Acylated chitooligomers are molecular signals that mediate the symbiotic interactions between nitrogen-fixing bacteria and their host plants

Jean-Claude Promé[a], Jean Denarié[b] and Georges Truchet[b]

[a]*Institut de Pharmacologie et de Biologie Structurale, CNRS, 205 Route de Narbonne,F-31077 Toulouse*

[b]*Laboratoire de Biologie Moléculaire des Relations Plantes-Microorganismes INRA-CNRS, Chemin de Borde-Rouge ,F-31326 Castanet-Tolosan*

Abstract: Rhizobia elicit the formation of new organs, called nodules, on their leguminous plant hosts, in which they fix nitrogen. The early steps of this symbiosis are mediated by the bacterial secretion of substituted lipochitooligomers, called Nod factors. Each rhizobial species has a characteristic set of nodulation genes that specifies the length of the chitobackbone and the nature of the substitutions. These structural characteristics are the main recognition elements for mediating plant developmental responses and the bacterial infection. It is likely that this new class of signaling molecules represents also a new family of general plant growth regulators.

INTRODUCTION

Rhizobia are soil bacteria that elicit the formation of nitrogen-fixing nodules on the roots of leguminous host plants. *Leguminosae* is a large family containing more than 15,000 species and these nitrogen-fixing associations have considerable ecological and economical importance. Rhizobia-plant associations are specific as a given bacterium can only nodulate a defined number of plants. However, the degree of specificity varies. Some rhizobia, such as *R. meliloti, R. leguminosarum* bv *viciae* exhibit a narrow host range of nodulation. In contrast, some strains have a broad host range, such as *R.sp NGR 234* that nodulates more than 70 legumes.

The early steps of the symbiosis are mediated by mutual exchanges of chemical signals between the plant and the bacteria. Root plant exsudates are able to activate the transcription of the rhizobial nod genes, and in response, bacteria synthesize and secrete in the external medium molecules called Nod factors that induce morphologic changes of the host plant root hairs , development of meristems in the root cortex and nodule organogenesis.

CHEMICAL NATURE OF NOD FACTORS

Using an alfalfa root hair deformation bioassay, extracellular Nod factors (NFs) were purified from cultures of *R.meliloti* engineered to overexpress *nod* genes (ref. 1). They were extracted by liquid/liquid extraction (butanol) followed by series of chromatographic steps. Purified Nod factors were analyzed by mass spectrometry (FAB ionization) , nuclear magnetic resonnance and chemical analysis and found to be O-sulphated and N-acylated chitooligomers. The chitooligomeric core has 4 glucosamine residues, the sulphate group is on O-6 of the reducing end, while the nitrogen atom at the non reducing end is amide-linked to different fatty acyl groups instead of an acetyl group. Part of the molecules are also 6-O-acetylated at the non reducing end. The most abundant fatty acyl group corresponds to the 2E,9Z-hexadecadienoic acid, but lesser amounts of 2E,4E,9Z-hexadecatrienoic and 9Z-hexadecenoic acids are found, together with series of (ω-1) fatty acids (ref.2).

This LCO mixture was able to induce root hair deformation at concentrations as low as $10^{-12}M$, and nodule formation at $10^{-9}M$(ref.3). Recent plant assays with chemically synthesized LCOs indicate that the sulphated chitotetramer bearing an acetyl and a $\Delta 2,9$-C16:2 fatty acyl group was the far most active product of the mixture (unpublished result). Suppression of the sulphate group by mild acid hydrolysis abolishes root hair deformation and nodule induction. Nodule formation was strongly reduced by hydrogenation of the two double bonds. Suppression of the O-acetyl group slighty diminished the number

of nodules. Opening the reducing end by borohydride reduction decreases the nodulation activity by a factor of 10 (ref.3).
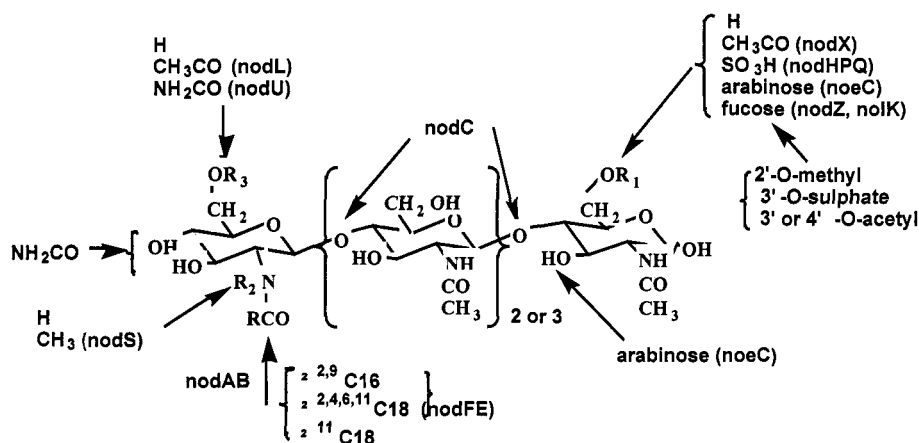
NOD FACTORS DIVERSITY

