the primary cause of crop loss more than 50% worldwide (Boyer, 1982; Bray et al., 2000). To survive, plants must respond and adapt to these stresses at the molecular and cellular levels, as well as at the physiological and biochemical levels. An increasing number of stress induced genes at transcription level have been identified in several plant species (Zhu, 2002; Bartels and Sunkar 2005; Yamaguchi-Shinozaki and Shinozak, 2006). Drought tolerance is a complex trait controlled by many genes, including those encoding functional proteins (e.g. late embryogenesis abundant (LEA) proteins, osmotin, molecular chaperones, and mRNA binding proteins) or whose products are the transcription factors (TFs) such as basic leucine zipper (bZIP), myelocytomatosis oncogene (MYC) and dehydration-responsive element-binding factor (DREB) (Liu et al., 2000). The TFs interacting with specific *cis*-acting elements in the promoters regions of various stress-related genes to up-regulate the expression of many downstream genes, thus imparting stress tolerance to environmental stress and pathogen attack (Chen and Zhu, 2004). DREB is one of the largest families of TFs that form an important role in signaling network which modulates many plant processes, such as abiotic stress tolerance (Liu et al. 1998). The DREB proteins induce a set of abiotic stress responsive genes and maintain water balance in plant systems thus imparting abiotic stress tolerance. The DREB transcription factors have been divided into two classes, DREB1 and DREB2, involved in two separate signal transduction pathways under low temperature and dehydration, respectively. They belong to the ethylene responsive factor (ERF) family of transcription factors. ERF proteins are the sub-family of the AP2 (APETALA2)/EREBP (Ethylene Responsive Element Binding Protein) TFs that are unique to plants (Sakuma et al., 2002). In *Arabidopsis*, the two subclasses namely DREB1/CBF and DREB2 are induced by cold and dehydration, respectively (Shinozaki et al., 2000). The *DREB2* genes have been reported to confer drought tolerance in several plant species such as rice, maize,

barley, wheat and *Arabidopsis* (Ehawa et al., 2006; Dubouz et al., 2003; Qin et al., 2007; Xue et al., 2004; Cong et al., 2008; Zhao et al., 2009; Lui et al., 1998). Though, *DREB1/CBF* and *DREB2* genes share a sequence similarity at AP2 domain and bind to the same DRE sequence, therefore up-regulated by low temperature (DREB1) or by drought/high salt concentration (DREB2). DREB transcription factors bind to the promoter region of genes such as *RD29*, *COR15* and *RD17*, thus indicating expression in response drought, salt and cold. Hence, it is very important to explore the resistance mechanism of plants to clone *DREB* gene and then study its sequence, structure and functions (Wang et al., 2004). In order to understand the transcriptional regulation in wheat, different classes of TF were isolated and characterized, including some of the AP2/ERF family (Galiab et al., 2009). Previously, Shen and co-workers have characterized one TF of DRE-binding protein (TaDREB1) from a drought induced cDNA library of *T. aestivum* (Shan et al., 2003). A DREB subfamily gene *WDREB2* was isolated from wheat and its expression was activated by cold, drought, salt and ABA treatment (Egawa et al., 2006). A pathogen-inducible AP2/ERF family gene, TaERF3, in wheat was isolated and found to be involved in defense response (Zhang et al., 2007). DRE, which has a core sequence TACCGACAT, is recognized by proteins of the DREB subfamily (Liu et al., 2000; Yamaguchi-Shinozaki and Shinozaki, 2005).The sequence CCGAC inside the DRE element is the minimum sequence motif for binding and C4, G5, and C7 are essential for specific interaction (Hao et al., 1998; Sakuma et al, 2002). Additionally, the DNA-binding specificity of *Arabidopsis DREB*s is well known. It has been shown that both DREB1 and DREB2 specifically bind to six nucleotides (A/GCCGAC) of DRE. This consensus sequence is generally referred as GCC-box, thus strongly suggesting that DREB proteins contain GCC-box binding (Agarwal et al., 2007). DREB2 transcription factor seems to be one of the most promising candidate gene involved in conferring drought tolerance in several crops. For better understanding of the molecular mechanism of target recognition and to visualize target genes for transcription factors at the genomic level, it is important to examine the relationship between the structure and function of transcription factors (Garg et al., 2008). Homology of DREB protein and its binding element have been done in the many plants. However, only a few reports exist that reveals the binding properties of this protein with the gene (s) promoter. In this study, we have isolated the *DREB2* homologue in a diverse set of Indian wheat genotypes (Table 1) with the aim to explore the structure-function relationships of *Triticum aestivum DREB2* gene. Using multiple sequence alignment, protein homology modeling and docking of TaDREB2 protein with the DNA containing GCC-box was carried out. In addition, a theoretical model of the tertiary structure prediction showed that TaDREB2 has a highly conserved GCC-box-binding N'-terminus and an acidic C'-terminus domain that acts as an activation domain for transcription. Furthermore, the protein has also been docked with the DNA double helix having GCC-box to explain the residues involved in the Protein-DNA interaction. To the best of our knowledge, this is the first report regarding structure function aspects of *T. aestivum* DREB2 using protein homology modeling which may provide the framework for understanding the functions of DREBs at the molecular level and can be used as a tool to investigate the regulatory mechanism. Besides, a large group of *T. aestivum* DREB ESTs was analyzed to gain insights with the *DREB* family gene expression among diverse wheat tissues. DREB is imparting important role in dehydration tolerance in crop plants, so there is a need to explore possible functions of *DREB* in Indian wheat genotype.

Results and Discussion

Isolation of DREB2 gene in Indian wheat genotypes

The *TaDREB2* was amplified using gene specific primer from a diverse set of Indian wheat genotypes (see Table 1) possessing different level of drought tolerance (Unpublished data). Fig. 1 shows a fragment of ~500 bp that represented *DREB*2 gene. The partial sequences were submitted to NCBI GenBank with JN191706-10, JF728296-JF728301, and HQ840652-55 accession numbers. Each isolated *DREB* gene was annotated using BLASTX, with cutoff expectation (e) value of $1e^{-8}$, against the UniProtKB-TrEMBL database and showed similarity with *DREB* gene from barley, sorghum, rice, and leguminous plants.

