Regulators and Regulation of Legume Root Nodule Development

Jens Stougaard*

Laboratory of Gene Expression, Department of Molecular and Structural Biology, University of Aarhus, Gustav Wieds Vej 10, 8000 C Aarhus, Denmark

Nitrogen is the nutrient plants require in the highest amount, and in agriculture, nitrogen availability has a major influence on both yield and product quality. In nature plants acquire nitrogen by assimilation of nitrate and ammonium or from dinitrogen through association with nitrogen-fixing bacteria. Symbiotic nitrogen fixation, where the plant supplies the carbon source for the energy-dependent reduction of dinitrogen and protects the oxygen-sensitive nitrogenase enzyme, is among the most effective fixation systems. To establish a symbiosis, the bacterial microsymbionts gain access to single plant cells and install themselves in compartments surrounded by a plant membrane. In Gunnera sp., the cyanobacterium Nostoc sp. invades pre-existing stem glands and forms nitrogen-fixing heterocysts in infected cells. In most other symbiotic interactions, a specialized plant organ, the root nodule, is developed to provide optimal conditions for the nitrogen-fixing bacteria. Among woody plant species belonging to eight different families, an interaction with the gram-positive genus Frankia leads to the development of actinorhizal root nodules. In legumes, gram-negative soil bacteria belonging to the family Rhizobiaceae (here collectively called Rhizobium) infect root tissue and induce the formation of the nitrogen-fixing nodules. Why certain plants are able to develop root nodules is unclear, but recent phylogenetic studies based on DNA sequence analysis place all plants involved in rhizobial or actinorhizal symbiosis in the same lineage and suggest that the predisposition for nodule formation evolved only once (Soltis et al., 1995; Doyle, 1998).

The relationship between Rhizobium and legume plants is selective. Individual species of rhizobia have a distinct host range allowing nodule formation of a particular set of legume plants. For example, Rhizobium leguminosarum bv. viciae nodulates pea and vetch, whereas Bradyrhizobium japonicum nodulates soybean. At the other extreme, the exceptionally broad host-range Rhizobium sp. NGR234 nodulates 353 legume species representing 122 genera (Pueppke and Broughton, 1999). Differences in both infection processes and organogenic programs are reflected in variations in root nodule morphology (Doyle, 1998), but overall there are pronounced developmental similarities as would be expected from a common ancestry. To cover most aspects of this unusual plant-prokaryote symbiosis, the study of nodulation is a multifaceted research area aiming to understand this plant-microbe interaction in a framework of physiological and developmental processes underlying infection and organogenesis. With this perspective, this Update draws on observations from different Rhizobium-legume interactions. The following sections focus on plant control of root nodule organ formation and sketch the way plant genetics and functional genomics are changing our thinking. The early signal exchange as well as the biosynthesis and properties of the bacterial Nod factor signal molecules have been reviewed extensively (Dénarié et al., 1996; Spaink, 1996; Downie and Walker, 1999, and refs. therein) and will be presented only briefly.

HISTOLOGY AND NODULE DEVELOPMENT

In the most studied legumes, infection occurs via an infection thread that takes the bacteria through the root hair into the root cortex and distributes them to cells, which become the infected cells of the nitrogen-fixing nodule (Fig. 1). The root zone susceptible to invasion is located behind the root tip where root hairs are still growing and competent for invasion. In response to attached bacteria, root hairs deform and curl setting up a pocket that provides a site for initiating the infection. The infection thread is a plant-derived structure originating from plasma membrane invagination accompanied by external deposition of cell wall material. In advance of the intracellular “inward” progressing thread, root cortical cells dedifferentiate and reenter the cell cycle. Cortical cells prepared for infection thread passage appear to be arrested in the G2 phase, whereas cells completing the cycle resume division to establish the nodule primordium. Later in the process, pattern formation and cell differentiation specify tissue and cell types. The bacteria are endocytosed into a subset of cells where they differentiate into nitrogen-fixing bacteria surrounded by the peribacteroid membrane of the symbiosome. In the mature functional nodule, peripheral vascular bundles are connected to the root vasculature and the main tissues/cell types can be distinguished cytologically and to some extent with

* E-mail stougaard@mbio.aau.dk; fax 45–8620–1222.
molecular markers. In determinate nodules such as soybean the meristematic activity ceases early, and the nodule grows by expansion giving a spherical shape (Fig. 1). All developmental stages from root hair curling to nodule senescence are consequently phased in time. Indeterminate nodules such as pea maintain an active meristem depositing cells that are subsequently infected. This results in a cylindrical shape with the organ formative developmental stages represented along the longitudinal axis. The type of nodule formed is specified by the plant host.

