Plant Physiol. (1996) 112: 1437–1446

Gene Structure and Differential Regulation of the *Rhizobium*-Induced Peroxidase Gene *rip1*[1]

Hui-Mei Peng, Dale A. Dreyer, Kathryn A. VandenBosch, and Douglas Cook*

Department of Plant Pathology and Microbiology (H.-M.P., D.C.), and Department of Biology (D.A.D., K.A.V.), Texas A&M University, College Station, Texas 77843

Primary expression of the *Rhizobium meliloti*-induced peroxidase gene *rip1* occurs prior to nodule morphogenesis, specifically at the site of impending rhizobial infection (D. Cook, D. Dreyer, D. Bonnet, M. Howell, E. Nony, K. VandenBosch [1995] Plant Cell 7: 43–55). We examined the distribution and structure of *rip1* transcript throughout nodule development. We determined that expression of *rip1* in root tips is correlated with the competence of this zone for symbiotic association, whereas after rhizobial infection *rip1* transcript is specifically associated with the zone of nodule development, including nascent nodule primordia. *rip1* transcripts are characterized by multiple polyadenylation sites distributed within 200 to 400 bp of the translation stop site, and a single major transcription initiation site in close proximity to the *rip1* open reading frame. Thus, *rip1* expression is likely to be mediated through effects on a single transcription unit. Immediately 5′ of the *rip1* transcription unit DNA sequence analysis identified a 377-bp DNA element containing extensive repeat structure that is widely distributed in the *Medicago truncatula* genome.

The legume root nodule is a specialized plant organ in which nodule-inducing bacteria, rhizobia, and their legume host conduct symbiotic nitrogen fixation. Development of the symbiosis can been divided into distinct phases based on morphological changes and changes in host and bacterial gene expression (Vincent, 1980; Scheres et al., 1990b; Mylona et al., 1995). Prior to infection, flavonoids or other plant metabolites contained in root exudates induce the expression of the bacterial *nod* gene regulon (Fisher and Long, 1992), the function of which is required for nodulation and for the synthesis of a rhizobial signal molecule, an acylated chitin oligomer called the Nod factor (Lerouge et al., 1990). Specific modifications to a common Nod-factor structure are correlated with the host range of the bacterial symbiont (Roche et al., 1991). In purified form, specific Nod factors can induce many of the host responses induced by compatible *Rhizobium*, including distortion of root-hair cell growth, mitotic activation of the root cortex, and induction of several nodulation-associated genes called nodulins (for review, see Mylona et al., 1995, and refs. therein).

Expression of nodule-specific genes occurs during all phases of nodule development. The earliest known nodulins include the peroxidase gene *rip1* (Cook et al., 1995) and the Pro-rich protein gene *ENOD12* (Scheres et al., 1990a; Pichon et al., 1992), both of which are expressed throughout the differentiating epidermis coincident with the root zone subsequently infected by the bacterium. Because the Nod factor is both necessary and sufficient for induction of these genes (Horvath et al., 1993; Journet et al., 1994; Cook et al., 1995), the signaling circuit between the symbionts must be complete within 3 h of inoculation, well in advance of infection and nodule morphogenesis. This evidence for a Nod-factor-induced regulatory program that precedes rhizobial infection both spatially and temporally is consistent with the suggestion that events critical to symbiotic development occur during the preinfection period (Bhuvaneswari et al., 1981; Pichon et al., 1992; Cook et al., 1995).

Despite the specific induction of certain nodulin genes prior to infection, high-level induction of most early nodulins, including *ENOD12*, is correlated with bacterial infection and the onset of nodule morphogenesis. Thus, rhizobial infection is accompanied by the induction of plant cyclins and other cell-cycle-related genes in a narrow band of cells that defines the path of impending infection thread growth (Yang et al., 1994). Upon infection, *ENOD12* (Scheres et al., 1990a) and *ENOD5* (Scheres et al., 1990b) are expressed in cells in the vicinity of the growing infection thread. Coincident with the initiation of infection, cells of the root inner cortex reenter the cell cycle, giving rise to the nascent nodule primordium. The nodule primordium is the site of multiple nodule-specific transcripts, including *ENOD12* (Scheres et al., 1990a), Gm93 (Kouchi and Hata, 1993), MtPRP4 (Wilson et al., 1994), and *ENOD40*, which is simultaneously induced in the adjacent pericycle tissue (Kouchi and Hata, 1993; Yang et al., 1993; Asad et al., 1994; Crespi et al., 1994).

Although the temporal induction of most early nodulin genes is similar and their tissue localization often overlaps, specific patterns of expression differ, indicating that synchronous regulation of early nodulin gene expression is likely to be complex. For example, both *ENOD12* and *ENOD5* are induced by purified Nod factors (Horvath et al., 1993). However, during nodule development *ENOD5* expression is limited to infected cells, whereas *ENOD12* is expressed in both infected cells and in uninfected cells in advance of infection (Scheres et al., 1990a). In fact, not all Nod-factor-inducible genes are equally sensitive to induction by the Nod factor (Cook et al., 1995), suggesting the

[1] This research was supported by U.S. Department of Agriculture grant no. 95-37305-2464.

* Corresponding author; e-mail drc1653@acs.tamu.edu; fax 1-409-862-4790.

Abbreviation: RACE, rapid amplification of cDNA ends.

presence of multiple Nod-factor-induced transcriptional mechanisms. Ardourel et al. (1994) have proposed the existence of at least two distinct Nod factor receptors; however, this diversity of gene expression patterns might also be explained by the existence of other primary and secondary signaling events. Examples of other signal molecules that have been implicated in establishing the symbiosis include bacterial exopolysaccharide (Battisti et al., 1992; Urzainqui and Walker, 1992), which is required for sustained infection of alfalfa by *Rhizobium meliloti*, and uridine, which is thought to be the "stele factor" (Libbenga et al., 1973; Smit et al., 1995) that influences activation of the nodule primordium in pea.

In contrast to other early nodulins, the *Rhizobium*-induced peroxidase gene *rip1* is most highly induced prior to rhizobial infection (Cook et al., 1995), when its transcript is localized to epidermal cells at the site of impending infection. Its decline at the onset of bacterial infection and nodule morphogenesis, when transcript levels for other Nod-factor-dependent genes such as *ENOD12*, *ENOD40*, *MtPRP4*, and *ENOD5* are increasing, serves to further distinguish *rip1* and indicates the presence of a heretofore unknown regulatory event. However, despite the largely transient nature of *rip1* induction, intermediate levels of transcript remain throughout early nodule morphogenesis, and low levels of transcript are associated with uninoculated roots.

The purpose of this study was 2-fold: to further examine *rip1* expression prior to rhizobial inoculation and subsequent to infection and to characterize the *rip1* transcript structure. We show here that *rip1* expression is consistently associated with nodulation or the competence for nodulation, and that *rip1* transcripts are characterized by a single major transcription initiation site and multiple polyadenylation sites. Furthermore, DNA sequence analysis identified a 377-bp DNA element adjacent to the transcription initiation site that may be a remnant of a transposable element.

