

## Two hemoglobin genes in *Arabidopsis thaliana*: The evolutionary origins of leghemoglobins

BEN TREVASKIS\*<sup>†‡</sup>, RICHARD. A. WATTS<sup>†‡</sup>, CAROL R. ANDERSSON<sup>†§</sup>, DANNY J. LLEWELLYN<sup>†</sup>, MARK S. HARGROVE<sup>¶</sup>, JOHN S. OLSON<sup>¶</sup>, ELIZABETH S. DENNIS<sup>†||</sup>, AND W. JAMES PEACOCK<sup>†</sup>

\*Cooperative Research Centre for Plant Cell Biology, Australian National University, Australian Capital Territory, 2601 Australia; <sup>†</sup>Division of Plant Industry, Commonwealth Scientific and Industrial Research Organisation, GPO Box 1600, Canberra, Australian Capital Territory, 2601 Australia; <sup>‡</sup>Department of Biochemistry and Molecular Biology, Australian National University, Australian Capital Territory, 2601 Australia; and <sup>¶</sup>Department of Biochemistry and Cell Biology, Rice University, 6100 Main Street, Houston, TX 77005-1892

Communicated by Marshall D. Hatch, Commonwealth Scientific and Industrial Research, Canberra, Australia, August 27, 1997 (received for review June 27, 1997)

**ABSTRACT** We cloned two hemoglobin genes from *Arabidopsis thaliana*. One gene, *AHB1*, is related in sequence to the family of nonsymbiotic hemoglobin genes previously identified in a number of plant species (class 1). The second hemoglobin gene, *AHB2*, represents a class of nonsymbiotic hemoglobin (class 2) related in sequence to the symbiotic hemoglobin genes of legumes and *Casuarina*. The properties of these two hemoglobins suggest that the two families of nonsymbiotic hemoglobins may differ in function from each other and from the symbiotic hemoglobins. *AHB1* is induced, in both roots and rosette leaves, by low oxygen levels. Recombinant *AHB1* has an oxygen affinity so high as to make it unlikely to function as an oxygen transporter. *AHB2* is expressed at a low level in rosette leaves and is low temperature-inducible. *AHB2* protein has a lower affinity for oxygen than *AHB1* but is similar to *AHB1* in having an unusually low, pH-sensitive oxygen off-rate.

Leghemoglobins, the first hemoglobins to be identified in higher plants (1), are found in the nodules of legumes where they are believed to transport oxygen to nitrogen-fixing endosymbiotic bacteria (2). Nodule-expressed hemoglobins also have been identified in *Casuarina glauca*, a nitrogen-fixing nonlegume (3). Collectively, these nodule-expressed hemoglobins are called “symbiotic hemoglobins” (4). In addition to symbiotic hemoglobins, these plants have another hemoglobin that shows only limited amino acid sequence similarity to the symbiotic hemoglobins. Genes encoding this type of hemoglobin also have been cloned from the nitrogen-fixing elm *Parasponia andersonii* (5, 6) and from plants that do not fix nitrogen, including monocots (6–8). These “nonsymbiotic” hemoglobins are typically expressed at low levels in root tissue. The *Parasponia* hemoglobin is unusual in that it appears to be bi-functional; it transports oxygen in nitrogen-fixing nodules and, like the nonsymbiotic hemoglobins, is expressed in root tissue (6). The function of nonsymbiotic hemoglobins in normal plant tissues has not been established. To examine the function of nonsymbiotic hemoglobins, we have cloned hemoglobins from *Arabidopsis thaliana*.

### MATERIALS AND METHODS

**Plant Growth and Stress Treatments.** *Arabidopsis* C24 ecotype seeds were surface sterilized, then grown on MS agar with 3% sucrose, 40–50 plants/plate, at 22°C (16/8 h light/dark cycle) until harvest. Plants used for *AHB1* induction analysis were grown on agar without sucrose. Dehydration,

chilling, and hypoxia treatments were performed as outlined in Dolferus *et al.* (9). All treatments, except chilling (24 h), were for 12 h unless specified. Sucrose treatment involved placing plantlets in MS with 1% sucrose overnight. RNA was extracted using the method of Logemann *et al.* (10) and separated on formaldehyde gels as described by Dolferus *et al.* (9); RNA gel blotting, hybridization, and posthybridization washes also were done as described by Dolferus *et al.* RNA loadings were calibrated with an *Arabidopsis* Ubiquitin riboprobe.

