The EMBO Journal vol.13 no.21 pp.5099–5112, 1994

enod40, a gene expressed during nodule organogenesis, codes for a non-translatable RNA involved in plant growth

Martin D.Crespi[1,5], Edouard Jurkevitch[1,5], Maryse Poiret[1], Yves d'Aubenton-Carafa[2], György Petrovics[1,3], Eva Kondorosi[1] and Adam Kondorosi[1,3,4]

[1]Institut des Sciences Végétales, Centre National de la Recherche Scientifique, F-91198 Gif-sur-Yvette Cedex, [2]Centre de Génetique Moléculaire, Centre National de la Recherche Scientifique, F-91198 Gif-sur-Yvette Cedex, France and [3]Institute of Genetics, Biological Research Center, Hungarian Academy of Sciences, Szeged PO Box 521, H-6701 Hungary
[4]Corresponding author
[5]Both authors have contributed equally to this work

Communicated by A.Kondorosi

Rhizobium meliloti **can interact symbiotically with** ***Medicago*** **plants, thereby inducing root nodules. However, certain** ***Medicago*** **plants can form nodules spontaneously, in the absence of rhizobia. A differential screening was performed using spontaneous nodule versus root cDNAs from** ***Medicago sativa*** **ssp.** ***varia*****. Transcripts of a differentially expressed clone,** ***Msenod40*****, were detected in all differentiating cells of nodule primordia and spontaneous nodules, but were absent in fully differentiated cells.** ***Msenod40*** **showed homology to a soybean early nodulin gene,** ***Gmenod40*****, although no significant open reading frame (ORF) or coding capacity was found in the** ***Medicago*** **sequence. Furthermore, in the sequences of cDNAs and a genomic clone (*Mtenod40*) isolated from** ***Medicago truncatula*****, a species containing a unique copy of this gene, no ORFs were found either.** ***In vitro*** **translation of purified** ***Mtenod40*** **transcripts did not reveal any protein product. Evaluation of the RNA secondary structure indicated that both** ***Msenod40*** **and** ***Gmenod40*** **transcripts showed a high degree of stability, a property shared with known non-coding RNAs. The** ***Mtenod40*** **RNA was localized in the cytoplasm of cells in the nodule primordium. Infection with** ***Agrobacterium tumefaciens*** **strains bearing antisense constructs of** ***Mtenod40*** **arrested callus growth of** ***Medicago*** **explants, while overexpressing** ***Mtenod40*** **embryos developed into teratomas. These data suggest that the** ***enod40*** **genes might have a role in plant development, acting as 'riboregulators', a novel class of untranslated RNAs associated with growth control and differentiation.**

Key words: differentiation-related RNA/early nodulin/*Medicago*/plant riboregulator/spontaneous nodules

Introduction

Soil bacteria of the genera *Rhizobium*, *Azorhizobium* and *Bradyrhizobium* can induce the formation of a new plant organ, the root nodule, on their leguminous host plants. In the nodule, nitrogen fixation occurs, rendering the plant independent of soil nitrogen. In a first step, the bacteria attach to and enter the root hair cells while the plant forms a tube-like infection thread through which the bacteria move into the root cortex. Simultaneously, cortical cell divisions occur and branches of the infection threads invade these dividing cells forming the nodule primordium. Growth and development of this primordium leads to the formation of the mature nitrogen-fixing nodule (Hirsch, 1992, and references therein).

This process is achieved by means of a complex series of developmental changes in both partners, involving the spatially and temporally regulated expression of specific genes (Nap and Bisseling, 1990). The bacterial genes involved in nodulation (*nod* genes) code for the synthesis of specific morphogenetic signals, the Nod factors, able to trigger nodule morphogenesis (Lerouge *et al.*, 1990; Spaink *et al.*, 1991; Truchet *et al.*, 1991). The Nod factor NodRmIV(S) of *Rhizobium meliloti* (Lerouge *et al.*, 1990) is a tetrasaccharide of β-1,4-D-glucosamine acylated at the non-reducing end with an unsaturated C16 fatty acid chain and a sulphate group in the reducing sugar residue. Similar molecules found in other rhizobia suggest that variations in the structure of this signal molecule could account for host specificity (Spaink, 1992). In contrast to the advanced progress in the study of bacterial genes, little is known about the responses in the plant counterpart. A few cDNAs, corresponding to plant genes expressed in the early stages of nodule development, have been characterized. It has been suggested that the encoded proteins ('early nodulins') are involved in the infection process and/or nodule morphogenesis (Franssen *et al.*, 1992). Certain *Medicago* plants form nodules spontaneously in the absence of rhizobia (NAR^+ phenotype; Truchet *et al.*, 1989), indicating that these plants can initiate the nodule organogenesis programme independent of the presence of nodulation signals. Early nodulin genes expressed in spontaneous nodules are, therefore, associated exclusively with this developmental programme.

There is growing evidence that certain RNA molecules have defined cellular roles per se (Cech, 1993). RNAs acting as ribozymes (e.g. nuclear introns) have been identified (Cech, 1993) and highly purified rRNA is able to sustain enzymatic reactions implicated in the translation process (Noller *et al.*, 1992). Even RNA motifs able to bind ATP have been obtained by '*in vitro*' evolution (Sassanfar and Szostak, 1993). Moreover, untranslated RNAs have recently been found to act as tumour suppressors in mammalian cells (Hao *et al.*, 1993; Rastinejad *et al.*, 1993). These RNAs, involved in growth control, promote differentiation and the term 'riboregulator' has therefore been proposed for them (Rastinejad *et al.*, 1993).

In this paper, we report that *Msenod40*, a *Medicago*

sativa homologue of a recently identified early nodulin gene from *Glycine max*, *Gmenod40* (Kouchi and Hata, 1993; Yang *et al.*, 1993), is expressed in spontaneous nodules and in differentiating cells of the nodule primordia, showing its association with the nodule organogenesis developmental programme. Surprisingly, this gene does not contain any significant ORF or coding capacity. Corresponding genomic and cDNA clones, designated *Mtenod40*, were identified from *Medicago truncatula*, a species which contains a single copy of this gene. Sequence analysis and evaluation of structural properties of the *enod40* transcripts in soybean and alfalfa suggest that these genes code for untranslated RNAs. Moreover, alteration of *Mtenod40* expression was shown to affect plant morphogenesis in transgenic *Medicago* plants. These data allow us to propose that *enod40* genes code for a novel class of RNAs involved in plant growth control and differentiation.

Results

Msenod40 is expressed in spontaneous nodules and in Nod factor-treated roots

Growth under continuous lack of combined nitrogen of certain *Medicago* plants in the absence of rhizobia results in the appearance of non-functional spontaneous nodules (NAR^+ phenotype; Truchet *et al.*, 1989). We therefore reasoned that RNA preparations of this tissue could serve as a good source of material to identify genes associated with nodule organ formation. To detect transcripts common to spontaneous and young growing nodules which are not expressed in roots, 200 000 phages of a cDNA library from young nodules of *M.sativa* ssp. *varia* were differentially screened using spontaneous nodule versus root cDNAs. One differentially expressed clone revealed a 700 bp transcript whose tissue specificity was analysed. Using Northern analysis (Figure 1A), the transcript was only found in nodules and was detected neither in other plant tissues nor in roots inoculated with a non-nodulating mutant of *R.meliloti* (ZB138; Kondorosi *et al.*, 1984). During nodule development, the highest transcript level was detected in young 7-day-old nodules induced by wild-type *R.meliloti*. In nodules induced by an Exo^- *R.meliloti* mutant where bacterial infection was highly reduced, as well as in spontaneous nodules, the transcript levels were similar to those of the young nodules, induced by wild-type bacteria. These data indicate that the expression of this gene is associated with nodule organogenesis and is independent of the infection process.

The sequence of this cDNA showed significant homology (see below) to a recently identified early nodulin from *G.max* (soybean), *Gmenod40* (Kouchi and Hata, 1993; Yang *et al.*, 1993), and the gene was therefore named *Msenod40*. The expression of *Msenod40* was further studied using reverse transcription-polymerase chain reaction (RT-PCR). With this technique, a low level of expression (after 25 cycles) was detected in roots, stems, flowers (Figure 1B) and in dividing cell suspensions of *Medicago* (data not shown).

To determine the appearance of the transcript during nodule development, pieces of alfalfa roots containing agar inocula of *R.meliloti* (see Materials and methods) were collected at different time points and used in RT-PCR assays (for 15 cycles). The level of the *Msenod40* transcript was enhanced at 2 days post-inoculation (d.p.i.), the earliest time point tested, and increased up to 5 d.p.i. (Figure 1C). Under this condition of inoculation, small white nodules became visible at 4–5 d.p.i. and detection of leghaemoglobin started at 6–7 d.p.i. (data not shown). To exclude genomic DNA contamination in the RNA sample, PCR assays were performed using oligos corresponding to a flavonoid-3-hydroxylase gene spanning an intron (B.Charrier *et al.*, in preparation). No genomic DNA contamination was detected in any cDNA samples (data not shown).

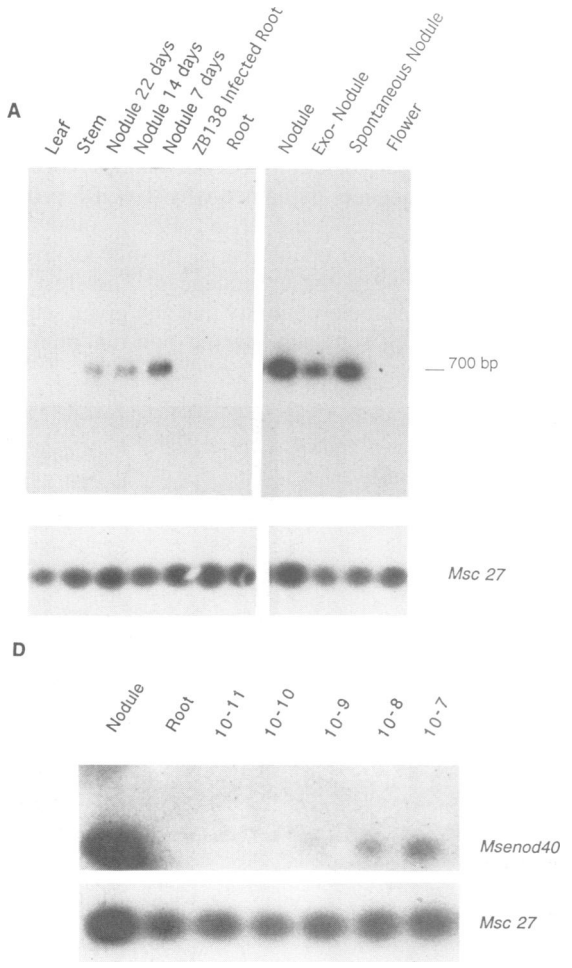
Induction of the expression of the *enod40* gene by Nod factor was tested 2 days after treatment of alfalfa roots with the Nod factor NodRm(IV,S) (Lerouge *et al.*, 1990) at different concentrations. The transcript level of the gene was induced at 10^{-9} M and significantly enhanced at 10^{-8} and 10^{-7} M final concentrations (Figure 1D). After treatment with the higher concentrations of Nod factor, several groups of dividing cortical cells were detected in the roots, as reported previously (Truchet *et al.*, 1991; Schultze *et al.*, 1992). These data indicate that the appearance of Nod factor-induced nodule primordia is correlated with *Msenod40* expression.