**MUTUAL SIGNAL EXCHANGE STARTS THE PROCESS**

The early plant host signals secreted into the rhizosphere can be (iso)flavonoids, stachydrines, or aldonic acids. Best studied are the flavonoids that, in conjunction with the rhizobial NodD transcriptional activator, induce expression of the nod gene regulon. In turn, nod gene products synthesize and transport the Nod factor (Fig. 1B), the major early signal molecule perceived by the host plant. A flexible interaction with the host is enabled by several mechanisms. Alternative NodD activators recognizing different plant flavonoids provide an extended host range in some bacterial strains. *Bradyrhizobium japonicum* even has an alternative two component regulatory pathway for activating its nod regulon, and in *Sinorhizobium meliloti* nod gene expression is fine-tuned by positive- and negative-control circuits.

Nod factors are low-Mr 1-4-linked N-acetylglucosamine compounds (Lerouge et al., 1990) typically carrying a fatty acid on the non-reducing sugar and sulfuryl, fucosyl, mannosyl, or arabinosyl groups at the reducing terminal sugar. Additional substitutions include carbamoyl, glycerol, and fucosyl derivatives (Dénaire et al., 1996; Spaink, 1996). When purified and applied in the absence of bacteria these lipochitoooligosaccharides (LCOs) function as mitogens or “morphogens” on some legume roots. For example, addition of nanomolar concentrations induces root hair deformation in most legumes. In more responsive plants, pre-infection threads (cytoplasmic bridges in G2 arrested cortical cells), cortical cell divisions, and empty nodule structures with an anatomy comparable to rhizobial-induced nodules also develop (Dénaire et al., 1996; Spaink, 1996, and references therein). These responses show that some legumes encode all functions necessary to develop the nodule once the process has been switched on. Spontaneous nodule development on certain alfalfa mutants grown axenically supports this idea.

*Rhizobium* strains differing in their repertoire of nod genes produce LCOs with different structural features. Biological activity is determined by the length of the chitin backbone, the structure of the lipid, and a suite of other substitutions on the oligosaccharidic backbone. Host specificity results at least partly from

---

**Figure 1.** A, Whole mount of a developing *Lotus japonicus* root nodule invaded by two infection threads. Fluorescence microscopy shows the autofluorescent nodule cells and the *Mesorhizobium loti* bacteria stained for β-galactosidase activity expressed from a lacZ reporter gene. The determinate root nodules of *L. japonicus* are initiated from cell division in the outer cortex, whereas indeterminate nodules (of pea, for example) are initiated from cell divisions in the inner cortex. Conditions or factors influencing nodule initiation or development are listed to the right and left of the root nodule. Where different responses to the factor or condition have been reported in individual legume species, this is indicated with the following: +, positive effect; −, negative effect; or 0, lack of response. EPS, Exopolysaccharides; LPS, Lipopolysaccharides. B, Structure of a bacterial LCO Nod factor. The acetylated fucosyl in blue results from NodZ and NolL modification of the *R. leguminosarum* LCO.
a two-way communication where both plant flavonoid activation of nod genes and plant perception of the resulting bacterial LCO signal molecules are required for nodulation to occur (Fig. 2A). For example, R. leguminosarum bv viciae can be genetically modified to produce an acetyl fucosylated LCO under control of a flavonoid-independent NodD activator (Figs. 1B and 2B). This modified LCO is now recognized by L. japonicus plants resulting in nodulation by the nod gene deregulated R. leguminosarum strain (Bras et al., 2000).

