ELUCIDATION OF A PERIBACTEROID MEMBRANE-BOUND bHLH TRANSCRIPTION FACTOR REQUIRED FOR LEGUME NITROGEN FIXATION

A THESIS SUBMITTED BY

PATRICK CHARLES LOUGHLIN

SCHOOL OF AGRICULTURE, FOOD AND WINE

UNIVERSITY OF ADELAIDE

NOVEMBER 2007

4. Genetic Transformation of Soybean and Silencing of GmSAT1 and GmNOD26

4.1 INTRODUCTION

4.1.1 GENETIC TRANSFORMATION OF PLANTS

Genetic manipulation of an organism commonly requires the ability to efficiently introduce foreign genetic material into its genome. Many plant transformation techniques are available to do this, allowing for both stable and transient transformation. Stable, heritable genetic transformation is often the most desirable outcome. Many plant species however, do not regenerate efficiently from transformed callus tissue and are not readily amenable to transformation using, for example, the floral dip method commonly used in *Arabidopsis* (Clough and Bent, 1998). Particle bombardment of embryonic tissue (Christou et al., 1989) and *Agrobacterium tumefaciens*-mediated transformation of meristematic explants (Somers et al., 2003) and subsequent whole plant regeneration have proven effective in a number of legume species, including soybean, *Lotus japonicus* and *Medicago truncatula,* however these techniques are relatively inefficient, time consuming and require considerable training and plant cell culture resources. Conversely, *Agrobacterium rhizogenes*-mediated plant transformation offers a quick, efficient means of obtaining transformed plant material. *A. rhizogenes* infection of plants produces a callus from which so-called 'hairy roots' emerge and, similar to *A. tumefaciens*, the bacteria have the ability to randomly insert DNA flanked by left and right T-DNA border sequences into the genome of the infected plant. This transformation method has been used effectively to generate composite plants, having wild-type aerial tissue and transformed roots,

allowing the examination of root specific events including nematode infection, mycorrhizal development and nodulation (Cho et al., 2000; Boisson-Dernier et al., 2001).

4.1.2 SELECTION OF TRANSFORMED PLANT TISSUE

Each root emerging from a callus induced by an *A. rhizogenes* infection is potentially a product of an independent transformation event and problems such as so called chimeric roots (half transformed and half untransformed) have been observed in *M. truncatula* (Limpens et al., 2004), making careful selection of transformed tissue essential. A number of different detection methods have been utilised in the past, including the use of selectable herbicide markers, selection using fluorescent proteins, detection of specific enzymatic activities e.g., GUS, or direct examination of the genomic DNA for the T-DNA insertion. Herbicide markers have been effectively used in *M. truncatula* hairy root transformation with no apparent effects associated with having both transformed (resistant) roots and untransformed (susceptible) aerial organs (e.g., Boisson-Dernier et al., 2001). Perhaps the simplest and most efficient selectable marker is a fluorescent protein such as the green fluorescent protein GFP from *Aequorea victoria*, whereby potentially transformed roots may be examined non-destructively under UV light for green fluorescence (Hraska et al., 2006). Unfortunately the roots of certain plant species, including *Nicotiana tabacum* (Molinier et al., 2000) and soybean (Cho et al., 2000) autofluoresce green under UV light making it difficult to observe GFP fluorescence. ß-glucoronidase (GUS) is another common reporter gene for plant transformation, the activity of which may be examined histologically or quantified using a simple fluorometric assay (Jefferson et

al., 1987). Direct analysis of the genomic DNA of potentially transformed roots is a destructive and relatively time consuming process, however this technique may be preferable in that the results are not dependent on expression of selectable marker genes (Lee et al., 1993; Schmidt et al., 2007).

4.1.3 GENE SILENCING IN PLANTS

Reverse genetics is a commonly used tool in determining gene function. In plants, genomic lesions or insertions may be introduced through irradiating seeds (Shikazono et al., 2005) or using transposon mutagenesis respectively (May and Martienssen, 2003). Large collections of *Arabidopsis* and rice T-DNA tagged insertion lines are readily available which cover much of their genomes and allow relatively easy selection of a line with a mutation in a gene of interest. Such collections of random integration mutants are a powerful tool, however the time and monetary costs associated with maintenance of such collections is prohibitive for non-model plant species such as soybean. Although there has been some work towards homologous recombination in plants as a means of gene-specific deletion (Li et al., 2007), other techniques including gene silencing have been put to extensive use, in both model and non-model plant species. Gene silencing is a phenomenon identified in all eukaryotic organisms whereby the presence of dsRNA in a cell leads to a sequence specific repression of gene expression. This repression may manifest at a number of levels, including genomic DNA methylation changes and protein translation, however the best characterised mode of action is through degradation of cognate endogenous mRNA (Small, 2007). Double stranded RNA may be produced in a plant by introducing an inverted repeat DNA sequence with an intron separating the repeats,

using a suitable transformation method. When transcribed, this construct generates a hairpin RNA structure (hpRNA) which efficiently activates silencing of the cognate gene through degradation of the native transcript (Waterhouse et al., 1998; Helliwell and Waterhouse, 2003). Antisense or dsRNA constructs have been used in conjunction with *A. rhizogenes*-mediated hairy root transformation to examine the effects of silencing of specific genes in the roots of *Vigna* spp., soybean, *Lotus japonicus*, *Medicago truncatula* and *Arabidopsis thaliana* (Gonzalez-Rizzo et al., 2006; Cheon et al., 1993; Lee et al., 1993; Karimi et al., 1999; Kumagai and Kouchi, 2003; Limpens et al., 2004; Wasson et al., 2006).

4.1.4 GmSAT1 AND OTHER NODULE-EXPRESSED GENES

In the soybean nodule, GmSAT1 is localised to the nucleus (Section 3.2.2) and peribacteroid membrane (Kaiser et al., 1998), the interface between nitrogen-fixing bacteroids and their soybean host, so is ideally situated to respond to symbiosis-related conditions. In this chapter, I sought to examine the functional importance of GmSAT1 through a reverse genetics approach. An efficient means of soybean transformation, using an adaptation of a previously published *A. rhizogenes*-mediated transformation technique (Boisson-Dernier et al., 2001) was exploited to transform soybean roots with RNAi silencing vectors specific to GmSAT1 and GmNOD26, a PBM localised aquaporin (Fortin et al., 1987). Silencing vectors specific to the PBM localised iron transporter GmDMT1 (Kaiser et al., 2003) and the nodule expressed ammonium transporter GmAMT1;3 (Kaiser, unpublished data) were also constructed however time constraints prevented their use in this study. Presented here are data describing the development of an *A. rhizogenes*-mediated soybean transformation

protocol using GUS-expression vectors from Cambia (www.cambia.org, Figure 4-1) and a description of the morphology of putatively GmSAT1- and GmNOD26-silenced nodules.

4.2 RESULTS

4.2.1 TRANSFORMED SOYBEAN ROOTS ARE EFFICIENTLY PRODUCED THROUGH INOCULATION OF RADICLE WOUND SITE WITH *A. rhizogenes* K599

A number of different *A. rhizogenes*-mediated hairy root transformation techniques were trialed in this study. Transformation of soybean roots by injection of *A. rhizogenes* at the cotyledonary node as described by Cheon et al., (1993) met with limited success. Attempts were also made to use the *ex-vitro* method described by Collier et al., (2005) where apical cuttings were taken and regenerated in 'oasis' medium soaked in *A. rhizogenes* however the survival rate was relatively low. Eventually a technique modified from Boisson-Dernier et al., (2001) was adopted, as outlined in section 4.4.3 and Figure 4-3. This involved the excision of the radicle tip (5 mm from the end) of germinated soybean seedlings and smearing the wound with *A. rhizogenes* K599 harbouring the construct of interest and placing the seedling on solid (0.7% (w/v) agar) minus nitrogen Herridge's medium as outlined in section 4.4.3. Boisson-Dernier et al., (2001) used 25 µg/ml kanamycin to select for transformed *M. truncatula* roots however attempts to use selection, either hygromycin for the pCAMBIA vectors (Figure 4-1) or kanamycin for the pHellsgate 8 vectors (Figure 4-2) met with little success. For example, hygromycin selection at 25µg/ml arrested growth of both pCAMBIA 1305.1 transformed and untransformed roots (Figure 4-4) and lower concentrations proved ineffective in preventing untransformed

root growth (data not shown). Despite these problems, careful visual selection of roots originating from the callus which developed at the base of transformed plants allowed for a relatively high selection of transformed roots (e.g., Figure 4-5B, GUS positive roots).

4.2.2 GUS EXPRESSION UNDER THE CaMV35S PROMOTER IS SYSTEMIC THROUGHOUT SOYBEAN ROOTS AND NODULES

Initially plants were transformed with the GUS expression vector pCAMBIA1305.1 (www.cambia.org) to verify the transformation of the roots and the efficacy of gene expression under the cauliflower mosaic virus (CaMV) 35S promoter (Figure 4-1A). Transformed plants were grown under low nitrogen conditions and inoculated with *Bradyrhizobium japonicum* USDA110 to induce nodulation. The morphology of transformed nodules appeared normal (Figure 4-5D), which is in contrast to the 15-20% of *A. rhizogenes*-transformed soybean nodules with indeterminate morphology observed by Bond and Gresshoff (1993). Initially tissue samples were vacuum infiltrated with GUS stain and incubated overnight at 37°C, however shorter staining periods (approximately 1 hour) and no vacuum infiltration were usually found sufficient to observe GUS staining. GUS was highly expressed throughout transformed roots under the CaMV35S promoter, apart from the outer cortex of the nodule excluding the lenticels associated with the nodule vasculature (Pankhurst and Sprent, 1975) Figure 4-5B-F). The outer cortex of the nodule was impermeant to the X-gluc substrate however sectioning of the nodule facilitated staining of the inner cortex (Figures 4-5E & F). Soybeans were also transformed with the GFP expression vector pCAMBIA 1302 and the GFP-GUS fusion expression vector pCAMBIA 1303 (Figure 4-1B & C). Unfortunately the soybean roots showed a high level of

autofluorescence under a GFP filter, which made it difficult to distinguish between transformed and untransformed roots (data not shown). This problem has previously been encountered in GFP-expressing tobacco roots (Molinier et al., 2000), however Schmidt et al., (2007) circumvented this problem in *Saponaria vaccaria* roots through narrow optical filtering of the green fluorescence. These results demonstrate that the CaMV35S promoter is active in all root and nodule tissue examined apart from the nodule outer cortex and that GUS may be used effectively to detect transformed soybean root and nodule tissue.

4.2.3 GUS EXPRESSION DRIVEN BY THE *Gmlbc3* PROMOTER IS PREDOMINANTLY RESTRICTED TO THE INFECTED INNER CORTEX OF THE NODULE

High expression, cell-specific promoters are useful tools when looking at overexpression or silencing of genes in particular cell types. Leghaemoglobins are haem-containing proteins involved in sequestering free oxygen in legume nodules, whilst allowing diffusion of the oxygen to the rapidly respiring bacteroids (Ott et al., 2005). Soybean have 4 symbiotic leghaemoglobin genes, *lba*, *lbc₁*, *lbc₂,* and *lbc₃*. The expression pattern of the *Gmlbc₃* promoter is the best characterised, with transcript detected early in the development of the nodule and high levels of expression continuing throughout maturity of the nodule (Marcker et al., 1984; Christensen et al., 1989). Previous analyses of the expression pattern of *Gmlbc₃* suggest expression is restricted to infected nodule cells (Nguyen et al., 1985; Franche et al., 1998). In this study, the 2kb genomic DNA region 5' of the leghaemoglobin start codon was cloned into the *Sal* I / *Eco* RI sites 5' of the promoter-less *gusA* gene in pCAMBIA 1391Z (Figure 4-1D). GUS was expressed at a high level under the *Gmlbc₃* promoter, and was generally restricted to the inner cortex of the soybean nodule (Figure 4-6).

Expression was detected from early in the development of the nodule primordium (Figure 4-6A) through to nodule maturity (Figure 4-6H). Weak GUS expression was also detected in the root vasculature immediately adjacent young nodules and root tips (Figure 4-6C, F & G). This may be due to weak activation of the *Gmlbc₃* promoter in these regions or in some cases may be an artefact of over-staining or transport of the glucuronidase enzyme out of the infected region of the young nodules. Nodules transformed with the empty pCAMBIA 1391Z, with a promoter-less *gusA* gene showed no GUS activity (data not shown).

4.2.4 VERIFICATION OF THE TRANSFORMATION OF SOYBEAN ROOTS WITH pHELLSGATE 8 SILENCING VECTORS

Several nodule enhanced and nodule specific soybean genes were chosen for RNAi silencing using the hpRNA silencing vector pHellsgate8 (Helliwell et al., 2002). GmSAT1 is expressed almost exclusively in the nodule, with some expression detectable in nodulated root tissue as well (Kaiser et al., 1998). Previous work, and work in this study, demonstrate GmSAT1 is localised to the peribacteroid membrane (Kaiser et al., 1998) and nucleus (Section 3.2.2) of infected cells in the soybean nodule. GmNOD26 is a well-characterised aquaporin highly expressed on the soybean PBM (Fortin et al., 1987; Weaver et al., 1994; Dean et al., 1999; Niemietz and Tyerman, 2000). pHellsgate8-silencing constructs using 3' non-coding gene fragments of the PBM-localised iron transporter GmDMT1 (Kaiser et al., 2003), the nodule enhanced ammonium transporter GmAMT1:3 (Kaiser, unpublished data), GmSAT1 and GmNOD26 were made as outlined in section 4.4.1 and Figure 4-2. Time constraints prevented the use of the GmDMT1 and GmAMT1;3 silencing vectors, however soybean roots were transformed with the pHellsgate8-GmSAT1 and

pHellsgate8-GmNOD26 vectors. Verification of putatively transformed roots was through PCR amplification of a 212bp fragment of the pHellsgate8-specific kanamycin resistance *nptII* gene from soybean genomic DNA. Control PCR reactions with primers designed across the left border of the T-DNA insert of pHellsgate 8 vectors were also used to verify the *nptII* PCR product was being amplified from DNA inserted in the soybean genome and not from, for example, residual *A. rhizogenes* harbouring the pHellsgate 8 construct (Figures 4-2 & 4-7).

4.2.5 GmSAT1 IS ESSENTIAL FOR NODULE DEVELOPMENT AND/OR MAINTENANCE

Soybean plants transformed with empty pHellsgate 8, pHellsgate8-GmSAT1 and pHellsgate8-GmNOD26 vectors were grown under low nitrogen conditions, and inoculated with *B. japonicum* USDA110 to induce nodulation. Plants with only roots transformed with pHellsgate8-GmSAT1 were severely chlorotic, indicative of nitrogen deficiency, when compared with control plants (Figure 4-8A). Other plants that had both untransformed and pHellsgate8-GmSAT1 transformed roots were not chlorotic, with healthy growth presumably being supported by fixed nitrogen supplied by untransformed nodules (data not shown). Examination of gross root morphology immediately identified a number of differences between vector controls and GmSAT1-silenced (*sat1*) roots. *sat1* roots had more nodules than control roots, however these nodules did not develop normally, being much smaller than vector transformed controls (Figure 4-8B). In some instances, phenotypically wild-type nodules developed on *sat1* roots (Figure 4-8D), possibly due to incomplete gene silencing, or chimeric root development although this was not examined further. Visual examination of the nodule inner cortical region demonstrated a small infected

zone, lacking the typical pink colour caused by leghaemoglobin in functional nodules
(Figure 4-8C). Infected cells were much smaller in *sat1* nodules and numerous
vacuoles remained (compare Figures 4-8E & 8F). Transmission electron microscopy
further elucidated these structural changes associated with silencing of GmSAT1. The
inner cortex of the *sat1* nodules contained rhizobia infected cells, however these cells
were much smaller than in a normal nodule (compare Figure 4-9A & 9B), being of
similar size to adjacent uninfected cells. The large central vacuole of the infected cells
of *sat1* nodules often remained, but in other cases was subdivided into smaller
vacuoles while the number of symbiosomes was also reduced. (Figure 4-9B to F).