Sequence alignment and conserved motifs of DREB in wheat

In this study, the deduced amino acid sequences of TaDREB2 were subjected to detailed functional analysis. Nucleotide alignment of DNA sequences of fifteen bread wheat cultivars was carried by Muscle software (Suppl. Fig. 1) indicating that our sequences showed ~98 % homology among each other. We noticed variation at a few nucleotides, which might be critical for amino acid changes. Recently, Mondini et al. (2012) observed a single nucleotide change in DREB1, HTK1 and WRKY1 genes in durum genotypes while exposed to various stresses. However, it remains to be explored in the present study. Protein alignment showed conserved N'-terminal nuclear localization signal (NLS) and AP2 domain (57 amino acids) from 56 to 113 amino acid (Fig. 2). Since, TFs only function in the nucleus, the regulation of their entry into the nucleus is mediated by NLS (Akhtar et al., 2012). Alignment of the AP2/ERF domain of published DREB2 proteins from *Poa pratensis* (Pp), *Festuca arundinacea* (Fa), *Avena sativa* (As), *Sorghum bicolor* (Sb) and *Triticum aestivum* (Ta) revealed a conserved AP2 domain consists of a three β-sheet strand and one α-helix strand running almost parallel to the β-sheet (Fig. 3). Two conserved functional amino acids (valine and glutamic acid) at the 14^{th} and 19^{th} residues were also observed in the DNA binding domain and thought to be crucial sites for the binding of DREB to the DRE core sequences (Liu et al. 1998). Ser/Thr-rich region adjacent to the AP2/ERF DNA-binding domain is considered to be responsible for phosphorylation of DREB proteins. Our study is in agreement with previous studies (Liu et al., 1998 ; Agarwal et al., 2006). The amino acids at 14^{th} position and 19^{th} position of the AP2/ ERF domain were highly distinctive as they distinguish the DREB (valine and glutamic acid) from the ERF (alanine and aspartic acid, respectively) gene classes (Agarwal et al. 2007). In DREB1 proteins, Ahktar et al. (2012) have reported DSAW (AP2 domain) and LWSY motifs (at the end of C'-terminal region). In this study, we could not find any such motifs in DREB2 of wheat. DREB/CBF genes containing AP2 domain have been reported from diverse plant species, including both monocots and dicots revealed same structural characteristic of *DREB* gene (Allen et al., 1998; Lata and Prasad 2011). Moreover, some residues in the β-sheet have been identified

Table 1. List of isolated *DREB* gene along with accession number.

S.No	Accession Number	*T. aestium* cultivars	Groups	References
1	HQ840651.1	AKW-381	LDT	Gawandel et al., 2004
2	HQ840652.1	HUW-234	LDT	-
3	HQ840654.1	WH-147	HDT	Dhanda and Sethi 2008
4	HQ840655.1	C-306	HDT	-
5	JF728296.1	HD-2932	LDT	-
6	JF728297.1	Raj 4037	LDT	-
7	JF728298.1	HD-2781	LDT	-
8	JF728299.1	NI-5439	HDT	Devi et al., 2011
9	JF728300.1	Raj-3765	MDT	-
10	JF728301.1	PBW-175	HDT	Devi et al., 2011
11	JN191706.1	MACS 2496	MDT	-
12	JN191707.1	HD 2733	MDT	-
13	JN191708.1	PBW 343	LDT	-
14	JN191709.1	HUW 468	MDT	-
15	JN191710.1	GW 322	MDT	-

*HDT- high drought tolerance, MDT-medium drought tolerance, LDT-low drought tolerance; -unpublished data

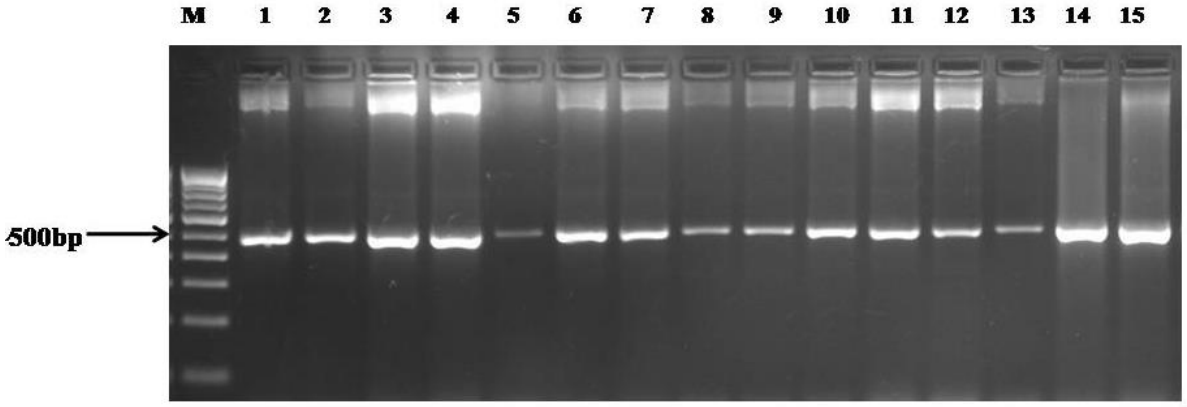

Fig 1. Gel electrophoresis of PCR product of DREB2 DNA in drought tolerant wheat genotypes amplified using specific primers. Lane 1= AKW381, Lane 2 = HUW 234, Lane 3 = WH 147, Lane 4 = C306, Lane 5 = HD2932, Lane 6 = Raj4037, Lane 7 = HD2781, Lane 8 = NI-5439, Lane 9 = Raj3765, Lane 10 = PBW175, Lane 11 = MACS2496, Lane 12 = HD2733 , Lane 13 = HUW468 , Lane 14 = GW322, and Lane15 = PBW343 M = 100bp DNA ladder, sizes of each fragment are shown

as key residues for DNA-binding activity of AP2/EREBP type protein. These data indicate that the structure of DREB family proteins is well conserved in plants, and also suggests these *TaDREB2* genes may function in a similar way.

Prediction of the DREB proteins folding state of T. aestivum

Based on ExPASy (http://www.expasy.org/) and some other softwares, physical and chemical properties such as number of amino acids, composition of deduced amino acid sequence, theoretical Mw, theoretical pI, and recombinant protein solubility (percent of insolubility) of *DREB* gene from *T. aestivum* were analyzed (Suppl. Table 1). Structural and biophysical studies are a prerequisite for determining protein solubility. The statistical solubility model in *Escherichia coli* predicts protein solubility assuming the protein is being over expressed in *E. coli* (Wilkinson et al. 1991). Results showed that the chance of insolubility for TaDREB2 protein when over expressed in *E. coli* ranged from 58.5% - 92.9%, respectively. The proteins were insoluble at pI and would be soluble both below and above this point, help in predicting the pH of the elution buffer system in purification process (Friedman et al. 2009) (Suppl. Table 1). The average value of theoretical pI of *T. aestivum* was greater than 9.0. Based on the calculated isoelectric point (pI), TaDREB2 was separated using isoelectric focusing method. The folded or unfolded structure of TaDREB2 protein has not been studied clearly till date. It is now possible to predict extended disordered regions in gene sequences with a high degree of precision (Prilusky et al., 2005). The percentage of disordered residue of TaDREB2 and AtDREB2 were 60.4% and 59.7%, respectively (Fig. 4). The results obtained in this study could be used to infer the local and general probability for the provided protein sequences, under specific conditions to fold. These differences in the folding state of TaDREB are our future research targets.

Phylogenetic analysis of DREBs

Phylogenetic analysis suggested a close relationship between DREB protein isolated from wheat and other crops like *Oryza sativa* (Os), *Zea mays* (Zm), *Sorghum bicolor* (Sb), *Glycine max* (Gm), *T. aestivum* (Ta) *Hordeum vulgare* (Hv), *Poa pratensis* (Pp), *Festuca arundinacea* (Fa), *Avena sativa* (As), and *A. thaliana* (At), *Nicotiana tabacum* (Nt), *Gossypium arboretum* (Ga) and *Brassica rapa* (Br). The results showed that isolated *TaDREB2* gene in this study forms a separate clade including barley and rice (Fig. 5). The discovery of these genes provided valuable information for understanding the whole DREB family of common wheat.

Table 2. Sum of the interaction involved in Protein-DNA docking. The hydrogen bonding and non bonded interaction were shown between nucleotide bases of DNA (GCC box) and amino acid residues of protein (DREB2).