MATERIALS AND METHODS

Plant Material and Growth Conditions

For nodulation experiments, germinated seedlings of *Medicago truncatula* genotype A-17 were sown in aeroponic chambers containing an inorganic nutrient medium (Lullien et al., 1987) lacking nitrate and ammonia. Five-day-old seedlings were used for the initial inoculation. Since 5-d-old seedlings contain only a single root tip and no lateral roots, every root tip in the population was a potential site for nodule development. Under these conditions nodulation was rapid and uniform, with macroscopic nodule primordia first evident at 72 h post-inoculation. The inoculum consisted of *Rhizobium meliloti* strain ABS7 grown to late log phase in TY broth cultures (5 g/L tryptone, 3 g/L yeast extract, and 1.47 g/L $CaCl_2 \cdot 2H_2O$; Somasegaran and Hoben, 1994), washed by centrifugation, and added directly to the nutrient medium in aeroponic chambers. Initial bacterial concentration in the aeroponic medium was approximately 2×10^6/mL. Plants were maintained with a 12-h photoperiod (18°C night/22°C day) and a light intensity of 300 μmol m^{-2} s^{-1}.

Nucleic Acid Isolation and Gel Blot Analysis

DNA was isolated from fresh or frozen tissue by extraction with hexadecyltrimethylammonium bromide according to the method of Rogers and Bendich (1988) and purified further by CsCl density gradient centrifugation. RNA was isolated as described previously (Cook et al., 1995). Control hybridizations for mRNA content were with the soybean actin gene clone pSac3 (Shah et al., 1982). General nucleic acid manipulations, including DNA gel electrophoresis, RNA formaldehyde gel electrophoresis, and transfer of nucleic acids to nylon membranes were as described by Sambrook et al. (1989). Radioactive probes were prepared by the incorporation of [α-^{32}P]dCTP using the oligolabeling procedure of Feinberg and Vogelstein (1983). Hybridization was conducted at 60°C in a solution of 7% SDS, 0.25 M NaH_2PO_4, pH 7.0, and 0.1 mM Na_2EDTA. Following overnight hybridization, filters were washed successively in solutions of 0.1% SDS and 2× SSC (1× SSC = 0.15 M NaCl, 0.015 M sodium citrate; 21°C for 15 min); 0.1% SDS and 0.1× SSC (21°C for 15 min); and 0.1% SDS and 0.1× SSC (60°C for 1 h).

DNA Sequencing

DNA sequence was determined by the dideoxy chain termination method (Sanger et al., 1977) adapted for use in the PCR according to the manufacturer's directions using a kit purchased from Epicentre Technologies (Madison, WI). Ambiguities in the DNA sequence were resolved by tailing prematurely terminated products with terminal deoxynucleotidyl transferase (as described in literature available from United States Biochemical). Templates were an ordered series of exonuclease III deletion clones, prepared as described by Sambrook et al. (1989).

In Situ Hybridizations

Microtome sections (8- to 10-μm thick) of paraffin-embedded root tissue were prepared and processed as described previously (Cook et al., 1995). Sense and antisense [α-^{35}S]UTP riboprobes were prepared by in vitro transcription of a cDNA fragment corresponding to exon III of the *rip1* peroxidase gene (pG2A11, Cook et al., 1995). ^{35}S-labeled transcripts were purified by passage over an RNase-free Sephadex G-50 column and added directly to hybridization solution. Hybridization and subsequent processing were based on the protocol of Cox and Goldberg (1988), with minor modifications according to de Billy et al. (1991). Rehydrated sections were treated successively with HCl, pronase E, and acetic anhydride to reduce the background on subsequent hybridization and to improve penetration of the ^{35}S-labeled riboprobe. Sections were prehybridized in 50% formamide hybridization buffer at 42°C followed by overnight hybridization with sense or antisense ^{35}S-labeled riboprobes under similar conditions. Nonhybridized riboprobe was removed by treatment with RNase A, followed by three 40-min, high-stringency washes (0.2% SSC, 1 mM DTT at 63°C). Slides were coated with Kodak NTB 2 emulsion and exposed for 3 to 10 d in the dark before developing. Specimens were observed by

bright-field, dark-field, or Nomarski microscopy using a Zeiss Axioskop or an Olympus BH-2 microscope. Selected sections were counterstained with 0.05% azur II and 0.05% methylene blue in 1% sodium *meta*-borate (w/v).

Primer-Extension Experiments

Synthetic oligonucleotides used as primers (P1 and P2) were complementary to the nucleotide sequences indicated in Figure 5. The primers were chosen based on their position relative to possible TATA elements predicted from the *rip1* DNA sequence. Primers were labeled to high specific activity by means of T4 polynucleotide kinase and [γ-^{32}P]ATP according to the method of Sambrook et al. (1989). Labeled primer (2×10^6 cpm; average specific activity for six independent experiments = $5.32 \times 10^6 \pm 1.4$ cpm/pmol) was co-precipated with 5 to 20 μg of total RNA, and processed essentially as described by Scheres et al. (1990a). Following reaction with avian myeloblastosis virus reverse transcriptase (GIBCO-BRL), samples were extracted with phenol chloroform and ethanol-precipitated with carrier tRNA. Denatured extension products were resolved on a 6% polyacrylamide-8M urea sequencing gel, and migration was compared with the products of DNA sequencing reactions (described above) performed on a *rip1* DNA template using the corresponding oligonucleotides as primers.

RACE

5′- and 3′RACE were performed as described by Frohman (1990). *rip1*-specific primers (5′-tcacttcactaccttgtccc-3′ for 5′RACE and 5′-ctggtggccacactatcggg-3′ for 3′RACE) were designed to yield overlapping cDNA products originating from exon III of *rip1* (Cook et al., 1995). The oligo $(dT)_{17}$ primer and the *rip*-specific primers also contained an *Eco*RI adapter sequence (5′-ctagtcgaattctggtgg-3′) at their 5′ end to facilitate cloning. Subsequent to first- and second-strand cDNA synthesis reactions, PCR was conducted using a primer corresponding to the adapter sequence (5′-ctagtcgaattctggtgg-3′). *Eco*RI-digested PCR products were cloned into similarly digested and dephosphorylated pBluescript SK+ vector (Stratagene). To visualize 3′RACE products on 5% polyacrylamide-8M urea sequencing gels, [^{35}S]dATP (3.7×10^5 cpm) was included in the PCR reaction of the standard 3′RACE protocol, except that PCR was discontinued after 30 cycles. Subsequent to electrophoresis the dried gels were subject to autoradiography and the sizes of 3′RACE products were determined based on migration relative to a known DNA sequence standard.

