The characterisation of hemoglobins from *Arabidopsis thaliana*.

Ben Trevaskis.

Submitted for the degree of Doctor of Philosophy at the Australian National University.
1997

This thesis does not contain any material that has been accepted for the award of any other degree or diploma at this or any other University, nor any material that has been published previously, except where due reference is made in the text. The research described is my own original work with the following exceptions. Janice Norman and Danny Llewellyn screened for, isolated and purified the λ-genomic clone from which *AHB1* was subloned (Chapter 3). A publication resulting from this work is currently in press:

Ben Trevaskis, Richard. A. Watts, Carol R. Andersson, Danny J. Llewellyn, Mark S. Hargrove, John S. Olson, Elizabeth S. Dennis, W. James Peacock. "Two hemoglobin genes in *Arabidopsis thaliana*: the evolutionary origins of leghemoglobins." *PNAS* , in press October 1997.

Ben Trevaskis

Acknowledgments.

A number of people have encouraged, supported or inspired me during the gestation of this thesis. First and foremost are my supervisors. I would like to thank Liz Dennis for her excellent supervision, and also Jim Peacock, Michael Udvardi and Cyril Appleby who have had significant inputs to this work.

Thanks to the Division of Plant Industry in general and to the following people in particular: Janice Norman, Richard "Refugee" Watts, Mark “Chocolate” Olive, Rudy Dooferus, Frank "Doc Holiday" Hoeren, Francoise "Fuzzwuzz" Jaudinaud, and Marc "Climber" Ellis. Also, labwork would have been a lot harder without the excellent work of Janice, and Alison Norman.

Thanks also to the following people for miscellaneous reasons. To Marc and Ian Watson for the useful discussions, to Guy de Bruxelles for the apologies. To Leanne Murdoch, Irene Lambiris and Diana Lane for all the parties. To the ANU Aikido club for all the bruises and all the fun. To the Simpsons, the SONY corporation, and the choughs who have been constant sources of inspiration. I would also like to aknowledge the Cooperative Research Centre for Plant Science for financial support.

And last, but not least, thanks to Gill for the support and for all the coffees in the morning.

Abstract.

Two families of hemoglobin genes have been identified in plants, the symbiotic hemoglobins and the non-symbiotic hemoglobins. These families are divided on the basis of gene sequence and expression pattern. Symbiotic hemoglobins are expressed exclusively in the nodules of nitrogen-fixing plants, at high levels, where they are thought to transport oxygen to nitrogen-fixing symbiotic microbes. Symbiotic hemoglobins genes have been found only in plants that fix nitrogen. Non-symbiotic hemoglobins are found in a wide range of plants, including both plants that do and do not fix nitrogen. These genes are expressed in normal plant tissues but the requirement for non-symbiotic hemoglobins is unclear. It has been suggested that non-symbiotic hemoglobins may act as either oxygen transporters or as oxygen sensors.

In order to further examine the function of non-symbiotic hemoglobins, hemoglobin genes were isolated from a model plant, *Arabidopsis thaliana*. This thesis describes the cloning and characterisation of two hemoglobin genes from *Arabidopsis thaliana*. One, *AHB1*, is related in sequence to the non-symbiotic hemoglobins previously identified in other plants. The other, *AHB2*, is novel in that it more similar in sequence to the symbiotic hemoglobins than it is to the non-symbiotic hemoglobins. The occurrence of these two different hemoglobin genes in *Arabidopsis* (a plant that does not fix nitrogen and is not closely related to nitrogen fixing plants) suggests that there are two distinct classes of non-symbiotic hemoglobin in plants. One class corresponds in sequence to the "non-symbiotic" hemoglobin genes, while the other class is related to the "symbiotic" hemoglobin genes that have previously been found only in nitrogen fixing plants. The symbiotic

hemoglobins may have evolved from an *AHB2* -like non-symbiotic hemoglobin gene family.

The expression patterns of the two hemoglobin genes cloned from *Arabidopsis* have been examined. *AHB1* is expressed at a low level in root tissue and is induced, in both roots and rosette leaves, by low oxygen levels. Expression of *AHB1* is also induced by sucrose. This pattern of expression is similar to the non-symbiotic hemoglobins of other plants. The promoter of *AHB1* shows sequence homology to the promoters of other non-symbiotic plant hemoglobins, and directs a pattern of GUS (β-glucuronidase) reporter gene expression in transgenic plants that is typical of non-symbiotic hemoglobin gene promoters. The promoter of *AHB1* is also able to confer hypoxic induction to the GUS reporter gene, and there is evidence that the expression of *AHB1* is subject to post-transcriptional regulation during hypoxia.

AHB2 is expressed at a low level in above ground tissue and transcript levels increase after chilling. Unlike *AHB1, AHB2* gene expression is not affected by exogenous sucrose levels or hypoxia. The expression pattern of *AHB2* is unique amongst plant hemoglobins, and the promoter of *AHB2* does not appear to contain any of the conserved sequence motifs found in the promoters of other plant hemoglobin genes.

The possible functions of the two hemoglobins of *Arabidopsis* are discussed. A number of lines of evidence suggest that these hemoglobins are unlikely to operate as oxygen sensors or as oxygen transporters. Instead, non-symbiotic hemoglobins may act as oxygen scavenging proteins that bind oxygen for delivery to oxygen consuming reactions, possibly through specific interactions with other proteins.

TABLE OF CONTENTS.

Contents. .... **Page.**

Title Page. .... **i**

Candidates statement. .... **ii**

Acknowledgments. .... **iii**

Abstract. .... **iv**

Contents. .... **vi**

List of figures. .... **x**

CHAPTER 1. .... **1**

Introduction.

1.1 Hemoglobins: 1
- Myoglobin: The model hemoglobin. 1
- Vertebrate hemoglobins. 3
- Variation in oxygen affinity. 5
- Invertebrate hemoglobins. 6
- Hemoglobin protein sequences. 8

1.2 Other types of hemoglobin. 8
- Proto-globins. 8
- Flavo-hemoglobins. 9
- *Vitreoscilla* hemoglobin. 10
- True hemoglobins? 10

1.3 Plant hemoglobins. 11
- Hemoglobins in nitrogen fixing symbioses. 11
- The leghemoglobins. 14
- *Casuarina* symbiotic hemoglobin. 16
- The hemoglobin of *Parasponia*. 17
- Non-symbiotic hemoglobins. 18
- *Trema* non-symbiotic hemoglobin. 18
- *Casuarina* non-symbiotic hemoglobin. 19
- Soybean non-symbiotic hemoglobin. 20
- Non-symbiotic hemoglobins in cereals. 20
- The promoters of plant hemoglobin genes. 24

1.4 The function of non-symbiotic hemoglobins? 27
Oxygen transporters? 27
Oxygen sensors? 28
Other functions? 29

1.5 Determining non-symbiotic hemoglobin function. 30

Arabidopsis hemoglobins. 30

CHAPTER 2. 31

Methods.

2.1 Plant growth. 31

2.2 Library screening and DNA manipulation. 31

2.3 RNA Gel Blot Analysis. 32

2.4 RT-PCR. 32

2.5 Database Search. 32

2.6 Expression of recombinant AHB1 and AHB2 in *E. coli.* 33

2.7 Western analysis of bacterial strains. 34

2.8 Western analysis of AHB1 levels in *Arabidopsis* tissues. 34

2.9 *AHB1* promoter-GUS fusion. 35

2.10 GUS staining. 35

CHAPTER 3. 36

The cloning and characterisation of *AHB1*.

3.1 Introduction. 36

3.2 Results. 37
Cloning of *AHB1*. 37
AHB1 is a single copy gene. 39
Analysis of *AHB1* sequence. 40
Analysis of *AHB1* gene expression. 46
Isolation of a *AHB1* cDNA clone. 49

3.3 Discussion. 51

CHAPTER 4. 55

Cloning and characterisation of *AHB2*.

4.1 Introduction. 55

4.2 Results. 58
Cloning of *AHB2*. 58
Sequence analysis of *AHB2*. 58
Southern analysis of *AHB2*. 67
Analysis of *AHB2* expression. 67
Cloning of a *AHB2* cDNA. 70

4.3 Discussion. 71

CHAPTER 5. 74

Control of *AHB1* gene expression.

5.1 Introduction. 74

5.2 Results. 77

5'*AHB1*-Gus Fusion. 77
GUS activity is sucrose inducible in GH1-4 plants. 79
Hypoxia does not induce GUS activity in GH1 transformed plants. 80
GUS transcription is induced by hypoxia in GH1-4 plants. 81
Expression of recombinant *AHB1* and *AHB2* in E. coli. 83
Cross reactivity of anti-Phb and anti-Cassym antisera with recombinant AHB1 and AHB2. 86
AHB1 protein levels increase during hypoxia treatment. 88

5.3 Discussion. 90

Chapter 6. 95

The evolution and function of plant hemoglobins.

6.1 Two classes of non-symbiotic hemoglobin in plants. 95
When did the two classes of plant hemoglobin diverge? 97
The evolutionary origins of the *Parasponia* hemoglobin. 102
The evolutionary origins of the Class 2 symbiotic hemoglobins. 104
Plant hemoglobin promoters. 107

Hemoglobins and the origins of nitrogen-fixing symbioses. 108

6.2 The oxygen binding kinetics of *Arabidopsis* hemoglobins. 111

6.3 The function of non-symbiotic plant hemoglobins? 114

AHB1 and AHB2 are unlikely to be oxygen sensors. 115

AHB1 may deliver oxygen to mitochondrial terminal oxidases. 115

AHB1 may scavenge oxygen in hypoxic tissues. 116

AHB1 and AHB2 may channel bound oxygen into biosynthetic reactions. 117

6.4 Determining the functions of plant hemoglobins. 117

Cloning of non-symbiotic Class 2 hemoglobins from other plants. 117

Localisation of *AHB1* and *AHB2* gene expression. 118

Localisation of AHB1 and AHB2 proteins. 118

Heterologous expression. 119

Reverse genetics. 120

6.5 Conclusion. 121

Bibliography. 122

Appendix. 134

LIST OF FIGURES

Figure 1. Heme molecule. 2

Figure 2. The globin fold. 4

Figure 3. Molecular phylogeny of seed plants. 13

Figure 4. Conserved intron positions 15

Figure 5. Similarity tree of plant hemoglobin protein sequences. 22

Figure 6. Conserved sequences in leghemoglobin promoters 24

Figure 7. Partial restriction map of Ara111 l clone 38

Figure 8. Genomic DNA blot probed with *AHB1*. 39

Figure 9. *AHB1* Genomic sequence. 41

Figure 10. Alignment of non-symbiotic hemoglobin protein sequences. 43

Figure 11. Plant hemoglobin protein similarity tree. 44

Figure 12a. *AHB1* Promoter sequence. 45

Figure 12b. Promoter motif sequence comparisons 46

Figure 13. *AHB1* expression in above and below ground tissues 48

Figure 14. *AHB1* expression during environmental stress. 48

Figure 15. Time course of the *AHB1* hypoxia response 49

Figure 16. Evolution of the symbiotic hemoglobins. 57

Figure 17. Partial restriction map of *AHB2* genomic sequence. 60

Figure 18. Sequence of *AHB2*. 61

Figure 19. Alignment of plant hemoglobins. 64

Figure 20. Protein similarity tree. 65

Figure 21. Sequence from the 5' of *AHB2*. 66

Figure 22. Genomic DNA blot probed with *AHB2*. 68

Figure 23. Induction of *AHB2* by low temperature. 69

Figure 24. *AHB2* expression during environmental stress. 69

Figure 25. GUS staining in the roots of GH1-4 plants. 78

Figure 26. GUS staining in the above ground tissues of GH1-4 plants. 78

Figure 27. GUS staining in roots of GH1-4 plants treated with sucrose. 80
Figure 28. GUS staining in above ground tissue of GH1-4 plants treated with sucrose.......... 80
Figure 29. Hypoxic induction of GUS transcript in GH1-4 plants. 82
Figure 30. Bacterial expression construct. 84
Figure 31. Recombinant hemoglobin expression. 85
Figure 32. Cross reactivity of anti-hemoglobin antisera with recombinant *Arabidopsis* hemoglobins. 87
Figure 33. Expression of AHB1 protein in hypoxia treated plants......... 88
Figure 34. The two classes of plant hemoglobin. 99
Figure 35. Phylogram of plant hemoglobin proteins.......... 100
Figure 36. Partial phylogeny of dicotyledonous plants. 104
Figure 37. The evolution of plant hemoglobin genes.......... 106

CHAPTER 1.

Introduction.

1.1 Hemoglobins:

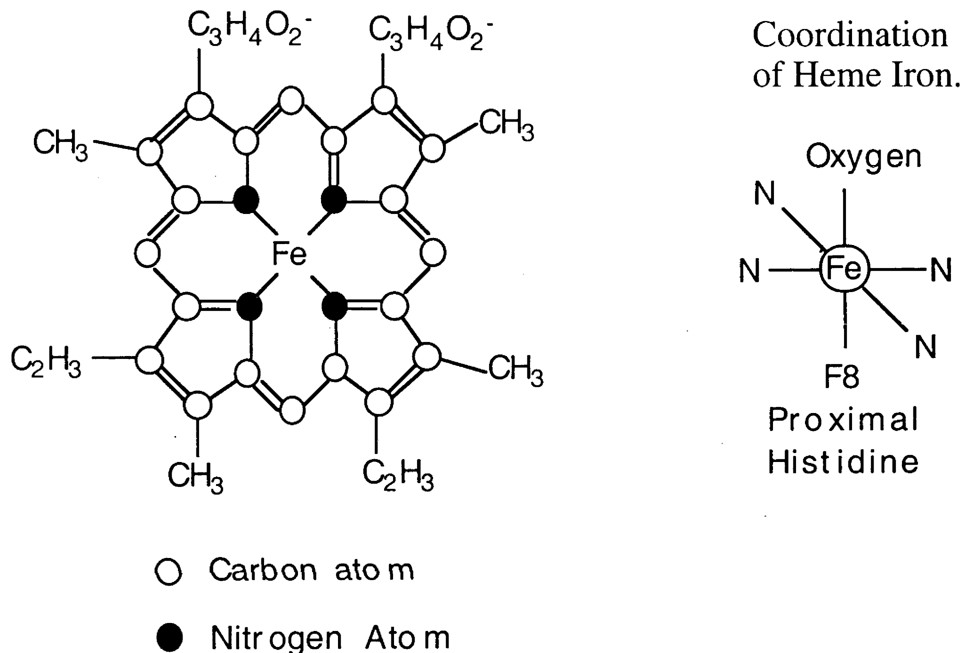
A hemoglobin consists of a globin apo-protein and a heme prosthetic group that together are able to bind oxygen reversibly. Heme is a planar protoporphyrin IX molecule with an iron atom (Fe^{2+}) coordinated in its center (Figure 1). The iron atom of the heme group is able to bind a range of ligands and is found in many proteins. Reversible binding of oxygen and a characteristic tertiary structure, the globin fold, are the defining features of hemoglobins that set them apart from other heme-proteins (Bashford et al. 1987, Zhu and Riggs 1992). Typically red in solution, hemoglobins also display distinctive absorption spectra that vary with the binding/release of oxygen and with the redox state of the iron atom of the heme group. Hemoglobin genes have been identified in a wide range of organisms and may share a common ancestor (Riggs 1991, Zhu and Riggs 1992, Moens et al. 1996, Hardison 1996).

Myoglobin: The model hemoglobin.

Myoglobin is a hemoglobin that is found in high concentrations in the muscle tissue of vertebrates. It was originally named myoglobin (*myo* being Latin for muscle) to distinguish it from the hemoglobins found in red blood cells. Myoglobin was the first protein structure to be solved and the relationships between the structure and function of myoglobin have been extensively examined (Dickerson and Geis 1983, Springer et al. 1994). For these reasons myoglobin is often viewed as the model hemoglobin.

Figure 1. Heme molecule.

Iron-Protoporphyrin IX. Of the six coordination sites available to the iron atom, four form bonds to the nitrogen atoms of the porphyrin group, one coordinates to the proximal histidine (F8) and the final site binds oxygen. (Stryer 1988)



Myoglobin is a monomeric hemoglobin protein. It consists of a globin apo-protein of around 153 amino acids (depending on the organism examined) and a heme prosthetic group. To form myoglobin, the globin apo-protein folds around the heme group forming a distinctive arrangement of eight α-helices (A-H). This structure has been called the "globin fold" (Figure 2). The heme group is attached to the globin apo-protein through the co-ordination of the iron atom to a critical amino acid residue, the proximal histidine. The interaction of the protein chain and the heme group modifies the properties of the heme iron favouring reversible binding of oxygen. Furthermore, the protein forms a "binding pocket" near the heme iron restricting the access of other ligands to the heme iron, increasing

specificity for oxygen binding (see Dickerson and Geis 1983, Stryer 1988, Springer et al. 1994).

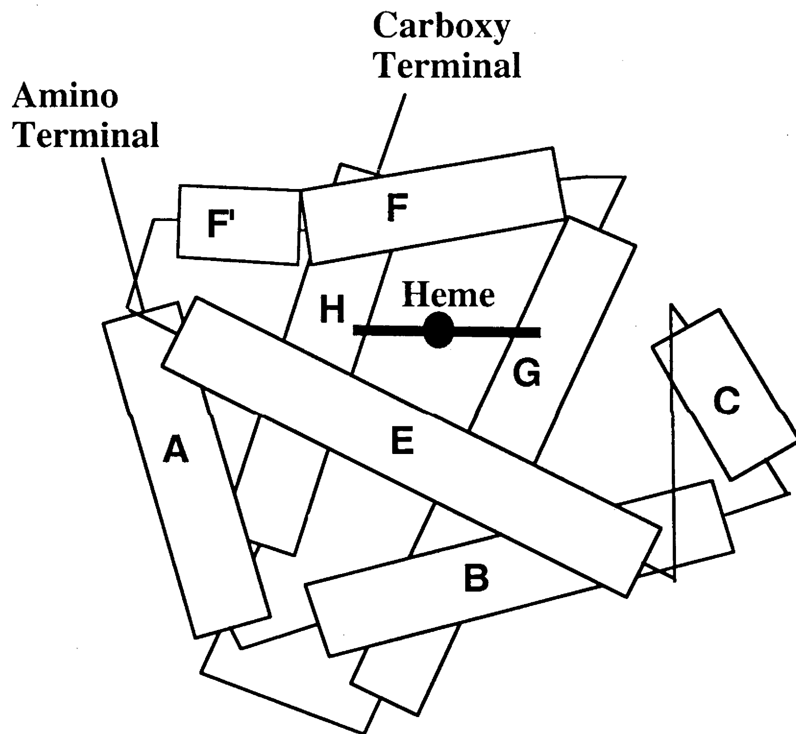
Myoglobin functions as an oxygen transporter. The high concentration of oxygen-binding myoglobin lowers the concentration of free oxygen in muscle tissue. This increases the rate at which oxygen can diffuse in from capillaries. Oxygen is then exchanged between myoglobin molecules and diffuses down an oxygen gradient to respiring mitochondria (Wittenberg and Wittenberg 1989). Myoglobin also functions as an oxygen storage molecule, providing an internal source of oxygen in the muscle tissue. The oxygen storage function of myoglobin is important during times of intense muscle activity when oxygen availability can become limited (see Dickerson and Geis 1983).

Vertebrate hemoglobins.

The blood of vertebrate animals contains erythrocytes with high concentrations of hemoglobin. These blood cells move through the blood stream carrying hemoglobin bound oxygen from areas of high oxygen concentration to areas of low concentration. Vertebrate hemoglobins are myoglobin-like proteins, but unlike myoglobin which functions as a monomer, vertebrate hemoglobins generally consist of tetramers formed from two different subunits. In adults, these subunits are called alpha and beta globin, and two of each subunit are found in a hemoglobin tetramer ($\alpha 2/\beta 2$). There is another form of the hemoglobin tetramer found in the blood stream of the vertebrate fetus. In fetal hemoglobin, gamma globin subunits replace the beta globin subunits (α_2/γ_2) (reviewed in Dickerson and Geis 1983).

Figure 2. The globin fold.

Representation of the three dimensional protein structure that is typical of hemoglobins, the "Globin fold". Helices are represented by cylinders and are labelled A-H (Bashford et al. 1987). The short D helix (absent from this diagram) is not found in all hemoglobins.



In vertebrate hemoglobin, complex interactions between hemoglobin subunits also allow alterations in oxygen affinity depending on the amount of oxygen available (see Perutz 1978, Dickerson and Geis 1983). Binding of oxygen to one subunit promotes the binding of oxygen to the other subunits of the hemoglobin tetramer. Thus, the binding of oxygen to hemoglobin tetramers is cooperative. The result of cooperative binding of oxygen is that when oxygen is available it tends to bind to all four subunits. Conversely, when oxygen is released from one subunit, it promotes the release of oxygen by other subunits. This effect is important for systemic oxygen transport, as it promotes binding of oxygen to hemoglobin tetramers in the oxygen rich environment of the lungs and release of oxygen in oxygen depleted active tissues (reviewed in Dickerson and Geis 1983).

Other molecules can affect hemoglobin oxygen affinity. Protons bind to hemoglobin and promote oxygen dissociation. Carbon dioxide has a similar effect. As active tissues tend to have a lower pH and higher carbon dioxide concentrations this promotes the release of oxygen in active tissues. Like the cooperative binding of oxygen, the effects of carbon dioxide and protons on hemoglobin oxygen affinity are mediated through complex subunit interactions (see Dickerson and Geis 1983). Thus, the tetrameric arrangement of myoglobin-like subunits allows more complex and specialised regulation of oxygen binding behaviour.

Variation in oxygen affinity.

Different hemoglobin molecules show different oxygen binding kinetics and have evolved oxygen affinities to suit different oxygen

environments. For example, the affinity of myoglobin for oxygen is higher than that of hemoglobin tetramers, allowing oxygen to diffuse from the hemoglobin in the bloodstream into muscle tissue. Likewise, fetal hemoglobin tetramers have a higher affinity for oxygen than adult hemoglobin, due to the replacement of the beta globin subunit with gamma globin. This maximises the transfer of oxygen from the maternal bloodstream into the fetus.

The overall oxygen affinity of a hemoglobin molecule will depend on how rapidly it binds and releases oxygen. The rates at which a hemoglobin binds or releases oxygen can also affect hemoglobin function. Vertebrate hemoglobins and myoglobins must bind and release oxygen rapidly, in order to function as efficient oxygen transporters, and therefore have large oxygen association and dissociation rate constants (see Dickerson and Geis 1983).

The differences in oxygen affinities shown by hemoglobins is generated in part by amino acid sequence variation. Further variation is generated through interactions between multiple hemoglobin subunits. Such interactions can also generate other complex behaviour, such as the cooperative binding of oxygen by vertebrate hemoglobins. The vertebrate hemoglobins and myoglobins provide a useful demonstration of how the oxygen affinity of a hemoglobin can evolve to suit the environment in which it functions (see Dickerson and Geis 1983).

Invertebrate hemoglobins.

Myoglobin-like hemoglobins are also found in invertebrate animals. The larvae of the midge *Chironomus* for example, contain large amounts of a hemoglobin. This hemoglobin is thought to act as an oxygen

transporter and storage molecule, allowing the larvae of this insect to survive in warm stagnant pools of water that often have low concentrations of dissolved oxygen (Osmulski and Leyko 1986). The amino acid sequence of the *Chironomus* hemoglobin protein contains all the residues that are known to be highly conserved in hemoglobins (Antoine and Niessing 1984) and the tertiary structure of the *Chironomus* hemoglobin is similar to that of myoglobin (see Bashford et al. 1987).

One interesting aspect of invertebrate hemoglobins in general is that they exhibit great diversity in quaternary structure. Some are found as single domain proteins, while others consist of two or more connected globin domains (Riggs 1991, Terwilleger and Terwilleger 1985). Further diversity is found when oligomerisation is examined. Examples of monomeric, dimeric or larger hemoglobin complexes exist within the invertebrate phyla (Riggs 1991). In some cases oligomerisation has been associated with cooperative binding of oxygen and other more complex regulation of oxygen binding behaviour (Riggs 1991).

The functions of most invertebrate hemoglobins are analogous to vertebrate hemoglobins or myoglobins. Invertebrate hemoglobins generally function as oxygen transporters, systemic or intracellular, or oxygen storage proteins* (Riggs 1991). Thus, the basic property of the hemoglobin monomer, which is reversible binding of oxygen, has been utilised for oxygen transport and storage functions across the animal kingdom.

---

* some molluscan hemoglobins transport H_2S from symbiotic bacteria (Terwilleger and Terwilleger 1985)

Hemoglobin protein sequences.

Despite the conserved tertiary structure of hemoglobin proteins there is little absolute amino acid sequence conservation between hemoglobin proteins, with identities as low as 16% for globins from distantly related organisms (Bashford et al. 1987). Indeed, there are only two absolutely conserved amino acids, and only a handful of strongly conserved residues. It seems that the globin protein can tolerate a range of amino acid sequences while still maintaining the globin fold structure. The key to maintaining this structure appears to be in the hydrophobic nature of internal residues that stabilise the overall globin fold structure (Bashford et al. 1987).

Structural and kinetic analysis of mutant globin proteins, either naturally occurring or artificially generated, has provided a great deal of information about the roles of the few highly conserved amino acid residues that do occur in hemoglobins. The absolutely conserved proximal histidine (F8) co-ordinates with the heme iron binding it to the globin apo-protein, while the phenylalanine (CD1) provides an essential site of protein-heme interaction. Other conserved residues are the distal histidine (E7) which stabilises the binding of oxygen to the heme iron and a proline (C2) that terminates an alpha helix and helps maintain the globin fold structure (Springer et al. 1994).

1.2 Other types of hemoglobin.

Proto-globins.

In the past decade a group of related, globin-like proteins have been identified in protozoans, a cyanobacterium and a green alga (Iwaasa et al.

1989, Potts et al. 1992, Couture et al. 1994). These putative globins are red in solution and some have been shown to bind oxygen reversibly (Potts et al. 1992, Couture et al. 1994). The amino acid sequences of these proteins are very similar to each other, but differ in many respects to those of animal hemoglobins and myoglobins. The distal histidine is replaced by a glutamine residue, and they are smaller than myoglobin averaging around 120 amino acids, compared to 150-160 amino acids for myoglobins. However, these globins still possess the conserved proximal histidine (F8) and phenylalanine (CD1) residues. This family of hemoglobins will be referred to as the "Proto-globins", as they were first identified in protozoans.

Flavo-hemoglobins.

Another type of protein has been found to consist of a globin domain linked to a second domain which shows homology to flavin binding oxido-reductase domains from other proteins. This type of chimeric protein has been termed flavohemoglobin (Zhu and Riggs 1992). Flavohemoglobins have been identified in yeast and bacteria (Cramm et al. 1994, Zhu and Riggs 1992, Vasudevan et al. 1991). The globin domains of these proteins are more similar to the animal hemoglobins and myoglobins than the proto-globins, and are approximately 145 amino acids in size. Despite stronger homology with classical globins, the distal E7 site is occupied by glutamine as in the proto-globins. Recently the crystal structure of the flavohemoglobin from the bacterium *Alcaligenes eutrophus* has been solved using X-ray crystallography and shown to conform to the globin fold structure (Ermler et al. 1995). Flavohemoglobins may function as oxygen regulated catalytic proteins, where the globin group regulates the activity of the catalytic domain (Zhu and Riggs 1992).

Vitreoscilla hemoglobin.

The hemoglobin of the obligate aerobe, gram positive bacterium, *Vitreoscilla* falls into a class of its own. The *Vitreoscilla* hemoglobin protein sequence shows stronger homology to globin domains of the flavohemoglobins than it does to the proto-hemoglobins, but consists of a single globin domain (Wakabayashi et al. 1986, Cramm et al. 1994). The *Vitreoscilla* hemoglobin may have evolved from a flavohemoglobin that was truncated to a single globin domain (Cramm et al. 1994). In *Vitreoscilla,* this hemoglobin is expressed at high levels under hypoxia and is thought to trap or scavenge oxygen, allowing survival and growth under these conditions (Wakabayashi et al. 1986). Recently the *Vitreoscilla* hemoglobin has been expressed in transgenic tobacco plants and found to increase plant growth rates (Holmberg et al. 1997). Interestingly, the content of some metabolites that require oxygen for biosynthesis, nicotine for instance, were found to increase. It has been suggested that these effects may be due to an general increase in oxygen availability (Holmberg et al. 1997). Alternatively, *Vitreoscilla* hemoglobin may interact with terminal oxidases, or other parts of respiratory electron transport chains, increasing metabolic rates (Chen et al. 1994).

True hemoglobins?

Comparison between the tertiary structures of the classical myoglobin-like hemoglobins and the globin domains of flavohemoglobins or the *Vitreoscilla* hemoglobin support the theory that these proteins are related. Proto-globins show less amino acid sequence homology to the classical hemoglobins, but do have sufficient homology to satisfy sequence alignment algorithms that take into account structural constraints associated with the globin fold structure (Moens et al. 1996). However, proto-globins

cannot be considered true hemoglobins while the tertiary structure of these proteins remain unknown.

1.3 Plant hemoglobins.

Hemoglobins have been identified in plants. The amino acid residues that are conserved in animal hemoglobins and myoglobins are also found in plant hemoglobins and some plant hemoglobins have been shown to have the globin fold structure (see Bashford et al. 1987). Thus plant hemoglobins are myoglobin-like proteins. Plant hemoglobins also have a gene structure similar to myoglobin (see below) suggesting that plant and animal hemoglobins evolved from a common ancestral hemoglobin gene.

Hemoglobins in nitrogen fixing symbioses.

