Available online at www.sciencedirect.com
ScienceDirect
Plant Physiology and Biochemistry xx (2008) 1–5
Plant Physiology and Biochemistry
www.elsevier.com/locate/plaphy

Research article

Expression and *in silico* structural analysis of a rice (*Oryza sativa*) hemoglobin 5

Verónica Garrocho-Villegas [a], Genoveva Bustos-Rivera [a], Julian Gough [b], Serge N. Vinogradov [c], Raúl Arredondo-Peter [a,*]

[a] *Laboratorio de Biofísica y Biología Molecular, Facultad de Ciencias, Universidad Autónoma del Estado de Morelos, Avenida Universidad 1001, Colonia Chamilpa, 62210 Cuernavaca, Morelos, México*
[b] *Genomic Sciences Center, RIKEN, Yokohama 230-0045, Japan*
[c] *Department of Biochemistry and Molecular Biology, Wayne State University, School of Medicine, Detroit, MI 48201, USA*

Received 29 October 2007

Abstract

This work reports the analysis of an additional hemoglobin (*hb*) gene copy, *hb5*, in the genome of rice. The amino acid sequence of Hb5 differs from the previously determined rice Hbs 1–4 in missing 11 residues in helix E. Transcripts of *hb5* were found to be ubiquitous in rice organs, and hormone- and stress-response promoters exist upstream of the rice *hb5* gene. Furthermore, the modeled structure of Hb5 based on the known crystal structure of rice Hb1 is unusual in that the putative distal His is distant from the heme Fe. This observation suggests that Hb5 binds and releases O_2 easily and thus that it functions as an O_2-carrier or in some aspects of the O_2 metabolism.

Keywords: Evolution; Function; Gene family; Modeling; Non-symbiotic; Promoter

1. Introduction

Three types of hemoglobins (Hbs) have been identified in plants: symbiotic, non-symbiotic and 2/2-like Hbs [4]. Symbiotic Hbs are specifically localized in nodules of N_2-fixing plants, and their apparent function is to facilitate the diffusion of O_2 to the symbiotic bacteria. Non-symbiotic Hbs (nsHbs) are widespread in land plants, as they have been detected in primitive bryophytes and gymnosperms and angiosperms. The nsHbs are localized in diverse plant organs, overexpress in plants subjected to stress conditions, and their apparent function *in vivo* is to modulate the levels of ATP and NO (nsHbs class 1) and to facilitate the diffusion of O_2 (nsHbs class 2). The 2/2-like plant Hbs have sequence similarity to microbial 2/2 (truncated) Hbs, have been detected in diverse organs of flowering plants, and their function remains unclear.

Non-symbiotic Hbs have been detected in a number of monocot species [1,2,8,16]. A family of the *nshb* gene, consisting of four (*hb1–4*) gene copies, was identified in the first fully sequenced plant genome, that of rice (*Oryza sativa*) [9]. The expression analysis showed that *hb1* and *hb2* are functional genes, that are expressed differentially [2], and that stress conditions up-regulate the synthesis of nsHbs in rice leaves and roots [10]. However, the existence of potential promoters upstream of the rice *hb1–4* genes indicates that the regulation of the *nshb* genes is complex [9]. The use of anti-rice Hb1 antibodies revealed that nsHb proteins are localized in specialized tissues of rice organs [12]. This observation suggests that nsHbs function in specific (not housekeeping) cell metabolisms.

Here, we report the identification of an additional *nshb* gene copy, *hb5*, in the rice genome using *in silico* tools from the SUPERFAMILY database, the cloning and characterization of rice Hb5, and the results of an analysis of its expression

Abbreviations: Hb, hemoglobin; nsHb, non-symbiotic hemoglobin.

* Corresponding author. Tel.: +52 (777) 329 7020; fax: +52 (777) 329 7040.
E-mail address: ra@buzon.uaem.mx (R. Arredondo-Peter).

0981-9428/$ - see front matter © 2008 Elsevier Masson SAS. All rights reserved.
doi:10.1016/j.plaphy.2008.05.004

in rice embryonic and vegetative organs and of modeling its structure based on the known crystal structure of rice Hb1.

