Towards functional analysis of *ENOD40*

Bert Compaan

Bert Compaan

Towards functional analysis of *ENOD40*

Promotor: prof. dr. A.H.J. Bisseling

Hoogleraar Moleculaire Biologie,
Wageningen Universiteit

Co-promotor: dr. H.G.J.M. Franssen

Universitair Docent Moleculaire Biologie,
Wageningen Universiteit

Samenstelling promotiecommissie:

Dr. Ir. A. Goverse (Wageningen Universiteit)

prof. dr. J.A.M. Leunissen (Wageningen Universiteit)

prof. dr. H.P. Spaink (Universiteit Leiden)

dr. ir. R. Heidstra (Universiteit Utrecht)

Dit onderzoek is uitgevoerd binnen de onderzoekschool Experimentele Plantwetenschappen (EPS)

Bert Compaan

Towards functional analysis of *ENOD40*

Proefschrift

ter verkrijging van de graad van doctor

op gezag van de rector magnificus

van Wageningen Universiteit,

Prof. dr. ir. L. Speelman,

in het openbaar te verdedigen

op woensdag 15 December 2004

des namiddags te half twee in de Aula.

Towards functional analysis of *ENOD40*

Bert Compaan

Thesis Wageningen University, The Netherlands

With references – with summary in Dutch

ISBN 90-8504-123-6

Table of contents

Scope

Legumes interact with mycorrhizal fungi and soil-borne rhizobia. In the latter case, this leads to the formation of a complete new organ on the roots of the host plant, the root nodule, in which the bacteria are hosted. Genetic screens have revealed that a limited number of legume genes is essential for initiation of the symbiotic interaction. However, the expression of many genes, formerly annotated as nodulin genes, is highly induced during nodulation. This suggests that a small number of symbiosis specific genes activates the expression of a large number of genes that are recruited from modules, like cell division, to establish a successful symbiosis. None of the known nodulin genes were identified in genetic screens, indicating that they are not essential for nodulation, that mutations in these genes are lethal or that the nodulin genes are redundant.

We have chosen to study the nodulin *ENOD40*, in order to clarify this problem and to elucidate in which module this gene operates during the early phases of nodule development. Chapter 1 gives a short introduction, in which I describe some recent advancements in symbiotic research and the knowledge on *ENOD40* known at the time I started my thesis research.

As *ENOD40* expression is highly elevated in legumes shortly after inoculation of plants with *Rhizobium*, *ENOD40* could be part of the mechanism by which the bacterium-secreted Nod factors trigger nodule initiation. If this would be true, *ENOD40* expression should precede cell division in the root cortex. In Chapter 2 the experiments are described investigating the timing of *ENOD40* expression during nodulation.

To determine whether *ENOD40* is crucial for nodule development, we used RNA interference to study the effect of downregulating *ENOD40* expression on nodule development in *Medicago truncatula* plants (Chapter 3).

Knowledge of the nature of the molecule that represents *ENOD40* activity could be informative about the module *ENOD40* is part of. Comparison of the nucleotide sequences of several *ENOD40* homologues, revealed that the *ENOD40* transcript lacks a long ORF. Instead, two domains conserved among *ENOD40* transcripts can be distinguished (Vijn *et al*., 1995). The first region (box1) contains a highly conserved ORF and the putative start codon of this ORF is the most 5'-localised one in most *ENOD40* genes. Box1 encodes for a putative small peptide of about 12 amino acids (Vijn *et al*., 1995). The second region (box2) however does not contain a conserved ORF. At the time the research described in this thesis started, the absence of a long ORF in the *ENOD40* transcript raised questions on whether the function of this region

was RNA or protein mediated. I therefore investigated whether a putative start codon in this region is able to initiate translation using *ENOD40-GFP* fusions. The experiments are described in Chapter 2 and 6.

Resolving the function of *ENOD40* would be greatly facilitated by the availability of knock-out mutants in *ENOD40*. Because systems that enable the identification of this mutant were not available in legumes, we used a non-legume that also contained *ENOD40* to identify such a mutant. I analysed the phenotype of an *ENOD40* knock out in *Zea mays*, the first described mutant in *ENOD40*, identified in a genetic screen (Chapter 4).

In the last chapter (Chapter 7), I will discuss the obtained results and place them in the context of the results obtained in parallel by others.

Chapter 1

General Introduction

Legumes have the unique ability to establish a symbiosis with nitrogen fixing *Rhizobium* bacteria. In the legume-*Rhizobium* symbiosis, bacteria are hosted in newly formed organs: the root nodules. The formation of these nodules requires developmental reprogramming of root cortical cells as well as the formation of infection threads through which the bacteria penetrate the root.

Nodulation starts when bacteria are attracted to plant roots: here rhizobia induce the curling of root hair cells and via a tubular structure formed within these root hairs, the so-called infection thread, the bacteria are able to reach the root cortex. Meanwhile differentiated cells in the root cortex, dedifferentiate and start to divide. These dividing cortical cells form a nodule primordium. When the infection threads reach the primordium, they ramify and the rhizobia are released in the primordium cells. In e.g. *Lotus japonicus* and soybean, shortly after infection, mitotic activity of primordium cells will cease and therefore these nodules have a determinate growth pattern and are called determinate nodules. In other plants, like pea and *Medicago* sp., cells at the base of the primordium are infected. The non-infected cells at the distal part of the primordium will form the nodule meristem, that remains mitotically active in the growing nodule. Basipetally, cells of the meristem differentiate into infected and non-infected cells of the central tissue and into cells of the pheripheral tissues that enclose the central tissue. Inward-outward the peripheral tissues are a boundary layer, the nodule parenchyma in which the vascular bundles are embedded, the nodule endodermis and at the outside the nodule cortex. As a consequence of persistent meristematic activity, these nodules have an indeterminate growth. Therefore, the tissues are of graded age and can be divided into zones. From the distal end to the proximal base; zone I represents the meristem, zone II is the infection zone. Here cells are infected and bacteria proliferate. At the proximal part of this zone both plant cells and bacteria are differentiating and in cells of zone III fully differentiated bacteria, called bacteroids, reduce nitrogen to ammonia, which can be used by the plant. In zone IV plant cells senesce.

To unravel the mechanisms underlying root nodule formation genetic approaches involving *Rhizobium* or the legume host have been followed. Mutants of *Rhizobium* impaired in establishing a proper symbiosis were used to identify the bacterial genes crucial for symbiosis. Among them, the *nodD* gene

codes for a protein involved in the recognition of plant secreted (iso)flavonoids. The (iso)flavonoid activated NodD protein, acts as transcriptional activator for the other bacterial *nod* genes that code for proteins involved in the production and secretion of lipochitooligosaccharides, the so-called Nod factors. These molecules are pivotal for all (early) steps of nodulation and are sufficient to induce responses like cortical cell divisions and root hair deformation.

In genetic screens, genes have been identified, which are indispensable for establishing the symbiosis between rhizobia and legumes. These have been cloned from the model legumes *Medicago truncatula* and *Lotus japonicus*. Several of the genes identified appear to be crucial as well for mycorrhisation of the roots, suggesting that the *Rhizobium*-legume symbiosis evolved from the symbiosis between roots and mycorrhizal fungi that is common to the majority of higher plants.

The identified genes encode LysM-domain containing protein kinases (*LYKs*: Limpens *et al.*, 2003; *NFR*1 and 2: Madsen *et al.*, 2003; Radutoiu *et al.*, 2003), leucine-rich-repeat (LRR) containing protein kinases (NORK, DMI2: Endre *et al.*, 2002), a Calcium-Calmodulin kinase (DMI3: Levy *et al.*, 2004), a putative symporter (DMI1: Ane *et al.*, 2004) and a transcription factor (NIN: Schauser *et al.*, 1999). As in both model legumes and in pea, the same mutants have been identified and also the number of mutants unable to establish a symbiosis is similarly small, it becomes clear that only a small number of legume genes is essential and specific for symbiosis, since the mutants do not show other defects apart from the nodulation phenotype. The notion that only a small number of genes seem to be symbiosis specific, implies that the symbiosis specific genes activate modules that are derived from non-symbiotic processes that are active in both legumes and non-legumes. Within these nodules genes are activated that are expressed at markedly higher levels than in non-symbiotic organs. Initially, such genes were even considered to be nodule specific (nodulins (van Kammen, 1984)). However, in all cases these genes are also expressed in non-symbiotic tissues. Apparently, the alteration of expression of these genes in symbiosis is required to adapt the non-symbiotic processes in which they are involved, for the symbiosis demands. The question then pops up whether the genes, formerly named nodulin, are essential for symbiosis and if so, why they have not been identified in genetic screens? One possibility is that these genes are not essential, as for instance has been demonstrated for *MsENOD12* (Csanadi *et al.*, 1994). A second possibility is that mutations in these genes do not only affect symbiotic but also non-symbiotic processes and this may be lethal. Therefore, mutants in these genes remained unnoticed. A third possibility is gene redundancy. This implies that a mutation in one member of a gene family is complemented by another member of the family. Hence, to study whether nodulins are essential and to identify which role

these genes have in symbiosis, reverse genetics has to be applied. We chose to analyse *ENOD40* as this gene is highly expressed at early and late stages in nodule development. The gene is also expressed in other parts of plants and even in non-legumes, but most notably not in the model plant *Arabidopsis thaliana*.

ENOD40

ENOD40 expression is induced in the root pericycle within hours after plants are inoculated with *Rhizobium*. In the dividing cortical cells that form a nodule primordium the gene is also expressed. In the mature nodule *ENOD40* is expressed in the uninfected cells of the central tissue (Crespi *et al.*, 1994, Matvienko *et al.*, 1994, Vijn *et al.*, 1995), in cells of the boundary layer, and in the pericycle of nodule vascular bundles (Yang *et al.*, 1993, Kouchi *et al.*, 1993). The rapid induction of *ENOD40* expression and the location of this expression indicate that the induction of this gene could be part of the mechanism by which Nod factors trigger nodule initiation. Also functional studies underline this hypothesis: over-expression of *MtENOD40* in *Medicago truncatula* induced extensive cortical cell division in roots in the absence of rhizobia (Charon *et al.*, 1997) and accelerated nodulation upon inoculation with rhizobia (Charon *et al.*, 1999). In *Medicago truncatula* lines with a reduced level of *ENOD40* transcript, as a result of co-suppression by an *ENOD40* transgene, nodulation was strongly reduced (Charon *et al.*, 1999). Furthermore, ectopic expression of *ENOD40* in roots by particle bombardment induced cortical cell divisions and the expression of *ENOD12* (Charon *et al.*, 1997).

If *ENOD40* is part of the mechanism by which Nod factors trigger nodule initiation, this would imply that *ENOD40* expression precedes cell division in the root cortex. Involvement of *ENOD40* in signal transduction has already been proposed based on the nucleotide sequence. Because no long open reading frames (ORF) are present in the transcript, it seems that *ENOD40* function is mediated by the RNA molecule or by a peptide that could be encoded by one of the small ORFs that are present in the transcript. Comparison of the nucleotide sequence of *ENOD40* genes from different plant species has revealed the presence of two conserved regions in this gene, the first contains a short ORF that possibly could encode a peptide. Such peptide might be involved in signalling. A putative start codon preceding this ORF has the capacity to initiate translation as has been shown by translation fusions between this putative start codon and the Green Fluorescence Protein (GFP; Van de Sande *et al.*, 1996). Furthermore, by the use of an antibody raised against a synthetic peptide encoded by the first conserved region the presence of an antigenic determinant within soybean nodule extracts could detected (Van de Sande *et al.*, 1996), providing support for the existence of the encoded peptide *in*

vivo.

The second conserved region has been proposed to fulfil a role in the regulation of translation of the box1 peptide. However, at the time I started the research, there were indications that also the second conserved region was translated (John *et al.*, 1997). Whereas the conservation of the peptide encoded within the first conserved region is high among all *ENOD40* genes known, the conservation of the peptide encoded within the second conserved region is very low. This could imply that the peptide is not translated in a limited number of species, or that the function of the peptide from the second conserved region does not fulfil the same function in *ENOD40* gene containing species. Therefore, more research is required to demonstrate whether this region can be translated. Furthermore, we tested the hypothesis that this region is involved in regulating the translation of the first peptide.

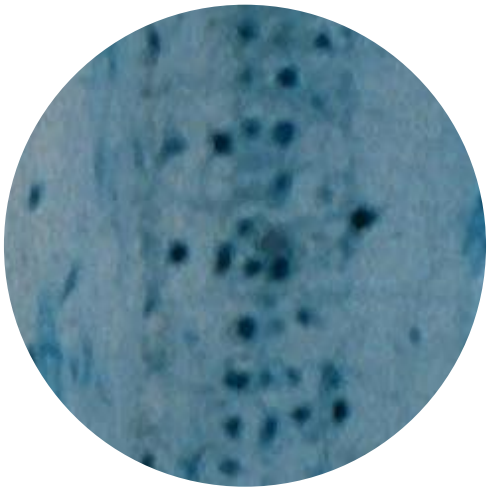
Although *ENOD40* was isolated from legumes, homologous genes were also identified in non-legumes, showing that its activity is not restricted to the nodulation process. Indeed expression in non-symbiotic tissues indicates a broader role for *ENOD40*. During plant development, *ENOD40* expression is observed in the vascular tissue of the root and stem of legumes and non-legumes (Kouchi *et al.*, 1993, Varkonyi-Gasic *et al.*, 2002), in the ovule of *Lotus japonicus* (Flemetakis *et al.*, 2000), in the tomato flower (Vleghels *et al.*, 2003) and during tomato seed germination (Vleghels *et al.*, 2003). These data show that the expression pattern of *ENOD40* is complex. Although a function for this gene can therefore not simply be assigned based on its expression pattern, yet several suggestions for *ENOD40* function have been made. For instance, based on expression patterns in *Lycopersicon esculentum* it was suggested that *ENOD40* has a role in counteracting ethylene provoked responses (Vleghels *et al.*, 2003), while based on the expression pattern in *Glycine max*, *Trifolium repens* and *Oryza sativa* it was suggested that *ENOD40* expression is associated with the formation of transfer cells (Kouchi *et al.*, 1993, Kouchi *et al.*, 1999, Varkonyi-gasic *et al.*, 2002).

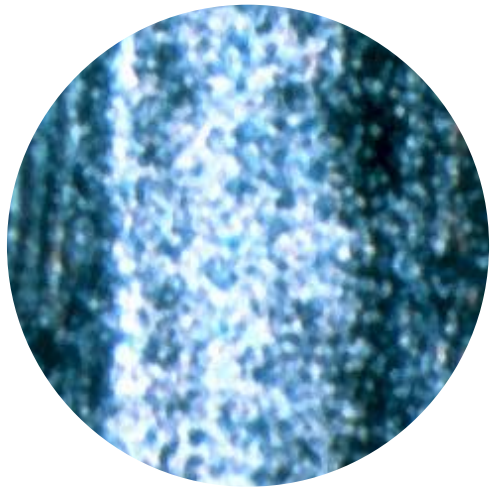
Reverse genetic approaches as over-expression and anti-sense expression have been used to identify the function of *ENOD40.* In *Medicago sativa*, expression of an anti-sense transcript of *MtENOD40* in calli that were grown from explants blocked embryo development. When the sense transcript was over-expressed, the embryos did develop, however, these embryos failed to develop into mature plants, instead of this, teratomas were formed (Crespi *et al.*, 1994). Over-expression of *GmENOD40* in *Nicotiana tabacum* caused the formation of adventitious shoots at the base of the main shoot, indicating that the apical dominance was reduced in these plants (Van de Sande *et al.*, 1996). In tobacco BY-2 cells, it was shown that *ENOD40* influenced the cross talk of the hormones

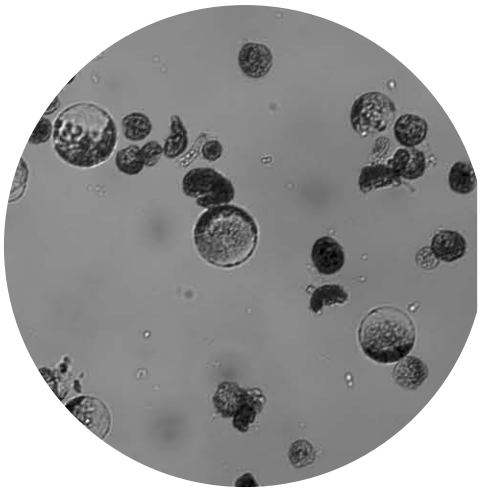
ethylene, auxin and cytokinin, resulting in a suppressed elongation growth in these cells (Ruttink *et al.*, 2003). These experiments show that *ENOD40* has a function in non-symbiotic tissues.

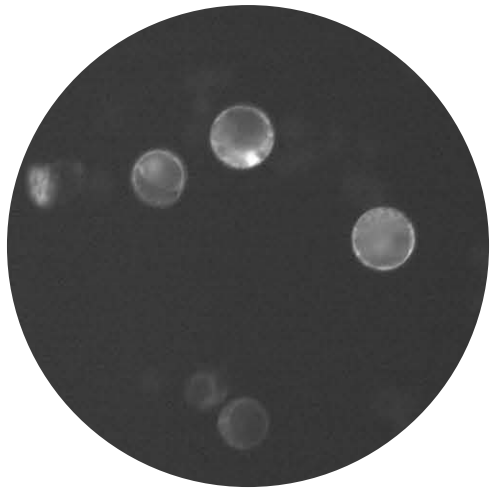
Ectopic expression of *ENOD40* in *Medicago* roots by particle bombardment induced cortical cell divisions and the expression of *ENOD12* (Charon *et al.*, 1997), supporting a role of *ENOD40* in symbiosis. The observed effects of mis-regulation of *ENOD40* expression therefore range from disturbed embryonic development to alterations in symbiotic as well as pathogenic interactions and interference with phytohormone signalling. Thus, it can be concluded that *ENOD40* is involved in many processes, yet the function of *ENOD40* in any of these processes is poorly understood.

So although extensive research has been done in the past, major questions still remain to be answered. Examples are: Is *ENOD40* expression essential for nodulation; in which module is *ENOD40* activity required and what is the biological nature of *ENOD40* activity? It seems essential then that more extensive genetic approaches will have to be used to solve these questions. Such analyses could be done in legumes as well as in non-legumes. In this thesis we describe experiments performed to obtain answers to these questions.









Chapter 2

ENOD40 expression in the pericycle precedes cortical cell division in Rhizobium-legume interaction and the highly conserved internal region of the gene does not encode a peptide

Bert Compaan, Wei-Cai Yang, Ton Bisseling and Henk Franssen

Published in: Plant and Soil **230**: 1-8, 2001

Abstract

In situ expression studies show that *MsENOD40* is expressed in pericycle cells of alfalfa roots within 3 hr after addition of *Sinorhizobium meliloti* 1021. This is about 15 hr before cortical cell divisions become apparent. All *ENOD40* clones isolated so far share two conserved regions, box1 and box2. Translational fusions of *GFP* to potential open reading frames within the conserved regions of *NtENOD40* have been expressed in Cowpea protoplasts to study which open reading frames are translated. In this way it was shown that only the first AUG of the transcript, present in box1, is used as a translational start in Cowpea protoplasts. The open reading frame in this conserved region in *ENOD40* transcripts encodes a peptide of 10–13 amino acids. In legumes, we suggest a role for the ENOD40 peptide as a signal molecule in the formation of the nodule primordium.

Introduction

In the symbiosis of legumes with rhizobia (reviewed in e.g. Geurts *et al.*, 1996; Hirsch *et al.*, 1992; Mylona *et al.*, 1995), bacteria are hosted in a new organ, the root nodule, where they fix nitrogen. Upon inoculation with compatible rhizobia, cortical cells located opposite the protoxyleme poles divide and form a nodule primordium. Centres of dividing cortical cells arise either in the outer cortex, like for instance in *Lotus japonicum* or in the inner cortex, as occurs in *Medicago* species. Several experiments indicate that a distortion of the auxin and cytokinin homeostasis in the root is a crucial early event in the formation of root nodules (Hirsch *et al.*, 1989; Hirsch *et al.*, 1997; Mathesius *et al.*, 1998). However, this step in nodule development is shared with several other plant developmental processes. Thus the question arises as to whether additional plant derived molecules, others then the classical phytohormones, act as growth regulators in nodulation. A good candidate for such a growth regulator might be the product of *ENOD40*, one of the plant genes that are specifically induced during nodule organogenesis (Kouchi *et al.*, 1993; Yang *et al.*, 1993), ectopic expression of this gene in legumes and non-legumes has a clear effect on plant development (Charon *et al.*, 1997; Crespi *et al.*, 1994; Van de Sande *et al.*, 1996).

If *ENOD40* indeed has a role in redirecting the root developmental program, its expression should precede or coincide with cortical cell divisions required for nodule primordium formation. Using *MsENOD40* promoter-*GUS* fusions, *ß*-glucoronidase activity could be detected in cortical and epidermal cells

A

```
Mt taaggatgaagcttctttgttggaaaaatcaatccatggttcttaaaacaa 91
Lj    tagacatgagattttgttggcaaaaatccatccatggctcttgaagaag 897
Nt          aaaaaatgcagtggatgaagcaatccatgggtcttaggaaaa 106
Zm    agaccatggaagatgcatggcttgagcatctgcatggttcttgagagtt 86
```

B

```
Mt     MKLLCWQKSIHGS
Lj      MRFCWQKSIHGS
Nt        MQWDEAIHGS
Zm      MEDAWLEHLHGS
```

C

```
Mt agaagcttttgttatagcatggcaaaccggcaagtcacaaaaaggcaatggattcctt  tttggagtcttaatggct 337
Lj agaagctttggctacagcctggcgaaccggcaagtcacagaaaggcaatggaccccat  taggggttcttatggcta 1190
Nt aaggatgttggcaaaagcatggcgaaccggcaagtcacgaaacggcaatggactccgt  tttggagtcttttcttggc 340
Zm tggtggtcagctgaggtgcgcacaaaccggcaagtcacaatacggcaatggcggactccattaggagtcgtccttttt 205
```

Figure 2.1 Sequence alignment of regions box1 and box2 conserved among *ENOD40* genes, sequences and position numbers as deposited in GenBank. The accession numbers are: *M. truncatula*: X80264, *L. japonicus*: AF013594, *N. tabacum*: X98716 and *Z. mays*: W21740 (Panel A, B), AI001271 (Panel C). Residues conserved among all sequences are presented in bold. **A.** Alignment of nucleotides in the conserved region located at the 5'-end of the transcript, box1; **B.** Alignment of the amino acid sequences of the open reading frame present in box1, representing the *ENOD40* peptides; **C.** Alignment of nucleotides in the second conserved region, box2, present in *ENOD40* transcripts.

of *M.sativa* before cortical cell divisions appear (Fang *et al.*, 1998). Since promoter-reporter fusions do not always reflect the expression of a gene accurately (Taylor *et al.*, 1997), we studied the time course of expression of *MsENOD40* in relation to cortical cell divisions, using *in situ* hybridisation. To date, *ENOD40* has been identified in both legumes and non-legumes indicating that *ENOD40* not only has a function in nodule formation. Alignment of available *ENOD40* nucleotide sequences shows that the gene contains two short conserved regions, box1 (Figure 2.1A) and box2 (Figure 2.1C), embedded within non homologous sequences. As a result of this, no long conserved open reading frame (ORF) occurs in *ENOD40* transcripts, suggesting that the gene might be functional as an RNA (Crespi *et al.*, 1994; Matvienko *et al.*, 1994). However, within box1, a small ORF is present, which is conserved among all species (Figure 2.1B). Translational fusions between this small ORF of *GmENOD40* and *GFP* showed that this ORF is recognised by the plant translational machinery (Van de Sande *et al.*, 1996). Furthermore, the presence of

the *GmENOD40* peptide in soybean nodule extracts was demonstrated with antibodies raised against the synthetic *GmENOD40* peptide (Van de Sande *et al.*, 1996). Thus, it is likely that this peptide is responsible for (at least some) biological activities of *ENOD40*. Strikingly though, ballistic introduction of DNA constructs lacking box1 into *M. truncatula*, led to the induction of cortical cell divisions and the induction of the *GUS* gene driven by the *MsENOD12* promoter (Charon *et al.*, 1997). Although the second region is highly conserved among all *ENOD40* genes at the nucleotide level, it is lacking a conserved ORF. Nevertheless it was claimed that a synthetic peptide derived from this region of *NtENOD40* had a biological activity (John *et al.*, 1997). Although this paper has been retracted (John *et al.*, 1998), it remains unsolved whether the activity rendering in the second conserved region (Charon *et al.*, 1997) is RNA or protein mediated. Therefore we explored whether a second peptide can be translated from *NtENOD40* transcripts. To facilitate identification of such a translation product, tags like GUS, GFP and myc are commonly used. Tagging for example is used in genetic screens in order to find genes with a specific expression pattern (Sundaresan *et al.*, 1995) or in studying protein-protein interactions. In this study GFP is used, which can be visualised easily in living Cowpea protoplasts.

Results

MsENOD40 expression in alfalfa roots precedes *S. meliloti 1021* induced cortical cell divisions

The time course of appearance of *MsENOD40* mRNA during nodule development was studied by spot inoculation of *M. sativa* roots with *S. meliloti 1021* (Cooper *et al.*, 1994). *ENOD40* expression was studied by *in situ* hybridisation of longitudinal sections of the harvested root segments. Already 3 hr after inoculation an increase in the concentration of the transcript in the pericycle could be detected (Figure 2.2, panel 3 hr). At prolonged post-inoculation times, the mRNA accumulated in pericycle cells (Figure 2.2, panels 5 hr–48 hr). Cortical cell division was first apparent around 20 hr of inoculation (Figure 2.2, panels 20 hr, arrowhead). From this time on, *MtENOD40* was expressed in the dividing cortical cells that make up the nodule primordium (Figure 2.2, panel 48 hr). Thus *MsENOD40* expression in pericycle cells precedes cortical cell division by about 18 hr.

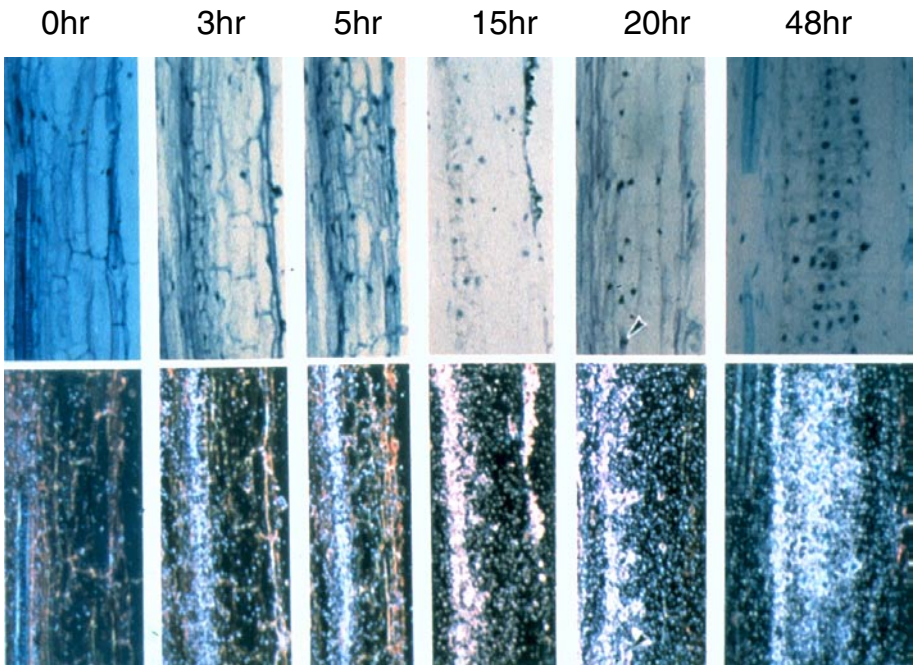


Figure 2.2 Time course of expression of *MsENOD40* transcripts in *M. sativa* roots after spot inoculated with *S. meliloti* 1021. Upper panels: bright field micrograph of a longitudinal section of a spot inoculated root segment. Time of harvesting after application is indicated above the panels. Lower panels: dark field micrographs corresponding to the longitudinal sections presented in the upper panels. Silver grains representing hybridisation signal are visible as white spots. Sections were hybridised with ^{35}S-labelled antisense RNA probes transcribed from the insert of a plasmid containing *MsENOD40*. The arrowhead in the panels, representing the stages harvested 20 hr after bacterium application, points to the first visible dividing cortical cell.