RESULTS

rip1 Is Differentially Regulated in Root Tips

Nodule development occurs within a narrow (approximately 2 cm) root zone that is first associated with the differentiating root tip prior to bacterial infection, and then with the mature root portion during bacterial infection and nodule morphogenesis. To determine the spatial correlation between *rip1* transcript and nodulation, RNA was isolated from the terminal 2-cm root-tip zone (referred to as "tip") and the remaining root portion (referred to as "main root zone") of inoculated roots throughout a 17-d time course of nodule development. Following inoculation with *Rhizobium*, average root growth was 1.98 ± 0.29 cm/d; therefore, the susceptible root zone was contained within the root tip sample during the first 24-h interval (representing the pre-infection period), and within the main root sample during subsequent 24-h intervals (including initial infection through nodule morphogenesis and function).

As shown in Figure 1A, the tips of uninoculated roots (0 h) contained a low level of *rip1* transcript. This level of *rip1* transcript is typical of root tips on uninoculated plants, independent of plant age (D. Cook, unpublished observation of plants from 3–21 d post-germination). Following inoculation *rip1* was transiently induced in root tips, reaching maximal levels during the pre-infection period (Fig. 1A, 1 d), and declining markedly by the onset of infection and early nodule morphogenesis (2–3 d). During the period of nodule morphogenesis in the main root (3–7 d post-inoculation), *rip1* transcript in root tips declined to levels below those observed in uninoculated roots (Fig. 1A). However, by d 8 *rip1* transcript was again detected in root tips at levels comparable to those associated with uninoculated roots. It is interesting that the recurrence of *rip1* transcript in root tips by d 8 was correlated with the reinitiation of nodulation on these same plants. Specifically, up to 20% of the plants sampled beginning on d 10 contained macroscopic nodule primordia in near proximity to the root tip. Control hybridization with the soybean actin gene demonstrated that RNA concentrations in each lane were approximately equal.

Nodule Primordia Are the Primary Post-Infection Locus of *rip1* Transcript

Prior to inoculation, *rip1* transcript was not detected in the main root zone of uninoculated roots (Fig. 1B, 0 h). However, by 1 d post-inoculation an intermediate level of transcript was evident in the main root zone, and was maintained at this level throughout early nodule morphogenesis (Fig. 1B, 2–4 d). Based on more precise dissection of roots containing nodule primordia, we determined that *rip1* transcript was restricted to the zone of active nodule morphogenesis (Fig. 2A, 72-h time point). By d 7, nearing full nodule differentiation, *rip1* transcript in the main root zone had declined to levels comparable to those observed in the main root tissue of uninoculated plants. Lateral roots, which were first evident at d 4, were bulked with the main root sample for RNA gel blot analysis. The absence of detectable *rip1* transcript in these root RNA samples after d 5 (Fig. 1B) indicates that lateral root tips are not a primary locus of *rip1* expression. Lateral roots were also not a locus of secondary nodulation in these experiments.

As a control for early nodulin gene expression, RNA gel blots were stripped and rehybridized with an ENOD12 gene probe. The transcript levels of ENOD12 were maximal at d 4 (Fig. 1B), consistent with its known high level of expression during early infection and nodule differentia-

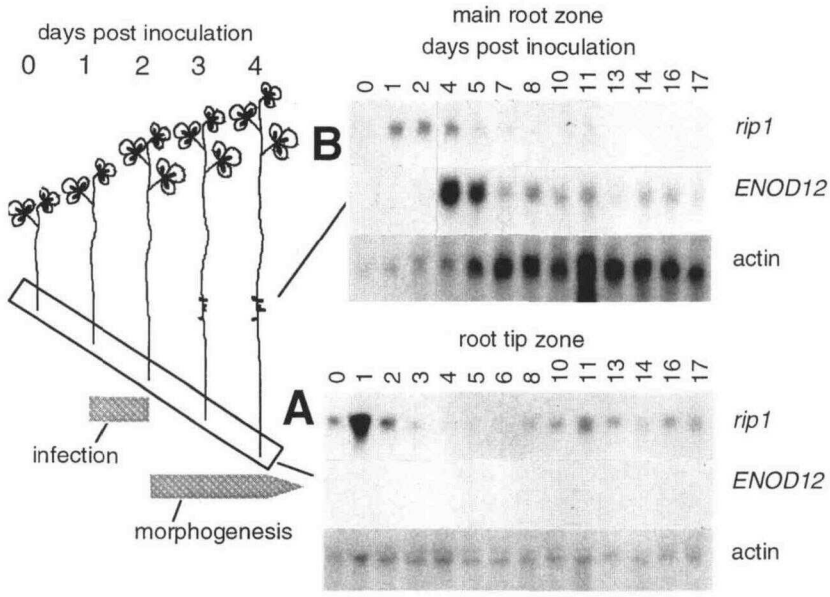

Figure 1. Induction of the early nodulins *rip1* and *ENOD12* in the tip and main root zone of inoculated *M. truncatula*. Autoradiographs of gel blots containing total RNA isolated from the 2-cm root tip zone (A) and main root zone (including laterals; B) of *M. truncatula* inoculated with *R. meliloti*. Each lane contains RNA isolated from 20 roots to allow direct comparison of transcript levels between samples on a per root basis irrespective of increasing root mass and RNA content. Control hybridization with the soybean actin gene confirmed that levels of both *ENOD12* and *rip1* transcripts vary independently of total root RNA content. Blots were probed successively with ^{32}P-labeled *rip1* exon III (clone pG2A11; Cook et al., 1995), Mt*ENOD12* (Pichon et al., 1992), and actin (Shah et al., 1982), as indicated. Samples were collected at regular 24-h intervals (days post-inoculation), with the exception of the 1-d sample, which corresponds to 16 h. As indicated in the accompanying diagram, the primary period of infection occurs from 24 to 48 h, and macroscopic nodule primordia are first evident near 72 h postinoculation. Results are typical of those obtained in three replicate experiments.