**Library Screening, Subcloning, and Sequencing.** *Arabidopsis* genomic libraries (9- to 15-kb fraction of *Sau3A*-digested C24 genomic DNA in  $\lambda$ EMBL4) were screened with random primed probes (11). Screening and subsequent subcloning involved standard conditions and DNA manipulation, as described in Ausubel *et al.* (11). Automated DNA sequencing, with both dye primer and dye terminator reactions, was performed according to company specifications (Perkin-Elmer).

**Reverse Transcriptase-PCR Analysis.** Reverse transcriptase-PCR was carried out on first strand cDNA that was synthesized from RNA with Moloney murine leukemia virus reverse transcriptase as per company specifications (Promega). PCR was then performed with specific primers (100 nM) and standard PCR conditions (11). The resulting reverse transcriptase-PCR products were cloned and sequenced, confirming that the predicted introns are absent from the final transcript.

**Southern Analysis.** Genomic DNA was extracted using the method of Taylor and Powell (12). Hybridization and post hybridization treatment was as described in Ausubel *et al.* (11).

**Database Search.** Database searches were conducted through key word searching of the Expressed Sequence Tag database on the National Center for Biotechnology World Wide Web site (<http://www.ncbi.nlm.nih.gov>).

**Expression, Purification, and Analysis Of Recombinant Proteins.** cDNA clones of *AHB1* and *AHB2* were fused to an efficient ribosomal binding site, as per Springer and Sligar (13), and inserted into the pET24(+) expression vector (Novagen). *Escherichia coli*, strain BL21(DE3), containing each expression vector was grown to OD<sub>600nm</sub> of 0.6, and then recombinant protein expression was induced with 1 mM isopropyl  $\beta$ -D-thiogalactoside. After 4 h at 37°C, cells were centrifuged and resuspended in 100 mM Tris-HCl (pH 8.0), sonicated, and cleared by centrifugation; 50 mg/ml hemin-HCl was added to the soluble fraction. Recombinant *AHB1* protein was purified

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. §1734 solely to indicate this fact.

© 1997 by The National Academy of Sciences 0027-8424/97/9412230-5\$2.00/0  
PNAS is available online at <http://www.pnas.org>.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database [accession numbers U94998 (*AHB1*) and U94999 (*AHB2*)].

<sup>§</sup>Present address: Department of Biology, Texas A & M University, College Station, TX 77843-2133.

<sup>||</sup>To whom reprint requests should be addressed. e-mail: liz@pican.pi.csiro.au.

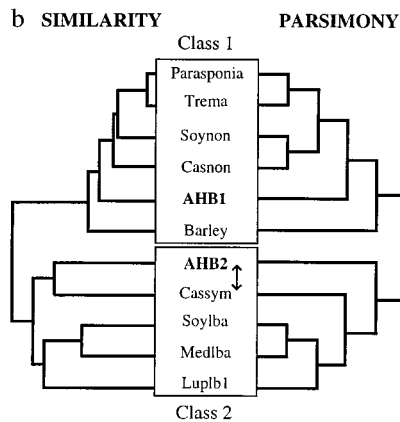
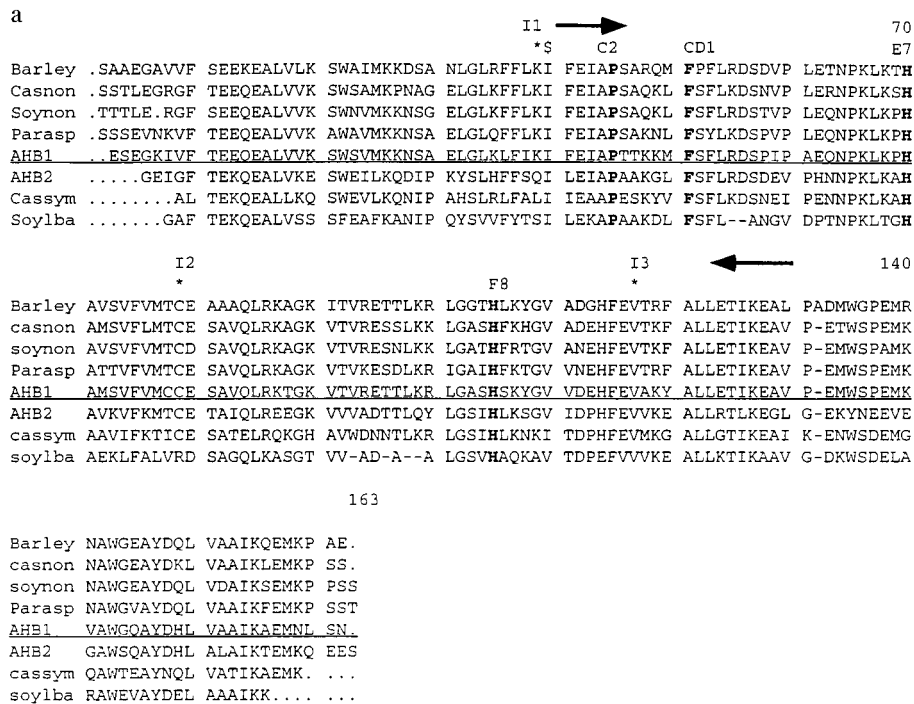


