Full Length Research Paper

Isolating Barley (*Hordeum vulgare* L.) B1 Hordein Gene Promoter and Using Sequencing Analaysis For The Identification of Conserved Regulatory Elements By Bioinformatic Tools

Kobra Nalbandi[1], Bahram Baghban Kohnehrouz[2]*, Khalil Alami Saeed[1] and Ashraf Gholizadeh[3]

[1]Ramin Agricultural and Natural Resources University, Mollasani, Ahwaz, Iran.
[2]Department of Plant Breeding and Biotechnology, University of Tabriz, Tabriz, Iran.
[3]Research Institute for Fundamental Sciences (RIFS), University of Tabriz, Tabriz, Iran.

Gene expression is a complex multi-step process. For the efficient expression of foreign genes in plants, it is essential to optimize every step of the process for the plant machinery, which includes choosing suitable promoters. Promoters play the most important role in determining the temporal and spatial expression pattern and transcript level of a gene. Some strong constitutive promoters, such as cauliflower mosaic virus 35s promoter, are widely used in plant genetic engineering research. However, the expression levels of the foreign genes in all tissues are often unsatisfied, but some of such problems can be may solved by using a strong seed-specific promoter to restrict gene expression to the seed only. The aim of this article was to characterize a B1 hordein-specific promoter. The promoter region of B1 hordein gene was isolated from the genomic DNA of Walfajre and Alger barley by polymerase chain reaction. Sequence analysis showed that the cloned fragment B hordein promoter (BHP) - contained motifs like TATA box, (CA)n box, ACGT motif, AAAG motif, GCN4-like motif and E-box, which constituted the seed-specific promoter activity. It was therefore concluded that B1 hordein promoters can be used to engineer and subsequently study stable seed specific gene expression in barley, and potentially to modify barley seeds through genetic engineering.

Key words: Barley, Seed Specific Promoter, Motif, Hordein.

INTRODUCTION

In most cereal species, the major seed storage proteins are prolamins, a complex group of alcohol-soluble polypeptides. The prolamin proteins in barley endosperm are usually termed hordein due to their high levels of proline and glutamine (Piston et al., 2004). The four main groups of hordein, namely the B, C, D and Y hordeins, account for approximately 50% of the total protein in the mature grain. Hordeins B, C and D are specified by separate compound genetic loci on chromosome 5 designated Hor-2, Hor-1 and Hor-3, respectively (Blake et al., 1982; Jensen et al., 1980; Shewry et al., 1980, 1983). Bhordeins are sulfur-rich prolamins, which account for 70 to 80% of the total hordein fraction in barley. The genes encoding hordein seed storage proteins are under seed-specific developmental regulation. Synthesis of these proteins is regulated at the transcription level and continues until it comprises of up to 60 to 80% of the total protein in mature seeds (Evanns et al., 1984; Gatehouse et al., 1986).

Promoters used in biotechnology are of different types according to the intended type of control of gene expression. Tissue-specific or development stage

---

*Corresponding author. Email: Bahramrouz@yahoo.com. Tel: +98-411-3392031. Fax:+98-411-3356003.

Abbreviations: PCR, polymerase chain reaction; **BHP,**B Hordein Promoter.
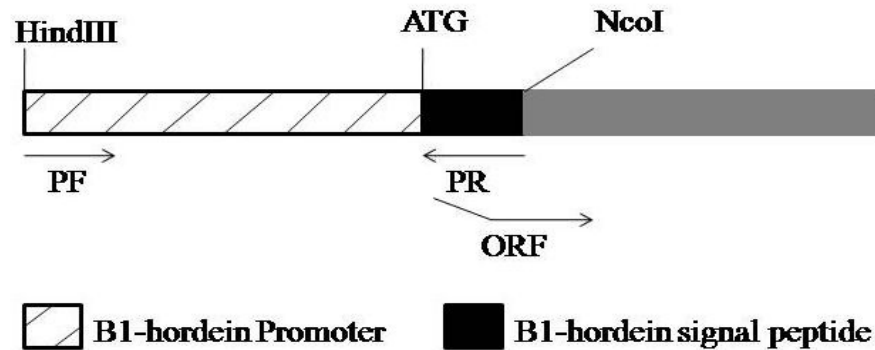
Figure 1. Schematic structure of B1 hordein gene promoter.

specific promoters could regulate the expression of a gene in specific tissue(s) or at certain stages of development. Thus, seed specific promoters provide an excellent system for studying the expression control of plant genes and are used in developing "transgenic" for improving seed quality. This study reports the isolation and charac-terization of a promoter sequence of B1 hordein in barley cultivated in Iran and presents a comparative analysis bybioinformatics tools.

MATERIALS AND METHODS

Hordeum vulgare L. cv. Walfajre and Alger seeds (Kindly provided by professor M.Mogadam) were grown in plastic trays at a glass house until two to three leave stages in order to extract their total cellular genomic DNA. *Escherichia coli* strain DH5α, was used for cloning purposes. The pGEMT-Easy PCR cloning kit and *Taq* DNA polymerase along with restriction enzymes were purchased from Promega and Takara corporations, respectively. The gel and plasmid DNA extraction kits were provided from Bioneer Corporation. The sequencing of the promoters was carried out by Millegen (Labège, France).