Only box1 of *NtENOD40* is translated

To test which reading frame(s) of *NtENOD40* are being used in plants, several fusions were made with *GFP*. *NtENOD40* was selected for these studies, since it has been claimed that two peptides are encoded by this particular gene (John *et al.*, 1997); one by box1 and the second by box2. Indeed within both boxes an AUG that might serve as a start codon for translation could be recognised. GFP was therefore cloned in frame with either of the AUG codons (Figure 2.4A). Subsequently Cowpea protoplasts were transfected with these fusion gene containing DNAs. Gene expression was analysed in Northern blots, containing RNA isolated from protoplasts transfected with the different constructs (Figure 2.4D). As is shown in Figure 2D, transfection of each construct resulted in detectable levels of GFP RNA. To check if GFP was produced, protoplasts were

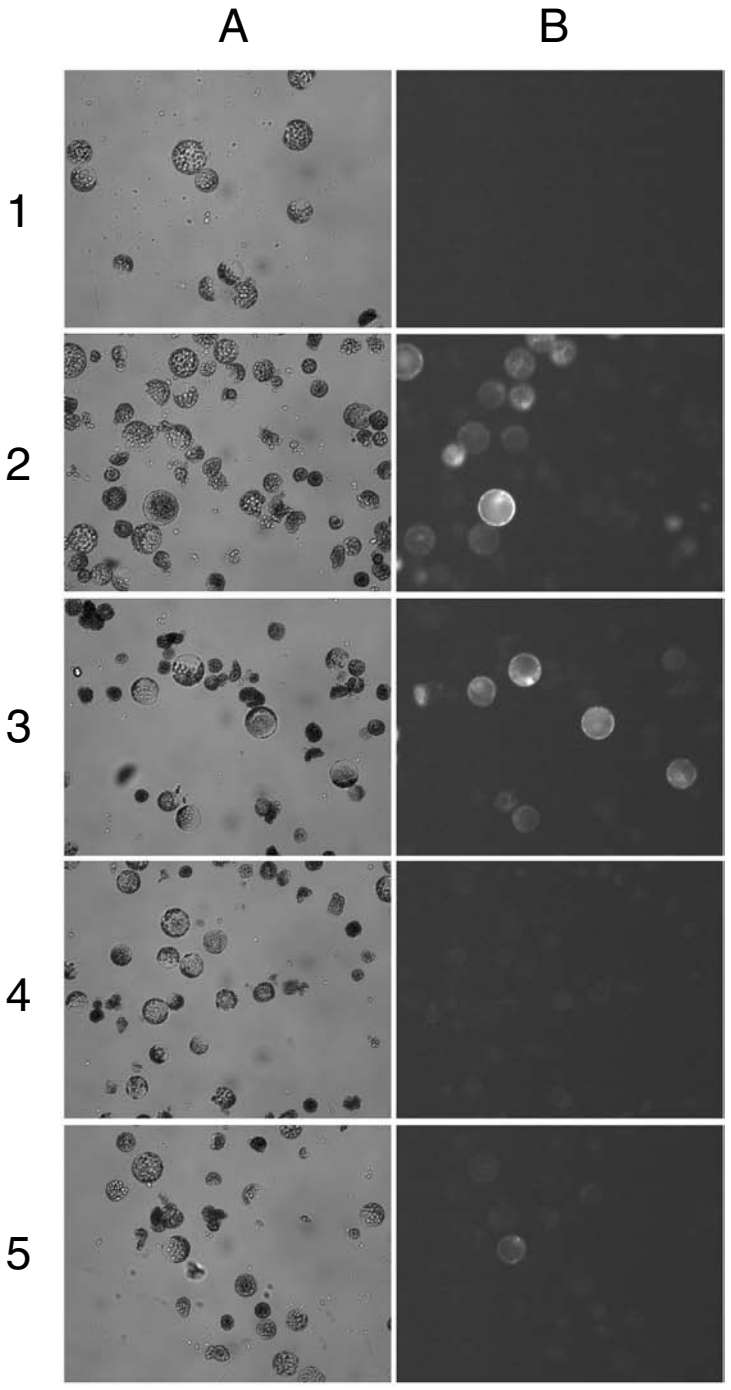


Figure 2.3 Detection of *NtENOD40-GFP* fusion products in Cowpea protoplasts. 24 hr after transfection, expression of GFP was scored by looking for GFP fluorescence. Photographs were taken under bright field (column A) or with the GFP filter (column B). The numbers at the left correspond to the numbers of the constructs (Figure 2A) used for transfection.

analysed by fluorescence microscopy. To reduce auto fluorescence, protoplasts were excitated at 470–500 nm and the emission below 505 nm and above 600 nm was filtered out. In this way auto fluorescence was very low and it could be demonstrated that 60–70% of the protoplasts show GFP fluorescence when transfected with construct 2 (Figure 2.3-2B) or construct 3 (Figure 2.3-3B). This confirmed the earlier observation, that box1 was translated (Van de Sande *et al.*, 1996). In the present experimental set up this construct functions as a positive control. In contrast, protoplasts transfected with construct 4, show no GFP fluorescence (Figure 2.3-4B), which is obviously not due to the amount of transcript present (Figure 2.4D). This shows that the second conserved region of *ENOD40* is not translated, when it is an integral part of the mRNA. To find out whether the AUG of box2 can function as a start of translation at all, protoplasts were transfected with construct 5, which is lacking box1. In this case about 10% of the protoplasts show fluorescence (Figure 2.3-5B), but GFP fluorescence intensity was very low compared to that of protoplasts transfected with construct 3. To quantify the translational efficiency the amount of GFP was determined in Western blots. The amount of GFP in protoplasts transfected with construct 2 and 3 was comparable (Figure 2.4B, lanes 2 and 3). The difference in molecular sizes of the GFPs, is a result of the fusion of the NtENOD40 peptide at the N-terminus of GFP. No GFP could be detected in protoplasts transfected with either construct 4 or 5 (Figure 2.4B, lanes 4 and 5). So the mRNA transcribed from construct 5 is translated far less efficient than that of mRNA transcribed from construct 2 or 3. However in a few protoplasts GFP can reach a concentration that is high enough to be visualised with a fluorescence microscope. These experiments confirm that the AUG of the ORF spanning box1 is used for initiation of translation. Furthermore, translation initiation from the AUG of the ORF of box2 occurs only in constructs in which box1 has been removed, albeit with a low efficiency compared to box1. However in the natural occurring situation where box1 is preceding box2, this efficiency is even more reduced as demonstrated in protoplasts transfected with construct 4, where GPF could not be detected, neither by fluorescence microscopy (Figure 2.3-4B) nor by Western blot analysis (Figure 2.4B, lane 4).

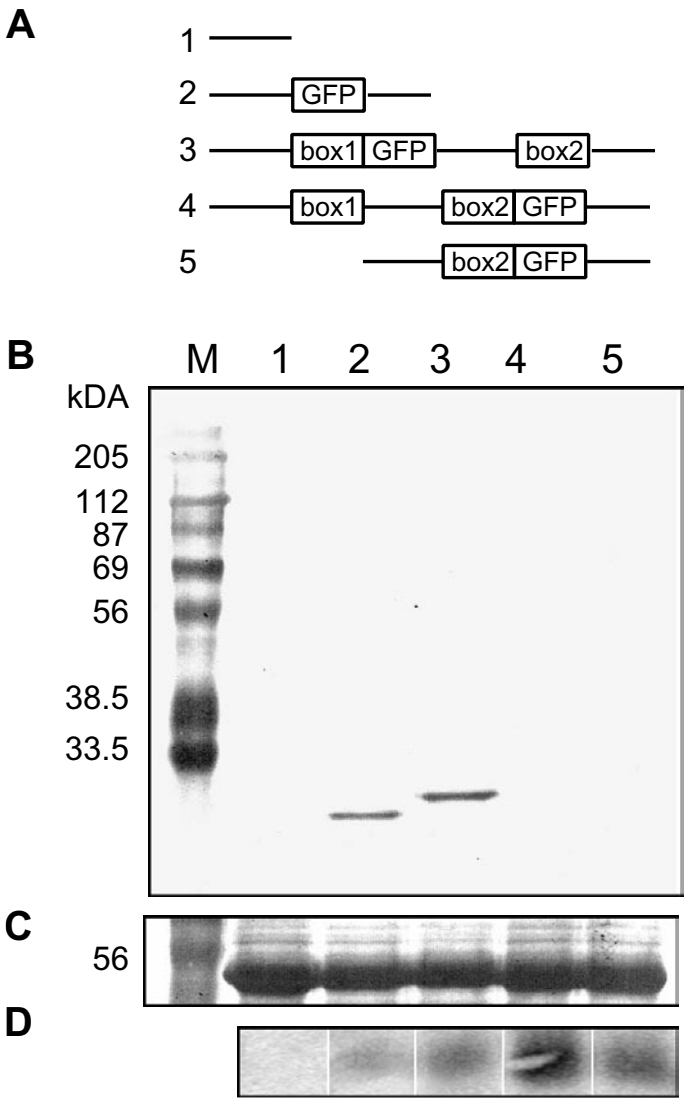


Figure 2.4 Analysis of GFP expression in Cowpea protoplasts. The use of putative AUG codons within *NtENOD40* was investigated by transfection of Cowpea protoplasts with constructs where *GFP* is cloned in frame to either of the codons. **A**. Schematic representation of the constructs used; **B**. Immunodetection of GFP after Western blotting of protein samples of transfected protoplasts. Numbers above the lanes correspond to the numbers of the constructs used and depicted in A. Numbers at the left refer to the molecular sizes of marker (M) proteins; **C**. Protein staining with Ponceau S of part of the Western blot, to verify equal loading of proteins. The numbers at the left refer to the size of the molecular marker; **D**. Northern blot containing RNA isolated from protoplasts transfected with the constructs as shown in A, and probed with ^{32}P-labelled *GFP* DNA.

Discussion

Here, we have shown that the expression of *MsENOD40* is induced within 3 hr in pericycle cells. This is markedly before the first cortical cell divisions become apparent at this site. Our data confirm and extend *MsENOD40* promoter-*GUS* studies. These data, together with the observation that ectopic expression of *ENOD40* leads to perturbation in plant growth (Charon *et al.*, 1997; Crespi *et al.*, 1994; Van de Sande *et al.*, 1996), suggest a role for this gene in the formation of the nodule primordium. Furthermore, our data show that it is very unlikely that in addition to the AUG at the beginning of the ORF of box1 the AUG of box2 of *NtENOD40* is recognised by the plants translational machinery. The notion that eukaryotic mRNAs are monocistronic, together with the fact that the nucleotide sequence of box2 among all *ENOD40* clones is very conserved, but not the putative start and stop codons, makes it unlikely that box2 sequences may function as a template for protein translation. Hence, this implies that the biological activity of *ENOD40* box1 resides within the peptide encoded by this region while the activity ascribed to box2 (Charon *et al.*, 1997) is RNA mediated. Within the various *ENOD40* genes, the peptide encoded by this region consists of 10–13 amino acids. In other organisms than plants, small peptides are often active as signal molecules. In analogy with this, we propose a role as a signal peptide for ENOD40. However, it will demand an integrated biochemical, molecular and genetic approach to dissect the role of *ENOD40* in nodule development and to define the role of this gene in plant development.

Materials and methods

In situ hybridisation

M. sativa plants were cultured and spot inoculated (Cooper *et al.*,1994). Roots were harvested at time points 0, 3, 5, 15, 20, 48 hr after spot inoculation with *S. meliloti 1021*. Roots were fixed in 4% paraformaldehyde supplemented with 0.25% glutaraldedyde in 10 mM sodium phosphate buffer for 4 hr (Van de Wiel *et al.*, 1990), dehydrated and embedded into paraffin by routine methods. Longitudinal sections of 7 μm thick were hybridised with [^{35}S]UTP (1000–l500 Ci mmol^{-1}Amersham) labelled antisense or sense RNA probes (Scheres *et al.*, 1990). Sections exposed for 2 to 4 weeks were stained with toluidine blue and photographed with a Nikon microscope equipped with epipolarisation.

Construction of *NtENOD40-GFP* fusions

For all constructions the smRS-GFP (soluble modified Red Shifted green fluorescent protein (Davis *et al.*, 1998) was used.

All constructs were cloned in the *XbaI*/*BamHI* site of pMON999 (Van Bokhoven *et al.*, 1993) behind the 35S promoter. The fusion constructs all consist of a *5'-ENOD40* part, the *GFP* coding sequence and a *3'-ENOD40* part. So in fact the *GFP* coding sequence was inserted in the *ENOD40* gene. The position of *GFP* within *NtENOD40* in the various constructs is depicted in Figure 2.4A.

The *GFP* coding sequence for these constructs, lacking its own start codon, was amplified with primers GFP-1 and GFP-2, including *HindIII* and *KpnI* sites. The *5'-ENOD40* part was obtained by PCR using primers with *XbaI* and *HindIII* sites and the *3'-ENOD40* part was amplified using primers supplied with *KpnI* and *BamHI* sites. For the box1-GFP construct (Figure 2.4A, 3), the *5'-ENOD40* part was amplified with primers 1 and 2 and the *3'-ENOD40* part was amplified with primers 3 and 8. To obtain the box2-GFP construct (Figure 2.4A, 4), the *5'-ENOD40* part was amplified with primers 6 and 1 and the *3'*-part was amplified with primers 7 and 8.

The box2-GFP minus box1 construct (Figure 2.4A, 5) was created by a PCR reaction on box2-GFP with primers 9 and 8, and after digestion of the PCR product with *XbaI* and *BamHI*, the DNA fragment was cloned into pMON999.

As a control, the GFP coding sequence, amplified with primers GFP-3 and GFP-4, was used. The original start codon of GFP was present in this construct.

All cloning steps have been verified by sequence analyses before being used in protoplasts transfections.

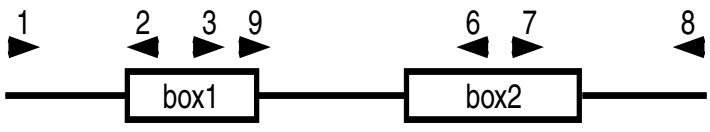


Figure 2.5 Primer scheme.

Primers

```
GFP-1:       5′-CCCAAGCTTGGGCAAAGGAGAAGAACTT-3′
GFP-2:       5′-GGGGTACCTTATTTGTATAGTTCATCCAT-3′
GFP-3:       5′-GCTCTAGAGCCATGGGCAAAGGAGAAGAACTTT-3′
GFP-4:       5′-CCGGATCCTTATTTGTATAGTTCATCCATGC-3′
NTENOD40-1: 5′-GCTCTAGACTAGCTTGTCTCAAGAAC-3′
NTENOD40-2: 5′-CCCAAGCTTTCATCCCACTGCATTTTTTC-3′
NTENOD40-3: 5′-GGGGTACCCATGGGTCTTAGAAAAG-3′
NTENOD40-6: 5′-CCCAAGCTTTGAGTCCATTGCCGTTTCGTG-3′
NTENOD40-7: 5′-GGGGTACCACGAAACGGCAATGGACT-3′
NTENOD40-8: 5′-CGGGATCCATGACAATCTTAACAACTCTAT-3′
NTENOD40-9: 5′-GCTCTAGAGCAATCCATCGGTCTTAGG-3′
```

The position of the various NTENOD40 primers within the cDNA clone is depicted below.

Transfection of protoplasts

Vigna unguiculata cv. California Blackeye was grown in growth chambers (16-hr light/8-hr dark periods at 25°C).

Mesophyll protoplasts were isolated from primary leaves of 9–l0 days old Cowpea and 10^6 protoplasts were transfected with 10 μg vector DNA using polyethylene glycol (Van Bokhoven *et al.*, 1993). Finally, protoplasts were incubated under continuous illumination at 25°C. After 24 hr of incubation, aliquots of each protoplast sample were used for RNA isolation, fluorescence microscopy and protein analysis.

Fluorescence microscopy

Fluorescence microscopy was performed with the FRIM system (Gadella *et al.*, 1997) based on a Leica DMR microscope (Leitz, Wetzlar, Germany) with Leitz fluotar 10x NA 0.3 air or fluotar 40x NA 0.5–1.0 oil immersion objectives. Images were captured by a Quantix CCD-camera (Photometrics, Tucson, AZ, USA) interfaced through a PCI-card to an Apple Macintosh PowerPC 8500/180 computer (Apple computer, Cupertino, CA, USA) and controlled by IPLab 3.1 software (Signal analytics, Vienna, VA, USA).

Fluorescence of GFP was acquired by excitation with a 100 W USH-102D mercury lamp (Fairlight, Rotterdam, The Netherlands) and an Omega 485DF15 nm bandpass filter. The emission was separated by a dichroic mirror (505 nm) and passed through an Omega 550DF50 bandpass filter.

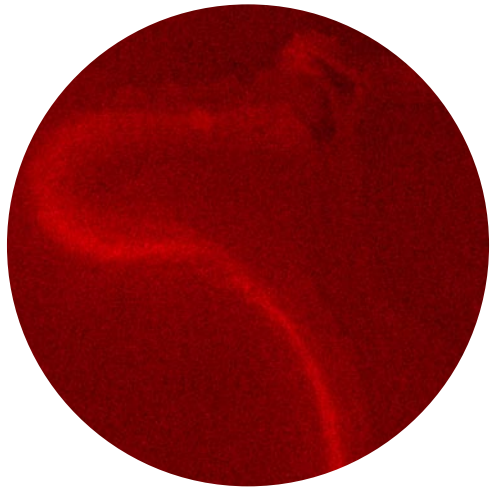
RNA isolation and Northern blotting

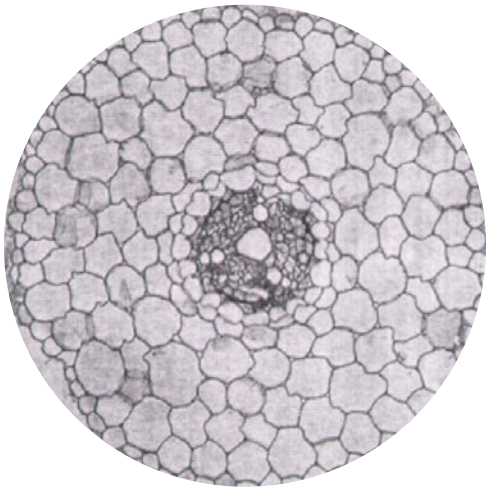
Protoplasts were harvested by centrifugation at 3000 rpm and RNA was isolated using the TRIzol®Reagent (GibcoBRL) according to the manufacturers protocol. RNA was separated on a 1% agarose gel after denaturation with glyoxal and DMSO (Sambrook *et al.*, 1989). After electrophoresis, the gel was blotted onto genescreen (Amersham) membranes, and the membranes were hybridised to ^{32}P-labelled GFP probe at 42°C for 16 hr, subsequently washed and exposed to a storage phosphor screen for one day. The screen was analysed with STORM 840 (Molecular Dynamics) and the ImageQuant software.

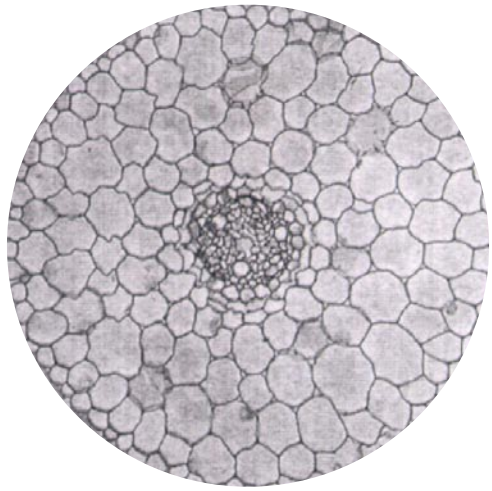
Western blotting

Protoplasts were lysed by adding 10 mM Tris-HCl (pH 7.5) containing Complete™ (Boehringer Mannheim) protease inhibitor cocktail. After centrifugation at 10.000 rpm (4°C) for 10 min. the supernatant, containing the cytoplasmatic proteins, was collected. 90 μg of denatured protein was separated on a 12% SDS-PAA gel and subsequently electroblotted onto nitrocellulose (0.45 micrometer, Schleicher & Schuell). Immunodetection was carried out (Walker *et al.*, 1986), using rabbit anti-GFP (Clonetech) and phosphatase-labelled goat anti-rabbit (GibcoBRL) respectively as first and second antibody.

Finally the blot was coloured with Ponceau-S (Sambrook *et al.*, 1989) to verify equal loading of all samples.







Chapter 3

Knock-down of *ENOD40* affects nodule initiation and development in *Medicago truncatula*

Bert Compaan, Jan Hontelez, Karel Theeuws, Erik Limpens, Carolien Franken, Rene Geurts, Ton Bisseling and Henk Franssen

Abstract

In order to analyse the function of *ENOD40* we used RNAi to down regulate the expression of the gene. The *ENOD40* constructs were introduced by *A. rhizogenes* mediated root transformation in the model legume *Medicago truncatula*. Knock-down of *MtENOD40* resulted in a 5-fold reduction of the amount of *ENOD40* transcript in the transformed roots. When these roots are inoculated with rhizobia, the number of nodules is reduced. Moreover, the nodules that are formed do not have the characteristic rod shaped structure of indeterminate nodules but instead have a spherical shape. Within these nodules the specific zonation of the central tissue is still laid down, but the size of the infection zone is slightly increased while the size of the fixation zone is dramatically reduced. We therefore conclude that *ENOD40* silencing affects nodule initiation and reduces nodule growth.

Introduction

The nodule

Root nodules are specialised organs on the roots of legumes in which nitrogen fixing rhizobia are hosted in an intracellular manner. The formation of these symbiotic organs requires a complex communication between the bacteria and their host plants. The bacteria present in the soil react on exudates from the plant root by colonising the root system and the production of signalling molecules, the so-called Nod factors. These Nod factors elicit alteration of gene expression in the pericycle, the root cortex and the root epidermis at the site of nodulation. As a consequence, deformation of root hairs and the initiation of cell division in the root cortex are induced. Rhizobia are able to enter the root via infection threads that penetrate the curled root hairs. When the infection threads reach cells of the nodule primordium, which is formed as a result of dividing cortical cells, the rhizobia are released into these cells.

After infection of plant cells, the nodules start to grow. Tissues of growing indeterminate nodules are of graded age. From the distal to the proximal end, these rod shaped nodules can be divided in distinct zones based on the morphology of rhizobia and host cells (Vasse *et al.*, 1990) and the expression of plant and rhizobial genes (Scheres *et al.*, 1990). Starting distally, zone I is the meristem which is devoid of rhizobia. Zone II is the infection zone. At the distal part of this zone threads penetrate cells that originate from dividing meristem cells and bacteria are released from the thread and become surrounded by a

plant membrane, the peribacteroid membrane. At this moment, differentiation of rhizobia as well as host cells starts. In the proximal part of zone II, plant cells elongate and the bacteria proliferate. In the infection zone *ENOD12* and *ENOD40* are expressed (Scheres *et al.*, 1990; Vijn *et al.*, 1995). Zone III, the fixation zone, has been defined as the zone where the rhizobia start to fix nitrogen. This was based on cytological analysis of nodules and the timing of nitrogen fixation (acetylene reduction) during nodule development. At the bases of the nodule a senescent zone can be present in which the bacteroids degenerate.

ENOD40

During nodulation several host genes are up regulated, among them *ENOD40* (Yang *et al.*, 1993). In the transcript of this gene no long open reading frames could be identified. Therefore it was suggested that *ENOD40* could encode a small peptide, involved in the signalling processes that operate during nodule development. *ENOD40* shows a complex expression pattern in legumes during symbiosis. The gene is expressed in the pericycle of the root and the expression of *ENOD40* can be induced by purified Nod factors. The expression in the pericycle cells precedes divisions of cortical cells (Minami *et al.*, 1996; Compaan *et al.*, 2001). *ENOD40* is expressed in the first cortical cell divisions and expression is high in cells in the nodule primordium (Kouchi *et al.*, 1993; Yang *et al.*, 1993; Vijn *et al.*, 1995). In indeterminate nodules *ENOD40* is expressed in cells of the infection zone (Crespi *et al.*, 1994; Vijn *et al.*, 1995). In this zone plant cells cease division and start to expand and the bacteria start to proliferate. Next to the expression in cells of the central tissue, *ENOD40* expression is also high in the pericycle of the nodule vascular bundle. The observed increase in expression of *ENOD40* at the onset of nodulation and in cells of the infection zone, where differentiation of host cells and rhizobia is initiated suggests that this gene could play an important role in the developmental program of nodulation.

Evidence for the functional involvement of *ENOD40* in the nodulation process, stems from the analysis of *Medicago truncatula* (Medicago) plants that contain stably integrated *MtENOD40* under transcriptional control of a constitutive *35S* promoter. In plants with an elevated level of *MtENOD40* transcript, the kinetics of nodulation is accelerated and extensive cell divisions of cortical cells in the region close the root tip are observed. In some lines co-suppression led to a reduction in the number of nodules formed and nodules that were formed went into early senescence (Charon *et al.*, 1997; Charon *et al.*, 1999). However, the steps in nodule development that are affected by *ENOD40* knock-down remain unknown. To study in which steps *ENOD40* reduction interferes with

normal nodule development, we used RNA-interference to silence *ENOD40* in the root system of Medicago.

Silencing

RNA silencing comprises a mechanism by which an organism can decrease the number of mRNA molecules in a sequence specific manner. An important step during RNA silencing is the processing of a dsRNA molecule into short interfering RNAs (siRNAs). These siRNAs can induce the degradation of a target RNA with sequence similarity to the siRNAs. RNA silencing can be induced in a cell by introduction of dsRNA or a DNA vector that will give rise to a dsRNA molecule. In plants, several methods have been described by which RNA silencing can be induced. Well known examples are Virus Induced Gene Silencing and the introduction of an antisense RNA molecule. A very effective way of inducing RNA silencing, is the genomic integration of DNA containing an inverted repeat of the target sequence. The RNA that is transcribed from this inverted repeat will form a hairpin of dsRNA, which subsequently can be processed into siRNAs.

Here we have introduced an inverted repeat of *MtENOD40* under control of a constitutive promoter in the genome of Medicago roots. We used *Agrobacterium rhizogenes* mediated root transformation, that does not require the regeneration of complete plants and therefore enables to study the effects of *ENOD40* knock-down solely on roots and nodules.

Using this system we were able to show that *ENOD40* knock-down leads to a decrease in the number of nodules and that the nodules formed are morphologically aberrant.

Results

Construction of the silencing vector

A DNA fragment of 330 bp of the 3'-UTR of *MtENOD40* was amplified with RT-PCR using primers Mt540 and Mt340 that are provided with restriction enzyme recognition sites at their 5' end to facilitate subsequent cloning. Medicago RNA from nodules that were harvested 2 weeks after inoculation with *Sinorhizobium meliloti* was used as template.

The amplified fragments were cloned in two orientations by two sequential cloning steps in pBS-d35S-RNAi (Limpens *et al.*, 2003), using *AscI* and *SwaI* in the first cloning step and *BamHI* and *SpeI* in the second step, respectively.

The inverted repeat preceded by the 35S promoter was released by digestion and subsequently cloned in the binary vector pBINplusgus using *KpnI*/*PacI* restriction enzymes delivering the pBINsil40 vector. The pBINplusgus vector contains DR5::*GUS* as a selectable marker. This pBINplusgus vector was also used as control vector.

Both pBINsil40 and pBINplusgus were transformed to *Agrobacterium rhizogenes* MSU 440, containing the helper plasmid pRiA4 (Jouanin *et al.*, 1989), by means of electroporation.

In roots containing sil40 DNA the expression level of *MtENOD40* is reduced

The Medicago plants that were transformed with *A. rhizogenes*, contain a *MtENOD40::dsRED* transgene (*cam40bc-1* plants)(Compaan *et al.*, this thesis). In these plants, *ds*RED is expressed in the pericycle of the root system reflecting the expression of endogenous *MtENOD40* (Fang *et al.*, 1998). The sil40 induced silencing will also be directed against the *ds*RED transcript because it contains the *MtENOD40* 3'UTR. Therefore, in these plants silencing of *MtENOD40* is accompanied by reduced *ds*RED fluorescence, which easily can be detected with a stereo-macroscope. The availability of these transgenic *Medicago* plants enabled the selection of silenced roots.

To analyse whether RNAi silencing can be induced in the roots of Medicago, 100 *cam40bc-1* plants were infected with *A. rhizogenes* containing pBINsil40 and 50 plants from the same line were infected with *A. rhizogenes* containing pBINplusgus and the latter plants are referred to as control plants. The roots

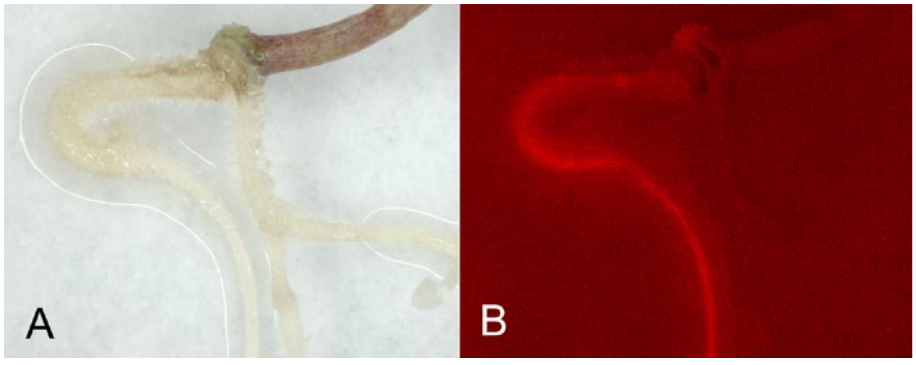


Figure 3.1 Example of a *A.rhizogenes-pBINsil40* transformed root system with two roots emerging from the callus. Panel B shows the *ds*RED fluorescence image indicating that only one root is silenced for the other root is still fluorescent.

that were formed were analysed for their fluorescence properties two weeks post infection. All control roots showed red fluorescence in the pericycle, whereas some of the sil40 transformed roots lacked red fluorescence (Figure 3.1).

To analyse whether the observed reduction in fluorescence is correlated with a reduction in *MtENOD40* expression, root cultures were initiated of 6 non-fluorescent sil40 roots. Such cultures allow the generation of sufficient plant material to detect the expression of the T-DNA selectable marker *DR5::GUS* mRNA level. To confirm co-transformation, ß-glucoronidase (GUS) activity staining was performed on cultured sil40 roots. The DR5 promoter is active in the meristematic zone of the root as has previously been shown in Arabidopsis. In 5 of the 6 root cultures, GUS activity could be detected in the root tips and these 5 cultures were further analysed.

In parallel, root cultures were also initiated from control plants. Because no pre selection based on *ds*RED silencing was performed on these roots, 24 instead of 6 root cultures were made from these plants. In 6 of these 24 root cultures, GUS activity could be detected in the root tips and these were used for further analysis.

To determine whether the *MtENOD40* mRNA levels are reduced in roots that showed reduced *ds*RED fluorescence, RNA was isolated from 2 control and 2 sil40 root cultures. Two microgram of this RNA was used in a RT-PCR reaction using *MtENOD40* specific primers. Analysis of the resulting hybridisation

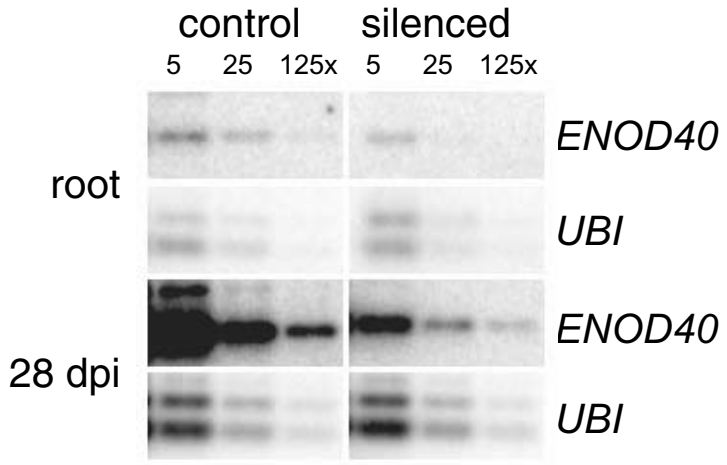


Figure 3.2 RT-PCR on roots and nodules (28 d.p.i.) of both silenced and non-silenced plants using *ENOD40* specific primers. *UBIQUITIN* is used as control. PCR was performed on a dilution range (5, 25 and 125 x) of the cDNA. *ENOD40* transcript levels are 5 fold reduced as a result of silencing.

