



Expression of globulin-2, a member of the cupin superfamily of proteins with similarity to known food allergens, is increased under high temperature regimens during wheat grain development

Susan B. Altenbach*, Charlene K. Tanaka, William J. Hurkman, William H. Vensel

USDA-ARS Western Regional Research Center, 800 Buchanan Street, Albany, CA 94710, USA

ARTICLE INFO

Article history:

Received 27 February 2008

Received in revised form 28 June 2008

Accepted 8 July 2008

Keywords:

Flour quality

Gene expression

Proteomics

Quantitative RT-PCR

ABSTRACT

Twenty-three expressed sequence tags (ESTs) from the US spring wheat Butte 86 were identified that encode proteins similar to a globulin-2 protein from maize embryos. The ESTs assembled into three contigs, two of which include the entire coding region for the mature protein. The encoded proteins contain two cupin domains and show significant identities with 7S seed proteins from other species that are known or putative food allergens. Quantitative reverse-transcriptase polymerase chain reaction (qRT-PCR) with primers specific for two of the sequences demonstrated that the globulin-2 genes are expressed late in grain development and that transcript levels increase when grain is produced under high temperature conditions. Transcripts were detected in both whole grain and endosperm, but levels were significantly higher in whole grain and highest in embryo. In wheat flour, at least 17 protein spots that differ in both size and *pI* were identified as globulin-2 by 2-DE/MS. Seven of the spots increased more than 2-fold in relative proportion when grain was produced under high temperature regimens. The data suggest that both transcriptional and post-translational mechanisms are involved in the response of globulin-2 to high temperatures.

Published by Elsevier Ltd.

1. Introduction

Wheat is one of the major crops grown throughout the world with a primary use in human nutrition. In 2005, wheat was grown on more than 200 million hectares and more than 600 million tons of grain were produced worldwide (<http://faostat.fao.org/>). A large variety of breads, noodles and baked goods are made from wheat flour because of the unique viscoelastic properties that result when the flour is mixed with water. These viscoelastic properties are determined largely by the gluten proteins, the major storage proteins that comprise 80–85% of wheat flour protein. The gluten proteins, consisting of gliadins and glutenin subunits, are characterized by unusually high levels of proline and glutamine and have very low solubilities in water or salt solutions. These proteins have been studied extensively because of their importance in flour functionality. The remaining 15–20% of the endosperm protein is

a heterogeneous mixture of proteins, most of which are soluble in water or dilute salt solutions. Within the salt-soluble protein fraction are a number of proteins related to the predominant storage proteins characterized in seeds from dicotyledonous plants. Some of these proteins have been shown to have sedimentation values similar to the 7S vicilins from pea and to share antigenic determinants with the pea vicilins (Robert et al., 1985). A 7S + 3S globulin fraction characterized by Robert et al. (1985) contained major proteins of ~75, 55, 36–40, 24, 17–20 and 10 kDa when analyzed by SDS-PAGE.

Recently, a 2-DE/MS approach was used to identify more than 200 proteins contained in a salt-soluble fraction of wheat endosperm from the US spring wheat Butte 86 (Vensel et al., 2005). Seven spots in the 55 kDa region of the 2-D gels were determined to be most similar to a 7S vicilin-like protein characterized from maize embryos, referred to as globulin-2 (Wallace and Kriz, 1991). These proteins were found in wheat endosperm at 36 DPA but not at 10 DPA. In a subsequent proteomic analysis of endosperm from developing grain, the globulin-2 proteins were shown to accumulate during late stages of wheat grain development. Interestingly, these proteins also were among a group that increased in relative proportion when grain development occurred under a high temperature regimen of 37/28 °C (day/night) rather than under a moderate temperature regimen of 24/17 °C (Hurkman et al., 2009). Since these data suggest biological roles for the globulin-2 proteins

Abbreviations: 2-DE, two-dimensional gel electrophoresis; 2-DE/MS, two-dimensional gel electrophoresis/mass spectrometry; DPA, days' post-anthesis; ESTs, expressed sequence tags; MS/MS, tandem mass spectrometry; NPK, nitrogen-phosphorous-potassium; qRT-PCR, quantitative reverse-transcriptase polymerase chain reaction; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis.

* Corresponding author. Tel.: +1 510 559 5614; fax: +1 510 559 5818.

E-mail address: altnbach@pw.usda.gov (S.B. Altenbach).

in the response of the grain to abiotic stress, we surveyed databases of wheat expressed sequence tags (ESTs) to identify the complement of globulin-2 genes from the US spring wheat Butte 86. Additionally, we examined the expression of two of the globulin-2 genes in developing grain and endosperm under different controlled environmental regimens using quantitative RT-PCR and evaluated the relative proportions of globulin-2 proteins in flour produced from grain subjected to moderate and high temperature regimens.

2. Experimental

2.1. Growth of plants and tissue collection

The US hard red spring wheat *Triticum aestivum* cv. Butte 86 was grown in a climate-controlled greenhouse as described by Altenbach et al. (2007). During grain development, plants were placed under either a moderate (24/17 °C day/night) or a high (37/28 °C day/night) temperature regimen. Levels of fertilizer (Plantex 20–20–20 NPK) were adjusted at anthesis so that plants received 300 mg per day (1×), 150 mg per day (0.5×) or no (0×) NPK during grain development. Developing grain and endosperm were collected at various intervals after anthesis under both temperature regimens. Embryo tissue was harvested from 30 DPA grain from plants grown under the 24/17 °C regimen. Awns, glumes, stems and leaves were obtained during middle stages of grain development under the moderate regimen. Roots were harvested from seedlings grown on water-saturated sand at room temperature for 20 days. All tissues were frozen in liquid nitrogen and stored at –80 °C until use.

2.2. RNA preparation and qRT-PCR

Total RNA was isolated from all tissues as described previously for endosperm (Altenbach, 1998), treated with RQ1 RNase-free DNase (Promega, Madison, WI), and further purified by extraction with phenol:chloroform:isoamylalcohol (24:24:1) and ethanol precipitation. RNA was reverse-transcribed using the QuantiTect Reverse Transcription Kit (QIAGEN, Valencia, CA) according to the manufacturer's directions. Amplification reactions were carried out in a volume of 25 µl containing cDNA, 0.3 µM of forward and reverse primers and SYBR Green Supermix (Biorad Laboratories, Hercules, CA) using a BioRad iCycler with an initial denaturation step of 95 °C for 3 min, followed by 40 cycles of 95 °C for 10 s and 55 °C for 45 s. At the end of the PCR cycles, a melting curve was generated and analyzed by the iCycler software.

Oligonucleotide primers for globulin-2 genes were based on expressed sequence tags (ESTs) from developing grain of Butte 86 available from NCBI. Primers were selected using Beacon Designer 4.0 (Premier Biosoft International, Palo Alto, CA) and synthesized by Operon (Huntsville, AL). Forward primer QF123 has the sequence 5' GGTCCAAATGGCTCCGC 3' and reverse primer QR123 has the sequence 5' CAAGAAGGACTGCGGC 3'. Forward primer QF126 has the sequence 5' GGTGCTCATGCTCTCAAC 3' and reverse primer QR126 has the sequence 5' GACGCTGAAGAAGGACTGTG 3'. Primers for an 18S rRNA used as a reference were reported in Altenbach et al. (2007). Amplification efficiencies for each primer pair were calculated from standard curves generated in three independent experiments using the iCycler software. Each standard curve had a minimum of five points and *R* values greater than 0.99. Amplification products were evaluated by gel electrophoresis on 4% Metaphor agarose gels (VWR International, Brisbane, CA).

For quantification of transcripts, PCR reactions were carried out in triplicate using cDNA from the equivalent of 10 ng RNA from each time point. The 18S rRNA served as a reference RNA and was amplified in parallel with target genes in all experiments. The *C_t*

value was determined for each RNA sample and primer pair by the iCycler software. Mean normalized expression and standard errors were determined for each target gene relative to the 18S reference using the Q-Gene Core Module (available at <http://www.gene-quantification.de/download.html#qgene>). Amplification efficiencies of each primer pair were taken into account for all calculations using equation #3 of Muller et al. (2002). Mean normalized expression was plotted as a function of chronological age for each experiment.