### SIGNAL RELAY IN NOD-FACTOR PERCEPTION

Bacterial LCO signals are first perceived by epidermal cells in the zone of root hair differentiation but the earliest responses to Rhizobium or to the addition of LCO has mainly been studied in more developed root hairs protruding from the root surface where they are easily accessible for microscopical and physiological analysis. Using a combination of electrophysiology and fluorescence microscopy with ion sensitive dyes, it was shown that LCO-induced ion fluxes precedes root hair deformation (Fig. 3). A rapidly induced Ca\(^{2+}\) influx and a transient plasma membrane depolarization associated with Cl\(^-\) and K\(^+\) effluxes occur within seconds. This is accompanied by alkalization of root hair cytoplasm and after some minutes of Ca\(^{2+}\) oscillation (Ehrhardt et al., 1996; for mechanism, see Felle et al., 1998). Subsequently rearrangement of actin filaments and redirection of root hair tip growth is observed. Although root hairs that remain uninfected (or are not competent for rhizobial infection) may also respond, it is conceivable that these rapid physiological changes are transmitted into a signal transduction pathway leading to activation of genes regulating nodulation. Absence of Ca\(^{2+}\) oscillation in the alfalfa MN-NN1008 non-nodulating mutant, and the overlap of LCO chemical structure requirements needed for activity on the plant and for eliciting changed root hair physiology favors this interpretation.

Pharmacological experiments using agonists and antagonists [e.g. mastoparan, pertussis toxin, EGTA, 2,5-di(t-butyl)-1,4-benzohydroquinone, and La\(^{3+}\)] of possible signal transduction components suggested that small trimeric GTPases together with phospholipase C and phosphoinositides are involved (Pingret et al., 1998). The causal relation between the various physiological changes now needs to be established and related to activation of downstream plant genes. For this work plant symbiotic mutants defining genetically separable steps and root hair expressed genes will be useful tools. It is interesting that a recently cloned LjCbp1 gene encoding a putative Ca\(^{2+}\)-binding protein is expressed in an LCO-dependent fashion in root epidermal cells (Webb et al., 2000).

### THE ELUSIVE NOD FACTOR RECEPTOR

The plant preference for particular rhizobial partners as well as the low operational concentrations and structural specificity of bacterial LCO signals suggest that signal perception is mediated by a plant receptor. At present a bona fide receptor for binding the LCO and amplifying the bacterial signal has not been identified. Two binding sites (NFBS1 and NFBS2) were described in microsomal fractions from alfalfa roots and tissue culture cells. NFBS1 has low substrate affinity, whereas NFBS2 binds LCO with higher affinity (Gressent et al., 1999). Both sites have specificity for LCO but will also bind derivatives without the sulfuryl substitution needed for in vivo activity. This discrepancy could result from a receptor that in vitro was depleted of a subunit adding specificity.

A novel lectin type protein was recently purified from Dolichos biflorus and shown to bind LCO with high affinity (Etzler et al., 1999). This lectin also has apyrase (nucleotide phosphohydrolase) activity, sug-

---

**Figure 2.** Schematic representation of the early communication between legume host and Rhizobium. A, Symbiotic development is initiated only when the correct plant and bacterial signal molecules are synthesized, presented, and perceived. B, Changes of the genetic repertoire of R. leguminosarum bv viciae leads to the synthesis of different LCOs and changes of host specificity. Stepwise addition of a flavonoid-independent NodD activator plus NodZ fucosyl transferase and NolL acetyl transferase allows R. leguminosarum bv viciae to nodulate L. japonicus. –, No nodulation; +, slow nodulation; +++, normal nodulation.
gesting a function at the start of a signal transduction (phosphorylation) pathway and thus adding a new twist to the study of lectins. Expression of pea or soybean genes encoding seed lectins with non-catalytic sugar-binding sites was previously shown to extend the host range of clover and Lotus corniculatus. This effect was attributed to the lectin’s ability to attach sufficient bacteria at the root hair tip prior to infection or enhancement of the LCO induced mitogenic response rather than being mediated by an LCO-binding/receptor function (van Rhijn et al., 1998; Díaz et al., 2000).

The question concerning the nature and localization of the Nod factor receptor is still open, leaving room for alternative models (Hirsch, 1992; Ardourel et al., 1994; Schultz and Kondorosi, 1998). Intriguing results from studies with bacterial mutants, non-specific LCOs, and chitin-oligomers suggest a perception mechanism using two receptors of different stringency or a receptor with a complex substrate interaction. Simple LCO derivatives (O-acetylated chitin oligosaccharides) can, for example, induce cortical cell divisions after microtargeting into Vicia roots and in soybean, transient induction of the early nodulin Enod40 gene can be obtained with an unsubstituted chitin pentamer. This points toward a hierarchy of plant responses requiring different LCO specificity for execution and a mechanism where LCO transport, localization, and receptor affinity may be influenced by plant chitinases or glycosyl hydrolases can be envisaged.