2. Methods

Homologs to the globin fold were searched in the rice (*Oryza sativa* var. japonica) genome using tools from the SUPERFAMILY database (version 1.69) [6]. Resulting scaffold (protein) sequences with similarity to globins were used as a query to identify the gene sequence that codes for each scaffold in the GenBank and rice (http://www.tigr.org) genome databases.

2.1. DNA isolation, PCR amplification, and gene cloning and sequencing

Primers for the amplification of rice *hb5* (see below) were synthesized from the above gene sequences. Rice (*O. sativa* var. Morelos) total DNA was isolated by the CTAB method [3] from 14 to 16 weeks old plants growing in a crop field near the city of Cuautla, state of Morelos, México. The full *hb5* gene was amplified by PCR using Forward and Reverse rice Hb5 primers (5′-ATGGGGTTCAGCGAGACGC-3′ and 5′-TTAGGCAGCCTTCTTCAT-3′, respectively) and total rice DNA as template; amplification was carried out for 35 cycles at 55 °C for annealing essentially as described by Arredondo-Peter et al. [2]. PCR fragments were analyzed by agarose gel electrophoresis, isolated from the gel by using the GenClean kit (Bio101), cloned into the pCR2.1 vector (Invitrogen), and fully sequenced in both directions.

2.2. Isolation of poly(A^+) RNA, transcript amplification by RNA-PCR and Southern blotting

Poly(A^+) RNA was isolated from rice organs using the QuickPrep micro mRNA purification kit (Amersham-Biosciences) and following the manufacturer's instructions. Expression of the rice *hb5* gene was examined by RNA-PCR using the Forward and Reverse rice Hb5 primers and poly(A^+) RNA as template under the same conditions as those described for the amplification of the rice *hb5* gene (see above). Transcripts for Hb5 (498 bp in length) were detected by gel electrophoresis and Southern blotted at high stringency (60 °C) using the rice *hb5* gene as probe as described by Arredondo-Peter et al. [2].

2.3. In silico analyses

Multiple sequence alignment and cluster analysis of rice Hbs and selected plant Hbs were performed by using the Neighbor Joining Method of the Clustal X program [18]. The tertiary structure of rice Hb5 was predicted by homology modeling using the crystal structure of rice Hb1 (PDB ID 1D8U) as template as described by Gopalasubramaniam et al. [5]. Potential promoter sequences were searched within a 1 Kb region located upstream of the rice *hb5* gene (GenBank accession number AC130603) by using the PLACE program (http://www.dna.affrc.go.jp/htdocs/PLACE/).

3. Results and discussion

3.1. Detection, PCR amplification and characterization of the rice hb5 gene

The search for homologs to the globin fold in the SUPERFAMILY database revealed that a protein sequence (scaffold874_6) with high similarity (~60%) to the rice Hbs 1–4 is coded into the rice (*O. sativa* var. japonica) genome. Sequence of scaffold874_6 is also similar to several plant Hbs, and contains distal and proximal His and PheB10 and PheCD1 (see below), which are highly conserved in plant and non-plant Hbs. Thus, we named scaffold874_6 as rice Hb5. *In silico* mapping in the GenBank database showed that the *hb5* gene is located at the rice chromosome 5 and that the *hb1–4* and *hb2* clusters [9] are located at the rice chromosome 3.