2.3. Protein extraction, 2-DE and MS/MS analyses

Endosperm was collected from developing grain and a KCl-soluble/methanol-insoluble protein fraction was isolated and quantified as described previously (Hurkman and Tanaka, 2004; Vensel et al., 2005). Equal amounts of protein (18 µg) from each time point were separated by 2-DE in triplicate. Gels were digitized with a calibrated scanner (UMAX PowerLook III, Dallas, TX) at 300 dpi using the same setting for all gels and analyzed using a computerized image analysis system (Progenesis PG240 version 2006, Non-Linear Dynamics Limited, Newcastle upon Tyne, UK). Normalized volumes (individual spot volume/total spot volume × 100) were determined for each spot, averaged over the three gels, and plotted against the chronological age of the grain in DPA.

Grain produced under different temperature regimens was tempered to 15% moisture and milled to flour by the USDA-ARS Wheat Quality Laboratory (Manhattan, KS) using a Brabender Quadrumat Jr. mill (Hackensack, NJ). Methods for extraction and analysis of proteins from flour were the same as those used for endosperm.

For identification of proteins by tandem mass spectrometry (MS/MS), spots were excised from a representative gel and reduction, alkylation, reagent removal, and tryptic digestion were carried out automatically by a DigestPro xyz robot (INTAVIS Bioanalytical Instruments AG, Bergisch Gladbach, Germany). The DigestPro collection tray containing the tryptic peptides was placed in an autosampler that was interfaced with a QSTAR pulsar I hybrid quadrupole-TOF instrument (Applied Biosystems/MDX Sciex, Toronto, Canada) configured with an ESI source. Acquisition of tandem mass spectrometry data was as follows. From an initial survey scan of mass range *m/z* 400–2000, the most abundant doubly or triply charged ion above a threshold of 20 counts was selected for fragmentation. Collision induced dissociation of the mass-selected ion was carried out using UHP nitrogen. Following the 3S MS/MS fragmentation period, the MS survey scan was repeated until another MS/MS period was triggered. Wiff data files were created for each sample by the Analyst QS version 1.1 software, converted to MGF text files using Mascot Daemon (<http://www.matrixscience.com/>) and submitted in batch mode to a locally installed copy of X!Tandem (Craig and Beavis, 2004) using a script provided by Jayson Faulkner (University of Michigan). X!Tandem was configured to search files containing sequences from the NCBI non-redundant collection: nr-*Arabidopsis-thaliana*.fasta, nr-other-Viridiplantae.fasta, nr-*Oryza-sativa*.fasta, in addition to entries from all proteins in the HarvEST:Wheat version 1.04 database (<http://harvest.ucr.edu/HWheat104.exe>), NCBI *Triticum aestivum*: UniGene Build #37, and wEST Database (<http://wheat.pw.usda.gov/wEST>). The results were visualized using a locally installed copy of the Global Proteome Machine (GPM) (<http://thegpm.org/>). Trypsin was selected as the cleavage enzyme. The results were searched with a fragment ion mass tolerance of 1.00 Da and a parent ion tolerance of 0.40 Da. Oxidation of methionine was specified as a variable modification (see also: Vensel et al., 2002, 2005). Identifications accepted as valid had log *e* values less than –4.0.

3. Results

3.1. Accumulation of globulin-2 in developing endosperm under different temperature regimens

A number of protein spots were identified as globulin-2 in proteomic maps of KCl-soluble/MeOH-insoluble proteins from developing endosperm from Butte 86 (Vensel et al., 2005). Six spots (referred to as 147, 155, 559, 561, 851, and 852 in Vensel et al., 2005) ranged from about 54.8 to 56.1 kDa by 2-DE and had pIs between 6.2 and 7.0. Another spot (631 in Vensel et al., 2005) was about 15 kDa and had a pI of about 5.7. Accumulation profiles for these seven globulin-2 proteins during grain development under moderate and high temperature regimens were summarized as part of a large-scale proteomic analysis by Hurkman et al. (2009) and are shown in more detail in Fig. 1. Under moderate temperatures (24/17 °C), these proteins accumulated very late in grain development, between 32 and 34 DPA, and each protein spot

encompassed less than 0.2% of the normalized volumes of all spots in the KCl-soluble/MeOH-insoluble fraction. When a high temperature regimen of 37/28 °C was applied from 10 DPA until maturity, the proteins accumulated earlier, between 22 and 24 DPA. In addition, maximum normalized spot volumes were 1.9- to 4.1-fold higher than in grain produced under the 24/17 °C regimen (Fig. 1 and Hurkman et al., 2009).

3.2. Identification of ESTs encoding globulin-2 from Butte 86

A collection of contigs assembled from more than 116,000 wheat ESTs by Chao et al. (2006) was surveyed to identify sequences that encode the wheat globulin-2 proteins. The EST collection includes 3639 ESTs from Butte 86 developing grain. Information about both contigs and ESTs is available at <http://wheat.pw.usda.gov/westsql/>. Three contigs, NSFT03P2_Contig18428, NSFT03P2_Contig17295 and NSFT03P2_Contig17366, encode proteins similar to a globulin-2 characterized previously in maize (Wallace and Kriz, 1991). Amino

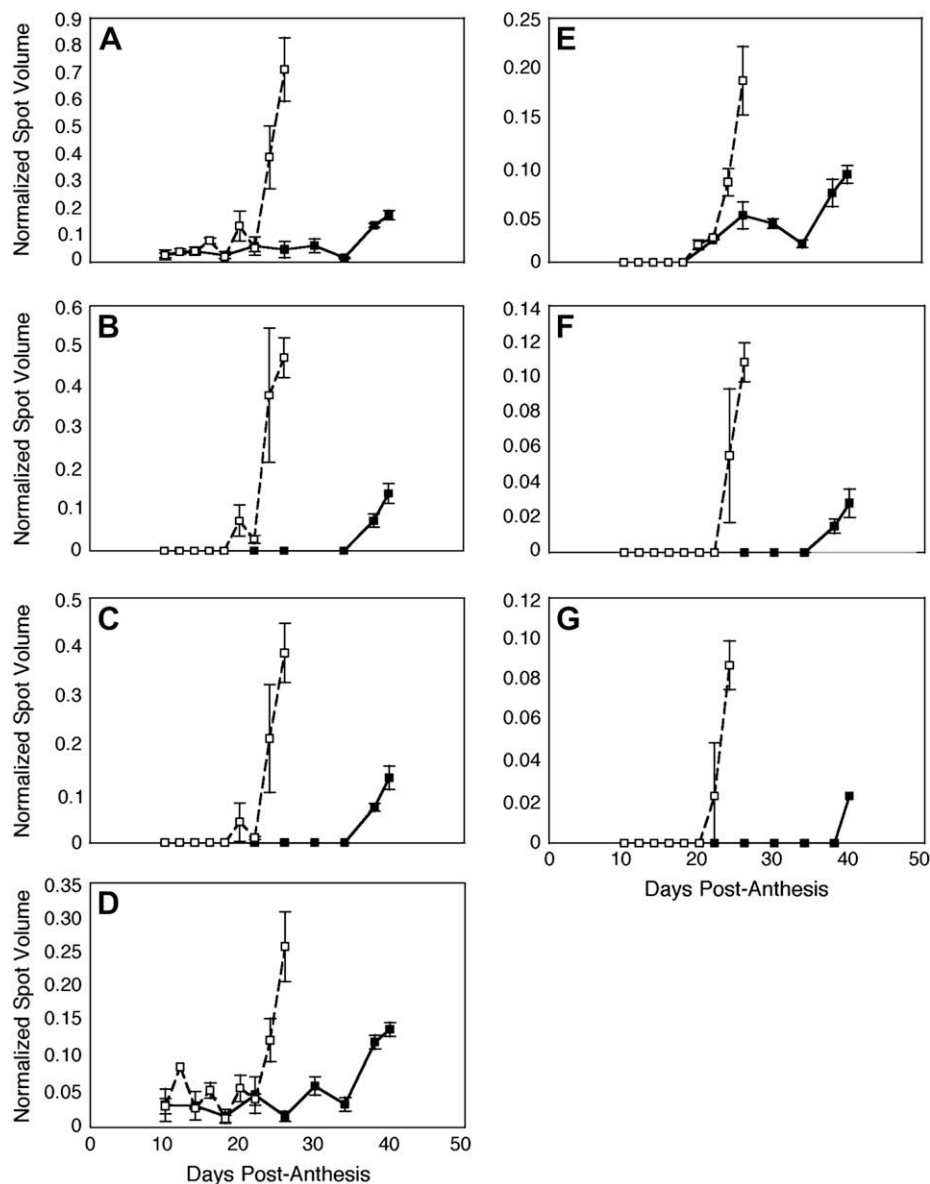


Fig. 1. Accumulation of globulin-2 proteins in endosperm from grain produced under moderate and high temperatures. In each panel the solid lines denote endosperm produced under the 24/17 °C regimen and the dashed lines denote endosperm produced under the 37/28 °C regimen. The high temperature regimen was imposed from 10 DPA to maturity and plants were supplied with $0.5 \times$ NPK. Bars indicate the standard error for each time point among triplicate gels. Accumulation data for spots 155, 559, 851, 147, 631, 852 and 561 from Vensel et al. (2005) are presented in panels A, B, C, D, E, F and G, respectively.