Plants are unable to fix atmospheric nitrogen directly. In some plants fixation of atmospheric dinitrogen is achieved indirectly, through symbiotic interactions with certain nitrogen fixing micro-organisms. These symbioses arise from a series of complex molecular communication events between plant and micro-organism resulting in the formation of specialised plant tissues. Such events have been characterised in detail for the legume/*Rhizobium* symbiosis where the resulting nitrogen fixing structures are called nodules (reviewed in Long 1996). In legumes, nodules generally form from modified root tissues* following infection by *Rhizobium,* a gram negative bacterium. Within these structures the *Rhizobia* bacterial endo-symbionts, now termed bacteroids, are able to fix nitrogen.

In addition to legume/*Rhizobium* symbioses, nitrogen-fixing symbioses have also been found to occur between a tropical tree

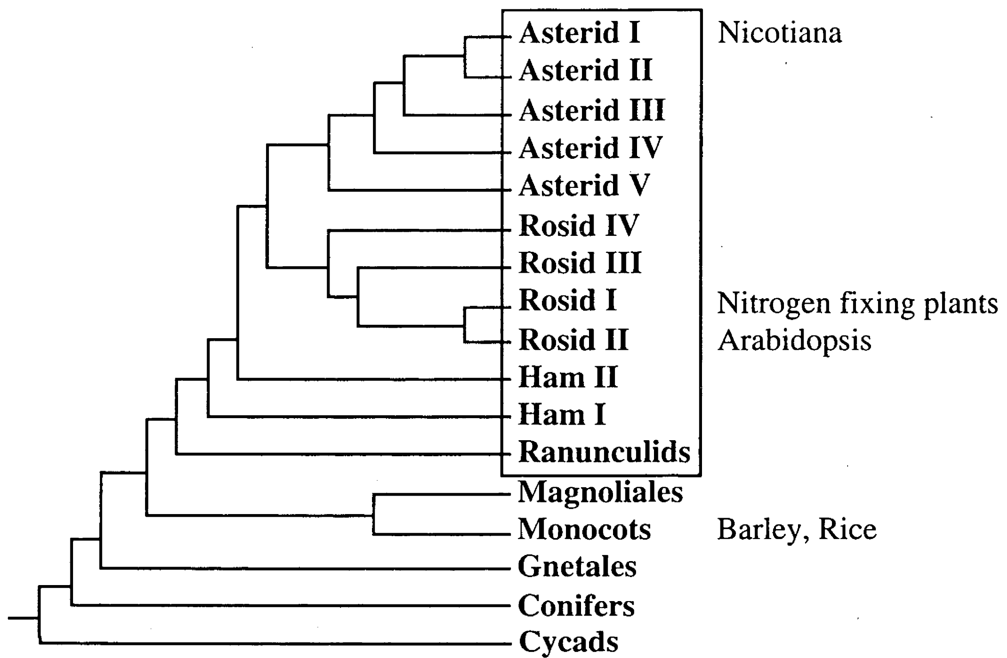
* with the exception of one line of legumes that forms stem nodules, (Drefus and Dommogue 1981)

Parasponia andersonii (Ulmaceae) and *Rhizobia* (Trinick 1973). Also, actinorhizal symbioses occur between a range of trees (*Casuarina* for example) and an actinomycete gram positive bacterium *Frankia* (Newcomb and Wood 1987). These systems have not been characterised to the extent that the legume/*Rhizobium* relationships have been. All plants* that are known to form nitrogen-fixing symbioses can be grouped into a single clade of dicots, Rosids I (Chase et al. 1993, Soltis et al. 1995) and it seems that some aspects of this lineage may have allowed multiple nitrogen-fixing symbioses to evolve (Figure 3).

---

* with the exception of Gunneraceae which forms a nitrogen-fixing cymbiosis with a cyanobacteria (see Soltis et al. 1995).

Figure 3. Molecular phylogeny of seed plants.



The major evolutionary lineages of seed plants identified by phylogenetic analysis of Ribulose-1,5-biphosphatase carboxylase (Rubisco) large subunit sequences (from Chase et al. 1993). Dicotyledonous (Eudicots) plants are boxed. Some plants discussed in this thesis are noted next to the clade in which they have been grouped.

One problem associated with biological nitrogen fixation is that nitrogenase, the enzyme that fixes nitrogen, is inactivated by oxygen (see Robson and Postgate 1980). For nitrogen fixation to occur nitrogenase must be kept in a low oxygen environment. Nitrogen fixation is also an extremely energy demanding process and high levels of respiration are required in nitrogen-fixing nodules. This creates the "nitrogen fixing paradox", with the need to keep oxygen levels low to maintain nitrogenase activity, but to also provide enough oxygen to allow the bacteroids to obtain the chemical energy required to power nitrogen fixation. In many nitrogen fixing plants the nitrogen fixing paradox is resolved (to a large degree) by "symbiotic" plant hemoglobins. These hemoglobins allow rapid transfer of oxygen to the bacteroids, but deliver oxygen in a bound form, thus preserving nitrogenase activity (see Appleby 1984).

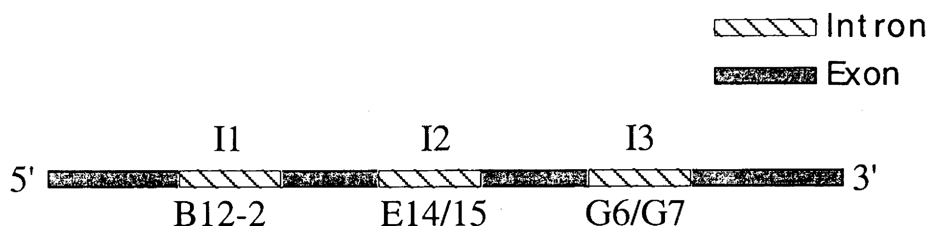
The leghemoglobins.

The symbiotic hemoglobins from the nitrogen fixing nodules of legumes, leghemoglobins, have been studied in great detail. Leghemoglobin is a myoglobin-like protein. The tertiary structure of leghemoglobin is remarkably similar to that of myoglobin (Ollis et al. 1983), and leghemoglobins retain the conserved residues found in animal hemoglobins (the proximal F8 and distal E7 histidine residues, proline C2 and phenylalanine CD1). Leghemoglobins show some overall amino acid sequence homology to vertebrate globins, and leghemoglobin genes have two introns in identical positions to the two introns found in all animal globin genes*. Leghemoglobin genes also have a third central intron (Jensen et al. 1981) that is absent in animal hemoglobins but found in all plant hemoglobin genes (Figure 4). The similarity in protein structure and

* With the exception of the Chironomid hemoglobin (Antoine et al. 1987)

in gene structure suggest that plant and animal hemoglobin genes are the descendants of a common ancestral hemoglobin gene.

Figure 4. Conserved intron positions.



Representation of plant hemoglobin gene structure. I1 and I3 occur in identical positions in both animal and plant genes. The central intron I2 is found in all plant haemoglobin genes (in the codon for amino acid E14/15). Central introns also occur, in different positions in some nematode hemoglobin genes, (at E8-1 in *Ascaris*, at E3-2 in *C elegans*). Central introns are thought to have appeared after the divergence of the plants and animal kingdoms (Stoltzfus and Doolittle 1993). Amino acid residue definitions refer to globin protein sequence alignment template of Bashford et al. 1987.

Leghemoglobins are typically encoded by a small family of genes and are expressed exclusively in nodules, at high levels (millimolar concentrations, see Appleby 1984). The oxygen affinities of many leghemoglobins are known and are consistent with an ability to deliver oxygen to respiring bacteroids. They have a high affinity for oxygen, and are thus able to maintain low free oxygen concentrations, but also have fast association and dissociation rates allowing rapid exchange of oxygen (Wittenberg et al. 1974, Gibson et al. 1989). Thus, the function of leghemoglobins is also myoglobin-like; intracellular transport of oxygen.

The biochemistry of leghemoglobins, and the role of these proteins in nitrogen fixation has been characterised in great detail (reviewed in Appleby 1984).

Leghemoglobins were the first hemoglobins to be identified in plants, a discovery that was facilitated by the high concentration of these proteins in the nitrogen-fixing nodules of legumes (Kubo 1939, cited in Appleby 1992). The observation that only one line of plants, the legumes, contained hemoglobins related to those in animals was difficult to explain. Leghemoglobins appeared to be related to animal hemoglobins, but the limited distribution of hemoglobins in plants suggested that hemoglobins had not been present in the ancestor of plants. Instead, horizontal gene transfer, from an animal or bacterial genome to the ancestor of legumes, was suggested to account for the limited occurrence of hemoglobins in plants (Hyldig-Nielsen et al. 1982).

***Casuarina* symbiotic hemoglobin.**

More recently, hemoglobins have been identified in non-leguminous plants. The actinorhizal nodules of *Casuarina* species have been shown to contain significant amounts of a symbiotic hemoglobin (Fleming et al. 1987). The gene encoding this hemoglobin, *Cas-sym*, has been cloned (Jacobsen-Lyon et al. 1995). The *Cas-sym* gene is interrupted by three introns found in positions identical to the introns of leghemoglobin genes (Jacobsen-Lyon et al. 1995). The *Casuarina* symbiotic hemoglobin amino acid sequence is highly similar to the sequences of leghemoglobins. The expression of *Cas-sym* is limited to nodules, where a high level of expression is found, and like leghemoglobins this gene appears to be a member of a small gene family with at least three different genes present (Jacobsen-Lyon et al. 1995). The oxygen affinity of the *Casuarina*

symbiotic hemoglobin is known and is also similar to that of leghemoglobins (Fleming et al. 1987).

Thus, the *Casuarina* symbiotic hemoglobin is related to and presumably fulfils a similar function to the leghemoglobins. The hemoglobins of *Casuarina* and legume nitrogen-fixing symbioses can be considered to belong to the same family on the basis of amino acid sequence similarity and shared function, and have been designated the symbiotic hemoglobin gene family (Jacobsen-Lyon et al. 1995).

The hemoglobin of *Parasponia*.

The tropical tree *Parasponia andersonii* forms a nitrogen-fixing symbiosis with *Rhizobium* (Trinick 1973). Nitrogen fixation occurs in modified lateral roots that are analogous to the nodules of legumes. These nodule-like structures contain a hemoglobin that is thought to fulfil a function similar to leghemoglobin (Appleby et al. 1983). Hemoglobin was purified from *Parasponia* nodules and amino acid sequence obtained (Appleby et al. 1983), allowing the corresponding gene to be cloned (Landsmann et al. 1986).

The *Parasponia* hemoglobin retains the conserved amino acid residues characteristic of hemoglobins, and the *Parasponia* hemoglobin gene is interrupted by three introns located in exactly the same positions as leghemoglobin introns (Landsmann et al. 1986). However, the *Parasponia* hemoglobin differs from both leghemoglobins and *Casuarina* symbiotic hemoglobin in a number of respects. Firstly, considerable amino acid sequence variation occurs in regions that are conserved in the symbiotic hemoglobins (Landsmann et al. 1986). The *Parasponia* hemoglobin protein has a N-terminal region that is 5 amino acids longer than symbiotic

hemoglobins. Furthermore, the *Parasponia* hemoglobin gene is expressed at a low level in root tissue, as well as being expressed at a high level in nodules (Landsmann et al. 1986), whereas other symbiotic hemoglobins are expressed exclusively in nodules.

Non-symbiotic hemoglobins.

The observation that the hemoglobin of *Parasponia* is expressed in root tissue, as well as nodules, lead to the suggestion that hemoglobins may be required in the roots of all plants (Landsmann et al. 1986). If hemoglobins, related in sequence to the *Parasponia* hemoglobin, are present in all plants then the evolution of the *Parasponia* hemoglobin would be easy to explain. A hemoglobin that functions in normal plant tissues, a "non-symbiotic hemoglobin", may have evolved high level nodule expression to fulfil a symbiotic function, while also retaining a function in root tissue. If this was the case, the occurrence of hemoglobins in legumes would also be explicable. Leghemoglobins may have diverged from the same ancestral plant hemoglobin that functions in normal plant tissues. If leghemoglobins diverged from the non-symbiotic hemoglobins earlier than the *Parasponia* hemoglobin then the degree of sequence variation between these hemoglobins would also be explained. As the leghemoglobins and *Casuarina* symbiotic hemoglobin do not appear to function outside nodules, it is also possible that the sequences of these proteins are not as constrained by selection as that of the *Parasponia* hemoglobin allowing more rapid sequence divergence.

***Trema* non-symbiotic hemoglobin.**

To test the hypothesis that plant genomes may contain a non-symbiotic hemoglobin gene, a hemoglobin gene showing homology to the

Parasponia hemoglobin gene was isolated from a closely related plant that does not fix nitrogen, *Trema tomentosa* (Bogusz et al. 1988). The predicted protein sequence of this hemoglobin has a N-terminal extension of 5 amino acids and, like the *Parasponia* hemoglobin, low levels of transcript were detected in root tissue RNA extracts (Bogusz et al. 1988).

Trema non-symbiotic hemoglobin gene was the first hemoglobin gene to be isolated from a plant that does not fix nitrogen. Following the cloning of the *Trema* non-symbiotic hemoglobin gene, similar genes have been isolated from a wide range of plants. *Parasponia* hemoglobin is grouped with these non-symbiotic hemoglobin genes on the basis of sequence homology, but is unusual in that it appears to be bi-functional, being expressed in nodules where it presumably has a role in nitrogen fixation and at a low level in root tissue where it may have a non-symbiotic function.

Casuarina non-symbiotic hemoglobin.

A non-symbiotic hemoglobin gene was cloned from *Casuarina* by using the *Parasponia* hemoglobin gene as probe under conditions of low stringency. The predicted amino acid sequence of the hemoglobin encoded by the gene isolated is more similar to the *Parasponia* and *Trema* hemoglobin protein sequences than those of the leghemoglobins, and it has a N-terminal extension of five amino acids (Christensen et al. 1991). Northern analysis detects expression of this gene in stem, leaves and roots (Christensen et al. 1991). It has, therefore, been called the *Casuarina* non-symbiotic hemoglobin (*Cas-nonsym*) (Jacobsen-Lyon et al. 1995). Thus, *Casuarina* has both symbiotic and non-symbiotic hemoglobin genes.

Soybean non-symbiotic hemoglobin.

The discovery that *Casuarina* has both symbiotic and non-symbiotic hemoglobins lead to the hypothesis that legumes may also retain a non-symbiotic hemoglobin gene (Andersson 1994, Andersson et al. 1996). In order to ascertain if non-symbiotic hemoglobins are present in legumes, PCR primers were designed on the basis of the exon sequences conserved between non-symbiotic hemoglobins (*Parasponia, Trema*, *Casuarina* and barley at that time) and used to amplify non-symbiotic hemoglobin gene fragments from soybean genomic DNA. A fragment of a non-symbiotic hemoglobin gene was isolated and this in turn was used to isolate both genomic and cDNA clones (Andersson et al. 1996). This gene is more closely related to non-symbiotic hemoglobins than it is to leghemoglobins, and has been designated the soybean non-symbiotic hemoglobin gene, *Soy-nonsym* (Andersson et al. 1996). *Soy-nonsym* is expressed in leaves, stems, roots and at a low level in nodules (Andersson et al. 1996), similar to the pattern of gene expression observed for the *Casuarina* non-symbiotic hemoglobin gene.

Non-symbiotic hemoglobins in cereals.

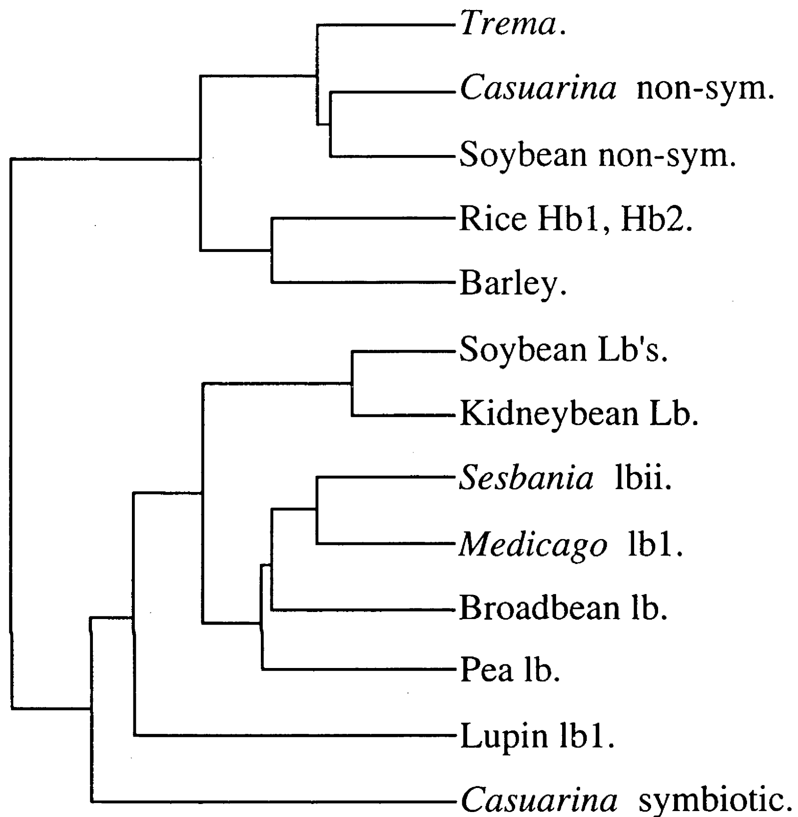
A hemoglobin cDNA was fortuitously isolated from the monocot barley, *Hordeum vulgare*, during the screening of an aleurone layer cDNA expression library with alpha-glucosidase antiserum (Taylor et al. 1994). This cDNA encodes a hemoglobin that, on the basis of sequence homology, groups with the non-symbiotic hemoglobins of other plants (Taylor et al. 1994).

Northern analysis of barley hemoglobin gene expression at different oxygen concentrations revealed a strong response to hypoxia. The barley

hemoglobin transcript is induced by hypoxia, in both roots and coleoptiles, with maximal response at 2% oxygen (Taylor et al. 1994). Furthermore, the kinetics of transcript accumulation are similar to known flooding response genes such as alcohol dehydrogenase (ADH). This observation suggests that the barley hemoglobin may be required under hypoxic conditions (Taylor et al. 1994). Expression of the barley hemoglobin in aleurone layers is also induced by a range a range of mitochondrial inhibitors (Nie and Hill 1997). Respiratory inhibitors that reduce oxygen consumption (Carbon Monoxide, Cyanide and Antimyacin A) enhanced expression of the barley hemoglobin transcript, as did an uncoupling agent that increases oxygen consumption (Di-nitro-phenol, DNP). As inhibition of ATP synthase (Oligomycin) also enhanced hemoglobin expression it was concluded that induction of the barley hemoglobin is controlled by the availability of ATP rather than demand for oxygen directly.

The barley hemoglobin sequence was used to probe genomic DNA from other cereal crop plants (wheat, rye and maize). Hybridising sequences were detected, suggesting that similar hemoglobins are present in these plants (Taylor et al. 1994). When RNA from control and anoxia treated maize plants was probed with the barley hemoglobin cDNA probe an anoxia inducible transcript of the appropriate size was detected (Taylor et al. 1994). It appears that non-symbiotic hemoglobins are present in a wide range of cereals. The observation that there are two different rice hemoglobins with homology to non-symbiotic hemoglobins (Arredondo-Peter et al. in press) further supports this idea.

Figure 5. Similarity tree of plant hemoglobin protein sequences.



Tree simplified from Andersson et al. 1996. The plant hemoglobin protein sequences clearly group into two clades, the non-symbiotic and the symbiotic hemoglobins.

Summary of hemoglobin gene expression.

Barley: Low level in root tissue. Hypoxia inducible in roots and coleoptiles. Expressed in aleurone layer of barley grains and responds to mitochondrial inhibition (Taylor et al. 1994 and Nie and Hill 1997).

Parasponia **:** Low level in root tissue. High level in nodule structures (Landsmann et al. 1986)

Trema **:** Low level in root tissue (Bogusz et al. 1988).

Cas-nonsym**::** Root tissue, stems and leaves. Low level in nodules. Induced by sucrose (Christensen et al. 1991, Jacobsen-Lyon et al. 1995).

Soy-nonsym **:** Root tissue, stems and leaves. Low level in nodules (Andersson et al. 1996).

Leghemoglobins: Nodules only (Hyldig-Nielsen et al. 1982).

Cas-sym: Nodules only (Jacobsen-Lyon et al. 1995).

The promoters of plant hemoglobin genes.

Two conserved sequence motifs (AAAGAT and CTCTT) occur in close proximity to each other in the promoters of leghemoglobins, the *Casuarina* symbiotic hemoglobin and other nodule expressed genes (Sandal et al. 1987, Metz et al. 1988, Szczyglowski et al. 1994) (Figure 6). The promoters of the soybean Lbc3 and Lba, and *Sesbania rostrata* glb3 genes have been fused to a reporter gene (β–glucuronidase or chloramphenicol acetyl transferase) and transformed into *Lotus corniculatus* (a transformable, nodulating legume) where high levels of nodule expression were observed. Deletion of the region containing the AAAGAT and CTCTT "nodulin motifs" from these leghemoglobin gene promoters eliminates reporter gene expression (Stougaard et al. 1987, She et al. 1993, Szczyglowski et al. 1994). Mutation of the CTCTT sequence motif abolishes reporter gene expression, while the effects of mutations in the AAAGAT motif vary between the promoters of different leghemoglobin genes (Ramlov et al. 1993, Szczyglowski et al. 1994).

Figure 6. Conserved sequences in leghemoglobin promoters.

-140 -75

TTGA**AAAGAT**GATTGT**CTCTT**CACCATACCATTGATCACCTCCTCCAACAAGCC**AAGAG**

The region of promoter from the LbC3 gene of soybean that contains the conserved "nodulin box" motifs (Stougaard et al. 1987). A second CTCTT motif also occurs in reverse orientation (AAGAG) in many leghemoglobin genes downstream from the AAAGAT and CTCTT motifs (all shown in bold).

The promoters of the *Trema* and *Parasponia* hemoglobin genes have also been studied in detail. Both have been fused to the β-

glucuronidase reporter gene (GUS) and transformed into *Lotus*. In transformed *Lotus* plants, staining was found in the nodules and in the vascular tissue and tips of roots (Bogusz et al. 1990). It seems that the promoters of both the *Trema* and *Parasponia* hemoglobin genes contain the necessary regulatory elements to direct both symbiotic and non-symbiotic patterns of gene expression. The observation that the *Trema* hemoglobin promoter is able to confer nodule expression to the GUS reporter gene suggests that *Trema* has lost the ability to nodulate but the *Trema* hemoglobin gene has retained the capacity for nodule expression.

The same constructs have been transformed into tobacco (Landsmann et al. 1988, Bogusz et al. 1990). In tobacco the promoters of the *Trema* and *Parasponia* hemoglobin genes produce similar patterns of reporter gene expression. Staining is found in the vascular tissue of roots and in the root tips. This is similar to the pattern of expression produced by these promoter-reporter gene fusions in the roots of transgenic *Lotus*. It seems that the promoter elements required for root tissue expression of the *Trema* and *Parasponia* hemoglobin genes can be recognised by both tobacco and *Lotus* despite the fact that these plants are not closely related (see Chase et al. 1993).

The *Cas-nonsym* promoter has also been linked to the β-glucuronidase reporter gene and transformed into *Lotus*. Staining is found in the vascular tissue of roots, and in root tips. A low level of staining was also found in nodules (Jacobsen-Lyon et al. 1995). This pattern is similar to the pattern of staining produced when the *Parasponia* and *Trema* promoter GUS fusions were transformed into *Lotus* (Bogusz et al. 1990). The *Cas-nonsym* GUS fusion was found to be induced by 1% sucrose solution (Jacobsen-Lyon et al. 1995), a treatment that has not been

examined with the *Parasponia* and *Trema* GUS fusions. The *Casuarina* promoter construct was inserted using hairy root *Rhizogenes* transformation (Jacobsen-Lyon et al. 1995), and as a result the exact location of leaf and stem expression, detectable on RNA gel blots of *Casuarina* RNA (Jacobsen-Lyon et al. 1995), is unknown.

The conserved sequence motifs (AAAGAT and CTCTT) that occur in the promoters of leghemoglobins have been identified in the promoters of the *Parasponia* and *Trema* hemoglobin genes, although the CTCTT motif is found in reverse orientation AAGAG. Deletion of the region containing the AAAGAT motif, or the region containing both motifs, in the *Parasponia* hemoglobin promoter abolishes GUS expression from promoter-reporter gene fusion constructs (Andersson et al. 1997). Mutation of the AAGAG motif also reduces GUS expression in both roots and nodules of transgenic *Lotus* (Andersson et al. 1997). However, mutation of the AAAGAT did not affect reporter gene expression despite the fact that deletion of this element did reduce reporter gene expression. A CT rich region, with some homology to the CTCTT motifs of leghemoglobins is also found between the AAGAT and AAGAG elements of the *Parasponia* and *Trema* hemoglobin gene promoters. The importance of this sequence has not been examined by mutagenesis.

Similar sequence motifs (AAGAT, CTCTT and AAGAG) are found in the promoter of the *Casuarina* non-symbiotic hemoglobin (Jacobsen-Lyon et al. 1995) but neither deletion nor mutagenesis analyses have been attempted with this promoter.

1.4 The function of non-symbiotic hemoglobins?

The occurrence of non-symbiotic hemoglobin genes in a wide range of plants, combined with the conserved sequences and the similarities in expression patterns of these hemoglobin genes suggest a conserved function for non-symbiotic hemoglobins.

Oxygen transporters?

A model that is able to account for all the data is that the non-symbiotic hemoglobins function as oxygen carriers (Appleby et al. 1988). If sufficient concentrations of hemoglobin are present in a tissue the amount of oxygen entering cells will be increased, as occurs in vertebrate muscle tissues. The constitutive low level of expression of non-symbiotic hemoglobins in root tissue may be required to increase oxygen availability, especially in the vascular bundles that are metabolically active, but often buried deep in the root. The expression in leaves and stems of some plants is also explicable if metabolic demand for oxygen outstrips supply. This may occur in young leaves where metabolic activity is high or around the vascular bundles where accumulation of sucrose into the phloem requires active metabolism. While the expression of these genes in roots has been localised to specific areas that are consistent with the "Oxygen Transport" model, the location of leaf and stem tissue expression remains unknown.

This "Oxygen Transport" model is supported by the hypoxia inducibility of the barley hemoglobin (Taylor et al. 1994). The accumulation of hemoglobin may increase oxygen diffusion rates and thus allow maintenance of respiration during hypoxia. Although hypoxia induction of non-symbiotic hemoglobin gene expression has not been

observed in dicots, it has never been analysed thoroughly using RNA gel blots and defined gas mixes.

The sucrose inducibility of the *Cas-nonsym* gene is also consistent with the oxygen model. Sucrose is known to have a number of effects at the cellular level (see Koch 1996) and has been shown to affect the expression of other plant genes such as sucrose synthase (Martin et al. 1993) and glyceraldehyde-phosphate dehydrogenase (Yang et al. 1993). Furthermore, sucrose may increase the rate of mitochondrial metabolism, and thus oxygen consumption, (Bingham and Stevenson 1993) and accumulation of hemoglobin may be required to increase the rate of oxygen flux in this situation.

The major problem with the oxygen transporter model is that high concentrations of hemoglobin protein are required for an increase in flux to occur (Wittenberg et al. 1974, Appleby 1985). Such high concentrations of protein may not be present. An estimate of 20μM hemoglobin in barley aleurone layers (Nie and Hill 1997) is well below the millimolar concentrations of leghemoglobin found in legume nodules. For the oxygen transport model to be valid, it is also necessary to assume that non-symbiotic hemoglobins have suitable oxygen binding kinetics.

Oxygen sensors?

It has been proposed (Appleby et al. 1988) that non-symbiotic plant hemoglobins may act as oxygen sensors, which would sense the onset of hypoxia in plants. This would account for the low level of expression in root tissues that is found for all these hemoglobin genes, as only a small amount of protein product would be required for an oxygen sensing signal transducer. An example of an oxygen sensing heme-protein is already

known, the Fixl protein of *Rhizobium meliloti* (Gilles-Gonzales et al. 1991). However, the Fixl protein is directly linked to a histidine kinase domain, and shows homology to the well characterised two component signal transduction systems of prokaryotes (Gilles-Gonzales et al. 1991). If non-symbiotic plant hemoglobins are functioning as a oxygen sensors then there must be an associated protein that can initiate a signal cascade (a kinase for instance). It is also necessary to assume that the binding of oxygen to non-symbiotic hemoglobins results in alterations of either protein conformation or redox potential that can be communicated to another protein. The varied expression patterns of non-symbiotic hemoglobin genes, with expression in the leaves and stems of some plants but not others, is also difficult to reconcile with this model, as is the hypoxia response of the barley hemoglobin gene and the sucrose response of the *Cas-nonsym* gene.

Other functions?

A slight variation on the oxygen transport theory is that non-symbiotic hemoglobins act as oxygen scavenging proteins. In hypoxic tissues, induction of hemoglobin expression could lead to more oxygen being bound for subsequent metabolic use. Bound oxygen could be channelled directly to respiring mitochondria or to other oxygen requiring processes. A role for non-symbiotic hemoglobins in supplying oxygen to oxygen consuming enzymatic reactions is also possible. It has been suggested that the hemoglobin of the nematode *Ascaris* may channel oxygen into the synthesis of a steroid hormone under conditions where oxygen concentrations are very low (see Goldberg 1995). It is possible that non-symbiotic hemoglobins may be required in oxygen linked enzymatic catalysis. For either of these theories to be valid, it must be assumed that non-symbiotic hemoglobins are able to interact with other proteins *in vivo*.