In order to experimentally verify the existence of sequences coding for scaffold874_6 in the rice genome the *hb5* gene was amplified by PCR from template rice (*O. sativa* var. Morelos) DNA. The resulting PCR fragment was 1,113 bp in length, and cloning and DNA sequencing analysis revealed that it has an open reading frame that is 97% similar to that coding for scaffold874_6. Thus, this fragment corresponds to the rice *hb5* gene. The *hb5* sequence amplified from the rice var. Morelos DNA is deposited in the GenBank database under the accession number EF061459. The *hb5* gene from rice var. japonica and rice var. Morelos codes for a predicted polypeptide 145 amino acids in length with a calculated molecular weight of 16,272 Da. Furthermore, it comprises three introns located in the same positions as all the known plant (symbiotic and non-symbiotic) *hb* genes. The amino acid sequences of Hb5 from rice var. japonica and rice var. Morelos differ at positions 58, 98, 103, and 123 (not shown), however these substitutions are mostly conservative: N ↔ S, Y ↔ H, K ↔ R, and K ↔ R, respectively. Sequence alignment of Hb5 with Hbs 1–4 shows conservation of the distal and proximal His and of PheB10 and PheCD1 (positions 73 for Hb5 and 79 for rice Hb1 to 4, 114, 44 and 58, respectively); however, there exists an 11 amino acid deletion in the helix E of Hb5 (position 75–85) (Fig. 1), which apparently generates an unusual predicted folding for Hb5 (see below). Of the five residues postulated to facilitate the formation of dimers in rice Hb1, Val50, Ser53, Phe129, Asp125 and Val126 [7], only the first three are conserved in Hb5, suggesting that *in vivo* rice Hb5 is a monomer. Phenetic analysis of rice Hb5 and selected plant Hbs, including to rice Hbs 1–4, showed that Hb5 is divergent from the rice Hbs 1–4 family (Fig. 2). This and the above observations suggest that the evolutionary events that occurred during the divergence of rice Hb5 involved a gene duplication and copy translocation from chromosome 3 to 5 and subsequent deletion of 11 amino acids in the helix E.

Please cite this article in press as: Veró. Garrocho-Villegas et al., Plant Physiol. Biochem. (2008), doi:10.1016/j.plaphy.2008.05.004

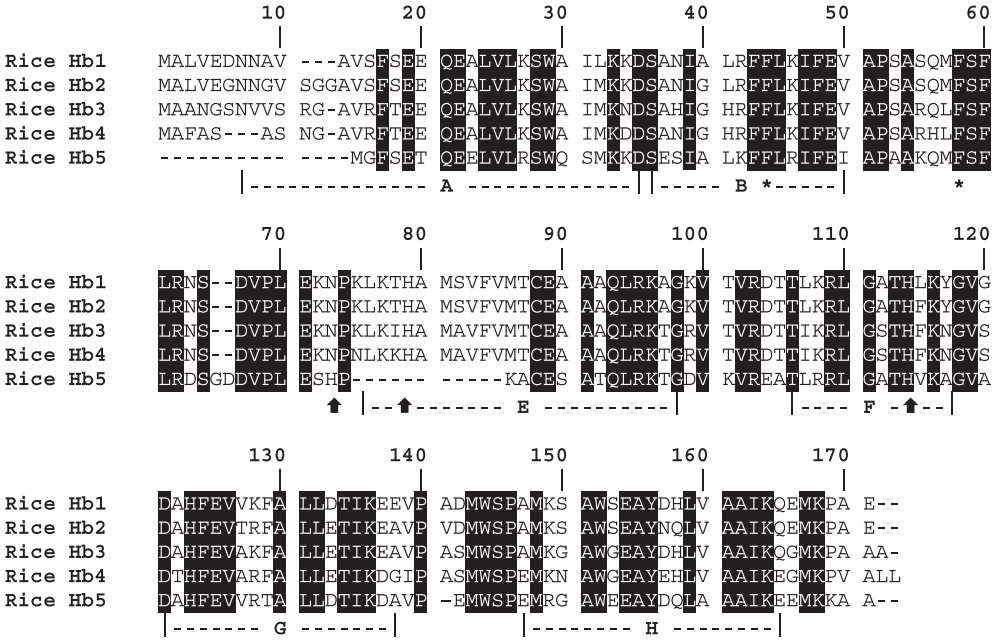

Fig. 1. Sequence alignment of rice nsHbs. Conserved amino acids are shown with black background. Arrows and asterisks show distal and proximal His at positions 79 (73 in rice Hb5) and 114, and Phe B10 and CD1 at positions 44 and 58, respectively. Note the 11 amino acids deletion in rice Hb5 at position 75–85.