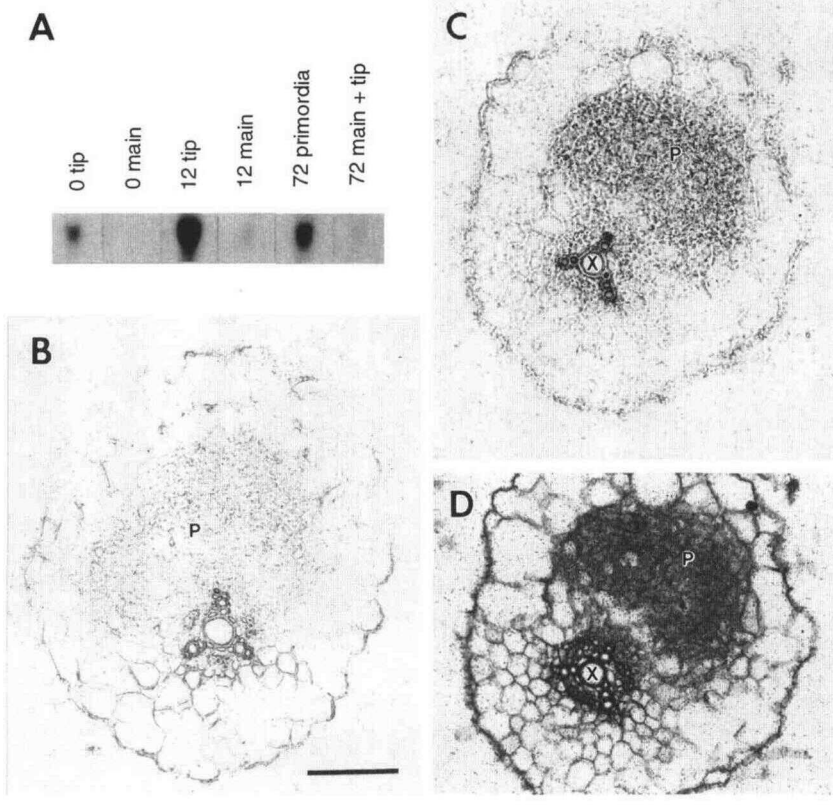

Figure 2. Spatial distribution of *rip1* transcript during early nodule morphogenesis. A, Gel blots are of total RNA isolated from specific root regions, as indicated. Zero- and 12-h tip samples correspond to the terminal 2-cm root zone, whereas the 72-h primordia sample corresponds to an approximately 2-cm zone from the main root, which contains macroscopic nodule primordia. For each time point, RNA isolated from the remaining root tissue is contained in the adjacent lanes labeled 0 main, 12 main, and 72 main + tip. B, C, and D, Bright-field light micrographs are of 8- to 10-μm tissue sections taken from the root zone containing nodule primordia at 72 h. View B was hybridized with a sense probe as a control. Views C and D show hybridization with the antisense *rip1* probe. D is the same section depicted in C after counterstaining with azur II/methylene blue. Note the intense hybridization to the nodule primordium (P), and an intermediate level of signal in the root epidermis. The high contrast associated with cell walls of xylem (X) is due to refraction, and is not related to hybridization. Views B, C, and D are 120×. Bar = 100 μm.

tion. The continued presence of ENOD12 transcript at later times is consistent with its known expression in the preinfection zone of mature nodules. In contrast to *rip1*, *ENOD12* transcript was never detected in total RNA isolated from root tips (Fig. 1A).

To determine more precisely the correlation between *rip1* expression and nodule morphogenesis, we used in situ hybridization to examine transcript distribution in plants at 72 h post-inoculation. Initially we determined that *rip1* transcript was confined to a narrow zone (approximately 2 cm) containing nascent nodule primordia (Fig. 2A, 72 h). In subsequent experiments we analyzed *rip1* transcript within this zone by in situ hybridization. Figure 2C shows that at 72 h, when macroscopic primordia were first evident, *rip1* transcript was abundant in the nascent nodule primordium, and sense controls (Fig. 2B) confirmed that the signal was specific to the coding strand of *rip1*. *rip1* transcript was also evident in the root epidermis (Fig. 2C). At 72 h post-inoculation there is no evident tissue differentiation within the primordium (Fig. 2D), and transcript appears to be uniformly distributed within all cells. The high contrast of the xylem evident in Figure 2, C and D, is a consequence of light refraction within the tissue, and not of probe hybridization.

Defining the *rip1* Transcription Unit

In a previous study we reported the isolation and characterization of a peroxidase gene genomic clone from *M. truncatula* (Cook et al., 1995) that was 98% identical to a 264-bp cDNA fragment obtained from a nodulin-enriched cDNA library. However, we were uncertain whether the 2% sequence divergence was an artifact of multiple rounds of PCR used in the cDNA enrichment or whether the cDNA clone was the product of a different but related peroxidase gene. To address this question we used 3′- and 5′RACE to isolate cDNA clones from *Rhizobium*-induced roots and compared their coding and noncoding regions with those of the peroxidase genomic clone. Sequence analysis of three 3′ cDNA clones and two 5′ cDNA clones revealed that each of these was 100% identical (in both coding and noncoding regions) to the peroxidase gene genomic clone, consistent with its assignment as the *Rhizobium*-induced peroxidase gene, *rip1*. The verified exon-intron junctions were highly homologous to the known splice site consensus sequence of 5′- . . . AG/GUAAGU . . . CAG/. . . -3′, and the high AT content of the two introns (69.9 and 74.7%) conforms to the AT-rich bias of introns from other dicotyledonous plant genes (Goodall and Filipowicz, 1989). As shown in Figure 3, the length of the 3′ untranslated region was variable, since each of the 3′ cDNAs possessed a different site of polyadenylation distributed over a 113-bp region. However, with the exception of one putative far upstream element (UUGUA; Mogen et al., 1992) adjacent to the proximal polyadenylation site, homology to predicted plant polyadenylation signals was not evident. In subsequent analysis, we deduced polyadenylation site distribution by electrophoretic separation of ^{35}S-labeled 3′RACE products on sequencing gels; the results confirmed that polyadenylation of *rip1* transcripts occurs at multiple additional sites

Figure 3. Nucleotide sequence of the *rip1* 3′ untranslated region. Numbering is given relative to the primary transcription initiation site (+1 in Fig. 5). The predicted translation stop codon "tag" at position 1016 is underlined. Polyadenylation sites determined by cDNA sequence analysis are indicated by arrowheads. Putative AURE elements are underlined in bold with the core "ATTTA" sequence in uppercase letters. A possible far upstream element "ttgta" is boxed.

clustered within 200 to 400 bp of the translation stop codon (data not shown). Two regions with homology to a class of mRNA destabilizing elements, the so-called AURE elements (Gillis and Malter, 1991), were also identified within the *rip1* 3′ untranslated region. The common feature of these elements is a central AUUUA motif within an otherwise AU-rich domain (Fig. 3).

To provide a more complete description of the *rip1* transcription unit, we sought to map sites of transcription initiation. Two hundred roots were harvested at 0, 12, and 72 h post-inoculation; each time point is representative of a distinct tissue specificity and/or level of *rip1* transcript (see Fig. 2A). Prior to extraction of total RNA, each root was dissected to enrich regions known to contain *rip1* transcript: 0- and 12-h roots were divided into tip and main root samples, and 72-h roots were dissected into the zone containing visible nodule primordia and a second sample containing the remainder of the root, including the tip. Primer-extension experiments with these RNAs identified a single major extension product, designated transcript T1, clustered with several minor extension products 28 bp upstream of the proposed translation start site (Fig. 4). T1 transcripts were detected only in the tips of 0- and 12-h plants, and only in the nodulation zone of 72-h plants, in agreement with the specificity of *rip1* expression determined previously by northern blot analysis (see Figs. 1 and 2A). A putative TATA-box sequence (5′-TATATAAA-3′) is located precisely 30 bp upstream of the T1 initiation site (Fig. 5), which itself bears homology to the weak consensus sequence, 5′-PyPyCA^{+1}PyPyPyPyPy-3′, common to the initiation site of many TATA-containing promoters, and may represent an initiator *cis* element (Smale and Baltimore, 1989). A second primer, P2, which hybridizes transcriptionally upstream of T1, identified multiple additional extension products (Fig. 4). This result was surprising, since primer P1 detected only the T1 transcript. We subsequently determined (see below) that primer P2 is complementary to a highly repeated DNA element, indicating that it has the potential to detect transcripts associated with other copies of the DNA repeat.