acid sequences of the proteins encoded by the contigs are shown in Fig. 2. Contig 18428 includes nine ESTs from Butte 86 (BQ806323, BQ807065, BQ806087, BQ805228, BQ806420, BQ805507, BQ806904, BQ804552, BQ804667) that represent the entire coding region. Contig 17295 includes four ESTs from Butte 86 (BQ805639, BQ839104, BQ805390, BQ807159) that cover the 5' portion of the coding region and one EST (BQ805520) that represents the 3' end. Nine ESTs from Butte 86 were included in contig 17366 (BQ839060, BQ805920, BQ807178, BQ838786, BQ839055, BQ838678, BQ807089, BQ806653, BQ805350) and all nine represent the 5' half of the coding region. It is likely that this contig is missing the 3' end of the coding region.

When evaluated with the Signal P algorithm (<http://www.cbs.dtu.dk/services/SignalP/>), proteins encoded by the three contigs are predicted to contain signal peptide cleavage sites between the alanine and the serine residues at positions 26 and 27 in contigs 17295 and 17366 and positions 20 and 21 in contig 18428 (Bendtsen et al., 2004), suggesting that these proteins enter the secretory pathway (Fig. 2). Contigs 18428 and 17295 encode proteins with predicted molecular weights of 53,495 and 53,773, respectively, excluding the signal peptide. Both proteins contain abundant arginine, glycine and glutamine residues. An InterProScan (<http://www.ebi.ac.uk/InterProScan/>) showed that each protein contains two cupin domains (Fig. 2) that are able to form barrel-like structures typical of the cupin superfamily of proteins.

Since many plant proteins within the cupin superfamily are food allergens (Mills et al., 2002), the protein encoded by contig 17295 was used to search the Food Allergy Research and Resource Program (FARRP) Protein AllergenOnline Database version 8.0 (<http://www.allergenonline.com/databasefasta.asp>). A sliding 80-mer search revealed matches of >35% identity between this globulin-2 and 18 7S proteins that are known or putative allergens (Table 1). These include proteins from sesame, various tree nuts (English walnut, black walnut, hazelnut, cashew) and legumes (lentil, pea, soybean, peanut). Current Codex Alimentarius guidelines indicate that proteins with identities to known allergens that are >35% over 80 amino acid regions may be allergenic (Goodman, 2006).

3.3. Selection of primers for qRT-PCR and analysis of gene expression

ESTs BQ806323, BQ805639 and BQ839060 from cv Butte 86 represent the first half of the coding sequences of contigs 18428, 17295 and 17366, respectively, and were used for the design of gene-specific primers. Primer QF123 was a perfect match with BQ806323, but had only 13/17 nucleotides identical to BQ839060 and BQ805639 (Fig. 3). Primer QR123 was a reverse complement of BQ806323, but had only 13/17 nucleotides that complemented either BQ839060 or BQ805639. QF123/QR123 amplified a 110 bp fragment that was verified to correspond to BQ839060 by digestion with AlwI, HaeII, and HpaII. The average amplification efficiency of QF123/QR123 was 96.7%. Primer QF126 was a perfect match with both BQ805639 and BQ839060 but had only 16/19 nucleotides in common with BQ806323 (Fig. 3). Primer QR126 was a reverse complement of BQ805639, but had only 18 and 17 of 20 residues that complemented BQ806323 and BQ839060, respectively. QF126/QR126 amplified a 100 bp fragment that was verified to correspond to BQ805639 by digestion with HpaII, MboII, AclI, AlwI and HaeII. The average amplification efficiency of QF126/QR126 was 93.7%.

Transcript accumulation for the two globulin-2 genes was assessed in developing grain produced under either moderate or high temperature regimens from anthesis to maturity (Fig. 4). In the absence of post-anthesis NPK (Fig. 4A, B), transcripts corresponding to both genes were first detected at 11 DPA under moderate temperatures and reached maximum levels at 26 DPA. Under the high temperature regimen, transcripts began to accumulate earlier in grain development, by 6 DPA, and reached maximum levels at 22 DPA. Changes in the timing of globulin-2 transcript accumulation were consistent with changes in the timing of developmental events due to high temperature (Altenbach et al., 2003) and with changes in the timing of expression of other genes (Altenbach and Kothari, 2004; Altenbach et al., 2007, 2008). Maximum levels of transcripts corresponding to BQ806323 and BQ805639 were 4.1- and 4.9-fold higher, respectively, under high temperatures than under moderate temperatures. In a second growth experiment,



Fig. 2. Comparison of protein sequences encoded by NSFT03P2_Contig18428, NSFT03P2_Contig17295 and NSFT03P2_Contig17366. Identical amino acid residues in each protein are enclosed in boxes. Brackets denote portions of the coding regions in each contig that correspond to ESTs from Butte 86. The arrow indicates putative signal peptide cleavage sites. Cupin domains in each protein are shown in bold type and putative N-linked glycosylation sites are indicated by asterisks.

Table 1

Comparison of globulin-2 encoded by NSFT03P2_Contig17295 with known and putative allergens in the FARRP Protein AllergenOnline Database Version 8.0 using a sliding 80-mer search

GenBank accession #	Source	Best % identity	# Hits > 35% ^a
13183177	<i>Sesame indicum</i> (sesame)	51.90	285
6580762	<i>Juglans regia</i> (English walnut)	48.10	269
31321944	<i>Juglans nigra</i> (black walnut)	47.50	251
19338630	<i>Corylus avellana</i> (hazelnut)	47.50	204
21666498	<i>Anacardium occidentale</i> (cashew)	43.80	173
21914823	<i>Anacardium occidentale</i> (cashew)	43.80	179
29539109	<i>Lens culinaris</i> (lentil)	42.51	112
46560474	<i>Arachis hypogaea</i> (peanut)	41.29	62
42414627	<i>Pisum sativum</i> (pea)	41.24	97
42414629	<i>Pisum sativum</i> (pea)	41.24	105
1168390	<i>Arachis hypogaea</i> (peanut)	40.70	75
46560472	<i>Arachis hypogaea</i> (peanut)	40.70	58
1168391	<i>Arachis hypogaea</i> (peanut)	40.70	79
46560476	<i>Arachis hypogaea</i> (peanut)	40.70	79
29539111	<i>Lens culinaris</i> (lentil)	40.00	123
256427	<i>Glycine max</i> (soybean)	38.79	45
18536	<i>Glycine max</i> (soybean)	36.29	11
169929	<i>Glycine max</i> (soybean)	35.03	2

Proteins with >35% identity over 80 amino acids to wheat globulin-2 are shown.

^a of 426 80-mers.

plants were supplied with post-anthesis NPK (Fig. 4C, D). In this experiment, transcripts for BQ806323 were first detected by 12 DPA and reached maximum levels at 34 DPA under moderate temperatures (Fig. 4C) while transcripts for BQ805639 were detectable at low levels at 7 DPA and reached maximum levels at 34 DPA (Fig. 4D). Under the high temperature regimen, transcripts for both genes were first detected at 7 DPA and reached maximum levels at 18 DPA, the last time point sampled. Maximum levels of BQ806323 and BQ805639 transcripts were 4.1- and 2.1-fold higher, respectively, under the high temperature regimen than under moderate temperatures.