Uncoupling of cell division and infection thread formation seen for example in sym5 mutants of pea and sym4 of sweetclover could lend support for a two receptor model, but the abundance of generally unresponsive plants mutants indicate that common steps or interactive pathways are involved. Future map-based cloning of loci believed to be involved in LCO perception, for example, Sym2 from pea and Sym5 from L. japonicus are likely to clarify some of the questions and may also explain why infection threads are only observed in the presence of bacteria.

DEVELOPMENTAL CROSS TALKING: CHECKS AND BALANCES

Development of functional root nodules relies on synchronized activation of particular gene sets in both symbionts. The stage-characteristic arrest observed with rhizobial as well as plant mutants indicates that the processes are highly coordinated, but information on “late” signal exchange is somewhat sparse. However, continued expression of nod genes in bacteria contained within infection threads and localization of internalized immunoreactive LCO in cells of maturing root nodules, indicate a connection to the early LCO signaling. After endocytosis, the synthesis of LCOs is down-regulated in bacteroids. Endocytosis and bacterial differentiation appear therefore to mark a shift in plant-rhizobial communication.

Bacterial surface polysaccharides are known to be involved in the infection process. In some symbiotic interactions, rhizobial mutants that are deficient in exo- or lipo-polysaccharides (EPS I, EPS II, and LPS) are defective in the infection process and may provoke increased host defense reactions suggesting that surface polysaccharides shield the bacteria. However, EPS mutants are partially rescued by exogenous application of picomolar concentrations of low-M_r polysaccharide fractions of EPS I or EPS II, suggesting that these function as signal molecules (González et al., 1996).

Secreted proteins also contribute to signaling. Strains of R. leguminosarum bv. viciae excrete NodO, a protein shown to form Ca^{2+}-transporting ion channels in vitro. In vivo a NodO-mediated enhancement of infection thread progression and nodulation was observed in partially compatible pea hosts. The NGR234 strain has a type-III protein secretion system known from pathogenic bacteria to export proteinaceous pathogenicity factors. Mutations preventing secretion of the NGR234 NolX and y4xl proteins (functions unknown) change the nodulation pattern.
on some but not all legume hosts, implying that aspects of protein signaling are shared between pathogens and symbionts (Viprey et al., 1998). It will be interesting to determine the targets of these “late” signals and address the possibility of positive roles in nodule development as well as defense avoidance.

SECONDARY SIGNALING THROUGH PHYTOHORMONES

Most of the plant architecture is formed by post-embryonic development, and changes in relative hormone concentrations under the influence of biotic and abiotic factors strongly influence the developmental fate of cells and organs. Formation of root nodules is no exception, and several lines of evidence suggest a role for phytohormones in secondary signaling. Incubation with auxin transport inhibitors results in development of empty nodule-like structures on roots of some legumes and expression of results in development of empty nodule-like structures. In Medicago truncatula ethylene-insensitive sickle mutant from Sesbania and alfalfa, the LCO responding genes Enod12 and Enod40 as well as the Enod2 gene are all up-regulated by cytokinin, and cell divisions are activated in these roots. Cytokinin and LCO may thus be part of or influence the same signal transduction during nodule development.

Although there are differences in legume responses, exogenous application of ethylene generally influences nodule formation, and agents inhibiting ethylene biosynthesis or perception [Iα-(2-aminoethoxyvinyl)-glycine and Ag+] increase nodule formation. A drastic increase in persistent infections and numbers of bacteria developing in the Medicago truncatula ethylene-insensitive sickle mutant shows that ethylene is involved in a local regulation of infection (Penmetsa and Cook, 1997). However, a comparable ethylene-insensitive mutant of soybean was unaffected in nodulation, illustrating the difficulties of these hormonal investigations without a firm knowledge of the involved mechanisms (Schmidt et al., 1999). Alfalfa has indeterminate nodules, whereas soybean has determinate ones. One could speculate that the ethylene-mediated switch from indeterminate to determinate nodules observed on Sesbania is another manifestation of a differential response in the two nodule types. In pea roots, nodules develop preferentially opposite protoxylem poles. Localization of ACC oxidase transcripts (and by extrapolation ethylene) in cells between protoxylem poles in combination with a gradient of the uridine “stele factor” from the poles may provide the positional information for this positioning.