Extracting total DNA of barley

The total DNA was extracted from fine powder of 24 h dark grown leaf blades by liquid nitrogen (N_2) after removing midribs using cetyltrimethylammonium bromide (CTAB) extraction method as described by Saghai-Maroof (1984). Its quality and quantity were determined by 0.8% agarose gel electrophoresis and its concentration was adjusted on 50 ng/μL.

Designing the primer

The sequence of *H. vulgare* B1 hordein gene (GeneBank accession number, X87232) consisting of a promoter and coding regions of B1 hordein gene, were retrieved from GeneBank in order to design a specific primer for the promoter region and signal peptide of B1 hordein gene (Figure 1), forward primer of 5′AT*AAGCTT*GTCGAGAAGAACCGTCCAC3′ and reverse primer of 5′TG*CCATGG*TACTTGTTGCCGCAATG 3′ were designed using Primer3 Online software and their validity were confirmed by Oligo5 and NTC tools. The restriction enzyme sites of *Hind*III and *Nco*I were added to the 5′ end of the forward and reverse primers for

further subcloning, respectively.

PCR amplification

The total genomic DNA was used as a template for the amplification of promoters at the concentration of 50 ng/μL. The amplification program was done as the primary denaturing using 94°C for 5 min, followed by 35 cycles of 94°C for 50 s, 63°C for 50 s and 72°C for 1 min, and a final extension step at 72°C for 7 min. The amplified DNA products were separated by 0.8% agarose gel electrophoresis and extracted using the DNA gel extraction kit. The concentration of the eluted fragment was adjusted to 50 ng/μL and ligation reaction was done at 4°C for 24 h for cloning in the *p*GEMT-Easy vector. The *E. coli* competent cells were used for transformation by 5 μL of ligation reaction and the transformants were selected on white-blue test media containing antibiotic ampicillin, isopropyl β-D-1-thiogalactopyranoside (IPTG) and 5-bromo-4-chloro-3-indolyl-beta-D-galacto-pyranoside (X-gal). The plasmid DNA was extracted from the positive colony PCR of white colonies using the extraction kit. The cloned fragments in *p*GEMT-Easy were confirmed by reamplification and restriction enzymes of *Hind*III and *Nco*I. The obtained clones harboring putative promoters were sequenced by a MILLGENE company and followed by *in silico* analysis which was done using any bioinformatics tools.

RESULTS AND DISCUSSION

Cloning seed-specific promoter fragment ProB-Hor.A and ProB-Hor.W

The promoter fragments ProB-Hor.A and ProB-Hor.W were obtained by PCR amplification via primers PF and PR, respectively. A 518-nt-long sequence of the barley B hordein promoter was PCR amplified, the result of which is shown in Figure 2. This fragment was ligated to pGEMT-Easy in order to obtain pT-ProB-Hor.W and pT-ProB-Hor.A sequences. The procedure to construct the pT- ProB-Hor.A and pT- ProB-Hor.W is shown in Figure 3. The cloned fragments in pGEMT-Easy were further confirmed by PCR and digested by restriction enzymes, *Hind*III and *Nco*I, as shown in Figure 4. Individual digestion with *Nco*I and *Hind*III enzymes indicated the fragment to be 3015 bp, which included the 518 bp B
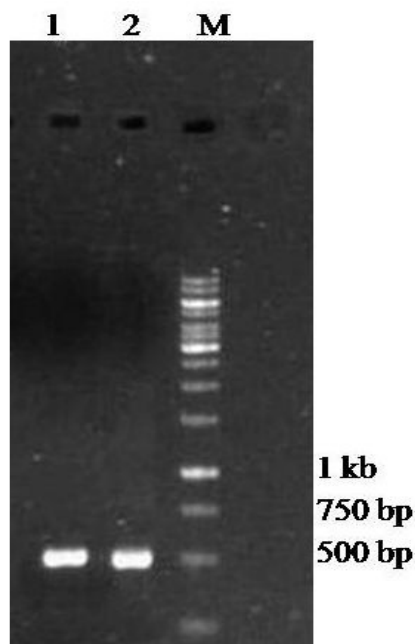
Figure 2. PCR product 1, Walfajre; 2, Alger; M, DNA marker.

hordein promoter. Double digestion with *NcoI* and *HindIII* enzymes indicated the 518 bp B hordein promoter fragment in the vector.

Sequencing analyses of the seed-specific promoter fragments ProB-Hor.A and ProB-Hor.W

To find regulatory elements in promoter sequences, the PLANTCARE (http://bioinformatics.psb.ugent.be/webtools/plantcare/html) and Softberry (http://linux1.softberry.com/berry.phtml) software were applied. Analysis of the sequences of the seed-specific promoter fragments ProB-Hor.A and ProB-Hor.W revealed several kinds of seed specific promoter motifs in 518 base pairs of these cloned promoter fragments as shown in Figures 5 and 6. The comparison of the two sequences with the sequences available in NCBI indicated the differences in the motifs (Figure 7). These kinds of seed specific promoter motifs in frag-ments ProB-Hor.A and ProB-Hor.W may be responsible for its seed-specific promoter activity. The results of the identified general transcription and potential regulatory elements are summarized in Table 1.

Based on these seed-specific motifs and the elements identified by the bioinformatics analysis, the isolated regions were selected as candidates for promoting gene expression in seed (Zavallo et al., 2010). The sequences important in gene expression were likely to be conserved among a group of genes with the same pattern of expression (Davidson et al., 1983). The 5′ noncoding regions were highly similar to those of the genes

encoding B Hordein polypeptides. Bioinformatics analyses of these sequences allowed for the identification of conserved motifs.