The accumulation of globulin-2 transcripts also was assessed in developing endosperm (Fig. 4E, F). Transcripts for both BQ806323 and BQ805639 were detected at low levels at both 7 and 14 DPA under moderate temperatures and reached maximum levels at 34 DPA. When high temperatures were applied from anthesis to maturity, transcripts for both genes were detected at the first time point sampled at 5 DPA and increased to maximum levels by 20 DPA. Transcript levels at the 20 DPA time point under the high temperature regimen were 3.9- and 5.3-fold higher for BQ806323

and BQ805639, respectively, than at the 34 DPA time point under moderate temperatures.

In the experiments shown in Fig. 4, globulin-2 transcript levels relative to the 18S reference RNA were higher in whole grain than in endosperm. Fig. 5 shows a direct comparison of globulin-2 transcript levels in whole grain at 30 DPA and in embryo and endosperm dissected from 30 DPA grain produced under the moderate temperature regimen. In this analysis, transcripts corresponding to BQ806323 and BQ805639 were 10.1- and 7.8-fold higher, respectively, in whole grain than in endosperm. BQ806323 and BQ805639 transcripts were most abundant in embryo, about 2.5- and 2.6-fold higher, respectively, in embryo than in whole grain. Consistent with these findings is the fact that a significant proportion of the ESTs that comprise the globulin-2 contigs were derived from an embryo cDNA library, including 9/22 for contig 18428, 4/11 for contig 17295 and 2/11 for contig 17366. The collection of ESTs assembled into contigs by Chao et al. (2006) included 4147 sequences from endosperm and 1916 sequences from embryo in addition to the 3639 from Butte 86 whole grain. RNA prepared from awns, glumes, stems, leaves and roots also was analyzed in the same experiment. Transcripts corresponding to BQ805639 were detected at very low levels in roots (Fig. 5B), but were not detectable in awns, glumes, stems or leaves. BQ806323 was not expressed in any of the non-grain samples (data not shown).

3.4. Globulin-2 proteins in flour produced under moderate and high temperatures

A proteomic analysis was undertaken to determine the levels of globulin-2 proteins in white flour milled from grain produced under moderate and high temperatures with and without post-anthesis NPK. This analysis revealed a complex collection of globulin-2 proteins in the KCl-soluble/MeOH-insoluble fraction (Fig. 6A–C). Seventeen spots in the 52–55 kDa region of the 2-D gels were identified as globulin-2. Spots 219, 209, 199, 200, 193, 190 and 258 in Fig. 6C correspond to spots 147, 155, 559, 561, 851, 852 and 175 identified in Vensel et al. (2005). MS/MS data used to identify the remaining spots (192, 202, 205, 221, 222, 228, 236, 239, 320, and 210) as globulin-2 proteins are summarized in Table S1. One of the spots (210) contained more than one protein and was excluded from further analysis. Collectively, the 16 remaining spots comprise about 2.9 and 3.1% of the total protein in the fraction when grain was produced under a moderate temperature regimen with or without post-anthesis NPK, respectively. When grain was produced under a high temperature regimen from anthesis to maturity, the 16 spots comprised a larger percentage of the total protein in the fraction, 5.2% when plants were supplied with NPK and 5.0% when plants did not receive post-anthesis NPK. Of the 16 spots, nine showed statistically significant increases with high temperatures ($P < 0.05$) under at least one of the fertilizer regimens. Seven spots increased more than 2-fold in response to high temperature under both fertilizer regimens (192, 193, 199, 202, 205, 209 and 320) (Fig. 6C, D). The largest change with high temperature was a 8.3-fold increase detected for spot 199 when grain was produced without NPK. When grain was produced with NPK, the largest response was a 4.8-fold increase for spot 320. Four spots (200, 221, 228 and 239) showed small decreases with high temperatures. The other three spots identified as globulin-2 proteins encompassed similar proportions of the KCl-soluble/MeOH-insoluble flour fraction under all four environmental regimens (222, 236 and 258). Five of the seven spots that increased more than 2-fold (193, 199, 202, 209 and 320) also increased in flour produced under high temperatures in a second growth experiment in which plants were supplied with $0.5\times$ NPK (data not shown).

QF123	BQ806323	481	GGTCCAAATGGCTCCGC	497
	BQ839060	511	GCTCCAAGTGGTTCGG	527
	BQ805639	516	GCTCCAAGTGGTTCGG	532
QR123	BQ806323	573	GCCCGCAGTCCTTCTTG	590
	BQ839060	603	AGCCGCGAGTCGTTCTTC	620
	BQ805639	608	AGCCGCGAGTCCTTCTTC	625
QF126	BQ805639	532	GGTCGTCATGCTCCTCAAC	550
	BQ806323	497	CGTCGTCATGCTCCTCAAC	515
	BQ839060	527	GGTCGTCATGCTCCTCAAC	545
QR126	BQ805639	612	CACAGTCCTTCTTCAGCGTC	631
	BQ806323	577	CGCAGTCCTTCTTCAGCGTC	596
	BQ839060	607	CGCAGTCCTTCTTCAGCGTA	626

Fig. 3. Comparison of the nucleotide sequences of BQ806323, BQ839060 and BQ805639 in regions used for gene-specific primers. QF123 is identical to nucleotides 481–497 of BQ806323 and QR123 is the reverse complement of nucleotides 573–590 of BQ806323. QF126 is identical to nucleotides 532–550 of BQ805639 and QR126 is the reverse complement of nucleotides 612–631 of BQ805639. Nucleotides that differ in other ESTs are underlined.

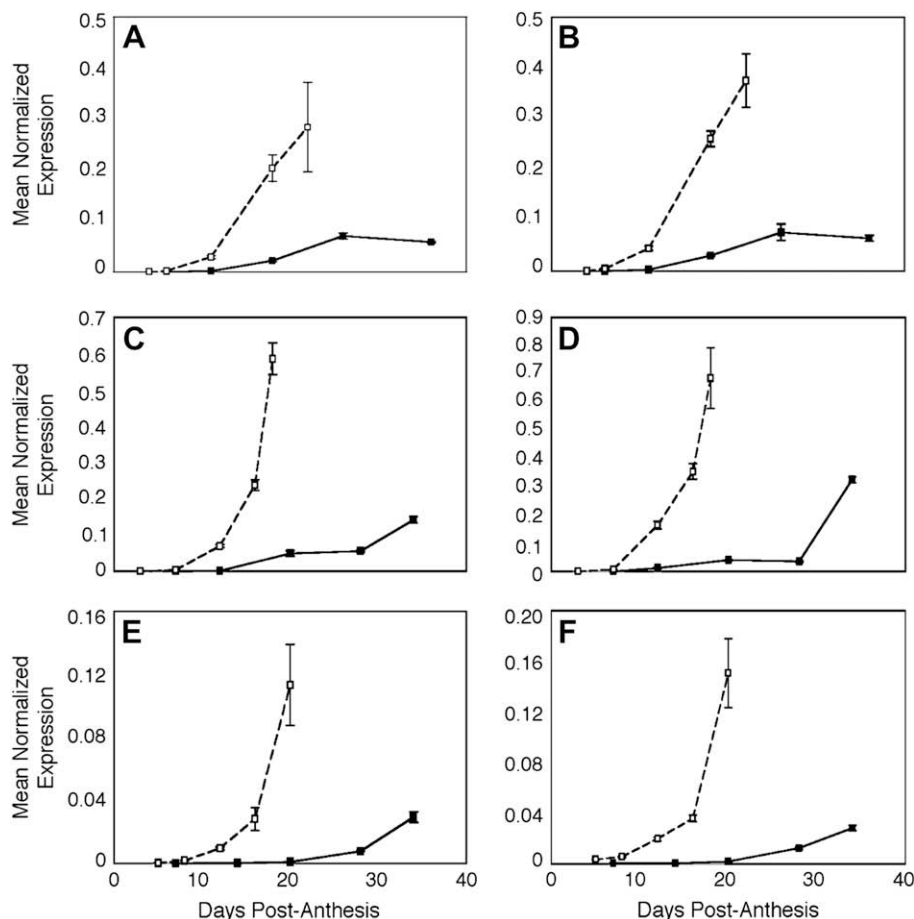


Fig. 4. Accumulation of transcripts corresponding to BQ806323 (A, C, E) and BQ805639 (B, D, F) in developing grain (A–D) or endosperm (E, F) produced under a 24/17 °C regimen (solid lines) or a 37/28 °C regimen (dashed lines). Plants in A and B did not receive post-anthesis NPK, while plants in C and D received 1× NPK and plants in E and F received 0.5× NPK. Bars indicated the standard error among triplicate reactions.