NITROGEN CONTROL AND AUTOREGULATION OF NODULATION

Successful nodulation occurs under nitrogen-limiting conditions where a fraction of the invasion events progress into functional nodules. Even under optimal conditions, most infection threads are arrested in hypodermal root cell layers and the actual nodule numbers are limited by the plant. An autoregulatory mechanism enables cortical cell divisions in older nodule primordia to suppress younger cell division foci systematically (Caetano-Anolles and Gresshoff, 1991). Hypernodulating mutants developing excess nodules escape autoregulation and lack the normal nitrate suppression, indicating that nitrate exerts its main local effect via the autoregulatory pathway. One report suggested that ethylene might be involved in nitrate repression in alfalfa, but data from pea did not support a role for ethylene. A model predicting transport of a nodule-derived compound to the shoot and return of an inhibitor was formulated on the basis of grafting experiments demonstrating shoot genotype control of the root phenotype (Caetano-Anolles and Gresshoff, 1991). The identity of these compounds remains as yet unknown, but both nitrate and autoregulation act independently of nutritional effects per se. Recent characterization of the hypernodulation har1 mutant from Lotus suggests that regulation of nodule numbers is integrated in the mechanisms controlling lateral root development (Wopereis et al., 2000).

NUMEROUS GENES ARE ACTIVATED DURING NODULE DEVELOPMENT

Formation of a new organ requires temporally and spatially controlled activity of genes and gene products participating in the organogenic process. Using biochemical or molecular techniques, many proteins (called nodulins) were found at elevated levels in root nodules, and many genes were highly induced in particular cell types or nodule tissues. The earliest expressed genes such as Enod12, Rip1, and LjChp1 are active in root hairs, whereas genes encoding proteins needed in the physiology and biochemistry of the mature active nodule tend to be strongly induced just
before onset of nitrogen fixation. Leghemoglobin, involved in oxygen protection of nitrogenase, is the classical example of this class of “late” nodulins. The complexity of gene regulation during nodulation is also exemplified by the leghemoglobin gene family where some members are already expressed in root hairs and primordial cells (Cvitanich et al., 2000). From the studies of induced genes so far only the Enod40 gene has emerged with a regulatory role in nodule initiation. The Enod40 was suggested to encode a small 12- to 13-amino acid peptide (Franssen, 1998) and a 3′ RNA proposed to control translation efficiency or location. In accordance with a regulatory role, Enod40 is transcriptionally up-regulated in root pericycle cells within hours after inoculation or LCO application and prior to division of cortical cells.

With the numerous expressed sequence tag (EST) sequences now available for various legumes, a comprehensive view of the genes active in root nodules will appear and probably change the nodulin concept. Already a complete overview of the genes (ESTs) is only possible with bio-informatics tools, and readers are therefore referred to databases and Web pages for access and mining of this information (www.kazusa.or.jp/en/plant/lotus/EST/, www.mbio.aau.dk/~chp/, www.bio-SRL8.stanford.edu, http://212.6.137.235/agowa/, and www.ncgr.org/research/mgi).

FUNCTIONAL ANALYSIS OF NODULIN GENES

The challenge is to assign function to genes of the EST inventories and to provide a detailed analysis of key genes governing nodule formation and function. Although similarity to known genes will help, the scope of this task is clear from the present short list showing nodulins studied using molecular genetics (Table I). Techniques for gene knockout have unfortunately proven difficult to establish in plants and functional studies in legumes often rely on sense/antisense studies in the absence of a null phenotype. For example, some plants overexpressing an Enod40 cDNA demonstrated accelerated nodulation, whereas other lines showing cosuppression of the endogenous Enod40 gene, developed modified or fewer nodules (Charon et al., 1999). Studies with the Enod12 gene were less complicated. In a progeny segregating null alleles, it was found that absence of Enod12 did not influence nodule development or function. Either the gene was dispensable or redundancy compensates for the null allele. The recent demonstration that double-stranded RNA interference can selectively reduce gene activity in Arabidopsis (Chuang and Meyerowitz, 2000) may in the future allow more effective in planta studies of gene function in legumes.