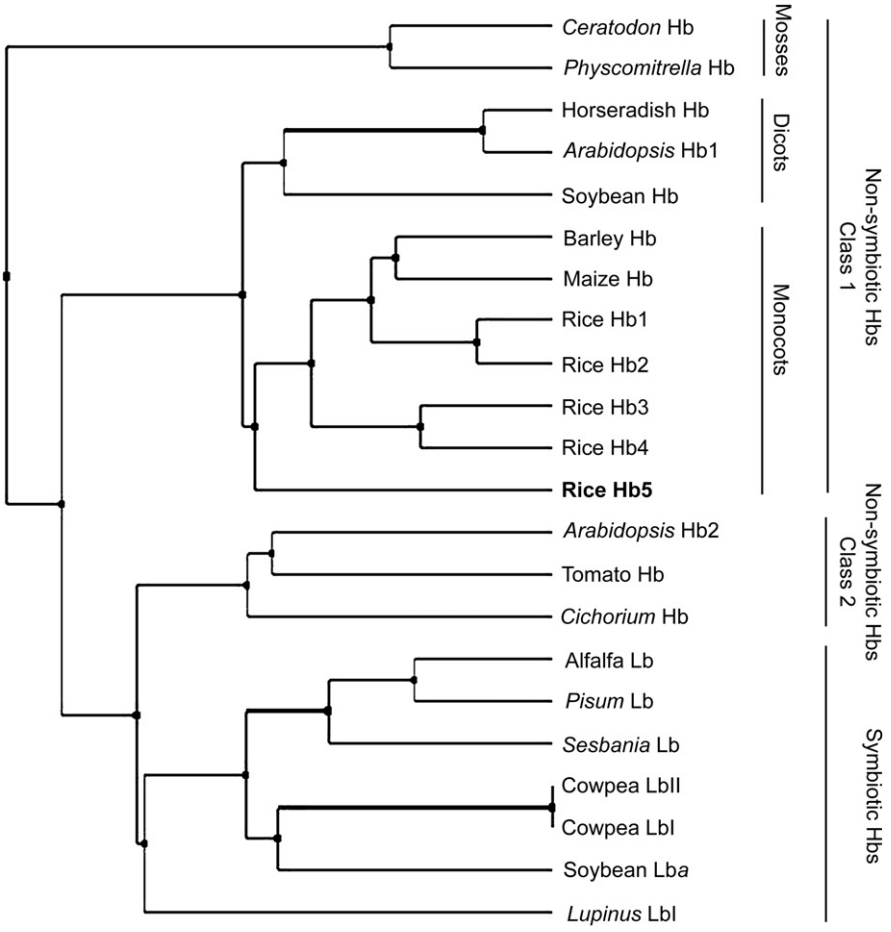

Fig. 2. Phenetic relationships of rice Hb5 and selected plant Hbs. The phenogram was constructed from Hb sequences aligned (not shown) using the Neighbor-Joining Method of the Clustal X program. Protein sequences were obtained from the GenBank database using the accession numbers reported by Garrocho-Villegas et al. [4].

Please cite this article in press as: Veró. Garrocho-Villegas et al., Plant Physiol. Biochem. (2008), doi:10.1016/j.plaphy.2008.05.004

3.2. Structural characterization of predicted rice Hb5

Sequences of rice Hb5 and Hb1 are 67% similar, thus the tertiary structure of rice Hb5 was predicted by homology modeling using the crystal structure of rice Hb1 as a template. Fig. 3A shows that predicted structure of rice Hb5 folds into the globin fold, however overlaying of rice Hb5 and Hb1 structures revealed that prehelix A and GH-loop are shorter in Hb5 than in Hb1. Moreover, the length of the CD-loop and helix E in the predicted Hb5 structure are unusually long and short, respectively, compared to other Hbs. We attribute the unusual structure of the predicted CD-loop and helix E to the deletion described above for the helix E of Hb5 (Fig. 1). Examination of the amino acid residues essential for ligand binding showed that, although the positions of the proximal His, PheB10 and PheCD1 are conserved within the predicted Hb5 structure relative to the native Hb1, the distal His is much farther away from the heme Fe within the predicted Hb5 structure (Fig. 3B), as would be predicted from the alignment shown in Fig. 1. This observation suggests that the O_2-binding properties of rice Hb5 are different from those of rice Hb1: specifically, we predict that the O_2-association and dissociation rate constants for Hb5 are high, and that consequently this protein functions in plant tissues as an O_2-carrier or in some aspects of the O_2 metabolism.