A-box motif found in *Petroselinum crispum*, is named CCGTCC-box in *Arabidopsis thaliana,* and is related to meristem specific activation (Logemann et al., 1995).

AAGAA and AC-II motifs were found in *Avena sativa* and *Phaseolus vulgaris*, respectively. G-Box in *Pisum sativum* and *A. thaliana* (Persson et al., 1998), GA-motif in *Helianthus annuus* (Waksman G et al., 1987), GAG-motif in *Spinacia oleracea* (Werneke et al., 1989), I-box in *Zea mays* (Sheen, 1991) and MRE motif in *Petroselinum crispum* in the MYB binding site (Feldbrugge et al., 1997) were involved in light responsiveness. ARE motif in *Z. mays* is essential for the anaerobic induction (Manjunath and Sachs, 2005). ABRE motif in *A. thaliana* is involved in the abscisic acid (ABA) responsiveness on the *rd29B* gene (Yamaguchi-Shinozaki and Shinozaki, 1993).

CAAT-box motif common cis-acting element in promoter and enhancer regions is often located in -80 positions, but much farther distances from the starting point and can also operate in both directions. CAAT box plays an important role in determining the efficiency of promoter (Lewin, 2009) and is found in *Brassica rapa, A. thaliana*, *Glycine max*, *Petunia hybrid* and *H. vulgare* (Shirsat et al., 1989). CGTCA-motif and TGACG-motif in *Hordeum vulgare* are cis-acting regulatory elements involved in the methyl jasmonate (MeJA) responsiveness (Rouster et al., 1997). GCN4-motif and Skn-1 motif *Oryza sativa* are required for endosperm expression (Takaiwa et al., 1991).

A TATA box sequence has been found in almost all plant genes in recent studies (Mesing et al., 1983). Many eukaryotic promoters between 10 and 20% of all genes (Gershenzon and Ioshikhes.,2005) contain a TATA box (sequence TATAAA), which in turn binds a TATA binding protein and assists the formation of the RNA polymerase transcriptional complex (Smale and Kadonaga, 2003). The TATA box typically lies very close to the transcriptional start site (often within 50 bases). The TATA box tends to be surrounded by GC rich sequences. A sequence which is rich in guanidine (G) and cytidine (C) nucleotides is usually found in multiple copies in the promoter region and normally surrounds the TATA box and CAP site. A TATA box is located at -79 base pairs from the start codon. The (CA) n box-like sequence has been designated a (CA) n box located at -138 bp from the initiation sites. TATA and (CA) n boxes within both promoters identified as well as other transcriptional enhancer boxes (Forde et al., 1985).

The sequence TGTAAAG, commonly named the pro-lamin box (p-box), was initially identified on the basis of both its highly conserved nucleotide sequence and location (-300 bp) relative to the transcription initiation site (TIS) of prolamin genes. It was recognized as a strong candidate for coordinating the expression of many seed specific expressions of many storage proteins
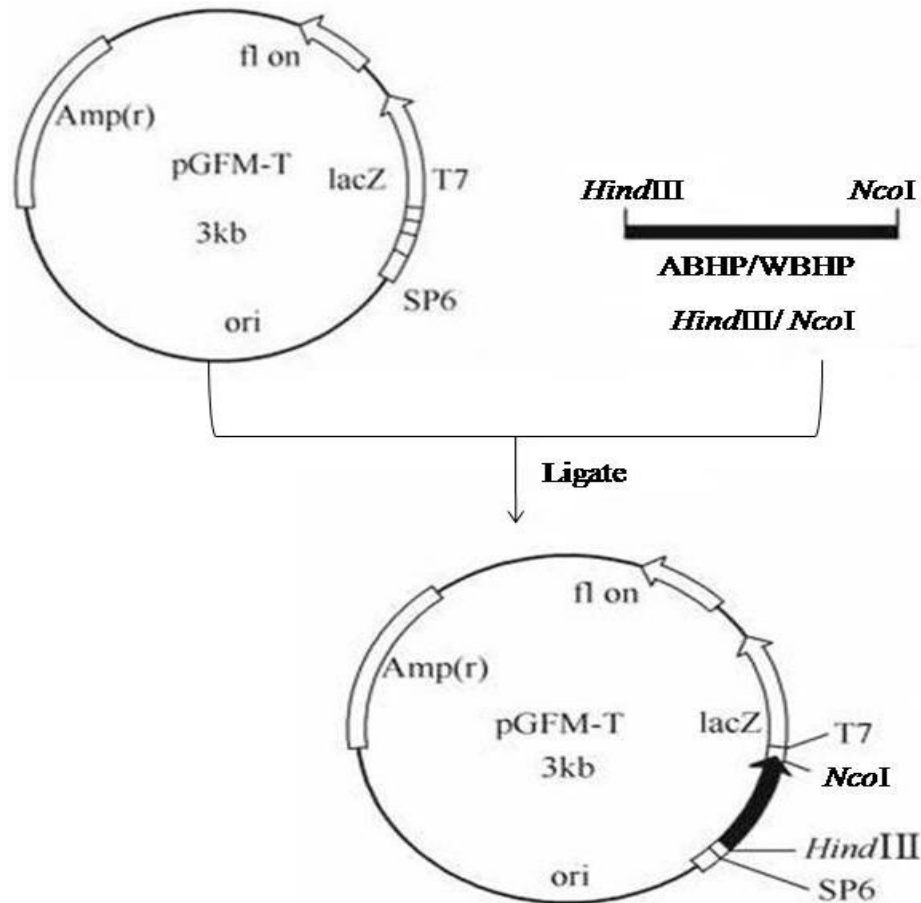
Figure 3. Construction of pT- ProB-Hor.A and pT- ProB-Hor.W
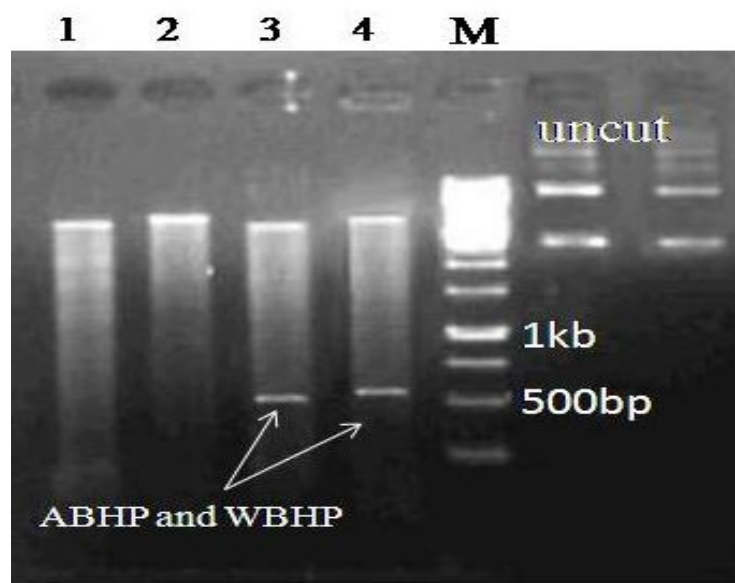
Figure 4. Construction and identification of ProB-Hor.A and ProB-Hor.W fragment M, DNA marker; 1, digestion product pT- ProB-Hor.A /*Hin*dIII; 2, digestion product pT- ProB-Hor.A /*Nco*I; 3, double digestion product pT-ProB-Hor.A /*Hin*dIII/*Nco*I; 4, double digestion product pT- ProB-Hor.W *Hin*dIII/*Nco*I.