Msenod40 is expressed in differentiating cells of the nodule primordium

In the determinate soybean and in the indeterminate pea nodules, the early expression of the *enod40* gene was localized in the cells of the primordium and in the pericycle of the infected region of the roots (Kouchi and Hata, 1993; Yang *et al.*, 1993). We analysed whether the spatial distribution of the *Msenod40* RNA was similar in *Medicago*. Roots were agar inoculated with *Rhizobium* (see Materials and methods) and cell division in the inner cortex started at 1–2 d.p.i. By using *in situ* hybridization to an antisense *Msenod40* RNA probe, *Msenod40* expression was detected in the dividing cortical cells (Figure 2A and B). At 4 d.p.i., *Msenod40* transcripts were detected in all differentiating cells of the growing nodule primordium from the cortex and the pericycle (note hybridization around a xylem pole from the vascular tissue Pc; Figure 2C), as previously reported for pea nodule initiation (Yang *et al.*, 1993).

The developing nodule primordium cells contain intensively stained nuclei and a dense cytoplasm. Once invaded, they enlarge and differentiate into cells of the symbiotic zone (zone III) where transcripts for *Msenod40* were no longer detected (Figure 2D and E). As shown in Figure 2F and G, in the mature nodule (20 d.p.i.) the *enod40* RNA was detected in the broad zone II (Vasse *et al.*, 1990), but its expression above background levels was not detected in the fully differentiated cells of zone III. Observation at higher magnification showed that the gene was not expressed in the apical meristem (zone I) nor in the interzone II–III containing amyloplasts (Figure 2H). The *enod40* transcript was detectable, however, in the region around the nodule vascular tissue (Figure 2N and O, see below), as reported for the determinate soybean nodule (Kouchi and Hata, 1993; Yang *et al.*, 1993).

In spontaneous nodules, a group of undifferentiated cells resembling those of the distal part of the nodule primordium give rise to derivative cells that differentiate into a central zone in the region adjacent to the root stele



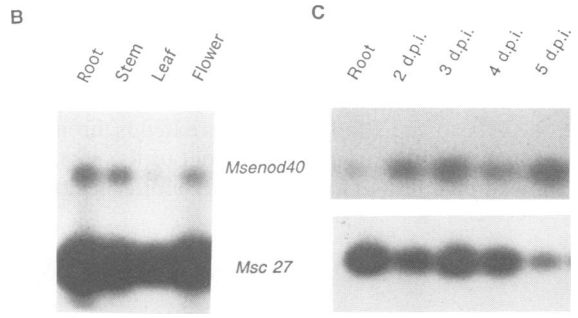


Fig. 1. Analysis of *Msenod40* expression. (**A**) Northern analysis. Left panel: total RNAs (10 μg) of leaves, stems, roots and nodules of different ages were subjected to electrophoresis on a 1% agarose gel, blotted and hybridized with labelled *Msenod40* cDNA. RNA from roots treated for 5 days with non-nodulating bacteria (lane ZB138) was also assayed. Right panel: Similarly prepared and analysed samples of Rm41-infected (7 days old), exo$^-$ and spontaneous nodules and flowers. Spontaneous and exo$^-$ nodules were obtained after 3 weeks of nitrogen deprivation, and after 2 weeks of infection with the mutant strain, respectively. The amount of RNA loaded was checked by probing with *Msc27*. (**B**) RNAs of roots, stems, leaves and flowers were assayed using RT-PCR (25 cycles). (**C**) Kinetic analysis (2–5 d.p.i.) of spot-inoculated roots using the RT-PCR assay. (**D**) Roots treated for 2 days with different concentrations of Nod factors. PCR was allowed to proceed for 15 cycles to analyse the induction of the gene in inoculated or Nod factor-treated root samples (C and D). In (B), (C) and (D), the integrity of each RNA preparation and the amount used in each reaction were checked by co-amplification with *Msc27* specific oligonucleotides, followed by probing with an *Msc27*-labelled probe.

(Joshi *et al.*, 1991; Truchet *et al.*, 1989). *Msenod40* mRNAs were found in the region containing differentiating cells (region D; Figure 2I and J), but no transcripts were detected either in the central type of cells containing amyloplasts or in the nodule cortex (region C), as shown at higher magnification in Figure 2M. Moreover, the first apical cell layers of these nodules did not contain *enod40* transcripts (Figure 2K and L), indicating that the gene is not expressed in the distal layers of the apical meristem, but very early after the dividing cells entered differentiation. Sense RNA probes did not yield hybridization signals above background levels (data not shown).

These results indicate that the *Msenod40* transcript is present in differentiating cells, but neither in the apical meristematic nor in the terminally differentiated cells of infected or uninfected nodules.

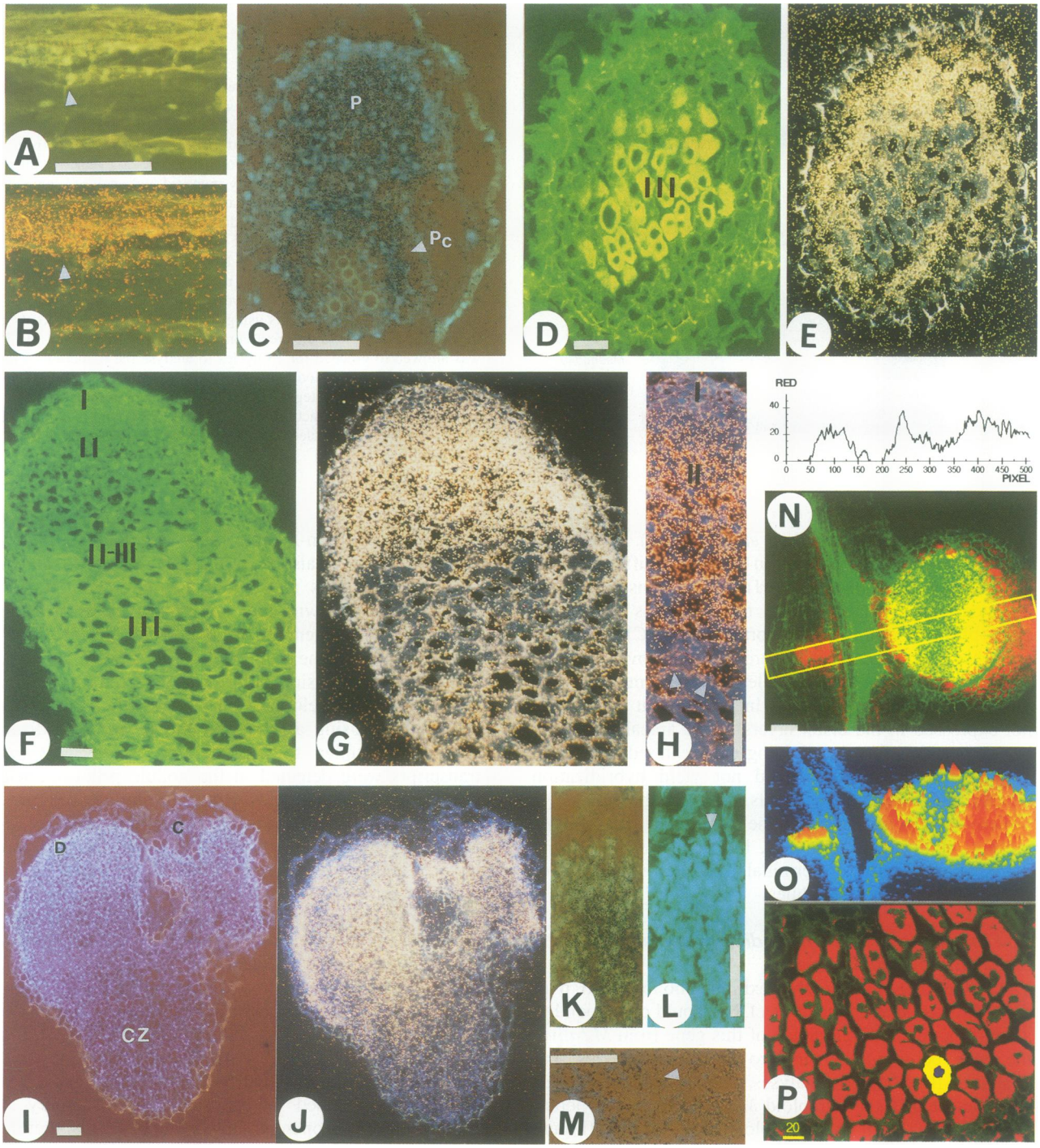
The enod40 genes from Medicago species do not show any coding capacity

The isolated *Msenod40* cDNA exhibited no long ORF over the entire sequence. Since this cDNA could correspond to a non-functional mutant allele of this gene from *M.sativa*, the gene copy number in two *Medicago* species was analysed. By Southern analyses, three hybridizing bands were detected in the tetraploid *M.sativa* species, whereas only one was seen in the diploid *M.truncatula* (Figure 3). Under low stringency, a cross-hybridizing band was detected in the non-legume plant, *Arabidopsis thaliana*. A single band was also seen in digestions of *M.truncatula* DNA with various restriction enzymes (*Bam*HI, *Bgl*II, *Xho*I; data not shown), indicating that there is probably a single gene in the genome of this plant. Therefore, cDNAs and a genomic clone from *M.truncatula* were isolated. A similar gene expression pattern was detected for *Mtenod40* during nodule development in *M.truncatula* plants (data not shown). Using a whole-mount *in situ* hybridization procedure coupled to confocal microscopy, *Mtenod40* transcripts were detected in the nodule primordium and in the zone II of mature *M.truncatula* nodules (Figure 2O and P), confirming the results obtained with *Msenod40*. Moreover, expression of the gene in the nodule parenchyma around the vascular tissue of the mature nodule was also detected. The amount of red fluorescence corresponding to the probe can be quantified in this section, as shown in Figure 2P. Sense RNA probes did not show any significant fluorescent signal (data not shown).