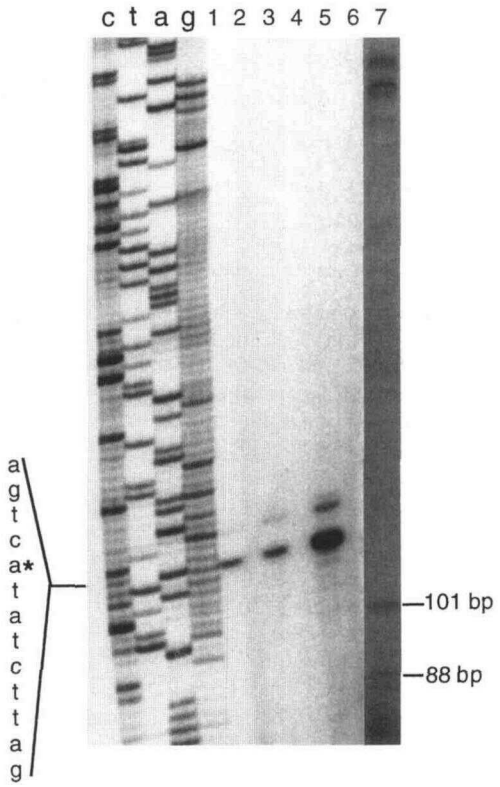

Figure 4. Primer-extension analysis of *rip1* transcripts. Results are presented for primer-extension experiments with two separate primers, P1 and P2, the annealing sites for which are underlined in Figure 5. Lanes 1 to 6 represent extension products obtained with primer P1. Lane 7 represents extension products obtained with primer P2. Primer-extension reactions were conducted with total RNA isolated from various portions of inoculated roots: lanes 1 and 2, 0-h tip and main root, respectively; lanes 3 and 4, 12-h tip and main root, respectively; lanes 5 and 6, 72-h zone with nodule primordia and combined main root and tip, respectively; and lane 7, 12-h tip. The DNA sequencing reaction in lanes c, t, a, and g was primed with P1 and provides a reference for extension products in lanes 1 to 6. The sizes of the extension products in lane 7 were estimated by comparison with a known DNA sequencing reaction (data not shown).

The *rip1* Transcription Initiation Site Is Preceded by a Transposable Element-Like Structure

In conjunction with the primer extension analysis described above, we sequenced a 2.6-kb region immediately 5′ of the proposed *rip1* translation start site. In close proximity to the peroxidase gene open reading frame, we identified a 377-bp region with multiple repeated DNA elements. As shown in Figure 5, the 377-bp region contains two terminal inverted repeats of 114 and 115 bp that differ by only 4 bp. The intervening 147-bp sequence shares no obvious homology with the terminal inverted repeats, except for the presence of two 17-bp repeats that occur a total of eight times throughout the 377-bp region. Although the context of these 17-bp repeats is not conserved, their spacing interval (in particular, combinations 1/2, 3/4/5/6, and 7/8) has a common denominator near 10.5 bp, consistent with the possibility that these 17-bp repeats lie on the same face of a B-form DNA helix. In addition to its internal structure, the 377-bp region is flanked by two 9-bp direct repeats. To determine the copy number of the *rip1* sequence element, Southern blots were hybridized with probes corresponding to the internal region of the 377-bp element or with an adjacent DNA fragment containing all of exon 1 and the first 54 bp of intron 1. As shown in Figure 6, the internal region of the sequence element hybridized to multiple bands in *Eco*RV-digested DNA, whereas a probe from the adjacent exon I region of *rip1* hybridized to only a single band.

DISCUSSION

In a previous study we determined that a distinguishing feature of the *R. meliloti* infection zone is the prior induction of *rip1* transcript throughout the differentiating *M. truncatula* root epidermis (Cook et al., 1995). Here we further analyzed the relationship between *rip1* expression and nodulation by examining the distribution and structure of *rip1* transcript in uninoculated roots and throughout bacterial infection and nodule morphogenesis. We determined that all phases of *rip1* expression are correlated with specific aspects of the *Rhizobium*-legume symbiosis, as indicated diagrammatically in Figure 7. For example, expression in root tips was correlated with the competence of this zone for nodulation. Thus, low levels of transcript were observed in uninoculated root tips, where the differentiating root zone was competent but not yet induced for symbiotic association, whereas after inoculation and before infection high levels of *rip1* transcript were observed in root tips (in the differentiating epidermis; Cook et al., 1995), coincident with the zone subsequently infected by *Rhizobium*. During early nodule morphogenesis, when root tips were suppressed for renewed nodulation, there was a corresponding absence of *rip1* transcript in tips, whereas the reinitiation of nodulation in previously nodulated roots was correlated with the reappearance of *rip1* transcript in the primary root tip but not in lateral root tips, which uniformly lacked nodules. These results indicate that both primary and secondary nodulation are preceded by the expression of *rip1*, specifically in root tips competent for nodulation. Furthermore, the absence of *rip1* transcript from root tips subsequent to primary nodulation is temporally correlated with feedback inhibition of nodulation (Bhuvaneswari et al., 1981; Caetano-Anollés and Bauer, 1988; Caetano-Anollés and Gresshoff, 1991).

In addition to its presence at sites of potential or impending infection, *rip1* transcript was also associated with the zone of active nodule morphogenesis, both in nascent nodule primordia and in the surrounding root epidermis. The high level of *rip1* transcript observed in nodule primordia suggests that *rip1* is either involved in or directly influenced by the development of the nodule primordium. Consistent with this interpretation, an *M. truncatula* mutant that exhibits a 10-fold increase in the number of nodules displays a correspondingly large increase in *rip1* transcript levels coincident with the appearance of microscopic primordia (R. Penmetsa and D. Cook, unpublished data). A distinguishing feature of nodule primordia is small, rapidly dividing cells, which presumably have a (temporarily)