A

1   GTCGAGAAGAACCGTCCACGTGTAAAGCTTTAACAACCCACACATGATTGCA

61 ACTTAGTCCTACACAAGTTTTCCATTCTTGTTTCAGGCTAACAACCTATACAAGGTTCCA

                         B   **A**   **C**
121 AAATCATTTGCAAAAGTGATGCTAGGTTGATAATTGTGTGACATGTAAAGTGAATAAGGT

         C
181 GAGTCATGCATACCAAACCTCGGGATTTCTATACTTTGTGTATGATCATGTGCACAACTA

         B   **B**
241 AAAAGCAACTTTGATTATCAATTCAAAAGTACGCTTGTAGCTTGTGCAACCTAACACAAT

                              D   **B**   **B**
301 GTACCAAAAATCCATTTGCAAAACATCCAAACACAATTGTTAAAGCTGATGCAAAGAAAG

         B   **E**   **F**
361 AAAGAGATGAAGCCCTGGCTACTATAAATAGGCATGTAGTATAGAGATCATCACAAACAC

                                                  G
421 AAGCATCAAAACCAAGAAACACTAGTTAACACCAATCCACTATGAAGACCTTCCTCATCT

481 TTGCACTCCTCGTCATTGCGGCAACAAGTACCATGGC

Figure 5. Sequence analysis of B1 hordein specific promoter fragment ProB-Hor.A A, E-box; B, Dof core recognition sequence; C, GLM motif; D, (CA)n box; E: N mtif; F, TATA box; G, start codon;–, ACGT motif.