FIG. 1. (a) Alignment of some plant hemoglobin protein sequences. Proximal and distal histidines, phenylalanine CD1, and proline C2 are shown in bold. The location of primers used to amplify *AHB1* PCR fragment are shown as arrows. \$ represents the 5' end of the *AHB2* cDNA (Clone ID 11161). Intron positions are marked with stars and labeled I1-I3. Class 1 and 2 hemoglobins are separated by the horizontal line. (b) Cladistic analysis of plant hemoglobins. The left tree is a similarity tree of plant hemoglobin protein sequences created using the Genetics Computer Group PILEUP program. Class 1 and 2 hemoglobins form two distinct clades. The right tree is one of the two most parsimonious trees (CI .855, RC .654) generated by an exhaustive search of all possible trees using the PAUP, Ver. 3.1.1 program. The two trees of maximum parsimony differed in whether they grouped the *Casuarina* symbiotic hemoglobin or AHB2 with the leghemoglobins (represented by the arrow). The bootstrap value for the basal branch of this tree is 100%. The two classes of plant hemoglobins are highlighted by boxes. Hemoglobins (and GenBank accession numbers): barley: barley nonsymbiotic hemoglobin (U01228); casnon: *Casuarina* nonsymbiotic hemoglobin (X53950); soynon: soybean nonsymbiotic hemoglobin (U47143); parasp: *Parasponia* hemoglobin (U27194); AHB1: This study (U94998); AHB2: this study (U94999); cassym: *Casuarina* symbiotic hemoglobin (77695); soylba: soybean leghemoglobin A (J01299); medslba: *Medico sativa* leghemoglobin A (X14311); luplb1: lupin leghemoglobin 1 (Y00401).

by 25–70% ammonium sulfate fractionation, resuspended in 20 mM Tris·HCl (pH 8.0), and dialyzed twice into 4 liters of the same buffer. Dialyzed extracts were concentrated (YM-10 membrane, Amicon) and passed over a cation exchange column (DE52 Whatman) in 20 mM Tris·HCl (pH 8.0). Red fractions were concentrated, adjusted to pH 6.0 with dilute acetic acid, and passed over an anion exchange column (CM52, Whatman) with 20 mM sodium phosphate buffer (pH 8.0). The red fraction was concentrated to 10 mg/ml and stored at –80°C. Recombinant AHB2 protein was purified as was described for AHB1, with the following modification: 25–90% ammonium sulfate fractionation. Elution from DE52 resin was achieved with 250 mM Tris·HCl, following 5 column volumes of 100 mM Tris·HCl wash buffer. After elution from CM52

resin, red fractions were adjusted to 25% ammonium sulfate saturation (pH 8.0) and passed over a phenol-Sepharose column (equilibrated with 20 mM Tris·HCl/25% ammonium sulfate) then eluted with 20 mM Tris·HCl/12.5% ammonium sulfate and concentrated to 10 mg/ml for storage. Rate constants for O<sub>2</sub> and CO binding were measured using rapid mixing and laser photolysis methods (14).