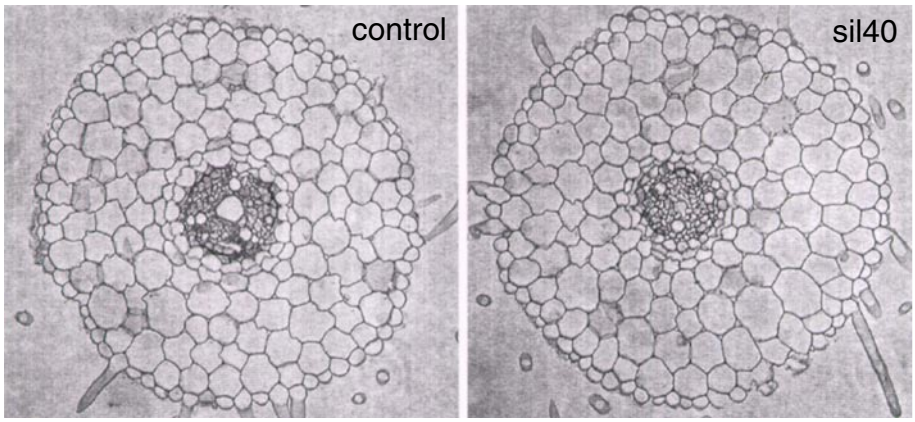


Figure 3.3 Cross section through a control and a silenced root, showing that the size of the cells and the radial organisation is not altered in silenced roots.

signals (Figure 3.2) shows that in the non-fluorescent roots the transcript level of *MtENOD40* was reduced about 5-fold as compared to the brightly fluorescent control roots (Figure 3.2).

To find out whether this reduction of *MtENOD40* transcript level in the pericycle influences the morphology of the root, sil40 and control roots are compared. No obvious differences could be detected in the length of the main root, nor in the number of lateral roots and the overall morphology of the root. For a closer examination, roots were embedded in plastic and 8 μm thick cross sections were made of the roots approximately 2 cm above the root tip (Figure 3.3). These sections show that the size of the cells and the radial organisation is similar in control and sil40 roots. Therefore we conclude that *ENOD40* knock-down in Medicago roots does not affect root morphology.

Reduced *ENOD40* expression in the pericycle correlates with reduced nodule initiation

During interaction of legumes with rhizobia, *ENOD40* is up-regulated in the pericycle at the site of primordium formation, markedly earlier than the division of cortical cells (Compaan *et al.*, 2001). To find out whether *MtENOD40* knock-down has an effect on nodule initiation, nodulation tests were performed on both control and sil40 plants.

Before inoculation with rhizobia red fluorescent roots were removed from sil40 transformed plants, as these roots are not co-transformed. Non-fluorescent sil40 roots could be obtained on 20 of 100 initially transformed plants.

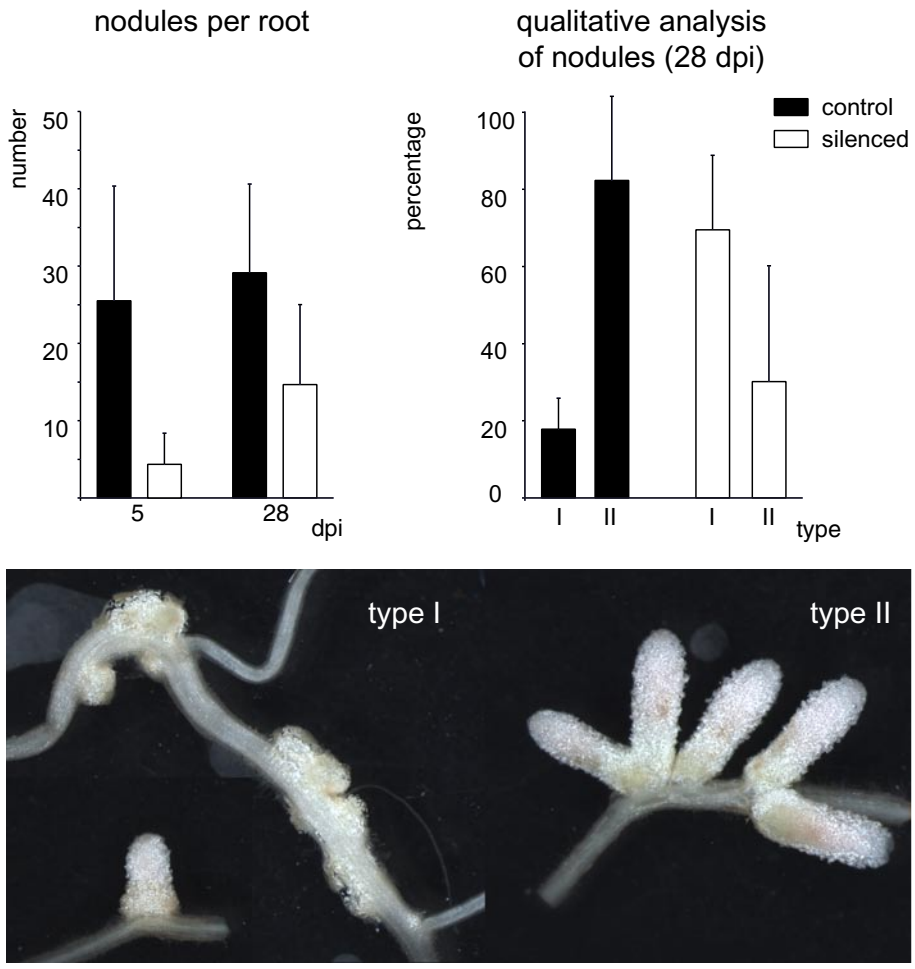


Figure 3.4 Upper right panel: The number of nodules that is initiated on silenced and control roots. Upper left panel: The amount of type I and type II nodules on silenced and control roots. Lower Panel: Examples of type I and type II nodules

For the nodulation test 20 sil40 plants that had formed non-fluorescent roots and 20 control plants, were transferred to gravel. After growing for one week on gravel, the plants were inoculated with GFP expressing *Sinorhizobium meliloti* RM41 (Material and Methods) enabling to determine whether roots were infected.

Five roots were harvested at 5 days post inoculation (d.p.i.) of both the sil40 and control plants. At this time nodules could be observed on the control roots both by bright red fluorescence in the pericycle and nodule primordium and green fluorescence pointing to the presence of rhizobia. The average number of nodules per root is 26 (±15) (Figure 3.4). In three out of five sil40 transformed plants, nodules on roots could be detected that are devoid of red fluorescence. In nodules on two of these three roots we detected green fluorescence

suggesting the presence of rhizobia. In the other two sil40 roots, *ds*RED and green fluorescence could be observed in nodules, although the *ds*RED fluorescence in the pericycle and nodule primordium is markedly lower in these plants than in the control plants. This indicates that the introduction of *ENOD40* in an RNAi fashion prevents induction of *MtENOD40* expression. The average number of nodules per sil40 root is 4 (±4). Compared to the average number of nodules formed on control roots (26 (±15)), this strong reduction in sil40 roots shows that silencing of *MtENOD40* leads to a dramatic reduction in the number of nodules. We next investigated what the effect of *ENOD40* silencing is on the morphology of the nodules that were formed on sil40 roots.

Reduced *ENOD40* expression affects nodule development

After 28 d.p.i., the remaining 15 roots were harvested of both control and sil40 plants. Nodules that are formed on the control roots show a clear red fluorescence that is mainly located in the vascular tissue of the nodules and a clear GFP fluorescence is present in the central part of the nodule, indicating the presence of bacteria. On average 29 (±11)(SD) nodules were formed on each control root, a number similar to the number of nodules on control roots present 5 d.p.i. In nodules that were formed on 8 of the sil40 roots, no red fluorescence could be observed. In nodules that are formed on the other sil40 roots, red fluorescence could be observed although the level of fluorescence is much lower as compared to control nodules. Hardly any GFP fluorescence could be detected in nodules formed on the sil40 roots.

Only 15 (±11) nodules on average were formed on each sil40 root, compared to 29 (±11)(SD) nodules on control roots ($P<0.001$) (Figure 3.4). Compared to the number of nodules present 5 d.p.i. on sil40 roots, the number of nodules increases on sil40 roots but not on control roots. Yet the number of nodules on sil40 roots remains less compared to control roots. This shows that less nodules are formed and that nodule formation is slowed down on sil40 roots, suggesting that autoregulation is becoming operational later in sil40 roots. No significant difference in the number of nodules could be observed between sil40 plants that do have a reduced red fluorescence or do not have. Thus *ENOD40* knock-down affects the number of nodules that are formed on the roots of Medicago plants, an observation in agreement with data published by Charon *et al.* (1999).

Next, we examined the nodules that were formed 28 d.p.i. in more detail. Whereas most of the nodules that are formed on the control roots are elongated rod shaped structures, most of the nodules on sil40 roots are small, sphere

shaped structures. Moreover, the nodules on control plants are pink-coloured, indicating the presence of leghaemoglobin, whereas most of the sil40 nodules are white. To quantify this morphological effect, nodules were classified in 2 types. Type 1 nodules were small sphere shaped white bumps and type 2 nodules were elongated rod shaped pink coloured nodules. This classification of the nodules showed that 70% of the nodules that were formed on sil40 roots, were classified as type I whereas only 18% of the nodules formed on control roots, were classified as type I (Figure 3.4). The observation that the nodules on sil40 plants did not elongate properly and presumably contained less leghaemoglobin indicated that they are hampered in their development.

To quantify the reduction of *ENOD40* mRNA level in nodules, RNA was isolated from several 28 d.p.i nodules harvested from 8 of the sil40 roots in which no red fluorescence could be observed. RT-PCR quantification was performed on *ENOD40* as was described above (Figure 3.2). Compared to the expression in nodules harvested from control plants, *ENOD40* transcript level is 5-fold reduced in nodules formed on sil40 roots.

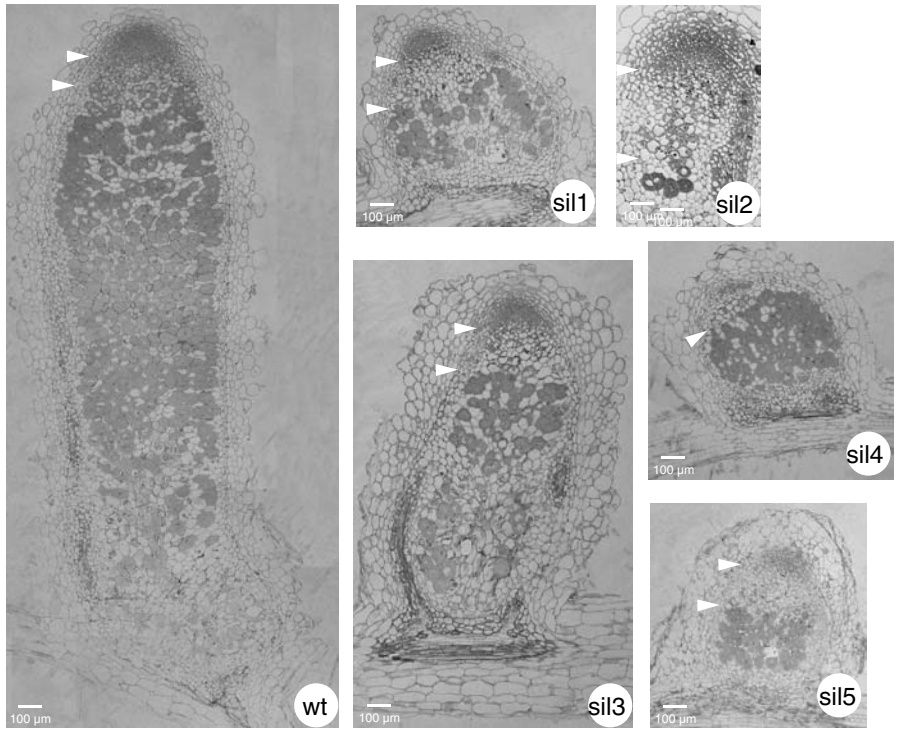


Figure 3.5 Comparison of longitudinal sections of several silenced nodules to a control nodule. Arrowheads indicate the transition from zone I to zone II and from zone II to zone III. In all the nodules the specific zonation can be recognised except for sil4 where the meristem seems to be lacking. The major difference between silenced and control nodules is the enormous difference in size of the nodules.

Microscopic analysis of silenced nodules

To analyse whether the morphological difference of sil40 nodules stems from a disturbed differentiation of cells within the nodules, longitudinal sections were prepared from control nodules and the remaining nodules of the sil40 roots in which no dsRED fluorescence had been observed (Figure 3.5). By comparing the peripheral tissue of both sil40 and control nodules, no differences in tissue organisation nor in cell sizes were observed, although *ENOD40* is normally highly expressed in the pericycle of the vascular bundle located within this peripheral tissue.

Small, uninfected cells that form the meristem are present at the distal end of both control and sil40 nodules. Adjacent to the meristem is the infection zone, containing cells with infection threads and in some of these cells the release of bacteria from the infection threads was observed. It seems that the infection zone of sil40 root contains more cell layers than the infection zone of control nodules. Adjacent to the infection zone, is the fixation zone. In nodules on control roots the fixation zone represents a large zone of cells in which infected cells have reached their maximum size and bacteroids fix nitrogen. However, in nodules on sil40 roots this zone is markedly smaller and is followed by a zone that is characterised by al large amount of dead cells. Compared to the zonation observed in nodules on control plants, we conclude that the zonation that is characteristic of the indeterminate nodule, still can be recognised in the central tissue of sil40 nodules. However, in comparison to control nodules, zone II seems to be enlarged whereas zone III is reduced in sil40 nodules.

Discussion

Here we showed that we could reduce *MtENOD40* expression in the roots of Medicago by an RNAi approach involving *A. rhizogenes* mediated root transformation. We used plants that contain a chimeric *dsRED-ENOD40* transcript and therefore *ds*RED fluorescence could be used as a direct reporter of RNA silencing of *MtENOD40* in the roots. In this way non-transformed, chimaeric and low *ds*RED expressing roots could be removed so that plants were obtained in which *ENOD40* was silenced in the entire root system.

We showed that in roots that were obtained in this manner the endogenous *MtENOD40* transcript levels were reduced 5-fold, implying that a residual transcript level of 20% is left. Silencing of *ENOD40* did not lead to aberrant root development and the root system has a similar tissue organisation as control roots.

Knock-down of *ENOD40* led to a 5-fold reduction in the amount of nodules initiated during the first 5 days after inoculation of plants with rhizobia. Assuming that during the interaction with *S. meliloti* the level of *ENOD40* expression, like in the non-inoculated root, is 20% of the level in non-silenced roots, this observation suggests a role for *ENOD40* in the initiation of nodule organogenesis. This is supported by the high expression of *ENOD40* directly after the application of rhizobia in the pericycle of the root. Since we cannot rule out the complete absence of the biological active *ENOD40* product, we can not be sure that expression of *ENOD40* is essential for nodule initiation. Whereas at 5 or 28 d.p.i the number of nodules on control roots is similar, this number is different on sil40 roots. At 28 d.p.i the number of nodules on sil40 roots is 2-fold less than on control roots. The observation that the nodule number increases in sil40 roots in time and not in control roots can be explained by the inhibition of nodule formation in control roots by auto regulation.

Our results are in line with the analysis of transgenic Medicago lines that show a reduced *ENOD40* expression (Charon *et al.*, 1999) by co-suppression. Also in these lines the number of nodules is strongly reduced. The nodules formed on roots of these lines went into early senescence. Nodules formed on the sil40 roots and harvested at 28 d.p.i., are small, white and spherical shaped in contrast to nodules on the control roots that are large, pink and elongated shaped. Here we show that the typical zonation of indeterminate nodules is still intact in sil40 nodules, but that only a small fixation zone has been formed and therefore sil40 nodules remain small.

As *ENOD40* during nodule development is expressed in cells of the infection zone where differentiation of plant cells and bacteria is initiated, it is likely that the lower expression level of *ENOD40* in these cells interferes with the differentiation of plant cells and bacteria. This may lead to an impaired development of the consecutive fixation zone. The miss-development of the fixation zone is also supported by the absence of the characteristic pink colouring in sil40 nodules indicating a reduced leghemoglobin level, which is supported by a markedly reduced *Lb* mRNA level in the nodule (unpublished results). Whether reduced *ENOD40* mRNA level affects differentiation of *S.meliloti* is not known since we did not investigate e.g. whether *nif* gene expression was reduced. Therefore it remains to be studied whether reduction of *ENOD40* expression in sil40 nodules affects differentiation of rhizobia.

The phenotype of *ENOD40* knock-down shows that *ENOD40* apparently has a regulatory role in nodule growth. Since the size of the cells in the sil40 and control nodules is on average the same, the difference in length between the two types of nodules is due to the presence of less cells in sil40 nodules which is either caused by a reduced cell division frequency or to a cessation

of cell division after initial activity. In indeterminate nodules both symbiotic partners have been shown to be involved in regulation of nodule growth. The reduction of *ENOD40* RNA levels reduces nodule growth. This relies on the activity of the nodule meristem in which *ENOD40* is not expressed. Therefore, a major question is now, how *ENOD40* influences the activity of the nodule meristem. *ENOD40* is expressed in the infection zone that is located in between the meristem and the fixation zone where plant cells and bacteria are fully differentiated. We hypothesize that knock-down of *ENOD40* interferes with the proper development of the infection zone and as a consequence, meristematic activity is ceased. This implies that cells in the infection zone and the meristem communicate to control meristem activity.

Materials and Methods

Primers

Cloning primers

```
Mt540 5′-CCACTAGTGGCGCGCCTTTTGTTATAGCATGGC-3′
Mt340 5′-AATGGATCCATTTAAATAGAAGAAAAGGAACAT-3′
```

MtENOD40 primers

```
fw    5′-GCTCTAGACCCTTTAAGCATCCTCTA-3′
rv    5′-CGGGATCCCACAAACAAACAAGCATAC-3′
```

UBIQUITIN primers

```
fw    5′-ATGCAGAT(C/T)TTTGTGAAGAC-3′
rv    5′-ACCACCACG(G/A)AGACGGAG-3′
```

DNA constructs

pBINplusgus is the *pBINPLUS* (van Engelen *et al.*, 1995) binary vector with the *DR5::GUS* (Sabatini *et al.*, 1999) promoter-reporter inserted in the MCS.

pBINsil40 is the *pBINplusgus* binary vector in which an *ENOD40* inverted repeat is inserted that was generated using *pBS-d35S-RNAi* (Limpens *et al.*, 2003).

RNA isolation

RNA was isolated using the hot phenol method for RNA isolation (Pawlowski *et al.*, 1994). After RNA isolation the samples were treated with DNaseI (Promega) to remove chromosomal DNA.

RT-PCR

RT-PCR was performed (Albrechts *et al.*, 1998) using the following PCR program: 30 sec at 94º C, 30 sec at 55º C and 30 sec at 72º C and 20 cycles.

hairy roots

The roots of 5 day-old *M. truncatula* seedlings were removed at the hypocotyl and the wound surface inoculated with *Agrobacterium* MSU440 (Sonti *et al.*, 1995) containing the *pBINsil40* or the *pBINplusgus vector*. The seedlings were co-cultivated with *Agrobacterium* for 5 days at 21°C (16h light-8 h darkness) and subsequently transferred to Emergence medium (3 mM MES pH 5.8 containing 2,5 g/L KNO3, 0.4 g/L MgSO4.7H2O, 0.3 g/L NH4H2PO4, 0.2 g/L CaCL2.2H2O, 10 mg/L MnSO4.4H2O, 5 mg/L H3BO3, 1 mg/L ZnSO4.7H2O, 1 mg/L KI, 0.2 mg/L CuSO4.5H2O, 0.1 mg/L NaMoO4.2H2O, 0.1 mg/L CoCL2.6H2O, 15 mg/L FeSO4.7H2O, 20 mg/L Na2EDTA, 100 mg/L Myoinositol, 5 mg/L Nicotinic acid, 10 mg/L Pyriodoxine HCl, 10 mg/L Thiamine HCl, 2 mg/L Glycine, 1% sucrose, 0.9% Daichin agar containing 300 μg/ml Cefotaxime (Duchefa)).

Plants were grown for 6-18 days on Emergence medium. In this period new roots formed that were potentially co-transformed with the T-DNA of the binary vector. Subsequently the roots were analysed for the presence of red fluorescence using a fluoresence stereomacroscope and red-fluorescent roots were removed. Subsequently plants were transferred to gravel and inoculated with *S.meliloti 41* as described.

fluorescence microscopy

GFP and dsRED fluorescence was captured using a stereo fluorescence macroscope (Leica MZIII) equiped with dsRED and GFP specific filters (Chroma).

GUS staining

GUS staining was performed according to the protocol of Kosugi *et al.*, (1990). Plant material was stained in GUS buffer (0,1 M Na_2PO_4 pH 7.2, containing 1 mM X-Gluc dissolved in DMF and 0.1% Triton X-100). Plant material was put 10 minutes under vacuum and was incubated at 37°C for 4 hours or longer until the blue GUS stain was visible.

tissue sectioning

Fixation of root material was performed with 5% glutaraldehyde under vacuum for 2 hours. Subsequently this material was embedded in plastic using the Technovit Glycol Methacrylate Embedding Kit 7100. The embedded material was sectioned in 8 μm thick slices. Images of these sections were captured using a Nikon optiphot-2 microscope equiped with a Nikon coolpix digital camera.

nodule initiation

For quantification and classification of nodules and also for the isolation of nodule RNA, plants were grown in gravel supplemented with Fåhreus medium (Fåhreus, 1957) in a growth chamber at 25^{o} C and a 16 hour photoperiod (7W/m^{2}). One week after they had been transferred, the plants were inoculated with *S.meliloti 41-GFP*.




Chapter 4

Identification and characterisation of a *Zea mays* line carrying a transposon tagged *ENOD40*

Bert Compaan, Tom Ruttink, Cathy Albrecht, Robert Meeley, Ton Bisseling and Henk Franssen

Published in: Biochim. Biophys. Acta. **1629**(1-3):84-91,203

Abstract

In *Zea mays*, two *ENOD40* homologous were identified, that show only 30% of sequence homology to each other. We identified line *e40-mum1,* carrying a Mu transposon inserted in *ZmENOD40-1*, the maize gene that has the highest homology to leguminous *ENOD40.* The insertion causes a dramatic reduction of the *ZmENOD40-1* transcript level. Irrespective of this, homozygous *e40-mum1* plants are still able to interact with mycorrhizal fungi. Furthermore, no phenotypic aberrations correlated to the presence of *e40-mum1* have been identified and therefore it is concluded that *Zea mays ENOD40* genes are functionally redundant despite their strikingly low homology.

Introduction

In plants, the reverse genetics approach to resolve gene function is greatly facilitated by the availability of transposon or T-DNA tagged libraries. When studying a certain gene in a plant species for which such a library is not available, it is an accepted approach to analyse a homologue of this gene in a species for which these libraries have been developed. Information that is gathered in such a species can subsequently be used in the functional analysis of the gene in the species of interest.

Here we use the mutator transposon system in *Zea mays* to functionally characterize the gene *ENOD40*, that was first described in legumes (Yang *et al.*, 1993). Homologs have also been found in many other plant families (van de Sande *et al.*, 1996; Kouchi *et al.*, 1999), although strikingly not in *Brassaceae* of which family *Arabidopsis thaliana* is a member.

The overall nucleotide sequence homology between transcripts of *ENOD40* is low. Yet, *ENOD40* genes can be identified, because of the presence of two regions that are highly conserved between distantly related plant species. The most 5'located conserved region, designated box1, is defined by a start and stop codon yielding an open reading frame (ORF) that can be translated into a peptide of 10-13 amino acids. All box1-derived peptides share the characteristic [W-X4-HGS*] signature.

The second conserved region, designated box2, consists of 51 nucleotides (Kouchi *et al.*, 1999). However, the start and stop codons for translation that flank box2 do not give rise to ORFs that code for peptides conserved among *ENOD40* genes. Several approaches have been followed to test which of these AUG codons can be used to initiate translation (van de Sande *et al.*, 1996; Compaan *et al.*, 2001; Sousa *et al.*, 2001; Rohrig *et al.*, 2002). All studies show

that the AUG of box1 can be used as a start for translation. However, if in addition to box1, other ORFs are translated is still in debate (Rohrig *et al.*, 2002; Sousa *et al.*, 2001; Crespi *et al.*, 1994; Compaan *et al.*, 2001). Elucidation of the function of *ENOD40* would facilitate resolving if, or which ORF is actively required for the execution of this function.

The involvement of *ENOD40* in symbiotic processes was concluded from studies showing that in *Medicago truncatula* plants that overexpress *MtENOD40*, nodulation induced by *Sinorhizobium meliloti* is accelerated, whereas in plants that show co-suppression of *MtENOD40*, both nodule initiation and nodule development are disturbed (Charon *et al.*, 1999). Furthermore, in *Medicago truncatula* plants that over express *ENOD40* an enhancing effect on arbuscular mycorrhizae (AM) fungal growth in the root cortex and an increase in the frequency of arbuscule formation is observed (Rhijn *et al.*, 1997; Sinvany *et al.*, 2002). Co-suppression of *ENOD40* leads to a reduction in the colonisation of the root by AM fungi (Sinvany *et al.*, 2002; Staehelin *et al.*, 2001).

As *ENOD40* expression is also observed in the stem (Asad *et al.*, 1994; Varkonyi Gasic *et al.*, 2002), lateral root primordia (Asad *et al.*, 1994; Papadopoulou *et al.*, 1996) and in embryonic tissue (Flemetakis *et al.*, 2000), it is likely that *ENOD40* is also involved in non-symbiotic processes. Indeed, this was shown by studies on *Medicago sativa*, where both overexpression and anti-sense expression of *MtENOD40* disturbed the capacity of calli to form embryos (Crespi *et al.*, 1994).

Zea mays is a plant for which several transposon tagged libraries are available. One of them, the 'mutator' system has already been successful in generating mutants for a variety of genes, for example *a1* (O'Reilly *et al.*, 1985), *bz2* (McLaughlin *et al.*, 1987) and *vp1* (McCarty *et al.*, 1989). Benefits of the 'mutator' system are the high mutation rate and the high stability of the obtained insertions. Furthermore, the *Mu* elements are preferentially inserted into single copy DNA that, in combination with the high number of *Mu* elements in the genome, increases the chance of targeting a particular gene. In the case of *ENOD40*, this is of importance, for the gene is only small and its functionality is presumably restricted to two small conserved domains, so that the chance of putting this gene out of action will be much lower compared to the chance of hitting a normal sized gene. The capacity of *Zea mays* to interact with AM fungi enables us to study both the symbiotic and non-symbiotic effects of a mutation in *ENOD40*.

Here we describe the finding of two *ENOD40* genes in *Zea mays*. Furthermore a line is described that carries a transposon tag in one of the two *ENOD40* genes.

Results

Identification of two distinct *ENOD40* genes in Zea *mays*

ENOD40 genes can be recognised on basis of the presence of two conserved domains, denoted box1 and box2 that show more than 50% homology at the nucleotide level, even between distantly related plant species. Box1 putatively encodes a 10-13 amino acids long peptide that contains the characteristic [W-X_4-HGS*] signature. Within box2, a motif can be recognised to which every *ENOD40* gene has a homology of more than 90%, namely ACCGGCTAGTCA-$(X)_6$-GGCAAT. In this study, a gene was regarded as *ENOD40* when the gene contains these two hallmarks.

Using the stretch of 51 nucleotides from box2 of *OsENOD40* as query (Kouchi *et al*. 1999) in a blastn search, several ESTs were identified in the monocotyledon EST collection of Genbank that contain the box2 specific motif (Table 4.1) among them 7 ESTs of *Zea mays*. For the latter, comparison of the nucleotide sequences of the ESTs reveals that 6 out of 7 are contiguous, while one EST does not align with the contig and therefore is different from the other ESTs. These ESTs therefore represent two different genes, *ZmENOD40-1* (6 ESTs) and *ZmENOD40-2* (1 EST). *ZmENOD40-1* box1 putatively encodes for a peptide of 12 amino acids while *ZmENOD40-2* box1 encodes for a peptide of 13 amino acids (Table 4.2). Although the two *Zea mays ENOD40* peptides

Table 4.1 ESTs from monocotolydenous *ENOD40* genes that were identified in the genbank database using box2 as a query.

Species	Gene	EST accession
Oryza sativa	*ENOD40-1*	CA755970; BI809851; BI809872
	ENOD40-2	AU101849; AU065939
Triticum aestivum	*ENOD40-1*	CA698600; CA653680; CA652176; CA652176; CA646320; CA631431; CA611942; CA602529; BJ283670; BJ278615; BJ227547; CA609454; CA601715; CA653281; CA642008; CA642816; CA615844; CA647865
Hordeum vulgare	*ENOD40-1*	BQ765935; BI777713; BF259183
Zea mays	*ENOD40-1*	BM340203; BM340027; BM080692; BI478774; BI478512; AI001271
	ENOD40-2	AI491369
Sorghum bicolor	*ENOD40-1*	BE362733; BE362667

Table 4.2 Comparison of the peptide sequences that are encoded by box1 of *ENOD40* genes from monocots

Oryza brachyanta	MEDEWLEHAHGS
Oryza sativa-1	MEDEWLEHAHGS
Zea mays-1	MEDAWLEHLHGS
Triticum aestivum	MEGAWLEHLHGS
Hordeum vulgaris	MEGAWLEHLHGS

differ in length and amino acid sequence, they both contain the [W-X_4-HGS*] domain. From this we conclude that at least two *ENOD40* genes are present in the *Zea mays* genome.

Among the ESTs from *Oryza sativa* also two different types could be recognised so that we conclude that also *Oryza sativa* contains at least two *ENOD40* genes.

The nucleotide sequence of *ZmENOD40-1* and *ZmENOD40-2* show an overall homology of 30%, while *ZmENOD40-1* and *OsENOD40-1* share an overall homology of 52%. This observation suggests that *ZmENOD40-1* is functionally more related to *OsENOD40-1* than to *ZmENOD40-2* (Figure 4.1). This raises the question which of these two genes is functionally related to the legume *ENOD40* function required in symbiosis.

We suggest that the *ZmENOD40* gene that has the highest homology to legume *ENOD40* would most probably be the ortholog.