4. Discussion

In dicots, the 7S globulins are embryo proteins that serve a storage function and are the primary source of amino acids for the developing seedling. In wheat, the 7S globulins make relatively small contributions to the storage reserves of the grain and the primary storage proteins are the gliadins and glutenin subunits. While Halford and Shewry (2007) suggested that globulin proteins

in cereals have no biological role other than storage, data presented here suggests that some of the wheat globulin-2 proteins may be involved in the response of the developing grain to high temperature stress. Transcript profiles demonstrate that two closely-related globulin-2 genes are expressed in whole grain and endosperm late in grain development and that levels of transcripts increase in response to high temperatures. Globulin-2 proteins also encompass a greater proportion of the total KCl-soluble/MeOH-insoluble

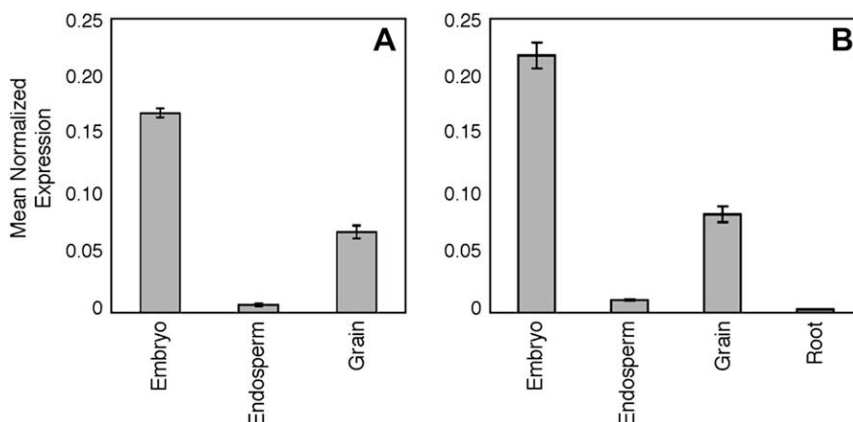


Fig. 5. Accumulation of transcripts corresponding to BQ806323 (A) and BQ805639 (B) in embryo, endosperm and whole grain, all at 30 DPA, and in root. Bars indicate the standard error among triplicate reactions.

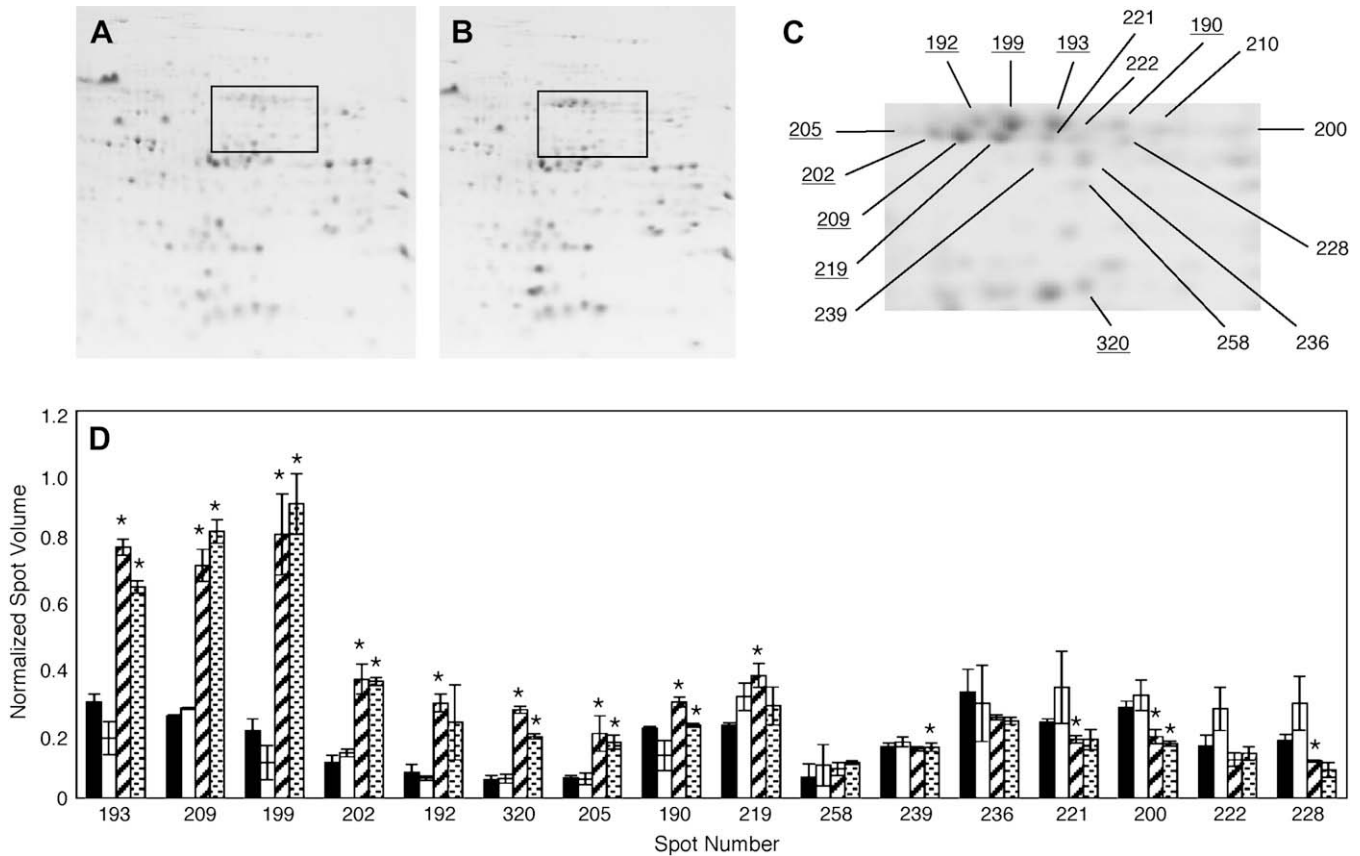


Fig. 6. 2-DE analysis of proteins from flour produced under a 24/17 °C (A) and a 37/28 °C (B) regimen with 1 × NPK. Boxes indicate regions of each gel that contain spots identified by MS/MS as globulin-2. The boxed region of the gel from panel B is enlarged in panel C and spots identified as globulin-2 are numbered. Numbers that are underlined refer to spots that increase in relative proportion under high temperature conditions. Panel D compares the average normalized volumes for each globulin-2 spot from plants subjected to the 24/17 °C and 37/28 °C regimens, with and without post-anthesis NPK. The solid and open bars denote the 24/17 °C regimens with or without NPK, respectively. The hatched and stippled bars denote the 37/28 °C regimen, with or without NPK, respectively. Asterisks denote changes with high temperature under each NPK regimen that are statistically significant ($P < 0.05$).

protein in endosperm and flour from grain produced under high temperature conditions.

These findings are consistent with previous work on the homologous globulin-2 protein from maize. In an early study, [Kriz and Wallace \(1991\)](#) reported maize variants lacking globulin-2 and suggested that this protein was not essential for seed development, maturation or germination. More recent studies using proteomic approaches have shown that globulin-2 proteins respond to stress in maize. [Kollipara et al. \(2002\)](#) found that globulin-2 was among the proteins up-regulated in desiccation-tolerant maize hybrids and [Labra et al. \(2006\)](#) noted that a number of proteins identified as globulin-2 were up-regulated in maize plantlets germinated in the presence of potassium dichromate.