	Interacting residues					Hydrogen bonds			
	Residue	Position		Residue	Position	M-M*	M-S*	S-S*	Non-bonded contacts
Residues:	DG	14	->	DC	13	0	4	0	6
Residues:	DT	16	->	DG	12	0	1	0	2
Residues:	DC	15	->	DG	12	0	2	0	8
Residues:	DG	14	->	DG	12	0	0	0	5
Residues:	DG	17	->	DA	11	0	0	0	7
Residues:	DT	16	->	DA	11	0	1	0	7
Residues:	DC	15	->	DA	11	0	0	0	1
Residues:	DG	17	->	DC	10	0	2	0	8
Residues:	DG	18	->	DC	10	0	1	0	2
Residues:	DG	18	->	DC	9	0	2	0	8
Residues:	DG	20	->	DG	8	0	0	0	6
Residues:	DC	19	->	DG	8	0	2	0	8
Residues:	DG	18	->	DG	8	0	0	0	1
Residues:	DG	20	->	DC	7	0	2	0	8
Residues:	DG	21	->	DC	7	0	1	0	2
Residues:	DG	21	->	DC	6	0	2	0	8
Residues:	DG	20	->	DC	6	0	1	0	1
Residues:	DT	23	->	DG	5	0	1	0	2
Residues:	DC	22	->	DG	5	0	4	0	6
Residues:	DG	21	->	DG	5	0	0	0	5
Residues:	DA	24	->	DA	4	0	0	0	7
Residues:	DT	23	->	DA	4	0	1	0	7
Residues:	DA	24	->	DT	3	0	1	0	7
Residues:	DT	3	->	TYR	96	0	0	0	3
Residues:	DA	4	->	ARG	80	0	0	0	5
Residues:	DA	4	->	ARG	73	0	0	0	8
Residues:	DT	3	->	ARG	73	0	0	0	9
Residues:	DT	3	->	ILE	72	0	0	0	3
Residues:	DT	3	->	GLU	71	0	0	0	2
Residues:	DA	4	->	ARG	61	0	1	0	5
Residues:	DT	3	->	ARG	61	0	0	0	3
Residues:	DT	3	->	GLY	59	0	0	0	10
Residues:	DT	3	->	ARG	58	0	0	0	13

* Main-chain donor to main-chain acceptor (MM), main-chain donor to side-chain acceptor (MS), side-chain donor to main-chain acceptor (SM) and side-chain donor to side-chain acceptor (SS), Deoxyadenosine (DA), Deoxythymidine (DT), deoxyguanosine (DG), deoxycytosine (DC)

The analysis implied that DREB subfamily can easily be classified on the basis of AP2 domain (Agarwal et al. 2007). It was also observed that AP2 domains can clearly distinct monocots from dicots. This functional conservation makes them important targets for crop improvement for abiotic stress tolerance through genetic engineering and plant breeding.

Expression profile analysis of TaDREB

As far as tissue-specific expression is concerned, only limited information is available in few plant species. Previous studies indicated that DREB1/CBFs were expressed in almost all tissues and organs. There were over 1, 2865, 282 ESTs from *T. aestivum* that have been deposited in GenBank and they were clustered into 56, 975 UniGene. Using a data mining tool known as Digital Differential Display (DDD), the expression profile was determined by analyzing the EST counts based on UniGene of *T. aestivum DREB* gene in the various tissues on the basis of transcript per million (Eujayl and Morris 2009). We executed gene expression analysis of the wheat DREB ESTs from nine different kinds of tissues. The DREB family genes from *T. aestivum* were detected in the callus, crown, flower, inflorescence, leaf, root, seed, sheath, and stem (Fig. 6). Transcripts of these DREB family genes were the most abundant in leaf tissue (17%), followed by callus and cell culture tissue (16%) and inflorescence (13%), and the least abundant in seed and stem tissue (7%). Some family *DREB* genes were detected in other tissues as well, callus (8%), crown (8%), flower and root (11%) (Fig.7). Recently, Rashid et al. (2012) have demonstrated the expression pattern of putative AP2/ERF family genes in as many as 12 different kinds of rice tissues using the RiceXPro database. Among the vegetative organs, the expression was found most abundant in root followed by leaf and stem. Wang et al. (2004) revealed that the expression was a little higher in leaf and stem and distinctly increased in root. Transcription analysis of rice in mature root tissue by massive parallel sequencing revealed previously unrecognized tissue-specific expression profiles and render an interesting platform to study the differential regulation of transcribed regions of root tissue (Kyndt et al. 2012). Using an experimental approach, Qin et al. (2007) highlighted that *OsDREB113* (a DREB subfamily gene) was expressed