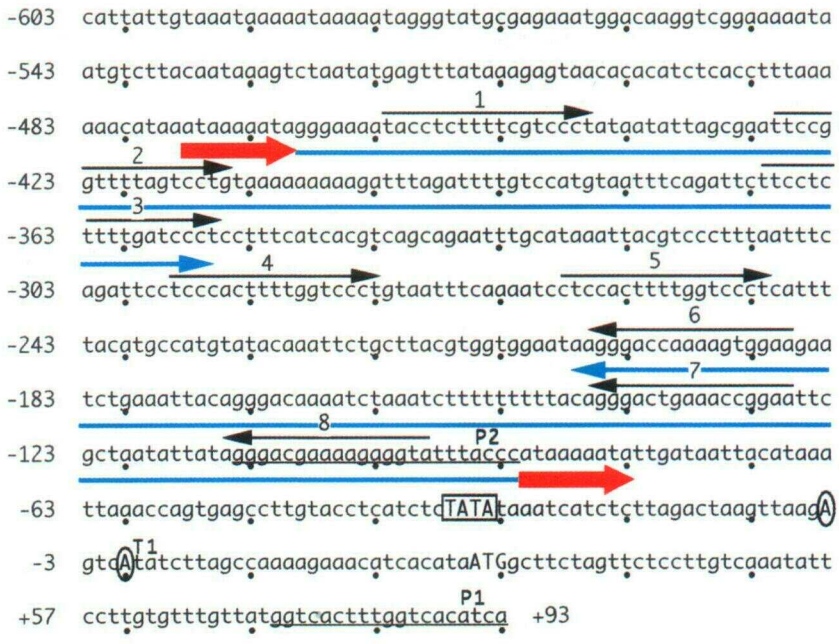

Figure 5. Nucleotide sequence 5′ of the *rip1* open reading frame. Blue and red arrows underline the large inverted repeats and the flanking direct repeats, respectively. Eight copies of a 17-bp repeat are overlined by black arrows and numbered 1 to 8. Sites complementary to primers P1 and P2 are underlined in black. Circled adenine residues at positions +1 and −4 correspond to extension products detected by primer P1 (Fig. 4). A putative TATA element is boxed. Numbering is relative to the major extension product, designated T1.

reduced capacity for cell expansion. Peroxidases have been implicated in limiting cell expansion in rapidly dividing embryogenic carrot cell cultures (Cordewener et al., 1991; van Engelen and de Vries, 1992), as well as during other periods of plant development (MacAdam et al., 1992; Zheng and van Huystee, 1992). It is possible, therefore, that Rip1 facilitates the development or maintenance of nodule primordia by restricting cell expansion.

Whatever the specific role of peroxidase in the nodule primordium, many models of cell-wall peroxidase function involve cross-linking of cell-wall proteins as a means to alter cell architecture. A possible substrate protein is the early nodulin MtPRP4 (Wilson et al., 1994), which is also highly expressed in nodule primordia and, like Rip1, is presumed to be localized to the cell wall. The repeating Pro-rich motifs of the predicted protein product of MtPRP4 are similar to the Pro-rich cell-wall protein PRP2 of soybean, which is known to be efficiently cross-linked by a peroxidase-dependent mechanism (Bradley et al., 1992). However, despite the coincidence of *rip1* and MtPRP4 transcripts in nodule primordia, primary expression of MtPRP4 occurs later, during nodule differentiation and in mature nodules, and thus is not coincident with *rip1*. Consequently, if the role of Rip1 and MtPRP4 is to alter cell-wall structure via oxidative cross-linking, then an alternative peroxidase(s) and substrate protein(s) (e.g. other repetitive, Pro-rich, cell-wall proteins) would be implicated during other phases of nodulation.

Analysis of *rip1* transcript structure indicates the existence of a single, major initiation site, transcripts from which exhibit the same general tissue specificity determined for *rip1* expression by northern blot analysis. Thus, the probable contribution of different regulatory elements to *rip1* expression is likely to be mediated through effects on a single transcription unit. Sequence homologies suggest the presence of a corresponding TATA box and *cis*

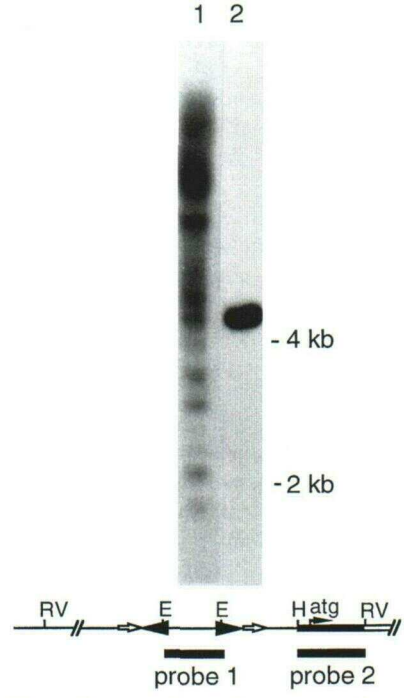

Figure 6. Gel blot of *Eco*RV-digested *M. truncatula* genomic DNA. Replicate filter membranes were hybridized with the 0.3-kb *Eco*RI fragment internal to the 377-bp DNA element (designated "probe 1," shown in lane 1) and with the 0.3-kb *Hinf-Eco*RI fragment containing all of exon 1 and the first 57 bp of intron 1 (designated "probe 2," shown in lane 2). The >4-kb band in lane 2 corresponds in size to the product expected from complete digestion by *Eco*RV at the *rip1* locus. The line diagram (bottom) shows the origin of the probe DNAs and the general structure of the 5′ region of the *rip1* locus. The inverted and direct repeats of the 377-bp DNA element are represented by the adjacent solid and open arrows, respectively. H, *Hin*fI; E, *Eco*RI; RV, *Eco*RV.

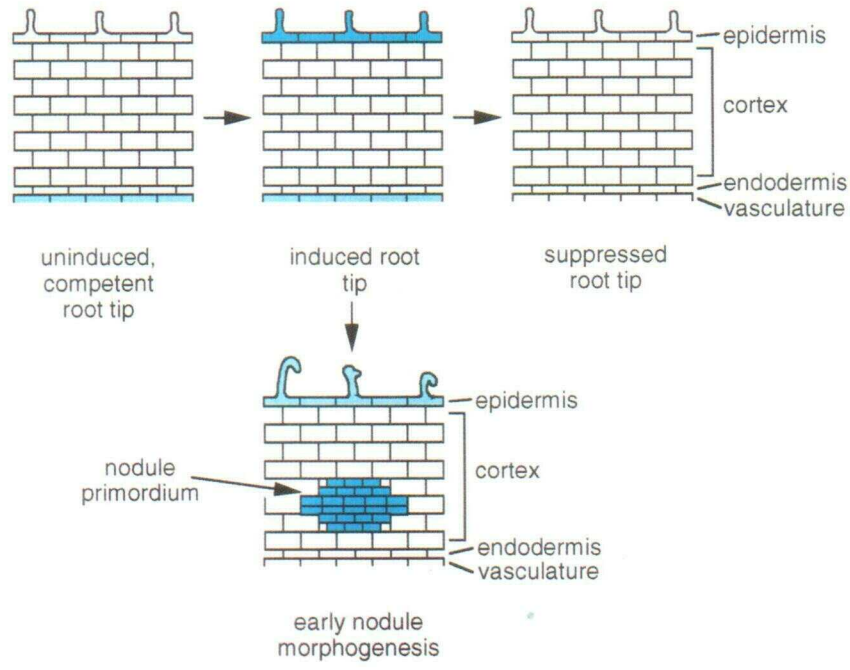

Figure 7. Schematic representation of *rip1* transcript distribution during early nodulation. Longitudinal views of *M. truncatula* root tips and the zone of active nodule development. The intensity of blue coloration represents the relative level of *rip1* transcript observed in various tissues according to Cook et al. (1995) and Figures 1 and 2 herein. Prior to infection, low levels of *rip1* transcript are associated with the differentiating vascular region. Subsequent to inoculation but prior to infection, *rip1* is strongly induced in the differentiating root epidermis. During early nodule morphogenesis (suppressed root tip), *rip1* transcript is absent from the root tip, but abundant at sites of successful infection (early nodule morphogenesis), particularly in nascent nodule primordia.

initiator element, which in many eukaryotic promoters are required for function and accurate initiation, respectively (Smale and Baltimore, 1989; Weis and Reinberg, 1992; Zhu et al., 1995). No evidence of other probable major sites of initiation was found.