<table>
<thead>
<tr>
<th>Gene or Protein</th>
<th>Function</th>
<th>In Planta, Effect, or Phenotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pea lectin, ps1</td>
<td>Mitogen enhancer?</td>
<td>Increases response to unspecific LCOs. Extends host range of clover to include R. leguminosarum by viciae, M. loti, S. meliloti.</td>
</tr>
<tr>
<td>Soybean lectin, Le 1</td>
<td>Cell attachment</td>
<td>Extends host range of L. corniculatus to allow B. japonicum to form empty nodule structures.</td>
</tr>
<tr>
<td>Vicia ENBP1</td>
<td>trans-Factor</td>
<td>Dispensable in alfalfa plants homozygous for null allele.</td>
</tr>
<tr>
<td>Alfalfa Enod40</td>
<td>Cell activation</td>
<td>Cell division in axenic roots and accelerated nodulation.</td>
</tr>
<tr>
<td>LjNin</td>
<td>Regulator of nodule initiation</td>
<td>nin mutants have root hair response but are non-nodulating.</td>
</tr>
<tr>
<td>Mszp2-1</td>
<td>Cell differentiation</td>
<td>Antisense suppression arrests differentiation of nodule central zone cells and inhibits bacterial release.</td>
</tr>
<tr>
<td>Soybean NAT2</td>
<td>trans-Factor</td>
<td>Binds a positive regulatory element and activates promoters in leaf and nodule.</td>
</tr>
<tr>
<td>Soybean CPP1</td>
<td>trans-Factor</td>
<td>Binds the lbc3 minimal promoter.</td>
</tr>
<tr>
<td>Alfalfa ccs52</td>
<td>Mitotic inhibitor</td>
<td>Antisense suppression decreases the number of endoreplication cycles during differentiation.</td>
</tr>
<tr>
<td>Pea rug4</td>
<td>Succ synthase</td>
<td>Ineffective nitrogen fixation.</td>
</tr>
<tr>
<td>GmRab1p, Rab7p</td>
<td>GTP-binding proteins</td>
<td>Involved in biogenesis of peribacteroid membrane.</td>
</tr>
</tbody>
</table>

LEGUME GENETICS WILL IMPROVE ANALYSIS OF REGULATORY MECHANISMS

The regulatory mechanisms controlling nodule development has so far been approached with studies of the early signaling or regulation of nodulin gene promoters. Conceptually these studies start at either end of the developmental process, aiming to meet in between. To fill the gap, a genetic approach offering a more direct identification of central players and assembly of pathways from epistasis studies is now gaining momentum. The potential contribution from plant genetics in gene identification and functional analysis was recently demonstrated by the isolation of the L. japonicus nodule inception (Nin) gene from a symbiotic locus tagged with the maize transposon Ac (Schauer et al., 1999). Symbiotic mutants have been known in soybean and pea for many years, but Nin is the first genetically defined symbiotic locus characterized at the molecular level and an example of the advantages offered by model legumes. Phenotypically, nin mutants are non-nodulating. Cortical cell divisions, the first step to nodule primordia formation,
appear not to be initiated. Like the wild-type plants, nin mutants exhibit root hair deformation after inoculation, indicating that LCO signal perception is functional. These observations point to a Nin function at the junction of signal transduction and gene activation. The presence of putative transcription factor domains in the NIN protein as well as regional similarity to the minus dominance (Mid) protein controlling gametogenesis in *Chlamydomonas* support this idea.

Figure 4 shows the NIN protein structure and outlines a working model for the role of NIN. Based on one mutant, it is obviously insufficient, but the strength of the model lies in the questions raised. Is NIN a DNA-binding central regulator or a modest transcriptional co-activator, and if so which genes are activated? What are the immediate upstream partners, and is NIN itself activated? To approach such questions, a careful investigation of the post-translational regulation of NIN implicated by the predicted transmembrane domains and identification of the first cells expressing active protein will be necessary. Clues may also come from the temporal-spatial expression patterns. The observed expression of *Nin* in uninfected roots makes it possible to test if NIN has a role either upstream or downstream of the LCO-induced auxin response of thenin mutants (Schauser et al., 1999). B, The present resolution of events during nodule inception predicts that this process. This signal may at the same time repress further root hair deformation explaining the excessive root hair development on the nin mutants, together with the apparent lack of cell division in the outer cortex, place the Nin gene in signal transduction or gene activation downstream of LCO perception. Domains in NIN have similarity to transcription factor domains, suggesting a function in activation of genes required for nodule initiation. To explain the lack of infection thread formation in nin mutants, a secondary positive signal is envisaged to be necessary for this process. This signal may at the same time repress further root hair deformation explaining the excessive root hair response of the nin mutants (Schauser et al., 1999).