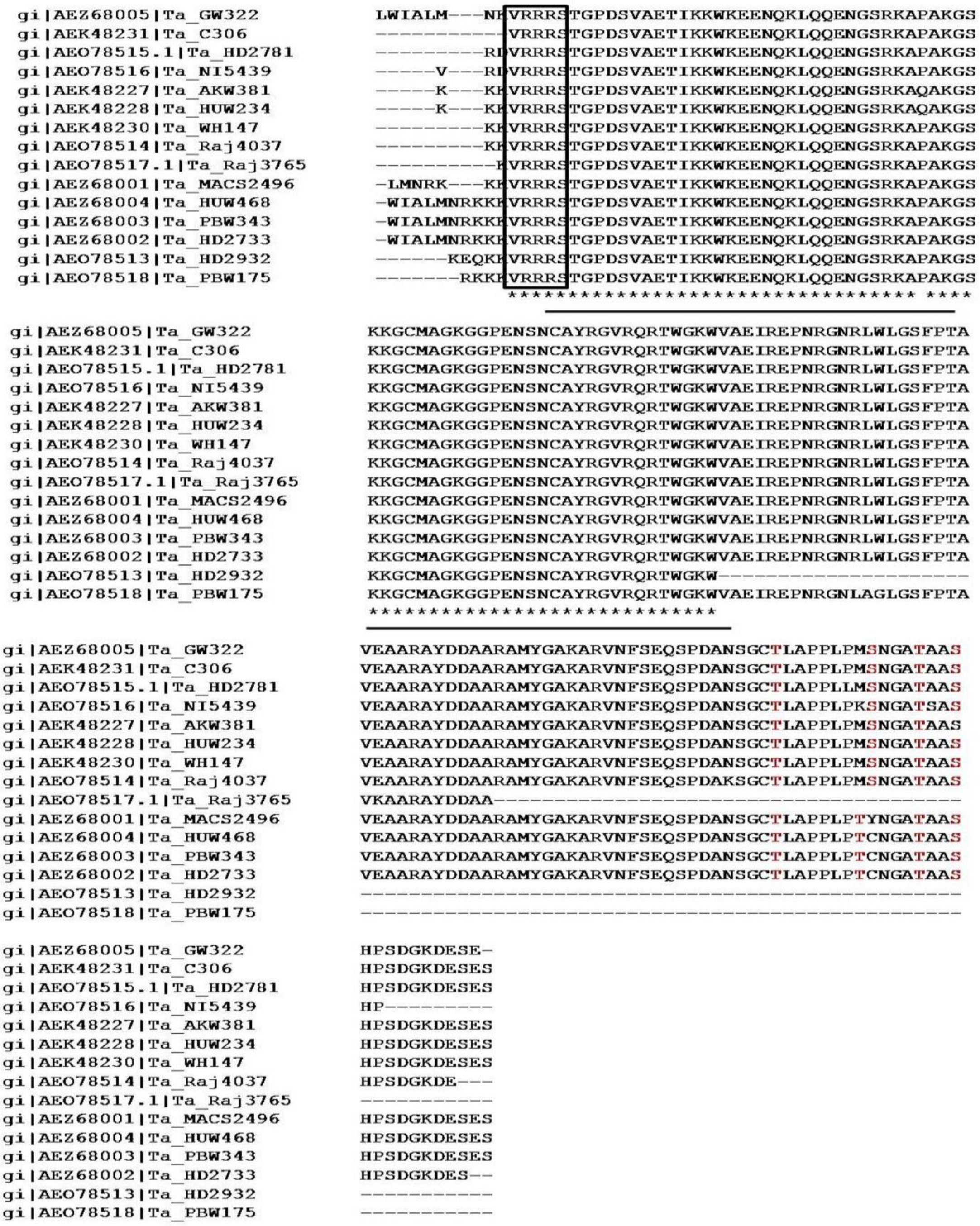

Fig 2. Analysis of the deduced amino acid sequence of the isolated TaDREB2 protein. The nuclear localization signal (VRRS) in the N-terminus is shown in bold box. The serine/threonine-rich regions are shaded in red. The AP2/ERF (57 amino acids) conserved domains of TaDREB2 are single over lined. The asterisk and underline indicate amino acids defined as AP2 domain and dashes indicate gaps in the amino acid sequences introduced to optimize alignment.

significantly in roots compared to leaves, shoots, growing points, mature seeds and other tissue of unstressed rice plants. The subfamily gene, OsAP211, is extremely prominent in shoot tips in mature rice and immature seeds at booting stage compared to other rice tissues (Gao et al., 2011). However, the *OsDREB1D* was not detected in any organs of rice with, or without stress treatments (Zhang et al. 2009). Similarly, Yang et al. (2009) revealed that *DmDREBa* and *DmDREBb* were accumulated more in leaves and stems than in roots and flower. Understanding *T. aestivum DREB2* functions will help to develop *T. aestivum* cultivars that may be adaptable to less favorable environmental conditions. Our results suggest that many wheat *DREB* genes have high sequence similarity and diverse roles as evidenced by analysis of sequences and their expression.

Homology modeling of TaDREB2

Homology modeling procedures are essential tools for conducting research when the experimental 3D-structure of the protein is not available (Sanchez and Sali 1997). The predicted 3D structure of TaDREB2 protein is shown in Fig. 7A

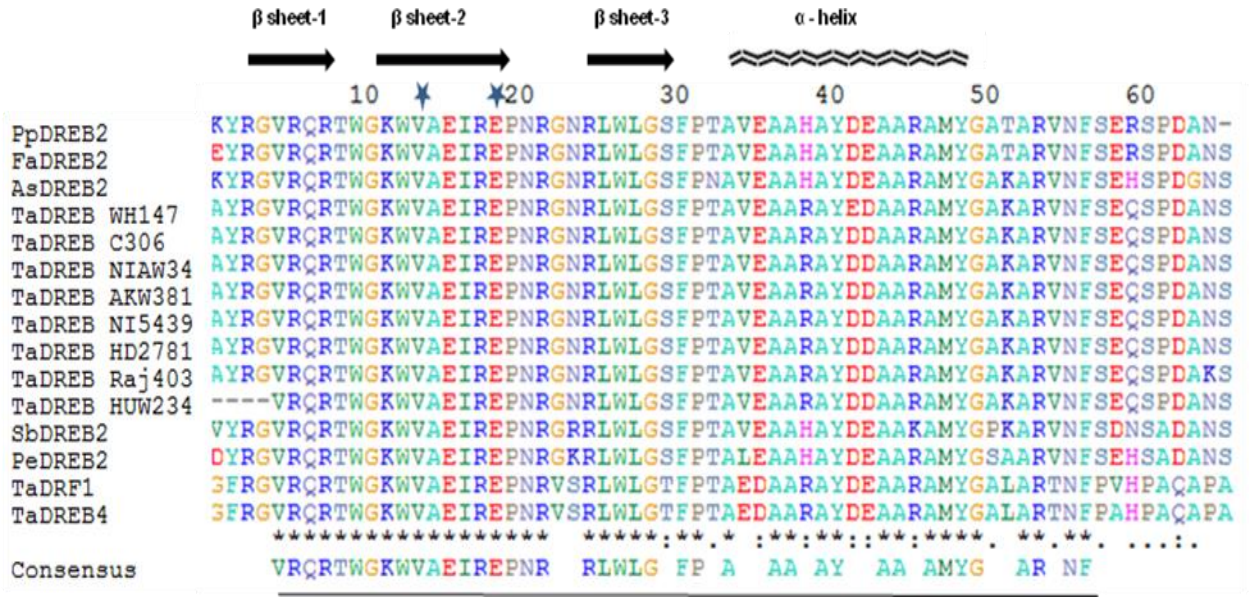

Fig 3. Features of AP2 domain: Multiple amino acid sequence alignment of highly conserved AP2 domains of DREB transcription factor in five plant species; *Poa pratensis DREB2* (Pp; Acc. No. AAS59530), *Festuca arundinacea DREB2 (*Fa, Acc. No. AAR11157), *Avena sativa DREB2* (As, Acc No. ABS11171), *Sorghum bicolor DREB2* (Sb, Accession No. AFI71292) and *Triticum aestivum DREB* (Ta, Acc. No. AY781345). The three β-sheets and a α-helix are marked above the corresponding sequences. Asterisks positions which have a single fully conserved residue; the colors indicate highly conserved positions; dots indicate weaker conserved in the amino acid sequences introduced to optimize alignment. The consensus sequence is at the bottom of the alignment.

Sequence identity between target and template (PDB: 1GCC) was 61%. The structure of the AP2 domain of *AtERF1* from *A. thaliana* (PDB code 1GCC) was taken as a template for the comparative modeling of the DNA-binding domain of TaDREB2 (Allen et al., 1998). Both template and predicted structures comprise of three stranded antiparallel *β*-sheet followed by *α*-helix and relatively unstructured C'- terminal. In *B. napus*, three stranded anti parallel *β*-sheet of DREB GBD contains strand1 (Arg58- Arg61), strand2 (Trp68- Glu74) and strand3 (Asn79- Phe86) while *α*-helix ranges from Thr88 to Gly105 residues. Despite a relatively less primary structure homology between *A. thaliana* and *T. aestivum* DREB2, their tertiary structure resembled to each other considerably as suggested by the root mean square deviation (RMSD) of Cα and the backbone was 0.84 Å and 0.88Å (Suppl. Fig. 2). Hussain et al. (2009) also noticed such differences or similarities among functionally related proteins when the studies conducted on DNA-Photolyase. In addition to this, only a single residue were in allowed and another one in generously allowed and disallowed regions in the Ramachandran plot (Suppl. Fig. 3), respectively and bear the free energy of –5312.458 KJ/mole, suggesting the structural and thermodynamic stability of the proposed structure of the TaDREB2. The above mentioned results suggest reliability of the modeling strategies used in the present study. The final protein structure was deposited in PMDB and is available under ID: PM0078617. To the best of our cognizance, no such 3-D model has been generated for TaDREB2 and this model will be used to understand how the activity changes.