1.5 Determining non-symbiotic hemoglobin function.

One way to examine the function of non-symbiotic hemoglobins is to clone a non-symbiotic hemoglobin from a model plant such as *Arabidopsis thaliana*, which is more amenable to molecular genetic analysis. The expression pattern of a hemoglobin gene could be examined in *Arabidopsis* with northerns and with promoter-reporter gene fusions. A focused study on the hemoglobin of one plant using a range of techniques may identify important features of non-symbiotic hemoglobin gene expression. The ease with which *Arabidopsis* can be transformed also allows a "reverse genetics" approach to be implemented. Strains that express more or less hemoglobin could be engineered (with sense or antisense constructs respectively) and their phenotype examined.

Arabidopsis hemoglobins.

A fragment of a non-symbiotic hemoglobin gene was amplified and cloned from *Arabidopsis thaliana* using the same PCR strategy used to isolate the soybean non-symbiotic hemoglobin gene (Andersson 1994, Andersson et al. 1996). This thesis will describe how the PCR fragment amplified by Andersson et al. was used to isolate a *Arabidopsis* non-symbiotic hemoglobin gene which has subsequently been characterised in some detail. Contrary to prior expectations a second hemoglobin gene that is more closely related to the symbiotic hemoglobins of nitrogen-fixing plants has been found in *Arabidopsis*. The cloning and characterisation of this second hemoglobin gene will also be described.

CHAPTER 2.

Methods.

2.1 Plant growth.

Arabidopsis C24 ecotype seeds were surface sterilised, then grown on MS agar with 2% sucrose, 40-50 plants/plate, at 22° (16/8hr light/dark cycle, 100 μEs^{-1}) until harvest. Exceptions were plants used for RNA extraction and RNA gel blot analysis, which were grown on agar without sucrose. Dehydration, chilling and hypoxia treatments were performed as outlined in Dolferus et al. (Dolferus et al. 1994). Treatments were for 12 hrs, with the exception of chilling which was for 24 hr . Sucrose treatment involved placing plantlets in MS with 1% sucrose, and incubating in the dark overnight. Wounding was achieved by squeezing leaves with blunt nosed forceps, and oxidative stress treatment was performed by placing the plants in 300 mM H_2O_2 for 1, 2, 4 and 12 hr time spans.

2.2 Library screening and DNA manipulation.

Arabidopsis genomic DNA libraries (λ-EMBL4 9-15 kb Sau3A digest or 8 kb BamH1 inserts, C24 ecotype) were screened with random primed probes (Amersham). λ YES (Elledge et al. 1991) cDNA library was also screened with the same probes. Screening and subsequent subcloning involved standard conditions and DNA manipulation, as described in Ausubel et al. (Ausubel et al. 1994).

2.3 RNA Gel Blot Analysis.

RNA was extracted using the method of Logemann et al. (Logemann et al. 1987) and separated on formaldehyde gels as described in Dolferus et al. (Dolferus et al. 1994), as was RNA gel blotting, hybridisation and post hybridisation washes. For RT-PCR template, two additional phenol extractions were included directly after lithium precipitation, and a carbohydrate precipitation (250 μl ethanol, 40 μl sodium acetate, 10 mins on ice, 10 mins at 13000g centrifugation) was included prior to the final sodium acetate/ethanol precipitation.

2.4 RT-PCR.

PCR was carried out on first strand cDNA, that was synthesised from hypoxia induced RNA with MMLV reverse transcriptase as per company specifications (Promega). PCR was then performed with specific primers (100 nM), and standard PCR reaction conditions (see Ausubel et al. 1994). Primers were annealed at 52° for 30 seconds, extended at 72° for 1 minute, then denatured at 95° for 30 seconds. This cycle was repeated 26 times on a Corbett Capillary PCR machine.

2.5 Database Search.

Database searches were conducted through key word searching of the EST database, compiled by the *Arabidopsis* EST sequencing project, on the National Centre for Biotechnology World Wide Web site (http://www.ncbi.nlm.nih.gov). Clone 11161 was detected using the key words "Hemoglobin" and "Arabidopsis". BLAST searches were conducted through the same internet site.

2.6 Expression of recombinant AHB1 and AHB2 in *E. coli.*

RT-PCR was performed using 5' primers containing a strong ribosomal binding site from the *Pseudomonas putida* P-450 gene (see Springer and Sligar 1987). The sequences of the 5' primers were as follows.

5' AHB1:

TTGGATCCATAACTAACTAAAGGAGAACAACAACA
A**ATG**GAGGTGAAGGAAAGATTGTGTTC

5' AHB2:

TTGGATCCATAACTAACTAAAGGAGAACAACAACA
ATGGGAGAGATTGGGTTTACAG

Primers include BamH1 sites (large text) to facilitate cloning, stop codons in all three reading frames to prevent translational read through from the LacZ ribosomal binding site, a strong ribosomal binding site (underlined) and then spacing sequence up to the ATG of the cDNA sequence (bold).

3' Primers included 20 bp of sequence designed to pair with the target gene (20 bp 5' from the stop codon up to and including the stop codon), with an additional stop codon and a BamH1 site to facilitate cloning. The RT-PCR reactions were performed as described above, and the resulting cDNA fragments were purified using Promega Wizard^{tm} columns (Promega), prior to ligation into pUC 119. Strains containing AHB1 or AHB2 expression vectors were grown at 37° in Luria broth with 1 mM IPTG for 16 hrs.

2.7 Western analysis of bacterial strains.

Total protein was extracted by pelleting 1ml of overnight culture and resuspending in 200 µl of water, then adding 200 µl of 2x SDS loading buffer and boiling for ten minutes, then cooling on ice. 10 µl of each expression strain extract, and extract from a strain that contained pUC 119, were then loaded onto a 15% denaturing polyacrylamide gel (with stacking gel), and separated at 50 mA. Gels were then electro-blotted overnight onto nitrocellulose and stained in india ink to check that loading was equivalent. Blots were then washed in water, 3 x 5mins, blocked in milk powder blocking mix for 2 hrs and then reacted for 2 hrs with anti-Phb (1/1800 dilution) or anti-cassym (1/1500) in blocking mix for 2 hrs. Blots were washed in NTE (sodium/Tris/EDTA), 5 x 20 mins, and then reacted with alkaline phosphatase conjugated secondary antibody, in blocking mix for 2 hrs, and washed again prior to color reaction with NBT/BCIP for the appropriate time span for the individual blots. All solutions and buffers were made according to Ausubel et al. (Ausubel et al. 1994).

2.8 Western analysis of AHB1 levels in *Arabidopsis* tissues.

Plant protein was extracted by grinding a gram of plant tissue in 1ml of native extraction buffer, pelleting the extract (3mins 10000g), and concentrating the supernatant by a factor of 4 using acetone precipitation and resuspension in 250 µl extraction buffer. 250 µl of 2x SDS running buffer was then added to the extracts, which were then boiled for ten minutes. 40 µl of extract were then separated on 15% polyacrylamide gels, as outlined above with the exception that 1.5 mm spacers were used instead of 0.75 mM spacers to allow the loading of greater extract volumes. Gels were then blotted and reacted with anti-sera as outlined above.

2.9 *AHB1* promoter-GUS fusion.

980 bp (Hind/Dra fragment) of AHB1 5' sequence was cloned into pBI121.1 (Clontech) to produce the GH1 construct. The GH1 construct was then electroporated into electrocompetent *Agrobacterium tumefaciens* (Aglil) strain, prepared as for *E coli* protocol in Ausubel et al. (Ausubel et al. 1994) but initial culturing, and post electroporation recovery steps were for twice the time recommended for *E. coli* to allow for slower growth rate of *Agrobacterium*. Plant transformation was then performed according to the tissue culture transformation protocol of Valvekans et al. (Valvekans et al. 1988). T1 plants were germinated on MS plates with kanamyacin, and kanamyacin resistant plants were transferred to compost until seeds were produced. Of the 20 T1 lines initially generated, 6 were taken to the T2 generation and analysed.

2.10 GUS staining.

Plants were germinated and grown on MS plates and subjected to stress treatments as described above, then stained using the staining protocol of Jefferson (Jefferson 1987).

CHAPTER 3.

The cloning and characterisation of *AHB1*.

3.1 Introduction.

Hemoglobin genes from organisms that are not closely related often show very low levels of sequence identity. This reflects the fact that hemoglobin proteins seem to have few conserved amino acid residues (Bashford et al. 1987) and that a wide range of amino acid sequences can produce the globin fold structure. As a result of this sequence diversity cloning hemoglobin genes can be difficult. Methods that rely on overall sequence homology, such as low stringency hybridisation, are unlikely to succeed. It is not surprising, therefore, that early attempts to detect hemoglobin sequences in non-legumes using low stringency hybridisation did not succeed (Baulcombe and Verma 1978). Similarly, a *Parasponia* hemoglobin probe did not detect homologous sequence in any plants that were not closely related (Bogusz et al. 1988). Thus, to clone non-symbiotic plant hemoglobins from soybean and *Arabidopsis* an alternative approach that does not rely on overall sequence homology was adopted.

To isolate non-symbiotic hemoglobin genes from soybean and *Arabidopsis* Andersson et al. (1996) chose a PCR based strategy. This method employs degenerate primers, designed on the basis of small stretches of highly conserved sequences to selectively amplify the gene of interest. Non-symbiotic hemoglobin genes had already been cloned from a range of plants, including both dicots and a monocot. This provided some indication of coding sequences conserved within these genes. and Andersson et al. (Andersson et al. 1996) designed degenerate PCR primers

on the basis of these conserved regions of non-symbiotic hemoglobin genes.

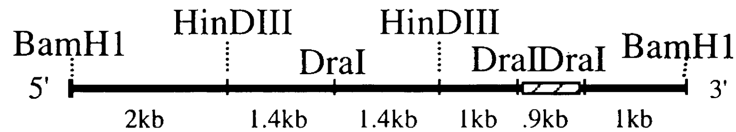
Non-symbiotic hemoglobin gene fragments (Andersson 1994, Andersson et al. 1996) were successfully amplified and isolated from both soybean and *Arabidopsis*. The isolation of a non-symbiotic hemoglobin gene fragment allowed an entire non-symbiotic hemoglobin gene to be cloned from *Arabidopsis*. This chapter describes the cloning and characterisation of this hemoglobin gene, *AHB1*.

3.2 Results.

Cloning of *AHB1*.

An *Arabidopsis* (C24 ecotype, 9kb fraction) genomic library was screened at high stringency using the 2f4r2 PCR fragment (the non-symbiotic hemoglobin PCR fragment, Andersson 1994) as a probe. Several positive plaques were isolated. One, λ111 was characterised in detail. Phage DNA from λ111 was isolated and shown to contain a 8kb insert that cross hybridises with the 2f4r2 probe. This fragment corresponds to the 8kb BamH1 fragment detected on genomic southerns using the 2f4r2 probe (Andersson 1994) The 8kb BamH1 insert was cloned into pUC119. A partial restriction map of this clone, Ara111, is shown (Figure 7).

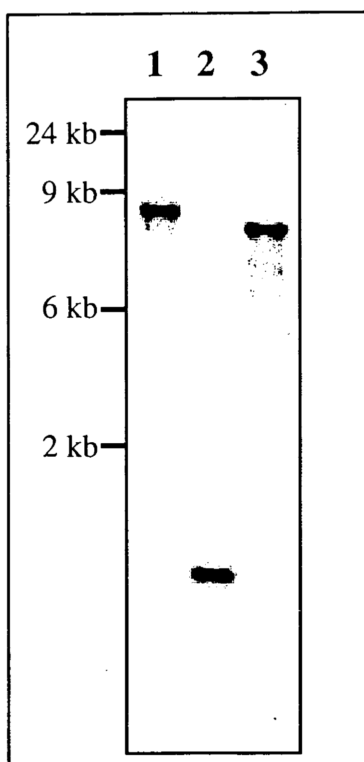
Figure 7. Partial restriction map of Ara111 λ clone.



A partial restriction map of Ara111 λ clone. Fragment sizes are shown in kilobases, coding sequence is represented as a shaded box. Figure is labelled 5' to 3' according to the orientation of the AHB1 coding sequence.

Several restriction fragments were subcloned and sequenced. In total, approximately 2kb of contiguous sequence from the region that hybridises with the 2f4r2 probe has been sequenced (Figure 7, Figure 9). The sequence from bases 272 to 681 is identical to that of the 2f4r2 probe confirming that this clone corresponds to the PCR fragment isolated by Andersson et al. (1994).

Figure 8. Genomic DNA blot probed with *AHB1*.



Arabidopsis (C24 ecotype) genomic DNA digested with BamH1, Dra1, and EcoR1 (lanes 1, 2 and 3), hybridised with *AHB1* cDNA probe.

***AHB1* is a single copy gene.**

An *AHB1* sequence was used to probe blots of *Arabidopsis* genomic DNA cut with a range of enzymes. This fragment hybridised, at high stringency, to either one or two bands of genomic DNA in all lanes (Figure 8). This suggests that there is a single copy of *AHB1* in the *Arabidopsis* genome. An identical banding pattern was also observed on low stringency southerns, with no extra bands evident. Thus, there is no

evidence for the presence of any closely related sequences in the *Arabidopsis* genome.

Analysis of *AHB1* sequence.

Sequence obtained from the Ara111 genomic clone contains an entire non-symbiotic hemoglobin gene, including upstream regulatory sequence. Putative translational initiation and termination codons are located at bases 1 and 739 respectively (Figure 9). Within this coding region four putative exons and three introns are predicted to exist on the basis of homology between the genomic sequence and other plant hemoglobin genes (Figure 8). These predicted introns are found in positions identical to the three introns found in all plant hemoglobins.

The predicted amino acid sequence of *AHB1* shows strong homology with the non-symbiotic hemoglobins of both monocots and dicots (Figure 10). Furthermore, cladistic analysis groups *AHB1* with the non-symbiotic hemoglobins of other plants (Figure 11). The presence of an extended N-terminal sequence (5 amino acids) is also characteristic of the non-symbiotic hemoglobin gene family.

Figure 9. *AHB1* Genomic sequence.

Stop and Start codons are shown in bold. Exon borders are underlined.

```
Protein sequence>                                           M
GGGGATCTACAAAACAGAGAGTTGTATACTTTAAATCATTTAGAGGTTGTGAAATATTAT
-58 ------+---------+---------+---------+---------+---------+2
CCCCTAGATGTTTTGTCTCTCAACATATGAAATTTAGTAAATCTCCAACACTTTATAATA

  E  S  E  G  K  I  V  F  T  E  E  Q  E  A  L  V  V  K  S  W
GGAGAGTGAAGGAAAGATTGTGTTCACAGAAGAGCAAGAGGCTCTTGTAGTGAAGTCTTG
---------+---------+---------+---------+---------+---------+62
CCTCTCACTTCCTTTCTAACACAAGTGTCTTCTCGTTCTCCGAGAACATCACTTCAGAAC

  S  V  M  K  K  N  S  A  E  L  G  L  K  L  F  I  K
GAGTGTCATGAAGAAAAACTCAGCTGAATTAGGTCTCAAACTCTTCATCAAGTAAGTAAT
---------+---------+---------+---------+---------+---------+122
CTCACAGTACTTCTTTTTGAGTCGACTTAATCCAGAGTTTGAGAAGTAGTTCATTCATTA

GATCCCATTGATCTCTCTCTATTTTCTTTTTATGTATATAGTCTGAGATATGAACTACTA
---------+---------+---------+---------+---------+---------+182
CTAGGGTAACTAGAGAGAGATAAAAGAAAAATACATATATCAGACTCTATACTTGATGAT

                 I  F  E  I  A  P  T  T  K  K  M  F  S  F  L
TTTTGAACTGTAGGATCTTTGAGATTGCACCAACAACGAAGAAGATGTTCTCTTTCTTGA
---------+---------+---------+---------+---------+---------+242
AAAACTTGACATCCTAGAAACTCTAACGTGGTTGTTGCTTCTTCTACAAGAGAAAGAACT

R  D  S  P  I  P  A  E  Q  N  P  K  L  K  P  H  A  M  S  V
GAGACTCACCAATTCCTGCTGAGCAAAATCCAAAGCTCAAGCCTCACGCAATGTCTGTTT
---------+---------+---------+---------+---------+---------+302
CTCTGAGTGGTTAAGGACGACTCGTTTTAGGTTTCGAGTTCGGAGTGCGTTACAGACAAA

F  V  M
TTGTCATGGTAATAATCAATATCAAATAACATGATTTTGCTTATATATTCGAATCAAAGA
---------+---------+---------+---------+---------+---------+362
AACAGTACCATTATTAGTTATAGTTTATTGTACTAAAACGAATATATAAGCTTAGTTTCT

                                C  C  E  S  A  V  Q  L  R  K  T
TTGTTGAGTTTTGGGGTTTATTATCAGTGTTGTGAATCAGCAGTACAACTGAGGAAAACA
---------+---------+---------+---------+---------+---------+422
AACAACTCAAAACCCCAAATAATAGTCACAACACTTAGTCGTCATGTTGACTCCTTTTGT

 G  K  V  T  V  R  E  T  T  L  K  R  L  G  A  S  H  S  K  Y
GGGAAAGTTACGGTGAGGGAGACTACTTTGAAGAGACTTGGAGCCAGCCATTCTAAATAC
---------+---------+---------+---------+---------+---------+482
CCCTTTCAATGCCACTCCCTCTGATGAAACTTCTCTGAACCTCGGTCGGTAAGATTTATG

 G  V  V  D  E  H  F  E  V
```

```
GGTGTCGTTGACGAACACTTTGAGGTTAGTAGTTATTTGTCATATCTCAAAATGTTCTTT
---------+---------+---------+---------+---------+---------+542
CCACAGCAACTGCTTGTGAAACTCCAATCATCAATAAACAGTATAGAGTTTTACAAGAAA


ATCATACAAATATGTTAACTTGATTTTTTTTGGTTGATGTAAAAATGATTTATAACTGCA
---------+---------+---------+---------+---------+---------+602
TAGTATGTTTATACAATTGAACTAAAAAAAACCAACTACATTTTTACTAAATATTGACGT


    A  K  Y  A  L  L  E  T  I  K  E  A  V  P  E  M  W  S  P
GGTGGCCAAGTATGCATTGTTGGAGACGATAAAGGAGGCAGTGCCGGAGATGTGGTCACC
---------+---------+---------+---------+---------+---------+662
CCACCGGTTCATACGTAACAACCTCTGCTATTTCCTCCGTCACGGCCTCTACACCAGTGG


 E  M  K  V  A  W  G  Q  A  Y  D  H  L  V  A  A  I  K  A  E
GGAGATGAAGGTGGCTTGGGGTCAGGCTTATGATCACCTTGTTGCTGCCATTAAAGCTGA
---------+---------+---------+---------+---------+---------+722
CCTCTACTTCCACCGAACCCCAGTCCGAATACTAGTGGAACAACGACGGTAATTTCGACT


 M  N  L  S  N  *
AATGAATCTTTCCAACTAAAAAATCATATACTATTATATAGTTGTAAACTTGTAATAAAT
---------+---------+---------+---------+---------+---------+780
TTACTTAGAAAGGTTGATTTTTAGTATATGATAATATATCAACATTTGAACATTATTTA


ATTTCATTTTGAATTGTTCTCATGACTGTTGTTCTATTTGGTTTGGTTTGATTTAGTGAC
---------+---------+---------+---------+---------+---------+840
TAAAGTAAAACTTAACAAGAGTACTGACAACAAGATAAACCAAACCAAACTAAATCACTG
```

Figure 10. Alignment of non-symbiotic hemoglobin protein sequences.

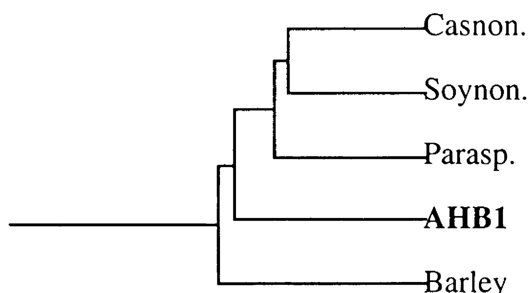
```
    1                                            *      C2     CD1        60
Bar .SAAEGAVVF SEEKEALVLK SWAIMKKDSA NLGLRFFLKI FEIAPSARQM FPFLRDSDVP
CN  .SSTLEGRGF TEEQEALVVK SWSAMKPNAG ELGLKFFLKI FEIAPSAQKL FSFLKDSNVP
Soy .TTTLE-RGF SEEQEALVVK SWNVMKKNSG ELGLKFFLKI FEIAPSAQKL FSFLRDSTVP
Par .SSSEVNKVF TEEQEALVVK AWAVMKKNSA ELGLQFFLKI FEIAPSAKNL FSYLKDSPVP
HB1 ..ESEGKIVF TEEQEALVVK SWSVMKKNSA ELGLKLFIKI FEIAPTTKKM FSFLRDSPIP

   61         E7      *                                F8          * 120
Bar LETNPKLKTH AVSVFVMTCE AAAQLRKAGK ITVRETTLKR LGGTHLKYGV ADGHFEVTRF
CN  LERNPKLKSH AMSVFLMTCE SAVQLRKAGK VTVRESSLKK LGASHFKHGV ADEHFEVTKF
Soy LEQNPKLKPH AVSVFVMTCD SAVQLRKAGK VTVRESNLKK LGATHFRTGV ANEHFEVTKF
Par LEQNPKLKPH ATTVFVMTCE SAVQLRKAGK VTVKESDLKR IGAIHFKTGV VNEHFEVTRF
HB1 AEQNPKLKPH AMSVFVMCCE SAVQLRKTGK VTVRETTLKR LGASHSKYGV VDEHFEVAKY

   121                                      163
Bar ALLETIKEAL PADMWGPEMR NAWGEAYDQL VAAIKQEMKP AE.
CN  ALLETIKEAV P-ETWSPEMK NAWGEAYDKL VAAIKLEMKP SS.
Soy ALLETIKEAV P-EMWSPAMK NAWGEAYDQL VDAIKSEMKP PSS
Par ALLETIKEAV P-EMWSPEMK NAWGVAYDQL VAAIKFEMKP SST
HB1 ALLETIKEAV P-EMWSPEMK VAWGQAYDHL VAAIKAEMNL SN.
```

Alignment includes: Bar: barley hemoglobin (Taylor et al. 1994), CN: *Casuarina* non-symbiotic (Christensen et al. 1991), Soy: soybean non-symbiotic (Andersson et al. 1996), Par: *Parasponia* hemoglobin (Landsmann et al. 1986). HB1: AHB1. Residues strongly conserved in all hemoglobins are shown in bold. Intron positions are shown with stars. Sequences were aligned with the pileup function of the GCG package.

Figure 11. Plant hemoglobin protein similarity tree.



Tree constructed with the Pileup function of the GCG package (Genetics Computer Group, see figure 10 for sequence references). Soybean leghemoglobin and *Casuarina* symbiotic hemoglobins were used as an outgroup but are not shown in this diagram. Barley: barley non-symbiotic hemoglobin (U01228). Casnon: *Casuarina* non-symbiotic hemoglobin (X53950). Soynon: soybean non-symbiotic hemoglobin (U47143). Parasp: *Parasponia* hemoglobin (U27194). AHB1: This study, (U94998).

In addition to regions that are predicted to be transcribed, the1kb region immediately 5' of the start codon has also been sequenced. This 5' region contains a number of sequences that may be important for transcriptional regulation of this gene. A putative TATA box, with 10 bp of sequence that is identical to that of the *Parasponia* hemoglobin TATA box, is found at -79 (Figure 12a). Furthermore, there are a number of 5' sequences with homology to motifs involved in the regulation of non-symbiotic hemoglobins and leghemoglobin genes (Figure 12b).

Figure 12a. ***AHB1*** **Promoter sequence.**

-1080 TTTTGAAGTG TTACGTGAGA CTACGACTTT GCTGGTGA**AA** **GAG**TTGATTA

-1030 TGAGAATGAT CTTTTACTTC CTCCTTATCC TGTCCATCCT CCTCTCTCCC

-980 TTGGTTTAGA TAATAATAAT GGTTACCTAT CTCACATACC TTCGAAGAAG

-930 AAATCGGATG CGAATGGGTT CGGCTCAGGA CATGTGAGGA ACGAGGCAGA

-880 GGCGAAAGTC CAACGGAGGA TCTGGAGGTT CGAGTCATGG GGTTATTCGC

-830 TTCTTAGGCT CTGTCGCAAG ATCGTGCTTC CAATCATCGG TGTTATCTCT

-780 CTATTGTCTG CATCAGGATA TGGTCCAGAG ATGAGGAAAA GAGGCGCTTC

-730 CTTGAACCTT TTTGGACTTC TCCGCACCGA GCTACCAGAG GGG**AAGAG**AA

-680 CACCTAACCA ATGTCCACCA GGGGAAGTTC TTGTGATTGA AGATGGAGAA

-630 GCACGGTGCC TTGTTAAGGA **AAGAG**TCGAA AATACCATTT GACTCTGTTG

-580 TAGCGAAACG CGATGTAACT TACGGATACG GTTGAGTTTG AAAGAAATTA

-530 AACACCCTCC TCTGTAATGT TCATTGCCTG ACATTCAGTG TGTAAAGTTT

-480 CAACCTTTTC TTACCTCTTA GATATGAAAG CCACATGGTG CTACATAACT

-430 CTCACAAGTT TTGTTCATAG TTGTGCAAAT AAATTGGATG AGGGAGAAGA

-380 TTCATTTGAG TAGTATAATC TCTGTTCTGC CAATTTGATC AATAGATTCA

-330 GTGTATCCCA ACTTATACCA CTTGGTGTGG TTGGCCAATA CATAAATAAG

-280 TAAATAATGA GAT**AAGAG**AT CTAAAAGACT TTTCTTTAGT GTTTTGACTA

-230 ATAATTGGTC AAGCCTACCA TTACAAACTA TGTTCCATTA CCAGTACAGA

-180 GCACATGGTT TCACTTTTTA ACCAAGCAAC TTTATCTTAA ACTTGCTTAT

-130 CAAGAACTGT CTCTTCGAAA ACACAAAAGG GTCTTTATCA AA**CCATATAT**

-80 **A**GGATACTTT TCATAATTGG AGTAAGATCT ACAAAACAGA GAGTTGTATA

-30 CTTTAAATCA TTTAGAGGTT GTGAAATATT +1 **(ATG)**

Sequence is numbered from the initiation of translation, designated +1. The putative TATA box is bold and underlined, 10bp identical to the putative TATA boxes of the promoters of *Parasponia* and *Trema* hemoglobin genes. The underlined region (which contains AAAGAT and CTCTT-like motifs) from -132 to -95 is compared to regions known to be important in the regulation of other plant hemoglobin promoters (see below). Other CTCTT (or AAGAG reverse complement) nodulin motifs are shown in bold. Possible Anaerobic Response Elements (AREs) are double underlined.

Figure 12b. Promoter motif sequence comparisons.

MOTIF		**AAAGAT**		**CTCTT**		**AAGAG**	
Sesb.[1]	-131	AAAGAT	6	CTCTT	30	AAGAG	
Cassym.[2]	-480	CAAGAT	6	CTCTT	21	AAGAT	
Casnon.[3]	-198	AAAGAA		CTTTC	7	AAGAG	
Para.[4]	-99	AAAGAT		CTCCC	8	AAGAG	
Trema[5]	-107	AAGGAG		CTCTC	8	AAGAG	
Soynon[6]	-184	AAATGG	6	CTCCC		GAGGG	
HB1	-132	ATCAAGAACTGTCTCTTCGAAAACACAAAAGGGTCTT					-95

Comparison of putative promoter motifs from the *AHB1* 5' region with the regulatory elements from the promoters of other plant hemoglobin genes. These elements occur in the same order in many plant hemoglobin genes. The numbers between motifs represent the spacing between these elements in the promoter in question if it is different from the spacing in the *AHB1* promoter. (1: Szczyglowski et al. 1994, 2: Christensen et al. 1991, 3: Jacobsen-Lyon et al. 1995, 4: Landsmann et al. 1986, 5: Bogusz et al. 1988, 6: Andersson et al. 1996)

MAIZE Adh1 B1	GCCCCGGTTT
MAIZE Adh1 B2	GCCGTGGTTT
MAIZE Adh2	CCCCTGGTTT
AHB1 -979	CCCTTGGTTT
AHB1 -175	GACATGGTTT

Comparison of putative AREs from the *AHB1* 5' region with the AREs from the Maize ADH genes (Walker et al. 1987, Olive et al 1990 Paul and Ferl 1991).

Analysis of *AHB1* gene expression.

The expression of the *AHB1* gene was analysed on RNA gel blots. A low level of *AHB1* transcript was detected in root tissue (Figure 13) but no transcript was detected in RNA extracted from above ground tissues of

plants growing under normal conditions. The effects of a range of environmental stress conditions on the transcription of *AHB1* were also examined. Hypoxia treatment was found to increase *AHB1* transcript levels. This induction was found in both above and below ground tissue (Figure 13). Hypoxic induction was observed with 5% oxygen treatment but stronger induction was observed after 0.1% oxygen treatment (Figure 14). The time course of *AHB1* transcriptional response was remarkably similar to that of the alcohol dehydrogenase gene (*ADH*). Maximal response was observed after 6-8 hrs treatment (Figure 15). Sucrose was also found to induce *AHB1* gene expression. Induction was maximal with 1% sucrose, but was also elicited to a lesser extent by 2% sucrose (data not shown). Chilling, dehydration, oxidative stress, heatshock, and wounding had no observable effect on *AHB1* transcript levels.