A

1   GTCGAGAAGAACCGTCCACATGTAAAGCTTTAACAACCCACACATGATTGCA

61 ACTTAGTCCTACACAAGTTTTCCATTCTTGTTTCACGCTAACAACCTATACAAGTTTCCA

                         B   **A**   **C**
121 AAATCATTTGCAAAAGTGATGCTAGGTTGATAATTGTGTGACATGTAACGTGAATAAGGT

         C
181 GAGCCATGCATACCAAACCTCGGGATTTCTGTACTTTGTGTATGATCATATGCACAACTA

         B
241 AAAAGCAACTTTGATTATCAATTCAGAAGTACGCTTGTAGCTTGTGCAACCTAACACAAT

                              D   **B**   **B**
301 GTACCAAAAATCCGTTTGCAAAACATCCAAACACAATTGTTAAAGCTGATGCAAAGAAAG

         B   **E**   **F**
361 AAAGAGATGAAGCCCTGGCTACTATAAATAGGCATGTAGTATAGAGATCATCACAAGCAC

                                                  G
421 AAGCATCAAAATCAAGAAACACTAGTTAACACCAATCCACTATGAAGACCTTCCTCATCT

481 TTGCACTCCTTGCCATTGCGGCAACAAGTACCATGGC

Figure 6. Sequence analysis of B1 hordein specific promoter fragment ProB-Hor.W A, E-box; B, Dof core recognition sequence; C, GLM motif; D, (CA)n box; E, N mtif; F, TATA box; G, start codon.

```
Pro-BH-Alger      GTCGAGAAGAACCGTCCACATGTAAAGCTTTAACAACCCACACAT-GATTGCAACTTAGT
Pro-BH-walfajre   GTCGAGAAGAACCGTCCACGTGTAAAGCTTTAACAACCCACACAT-GATTGCAACTTAGT
Pro-BH-Query      GTCGAGAAGAACCGTCCACATGTAAAGCTTTAACAACCCACACATTGATTGCAACTTAGT
                  *******************  *************************** *************

Pro-BH-Alger      CCTACACAAGTTTTCCATTCTTGTTTCACGCTAACAACCTATACAAGTTTCCAAAATCAT
Pro-BH-walfajre   CCTACACAAGTTTTCCATTCTTGTTTCAGGCTAACAACCTATACAAGGTTCCAAAATCAT
Pro-BH-Query      CCTACACAAGTTTTCCATTCTTGTTTCAGGCTAACAACCTATACAAGGTTCCAAAATCAT
                  ****************************  ******************  ************ ***********

Pro-BH-Alger      TTGCAAAAGTGATGCTAGGTTGATAATTGTGTGACATGTAACGTGAATAAGGTGAGCCAT
Pro-BH-walfajre   TTGCAAAAGTGATGCTAGGTTGATAATTGTGTGACATGTAAAGTGAATAAGGTGAGTCAT
Pro-BH-Query      --GCAAAAGTGATGCTAGGTTGATAAT-GTGTGACATGTAAAGTGAATAAGGTGAGTCAT
                    *************************** ************** ************** ***

Pro-BH-Alger      GCATACCAAACCTCGGGATTTCTGTACTTTGTGTATGATCATATGCACAACTAAAAAGCA
Pro-BH-walfajre   GCATACCAAACCTCGGGATTTCTATACTTTGTGTATGATCATGTGCACAACTAAAAAGCA
Pro-BH-Query      GCATACCAAACCTCGGGATTTCTATACTTTGTGTATGATCATATGCACAACTAAAAGGCA
                  **********************  ******************* ************* ***

Pro-BH-Alger      ACTTTGATTATCAATTCAGAAGTAC-GCTTGTAGCTTGTGCAACCTAACACAATGTACCA
Pro-BH-walfajre   ACTTTGATTATCAATTCAAAAGTAC-GCTTGTAGCTTGTGCAACCTAACACAATGTACCA
Pro-BH-Query      ACTTTGATTATCAATTGAAAAGTACCGCTTGTAGCTTGTGCAACCTAACACAATGT-CCA
                  ***************** * ****** ********************************* ***

Pro-BH-Alger      AAAATCCGTTTGCAAAA-CATCCAAACACAATTGTTAAAGCTGATGCAAAGAA--AGAAA
Pro-BH-walfajre   AAAATCCATTTGCAAAA-CATCCAAACACAATTGTTAAAGCTGATGCAAAGAA--AGAAA
Pro-BH-Query      AAAATCCATTTGCAAAAGCATCCAAACACAATTGTTAAAGCTGTTCAAACAAACAAAGAA
                  *******  ********  *********************** *  **  **  *  **

Pro-BH-Alger      GAGATGAAGCCCTGGCTACTATAAATAGGCATGTAGTATAGAGATCATCACAAGCACAAG
Pro-BH-walfajre   GAGATGAAGCCCTGGCTACTATAAATAGGCATGTAGTATAGAGATCATCACAAACACAAG
Pro-BH-Query      GAGATGAAGCC-TGGCTACTATAAATAGGCAGGTAGTATAGAGATCTACACAAGCACAAG
                  ***********  ******************  *************  ***** ******

Pro-BH-Alger      CATCAAAATCAAGAAACACTAGTTAACACCAATCCACTATGAAGACCTTCCTCATCTTTG
Pro-BH-walfajre   CATCAAAACCAAGAAACACTAGTTAACACCAATCCACTATGAAGACCTTCCTCATCTTTG
Pro-BH-Query      CATCAAAACCAAGAAACACTAGTTAACACCAATCCACTATGAAGACCTTCCTCATCTTTG
                  ********  **************************************************

Pro-BH-Alger      CACTCCTTGCCATTGCGGCAACAAGTACCATGGCA
Pro-BH-walfajre   CACTCCTCGTCATTGCGGCAACAAGTACCATGGCA
Pro-BH-Query      CACTCCTCGCCATTGCGGCAACAAGTACGATTGCA
                  ******* * ***************** ** ***
```

Figure7. Comparing the two sequences with sequences available in NCBI

(SSPs) because of its presence within the promoters of all zein genes in maize (Thompson and Larkins, 1989), as well as many storage protein genes from the related cereals. Prolamin box (p-box or E box), a conserved motif (5′ TGTAAAG 3′), is present in cereal storage protein gene promoters (Forde et al., 1985). In many prolamin genes, p-box and GCN4 motifs are coupled with each other with only a few nucleotides which are separating them. This tandem module is designated as the bifactorial endosperm box (Marzabel et al., 1998). Two endosperm box (E-box) and GCN4-like motifs (GLM-motif

5′GTGAGTCAT 3′) are present in these promoters (Muller et al., 1995). Ebox is also named an endosperm box and is the binding site of endosperm-specific nuclear factors named barley-prolamin-box binding factor (BPBF) and binding was prevented when the pbox motif was mutated to 5′TGTAgAc 3′(Mena et al., 1998). In the second piece of E-box in position 164, the PROB-HOR.W fragment is a cytosine to adenine mutation.