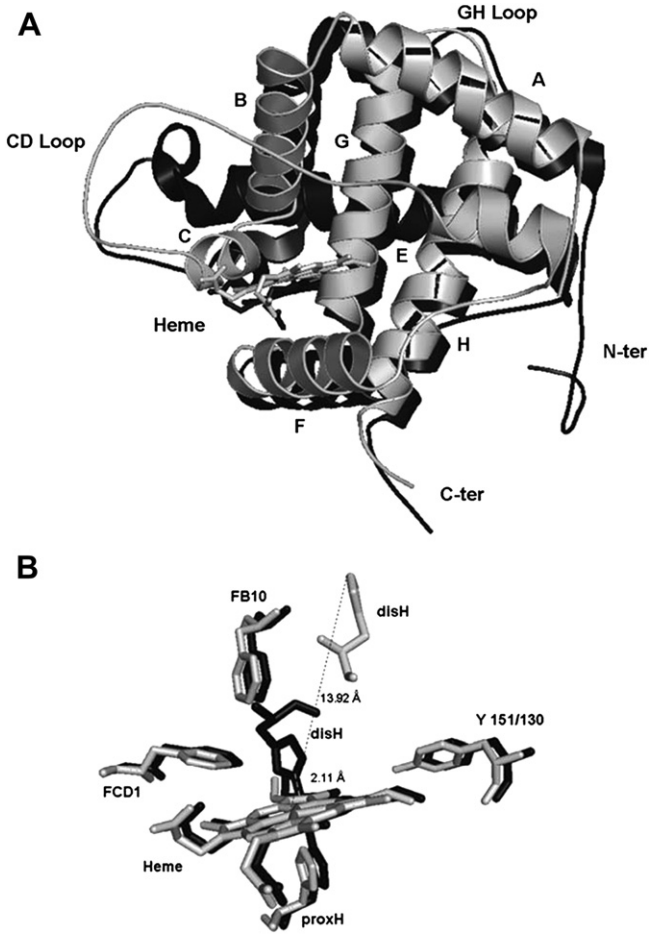

Fig. 3. (A) Overlay of predicted rice Hb5 (gray) and native Hb1 (black) tertiary structures. Helices (including pre-helix A) are indicated with letters A–H. (B) Overlay of selected amino acids in the rice predicted Hb5 (gray) and native Hb1 (black) heme pocket. DisH and proxH are distal and proximal His, respectively. The distance from distal His Nε to the heme Fe is shown with dashed lines.

3.3. Expression analysis and identification of promoter sequences upstream to the rice hb5 gene

In order to function *in vivo* *hb5* must be expressed in rice organs. Expression analysis by RNA-PCR and Southern blotting showed that transcripts for Hb5 exist in rice embryonic (embryos, coleoptiles and seminal roots) and vegetative (leaves and roots) organs (Fig. 4A). This evidence indicates that the expression of *hb5* is ubiquitous in rice plants, and thus suggests that Hb5 plays a role in the physiology of rice, probably by transporting O_2 for the aerobic metabolism. Also, the search for promoter sequences using the PLACE database revealed that a canonical TATA box and hormone- and stress-response promoters exist upstream of the rice *hb5* gene (Fig. 4B). Specifically, this observation suggests that the expression of the *hb5* gene is modulated by nitrates (GATA box [17]), phosphates (P1BS box [14]), giberellic acid (pyrimidine box [11,19]), cytokinines (ARR1 and GATCTT boxes [13]), and drought conditions (DRE box [20]); also, a nodulin promoter (the CTCTT box [15]), such as those that regulate the expression of symbiotic Hbs into the nodules of N_2-fixing plants, were identified upstream of the rice *hb5* gene.