Figure 13. ***AHB1*** **expression in above and below ground tissues.**

Total RNA extracted from rosette leaves (lane 1), rosette leaves treated with 5% oxygen (lane 2), roots (lane 3), roots treated with 5% oxygen (lane 4), hybridised with an *AHB1* antisense ribo-probe.

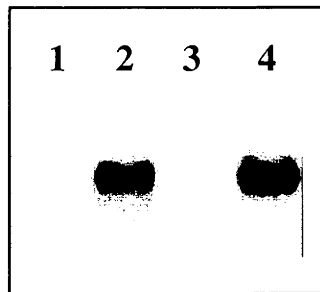


Figure 14. ***AHB1*** **expression during environmental stress.**

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) hybridised with an *AHB1* antisense ribo-probe.

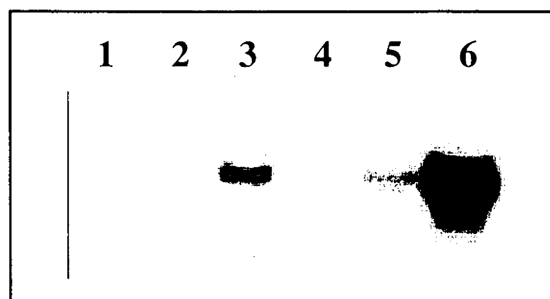
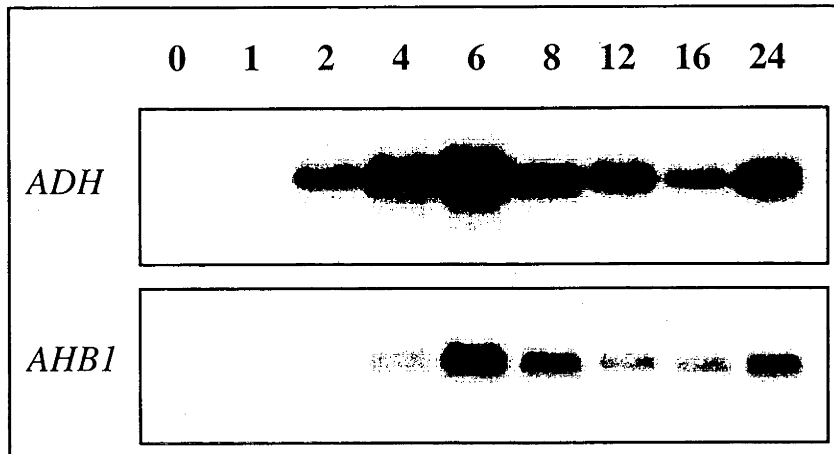


Figure 15. Time course of the *AHB1* hypoxia response.

Total RNA from whole plants treated with 5% oxygen for varying times hybridised with either *AHB1* or *ADH* antisense ribo-probes.



Isolation of a *AHB1* cDNA clone.

To allow further characterisation of *AHB1* attempts were made to isolate a cDNA clone. Initially a λ-YES cDNA library, with cDNA derived from a wide range of tissues, was screened with the 2f4r2 fragment as a probe. Over 2 x10^5 plaques were screened, at high stringency, but no hybridising clones were detected. The failure to isolate a *AHB1* cDNA was probably due to the low abundance of *AHB1* transcript in untreated plants, which is the source of the RNA used in the synthesis of this λ-YES cDNA library.

As there were no suitable cDNA libraries available (ie that were synthesised from hypoxia induced mRNA) an alternative approach to clone the *AHB1* cDNA was implemented. Primers specific to *AHB1* were used

to amplify a fragment from first strand cDNA template, reverse transcribed from hypoxia induced root RNA. This Reverse Transcriptase-Polymerase Chain Reaction (RT-PCR) approach allowed a *AHB1* cDNA fragment to be amplified and cloned. The sequence of this fragment matches that of the predicted exonic regions of the genomic clone, and confirms the positions of the three putative introns which are absent from the cloned cDNA sequence.

3.3 Discussion.

The predicted protein sequence of *AHB1* contains amino acid residues that are highly conserved in plant hemoglobins, and hemoglobins in general. The three introns, which are found in identical positions in all plant hemoglobins, are found in the same positions in the *AHB1* gene. Thus, *AHB1* encodes a hemoglobin protein. Furthermore, *AHB1* is predicted to encode a hemoglobin which shows high levels of amino acid sequence similarity to non-symbiotic plant hemoglobins and has a N-terminal region that is longer than those of symbiotic hemoglobins, a feature characteristic of non-symbiotic hemoglobins.

Most non-symbiotic hemoglobin genes identified in other plants are found as single copy genes, with the exception of rice which appears to have two closely related genes (Arrendondo et al. in preparation). The presence of two closely related non-symbiotic hemoglobin genes in rice may be an unusual situation that has resulted from extensive crossing of wild strains and subsequent genome rearrangements during the breeding of rice as an agricultural crop. The *Arabidopsis* hemoglobin, like most other non-symbiotic hemoglobins, appears to be a single copy gene. It also appears, on the basis of low stringency hybridisation analysis, that there are no sequences in the *Arabidopsis* genome that are closely related to *AHB1*. Thus, we would predict that the *Arabidopsis* genome contains just this single non-symbiotic hemoglobin gene.

The expression of the *Arabidopsis* non-symbiotic hemoglobin gene, *AHB1*, shows a number of similarities with the expression patterns of non-symbiotic hemoglobin genes from other plants. The low level of expression in root tissue that seems to be a common feature of this family of genes is also found for *AHB1*. Furthermore, the hypoxia inducibility of

the *AHB1* transcript, in above and below ground tissues, is similar to that of the barley hemoglobin (Taylor et al. 1994). Hypoxia response may be a conserved feature of all non-symbiotic plant hemoglobins, but has not been examined thoroughly using RNA gel blots and defined gas mixes for the non-symbiotic hemoglobin genes of *Casuarina*, soybean and *Trema*, or for the *Parasponia* hemoglobin gene.

Induction of gene expression by sucrose, similar to that of *AHB1*, has been observed for the *Casuarina* non-symbiotic hemoglobin promoter-GUS fusion in transgenic *Lotus* (Jacobson-Lyon et al. 1995). Sucrose is known to affect the transcription of a wide range of genes, especially those connected with carbohydrate metabolism and may increase respiratory rates (Koch 1996). Induction by sucrose has not been examined for other non-symbiotic hemoglobins, and like hypoxia induction, may be a general feature of non-symbiotic plant hemoglobins.

The promoter sequence of *AHB1* shows some homology to the promoters of other non-symbiotic hemoglobins. The putative TATA box region shows strong homology to the equivalent area of the *Parasponia* and *Trema* hemoglobin promoters. There are also a number of sequences that resemble the nodulin motifs from other plant hemoglobin gene promoters. The 5' region of *AHB1,* from -132 to -95, contains three sequences with homology to the AAAGAT and CTCTT motifs (Figure 12a and 12b). These putative motifs occur in the same order and orientation as in other plant hemoglobin promoters.

Anaerobic Response Elements (AREs) have been characterised in the promoters of hypoxia inducible genes (Walker et al. 1987, Olive et al. 1990, Paul and Ferl 1991, Dolferus et al. 1994). Typically AREs consist of a GC rich motif followed by a GT rich motif (Figure 12b), and can function

in either orientation (Olive et al. 1990). The GT rich motif (the TGGTTT core sequence) has been shown to be necessary for anaerobic induction (Olive et al. 1990). As the hypoxia response of the *AHB1* gene is remarkably similar to that of the *ADH* gene, the 5' region of *AHB1* was examined for AREs that may be involved in the anaerobic response of *AHB1*. Two such motifs are present in the *AHB1* 5' region (-979 and -175). Furthermore, the TGGTTT motif at -979 also has a neighbouring GC rich element (Figure 12b). Further analysis of the *AHB1* promoter is required to determine if these motifs have any functional significance.

The promoter sequence motifs involved in the regulation of sucrose inducible gene expression have not been characterised to the same extent as AREs. Promoter deletion analysis has been carried out on the promoters of some sucrose responsive genes, including the patatin I and the proteinase inhibitor II genes from potato, and the malate synthase gene from cucumber. The sequences required for sucrose response from these promoters have been localised to small regions (Liu et al. 1990, Kim et al. 1991, Sarah et al. 1996), but there is no obvious homology between these regions. Thus, it is difficult to identify the sequences responsible for the sucrose response of the promoter of *AHB1* by sequence data alone, as no core sequence elements required for sucrose induction have yet been identified in the promoters of other plant genes. Promoter deletion or promoter mutation analysis of the *AHB1* 5' region may help localise sucrose response motifs and determine if any of the ARE-like or nodulin-like motifs have any functional importance.

The cloning of a non-symbiotic hemoglobin gene from *Arabidopsis* supports the idea that non-symbiotic hemoglobins may be present in all plants. This gene also shares a number of features with other non-symbiotic hemoglobins. It is a single copy gene that shows strong amino

acid sequence similarity to other non-symbiotic hemoglobin and has a similar pattern of gene expression. Thus, *AHB1* should provide a good model of non-symbiotic hemoglobin function and further analysis of *AHB1* may help to define the function of non-symbiotic plant hemoglobins in general.

CHAPTER 4.

Cloning and characterisation of *AHB2*.

4.1 Introduction.

Genes that are expressed exclusively in the nodules of legumes often show considerable homology to other genes that are expressed in normal tissues of legumes and other plants. It seems that gene duplication, and subsequent recruitment of one of the resulting genes into nodule function, was a common theme during the evolution of the legume-*Rhizobium* nitrogen-fixing symbiosis. For example, there are nodule specific forms of glutamine synthase (Roche et al. 1993) and sucrose synthase (Thummler and Verma 1987), both enzymes that have important functions in normal plant metabolism.

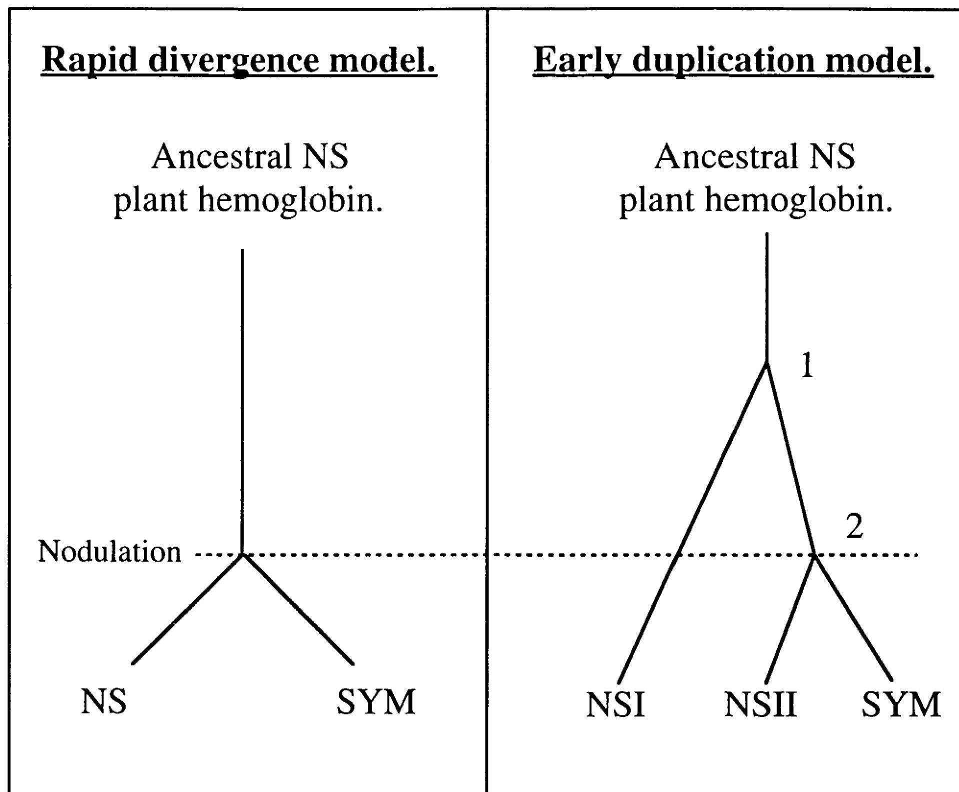
The evolutionary origin of the symbiotic hemoglobins of legumes and *Casuarina* has previously been attributed to a duplication of a non-symbiotic hemoglobin gene in the ancestor of nitrogen-fixing plants. This duplication is presumed to have occurred during the early stages of the evolution of nitrogen fixation. Subsequently the sequences of the duplicated genes would have diverged as one was recruited into a symbiotic role while the other retained a non-symbiotic function (Bogusz et al. 1988, Andersson et al. 1996). The high degree of divergence that is observed between the non-symbiotic and the symbiotic gene families was attributed to a rapid evolution of symbiotic hemoglobins following recruitment into a symbiotic function. The recruitment into symbiotic function may have removed sequence constraints associated with non-symbiotic hemoglobin gene function, allowing rapid sequence alterations to occur as these hemoglobins became specialised symbiotic proteins.

The alternative possibility, that the degree of divergence between symbiotic and non-symbiotic hemoglobins is due to an earlier gene duplication event which occurred before the evolution of nodulation seemed unlikely as no sequence resembling symbiotic hemoglobins had been identified in plants that do not fix nitrogen (Figure 16). Attempts to detect sequences with homology to leghemoglobins, using leghemoglobins as probes on low stringency southerns (Baulcombe and Verma 1978), failed to convincingly detect any related genes in non-legumes.

At present considerable resources are being invested in sequencing cDNAs from a range of *Arabidopsis* tissues as part of the effort to sequence the entire *Arabidopsis* genome. Single sequencing reactions are run from the 5' end of the cDNAs and the resulting sequences are referred to as Expressed Sequence Tags (ESTs) (Newman et al. 1994). Each EST is then used in a BLAST (Basic Local Alignment Search Tool) search of existing sequences to determine if it is similar in sequence to any known genes or protein sequences. The resulting information and sequence of each EST is stored in a database, which can be searched for genes of interest. One strength of this approach is that it is not biased by any presumption. The chance of a cDNA being sequenced is dependent only upon the level at which it is present in the cDNA libraries used.

During the characterisation of *AHB1*, the EST database was searched for the presence of an *AHB1* cDNA clone. An EST showing homology to plant hemoglobin genes was found in the *Arabidopsis* EST database. However, this EST does not correspond to *AHB1*. As this chapter describes, this hemoglobin gene, *AHB2*, is more closely related in sequence to the symbiotic hemoglobins of legumes and *Casuarina*.

Figure 16. Evolution of the symbiotic hemoglobins.



Two models of plant hemoglobin gene evolution. The rapid divergence model accounts for the high level of sequence divergence between the symbiotic (SYM) and non-symbiotic (NS) hemoglobin gene families through rapid gene divergence. Following the duplication of plant hemoglobin genes, one of the resulting gene families was recruited into symbiotic function leading to rapid sequence divergence as these hemoglobins became specialised symbiotic proteins. The alternative model is that a duplication event gave rise to two non-symbiotic hemoglobin gene families before the appearance of nodulation. One family evolved into the non-symbiotic hemoglobin gene family (NSI), while a second non-symbiotic family evolved into hemoglobins similar in sequence to the symbiotic hemoglobins (NSII). During the evolution of nodulation symbiotic hemoglobins (SYM) arose from a second gene duplication (of the NSII gene family). The earlier divergence model accounts for the large amount of sequence divergence between the non-symbiotic and the symbiotic hemoglobin gene families.

4.2 Results.

Cloning of *AHB2*.

A single EST showing homology to plant hemoglobins was detected during a key word search of the *Arabidopsis* EST database. This partial length cDNA (clone ID 11161) does not correspond to the *AHB1* gene, with only 69% identity at the nucleotide level. Indeed, this EST shows greater homology, as indicated by BLAST scores, to the symbiotic hemoglobin of *Casuarina* and the leghemoglobins than it does to the non-symbiotic hemoglobins.

The cDNA insert of clone 11161 was entirely sequenced, (Figure 18), and used to screen a λ EMBL genomic library (C24 ecotype). Two positive clones were isolated. One of these λB111 was partially restriction mapped (Figure 17) and several of the resulting fragments were subcloned and sequenced. λB111 was found to contain almost the entire sequence of *AHB2*, but was missing 100 bp of 3' coding sequence. The entire sequence of *AHB2* was constructed by including the 3' cDNA sequence from clone 11161. The 3' region missing from the genomic clone was also amplified from *Arabidopsis* genomic DNA using PCR. The resulting fragment was then cloned and sequenced to ensure that no introns occur in this region and that the 3' end of the genomic sequence corresponds to the 3' region of clone 11161.

Sequence analysis of *AHB2*.

The cDNA insert of clone 11161 is identical to the genomic sequence isolated but lacks two introns present in the genomic clone in positions identical to I2 and I3 of other plant hemoglobins. A third plant intron, I1,

is also predicted by homology with other plant hemoglobins. This putative intron is in the 5' region of the genomic sequence, a region that is absent from the partial length cDNA clone 11161 (Figure 18). Comparison of the genomic sequence with a full length cDNA clone shows that this intron is present in the genomic sequence but spliced out of the mRNA (see below). Thus, the *AHB2* genomic sequence is interrupted by the three introns found in positions equivalent to the three introns found in all plant hemoglobin genes. Interestingly, the central intron of *AHB2* is longer (760 bp) than the central intron of non-symbiotic hemoglobins (125 for *AHB1*), as is the central intron of the *Casuarina* symbiotic hemoglobin (570 bp). The central introns of the soybean leghemoglobin genes are shorter than those of the *AHB2* and *Casuarina* symbiotic hemoglobin (~225 bp) genes, but are still longer than the central introns of most non-symbiotic hemoglobin genes

The predicted amino sequence of *AHB2* shows a high degree of similarity to the symbiotic hemoglobins of legumes and *Casuarina* (Figure 19). It also shows significant, although weaker, similarity to the non-symbiotic hemoglobins. A protein similarity tree placed the predicted protein sequence of *AHB2* in a clade with the symbiotic hemoglobins, separated from the non-symbiotic hemoglobin clade (Figure 20). The N-terminal of the predicted AHB2 protein is intermediate in length to the N-terminals the symbiotic of non-symbiotic hemoglobins (Figure 19).

Figure 17. Partial restriction map of *AHB2* genomic sequence.

Partial restriction map of *AHB2* genomic sequence. Location of cDNA sequence is also shown in comparison to λB111 clone. Gaps in alignment of the cDNA correspond to two of the three introns in the *AHB2* gene. The translated sequence is shaded. Approximate fragment sizes are shown in kilobases.

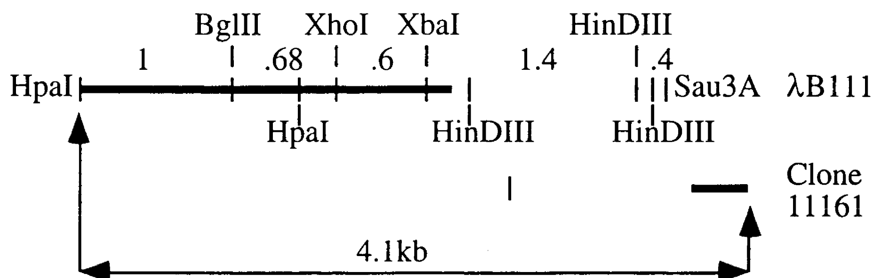


Figure 18. Sequence of *AHB2*.

Sequence of the coding region of the AHB2 gene.
Stop and start codons are shown in bold. Exon borders are underlined.
$ Denotes the start of the EST (Clone 11161), which (with exception of the absence of intron) is identical to the genomic sequence.

```
                                                  M  G  E
AATATTCGTGTTTTTTTCAAACTGTGAGAGAAAAAGAAAGAGAGAAAGAGATGGGAGAGA
---------+---------+---------+---------+---------+---------+10
TTATAAGCACAAAAAAAGTTTGACACTCTCTTTTTCTTTCTCTCTTTCTCTACCCTCTCT


I  G  F  T  E  K  Q  E  A  L  V  K  E  S  W  E  I  L  K  Q
TTGGGTTTACAGAGAAGCAAGAAGCTTTGGTGAAGGAATCGTGGGAGATACTGAAACAAG
---------+---------+---------+---------+---------+---------+ 70
AACCCAAATGTCTCTTCGTTCTTCGAAACCACTTCCTTAGCACCCTCTATGACTTTGTTC


D  I  P  K  Y  S  L  H  F  F  S  Q
ACATCCCCAAATACAGCCTTCACTTCTTCTCACAGTAACCATATATACTTAGTTATATAT
---------+---------+---------+---------+---------+---------+ 130
TGTAGGGGTTTATGTCGGAAGTGAAGAAGAGTGTCATTGGTATATATGAATCAATATATA

                                                         $
                                                          I
AAGCTCTTTACATGTTGTTTATATATGCGAGCTAATGAACAATATAATTGTGATAGGATA
---------+---------+---------+---------+---------+---------+ 190
TTCGAGAAATGTACAACAAATATATACGCTCGATTACTTGTTATATTAACACTATCCTAT


 L  E  I  A  P  A  A  K  G  L  F  S  F  L  R  D  S  D  E  V
CTGGAGATAGCACCAGCAGCAAAAGGCTTGTTCTCTTTCCTAAGAGACTCAGATGAAGTC
---------+---------+---------+---------+---------+---------+ 250
GACCTCTATCGTGGTCGTCGTTTTCCGAACAAGAGAAAGGATTCTCTGAGTCTACTTCAG


 P  H  N  N  P  K  L  K  A  H  A  V  K  V  F  K  M
CCTCACAACAATCCTAAACTCAAAGCTCATGCTGTTAAAGTCTTCAAGATGGTAATTACT
---------+---------+---------+---------+---------+---------+ 310
GGAGTGTTGTTAGGATTTGAGTTTCGAGTACGACAATTTCAGAAGTTCTACCATTAATGA


TACTTTCCGATTTTCCACATCTACATATATGTGAATCACTTGCATATACTGTATCATTAT
---------+---------+---------+---------+---------+---------+ 370
ATGAAAGGCTAAAAGGTGTAGATGTATATACACTTAGTGAACGTATATGACATAGTAATA


CTTACCATTCCTTAAAATTGAAAGTAGAATGTTTCATTATTTACAGCTAAGAATCTTTAT
---------+---------+---------+---------+---------+---------+ 430
GAATGGTAAGGAATTTTAACTTTCATCTTACAAAGTAATAAATGTCGATTCTTAGAAATA


TTACTCATTATATACCATTTATATATAATAGAAAATAGTAGTCTGAATTAACTTTTCTTG
---------+---------+---------+---------+---------+---------+ 490
AATGAGTAATATATGGTAAATATATATTATCTTTTATCATCAGACTTAATTGAAAAGAAC
```

```
TCATTTATTGACGCAGCCATGTGCACACAAGTTGCGATTTTGTTTGAACTTGGCTAGTTG
---------+---------+---------+---------+---------+---------+ 550
AGTAAATAACTGCGTCGGTACACGTGTGTTCAACGCTAAAACAAACTTGAACCGATCAAC

GCTTTGTCTTCTTCTTTGAGAATAAAAATCTCATACTAGTAAAGAATACTCTGTGATATT
---------+---------+---------+---------+---------+---------+ 610
CGAAACAGAAGAAGAAACTCTTATTTTTAGAGTATGATCATTTCTTATGAGACACTATAA

TTATTTTTAAGAACAAACATAGATTTCTCTGTCAATAAAGAATTGTTACTGAAGAATCCA
---------+---------+---------+---------+---------+---------+ 670
AATAAAAATTCTTGTTTGTATCTAAAGAGACAGTTATTTCTTAACAATGACTTCTTAGGT

AGTGGTTCGGGTCGCTTTATGGATTTTTACTTTTTGCTAATCTTATTATAATAGAACCA
---------+---------+---------+---------+---------+---------+ 730
TCACCAAGCCCAGCGAAATACCTAAAAATGAAAAACGATTAGAATAATATTATCTTGGT

TATAAACCAAATTCCGTTTACTTTTTAAATTTGGGTTTATGACTTGGTTTGGTTCAACTC
---------+---------+---------+---------+---------+---------+ 790
ATATTTGGTTTAAGGCAAATGAAAAATTTAAACCCAAATACTGAACCAAACCAAGTTGAG

ACTTTTGGCTTCTAAGACTTTGCATAACATGTTTAGACAGACAAAAAAGAAAAAGACTT
---------+---------+---------+---------+---------+---------+ 850
TGAAAACCGAAGATTCTGAAACGTATTGTACAAAATCTGTCTGTTTTTTCTTTTTCTGAA

                                                T  C  E  T  A  I
GCATAACATGTATGAATTTTTATTTTATTTTGTTTGTGTGTAGACATGTGAAACAGCTAT
---------+---------+---------+---------+---------+---------+ 910
CGTATTGTACATACTTAAAAATAAAATAAAACAAACACACATCTGTACACTTTGTCGATA

  Q  L  R  E  E  G  K  V  V  V  A  D  T  T  L  Q  Y  L  G  S
ACAGCTGAGGGAGGAAGGAAAGGTGGTAGTGGCTGACACAACCCTCCAATATTTAGGCTC
---------+---------+---------+---------+---------+---------+ 970
TGTCGACTCCCTCCTTCCTTTCCACCATCACCGACTGTGTTGGGAGGTTATAAATCCGAG

  I  H  L  K  S  G  V  I  D  P  H  F
AATTCATCTCAAAAGCGGCGTTATTGACCCTCACTTCGAGGTCTGTTATGTTAAAAAAAA
---------+---------+---------+---------+---------+---------+ 1030
TTAAGTAGAGTTTTCGCCGCAATAACTGGGAGTGAAGCTCCAGACAATACAATTTTTTTT

ATATATATACACATTAATTTTGGCTGATTTTGATTTTCGATTTGAACGCATTTTAATAAG
---------+---------+---------+---------+---------+---------+ 1090
TATATATATGTGTAATTAAAACCGACTAAAACTAAAAGCTAAACTTGCGTAAAATTATTC

                  E  V  V  K  E  A  L  L  R  T  L  K  E  G  L
GTGTGAATGTGAAAGCAGGTGGTGAAAGAAGCTTTGCTAAGGACATTGAAAGAGGGGTTG
---------+---------+---------+---------+---------+---------+ 1150
CACACTTACACTTTCGTCCACCACTTTCTTCGAAACGATTCCTGTAACTTTCTCCCCAAC

  G  E  K  Y  N  E  E  V  E  G  A  W  S  Q  A  Y  D  H  L  A
GGGGAGAAATACAATGAAGAAGTGGAAGGTGCTTGGTCTCAAGCTTATGATCACTTGGCT
---------+---------+---------+---------+---------+---------+ 1210
CCCCTCTTTATGTTACTTCTTCACCTTCCACGAACCAGAGTTCGAATACTAGTGAACCGA
```

```
 L  A  I  K  T  E  M  K  Q  E  E  S  *
TTAGCCATCAAGACCGAGATGAAACAAGAAGAGTCATAAAACCCTATTGATCATTGGGTA
---------+---------+---------+---------+---------+---------+ 1270
AATCGGTAGTTCTGGCTCTACTTTGTTCTTCTCAGTATTTTGGGATAACTAGTAACCCAT

TCGCATACATGAATCTATTCCACAT
---------+---------+-----1295
AGCGTATGTACTTAGATAAGGTGTA
```

Figure 19. Alignment of plant hemoglobins.