Analysis of wheat ESTs indicates that the wheat globulin-2 proteins are encoded by a simple gene family consisting of three members. Yet at least 17 protein spots have been identified as globulin-2 in wheat flour using 2-DE/MS, suggesting that the proteins undergo post-translational modifications that result in changes in size and *pI*. It is likely that the wheat globulin-2 proteins, like 7S seed storage proteins from legumes, are glycosylated. [Robert et al. \(1985\)](#) noted charge heterogeneity among proteins isolated from a 7S + 3S fraction of globulins from wheat grain in the 55 kDa region of 2-D gels and observed that 55 kDa proteins also were bound to a concanavalin A sepharose column. Additionally, proteins encoded by contigs 18428, 17295 and 17366 contain putative sites for N-linked glycosylation, defined as N-X-T, where X can be any amino acid except P ([Fig. 2](#)). Using affinity blotting, [Kriz \(1989\)](#) also noted that the homologous globulin-2 protein from maize was able

to bind Con A. Some 7S globulins also undergo proteolytic processing (reviewed by [Müntz, 1996](#)). It is interesting that not all of the globulin-2 protein spots identified by 2-DE increased in response to high temperatures. In fact, decreases in the levels of several globulin-2 proteins were detected in flour. Thus, both transcriptional and post-translational mechanisms may be involved in the response of globulin-2 genes to high temperatures in the developing grain.

Transcripts for globulin-2 were present at lower levels in the endosperm than in whole grain. [Halford and Shewry \(2007\)](#) noted that 7S globulins from cereals are expressed predominantly in the embryo and in the aleurone layer of the endosperm. Our analyses did not distinguish the aleurone tissue, which is a single cell layer in wheat, from the starchy endosperm tissue. Nonetheless, transcript levels clearly increased in both the whole grain and the endosperm samples. Globulin-2 proteins were present in endosperm tissue extruded from grain from which the embryo had been removed. In addition, globulin-2 proteins were present in milled white flour that generally contains minimal amounts of germ and bran.

Increases in the relative proportion of globulin-2 proteins in flour produced under high temperatures may have important implications given the potential allergenicity of these proteins. The KCl-soluble/MeOH-insoluble proteins account for about 11% of the total flour protein ([Hurkman and Tanaka, 2004](#)) and globulin-2 proteins comprise as much as 6% of this fraction under high temperature conditions. Wheat is among the top eight foods responsible for human food allergies. Allergies to wheat are characterized by a broad spectrum of symptoms including urticaria,

gastrointestinal distress, asthma and anaphylaxis and are due to sensitivities to a variety of proteins in both the gluten protein and albumin/globulin fractions (Battais et al., 2005). Since the wheat globulin-2 shows significant identity to other proteins that are known or putative allergens, further studies on the allergenicity of these proteins are warranted. IgE binding studies using sera from patients with characterized wheat allergies would be key to determining whether the globulin-2 proteins should be classified as food allergens.

An additional question is whether changes in the relative proportions of the globulin-2 proteins might alter the functional properties of the flour. It is generally accepted that the gluten proteins contribute elasticity and extensibility properties essential for the formation of wheat flour doughs. Interactions between the gluten protein matrix and other flour components also may affect rheological properties but have been little studied. Recently, the globulin-2 proteins were identified in a proteomic study of a dough liquor fraction prepared from wheat flour (Salt et al., 2005). It is interesting that the levels of the globulin-2 proteins in the dough liquor increased when salt or salt plus ascorbate was included in the dough mixture. Salt is frequently included in commercial processes to improve breadmaking properties and has been shown to influence water absorption, dough development time, dough stability and loaf volume. It is not known whether the globulin-2 proteins are associated with the gluten matrix in the absence of salt. Further study is required to determine whether these proteins play roles in flour functionality.

Acknowledgements

The authors thank Drs. Frances DuPont and Ann Blechl for critical review of the manuscript. Mention of a specific product does not constitute an endorsement and does not imply a recommendation over other suitable products.

Appendix. Supplementary data

Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.jcs.2008.07.005](https://doi.org/10.1016/j.jcs.2008.07.005).

References

- Altenbach, S.B., 1998. Quantification of individual low-molecular-weight glutenin subunit transcripts in developing wheat grains by competitive RT-PCR. *Theoretical and Applied Genetics* 97, 413–421.
- Altenbach, S.B., DuPont, F.M., Kothari, K.M., Chan, R., Johnson, E.L., Lieu, D., 2003. Temperature, water and fertilizer influence the timing of key events during grain development in a US spring wheat. *Journal of Cereal Science* 37, 9–20.
- Altenbach, S.B., Kothari, K.M., 2004. Transcript profiles of genes expressed in endosperm tissue are altered by high temperature during wheat grain development. *Journal of Cereal Science* 40, 115–126.
- Altenbach, S.B., Kothari, K.M., Tanaka, C.K., Hurkman, W.J., 2007. Genes encoding the PR-4 protein wheatwin are developmentally regulated in wheat grains and respond to high temperatures during grainfill. *Plant Science* 173, 135–143.
- Altenbach, S.B., Kothari, K.M., Tanaka, C.K., Hurkman, W.J., 2008. Expression of 9-kDa non-specific lipid transfer protein genes in developing wheat grains is enhanced by high temperatures but not by post-anthesis fertilizer. *Journal of Cereal Science* 47, 201–213.
- Battais, F., Courcoux, P., Popineau, Y., Kanny, G., Moneret-Vautrin, D.A., Denery-Papini, S., 2005. Food allergy to wheat: differences in immunoglobulin E-binding proteins as a function of age or symptoms. *Journal of Cereal Science* 42, 109–117.
- Bendtsen, J.D., Nielsen, H., von Heijne, G., Brunak, S., 2004. Improved prediction of signal peptides: SignalP 3.0. *Journal of Molecular Biology* 340, 783–795.
- Chao, S., Lazo, G.R., You, F., Crossman, C.C., Hummel, D.D., Lui, N., Laudencia-Chingcuanco, D., Anderson, J.A., Close, T.J., Dubcovsky, J., Gill, B.S., Gill, K.S., Gustafson, J.P., Kianian, S.F., Lapitan, N.L.V., Nguyen, H.T., Sorrells, M.E., McGuire, P.E., Qualset, C.O., Anderson, O.D., 2006. Use of a large-scale Triticeae expressed sequence tag resource to reveal gene expression profiles in hexaploid wheat (*Triticum aestivum* L.). *Genome* 49, 531–544.
- Craig, R., Beavis, R.C., 2004. TANDEM: matching proteins with mass spectra. *Bioinformatics* 20, 1466–1467.
- Goodman, R.E., 2006. Practical and predictive bioinformatics methods for the identification of potentially cross-reactive protein matches. *Molecular Nutrition and Food Research* 50, 655–660.
- Halford, N.G., Shewry, P.R., 2007. The structure and expression of cereal storage protein genes. *Plant Cell Monographs* 8, 195–218.
- Hurkman, W.J., Tanaka, C.K., 2004. Improved methods for separation of wheat endosperm proteins and analysis by two-dimensional gel electrophoresis. *Journal of Cereal Science* 40, 295–299.
- Hurkman, W.J., Vensel, W.H., Tanaka, C.K., Whitehand, L., Altenbach, S.B., 2009. Effect of high temperature on albumin and globulin accumulation in the endosperm proteome of the developing wheat grain. *Journal of Cereal Science* 49, 12–23.
- Kolipara, K.P., Saab, I.N., Wych, R.D., Lauer, M.J., Singletary, G.W., 2002. Expression profiling of reciprocal maize hybrids divergent for cold germination and desiccation tolerance. *Plant Physiology* 129, 974–992.
- Kriz, A.L., 1989. Characterization of embryo globulins encoded by the maize *Glb* genes. *Biochemical Genetics* 27, 239–251.
- Kriz, A.L., Wallace, N.H., 1991. Characterization of the maize globulin-2 gene and analysis of two null alleles. *Biochemical Genetics* 29, 241–254.
- Labra, M., Gianazza, E., Waitt, R., Eberini, I., Sozzi, A., Regondi, S., Grassi, F., Agradi, E., 2006. *Zea mays* L. protein changes in response to potassium dichromate treatments. *Chemosphere* 62, 1234–1244.
- Mills, E.N.C., Jenkins, J., Marigheto, N., Belton, P.S., Gunning, A.P., Morris, V.J., 2002. Allergens of the cupin superfamily. *Biochemical Society Transactions* 30, 925–929.
- Müntz, K., 1996. Proteases and proteolytic cleavage of storage proteins in developing and germinating dicotyledonous seed. *Journal of Experimental Botany* 47, 605–622.
- Muller, P.Y., Janovjak, H., Miserez, A.R., Dobbie, Z., 2002. Processing of gene expression data generated by quantitative real-time RT-PCR. *BioTechnology* 32, 1372–1379.
- Robert, L.S., Adeli, K., Altosaar, I., 1985. Homology among 3S and 7S globulins from cereals and pea. *Plant Physiology* 78, 812–816.
- Salt, L.J., Robertson, J.A., Jenkins, J.A., Mulholland, F., Mills, E.N.C., 2005. The identification of foam-forming soluble proteins from wheat (*Triticum aestivum*) dough. *Proteomics* 5, 1612–1623.
- Vensel, W.H., Harden, L., Tanaka, C.K., Hurkman, W.J., Haddon, W.F., 2002. Identification of wheat endosperm proteins by MALDI mass spectrometry and LC-MS/MS. *Journal of Biomolecular Techniques* 13, 95–100.
- Vensel, W.H., Tanaka, C.K., Cai, N., Wong, J.H., Buchanan, B.B., Hurkman, W.J., 2005. Developmental changes in the metabolic protein profiles of wheat endosperm. *Proteomics* 5, 1594–1611.
- Wallace, N.H., Kriz, A.L., 1991. Nucleotide sequence of a cDNA clone corresponding to the maize globulin-2 gene. *Plant Physiology* 95, 973–975.