Docking studies

DNA-binding domain, which is highly positively charged, docked with the negatively charged DNA double helix. *T. aestivum* DREB2 binds with major groove of DNA via its three stranded anti parallel *β*-sheet. Similarly, the N'- to C'-terminals of the protein corresponds to the 5'- to 3'- terminals of the DNA coding strand. Previously, the Protein-DNA interactions were established through *α*-helices of zinc finger containing proteins (Dutnall et al. 1996; Tan et al. 2003). Ligplot analysis depicted the involvement of eight amino acids in Protein-DNA interaction (Table 2). Amino acid of DNA-binding domain as an acceptor (TYR96, ARG80, ARG73, ILE72, GLU71, ARG61, GLY59, ARG58 and GLN26) formed hydrogen bond and hydrophobic contact with nitrogenous base Deoxyadenosine (DA4) and Deoxythymidine (DT3) of DNA double helix having GCC-box as donor (Table1). The structure was stabilized by extensive sixty one hydrophobic contacts (Table 2). This showed that the complex was stabilized by both hydrogen bonding and hydrophobic interactions. Out of several amino acids, ARG 61 were found to be crucial. ARG 61 donates hydrogen to DA 4 and form hydrogen bond with bond length 2.49Å. It also forms three hydrophobic contacts with DT 3 (Table 2) and we speculate that these residues were involved in the structural stability of the molecule. Any alteration in those residues may cause changes in the configuration of the protein. Thus, these outcomes suggest that the interaction complex is feasible and useful for understanding the binding mechanism between protein binding domain and DNA GCC-box of DREB2 protein in wheat. Electrostatic surface-to-surface contacts reveal complete accommodation of both molecules into each other. Electrostatic surface potential of DNA-binding domain showed large positive and small negative regions (Fig. 7B). The DNA-binding domain complemented the negatively charged DNA double helix (Fig. 8). It is noteworthy that all the earlier and later mentioned residues were located in three stranded anti parallel sheets. Our findings are in tune with earlier studies

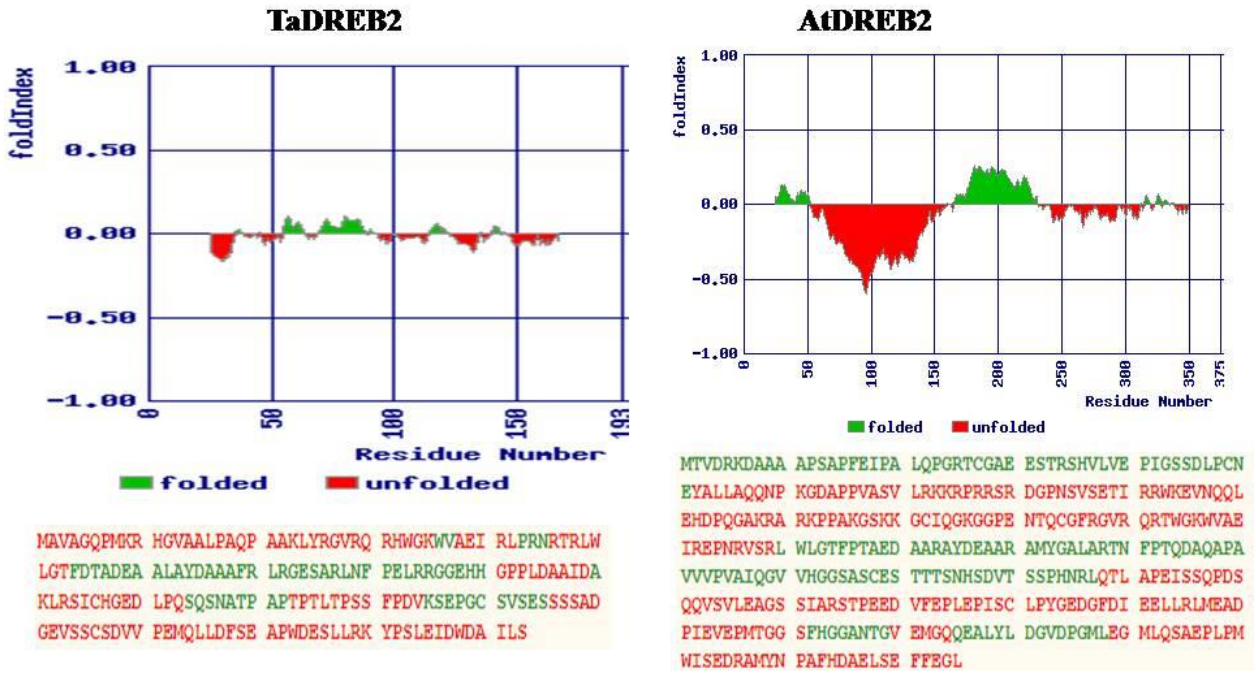

Fig 4. FoldIndex plotted with window size 51, the plot showed that TaDREB2 and AtDREB2 contain unstructured regions in its native state. Positive and negative numbers represent ordered and non-ordered protein, respectively. Amino acids suggested as being ordered is shown in green and unordered in red respectively.

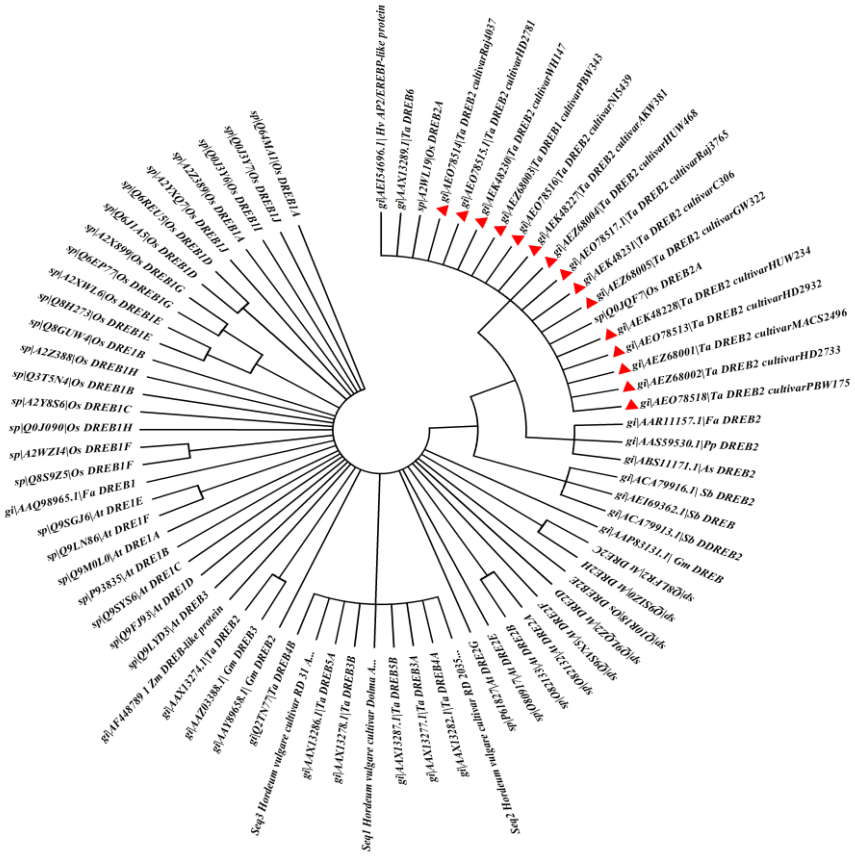

Fig 5. Phylogenetic analysis of AP2/ERF domains of published AP2/EREBP proteins in the NCBI database. The phylogenetic tree was generated by MEGA 4.0 software. Branch lengths indicate the distance. Isolated *TaDREB2* are indicated by a red triangle.