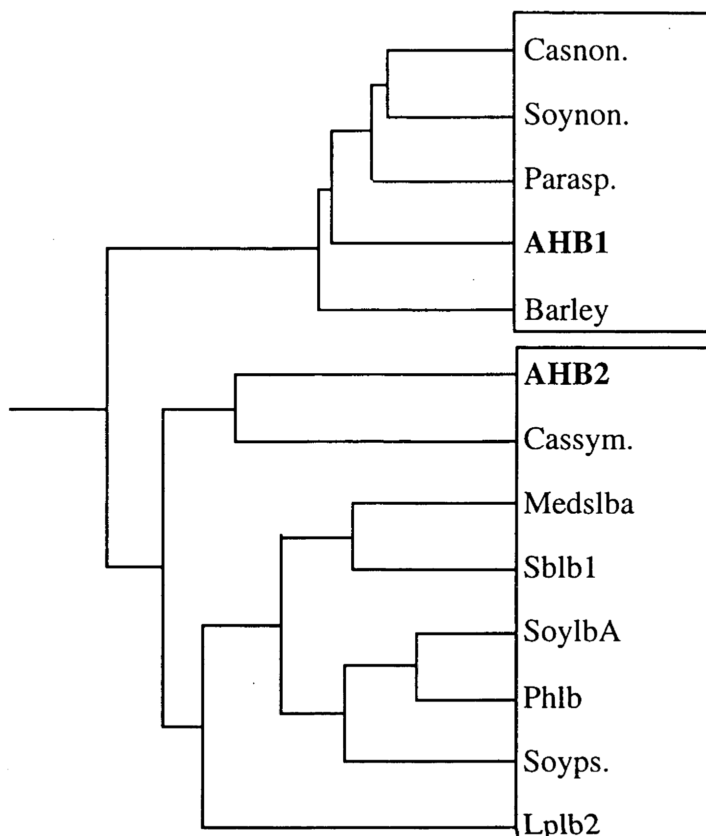
```
    1                                              *-----C2      CD1      60
Bar .SAAEGAVVF SEEKEALVLK SWAIMKKDSA NLGLRFFLKI FEIAPSARQM FPFLRDSDVP
CN  .SSTLEGRGF TEEQEALVVK SWSAMKPNAG ELGLKFFLKI FEIAPSAQKL FSFLKDSNVP
Soy .TTTLE-RGF SEEQEALVVK SWNVMKKNSG ELGLKFFLKI FEIAPSAQKL FSFLRDSTVP
Par .SSSEVNKVF TEEQEALVVK AWAVMKKNSA ELGLQFFLKI FEIAPSAKNL FSYLKDSPVP
HB1 ..ESEGKIVF TEEQEALVVK SWSVMKKNSA ELGLKLFIKI FEIAPTTKKM FSFLRDSPIP
HB2 .....GEIGF TEKQEALVKE SWEILKQDIP KYSLHFFSQI LEIAPAAKGL FSFLRDSDEV
CS  ........AL TEKQEALLKQ SWEVLKQNIP AHSLRLFALI IEAAPESKYV FSFLKDSNEI
lba .......GAF TEKQEALVSS SFEAFKANIP QYSVVFYTSI LEKAPAAKDL FSFL--ANGV

   61        E7      *                                F8          * 120
Bar LETNPKLKTH AVSVFVMTCE AAAQLRKAGK ITVRETTLKR LGGTHLKYGV ADGHFEVTRF
CN  LERNPKLKSH AMSVFLMTCE SAVQLRKAGK VTVRESSLKK LGASHFKHGV ADEHFEVTKF
Soy LEQNPKLKPH AVSVFVMTCD SAVQLRKAGK VTVRESNLKK LGATHFRTGV ANEHFEVTKF
Par LEQNPKLKPH ATTVFVMTCE SAVQLRKAGK VTVKESDLKR IGAIHFKTGV VNEHFEVTRF
HB1 AEQNPKLKPH AMSVFVMCCE SAVQLRKTGK VTVRETTLKR LGASHSKYGV VDEHFEVAKY
HB2 PHNNPKLKAH AVKVFKMTCE TAIQLREEGK VVVADTTLQY LGSIHLKSGV IDPHFEVVKE
CS  PENNPKLKAH AAVIFKTICE SATELRQKGH AVWDNNTLKR LGSIHLKNKI TDPHFEVMKG
lba DPTNPKLTGH AEKLFALVRD SAGQLKASGT VV-AD-A--A LGSVHAQKAV TDPEFVVVKE

   121-------- -                                   163
Bar ALLETIKEAL PADMWGPEMR NAWGEAYDQL VAAIKQEMKP AE.
CN  ALLETIKEAV P-ETWSPEMK NAWGEAYDKL VAAIKLEMKP SS.
Soy ALLETIKEAV P-EMWSPAMK NAWGEAYDQL VDAIKSEMKP PSS
Par ALLETIKEAV P-EMWSPEMK NAWGVAYDQL VAAIKFEMKP SST
HB1 ALLETIKEAV P-EMWSPEMK VAWGQAYDHL VAAIKAEMNL SN.
HB2 ALLRTLKEGL G-EKYNEEVE GAWSQAYDHL ALAIKTEMKQ EES
CS  ALLGTIKEAI K-ENWSDEMG QAWTEAYNQL VATIKAEMK. ...
lba ALLKTIKAAV G-DKWSDELA RAWEVAYDEL AAAIKK.... ...
```

Alignment includes: Bar: barley hemoglobin (U01228) (Taylor et al. 1994), CN: *Casuarina* non-symbiotic (X53950) (Christensen et al. 1991), Soy: soybean non-symbiotic (U47143) (Andersson et al. 1996), Par: *Parasponia* hemoglobin (U27194) (Landsmann et al. 1986). HB1: AHB1 (U94998). HB2: AHB2 (U94999). CS: *Casuarina* symbiotic (X77695) (Jacobsen-Lyon et al. 1995). LbA: soybean leghemoglobin A (X14311) (Hyldig-Nielsen et al. 1982). Residues strongly conserved in all hemoglobins are shown in bold. Intron positions are shown with stars. Dashed line indicates sites targeted by the PCR primers used to amplify *AHB1*. Sequences were aligned with the pileup function of the GCG package.

Figure 20. Protein similarity tree.



Protein sequence similarity tree, constructed with the pileup function of the GCG package. Hemoglobins, with genbank accesion numbers. Barley: barley non-symbiotic hemoglobin (U01228). Casnon: *Casuarina* non-symbiotic hemoglobin (X53950). Soynon: soybean non-symbiotic hemoglobin (U47143). Parasp: *Parasponia* hemoglobin (U27194). AHB1: This study, (U94998). AHB2: This study, (U94999) Cassym: *Casuarina* symbiotic hemoglobin (X77695). Soylba: soybean leghemoglobin A (J01299). Medslba: *Medico sativa* leghemoglobin A (X14311). Luplb2: lupin leghemoglobin 1 (Y00401). Sblb1: *Sesbania rostratum* leghemoglobin 1 (J01299). Phlb1: *Phaseolus* leghemoglobin 1 (K03152). Soyps: soybean leghemoglobin pseudogene (V00451). Luplb2: lupin leghemoglobin 2 (M17893).

Figure 21. Sequence from the 5' of *AHB2*.

440 bp of the *AHB2* promoter, 1.5 kb of which has been sequenced. A possible TATA box is found at -79 but no other putative promoter elements have been identified.

-439 GCTGATTTTG TGTGAAAGTA GAAAAGAACA AATGATTGAA GCTATGCGGG

-389 ATGGAGATTT TATACGCAGA AGACAATGCA AGTTTTTATT TATCTTTGTT

-339 TGTTCCTTTT AATCTTAACT TATTTGTATC AATCTAACTC ATGTATTATC

-289 TACGTCTATC TAGACTGATC TCTGATCGGA CGGACATAAT GTGTATCCAT

-239 ATTTCTTGGC TACGCGTGTC CACCTTTTAG AGACTATGCC TTTAGGTAGT

-189 AGATGTTTTA CTACAAAATA AACATATTTA GTCAAATAAA ATAAAATTTG

-139 AGAGAATCTT CTACAAAGTA GCTCACAGAC CCAACCAAAG GACCATTGAA

-89 TACCA**TATAT** **ATATA**GATAC ACAGACATAT AAACACACAA ATATTCGTGT

-39 TTTTTTCAAA CTGTGAGAGA AAAAGAAAGA GAGAAAGAG**A** **TG**

Sequences that may be involved in translational and transcriptional regulation have been identified. Putative translational start and stop codons are found at bases 1 and 1250, respectively (Figure 18). A possible TATA box is found at -74 in the 5' sequence of the *AHB2* (Figure 21). Unlike *AHB1*, there are no 5' sequences with homology to the regulatory sequences of nodulins or other plant hemoglobins. Indeed, the promoter of *AHB2* shows little homology to the promoters of any other plant genes, other than the homologous gene from *Brassica napus* (Richard Watts pers. comm).

Southern analysis of *AHB2*.

An *AHB2* cDNA probe was used to probe blots of C24 genomic DNA cut with a range of enzymes. At high stringency, this probe detected one or two band sizes in all lanes, consistent with a single copy gene (Figure 22). No additional bands were detected under low stringency conditions. It seems that *AHB2*, like *AHB1*, is a single copy gene in the *Arabidopsis* genome. Furthermore, the failure to detect additional bands on low stringency southern blots suggests that there are no other genes with significant homology to *AHB2* in the *Arabidopsis* genome. This result also implies that *AHB1* and *AHB2* are unable to cross hybridise, even under low stringency conditions.

Probing blots of genomic DNA from other plants (Richard Watts pers. comm) shows that *AHB2* like sequences may be present in the genome of *Brassicas napus** and possibly pumpkin, but no cross hybridisation with DNA from cotton, soybean, rice, *Casuarina* or *Parasponia* was observed.

Analysis of *AHB2* expression.

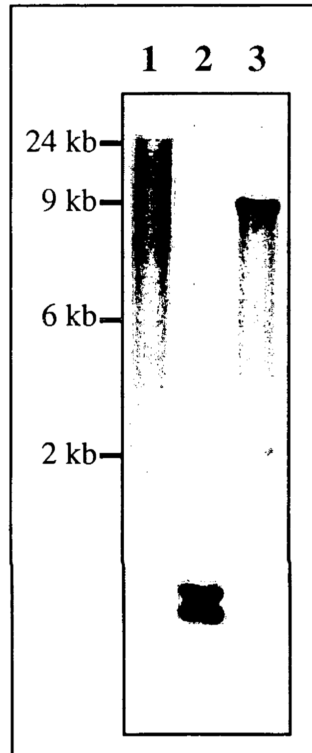
The expression of *AHB2* was examined on RNA gel blots. A low level of transcript was detectable in the rosette leaves, but there was no detectable expression in root tissue (Figure 23). The effects of a range of environmental stress conditions on *AHB2* transcript levels were also examined. Wounding, dehydration, heatshock and oxidative stress had no effect on *AHB2* expression. Furthermore, unlike *AHB1*, no increase in transcript levels in response to hypoxia or sucrose were observed (Figure 24). Chilling treatment (4° for 24 hrs) does increase *AHB2* transcript

---

* AHB2-like gene has recently been cloned from Brassica (Richard Watts unpublished results)

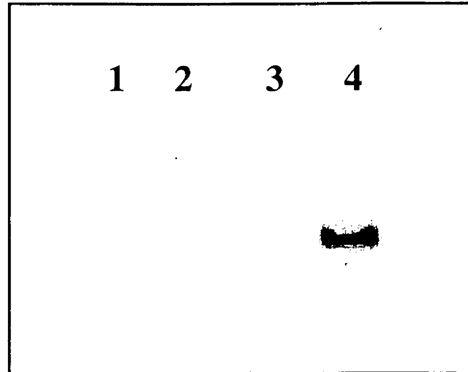
levels, and this effect seems to be limited to above ground tissue (Figure 23). The level of *AHB2* gene expression, even after chilling treatment, is low when compared to *AHB1* gene expression.

Figure 22. Genomic DNA blot probed with *AHB2*.



Arabidopsis (C24 ecotype) genomic DNA digested with BamH1, Dra1, EcoR1 (lanes 1, 2, and 3), hybridised with *AHB2* cDNA probe.

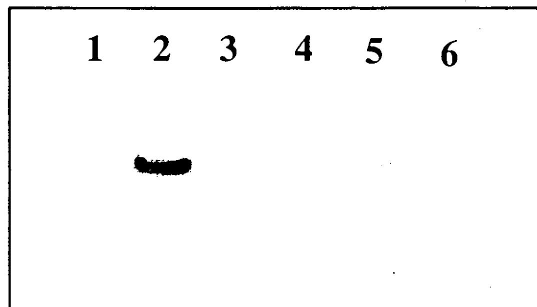
Figure 23. Induction of *AHB2* by low temperature.



Total RNA from root tissue (lane 1), rosette leaves (lane 2), root tissue after chilling treatment (lane 3), rosette leaves after chilling treatment (lane 4) hybridised with an *AHB2* antisense ribo-probe.

Figure 24. *AHB2* expression during environmental stress.

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) hybridised with an *AHB2* antisense ribo-probe.



Cloning of a *AHB2* cDNA.

To enable AHB2 protein to be expressed in bacteria, and to generate plants that over or under express AHB2 protein, a full length cDNA clone was required. The EST used to clone *AHB2* was lacking the sequence encoding the N-terminal region of the AHB2 protein. A cDNA clone spanning the entire coding sequence was isolated using the same RT-PCR approach that was used to isolate the *AHB1* cDNA clone. Specific primers from the start and stop codon regions of *AHB2* were used to amplify and clone an *AHB2* cDNA that contains the entire coding sequence. Comparison of the sequence of this clone with the *AHB2* genomic sequence confirms that the intron (I1) predicted on the basis of homology with other plant hemoglobin genes is spliced out of the *AHB2* transcript (as well as the I2 and I3 introns).

4.3 Discussion.

AHB2 is clearly a gene which encodes a hemoglobin. The predicted amino acid sequence of the AHB2 protein shows a high level of amino acid sequence similarity to other plant hemoglobins and all the amino acid residues that are conserved in hemoglobins are present in the predicted protein sequence of *AHB2*. Furthermore, the *AHB2* gene is interrupted by three introns, found in identical positions to those found in all other plant hemoglobin genes.

The *AHB2* gene sequence shows more homology to the symbiotic hemoglobins of legumes and *Casuarina* than it does to the non-symbiotic hemoglobin genes. This explains why *AHB1* probes failed to detect *AHB2* when hybridised with *Arabidopsis* genomic DNA at low stringency. These genes share only a moderate level of nucleotide sequence identity (66%) and do not cross hybridise even at reduced stringency. It also explains why *AHB2* was not amplified by the PCR protocol used to clone *AHB1*. The PCR primers used to isolate *AHB1* were designed to preferentially amplify non-symbiotic hemoglobins from the genome of soybean. The soybean genome also contains symbiotic hemoglobin genes so the PCR primers used to amplify the soybean non-symbiotic hemoglobin gene were deliberately biased against the amplification of sequences with homology to the symbiotic hemoglobins (Andersson et al. 1996). Indeed, the sequence of *AHB2* in the regions covered by the PCR primers does differ from the equivalent regions of non-symbiotic hemoglobin genes (see Figure 19).

AHB2 -like sequences may be widely distributed in plants. An *AHB2* probe detects homologous sequences in *Brassica napus* and a *AHB2* homologue has recently been cloned from *Brassica*, confirming this

observation (Richard Watts pers. comm). If a *AHB2* -like family of non-symbiotic hemoglobin genes is widely distributed in plants, a *AHB2* -like gene would be the obvious candidate for the ancestor of symbiotic hemoglobins. The similarity between the predicted amino acid sequence of *AHB2* and the amino acid sequences of the symbiotic hemoglobins of *Casuarina* and legumes strongly suggests that these hemoglobins share a common ancestor.

Arabidopsis is distantly related to nitrogen-fixing plants (Chase et al. 1993), and does not enter into any symbiotic nitrogen-fixing associations itself. It is highly unlikely, therefore, that the *AHB2* functions in nitrogen-fixing symbiosis. Instead, the expression of *AHB2* in above ground tissues suggests that *AHB2* has a function in normal plant tissues. Thus, *AHB2* represents a new class of plant hemoglobin. A non-symbiotic hemoglobin by function, that is related in sequence to the symbiotic hemoglobin family.

The requirement for a second hemoglobin in above ground tissues is unclear. The low level of expression of *AHB2* does not seem consistent with an oxygen transport function. Unless confined to a small number of cells, such a low level of expression is unlikely to lead to levels of protein high enough to increase the rates of oxygen diffusion into plant cells. It is possible that the AHB2 protein is required to provide bound oxygen for an enzymatic reaction. Such a function has been suggested for the hemoglobin of the nematode, *Ascasris suum* which is thought to bind oxygen and squalene, and may be involved in epoxidation of this steroid hormone precursor (Sherman et al. 1992). Another possible function for AHB2 is in sequestering oxygen in order to prevent free oxygen interfering with an oxygen sensitive reaction, but this may require higher levels of gene expression than observed for *AHB2*. AHB2 may be an oxygen sensing protein, but this seems unlikely as oxygen sensing would be

expected to be most important in below ground organs, where waterlogging is likely to result in hypoxic conditions, and *AHB2* expression has not been detected in root tissue.

The induction of *AHB2* following chilling treatment is also difficult to explain. It is possible that chilling creates a demand for oxygen in some areas of plant metabolism, such as an oxygen utilising enzymatic process. Alternatively, the increase in transcript levels may be due to an increase in transcript stability rather than an actual increase in the rate of *AHB2* transcription, and may have no functional significance at all. Increased mRNA stability following chilling has been shown to occur for some plant genes (Adamski and Kloppstech 1994). Comparing the levels of AHB2 protein before and after chilling treatment, using western blot analysis, may help determine if the increase in transcript levels found after chilling treatment results in increased levels of protein.

A better understanding of AHB2 function may come from detailed biochemical analysis. Expression of recombinant AHB2 in *E. coli* could provide a source of protein for oxygen binding studies. Accurate localisation of AHB2, to cell types and to subcellular compartments may also help elucidate the function of this hemoglobin. If AHB2 was found to be associated with a particular subcellular compartment of specialised tissue for instance, a specific function may be implicated. Finally, as discussed for AHB1, analysis of plants that over or under express AHB2 may also help understand the requirement for this second class of non-symbiotic hemoglobin.

CHAPTER 5.

Control of *AHB1* gene expression.

5.1 Introduction.

Flooded plants can encounter extreme oxygen shortages. In warm conditions waterlogged soil can become almost totally anaerobic within 24 hours of flooding (Drew 1997). Consequently plants have evolved a number of strategies to survive hypoxia and flooding tolerant plants can survive and grow in complete submergence for long periods of time (Kennedy et al. 1992). Some adaptations to flooding are long term physical alterations, such as the production of aerenchyma tissue which allows air to diffuse from aerial tissues to submerged tissues (see Drew 1990, Drew 1997). Other changes are at the level of cellular metabolism. These changes are relatively rapid, occurring within hours, and may increase the chances of survival when plants encounter short periods of hypoxia. The major metabolic response is a shift towards alcoholic fermentation which, although less efficient than respiration, does not require oxygen and can continue to provide ATP during short periods of severe hypoxia (reviewed in Drew 1997).

The switch in metabolism of hypoxia treated plants is mediated through rapid alterations in gene expression. The synthesis of "normal proteins" ceases and a set of proteins known as the "Anaerobic polypeptides" (ANPs) is selectively produced. Experiments where plants have been exposed to a mild hypoxia pretreatment, inducing synthesis of ANPs, and then exposed to severe hypoxia have shown that expression of ANPs is associated with an increase in hypoxia tolerance (see Drew 1997). Many ANPs have been identified as enzymes involved in glycolytic

fermentation (Sachs et al. 1980). Pyruvate decarboxylase (PDC) and alcohol dehydrogenase (ADH) are the classic examples of such enzymes. PDC and ADH act sequentially to catalyse the conversion of pyruvate, produced from glycolysis, to acetaldehyde and then ethanol. This conversion removes the end product of glycolysis and produces NAD^+ from NADH, which is essential for the maintenance of glycolysis.

During hypoxia the levels of the ADH and PDC proteins increase dramatically. This has been observed in a range of plant species (see Sachs et al. 1996, Dolferus et al. 1997). The increase in protein levels observed are due in part to greatly increased levels of ADH and PDC transcripts. Transcription of *ADH* in root tissue is strongly induced by hypoxia (Gerlach et al. 1982) and promoter elements required for hypoxia induction, Anaerobic Response Elements (AREs), have been identified in the promoter of the maize *ADH1* gene (Walker et al. 1987). Similar promoter motifs have also been found in the promoters of other ANP genes (see Drew 1997).

Hypoxia is also known to have a strong effect on the translation of proteins. When maize seedlings are subjected to hypoxia, translation of normal proteins ceases and the ANPs are the main proteins produced. However, if mRNA is extracted from hypoxia treated maize, and then translated *in vitro* many more proteins are synthesised (Gerlach et al. 1982, Fennoy and Bailey-Serres 1995). Thus, it appears that only a small proportion of mRNAs present are translated in hypoxic maize cells. This "selective" translation seems to be connected with the association of ribosomes with transcripts and may involve translational initiation factors (Vayda et al. 1995). Regulation of translation may be a mechanism to ensure that energy is not wasted in the production of proteins that are of no benefit during hypoxic stress.

The hypoxia response exhibited by *AHB1* and the barley hemoglobin are remarkably similar to that of *ADH*. The sequence of the promoter of *AHB1* is available, but the sequence of the barley hemoglobin promoter is not known and thus comparison of the promoter sequences from these two genes is not possible. Two putative ARE motifs occur within the *AHB1* 5' region (figure 12a/12b) and these may be involved in the hypoxia response of *AHB1*. To determine if the promoter of *AHB1* is able to confer anaerobic response, 1 kb of sequence from the 5' region of *AHB1* was fused to the B-glucuronidase reporter gene (*uidA*) (Jefferson 1987), and the resulting construct was transformed into *Arabidopsis*. This also allows the localisation of *AHB1* expression to be observed and compared with other plant hemoglobin or hypoxia inducible genes.

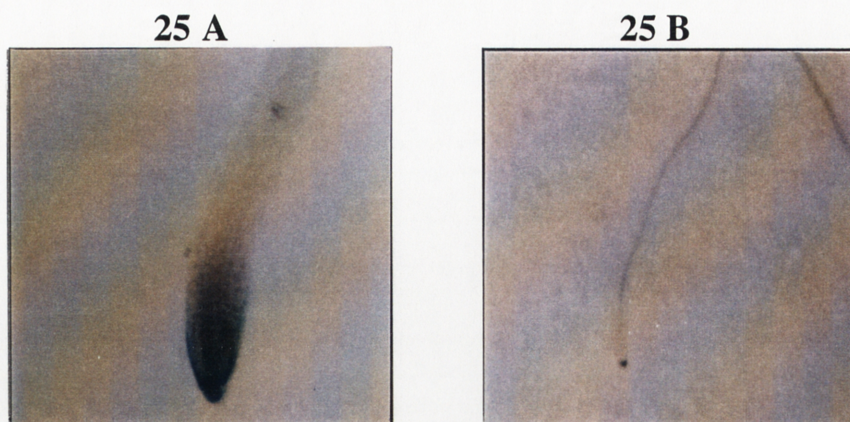
5.2 Results.

5'*AHB1*-Gus Fusion.

A 1 kb fragment (Hind/DraI -1090 to -25) from the *AHB1* 5' region was fused to GUS (β-Glucuronidase, *uidA* gene) in the commercially available pBI 101.1 vector. The resulting construct, GH1 was transferred into *Agrobacterium* and transformed into plants. Of the dozen independent GUS expressing transgenic lines generated six were analysed. These six lines displayed identical patterns of GUS expression, and one (GH1-4) is used as an example in the figures of this thesis.

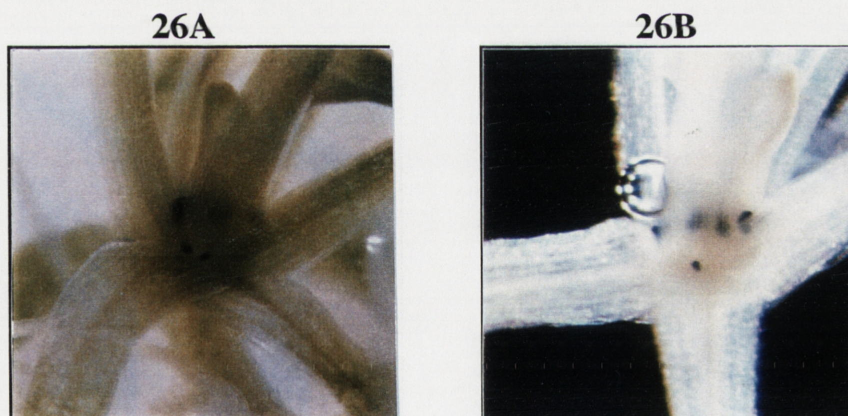
In untreated GH1-4 plants GUS staining was detected in the root tips (Figure 25a), around the root cap and root meristem tissue. Staining was also detected in the lateral root primordia and in the vascular tissue of roots in some plants (Figure 25b). In the above ground tissues, expression of GUS was detected in the hydatodes (Figure 26). Expression was also found in the petioles and vascular tissue of senescing leaves in some plants.

Figure 25. GUS staining in the roots of GH1-4 plants.



A. GUS staining in the roots of GH1-4 plants (normal growth conditions, 6x magnification). Staining occurs around root tips, lateral root primordia.
B. GUS staining also occurs in the vascular tissue (2.5x).

Figure 26. GUS staining in the above ground tissues of GH1-4 plants.

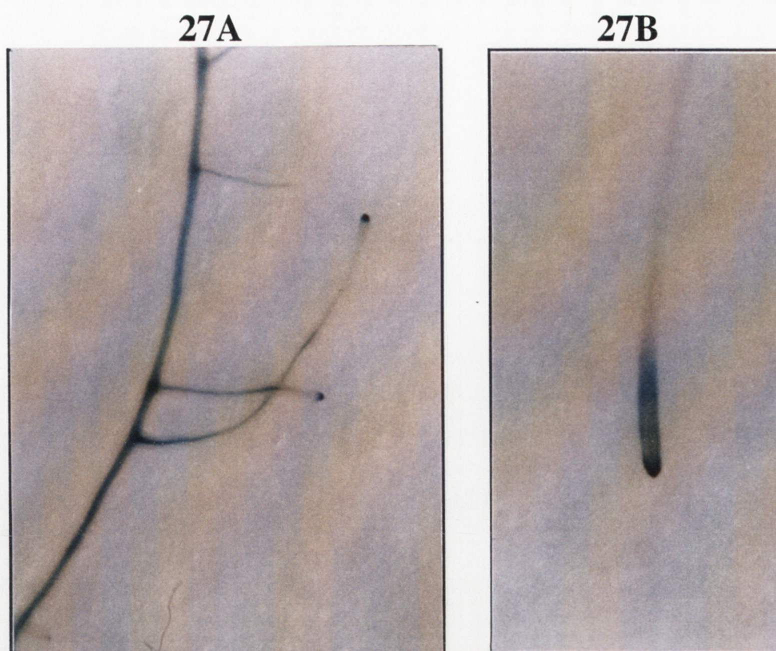


A. GUS staining in rosettes of GH1-4 plants. Staining occurs in the hydatodes (2.5x).
B. Same as 26A, but tissue has been cleared with 70% ethanol overnight.

GUS activity is sucrose inducible in GH1-4 plants.

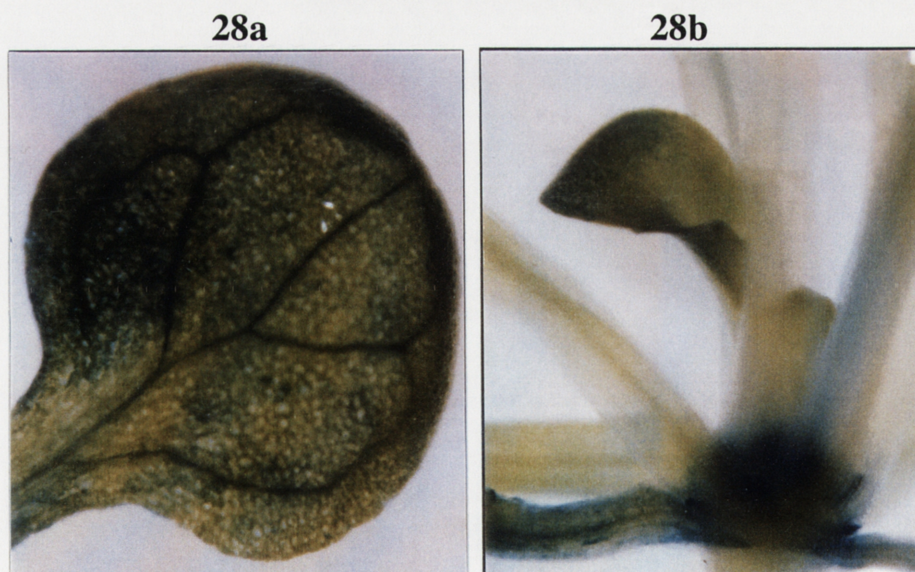
GUS activity was examined in plants that had been incubated in 1% sucrose for 24 hrs in the dark. Strong staining was found in the vascular tissue and tips of roots (Figure 27 a & b), in the center of rosettes (Figure 28 a) and in the vascular tissue of some leaves (Figure 28 b). Staining in root tips extended past the meristem area into vascular tissue (Figure 27 b), unlike the staining in the control plants which was confined to root tips in the majority of plants. The degree of staining in lateral root primordia and root vascular tissue was also stronger than control plants, and the vascular tissue of the older leaves showed staining in a majority of plants that had been treated with sucrose. Thus, the 1 kb promoter region from the *AHB1* gene has conferred sucrose induction to the GUS reporter gene.

Figure 27. GUS staining in roots of sucrose treated GH1-4 plants.



A. GUS staining in root tissue of GH1-4 plants treated with sucrose (2.5x).
B. GUS staining in root tips of sucrose treated GH1-4 plants (2.5x).

Figure 28. GUS staining in above ground tissue of GH1-4 plants treated with sucrose.



A. Staining in the vascular tissue of rosette leaves of GH1-4 plants that have been treated with sucrose (2.5x).
B. Staining in the rosettes of sucrose treated GH1-4 plants (2.5x).

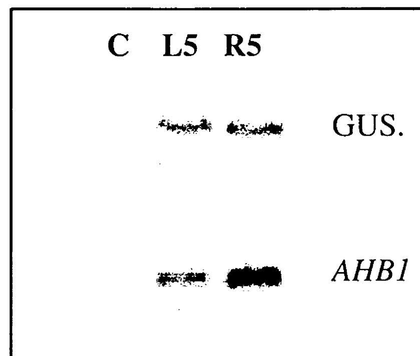
Hypoxia does not induce GUS activity in GH1 transformed plants.

In order to localise AHB1 expression during hypoxia, plants that had been treated with 0.1%, 5.0% or normal oxygen (~20% O_2) atmospheres were examined for GUS activity. No increase in GUS expression, as indicated by staining intensity, was detected when plants treated with 5% or 0.1% oxygen treatments were compared to the control plants. The pattern and degree of staining in hypoxia treated plants was identical to that observed in control plants. This is not due to a decrease in GUS activity in hypoxia treated plants, as positive control plants, transformed with *ADH* promoter-GUS fusions, have hypoxia inducible GUS activity (data not shown). Furthermore, the GH1 construct appears to produce functional GUS protein under normal conditions and the lack of increase in staining following hypoxia treatment is unlikely to be due to a defect in the GH1 construct. Therefore, it seems possible that, despite the presence of two putative AREs in the *AHB1* promoter, the sequences required for hypoxia response may be absent from the 1kb promoter fragment used in the GH1 construct.

GUS transcription is induced by hypoxia in GH1-4 plants.