4.2.6 SILENCING OF GmNOD26 IN SOYBEAN ROOTS

GmNOD26-silenced composite plants (*nod26*) were also generated in this study and
the nodule phenotype examined. GmNOD26 is a PBM associated aquaporin which
has been implicated in the transport of compounds including NH_3, malate, glycerol
and H_2O (Weaver et al., 1994; Dean et al., 1999; Niemietz and Tyerman, 2000).
Although *nod26* plants demonstrated some chlorosis (Figure 4-10A), they were never
as chlorotic as *sat1* plants and examination of the roots and nodules revealed an
essentially wild type phenotype (Figure 4-10B, 10C & 10D). Nonetheless, *nod26*
plants were distinct from vector transformed plants, and although the physiology of
these plants was not thoroughly examined, the silencing of GmNOD26 may have
been affecting malate (Weaver et al., 1994) or NH_3 (Niemietz and Tyerman, 2000)
transport across the PBM, causing the chlorosis observed. The weaker phenotype was
somewhat surprising, considering the high expression level of GmNOD26 on the
PBM of the soybean nodule. However the degree of gene silencing was not measured
in this study, so it may be that GmNOD26 silencing was inefficient or alternatively

that GmNOD26 is functionally redundant and other aquaporins can compensate for the reduced GmNOD26 expression in the soybean nodule.

4.3 DISCUSSION

The study of phenotypically mutant plants and the genetics behind these phenotypes has a long history going back to the earliest genetic studies of pea plants by Gregor Mendel during the mid-19th century. Although the tools used to study mutant phenotypes and to generate them have advanced considerably since then, the general principle of analytical examination and interpretation of these results remains. The development of various technologies has allowed a shift from the study of naturally occurring mutant plant forms to random mutagenesis through irradiation or other mutagens, and directed mutagenesis through RNA interference as described in this chapter.

4.3.1 DEVELOPMENT OF AN EFFICIENT SOYBEAN TRANSFORMATION PROTOCOL

One of the major hurdles encountered in the study of soybean genetics is the lack of an efficient transformation protocol to generate large numbers of stably transformed whole plants. The relative ease of transformation of *Arabidopsis* has allowed exploitation of numerous molecular biology techniques to examine gene expression patterns, both temporally and spatially and the effects of over-expression and silencing of these genes. The worldwide agricultural importance of soybean has led to intense research into numerous techniques to stably transform soybean including *Agrobacterium*-mediated and direct DNA transfer transformation techniques

(Christou et al., 1989; Somers et al., 2003). Although these protocols are effective, they are also generally inefficient and time consuming, often requiring > 6 months for plant regeneration and specialised plant cell culture techniques. Although it is possible to regenerate a whole plant from an *A. rhizogenes*-derived hairy root culture in some plant species (Cho et al., 1998; Crane et al., 2006), composite plants as described in this chapter are commonly used for research purposes.

Three different protocols for genetically transforming soybean using *A. rhizogenes* were trialed in this study; injection of *A. rhizogenes* into the coteledonary node of young plants (Cheon et al., 1993), transformation of apical cuttings from young soybean plants (Collier et al., 2005), and excision of the radicle tip and inoculation of the wound site with *A. rhizogenes* (Boisson-Dernier et al., 2001). Although all three techniques led to the generation of transformed roots (data not shown), an adaptation of the technique described by Boisson-Dernier et al. (2001) was chosen as the most efficient. The major hurdle encountered in this technique was the selection of transformed roots. In *M. truncatula*, Boisson-Dernier et al., 2001 found 25 µg/ml kanamycin was a sufficient selection pressure to increase the percentage of GUS positive (transformed) roots from 59 to 95%. Both hygromycin and kanamycin selection for hairy roots transformed with pCAMBIA and pHellsgate8 vectors respectively were examined, albeit briefly, in this study. Kanamycin and hygromycin concentrations ranging from 50-400 µg/ml and 2.5-40 µg/ml respectively were used in an attempt to select for the growth of transformed roots (data not shown). In both instances, lower concentrations of the selective agent did not affect root growth and at higher concentrations, all root growth, even that of transformed tissue, was suppressed (data not shown). For transformed soybean callus propagation, selection of

transformants using 50 µg/ml kanamycin is not begun until at least 2 weeks after transformation (Dhir et al., 1992), so potentially an intermediate antibiotic concentration or a change in the timing of its usage may have facilitated selection of transformed roots, however time constraints prevented further examination of this. Roots emerging from calli at the base of the soybean plants were considered putatively transformed. This was then verified by GUS staining or PCR analysis of extracted genomic DNA for pCAMBIA and pHellsgate8 vectors respectively.

4.3.2 BOTH THE CaMV35S AND *Gmlbc3* PROMOTERS ARE ACTIVE IN *A. rhizogenes* TRANSFORMED SOYBEAN TISSUE

Generally GUS expression under the CaMV35S promoter in root and nodule tissue transformed with pCAMBIA 1305.1 was very high and systemic. An exception to this was the outer cortex of the mature nodule, which did not appear to have any GUS expression excluding the lenticels associated with the nodule vasculature (Figure 4-5D). Contrary to reports of low GUS expression in the infected region of the soybean nodule (Bond and Gresshoff 1993), in this study the infected region of immature nodules stained well under both the CaMV35S (data not shown) and *Gmlbc$_3$* promoters (Figure 4-6) and the inner cortex of mature nodules stained well when dissected (Figure 4-5F). The staining and microscopy techniques used in this study were not sensitive enough to determine whether GUS expression under the *Gmlbc$_3$* was restricted to rhizobia-infected cells or was expressed throughout the entire infected region.

4.3.3 SILENCING OF GmSAT1 AND GmNOD26 HAVE DIFFERENT EFFECTS ON THE

SOYBEAN-*BRADYRHIZOBUM* SYMBIOSIS

Having an efficient transformation protocol available and demonstrating essentially systemic gene expression under the CaMV35S promoter allowed the use of pHellsgate 8 vectors to examine the effect of silencing or knocking down of GmSAT1 and GmNOD26 expression. Although the level of silencing of the GmSAT1 transcript was not determined in this study, obvious phenotypic differences were apparent in roots transformed with pHellsgate8-GmSAT1 when compared with roots transformed with an empty pHellsgate8 vector or untransformed roots. In contrast, differences between pHellsgate8-GmNOD26 transformed and vector transformed roots and nodules were less apparent. Despite this, *nod26* plants were distinct from vector transformed plants, suggesting GmNOD26 may have a supportive role in the symbiosis, perhaps, as has been suggested in isolated symbiosome and lipid bi-layer studies, catalysing the translocation of various nutrient and/or H_2O across the PBM. Future work will look at examining the PBM of *nod26* nodules using a patch clamp technique to determine if the GmNOD26-deficient PBM is electrophysiologically altered when compared with the wild-type PBM.

GmSAT1-silenced roots (*sat1*) had a disproportionately high number of small nodules and these nodules had abnormal structural features (Figure 4-8 & 4-9). The fact that these nodules are ineffective is probably preventing the inhibitory signal which would normally prevent the development of such a high number of nodules and is unlikely to be related to a defect in the autoregulation of nodule development as observed in soybean with mutations of the receptor-like kinase GmNARK (<u>n</u>odule <u>a</u>utoregulation <u>r</u>eceptor <u>k</u>inase (Searle et al., 2003). The pink colour associated with leghaemoglobin and effective nodules was absent in the infected inner cortex of *sat1* nodules and

many of the *sat1* plants exhibited varying degrees of chlorosis. As bacteroid-derived nitrogen was the only nitrogen source available to the plant, silencing of GmSAT1 is either affecting the ability of the plant to access bacteroid fixed nitrogen or affects the ability of the bacteroids to fix the nitrogen in the first place. Microscopic analysis of the nodules suggests the latter is the case, as there are far fewer bacteroids present in *sat1* nodules (Figure 4-9). In contrast to wild-type infected cells, large central vacuoles which usually breakdown during the formation of the infected cell remained and infected cells failed to expand in size. Previous work in this thesis (Chapters 2 and 3) demonstrates GmSAT1 probably acts as a membrane bound transcription factor. Taken together with the phenotype of *sat1* nodules presented in this chapter, a number of potential roles for GmSAT1 in the soybean nodule are possible including its involvement in signalling the break down of the vacuole, maintaining increased synthesis and correct folding of proteins destined for the PBM or the regulation of the delivery of protein to the PBM. The *Arabidopsis* bHLH transcription factor AtNAI1 (At2g22770) is one of three GmSAT1-like *Arabidopsis* genes and is the best characterised. AtNAI1 T-DNA knock-out plants are phenotypically indistinct from wild type plants (Matsushima et al., 2004), however they do lack endoplasmic reticulum structures termed 'ER bodies,' which are spindle-shaped structures, typically 10 μm long and 1μm wide (Matsushima et al., 2004). These ER bodies are observed in young plant tissue including the cotyledons, hypocotyls and roots of seedlings however they can be induced in mature rosette leaves by wounding or application of the wound-response plant hormone methyl jasmonate (MeJA) (Matsushima et al., 2002). Transcription of AtNAI1 is also induced by the application of MeJA (Matsushima et al., 2004). It is tempting to speculate on a role for GmSAT1 that is similar to that of AtNAI1, with the knowledge of the close association between

the ER and PBM development in nodules. On the one hand, the *Arabidopsis nai1* mutant does not form ER bodies, apart from aberrant ER-body-like structures which only formed when rosette leaves were exposed to MeJA (Matsushima et al., 2004). On the other hand, soybean *sat1* nodules were also aberrant in their endomembrane structures, with the synthesis and/or maintenance of PBM being compromised and the central vacuole remaining in rhizobia-infected cells. Potentially GmSAT1 has a role in the transcription of genes involved in the development of effective PBM, which requires the synthesis and correct trafficking of a unique complement of proteins from the ER and Golgi body. Alternatively GmSAT1 may be involved in the maintenance of the symbiosis and that in the absence of GmSAT1 symbiosomes are degraded, leading to the re-formation of the central vacuole and reduced number of bacteroids in the infected cell.

4.4 METHODS

4.4.1 VECTOR CONSTRUCTION

Standard DNA manipulations were carried out essentially as described by Sambrook and Russell (2001). 2 kb of soybean genomic DNA 5' of the leghaemoglobin c gene (accession X15061) was PCR amplified (Gmlbc3promSalI_FW 5'-gccgtcgacg-attcacccgtatcgatcgc-3' and Gmlbc3promEcoRI_RV 5-gccgaattcatttcttttctacttttctg-3', restriction enzyme sites underlined) and ligated into the pGemT-easy vector (Promega). This leghaemoglobin c (lbc3) promoter was sequenced and then cloned into the *Sal* I and *Eco* RI sites of the pCAMBIA 1391Z vector, immediately 5' of *gusA* (Figure 4-1D).

Primers specific to the 3' UTRs of *GmSAT1* (accession AAC32828, 1113-1472 nt), *GmNOD26* (accession P08995, 900-1103 nt), *GmDMT1* (accession AAO39834, 1674-1839 nt) and *GmAMT1;3* (B. Kaiser, unpublished, 1453-1792 nt) were designed with *attB* sequences engineered at the 5' end of each (Table 1). Products were amplified from a nodule cDNA library (Kaiser et al, 1998) and recombined with the pDONR222 vector in a BP clonase reaction (Invitrogen). The 'pDONR-Gene Fragment of Interest' construct was then recombined with pHellsgate8 (Helliwell et al., 2002) using LR clonase (Invitrogen) to insert two copies of the gene fragments into pHellsgate8 in an inverted repeat orientation (Figure 4-2). Inserts were sequenced to verify no PCR derived errors had occurred.

4.4.2 PREPARATION AND TRANSFORMATION OF CHEMICALLY COMPETENT *A. rhizogenes* K599

Agrobacterium rhizogenes K599 cells were made competent using a freeze-thaw method modified from Holfgen and Wilmitzer, 1988. A 5 mL *A. rhizogenes* culture was grown overnight in YEM media (Mathis et al., 1986) at 30°C with shaking and used to inoculate 200 mL of fresh YEM media. After 3-4 hours of logarithmic growth, bacteria were spun down at 3,000 x g for 20 min at $4^{°}$C. The resulting pellet was washed once with 10 mL pre-cooled TE (10 mM Tris-HCl, pH7.5, 1 mM EDTA) and resuspended in 20 mL YEM. 500 µl aliquots were either used directly for transformation or frozen in liquid N_2 and stored at $-80^{°}$C.

Aliquots of competent *A. rhizogenes* were thawed on ice and 0.5-1.0 µg of plasmid DNA added. Cells were incubated on ice for a further 5 min then immersed in liquid

N$_2$ for 5 min and transferred to 37°C for 5 min. The transformation mix was diluted with 1.0 mL YEM, transferred to a 15 mL falcon tube and incubated for 2-4 hours at 30°C with shaking. 50-200 µl was plated on YEM plates supplemented with a suitable antibiotic for selection, and incubated at 30°C for 2 days. Colonies were restreaked and grown for a further 2 days before being used in the soybean transformation.

4.4.3 'HAIRY ROOT' *A. rhizogenes*-MEDIATED TRANSFORMATION OF SOYBEAN

Soybean (*Glycine max* L. cv. Boyer) were sown in river sand 5 cm deep. After 2 days, plants were recovered, rinsed 3x in dH$_2$O and radicles were excised (approx. 3 mm from the root tip). The wound site was smeared with *Agrobacterium rhizogenes* K599 harbouring the construct of interest, and the germinated seed placed on solid modified Herridge's (-N) media (500 µM MgSO$_4$, 500 µM KH$_2$PO$_4$, 500 µM K$_2$SO$_4$, 50 µM KCl, 20 µM Fe-EDTA, 250 µM CaSO$_4$, 25 µM H$_3$BO$_3$, 2 µM MnSO$_4$, 2 µM ZnSO$_4$, 0.5 µM CuSO$_4$, 0.5 µM NaMoO$_4$ plus 0.7% w/v agar) inside an enclosed clear plastic container to maintain humidity. Plants within the plastic containers were grown in temperature controlled growth chambers under a 16/8h day /night regime at 26°C. Photosynthetic active radiation (PAR) at container/pot level ranged between 150 – 400 µmole m^2 s^{-1}, where plants initially were provided low light (~100 PAR) up to 10 days post transformation then gradually increased to 400 PAR as plants developed hairy roots.

After 3-4 days, adventitious roots began to emerge along the hypocotyl and immediately adjacent to the wound site. These roots were excised and plants transferred to solid Herridge's (-N) supplemented with 100 µg/ml cefotaxime to kill

any remaining *A. rhizogenes*. After 2-3 days, plants were transferred to solid Herridge's (-N) plus selection as required. As discussed in section 3.2.1, no antibiotic selection was used and all roots not emerging from the developing callus at the base of the plant were removed during growth on the solid Herridge's (-N) medium. Plants rapidly out-grew their plant tissue culture containers and were transferred to river sand medium after approximately 2 weeks growth on solid Herridge's medium and inoculated with 10 ml of a saturated *Bradyrhizobium japonicum* USDA110 culture (grown for 3 days at 28°C in YEM medium (Mathis et al., 1986) to induce nodulation). Plants were watered every 2 to 3 days and fertilised with Herridge's (-N) nutrient solution once per week.