Supplementary Table 1. MS/MS identification of protein spots shown in Figure 6 that were not reported previously. For most spots, the best matches of MS/MS spectral data were translated sequences of wheat contigs available at <http://wheat.pw.usda.gov/westsql/>. The best match of the translated sequence from each wheat contig to the NCBI non-redundant protein database using the BLASTP algorithm is indicated. Analytical details for each protein spot are shown and include the spectrum number, the expectation score for the spectrum (log(e)), the log of the intensity for the spectrum (log(I)), the mass of the protonated parent peptide (m+h), the deviation of the parent peptide from the matched peptide, and the peptide sequence (in caps) along with its position in the protein and flanking amino acids (lowercase). Within peptide sequences, a highlight at C indicates carbamidomethyl derivative, and at M indicates oxidation.

Spot #	Best GPM Match			% Coverage	log(e)	Best BLASTP Match	
192	NSFT03P2_Contig18428			21	-95.7	1802402A Globulin-2 (maize)	
spectrum	log(e)	log(I)	m+h	delta	z	sequence	
32.1	-1.1	1.80	1034.454	0.015	2	lsfa ²⁸ SWDAEDVGR ³⁶ gsrr	
22.1	-1.4	0.51	1189.535	0.025	2	rgsr ⁴⁰ RWQEGGDEGR ⁴⁹ srgs	
30.1	-2.6	1.05	1784.810	0.037	3	degr ⁵⁰ SGGSGRPYHF GQESYR ⁶⁵ ewak	
42.1	-3.6	1.39	1952.962	0.085	3	aapr ¹⁰⁴ AFLQPSHHDA DEIAFVR ¹²⁰ egeg	
46.1	-2.6	1.82	1084.637	0.023	2	afvr ¹²¹ EGEGVLVLLR ¹³⁰ ngkr	
45.1	-2.7	1.22	2098.093	0.038	3	fcir ¹⁴¹ EGDVIVIPAG SIVYSANTHR ¹⁶⁰ skwl	
43.1	-3.3	1.57	2310.121	0.014	3	fesk ²²³ SKGEGEIYEA SEEQIRELSR ²⁴² scsr	
33.1	-3.3	1.90	1824.861	0.023	3	fesk ²²³ SKGEGEIYEA SEEQIR ²³⁸ elsr	
20.1	-1.7	0.98	1743.8373	-0.0035	3	scsr ²⁴⁷ GGRGGGGSGS EKEDIQPR ²⁶⁴ sltg	
19.1	-6.3	1.13	1473.6933	0.0045	2	rggr ²⁵⁰ GGGGSGSEKE DIQPR ²⁶⁴ sltg	
38.1	-7.5	2.29	1641.8348	-0.0048	2	rvir ⁴⁸³ AQPESVVFVAG PQQQR ⁴⁹⁷ rvad	
202	NSFT03P2_Contig17366			26	-60.6	1802402A Globulin-2 (maize)	
spectrum	log(e)	log(I)	m+h	delta	z	sequence	
18.1	-1.4	2.66	1841.821	-0.061	3	derr ⁵⁹ SGESGRPYHF GEESFR ⁷⁴ ewak	
21.1	-4.0	2.90	1978.966	-0.054	3	aapr ¹¹³ AFLQPSHYDA DEIAFVR ¹²⁹ egeg	
23.1	-2.9	3.11	1084.637	-0.049	2	afvr ¹³⁰ EGEGVLVLLR ¹³⁹ ngkr	
22.1	-5.4	2.80	2132.078	-0.023	3	fcvr ¹⁵⁰ EGDVVFVIPAG SIVYSANTHR ¹⁶⁹ skwf	
6.1	-4.3	2.65	1761.837	-0.027	3	fetk ²³³ SRGEGQISEG SEEQIR ²⁴⁸ elsr	
4.1	-5.6	2.90	1518.703	-0.041	2	tksr ²³⁵ GEGQISEGSE EQIR ²⁴⁸ elsr	
1.1	-3.1	2.65	1530.7147	0.0082	3	rggr ²⁶⁰ GGGGSGSEK EDIQPR ²⁷⁵ sltg	
205	NSFT03P2_Contig17366			14	-30.1	1802402A Globulin-2 (maize)	
spectrum	log(e)	log(I)	m+h	delta	z	sequence	
2.1	-1.7	2.42	972.472	0.015	2	gdyr ¹⁰⁴ VA C LDAAPR ¹¹² afllq	

13.1	-4.2	2.54	1978.966	-0.070	3	aapr ¹¹³ AFLQPSHYDA DEIAFVR ¹²⁹ eg
14.1	-3.5	2.86	1084.637	-0.064	2	afvr ¹³⁰ EGEGVLVLLR ¹³⁹ ngkr
1.1	-5.6	2.29	1518.703	-0.151	2	tksr ²³⁵ GEGQISEGSE EQIR ²⁴⁸ elsr

221	NSFT03P2_Contig17366	20	-69.1	1802402A Globulin-2 (maize)
-----	----------------------	----	-------	--------------------------------

spectrum	log(e)	log(I)	m+h	delta	z	sequence
25.1	-3.3	2.40	1841.821	-0.027	3	derr ⁵⁹ SGESGRPYHF GEESFR ⁷⁴ ewak
10.1	-1.6	2.84	972.4721	-0.0052	2	gdyr ¹⁰⁴ VAQLDAAPR ¹¹² aflq
33.1	-5.6	2.96	1978.966	-0.068	3	aapr ¹¹³ AFLQPSHYDA DEIAFVR ¹²⁹ eg
36.1	-3.3	3.27	1084.637	-0.062	2	afvr ¹³⁰ EGEGVLVLLR ¹³⁹ ngkr
8.1	-4.2	2.91	1761.837	-0.024	3	fetk ²³³ SRGEGQISEG SEEQIR ²⁴⁸ elsr
7.1	-6.1	3.06	1518.703	-0.045	2	tksr ²³⁵ GEGQISEGSE EQIR ²⁴⁸ elsr
3.1	-4.0	2.91	1530.715	-0.048	3	rggr ²⁶⁰ GGGGGSGSEK EDIQPR ²⁷⁵ sltg
2.1	-1.8	1.97	1737.743	-0.016	3	ehgr ³⁸⁴ RSEEEHGHG GEQEK ³⁹⁸ srgy