The sequence of *Mtenod40*, the longest cDNA clone from *M.truncatula*, is shown in Figure 4A. All translation possibilities are indicated and the longest ORF (27 amino acids) is underlined. The initiation of translation for this small peptide requires that five start AUGs and 15 stop codons be bypassed by the ribosome machinery. The translation of this open reading frame (ORF) is, therefore, highly improbable. *Mtenod40* showed 99% homology with

the *M.sativa* cDNA (Figure 4B), a striking degree of conservation for non-coding sequences, even for two species of the same genus. Polyadenylated tails have been found for the transcripts of both species at almost the same position (1 bp difference).

The sequence of the *M.truncatula* genomic clone was found to be identical to the corresponding cDNA (Figure 4B), ruling out cloning artifacts due to reverse transcription or other post-transcriptional modification. Conservation between soybean and alfalfa sequences extends throughout these sequences (Figure 4B). Alignment of the sequences also showed that although they share a similar length, the *Medicago* clones are extended by 50 bp at the 5′ end while, in the 3′ region, they are shorter than the *G.max* clones by 100 bp. In addition, a 75 bp insertion was detected in the *Medicago* sequences (nucleotides 415–480, from the numbering of *Mtenod40*). This insertion and other base pair deletions between the soybean and alfalfa cDNA sequences explained why the ORF proposed for the soybean gene (Yang *et al.*, 1993) could not be found in its *Medicago* counterparts, in spite of the high conservation between these sequences. The Test Code program calculates the probability that any given sequence could be a coding sequence by the non-randomness of



every third base in a codon (lack of 'wobble'). Analysis of the *enod40* sequences (Figure 5A, B and C) showed that none of them possess significant coding features, while this program revealed the appropriate ORF in other early nodulin genes (such as *Msenod12B*; Figure 5D).

Purified transcripts of *Mtenod40* cDNAs were assayed for *in vitro* translation, with reticulocyte and wheat germ extracts, but no translation products were detected (data not shown), while other nodulins have been translated (Scheres *et al.*, 1990a; our unpublished results). Codon usage analysis of *Medicago* coding sequences did not reveal any particular alteration of the genetic code (data not shown). These data together suggest that the RNA product of the *Mtenod40* gene is not translated.

enod40 gene transcripts possess low free energies of folding

The observation that the *enod40* gene might be a developmentally regulated RNA prompted us to search for properties which would further support this hypothesis. RNA molecules folded by multiple interactions adopt structures whose stability is given mainly by Watson–Crick base pairing. This stability can be estimated by computation, but the calculated free energy of folding for a given RNA sequence depends mainly on its length and its GC content. This free energy of formation also reflects the capacity of the different nucleotides of a sequence to interact among each other. Therefore, the likely existence of a relevant structure can be proposed if this free energy value is significantly lower than that obtained by folding the same sequence after randomization.

A computation of 100 randomization events (keeping constant the total numbers of A, T, C and G nucleotides), as well as calculation of free energies and analysis of their distribution, was carried out for several coding and non-coding sequences (Table I). Based on the obtained Gaussian distribution of free energies, the average free energy (AVG) and standard deviation (σ) were determined for each sequence. The value of free energy of folding (FLD) of the actual sequence could then be compared to this average value. The relative difference was calculated by the formula $(AVG - FLD)/\sigma$ which therefore allows the comparison between different RNAs independently of the length and nucleotide composition of the sequences. All coding RNAs analysed showed a value of free energy of folding close to or higher than the average value and, consequently, their relative differences are zero or positive values (Table I). Ten additional coding sequences were tested (from soybean and alfalfa) and also showed positive values for the relative difference (data not shown). In contrast, all non-coding RNAs tested showed values well

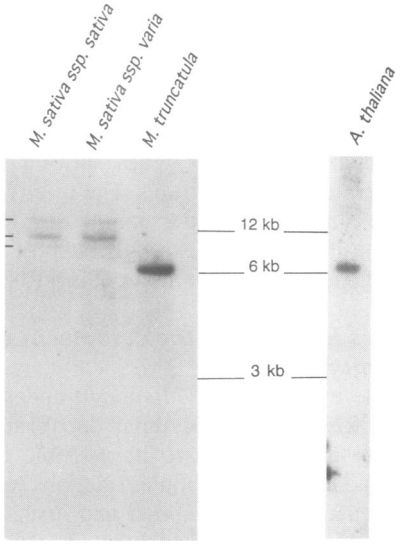


Fig. 3. Detection of *enod40* in tetraploid and diploid *Medicago* species and in *Arabidopsis thaliana* by Southern analysis. *Medicago sativa* ssp. *sativa*, *M.sativa* ssp. *varia* (tetraploids), *M.truncatula* (diploid) and *A.thaliana* DNAs (10 μg) were digested with *Eco*RI and hybridized to a *Msenod40* ^{32}P-labelled probe. The filter containing *A.thaliana* DNA was hybridized under low stringency.

Fig. 2. Localization of *enod40* expression during nodule development in *Medicago* plants by *in situ* hybridization. *Medicago* roots were inoculated with *R.meliloti* using the agar block technique (see Materials and methods) to assay for expression of *Msenod40* during nodule development. (**A**) Longitudinal section of a root segment 18 h d.p.i. under UV epifluorescence. Initiation of a nodule primordium is seen in the inner cortex (arrow). (**B**) Same section as shown in (A) observed by dark field microscopy. Hybridization signals (silver grains) appear as yellow dots. Transcripts are detected in the dividing cortical cells of the nodule primordium. (**C**) Transverse section of a growing nodule primordium at 4 d.p.i. Hybridization (dark grains) is seen in all cells of the developing nodule primordium (P) and in the pericycle cells around the vascular tissue (Pc arrow). (**D**) Transverse section of a nodule 6 d.p.i. with differentiating nodule zones under UV epifluorescence. III: Differentiated central zone. (**E**) Same as (D), observed by dark field microscopy. Expression of *Msenod40* is no longer detectable in the differentiated symbiotic cells (III) or in the nodule cortex (C). (**F**) Longitudinal section of a mature nodule (14 d.p.i.) showing the characteristic nodule zones: I, meristem; II, invasion zone; interzone II–III and zone III or the nitrogen-fixing zone (Vasse *et al.*, 1990). (**G**) *Msenod40* transcripts are seen in zone II. No signals above background are detected in the symbiotic zone III. (**H**) High magnification of the section shown in (G) indicates that no expression is detected in the meristem (zone I) nor in the amyloplast-containing interzone II–III. The arrow points to amyloplasts. Spontaneous nodules were obtained by nitrogen starvation of *M.sativa* ssp. *varia* plants (see Materials and methods) and collected after 3 weeks. (**I**) Longitudinal section of a multilobed spontaneous nodule photographed under UV microscopy. CZ: central zone (amyloplast-containing cells; Truchet *et al.*, 1989). D: differentiating cells. (**J**) Same section as in (I), observed by dark field microscopy. Transcripts are localized in the differentiating cells of the spontaneous nodule, but not in the central zone. (**K**) Detail of a longitudinal section showing that the two apical cell layers of a spontaneous nodule do not express the *Msenod40* transcript (arrow). (**L**) Same as (K), under UV microscopy, allowing for localization of *enod40* expression in the cells showing active nuclei. (**M**) Detail of longitudinal section observed using normal microscopy. Note that the amyloplast (arrows)-containing cells had fewer dark grains. All sections were hybridized with an antisense *Msenod40* probe. Sense RNA probes on similar samples did not yield any positive signal (data not shown). Bars: 40 μm. (**N**) Whole-mount *in situ* hybridization on a sample showing two *M.truncatula* nodules at different developmental stages. The presence of *enod40* transcripts (red fluorescence) is detected in the emerging nodule primordium (left) and in the invasion zone of a mature nodule (right). Note *Mtenod40* expression in the periphery of the mature nodule. Green fluorescence shows the tissue structure. The sample was observed with a confocal microscope (see Materials and methods). Below: graph of the intensity of red fluorescence along the indicated yellow rectangle. (**O**) The intensity of red fluorescence is depicted on a third axis on the same section as N to visualize the distribution of *Mtenod40* transcripts. (**P**) Optical section of a nodule primordium hybridized with a *Mtenod40* antisense probe. Red fluorescence intensities, indicating the presence of the transcript, were integrated separately in 45 cells (from nine sections) for the nuclear zone and the whole cell. These measurements are illustrated for one cell (area for a nuclear zone: blue; for a whole cell: yellow). Bar: 20 μm.

below the average, indicating a greater stability of the actual structure compared to all the other possibilities for sequences with the same base composition. Therefore, their relative differences were all negative (the highest negative value found was −1.27 for the 5S RNA). The soybean and alfalfa *enod40* genes showed significantly low free energies ($P < 3\times10^{-3}$). The relative difference values were similar to all the non-coding RNAs tested (Table I). These correlations strongly support the idea that the *enod40* gene product is a non-coding RNA.

Mtenod40 is localized in the cytoplasm of the nodule primordia cells

Based on the evidence that *Mtenod40* codes for a non-translatable RNA, the subcellular localization of this molecule can give hints as to its possible cellular role (Cech, 1993). Different subcellular fractions were analysed for the presence of the *Mtenod40* and *Msc27* transcripts by RT-PCR. *Msc27* is a polyadenylated coding transcript (Pay *et al.*, 1992). The two transcripts co-migrated during fractionation and were detected in the pelleted cytoplasm (probably corresponding to membrane microsomes) and in the cytoplasmic fractions (Figure 6; 20 000 *g* pellet, cytoplasm). No *Mtenod40* transcript was detected in the RNA corresponding to the nuclear fraction. In addition, whole-mount *in situ* hybridizations were performed on young nodule primordia and a fluorescent substrate was used to detect the presence of the transcript (Figure 2O and P). The use of confocal microscopy allowed the optical dissection of a nodule primordium and the integration of the amount of red fluorescence in the cytoplasm and the nuclei of these primordium cells (Figure 2Q). For 45 cells from nine sections, the nucleus contained on average 4.7 ± 1.2% of the total cellular fluorescent signal; the highest single value was 7.8%, the remaining 95% being in the cytoplasm. It therefore appears that *Mtenod40* encodes an RNA accumulating in the cytoplasm.