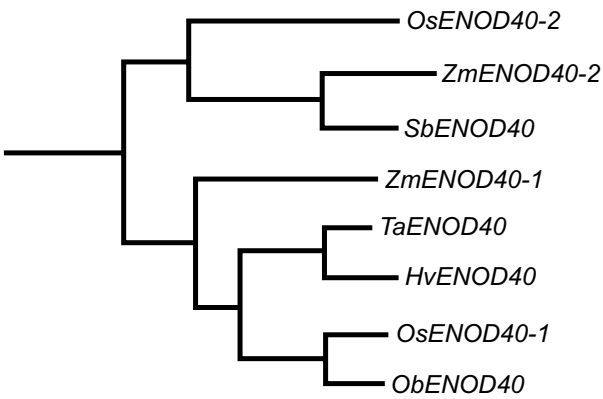


Figure 4.1 Phylogenetic tree that represents the homology between nucleotide sequences of *ENOD40* genes

Table 4.3 Comparison of the nucleotide sequences of box2

Os-1	GCACCAAACCGGCAAGTCACAAAACGGCAATGGTGAAACTCCAGCA-GGAGT
Zm-1	CGCACAAACCGGCAAGTCACAATACGGCAATGGCGG-ACTCCATTA-GGAGT
Mt	ATGGCAAACCGGCAAGTCACAAAAAGGCAATGG----ATTCCTTTTTGGAGT
Lj	CTGGCAAGCCGGCAAGTCACAAAAAGGCAATGG----ACTCCATTAGGGGTT
Nt	ATGGCGAACCGGCAAGTCACGAAACGGCAATGG----ACTCCGTTTTGGAGT
Os-2	TGAACAAACCGGCAAGTCATAAATCGGCAATAGTGA-GCTCCGACT-GGATC
Zm-2	GACAGAGACCGGCAAGTCAGAGATCGGCAATAGTGA-GCTCCAGCA-GGAGC
Mask 1	********** ****** * *** **
Mask 2	* * * * * * * *

The stars of Mask 1 indicate the nucleotides that are conserved in all the sequences. The stars of Mask 2 indicate the nucleotides in which *ZmENOD40-2* in contrast to *ZmENOD40-1* diverges from the leguminous *ENOD40* sequences.

Outside the conserved domains within *ENOD40*, the homology between distantly related plants is to low to make a proper alignment. Therefore, to identify the putative ortholog we used only box2 sequences of monocots, legume and *solanaceae ENOD40s* in a multiple sequence alignment (Table 4.3). From this, it appears that *ZmENOD40-1* shares 70% homology with the legume *Medicago truncatula*, while for *ZmENOD40-2* the homology is 60%. Thus *ZmENOD40-1* is more homologous to legume and *solanaceae ENOD40s* than *ZmENOD40-2*. In plants many multi-gene families occur. Although family members can be functionally redundant, in many cases they have a specific function by which a mutation in a single member leads to a specific phenotype. With this in mind we started the functional characterisation of *ZmENOD40-1* as the most probable ortholog of legume *ENOD40,* in *Zea mays*. To this end we searched for a *ZmENOD40-1* knockout mutant, to analyse whether it exhibits any specific phenotypes.

Characterisation of *ZmENOD40-1*

In order to determine the copy-number of *ZmENOD40-1*, a *Zea mays* EST clone, *MEST5-D3*, containing approximately 75% of *ZmENOD40-1* was used as a probe on a blot containing digested genomic DNA of *Zea mays*. Only one hybridising fragment was identified which in combination with the EST sequence analysis shows that there is only one copy of *ZmENOD40-1* in the genome of *Zea Mays* (Figure 4.2). Subsequently, the same blot was probed with a *ZmENOD40*-2 fragment that was generated by PCR on genomic DNA of *Zea Mays* using primers Zm40b-fw and Zm40b-rv. Also *ZmENOD40-2* hybridised to one DNA fragment and thus is most likely also a single-copy gene. Apparently

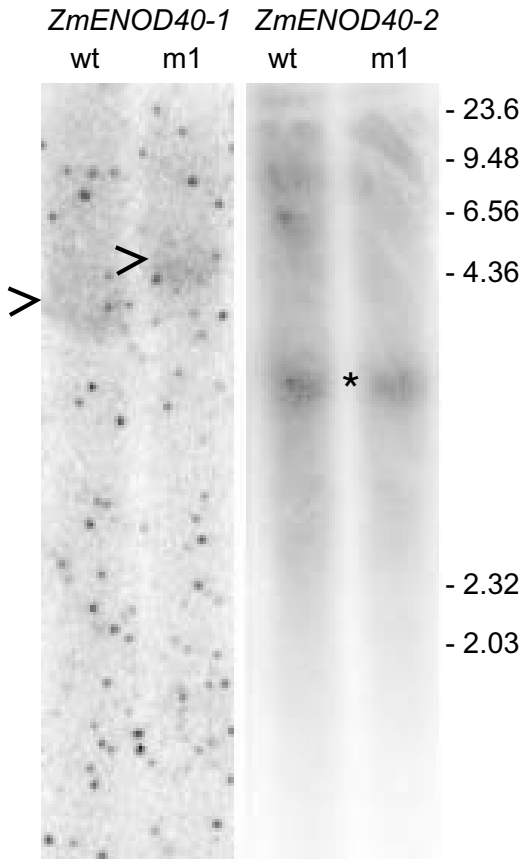


Figure 4.2 Blot of genomic DNA from wild type (wt) and homozygous *e40-mum1* plants (m1), that is probed with *ZmENOD40-1* and *ZmENOD40-2* respectively.

the two genes did not have enough homology to each other to cross-hybridise under these conditions.

To further analyse *ZmENOD40-1*, the *MEST5-D3* probe was used to screen a *Zea Mays* BAC library (ZMMBBb; J.P. Tomkins *et al.*, 2000). Filters containing 247,680 clones with a genome coverage of 13.6, were obtained from Clemson University Genomics Institute. Six BACs containing *ZmENOD40-1* sequences were identified and the presence of *ZmENOD40-1* was confirmed by hybridisation (data not shown). One of the six BACs, BAC ZMMBBb003F16, was used in further studies. BAC DNA was digested with *EcoRI* and *KpnI* and the fragments were cloned in pBluescript, digested with *EcoRI* and *KpnI*. Subsequently the obtained transformants were screened for the presence of *ZmENOD40-1* sequences. The insert of a sub-clone that hybridised to *MEST5-D3* was partially sequenced (Figure 4.3). Comparison of the nucleotide sequence from the obtained genomic sequence to the EST sequences revealed that *ZmENOD40-1* is lacking introns, as was also found for other *ENOD40* genes.

To determine the expression of *ZmENOD40-1*, plants were grown for two weeks and RNA, isolated from roots, stem, leaf and leaf veins, was separated on an agarose gel, blotted and hybridised with the *MEST5-D3* probe (Figure 4.4A). From this it can be seen that *ZmENOD40-1* is expressed in all tissues tested, with the highest expression in the stem. This expression pattern is reminiscent of expression patterns of *ENOD40* described in other plants (e.g.Varkonyi gasic *et al.*, 2002).

Screening the *Mu*-tagged population

In order to find a *ZmENOD40-1* mutant, a *Mu*-tagged population of *Zea mays* (cv. Cecilia) consisting of 40000 plants was screened for the presence of a *Mu*-transposon in *ZmENOD40-1* using the TUSC procedure (Meeley and Briggs, 1995; Chuck *et al.*, 1998). The PCR screening was performed with the Mutator TIR primer 9242 in combination with either of two other primers 35295F and 35296R that span the region of the gene encoding the transcript (Figure 4.3). Primers 35295F as well as 35296R were specific for *ZmENOD40-1* and do not have homology with *ZmENOD40-2*.

A fragment could be amplified that hybridised to *MEST5-D3* from 6 of the 40000 plants. The progeny of these plants was propagated; DNA was isolated from individual plants and tested for the presence of *Mu* insertion in *ZmENOD40-1* by PCR using primer 9242 and 35295F. Only one plant was found that delivered a 275 bp fragment in a PCR reaction using primers 35295F and 9242 (Figure 4.3). This fragment hybridised to *MEST5-D3* and was subsequently cloned in pGEM-T. Elucidation of the nucleotide sequence of this fragment revealed an insertion of *Mu 1.4* within box2 of *ZmENOD40-1*

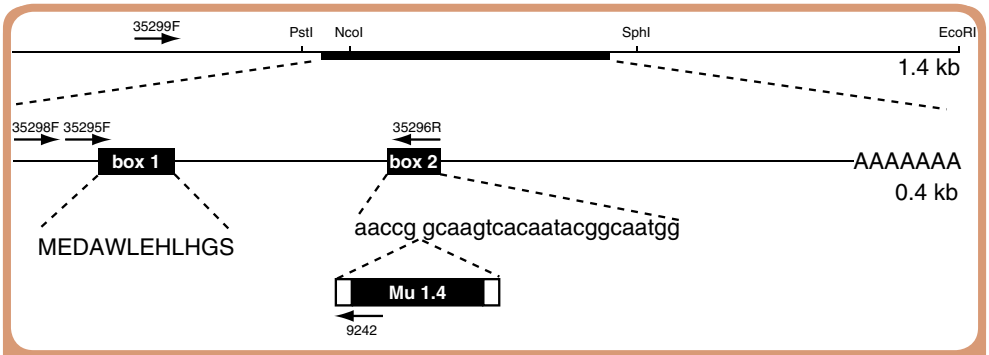


Figure 4.3. Schematic representation of the genomic region of *ZmENOD40-1* and the transcript that is encoded by this gene. Furthermore the exact location of the *Mu* transposon insertion is indicated. The arrows indicate the location and orientation of the primers that were used to detect the presence or absence of the *Mu* transposon within the gene.

(Figure 4.3). We named this allele *e40-mum1*.

To analyse *Zea mays* plants for the presence of the *e40-mum1* and *ENOD40-1* allele, PCR analyses were performed on genomic DNA from individual plants using primer 35299F in combination with either 9242 or 35296R. The combination 35299F and 35296R amplified in a PCR the *ENOD40-1* allele, while the combination 35299F and 9242 amplified the *e40-mum1* allele.

To determine whether the insertion did affect the expression of the mutated gene, homozygous *e40-mum1* plants were selved and the progeny was grown for two weeks. Cecilia (wild-type) plants were used as control. RNA was isolated from the stem of each individual plant and 20 microgram of RNA was separated on an agarose gel and subsequently blotted and hybridised with the *MEST5-D3* probe (Figure 4.4B). As compared to Cecilia plants, the expression of *ZmENOD40-1* was found to be dramatically reduced in homozygous *e40-*

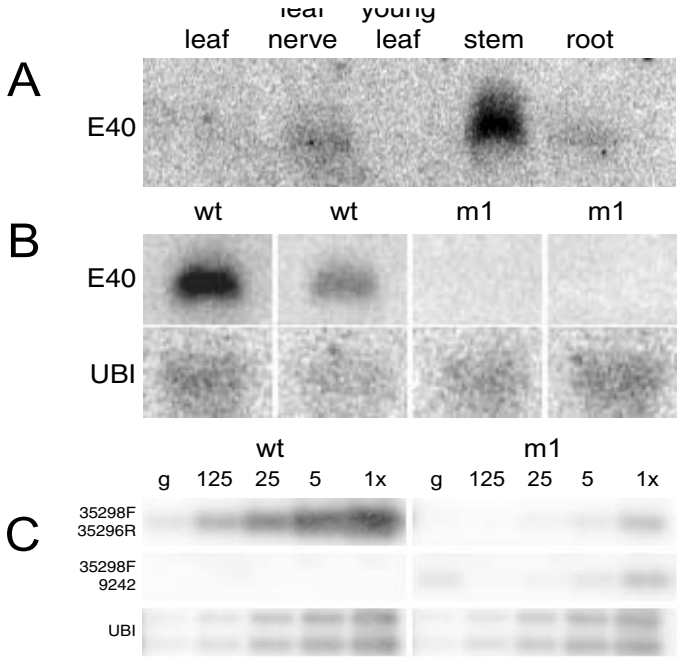


Figure 4.4 *ZmENOD40-1* expression in homozygous *e40-mum1* (m1) and wild-type (wt) plants. **A.** Northern analysis shows expression in the leaf vein, the root and a relative high expression in the stem. **B.** Comparison of two *e40-mum1* and two wild type plants using Northern analysis of RNA samples from stem, showed a severe reduction in expression of *ZmENOD40-1* in the *e40-mum1* plants. **C.** Quantification of the reduced expression using RT-PCR. Transcipts were detected without the *Mu* element (35298F- 35296R) and with the Mu element (35298F-9242). As standarisation an RT-PCR on *UBIQUITIN* (UBI) was used. The reaction was performed on a dilution range of the cDNA sample (125, 25, 5 and 1x dilution) and un-transcribed RNA samples (g).

mum1 plants, as no transcript in these plants could be detected in this way.

To determine the degree of reduction of *ZmENOD40-1* derived transcripts, quantitative RT-PCR was performed on stem RNA samples. In homozygous *e40-mum1* plants, two classes of transcripts could be expected: transcripts that are chimera of *ZmENOD40-1* and transposon-derived sequences and transcripts encoded by *ZmENOD40-1* from which the transposon had been excised. Somatic transposon excision is a common process in *Mu* tagged mutants, locally occurring in small sectors in plant tissues and which is correlated with the cessation of cell division (Raizada *et al.*, 2001). Therefore, one primer combination was used allowing the detection of *ZmENOD40-1* transcripts from which the *Mu* element was excised (primers 35298F and 35296R). A second primer combination was used allowing the detection of chimeric transcripts (primers 35298F and 9242). Among RNA isolated from homozygous *e40-mum1* plants both types of transcripts could be detected in an RT-PCR, albeit at strongly reduced levels when compared to the level of *ZmENOD40-1* transcripts in wild-type plants (Figure 4.4C). Quantification of the RT-PCR obtained signal indicates that the amount of *ZmENOD40-1* transcripts in homozygous *e40-mum1* plants is 200-fold lower as compared to the amount of *ZmENOD40-1* transcripts in wild-type plants. The use of two different primer combinations enabled the quantification of the two classes of transcripts separately. Among the detected transcripts, 99% contained *ZmENOD40-1* and transposon-derived sequences whereas 1% of the transcripts seems to be the result of a gene from which the transposon was excised.

Mycorrhizal fungi can interact with *e40-mum1* plants

To find out if the *e40-mum1* plants were affected in their ability to interact with mycorrhizal fungi, the allele was crossed into an *Accent* background, one that is commonly used in our greenhouse. For this, seven *ENOD40-1*/*e40-mum1* plants were pollinated with pollen from *Accent* wild-type plants. The F1 generation was genotyped by performing PCR on genomic DNA as described above. For further studies, 8 *ENOD40-1*/*e40-mum1* plants were selved, yielding the seeds of the F2 population.

Fifty kernels of this F2 population were sawn and the plants were grown in vermiculite containing a mixture of mycorrhizal fungi. After 2 weeks leaf punches from these plants were taken in order to genotype the plants by genomic PCR. Twenty seven percent of the population were *e40-mum1*/*e40-mum1*, 49% were *ENOD40-1*/*e40-mum1* and 24% were *ENOD40-1*/*ENOD40-1* plants, reflecting a normal Mendelian segregation. After 4 weeks roots were harvested and stained for mycorrhizal fungi infection. Microscopic analysis of

these stained roots revealed the presence of both arbuscules and vesicles in the roots of all plants that were analysed (Figure 4.5) and *e40-mum1* homozygotes could not be distinguished from wild-type. From this we conclude that homozygous *e40-mum1* plants are capable of interacting with mycorrhizal fungi.

No phenotypic aberrations could be found that are linked to the *e40-mum1* allele

To find phenotypes that were caused by the disturbed *ZmENOD40-1* expression in the mutant *Zea mays* plants, five homozygous *e40-mum1 Cecilia* plants were grown next to 5 wild-type *Cecilia* plants. All *e40-mum1* plants looked similar to each other but differ from the wild-type plants in at least three aspects.

First, the plants didn't have prop roots and the rest of the root system was strongly reduced in size. Second, the leaves did show a wilting effect, which could easily be observed from the strongly reduced leaf surface. Third, the development of the ear after self-pollination was affected; only a few kernels were formed.

To find out if one of these phenotypes was linked to the presence of the *e40-mum1* allele, the segregation of the *e40-mum1* allele and the phenotypes observed in the parental line were analysed. For this, the segregating population that resulted from the cross with *wt Accent* plants that is described above was used.

Three hundred and fifty plants were grown and genotyped for the presence of the *e40-mum1* allele using PCR as described. Twenty seven percent of the population were *e40-mum1*/*e40-mum1*, 49% were *ENOD40-1*/*e40-mum1* and 24% were *ENOD40-1*/*ENOD40-1* plants, reflecting a normal Mendelian segregation. Probably due to mutator activity about 10% of the plants did show growing disabilities as albinism and dwarfism. However none of these phenotypes was linked to the presence of the *e40-mum1* allele and therefore these plants were discarded. After two weeks, 250 plants were transferred to the greenhouse and these plants were analysed for the occurrence of three phenotypes, observed in the homozygous *e40-mum1* plants: disturbed prop root formation, wilting and disturbed ear formation. To be able to quantify these phenotypes, several measurements were performed on various time points.

Prop root formation was quantified by counting the number of prop roots after 10 weeks. The average number of prop roots per plant in the population was 11. Within the population, 22% of the plants did show a prop root phenotype in having less than 3 prop roots. However this phenotype did not co-segregate with the presence *e40-mum1* allele, for the average number of prop roots did

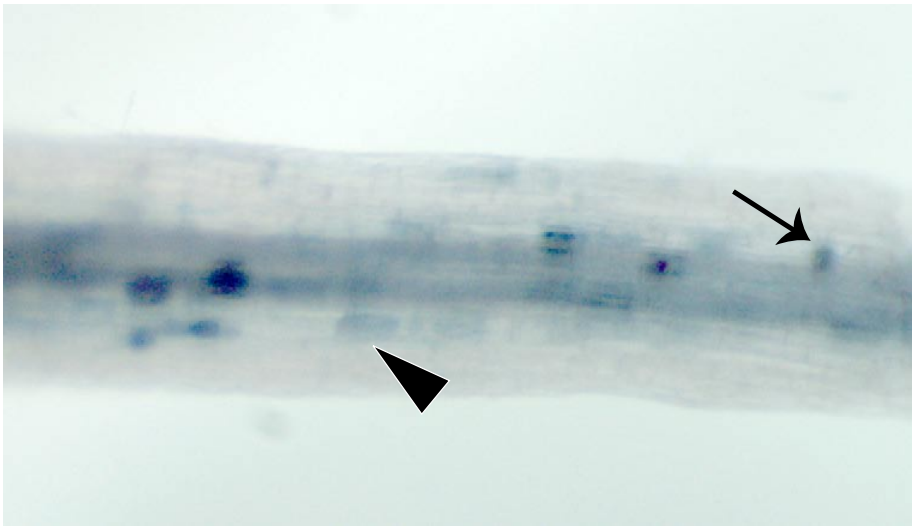


Figure 4.5 Small fragment of a homozygous *e40-mum1* root that is colonised by mycorrhizal fungi. Symbiotic structures are indicated, the arrow points at a vesicle and the arrowhead at an arbuscule.

not significantly differ between *e40-mum1* and wild-type plants.

To measure wilting of the leaves, the surface and the dry-weight of each 5th and 6th leave was measured. The average leaf surface being 480 cm^2 and dry weight (4.0 g) did not significantly differ between *e40-mum1* and wild-type plants. Within the population, 11% of the plants had a leaf surface that was smaller than 300 cm^2, however there was no linkage of this phenotype to the *e40-mum1* allele.

To analyse the ear phenotype, all plants were emasculated and pollinated with pollen from *Accent wt* plants. For 56% of the 250 plants the flowering time was delayed as compared to the *Accent wt* plants and therefore it was only possible to pollinate 44% of the plants. This late flowering was not linked to the *e40-mum1* allele.

After ripening the ears of the plants were harvested. Between individual plants clear differences could be observed regarding the shape of the ear. However, no strict correlation could be found between malformed ears and the presence of the *e40-mum1* allele.

Discussion

Here we describe the identification of two *ENOD40* genes of *Zea mays*.

Comparative analysis of these two genes and the *OsENOD40* gene learns that the homology between *ZmENOD40-1* is higher to *OsENOD40-1* than to *ZmENOD40-2*. This indicates that the duplication of the ancient *ENOD40* gene into *ZmENOD40-1* and *ZmENOD40-2* occurred before the division of *Zea mays* and *Oryza sativa*. Furthermore, it suggests that *ZmENOD40-1* and *OsENOD40-1* are in a subgroup, different from *ZmENOD40-2*. The ancient origin of the subgroups, preceding the *Zea mays – Oryza sativa* speciation, in combination with the high homology between members belonging to one group strongly suggest the existence of selective pressure on these genes and therefore indicates a functional separation. Comparative analysis of the two maize *ENOD40* genes to leguminous *ENOD40*, revealed that *ZmENOD40-1* is most homologous to legume *ENOD40*. For this reason *ZmENOD40-1* was chosen to be analysed in this study as the most probable ortholog of leguminous *ENOD40*.

We have analysed one *Zea mays* line in which a *Mu* element was inserted in the middle of the gene. This insertion resulted in a dramatic reduction of *ZmENOD40-1* transcript level**.** Although in the initial Cecilia *e40-mum1* line several growth and developmental abnormalities were observed, we could not determine a strict coupling between the presence of any of these abnormalities and the presence of the *e40-mum1* allele. For instance, irrespective of the reduced levels of *ZmENOD40-1* transcript plants are still capable to interact with mycorrhizal fungi, for both arbuscules and vesicles are found in *e40-mum1* and wild-type plants.

One plausible explanation for the failure to observe phenotypic defects in the mutant plants could be that *ZmENOD40-1* function is not completely abolished in these plants. *ZmENOD40-1* derived transcripts could still be detected in *e40-mum1* plants, although with a reduction in expression of at least 200-fold. However, from these transcripts the peptide can be translated and the amount made could be sufficient for proper *ENOD40* action. Another explanation for the faillure to detect a phenotype in *e40-mum1* plants could be that this phenotype is conditional or subtle. A third explanation can be that the function of *ZmENOD40-1* and *ZmENOD40-2* are redundant. This would imply that the *ZmENOD40-2* gene that is only 30% homologous to *ZmENOD40-1* substitutes for the function of *ZmENOD40-1*. Recent experiments in yeast show that gene redundancy even might exist for ancient duplicated genes (Gu *et al.*, 2003). This could imply that even ancient duplication events within the class of monocots, that in the case of *ENOD40* putatively predate the speciation of *Oryza* and *Zea,* still can generate functional redundancy in these species. The two regions in *ENOD40* that are conserved for more than 50% could be sufficient for this redundancy. Notably, a similar low extent of homology exists among members of the *ENOD40* gene family in *Lotus japonicus* (Flemetakis *et al.*,

2000) or *Trifolium repens* (Varkonyi-Gasic *et al.*, 2002). Therefore it can be anticipated that studies on knock-outs in these plant species will suffer from gene redundancy as well. Ideally, therefore, studies on the effect of a knock-out of *ENOD40* should be focused on plants carrying just one *ENOD40* gene. However, one might wonder whether the identification of one *ENOD40* gene in *Medicago truncatula* and pea so far is representing the actual number of genes in the genome of these species.

For *Zea mays* it would be worthwhile to repeat the procedure that was so succesful in the finding of a mutant for *ZmENOD40-1*, for *ZmENOD40-2*. The description of such a mutant and the double mutant would be informative on the functional overlap of these genes and lead to the identification of the function of *ZmENOD40*.

Materials and Methods

Primers

35295F	5′-TTGTCCACAGCCCCCGTAGACCA-3′
35298F	5′-TCCTCCTTCTCCCCTCCCC-3′
35299F	5′-TGGCCTGGTTCCTCGTGTTG-3′
35296R	5′-TCCGCCATTGCCGTATTGTGACTTGC-3′
9294	5′-AGAGAAGCCAACGCCAWCGCCTCYATTTCGTC-3′
Zm40-2-fw	5′-GCTGCAGGCAAGTAAAGGTA-3′
Zm40-2-rv	5′-GGCAGGAGGGATCAGACAAG-3′

RNA isolation

RNA was isolated from different organs of *Zea mays* using the hot phenol method for RNA isolation (Pawlowski *et al.*, 1994). After RNA isolation the samples were treated with DNaseI (Promega) to remove chromosomal DNA (Albrechts *et al.*, 1998).

RT-PCR

RT-PCR was performed (Albrechts *et al.*, 1998) using the following PCR program: 30 sec at 94º C, 30 sec at 58º C and 30 sec at 72º C.

Identification of mutants

A segregating population of *Zea mays* plants from a cross of inbred line

Cecilia x *Mu,* containing 40,000 plants was screened using the TUSC procedure (Meeley *et al.*, 1995). Pool screening was initiated with gene primers 35295F and 35296R, each in combination with the Mutator TIR primer 9242. F2 transmission testing was performed on 24 individuals, for one F2 individual a stable Mu insertion in *ZmENOD40-1* could be identified.

Plant material

Two wild-type inbred lines were used in these experiments: *Accent* and *Cecilia*.

Kernels were sawn in pots with a diameter of 15 cm, leaf punches were isolated after 2 weeks and after 3 weeks the plants were transferred to greenhouse soil and grown at a temperature of 25° C and a 16 hour photoperiod.

Northern blotting

10 μg RNA was denatured in DMSO and glyoxal and separated on an agarose gel. Subsequently the denatured RNA was transferred to Hybond-N^+ membranes (Amersham) (Sambrook *et al.*, 2001). Hybridisation of the membranes was performed using ^{32}P labeled DNA probes in a roller bottle, (Sambrook *et al.*, 2001).

Southern blotting

Genomic DNA was isolated (Kozik *et al.*, 1996) and 10 µg of genomic DNA was digested with *EcoRI* and *HindIII* and separated on a 0.8% agarose gel. Capillary transfer of digested genomic DNA to Hybond-N^+ membranes (Amersham) was performed as described (Sambrook *et al.*, 2001). Hybridisation of these membranes was performed using ^{32}P labeled DNA probes (Sambrook *et al.*, 2001).

Analysing hybridised membranes

Membranes were exposed to Storage Phospor Sreens (Molecular Dynamics) for one night and subsequently these screens were analysed using a STORM 840 Phosphor Imager (Molecular Dynamics). Results were analysed with ImageQuant for Macintosh Version 1.2 (Molecular Dynamics).

Inoculation with mycorrhizal fungi

Zea mays plants were inoculated with a mixture of mycorrhizal fungi (Albrechts *et al.*, 1998). The mixture contained the following fungi: *Glomus intraradices*, *Glomus claroideum*, *Glomus geosporum*, *Glomus mosseae*, *Glomus etunicatum*, *Glomus clarum*, *Scutellospora heterogama, Scutellospora pellucida, Gigaspora rosea*, *Gigaspora candida*, *Gigaspora margarita*, *Acaulospora laevis*.

To check mycorrhizal arbuscule and vesicles formation, root segments were cleared stained and assessed microscopically (Albrechts *et al.*, 1998).

Software

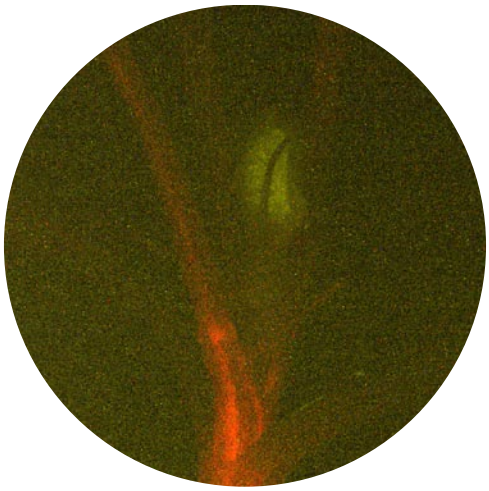
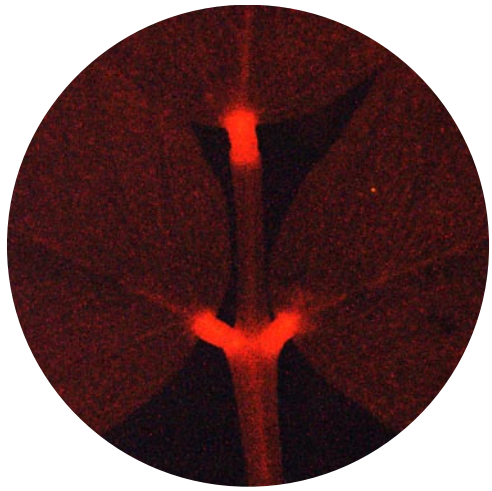
DNA sequences were aligned using ClustalX (Thompson *et al.*, 1997) and these alignments were edited using BioEdit (Hall *et al.*, 1999). A phylogenetic tree was drawn using TreeView (Page *et al.*, 2001).

PCR on Genomic DNA

Leaf punches (1 cm^2) were grinded by shaking in an eppendorf tube with glass beads. Then, 500 µl of Extraction Buffer was added (0.3 M NaCl, 50 mM Tris-HCl pH 7.5, 20 mM EDTA, 2% sarkosyl, 0.5% SDS, 5 M Urea and 5% Phenol). After vigorous shaking, 400 µl Phenol/Chloroform (1:1) was added and the samples were centrifuged for 15 minutes at 14000 rpm. The supernatant was transferred to a clean tube and genomic DNA was precipitated by adding 0.8x volume isopropanol and centrifugation for 10 minutes at 14000 rpm. After washing with 70% ethanol, the genomic DNA was dissolved in 100 µl TE. One µl of this solution was used in a PCR reaction (10 mM Tris-HCl pH 8.3, 50 mM KCl, 2.5 mM $MgCl_2$, 100 µM dNTPs, 100 ng of each primer and 1 U Taq polymerase (Boehringer Mannheim, USA) in a total volume of 50 µl).

For these reactions the following cycling program was used: 25 cycles of: denaturation 94º C 30 sec; annealing 55º C 30 sec; extension 72º C 30 sec. Amplified DNA samples were separated on an 1% agarose gel. The DNA was blotted to Nytran-N^+ membrane sodium hydroxide/chloride and hybridised to radiolabeled inserts of the corresponding cDNA clones.



Chapter 5

Expression analysis of *ENOD40* in relation to its conserved downstream neighbour

Bert Compaan, Tom Ruttink, Erik Limpens, Ton Bisseling and Henk Franssen

Abstract

Elucidation of the nucleotide sequence of the genomic context of *ENOD40* in *Medicago truncatula*, *Pisum sativum*, *Lycopersicon esculentum* and *Oryza sativa*, showed that the gene located directly downstream of *ENOD40* (*DN40*) turned out to be conserved in all these species. This linkage exceeds the monocot-dicot gap and suggests that there is selective pressure on the conservation of the linkage, which could be reflected in co-expression of both genes as has been observed in other organisms. To investigate, we analysed the expression of both genes in *Medicago truncatula* using the fluorescent reporter couple, *ds*RED-eYFP or GUS. We found that *ENOD40::dsRED* shows distinct expression during leaf, root and root nodule development from *DN40::eYFP* or *DN40*::*GUS*.

Introduction

Comparison of the nucleotide sequences of complete genomes, that are available nowadays, has revealed that the gene order on the chromosomes of related species can be similar. This so-called synteny has been observed in both prokaryotes and eukaryotes. Among the genomes of eukaryotes synteny is often disturbed by gene insertions and inversions (Devos *et al.*,1999; van Dodeweerd *et al.*, 1999; Liu *et al.*, 2001; Mayer *et al.*, 2001; Vandepoele *et al.*, 2002). Therefore, strict colinearity between divergent species is mainly confined to gene couples as has been shown for instance for the genomes of bacteria, archea (Dandekar *et al.*, 1998).