Using DNA sequence analysis, we determined that the *rip1* transcription unit is preceded by a sequence element with features reminiscent of a defective plant transposable element, including small size (377 bp), lack of extended open reading, and high copy number within the *M. truncatula* genome. The two 115-/114-bp terminal inverted repeats that constitute the majority of the sequence element are flanked by 9-bp direct repeats typical of target-site duplications generated during transposable element insertion (Doring and Starlinger, 1984). Furthermore, an 11-bp sequence at the 5′ flank (aaataaaaata) is repeated approximately 100 bp farther upstream (see Fig. 5, positions −590 to −580), consistent with preexisting, short, sequence duplications often found near transposable element insertion sites. Many normal eukaryotic genes, particularly plant genes, contain the remnants of transposable elements or retrotransposons in their 5′ flanking regions (Masson et al., 1987; White et al., 1994). With our present data we are unable to determine the contribution (if any) of this sequence element to *rip1* regulation; however, in other gene systems transposon-like sequences often contain known or suspected *cis* regulatory elements necessary for correct regulation of the adjacent gene, including both positive and negative regulatory effects (see White et al., 1994, and refs. therein).

The 3′ untranslated regions of several eukaryotic transcripts have also been shown to contain *cis* regulatory elements (Sachs, 1993). In *rip1* the variable polyadenylation site usage is typical of many plant genes in which, in the absence of a conserved polyadenylation signal, the high AT content of the 3′ untranslated region may be sufficient to direct variable polyadenylation (Luehrsen and Walbot, 1994). The 3′ untranslated region of the *rip1* transcript also contains sequence elements typical of certain unstable mRNAs in both plant (Sachs, 1993; Zhang et al., 1993; Zhang and Mehdy, 1994) and animal systems (Shaw and Kamen, 1986; Shyu et al., 1991). These instability elements consist of reiterated AUUUA motifs or a single AUUUA motif contained within an otherwise AU-rich region (Gillis and Malter, 1991). As shown in Figure 3, the *rip1* 3′ untranslated region contains two separate AUUUA motifs within regions of 95 and 84% AU, respectively, where the overall AU content of the 3′ untranslated region is 72%. It is intriguing that the position of polyadenylation affects the number of AUUUA motifs that remain in the processed message (see Fig. 3), and could theoretically provide a mechanism to indirectly affect message stability.

Our previous results suggested that the induction of *rip1* in differentiating epidermal cells is controlled by a Nod-factor-induced regulatory mechanism (Cook et al., 1995). However, the decline in overall *rip1* transcript levels observed at the onset of infection (Fig. 1, A and B, 24- to 48-h transition) and then again at the onset of primary nodule differentiation (Fig. 1B, 4- to 5-d transition) indicates that the loss of *rip1* transcript at specific points in the symbiosis is also likely to be highly regulated. In the present analysis we have identified several putative *cis* regulatory elements, including the transposon-like element 5′ of the *rip1* transcription initiation site and possible AURE instability elements in the *rip1* 3′ untranslated region. Furthermore, we have recently determined that the *rip1* promoter region is highly methylated in DNA extracted from leaves, where *rip1* is inactive (H.-M. Peng and D. Cook, unpublished data). Thus, the detailed analysis of transcriptional as well as pre- and posttranscriptional processes involved in *rip1*

regulation will be a significant component of these ongoing studies.

Received June 6, 1996; accepted August 27, 1996.
Copyright Clearance Center: 0032–0889/96/112/1437/10.
The accession number for the sequence reported in this article is U16727.

LITERATURE CITED

Ardourel M, Demont N, Debellé F, Maillet F, de Billy F, Promé J-C, Dénarié J, Truchet G (1994) *Rhizobium meliloti* lipooligosaccharide nodulation factors: different structural requirements for bacterial entry into target root hair cells and induction of plant symbiotic developmental responses. Plant Cell **6:** 1357–1374

Asad S, Fang Y, Wycoff KL, Hirsch AM (1994) Isolation and characterization of cDNA and genomic clones of Ms*ENOD40*: transcripts are detected in meristematic cells of alfalfa. Protoplasma **183:** 10–23

Battisti L, Lara JC, Leigh JA (1992) Specific oligosaccharide form of the *Rhizobium meliloti* exopolysaccharide promotes nodule invasion in alfalfa. Proc Natl Acad Sci USA **89:** 5625–5629

Bhuvaneswari TV, Bhagwat AA, Bauer WD (1981) Transient susceptibility of root cells in four common legumes to nodulation by rhizobia. Plant Physiol **68:** 1144–1149

Bradley DJ, Kjellbom P, Lamb CJ (1992) Elicitor- and wound-induced oxidative cross-linking of a proline-rich plant cell wall protein: a novel, rapid defense response. Cell **70:** 21–30

Caetano-Anollés G, Bauer WD (1988) Feedback regulation of nodule formation in alfalfa. Planta **175:** 546–557

Caetano-Anollés G, Gresshoff PM (1991) Alfalfa controls nodulation during the onset of *Rhizobium*-induced cortical cell division. Plant Physiol **95:** 366–373

Cook D, Dreyer D, Bonnet D, Howell M, Nony E, VandenBosch K (1995) Transient induction of a peroxidase gene in *Medicago truncatula* precedes infection by *Rhizobium meliloti*. Plant Cell **7:** 43–55

Cordewener J, Booij J, van der Zandt H, van Engelen F, van Kammen A, deVries S (1991) Tunicamycin-inhibited carrot somatic embryogenesis can be restored by secreted cationic peroxidase isoenzymes. Planta **184:** 478–486

Cox KH, Goldberg RB (1988) Analysis of plant gene expression. *In* CH Shaw, ed, Plant Molecular Biology: A Practical Approach. IRL Press, Oxford, UK, pp 1–34

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 non-translatable RNA involved in plant growth. EMBO J **13:** 5099–5112

de Billy F, Barker DG, Gallusci P, Truchet G (1991) Leghaemoglobin gene transcription is triggered in a single cell layer in the indeterminate nitrogen-fixing root nodule of alfalfa. Plant J **1:** 27–35

Doring HP, Starlinger P (1984) Barbara McClintock's controlling elements: now at the DNA level. Cell **39:** 253–259

Feinberg AP, Vogelstein B (1983) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal Biochem **132:** 6–13

Fisher RF, Long SR (1992) *Rhizobium*-plant signal exchange. Nature **357:** 655–660

Frohman MA (1990) RACE: rapid amplification of cDNA ends. *In* MA Innis, DH Gelfand, JJ Sninsky, TJ White, eds, PCR Protocols: A Guide to Methods and Applications. Academic Press, San Diego, CA, pp 28–38