Alteration of Mtenod40 expression affects plant growth and differentiation

The possible role of *enod40* was assessed by altering its expression using full-length cDNA sense and antisense constructs cloned under the control of a strong constitutive 35S promoter. *Agrobacterium tumefaciens* LBA4404 strains bearing (i) the binary vector pCP60 without insert, (ii) the antisense and (iii) the overexpressing sense *Mtenod40* constructs were used to infect *M.sativa* ssp. *varia* A2 explants. A summary of all the transformation experiments is given in Table II. The efficiency of transformation and regeneration was 12% for plants containing the control vector pCP60 without insert. Using the *A.tumefaciens* strain carrying antisense *enod40* constructs, 500 *Medicago* explants were infected (in several independent experiments); however, no embryos were formed and no

A

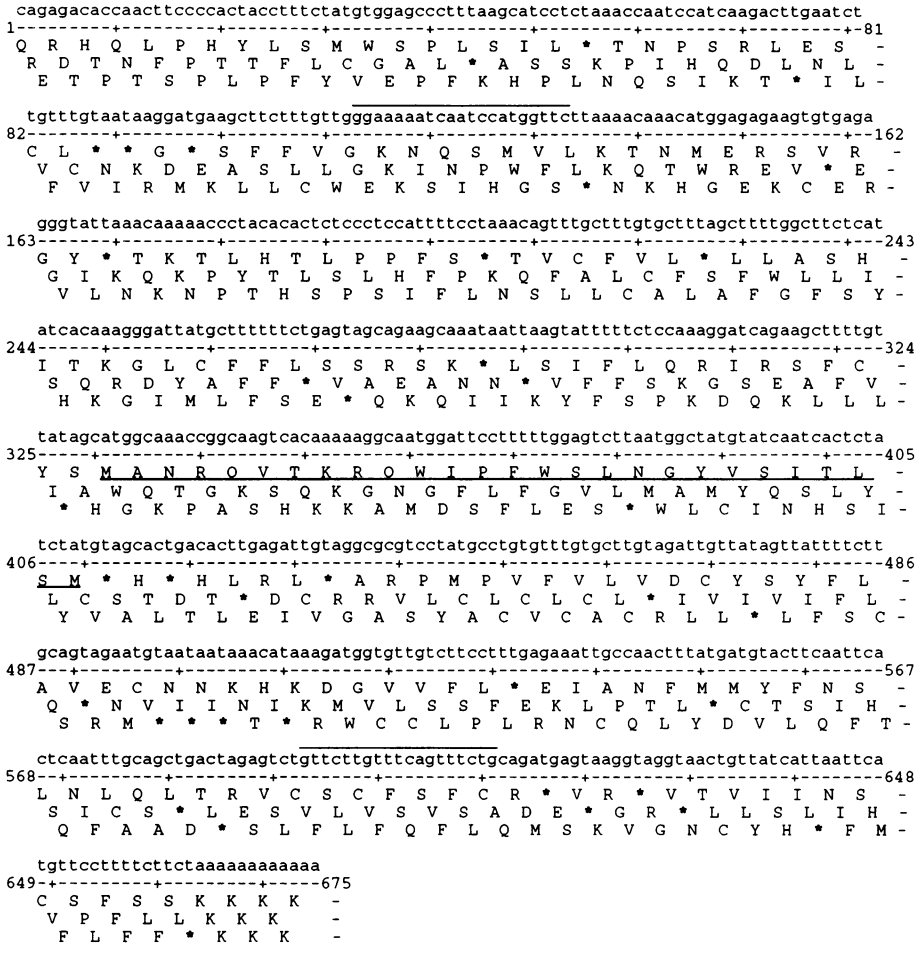


Fig. 4. Sequence analysis of *enod40* cDNAs and genomic clones from *Medicago* species. (**A**) Nucleotide sequence, putative ORFs and stop codons of the longest *M.truncatula* cDNA isolated, *Mtenod40*. The longest ORF is underlined. A poly(A) tail was detected on the 3′ end of the transcript. Oligonucleotides used in RT-PCR assays are overlined. (**B**) Alignment of the nucleotide sequences of the identified *enod40* genes. Gaps (points) were inserted to optimize the alignment (PILEUP program). Identical nucleotides are boxed. *Mtenod40* is the cDNA depicted in (A). *Mtenod40*gen is the genomic sequence corresponding to this cDNA. *Msenod40* is the sequence of the cDNA from *M.sativa*. The soybean sequences were obtained from GenBank: *Gmenod401* (accession number X69154), *Gmenod402* (X69155), *soya36A* (D13503) and *soya36B* (D13504).

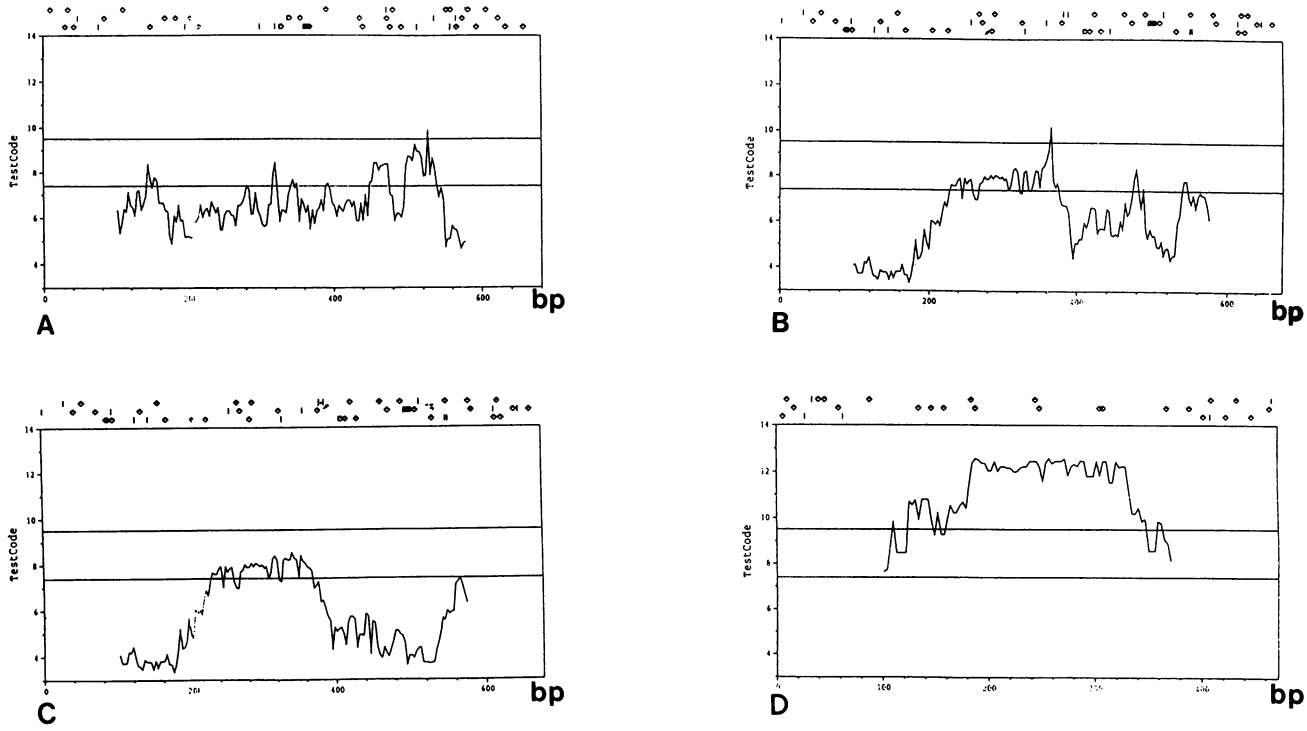


Fig. 5. *enod40* genes have no coding capacity. Evaluation of the coding capacity of *enod40* genes was performed using the Test Code program (GCG software package) for *G.max*, *Gmenod401* (accession number X69154) (**A**), *M.truncatula* (**B**) and *M.sativa* (**C**) cDNA clones. No coding regions were predicted in these sequences. The *Msenod12B* genomic sequence (Allison *et al.*, 1993) clearly indicated a coding sequence (**D**). The top region (above the upper line) predicts coding regions to a 95% level of confidence. Potential start codons are indicated by short vertical lines, stop codons by small diamonds.

Table I. Free energies of randomized and actual RNA sequences

	Average	Sigma	Fold value	Difference[a]	Relative difference ($n\sigma$)[b]
Coding RNAs					
Leghaemoglobin (559)[c]	−116.20	5.29	−116.30	−0.10	−0.02
Cold acclimation protein (1001)	−190.30	5.97	−178.40	+12.10	+2.02
Stress inducible protein (569)	−129.15	5.63	−127.00	+2.10	+0.37
Glutamine synthase (276)	−60.68	5.05	−52.20	+8.48	+1.67
Initiation factor (741)	−176.10	6.57	−175.40	+0.70	+0.11
Non-coding RNAs					
Rybozyme (517)	−123.72	4.70	−137.00	−13.28	−2.83
RNaseP RNA (402)	−128.41	5.20	−149.50	−21.09	−4.05
5S RNA (120)	−37.20	3.53	−41.70	−4.50	−1.27
3′ UTR tropomyosin (397)	−75.10	5.02	−81.80	−6.70	−1.33
Viroid RNA (358)	−113.13	5.73	−148.50	−35.37	−6.17
enod40 RNAs					
Gmenod40[d] (617)	−146.84	5.94	−169.3	−22.46	−3.78
Mtenod40[e] (662)	−140.95	5.79	−160.6	−19.65	−3.39

[a]Difference: Average (AVG)−Fold (Fld).
[b]Relative difference (Rel. Dif.): (AVG−FLD)/σ. $n\sigma$ = number of sigmas.
[c]Numbers in parentheses indicate the number of base pairs for each sequence.
[d]This sequence corresponds to *Gmenod401* (X69154).
[e]This is the cDNA sequence depicted in Figure 4.

plants could be regenerated (Table II). In contrast to the typically embryogenic calli obtained after infection with the control *A.tumefaciens* strain containing the pCP60 vector (as illustrated for one explant in Figure 7A), these explants showed inhibited callus formation and did not develop embryos under kanamycin selection (Figure 7B), even after 5 months. Explants from the same transformation were kept without antibiotic selection and gave rise to normal embryogenic calli, probably derived from untransformed cells (Figure 7C), ruling out the possibility of any pathogenic effect associated with infection by this strain. On the other hand, explants infected with the *Agrobacterium* strain harbouring the sense construct developed embryos earlier and with higher frequency

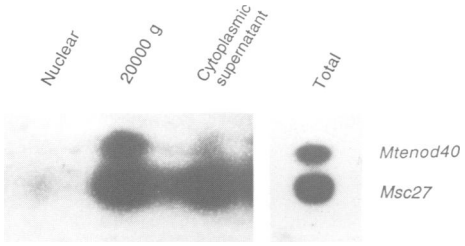


Fig. 6. Subcellular localization of *Mtenod40* mRNA. Cell fractionation of *M.truncatula* young nodules was done as described in Materials and methods. Total RNAs of the different subcellular fractions (nuclear, 20 000 *g* pellet, and cytoplasmic supernatant) and of unfractionated young nodules (total) were used in RT-PCR assays to detect *Mtenod40* and *Msc27* transcripts.