Microcolinearity in a specific genomic region of highly divergent species, suggests that a functional selective pressure acts on the gene order and content of that region, preventing the distortion of gene order in the course of speciation. Studies in yeast and *Drosophila* suggest that gene couples are frequently co-expressed. However, the adjacent genes are not necessarily functionally related (Cohen *et al.*, 2000; Cho *et al.*, 1998; Spellman *et al.*, 2002; Ueda *et al.*, 2002). In *Arabidopsis thaliana* for a set of co-regulated genes chromosomal clustering has been uncovered and co-regulated genes are most likely functionally related as they are e.g. co-expressed in the same domain of the root or in response to auxin (Birnbaum *et al.*, 2003).

Synteny among plant species belonging to the same family can be very high as has been shown for the *sh2*/*a1* and *adh1*/*adh2* regions of monocots (Chen *et al.*, 1997; Tarchini *et al.*, 2000), but only a few syntenic regions can be identified between evolutionary more distant species e.g. monocots and dicots.

Consequently, conservation of gene order (or microcolinearity), in the plant kingdom in divergent clades is exceptional.

ENOD40 is highly induced in legumes at the onset of nodule formation in the root pericycle at the site of nodulation only three hours after the application of nodule initiating rhizobia (Compaan *et al.*, this thesis). Genes that are expressed at this location and time are expected to be involved in the dedifferentiation of cortical cells and the initiation of a nodule primordium. Unlike its apparent involvement in the nodulation process, *ENOD40* also has been found in non-legumes, like tomato and the monocots rice (Vleghels *et al.*, 2003; Kouchi *et al.*, 1999) and maize (Compaan *et al.*, this thesis), although strikingly not in the model plant *Arabidopsis thaliana*. As one of the approaches to understand the mode of action of *ENOD40*, we searched for microcolinearity around the *ENOD40* locus by comparing sequences of genomes of divergent plant species that contain *ENOD40*. Here we present evidence of microcolinearity around the *ENOD40* locus. To compare the expression pattern of *ENOD40* and its neighbour gene, we used *ds*RED and *e*YFP to localise promoter activity of the 2 genes in one plant. Our studies show that the two adjacent genes are not co-expressed.

Results

Genomic sequence analysis of *ENOD40* region

To obtain genomic sequences in the vicinity of *ENOD40* from different plant species, several approaches were followed.

For *Medicago truncatula*, a BAC library containing 30.10^3 clones (Nam Y-W *et al.*, 1999) was screened with a cDNA fragment of *MtENOD40*. In this way 2 BACs (MT_ABa39I20 and MT_Aba23A06) were identified that hybridised to *MtENOD40*. A physical map was generated revealing that the two BACs form a contig of about 30 KB. The BACs were sub-cloned and the subclones were sequenced. In this manner, approximately 80% of the contig was sequenced.

For *Pisum sativum*, a genomic library in λDASH was screened with a *PsENOD40* cDNA fragment. Out of 10^6 plaques, one positive plaque was identified from which phage DNA was isolated. The 9 KB insert was sub-cloned and completely sequenced.

For *Lycopersicon esculentum*, a genomic library in λEMBL3 was screened with a *LeENOD40* cDNA fragment. Out of 10^6 plaques, 51 positive plaques were identified. Two plaques were chosen from which DNA was isolated. A physical map of the inserts was generated revealing that the two phages

form a contig of about 18 KB. The inserts were sub-cloned and the sub-clones were sequenced. The obtained sequence information was assembled using the physical map. From about 80% of the genomic region surrounding *LeENOD40* the nucleotide sequence was determined.

For *Oryza sativa*, the genomic sequence of *OsENOD40* (AL662969) was identified in the genbank database by a *blastn* search using *OsENOD40* as query.

To identify the presence of (putative) genes within the regions for which the nucleotide order had been determined, the genomic DNA sequence of the region surrounding *ENOD40*, was used as a query in a *blastx* search against the genbank database (Table 5.1). In all four genomic regions, directly downstream of *ENOD40* a region was identified with high homology to the *Arabidopsis thaliana* hypothetical protein At1g60790 (Figure 5.1). In the genomic region of *Medicago truncatula* downstream of the At1g60790-like gene a region was identified with high homology to At1g60800, a predicted protein from *Arabidopsis thaliana* that shows a high degree of homology to *the Dc*SERK protein. Notably, in *Arabidopsis thaliana* the genes At1g60790

Table 5.1 Proteins identified in a blastx search using the 4 genomic sequences surrounding *ENOD40* as query.

Species	homology	protein	expect value
Medicago		*MtENOD40*	0
truncatula	At1g60790	hypothetical protein	$1e^{-49}$
	At1g60800	receptor like kinase	$1e^{-63}$
Pisum		*PsENOD40*	0
sativum	At1g60790	hypothetical protein	$1e^{-49}$
Lycopersicon		*LeENOD40*	0
esculentum	At1g60790	hypothetical protein	$3e^{-71}$
	At1g60800	kinase domain fragment	$1e^{-58}$
	AF220603_4	Prf	0
Oryza	At3g04290	lipase	$1e^{-54}$
sativa	AAH03123	unknown protein	$1e^{-26}$
		OsENOD40	0
	At1g60790	hypothetical protein	$3e^{-30}$
	AAB39248.1	NADP-isocitrate dehydrogenase	$4e^{-42}$
	AAK08983	Myb transcription factor	$3e^{-19}$

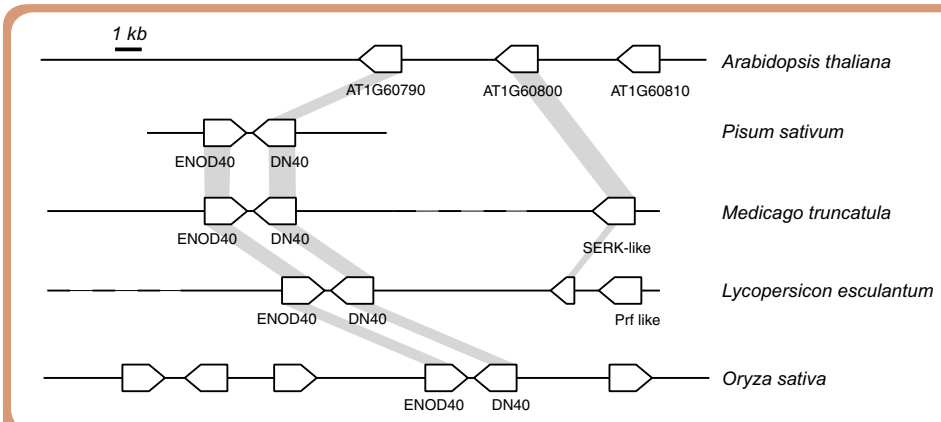


Figure 5.1 Schematic overview of the four genomic sequences surrounding *ENOD40* in comparison to a syntenic region on *Arabidopsis thaliana* chromosome 1.

and At1g60800 are adjacent, and this might also be the case in the *Medicago truncatula* genome. However, this region of the BAC has not completely been sequenced and therefore, we can not exclude that there might be an additional gene located in between these two genes. Thorough analysis using the *blastn* algorithm revealed that no *ENOD40* homologous sequences could be identified in this region of the *Arabidopsis thaliana* genome.

When using the genomic sequence of *Lycopersicon esculentum* as a query, in addition to At1g60790, further downstream a fragment was identified that has the highest homology to the kinase domain of At1g60800, but a region homologous to the extra-cellular part of At1g60800 is lacking in the tomato DNA. This is in contrast to *Medicago truncatula* sequences, where homology over the full length of this protein was identified. Directly adjacent to this region in *Lycopersicon esculentum*, a region was identified with high homology to the disease resistant protein PRF, which could not be identified in the genomic sequences of the other species.

In the region of 50 kb surrounding *OsENOD40*, apart from a region that shows homology to At1g60790, 3 regions were identified that show homology to known proteins (table 1), but none of these had homology to the kinase shown to be present in tomato and *Medicago* genomic DNA.

Thus, a region with homology to At1g60790 is lying directly downstream of *ENOD40* in all the sequences that were tested. We propose that this region represents a gene that we denote *DN40* for downstream neighbour of *ENOD40*. The coupling of these genes in both monocots and dicots suggests that there is a functional relation between these genes.

Description of *DN40*

To determine the structure of *DN40* in the analysed plant species, we conducted blastn searches against the NCBI database to identify ESTs of *DN40*. In this way, several ESTs were identified that are identical to genomic *DN40* sequences from *M. truncatula* and *L. esculentum*, which indicates that *DN40* is transcribed in these species. For *O. sativa* and *P. sativum* no ESTs could be identified.

Furthermore, to obtain the complete cDNA sequence of *DN40*, RACE was performed on RNA samples from *M. truncatula* leaf. For the 3'-RACE, cDNA was prepared using the RACE-T primer and two subsequent PCRs were performed using primers mtf1 and mtf2 in combination with the RACE-A primer. For the 5'-RACE, A-tailing was performed on the cDNA that was generated using the emtr primer. Two subsequent PCRs were performed on this tailed cDNA, using primers mtpr1 and mtpr2 in combination with the RACE-A primer.

The RACE products were cloned in pGEM-T and the nucleotide sequence of the inserts was determined. Combining all sequence data from *MtDN40* showed that the transcript was about 1.9 kb long, including the 5' UTR. Alignment of the cDNA sequences to the genomic sequence showed that *MtDN40* consisted of 5 exons and 4 introns (Figure 5.2).

Next, the nucleotide sequence of the *MtDN40* transcript, *DN40* derived sequences collected from the genbank EST database and splice prediction programs (NetGene2 and SplicePredictor) were used to predict the location of the *DN40* transcript within the genomic sequences available from pea, tomato and rice. Comparison of the predicted *DN40* transcripts and the genomic DNA shows that all genes consist of 5 exons. Only in pea, the first exon is interrupted

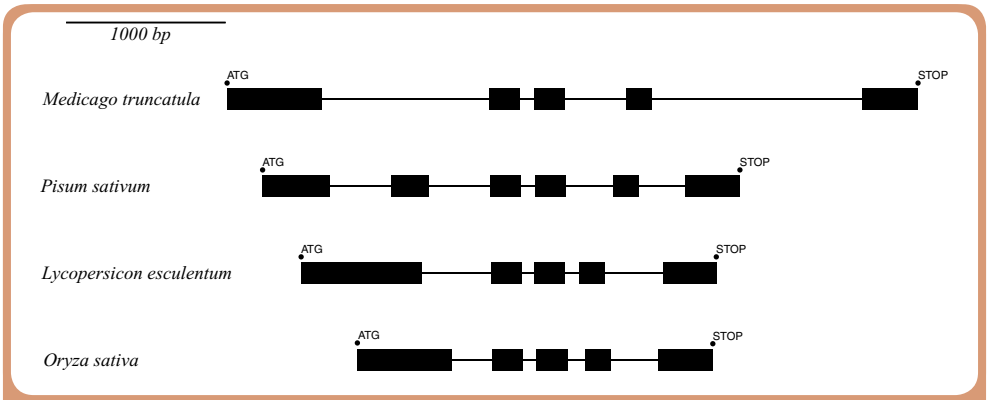


Figure 5.2 Schematic overview of the four identified *DN40* genes, showing that the intron-exon boundaries are in general conserved.

by an extra intron. Furthermore, we conclude that the nucleotide sequence of the predicted exons 2 to 5 is conserved as are the positions of the intron-exon boundaries. However, the nucleotide sequence of the first exon is hardly conserved among *Leguminoseae*, *Solanacea* and *monocots*. Translation of the longest ORF that can be identified in the predicted *DN40* transcripts results in a putative protein of about 65 kDa. Alignment of the proteins derived from the various plant species shows that there is a high degree of homology in the part of the proteins that is encoded by exons 2-5. However, the homology among the amino-terminal part of proteins, encoded by exon 1 is low. This could also indicate that *in vivo* translation does not start at the AUG present within exon 1, but instead at the AUG in exon 2, which then gives rise to proteins that show a high overall homology. As a consequence, *DN40* transcripts would have a large 5'-UTR. Notably, in all species studied *DN40* and *ENOD40* are transcribed in opposite direction.

To get a clue of the function of this gene, we searched for homologous proteins with a known function in the genbank database. Several *DN40* homologues were found and all of them were of plant origin and for none a function was known. The *Arabidopsis thaliana* genome for example contains at least 40 proteins that share homology to *DN40*. Thus our database search revealed that *DN40* is a member of a large, plant-specific, gene family for which no function has yet been determined. We used PSORT, and motif search programs to identify known domains in the predicted DN40 protein but these programs delivered conflicting results.

Expression patterns of *ENOD40* and *DN40*

If the observed microcolinearity implies functional relationship between *ENOD40* and *DN40*, it is likely that both genes are expressed simultaneously in the same cells. To analyse whether *ENOD40* and *DN40* show syn-expression, the expression of these genes was analysed in *M. truncatula*. To this end RT-PCR was performed with *DN40* primers mtr1 and mtf1 and with *ENOD40* specific primers e40f and e40r. RNA was isolated from roots, nodules, leafs and flowers of *M. truncatula* plants, cDNA was prepared and subsequently PCR was performed on a dilution range of these cDNAs. The products were separated on an agarose gel, blotted and hybridised with a DNA fragment of *MtDN40* or *MtENOD40*, respectively (Figure 5.3). For *DN40*, the signal intensity is similar in all the RNA samples indicating that *MtDN40* is expressed at an equal level in these four organs. Also *ENOD40* could be detected in all samples, although the expression in the nodule is notably higher than anywhere else.

To analyse whether *ENOD40* and *DN40* also show a similar expression pattern

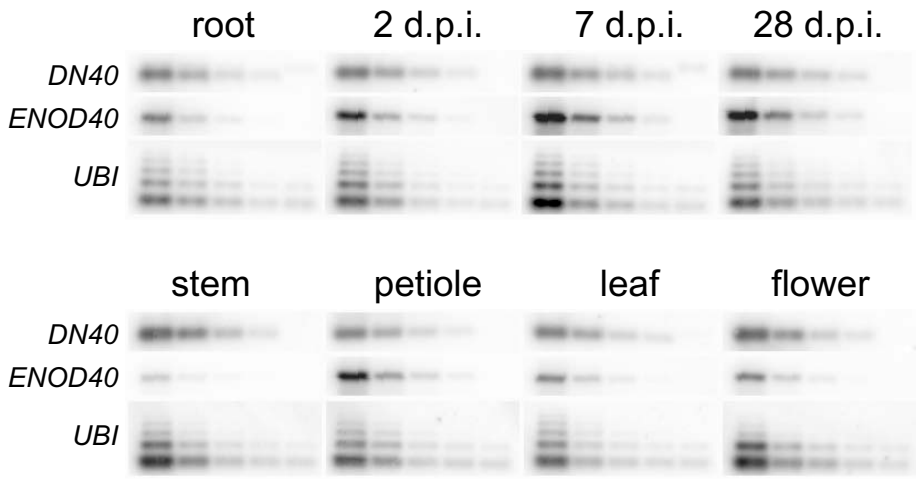


Figure 5.3 RT-PCR on RNA isolated from various *Medicago truncatula* tissues, showing that *ENOD40* and *DN40* are both ubiquitously expressed. The first panel shows the expression of both genes in roots that were inoculated with *Sinorhizobium meliloti* for 2, 7 or 28 days.

at the cellular level, a T-DNA construct was prepared that enabled the visualisation of the expression of both genes in the same plant (Figure 5.4). Therefore, the complete genomic region of *MtENOD40*, containing the promoters of *ENOD40* and *DN40* as well as the 5'- and 3'-UTR's of the two genes and the intergenic region was cloned in pCambia1300. To be able to visualise the expression of the genes, the ENOD40 peptide-coding ORF was replaced by *dsRED* and the protein coding part of *DN40*, including introns, was replaced by *eYFP*. Both eYFP and *ds*RED have long wavelength excitation spectra and therefore the detection of these reporters is not disturbed by auto-fluorescence, which is a common problem in plant roots. This construct was stably integrated into the genome of *Medicago truncatula cv.* R108. Experiments were performed with plants from the lines *cam40bc-1* and *cam40bc-6* from the T3 generation.

Plants of the lines *cam40bc-1* and *cam40bc-6* were grown on agar plates, enabling the visualisation of fluorescence as a result of the expression of the genes encoding *ds*RED and eYFP using the stereo macroscope. Both lines show the same expression patterns and therefore only the studies with line

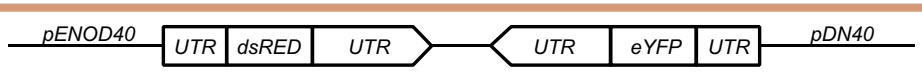


Figure 5.4 Schematic representation of the reporter construct that was used to visualise the expression of *ENOD40* and *DN40* simultaneously.

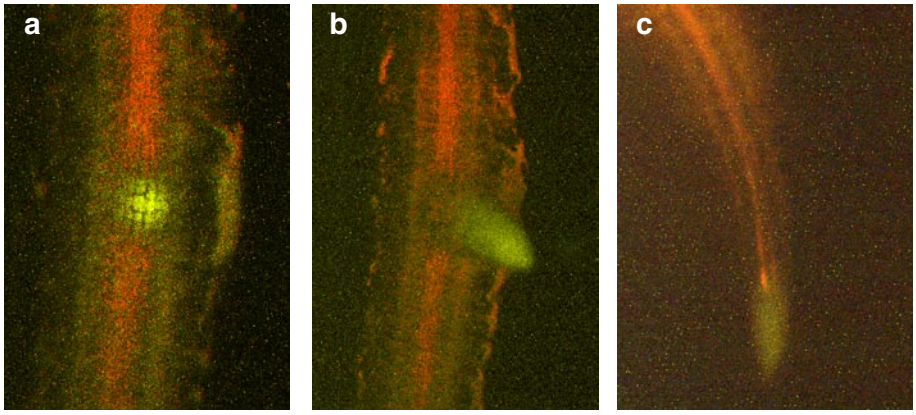


Figure 5.5 Expression of *ds*RED and eYFP in the root of a *cam40bc-1* plant. Images were captured from a lateral root primordium just before emergence (day 1, panel a), after emergence (day 2, panel b) and from the growing lateral root (day 5, panel c). *ds*RED is expressed in the vascular tissue of the root. The localisation of the lateral root primordium is marked by a local reduction in *ds*RED expression and an induction of eYFP expression.

cam40bc-1 are described.

In roots, *ds*RED fluorescence (*ENOD40*) is observed in the vascular tissue (Figure 5.5). To determine the location of this fluorescence more precisely, roots were 'embedded' in agarose and sectioned in 100 μm thick slices. These slices were analysed by confocal microscopy, showing that *ds*RED occurs in the pericycle and in the root vascular cambium (Figure 5.6). This location of *ENOD40*::*ds*RED expression is in agreement with data of other *ENOD40*

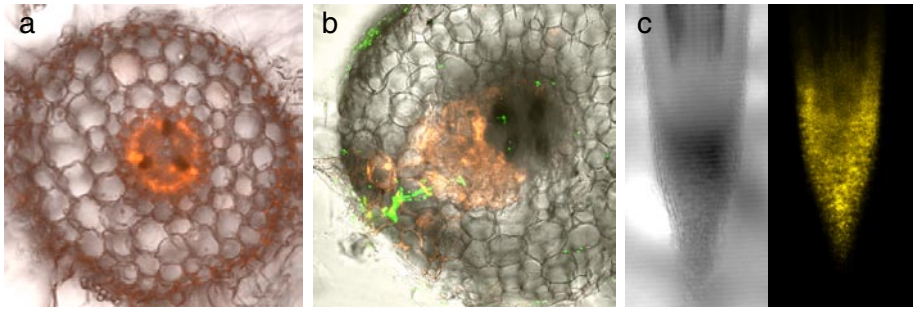


Figure 5.6 Confocal images from sectioned roots of *cam40bc-1* plants. *ds*RED expression is confined to the pericycle and vascular cambium (panel a). *ds*RED expression is induced in the inner cortex upon inoculation with *Sinorhizobium meliloti*. These bacteria harbour a GFP expression vector, so that infection threads could be visualised (panel b). eYFP expression could be visualised in the root meristem (panel c).

expression studies in legumes (Asad *et al.*, 1994; Corich *et al.*, 1998) and non-legumes (Vleghels *et al.*, 2003).

In roots of *cam40bc-1*, eYFP (*DN40*) fluorescence is not present in the vascular tissue but in cells of the root tip (Figure 5.5). However, the cell type in which it is present could not be determined despite of the use of confocal microscopy (Figure 5.6).

In the aerial parts of the plant, *ds*RED fluorescence occurs in the vascular bundles of stem and leaf. Furthermore a very high expression was observed specifically in the pulvinus of the leaf (Figure 5.7).

In contrast, *DN40::eYFP* fluorescence was not present in vascular tissue of leaves or stems. Instead, *DN40::eYFP* fluorescence was only detectable in epidermal or mesophyl cells of the developing leaf while the leaf is still folded and the fluorescence was absent in those cells in the mature leaf. So neither in root nor in aerial tissue *ENOD40::dsRED* and *DN40::eYFP* expression co-localise.

In legumes, *ENOD40* expression is induced in roots after interaction of plants

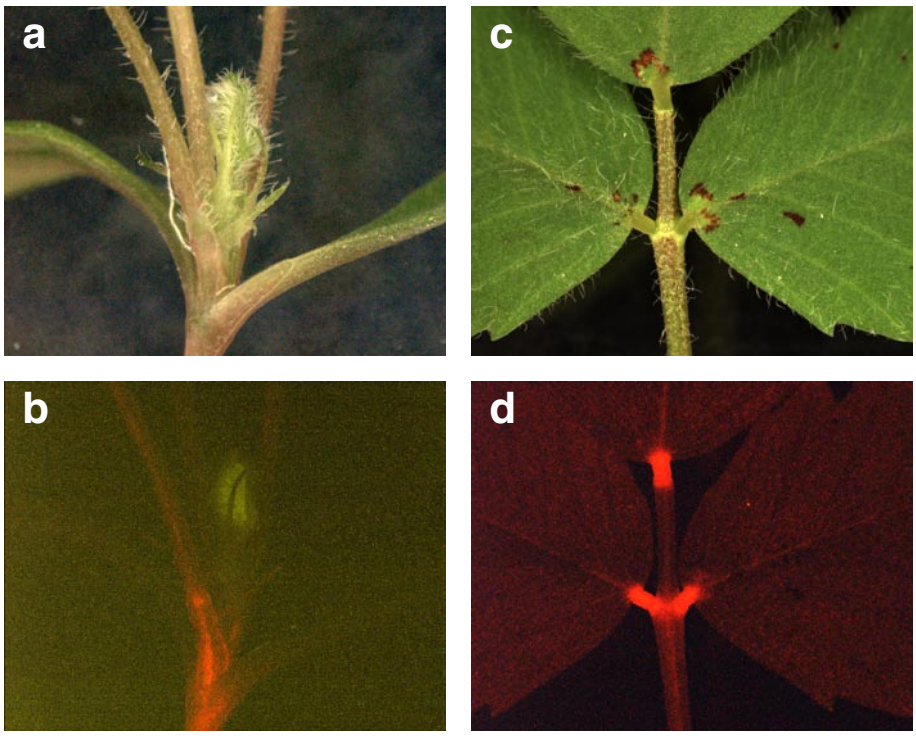


Figure 5.7 Expression of *ds*RED and eYFP in the aerial parts of the *cam40bc-1* plants. Expression of *ds*RED is localised to the vascular tissue, especially in the pulvinus of the leaf (panel d). eYFP fluorescence is localised to the non-vascular tissue of the young developing leaf (panel b).

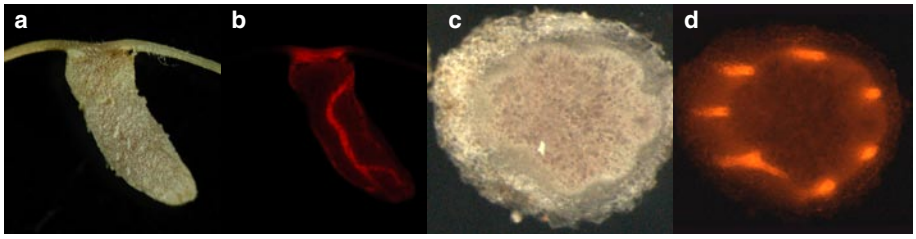


Figure 5.8 Expression of *ENOD40* in vascular tissue of the mature nodule. Panel d shows red fluorescence in a cross section of a nodule, clearly showing localisation to the vascular tissue.

with rhizobia leading to root nodule formation. Expression is first highly elevated in root pericycle cells and later in cortical cells that have dedifferentiated and subsequently started to divide forming the nodule primordium (Compaan *et al.*, this thesis). To analyse the expression of the two reporters during nodule development, plants were grown on agar plates under nitrogen limiting conditions and inoculated with *Sinorhizobium meliloti (41)*. Two days after inoculation, *ENOD40::dsRED* expression was observed in the pericycle and in the dividing cortical cells that form the nodule primordium, like has been shown for the expression of *ENOD40* by *in situ* hybridisation (Compaan *et al.*, this thesis) (Figure 5.6). A high level of expression in the infection zone in the central part of a developing nodule at 7 days post inoculation is visible. In 4-week old nodules expression is confined to nodule vascular tissue and to cells located at the base of the nodule (Figure 5.7). During nodule development *DN40::eYFP* expression was only observed between 2-3 days after inoculation, albeit at a low intensity. However, it can not be excluded that the low intensity obtained with YFP filter setting is due to the rather high level of *ds*RED as the excitation spectra of dsRed and eYFP partly overlap.

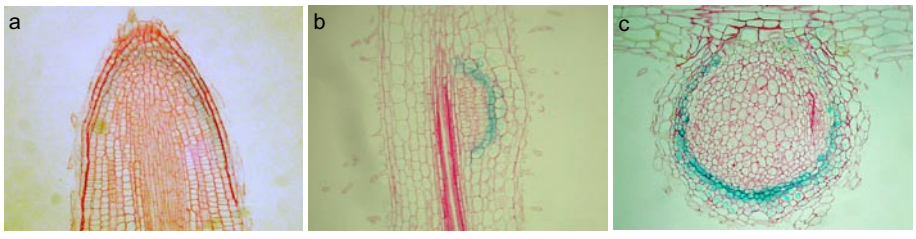


Figure 5.9 Localisation of the DN40 promoter driven GUS expression in *Medicago* root and nodules. Thin sections of a root tip (a), a 3 d.p.i (c) and a 7 d.p.i. nodule (d). In a 3 d.p.i. nodule, expression was detected in the periphery of the primordium (c) and in a 7 d.p.i. nodule in the nodule endodermis and the adjacent outer cortex cell layer (d). In the root tip GUS expression is detected in in a few cells of the protoderm and of the lateral root cap.

Since the expression of *DN40::eYFP* could not be localised in roots and nodules at the cellular level, we decided to study *DN40* promoter activity in hairy roots on *M. truncatula* involving the β-glucoronidase gene (GUS) as reporter. Therefore *M. truncatula* roots were transformed with *A rhizogenes* harbouring pCambia 1381Z from which the hygromycin gene including the 35S promoter and *nos* terminator were deleted and in which the *MtDN40* promoter was introduced to drive *GUS*.

Like in *DN40::eYFP* expressing plants, *DN40::GUS* expression was confined to cells in the root tip. To determine which cells express GUS activity, root tips were embedded and sections were analysed (Figure 5.9). From this it is clear that *DN40::GUS* is not expressed in a specific cell type, but in a few cells of the protoderm and of the lateral root cap. The expression of *DN40* is not detected in root epidermis cells and therefore, expression is down regulated when protoderm cells become epidermis cells.

To determine in which cells *DN40* is active in nodules, hairy roots were inoculated with *S. meliloti* RM2011. Roots and nodules were investigated for GUS activity 3 and 7 days after inoculation. Regions of the root or nodules that stained blue, were embedded and sections were made to determine in which cells *DN40::GUS* is active. Sections of 7-day-old nodules learned that GUS is active in two cell layers outside of the central tissue of the nodule. The *DN40* expressing layers are the nodule endodermis and the adjacent layer that is part of the outer cortex. In 3-day-old nodules *DN40::GUS* activity was also detectable at the periphery of the nodule primordium (Figure 5.9). The *DN40::GUS* expressing cells are in a file and this file of cells is connected to the root endodermis. At this stage of development *ENOD40* expression occurs in cells in the central part of the nodule primordium. Thus in none of the stages a co-expression for *ENOD40* and *DN40* has been observed.

Discussion

Analyses of the genomic context of *ENOD40* in 4 plant species revealed a strict gene coupling between *ENOD40* and a gene we named *DN40* that is highly homologous to the *Arabidopsis thaliana* gene coding for protein At1g60790. The chance to find microcolinearity between monocots and dicots is exceptionally low as deduced from studies in *A. thaliana* and *O. sativa* (Liu *et al.*, 2001). Therefore, we consider the finding of the strict linkage between *ENOD40* and *DN40* as present between *O. sativa* and *M. truncatula* a good indication that the coupling has a biological significance. In eukaryotes, prokaryotes and archea the extensive analysis of a number of completely sequenced genomes has revealed that strict colinearity between divergent species is mainly confined

to gene couples.

It has been shown by studies in yeast and *Drosophila melanogaster* that gene couples have a high probability to be co-expressed (Cohen *et al.*, 2000; Cho *et al.*, 1998). In *Arabidopsis thaliana* for a set of co-regulated genes chromosomal clustering has been uncovered and co-regulated genes are most likely functionally related as they are e.g. co-expressed in the same domain of the root or in response to auxin (Birnbaum *et al.*, 2003).

Along this line, we postulated that the strict linkage between *ENOD40* and *DN40* could point to a co-regulated expression of these two genes. RT-PCR analysis showed that *ENOD40* and *DN40* were both expressed in all organs tested. By using *eYFP* and *ds*RED reporters we showed that *DN40* promoter activity is confined to non-vascular cells of young leaves and to cells in the root tip, while the *ENOD40* promoter is active in vascular tissue, which observation is in agreement with *in situ* hybridisation studies described for *ENOD40* expression earlier. This implies that the 2 genes are expressed in different cells within the root and the leaf. The use of vital reporters however, did not allow a precise localisation of *ENOD40* and *DN40* expression at the cellular level. Therefore, we determined *DN40* expression in hairy roots transformed with *A.rhizogenes* carrying a plasmide containing *DN40::GUS*. In this way we showed that *DN40* is active in protoderm cells and adjacent lateral root cap cells. Therefore, the observed *DN40::eYFP* expression most likely is present in these cells in the root tip. In nodules *DN40::GUS* activity occurs in endodermis and an outer cortical cell layer, cell types devoid of *ENOD40* expression. As *DN40* expression is not observed in root endodermis, but in protoderm and a subset of lateral root cap cells, *DN40* is not expressed in a tissue specific manner. The observation that expression of *ENOD40* and *DN40* does not co-localise in roots, leaves and nodules suggests that their strict coupling has another biological significance that remains to be elucidated.