4. Conclusions

The results reported in this work show that the rice *hb* gene family is formed by five copies (*hbs1–5*), and that rice *hb5* is ubiquitously expressed in embryonic and vegetative plant organs, although it might be (up-) regulated by plant hormones and stress conditions. From the structural predictions described above, we hypothesize that a probable function for

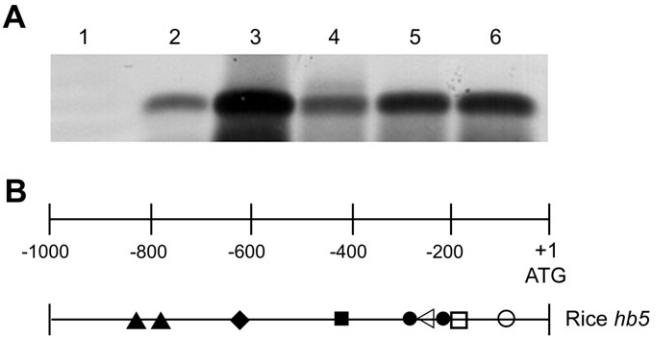

Fig. 4. (A) Expression of the *hb5* gene in rice embryonic and vegetative organs. Southern blot of the rice Hb5 transcripts (498 bp in length) amplified by RNA-PCR using the rice *hb5* gene as a probe. Lane 1, negative control (1 kb ladder (Invitrogen)); lane 2, coleoptiles; lane 3, seminal roots; lane 4, embryos; lane 5, leaves; and lane 6, roots. (B) Promoter sequences located upstream of the rice *hb5* gene. ■, TATA box; ○, DRE; □, pyrimidine box; ●, ARR1; ◁, GATA box; ◆, P1BS; ▲, nodulin CTCTT.

Please cite this article in press as: Veró. Garrocho-Villegas et al., Plant Physiol. Biochem. (2008), doi:10.1016/j.plaphy.2008.05.004

Hb5 is to carry O_2 for the aerobic metabolism of normal and stressed plants.

Acknowledgements

This work was funded by PROMEP (project number UAE-Mor-PTC-01-01/PTC-23) and Consejo Nacional de Ciencia y Tecnología (CoNaCyT, project number 42873-Q), México. V.G.-V. is a postdoctoral fellow financed by CoNaCyT (IdAP 9272); G.B.-R. is a recipient of an undergraduate fellowship from PROMEP and CoNaCyT (IdAP 9138), México.

References

[1] E. Aréchaga-Ocampo, J. Sáenz-Rivera, G. Sarath, R.V. Klucas, R. Arredondo-Peter, Cloning and expression analysis of hemoglobin genes from maize (*Zea mays* ssp. *mays*) and teosinte (*Zea mays* ssp. *parviglumis*), Biochim. Biophys. Acta: Gene Struct. Expr 1522 (2001) 1–8.

[2] R. Arredondo-Peter, M.S. Hargrove, G. Sarath, J.F. Moran, J. Lohrman, J.S. Olson, R.V. Klucas, Rice hemoglobins: gene cloning, analysis and oxygen-binding kinetics of a recombinant protein synthesized in *Escherichia coli*, Plant Physiol 115 (1997) 1259–1266.

[3] J. Crose, B. Amorese, Isolation of plant DNA from fresh tissue, Focus 9 (1978) 3–6.

[4] V. Garrocho-Villegas, S.K. Gopalasubramaniam, R. Arredondo-Peter, Plant hemoglobins: what we know six decades after their discovery, Gene: Funct. Genom. 398 (2007) 78–85.

[5] S.K. Gopalasubramaniam, V. Garrocho-Villegas, G. Bustos, N. Pastor, R. Arredondo-Peter, Use of *in silico* (computer) methods to predict and analyze the tertiary structure of plant hemoglobins, Meth. Enzymol 436 (2008) 393–410.