Table II. Effects of sense and antisense *Mtenod40* constructs on embryogenesis and regeneration of *Medicago sativa* A2 explants

Number of explants	Construct	Number of independent regenerated embryos	Growth characteristics
100	control[a]	12[b]	normal plants
500	antisense[c]	0	inhibited callus growth
100	sense[d]	25	teratomas

[a]Explants (pieces of roots, stems and leaves) were infected with an *A.tumefaciens* disarmed strain LBA4404 containing the control vector pCP60 without insert, and selected on kanamycin-containing media. Proliferating shoots were obtained after 1 month and plated on kanamycin medium without hormones for plant regeneration.
[b]Each group of independent embryos gave rise to proliferating shoots and yielded one (or more) Kan^R plant. It was counted as one independent regenerated embryo, although after propagation on rooting medium several identical transgenic plants were obtained.
[c]Idem as in (c), but the full-length *Mtenod40* sequence was inserted in the antisense orientation. No green embryo was observed.
[d]Embryos as in (a), but using a sense construct containing a full-length *Mtenod40* cDNA sequence inserted into pCP60 in the sense orientation.

(25% of the explants; Table II) than the controls. In order to obtain plants ectopically overexpressing the *Mtenod40* gene, the embryos were transferred to regenerating medium for 1 week and then to rooting medium. Instead of developing whole plants, however, polyembryogenic calli containing both undifferentiatied tissue and developing embryos appeared (Figure 7D). Longer cultivation resulted in plant teratomas with limited root development and calli containing developing shoots (Figure 7E and F, right). In contrast, control transformations with the binary vector pCP60 (without insert) resulted in fully developed plants after 1 month in the rooting media. Two representative control plants are depicted in the left part of Figure 7F.

The teratoma tumour tissues were tested for the presence of *enod40* transcripts (Figure 8). Using Northern analysis, a high level of *enod40* expression was detected in four independently developed teratomas, compared to a control transgenic plant. In the latter case, *enod40* transcripts were only detectable by the RT-PCR technique (Figure 1B). RNA loading was quantified by hybridization to a *Msc27* probe.

These data indicate that the antisense transcripts of *Mtenod40* severely impaired callus growth, probably by blocking the action of *enod40* in *Medicago* explants, while the ectopic overexpression of *Mtenod40* induced uncontrolled proliferation of the transgenic *Medicago* tissue, leading to the formation of teratomas. These data suggest that the *enod40* RNA plays a role in plant growth.

Discussion

In this paper, we have provided evidence that *enod40*, a gene expressed in differentiating cells of the nodule primordium, codes for a novel type of non-translatable RNA involved in plant growth control.

Several results indicate that *enod40* gene transcripts are associated with the developmental programme of nodule organogenesis. After infection of roots with *Rhizobium*, *enod40* transcripts were detectable from the onset of nodule primordium formation (Kouchi and Hata, 1993; Yang *et al.*, 1993; our results), well before the bacteria reached the inner cortex. These experiments did not, however, clearly differentiate between events related to the activation of the genetic programme of nodule organogenesis and those related to the infection of plants by rhizobia. Although the gene is expressed in nodules formed by exo^- mutants of *Bradyrhizobium japonicum* (Yang *et al.*, 1993) and *R.meliloti*, where infection processes are highly reduced, the present analysis of axenic spontaneous nodules provides evidence that *enod40* expression is indeed uncoupled from bacterial infection. Transcripts of the *Msenod2* gene, a well-characterized marker of the nodule parenchyma, have been detected in these nodules (Truchet *et al.*, 1989; Hirsch *et al.*, 1992), whereas infection-related early nodulin genes were not (e.g. *Msenod12A* and *B*; Bauer *et al.*, 1994). In addition, purified Nod factors are able to enhance the expression of *Msenod40* in the host roots at concentrations necessary for the induction of cortical cell division and in the same place as *Rhizobium* does (Vijn *et al.*, 1993; our results). On the other hand, it has recently been shown that lower concentrations of Nod factors, sufficient to induce root hair curling but not to initiate nodule primordium formation (Truchet *et al.*, 1991; Schultze *et al.*, 1992), are able to induce certain early nodulin genes associated with nodule infection processes (*Psenod12* and *Msenod12B;* Horvath *et al.*, 1993; Bauer *et al.*, 1994, respectively).

enod40 transcripts were detected not only in the dividing cortical cells, but also in the pericycle cells of the vascular tissue of the root adjacent to the infection point (Yang *et al.*, 1993). These data suggest that the expression of *enod40* is induced very early, before these cells formed the nodule primordium. Moreover, the localization of the expression in infected and uninfected indeterminate nodules, showing that *enod40* expression is associated with all differentiating cells of the nodule organ, provides further support that it is involved in the differentiation process that transforms the cells arising from the nodule meristem into the fully differentiated central cells. However, the low transcript levels found in other tissues indicate that *enod40* expression is not exclusively related to nodule formation events, as is also true for other identified early nodulin genes (Scheres *et al.*, 1990b; Dehio and de Bruijn, 1992). This is also supported by the observed effect of the alteration of *enod40* expression on plant growth (see below) and the possible presence of homologous sequences in non-legume plants.

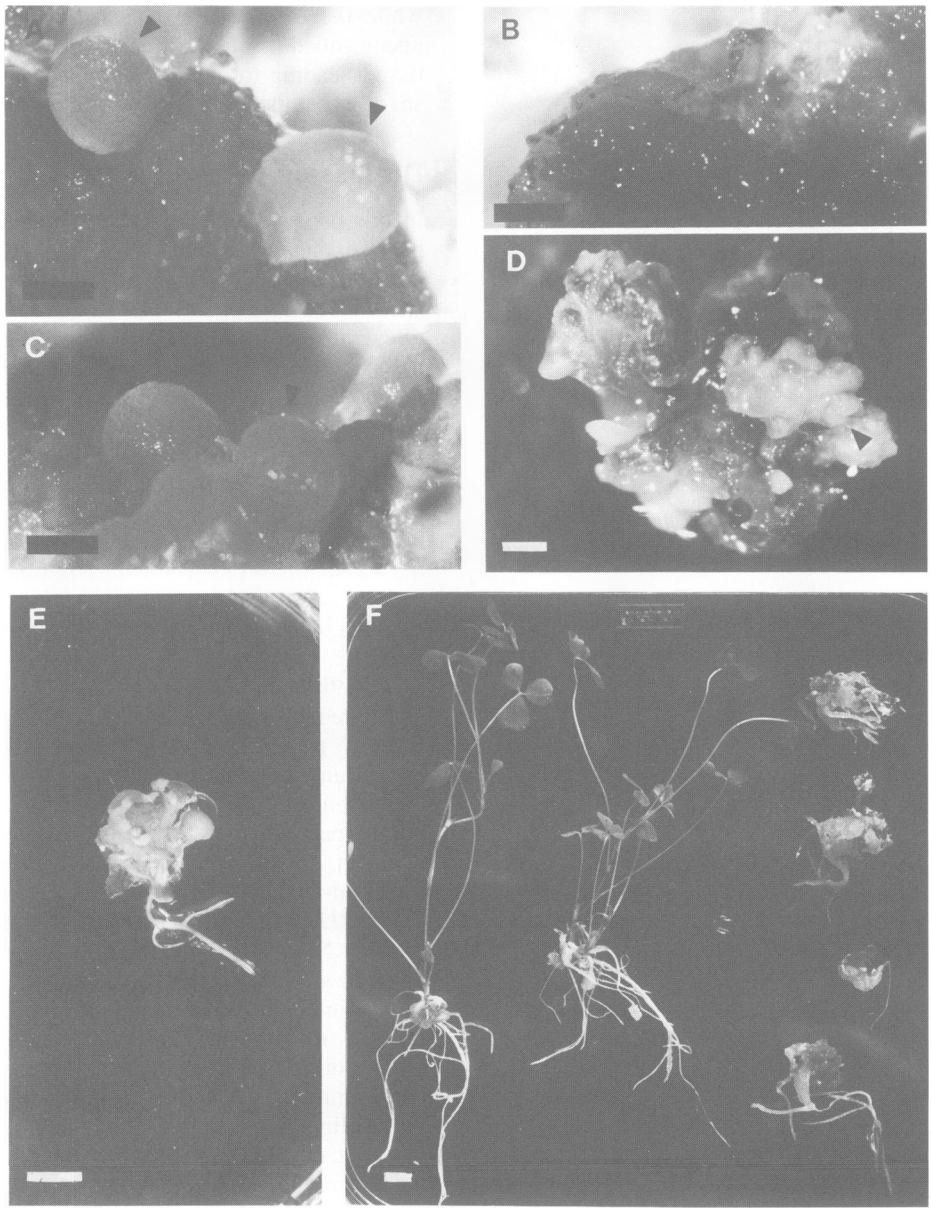


Fig. 7. Altered *Mtenod40* expression affects plant development in *Medicago*. (**A**) Two globular embryos formed on a callus (arrows) obtained from explants infected with an *A.tumefaciens* strain harbouring a control pCP60 vector. (**B**) Poor callus growth and no embryo formation (only brown tissue) on explant treated with *A.tumefaciens* bearing an antisense *enod40* construct after kanamycin selection (50 μg/ml). (**C**) Normal proliferation and embryogenesis (arrows) on the same explants as in (B) grown on non-selective medium. (**D**) A polyembryogenic callus (developing embryos are indicated by arrows) formed on explants treated with *A.tumefaciens* strains harbouring a sense *enod40* construct. Bars in (A–D): 1 mm. (**E**) The polyembryogenic callus as shown in (D), after transfer to rooting medium, gave rise to teratomas: a callus containing developing shoots and short roots. (**F**) On the left: two transgenic *Medicago* plants transformed with the control vector (pCP60) with normal root and stem development. On the right: four teratoma tumours obtained after transformation with overexpressing sense *enod40* constructs; several embryogenic shoots and proliferating calli can be observed with limited root development. Bars in (E) and (F): 1 cm.