Nod factors have been isolated from the culture medium of different bacteria of the rhizobium family. All possess a chitin oligomeric backbone which is diversely substituted. Its length varies from 3 to 6 residues. All substitutions are located at both end of the molecules (for reviews, see ref.4 and 5) except for R.galegae which bears a 3-O-acetyl group on the antepenultimate glucosamine residue (unpublished result). At the non reducing end, the most common O-substitutions are 6-O-acetyl , 4- or 6-O-carbamoyl or bis 4,6-O-carbamoyl groups . The amine group may be N-methylated or not, and is always N-acylated by a variety of fatty acids , either "common" ones, as found in membrane lipids, or " specific" ones, only found in these LCOs. The reducing end is often 6-O-substituted either by sulphate, or fucose , or 2-O-methyl fucose (that can be 4-O-acetylated or 3-O-sulphated). In bacteria nodulating the Sesbania plant, an additionnal 3-O-arabinose is present. Finally, although most of Nod factors have a free reducing end, some of them are glycosidically linked to glycerol or to mannose. It must be pointed out that Nod factors are always produced as more or less complex LCO mixtures , differing by the fatty acid nature , the length of the chitobackbone and the presence or the absence of some substituents. As pure compounds are not easy to isolate, it would be somewhat difficult to characterize which are the active components in the mixture. To estimate the role of a peculiar substitution on the biological activity of NFs, genetic transformations of the bacteria were used in order to supress or to introduce a gene encoding for the introduction of a given substitution.

NOD FACTORS AND HOST SPECIFICITY

The biosynthetic pathway of Nod factors was progressively dissected. When mutations inactivate species-specific genes, modified NFs were produced and it was generally found that biological activities of these molecules parallel the symbiotic behaviour of the mutants.

The *nod* genes may be classified into three categories. The regulatory genes (*nod D, SyrM*) encode for regulatory proteins that , in association with inducers from root plants secretions (flavonoïds, betaïns) , promote the transcription of the other *nod* genes. The"common *nod* " genes (*nodABC*) are present in all rhizobium species and play a pivotal role as their inactivation results in the complete loss of the ability to elicit any plant response. The "species-specific" genes are present in different combinations in the different species, and their mutation generally produces alterations of the nodulation host range.

The role of the *nodABC* "common" genes is to specify the synthesis of enzymes building the lipo-chitooligomeric backbone. The NodC protein synthesizes chito-oligomers of up to five residues. The chitooligomers are specifically desacylated at their non reducing end by NodB and then N-acylated by NodA. It seems that the NodC protein, a chitin synthase-like membrane-bound enzyme, is able to control

the length of the chitooligomer, at least partly (up to 5 residues)(ref.6,7), but NodB is also involved to give a more accurate control (R.Geremia, personal communication). Some NodA proteins exhibit a specifity toward the structure of the transfered fatty acid, thus playing a role in the synthesis of specific structures. Thus, allelic variations of the *nodA* and *nodC* genes induce variations in the host specificy of the corresponding bacteria (ref.8).

Sulphation of Nod factors is specified by the *nodHPQ* genes and is required to induce root hair deformation and nodulation of alfalfa, the normal host of *R. meliloti* (ref.9). NodH is a sulfotransferase,and mutation of the corresponding gene induces a shift of the host specificity to another plant vetch, which needs non-sulphated Nod factors to nodulate. NodPQ proteins are involved in the activation of sulphate as PAPS (phospho adenosine phospho sulphate), which is the sulphate donor for the sulfotransferase (ref.10). There is two copies of the *nodPQ* genes in R.meliloti : mutation of one of them results in the extention of the host range to both alfalfa and vetch , and, correlatively, a mixture of sulphated and non sulphated Nod factors is produced because of a reduced amount of activated sulphate. Mutation of both copies produces the same effect as a mutation of nodH : no sulphated NFs are synthesized.