4.4.4 GENOMIC DNA ISOLATION AND TRANSFORMATION VERIFICATION

Genomic DNA was isolated from soybean roots using an alkaline lysis/boiling technique. 2-3 cm pieces of root were excised and placed into 1.5 mL microfuge tubes. Tissue samples were crushed using a teflon pestle and 40 µl 0.25 M NaOH added. Samples were heated to 100°C for 30 sec and then placed on ice. 40 µl 0.25 M HCl was added to neutralise the samples, and to this, 20 µl 0.5 M Tris, pH 8.0 / 0.25% (v/v) Triton X-100 was added. The genomic DNA samples were mixed well and heated to 100°C for 2 min, and put on ice. PCRs with 1.0 µl genomic DNA as template and primers specific to the *nptII* gene (NPTII_FW 5'-AGACAATCGGCTGCTCTGAT-3' and NPTII_RV 5'CAATAGCAGCCAGT CCCTTC-3') and across the left border of the T-DNA insert of pHellsgate 8 (Across_LBFW 5'-GCGCGGTGTCATCTATGTTA-3' and Across_LBRV 5'-

GGTGCGAATAAGGGACAGTG-3') were performed to verify transformation of the roots with the pHellsgate 8 vectors and insertion of the T-DNA.

4.4.5 GUS STAINING OF SOYBEAN ROOT AND NODULE TISSUE

Root and nodule tissue was excised from the soybean plants and placed in ice cold 100% acetone. Samples were then rinsed in GUS staining buffer (0.2% (v/v) Triton X-100, 2 mM potassium ferricyanide, 2 mM potassium ferrocyanide, 50 mM sodium phosphate buffer, pH 7.2) before being incubated at 37°C in GUS staining buffer plus 2 mM X-Gluc (5-bromo-4-chloro-3-indolyl-β-d-glucoronic acid). Depending on the strength of the signal, samples were incubated in the staining solution for 1-4 hours. For mature nodules, tissue was dissected prior to incubation in the X-Gluc staining solution to allow penetration of the stain into the inner cortex.

4.4.6 MICROSCOPY

Unless otherwise specified, all incubations were performed with gentle mixing at room temperature. Either whole nodules or small pieces of larger nodules (~2 mm^2) were fixed overnight at 4°C in fixative (4% v/v paraformaldehyde, 0.5% v/v glutaraldehyde and 4% w/v sucrose, made up in PBS (137 mM NaCl, 2.7 mM KCl, 4.3 mM Na$_2$HPO$_4$, 1.47 mM KH$_2$PO$_4$, pH7.3). Nodule specimens were washed 2 x 10 min in PBS plus 4% (w/v) sucrose before being dehydrated in an ethanol/water series of 70%, 90%, 95%, and 100% ethanol for 3 x 20 min each. Specimens were then incubated overnight in a 1:1 mix of ethanol : LR White resin (London Resin Company Ltd). The following morning specimens were transferred to LR White resin and incubated for approximately 36 hours, with 3 changes of the resin. Specimens were

transferred to small gelatin capsules filled with LR White resin and embedded at 50^oC for 1-3 days. Sections of either 1 µm or 90 nm were taken for light and transmission electron microscopy respectively. The 1 µm sections were stained with toluidine blue and examined using a Zeiss Axiophot light microscope. The 90 nm specimens placed onto nickel grids, stained for 5 min with 3% (v/v) uranyl acetate in 70% (v/v) ethanol and washed, followed by staining with lead citrate (Venable and Coggeshall, 1965) and examined using a Phillips CM100 transmission electron microscope.

NAME	SEQUENCE (5'→3')
GmSAT1HellsFW	**GGGGACAAGTTTGTACAAAAAAGCAGGCT**TCCCCGAAAGGTACATGAAG
GmSAT1HellsRV	**GGGGACCACTTTGTACAAGAAAGCTGGGT**CAACAAAGGCCTGTCTGTCA
GmNOD26HellsFW	**GGGGACAAGTTTGTACAAAAAAGCAGGCT**TGAGCAAATGGAGACACTCAA
GmNOD26HellsRV	**GGGGACCACTTTGTACAAGAAAGCTGGGT**TCCACCAACAATGTCAAAAATC
GmAMT1;3HellsFW	**GGGGACAAGTTTGTACAAAAAAGCAGGCT**AGCAGGGGAGTAGGGTTCAT
GmAMT1;3HellsRV	**GGGGACCACTTTGTACAAGAAAGCTGGGT**ACGTGACAACAGTCGACCAA
GmDMT1HellsFW	**GGGGACAAGTTTGTACAAAAAAGCAGGCT**GGAATAGGAATGCCATCCA
GmDMT1HellsRV	**GGGGACCACTTTGTACAAGAAAGCTGGGT**TTTTGACAACAAAATTTGAAAGTA

Table 4-1. PCR primers used to amplify gene fragments for pHellsgate8 silencing vector. *AttB* sequences for recombination into the pDONR222 vector (Invitrogen) are in bold.
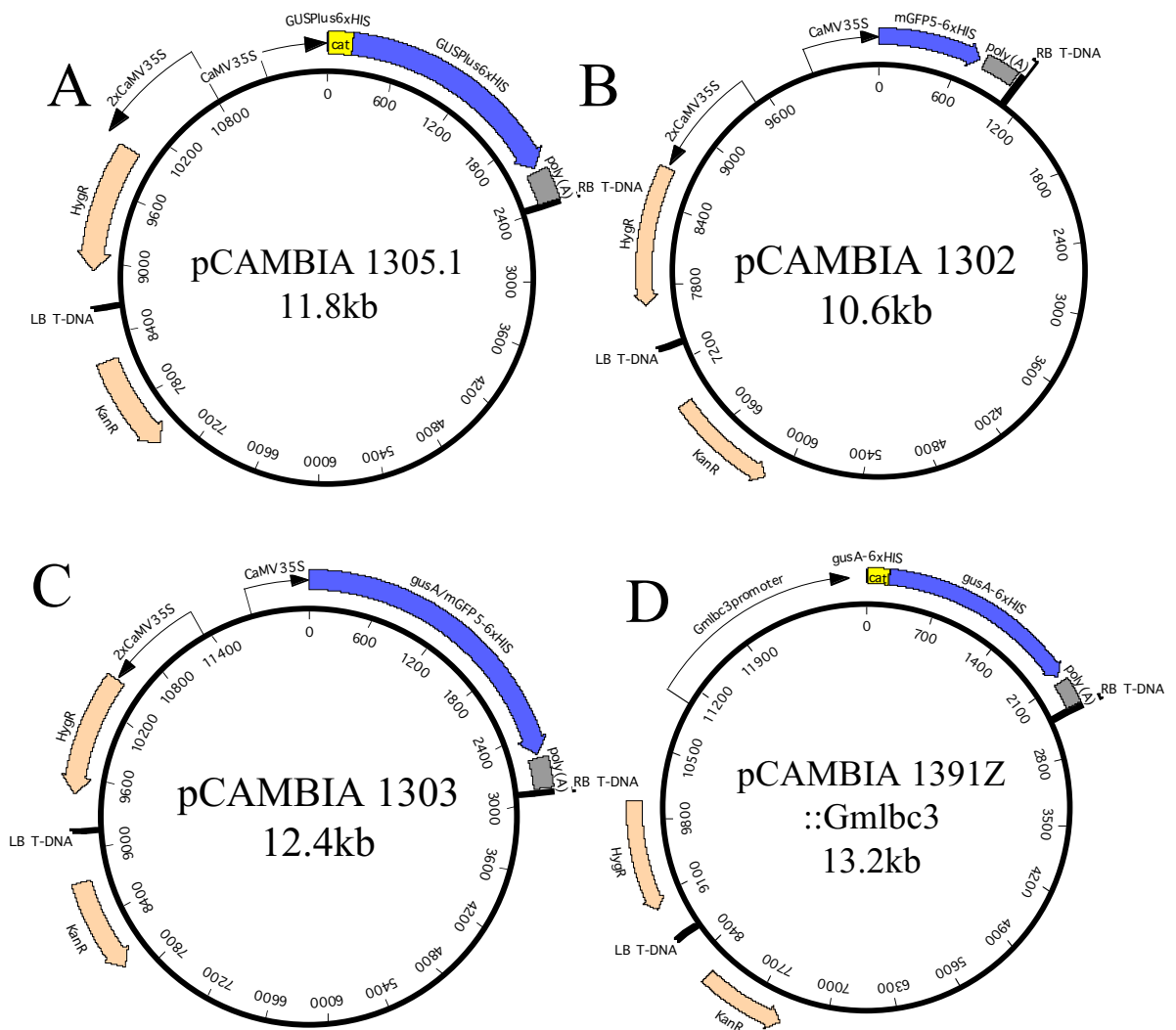
Figure 4-1. pCAMBIA (www.cambia.org) binary vectors (A-D) used in this study. Vectors **A-C** were obtained from Cambia (Canberra, Australia). pCAMBIA 1391Z::Gmlbc3 **(D)** was constructed from pCAMBIA 1391Z with the 2kb *Gmlbc3* promoter cloned into the *Sal* I/ *Eco* RI sites upstream of the gusA gene. HygR – hygromycin resistance, KanR – kanamycin resistance, LB/RB T-DNA – left/right border T-DNA sequences, (2x)CaMV35S – (2 copies) Cauliflower Mosaic Virus 35S eukaryotic promoter, cat - catalase intron. *gusA*-6xHIS - His-tagged glucoronidase gene. mGFP5-6xHIS - His-tagged GFP
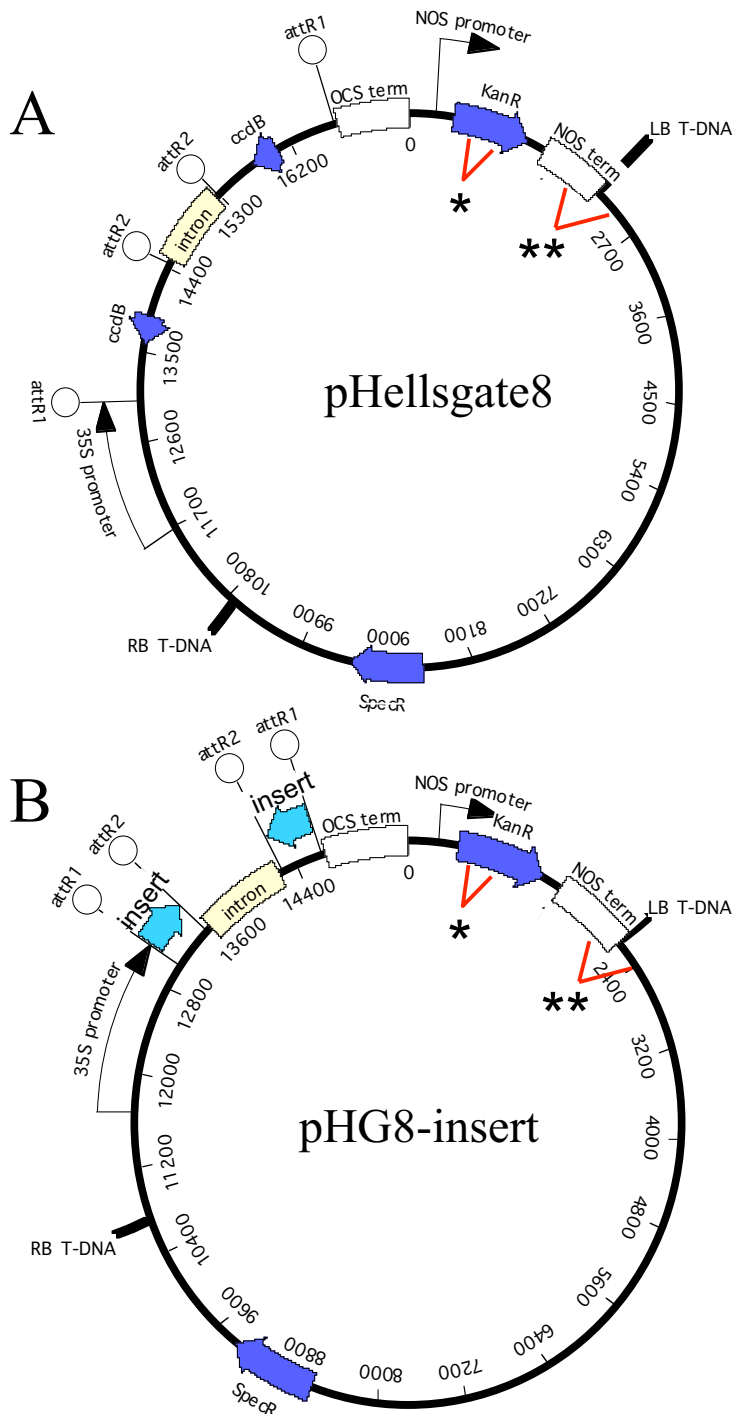
Figure 4-2. RNAi silencing vectors used in this study. Inserts of interest were recombined into the attR1 and attR2 recombination sites of the pHellsgate8 vector (Helliwell et al., 2002), (**A**) as described in section 3.4.1 to give an inverted repeat of the insert of interest (insert), separated by an intron (**B**). Expression of the construct was under the CaMV35S promoter. SpecR - bacterial spectinomycin resistance, ccdB - cytotoxic positive bacterial selection marker, KanR - plant kanamycin resistance, OCS-octopine synthetase, NOS - nopaline synthetase. LB/RB-TDNA - left/right borders of T-DNA. (*) and (**) - approximate region of plasmid amplified with '*nptII*' and 'Across_LB' primers respectively, as described in section 2.4.4.
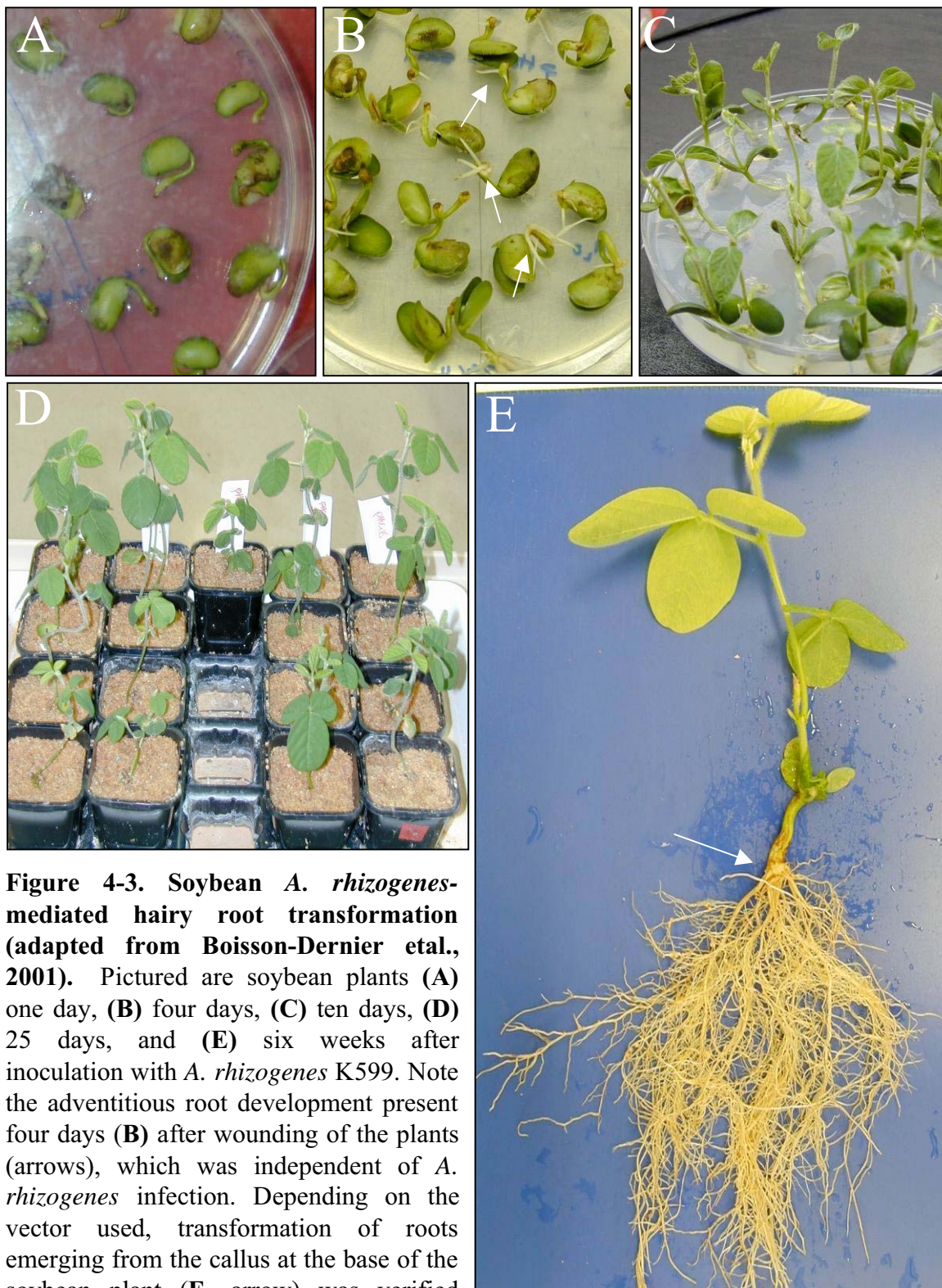
Figure 4-3. Soybean *A. rhizogenes*-mediated hairy root transformation (adapted from Boisson-Dernier etal., 2001). Pictured are soybean plants **(A)** one day, **(B)** four days, **(C)** ten days, **(D)** 25 days, and **(E)** six weeks after inoculation with *A. rhizogenes* K599. Note the adventitious root development present four days **(B)** after wounding of the plants (arrows), which was independent of *A. rhizogenes* infection. Depending on the vector used, transformation of roots emerging from the callus at the base of the soybean plant **(E**, arrow) was verified through GUS staining or PCR analysis.