To determine if the 1kb of promoter present in the GH1 construct is able to confer hypoxia induced transcription to the GUS reporter gene, the steady state levels of GUS transcript were examined in GH1-4 plants that had been treated with normoxic or 5.0% oxygen atmospheres. RNA gel blots show that the levels of GUS transcript in GH1-4 plants increase in response to hypoxia and are found at a similar levels to the *AHB1* transcript (Figure 29). Thus, despite the lack of response in GUS activity, the promoter of *AHB1* does confer transcriptional response to hypoxia.

Figure 29. Hypoxic induction of GUS transcript in GH1-4 plants.



Comparison RNA gel blots of separated RNA from whole GH1-4 plants from a normal atmosphere (C), leaves treated with 5% oxygen (L5) and roots treated with 5% oxygen (R5) probed with GUS and *AHB1* antisense riboprobes.

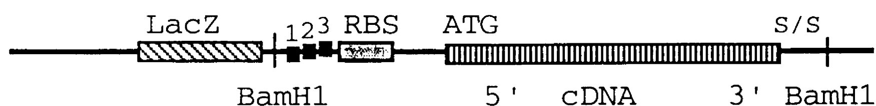
It is possible that during hypoxia, because of the translational regulation of gene expression, increased transcription may not lead to increased levels of protein. A hypoxia induced transcript may not be translated in hypoxic plant cells. This could explain the lack of increase in GUS staining in hypoxia treated GH1-4 plants. This also raises the possibility that *AHB1* transcripts are not translated in hypoxic cells. The 2-dimensional protein gels that have been used to visualise ANPs in *Arabidopsis* did not include the size range of AHB1 and therefore do not resolve this question (Dolferus et al 1985). In order to compare the levels of AHB1 protein in plants grown in normal conditions and plants subjected to hypoxia, antibodies specific for the AHB1 protein were required. AHB1 or AHB2 proteins had not been purified so no antibodies against these proteins were available. Poly-clonal sera raised against the *Parasponia* hemoglobin (anti-Phb) (Trinick et al. 1989) and the *Casuarina* symbiotic hemoglobin (anti-Cassym) (Fleming et al. 1987) were available.

While the anti-Cassym anitserum was known to cross react with soybean leghemoglobin (Fleming et al. 1987), whether or not either of these anti-sera would cross react with AHB1 or AHB2 was not known. To determine the affinity of existing anti-sera for AHB1 and AHB2, and to allow the production of poly clonal anti-sera against each of these proteins, heterologous expression of AHB1 and AHB2 cDNA clones, in *E. coli*, was attempted.

Expression of recombinant AHB1 and AHB2 in *E. coli*.

Sperm whale myoglobin has been expressed at high levels in *E coli* by fusing a cDNA sequence to the strong ribosomal binding site of the *Pseudomanas putida* cytochrome P-450 gene and driving expression from the LacZ promoter (Springer and Sligar 1987). Soybean leghemoglobin A has also been expressed in *E. coli*, using a stronger promoter, and when purified the recombinant soybean LbA was found to be indistinguishable from native soybean leghemoglobin (Hargrove et al. 1997). It seems that *E. coli* can be used for the heterologous expression of plant hemoglobins.

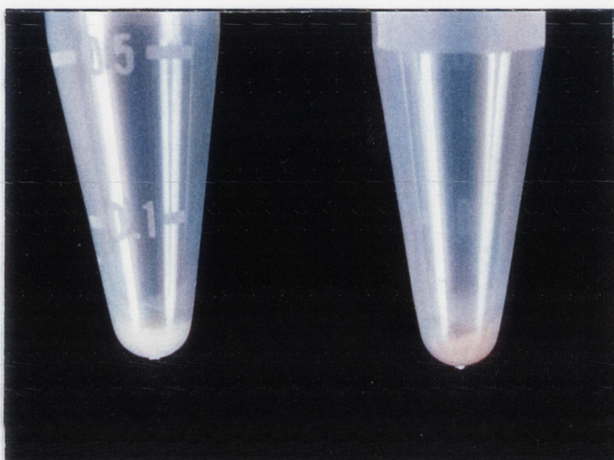
Figure 30. Bacterial expression construct.



Map of the bacterial expression construct used to express *Arabidopsis* hemoglobin genes in *E. coli*. RT-PCR was used to fuse stop codons in all three reading frames (1-3), a strong ribosomal binding site (RBS) and spacer sequence to the 5' of *AHB1* and *AHB2* cDNAs. Two in frame stop codons were also fused to the 3' ends (S/S) and BamH1 sites were included to facilitate cloning into the pUC119 plasmid.

The ribosomal binding site used by Springer and Sligar (Springer and Sligar 1987) was fused to the 5' of AHB1 and AHB2 cDNA sequences by including it in 5' primer sequences used to amplify from first strand cDNA template (Figure 30). The resulting fragments were cloned and then sequenced to ensure that no amplification errors had occurred during RT-PCR. Cell strains containing the resulting expression vectors had a reddish color (Figure 31) and had whole cell carbon monoxide (CO) difference spectra typical of cells containing high levels of hemoglobin (data not shown).

Figure 31. Recombinant hemoglobin expression.

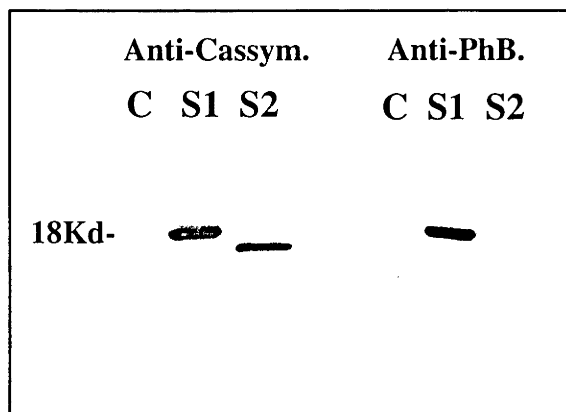


Comparison of pelleted cell cultures from a control strain containing pUC119 (left), and a strain containing the AHB1 expression construct (right).

Cross reactivity of anti-Phb and anti-Cassym antisera with recombinant AHB1 and AHB2.

Cytoplasmic proteins were extracted from bacterial strains expressing recombinant AHB1 and AHB2, and from a control strain containing pUC119 with an antisense *AHB1* cDNA insert. These protein extracts were separated on denaturing polyacrylamide gels and blotted onto nitrocellulose for western analysis. Duplicate blots were incubated with dilutions of the anti-Phb or anti-Cassym antisera. Another blot was incubated with the secondary antibody alone as a control. As shown (Figure 32), the anti-Phb antiserum cross reacts with a band of the expected size in strains containing the AHB1 expression vector but not the control strain containing pUC119. The anti-Phb antiserum does not cross react with recombinant AHB2. The anti-Cassym antiserum cross reacts with the appropriate size band in the AHB1 expressing strain and also appears to bind to recombinant AHB2. The apparent size of the recombinant AHB2 protein is slightly smaller than expected. This appears to be due to only partial denaturation of the AHB2 protein prior to electrophoresis. When the cell extracts are boiled for longer than 10 minutes the apparent size of the protein shifts to the predicted size (Richard Watts pers. comm). The affinity of the anti-Phb antiserum for AHB1, and the observation that it is specific for AHB1 and not AHB2, suggests that it is suitable for use in western analysis with *Arabidopsis* tissues. The reactivity of the anti-Cassym antiserum with both *Arabidopsis* hemoglobins would make it difficult to specifically detect AHB2 using this antiserum.

Figure 32. Cross reactivity of anti-hemoglobin antisera with recombinant *Arabidopsis* hemoglobins.

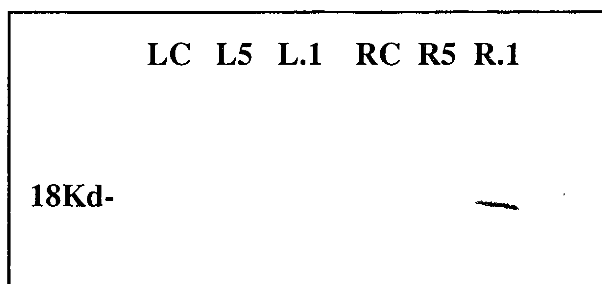


Bacterial protein extracts from a control strain (C) containing an *AHB1* antisense insert, an AHB1 expressing strain (S1) and an AHB2 (S2) expressing strain, incubated with either anti-*Casuarina* symbiotic hemoglobin (Anti-Cassym) or anti-*Parasponia* hemoglobin (Anti-Phb) antisera. Apparent size of the AHB1 protein is marked on the left hand side of the figure (18kD as determined from the migration of size marker proteins and the *Parasponia* hemoglobin protein which was used as a positive control but not shown here).

AHB1 protein levels increase during hypoxia treatment.

Total protein was extracted from rosette leaves and from root tissue of plants that had been exposed to 0.1%, 5%, or 20% oxygen concentrations. The resulting protein extracts were separated on denaturing polyacrylamide gels, blotted onto nitrocellulose and incubated with the anti-Phb antiserum. A protein band of the size expected for AHB1, was detected in the root tissue protein extracts, but not the leaf tissue protein extracted from untreated plants (Figure 33). The signal from this protein was more intense in root extracts from plants treated with 5% oxygen and was more intense again in root extracts from plants treated with 0.1% oxygen. This pattern matches that of the *AHB1* transcript observed with northern analysis suggesting that this band does correspond to the AHB1 protein. AHB1 protein was also detected in extracts from leaves of plants subjected to 5% and 0.1% oxygen treatments, but was present at levels lower than in the root tissue extracts of plants from the same treatment. Thus, *AHB1* is translated in hypoxic tissues.

Figure 33. Expression of AHB1 protein in hypoxia treated plants.



Levels of AHB1 protein in a leaf or root tissues from control plants or plants subjected to 5%, 0.1% oxygen atmospheres. Loading order: Leaf tissue control (LC), Leaf 5% oxygen (L5), Leaf 0.1% oxygen (L.1), root tissue control (RC), root 5% oxygen (R5), root 0.1% oxygen (R.1). AHB1 bands migrated at an apparent size of 18kd.

5.3 Discussion.

The pattern of GUS staining in GH1 transformed lines, grown under normal growth conditions, corresponds with the results of RNA gel blot analysis of *AHB1* gene expression (see Chapter 3). A low level of GUS activity is present in root tissue and staining is not detected in above ground tissue for the majority of plants examined (other than the small number of cells in the hydatodes). Furthermore, the patterns of GUS reporter gene expression in GH1 transformed plants are similar to the patterns of GUS gene expression that have been observed in plants transformed with non-symbiotic hemoglobin promoter-GUS fusions. Staining in the tips and vascular tissue of roots has been observed in plants transformed with *Casuarina* non-symbiotic (Jacobsen-Lyon et al. 1995) and *Parasponia/Trema* promoter-GUS gene fusions (Bogusz et al. 1990, Andersson et al. 1997). The increased GUS activity observed in GH1 plants that have been treated with 1% sucrose also correlates with RNA gel blot analysis of *AHB1* gene expression. The pattern of sucrose induced GUS expression is similar to that produced by the *Casuarina* non-symbiotic hemoglobin promoter-GUS fusion in transgenic *Lotus*, which also confers sucrose responsive expression to the GUS reporter gene (Jacobsen-Lyon et al. 1995).

If these results are a true indication of AHB1 expression it appears that, under normal growth conditions, AHB1 expression is localised to tissues with high metabolic activity. The root tips are sites of rapid cell division and growth, and the vascular tissue is a site of active transport processes that consume metabolic energy. The location of the sucrose induced GUS staining in the vascular tissue of GH1-4 plants may also be connected with metabolic activity. Active transport of sucrose into, or through, phloem sieve tube elements may require an increase in metabolic

activity in surrounding cells. The staining sometimes observed in older leaves, particularly after sucrose induction, could also be associated with transport processes as metabolites are exported from these senescing tissues. Another possibility is that sucrose itself increases respiratory rates (Koch 1996), and therefore, induction of *AHB1* by sucrose could be connected with increased rates of respiration as sucrose is metabolised in some tissues.

Expression of the *AHB1* gene in metabolically active tissues deep inside the root tissue may be linked with oxygen availability. Direct measurements of oxygen concentrations in metabolically active root tissues, using microelectrodes, have shown that tissues inside the root can encounter hypoxia even in well aerated soils (Drew 1990). The constitutive expression of *AHB1* in root tissues may be a means of maintaining oxygen supply in tissues where oxygen availability may become limited, even under normal growth conditions.

The observation that there is no increase in GUS activity in hypoxia treated GH1 plants is surprising. When RNA is extracted from GH1-4 plants that have, or have not, been incubated in a 5% oxygen atmosphere the increase in the level of GUS transcript observed is similar to the increase in levels of the *AHB1* transcript. Thus, the 1 kb of 5' region from *AHB1* that is present in the GH1 construct is sufficient to confer transcriptional regulation to the GUS reporter gene. This transcriptional response could be mediated through the two putative AREs present in the *AHB1* promoter region that was included in the GH1 construct.

The lack of any increase in GUS activity in GH1 plants following hypoxia treatment may be due to translational controls. The "selective" translation of ANPs that appears to occur in hypoxia treated plant tissues

may be conferred by a sequence motif present in the transcripts encoding ANPs. As the majority of the AHB1 transcript is not present in the GH1 fusion construct it is possible that the regions required for translation of *AHB1* transcripts in hypoxic cells are absent from the GH1 construct. The introns and 3' untranslated region of the maize ADH gene, for instance, have been shown to increase the hypoxia response of ADH promoter fusions (Bailey-Serres and Dawe 1996).

A range of fusion constructs have been engineered to determine if sequences present in the *AHB1* transcript can confer the ability to be translated under hypoxic conditions to the GUS reporter gene. These constructs contain either the entire coding region from the genomic sequence (up to the penultimate codon), translationally fused to the GUS sequence (GH1C+), the 3' untranslated sequence of AHB1 instead of the nopaline synthase terminator (GH1T12), or both (GH1C+T12) (Appendix). These are currently being transformed into *Arabidopsis* C24 ecotype plants. Similar approaches have been used to analyse postranscriptional regulation of other plant genes (Bailey-Serres and Dawe 1996, Sieburth and Meyerowitz 1997). One potential pitfall of this approach is that fusing proteins to GUS may decrease, or even abolish GUS activity.

Regardless of the reasons, the lack of hypoxia responsive GUS activity in GH1 transformed lines is an example of why GUS constructs alone cannot be relied upon for expression analysis. Increasingly, sequences involved in regulating the expression of plant genes are being found to occur outside the 5' promoter region (see Taylor 1997). It also raises questions about the promoter-reporter gene fusions that have been used to analyse the expression of other non-symbiotic hemoglobin genes. Perhaps the lack of hypoxia response from the *Casuarina* and *Parasponia* promoter-GUS fusion constructs in transgenic *Lotus* is due to similar

translational effects rather than an absence of transcriptional response to hypoxia. Re-examining the hypoxia response of other non-symbiotic hemoglobin genes, by incubating plants in defined gas mixes and using RNA gel blots to examine gene expression, may be necessary to resolve this question.

The importance of the putative regulatory elements found in the 5' region of *AHB1* can still be examined despite the apparent translational regulation in hypoxia treated plants. RNA gel blot analysis of GUS transcript levels (rather than staining activity) could be used to examine GUS expression from promoter deletion or mutation constructs in transgenic *Arabidopsis*. This could determine if the putative AREs in the *AHB1* promoter, at -979 and -175 (Figures 12a&b), are involved in the induction of *AHB1* gene expression by hypoxia.

The results of the western analysis confirm that *AHB1* transcripts are translated in hypoxic tissues. Therefore, AHB1 is an ANP. The increased levels of AHB1 protein present in hypoxia treated or untreated plants is in agreement with northern data. The only difference between the data obtained from northern and western analysis is that more AHB1 protein is detected in extracts from roots than in extracts from leaves, when the amount of transcript in these tissues is roughly equivalent. This may reflect a difference between the efficiency of translation of *AHB1* transcripts in root and leaf tissue under hypoxic conditions.

It is possible that the function of AHB1 is as an oxygen scavenging protein. In cells where oxygen availability is limited, AHB1-bound oxygen could be delivered to the terminal oxidases of mitochondria. This could allow oxidative metabolism to continue during mild hypoxia. The increased levels of AHB1 protein in hypoxia treated tissues would be

consistent with such a function. The induction of barley hemoglobin gene expression when mitochondrial energy metabolism is disrupted (Nie and Hill 1997), also suggests a connection between hemoglobin gene expression and mitochondrial metabolism. Likewise, the induction of *AHB1* (and *Casnonsym*) gene expression by sucrose may be connected with increased mitochondrial respiratory metabolism creating an increased demand for oxygen. The basal expression of *AHB1* in cells that may encounter low oxygen availability during normal growth is also consistent with an oxygen scavenging function for AHB1. Thus, a possible function for AHB1, and other non-symbiotic Class 1 hemoglobins, is to deliver oxygen to mitochondria terminal oxidases when oxygen availability becomes limited.

Chapter 6.

The evolution and function of plant hemoglobins.

6.1 Two classes of non-symbiotic hemoglobin in plants.

Two families of plant hemoglobin gene have previously been identified in plants. These families have been differentiated on the basis of gene sequence and expression patterns, and have been referred to as the "symbiotic" and "non-symbiotic" hemoglobins (Jacobson-Lyon et al 1995). The two hemoglobin genes from *Arabidopsis* correspond in sequence to these two families. The predicted amino acid sequence of *AHB1* shows a high degree of similarity to the non-symbiotic hemoglobins of other plants, whereas the predicted amino acid sequence of *AHB2* is more similar to the symbiotic hemoglobins of legumes and *Casuarina.*

A number of methods of sequence analysis support this conclusion. Similarity trees clearly illustrate the division of plant hemoglobins into two classes (Figure 34, left side), with both nucleotide and amino acid sequence data sets. Parsimony analysis, which finds the tree(s) that describe(s) the inter-relationship of sequences with the least number of sequence transitions (the shortest tree), shows that the same clade occurs in the trees of maximum parsimony for both nucleotide and amino acid sequence data sets (Figure 34, right side). Both methods group *AHB1* with the non-symbiotic hemoglobins and *AHB2* with the symbiotic hemoglobins.

Furthermore, when *AHB2* nucleotide and predicted amino acid sequences were used to perform a BLAST search of known sequences the

results suggest that *AHB2* is more similar to symbiotic hemoglobins than non-symbiotic hemoglobins (Table 1). This is in contrast to the results of a BLAST search carried out with the predicted amino acid sequence of *AHB1* which is more similar to non-symbiotic hemoglobins (Table 1). Although the rice non-symbiotic hemoglobin genes occur in the ten sequences most closely related to *AHB2* when the predicted amino acid sequence is used on a BLAST search, the results of a BLAST search using the *AHB2* nucleotide sequence shows that *AHB2* is much more closely related to the *Casuarina* symbiotic hemoglobins than it is to any of the non-symbiotic hemoglobin genes.

While *AHB2* is related in sequence to the symbiotic hemoglobins, it is unlikely to function in nitrogen-fixing symbioses. *Arabidopsis* does not fix nitrogen and is not closely related to nitrogen-fixing plants. *AHB2* is more likely to have a non-symbiotic function, as indicated by gene expression in rosette leaves. Therefore, *AHB2* represents a new class of non-symbiotic plant hemoglobin.

The discovery of *AHB2* shows that the "symbiotic" hemoglobin genes are not confined to nitrogen-fixing plants. Instead, it seems that there are two classes of non-symbiotic plant hemoglobin gene (*AHB1* -like hemoglobins and the *AHB2* -like) and that the symbiotic hemoglobins of legumes and *Casuarina* may have evolved from an *AHB2* -like non-symbiotic hemoglobin gene. The use of the terms "symbiotic" and "non-symbiotic" to describe the two families of plant hemoglobins is therefore no longer valid. Instead, plant hemoglobins can be reclassified on the basis of predicted protein sequence as either Class 1, for the *AHB1* -like hemoglobins, or Class 2, for the *AHB2* -like hemoglobins. Hemoglobins from each class could then be further divided on the basis of whether they function in symbiotic or non-symbiotic tissues. For instance, *AHB2* is a

Class 2 non-symbiotic plant hemoglobin, while the leghemoglobins are Class 2 symbiotic hemoglobins. This also resolves any confusion associated with the classification of the *Parasponia* hemoglobin which was previously described as a non-symbiotic hemoglobin that functions in nitrogen-fixing symbiosis. The *Parasponia* hemoglobin can be reclassified as a Class 1 symbiotic hemoglobin.

When did the two classes of plant hemoglobin diverge?

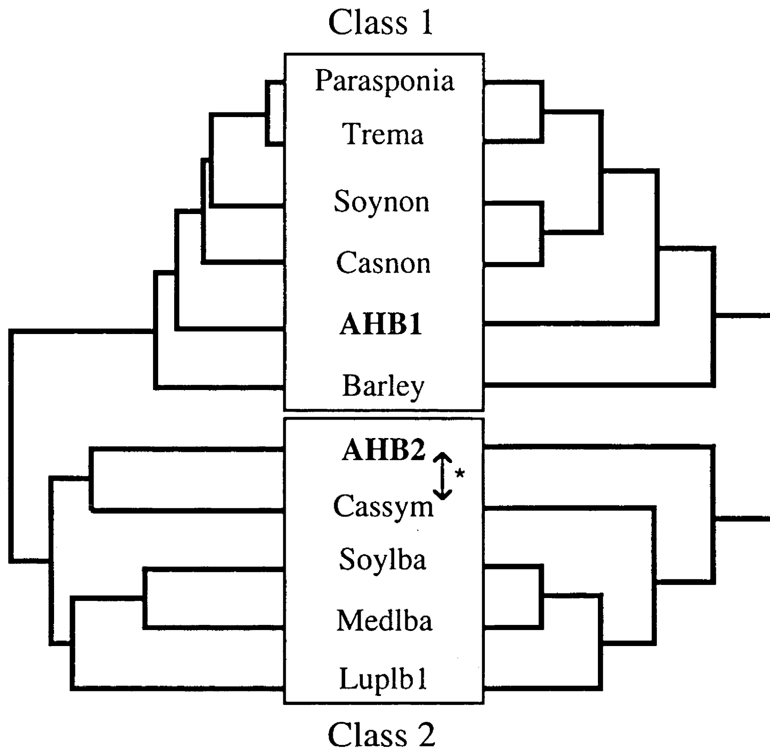
If we assume that non-symbiotic hemoglobin genes from each class of plant hemoglobin have an equal rate of sequence evolution then the two classes of plant hemoglobin probably diverged before the separation of monocots and dicots, as there is less similarity between the amino acid sequences of the two hemoglobins of *Arabidopsis* than there is between the Class 1 hemoglobins from *Arabidopsis* and barley. Following the divergence of plant hemoglobins gene families, Class 2 hemoglobins have been retained in the Brassicaceae, legumes and Casuarinaceae, at least. Thus, it seems likely that Class 2 hemoglobins will be present in other dicots and probably monocots.

The situation may not be so simple, however. The protein similarity tree of plant hemoglobins suggests that Class 2 hemoglobins may be evolving more rapidly than the Class 1 hemoglobins. There seems to be more variation amongst the Class 2 hemoglobins, indicated by longer branch lengths on the protein similarity tree, than there is for the Class 1 hemoglobins (Figure 34). Pairwise comparison of symbiotic and non-symbiotic hemoglobins from the same plants clearly illustrate this. The leghemoglobins and *Casuarina* symbiotic hemoglobins are less similar to each other (67% identical at the amino acid sequence level) than are the soybean and *Casuarina* non-symbiotic hemoglobins (88% identical at the

amino acid sequence level) suggesting a faster rate of divergence for the Class 2 hemoglobins from these plants. A phylogram of the most parsimonious tree also suggests that Class 2 hemoglobin protein sequences are diverging more rapidly than the those of the Class 1 hemoglobins, again indicated by longer branch lengths for the Class 2 clade (Figure 35). As a result of the difference in divergence rates for the two classes of plant hemoglobins it is difficult to determine when the divergence between these two classes of hemoglobin occurred.

The rapid divergence of legume and *Casuarina* symbiotic hemoglobins may reflect alterations in selective pressures and rapid sequence divergence after these hemoglobins were recruited into a symbiotic function, rather than faster evolution of Class 2 hemoglobins in general. In fact, there is some evidence that non-symbiotic Class 2 hemoglobins may be evolving at a similar rate to the Class 1 hemoglobins. Low stringency hybridisation analysis shows that both *Arabidopsis* hemoglobins cross hybridise with sequences from *Brassica napus*, but not those from the genomes of more distantly related plants. Furthermore, the Class 2 hemoglobins of *Arabidopsis* and *Brassica napus* are highly homologous (Richard Watts pers. comm.).

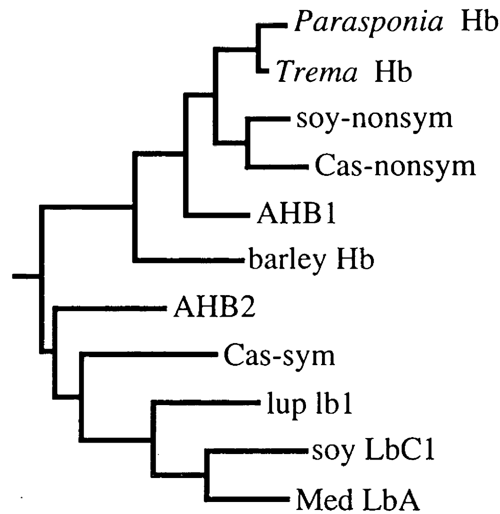
Figure 34. The two classes of plant hemoglobin.



The left hand tree is a similarity tree of plant hemoglobin protein sequences created using the GCG Pileup program. Class 1 and Class 2 hemoglobins form two distinct clades. Branch lengths represent relative similarity.

The right hand 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 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. The branch lengths are non-informative.

Figure 35. Phylogram of plant hemoglobin proteins.



Phylogram of plant hemoglobin protein sequences created using the PAUP 3.1.1 program. This tree demonstrates the more rapid divergence of the symbiotic Class 2 hemoglobins with relatively long branch lengths between symbiotic Class 2 hemoglobins from closely related species. The branch lengths are proportional to the degree of amino acid sequence divergence.

Key (Figure 34 and 35):
Hemoglobins, with genbank accesion numbers.
Barley: barley non-symbiotic hemoglobin (U01228). Casnon: *Casuarina* non-symbiotic hemoglobin (X53950). Soynon: soybean non-symbiotic hemoglobin (U47143). Parasp: *Parasponia* hemoglobin (U27194). AHB1: This study, (U94998). AHB2: This study, (U94999) Cassym: *Casuarina* symbiotic hemoglobin (X77695). Soylba: soybean leghemoglobin A (J01299). Medslba: *Medico sativa* leghemoglobin A (X14311). Luplb1: lupin leghemoglobin 1 (Y00401).

It is difficult to determine the rate at which Class 2 hemoglobins are diverging, and therefore predict when the two classes of plant hemoglobin genes diverged from a common ancestral plant hemoglobin gene, when non-symbiotic class 2 hemoglobins have only been cloned from two closely related plants (*Arabidopsis* and *Brassica*). The cloning of more Class 2 hemoglobins, from a broader range of plant species, would allow a

Table 1. BLAST search results.

The top ten BLAST search results for searchs (from most related to least related) carried out with AHB1 and AHB2 predicted amino acid sequences, and with the *AHB2* cDNA nucleotide sequence.

Query: AHB1 amino acid sequence

Accession number	Sequence	P(N)*
gb|U94968|HVU94968	Barley Hb (cDNA)	4.1e-77
gb|U01228|U01228	Barley Hb (gen)	3.2e-75
emb|Y00296|TTHB	Trema Hb	1.8e-74
gb|U76031|OSU76031	Rice Hb 2 (cDNA)	1.2e-73
gb|U76030|OSU76030	Rice Hb 1 (cDNA)	5.6e-73
gb|U27194|PAU27194	Parasponia Hb	5.1e-70
emb|X53950|CGHBIIG	Cas-nonsym.	3.7e-69
gb|U76029|OSU76029	Rice Hb 1 (gen)	3.6e-66
gb|U47143|GMU47143	Soy-nonsym.	2.4e-52

Query: AHB2 amino acid sequence.

Accession number	Sequence	P(N)
emb|X77695|CGSYMB	Cas-sym B (cDNA)	1.2e-57
emb|X77694|CGSYMA	Cas-sym A (cDNA)	5.6e-57
emb|X14311|MSLEGHE2	M. sativa Lb2	8.5e-56
emb|X13375|MSLEGHEM	M. sativa Lb	1.2e-55
gb|M91077|ALFLEGHEMO	M. sativa Lb3	2.9e-55
emb|X77042|LLLHGII	L.luteus LbII	3.5e-55
gb|U76029|OSU76029	Rice Hb 1 (gen)	4.1e-55
emb|Z54158|VFLBKMR	V.faba LbK	2.4e-54
gb|U76030|OSU76030	Rice Hb 1 (cDNA)	5.4e-54
gb|M23312|SESLBDRLA	S.rostrata Lb1)	7.1e-54

Query: AHB2 cDNA nucleotide sequence

Accession number	Sequence	P(N)
emb|X77695|CGSYMB	Cas-sym B (cDNA)	7.9e-69
emb|X77694|CGSYMA	Cas-sym A (cDNA)	2.6e-67
emb|X77696|CGSYMC	Cas-sym C (cDNA)	2.3e-58
emb|Z54157|VFLBBMR	V.faba LbB	3.8e-46
gb|M91077|ALFLEGHEMO	M. sativa Lb3	2.2e-40
gb|U33205|VUU33205	V. unguiculata Lb II	1.7e-35
emb|X77042|LLLHGII	L.luteus LbII	5.3e-35
gb|U76030|OSU76030	Rice Hb 1 (cDNA)	8.2e-35
gb|U76031|OSU76031	Rice Hb 2 (cDNA)	2.0e-31

* P(N) is the smallest sum probability value. This represents the chance that the sequence found is related to the query sequence by chance alone. The more similar the sequence identified the smaller the P(N) value.

more informed picture of plant hemoglobin evolution to be formulated. As discussed earlier (Chapter 3 introduction) there are often low levels of sequence homology between hemoglobins from distantly related plants and cloning homologues using DNA hybridisation techniques can be difficult. PCR using degenerate primers that target the small stretches of sequence that are conserved in these genes may prove successful for cloning Class 2 hemoglobins, as has been the case with Class 1 hemoglobins. Antibodies raised against recombinant AHB1 and AHB2 could also be used to screen cDNA expression libraries (synthesised from mRNA from the appropriate plant tissues) for AHB2 homologues. This approach would target the tertiary structure of hemoglobins, another feature which is strongly conserved between hemoglobin proteins.