**RESULTS**

**A Nonsymbiotic Hemoglobin Gene from *Arabidopsis*.** Using degenerate primers that amplified a nonsymbiotic hemoglobin gene fragment from soybean (7), we amplified a 459-bp fragment of a nonsymbiotic hemoglobin gene from *Arabidop-*

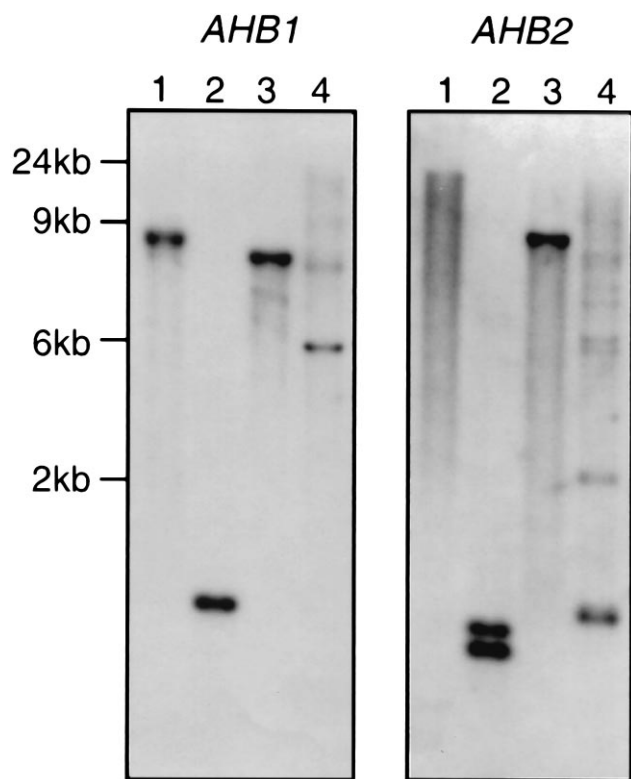


FIG. 2. Southern blots of *AHB1* and *AHB2*. *Arabidopsis* (C24 ecotype) genomic DNA digested with *Bam*HI, *Dra*I, *Eco*RI, and *Xba*I (lanes 1, 2, 3, and 4, respectively) and hybridized with *AHB1* and *AHB2* cDNA probes. (Extra bands on the *Xba*I lanes are due to partial digestion.)

*sis*. This gene fragment was used to isolate a full-length hemoglobin gene, *AHB1*, with a predicted amino acid sequence that shows strong homology to the sequences of nonsymbiotic hemoglobins from other plants (Fig. 1 *a* and *b*), including predicted ORFs containing amino acid residues that are conserved in all plant hemoglobins [the proximal F8 and distal E7 histidines, phenylalanine CD1 and proline C2 (15)]. There are three introns in positions identical to the introns occurring in all other plant hemoglobins (Fig. 1*a*). Southern analysis, at high stringency, suggests this gene is present in single or low copy number in the *Arabidopsis* genome (Fig. 2).

**A Leghemoglobin-Like Gene from Arabidopsis.** We identified a second hemoglobin gene (*AHB2*) by searching the *Arabidopsis* Expressed Sequence Tag database. The single partial length cDNA showing homology to plant hemoglobins (clone ID 11161) does not correspond in sequence to the *AHB1* gene, with only 69% identity at the nucleotide level. This gene, *AHB2*, has three introns in identical positions to those of other plant hemoglobins and, like *AHB1*, appears to be a single copy gene according to Southern blot data (*Dra*I cuts within the

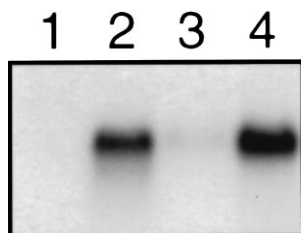


FIG. 3. *AHB1* expression in root and rosette tissues. Total RNA extracted from rosette leaves (lane 1), rosettes treated with 5% oxygen (lane 2), roots (lane 3), and roots treated with 5% oxygen (lane 4) hybridized with an *AHB1* antisense riboprobe.

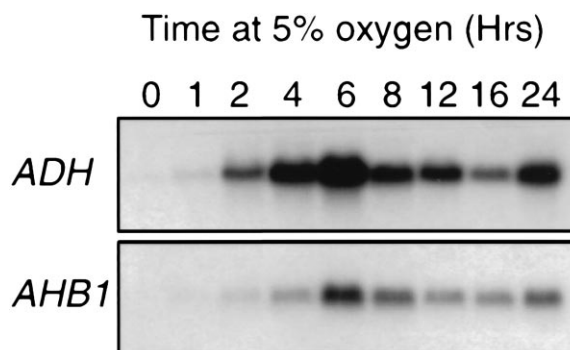


FIG. 4. Time course analysis of the *AHB1* hypoxia response. Total RNA from whole plants treated with 5% oxygen for varying times hybridized with either *AHB1* or *ADH* antisense riboprobes.