-Hygromycin   + Hygromycin
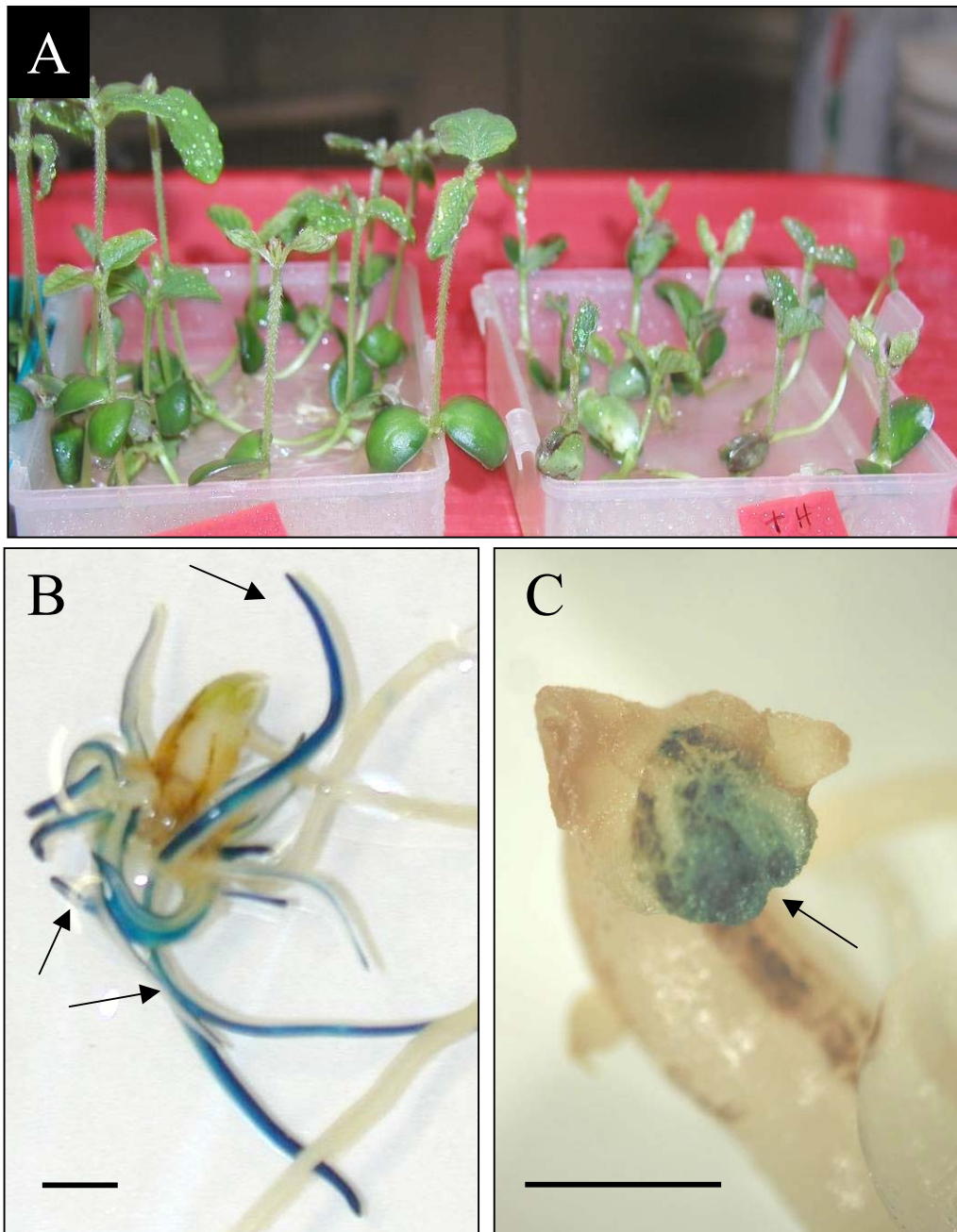
Figure 4-4. Inhibitory effects of hygromycin on soybean growth. Soybean plants were transformed with pCAMBIA1305.1 as described in section3.4.3 and grown on –N Herridge's with or without 25µg/ml hygromycin B. **(A)** Development of soybean aerial tissue was retarded with the addition of hygromycin. (**B & C**) GUS staining of 10 day old root systems demonstrated that hygromycin also prevented development of transformed (GUS positive, arrows) roots **(C)** when compared with roots of plants grown without hygromycin **(B)**. Bar = 5mm
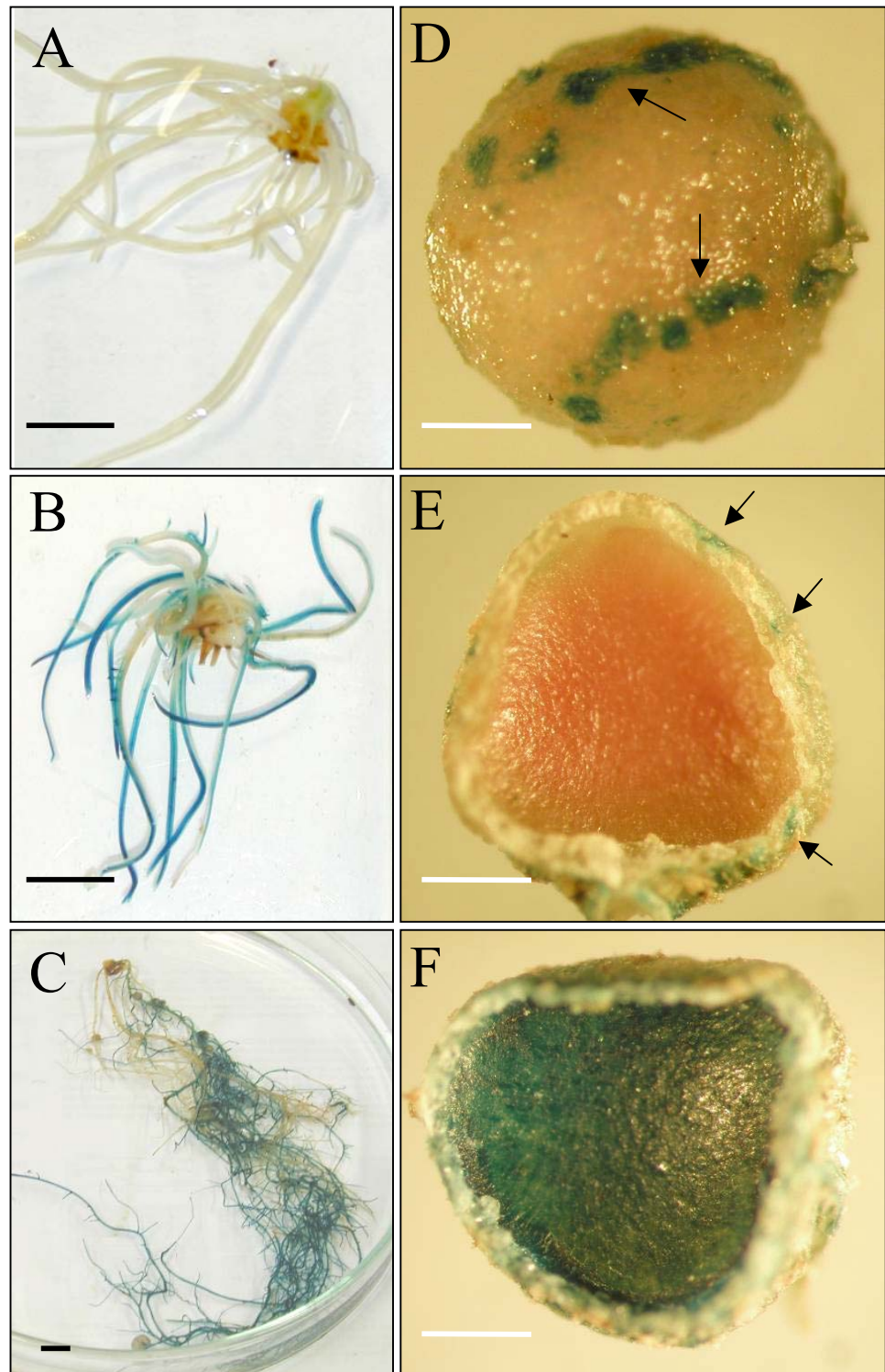
Figure 4-5. CaMV35S promoter-driven GUS expression in soybean roots and nodules. Plants were transformed as described in section 2.4.3 with the CaMV35S promoter-driven GUS expression vector pCAMBIA 1305.1, and roots and nodules GUS stained (section 2.4.5). **(A)** No glucoronidase activity was observed in untransformed root material. GUS expression was detected throughout the transformed root tissue **(B & C)** and the lenticels of the outer cortex of mature nodules (arrows**, D, E**). The mature nodule outer cortex was impermeant to the GUS stain **(E),** so mature nodules were sectioned prior to staining **(F)** to observe GUS staining of the inner cortical region. Black bar = 1cm, white bar = 1mm
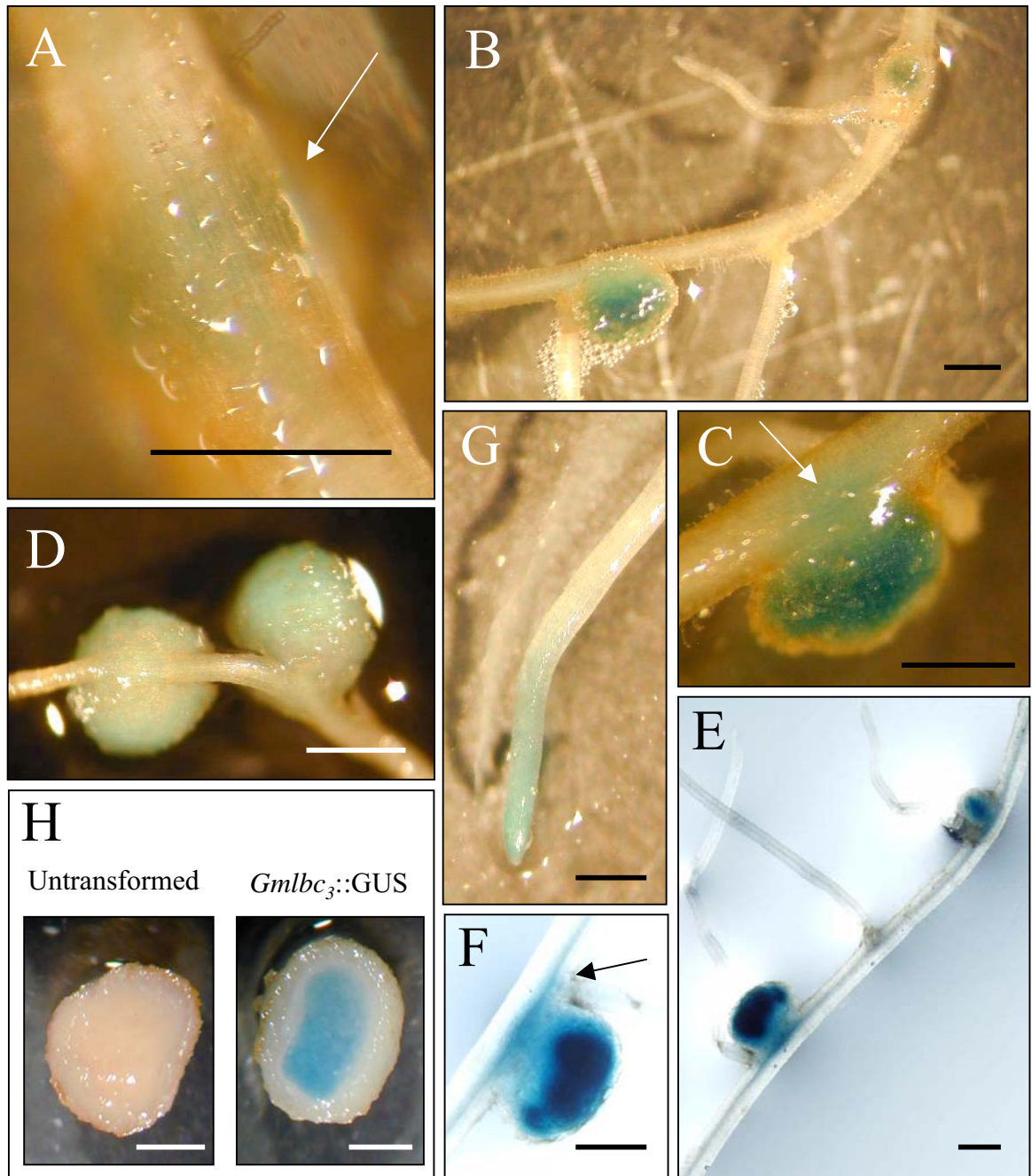
Figure 4-6. GUS expression under the soybean leghaemoglobin promoter (*Gmlbc₃*). 2 kb of soybean genomic DNA 5' of the lbc₃ gene was cloned upstream of the GUS open reading frame in the pCAMBIA 1391Z vector and plants transformed with the resulting construct (pCAMBIA 1391Z::Gmlbc3, Figure 1D). **(A,** arrow**)** GUS expression was first observed in the nodule primordium and subsequently throughout nodule development **(B-F, H)**, primarily in the infected inner cortex. (**arrows** in **C, F)** GUS staining was also observed in the vasculature immediately adjacent nodules and **(G)** occasionally in the root tip. **(H,** right**)** Mature pCAMBIA 1391Z::Gmlbc3 transformed nodules (Gmlbc3::GUS) were sectioned prior to GUS staining. **(H,** left**)** A GUS stained untransformed nodule is pictured for comparison. Bar = 2mm
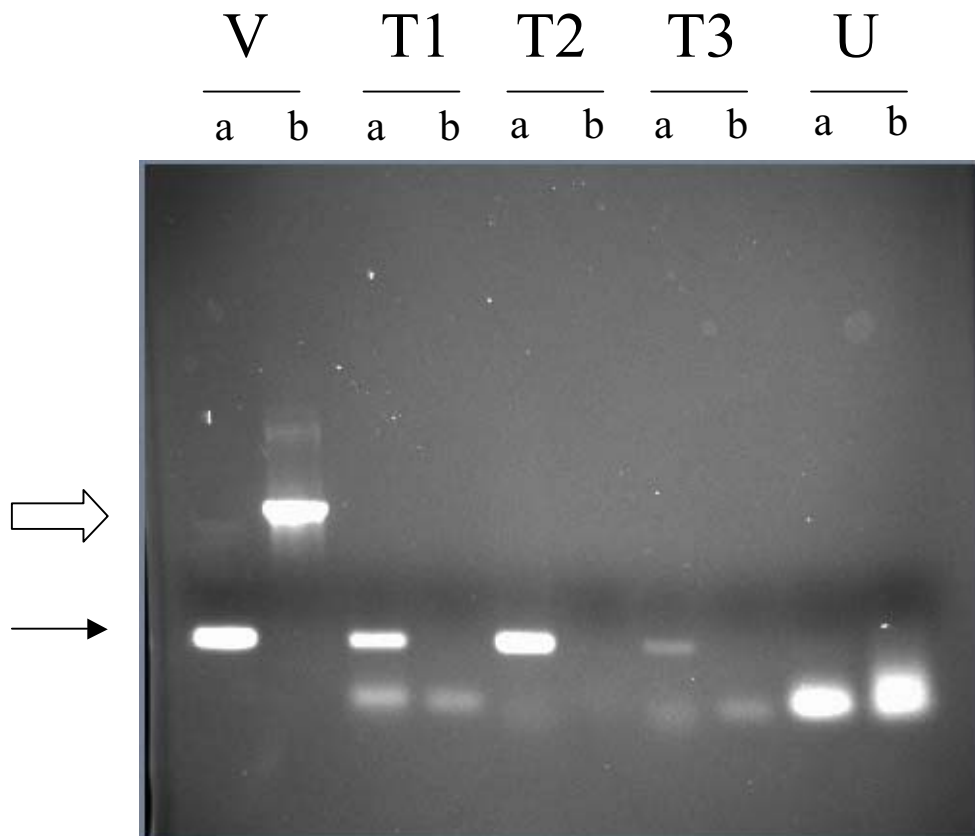
Figure 4-7. Confirmation of transformation of roots with pHellsgate8-derived vectors. PCRs using primers described in section 3.4.4 specific to the *NptII* kanamycin resistance gene within the T-DNA of pHellsgate 8 vector (a) and across the left border of the T-DNA of pHellsgate 8 (b) (see also Figure 2 in this chapter) were performed using a pHellsgate 8 plasmid (V) as template or genomic DNA from three independent putatively pHellsgate8-SAT1 transformed roots (T1, T2 and T3) or untransformed root (U). Products specific for *NptII* are marked with filled arrow (212bp) and across the left border marked with an open arrow (793bp).