The evolutionary origins of the *Parasponia* hemoglobin.

The hemoglobin from the nitrogen-fixing elm *Parasponia* (Ulmaceae) clearly groups with the Class 1 hemoglobin gene family (Figure 34). However, unlike other Class 1 hemoglobins it is expressed in nodules and appears to fulfil a symbiotic function analogous to the role of the nodule expressed hemoglobins of legumes and *Casuarina*. Thus, a Class 1 hemoglobin has acquired a symbiotic function during the evolution of the *Parasponia-Rhizobium* nitrogen-fixing symbiosis. In addition to high level gene expression in nodules, the *Parasponia* hemoglobin gene retains an expression pattern typical of Class 1 hemoglobins with a low level of expression in root tissue. The *Parasponia* hemoglobin may fulfil both symbiotic and non-symbiotic functions (Bogusz et al. 1988, Bogusz et al. 1990).

There are two explanations for the evolution of dual functions for the *Parasponia* hemoglobin. The promoter of the *Parasponia*

hemoglobin gene may have mutated leading to high levels of gene expression in nodules, while maintaining the sequences required for a low level of gene expression in root tissue. Thus, the *Parasponia* hemoglobin may have acquired a function in nitrogen-fixing nodules, and retained a non-symbiotic function in roots. It is also possible that a gene duplication has occurred during the evolution of *Parasponia*, allowing one hemoglobin gene to evolve a symbiotic function while the other gene retained a non-symbiotic function. If this second theory is correct the expression of the *Parasponia* hemoglobin in root tissue is an evolutionary artefact and a second, undetected, hemoglobin gene may be present in the *Parasponia* genome. Since a *Parasponia* hemoglobin gene probe detected only single bands on blots of *Parasponia* genomic DNA (Landsmann et al. 1986) this seems unlikely.

The evolutionary origins of the Class 2 symbiotic hemoglobins.

In contrast to the situation in *Parasponia*, the coding sequences of the symbiotic hemoglobins from legumes and *Casuarina* are clearly *AHB2* -like (Figure 34) It seems that in legumes and *Casuarina*, Class 2 hemoglobins have been recruited into a symbiotic function. Thus, there are examples of both Class 1 and Class 2 hemoglobins being recruited into symbiotic roles during the evolution of nitrogen-fixing symbioses.

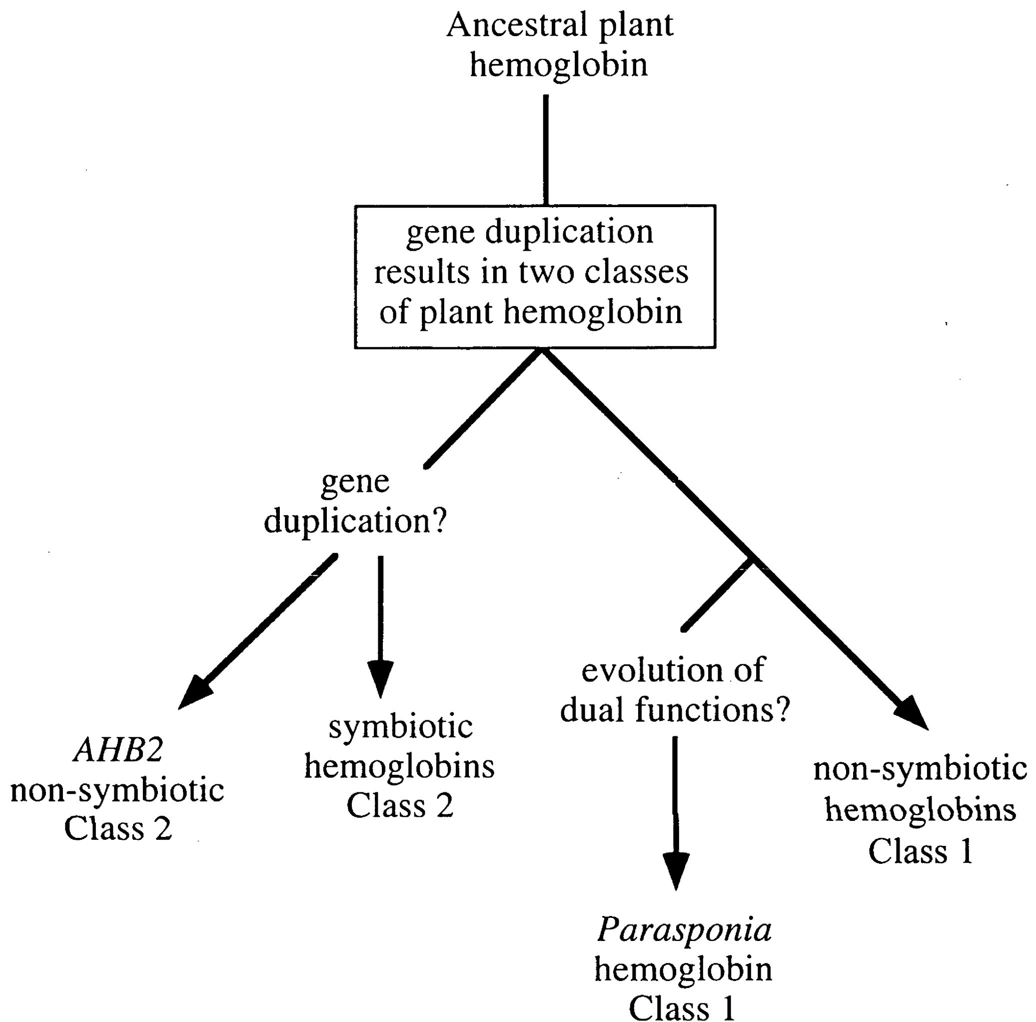
Figure 36. Partial phylogeny of dicotyledonous plants.

Molecular phylogeny created by parsimony analysis of Rubisco Lsu sequence data (simplified from Chase et al. 1993). Some plants from the Rosid I and Rosid 2 clades are shown. For nitrogen-fixing plants (shown in bold text) the class of hemoglobin expressed in nodules is also indicated (numbers in brackets indicate the class of symbiotic hemoglobin). A (?) denotes a plant that has been shown to have a symbiotic hemoglobin, the sequence of which is not known.

The leghemoglobins and the *Casuarina* symbiotic hemoglobins may have evolved from the duplication of a Class 2 hemoglobin gene. This would have allowed a symbiotic Class 2 hemoglobin to evolve nodule specific expression, while a non-symbiotic Class 2 hemoglobin retained a function in normal plant tissues. If this was the case, it is probable that soybean and *Casuarina* retain non-symbiotic Class 2 hemoglobin genes. Cloning a non-symbiotic Class 2 hemoglobin from soybean would support this hypothesis and may contribute a great deal to our understanding of the evolution of the leghemoglobins.

Symbiotic hemoglobins have been purified from two other nitrogen fixing trees that form actinorhizal symbioses, *Myrica* (Pathirana and Tjepkema 1995) and *Alnus* (Suharjo and Tjepkema 1995). Given that *Myrica* and *Alnus* seem to be closely related to *Casuarina* (Chase et al. 1993), these symbiotic hemoglobins are likely to belong to the Class 2 group (Figure 36). Sequences that cross hybridise strongly to a soybean leghemoglobin probe have also been detected in the genomic DNA from the actinorhizal plants *Alnus*, *Caenothus americanus* and *Elaeagnus pungens* (Roberts et al. 1985) suggesting that Class 2 hemoglobin genes are present in these plants. Sequence from these genes or the corresponding proteins could determine the origins of symbiotic hemoglobins expressed in the nodules of *Myrica* and *Alnus*.

Figure 37. The evolution of plant hemoglobin genes.



A schematic representation of plant hemoglobin gene evolution. Following a gene duplication event, two divergent families of plant hemoglobin gene have been present in plants. A second gene duplication event could have given rise to the symbiotic hemoglobins of legumes and *Casuarina*, while the *Parasponia* hemoglobin may have evolved a symbiotic function in nitrogen-fixing nodules while still retaining a non-symbiotic function in roots.

Plant hemoglobin promoters.

The promoters of the symbiotic hemoglobin genes from legumes and *Casuarina* contain a number of motifs with homology to those found in the promoters of Class 1 hemoglobin promoters. Furthermore, leghemoglobin promoters can function in a similar manner to the promoters of the Class 1 hemoglobins. When the promoter of the *Sesbania rostrata* leghemoglobin is fused to GUS and transformed into tobacco it directs staining activity to the root tips (Szczyglowski et al. 1996), as do the Class 1 hemoglobin promoters from *Arabidopsis* (this thesis), *Casuarina* (Jacobsen-Lyon et al. 1995) and *Parasponia* (Bogusz et al. 1990). The Class 1 hemoglobin promoter, containing the AAAGAT and CTCTT motifs may be a general plant hemoglobin promoter.

There is no obvious homology between the promoter of *AHB2* and the promoters of the other plant hemoglobins. The AAAGAT and CTCTT motifs do not occur in the *AHB2* promoter sequence. Furthermore, it appears that the promoter of *AHB2* functions in a different manner to the promoters of other plant hemoglobin genes, directing gene expression to rosette leaves. There are two possible explanations for the differences between the promoter of *AHB2* and those of other plant hemoglobins. A sequence rearrangement may have occurred in the ancestor of *Arabidopsis* resulting in the promoter from another gene controlling the expression of a Class 2 hemoglobin in this plant. Alternatively, if the promoter of *AHB2* is typical of Class 2 hemoglobins, a rearrangement may have occurred in the genome of an ancestor of legumes and *Casuarina,* resulting in the acquisition of a Class 1 promoter by Class 2 hemoglobin genes. Theoretically, this could occur through a recombination event between genes encoding the two different classes of hemoglobin. Perhaps the acquisition of a Class 1 hemoglobin promoter allowed a Class 2

hemoglobin to evolve high level nodule-specific gene expression in the ancestor of legumes and *Casuarina*.

In order to determine whether the "Class 1" hemoglobin promoter is the general plant hemoglobin promoter, or whether other non-symbiotic Class 2 hemoglobins will have 5' regions homologous to that of AHB2, non-symbiotic Class 2 hemoglobins must be isolated from other plants (particularly plants from other dicot clades) and the promoters from such genes compared with the promoter of *AHB2*.

Hemoglobins and the origins of nitrogen-fixing symbioses.

All nitrogen-fixing plants group to a single dicot clade (Chase et al. 1993, Soltis et al. 1995) suggesting a single origin of plants that are able to enter into nitrogen-fixing symbioses, and despite differences in nodule ontogeny and structure, there are extensive similarities in the infection processes of legume-*Rhizobium*, *Parasponia-Rhizobium* and actinorhizal symbioses (reviewed in Pawlowski and Bisseling 1996). A common ancestor of these plants may have possessed a trait that allowed it to enter into symbiotic nitrogen-fixing relationships. Subsequently a number of nitrogen fixing symbioses have evolved.

The distribution of nitrogen fixing plants in the Rosid I clade is patchy. Many nitrogen fixing plants have close relatives that are unable to form nitrogen fixing symbioses. Furthermore, there is no phylogenetic division between actinorhizal plants and the legumes, with some actinorhizal plants being more closely related to legumes than to other actinorhizal species (Soltis et al. 1995). Conversely, *Parasponia* is more closely related to actinorhizal plants but, like legumes, is a host for *Rhizobium* .

The patchy and apparently random distribution of the different types of nitrogen-fixing symbioses could result from the combined effects of a number of evolutionary processes. Nitrogen-fixing symbiosis may have evolved in the common ancestor of nitrogen-fixing plants. Subsequently, the symbiotic relationships may have been modified in some descendants of the original nitrogen fixing plant and lost completely from others. *Frankia*, which has the broadest distribution of host plants (Soltis et al. 1995), may have been the original microsymbiont but appears to have been displaced by *Rhizobia* in legumes. This would explain the similarity between the symbioses of actinorhizal and leguminous plants.

Once microbes such as *Frankia* and *Rhizobium* acquired the ability infect susceptible plants and form symbiotic relationships it is also possible that they entered into symbiotic relationships with other related plant species that did not enter into nitrogen-fixing symbioses. This lateral transfer of symbiotic microbes from one host plant to another may explain the origins of the *Parasponia-Rhizobium* symbiosis. *Rhizobium*, already able to establish symbiotic relationships with legumes, may have entered into a symbiotic relationship with *Parasponia*. The *Rhizobium* strain that infects *Parasponia* is able to infect a wide range of legume species supporting this idea (see Frelberg et al. 1997). A recent "lateral" origin for the *Parasponia-Rhizobium* symbiosis would also explain why this symbiosis occurs in only one genus of plants.

The analysis of hemoglobin genes from nitrogen fixing plants supports such a model. *Casuarina*, an actinorhizal plant, has a Class 2 symbiotic hemoglobin with similar protein sequence, expression patterns and oxygen binding properties to the symbiotic hemoglobins of legumes supporting a common origin for these symbioses. The Class 1 symbiotic

hemoglobin of *Parasponia* suggests a separate origin for the *Parasponia-Rhizobium* symbiosis. Extensive analysis of other nodulin genes has been carried out in legumes and progress is currently being made in the analysis of nodulin genes from the actinorhizal plant *Alnus* (see Pawlowski and Bisseling 1996). It will be interesting to see if comparison of other nodulin genes involved in the various symbiotic relationships supports this model.

Mycorrhizal symbioses, where fungi growing in association with plant root tissues protect the plant from other microbes and aid in nutrient uptake, occur in a wide range of land plant taxa (Gianinazzi-Pearson 1996). Recently a number of physical similarities between mycorrhizal symbioses and nitrogen-fixing symbioses have been demonstrated. Furthermore, the formation of both types of symbioses requires a common set of genes. Many mutations that block nodulation in legumes also block the formation of endo-mycorrhizal symbioses (Gianinazzi-Pearson 1996). Some nodulin genes are also expressed in response to infection by mycorrhizal fungi (Gianinazzi-Pearson 1996), as well as being expressed in nodules. It may have been an alteration in some aspect of endo-mycorrhizal symbiosis of an ancestral Rosid I plant that allowed nitrogen-fixing symbioses to evolve in this group of plants.

A leghemoglobin gene, *VfLb29*, that is induced by infection with endo-mycorrhizal funghi (and also expressed at high levels in nodules) has been cloned from *Vicia faba* (Fruhling et al. 1997). The *VfLb29* protein sequence forms an outgroup to the other leghemoglobins of *Vicia*, but is more closely related to leghemoglobins than any other plant hemoglobins (Fruhling et al. 1997). The induction of *VfLb29* by infection with mycorrhizal fungus may be evidence for a broader distribution of symbiotic hemoglobins. If hemoglobins have a function in mycorrhizal

symbiosis, then symbiotic hemoglobins may also occur in a range of plants that is wider than currently thought (Fruhling et al. 1997). As mycorrhizal endosymbionts do not fix nitrogen it is unlikely that *VfLb29* is required to buffer oxygen in order to protect nitrogenase. Instead, it has been suggested that *VfLb29* may transport oxygen in infected cells (Fruhling et al. 1997).

6.2 The oxygen binding kinetics of *Arabidopsis* hemoglobins.

The oxygen binding kinetics of AHB1 and AHB2 have been examined (Richard Watts pers. comm.). Both hemoglobins have extremely slow, pH sensitive, oxygen dissociation rates (Table 2). The same has been found for the barley (Duff et al. 1997), rice (M. Hargrove pers. comm.) and soybean (Richard Watts pers. comm.) Class 1 hemoglobins. These off rates are significantly lower than those of symbiotic plant hemoglobins. As a result, the overall affinity of AHB1 (which has a fairly rapid oxygen association rate) for oxygen is extremely high, much higher than the symbiotic hemoglobins. AHB2 has a lower oxygen association rate constant than AHB1, apparently due to competition between oxygen binding and the formation of a hemochrome complex in which the distal histidine binds the heme iron (Richard Watts pers. comm.). This results in AHB2 having a lower overall oxygen affinity (P_{50}) than AHB1 (Table 2). The slow, pH sensitive, off rates of AHB1 and AHB2 may be a conserved feature of non-symbiotic plant hemoglobins.

Table 2. The oxygen binding to properties of various hemoglobins.

Protein	P_{50} nM	kOff (S^{-1})
AHB1	1.6	0.12
Barley	2.8	0.02
Parasponia	89	13.8
AHB2	130	0.14
Soy lbA	44	5.6
Cassym	135	6
SwMb.	857	12

* k_{Off} is the bimolecular dissociation rate constant for bound O_2. The dissolved O_2 concentration at which hemoglobin is half saturated (P_{50}) is equivalent to the dissociation equilibrium constant for O_2 ($K_D = k_{on}/k_{off}$) and gives an indication of the relative affinity of different hemoglobins for oxygen. Cassym: *Casuarina* symbiotic hemoglobin. Soy lba: recombinant soybean leghemoglobin A. SwMb: spermwhale myoglobin (see Hargrove et al. 1997, Gibson et al. 1989).

AHB1 has too high an affinity for oxygen to allow release of oxygen to terminal oxidases. The P_{50} of AHB1 is lower than any known terminal oxidase (Millar et al. 1994). Both hemoglobins have such slow oxygen dissociation rates that it seems neither would allow rapid oxygen exchange. Thus, it is unlikely that either AHB1 or AHB2 is able to act as oxygen transporters in the manner of the symbiotic hemoglobins.

Members of each class of non-symbiotic plant hemoglobin have been recruited into symbiotic function. This suggests that the alterations in protein sequence required to convert the oxygen binding kinetics of either type of non-symbiotic hemoglobin into those suitable for a symbiotic oxygen transport/buffering function may occur with relatively few amino acid residue substitutions. The *Parasponia* hemoglobin, with amino acid sequence very similar to other Class 1 hemoglobins, but oxygen binding kinetics typical of a symbiotic hemoglobin supports this theory. The residues required to generate the pH sensitive, extremely slow off rates observed in the non-symbiotic plant hemoglobins may be identified by comparing the *Parasponia* hemoglobin, which does not show this oxygen binding behaviour, to other non-symbiotic plant hemoglobins. There are few unique amino acid residue substitutions in the *Parasponia* sequence (eg Phe59->Tyr, see Figure 10) and the importance of these residues could be examined through *in vitro* mutagenesis and oxygen binding studies of recombinant AHB1 or *Parasponia* hemoglobin.

The oxygen binding properties of the *Parasponia* hemoglobin are so different to those of the non-symbiotic Class 1 hemoglobins that it seems possible that any non-symbiotic function has been lost. The *Parasponia* hemoglobin may therefore only function as a symbiotic hemoglobin. The low level of expression in root tissue might represent a redundant feature of the Class 1 hemoglobin promoter, similar to the root tip expression from the *Sesbania* promoter-GUS fusion expressed in tobacco. If this is the case then either a second Class 1 hemoglobin has been retained in *Parasponia* or the function fulfilled by Class 1 hemoglobins is not essential for the survival of this plant.

6.3 The function of non-symbiotic plant hemoglobins?

Understanding the evolution of plant hemoglobins and the recruitment of these proteins into nitrogen-fixing symbioses is difficult in the absence of any understanding of the functions of non-symbiotic plant hemoglobins. At present there are three lines of evidence that may offer some insight into the possible functions of non-symbiotic plant hemoglobins; the protein/gene sequences, the patterns of gene expression and the oxygen binding kinetics. All three lines of evidence suggest that there are two distinct types of hemoglobins in plants and this in turn implies that there are probably two distinct requirements for hemoglobins in plants.

AHB1 and AHB2 are unlikely to be oxygen transporters.

The oxygen binding properties of AHB1 and AHB2 suggest that neither operates as an oxygen transporter in the manner of myoglobin or the leghemoglobins. The P_{50} of AHB1 is lower than that of any known plant terminal oxidase (Millar et al. 1994) suggesting that the oxygen affinity of AHB1 is too high for release of oxygen to cytochrome oxidases. Furthermore, both hemoglobins exhibit stable binding of oxygen, as indicated by low dissociation rates, suggesting that they would be inefficient oxygen transporters. This kinetic data is supported by the observation that the levels of transcript produced from these genes are low and are therefore unlikely to produce sufficient amounts of hemoglobin protein to affect oxygen flux rates.

AHB1 and AHB2 are unlikely to be oxygen sensors.

It has been suggested previously that non-symbiotic hemoglobins (Class 1) could act as oxygen sensors. In the case of Class 1 hemoglobins this appears unlikely. The affinity of recombinant AHB1 (and other non-symbiotic Class 1 hemoglobins) for oxygen is so high that it is unlikely to be deoxygenated even under extremely hypoxic conditions. Thus, AHB1 would not be a suitable oxygen sensor as it is unable to perceive moderate changes in oxygen tension . The oxygen affinity of AHB2 is lower than those of non-symbiotic Class 1 hemoglobins, possibly due to competition between the distal histidine and oxygen for access to the heme iron. AHB2, with a lower affinity for oxygen, could act as an oxygen sensor. However, the observation that *AHB2* gene expression is confined to above ground tissue appears to rule out a function in oxygen sensing. Hypoxia, resulting from waterlogging, would mainly affect below ground tissues and an oxygen sensor would be expected to be expressed in the roots of a plant.

AHB1 may deliver oxygen to mitochondrial terminal oxidases.

The observations that the both *AHB1* and the barley hemoglobin (Class 1) are induced by hypoxia, and that the barley hemoglobin is induced by lowering ATP/ADP ratios, suggest an involvement of Class 1 hemoglobins in mitochondrial respiration.

Despite the high oxygen affinity of recombinant AHB1, which suggests that oxygen would not be released to mitochondrial terminal oxidases, this hemoglobin may be able to deliver oxygen to mitochondrial terminal oxidases. It is possible that the *in vitro* oxygen binding properties of plant hemoglobins do not apply *in vivo*. Specific interactions

with other proteins or ligands may modify oxygen binding behaviour such that delivery of oxygen to terminal oxidases can occur. It has been suggested that myoglobin may interact specifically with another protein, an unidentified component of mitochondria (Wittenberg and Wittenberg 1989). A function for AHB1 in delivering oxygen to terminal oxidases cannot be ruled out completely on the basis of kinetic data alone. Examining rates of oxygen consumption when recombinant AHB1 or AHB2 preparations are incubated with intact mitochondria may provide a more realistic test for any effects that these hemoglobins have on mitochondrial respiration.

AHB1 may scavenge oxygen in hypoxic tissues.

It is possible that Class 1 hemoglobins scavenge oxygen in hypoxic tissues to provide oxygen for other oxygen consuming processes. Recently it has been shown that *Arabidopsis* plants have an absolute requirement for oxygen. Plants that have been pretreated with 5% oxygen (inducing ANP synthesis) can survive 24hrs in extreme hypoxia (0.1% oxygen/nitrogen mixture) but not 24hrs in complete anoxia (100% nitrogen), (Marc Ellis personal communication). The requirement for such small amounts of oxygen is unlikely to be connected with mitochondrial respiration (as the P_{50} of cytochrome oxidase suggests that it is unlikely to bind oxygen at such low oxygen tensions) but the exact nature of this need for oxygen is unclear. It is possible that Class 1 hemoglobins scavenge oxygen to help meet this absolute requirement for oxygen. The strong induction of *AHB1* by extreme hypoxia (0.1% oxygen atmosphere), and the extremely high affinity of recombinant AHB1 for oxygen support the theory that AHB1 scavenges oxygen in hypoxic tissues for use in oxygen consuming reactions.

AHB1 and AHB2 may channel bound oxygen into biosynthetic reactions.

Many biosynthetic reactions require molecular oxygen. It is possible that either class of non-symbiotic plant hemoglobin interact with oxygen consuming catalytic proteins *in vivo*. The slow oxygen dissociation rates would be consistent with a function in binding oxygen tightly until interaction with another protein occurs. Hemoglobin proteins are known to interact with electron transport enzymes. Ferric leghemoglobin reductase and methemoglobin reductase maintain leghemoglobins and vertebrate hemoglobins respectively in the ferrous state (Ji et al. 1994, Kuma 1981) by transferring electrons to the heme iron. This interaction may be of functional significance in the case of non-symbiotic plant hemoglobins. It is also possible that plant hemoglobins may have an undetected catalytic property themselves. The exact nature of such reactions is difficult to envisage.

6.4 Determining the functions of Class 1 and Class 2 plant hemoglobins.

Cloning of non-symbiotic Class 2 hemoglobins from other plants.

It is likely that Class 2 non-symbiotic hemoglobins will be present in a wide range of plants, as are the Class 1 non-symbiotic hemoglobins. These genes may have been retained in the genomes of legumes, and *Casuarina*, and cloning a non-symbiotic Class 2 hemoglobin from these plants may provide a great deal of information about the origins of the leghemoglobins. The cloning and sequencing of Class 2 hemoglobin genes from a wider range of plants may also resolve the disparity between the AHB2 and leghemoglobin promoters. Examining the expression

patterns of *AHB2* -like genes in other plants may also increase our understanding of the function of these non-symbiotic hemoglobins. However, as seems to be the case for Class 1 hemoglobins more direct analysis of hemoglobin function may be required.

Localisation of *AHB1* and *AHB2* gene expression.

Localising the expression of each hemoglobin to specific cell types is an essential part of determining the functions of these proteins. The promoter-reporter gene fusion analysis of *AHB1* and other Class 1 hemoglobins may have given some indication of what types of cell these hemoglobins are expressed in but (as discussed in Chapter 5) the location of hypoxia induced *AHB1* gene expression is unknown. Even less is known of *AHB2* gene expression patterns, although *AHB2* promoter-GUS fusions are currently being transformed into *Arabidopsis* (Appendix). The apparent translational regulation of AHB1 during hypoxia may complicate the promoter-reporter gene approach. An alternative method that could be used to localise expression of either hemoglobin to specific cells would be *in situ* hybridisation. The levels of *AHB1* transcript detected by probing RNA gel blots should be detectable with *in situ* hybridisation. However, the levels of *AHB2* transcript, as observed on RNA gel blots, may be too low for detection by *in situ* hybridisation.

Localisation of AHB1 and AHB2 proteins.

The subcellular location of AHB1 and AHB2 proteins is another area that can be examined. There are no obvious nuclear, chloroplast or mitochondrial import sequences in either hemoglobin. However, given the apparent association between expression of the barley hemoglobin and mitochondrial energy metabolism, it would be interesting to see if AHB1

is found in high concentrations around active mitochondria. Antibodies raised against these proteins may allow immuno-histochemical localisation. This technique has been used to localise *Parasponia* hemoglobin in infected cells (Trinick et al. 1989). A potential problem with this approach is that it may not be sensitive enough to detect low levels of AHB2. Also, if the anti-hemoglobin antibody preparations show background binding to other plant proteins then specificity may be a problem. An alternative approach could be to extract proteins from subcellular fractions (eg cytoplasmic, nuclear, plastid and mitochondrial) of different *Arabidopsis* tissues and then use western analysis to compare the levels of each hemoglobin in the various subcellular fractions.

Proteins that interact with AHB1 and AHB2?

In order to understand the function of these hemoglobins it may be necessary to determine if they interact with any other proteins *in vivo*. A powerful method that can be used to examine protein-protein interactions is the "Two Hybrid" system in yeast. This system allows cDNA expression libraries to be screened for genes encoding proteins that can bind a target protein, using expression of a reporter gene as a detection system (Ausubel et al. 1994). This system has the advantage that genes encoding any proteins able to bind a target protein are also isolated. The Two-Hybrid system could be used to search for proteins that interact with either AHB1 or AHB2.

Heterologous expression.

The ability to express functional hemoglobins in bacteria and yeast allows another genetic approach to be implemented. Heterologous expression of plant hemoglobins could be used to test for any effects on

cellular metabolism and for complementation of a range of mutations. For instance, expressing hemoglobins in yeast and comparing rates of oxygen consumption with untransformed strains may help determine if either type of hemoglobin is able to deliver oxygen to mitochondrial respiratory metabolism.

Reverse genetics.

Perhaps the most direct and powerful method that can be employed to analyse the function of plant hemoglobins is to generate *Arabidopsis* plants that over or under express AHB1 and AHB2. Plants transformed with hemoglobin over-expression or anti-sense constructs may show a dramatic phenotype that clearly indicates a specific function. Transgenic lines that are known to under or overexpress either hemoglobin could also be examined under a range of environmental conditions for deviations from the phenotype of wild type plants. For example, plants over or under expressing AHB1 could be assayed for hypoxia tolerance using defined gas mixes and compared with untransformed plants. Transformation of *Arabidopsis* with over-sense and anti-sense constructs for *AHB1* and *AHB2* is currently under way (Appendix).