coding sequence of *AHB2*, producing two bands in this lane; Fig. 2). When the amino acid sequence of *AHB2* is aligned with other plant hemoglobin protein sequences, it appears to be more similar to the symbiotic hemoglobins of *Casuarina* and legumes than it does to nonsymbiotic plant hemoglobins (Fig. 1*a*). If the *AHB2* protein sequence is used in a BLAST (basic local alignment search tool) search, it consistently detects symbiotic hemoglobins with higher BLAST scores than most nonsymbiotic hemoglobins. A similarity tree of plant hemoglobin protein sequences, constructed with the PILEUP function of the Genetics Computer Group package, also shows *AHB2* to be more similar to the symbiotic hemoglobins of legumes and *Casuarina* (Fig. 1*b*). The same clade occurs in the two trees of maximum parsimony constructed using the PAUP program, Ver. 3.1.1, one of which is presented (Fig. 1*b*). The bootstrap value for the basal branch of this clade is high (100%), supporting the grouping of *AHB2* with the symbiotic hemoglobins of legumes and *Casuarina*.

***AHB1* Is Expressed in Root Tissue and Is Hypoxia-Inducible.** *AHB1* expression, examined with RNA gel blots, is detectable at low levels in root tissue and is strongly induced in both the roots and rosette leaves of plants subjected to hypoxic conditions (Fig. 3). The time course of the *AHB1* response to low oxygen levels is similar to that of the alcohol dehydrogenase gene (*ADH*), a well characterized hypoxic response gene (ref. 16; Fig. 4). The maximum level of *AHB1* transcript was observed after 6–8 h of hypoxia treatment. Transcription of *AHB1* is induced under a 5% oxygen environment, but stronger induction occurs at 0.1% oxygen (Fig. 5). A transcriptional response also was elicited by a 1% sucrose solution. Cold, dehydration, heat shock, oxidative stress, and wounding treatments did not affect *AHB1* transcript levels.

***AHB2* Is Expressed in Rosette Leaves.** A low level of *AHB2* expression was found in rosette leaves, but unlike *AHB1*, there

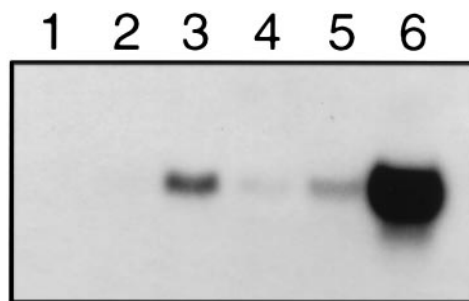


FIG. 5. *AHB1* expression during environmental stress treatments. Total RNA from control plants (lane 1) or plants subjected to chilling (lane 2), 1% sucrose (lane 3), dehydration (lane 4), 5% oxygen (lane 5), or 0.1% oxygen (lane 6) hybridized with an *AHB1* antisense riboprobe.

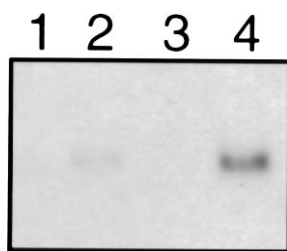


Fig. 6. Induction of *AHB2* by low temperature. Total RNA from root tissue (lane 1) or rosette leaves (lane 2) and root tissue after chilling treatment (lane 3) or rosette leaves after chilling treatment (lane 4) hybridized with an *AHB2* antisense riboprobe.

was no detectable expression in root tissue. *AHB2* expression was not affected by sucrose or hypoxia treatment, but low temperature did increase *AHB2* transcript levels (Fig. 6). The low temperature response seems to be confined to the rosette leaves with no response evident in root tissue (Fig. 6). Wounding, heat shock, oxidative stress, and dehydration had no effect on the level of *AHB2* transcript.