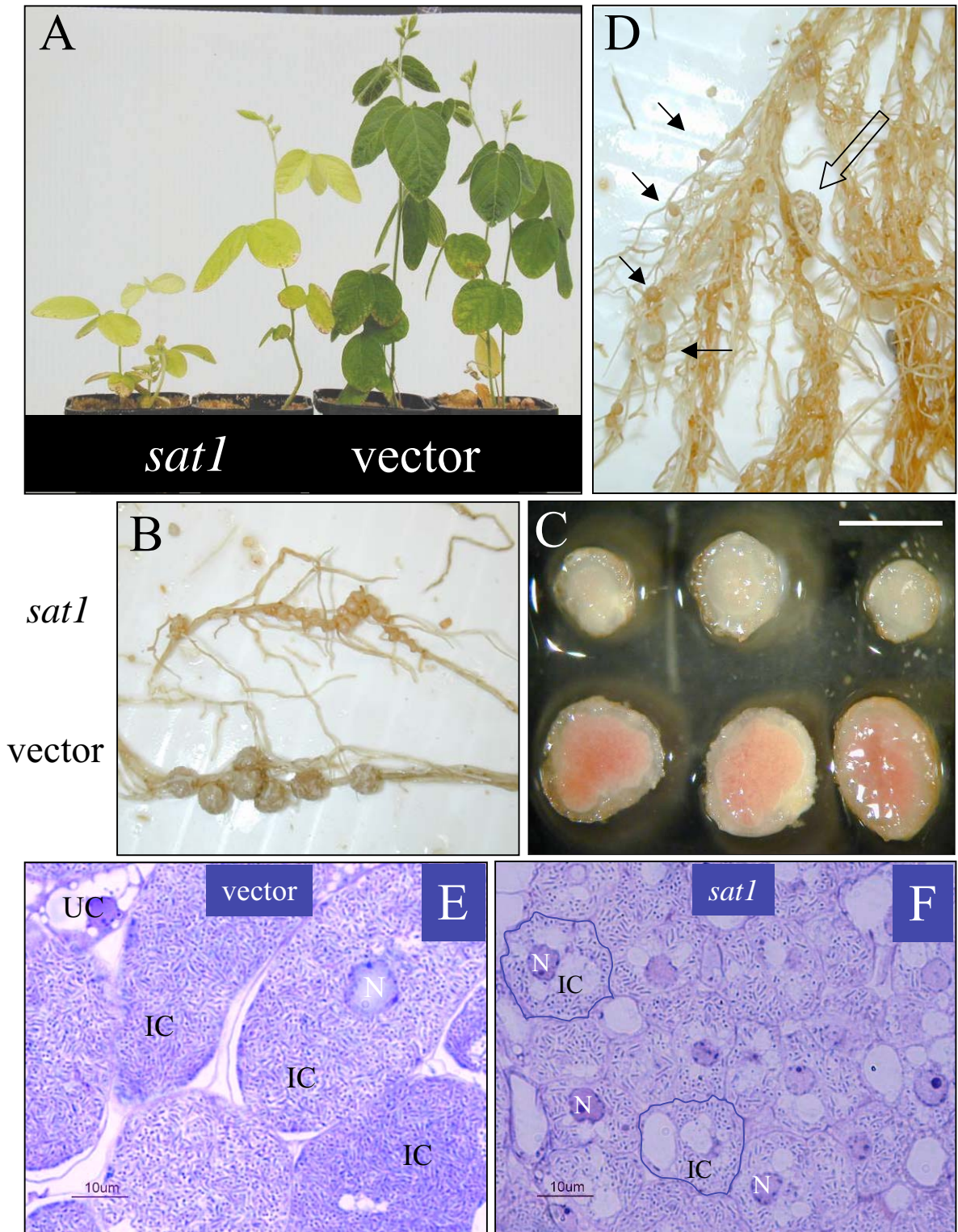
Figure 4-8. Phenotype of GmSAT1-silenced (*sat1*) roots and nodules. Soybean plants were transformed as described in section 2.4.3 with pHellsgate8-SAT1 (*sat1*) or pHellsgate 8 (vector) and transformed roots selected for using PCR analysis. **(A)** The leaves of *sat1* plants were typically chlorotic and plant growth was stunted. **(B)** *sat1* roots were generally covered in small, ineffective nodules with a very small brownish infected region **(C)**, which was distinct from empty vector transformed nodules. (**D,** open arrow) Occasionally phenotypically wild type nodules developed on *sat1* roots as well as the typically small nodules (**D**, filled arrows). Toluidine blue stained inner cortical nodule sections of **(E)** vector and **(F)** *sat1* nodules revealed anatomical differences caused by silencing of GmSAT1. Note the size difference between infected cells (IC, outlined) of *sat1* nodules and vector controls. IC -infected cell, UC -uninfected cell, N - nucleus White bar = 5mm
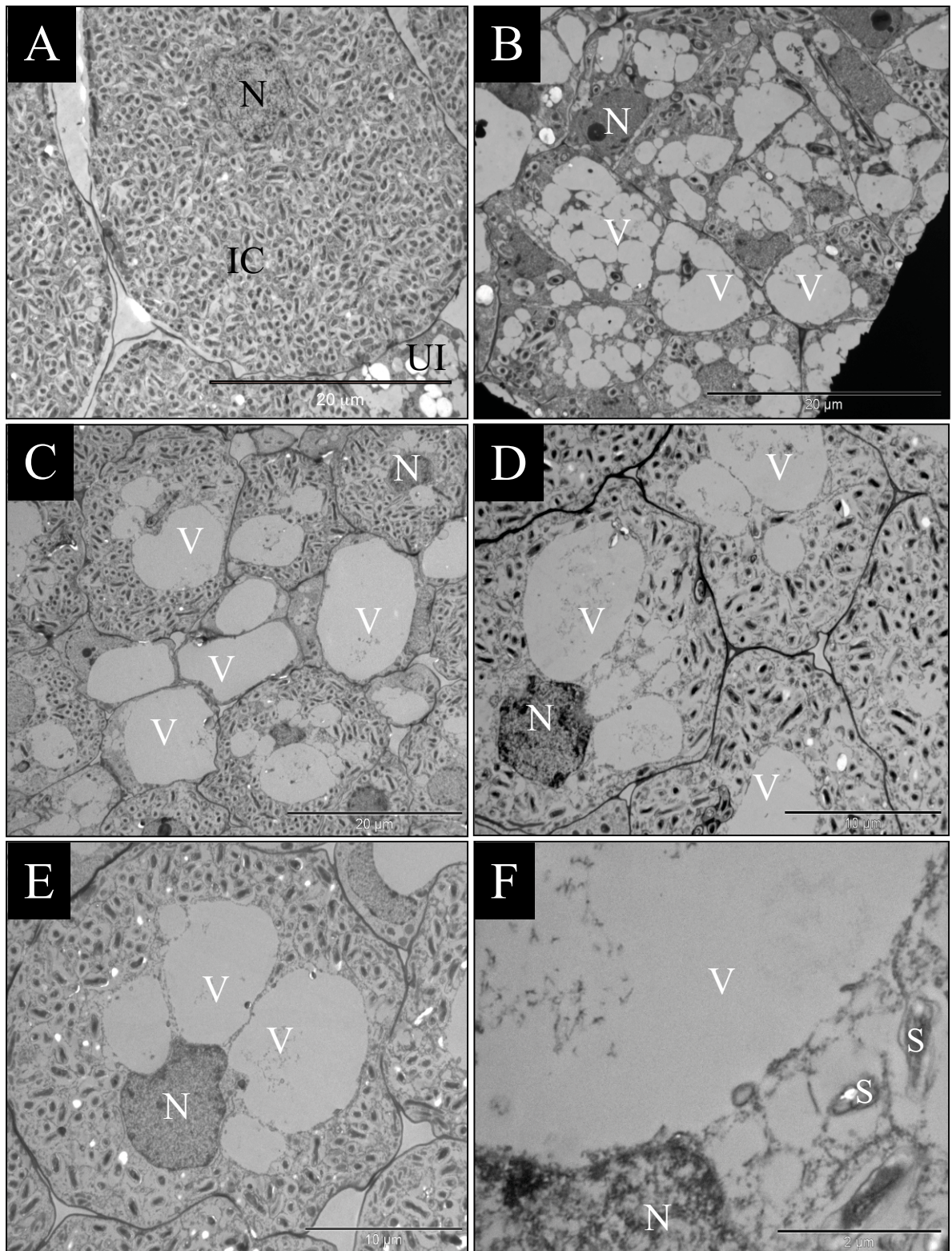
Figure 4-9. Microscopic analysis of structural aberrations of *sat1* nodules. Nodules from **(A)** empty vector-transformed and **(B-F)** pHG8-SAT1 transformed roots were harvested, fixed, sectioned and stained for transmission electron microscopy as described in section 2.4.6. **(A)** Vector transformed infected cells were at least twice the size of adjacent uninfected cells and packed full of bacteroids. Only small vacuoles were observed. Conversely, *sat1* infected cells **(B-F)** had varying numbers of bacteroids, and almost always retained at least some of the central vacuole. In some extreme cases **(B)**, very few bacteroids were present in the infected cells, and the majority of the cell remained highly vacuolated. **(F)** Symbiosomes were also observed fusing with the central vacuole of *sat1* infected cells. N - nucleus, V- vacuole, UI - uninfected cell, IC - infected cell, S - symbiosome
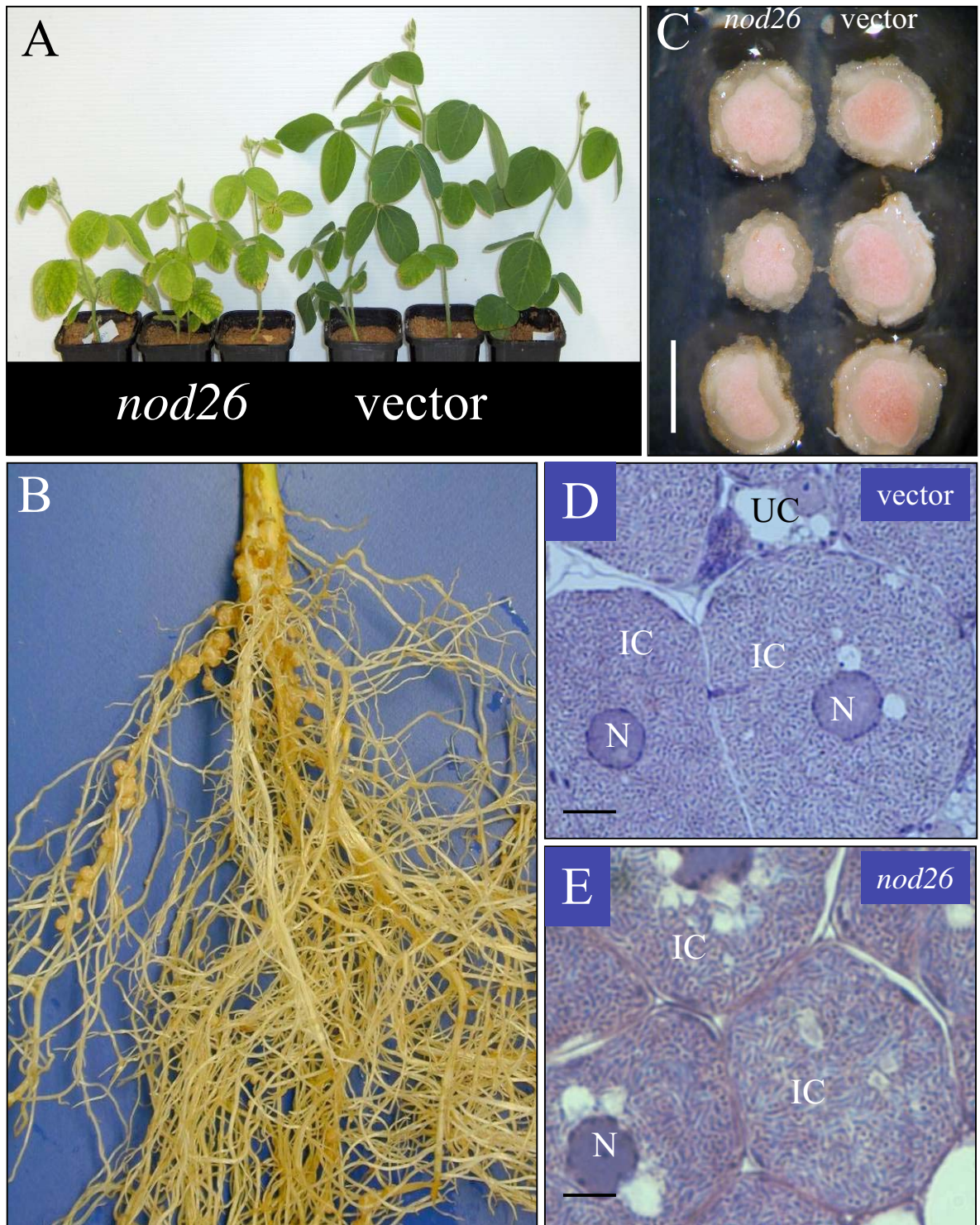
Figure 4-10. The phenotype of putatively GmNOD26-silenced soybean plants. Soybean plants were transformed as described in section 2.4.3 with pHellsgate8-NOD26 (*nod26*) or pHellsgate 8 (vector) and transformed roots selected for using PCR analysis (section 2.4.4). (**A**) *Nod26* plants showed varying degrees of chlorosis however root architecture appeared unaffected (**B**). Nodules were essentially indistinguishable from vector transformed control nodules (**C**) although the inner cortex of *nod26* nodules was slightly less pink than control nodules. (**E**) Microscopic analysis failed to detect any obvious ultrastructural differences in *nod26* nodules when compared with (**D**) empty vector control nodules. N - nucleus, IC - rhizobia-infected cell, UC - uninfected cell. White bar = 5mm, black bar = 10µm