We propose that the gene function of *enod40* is not associated with any protein, but with the RNA transcript. This is supported by the following observations: (i) there are no significant ORFs in the cDNAs nor in the genomic sequence of *Mtenod40;* (ii) no detectable translation products have been found for these transcripts; (iii) there exists a high degree of overall sequence conservation between different species; (iv) the soybean and *Medicago* gene transcripts showed very stable secondary structures. In spite of the fact that *in vitro* translation and examination of the nucleotide sequences of other nodulins did not reveal any alteration of the universal genetic code (Scheres *et al.*, 1990a; Franssen *et al.*, 1992), *Medicago enod40* sequences do not exhibit any long ORF. No translation products were either predicted or detected for these cDNAs. The correspondence between the *Mtenod40* genomic and cDNA sequences from *M.truncatula* eliminates any possibility of cloning artifacts or post-transcriptional modification of the RNA transcript. In addition, the degree of conservation of the nucleotide sequences between soybean and alfalfa species is relatively high for a non-coding gene. During evolution, non-functional sequences diverge faster between different species than coding ones. An inserted sequence was identified in the

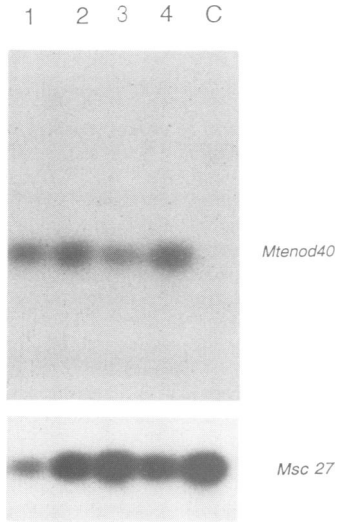


Fig. 8. Northern analysis of the overexpression of *Mtenod40* in transgenic *Medicago* tissue. Total RNAs (20 μg) from four transgenic teratomas (1–4, depicted in Figure 7) and from a whole control transgenic plant (C) were blotted for Northern analysis as described in Figure 1A.

Medicago genes by alignment with the soybean counterparts, and it is interesting that this region also shows the highest degree of divergence between the two *Medicago* sequences. This also explains why the ORF that had been proposed in a previous report (Yang *et al.*, 1993) for the soybean gene was not found in the *Medicago* sequences. Indeed, no coding capacity was found for any *enod40* cDNA sequence.

Analysis of the stability of the secondary structure for the *enod40* transcripts also indicates that the functional *enod40* gene product is a non-translated RNA. The most stable secondary structure of a given RNA sequence can be predicted with moderate accuracy by computation without additional information (Pace *et al.*, 1989). The free energy of folding estimates the ability of a given sequence to generate relevant secondary structures. The difference between the free energy values of RNA structures of regulatory signals and randomized sequences has already been used to predict transcription termination sites in prokaryotes (d'Aubenton-Carafa *et al.*, 1990). Therefore, sequence randomization and analysis of the distribution of the resulting free energies of folding were used to evaluate the stability of different RNAs. The *enod40* sequences showed significantly lower free energies of folding compared to the calculated averages, sharing this structural property with all non-coding RNAs tested and indicating that they are strongly structured RNAs. However, it remains to carry out structure–function studies with this gene to conclusively demonstrate its molecular nature.

To have an indication of a possible cellular role for the *enod40* RNAs, we determined the subcellular localization of the RNA transcript. Using *in situ* hybridization, it has previously been possible to localize transcripts at the subcellular level, specifically in the cell nucleus (Cumberledge *et al.*, 1990; Brown *et al.*, 1992). Using both cell fractionation and *in situ* hybridization techniques (coupled to confocal microscopy), we showed that the *Mtenod40* RNA is mainly found in the cytoplasm, co-migrating with other polyadenylated RNAs such as *Msc27*. Although the presence of a poly(A) tail in the *enod40* RNAs differentiates them from the well-defined RNAs possessing catalytic activity (nuclear introns, rRNA; Cech, 1993), polyadenylated non-coding RNAs have recently been identified as tumour suppressors in mammalian cells. Rastinejad *et al.* (1993) provided evidence that the constitutive expression of a polyadenylated non-translated region of the tropomyosin gene was sufficient to induce *in trans* the differentiation of a myogenic line, otherwise capable of inducing rhabdomyosarcomas in mice. This RNA can also control the growth response of other non-myogenic lines (e.g. fibroblasts; Rastinejad *et al.*, 1993). Furthermore, the gene H19, coding for a 2.5 kb transcript with no apparent ORF, was shown to be expressed in differentiating fetal cells during the transition of the mammalian embryo to the adult organism (Pachnis *et al.*, 1988). The H19 RNA also showed tumour suppressor activity on two embryonic tumour cell lines (Hao *et al.*, 1993) and overexpression of the H19 RNA severely affected the growth of mammalian embryos, leading to prenatal lethality of transgenic mice (Brunkow and Tilghman, 1991). Untranslated polyadenylated RNAs can, therefore, act *in trans* to suppress tumour formation and might have a role in differentiation and growth control. Several mechanisms have been proposed for this action, such as antisense action on specific transcripts, modulation of cellular kinases, translation perturbation or effects on RNA metabolism. The term 'riboregulator' has been proposed for RNAs with a regulatory effect not mediated by an encoded protein (Rastinejad and Blau, 1993).

We showed here that alteration of *enod40* expression affected plant growth and development. Antisense constructs of this gene impaired plant morphogenesis and regeneration of *Medicago* explants. No embryo formation was detected on these calli. The use of inducible plant promoters may help us further elucidate these antisense effects. Overexpression of the *Mtenod40* gene in *Medicago* transgenic plants induced the formation of embryogenic tumours or teratomas, indicating a role of the *enod40* gene in the induction of plant cell proliferation. In proliferating explants, these phenotypes can be mimicked by alteration of the auxin/cytokinin ratio (reviewed in Klee and Estelle, 1991). It has been suggested that the Nod factor is capable of inducing nodule morphogenesis by modifying the phytohormonal balances inside the root cortex and pericycle (Hirsch *et al.*, 1992). Application of auxin transport inhibitors to legume roots induced the formation of 'pseudonodules' (Hirsch *et al.*, 1989) and *R.meliloti* nod mutants can also form pseudonodules after the introduction of a gene coding for cytokinin synthesis and secretion (*tzs;* Cooper and Long, 1994). Based on the observed phenotypes after overexpression of *Mtenod40*, we hypothesize that induction of *enod40* by Nod signals in the responding cortical and pericycle cells is associated with the alteration of the hormonal balances inside these cells, leading to their division and dedifferentiation into a nodule primordium. Furthermore, this implies that apart from being induced in the nodule primordium, *enod40* function may then be linked to other organogenetic processes in plants where changes of the hormonal status of the plant tissues are involved, such as somatic embryogenesis (Liu *et al.*, 1993). *enod40* expression can be detected in

dividing cells and even in somatic embryos by RT-PCR (E.Jurkevitch, unpublished data) and this might explain the antisense effects on callus growth and embryo formation. Moreover, the efficiency of nodule formation correlated with the embryogenic properties of several tested *Medicago* cultivars (Kondorosi *et al.*, 1993), also suggesting a relationship between cell division leading to nodule initiation and the embryogenic potential of *Medicago*.

enod40 genes are expressed in the differentiating cells of the nodule formed by dividing pericycle and cortical cells at the site of initiation of the organ primordium and along the developmental gradient of indeterminate nodules. Their sequences strongly suggest that these genes act as RNA molecules, and alteration of *enod40* expression affects plant growth and differentiation. It is, therefore, attractive to propose that the *enod40* RNA is a first plant riboregulator. The analysis of *enod40* function may, then, have profound implications in the understanding of plant organogenesis and the development of a differentiated state, a crucial point considering the differences in the establishment of a determined cell fate between plants and animals.

Materials and methods

Plant material and plant growth

Alfalfa seeds (*M.sativa* ssp. *sativa* cv. Sitel) were surface sterilized by immersing the seeds for 10 min in 70% ethanol, rinsed with sterile water then treated with a 2.5% $HgCl_2$ solution for 15 min, followed by thorough washing with sterile water. *Medicago truncatula* cv. Ghor seeds were submitted to three cycles of freezing in liquid nitrogen and thawing prior to sterilization. The seeds were germinated on 1% agar plates for 1 day, then the seedlings were transferred and placed in a row in 12×12 cm Petri dishes containing nitrogen-free Gibson medium. The plates were incubated in a growth chamber at 24°C under a 16 h light period as described previously (Schultze *et al.*, 1992). The roots were inoculated 2 days later with *R.meliloti* or treated with Nod factors.

To obtain larger quantities of nodules, an aeroponic culture system was used (Allison *et al.*, 1993): 3-day-old seedlings or rooted shoot cuttings of *M.sativa* ssp. *varia* genotype A2 (Deak *et al.*, 1986) were transferred into aeroponic tanks and grown for 5 days in nitrogen-containing medium. Two days before infection with *R.meliloti* (see below), the medium was changed to a nutritive solution containing limited amounts of nitrogen (0.25 mM KNO_3). Roots were harvested from both inoculated and uninoculated plants, whereas stems, leaves and flowers were collected from infected plants.

Infection and treatment of plants

Plants grown in the aeroponic system were infected with rhizobia (8×10^5 cells/ml) and collection of wild-type nodules at different stages was performed as described previously (Allison *et al.*, 1993). Spontaneous nodules were obtained on selected genotypes of the A2 plants (E.Kondorosi, unpublished data) after 3 weeks of nitrogen starvation. White exo$^-$ nodules were collected 2 weeks after infection of A2 with the *exo* strain PP553 (Putnoky *et al.*, 1990).

Bacterial inoculation with agar blocks. Wild-type *R.meliloti* strain 41 (Rm41) and the non-nodulating mutant ZB138 (Kondorosi *et al.*, 1984) were grown overnight in minimal GTS medium in the presence of 10 μM luteolin, centrifuged, and resuspended at an OD_{540} of 0.35 in liquified 0.8% agarose, 10 mM $MgSO_4$. Ten microlitres of either inoculum were applied onto roots in the region containing the growing root hairs of 3-day-old alfalfa seedlings grown on Gibson plates. The suspension solidified on the root and the spot marked the site of infection. Small white nodules were first visible 4 d.p.i. with Rm41 in 60–70% of the plants. The region of the root where the agar spot was applied was dissected at appropriate time points.

Nod factor treatment. Roots of 3-day-old seedlings were immersed for 30 min in Nod factor [Nod RmIV(S)] solutions of concentrations ranging from 10^{-7} to 10^{-11} M. The Nod factor solutions were thereafter applied onto the roots at the same final concentration using the agar block technique. The appropriate root zone was cut and analysed 2 days later.

The collected tissues were either frozen in liquid nitrogen for RNA preparation or fixed for *in situ* hybridization.

Nucleic acid techniques and in vitro translation

Genomic DNA was prepared from 1 g of leaf material according to Dellaporta *et al.* (1983). Ten micrograms of *Eco*RI-digested DNA were used in each sample for Southern hybridizations. RNA was extracted from 0.2–1.0 g of frozen plant material with guanidinium thiocyanate, followed by caesium chloride gradient centrifugation according to standard procedures (Sambrook *et al.*, 1989). For Northern analysis, 10–20 μg RNA samples separated in formaldehyde gels were transferred onto Hybond N filters (Amersham). Northern and Southern hybridizations were carried out under high stringency in 5×SSC at 65°C as described previously (Sambrook *et al.*, 1989) and washed twice in 0.2×SSC, 0.1% SDS at 65°C for 10 min. Southern blots at low stringency were hybridized at 55°C in 5×SSC for 24–48 h and then washed twice in 5×SSC at 55°C for 30 min. All probes were labelled with ^{32}P by random priming.