**Recombinant AHB1 and AHB2 Proteins Have Different Oxygen Binding Properties.** Recombinant AHB1 protein, expressed in *E. coli*, was purified, and the kinetics of O<sub>2</sub> and CO binding were measured. AHB1 shows a remarkably high O<sub>2</sub> affinity, with a P<sub>50</sub> value equal to 1.6 nM at pH 7, 20°C (Table 1). Similar oxygen binding parameters have been observed for the nonsymbiotic hemoglobins of soybean and rice (R.A.W., unpublished data, and M. Hargrove, personal communication). Lowering the pH to 5 raises the P<sub>50</sub> to 6.7 nM (Table 1), but this value is still smaller than that of any known plant cytochrome oxidase (17), making it unlikely that AHB1 could transport oxygen to these terminal oxidases. AHB1 also has a low O<sub>2</sub> dissociation rate constant compared with that of soybean leghemoglobin A, a known oxygen transporter (Table 1). Thus, AHB1 would appear to be an oxygen-scavenging protein rather than an oxygen transporter.

In contrast, recombinant AHB2 has a much lower affinity for O<sub>2</sub>, with P<sub>50</sub> values of 130 nM at pH 7 and 2,700 nM at pH 5 (Table 1). The dissociation rate constant is low at pH 7 but increases 20-fold at pH 5, approaching the value for soybean leghemoglobin A under the same acidic conditions. The low association rate constant for O<sub>2</sub> binding to AHB2 appears to be caused by the reversible formation of a hexacoordinate complex (hemochrome) in deoxygenated AHB2. The imidazole side chain of the distal histidine (E7), the only polar residue in the distal pocket, may coordinate directly to the iron atom of the heme group, thereby inhibiting rapid O<sub>2</sub> binding and facilitating O<sub>2</sub> dissociation.

Table 1. Comparison of parameters\* for O<sub>2</sub> binding to recombinant hemoglobins at 20°C

| Protein              | pH  | P <sub>50</sub> , nM | k <sub>O<sub>2</sub></sub> , s <sup>-1</sup> | M K <sub>CO</sub> /K <sub>O<sub>2</sub></sub> |
|----------------------|-----|----------------------|--|---|
| AHB1                 | 7.0 | 1.6                  | 0.12   | 2.0   |
|                      | 5.0 | 6.7                  | 0.38   |   |
| AHB2                 | 7.0 | 130                  | 0.14   | 26  |
|                      | 5.0 | 2700                 | 2.1  |   |
| Soy lbA <sup>†</sup> | 7.0 | 44                   | 5.6  | 87  |
|                      | 5.0 | 20                   | 2.7  |   |

\*Where k<sub>O<sub>2</sub></sub> and k'<sub>O<sub>2</sub></sub> are the bimolecular dissociation and association rate constants for O<sub>2</sub> binding, respectively. The dissolved O<sub>2</sub> concentration at which hemoglobin is half saturated (P<sub>50</sub>) is equivalent to the dissociation equilibrium constant for O<sub>2</sub> (K'<sub>O<sub>2</sub></sub> = k<sub>O<sub>2</sub></sub>/k'<sub>O<sub>2</sub></sub>). K<sub>O<sub>2</sub></sub> and K<sub>CO</sub> are the association equilibrium constants for O<sub>2</sub> and CO, respectively. M is the partition coefficient expressing the relative affinities for O<sub>2</sub> and CO.

<sup>†</sup>Recombinant soybean leghemoglobin A.

## DISCUSSION

*AHB1* is related to the nonsymbiotic hemoglobin genes present in other plants. The identification of this gene in *Arabidopsis* supports the idea that nonsymbiotic hemoglobin genes may be present in all plants (6). The expression pattern of *AHB1* is similar to that of other nonsymbiotic hemoglobins; low levels of expression in root tissue are common to all nonsymbiotic hemoglobins (3–8), sucrose inducibility has been observed for the *Casuarina* nonsymbiotic hemoglobin (4), and the hypoxic induction of *AHB1* resembles that of the barley hemoglobin (8).

*AHB2* represents a new class of nonsymbiotic hemoglobin gene. It shows limited homology to the other nonsymbiotic hemoglobins and has a different pattern of gene expression. *AHB2*-like genes may be present in a wide range of plants. In support of this theory, we have cloned an *AHB2* homologue from *Brassica napus* (R.A.W., unpublished data). We suggest that the *AHB1*-like “nonsymbiotic hemoglobin” gene family be referred to as class 1 plant hemoglobin genes and that *AHB2*-like genes be referred to as class 2 plant hemoglobin genes.