Materials and Methods

Primers

RACE-A	5′-CATCTAGAGGATCGAATTC-3′
RACE-T	5′-CATCTAGAGGATCGAATTC-$T_{(16)}$-3′
emtr	5′-GGACTCACTCACTCTCTCA-3′
mtpf1	5′-GATTCACTGAACCGGAACAT-3′
mtpf2	5′-CATTGGTGGACCCACGAGAAAACA-3′
mtpr1	5′-TGTTTTCTCGTGGGTCCACCAATG-3′
mtpr2	5′-ATGTTCCGGTTCAGTGAATC-3′
mtprom1	5′-GGAATTCGTAAATTGTCAGTC-3′

startkpnI	5′-GGGGTACCTTATTACAAACAAGATTCAAGTC-3′
red1kpnI	5′-GGGGTACCACAATGGCGCGCTCCTC-3′
red2xhoI	5′-CCCTCGAGTACAGGAACAGGTGGTG-3′
mt564xhoI	5′-CCCTCGAGGTGAGAGGGTATTAAACAAAAAC-3′
mt564xbaI	5′-GCTCTAGAGGGCATTGGAAAAGTTGAGC-3′
yfpsstI	5′-GAGCTCACAATGGCAGTGAGCAAGGGCGAGGAGC-3′
yfpnheI	5′-GCTAGCACTTGTACAGCTCGTCCAT-3′
promdn40sstI	5′-GAGCTCTGAAGAAGATGAAGAAGTGAAAG-3′
promdn40salI	5′-GTCGACTTTAAAAAATACAAAGTGGTGAATG-3′
e40f	5′-GCTCTAGACCCTTTAAGCATCCTCTA-3′
e40rv	5′-CGGGATCCCACAAACAAACAAGCATAC-3′
ubifw	5′-ATGCAGAT(C/T)TTTGTGAAGAC-3′
ubirv	5′-ACCACCACG(G/A)AGACGGAG-3′

Genomic sequences

Filters spotted with a *Medicago truncatula* BAC library (Nam Y-W *et al.*, 1999) were obained from Clemson University Genomics Institute and screened with a radiolabelled *MtENOD40* cDNA fragment as a probe. Two hybridising BACs were obtained from the same institute. These BACs were subsequently subcloned in cloning vector pBluescript II KS+ (Stratagene), using the following restriction enzyme combinations: *Bam*HI-*Hin*dIII, *Sal*I-*Cla*I, *Sal*I-*Nsi*I and *Hin*dIII- *Eco*RI. The subclones were sequenced using the T3 and T7 primers and the sequences were assembled using SeqMan II software (DNAstar Inc.).

A genomic library of *Pisum sativum* in the lambda DASH vector (Stratagene) was screened using a radiolabeled *PsENOD40* cDNA fragment as a probe. One hybridising clone was identified and its DNA was subcloned in cloning vector pBluescript II KS+ (Stratagene) using restriction enzyme combination *Hin*dIII-*Eco*RI. The subclones were sequenced using the T3 and T7 primers. Sequencing was completed by direct sequencing on the lambda clone with primers designed on obtained sequences. The sequences were assembled using SeqMan II software (DNAstar Inc.).

A genomic library of *Lycopersicon esculentum* VFN8 in the EMBL3 vector (Clontech) was screened using radiolabelled *LeENOD40* as a probe. This screen resulted in 51 positive clones, two were selected for subcloning. The clones were subcloned in cloning vector pBluescript II KS+ (Stratagene), using the following restriction enzyme combinations: *Eco*RI-*Sal*I and *Hin*dIII-*Pst*I. The subclones were sequenced using the T3 and T7 primers and the sequences were assembled using SeqMan II software (DNAstar Inc.).

RNA isolation

RNA was isolated using the hot phenol method for RNA isolation (Pawlowski *et al.*, 1994). After RNA isolation the samples were treated with DNaseI (Promega) to remove chromosomal DNA.

RACE

To determine the 5' end of the *MtDN40* mRNA a 5' RACE was performed on total RNA of *Medicago truncatula* cv. A17. For this, cDNA was constructed from 20 μg of total RNA (Albrecht *et al.*, 1998) using the primer emtr (100 ng). The sample volume was adjusted to 100 μl and purified using a PCR Purification Kit (Boehringer). The sample was eluted from the PCR purification column using 50 μl 1x TdT buffer. TdT tailing was performed in a total volume of 55 μl containing 1x TdT buffer, 25 μM dATP and 15 U TdT. This mixture was incubated for 30 minutes at 37° C and 5 minutes at 70° C to inactivate the TdT enzyme. 5 μl of the tailed cDNA was used for the first PCR for which the primers mtpr1 (100 ng) and RACE-T (200 ng) were used. A hot start PCR program was used: 5 min 94° C, 3 min 80° C (in this step the Tag polymerase was added); 94° C 15 sec, 55° C 30 sec, 72° C 40 sec, for 10 cycles; 94° C 15 sec, 55° C 30 sec, 72° C 40 sec with 20 sec extension, for 20 cycles; 5 min 72° C. A second PCR was performed with mtpr2 (100 ng) and the RACE-A primer (200 ng). The products were cloned in the pGEM-T vector (Promega) and the nucleotide sequence was determined.

RT-PCR

RT-PCR was performed (Albrecht *et al.*, 1998) using the following PCR program: 30 sec at 94° C, 30 sec at 50° C and 30 sec at 72° C and 20 cycles. In the RT-PCR, *UBIQUITIN* was used as a standarisation control using the primers *ubifw* and *ubirv*. To detect *ENOD40* the primers e40f and e40r were used and to detect *DN40* the primers mtpf1 and mtpr1 were used.

generation of the T-DNA construct

The promoter of *MtENOD40* including the start codon of the box1 peptide was isolated by PCR amplification on BAC MT_ABa39I20 with primers *mtprom1* and *startkpnI* and cloned in pCAMBIA1300 (Cambia, Australia) using *Eco*RI and *Kpn*I generating pCAM-PROM1.

The *ds*RED coding sequence was isolated by PCR amplification with primers

red1kpnI and *red2xhoI* and cloned in pGEM-T (Promega), clones with the right orientation (T7→RED←SP6) were selected and denoted: pGEM-RED.

The intermediate fragment (IF) from the stop codon of the box1 peptide to the stop codon of DN40 was isolated by PCR amplification on BAC MT_ABa23A06 with primers *mt564xhoI* and *mt564xbaI* and cloned in pGEM-T. Clones with the right orientation (T7→IF←SP6) were selected and denoted: pGEM-IF.

The RED fragment was released from pGEM-RED by digestion with SphI and XhoI and cloned in the *SphI*/*XhoI* sites of pGEM-IF, generating pGEM-RED-IF.

The RED-IF fragment was released from this vector by digestion with *Kpn*I and *Xba*I and cloned in the *Kpn*I/*Xba*I sites of pCAM-PROM1 generating pCAM-PROM1-RED-IF.

The eYFP coding sequence was isolated by PCR amplification on pEYFP-C1 (Clontech) using primers *yfpsstI* and *yfpnheI* and cloned in pGEM-T, clones with the right orientation (SP6→YFP←T7) were selected and denoted: pGEM-YFP.

The promoter of *DN40* was isolated by PCR amplification on BAC MT_ABa23A06 with primers *promdn40sstI* and *promdn40salI* and cloned in pGEM-YFP using SalI and SstI generating pGEM-PROM2-YFP.

Finally the PROM2-YFP fragment was released by digestion with SalI and NheI and cloned in the *SalI*/*XbaI* sites of pCAM-PROM1-RED-IF, generating CAM40BC.

This construct was transformed to *A.tumefaciens* GV3101 containing helper plasmid C53C1 by means of electroporation.

Medicago truncatula R108-1(c3) plants were transformed using regeneration-transformation (Trinh *et al.*, 1998).

Plant growth conditions

Plants were grown aseptically on agar plates containing MS (Murashige and Skoog medium, Duchefa) in a growth chamber at 25° C and a 16 hour photo-period (7W/m^2).

Nodule initiation

For microscopic analysis, the plants were grown aseptically on agar plates containing Fåhreus medium (Fåhreus, 1957) in a growth chamber at 25° C

and a 16 hour photoperiod (7W/m^2). The roots of the plants were sprayed with *Sinorhizobium meliloti* 41 harbouring GFP expression cassette pHC60 (Cheng *et al*., 1998). Rhizobia were grown in YEM (0.5% manitol, 0.05% yeast extract, 0.02% $MgSO_4 \cdot 7H_2O$, 0.01% NaCl, 0.05% K_2HPO_4, 0.5% NaGluconate, 0.016% $CaCl_2$) and diluted with water to OD 0.01.

For quantifiation and classification of nodules and also for the isolation of nodule RNA, plants were grown in gravel supplemented with Fåhreus medium in a growth chamber at 25º C and a 16 hour photoperiod (7W/m^2). One week after germination, the plants were inoculated with *S.meliloti 41-GFP*.

Microscopy

YFP and *ds*RED fluorescence was captured using a stereo fluorescence macroscope (Leica MZIII) equiped with *ds*RED, eYFP and GFP specific filters (Chroma).

For detailed analysis a confocal microscope was used: LSM 510 (Zeiss). YFP was excitated with an argon ion laser at 514 nm and RFP with a helium neon laser at 543 nm.

GUS staining

GUS staining was performed according to the protocol of Kosugi *et al.* (1990). Plant material was stained in GUS buffer (0,1 M Na_2PO_4 pH 7.2, containing 1 mM X-Gluc dissolved in DMF and 0.1% Triton X-100). Plant material was put 10 minutes under vacuum and was incubated at 37°C for 4 hours or longer until the blue GUS stain was visible.

Embedding of root tissue

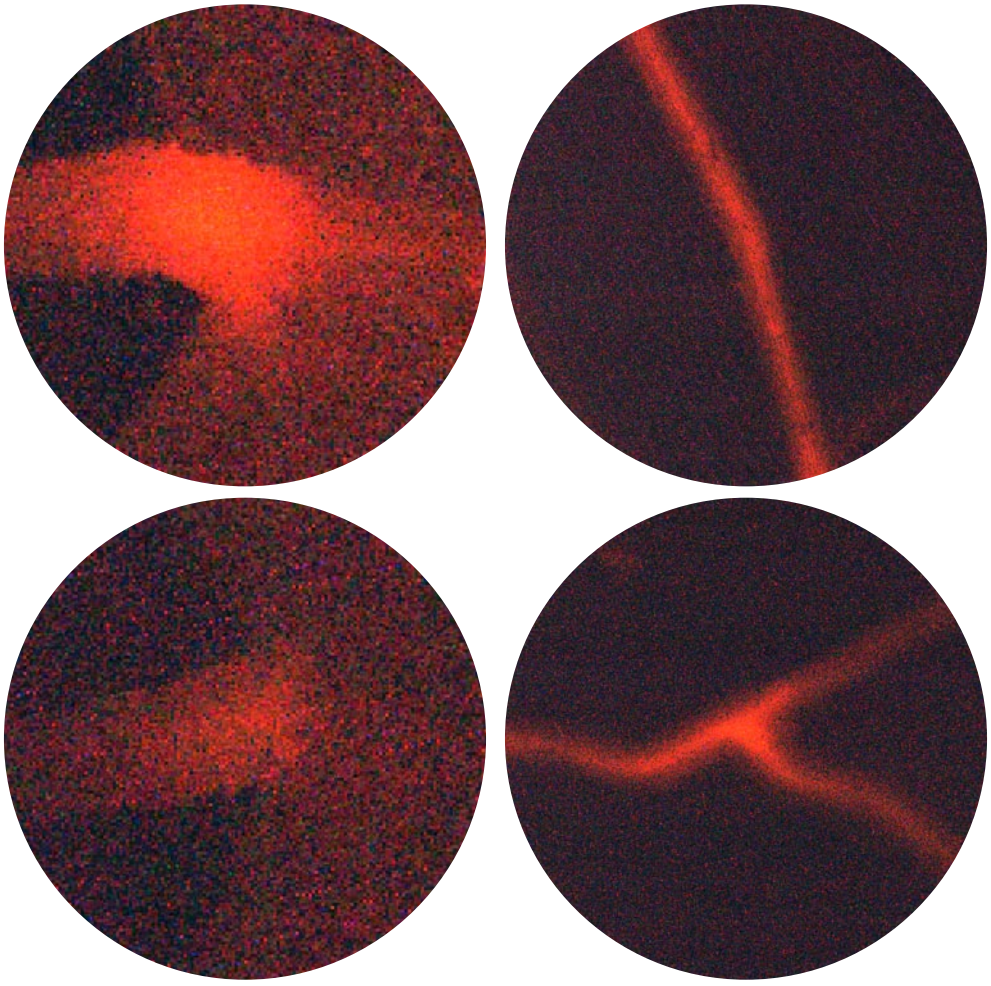
Fixation of root material was performed with 5% glutaraldehyde under vacuum for 2 hours. Subsequently this material was embedded in plastic using the Technovit Glycol Methacrylate Embedding Kit 7100. The embedded material was sectioned in 8 μm thick slices. Images of these sections were captured using a Nikon optiphot-2 microscope equiped with a Nikon coolpix digital camera.

Hairy root transformation

The roots of 5 day-old *M. truncatula* seedlings were removed at the hypocotyl

and the wound surface inoculated with *Agrobacterium* MSU440 (Sonti *et al.*, 1995) containing the *pDN40::GUS* vector. The seedlings were co-cultivated with *Agrobacterium* for 5 days at 21°C (16h photoperiod) and subsequently transferred to Emergence medium (3 mM MES pH 5.8 containing 2,5 g/L KNO3, 0.4 g/L MgSO4·7H2O, 0.3 g/L NH4H2PO4, 0.2 g/L CaCL2·2H2O, 10 mg/L MnSO4·4H2O, 5 mg/L H3BO3, 1 mg/L ZnSO4.7H2O, 1 mg/L KI, 0.2 mg/L CuSO4.5H2O, 0.1 mg/L NaMoO4·2H2O, 0.1 mg/L CoCL2·6H2O, 15 mg/L FeSO4·7H2O, 20 mg/L Na2EDTA, 100 mg/L Myoinositol, 5 mg/L Nicotinic acid, 10 mg/L Pyriodoxine HCl, 10 mg/L Thiamine HCl, 2 mg/L Glycine, 1% sucrose, 0.9% Daichin agar containing 300 μg/ml Cefotaxime (Duchefa)).

Plants were grown for 6-18 days on Emergence medium. In this period new roots formed that were potentially co-transformed with the T-DNA of the binary vector. Subsequently plants were transferred to gravel and inoculated with *S.meliloti 41* as described.



Chapter 6

A highly conserved region in the 3'UTR of *ENOD40* is involved in translational regulation.

Bert Compaan, Ton Bisseling and Henk Franssen

Abstract

ENOD40 genes are characterised by two conserved regions. Several experiments have indicated the presence of biological activity in these 2 domains, that either can act independently from each other or in a yet unknown complex interaction (Sousa *et al.*, 2001; Ruttink *et al.*, 2003). Whereas region 1 encodes a small peptide the nature of the biological activity residing in the second conserved region however is still unclear. We have used stably transformed *Medicago truncatula* plants to show that deletion of the second conserved region stimulates the translation of the *ds*RED reporter, replacing the first conserved region. The concentration of transcript however is not influenced by the deletion and therefore we conclude that the second conserved region of *ENOD40* has a repressive effect on the translation of the ORF present in the first conserved region.

Introduction

Bacteria of the genera *Rhizobium*, *Bradyrhizobium* and *Azorhizobium* are able to induce the formation of a new organ on the roots of legumes, the root nodule. The steps that lead to the formation of this symbiotic organ are marked by the expression of specific genes. One of these genes is *ENOD40*. After inoculation with rhizobia this gene is upregulated in the pericycle cells opposite cortical cells that are activated to divide to form the nodule primordium. Comparison of *ENOD40* genes from different plant species has revealed that sequence homology is restricted to two regions, designated box1 and box2, while the overall sequence homology is low. Strikingly, the tertiary structure of *ENOD40* RNA is well conserved (Girard *et al.*, 2003). Within leguminous *ENOD40* mRNAs five conserved structural domains can be identified of which three are also present in *ENOD40* mRNAs from non-leguminous species. Furthermore, an additional conserved structural domain is present in the RNA from legumes on which indeterminate nodules are formed (Girard *et al.*, 2003). Intriguingly, the nucleotide sequences of box1 and box2 are not part of the structural conserved domains, raising the question on the nature of the biological activity enclosed within *ENOD40*; is it enclosed in the tertiary structure or in the conserved nucleotides in box1 and box2 or a combination of both? Box1 spans the most 5'-located ORF in the transcript and is in all known *ENOD40* genes bordered by a start and a stop codon, indicating that this region is translated (Compaan *et al.*, 2001; Sousa *et al.*, 2001; Rohring *et al.*, 2002; Ruttink *et al.*, 2003). The putative translation product of box1 is a small peptide of 10-13 amino acids and the peptides contain the motif W-(X4)-HGS

(Compaan *et al.*, 2001, 2003; Sousa *et al.*, 2001; Rohrig *et al.*, 2002; Ruttink *et al.*, 2003). Box2 consists of 51 nucleotides and is located in the centre of the gene between the structural domains 2 and 3 (Girard *et al.*, 2003). For 50% of the *ENOD40* genes known, the reading frame spanning this region does not include an AUG, thus it is unlikely that in these *ENOD40* genes this region is translated into a protein. For the other 50%, the position of start and stop codons flanking the second conserved region are not conserved among the *ENOD40* genes. Consequently, translation of this region into a protein would not lead to proteins highly homologous among all *ENOD40* genes. It is therefore likely that this region is not translated, although it cannot be ruled out completely. Cowpea protoplasts that are transfected with *ENOD40* in which *GFP* was cloned in frame with a putative box2 start codon, fail to express the GFP protein, indicating that box2 is not translatable (Compaan *et al.*, 2001). However, bombardment of *Medicago truncatula* roots with DNA constructs in which GUS was cloned in frame with a putative box2 start codon results in detectable GUS activity in these roots, indicating that the AUG in this gene is recognised as a start of translation. Moreover, using the same approach, it has been reported that transient introduction of transcripts containing only box2 can induce cortical cell divisions, whereas a mutation in the putative start codon leads to a failure in inducing such cortical cell divisions (Sousa *et al.*, 2001). This strongly indicates that a biological activity resides in box2 and that this activity is mediated by a protein.

Alternative to a protein encoding function of box2, it has been proposed that box2 is involved in regulating the translation of box1 (Van de Sande *et al.*, 1996). Such a regulatory mechanism is a common function of UTRs in eukaryotic messengers, allowing the regulation of expression of genes on the level of translation. Also in plants, sequences within 3'UTRs of messengers were identified that modulate protein expression (An *et al.*, 1989; Dietrich *et al.*, 1992).

To analyse whether *ENOD40* box2 has a regulating effect on translation of box1, we tested whether the deletion of box2 specific sequences influences the expression of the *ds*RED reporter that is replacing box1.

Results

Generation of the del box2 line

To determine whether box2 has an effect on translational or transcriptional efficiency of *ENOD40*, two T-DNA constructs were created. CAM40BC as described in chapter 5 contains the *MtENOD40* gene in which the region

of 39 nucleotides that encode the peptide of box1 is replaced by the protein coding part of the *ds*RED reporter gene. A second construct CAM40BCDEL2 was derived from CAM40BC, by deleting a 62 nucleotide long region (hereafter called box2) containing the 51 nucleotides of box2 (Figure 6.1) and the conserved structural domain 3 (Girard *et al.*, 2003). Both constructs were transformed to *Medicago truncatula cv. R108* plants. The two sets of transformed lines were subsequently called: *cam40bc* and *cam40bcdel2*, respectively.

```
ENOD40      CAGAGACACC AACTTCCCCA CTACCTTTCT ATGTGGAGCC CTTTAAGCAT 50
del 2       CAGAGACACC AACTTCCCCA CTACCTTTCT ATGTGGAGCC CTTTAAGCAT

ENOD40      CCTCTAAACC AATCCATCAA GACTTGAATC TTGTTTGTAA TAAGGATGAA 100
del 2       CCTCTAAACC AATCCATCAA GACTTGAATC TTGTTTGTAA TAAGGATG--

ENOD40      GCTTCTTTGT TGGGAAAAAT CAATCCATGG TTCTTAAAAC AAACATGGAG 150
del 2       ---------- dsRED coding sequence ----TAAAAC AAACATGGAG

ENOD40      AGAAGTGTGA GAGGGTATTA AACAAAAACC CTACACACTC TCCCTCCATT 200
del 2       AGAAGTGTGA GAGGGTATTA AACAAAAACC CTACACACTC TCCCTCCATT

ENOD40      TTCCTAAACA GTTTGCTTTG TGCTTTAGCT TTTGGCTTCT CATATCACAA 250
del 2       TTCCTAAACA GTTTGCTTTG TGCTTTAGCT TTTGGCTTCT CATATCACAA

ENOD40      AGGGATTATG CTTTTTTCTG AGTAGCAGAA GCAAATAATT AAGTATTTTT 300
del 2       AGGGATTATG CTTTTTTCTG AGTAGCAGAA GCAAATAATT AAGTATTTTT

ENOD40      CTCCAAAGGA TCAGAAGCTT TTGTTATAGC ATGGCAAACC GGCAAGTCAC 350
del 2       CTCCAAAGGA TCAGAAGCTT ---------- ---------- ----------

ENOD40      AAAAAGGCAA TGGATTCCTT TTTGGAGTCT TAATGGCTAT GTATCAATCA 400
del 2       ---------- ---------- ---------- --ATGGCTAT GTATCAATCA
                                              AAGCTT

ENOD40      CTCTATCTAT GTAGCACTGA CACTTGAGAT TGTAGGCGCG TCCTATGCCT 450
del 2       CTCTATCTAT GTAGCACTGA CACTTGAGAT TGTAGGCGCG TCCTATGCCT

ENOD40      GTGTTTGTGC TTGTAGATTG TTATAGTTAT TTTCTTGCAG TAGAATGTAA 500
del 2       GTGTTTGTGC TTGTAGATTG TTATAGTTAT TTTCTTGCAG TAGAATGTAA

ENOD40      TAATAAACAT AAAGATGGTG TTGTCTTCCT TTGAGAAATT GCCAACTTTA 550
del 2       TAATAAACAT AAAGATGGTG TTGTCTTCCT TTGAGAAATT GCCAACTTTA

ENOD40      TGATGTACTT CAATTCACTC AATTTGCAGC TGACTAGAGT CTGTTCTTGT 600
del 2       TGATGTACTT CAATTCACTC AATTTGCAGC TGACTAGAGT CTGTTCTTGT

ENOD40      TTCAGTTTCT GCAGATGAGT AAGGTAGGTA ACTGTTATCA TTAATTCATG 650
del 2       TTCAGTTTCT GCAGATGAGT AAGGTAGGTA ACTGTTATCA TTAATTCATG

ENOD40      TTCCTTTTCT TCT 663
del 2       TTCCTTTTCT TCT
```

Figure 6.1 Nucleotide sequence of '*MtENOD40 del box2*' aligned to the nucleotide sequence of *MtENOD40*. The deleted nucleotides are indicated with (-), the nucleotide sequence that was used for the primer is single underlined and the *HindIII* recognition sequence within *MtENOD40* and the one that was adapted to the primer are double underlined. The box2 motif is indicated in bold.

To obtain homozygous lines that contain a single T-DNA insertion, from each set 3 T1 plants were selected. Offspring for which in the T2 the *ds*RED fluorescence segregated in a 3:1 ratio as expected for a single insertion line was selected. For two lines per set, the presence of a single T-DNA insertion in the genome of these plants was subsequently confirmed by Southern analysis of genomic DNA isolated from individual plants of these lines (data not shown). These 4 lines, designated *cam40bc-1*, *cam40bc-6*, *cam40bcdel2-8* and *cam40bcdel2-20* were selected for further studies. The selected lines showed a fluorescence intensity that did not deviate from the average *ds*RED expression of other lines from the same set.

Deletion of box2 leads to an increase of dsRED fluorescence

To analyse whether the deletion of box2 influences the pattern or the level of *ds*RED expression, *cam40bcdel2* lines were compared with *cam40bc* lines. Plants of the 4 chosen lines were grown on agar plates containing 1x MS. After 10 days, the roots and leaves of the plants were analysed with the stereomacroscope for *ds*RED fluoresence. In all the plants, *ds*RED fluorescence could be detected in the root vascular tissue and in the pulvinus of the leaf. This indicates that deletion of box2 does not affect the pattern of *ENOD40* expression. However the intensity of fluorescence is higher in all plants of the 2 *cam40bcdel2* lines (Figure 6.2), when compared to 2 *cam40bc* lines. Although the number of analysed lines is small this suggests that deletion of box2 has an effect on the quantity of *ds*RED produced in the transgenic lines.

ENOD40 is expressed highest in nodules, therefore the intensity of *ds*RED fluorescence in the *cam40bc* and *cam40bcdel2* lines was also analysed in root nodules. Five plants from each line were grown in vermiculite containing Fahreus medium. Five days after germination, the plants were inoculated with *Sinorhizobium meliloti 41*. After 2 weeks the induced nodules were embedded in agarose and 100μM thick sections were made and analysed with the stereomacroscope for the pattern and intensity of *ds*RED fluorescence *ds*RED occurs in the nodule vascular tissue and the infection zone of the central tissue in nodules of *cam40bc-1* as well as *cam40bcdel2-20* lines. This pattern is similar to that of *ENOD40* expression (Crespi *et al.*, 1994). Also within nodules, the fluorescence intensity in *cam40bcdel2-20* nodules is higher than in *cam40bc* nodules (data not shown).

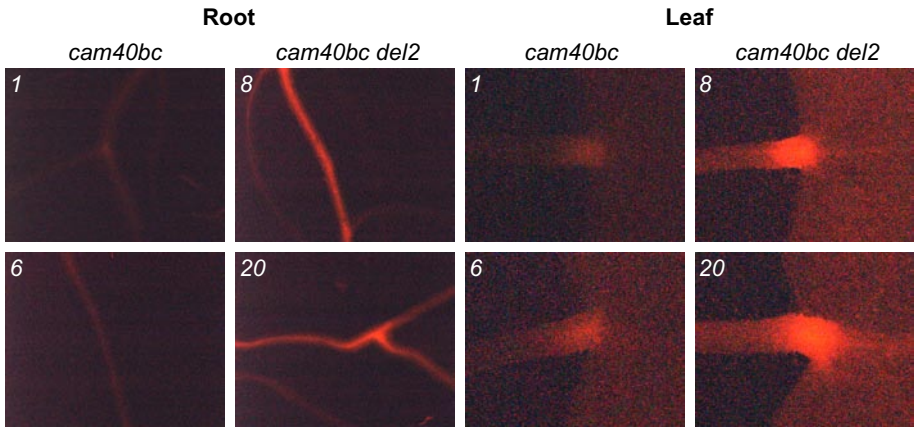


Figure 6.2 Comparison of the root and leaf pulvinus of *M.truncatula* plants from lines *cam40bc-1*, *cam40bc-6*, *cam40bcdel2-8* and *cam40bcdel2-20*. In the lines that contain the construct in which box2 was deleted, a clear increase in the amount of *ds*RED fluorescence can be observed.

Box2 acts as a negative regulator of translation *in cis*

The increased *dsRED* levels in the absence of box2 could be due to an increased transcript level, e.g. by a higher transcription level or transcript stability. Alternatively, box2 has a negative effect on the translation efficiency. To distinguish between these two options the ratio of the transcript and protein levels was determined in the 4 lines.

In order to quantify the intensity of *ds*RED fluorescence in *cam40bcdel2* and *cam40bc* lines, 10 plants of each line and 10 *Medicago truncatula R108 wt* plants were grown on plates for 10 days and protein extracts were made from these plants. Total protein concentrations of these extracts were determined using the Biorad method and all the samples were diluted with FPS buffer to a final concentration of 0.5 μg/μl.

From these diluted samples, the fluorescence intensities were measured using a spectrofluorometer. The samples were excitated at 550 nm and the emission spectrum between 565 to 650 nm was analysed. The height of the obtained emission peak at 583 nm was determined for each sample. The amount of autofluorescence as measured in the protein extract from the *wt* plants was subtracted from this value. The in this way obtained quantification of *ds*RED fluorescence shows that the amount of *ds*RED in plants from 2 different *cam40bc* lines is almost equal (Figure 6.3). The *ds*RED fluorescence in plants from two *cam40bcdel2* lines is not equal when compared to each

other. However, when compared to the fluorescence in *cam40bc* plants the *ds*RED fluorescence in extracts from *cam40bcdel2* lines is at least two times higher than in the *cam40bc* lines (Figure 6.3).