6.5 Conclusion.

The cloning of two divergent hemoglobin genes from *Arabidopsis* suggests that there are two classes of non-symbiotic hemoglobin gene in a wide range of plants. Furthermore, it seems that hemoglobins from each class have been recruited into the symbiotic function of transporting oxygen in nitrogen-fixing symbioses. While the functions of these two classes of non-symbiotic hemoglobin are unclear, it is unlikely that either functions as an oxygen transporter or as an oxygen sensor. Analysis of transgenic plants that over or under express these hemoglobin genes may help elucidate the functions of non-symbiotic plant hemoglobins.

Bibliography.

Adamski, I., Kloppstech, K. (1994) Low temperature increases the abundance of early light inducible transcript under light stress conditions. J. Biol. Chem., **269**, 30221-30226.

Andersson, C.A. (1994) PhD Thesis, Australian National University.

Andersson, C.A., Jensen, E.O., Llewellyn, D.J., Dennis, E.S., Peacock, W.J. (1996) A new hemoglobin in soybean: a role for hemoglobin in all plants. PNAS, **93**, 5682-5687.

Andersson, C.A., Llewellyn, D.J., Peacock, W.J., Dennis, E.S. (1997) Cell-specific expression of the promoters of two nonlegume hemoglobin genes in a transgenic legume, *Lotus corniculatus*. Plant Phys., **113**, 45-57.

Antoine, M., Niessing, J. (1984) Intron-less globin genes in the insect *Chironomus thummi thummi.* Nature, **340**, 795-798.

Appleby, C.A., Tjepkema, J.D., Trinick, M.J. (1983) Hemoglobin in a Non-leguminous plant, *Parasponia* : Possible genetic origin and function in Nitrogen fixation. Science, **220**, 951-953.

Appleby, C.A. (1984) Leghaemoglobin and *Rhizobium* respiration. Ann. Rev. Plant. Phys., **35**, 443-478.

Appleby, C.A. (1985) Plant Haemoglobin Properties, Function, and Genetic Origin. In *Nitrogen Fixation and CO2 Metabolism*, Elsevier, New York.

Appleby, C.A., Bogusz, D., Dennis, E.S., Peacock, W.J. (1988) A role for haemoglobins in all plant roots? Plant, Cell and Environment., **11**, 359-367.

Appleby, C.A. (1992) The origin and function of haemoglobin in plants. Scientific Progress, **76**, 365-398.

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. New York.

Bailey-Serres, J., Dawe, R.K. (1996) Both 5' and 3' sequences of maize *Adh1* mRNA are required for enhanced translation under low oxygen conditions. Plant Phys., **112**, 685-695.

Bashford, D., Chothia, C., Lesk, A.M. (1987) Determination of protein fold, Unique features of globin amino acid sequences. J. Mol. Biol., **196**, 199-216.

Baulcombe, D., Verma, D.P.S. (1978) Preparation of a complementary DNA for leghaemoglobin and direct demonstration that leghaemoglobin is encoded by the soybean genome. Nuc. Acid. Res., **5**, 4141-4153.

Bingham, I.J., Stevenson, E.A. (1993) Control of root growth: effects of carbohydrates on the extension, branching, and rate of respiration of different fractions of wheat roots. Physiologica Plantarium, **88**, 149-158.

Bogusz, D., Appleby, C.A., J, L., Dennis, E.S., Peacock, W.J. (1988) Functioning haemoglobin genes in non-nodulationg plants. Nature, **331**, 178-180.

Bogusz, D., Llewllyn, D.J., Craig, S., Dennis, E.S., Appleby, C.A., Peacock, W.J. (1990) Nonlegume hemoglobin genes retain organ specific expression in heterologous transgenic plants. The Plant Cell, **2**, 663-641.

Chang, C., Meyerowitz, E.M. (1986) Molecular cloning and DNA sequence of the *Arabidopsis thaliana* alcohol dehydrogenase gene. PNAS, **83**, 1408-1412.

Chase, M.W., Soltis, D.E., Olmstead, R.G., D, M., Les, D.H., Mishler, B.D., Duvall, M.R., Price, R.A., Hills, H.G., Qiu, Y.L., Kron, K.A., Rettig, J.H., Conti, E., Palmer, J.D., Manhart, J.R., Sytsma, K.J., Miachaels, H.J., Kress, W.J., Karol, K.G., Clark, W.D., Hedrem, M., Gaut, B.S., Jansen, R.K., Kim, K.J., Wimpee, C.F., Smith, J.F., Furnier, G.R., Strauss, S.H., Xiang, Q.Y., Plunkett, G.M., Soltis, P.S., Golenberg, E., Learn, G.H.J., Graham, S.W., Barret, S.C.H.,

Dayanandan, S., Albert, V.A. (1993) Phylogenetics of seed plants: analysis of nucleotide sequences from the plastid *rbcl* gene. Ann. Missouri Bot. Gard., **80**, 528-580.

Chen, W., Hughes, D.E., Bailey, J.E. (1994) Intracellular expression of *Vitreoscilla* hemoglobin alters aerobic metabolism of *Saccharomyces cerevisiae.* Biotechnol. Prog., **10**, 308-313.

Christensen, T., Dennis, E.S., Peacock, J.W., Landsmann, J., Marcker, K.A. (1991) Hemoglobin genes in non-legumes: cloning and characterisation of a *Casuarina glauca* hemoglobin gene. PMB, **16**, 339-344.

Couture, M., Chamberland, H., St Pierre, B., Lafontaine, J., Guertin, M. (1994) Nuclear genes encoding chloroplast hemoglobins in the green alga *Chlamydomonas eugametos.* MGG, **243**, 185-197.

Cramm, R., Siddiqui, R.A., Friedrich, B. (1994) Primary sequence and evidence for a physiological function of the flavoprotein of *Alcaligenes eutrophus.* J. Biol. Chem., **269**, 7349-7354.

Davidowitz, E.J., Lang-Unnasch, N. (1989) Nucleotide sequence of a cDNA clone from *Medico sativa*. Nuc. Acid. Res., **17**, 3307.

Dickerson, R.E., Geis, I. (1983) Hemoglobin: Structure, Function, Evolution and Pathology. Bejamin/Cummings. Menlo Park, California.

Dolferus, R., Marbaix, G., Jacobs, M. (1985) Alcohol dehydrogenase in arabidopsis: Analysis of the induction phenonomen in plantlets and tissue cultures MGG, **199**, 256-264.

Dolferus, R., Jacobs, M., Peacock, W.J., Dennis, E.S. (1994) Differential interactions of promoter elements in the stress response of the *Arabidopsis* alcohol dehydrogenase gene. Plant Phys., **105**, 1075-1087.

Dolferus, R., Ellis, M., de Bruxelles, G., Trevaskis, B., Hoeren, F., Dennis, E.S., Peacock, W.J. (1997) Strategies of gene action in Arabidopsis during hypoxia. Annals of Botany, (supplement A), **79**, 21-31.

Drew, M.C. (1990) Sensing Soil Oxygen. **13**, 681-693.

Drew, M.C. (1997) Oxygen defficiency and root metabolism: injury and acclimation under hypoxia and anoxia. Ann. Rev. Plant. Phys., **48**, 223-250.

Duff, A.M., Wittenberg, J.B., Hill, R.D. (1997) Expression, purification and properties of recombinant barley hemoglobin: optical spectra and reactions with gaseous ligands. J. Biol. Chem., **In press**.

Elledge, S.J., Mulligan, J.T., Ramer, S.W., Spottswood, M., Davis, R.W. (1991) l-Yes: a multifunctional cDNA expression vector for the isolation of genes by complementation of yeast and E. coli mutations. PNAS, **88**, 1731-1735.

Ermler, U., Siddiqui, R.A., Cramm, R., Friedrich, B. (1995) Crystal Structure of the flavohaemoglobin from Alcaligenes eutrophus at a 1,75Å resolution. EMBO, **14**, 6067-6077.

Fennoy, S.F., Bailey-Serres, J.B. (1995) Post transcriptional regulation of gene expression in oxygen deprived roots of maize. Plant Journal, **7**, 287-295.

Fleming, A.I., Wittenberg, J.B., Wittenberg, B.A., Dudman, W.F., Appleby, C.A. (1987) The purification, characterisation and ligand binding kinetics of hemoglobins form root nodules of the non-leguminous *Casuarina-frankia* Symbiosis Biochemica et Biophysica Acta, **911**, 209-220.

Frelberg, C., Fallay, R.,Balroch, A., Broughton, W.J. Rosenthal, A., Perret, X. (1997) Molecular basis of symbiosis between Rhizobium and legumes. Nature, **387**, 394-400.

Fruhling, M., Roussel, H., Gianinazzi-Pearson, V., Puhler, A., Perlick, A.M. (1997) The *Vicia faba* leghemoglobin gene *Vflb29* is induced in root nodules and in roots colonised by the Arbuscular mycorrhizal fungus *Glomus fasciculatum*. Molecular Plant-Microbe Interactions, **10**, 124-131.

Gerlach, W.L., Pryor, A,J., Dennis, E.S., Ferl, R.J., Sachs, M.M. (1982) cDNA cloning and induction of alcohol dehydrogenase gene of maize. PNAS **79**, 2891-2895.

Gianinazzi-Pearson, V. (1996) Plant cell responses to Arbuscular Mycorrhizal fungi: getting to the roots of symbioses. The Plant Cell, **8**, 1871-1883.

Gibson, Q., Wittenberg, J.B., Wittenberg, B.A., Bogusz, D., Appleby, C.A. (1989) The Kinetics of ligand binding to plant haemoglobins. J. Biochem., **264**, 100-107.

Gilles-Gonzales, M., Ditta, G.S., Helinski, D.R. (1991) A Haemoprotein with kinase activity encoded by the oxygen sensor of *Rhizobium Meliloti.* Nature, **350**, 170-172.

Goldberg, D.E. (1995) The enigmatic oxygen avid haemoglobin of *Ascaris.* Bioessays, **17**, 177-182.

Hardison, R., C (1996) A brief history of hemoglobins: Plant, animal, protist, and bacteria. PNAS, **93**, 5675-5679.

Hargrove, M.S., Barry, J.K., Brucker, E.A. Phillips G.N., Olson, J.S. , Arredondo-Peter, R., Klucas, R.V., Sarath, G. (1997) Characterisation of recombinant soybean leghemoglobin a and apolar distal histidine mutant. JMB., **266**, 1032-1042.

Holmberg, N., Lillius, G., Bailey, J.E., Bulow, L. (1997) Trasngenic tobacco expressing *Vitreoscilla* hemoglobin exhibits enhanced growth and altered metabolite production. Nature Biotech., **15**, 244-247.

Hyldig-Nielsen, J.J., Jensen, E.L., Paludan, K., Wiborg, O., Garret, R., Jorgensen, P., Marcker, K.A. (1982) The primary structure of two leghemoglobin genes from soybean. Nuc. Acid. Res., **10**, 689-701.

Iwaasa, H., Takagi, T., Shikama, K. (1989) Protozoan myoglobin from *Paramecium caudatum.* J. Mol. Biol., **208**, 355-358.

Jacobsen-Lyon, k., Jensen, E.Ø., Jorgensen, J., Marcker, K.A., Peacock, W.J., Dennis, E.S. (1995) Symbiotic and Non-symbiotic Hemoglobin genes of *Casuarina glauca.* The Plant Cell, **7**, 213-223.

Jefferson, R.A. (1987) Assaying chimeric genes in plants: the GUS gene fusion system. Plant Mol. Biol. Rep., **5**, 387-405.

Jensen, E.O., Paludan, K., Hyldig-Nielsen, J.J., Jorgensen, P., Marcker, K.A. (1981) The structure of a chromosomal leghaemoglobin gene from soybean. Nature **291**, 677-679.

Ji, L., Becana, M., Sarath, G., Klucas, R., V (1994) Cloning and sequence analysis of a cDNA encoding ferric leghemoglobin reductase from soybean nodules. Plant Phys., **104**, 453-459.

Kennedy, R.A., Rumpho, M.E., Fox, T.C. (1992) Anaerobic metabolism in plants. Plant Phys., **100**, 1-6.

Kim, S.-R., Costa, M.A., Gyheung, A. (1991) Sugar response element enhances wound response of potato proteinase inhibitor II in transgenic tobacco. PMB, **17**, 973-983.

Koch, K.E. (1996) Carbohydrate modulated gene expression in plants. Ann. Rev. Plant. Phys., **47**, 509-540.

Konieczny, A.E., Jensen, E.O., Marcker, K.A., Legocki, A.B. (1987) Molecular cloning of lupin leghemoglobin cDNA. Plant Mol. Biol. Rep., **12**, 61-68.

Kuma, F. (1981) Properties of methemoglobin reductase and kinetic study of methemoglobin reduction. J. Biol. Chem., **256**, 5518-5523.

Landsmann, J., Dennis, E.S., Higgins, T.J., Appleby, C.A., Kortt, A.A., Peacock, W.J. (1986) Common evolutionary origin of legume and non-legume plant haemoglobins. Nature, **324**, 166-168.

Landsmann, J., Llewellyn, D., Dennis, E.S., Peacock, W.J. (1988) Organ regulated expression of the *Parasponia andersonii* haemoglobin gene in transgenic tobacco plants. MGG, **214**, 68-73.

Lehtovaara, p., Ellfolk, N. (1975) The Amino Acid sequence of Leghemoglobin component a from *Phaseolus vulgaris* (Kidney Bean). Jornal of Biochemistry, **54**, 577-584.

Liu, X.J., Prat, S., Willmitzer, L., Frommer, W.B. (1990) *Cis* regulatory elements directing tuber specific and sucrose expression of a chimeric class 1 patatin promoter/GUS gene fusion. MGG, **223**, 401-406.

Logemann, J., Schell, J., Willmitzer, L. (1987) Improved method for the isolation of RNA from plant tissues. Anal. Biochem, **163**, 16-20.

Long, S.R. (1996) Rhizobium Symbiosis: Nod factors in perspective. The Plant Cell, **8,** 1885-1898.

Martin, T., Frommer, W.B., Salanoubat, M., Willmitzer, L. (1993) Expression of an *Arabidopsis* sucrose synthase gene indicates a role in phloem loading and in sink organs. Plant Journal, **4**, 367-377.

Metz, B.A., Welters, B.A., Hoffman, H.J., Jensen, E.O., Schell, J., de Bruijn, F.J. (1988) Primary structure and promoter analysis of leghemoglobin genes of the stem nodulated tropical legume *Sesbania rostrata* : conseerved coding sequences, cis-elements and trans-acting factors. MGG, **214**, 181-191.

Millar, A.H., Bergerson, F.J., Day, D.A. (1994) Oxygen affinity of terminal oxidases in plant mitochondria. Plant Physiol. Biochem, **32**, 847-852.

Moens, L., Vanfleteren, J., Van de Peer, Y., Peeters, K., Kapp, O., Czeluzniak, J., Goodman, G., Blaxter, M., Vinogradov, S. (1996) Globins in nonvertebrate species: dispersal by horizontal gene transfer and the evolution of structure-function relationships. Molecular biology and evolution, **13**, 324-333.

Newcomb, W., Wood, S.M. (1987) Morphogenisis and fine structure of *Frankia* (Actinomycetales): the microsymbiont of nitrogen fixing actinorhizal root nodules. International review of cytology., **109**, 1-88.

Newman, T., de Bruijn, F.J., Green, P., Keegstra, K., Kende, H., McIntosh, L., Ohlrogge, J., Raikhel, N., Somerville, S., Tomashow, M. Retzel, E., Sommerville, C. (1994) "Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous *Arabidopsis* cDNA clones. Plant Physiol. **106**, 1241-1255.

Nie, X., Hill, R. (1997) Mitochondrial respiration and hemoglobin gene expression in barley aleurone tissue. Plant Phys., **114**, 835-840.

Olive, M.R., J, W.W., Singh, K., Peacock, W.J., Dennis, E.S. (1990) Functional properties of the anaerobic response element of the maize *Adh1* gene. PMB, **15**, 593-604.

Ollis, D., Appleby, C.A., Colman, P.M., Cutten, A.E., Guss, J.M., Venkatappa, Freeman, H.C. (1983) Crystal structure of soybean ferric leghaemoglobin A nicotinate at a resolution of 3.3Å. Australian Journal of Chemistry, **36**, 451-468.

Osmulski, P.A., Leyko, W. (1986) Structure function and physiological role of Chironomus hemoglobin. Comparative Biochemistry and Physiology., **85B**, 701-722.

Pathirana, S., Tjepkema, J.D. (1995) Purification of Haemoglobin from the actinorhizal root nodules of *Myrica gale.* Plant Phys., **107**, 827-831.

Paul, A.-L., Ferl, R.J. (1991) In vivo footprinting reveals unique cis-elements and differential modes of hypoxic inducton in maize *Adh1* and *Adh2*. The Plant Cell, **3**, 159-168.

Pawlowski, K., Bisseling, T. (1996) Rhizobial and Actinorhizal symbioses: What are the shared features? The Plant Cell, **8**, 1899-1913.

Perutz, M.F. (1978) Hemoglobin structure and respiratory transport. Scientific American, **239**, 92-125.

Potts, M., Angeloni, S.V., Ebel, R.E., Bassam, D. (1992) Myoglobin in Cyanobacterium. Science, **256**, 1690-1691.

Ramlov, K.B., Laursen, N.B., Stougaard, J., Marcker, K.A. (1993) Site directed mutagenisis of the organ specific element in the leghaemoglobin lbc3 gene promoter. Plant Journal, **4**, 577-580.

Riggs, A.F. (1991) Aspects of the origin and evolution of non-vertebrate hemoglobins. Amer. Zool., **31**, 535-545.

Roberts, P.M., Jafar, S., Mullin, B.C. (1985) Leghemoglobin-like sequences in the DNA of four actinorhizal plants. PMB, **5**, 333-337.

Robson, R.L., Postgate, J.R. (1980) Oxygen and hydrogen in biological nitrogen fixation. Ann. Rev. Micro., **34**, 183-207.

Roche, B.G., Temple, S.J., Segupta-Gopalan, C. (1993) Two classes of differentially regulated glutamine synthase genes expressed in the soybean nodule: a nodule specific class and a constituitively expressed class. PMB, **22**, 971-983.

Sachs, M.M., Freeling, M., Okimoto, R. (1980) The anaerobic proteins of maize. Cell, **20**, 761-767.

Sachs, M.M., Subbaiah, C.C., Saab, N. (1996) Annaerobic gene expression and flooding tolerance in maize. J. Exp. Bot., **47**, 1-15.

Sandal, N.N., Bojsen, K., Marcker, K.A. (1987) A small family of nodule specific genes from Soybean. Nuc. Acid. Res., **15**, 1507-1519.

Sarah, C.J., Graham, I.A., Reynolds, S.J., Leaver, C.J., Smith, S.M. (1996) Distinct cis-acting elements direct the germination and sugar responses of the cucumber malate synthase gene. MGG, **250**, 153-161.

She, Q., Sandal, N.N., J, S., Marcker, K.A. (1993) Comparative sequence anlysis of cis elements present in *Glycine max* l leghemoglobin and lbc3 genes. Plant Molecular Biology., **22**, 931-953.

Sherman, D.R., Guinn, B., Perdok, M.M., Goldberg, D.E. (1992) Components of sterol biosynthesis assemble on the oxygen avid haemoglobin of *Ascaris.* Science, **258**, 1930-1932.

Sieburth, L.E., Meyerowitz, E.M. (1997) Molecular dissection of the AGAMOUS control region shows that *cis* elements for spatial regulation are located intragenically. The Plant Cell, **9**, 355-365.

Soltis, D.E., Soltis, P.S., Morgan, D.R., Swensen, S., M, Mullin, B.C., Dowd, J.M., Martin, P.G. (1995) Chloroplast gene sequence data suggests a single origin of the predisposition for symbiotic nitrogen fixation in angiosperms. PNAS, **92**, 2647-2651.

Springer, B.A., Sligar, S. (1987) High level expression of sperm whale myoglobin in *Esherichia coli.* PNAS, **84**, 8961-8965.

Springer, B.A., Sligar, S.G., Olson, J., S, Phillips, G.N.J. (1994) Mechanism of ligand recognition in myoglobin. Chemical reviews, **94**, 699-714.

Stoltzfus, A., Doolittle, W.F. (1993) Slippery introns and globin gene evolution. Current Biology, **3**, 215-217.

Stougaard, J., Sandal, N.N., Gron, A., Kuhle, A., Marcker, K.A. (1987) 5' analysis of the soybean leghemoglobin lbc3 gene: regulatory elements required for promoter activity and organ specificity. EMBO, **6**, 3565-3569.

Stryer, L. (1988) Biochemistry. W.H. Freeman and company, New York,

Suharjo, U.K., Tjepkema, J.D. (1995) Occurrence of hemoglobin in the nitrogen fixing root nodules of *Alnus glutinosa*. Physiologica Plantarium, **95**, 247-252.

Swensen, S. (1996) The evolution of actinorhizal symbioses: Evidence for multiple origins of the symbiotic association. Am. J. Bot., **83**, 1503-1512.

Szczyglowski, K., Fujimoto, S., Szabados, L., Silver, D., de Bruijn, F.J. (1994) Site directed mutagenisis of the nodule infected cell expression (NICE) element and the AT rich element ATRE-BS2 of the *Sesbania rostrata* leghemoglobin glb3 promoter. The Plant Cell, **6**, 317-332.

Szczyglowski, K., Potter, T., Stoltzfus, J., Fujimoto, S., de Bruijn, F.J. (1996) Differential expression of the *Sesbania rostrata* mleghemoglobin

glb3 gene promoter in transgenic legume and non-legume plants. PMB, **31**, 931-935.

Taylor, C.B. (1997) Promoter fusion analysis: an insufficient measure of gene expression. The Plant Cell, **9**, 273-275.

Taylor, E.R., Nie, X.Z., MacGregor, A.W., Hill, R.D. (1994) A cereal haemoglobin gene is expressed in seed and root tissues under anaerobic conditions. PMB, **24**, 853-862.

Terwilleger, R.C., Terwilleger, N.B. (1985) Molluscan hemoglobins. Comparative Biochemistry and Physiology, **81B**, 255-261.

Thummler, F., Verma, D.P.S. (1987) Nodulin 100 of soybean is the subunit of sucrose synthase that is regulated by free heme availability. J. Biol. Chem., **262**, 14730-14736.

Trinick, M.J. (1973) Symbiosis between *Rhizobium* and the non-legume, *Trema aspera.* Nature, **244**, 459-460.

Trinick, M.J., Goodchild, D.J., Miller, C. (1989) Localisation of bacteria and hemoglobin in root nodules of *Parasponia andersonii* containing both *Bradyrhizobium* strains and *Rhizobium leguminosarum* bivar *trifolii.* Appl. Environ. Micorbiol., **55**, 2046-2055.

Valvekans, D., Van Montagu, M., Van Lijsebettens, M. (1988) *Agrobacterium tumefaciens* mediated transformation of *Arabidopsis thaliana* root explants by using kanamyacin selection. PNAS, **85**, 5536-5540.

Vasudevan, S.G., Amarego, W.L., Shaw, D.C., Lilley, P.E., Dixon, N.E., Poole, R.K. (1991) Isolation and nucleotide sequence of the *hmp* gene that encodes a haemoglobin-like protein in *Escheria coli* K-12. MGG, **226**, 49-58.

Vayda, M.E., Shewmaker, C.K., Morelli, J.K. (1995) Translational arrest in hypoxic potato tubers is correlated with the aberrant associaton of elongation factor EF1-alpha with polysomes. PMB, **28**, 751-757.

Wakabayashi, S., Matsubara, H., Webster, D.A. (1986) Primary sequence of a dimeric bacterial hemoglobin from Vitreoscilla. Nature, **322**, 481-483.

Walker, J.C., Howard, E.A., Dennis, E.S., Peacock, W.J. (1987) DNA sequences requiured for anaerobic expression of the maize *Adh1* gene. PNAS, **84**, 6624-6629.

Wiborg, O., Jorgen, J., Jensen, E.Ø., Paludan, K., Marcker, K.A. (1982) The nucleotide sequeces of two leghemoglobin genes from soybean. Nuleic Acids Research, **10**, 3487-3491.

Wittenberg, J.B., Bergersen, F.J., Appleby, C.A., Turner, G.L. (1974) The kinetics of the reactions of leghemoglobins with oxygen and carbon monoxide. J. Biol. Chem., **249**, 4057-4066.

Wittenberg, J.B., Wittenberg, B.A. (1989) Transport of oxygen in muscle. Annual review of physiology, **51**, 857-878.

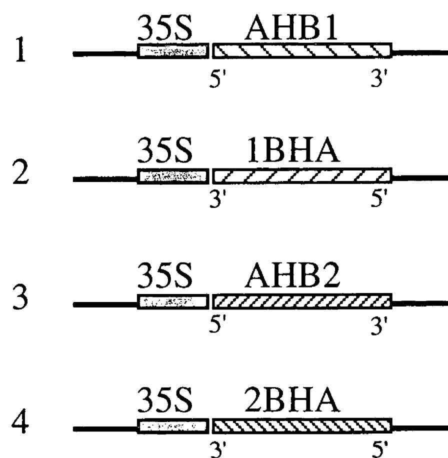
Yang, Y., Kwon, H.B., Peng, H.P., Shih, M.C. (1993) Stress response and metabolic regulation of glyceraldehyde-3-phosphate dehydrogenase genes in *Arabidopsis*. Plant Phys., **101**, 209-216.

Zhu, H., Riggs, A.F. (1992) Yeast flavohemoglobin is an ancient protein related to globins and a reductase family. PNAS, **89**, 5015-5019.

Appendix.

Several constructs are currently being transfomed into *Arabidopsis* plants. The reasons for the design of these constructs have been detailed in the main body of this thesis. The maps of the constructs and other technical details are presented here.

Over-expression and Anti-sense constructs.



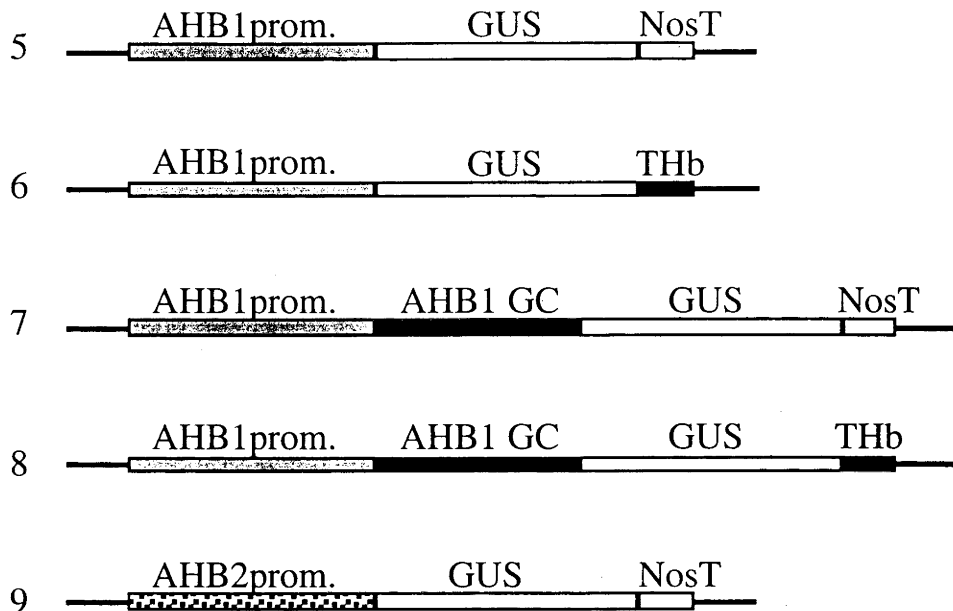
1. OxHB1. Over-expression construct for *AHB1*. *AHB1* cDNA cassette ligated in front of the 35S promoter (BamH1) in the plasmid pJ35SN. A SalI/ScaI fragment was then ligated into pBIN19 (SalI/SmaI). The resulting construct has been electroplorated into *Agrobacterium* and is currently being transformed into *Arabidopsis*.

2. AxHB1. Anti-sense construct for *AHB1*. *AHB1* cDNA cassette ligated in front of the 35S promoter (BamH1) in the plasmid PJ35SN. A SalI/ScaI fragment was then ligated into pBIN19 (SalI/SmaI). The resulting construct has been electroporated into *Agrobacterium* and is currently being transformed into *Arabidopsis*.

3. OxHB2. Over-expression construct for *AHB2*. *AHB2* cDNA cassette ligated in front of the 35S promoter as per OxHB1. Is currently being transformed into *Arabidopsis*.

4. **AxHB2.** Anti-sense construct for *AHB2*. *AHB2* cDNA cassette ligated in front of the 35S promoter as per AxHB1. Is currently being transformed into *Arabidopsis*.

Promoter-reporter gene fusions.



5. GH1. 1kb of *AHB1* promoter sequence fused (transcriptional fusion) to the B-gus reporter gene in the pBI 101.1 plasmid (Clontech). See Chapter 5.

6. GH1T12. GH1 with the Nos terminator replaced by 300bp of the 3' untranslated sequence from *AHB1* (THb, bases 742-1042). The THb fragment was PCR amplified from the Ara111 plasmid with EcoRI and SacI linker sites. Currently being transformed into *Agrobacterium* .

7. GH1C+. The promoter (1 kb) and entire genomic coding region (not including stop codon) of *AHB1* translationally fused to the B-Gus sequence in pBI 101.1. Currently being transformed into *Agrobacterium* .

8. GH1C+T12. GH1C+ with the Nos terminator replaced by THb. Currently being transformed into *Agrobacterium* .

9. GH2. 1.2 kb of the AHB2 promoter fused (transcriptional fusion) to the B-gus reporter gene in the pBI101.1 plasmid (Clontech). Is currently being transformed into *Arabidopsis*.