Sequencing of cDNA and genomic clones was carried out by the dideoxy chain termination method using a Pharmacia Kit.

In vitro transcription of inserts cloned in pSK+ with T3 and T7 polymerases were performed in conditions for the generation of capped RNA using the Trans probe kit (Pharmacia). *In vitro* translation of purified transcripts was done using rabbit reticulocyte lysates and wheat germ extracts (Combination System, Promega).

Library constructions and screening

Total RNA was prepared from young, white, 4- to 5-day-old nodules induced on *M.sativa* ssp. *varia* A2 roots by *R.meliloti* Rm41. Poly(A)$^+$ RNA was prepared (Sambrook *et al.*, 1989) and used for the construction of a cDNA library in λ zap, according to the manufacturer's instructions (λzap kit, Stratagene). The library was plated (200 000 clones) and replicate filters were prepared on Hybond-N. cDNA probes from spontaneous nodule RNA (10^6 c.p.m./ml) or non-inoculated root RNA (2×10^6 c.p.m./ml) were made by reverse transcription of 20 μg of total RNA with oligo dT_{12-18} and the Superscript enzyme (Gibco BRL) for 2 h at 37°C. Labelled cDNAs were treated with 1 N NaOH for 10 min and neutralized with 1 N HCl. After heating at 80°C for 5 min, the probes were hybridized to replicate filters at high stringency (65°C, 5×SSC overnight; Sambrook *et al.*, 1989). Plaques showing differential expression were re-screened with the same probes. Inserts of isolated positive phages were amplified by PCR with T3/T7 primers (1 min, 94°C; 2 min, 55°C; 3 min, 72°C, 30 cycles). Duplicate samples of the amplification products were run in separate gels and transferred to Hybond-N filters. Newly prepared cDNA probes were used to screen for differential hybridization with the amplified inserts. Positive inserts were subcloned in pBluescript II SK+ (pSK+, Stratagene) for further characterization using an *in vivo* excision procedure.

Medicago truncatula cv. Ghor 13-day-old nodules were collected and processed to prepare poly(A)$^+$ RNA. This material was used to construct a cDNA library in λgt 10 using an Amersham kit according to the manufacturer's instructions. The insert of *Msenod40* was used to screen this cDNA library, as well as a λ-EMBL4 genomic library of the same species (E.Kondorosi and H.Sommer, unpublished). Two cDNAs and one genomic clone were isolated. *Mtenod40* cDNAs and the corresponding genomic clone were subcloned in pSK+ and M13 vectors, respectively, for characterization and sequencing (Sambrook *et al.*, 1989).

Reverse transcription and polymerase chain reaction

Total RNA (5–10 μg) was treated for 30 min at 37°C with 10 U of RNase-free DNase I in a volume of 24 μl in the presence of 40 mM Tris−HCl (pH 7.5), 6 mM $MgCl_2$ and 12 U RNAguard (Pharmacia). After heat inactivation of the DNase I, the RNA was precipitated with ethanol and resuspended in DEPC-treated water. DNase-treated RNA (2–5 μg) was reverse transcribed by 50 U MMLV H$^-$ superscript reverse transcriptase (Gibco BRL) in a 30 μl reaction mixture containing 100 pmol oligo dT_{12-18}, 1×HRT buffer (Gibco BRL), 10 mM DTT, 0.8 mM dNTP and 14.4 U RNAguard (Pharmacia) for 1.5 h at 37°C. After heat inactivation of the enzyme, equal amounts of cDNAs (in general one-tenth of the reaction mix) were used for amplification in 100 μl of 1×Taq buffer (Promega), 1.5 mM $MgCl_2$, 0.12 mM dNTP, 150 pmol 5′ and 3′ *Msenod40* primers (described in Figure 2), 10 pmol *Msc27* primers (Allison *et al.*, 1993) and 1.5 U of Taq polymerase (Promega). Amplification was performed during 15–25 cycles as follows: 1 min, 92°C; 1 min, 55°C; 1 min, 72°C. The amplification rate was in a linear range for all PCR products (data not shown). The concentration

of the *Msc27* control primers was reduced to a tenth of the optimum to avoid an out-titration of *enod40* PCR fragments by the highly abundant *Msc27* PCR products (Bauer *et al.*, 1994). Detection of *Msc27* transcripts was used as a control for RNA loading. Expression of this gene is constitutive in several *Medicago* tissues (Allison *et al.*, 1993). The technical aspects of the use of the semi-quantitative RT-PCR assay and the necessary controls have been developed in a previous work on *Msenod12* expression (Bauer *et al.*, 1994). One-tenth of the PCR products were separated on a 2% TBE agarose gel, transferred onto Hybond-N membranes and hybridized at high stringency with both *Mtenod40* (shown in Figure 3A) and *Msc27* (Allison *et al.*, 1993) probes. Oligonucleotides spanning an intron of the flavonoid-3-hydroxylase gene were used in a PCR reaction (30 cycles) aimed at checking genomic contamination of the cDNA samples (B.Charrier *et al.*, in preparation); the product of the reaction yields a 250 bp product amplified from cDNA as compared to a 500 bp band corresponding to the genomic DNA. No genomic DNA contamination was detected in the cDNA samples.

In situ hybridization

Preparation of sections, fixation and *in situ* hybridization of nodule and root tissues to RNA probes were carried out as described previously (Grosskopf *et al.*, 1993). *In vitro* transcription of inserts cloned in pSK+ with T3 and T7 polymerases were performed to obtain sense and antisense RNA probes, respectively, using a Pharmacia kit. Transverse and longitudinal sections were used for *in situ* hybridization and exposed for 1–3 weeks.

Whole-mount in situ hybridization and confocal microscopy

An *in situ* whole-mount procedure was adapted from a protocol provided by G.Engler and J.Almeida Engler, Laboratory of Genetics, University of Gent, Belgium, and described by Hemerly *et al.* (1993). Briefly, nodule and root tissues were fixed in 6% formaldehyde and 10% dimethyl sulphoxide (DMSO), and treated with 50% heptane to weaken cell walls. The tissues were then washed, post-fixed with 5% formaldehyde, digested with proteinase K for 10 min, rinsed, similarly post-fixed again, washed with hybridization solution (50% formamide, 5×SSC, 50 µg/ml heparin, 100 µg/ml salmon sperm DNA) and hybridized at 55°C for 18 h. Sense and antisense RNA probes were prepared from linearized pSK+ containing the *Msenod40* insert using digoxigenin DIG-11-rUTP (Boehringer Mannheim). To allow efficient penetration of the probes into the permeabilized tissue, partial hydrolysis of the transcripts was performed for 60 min, to yield 50–100 bp fragments, and then used for hybridization.

Hybridized samples were washed twice for 1 h in 2×SSC at room temperature, once in 1×SSC for 1 h at room temperature and once in 0.2×SSC for 10 min at 55°C. The samples were then washed in antibody dilution buffer [0.3% Tween 20, 1% bovine serum albumin (BSA), 150 mM NaCl, 100 mM Tris−HCl (pH 7.5)] for 5 min and incubated overnight at 4°C with diluted anti-digoxigenin−alkaline phosphatase (calf intestine) conjugate (1:500) in the same buffer, followed by two washes in the same buffer for 10 min and three washes in developing buffer [100 mM NaCl, 50 mM $MgCl_2$, 100 mM Tris−HCl (pH 9.5)]. Samples were covered with a coverslip and incubated with the fluorescent Fast Red substrate (Fast Red RC plus naphthol AS-TR; Sigma) for alkaline phosphatase in the presence of 3.3 mM levamisole to inhibit endogenous phosphatases. The reaction was stopped with 100 mM Tris−HCl (pH 7.0), 10 mM EDTA after 3–4 h. Slides were mounted in citifluor (Citifluor Ltd, London). Samples were examined using a confocal microscope (Sarastro 2000, Molecular Dynamics) configured for autofluorescence (green) versus phosphatase (red) excitation at 488 nm, dichroic 595 nm, green channel 530 LP, red 600 LP. Some sections were counterstained with mithramycin (10 µg/ml in 50 mM $MgCl_2$). In this case, the microscope was set with excitation 457 nm, and the yellow emission marked both cell walls and the nuclei. Dual images were taken simultaneously and corrected for spectral overlap using the Enhance algorithm (ImageSpace, Molecular Dynamics).

Cell fractionation

Pieces (up to 2 g) of *M.truncatula* roots containing 5-day-old nodules were collected from aeroponic tanks, homogenized in liquid nitrogen and resuspended in isolation buffer [1 M sucrose, 5 mM $MgCl_2$, 2 mM β-mercaptoethanol, 10 mM Tris (pH 7.5)]. After filtering through eight layers of gauze, the suspension was centrifuged at 5000 *g* for 10 min. The supernatant was kept at 4°C, the pellet was washed twice with isolation buffer and then resuspended in the same buffer with 0.5% Triton X-100 for 2–3 min. The suspension was layered onto a 35–80% bilayer of Percoll (Pharmacia) in isolation buffer. A band was recovered between the two layers of Percoll after centrifugation for 30 min at 8000 r.p.m. This nuclear fraction was then washed in isolation buffer and centrifuged at 5000 *g* for 10 min. The pellet was used for RNA preparation as described.

The supernatant from the first centrifugation step was centrifuged at 20 000 *g* in a swing-out rotor for 30 min. The pellet was then dissolved in guanidinium buffer, layered onto a CsCl cushion and centrifuged. The supernatant was directly layered onto the CsCl cushion and centrifuged. RNAs obtained from the different fractions were assayed by RT-PCR to determine the subcellular localization of *Mtenod40* and *Msc27*.

Computation

The coding capacity of selected sequences was monitored with the Test Code and Codon Preference programs (GCG software package). Computational analysis and alignment of sequences was done using the RNA Fold, the Shuffle program and the PileUP program (GCG software package; Zucker, 1989). Free energy data were statistically analysed with the Excel 3.0 program.

Coding sequences from *M.sativa* were: leghaemoglobin (accession number X13375), a cold acclimation protein (L07516), a stress-inducible protein (M74190), glutamine synthase (K03282) and a homologue of eukaryotic initiation factor 4d (X59441).

cDNAs for non-coding RNAs were: *Tetrahymena* ribozyme (V01416), *Agrobacterium tumefaciens* RNaseP RNA (M59354), potato spindle tuber viroid (PSTVF), 3′ untranslated region of the human tropomyosin gene (M12127) and the *Xenopus* 5S RNA (J01010).