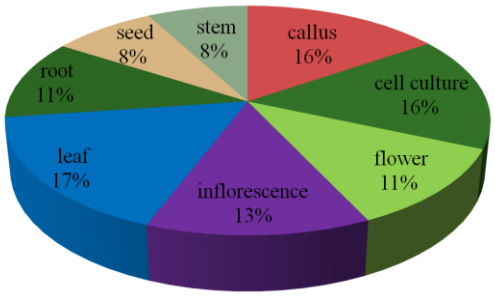

Fig 6. Distribution of *T. aestivum* DREB2 family in various tissues. The expression profile was determined by analyzing the EST counts of eight different kinds of tissues.

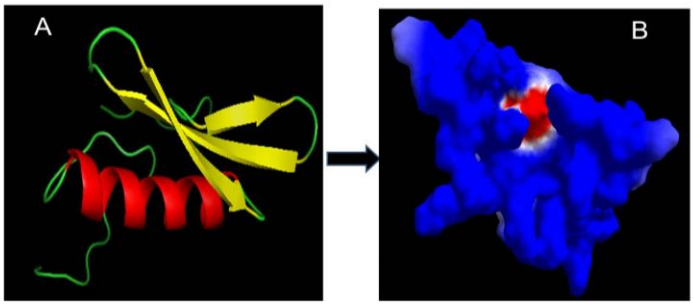

Fig 7. Homology modeling of TaDREB2 protein (A) 3D structure of the full-length TaDREB2 protein (Accession no AAX13274.1). Shown as a cartoon, three strands of anti parallel β sheet are respectively represented by yellow arrows. The DNA-binding domain which spans from 56 to 113 amino acid residues in the peptide chain is colored blue (B) the electrostatic surface potential is shown in a color gradient from positive (blue) to negative (red).

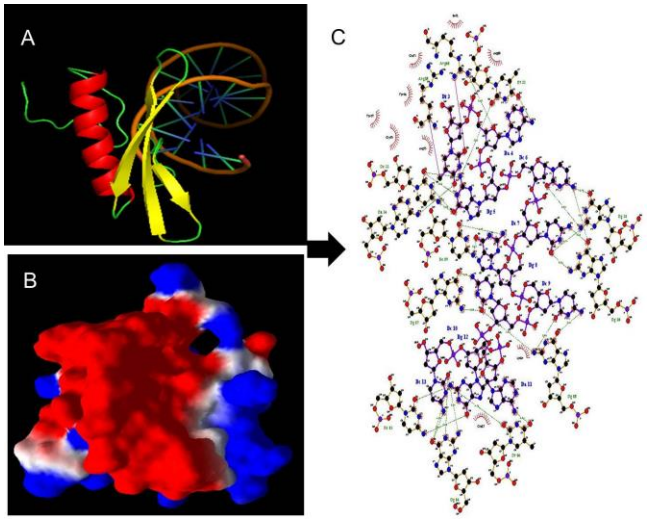

Fig 8. Results showing interaction between 3D structure of TaDREB2 protein with GCC box (A) DNA binding domain docked with DNA containing GCC box (B) The electrostatic surface potential is shown in a color gradient from positive (blue) to negative (red). The DNA-binding domain is highly positively charged which complements the negatively charged DNA double helix. (C) Ligplot shows the contacts between the residues of DNA binding protein and DNA GCC box and their bond length.

(Allen et al., 1998). Overall, the residue interactions directly covered six base pairs in the conserved AGCCGCC sequence, validated the presence of DNA binding domain in the *T. aestivum* DREB2. Previous studies suggested the presence of with Val and Glu for the effective interaction of *β*-sheet DNA (Sakuma et al. 2002; Allen et al. 1998), however, no such interactions were observed in the present study. The structure determination of the gene and the DNA-binding domain may provide the framework for understanding the functions of DREBs at the molecular level.

Materials and Methods

Plant materials

In total, fifteen diverse Indian bread wheat genotypes were selected, as these were evaluated for drought tolerance (Table 1). The seeds were obtained from the germplasm Unit, Directorate of Wheat Research, Karnal, Haryana. List of genotypes used in this study is summarized in Table 1. The seeds were germinated separately under aseptic conditions. One week old seedlings were harvested and used for DNA isolation.

PCR amplification

About 200 mg of leaf tissue was ground to a fine powder in liquid nitrogen and DNA was extracted using modified CTAB method (Sumer et al., 2003). Primers were designed by Primer 2 software and consensus of *DREB2* genes were aligned from NCBI GenBank. Specific primer pairs (PsDREB-F 5'-TATGGATTGCCTTGATGAACA-3' and PsDREB-R 5'-GACTCCGATTCATCCTTCCC-3) were designed and PCR amplification was performed in a total volume of 25 µl volume containing 80 ng of genomic DNA, 2.5 µl of 10X PCR buffer, 200 µM of each dNTP, 0.2 µM of each primer and 1.0 unit of *Taq* DNA polymerase (Toyobo, Japan). The PCR amplifications were carried out using My® Thermal cycler (Bio-Rad, USA) as follows: initial denaturation at 94°C for 5 min. followed by 35 amplification cycles (94°C for 1 min, 58°C for 1 min and 72°C for 1 min) and final extension at 72°C for 10 min. The amplified products were electrophoresed on 2% (w/v) agarose gel, stained with ethidium bromide and visualized under Gel Documentation (Alpha Innotech, USA) system.

PCR product purification and sequencing

PCR products were excised from the gel and purified using a PCR purification kit (Qiagen, USA) according to manufacturer's instructions. The purified products were then sequenced. The sequencing was determined by using the Big Dye Terminator kit (ABI Foster City, USA) on an ABI 310 DNA analyzer.

Database search and Bioinformatics analysis

We make BLAST (http://blast.ncbi.nlm.nih.gov/) search to align our isolates with the already existing *DREB* genes in the NCBI. In order to study the sequence variation at protein level, the nucleotide sequences were translated into amino acid using an online EMBOSS transeq program (http://www.ebi.ac.uk/Tools/st/emboss_transeq/). The amino acid sequences were searched for homology with cereal and leguminous plant. The sequence analyzed by a consensus approach using ExPASy proteomics server of the Swiss

Institute of Bioinformatics (SIB) (http://cn.expasy.org/) for physicochemical and secondary structure characterization of deduced amino acid sequence of *DREB* gene from *T. aestivum*. Recombinant protein solubility was predicted using the statistical model assuming the protein is being over expressed in *Escherichia coli* at the web site of the University of Oklahoma (http://biotech.ou.edu/) (Harrison, 2000). The isolated DREB sequences were subjected to programs INTERPROSCAN (http://www.ebi.ac.uk/Tools/pfa/iprscan/) and SMART (http://smart.embl-heidelberg.de/) was employed to detect conserved domains. The folding states of DREB protein sequences from wheat and barley were predicted by the FoldIndex program (http://bioportal.weizmann.ac.il/fldbin/findex)

Phylogenetic tree construction

A systematic phylogenetic analysis was carried out based on the similarities of AP2 domains in the proteins isolated from *Arabidopsis*, rice, maize, wheat, sorghum, barley, *Festuca*, *Poa*, *Festuca*, *Avena* and *Brassica* species. The protein multiple sequence alignment was performed using MUSCLE (Edgar, 2004) and T-coffee (Notredame et al., 2008) softwares. The final consistency of the alignment was obtained through manual editing. Phylogenetic trees were constructed using the neighbor-joining (NJ) method and the pictures of phylogenetic trees were drawn with program MEGA4 (http://www.megasoftware.net/mega.html) (Tamura et al., 2007).