ACGT motif is also a sequence motif required for seed-specific expression (Vincentz et al., 1997). In barley, nitrogen regulation is mediated by the endosperm

Table 1. Regulatory elements in promoter sequence.

Site	Sequence		Position	Strand	Function
	ProOB-Hor.A	ProB-Hor.W			
A-Box	CCGTCC	CCGTCC	12	+	Cis acting regulatory element
AAGAA-Motif		GAAAGAA	348	+	
ABRE	CACGTG	-	17	+	cis-acting element involved in the abscisic acid responsiveness
AC-II	(C/T)T(T/C)(C/T)(A/C)(A/C)C(A/C)A(A/C)C(C/A)(C/A)C	(C/T)T(T/C)(C/T)(A/C)(A/C)C(A/C)A(A/C)C(C/A)(C/A)C	28	+	-
ACE	ACGTGGA	-	15	-	cis-acting element involved in light responsiveness
ARE	TGGTTT	-	421	-	cis-acting regulatory element essential for the anaerobic induction
CAAT-box	CAAT	CAAT	445 289	+	common cis-acting element in promoter and enhancer regions
CCGTCC-box	CCGTCC	CCGTCC	12	+	cis-acting regulatory element related to meristem specific activation
CGTCA-motif	CGTCA	-	483	+	cis-acting regulatory element involved in the MeJA-responsiveness
G-Box	CACGTG	CACGTG	17 159	+ -	cis-acting regulatory element involved in light responsiveness
GA-motif	AAAGATGA	AAAGATGA	467	-	part of a light responsive element
GAG-motif	AGAGATG	AGAGATG	355	+	part of a light responsive element
GCN4-like or GSN;hor1-box	GTGAGTCAT	-	171	+	BF:bZIP transcription factor, cis-regulatory element involved in endosperm expression
I-box	gGATAAGGTG	gGATAAGGTG	164	+	part of a light responsive element
MRE	AACCTAA	AACCTAA	280	+	MYB binding site involved in light responsiveness
Prolamin-box (E-box)	tgagaTGTAAAGtgaat	tgacaTGTAAAGtgaat	151	+	Conserved in cereal seed storage protein gene promoters

Table 1. Contd.

Site	Sequence		Position	Strand	Function
	ProOB-Hor.A	**ProB-Hor.W**			
Skn-1_motif	GTCAT	-	175 484	+	cis-acting regulatory element required for endosperm expression
TATA box	TATAAATA	TATAAATA	375	+	core promoter element around -30 of transcription start, Important for recognition by RNA polymeraseII,
TGACG-motif	TGACG	-	483	-	cis-acting regulatory element involved in the MeJA-responsiveness
Unnamed__4	CTCC	CTCC	478	+	-
(CA)n	CATCCAAACACAA	CATCCAAACACAA	315	+	Seed specificity
ACGT motif	ACGT	-	17	+	Seed development
N motif	GATGAAGCCCT	GATGAAGCCCT	357	+	
Dof core recognition sequence	AAAG	AAAG	Very frequent	+	Light regulation, tissue specific gene expression
GA-2, Arabidopsis (*A. thaliana*)	AGAAAGAGA	AGAAAGAGA	350	+	BF: BPC
GA-1, Arabidopsis (*A. thaliana*)	-	AGAAAGAAA	346	+	BF: BPC1
EM1,wheat (T. *aestivum*)	TGTAAAGTG	-	155	+	BF: Unknown nuclear factor
GLM1, wheat (*T. aestivum*)	GGTGAGTCAT	-	169	+	BF: Unknown nuclear factor
GLM2, wheat (*T. aestivum*)	TGTGTGACAT	TGTGTGACAT	146	+	BF: Unknown nuclear factor

box via the interaction of endosperm and GCN4-like motif (Muller et al., 1995). The N motif has the consensus sequence G (A/G) TGAAGTCAT (Shewry and Halford, 2002). In these fragments, N motifs had only two nucleotide mutations that had occurred by replacing nucleotide T with C and A with C. This motif requires additional investigations. AAAG is Dof core recognition sequence and is so frequent that is regulated by light and tissue specific gene expression (Vicente-Carbajosa et al., 1997). The alignment between ProB-Hor.A and ProB-Hor.W showed that there were nine mismatches between the two sequences. Nucleotides 462 to 518 constituted a signal peptide that began with one ATG codon

and encoded a signal peptide (Watson. 1984). The presence of a signal peptide would be consistent with the evidence that the B hordeins were synthesized on the rough endoplasmic reticulum and were deposited in protein bodies (Matthews and Miflin., 1980; Cameron-Mills., 1980).

A typical signal peptide comprises of three distinct regions: a polar N terminal end (n-region) that may have a net positive charge, a central hydrophobic core (h-region) that consists of 6±15 hydrophobic amino acids and a polar C-terminal (c-region) end that contains prolines and glycines (von Heijne, 1985). A signal peptide containing the consensus sequence and proper cleavage site

ensures that proteins are inserted into the endoplasmic reticulum (ER) membrane. Mutations within the sequence immediately downstream the signal peptide affected protein processing. The nucleotide sequence of the signal peptide genotype Alger and Walfajr into amino acid sequence was converted by vector NTI 9.1 software.