Plant transformation

A full-length *Mtenod40* cDNA (Figure 4) was inserted in the sense and antisense orientation in the *Bam*HI site of the binary vector pCP60 (C.Coronado and P.Ratet, unpublished) between a 35S promoter and the nopaline synthase (NOS) 3′ terminator. This vector, derived from the pBin19 vector (Bevan, 1984), allowed the selection for kanamycin resistance (Kan^R) in the transformed plant cells. Binary plasmids containing sense and antisense constructs were transferred to *A.tumefaciens* LBA4404 by triparental mating (Ooms *et al.*, 1982). Root and stem explants from *M.sativa* ssp. *varia* A2 (Deak *et al.*, 1986) were transformed as described previously (d'Halluin *et al.*, 1990) with minor modifications. In brief, kanamycin-resistant calli were selected at 50 µg/ml and then somatic embryogenesis was induced by an auxin shock [10 mg/l 2,4-dichlorophenoxyacetic acid (2,4D) for 3 days]. Light-grown green embryos (obtained after 1 month) were transferred to regenerating medium [MS: Murashige and Skoog complete medium (Sigma), 2% sucrose, 50 µg/ml kanamycin, 0.5 mg/l 2,4D and 5 mg/l benzylaminopurine (BAP)] for 1 week and then transferred to rooting medium (MS, 1% sucrose and 50 µg/ml kanamycin). Transgenic plants were obtained 2–3 months after the auxin shock.

Control transformations were done with *A.tumefaciens* LBA4404 strain containing the binary vector pCP60 without insert. These plants, together with teratomas obtained from the transformation with sense constructs, were processed for DNA and RNA preparation as described above. Southern analysis was used for confirming the transgenic nature of these plants. Northern blots of total RNA were done (Sambrook *et al.*, 1989) to confirm the expression levels of the transgene.

Acknowledgements

We thank Drs H.McKhann and S.Brown for careful reading of the manuscript, constant support and critical discussions. We thank P.Bruhns, T.Coba, J.P.Prat and S.Brown for their help in the use of the confocal microscope. The pCP60 vector and helpful suggestions of Dr P.Ratet, as well as the help of J.Szécsi in the *in situ* hybridization experiments, are also acknowledged. We thank Dr M.Schultze for supplying the Nod factor and P.Bauer for help in the Nod factor experiments, and J.C.Barbet for writing a computer program. Research support was provided by the Commission of the European Countries (BRIDGE BIOT-900159-C) contract to A.K. M.C. was the recipient of an EEC postdoctoral fellowship of the ECLAIR Programme, E.J. was recipient of an EMBO long-term postdoctoral fellowship and G.P. was supported in the framework of the Jumelage Program between the Gif and Szeged laboratories.

References

Allison,L.A., Kiss,G.B., Bauer,P., Poiret,M., Pierre,M., Savouré,A., Kondorosi,E. and Kondorosi,A. (1993) *Plant Mol. Biol.*, **21**, 375–380.

Bauer,P., Crespi,M., Szécsi,J., Allison,L.A., Schultze,M., Ratet,P., Kondorosi,E. and Kondorosi,A. (1994) *Plant Physiol.*, **105**, 585–592.
Bevan,M. (1984) *Nucleic Acids Res.*, **12**, 8711–8721.
Brown,C., Hendrich,B., Rupert,J., Lafrenière,R., Xing,Y., Lawrence,J. and Willard,H. (1992) *Cell*, **71**, 527–542.
Brunkow,M. and Tilghman,S. (1991) *Genes Dev.*, **5**, 1092–1101.
Cech,T.R. (1993) *Nature*, **365**, 204–205.
Cooper,J. and Long,S. (1994) *Plant Cell*, **6**, 215–225.
Cumberledge,S., Zaratzian,A. and Sakonju,S. (1990) *Proc. Natl Acad. Sci. USA*, **87**, 3259–3263.
d'Aubenton-Carafa,Y., Brody,E. and Thermes,C. (1990) *J. Mol. Biol.*, **216**, 835–858.
Deak,M., Kiss,G.B., Koncz,C. and Dudits,D. (1986) *Plant Cell Rep.*, **5**, 97–100.
Dehio,C. and de Bruijn,F. (1992) *Plant J.*, **2**, 117–128.
Dellaporta,S.L., Wood,J. and Hicks,J.B. (1983) *Plant Mol. Biol. Rep.*, **4**, 19–21.
D'Halluin,K., Botterman,J. and De Greef,W. (1990) *Crop Sci.*, **30**, 866–871.
Franssen,H., Vijn,I., Yang,W.C. and Bisseling,T. (1992) *Plant Mol. Biol.*, **19**, 89–107.
Grosskopf,E., Ha,D.T.C., Wingender,R., Röhrig,H., Szécsi,J., Kondorosi,E., Schell,J. and Kondorosi,A. (1993) *Mol. Plant Microbe Interact.*, **6**, 173–181.
Hao,Y., Crenshaw,T., Moulton,T., Newcomb,E. and Tycko,B. (1993) *Nature*, **365**, 764–767.
Hemerly,A., Ferreira,P., Almeida Engler,J., Van Montagu,M., Engler,G. and Inzé,D. (1993) *Plant Cell*, **5**, 1711–1723.
Hirsch,A.M. (1992) *New Phytol.*, **122**, 211–237.
Hirsch,A.M., Bhuvaneswari,T., Torrey,J.G. and Bisseling,T. (1989) *Proc. Natl Acad. Sci. USA*, **86**, 1244–1248.
Hirsch,A.M., McKhann,H. and Lobler,M. (1992) *Int. J. Plant Sci.*, **153**, 5171–5181.
Horvath,B., Heidstra,R., Lados,M., Moerman,M., Spaink,H., Promé,J.-C., van Kammen,A. and Bisseling,T. (1993) *Plant J.*, **4**, 727–733.
Joshi,J., Caetao-Anolles,G., Graham,E. and Gresshoff,P. (1991) *Protoplasma*, **162**, 1–11.
Klee,H. and Estelle,M. (1991) *Annu. Rev. Plant Physiol. Plant Mol. Biol.*, **42**, 529–551.
Kondorosi,E., Banfalvi,Z. and Kondorosi,A. (1984) *Mol. Gen. Genet.*, **193**, 445–452.
Kondorosi,E., Schultze,M., Savouré,A., Hoffman,B., Dudits,D., Pierre,M., Allison,L., Bauer,P., Kiss,G.B. and Kondorosi,A. (1993) In Nester,E.W. and Verma,D.P.S. (eds), *Advances in Molecular Genetics of Plant–Microbe Interactions*. Kluwer, Dordrecht, Vol. 2, pp. 143–150.
Kouchi,H. and Hata,S. (1993) *Mol. Gen. Genet.*, **238**, 106–119.
Lerouge,P., Roche,P., Faucher,C., Maillet,F., Truchet,G., Promé,J.C. and Dénarié,J. (1990) *Nature*, **344**, 781–784.
Liu,C., Xu,Z. and Chua,N. (1993) *Plant Cell*, **5**, 621–630.
Nap,J.P. and Bisseling,T. (1990) *Science*, **250**, 948–954.
Noller,H.F., Hoffarth,V. and Zimniak,L. (1992) *Science*, **256**, 1416–1419.
Ooms,G., Hoykaas,P., Van Veen,R., Van Beelen,P., Regensburg-Tunik,T. and Schilperoort,R. (1982) *Plasmid*, **7**, 15–29.
Pace,N.R., Smith,D.K., Olsen,G.J. and James,B.D. (1989) *Gene*, **82**, 65–75.
Pachnis,V., Belayew,A. and Tilghman,S.M. (1988) *EMBO J.*, **7**, 673–681.
Pay,A., Heberle-Boers,E. and Hirt,H. (1992) *Plant Mol. Biol.*, **19**, 501–503.
Putnoky,P., Petrovics,G., Kereszt,A., Grosskopf,E., Ha,D.T.C., Banfalvi,Z. and Kondorosi,A. (1990) *J. Bacteriol.*, **172**, 5450–5458.
Rastinejad,F. and Blau,H.M. (1993) *Cell*, **72**, 903–917.
Rastinejad,F., Conboy,M., Rando,T. and Blau,H.M. (1993) *Cell*, **75**, 1107–1117.
Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Sassanfar,M. and Szostak,J. (1993) *Nature*, **364**, 550–553.
Scheres,B., van Engelen,F., van de Knaap,E., van de Wiel,C., van Kammen,A. and Bisseling,T. (1990a) *Plant Cell*, **2**, 687–700.
Scheres,B., van de Wiel,C., Zalensky,A., Horvath,B., Spaink,H., van Eck,H., Zwartkruis,F., Wolters,A.M., Gloudemans,T., van Kammen,A. and Bisseling,T. (1990b) *Cell*, **60**, 281–294.
Schultze,M., Quiclet-Sire,B., Kondorosi,E., Virelizier,H., Glushka,J.N., Endre,G., Géro,S.D. and Kondorosi,A. (1992) *Proc. Natl Acad. Sci. USA*, **89**, 192–196.
Spaink,H.P. (1992) *Plant Mol. Biol.*, **20**, 977–986.
Spaink,H.P., Sheeley,D.M., van Brussel,A.A.N., Glushka,J., York,W.S., Tak,T., Geiger,O., Kennedy,E.P., Reinhold,V.N. and Lugtenberg,B.J.J. (1991) *Nature*, **354**, 125–130.
Truchet,G., Barker,D.G., Camut,S., Billy,F., Vasse,J. and Huguet,T. (1989) *Mol. Gen. Genet.*, **219**, 65–68.
Truchet,G., Roche,P., Lerouge,P., Vasse,J., Camut,S., de Billy,F., Promé,J.C. and Dénarié,J. (1991) *Nature*, **351**, 670–673.
Vasse,J., de Billy,F., Camut,S. and Truchet,G. (1990) *J. Bacteriol.*, **172**, 4295–4306.
Vijn,I., das Neves,L., van Kammen,A., Franssen,H. and Bisseling,T. (1993) *Science*, **260**, 1764–1765.
Yang,W.C., Katinakis,P., Hendriks,P., Smolders,A., de Vries,F., Spee,J., van Kammen A., Bisseling,T. and Franssen,H. (1993) *Plant J.*, **3**, 573–585.
Zucker,M. (1989) *Science*, **244**, 48–52.

Received on March 11, 1994; revised on July 18, 1994

Note added in proof

The nucleotide sequence data reported here have been deposited in the EMBL data bank under the following accession numbers: X80262 (*Mtenod'40*), X80263 (*Msenod40*), X80264 (*Mtenod40*gen).