222	NSFT03P2_Contig17366	26	-70.6	1802402A Globulin-2 (maize)
-----	----------------------	----	-------	--------------------------------

spectrum	log(e)	log(I)	m+h	delta	z	sequence
17.1	-5.7	2.38	1841.821	-0.052	3	derr ⁵⁹ SGESGRPYHF GEESFR ⁷⁴ ewak
24.1	-6.6	2.86	1978.966	-0.042	3	aapr ¹¹³ AFLQPSHYDA DEIAFVR ¹²⁹ eg
27.1	-3.1	3.00	1084.637	-0.092	2	afvr ¹³⁰ EGEGVLVLLR ¹³⁹ ngkr
26.1	-5.7	2.94	2132.078	-0.080	3	fcvr ¹⁵⁰ EGDVFVIPAG SIVYSANTHR ¹⁶⁹ skwf
5.1	-5.3	2.58	1761.837	-0.016	3	fetk ²³³ SRGEGQISEG SEEQIR ²⁴⁸ elsr
4.1	-4.9	2.78	1518.703	-0.038	2	tksr ²³⁵ GEGQISEGSE EQIR ²⁴⁸ elsr
1.1	-6.0	2.67	1530.715	-0.028	3	rggr ²⁶⁰ GGGGGSGSEK EDIQPR ²⁷⁵ sltg

228	NSFT03P2_Contig17366	26	-65.8	1802402A Globulin-2 (maize)
-----	----------------------	----	-------	--------------------------------

spectrum	log(e)	log(I)	m+h	delta	z	sequence
22.1	-6.0	2.63	1841.821	-0.099	3	derr ⁵⁹ SGESGRPYHF GEESFR ⁷⁴ ewak
30.1	-5.6	3.03	1978.966	-0.037	3	aapr ¹¹³ AFLQPSHYDA DEIAFVR ¹²⁹ eg
33.1	-2.8	3.04	1084.637	-0.057	2	afvr ¹³⁰ EGEGVLVLLR ¹³⁹ ngkr
32.1	-3.1	3.04	2132.078	-0.072	3	fcvr ¹⁵⁰ EGDVFVIPAG SIVYSANTHR ¹⁶⁹ skwf
6.1	-5.6	2.80	1761.837	-0.060	3	fetk ²³³ SRGEGQISEG SEEQIR ²⁴⁸ elsr
5.1	-5.8	2.94	1518.703	-0.065	2	tksr ²³⁵ GEGQISEGSE EQIR ²⁴⁸ elsr
1.1	-3.8	2.86	1530.715	-0.052	3	rggr ²⁶⁰ GGGGGSGSEK EDIQPR ²⁷⁵ sltg

236	NSFT03P2_Contig17366	26	-70.8	1802402A Globulin-2 (maize)
-----	----------------------	----	-------	--------------------------------

spectrum	log(e)	log(I)	m+h	delta	z	sequence
----------	--------	--------	-----	-------	---	----------

15.1	-4.4	1.32	1754.789	0.052	3	errs ⁶⁰ GESGRPYHFG EESFR ⁷⁴ ewak
22.1	-7.6	2.29	1978.966	0.110	3	aapr ¹¹³ AFLQPSHYDA DEIAFVR ¹²⁹ egeg
25.1	-3.3	2.10	1084.637	0.062	2	afvr ¹³⁰ EGEGLVLLR ¹³⁹ ngkr
24.1	-7.6	2.20	2132.078	0.114	3	fcvr ¹⁵⁰ EGDVFIIPAG SIVYSANTHR ¹⁶⁹ skwf
3.1	-5.8	1.75	1761.837	0.067	3	fetk ²³³ SRGEGQISEG SEEQIR ²⁴⁸ elsr
4.1	-4.0	1.87	1518.703	0.059	2	tksr ²³⁵ GEGQISEGSE EQIR ²⁴⁸ elsr
2.1	-4.1	1.58	1530.715	0.084	3	rggr ²⁶⁰ GGGGGSGSEK EDIQPR ²⁷⁵ sltg

239 NSFT03P2_Contig17366 24 -65.8 1802402A
Globulin-2 (maize)

spectrum	log(e)	log(I)	m+h	delta	z	sequence
8.1	-1.3	2.89	972.472	-0.024	2	gdyr ¹⁰⁴ VAQLDAAPR ¹¹² aflq
25.1	-6.9	3.07	1978.966	-0.117	3	aapr ¹¹³ AFLQPSHYDA DEIAFVR ¹²⁹ egeg
29.1	-3.5	3.05	1084.637	-0.042	2	afvr ¹³⁰ EGEGLVLLR ¹³⁹ ngkr
27.1	-5.0	3.17	2132.078	-0.133	3	fcvr ¹⁵⁰ EGDVFIIPAG SIVYSANTHR ¹⁶⁹ skwf
7.1	-4.2	3.01	1761.837	-0.072	3	fetk ²³³ SRGEGQISEG SEEQIR ²⁴⁸ elsr
6.1	-5.8	3.04	1518.703	-0.070	2	tksr ²³⁵ GEGQISEGSE EQIR ²⁴⁸ elsr
4.1	-5.3	2.39	1530.715	-0.056	2	rggr ²⁶⁰ GGGGSGSEK EDIQPR ²⁷⁵ sltg

320 NSFT03P2_Contig17295 11 -52.4 1802402A
Globulin-2 (maize)

spectrum	log(e)	log(I)	m+h	delta	z	sequence
13.1	-5.6	1.19	1856.831	0.178	3	degr ⁵⁹ SGSGSRPYH FGEESFR ⁷⁵ ewak
20.1	-3.3	1.84	1084.637	0.104	2	afvr ¹³¹ EGEGLVLLR ¹⁴⁰ ngkr
15.1	-2.7	1.36	1972.957	0.177	3	rksr ²³⁶ GEGPISEGSE EQIRELSR ²⁵³ scsr
6.1	-7.7	1.70	1487.698	0.131	2	rksr ²³⁶ GEGPISEGSE EQIR ²⁴⁹ elsr
2.1	-4.7	1.66	1530.715	0.145	3	rggr ²⁶¹ GGGGGSGSEK EDIQPR ²⁷⁶ sltg
14.1	-1.1	1.71	948.552	0.094	2	rstr ³³² IYIVVEGR ³³⁹ dgyf

210^a NSFT03P2_Contig18428 20 -75.3 1802402A
Globulin-2 (maize)

spectrum	log(e)	log(I)	m+h	delta	z	sequence
13.1	-2.0	1.26	1033.4338	0.0010	2	gsrr ⁴¹ WQEGGDEGR ⁴⁹ sogs
27.1	-1.2	0.14	1784.810	0.019	2	degr ⁵⁰ SGSGSRPYHF GQESYR ⁶⁵ ewak
41.1	-4.3	1.59	1952.962	0.085	3	aapr ¹⁰⁴ AFLQPSHHDA DEIAFVR ¹²⁰ egeg
47.1	-4.8	1.61	2098.093	0.038	3	fcir ¹⁴¹ EGDVFIIPAG SIVYSANTHR ¹⁶⁰ skwl
49.1	-4.8	1.10	1416.7672	-0.0022	2	kwlr ¹⁶⁶ VVMFINPVST PGR ¹⁷⁸ fgef
42.1	-4.7	1.74	2310.121	0.014	3	fesk ²²³ SKGEGEIEYA SEEQIRELSR ²⁴² scsr
12.1	-6.2	1.27	1473.6933	0.0045	2	rggr ²⁵⁰ GGGGGSGSEKE DIQPR ²⁶⁴ sltg
37.1	-7.5	2.27	1641.835	0.021	2	rvir ⁴⁸³ AQPESVFIIVAG PQQQR ⁴⁹⁷ rvad

210^a gi|804656| 8.1 -19.8 Beta-glucosidase (barley)

spectrum	log(e)	log(I)	m+h	delta	z	sequence
----------	--------	--------	-----	-------	---	----------

28.1	-1.2	0.73	1082.5555	0.0068	2	eyhr ⁹⁹	YKEDVGIMK	¹⁰⁷	nmgf
46.1	-7.7	1.91	2233.209	0.025	3	gdsr ²³⁵	TEPYIVTHNI ILSHAAAVQR	²⁵⁴	yrek
53.1	-2.7	1.33	1809.8684	0.0089	2	igpr ³⁸³	ANSDWLYIVP WGM	³⁹⁷	avty

^a Spot contains more than one protein.