5. General Discussion

5.1 THE LEGUME-RHIZOBIA SYMBIOSIS

The symbiosis between legume and N_2-fixing rhizobia is significant for a number of reasons. First and foremost, this symbiosis allows legumes to grow in marginally fertile, nitrogen-deficient soils that would be considered hostile for other plant species. In a practical sense, it is agriculturally important, with the provision of fixed nitrogen to the legume by the rhizobia allowing reduced costs to the grower in terms of nitrogen fertiliser addition. Legumes are grown as a 'green manure' to fertilise the soil for future crops, as protein rich fodder for livestock, or as seed for human consumption. Fixation of N_2 to a more biologically useful molecular species such as ammonium has only evolved in prokaryotes and can only be accessed by eukaryotes through symbiotic interactions. As such, this symbiosis may represent an intermediate step between two independent organisms and the formation of a novel intracellular organelle akin to mitochondria or plastids (Kneip et al., 2007). The symbiosis can also act as a model for other plant-microbe interactions, both mutualistic and parasitic, with studies into this particular symbiosis proving to be paradigms applicable to other interactions. Already it is clear there are overlapping signalling pathways between nodulation and mycorrhizal development and it is likely that the rhizobia-legume symbiosis developed from the already established mycorrhizal symbiosis (Oldroyd et al., 2005).

The development and maintenance of the symbiosis between legume and rhizobia is also of scientific interest. Initiation of meristematic tissue in the otherwise quiescent cortical tissue provides researchers with a unique system in which to study cell cycle initiation and the role of cytokinins (Gonzalez-Rizzo et al., 2006; Murray et al., 2007; Tirichine et al., 2007). The organogenesis of the nodule, including the synthesis and trafficking of large amounts of phospholipid and protein material to the novel PBM in infected cells and concomitant transcriptional reprogramming, is also of great interest Maintaining the intracellular presence of the bacteroids in infected nodule cells and the mutually beneficial exchange of nutrients without the bacteroids triggering a pathogen response from the legume is a balancing act, requiring exquisite control, especially at the membrane interface between the two organisms.

5.2 INITIAL IDENTIFICATION AND CHARACTERISATION OF GmSAT1

The soybean nodule-specific gene GmSAT1 (*Glycine max* symbiotic ammonium transporter 1) was isolated from a nodule cDNA library in a complementation screen of a *Saccharomyces cerevisiae* (yeast) ammonium transport mutant 26972c (Kaiser et al., 1998). Expression of GmSAT1 allows growth of 26972c on medium containing 1 mM ammonium as the sole nitrogen source and causes toxicity when grown on 100 mM methylammonium (MA), a toxic ammonium analogue. Patch clamping of 26972c spheroplasts from yeast expressing GmSAT1 also detected a time dependent voltage-sensitive inward cation current, providing further evidence of GmSAT1 acting as an ammonium channel (Kaiser et al., 1998). GmSAT1 is located on the PBM, which surrounds the nitrogen-fixing bacteroid. The physical location of GmSAT1 on the PBM and its apparent activity in yeast made it a prime candidate

involved in the flux of ammonium from the bacteroid to the infected plant cell cytosol (Tyerman et al., 1995). However its unique structure (predominantly hydrophilic with a C-terminal membrane anchor) which includes a highly conserved bHLH DNA binding domain suggested a different role than that of a typical membrane transport protein. This was highlighted in a subsequent study in yeast where the complement of ammonium transporters in 26972c was clarified. *S. cerevisiae* have three high affinity ammonium transporters, called Meps, and subsequent characterisation of the EMS-derived 26972c revealed this strain has a deletion of Mep2 and a point mutation of a highly conserved glycine residue (G413$\rightarrow$D) in the C-terminus of Mep1 which inhibits its own activity and also trans-inhibits the third Mep, Mep3 (Marini et al., 2000). This trans-inhibition is thought to be due to a direct protein-protein interaction between the Mep proteins, as it has been shown that other ammonium transporters of the Mep/Amt family function as homo- and hetero-trimers and that mutation of glycine residues corresponding to Gly413 in Mep1 also have a dominant negative effect on their function (Ludewig et al., 2003; Loque et al., 2007). The authors also demonstrated that elevated expression of Mep3 (using a high copy number plasmid) overcomes this trans-inhibition and allows growth of 26972c on 1 mM ammonium medium. Furthermore, GmSAT1 was found unable to complement growth of a yeast mutant where all three Meps were deleted (strain 31019b) and that in 26972c (but not in the parent strain Σ1278b), GmSAT1 increases the amount of MEP3 protein located on the yeast plasma membrane (Marini et al., 2000). Thus questions remain as to how GmSAT1 expression mediates complementation of 26972c on low ammonium and, more importantly, what its role in the soybean nodule is.

5.3 GmSAT1 ACTS AS A bHLH TRANSCRIPTION FACTOR

The two most prominent features of the protein sequence of GmSAT1 are its homology with the basic Helix-Loop-Helix (bHLH) DNA binding transcription factor family (amino acids 169-218) and a predicted C-terminal transmembrane domain (amino acids 306-327). Mutation analysis of some of the highly conserved bHLH amino acids found in GmSAT1 have highlighted the functional importance of the bHLH domain and provided evidence of a transcriptional activation role for the protein (Chapter 2). The glutamate residue at position 9 of the bHLH domain (E177 in GmSAT1) is invariant among DNA-binding bHLH proteins and is believed to be essential for DNA-binding, with the exception of an artificial Glu9 $\rightarrow$ Gln substitution in ScPHO4 which retains its DNA-binding ability (Fisher and Goding, 1992). Crystal structures of bHLH proteins show that this glutamate makes contacts with the C and A base groups of their cognate 'CANNTG' DNA sequence. Conservative substitution of Glu177 in GmSAT1 to either Asp or Gln yielded an unstable protein product and disrupted GmSAT1 activity (Figure 2-5). However conversion of this Glu to Ala produced a stable protein, which was toxic to yeast even when grown on the alternative nitrogen source proline (Figure 2-5). A class of bHLH proteins which lack a DNA-binding basic domain have been characterised in mammals (Benezra et al., 1990; Ohtani et al., 2001) and more recently in *Arabidopsis* (Hyun and Lee, 2006). Rather than acting as transcription factors, these proteins act as negative regulators of other, DNA-binding bHLH TFs, heterodimerising with them and preventing them from binding DNA. The toxicity associated with expression of the E177A GmSAT1 mutant may be caused by its dimerising with endogenous yeast bHLH TFs and preventing their DNA binding activity. This would conceivably

disrupt normal cellular activities in yeast, leading to toxicity. If this is the case, then it implies that GmSAT1 may heterodimerise with yeast bHLH TFs *in vivo* in a positive or negative regulatory fashion. Overexpression of selected yeast bHLH TFs failed to result in a distinctive GmSAT1-like phenotype where 26972c growth could be rescued on low ammonium medium although overexpression of Ino2 did cause MA-associated toxicity (Figure 2-32). As only six of the ten bHLH and bHLH-like yeast proteins were examined in this study (Figure 2-11) it is not possible to rule out that overexpression of one of the remaining bHLH proteins could mediate a GmSAT1-like phenotype in 26972c.

5.4 THE MEMBRANE LOCALISATION OF GmSAT1

Prior to the nuclear importation of GmSAT1 or a portion of the protein, GmSAT1 associates with membrane systems in both yeast and soybean. There is compelling evidence which demonstrates GmSAT1 is located at the yeast plasma membrane and on the PBM of soybean rhizobia-infected cells (Kaiser et al., 1998; Chapter 3). For example, an N-terminal GFP-GmSAT1 construct was found to be localised to punctate structures around the periphery of the yeast cell next to or part of the yeast plasma membrane (Figure 3-7). The biochemical identity of the punctate bodies was not determined, however recent research may shed some light on their possible identity. This punctate expression pattern is similar to that observed for endosome or pre-vacuolar localised proteins (Hicke et al., 1997). Endosomes are small vesicles involved in the trafficking of plasma membrane proteins from the plasma membrane to the vacuole for degradation (Dupre et al., 2001). The presence of GmSAT1 in these vesicles may represent the endocytosed protein enroute from the plasma membrane to

the vacuole. Alternatively, it has been found that similar to mammalian lipid rafts, yeast have stable microdomains within the plasma membrane, with distinct protein and lipid components (Malinska et al., 2003). One particular 300nm microdomain in yeast termed the MCC (membrane compartment occupied by Can1p; previously RMC C) is characterised by the presence of Can1p and Sur7p and bears a striking resemblance to the cell-peripheral punctate localisation of GmSAT1 (Grossmann et al., 2007). Recently these MCC sites have been co-localised with 'eisosomes' which appear to be stationary sites of endocytosis of yeast plasma membrane proteins (Walther et al., 2006). Future studies will be looking at determining if GmSAT1 is also localised to these MCCs and if so, what the functional significance of this localisation is. Furthermore I will be examining whether the MCCs show different properties across the ammonium transport mutants (26972c, 31019b) and those of wild-type parental cells Σ1278b because, as discussed below, it may be that the yeast strain 26972c is defective in correct trafficking of plasma membrane proteins.

5.5 GmSAT1: A MEMBRANE BOUND TRANSCRIPTION FACTOR

In light of the detrimental effects caused by conservative mutation of key amino acids in the bHLH DNA-binding domain of GmSAT1, it was not surprising to observe using both immunocytochemistry and GFP fusion experiments that GmSAT1 or a component of it is localised in the nucleus. When expressed in yeast, fusion of both an artificial transcription factor (Figure 3-10) and GFP (Figure 3-8) to the N-terminus of GmSAT1 localised the protein to the nucleus. Immunogold studies also localised the GmSAT1 protein to the nucleus of rhizobia infected cells in the soybean nodule (Figure 3-3 & 3-4A), whereas previous localisation studies of GmSAT1 localised it to

the plasma and peribacteroid membranes of yeast and soybean respectively (Kaiser et al., 1998). Western blotting of protein fractions from yeast expressing GmSAT1 and soybean nodules (Figure 3-1) helped to resolve this apparent discrepancy. Using anti-GmSAT1 antiserum, an immunogenic band approximately corresponding to the expected MW of the full length GmSAT1 protein was identified in the insoluble nodule protein fraction whereas in the soluble fraction a 5 kDa smaller immunogenic band was recognised by the antibody (Figure 3-1A). It is possible that similar to other transcription factors such as SREBP (Duncan et al., 1997) and ATF6 (Ye et al., 2000), GmSAT1 is a membrane bound transcription factor which undergoes proteolysis, liberating the N-terminal, DNA-binding component of the protein, which translocates to the nucleus and activates target gene expression.

5.5.1 THE ENDO-PROTEOLYSIS OF GmSAT1

Western blotting of yeast and soybean nodule protein extracts indicate that GmSAT1 is being proteolytically cleaved in both a yeast expression system and in soybean nodules (Figure 3-1). This cleavage is very unlikely to be occurring after yeast cell lysis as yeast cultures were spun down and washed once in 1 mM EDTA before being lysed in ice-cold 2 M NaOH and total protein precipitated in ice-cold 25% (w/v) trichloroacetic acid. Furthermore, only fusion of an artificial transcription factor to the N-terminus and not the C-terminus of GmSAT1 self-activated reporter gene expression in a modified yeast two-hybrid system, indicating that GmSAT1 is being cleaved *in vivo* (Figure 3-10). The size difference between the full length and truncated GmSAT1 proteins, as estimated from Western blotting, suggests that GmSAT1 is being cleaved within, or close to its putative C-terminal transmembrane domain. Interestingly, the one point mutation that did appear to reduce the presence of

the truncated GmSAT1 product was the substitution of Ile318→Gln in the middle of the predicted transmembrane domain (Figure 2-4). It is unclear how this mutation may be affecting the cleavage of GmSAT1 and we are currently attempting to definitively map the proteolytic cleavage site in GmSAT1 using Western blotting of C-terminal FLAG fusions and introducing further mutations within and adjacent the predicted transmembrane domain.

A potential proteolytic cleavage recognition sequence, 'RXXL', which is recognised by the subtilisin-like site-1 protease (S1P) in organisms as diverse as mammals, plants and yeast (Duncan et al., 1997; Hughes et al., 2005; Liu et al., 2007), is also present in GmSAT1, approximately 30 amino acids away from the N-terminal side of its TMD. This cleavage site was mutated in GmSAT1 to determine its role in the proteolytic cleavage of the protein. Mutations to both the Arg and Leu in the RXXL motif of GmSAT1 prevented self-activation of reporter gene expression using a modified yeast two-hybrid assay (Figure 3-12), but did not affect cleavage of GmSAT1 as observed by Western blotting (Figure 3-13B). As mentioned previously, the mutant GmSAT1 were more highly expressed in yeast used for the Western blotting experiments when compared with the yeast two-hybrid experiments, and this overexpression may mask the effect that mutation of the RXXL site is having on proteolysis. The fact that similar full length and truncated products of GmSAT1 are present in soybean nodules and yeast suggests a similar proteolytic cleavage event is occurring in both organisms. At present I can only speculate as to the identity of the protease(s) involved. A BLAST search of the soybean EST database with the human S1P identified two similar partial ESTs (accessions BM095107 and CF806874) indicating the presence of a S1P-like protein in soybean. Another soybean subtilisin-like protease, SLP-1 is

thought to be involved in the development of the seed coat (Beilinson et al., 2002). Subtilisin-like proteases have also been implicated in symbiosis formation. For example, the subtilisin-like protease ag12 (accession X85975) is specifically expressed during the early stages of actinorhizal nodule development in *Alnus glutinosa* (Ribeiro et al., 1995). As ag12 transcript is restricted to *Frankia*-infected actinorhizal nodule cells, the authors speculate it may be important in the processing of undetermined proteins at the interface between the bacteria and plant (Ribeiro et al., 1995). Interestingly, a soybean proteome study has also identified a subtilisin-like serine protease on the PBM (Panter et al., 2000) which conceivably may be involved in cleavage of GmSAT1 from the PBM. Unlike the yeast *S. pombe, S. cerevisae* does not have a hypoxia response mechanism which involves the cleavage of a SREBP-like transcription factor by site-1 (S1) and site-2 (S2) proteases (Hughes et al., 2005). However *S. cerevisiae* do posses other membrane associated proteases which may potentially mediate cleavage of GmSAT1. For example, the trans-Golgi localised serine protease Kex2 is a membrane bound protease found in yeast, which endo-proteolytically cleaves the mating factor precursor pro-α-factor (Julius et al., 1984). Its punctate localisation pattern (Redding et al., 1991), which is similar to that of GmSAT1 when expressed in yeast may place the two together, allowing cleavage of GmSAT1 by Kex2, although the biochemical identity of the GmSAT1-containing vesicles is yet to be determined (see above). Another example of a membrane associated protease in *S. cerevisae* is the serine-protease Ssy5p, which is a component of the amino acid sensor complex at the yeast plasma membrane (Abdel-Sater et al., 2004). Upon external amino acid sensing, Ssy5p proteolytically cleaves a 10 kDa negative regulatory element from the N-terminus of the transcription factors Stp1p

and Stp2p, allowing their import into the nucleus and activation of amino acid transporter gene transcription (Andreasson et al., 2006).