**Figure 4.** Structure of the NIN protein and a working model for the role in early phases of root nodule initiation. A, The LCO-dependent root hair deformation observed on the nin mutants, together with the apparent lack of cell division in the outer cortex, place the *Nin* gene in signal transduction or gene activation downstream of LCO perception. Domains in NIN have similarity to transcription factor domains, suggesting a function in activation of genes required for nodule initiation. To explain the lack of infection thread formation in nin mutants, a secondary positive signal is envisaged to be necessary for this process. This signal may at the same time repress further root hair deformation explaining the excessive root hair response of the nin mutants (Schauser et al., 1999). B, The present resolution of events during nodule inception predicts that hormonal changes, Enod40, and Nin are involved at about the same developmental stage. Vertical hatching. Putative transmembrane domains; black, acidic activation domains; gray, putative DNA-binding/dimerization domain.

**NODULIN PROMOTERS: CONSERVATION OF REGULATORY ELEMENTS**

In an attempt to unravel the regulatory cascade controlling genes activated during nodule development and at the same time describe putative targets for signal transduction pathways, promoter regions of nodulin genes expressed early and late in the developmental process were characterized. Promoters from the soybean N23 and leghemoglobin *lbc3* genes (Stougaard et al., 1990) together with the Sesbania *Srglb3* (Szczyglowski et al., 1994) were most extensively analyzed. Several cis-acting regulatory sequences were delineated by deletion, hybrid promoter, and point mutational studies of promoters fused to reporter genes. The type and localization of cis-acting elements in the leghemoglobin promoters were remarkably similar as exemplified by the *lbc3* promoter in Figure 5. The *lbc3* promoter has a strong positive element (enhancer), a weak-positive element (WPE), and an organ-specific element containing the highly conserved AAGAT-taTTGT-CTCTT box within a 2-kb promoter region (Ramlov et al., 1993; Szczyglowski et al., 1994).

Trans-acting proteins binding at or in the vicinity of these DNA elements were identified either by direct-binding studies using gel retardation or in South-western hybridization. Two high-mobility group 1-like proteins (NAT1, LAT1) were found in nodules and leaves. A similar nodule NAT2 protein binds at AT-rich sequences bordering the soybean *lbc3* WPE and activates a minimal *lbc3* promoter in nodules as well as the tobacco *rbcS-8B* promoter in leaves. Hence, NAT2 is a general transactivator. Another soybean protein, CPP1, binding to a minimal 1280 *lbc3* promoter contains two Cys-rich domains present in polycomb proteins like *Drosophila* E(z) (Cvitanich et al., 2000). Polycomb proteins usually restrict expression of developmental
regulators, and there is some support for negative regulation of lbc3 expression by CPP1 (Cvitanich et al., 2000). Additional lbc3 trans-factors currently analyzed at the functional levels are homeobox-like proteins and pathogenesis-related activator proteins (Fig. 5).

The pea early nodulin promoter from Enod12B was analyzed by a similar approach. A short −139 promoter region was found to suffice for both nodule expression, activation in LCO-induced primordia, and interaction with an early nodulin-binding protein (ENBP1) protein. When mutated binding site motifs were analyzed for ENBP1 binding in vitro and in vivo a correlation between binding and promoter activity was seen (Hansen et al., 1999), but ENBP1 activity was seen (Hansen et al., 1999), but ENBP1 was not limiting for Enod12 expression in transgenic plants. The ENBP1 protein contains typical AT hooks shown to be required for promoter binding and a zinc finger domain. Both the MsEnod40-1 and MsEnod12A gene promoters respond to LCO and cytokinins, but well-defined regulatory DNA elements operating in phytohormone controlled gene expression were not reported (Bauer et al., 1996; Fang and Hirsch, 1998). Additional promoter analysis may eventually draw a direct link between phytohormone physiology and gene expression during nodule organogenesis.