The sequence of the signal peptide region of Alger genotype was compared with the signal peptide recorded by basic local alignment search tool program (BLASTP). Their similarity was in the range of 90 to 95%. The results indicate that the amino acid methionine that mutated the isoleucine in the Alger genotype did not prevent any transfer of proteins in vacuoles (Figure 8A). The sequence

```
Query 1 MKTFLIFALLVIAATSTMA 19        Query    1 MKTFLIFALLAIAATSTMA 19

        MKTFLIFALLVIAATST A                    MKTFLIFALLAIAATST A

Alger1 MKTFLIFALLVIAATSTIA  19        Walfajre 1 MKTFLIFALLAIAATSTIA 19
```

A   **B**

```
Alger    1 MKTFLIFALLVIAATSTMA  19

           MKTFLIFALL IAATSTMA

Walfajre 1 MKTFLIFALLAIAATSTMA  19
```

C

```
                        n     h         c
Sp.B-Hordein,Alger    MKTFLIFALLVIAATSTMA

Sp.B-Hordein,Walfajre MKTFLIFALLAIAATSTMA
                        n     h         c
```

D

Figure 8. A, Comparison of the signal peptide sequence in Alger genotype; B, comparison of the signal peptide sequence in walfajre genotype, C: Results of BLAST P signal peptide in Alger and walfajre; D, various section of the signal peptide and cleavage site is shown with↑

of the signal region of Alger genotype in comparison with the signal peptide recorded by BLASTP had 95% similarity (Figure 8B). The amino acid similarity of signal peptide of two genotypes Alger and Walfajr was 95%. The comparison of DNA blast and peptide signals (Figure 8C) showed that mutations in sequences coding of 10th amino acids (leucine) did not change it because of the 3rd nucleotides changes. However, the changes in the second nucleotide in the 11th amino acid led to the change of valine in Alger to Alanin in Walfajre.

Signal peptide was investigated by the online applicationof: www.cbs.dtu.dk/services/SignalP/index.php. Cleavage site in the signal peptide of Alger and Walfajre genotypes was located between the 13th and 14th amino acids with probability of 97% (Figure 8D). Proteins transfer to the vacuoles was not conducted if the cleavage site was located in other areas.

Conclusion

The aim of this research was to isolate one of the seed specific promoters of *H. vulgare.* Two barley varieties that were common in Iran were selected and their genomes were extracted using the specific designed primers for amplifying the B hordein promoter. The amplified fragment of the insert was cloned in an appropriate vector and then was transformed to *E. coli*. At last, for the final admission of accuracy, the cloned fragments were sent for sequencing. The results are compared with the sequences existing in data banks. The promoter region contained motifs; like TATA box, (CA) $_n$ box, ACGT motif, AAAG motif, GCN4-like motif and E-box, which constituted the seed-specific promoter activity.

The results obtained from the neural network algorithm

show three scores in signal peptide: the high S-score, C-score and the Y-score. The amino acid similarity of signal peptide of two genotypes (Alger and Walfajr) was 95%.

REFERENCES

BlakeL TK, Ullrich SE, Nilan RL (1982). Mapping of Hor-3 locus encoding D hordein in barley.Theor. Appl. Genet. 53: 367-371.

Cameron-Mills V (1980). The structure and composition of protein bodies purified from barley endosperm by silica sol density gradients. Carlsberg Res. Commun. 45: 557-576.

Davidson EH, Jacobs HT, Britten RJ (1983). Eukaryotic gene expression: Very short repeats and coordinate induction of genes. Nature, 301: 468-470.

Evans IM, Gatehouse JA, Croy RD, Boulter D (1984). Regulation of the transcription of storage protein mRNA in nuclei isolated from developing pea (*Pisum sativum*) cotyledons. Planta, 160: 559-568.

Feldbrugge M, Sprenger M, Hahlbrock K, Weisshaar B (1997). PcMYB1, a novel plant protein containing a DNA-binding domain with one MYB repeat, interacts in vivo with a light-regulatory promoter unit. Plant J. 11(5): 1079-1093.

Forde BG, Heyworth A, Pywell J, Kreis M (1985). Nucleotide sequence of a B1 hordein gene and the identification of possible upstream regulatory elements in endosperm storage protein genes from barley, wheat and maize. Nucleic Acids Res. 13: p. 20.

Gatehouse JA, Evans IM, Croy RRD, Boulter D (1986). Differential expression of genes during legume seed development. Philos Trans. R. Soc. Lond (B) Biol. Sci. 314: 367-384.

Gershenzon NI, Ioshikhes IP (2005). Synergy of human Pol II core promoter elements revealed by statistical sequence analysis. Bioinformatics, 21(8): 1295-300.

Jensen J, Jorgensen JH, Jensen HP, Giese H, Doll H (1980). Linkage of the hordein loci, Hor 1 and Hor 2 with the powdery mildew resistance loci MI-k and Mt-a on barley chromosome 5. Theor. Appl. Genet. 58: 27-31.

Lewin B (2009).GENES VIII. Pearson Prentice Hall. p. 1030.

Logemann E, Parniske M, Hahlbrock K (1995). Modes of expression and common structural features of the complete phenylalanine ammonia-lyase gene family in parsley. Proc Natl Acad. Sci. USA. 92: 5905-5909.

Manjunath S, Sachs MM (2005). Molecular characterization and

promoter analysis of the maize cytosolic glyceraldehyde 3-phosphate dehydrogenase gene family and its expression during anoxia. Unpublished. AC(NCBI). pp. U45858.