Analysis of expression profile

The expression profile was determined by analyzing the EST counts based on UniGene of *T. aestivum* for the various tissues (http://www.ncbi.nlm.nih.gov/UniGene). The EST pool was derived from a variety of different cDNA libraries, capturing genes expressed in different tissue types and developmental stages or expressed during pathogen-elicited responses. Libraries were normalized, the subtracted expression profile was calculated by
Transcript per million=Gene EST/ ∑Total EST in the pool

Protein modeling and structure analysis:

Automated homology model building of the wheat DNA-binding domain was performed using the protein structure modeling program modeler 9v8 (http://www.salilab.org/modeller/), which model protein tertiary structure by spatial restraints satisfaction using a GCC-box binding domain from *Arabidopsis thaliana* (PDB code 2GCC) as a template. The model with the lowest discrete optimized protein energy (DOPE) score was selected. Finally, the model was minimized by the energy minimization performed by the steepest descent method GROMACS (Hess et al., 2008). Ramachandran plot statistics was used to evaluate the best model. The root mean square deviation (RMSD) values were calculated using the SUPERPOSE program (Maiti et al., 2004) by fitting the carbon backbone of the predicted model onto the template structure.

Protein –DNA docking

The final 3D structure of *T. aestivum* DREB2 protein was submitted to the Protein Model Database (PMDB) (Castrignano et al., 2006). The refined protein model (TaDREB2) was used to study its ligand binding mechanism. Docking analysis was performed by the HEX molecular docking tool. The selected model of DREB2 was docked against the GCC-box containing double helical DNA. The atomic coordinates of GCC-box binding domain from *Arabidopsis thaliana* (PDB code 1GCC) were retrieved from Protein Data Bank (PDB) (Berman et al., 2000). One hundred models were constructed with defined global scoring covering electrostatic, hydrophobic; solvation energy and side chain contacts attribute, hydrophobic and electrostatic restrains. Out of these the best hydrophobic and electrostatic models were chosen for more detailed analysis. Hydrogen and hydrophobic interactions between these subunits were analyzed by Ligplot (Wallace et al., 1995). Figure representation (Fig. 7 and Fig. 8) was generated with Pymol (DeLano, 2002).

Conclusion

In the present study, we have isolated and computationally characterized *TaDREB2* gene from 15 diverse Indian bread wheat cultivars. TaDREB2 transcription factor is one of the most promising candidate genes, involved in conferring drought tolerance. Furthermore, multiple alignment of corresponding amino acid sequences showed conserved AP2 domain. It is apparent that DREB2 transcription factor was more conserved within species, phylogenetic analysis of TaDREB2 indicated that they were conserved across the dicot and monocot species. In addition, we discovered an expression profile pattern of *T. aestivum* DREB2 family in various tissues using the available EST information. The structure determination of the *TaDREB2* gene and the DNA binding domain may provide the foundation for understanding the functions of DREBs at the molecular level. To the best of our knowledge this is the first report of molecular modeling and protein–DNA interaction study of a *DREB2* gene in wheat.

Acknowledgements

The authors thankfully acknowledge the financial support for Indian Council of Agricultural Research (No. DWR/RP/10-5.3), and Grant-in-Aid under LBSA scheme (No.7/33/FFC) Government of India, New Delhi. This is DWR contribution No. 59/DWR/2012.

References

Agarwal P, Agarwal PK, Nair S, Sopory SK, Reddy MK (2007) Stress-inducible DREB2A transcription factor from *Pennisetum glaucum* is a phosphoprotein and its phosphorylation negatively regulates its DNA. Mol Genet Genomics 277: 189-98

Agarwal PK, Agarwal P, Reddy MK, Sopory SK (2006) Role of DREB transcription factors in abiotic and biotic stress tolerance in plants. Plant Cell Rep 25: 1263-1274

Akhtar M, Jaiswal A, Taj G, Jaiswal JP, Qureshi MI, Singh NK (2012) DREB1/CBF transcription factors: their structure, function and role in abiotic stress tolerance in plants. J Genet 91:385-95

Allen MD, Yamasaki K, Ohme-Takagi M, Tateno M, Suzuki M (1998) A novel mode of DNA recognition by a β-sheet revealed by the solution structure of the GCC-box binding domain in complex with DNA. EMBO J 17:5484-5496

Bartels D, Sunkar R (2005) Drought and salt tolerance in plants. Crit Rev Plant Sci 24: 23-58

Berman HM, Westbrook J, Feng Z, Gililand G, Bhat TN, Weissig H, Shindylaov IN, Bourne PE (2000) The Protein Data Bank. Nucleic Acid Res 28: 1235-1242

Boyer JS (1982) Plant productivity and environment. Science 218: 443-8

Bray EA, Bailey-Serres J, Weretilnyk E (2000) Responses to abiotic stresses. In: Gruissem W, Buchannan B, Jones R. (eds), Biochemistry and Molecular Biology of Plants, American Society of Plant Physiologists, Rockville MD, p. 1158.

Castrignano T, De Meo PD, Cozzetto D, Talamo IG, Tramontano A (2006) The PMDB Protein Model Database. Nucleic Acids Res. 34: 306-309.

Chen WJ, Zhu T (2004) Networks of transcription factors with roles in environmental stress response. Trends Plant Sci 9: 591-596

Cong L, Chai TY, Zhang YX (2008) Characterization of the novel gene *BjDREB1B* encoding a DRE-binding transcription factor from *Brassica juncea*. Biochem Biophy Res Comm 371: 702-706

DeLano W L (2002) PyMOL molecular graphics system. http://www.pymol.org

Devi R, Karu N, Gupta AK (2011) Potential roles of antioxidant enzymes in depinditing drought tolerance of wheat. Indian J Biochem Biophy 49: 257-265

Dhanda SS, Sethi GS (2008) Drought tolerance among 30 cultivars of wheat for water relation parameters. Agric Sci digest 28: 182-185

Dutnall RN, Neuhaus D, Rhodes D (1996) The solution structure of the first zinc finger domain of SWI5: a novel structural extension to a common fold. Structure 4: 599-611

Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32: 1792-1797

Egawa C, Kobayashi F, Ishibashi M, Nakamura T, Nakamura C, Takumi S (2006) Differential regulation of transcript accumulation and alternative splicing of a DREB2 homolog under abiotic stress conditions in common wheat. Genes Genet Syst 81:77-91

Eujayl I, Morris C (2009) Identification of differentially expressed UniGenes in developing wheat seed using Digital Differential Display. J Cereal Sci 49: 316-318

Friedman DB, Hoving S, Westermeier R (2009) Isoelectric focusing and two-dimensional gel electrophoresis. Methods Enzymol 463:515-40

Galiba G, Vagujfalvi A, Li C, Soltesz A, Dubcovsky J (2009) Regulatory genes involved in the determination of frost tolerance in temperate cereals. Plant Sci 176:12-19

Gao FC, Xiong AS, Peng RH, Liu JG, Cai B, Yao QH (2009) Isolation and characterization of a novel AP2/EREBP-type transcription factor OsAP211 in *Oryza sativa*. Biol Plantarum 53:643-649

Garg N, Pundhir S, Prakash A, Kumar A (2008). PCR Primer Design: DREB Genes. J Comput Sci Syst Biol 1: 21-40

Gawandel ND, Mahurkar DG, Dudhe MY, Ratnaparkhi RD (2004) In vitro screening of wheat genotypes for drought tolerance on the basis of callus induction. Journal of Soils and Crops 14:292-296: 292-296.