Soybean lbc3 promoter: weak in nodule primordia, strong in infected cells

Pea ENOD12B promoter, LCO induced tissue specific

Figure 5. Schematic representation of the cis-acting promoter elements and DNA-binding proteins of the soybean lbc3 and the pea Enod12B promoters. The lbc3 strong positive element (SPE) contains three half-sites of an inverted repeat (invXny) suggested to interact with trans-factors, whereas the WPE has two binding sites for the general trans-activator NAT2. Highly conserved regions important for promoter function are located in the organ-specific elements (OSE) and negative elements (NE). The three proteins, Cys-rich polycomb-like protein (CPP1), nodule homeo-domain-like (NDX), and leghemoglobin-binding factor (LBF) bind at least one defined sites of the −280 minimal promoter. In the Enod12B promoter, a short positive element binding the ENBP1 protein is sufficient for LCO tissue specific expression.

PERSPECTIVES AND CONCLUSIONS

The description of the LCO Nod factor molecules and their morphogenic effects on legume roots raised the possibility of a more general role as a hitherto undiscovered class of endogenous plant growth regulators. Now, 10 years later, it is still an open question whether nodule induction is a specialized effect or a unique event. LCOs similar to the compounds synthesized by Rhizobium remain undescribed in plants and the evidence of their existence circumstantial. As a recent example, a study of pea Enod12 promoter activity in rice concluded that the LCO perception mechanism is functional in a monocot plant. (Reddy et al., 1998; for additional observations, see Spaink, 1996). Arabidopsis, which has contributed new firm evidence for the role of brassinosteroids in plant development, has remained unusually silent in relation to LCOs. Genes with similarity to bacterial nod genes so far were not reported from the Arabidopsis genome sequencing program even though a gene with similarity to nodC was found in Xenopus (Spaink, 1996; Schultz and Kondorosi, 1998). Arabidopsis may still contribute but we may also have to await the characterization of the legume LCO receptor(s) and signal transduction pathway(s) before the question can be fruitfully approached in non-legumes. The promiscuous mycorrhiza-plant interaction is another opening for approaching symbiotic functions and genes with a more general role. Several non-nodulating plant mutants are also impaired in the mycorrhizal colonization. This, together with expression of Enod12, Enod40, and Enod2 in roots colonized by mycorrhiza, points at overlaps in the programs governing endosymbiosis. In this broader context comparative genomic analysis of symbiotic genes is bound to add significantly to the understanding of plant development. Taking Nin as an example, several uncharacterized Arabidopsis genes similar to the Nin gene can be identified from the genome sequence and the functional analysis of these genes in Arabidopsis is now an obvious task. This analysis can subsequently be taken back to help in determining the function of Nin and its paralogs in legume development and nodulation.

With the emerging new tools for functional genomics in model legumes, analysis of the plant contribution to symbiosis will gain speed. Transposon and T-DNA tagging together with map-based cloning of symbiotic loci are approaches bound to identify novel genes in the root nodule regulatory system. The combination of sequence information generated by EST sequencing and expression analysis using microarrays or DNA chips will provide information of global gene activities under various conditions, improve mutant characterization, and allow pathway members to be identified from co-expression data. Proteomics will complement this analysis and add to the biochemistry. In the coming years, mechanisms will be added to many of the detailed and careful
descriptive studies of the past. Functional and comparative genomics will be major contributors in elucidation of signaling, gene regulation, and gene function in symbiosis. With a more comprehensive understanding of legume symbiosis, the puzzling question why only *Parasponia andersonii* (Ulmaceae) among the non-legumes, forms nodules with *Rhizobium*, and the equally intriguing question of the genetic differences between legumes and other plants could be approached.

**ACKNOWLEDGMENTS**

I am grateful to the colleagues who generously provided reprints or preprints for this paper. I apologize to the many colleagues whose publications I was unable to include or cite. I thank Dr. Leif Schausser for the pictures used in Figures 1 and 2 and collaborators at the Laboratory of Gene Expression for comments on the manuscript.

Received May 3, 2000; accepted June 22, 2000.

**LITERATURE CITED**