Gillis P, Malter JS (1991) The adenosine-uridine binding factor recognizes the AU-rich elements of cytokine, lymphokine, and oncogene mRNAs. J Biol Chem **266:** 3172–3177

Goodall GJ, Filipowicz W (1989) The AU-rich sequences present in the introns of plant nuclear pre-mRNAs are required for splicing. Cell **58:** 473–483

Horvath B, Heidstra R, Lados M, Moerman M, Spaink HP, Promé JC, van Kammen A, Bisseling T (1993) Lipooligosaccharides of *Rhizobium* induce infection-related early nodulin gene expression in pea root hairs. Plant J **4:** 727–733

Journet EP, Pichon M, Dedieu A, de Billy F, Truchet G, Barker DG (1994) *Rhizobium meliloti* Nod factors elicit cell-specific transcription of the *ENOD12* gene in transgenic alfalfa. Plant J **6:** 241–249

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

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

Libbenga KR, Van Iren F, Bogers RJ, Schraag-Lamers MF (1973) The role of hormones and gradients in the initiation of cortex proliferation and nodule formation in *Pisum sativum* L. Planta **114:** 29–39

Luehrsen KR, Walbot V (1994) Intron creation and polyadenylation in maize are directed by AU-rich RNA. Genes Dev **8:** 1117–1130

Lullien V, Barker DG, de Lajudie P, Huguet T (1987) Plant gene expression in effective and ineffective root nodules of alfalfa (*Medicago sativa*). Plant Mol Biol **9:** 43–48

MacAdam JW, Nelson CJ, Sharp RE (1992) Peroxidase activity in the leaf elongation zone of tall fescue. Plant Physiol **99:** 872–878

Masson P, Surosky R, Kingsbury JA, Fedoroff NV (1987) Genetic and molecular analysis of the *Spm*-dependent *a-m2* alleles of the maize *a* locus. Genetics **177:** 117–137

Mogen BD, MacDonald MH, Leggewie G, Hunt AG (1992) Several distinct types of sequence elements are required for efficient mRNA 3′ end formation in a pea *rbcS* gene. Mol Cell Biol **12:** 5406–5414

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

Pichon M, Journet EP, Dedieu A, de Billy F, Truchet G, Barker DG (1992) *Rhizobium meliloti* elicits transient expression of the early nodulin gene *ENOD12* in the differentiating root epidermis of transgenic alfalfa. Plant Cell **4:** 1199–1211

Roche P, Debellé F, Maillet F, Lerouge P, Faucher C, Truchet G, Dénarié J, Promé J (1991) Molecular basis of symbiotic host specificity in *Rhizobium meliloti*: *nod*H and *nod*PQ genes encode the sulfation of lipo-oligosaccharide signals. Cell **67:** 1131–1143

Rogers SO, Bendich AJ (1988) Extraction of DNA from plant tissues. *In* SB Gelvin, RA Schilperoort, DP Verma, eds, Plant Molecular Biology Manual. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp 1–10

Sachs AB (1993) Messenger RNA degradation in eukaryotes. Cell **74:** 413–421

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain terminating inhibitors. Proc Natl Acad Sci USA **74:** 5463–5467

Scheres B, van de Weil C, Zalensky A, Horvath B, Spaink H, van Eck H, Zwartrkruis F, Wolters A, Gloudemans T, van Kammen A, Bisseling T (1990a) The *ENOD12* gene product is involved in the infection process during the pea- *Rhizobium* interaction. Cell **60:** 281–294

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

Shah DM, Hightower RC, Meagher RB (1982) Complete nucleotide sequence of a soybean actin gene. Proc Natl Acad Sci USA **79:** 1022–1026

Shaw G, Kamen R (1986) A conserved AU sequence from the 3′ untranslated region of GM-CSG mRNA mediates selective mRNA degradation. Cell **46:** 659–667

Shyu AB, Belasco JG, Greenberg ME (1991) Two distinct desta-

bilizing elements in the c-fos message trigger deadenylation as a first step in rapid mRNA decay. Genes Dev **5:** 221–231

Smale ST, Baltimore D (1989) The "initiator" as a transcription control element. Cell **57:** 103–113

Smit G, Dekoster CC, Schripsema J, Spaink HP, Vanbrussel AA, Kijne JW (1995) Uridine, a cell division factor in pea roots. Plant Mol Biol **29:** 869–873

Somasegaran P, Hoben HJ (1994) Handbook for Rhizobia. Springer-Verlag, New York

Urzainqui A, Walker GC (1992) Exogenous suppression of the symbiotic deficiencies of *Rhizobium meliloti exo* mutants. J Bacteriol **174:** 3403–3406

van Engelen FA, de Vries SC (1992) Extracellular proteins in plant embryogenesis. Trends Genet **8:** 66–70

Vincent JM (1980) Factors controlling the legume-*Rhizobium* symbiosis. *In* WE Newton, WH Orme-Johnson, eds, Nitrogen Fixation II. University Park Press, Baltimore, MD, pp 103–109

Weis L, Reinberg D (1992) Transcription by RNA polymerase II: initiator-directed formation of transcription-competent complexes. FASEB J **6:** 3300–3309

White SE, Habera LF, Wessler SR (1994) Retrotransposons in the flanking regions of normal plant genes: a role for *copia*-like elements in the evolution of gene structure and expression. Proc Natl Acad Sci USA **91:** 11792–11796

Wilson RC, Long F, Maruoka EM, Cooper JB (1994) A new proline-rich early nodulin from *Medicago truncatula* is highly expressed in nodule meristematic cells. Plant Cell **6:** 1265–1275

Yang W, Katinakis P, Hendriks P, Smolders A, de Vries F, Spee J, van Kammen A, Bisseling T, Franssen H (1993) Characterization of Gm*ENOD40*, a gene showing novel patterns of cell-specific expression during soybean nodule development. Plant J **3:** 573–585

Yang WC, de Blank C, Meskiene I, Hirt H, Bakker J, van Kammen A, Franssen H, Bisseling T (1994) *Rhizobium* Nod factors reactivate the cell cycle during infection and nodule primordium formation, but the cycle is only completed in primordium formation. Plant Cell **6:** 1415–1426

Zhang S, Mehdy MC (1994) Binding of a 50-kD protein to a U-rich sequence in an mRNA encoding a proline-rich protein that is destabilized by fungal elicitor. Plant Cell **6:** 135–145

Zhang S, Sheng H, Lui Y, Mehdy MC (1993) Fungal elicitor-induced bean proline-rich protein mRNA down-regulation is due to destabilization that is transcription and translation dependent. Plant Cell **5:** 1089–1099

Zheng X, van Huystee RB (1992) Peroxidase-regulated elongation of segments from peanut hypocotyls. Plant Sci **81:** 47–56

Zhu Q, Dabi T, Lamb C (1995) TATA box and initiator functions in the accurate transcription of a plant minimal promoter in vitro. Plant Cell **7:** 1681–1689