5.5.2 CLEAVAGE OF GmSAT1 BY A *B. japonicum* ENCODED PROTEASE?

The orientation of GmSAT1 on the PBM of infected cells is yet to be determined. Although unlikely, GmSAT1 may be oriented with its N-terminal TF domain facing into the peribacteroid space, allowing a rhizobia-secreted protease to mediate cleavage of the protein, as there are numerous annotated membrane associated and secreted proteases in the genome of the soybean rhizobial symbiont, *B. japonicum* (http://bacteria.kazusa.or.jp/rhizobase/Bradyrhizobium). The released TF would then need to be endocytosed to allow its entry into the cytosol and eventual import into the soybean nucleus to activate target gene transcription and there is limited evidence from immunogold labelling experiments to suggest this might occur.

5.6 GmSAT1 AND ITS ROLE IN AMMONIUM / METHYLAMMONIUM TRANSPORT IN 26972c

5.6.1 GmSAT1 INCREASES Mep3 EXPRESSION IN 26972C

It still remains unclear what specific role GmSAT1 plays in its original complementation of 26972c on 1 mM ammonium medium. Apart from improving cell growth on low concentrations of ammonium GmSAT1 also allows for the uptake of ^{14}C-MA and induces a MA toxicity phenotype, which is commonly associated with the activity of numerous members of the MEP/AMT class of high affinity ammonium transport proteins. As discussed above, GmSAT1 is located at the plasma membrane

in confined punctate vesicles as well as in the nucleus following a cleavage event where the N-terminal (bHLH) component separates from the C-terminal TMD anchor. In light of the large body of evidence presented in this thesis it is becoming clearer that a transcriptional response is fundamentally involved in GmSAT1's activity in the yeast ammonium transport mutant 26972c. Marini et al., (2000) demonstrated that when Mep3 was expressed from a high copy plasmid, the trans-inhibition by *mep1-1* was overcome. Indeed, elevated levels of Mep3 protein were observed in microsomal fractions of the 26972c mutant when GmSAT1 was expressed (Marini et al., 2000). This result suggests Mep3 expression may be enhanced in cells containing GmSAT1. We have demonstrated that Mep3 expression increases approximately 1.5-fold and *mep1-1* expression decreases slightly in 26972c cells containing GmSAT1 over that of control cell lines (Appendix A, Figure A1). It is unlikely that GmSAT1 is directly regulating expression of either Mep as there are no E-box sequences (CACGTG) within the promoter regions of either to which GmSAT1 is predicted to bind. Whether this modest increase in Mep3 transcript is sufficient to explain the ability of GmSAT1 to complement growth of 26972c on 1 mM ammonium is not known. The fact that this complementation phenotype of GmSAT1 is dependent on the presence of a functional Mep3 (Marini etal., 2000) implies that the phenotype is related to an increase in Mep3 activity, possibly through an upregulation in Mep3 transcription leading to an increase of the protein at the plasma membrane.

5.6.2 THE ROLE OF GmSAT1 IN MA ACCUMULATION AND TOXICITY IN YEAST

Mep3 is an ammonium transport protein with a predicted K_M for NH_4^+ between 1.4 – 2.1 mM (Marini et al., 1997). However one of the confusing aspects of GmSAT1 activity in 26972c is associated with its ability to accumulate ^{14}C-MA over that of

empty vector controls cells and induce a toxic phenotype when grown in the presence of 100 mM MA (Kaiser et al., 1998, Figure 2-5A & 5B). Due to the lower transport capacity of Mep3 it is not considered an effective transporter of MA or capable of accumulating MA to induce a toxicity response (Marini et al., 1997, 2000). This phenotype has been confirmed recently (Kaiser, B. unpublished results) where yeast strains containing Mep3 alone (haploid; 26972c1, diploid; MLY115, see Appendix A, Table A1) or strains devoid of any functional Mep transporters (31019b, 26972c2, see Appendix A, Table A1) display equal levels of ^{14}C-MA accumulation across a range of concentrations covering both high and low-affinity systems (Appendix A, Figure A2). Correspondingly, cells where Mep3 is the only active Mep grow well in the presence of 100 mM MA (Appendix A, Figure A3 A & B). Clearly Mep3 in its native state does not transport MA effectively or enhance its accumulation within the cell over that of cells devoid of any Mep protein. However these phenotypes change when GmSAT1 is expressed alongside Mep3, where the level of ^{14}C-MA uptake increases and cells become sensitive to 100 mM MA (Appendix A, Figure A3 B, e.g., Figure 2-5A & B).

At this stage we can still only propose hypothetical models to explain how GmSAT1 modifies MA uptake and sensitivity in yeast. We suggest GmSAT1 may influence MA transport in one or more of the following ways (Figure 5-1):

1. **Derepression of Mep3 through transcriptional activation by GmSAT1 which delivers greater levels of protein to the yeast plasma membrane.** Expression of GmSAT1 causes a modest but consistent increase in Mep3 transcript which may increase the density of Mep3 at the plasma membrane

and alter its native ammonium/MA transport capacity, causing it to become complicit in letting sufficient MA to enter the cell eventually causing cell toxicity (Figure 5-1A). In support of this hypothesis is the fact that kinetic properties of MA transport in cells containing GmSAT1 are similar to that of Mep3 where both show low-affinity transport kinetics, although cells expressing GmSAT1 exhibit a higher V_{MAX} than native transport (Marini et al., 1997; Kaiser et al., 1998). However this is unlikely to be the only mechanism whereby GmSAT1 increases MA accumulation as MA toxicity on 100 mM MA media is independent of the presence of Mep3 (Appendix A, Figure A3 B).

2. **GmSAT1 activates a second transport system in conjunction with Mep3 or which works independently.** Suggestion of a second system has developed from results obtained with GmSAT1 mutagenesis. Two GmSAT1 mutations, Leu191 → Ile in the dimerisation region of the bHLH domain and Ile318 → Gln in the predicted TMD abolished MA-associated toxicity ie., 26972c grew on 100 mM MA and ^{14}C-MA uptake was abolished, without affecting the ability of GmSAT1 to complement 26972c growth on 1 mM ammonium (Figures 2-4 & 2-6). From this separation of ammonium and MA phenotypes, it would appear that both MA toxicity and its transport into yeast is independent of Mep3 activity. Non-selective cation channels (NSCCs) are known to be active in yeast (Bihler et al., 2002) with similar electrophysiological signatures to that observed in 26972c cells containing GmSAT1 (Kaiser et al., 1998). It is conceivable that overexpression of

GmSAT1 is upregulating expression of a native NSCC(s) or other yeast genes involved in cation transport (Figure 5-1B).

3. **GmSAT1 expression regulates the yeast protein secretory pathway, altering the delivery or removal of transporters to and/or from the yeast plasma membrane.** Strain 26972c was selected after EMS mutagenesis of Σ1278b followed by selection on 100 mM MA (Dubois and Grenson, 1979), which would select against any transport systems that would allow the accumulation of MA at 100 mM. From genetic analysis of 26972c, modifications to the Mep system are evident, however it is possible other, lower affinity MA carrier(s) have also been mutated in 26972c. Alternatively, systems which facilitate normal trafficking and removal of proteins to and from the plasma membrane may have been mutated. Specific ER, vacuolar and endosomal proteins are important for the correct folding, trafficking and maintenance of yeast plasma membrane transporters such as amino acid permeases and phosphate transporters (Lau et al., 2000; Kota et al., 2007) although as yet, similar ancillary proteins have not been identified for Meps. It is interesting to note that when GmSAT1 is expressed in Σ1278b (wild type) cells, Mep3 levels at the plasma membrane are not at elevated levels as in the 26972c mutant (Marini et al 2000). This result may indicate that in 26972c, normal post-translational processing (trafficking and/or turnover) of proteins enroute to or at the plasma membrane is disrupted. In light of the localisation of GmSAT1 in yeast (punctate vesicle bodies at the plasma membrane), in soybean nodules (on the PBM) and the role of a homologue (AtNAI1) in *Arabidopsis* in the development of ER bodies (Matsushima et al., 2004), there

is an increasing body of evidence to suggest GmSAT1 may be involved in the signalling events controlling trafficking of membrane vesicles to cellular membranes such as the plasma membrane (Figure 5-1C) and peribacteroid membrane.

5.7 GmSAT1 IS ESSENTIAL FOR SYMBIOSME DEVELOPMENT

Expression of GmSAT1 is essential for effective soybean nodule development. GmSAT1-silenced nodules were small and ineffective at providing fixed nitrogen to the plant (Figure 4-8). Although the degree of GmSAT1 silencing was not determined due to the lack of sufficient transformed nodule material, it is unlikely that the nodule phenotype observed was due to a hairy root-induced aberration, as the various pCAMBIA and empty vector transformed plants displayed no observable abnormalities. It is possible that the hpRNA transcribed from the pHellsgate8-GmSAT1 vector was silencing other genes, however this is unlikely as the 3' UTR of GmSAT1 was chosen for silencing and a BLAST search using this gene fragment against the soybean genome and EST databases failed to give any significant hits other than GmSAT1. Silencing of GmSAT1 did not affect any of the earlier recognition and signal transduction events, which occur during the initial interaction between the *B. japonicum* and soybean roots. Nodules, although aberrant, developed extensively on *sat1* roots.

The symbiosis in *sat1* nodules appeared to break down following the release of bacteroids from the infection thread into the infected cells although it is difficult to speculate at which point precisely the absence of GmSAT1 activity disrupts the

symbiosis. A model outlining possible roles of GmSAT1 in the soybean nodule is presented in Figure 5-2, and discussed in detail in the following sections. Across the range of silenced nodules, symbiosomes were absent or at a low number per infected cell and infected cell size was markedly reduced. The cell remained highly vacuolated with small and large vesicles distributed across the cell. It has been previously observed in soybean that normal sized 'infected' cells form in the presence of a *B. japonicum* mutant which is unable to be endocytosed and proliferate within the nodule inner cortex (Morrison and Verma, 1987), which suggests infected cell expansion is independent of symbiosome formation. However in this study, reducing GmSAT1 expression appears to affect both the size of the infected cell and the number of symbiosomes, suggesting co-ordination of the two phenomena.

5.7.1 A ROLE FOR GmSAT1 IN TRAFFICKING OF PROTEINS TO THE PBM?

GmSAT1 may be required for the trafficking of the correct complement of proteins to the PBM, without which, the symbiosis breaks down. A similar nodule phenotype to *sat1* nodules is observed when either of two small GTP-binding proteins (Rab1 and Rab7) are silenced in soybean roots (Cheon et al., 1993). Rab proteins are involved in the sorting and regulation of post-Golgi transport vesicles (Hanton et al., 2007) hence it is likely that silencing of Rab1 and Rab7 in the nodule affected trafficking of proteins, potentially to the PBM. Similar to *sat1* nodules, infected cells of Rab-silenced soybean nodules failed to expand in size, and the central vacuole remained (Cheon et al., 1993). Although not definitive, these similarities in nodule phenotypes suggest the possibility that GmSAT1 may have a regulatory role associated with protein trafficking. Additionally, expression of a GmSAT1-like bHLH protein from *Arabidopsis* (AtNAI1) is required for formation of ER structures termed 'ER bodies'

(Matsushima et al., 2004), providing further evidence that GmSAT1 is in some way involved in regulation of the trafficking of proteins through the secretory pathway.

5.7.2 GmSAT1 AND SOYBEAN NODULE NITROGEN BALANCE

GmSAT1 was initially isolated from a soybean nodule cDNA library based on its ability to complement an ammonium transport yeast mutant. Taken together with the fact that the *Arabidopsis* gene locus where the three closest homologues of GmSAT1 are located has been implicated in the nitrogen status of the plant (Rauh et al., 2002), it is feasible that GmSAT1 has a role in nitrogen relations in the nodule. A deficit in accessible soil nitrogen is critical for the formation and maintenance of the legume-rhizobia symbiosis. For example, if soybeans are grown in nitrogen replete conditions, nodulation is either reduced or absent. Furthermore, application of nitrogen to the soybean growth medium following nodulation ends the symbiosis (by the plant) and the nodules senesce. It has also been shown that soybean plants can essentially 'sanction' rhizobia which do not provide the plant with fixed nitrogen, withholding nutrients from the rhizobia and reducing their replication within the nodule (Kiers et al., 2003). If GmSAT1 is directly or indirectly involved in ammonium transport across the PBM, knocking down its expression may cause the observed abnormal infection development and reduced symbiosome number. For example, as described above, GmSAT1 may be required for the correct trafficking of proteins, including ammonium carriers, to the PBM.

5.8 IDENTIFICATION OF A GmSAT1-LIKE GENE FROM SOYBEAN

The recent public release of the first draft of the soybean genome (Soybean Genome Project, DoE Joint Genome Institute; www.phytozome.net/soybean) has revealed the presence of at least one other GmSAT1-like gene, tentatively termed GmSAT2. GmSAT1 and GmSAT2 are 86.4% identical at the protein level and 100% identical across the bHLH and predicted transmembrane domains (Figure 5-3). The role of GmSAT2 in soybean is unknown as are its expression pattern and localisation. Experiments are currently underway to examine the function of this protein, particularly with respect to symbiosis, and whether it is able to complement growth of 26972c on low ammonium medium.