To determine the amount of transcript, 50 plants from the lines *cam40bc-1* and *-6* and from lines *cam40bcdel2-8* and *-20* were grown on plates. After 10 days, RNA was isolated from the seedlings. RNA was denatured, separated by electrophoresis in agarose and the gel was blotted subsequently. Hybridisation of this blot to ^{32}P labeled *dsRED* (Figure 6.3) shows that *DsRED* expression is higher in *cam40bcdel2* lines than in *cam40bc* lines. To be able to quantify

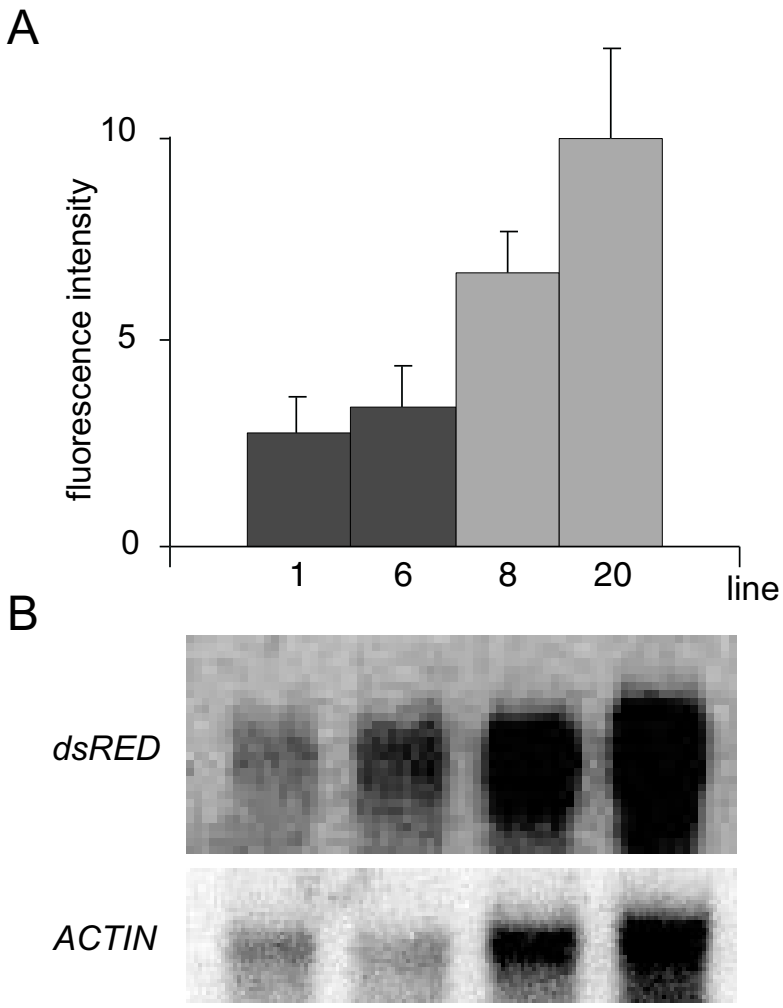


Figure 6.3 *ds*RED is higher expressed in lines 8 and 20 as compared to lines 1 and 6 (A). Quantification of the transcript levels by Northernblotting (B) however reveals that the *dsRED-ACTIN* transcript-ratio is the same in these four lines.

the level of *dsRED* transcript in the various lines, the blot was subsequently hybridised to ^{32}P labeled *Mtactin* probe (Figure 6.3). Also the expression level of actin is higher in *cam40bcdel2* lines. Quantification of the hybridisation signals revealed that the ratio of *dsRED* and *actin* mRNA is almost identical (Figure 6.3) in all RNA samples showing that, although not equal amounts of RNA had been loaded, the amount of *dsRED* containing transcripts is comparable between *cam40bc* and *cam40bcdel2* lines.

Based on the differences in *ds*RED fluorescence between *cam40bc* and *cam40bcdel2* lines, while the RNA levels of *dsRED* transcripts are similar in all lines, the ratio *ds*RED protein *dsRED* transcript is higher for c*am40bcdel2* plants, suggesting that box2 has a negative effect on the translation efficiency.

Discussion

Here we have used 2 lines of *Medicago truncatula*, *cam40bc* and *cam40bcdel2* lines, to determine whether box2 has an effect on translational or transcriptional efficiency of *ENOD40 in planta*. RNA analysis showed that the amount of *dsRED* mRNA was equal in *cam40bc* and *cam40bcdel2* plants, whereas fluorescence intensity measurements showed that the *ds*RED intensity in *cam40bc* is lower compared to *cam40bcdel2* seedlings. These observations fit with the assumption that deletion of box2 sequences has a stimulatory effect on the translation of the ORF upstream in the mRNA rather than on transcription of the RNA or on the stability of the mRNA. The latter is supported by RNA stability calculations on the *ENOD40* messenger that show that regions others than the two conserved domains including the deleted region from *ENOD40* that has been introduced into *cam40bcdel2* plants attribute to mRNA stability (Sousa *et al.*, 2001; Girard *et al.*, 2003).

The presence of regulatory sequences within 3'UTR is widespread among plant and animal kingdom. The regulatory activity of these elements often involves the binding of regulatory proteins. These proteins might bind directly to box2 sequences, which is attractive as box2 sequences are highly conserved among *ENOD40* genes. Alternatively, the deletion of box2 involves the disruption of a protein that is involved in translational control of the box1-encoded ORF.

The indication for the existence of a mechanism that controls the translation of box1-encoded peptide suggests that the production of the peptide is controlled. However, since no biological function has been assigned yet to the box1-encoded peptide, the biological implication of the observation made remains also unclear.

Taken together, our data show that the presence of box2 that is conserved at the nucleotide level between *ENOD40* genes and structural domain 3 conserved among *ENOD40* genes is involved in repression of translation of *dsRED* placed in the position of box1 and suggests a role for the peptide encoded within box1.

Materials and Methods

Primers

mt450hind	5′-CCCAAGCTTATGGCTATGTATCAATCACTC-3′
mt564xbaI	5′-GCTCTAGAGGGCATTGGAAAAGTTGAGC-3′

Generation of the T-DNA constructs

The generation of constuct CAM40BC is described in Chapter 5. For the generation of CAM40BCDEL2, plasmid pGEM-IF was used as a starting point. By PCR with primers *mt450hind* and *mt564xbaI* a genomic DNA fragment was amplified spanning the region between box2 and the stop codon of *DN40*. The primers that were used, contained restriction sites for HindIII and XbaI. Using these restriction enzymes, a fragment containing box2 was released form plasmid pGEM-IF and replaced with the PCR fragment, generating pGEM-IFDEL2. Subsequently the same cloning procedure was followed as described for CAM40BC in chapter 5, with this exception that instead of pGEM-IF, pGEM-IFDEL2 was used. In this manner CAM40BCDEL was obtained. This construct was transformed to *A.tumefaciens* GV3101 containing helper plasmid C53C1 by means of electroporation.

Medicago truncatula R108-1(c3) plants were transformed using regeneration-transformation (Trinh *et al.*, 1998).

Microscopy

YFP and RFP fluorescence was captured using a stereo fluorescence macroscope (Leica MZIII) equiped with *ds*RED, eYFP and GFP specific filters (Chroma).

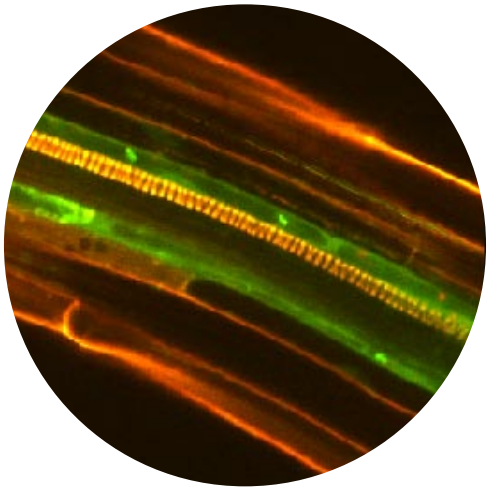
Quantification of fluorescence

For the quantification of fluorescence, plant material was collected and frozen

in liquid nitrogen. Frozen samples were grinded and subsequently suspended in FPS buffer (120 mM KCl, 50 mM Tris 8,0, 10% glycerol and 1 tablet per 10 ml protease inhibitor (Boehringer)). Total protein concentration was quantified using BioRad reagent and all samples were diluted to a final protein concentration of 0,5 μg/μl. From these samples a fluorescence spectrum was obtained by excitation at 550 nm and detection of fluorescence intensity from 565 nm to 650 nm. In all samples a *ds*RED peak was observed around 583 nm, from which the height was used as a representative of *ds*RED concentration. Protein extracts from non-transformed plant tissue were used to determine the base line of fluorescence emission.

Northern Blotting

RNA was isolated using the hot phenol method for RNA isolation (Pawlowski *et al.*, 1994). After RNA isolation the samples were treated with DNaseI (Promega) to remove chromosomal DNA (Albrechts *et al.*, 1998). RNA was denatured in DMSO and glyoxal and separated on an agarose gel. Subsequently the denatured RNA was transferred to Hybond-N^+ membranes (Amersham) (Sambrook *et al.*, 2001). Hybridisation of the membranes was performed using ^{32}P labeled DNA probes in a roller bottle. (Sambrook *et al.*, 2001).



Chapter 7
General Discussion

ENOD40 is involved in the induction of cortical cell divisions

Previously, it had been shown that Nod factors are sufficient to induce the formation of nodule primordia in which *ENOD40* is active. *ENOD40* was also active in the region of the pericycle opposite the primordium (Vijn *et al.*, 1995; Minami *et al.*, 1996). To test whether *ENOD40* could be part of a Nod factor activated process that leads to cortical cell division we studied the timing of *ENOD40* induction. We showed in chapter 2 that *ENOD40* is induced within 3 hours after spot inoculation with *S. meliloti*, in the root pericycle of *M. sativa* at the site of inoculation. This indicates that *ENOD40* is part of a Nod factor induced module/process, that precedes a morphological response in the cortex. The induction of *ENOD40* expression is at least 17 hours earlier than the first cortical cell divisions and therefore it might be part of a mechanism inducing these cortical cell divisions. The studies of Charon *et al.* (1997) strongly support this hypothesis as they showed that ectopic *ENOD40* expression in root cortical cells induces divisions in other cortical cells. So ectopic expression of *ENOD40* appears to be sufficient to induce cortical cell divisions in a non-cell autonomous manner. Therefore we hypothesise that the induction of *ENOD40* in the pericycle is an essential part of the mechanism by which Nod factors trigger division in the cortex. An important role for *ENOD40* in such mechanism is consistent with its sequence indicating that *ENOD40* encodes a peptide. Further, the strong conservation of a domain in the non-coding part of the mRNA suggests that (also) the RNA has a biological activity.

Box2 is not translated and is involved in the regulation of translation of Open Reading Frame present in box1.

Comparison of the nucleotide sequences of all studied *ENOD40* genes learned that the highest sequence conservation occurs in two regions, designated box1 and box2 that encompass 20-30% of the RNA. In all *ENOD40* genes box1 sequences encode a peptide of 10-13 amino acids. Further, it has been demonstrated (van de Sande *et al.*, 1996) that the ATG present in box1 is used as a start for translation. Antibodies that were raised against the peptide encoded within box1 could recognise an antigen in nodule extracts of *G. max* and

M. sativa (van de Sande *et al.*, 1996; Sousa *et al.*, 2001). These observations strongly argue that box1 encodes a peptide.

The number of ORFs within box2 varies among plant species from 0 to 3. Further, in several plant species one or more ATGs are present in box2 of *ENOD40* giving rise to non-overlapping ORFs. These observations, together with the presence of frame-shift inducing polymorphisms at the nucleotide level imply that the putative peptides encoded within box2 sequences of different plant species can share only a low degree of similarity. Therefore, the question arose whether box2 is translated in a peptide like box1 (Charon *et al.*, 1997; John *et al.*, 1997). To analyse whether a peptide could be translated from box2, we prepared translational fusions of ATGs with GFP in either box1 or 2 of *NtENOD40*. We showed that box1 is translated in Cowpea protoplasts, and also that box2 is translated but at a markedly lower level compared to box1. Further, when the putative start codon of the box2 peptide of *NtENOD40* is preceded by box1, as in the natural situation, it was not used as translation start (Chapter 2). Therefore, these experiments show that the ATG codon in box1 is used as translation start, whereas the one in box2 is not.

In *ENOD40-GUS* reporter fusion studies in *M. sativa* it was shown that the translational machinery uses the various ATGs within the transcript of *MtENOD40* as start codon with different efficiency (Sousa *et al.*, 2001). The ATG of box1 is used as a translation start with the highest efficiency. The start codon present in box2 giving rise to a peptide highly homologous to the NtENOD40- box2 encoded peptide is used with 3-fold less efficiency. However, after removal of box1, this codon is now used with a similar efficiency as the codon in box1 (Sousa *et al.*, 2001). So these studies indicate that the ATG in box1 is more efficiently used as start of translation than the one in box2, although, our studies indicate a larger difference in efficiency than these studies. Further, Sousa *et al.* showed that mutation of an ATG within box2, can significantly reduce the biological activity of the introduced DNA in a transient expression assay, suggesting that box2 activity is mediated by a peptide.

In case this box is translated, in such species it can encode for mutually divergent peptides. Box2 is strongly conserved at the nucleotide level even if it does not encode a peptide. Therefore, it seems probable that it is at least active at the nucleotide level.

Van de Sande *et al.* (1996) hypothesised that box2 is involved in regulating the translation of the peptide encoded in box1. To test this hypothesis we prepared transgenic plants carrying the *dsRED* reporter, inserted in-frame with the ATG of box1. This enabled us to monitor the effect of deleting box2 on the efficiency of translation of box1. In plants carrying a *dsRED-ENOD40* chimeric construct in which box2 was deleted a higher dsRED expression level

was observed compared to plants carrying this chimeric construct containing box2. Because the level of the dsRED encoding mRNA is not influenced by the presence of the second conserved region, we concluded that the presence of box2 in *cis* has a repressive effect on translation. Because box2 has an effect on the translation of the peptide, this part of the mRNA could be involved in the spatial and temporal regulation of peptide expression. Whether box2 has a regulatory role in translation of box1 remains unclear, because we have no indication that box2 is used to obtain differential levels of ENOD40 peptide. Charon *et al.* (1997) and Sousa *et al.* (2001) showed that introduction of DNA containing either a box1 or a box2 sequence in *Medicago* roots resulted in cortical cell divisions. In these dividing cells the *ENOD12* promoter is activated, suggesting either that both regions have a similar biological function or that box2 is also active in *trans* and regulates in *trans* the translation efficiency of box1

Only a small part of *ENOD40* encompass protein encoding sequences, implying that the major part of the RNA is not translated. Recently, analysis of the RNA structure of *ENOD40* messengers showed that these have markedly more structure (Girard *et al.*, 2003) than RNAs that have a protein encoding function, suggesting that *ENOD40* could also be active at the RNA level. Taken together, box1 encodes a peptide and the protein encoding capacity of box2 might be restricted to some plant species. Based on the high level of conservation of the box2 nucleotide sequence among all *ENOD40* genes, the RNA might be biologically active in addition to the peptide.

Search for *ENOD40* interactors

In order to identify partners of *ENOD40* within the molecular network by which Nod factors trigger cell division in the cortex, several approaches have been followed to identify molecules that interact with the *ENOD40* box1 encoded peptide. In *P. sativum* a putative interacting protein was identified in a two-hybrid screen (Vleghels *et al.*, 2002). This protein, PsP40, has homology to ribosomal associated protein P40 and its expression in pea nodules partly co-localises with that of *ENOD40*. Therefore, PsP40 and PsENOD40 can only interact in the cells where both genes are expressed. In cells devoid of PsP40 expression, PsENOD40 may interact with other proteins.

Affinity purification with a synthesised box1 peptide on soybean nodule extracts led to the identification of sucrose synthase (Röhrig *et al.*, 2002) as a binding protein. The expression of sucrose synthase in soybean nodules partly co-localises with *ENOD40* expression. This shows that a molecular interaction between box1 peptide and sucrose synthase is possible. It has

been shown that ENOD40 peptide could prevent the *in vitro* phosphorylation of certain amino acids in sucrose synthase by a calcium dependent protein kinase (CDPK; Hardin *et al.*, 2003). However, the amount of peptide that is needed for this inhibition is very high. Therefore, the biological significance of *ENOD40* peptide as an inhibitor of phosphorylation of sucrose synthase by CDPK is questionable.

Recently, also a protein, MtRBP1, was identified, of which the translocation from the nucleus is mediated by *MtENOD40* RNA (Campalans *et al.*, 2004). This observation shows that the RNA has the capacity to bind proteins and that it can function as protein translocator. However, the biological significance of this translocation is not known and further it is not known which *ENOD40* sequences are essential for this translocation. In case the conserved regions are essential for binding this could be a common function for *ENOD40* RNAs.

As one of the approaches to identify putative members of the network *ENOD40* is part of, we chose to analyse the genomic context of *ENOD40*. In several species genomic coupling has been shown for genes with a similar expression or that operate in the same functional module. The conservation of gene coupling (colinearity) between divergent species can be an indication for such a functional association. We found that the gene *DN40* is tightly linked to *ENOD40*, in divergent species as *O. sativa* and *M. truncatula* (Chapter 5) However, because the expression of *ENOD40* and *DN40* does not co-localise, it seems unlikely that these genes would operate in the same cellular process. Therefore, it remains to be elucidated whether the conserved linkage within the genome has a function.

Knock-down of *ENOD40-1* in *Z.mays* does not result in phenotypic alterations

Functional analysis of *ENOD40* would be facilitated by mutations in this gene. As such mutants are not available in the model legumes *Medicago truncatula* and *Lotus japonicus*, we chose another plant species that has *ENOD40* genes and for which large mutant collections are accessible.

Z. mays was chosen because of the availability of the mutator transposon system and because it contained *ENOD40* homologs, this in contrast to a model plant as *A. thaliana*. Furthermore, *Z. mays* is able to interact symbiotically with *mycorrhizal* fungi, a symbiotic interaction that like the *Rhizobium*-legume symbiosis is marked by an elevated *ENOD40* expression at the site of symbiosis (Rhijn *et al.*, 1997). A *Z. mays ENOD40* mutant would therefore enable us to study the role of *ENOD40* in the development of this plant as well as its ability to interact with mycorrhizal fungi.

In the EST collection of *Z. mays*, two genes were identified that have the characteristics of *ENOD40* but nonetheless show little homology to each other. We expected that these genes were also functionally divergent and therefore we chose to search for a mutation in *ZmENOD40-1* that has highest homology to legume *ENOD40s*.

By screening a transposon tagged library of *Z. mays* we were able to identify a plant with a transposon insertion in box2 of *ZmENOD40-1* (Chapter 4). Northern analysis revealed that this *Mu*-insertion caused a dramatic reduction in *ENOD40* transcript level. However, plants that were homozygous for *e40-mum1* were still capable to interact with mycorrhizal fungi. Neither could we identify a phenotypic alteration in plant shape or growth that co-segregated with the presence of the *e40-mum1* allele. Our experiments indicate that *ZmENOD40-1* is not essential for plant growth and mycorrhizal interaction in *Z. mays*.

Several studies in legumes indicate that the down regulation of its homolog has profound effects. In *M. sativa* explants, anti sense expression of *MtENOD40* severely impaired callus growth. In *M. truncatula* reduced amounts of transcript, probably due to co-suppression, disturbed nodulation (Charon *et al.*, 1999) as well as arbusculus mycorrhizal formation (Staehelin *et al.*, 2001). These experiments strongly suggest that in legumes, *ENOD40* is essential for plant growth and symbiotic interactions.

One probable reason for the discrepancy between the observations made in legumes and *Z.mays*, could be that *ZmENOD40-2* is functional redundant. Strikingly, *ZmENOD40-2* and *ZmENOD40-1* only share an overall homology of 30%, indicating that even strongly divergent genes could have redundant functions.

Also legumes contain more *ENOD40* genes; two paralogs with different expression pattern were described in *M. sativa* (Fang *et al.*, 1998) and the EST collection of *M. truncatula* contains at least 2 divergent *ENOD40* homologs. The reason that the existence of paralogs does not interfere with functional studies in *Z.mays*, is probably due to the use of transposon tagging to inactivate genes instead of techniques that are based on the formation of siRNA's, by which also slightly homologous genes can be targeted. Therefore studies on *ENOD40* action involving an RNAi approach would be preferable over transposon tagging approaches.

ENOD40 silencing disturbs nodule initiation and development

Previously, it has been shown that a dramatic drop in the level of *ENOD40* transcript as a result of co-suppression, led to a 50 % reduction in the number of nodules on these plants. Furthermore, down regulation of *ENOD40* caused severe morphological aberrations in the root nodules formed. Since the co-suppression phenomenon is triggered during the nodulation process itself (Charon *et al.*, 1999), these studies did not provide information about the role of *ENOD40* in e.g. formation of nodule primordia.

Application of an RNAi approach offers the opportunity to test whether, the induction of *ENOD40* in the pericycle is an essential part of the mechanism by which Nod factors trigger cell division in the cortex. To this end the hairy root system on *M. truncatula* was used to express an *ENOD40* RNA-hairpin that would be degraded into siRNAs and then lead to a knock down of endogenous *ENOD40* expression prior to the inoculation with *S.meliloti* as well as during all steps of nodulation (Chapter 3). Because the experiments were performed in the transgenic lines that contain the *ds*RED coding sequences (Chapter 5) embedded in *ENOD40* mRNA, *ENOD40* silencing coincided with a decrease in *ds*RED fluorescence and thus could be monitored *in vivo*.

Using this strategy we were successful in reducing the expression level of *ENOD40* to 20% of the normal amount in roots and nodules. Although *ENOD40* is normally transcribed in root tissue (Kouchi *et al.*, 1993), silencing of *ENOD40* did not result in visible morphological aberrations in the roots. However, the number of nodules that were formed on the silenced roots was reduced 2-5 fold, suggesting a role for *ENOD40* in nodule initiation. The nodules on silenced roots were mainly spherical in contrast to nodules on control roots that are rod shaped.

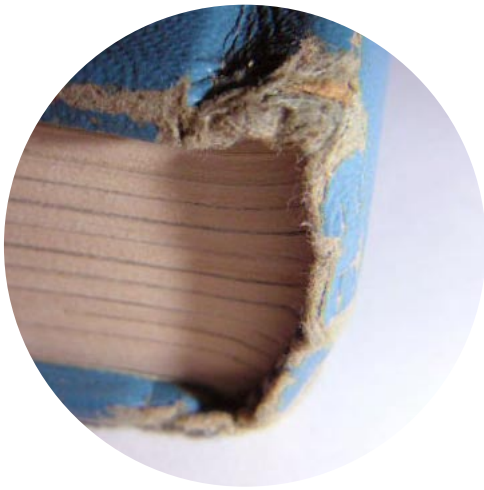
The altered shape of the silenced nodules can be caused by a reduced nodule growth or a premature inhibition of growth. These results confirm earlier findings of Charon *et al.* (1999). However in contrast to these experiments, we still could observe the specific zonation of the central tissue within the silenced nodules. In both experimental systems, the silenced nodules contain infected cells, indicating that differentiation of cells in the central zone is still possible.

Thus, *ENOD40* is involved in nodule initiation and growth although it is not clear whether the expression of the gene is essential for these processes.

Several experiments have indicated a putative involvement of *ENOD40* in plant hormone signalling. Expression studies in tomato indicate that *ENOD40* expression is associated with ethylene action (Vleghels *et al.*, 2003). Ruttink

et al. (2003) showed that *ENOD40* over-expression could affect the response of tobacco BY-2 cells to auxin and cytokinin. These experiments provided evidence that *ENOD40* affects hormonal cross-talk between auxin, cytokinin and ethylene.

To analyse if *ENOD40* has a similar function in nodule initiation and growth, it would be interesting to monitor the concentration of these hormones in *ENOD40* silenced roots.



References

Albrecht C, Geurts R, Lapeyrie F, Bisseling T. (1998) Endomycorrhizae and rhizobial Nod factors both require *Sym8* to induce the expression of the early nodulin genes *PsENOD5* and *PsENOD12A*. *Plant Journal* 15:67-82.

An G, Mitra A, Choi HK, Costa MA, An K, Thornburg RW, Ryan CA. (1989) Functional analysis of the 3' control region of the potato wound-inducible proteinase inhibitor II gene. *Plant Cell*. **1(1):115-22.**

Ané JM, Kiss GB, Riely BK, Penmetsa RV, Oldroyd GE, Ayax C, Levy J, Debelle F, Baek JM, Kalo P, Rosenberg C, Roe BA, Long SR, Dénarié J, Cook DR. (2004) Medicago truncatula DMI1 required for bacterial and fungal symbioses in legumes. *Science*. **27: 1364-7.**

Asad S, Fang Y, Wycoff KL, Hirsch AM. (1994) Isolation and characterisation of cDNA and genomic clones of MsENOD40; transcripts are detected in meristematic cells of alfalfa. *Protoplasma.* **183:10-23.**

Barker DG, Bianchi S, Blondon F, Datté Y, Duc G, Essad S, Flament P, Gallusci P, Génier G, Guy P, Muel X, Tourneur J, Dénarié J, Huguet T. (1990) *Medicago truncatula*, **a model plant for studying the molecular genetics of the Rhizobium-legume symbiosis.** *Plant Mol Biol Rep* **8: 40-49.**

Birnbaum K, Shasha DE, Wang JY, Jung JW, Lambert GM, Galbraith DW, Benfey PN. (2003) A gene expression map of the Arabidopsis root. *Science.* **12;302(5652):1956-60.**

Campalans A, Kondorosi A, Crespi M. (2004) *Enod40*, **a Plant Short ORF-Containing mRNA, Induces Cytoplasmic Localization of a Nuclear RNA-Binding Protein.** *Plant Cell* **16**

Charon C, Johansson C, Kondorosi E, Kondorosi A, Crespi M. (1997) Enod40 induces dedifferentiation and division of root cortical cells in legumes. *Proc. Natl. Acad. Sci. USA* **94: 8901–8906.**

Charon C, Sousa C, Crespi M, Kondorosi A. (1999) Alteration of enod40 expression modifies Medicago truncatula root nodule development induced by Sinorhizobium meliloti. *Plant Cell* **11: 1953-1965.**

Chen M, SanMiguel P, de Oliveira AC, Woo SS, Zhang H, Wing RA, Bennetzen JL. (1997) Microcolinearity in sh2-homologous regions of the maize, rice, and sorghum genomes. *Proc Natl Acad Sci U S A.* **1;94(7):3431-5.**

Chen R, Silver DL, de Bruijn FJ. (1998) Nodule parenchyma-specific expression of the *Sesbania rostrata* **early nodulin gene** *SrEnod2* **is mediated by its 3' untranslated region** *Plant Cell*. **10:1585-602.**

Cheng H-P, Walker GC. (1998) Succinoglycan Is Required for Initiation and Elongation of Infection Threads during Nodulation of Alfalfa by Rhizobium meliloti. *Journal of Bacteriology*. **180(19):5183-5191.**

Chian R-J, Gresshoff PM. (1997) Isolation and primary characterization of genomic enod40 gene from Lotus japonicus cultivar 'Gifu'. *Plant Physiol.* **115: 863.**

Cho RJ, Campbell MJ, Winzeler EA, Steinmetz L, Conway A, Wodicka L, Wolfsberg TG, Gabrielian AE, Landsman D, Lockhart DJ, Davis RW. (1998) A genome-wide transcriptional

analysis of the mitotic cell cycle. *Mol Cell.* **2(1):65-73.**

Chuck G, Meeley RB, Hake S. (1998). The control of maize spikelet meristem fate by the APETALA2-like gene indeterminate spikelet1. *Genes Dev* **12: 1145-54.**

Cohen BA, Mitra RD, Hughes JD, Church GM. (2000) A computational analysis of whole-genome expression data reveals chromosomal domains of gene expression. *Nat Genet.* **26(2):183-6.**

Compaan B, Yang W-C, Bisseling T, Franssen H. (2001) *ENOD40* **expression in the pericycle precedes cortical cell division in Rhizobium-legume interaction and the highly conserved internal region of the gene does not encode a peptide,** *Plant and Soil* **230 1-8.**

Cooper JB, Long SR. (1994) Morphogenetic rescue of *Rhizobium meliloti* **nodulation mutants by** *trans***-zeatin secretion.** *Plant Cell* **6: 215–225.**

Corich V, Goormachtig S, Lievens S, Van Montagu M, Holsters M. (1998) Patterns of ENOD40 gene expression in stem-borne nodules of Sesbania rostrata. *Plant Mol Biol.* **37(1):67-76.**

Crespi MD, Jurkevitch E, Poiret M, d'Aubenton-Carafa Y, Petrovics G, Kondorosi E, Kondorosi A. (1994) Enod40, a gene expressed during nodule organogenesis, codes for a nontranslatable RNA involved in plant growth. *EMBO J.* **13: 5090–5112.**

Csanadi G, Szecsi J, Kalo P, Kiss P, Endre G, Kondorosi A, Kondorosi E, Kiss GB. (1994) *ENOD12***, an early nodulin gene, is not required for nodule formation and efficient nitrogen fixation in alfalfa.** *Plant Cell* **6: 201-13.**

Dandekar T, Snel B, Huynen M, Bork P. (1998) Conservation of gene order: a fingerprint of proteins that physically interact. *Trends Biochem Sci.* **23(9):324-8.**

Davis SJ ,Vierstra D. (1998) Soluble, highly fluorescent variants of green fluorescent protein (GFP) for use in higher plants. *Plant Mol. Biol.* **36: 521–528.**

Denarie J, Cullimore J. (1993) Lipo-Oligosaccharide nodulation factors: a minireview. New class of signaling molecules mediating recognition and morphogenesis. *Cell* **74: 951–954.**

Devos KM, Beales J, Nagamura Y, Sasaki T. (1999) Arabidopsis-rice: will colinearity allow gene prediction across the eudicot-monocot divide? *Genome Res.* **9(9):825-9.**

Dietrich RA, Radke SE, Harada JJ. (1992) Downstream DNA sequences are required to activate a gene expressed in the root cortex of embryos and seedlings. *Plant Cell***. 4(11):1371-82.**

Endre G, Kereszt A, Kevei Z, Mihacea S, Kalo P, Kiss GB. (2002) A receptor kinase gene regulating symbiotic nodule development. *Nature* **27: 962-6.**

Fåhreus G. (1957) The infection of clover root hairs by nodule bacteria studied by a simple glass technique. *J General Microbiol* **16:374-381.**

Fang Y, Hirsch AM. (1998) Studying early nodulin gene *ENOD40* **expression and induction by nodulation factor and cytokinin in transgenic alfalfa.** *Plant Physiol.* **116:53-68.**

Favery B, Complainville A, Vinardell JM, Lecomte P, Vaubert D, Mergaert P, Kondorosi A, Kondorosi E, Crespi M, Abad P. (2002) The endosymbiosis-induced genes *ENOD40* **and** *CCS52a* **are involved in endoparasitic-nematode interactions in** *Medicago truncatula***.** *Mol Plant Microbe Interact***. 15:1008-13.**

Flemetakis E., Kavroulakis N., Quaedvlieg N.E.M., Spaink H.P., Dimou M., Roussis A., Katinakis P. (2000). *Lotus japonicus* **contains two distinct ENOD40 genes that are expressed in symbiotic, nonsymbiotic, and embryonic tissues.** *Molecular Plant Microbe Interactions* **13: 987-994.**

Gadella Jr. TWJ, Vereb G, Röhrig H, Schmidt J, John M, Schell J, Bisseling T. (1997) Microspectroscopic imaging of nodulation factor-binding sites on living *Vicia sativa* **roots using a novel bioactive fluorescent nodulation factor.** *Biophys. J.* **72: 1986–1996.**

Geurts R, Franssen H. (1996) Signal transduction in Rhizpbiuminduced nodule formation. *Plant Physiol.* **112: 447–453.**

Girard G, Roussis A, Gultyaev AP, Pleij CW, Spaink HP. (2003) Structural motifs in the RNA encoded by the early nodulation gene enod40 of soybean. *Nucleic Acids Res.* **1;31(17):5003-15.**

Gu Z., Steinmetz L.M., Gu X., Scharfe C., Davis R.W., Li W.H. (2003). Role of duplicate genes in genetic robustness against null mutations. Nature 421: 63-6.