[6] J. Gough, K. Karplus, R. Hughey, C. Chothia, Assignment of homology to genome sequences using a library of hidden Markov models that represent all proteins of known structure, J. Mol. Biol. 313 (2001) 903–919.

[7] M. Hargrove, E.A. Brucker, B. Stec, G. Sarath, R. Arredondo-Peter, R.V. Klucas, J.S. Olson, G.N. Philips Jr., Crystal structure of a non-symbiotic hemoglobin, Structure 8 (2000) 1005–1014.

[8] K. Larsen, Molecular cloning and characterization of cDNAs encoding hemoglobin from wheat (*Triticum aestivum*) and potato (*Solanum tuberosum*), Biochim. Biophys. Acta 1621 (2003) 299–305.

[9] V. Lira-Ruan, E. Ross, G. Sarath, R.V. Klucas, R. Arredondo-Peter, Mapping and analysis of a hemoglobin gene family from rice (*Oryza sativa*), Plant Physiol, Biochem 40 (2002) 199–202.

[10] V. Lira-Ruan, G. Sarath, R.V. Klucas, R. Arredondo-Peter, Synthesis of hemoglobins in rice (*Oryza sativa* var. Jackson) plants growing in normal and stress conditions, Plant Sci. 161 (2001) 279–287.

[11] M. Mena, F.J. Cejudo, I. Isabel-Lamoneda, P. Carbonero, A role for the DOF transcription factor BPBF in the regulation of giberellin-responsive genes in barley aleurone, Plant Physiol 130 (2002) 111–119.

[12] E.J.H. Ross, L. Shearman, M. Mathiesen, J. Zhou, R. Arredondo-Peter, G. Sarath, R.V. Klucas, Non-symbiotic hemoglobins are synthesized during germination and in differentiating cell types, Protoplasma 218 (2001) 125–133.

[13] E.J.H. Ross, J.M. Stone, C.G. Elowsky, R. Arredondo-Peter, R.V. Klucas, G. Sarath, Activation of the *Oryza sativa* non-symbiotic haemoglobin-2 promoter by the cytokinin-regulated transcription factor, ARR1, J. Exp. Bot 55 (2004) 1721–1731.

[14] V.L.F. Rubio, R. Solano, A.C. Martin, J. Iglesias, A. Leyva, J. Paz-Ares, A conserved MYB transcription factor involved in phosphate starvation signaling both in vascular and in unicellular alga, Genes Dev. 15 (2001) 2122–2133.

[15] J. Stougaard, J. Jorgensen, T. Christesen, A. Kühle, K.A. Marcker, Interdependence and nodule specificity of *cis*-acting regulatory elements in the soybean leghemoglobin lbc_3 and N23 gene promoters, Mol. Gen. Genet. 220 (1990) 353–360.

[16] E.R. Taylor, X.Z. Nie, A.W. MacGregor, R.D. Hill, A cereal haemoglobin gene is expressed in seed and root tissues under anaerobic conditions, Plant Mol. Biol. 24 (1994) 853–862.

[17] W.B. Terzaghi, Light regulated transcription, Annu. Rev. Plant Physiol, Plant Mol. Biol. 46 (1995) 445–474.

[18] J.D. Thompson, T.J. Gibson, F. Plewniak, F. Jeanmougin, D.G. Higgins, The clustal X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools, Nucl. Acids Res. 24 (1997) 4876–4882.

[19] K. Washio, Functional dissections between GAMYB and Dof transcription factors suggest a role for protein-protein associations in the giberellin-mediated expression of the *RAmy1A* gene in the rice aleurone, Plant Physiol 133 (2003) 850–863.

[20] L. Xiong, K.S. Shumaker, J.K. Zhu, Cell signaling during cold, drought, and salt stress, Plant Cell Supp 2002 (2002) S165–S183.

Please cite this article in press as: Veró. Garrocho-Villegas et al., Plant Physiol. Biochem. (2008), doi:10.1016/j.plaphy.2008.05.004