Marzabal P, Busk PK, Ludevid M D, Torrent M (1998). The bifactorial endosperm box of g-zein gene: characterisation and function of the Pb3 and GZM cis-acting elements. Plant J. 16: 41-52.

Matthews JA, Miflin BJ (1980). *In vitro* synthesis of barley storage proteins. Planta, 149: 262-268.

Mena M, Vicente-Carbajosa JJ, Schmidt R, Carbonero P (1998). An endosperm-specific DOF protein from barley, highly conserved in wheat, binds to and activates transcription from the prolamin-box of a native B-hordein promoter in barley endosperm. Plant J. 16(1): 53-62.

Messing J, Geraghty D, Heidecker G, Hu NT, Kridl J, Rubsenstein I (1983) in Genetic engineering of Plants, Kosuge, T., Meredith, C.P. and Hollaender, A. Eds., Plenum Press, New York. pp. 211-227.

Muller M , Muth J R, Gallusci P, Knudsen S, Maddaloni M, Motto M, Schmitzd D, Sorensen M B, Salamini F, Von Wettstein D, Thompson RD (1995). Regulation of storage protein synthesis in cereal seeds: developmental and nutritional aspects. Plant Physiol. 145: 606-613.

Persson EL, Brosche M, Strid A (1998). Cultivar differences in the sequences of chalcone synthase promoters in pea. Unpublished. AC (NCBI): pp. AF060237.

Piston F, Martin A, Dorado G, Barro F (2005). Cloning and molecular characterization of B-hordeins from Hordeum chilense (Roem. et Schult). Theor. Appl. Genet. 111: 551-560.

Rouster J, Leah R, Mundy J, Cameron-Mills V (1997). Identification of a methyl jasmonate responsive region in the promoter of a lipoxygenase 1 gene expressed in barley grain. Plant J. 11(3): 513-523.

Saghai-Maroof MA, Soliman K, Jorgensen RA, Allard RW (1984). Ribosomal DNA spacer length polymorphisms in barley: Mendelian inheritance, chromosomal location, and population dynamics. PNAS, 81: 8014-8018.

Sheen J (1991). Molecular mechanisms underlying the differential expression of maize pyruvate, orthophosphate dikinase genes. Plant Cell, 3(3): 225-245.

Shewry PR, Faulks AJ, Pickering RA, Jones IT, Finch RA, Miflin BJ (1980). The genetic analysis of barley storage protein. Heredity, 44: 383-389.

Shewry PR, Finch RA, Parmar SJ, Frankin J, Miflin BJ (1983). Chromosomal location of Hor3, a new locus governing storage proteins in barley. Heredity, 50(2): 179-189.

Shewry PR, Halford NG (2002). Cereal seed storage proteins: structures, properties and role in grain utilization. J. Exp. Bot. 53(370): 947-958.

Shirsat A, Wilford N, Croy R, Boulter D (1989).Sequences responsible for the tissue specific promoter activity of a pea legumin gene in tobacco. Mol. Gen. Genet. 215: 326-331.

Smale T, Kadonaga T (2003). The RNA polymerase II core promoter. Ann Rev. Biochem. 72: 449-479.

Takaiwa F, Oono K, Wing D, Kato A (1991). Sequence of three members and expression of a new major subfamily of glutelin genes from rice. Plant Mol. Biol. 17(4): 875-885.

Thompson GA, Larkins BA (1989). Structural elements regulating zein gene expression. Bioassayse, 10: 108-113.

Vincentz M, Leite A, Neshich G, Vriend G, Mattar C, Barros L, Weinberg D, De Almeida ER, De Carvalho MP, Aragao F, Gander ES (1997). ACGT and vicilin core sequences in a promoter domain required for seed-specific expression of a 2S storage protein gene are recognized by the opaque-2 regulatory protein. Plant Mol. Biol. 34: 879-889.

Vicente-Carbajosa J, Moose SP, Parsons RL, Schmidt RJ (1997). A maize zinc-finger protein binds the prolamin box in zein gene promoters and interacts with the basic leucine zipper transcriptional activator Opaque2. Proc. Natl. Acad. Sci. USA. 94: 7685-7690.

Von Heijne G (1985). Signal sequences. The limits of variation. J. Mol. Biol. 184: 99-105.

Waksman G, Lebrun M, Freyssinet G (1987). Nucleotide sequence of a gene encoding sunflower ribulose-1,5-bisphosphate carboxylase/oxygenase small subunit (rbcs) Nucleic Acids Res. 15(17): 7181-7181.

Watson MEE (1984). Compilation of published signal sequences Nucleic. Acids Res. 12: 5145-5164.

Werneke JM, Chatfield JM, Ogren WL (1989). Alternative mRNA splicing generates the two ribulosebisphosphate carboxylase/oxygenase activase polypeptides in spinach and Arabidopsis Plant Cell, 1(8): 815-825.

Yamaguchi-Shinozaki K, Shinozaki K (1993). Arabidopsis DNA encoding two desiccation-responsive rd29 genes. Plant Physiol. 101(3): 1119-1120.

Zavallo D, Lopez BM, Esteban Hopp H, Heinz R (2010). Isolation and functional characterization of two novel seed-specific promoters from sunflower (*Helianthus annuus* L.). Plant Cell Rep. 29: 239-248.