Handberg K, Stougaard J. (1992) *Lotus japonicus,* **an autogamous diploid legume species for classical and molecular genetics.** *Plant Journal* **2: 5273-5277.**

Hardin SC, Tang GQ, Scholz A, Holtgraewe D, Winter H, Huber SC. (2003) Phosphorylation of sucrose synthase at serine 170: occurrence and possible role as a signal for proteolysis. *Plant J.* **35: 588-603.**

Hirsch AM, Bhuvaneswari TV, Torrey JG, Bisseling T. (1989) Early nodulin genes are induced in alfalfa outgrowths elicited by auxin transport inhibitors. *Proc. Natl. Acad. Sci. USA.* **86: 1244–1248.**

Hirsch AM, Fang Y, Asad S, Kapulnik Y (1997) The role of phytohormones in plant-microbe symbioses. *Plant Soil* **194: 171–184.**

Hirsch AM. (1992) Developmental biology of legume nodulation. *New Phytol.* **122: 211–237.**

John M, Schmidt J,Walden R, Czaja I, Dülz M, Schell J, Röhrig H. (1997) Lipochitooligosaccharide-induced tobacco cells release a peptide as mediator of the glycolipid signal. *Proc. Natl. Acad. Sci. USA.* **94: 10178–10182.**

Jouanin L, Bouchez D, Drong RF, Tepfer D, Slightom JL. (1989) Analysis of TR-DNA / plant junctions in the genome of a Convolvulus arvensis clone transformed by Agrobacterium rhizogenes strain A4. *Plant Mol. Biol.* **12: 75-85.**

Kosugi S, Ohasi Y, Nakajima K, Arai Y (1990) An improved assay for beta-glucuronidase in transformed cells: Methanol alsmost completely suppresses a putative endogenous beta-glucuronidase activity. *Plant Science* **70:133-140.**

Kouchi H, Hata S (1993) Isolation and characterization of novel nodulin cDNAs representing genes expressed at early stages of soybean nodule development. *Mol. Gen. Genet.* **238: 106–119.**

Kouchi H, Takane K, So RB, Ladha JK, Reddy PM. (1999) Rice ENOD40: isolation and expression analysis in rice and transgenic soybean root nodules. *Plant J.* **18(2):121-9.**

Kozik A., Matvienko M., Scheres B., Paruvangada V.G., Bisseling T., van Kammen A., Ellis T.H., LaRue T., Weeden N. (1996). The pea early nodulin gene PsENOD7 maps in the

region of linkage group I containing sym2 and leghaemoglobin. *Plant Molecular Biology* **31: 149-56.**

Krusell L, Madsen LH, Sato S, Aubert G, Genua A, Szczyglowski K, Duc G, Kaneko T, Tabata S, de Bruijn F, Pajuelo E, Sandal N, Stougaard J. (2002) Shoot control of root development and nodulation is mediated by a receptor-like kinase. *Nature* **28: 422-6.**

Larsen K. (2003) Molecular cloning and characterisation of a cDNA encoding a ryegrass (*Lolium perenne*) ENOD40 homologue. *J. Plant Physiology* **160: 675-687.**

Lerouge P, Roche P, Faucher C, Maillet F, Truchet G, Promé JC, Dénarié J. (1990) Symbiotic host-specificity of Rhizobium meliloti is determined by a sulphated and acylated glucosamine oligosaccharide signal. *Nature* **344: 781-784.**

Lévy J, Bres C, Geurts R, Chalhoub B, Kulikova O, Duc G, Journet EP, Ane JM, Lauber E, Bisseling T, Dénarié J, Rosenberg C, Debelle F. (2004) A putative Ca2+ and calmodulin-dependent protein kinase required for bacterial and fungal symbioses. *Science* **27: 1361-4.**

Limpens E, Franken C, Smit P, Willemse J, Bisseling T, Geurts R. (2003) LysM domain receptor kinases regulating rhizobial Nod factor-induced infection. *Science*. **24:630-3.**

Liu H, Sachidanandam R, Stein L. (2001) Comparative genomics between rice and Arabidopsis shows scant collinearity in gene order. *Genome Res.* **11(12):2020-6.**

Long SR. (1989) *Rhizobium*-legume nodulation: life together in the underground. Cell 56: 203-214.

Madsen EB, Madsen LH, Radutoiu S, Olbryt M, Rakwalska M, Szczyglowski K, Sato S, Kaneko T, Tabata S, Sandal N, Stougaard J. (2003) A receptor kinase gene of the LysM type is involved in legume perception of rhizobial signals. *Nature* **9: 637-40.**

Mathesius U, Schlaman HRM, Spaink HP, Sautter C, Rolfe BG, Djordjevic MA (1998) Auxin transport inhibition precedes root nodule formation in white clover roots and is regulated by flavonoids and derivates of chitin oligosaccharides. *The Plant Journal* **14: 23–34.**

Matvienko M, van de Sande K, Yang W-C, van Kammen A, Bisseling T, Franssen H. (1994) Comparison of soybean and pea ENOD40 cDNA clones representing genes expressed during both early and late stages of nodule development. *Plant Mol.Biol.* **26: 487–493.**

Mayer K, Murphy G, Tarchini R, Wambutt R, Volckaert G, Pohl T, Dusterhoft A, Stiekema W, Entian KD, Terryn N, Lemcke K, Haase D, Hall CR, van Dodeweerd AM, Tingey SV, Mewes HW, Bevan MW, Bancroft I. (2001) Conservation of microstructure between a sequenced region of the genome of rice and multiple segments of the genome of Arabidopsis thaliana. *Genome Res.* **(2001) 11(7):1167-74.**

McCarty D.R., Carson C.B., Stinard P.S., Robertson D.S. (1989). Molecular Analysis of viviparous-1: An Abscisic Acid-Insensitive Mutant of Maize. *Plant Cell* **1: 523-532.**

McLaughlin M., Walbot V. (1987). Cloning of a mutable bz2 allele of maize by transposon tagging and differential hybridization. *Genetics* **117: 771-6.**

Meeley R.B., Briggs S.P. (1995). Reverse genetics for maize. *Maize Genetics Cooperation Newsletter* **69: 67-82.**

Minami E, Kouchi H, Cohn J, Ogawa T, Stacey G. (1996) Expression of the early nodulin, ENOD40, in soybean roots in response to various lipo-chitin signal molecules. *The Plant*

Journal. **10: 23-32.**

Mylona P, Pawlowski K, Bisseling T. (1995) Symbiotic nitrogen fixation. *Plant Cell* **7: 869–885.**

Nam Y-W, Penmetsa RV, Endre G, Uribe P, Kim D, Cook DR. (1999) Construction of a bacterial artificial chromosome library of *Medicago truncatula* **and identification of clones containing ethylene-response genes.** *Theor Appl Genet.* **98:638-646.**

O'Reilly C., Shepherd N.S., Pereira A., Schwarz-Sommer Z., Bertram I., Robertson D.S., Peterson P.A., Saedler H. (1985). Molecular cloning of the a1 locus of Zea mays using the transposable elements En and Mu1. *Embo Journal* **4: 877-882.**

Page R.D. (1996) TreeView: an application to display phylogenetic trees on personal computers. *Comput Appl Biosci* **12: 357-8.**

Papadopoulou K., Roussis A., Katinakis P. (1996) Phaseolus ENOD40 is involved in symbiotic and non-symbiotic organogenetic processes: Expression during nodule and lateral root development. *Plant Molecular Biology* **30: 403-417.**

Pawlowski K, Kunze R, de Vries S, Bisseling T. (1994) *in: Plant Mol biol Manual* **(Gelvin, S.B. and Schilderpoort, R.A., eds.), pp. 1-13, Kluwer Academic Publishers, Dordrecht.**

Radutoiu S, Madsen LH, Madsen EB, Felle HH, Umehara Y, Gronlund M, Sato S, Nakamura Y, Tabata S, Sandal N, Stougaard J. (2003) Plant recognition of symbiotic bacteria requires two LysM receptor-like kinases. *Nature* **9: 585-92.**

Raizada MN, Benito MI, Walbot V. (2001). The MuDR transposon terminal inverted repeat contains a complex plant promoter directing distinct somatic and germinal programs. *Plant J* **25: 79-91.**

Rohrig H, Schmidt J, Miklashevichs E, Schell J, John M. (2002) Soybean *ENOD40* **encodes two peptides that bind to sucrose synthase.** *Proc Natl Acad Sci U S A.***19;99(4):1915-20.**

Ruttink T. (2003) *ENOD40* **affects phytohormone cross-talk.** *thesis*, **Wageningen University, The Netherlands.**

Ryan CA, Pearce G. (2003) Systemins: a functionally defined family of peptide signals that regulate defensive genes in Solanaceae species. *Proc Natl Acad Sci U S A.* **25: 100.**

Sabatini S, Beis D, Wolkenfelt H, Murfett J, Guilfoyle T, Malamy J, Benfey P, Leyser O, Bechtold N, Weisbeek P, Scheres B. (1999) An auxin-dependent distal organizer of pattern and polarity in the Arabidopsis root. *Cell*. **24: 463-72.**

Sambrook J, Fritsch EF, Maniatis T. (1989), Molecular cloning: A Laboratory Manual 3th ed, Cold Spring Harbor, New York.

Schauser L, Roussis A, Stiller J, Stougaard J. (1999) A plant regulator controlling development of symbiotic root nodules. *Nature* **11: 191-5.**

Scheres B, Van Engelen F, Van der Knaap E, Van der Wiel C, Van Kammen A, Bisseling T (1990) Sequential induction of nodulin gene expression in developing pea nodule. *Plant Cell* **2: 687–700.**

Sinvany G., Kapulnik Y., Wininger S., Badani H., Jurkevitch E. (2002). The early nodulin enod40 is induced by, and also promotes arbuscular mycorrhizal root colonization. *Physiological and Molecular Plant Pathology* **60: 103-109.**

Sonti RV, Chiurazzi M, Wong D, Davies CS, Harlow GR, Mount DW, Signer ER. (1995) Arabidopsis mutants deficient in T-DNA integration. *Proc Natl Acad Sci U S A.* **92(25):11786-90.**

Sousa C, Johansson C, Charon C, Manyani H, Sautter C, Kondorosi A, Crespi M. (2001) Translational and structural requirements of the early nodulin gene enod40, a short-open reading frame-containing RNA, for elicitation of a cell-specific growth response in the alfalfa root cortex. *Mol Cell Biol.* **21(1):354-66.**

Spaink HP. (1995) The molecular basis of infection and nodulation by rhizobia: The ins and outs of sympathogenesis. *Annu. Rev. Phytopathol.* **33: 345–368.**

Spellman PT, Rubin GM. (2002) Evidence for large domains of similarly expressed genes in the Drosophila genome. *J Biol.* **1(1):5.**

Staehelin C., Charon C., Boller T., Crespi M., Kondorosi A. (2001). Medicago truncatula plants overexpressing the early nodulin gene enod40 exhibit accelerated mycorrhizal colonization and enhanced formation of arbuscules. *Proceedings of the National Academy of Sciences of the United States of America* **98:15366-15371.**

Sundaresan V, Springer P, Volpe T, Haward S, Jones JDG, Dean C, Ma H and Martienssen R. (1995) Patterns of gene action in plant development revealed by enhancer trap and gene trap transposable elements. *Genes and Dev.* **9: 1797–1810.**

Tarchini R, Biddle P, Wineland R, Tingey S, Rafalski A. (2000) The complete sequence of 340 kb of DNA around the rice Adh1-adh2 region reveals interrupted colinearity with maize chromosome 4. *Plant Cell.* **12(3):381-91.**

Taylor CB (1997) Promoter Fusion Analysis: An insufficient measure of gene expression. *The Plant Cell* **9: 273–275.**

Thompson J.D., Higgins D.G., Gibson T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* **22: 4673-80.**

Tomkins J.P., Frisch D.A., Wing R.A. (2000). Construction and characterization of a Maize BAC library for the inbred line B73. *Maize Genetics Cooperation Newsletter* **74: 18-19.**

Trinh TH, Ratet P, Kondorosi E, Durand P, Kamaté K, Bauer P, Kondorosi A. (1998) Rapid and efficient transformation of diploid *Medicago truncatula* **and** *Medicago sativa* **ssp.** *falcata* **lines improved in somatic embryogenesis.** *Plant Cell Reports* **17:345-355.**

Ueda HR, Matsumoto A, Kawamura M, Iino M, Tanimura T, Hashimoto S. (2002) Genome-wide transcriptional orchestration of circadian rhythms in Drosophila. *J Biol Chem.* **19;277(16):14048-52.**

Van Bokhoven H, Verver J, Wellink J, Van Kammen A. (1993) Protoplasts transiently expressing the 200K coding sequence of Cowpea mosaic virus B-RNA support replication of M-RNA. *J.Gen. Virol.* **74: 2233–2241.**

Van de Sande K, Pawlowski K, Czaja I, Wieneke U, Schell J, Schmidt J, Walden R, Matvienko M, Wellink J, van Kammen A, Franssen H, Bisseling T. (1996) Modification of phytohormone response by a peptide encoded by ENOD40 of legumes and a nonlegume. *Science* **273: 370–373.**

Van de Wiel C, Scheres B, Franssen H, Van Lierop MJ, Van Lammeren A, Bisseling T. (1990) The early nodulin transcript ENOD2 is located in the nodule parenchyma (inner

cortex) of pea and soybean root nodules. *EMBO J.* **9: 1–7.**

van Dodeweerd AM, Hall CR, Bent EG, Johnson SJ, Bevan MW, Bancroft I. (1999) Identification and analysis of homoeologous segments of the genomes of rice and Arabidopsis thaliana.*Genome* **42(5):887-92.**

van Engelen, F.A., Molthoff, J.W., Conner, A.J., Nap, J.P., Pereira, A., Stiekema, W.J., (1995). pBINPLUS, an improved plant transformation vector based on pBIN19. *Transgenic Res.* **4: 288–290.**

van Kammen A. (1984) Suggested nomenclature for plant genes involved in nodulation and symbiosis. *Plant Mol Biol Rep* **2: 43-45.**

van Rhijn P., Fang Y., Galili S., Shaul O., Atzmon N., Wininger S., Eshed Y., Lum M., Li Y., To V.V., Fujishige N., Kapulnik Y., Hirsch A.M. (1997) Expression of early nodulin genes in alfalfa mycorrhizae indicates that signal transduction pathways used in forming arbuscular mycorrhizae and Rhizobium-induced nodules may be conserved. *Proceedings of the National Academy of Sciences of the United States of America* **94: 5467-5472.**

Vandepoele K, Saeys Y, Simillion C, Raes J, Van De Peer Y. (2002) The automatic detection of homologous regions (ADHoRe) and its application to microcolinearity between Arabidopsis and rice. *Genome Res.* **12(11):1792-801.**

Varkonyi Gasic E., White D.W.R. (2002). The white clover enod40 gene family. Expression patterns of two types of genes indicate a role in vascular function. *Plant Physiology* **129: 1107-1118.**

Vasse J, de Billy F, Camut S, Truchet G. (1990) Correlation between ultrastructural differentiation of bacteroids and nitrogen fixation in alfalfa nodules. *J Bacteriol.* **172: 4295-4306.**

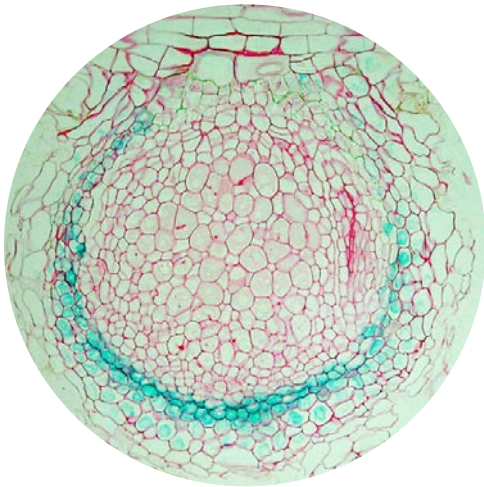
Vijn I, Martinez-Abarca F, Yang W-C, des Neves L, Brussel A, van Kammen A, Bisseling T. (1995) Early nodulin gene expression during Nod factor-induced processes in *Vicia sativa*. *The Plant Journal* **8: 111-119.**

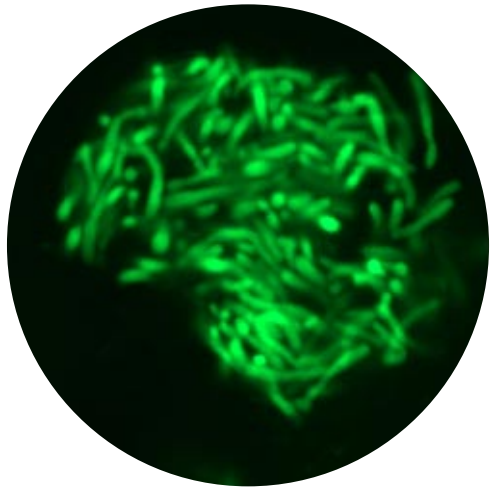
Vleghels I, Hontelez J, Ribeiro A, Fransz P, Bisseling T, Franssen H. (2003) Expression of ENOD40 during tomato plant development. *Planta.* **218(1):42-9.**

Walker JM. (1986) Protein (Western) blotting with immunodetection. In *Experiments in Molecular Biology.* **Ed. R J Slagter. pp 171–178. Humana Press, New Jersey, USA.**

Yang H, Matsubayashi Y, Nakamura K, Sakagami Y. (2001) Diversity of Arabidopsis genes encoding precursors for phytosulfokine, a peptide growth factor. *Plant Physiol.* **127: 842-51.**

Yang W-C, Katinakis P, Hendriks P, Smolders A, De Vries F, Spee J, Van Kammen A, Bisseling T, Franssen H (1993) Characterization of GmENOD40, a gene showing novel patterns of cell-specific expression during soybean nodule development. *The Plant Journal* **3: 573–585.**





Samenvatting

Vlinderbloemige planten hebben een bijzondere manier om in grond die weinig gebonden stikstof bevat toch hiervan voorzien te worden. In een speciaal orgaan, de wortelknol, kunnen ze bepaalde bodem bacteriën huisvesten, dit zijn de zogenaamde *Rhizobium* bacteriën. Deze rhizobia zijn op hun beurt in staat stikstof uit de lucht te reduceren tot ammonium, dat in deze vorm gebruikt kan worden door de plant. Dit is een erg goedkope en milieuvriendelijke vorm van bemesting en vanuit dit oogpunt de moeite waard om te onderzoeken, daarnaast kan onderzoek aan deze knolletjes waardevolle informatie opleveren over orgaanvorming in het algemeen.

Het proces van knolvorming begint wanneer de rhizobia door middel van chemotaxis naar de wortels gaan. Rhizobia moeten daartoe eerst flavonoiden herkennen die door de plant uitgescheiden worden. Als reactie scheiden de rhizobia signaalstoffen uit die herkent worden door de plant. Wortelharen reageren het eerst door te gaan krullen en vormen op die manier een zakje waarin de rhizobia opgesloten raken. Van hieruit dringen de bacteriën de wortelhaar binnen via een buisje, de infectiedraad. Deze draden groeien naar de wortelcortex. Ondertussen zijn sommige cellen in de wortelcortex gaan delen en vormen een knol-primordium. Wanneer de rhizobia dit primordium bereiken, dringen ze de cellen binnen, waarna het primordium verder kan uitgroeien en differentiëren tot een wortelknol.

Met behulp van genetische technieken zijn verschillende Rhizobium genen geïdentificeerd, die betrokken zijn bij het proces van knolvorming. Zo zijn er nu genen (dit is er vaak maar één) bekend van rhizobia die betrokken zijn bij de herkenning van signaalstoffen van de plant. Verder zijn er genen geïdentificeerd die verantwoordelijk zijn voor de vorming van signaalstoffen (Nod-factoren) die rhizobia maken om door te plant herkend te worden.

Voor het vinden van plantengenen die bij dit proces zijn betrokken zijn twee verschillende benaderingen gebruikt. Enerzijds zijn er mutanten gemaakt die verstoord zijn in de vorming van wortelknollen en een aantal hiervan zijn nu gekloneerd maar daarnaast is ook gekeken naar genproducten die in de knol in hogere concentratie aanwezig zijn dan elders.

Hoewel mogelijk verwacht kon worden dat beide benaderingen tot dezelfde resultaten zouden moeten leiden, bleek dit allerminst het geval: geen van de genen die geïnduceerd worden tijdens knolvorming (nodulines), zijn teruggevonden in de genetische analyses. Om de rol van deze nodulines te bestuderen en om te kijken of ze wel essentieel zijn tijdens knolvorming, hebben wij er in deze studie één geselecteerd, *ENOD40*. Dit gen komt in een zeer vroeg stadium

van het hele knolvormingsproces tot expressie en codeert bovendien voor een peptide, hetgeen weinig voorkomt in planten.

Om de expressie van het gen precies te kunnen bepalen, hebben we in hoofdstuk 2 de expressie van *ENOD40* in de tijd gevolgd. Hieruit blijkt dat het gen actief is zelfs voordat het primordium gevormd wordt. Het lijkt er dus op dat *ENOD40* betrokken kan zijn bij het doorgeven van het signaal dat tot celdeling leidt. Een duidelijke aanwijzing hiervoor is ook dat het kunstmatig tot expressie brengen van *ENOD40* in de wortel extra celdeling geeft. Omdat de expressie van *ENOD40* al geïnduceerd wordt door het toedienen van gezuiverde Nod-factoren (de bacteriële signaalstoffen), lijkt het mogelijk dat *ENOD40* onderdeel is van het mechanisme waardoor Nod-factoren celdeling in de wortelcortex kunnen induceren.

ENOD40 is ook interessant omdat het codeert voor een peptide. In de meeste gevallen bevat het boodschapper RNA van een gen een zogeheten open leesraam dat de code bevat voor een eiwit. In het transcript van *ENOD40* zijn echter geen lange open leesramen te vinden, zodat er geen grote eiwitten van afgelezen kunnen worden. Wanneer *ENOD40* genen van verschillende planten worden vergeleken, blijkt dat slechts twee kleine delen van dit gen geconserveerd zijn. Deze hebben we box1 en box2 genoemd. Door chimaere constructen te maken met een gen dat codeert voor een groen fluorescerend eiwit (GFP), hebben we laten zien dat box1 wel en box2 niet vertaald wordt in een eiwit. Ook andere experimenten hebben laten zien dat box1 vertaald wordt.

Eerder werd al gesuggereerd dat box2 dan mogelijk een functie heeft in het reguleren van de vertaling van box1. Om deze hypothese te testen, hebben we twee typen transgene planten geconstrueerd. In één type is een artificieel *ENOD40* gen gebracht waarin een rood fluorescerend eiwit (RFP) is gefuseerd met het box1-peptide en box2 verwijderd is. In het andere type is in het artificiële *ENOD40* gen alleen het RFP eiwit gefuseerd met het box1-peptide. Vervolgens hebben we gekeken of het wel of niet aanwezig zijn van box2, invloed had op de hoeveelheid RFP die gevormd werd in de plant. Het bleek dat het verwijderen van box2 een toename gaf van box1-RFP in de plant. Omdat de hoeveelheid transcript niet veranderde door deze deletie, komen we tot de conclusie dat box2 een represserend effect heeft op de translatie van box1. Deze regulerende werking van box2 is opnieuw een aanwijzing dat box1 inderdaad codeert voor een peptide.

Om te kijken wat de functie van *ENOD40* is, zijn verschillende onderzoekers op zoek gegaan naar moleculen waarmee *ENOD40* in contact komt in de plantencel. Met behulp van een verscheidenheid aan technieken zijn al een aantal van deze moleculen geïdentificeerd. Zo is er in erwtenknollen een gen

gevonden waarvan het product in een two-hybrid screen een interactie laat zien met het peptide. Verder is er in soya aangetoond dat sucrose synthase kan binden aan het peptide en dat deze binding phophorilatie van sucrose synthase kan verhinderen. Daarnaast is in *Medicago truncatula* een eiwit gevonden dat mogelijk bindt aan het transcript van *ENOD40*.

Wij beschrijven in hoofdstuk 5 een andere benadering om een mogelijke partner van *ENOD40* te vinden. Door de genomen van uiteenlopende planten die een ENOD40 gen bevatten te vergelijken vonden we dat het gen *DN40* altijd naast *ENOD40* ligt. Een mogelijke reden voor zo'n koppeling kan een functionele interactie zijn. Echter, omdat de expressie van beide genen niet overeenkomt, lijkt het niet logisch dat ze een interactie met elkaar kunnen aangaan. Of de koppeling veroorzaakt wordt door een functionele interactie is dus niet duidelijk.

Er zijn al diverse studies verricht naar de functie van *ENOD40* tijdens knolvorming maar ook in andere processen. Artificiële expressie van *ENOD40* in plantenwortels leidt tot celdelingen en de expressie van andere noduline genen, bovendien versnelde zo'n expressie het proces van knolvorming. Verder werd aangetoond dat uitschakeling van het gen het knolvormingsproces juist verhindert. Met soortgelijke technieken werd ook in andere processen dan knolvorming een activiteit van *ENOD40* waargenomen. Zo lijkt *ENOD40* betrokken bij de regulatie van zijscheut vorming in planten, hormoon interacties in tabakscellen, Mycorrhiza symbiose en ook bij de vorming van het embryo.

Een verdere functionele karakterisering van het gen zou zeer gebaat zijn met de beschikbaarheid van *ENOD40* mutanten. Door gebruik te maken van het 'mutator transposon system', konden we zo'n *ENOD40* mutant identificeren in maïs. Natuurlijk had een mutant in een vlinderbloemige de voorkeur, maar omdat dergelijke 'transposon' systemen niet beschikbaar waren voor vlinderbloemigen, kozen we voor maïs. Hoewel maïs geen knollen maakt, kan het wel een symbiose aangaan met een bodemschimmel, Mycorrhiza. Bovendien konden we het effect van het uitschakelen van *ENOD40* op algehele plantontwikkeling bestuderen. Hoewel de expressie van het gen nagenoeg afwezig was in de mutant, ontwikkelde de plant zich normaal en kon deze ook nog steeds een symbiose aangaan met Mycorrhiza. Een mogelijke oorzaak hiervan is dat de functie wordt overgenomen door een tweede *ENOD40* gen een andere oorzaak kan zijn dat er wel een effect is maar dat wij dat niet gevonden hebben.

Om ook de functie van *ENOD40* in vlinderbloemigen te kunnen bestuderen, hebben we gebruik gemaakt van het RNAi systeem om de expressie van *ENOD40* in *Medicago truncatula* uit te schakelen. Met deze techniek konden we de expressie van het gen tot ongeveer 20% van de normale intensiteit terugbrengen. Hoewel *ENOD40* normaal ook in wortels tot expressie komt,

leidt een reductie in *ENOD40* expressie niet tot afwijkingen in wortelgroei structuur. Echter, het aantal knolletjes dat gevormd werd op de wortel nadat deze was geïnoculeerd met rhizobia, was ongeveer 2 tot 5 keer minder in vergelijking met wortels waar de expressie van *ENOD40* niet was gereduceerd. Bovendien waren de knollen die gemaakt werden anders van vorm, normaal maakt *Medicago* langwerpige knollen, terwijl deze knolletjes juist meer bolvormig waren en veel kleiner. Toch konden we in deze knollen nog wel de specifieke weefsel organisatie van knollen herkennen. Dit duidt erop dat *ENOD40* betrokken is bij het initiëren van knollen en de groei ervan.

Het onderzoek dat hier beschreven is en de experimentele systemen die hiervoor zijn ontwikkeld, zijn een goed uitgangspunt om meer te weten te komen over de functie van *ENOD40* tijdens wortelknolvorming en ook in andere processen.

Curriculum Vitae

Egbert Compaan werd geboren op 5 April 1975 te Ruinerwold. Na het behalen van het VWO diploma aan het Gomarus College in Groningen, startte hij in 1993 zijn studie Moleculaire Wetenschappen aan de Landbouw Universiteit Wageningen. Afstudeerprojecten werden uitgevoerd bij de vakgroep Moleculaire Biologie van de Landbouw Universiteit Wageningen en bij het Hubrecht Laboratorium in Utrecht. In 1998 begon hij met een promotie onderzoek aan bij de vakgroep Moleculaire Biologie van Wageningen Universiteit, dit boekje is daar een resultaat van.

Sinds 2003 werkt hij als onderzoeker zaadfysiolgie bij het veredelingsbedrijf ENZA zaden te Enkhuizen.

Dankwoord

En dan nu, helemaal aan het eind, de gelegenheid om mensen die betrokken waren bij het werk van de afgelopen jaren te bedanken, ere wie ere toekomt! Heel veel mensen hebben meegedacht, interesse getoond en meegeholpen, ik ben daar heel erg blij mee. Vooral als h*et al*lemaal even niet mee zit, besef je pas goed hoe belangrijk dit is.

Allereerst Henk, bedankt! Je was een geweldige begeleider. Wat heb ik veel van jou geleerd. Al in 1996 ben ik bij je begonnen als student en in de volgende jaren hebben we veel afgeboomd over *ENOD40* en vele andere zaken. Ik waardeer je inzet en die enorme hoeveelheid aan ideeën die telkens weer opborrelt uit je brein. Veel succes met het verdere onderzoek aan *ENOD40*.

Ton Bisseling, hartelijk bedankt voor je hulp bij het meedenken en het verwoorden van alle bevindingen.

En natuurlijk Tom en Ingrid, waarde paranimfen, wat hadden een goede tijd gehad daar, in onze riante kamer met uitzicht op het arboretum. Ik denk met erg veel plezier terug aan alle goede momenten, vooral ook natuurlijk aan die supervakantie in Canada, laten we elkaar vooral niet uit het oog verliezen. Ook Jan Ho. en Xi, medeleden van 'de club van Henk', bedankt voor de goede tijd.

En dan de studenten die zich hebben ingezet voor dit onderzoek: Evert-Jan, Pieter, Martijn, Wilma en Rays bedankt voor jullie hulp! Verder zijn er vele collega's geweest die me op velerlei manieren hebben geholpen en ondersteund. Gerard jou wil ik nog speciaal noemen, beter een goeie lab-buur dan een verre vriend. Tony, bedankt voor alle sequenties. Bert, je groene vingers waren onmisbaar in de kas. Marnix, bedankt voor de tijd en creativiteit die je beschikbaar stelde voor het lay-outen van dit boekje.

Herma, voor jou deze speciale plaats! De tijd dat ik thuis zat te werken, was jij m'n collega en dat beviel uitstekend. Ik ben blij dat we samen door het leven kunnen.

Design and lay-out by Marnix Bras, Wageningen, www.marnixbras.nl
Printed by ProPress, Wageningen, www.propress.nl

This research described in this thesis was carried out at the Labatory of Molecular Biology, Wageningen University, The Netherlands. This work was supported by a research grant from the Netherlands Organisation for Scientific Research (NWO 805-18-284).