5.9 SUMMARY

Transcription factors are an essential component in the transcriptional (re)programming of cellular activities in response to particular stimuli and the determination of cell fate. The development and maintenance of the symbiosis between legumes and rhizobia requires a large transcriptional shift in receptive legume root cells, to initiate and maintain the intracellular presence of rhizobia within the specialised organ called the nodule. The legume nodule has a unique physico-chemical environment, necessary to allow efficient nitrogen fixation by the rhizobia and the exchange, metabolism and redistribution of nutrients between the symbionts and throughout the plant. GmSAT1 is a soybean nodule-specific gene, with strong homology with members of the bHLH transcription factor family, as well as a single, C-terminal transmembrane domain. In this study I have expanded on the body of published work concerned with the activity of GmSAT1 in a yeast expression system

(Kaiser et al., 1998; Marini et al., 2000), re-examining its function and localisation in both soybean nodules and yeast. Substitution of key amino acid residues in the bHLH motif of GmSAT1 suggest that it acts as a transcriptional regulator and the demonstration of the nuclear localisation of at least the N-terminal part of the protein is consistent with this role. The sequestration of GmSAT1 outside the nucleus by its transmembrane domain most likely represents a post-translational regulatory step, allowing rapid import into the nucleus and target gene activation upon proteolytic cleavage of the bHLH transcription factor domain. Presented in this thesis is evidence suggesting the cleavage event is taking place, however it was not possible to determine the exact site or stimulus for this cleavage, or the protease(s) responsible. Additionally, the identity of genes transcriptionally regulated by GmSAT1 in either soybean or yeast was not examined. Knocking down GmSAT1 expression in soybean roots disrupted the symbiosis between the soybean and rhizobia, implicating GmSAT1 as important in the transcriptional response required by the soybean to maintain the symbiosis. Furthermore the peribacteroid membrane localisation of GmSAT1 prior to nuclear import provides further evidence that GmSAT1 is intimately involved in this symbiosis.
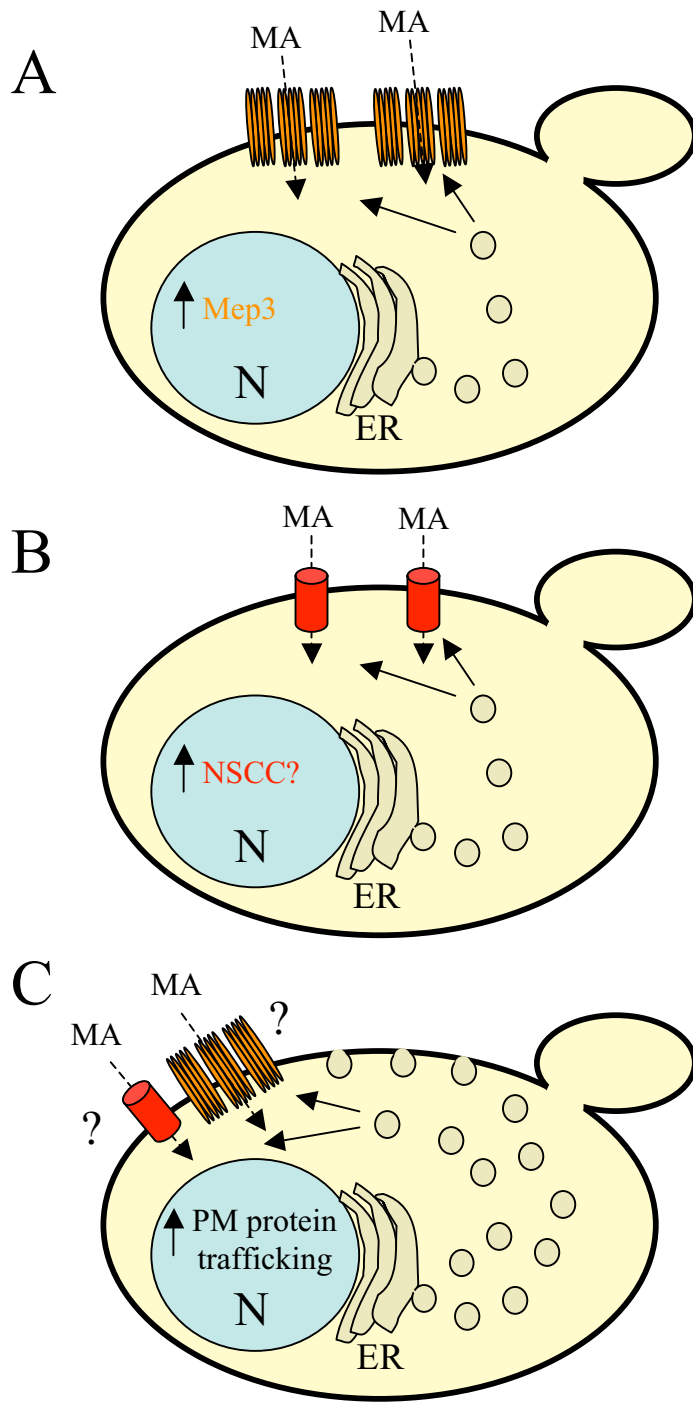
Figure 5-1. Three possible mechanisms by which GmSAT1 expression causes MA toxicity in 26972. (A) GmSAT1 increases Mep3 expression which in turn causes an increased density of Mep3 protein at the membrane and a change in Mep3 MA flux capacity. **(B)** GmSAT1 may upregulate expression of an unidentified non-selective cation channel (NSCC) which allows influx of MA and cell toxicity. **(C)** Secretory trafficking of membrane proteins, including potentially both Mep3 and NSCCs is increased by GmSAT1 expression, allowing increase MA influx and toxicity. For more details refer to Section 5.6. N, nucleus ER, endoplasmic reticulum MA, methylammonium
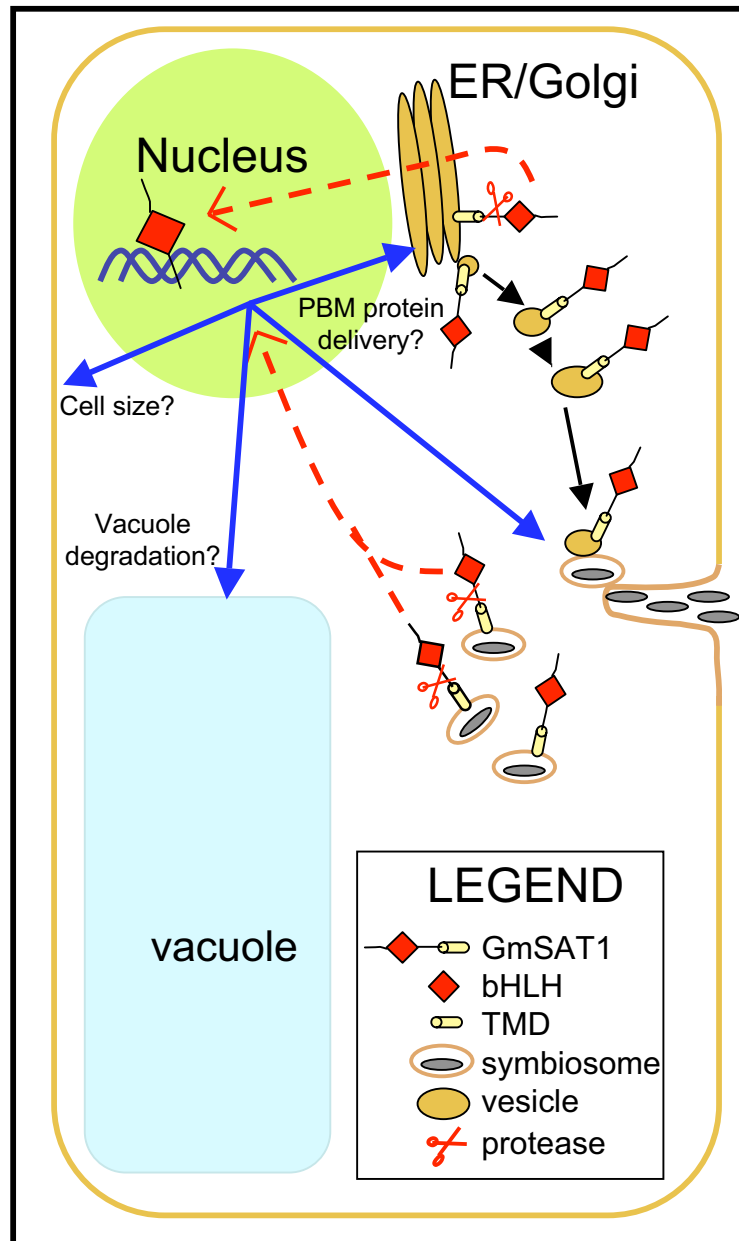
Figure 5-2. A model depicting the role of GmSAT1 in the soybean infected nodule cell. Immunogold localisation of GmSAT1 suggests it follows the secretory pathway, being synthesised in the ER, and trafficked to the peribacteroid membrane (PBM) via the golgi apparatus in vesicles. Either enroute to the PBM, or at the PBM, the 37 kDa N-terminal part of the protein, which includes the basic Helix-Loop-Helix (bHLH) transcription factor domain, is released from its transmembrane domain (TMD) anchor. The signaling event which leads to this proteolysis, and the protease(s) responsible are unknown. The soluble N-terminal part of GmSAT1 translocates to the nucleus where it acts as a DNA-binding transcription factor. The gene target(s) of GmSAT1 are unknown however silencing of GmSAT1 in soybean nodules suggest it is involved in the maintenance of the legume-rhizobia symbiosis. Refer to Section 4.7 for more detail as to possible roles for GmSAT1 *in planta.*

```
              10              20              30
GmSAT2  M I D F N A C T W D V V L F D I L I R L I A S A V F H F M R
GmSAT1  - - - - - - - - - - - - - - - - - - - - - - - - - - M R
        M I D F N A C T W D V V L F D I L I R L I A S A V F H F M R

                      40              50              60
GmSAT2  T S H M E I S S I R G L P E M G I I E D P N F L H H W Q L N
GmSAT1  S S H M E I S S I R G L P E L G I I E D P N F L H Q W Q L N
        . S H M E I S S I R G L P E   G I I E D P N F L H   W Q L N

                      70              80              90
GmSAT2  S I D T T S L T G A P F G E I L Q K H S F S D N S N F N P K
GmSAT1  S I D T T S L K G A A F G D I L Q K H S F S D N S N F N P K
        S I D T T S L   G A   F G . I L Q K H S F S D N S N F N P K

                      100             110             120
GmSAT2  T S M E T S P T G I E R P A K Q L R N N S W N H N K S Q Q Q
GmSAT1  T S M E T S Q T G I E R Y A K Q L G D N S W N H N K S Q Q Q
        T S M E T S   T G I E R   A K Q L     N S W N H N K S Q Q Q

                      130             140             150
GmSAT2  T P E T Q F A S C S N L L S F V N S N Y T S E L G L V K P K
GmSAT1  T P E T Q F A S C S N L L S F V N T N Y T S E L G L V K P K
        T P E T Q F A S C S N L L S F V N . N Y T S E L G L V K P K

                      160             170             180
GmSAT2  V E M V C P K I N N N T L A D M L I S Q G N L G N Q N Y L F
GmSAT1  V E M A C P K I D N N A L A D M L I S Q G T L G N Q N Y I F
        V E M . C P K I   N N   L A D M L I S Q G   L G N Q N Y . F

                      190
GmSAT2  K A S Q E A K K I E T R P K L S Q P Q D H I I A E R K R R E
GmSAT1  K A S Q E T K K I K T R P K L S Q P Q D H I I A E R K R R E
        K A S Q E   K K I   T R P K L S Q P Q D H I I A E R K R R E
```

bHLH

```
GmSAT2  K L S Q R F I A L S A L V P G L K K M D K A S V L G E A I K
GmSAT1  K L S Q R F I A L S A L V P G L K K M D K A S V L G E A I K
        K L S Q R F I A L S A L V P G L K K M D K A S V L G E A I K

                      250             260             270
GmSAT2  Y L K Q M Q E K V S A L E E E Q N R K R T V E S V V I V K K
GmSAT1  Y L K Q M Q E K V S A L E E E Q N R K R T V E S V V I V K K
        Y L K Q M Q E K V S A L E E E Q N R K R T V E S V V I V K K

                      280             290             300
GmSAT2  S R L S S D A E D S S S S E T G D T F D E A L P E I E A R F
GmSAT1  S Q L S S D A E D S S S - E T G G T F V E A L P E I E A R F
        S   L S S D A E D S S S S E T G   T F   E A L P E I E A R F

                      310             320             330
GmSAT2  Y E R N V L I R I H C E K N K G V I E K T I S E I E K L H L
GmSAT1  W E R N V L I R I H C E K N K G V I E K T I S E I E K L H L
        . E R N V L I R I H C E K N K G V I E K T I S E I E K L H L
```
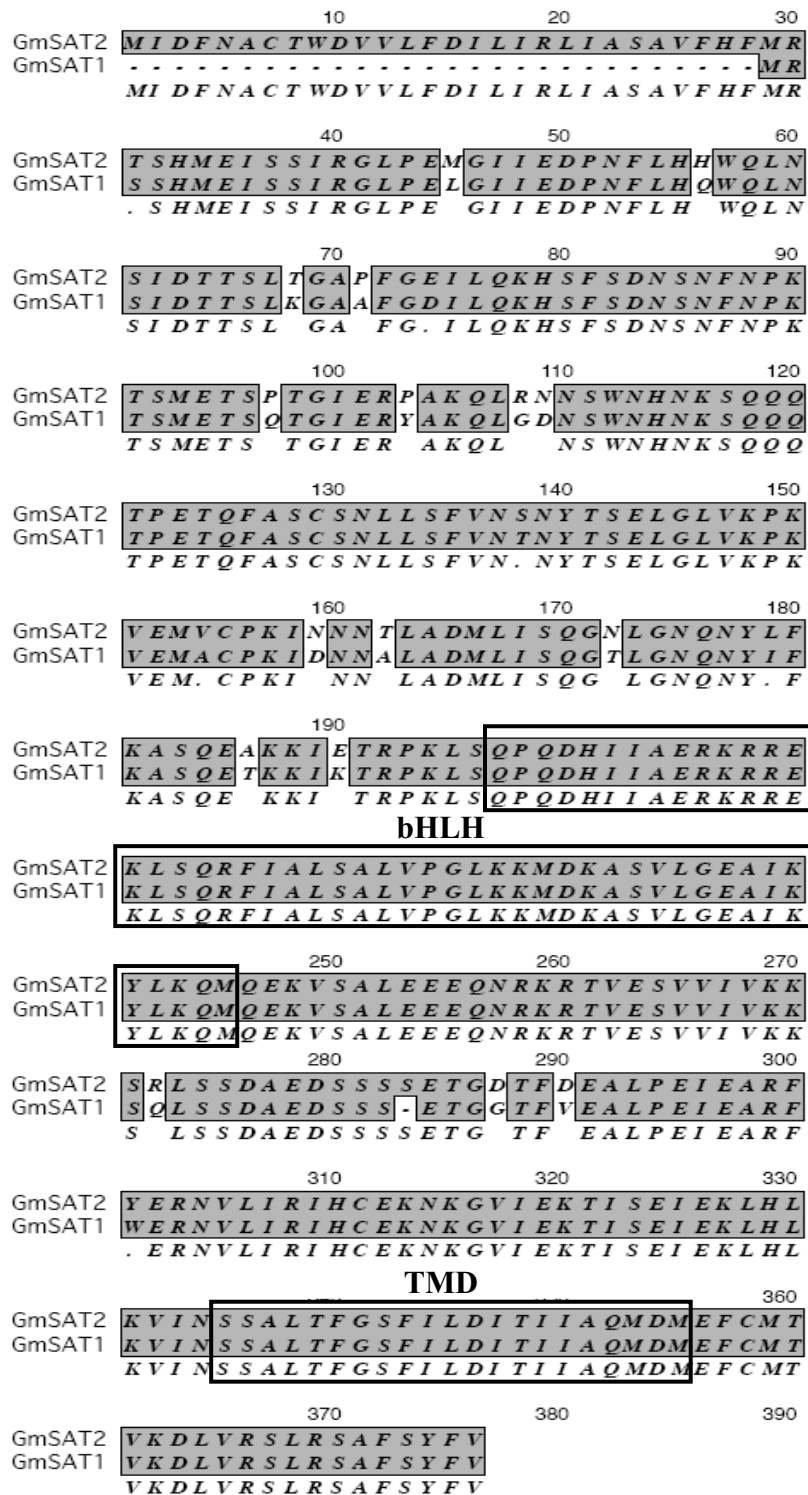
TMD

```
                                                      360
GmSAT2  K V I N S S A L T F G S F I L D I T I I A Q M D M E F C M T
GmSAT1  K V I N S S A L T F G S F I L D I T I I A Q M D M E F C M T
        K V I N S S A L T F G S F I L D I T I I A Q M D M E F C M T

                      370             380             390
GmSAT2  V K D L V R S L R S A F S Y F V
GmSAT1  V K D L V R S L R S A F S Y F V
        V K D L V R S L R S A F S Y F V
```

Figure 5-3. Protein sequence alignment of GmSAT1 and GmSAT2.
The predicted basic helix-loop-helix (bHLH) and transmembrane
(TMD) domains are indicated.