It has been suggested that symbiotic hemoglobin genes of legumes and *Casuarina* arose from gene duplication of a nonsymbiotic (class 1) hemoglobin gene in the ancestor of nitrogen-fixing plants (5, 6). The similarity between *AHB2* and symbiotic hemoglobins suggests an alternative explanation; a class 2 hemoglobin gene may have been the direct ancestor of the symbiotic hemoglobin genes. In contrast, it appears that a class 1 hemoglobin has been recruited into symbiotic function in *Parasponia* nodules. The recruitment of either class 1 or class 2 plant hemoglobins into symbiotic roles seems to have been associated with changes in gene expression, leading to high levels of expression in nodules, and with alterations in the oxygen binding properties of the protein.

Leghemoglobins facilitate oxygen flux down an oxygen gradient without the involvement of other proteins. It is unlikely that either AHB1 or AHB2 acts as an oxygen transporter in this manner. The oxygen affinity of AHB1 is too high for release of oxygen to cytochrome oxidase, and the oxygen dissociation rates of both proteins are too slow for efficient oxygen transport. Instead, these proteins may act as oxygen scavengers, binding oxygen until interactions with other proteins occur. The eventual destination for bound oxygen could be mitochondrial metabolism or oxygen-consuming enzymatic reactions. It is also possible that AHB1 or AHB2 have unidentified catalytic properties. Analysis of transgenic plants that over- or underexpress AHB1 or AHB2, combined with cellular localization of these proteins, should lead to a better understanding of the functions of these and other plant hemoglobins.

We thank Janice Norman for her excellent technical assistance and Cyril Appleby for his helpful discussions.

1. Kubo, H. (1939) in Appleby, C. A. (1992) *Sci. Prog.* **76**, 365–398.
2. Appleby, C. A. (1984) *Annu. Rev. Plant Physiol.* **35**, 443–478.
3. Fleming, A. I., Wittenberg, J. B., Wittenberg, B. A., Dudman, W. F. & Appleby, C. A. (1987) *Biochim. Biophys. Acta* **911**, 209–220.
4. Jacobsen-Lyon, K., Jensen, E. Ø., Jørgensen, J., Marcker, K. A., Peacock, W. J. & Dennis, E. S. (1995) *Plant Cell* **7**, 213–223.
5. Landsmann, J., Dennis, E. S., Higgins, T. J., Appleby, C. A., Kortt, A. A. & Peacock, W. J. (1986) *Nature (London)* **324**, 166–168.
6. Bogusz, D., Appleby, C. A., Landsmann, J., Dennis, E. S. & Peacock, W. J. (1988) *Nature (London)* **331**, 178–180.
7. Andersson, C. R., Jensen, E. Ø., Llewellyn, D. J., Dennis, E. S. & Peacock, W. J. (1996) *Proc. Natl. Acad. Sci. USA* **93**, 5682–5687.
8. E. R. Taylor, E. R., Nie, X. Z., MacGregor, A. W. & Hill, R. D. (1994) *Plant Mol. Biol.* **24**, 853–862.
9. Dolferus, R., Jacobs, M., Peacock, W. J. & Dennis, E. S. (1994) *Plant Physiol.* **105**, 1075–1087.

10. Logemann, J., Schell, J. & Willmitzer, L. (1987) *Anal. Biochem.* **163**, 16–20.
11. Ausubel, F. M., Brenton, R., Kingston, R. E., Moore, D. D., Siedman, J. G., Smyth, J. A. & Struhl, K. (1994) *Current Protocols in Molecular Biology* (Wiley, New York).
12. Taylor, A. & Powell, B. (1982) *BRL Focus* **4**, 4.
13. Springer, B. A. & Sligar, S. G. (1997) *Proc. Natl. Acad. Sci. USA* **84**, 8961–8965.
14. Hargrove, M. S., Barry, J. K., Bruckner, E. A., Berry, M. B., Phillips G. N., Jr., Olson. J. S., Arrendo-Peter, R., Dean, J. M., Klucas & R. V. Sarath. G. (1997) *J. Mol. Biol.* **266**, 1032–1042.
15. Chang, C. & Meyerowitz, E. M. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 1408–1412.
16. Bashford, D., Chothia, C. & Lesk A. M. (1987) *J. Mol. Biol.* **196**, 199–216.
17. Millar, A. H., Bergerson, F. J. & Day, D. A. (1994) *Plant Physiol. Biochem.* **32**, 847–852.
18. Wittenberg, J. B. & Wittenberg, B. A (1990) *Annu. Rev. Biophys. Biophys. Chem.* **19**, 217–241.