Hao DY, Ohme-Takagi M, Sarai A (1998) Unique mode of GCC box recognition by the DNA-binding domain of ethylene-responsive element-binding factor (ERF domain) in plant. J Biol Chem 273: 26857-26861

Harrison RG (2000) Expression of soluble heterologous proteins via fusion with NusA protein. Innovations11: 4-7

Hess B, Kutzner C, Van Der Spoel D, Lindahl E (2008) GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. J Chem Theory Comput 4: 435-447

Hussain MS, Qamarunissa S, Raza JA, Qureshi A, Wajid SA, Rasool (2009) DNA Photolyase: Possible explanation of low sunlight inactivation of Enterococci. Biologia Sec Cell Mol Biol 64: 852-858

Kyndt T, Denil S, Haegeman A, Trooskens G, De Meyer T, Van Criekinge W, Gheysen G. (2012) Transcriptome analysis of rice mature root tissue and root tips in early development by massive parallel sequencing. J Exp Bot 63: 2141-2157

Lata C, Prasad M (2011) Role of DREBs in regulation of abiotic stress responses in plants. J Exp Bot 62: 4731-4748

Liu Q, Kasuga M, Sakuma Y, Abe H, Miura S, Goda H, Yamaguchi-Shinozaki K, Shinozaki K (1998) Two transcription factors, DREB1 and DREB2, with an EREBP/ AP2 DNA binding domain separate two cellular signal transduction pathways in drought- and low-temperature-responsive gene expression, respectively, in *Arabidopsis*. The Plant Cell 10: 391-406

Maiti R, Van Domselaar GH, Zhang H, Wishart DS (2004) SuperPose: a simple server for sophisticated structural superposition. Nucleic Acids Res 32: W590-W594

Notredame C, Higgins DG, Heringa J (2008) T-Coffee: A novel method for fast and accurate multiple sequence alignment. J Mol Biol 302: 205-17

Prilusky J, Felder CE, Zeev-Ben-Mordehai T, Rydberg EH, Man O, Beckmann JS, Silman I, Sussman JL (2005) FoldIndex: a simple tool to predict whether a given protein sequence is intrinsically unfolded. Bioinformatics 121: 3435-3438

Qin F, Kakimoto M, Sakuma Y, Maruyama K, Osakabe Y, Tran LS, Shinozaki K, Yamaguchi-Shinozaki K (2007) Regulation and functional analysis of *ZmDREB2A* in response to drought and heat stresses in *Zea mays* L. Plant J 50:54-56

Rashid M, Guangyuan H, Guangxiao Y, Hussain J, Xu Y. (2012) AP2/ERF transcription factor in rice: Genome-wide canvas and syntenic relationships between monocots and eudicots. Evol Bioinf 8: 321-355

Sakuma Y, Liu Q, Dubouzet JG, Abe H, Shinozaki K, Yamaguchi-Shinozaki K (2002) DNA-binding specificity of the ERF/AP2domain of *Arabidopsis* DREBs, transcription factors involved in dehydration- and cold-inducible gene expression. Biochem Biophys Res Commun 290: 998-1009

Sanchez R, Sali A (1997) Advances in comparative protein-structure modeling. Curr Opin Struct Biol 7: 206-214

Shen YG, Zhang WK, He SJ, Zhang JS, Liu Q, Chen SY (2003) An EREBP/AP2-type protein in *Triticum aestivum* was a DRE binding transcription factor induced by cold, dehydration and ABA stress. Theor Appl Genet 106:923-930

Shinozaki K, Yamaguchi-Shinozaki K (2000) Molecular responses to dehydration and low temperature: Differences and cross talk between two stress signaling pathways. Curr Opin Plant Biol 3: 217-223

Sumer A, Ahmet D, Gulay Y (2003) Isolation of DNA for RAPD analysis from dry leaf material of some Hesperis L. Specimens. Plant Mol Biol Reptr. 21: 461a-461f.

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol 24: 1596- 1599

Tan S, Guschin D, Davalos A, Lee YL, Snowden AW, Jouvenot Y, Zhang HS, Howes K, McNamara AR, Lai A, Ullman C, Reynolds L, Moore M, Isalan M, Berg LP,

Campos B, Qi H, Spratt SK, Case CC, Pabo CO, Campisi J, Gregory PD (2003) Zinc-finger protein-targeted gene regulation: genome wide single-gene specificity. Natl Acad Sci USA 100:11997-12002

Wallace AC Laskowski RA, Thornton JM (1995) LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. Protein Eng 8:127-134

Wang SX, Wang ZY, Peng YK (2004) Dehydration responsive element binding (DREB) transcription activator and its function in plant tolerance to environment stress. Plant Physiol Commun 40: 7-13

Wilkinson DL, Harrison RG (1991) Predicting the solubility of recombinant proteins in *Escherichia coli*. Nature Biotech 9: 443-448

Yamaguchi-Shinozaki K, Shinozaki K (2005) Organization of *cis*-acting regulatory elements in osmotic- and cold-stress-responsive promoters. Trends Plant Sci 10: 88-94

Yamaguchi-Shinozaki K, Shinozaki K (2006) Transcriptional regulatory networks in cellular responses and tolerance to dehydration and cold stresses. Annu Rev Plant Biol 57: 781-803

Yang Y, Wu J, Zhu K, Liu L, Chen F, Yu D (2009) Identification and characterization of two chrysanthemum (*Dendronthema moriforlium*) *DREB* genes, belonging to the AP2/EREBP family. Mol Biol Rep 36:71-81

Zhang Y, Chen C, Jin XF, Xiong AS, Peng RH, Hong YH, Yao QH, Chen JM (2009) Expression of a rice DREB1 gene, OsDREB1D, enhances cold and high-salt tolerance in transgenic *Arabidopsis*. BMB Rep 42: 486-492

Zhang Z, Yao W, Dong N, Liang H, Liu H, Huang R (2007) A novel ERF transcription activator in wheat and its induction kinetics after pathogen and hormone treatments. J Exp Bot 58: 2993-3003

Zhao-Shi X, Zhi-Yong N, Zhi-Yong L, Lian-Cheng L, Ming C, Dong-Yao G, Xiu-Dao Y, Pei L, You-Zhi M (2009) Isolation and functional characterization of HvDREB-a gene encoding a dehydration-responsive element binding protein in *Hordeum vulgare*. Plant Res 122: 121-130

Zhu JK (2002) Salt and drought stress signal transduction in plants. Annu Rev Plant Biol 53: 247-273