Similarly, it has been found that acetylation of Nod factors at its reducing end , specified by *nodX*, is necessary for nodulation of the Afghan pea by *R. leguminosarum* (ref.11). The transfer of the *nodS* or *nodU* genes in *R.fredii* confers nodulation of *Leucaena* species (ref.12). These genes specify the N-methylation and the 6-O-carbamoylation of the non-reducing end. In *B. japonicum, a* symbiont of soybean, *nodZ* mutants both make NFs lack of methyl fucose and have defective nodulation of siratro (ref.13). However, despite all strains isolated from soybean synthesize NFs with fucose or methyl fucose, this substitution is not necessary for nodulating soybean. Using pure lipochitooligomers (LCOs) produced by chemical synthesis, it was found that root hair deformations of soybean can be induced, either by a10^{-15} M solution of a pentameric LCO acylated by vaccenic acid and fucosylated at the reducing end or by a10^{-12} M solution of a non substituted tetrameric LCO acylated by palmitic acid. Interestingly, neither fucosylated tetrameric nor non -fucosylated pentameric species are active.The formation of nodule primordia needs the same structural requirements (ref.14).

The *nodE* genes of *Rhizobium leguminosarum* bvs. *trifolii* and *viciae* strains are the major determinants of the host range of nodulation. These strains nodulate clover and vetch, respectively. The *nodE* genes (encoding a β-keto-acyl ACP-synthase) direct the synthesis of unsaturated fatty acids with carboxyl-conjugated *trans* double bonds with are part of the Nod factor structures.The only difference in NFs structures produced by the *trifolii* and *viciae* strains is the absence of a remote *cis* double bond in the former, leading to more hydrophobic NFs. It was demonstrated that a central domain of the NodE protein, containing only 44 non-conserved residues, mediates the host specificity by determining the hydrophobicity of NFs (ref.15).

The *nodF* gene codes for an acyl carrier protein. The presence of this protein modifies the fatty acid synthase cycles, lowering or suppressing the activity of the enoyl reductase, and thus giving rise to αβ-unsaturated fatty acids. When the *nodF* gene of *R.meliloti* is mutated, the C16:2 fatty acids are no longer synthesized and the NodA protein is still able (but at a reduced rate) to introduce "common" membrane fatty acids in deacetylated chitooligomers. The production of NFs devoid of αβ-unsaturated fatty acyl groups results in a nodulation delay, but only to a slighty reduced number of nodules. The same phenotype is observed with the *nodL* mutants from which NFs lack the O-acetyl group. However, the double *nodF/nodL* mutants are completely unable to infect and nodulate alfalfa, but they still produce morphologic changes in the root hairs . Cytological examinations reveal that each mutation decreases strongly the number of infection threads by which bacteria penetrate into the plant tissues, and the formation of infection threads is completely abolished with the double mutant. Thus, this mutant cannot infect the plant . This observation let us to propose a multiplicity of NF receptors, some of them involved in recognition, needing an acylated and sulphated chitotetramer, others needing more stringent structural features to allow bacterial entry, such as conjugated double bonds and an acetyl group (ref.16)..

Rhizobium loti and *Rhizobium etli* are two clearly distinct species having distinct host ranges, Lotus and Phaseolus plants, respectively. However, Nod factors from these strains appear identical (4-O-acetyl-fucosyl at the reducing end, 4-O- carbamoyl and N-methyl at the other end) (ref.17,18). The host determinants are *nodD* regulatory genes : a derivative of *R. etli* harbouring a *nodD* gene which activate the transcription of *nod* genes in the absence of flavonoïds efficiently nodulates Lotus plants. It is likely that, in vivo, root secretions of the plant are able to monitor a differential activation of the *nod* genes ,

modifying to some extent the enzyme balance and leading to some discrete modifications in the synthesized NFs. Purified NFs from *R. etli* were much less active in inducing nodule primordia on Phaseolus than crude NFs mixtures . This might indicate a synergistic effect or, alternatively, the loss of minor, biologically very active NFs, during the purification steps (ref.19). Phaseolus plants are nodulated by *R. etli* and by a variety of broad host range strains. NFs from these strains differ in the variety of the acyl groups and the nature of the substitutions. From experiments with purified NFs, it can be concluded that for *P. vulgaris* and for some Acacia plants, the presence of an N-methyl group on a chitopentamer is sufficient to induce nodule primordia .

Nod factors are inducers of plant chitinases. These enzymes are able to hydrolyse the chitooligomeric backbones, but substitutions borne by both ends protect glycosidic bonds located close to the substituted residues. By that way, the plant can modify the composition of the Nod factor mixture (ref.20) and components resistant to chitinase action, even low abundant in the initial mixture, may be those that are the biologically active ones. However, as the earliest plant responses occur within a 5 minutes time range, it seems unlikely that this selection process occurs so fast, but it can be operative for monitoring the subsequent steps of the symbiosis.

RELATIONS OF THE NOD FACTORS STRUCTURES WITH BACTERIA OR PLANT TAXONOMY

Indications that NFs structures are more related to the plant , the signal receiver, than to the rhizobia, the signal donor comes from the analogy of NFs structures produced by phylogenetic distant bacteria nodulating the same plant.. As a striking exemple, *Sinorhizobium Saheli* and *Azorhizobium caulinodans,* quite distant phylogenetically, and nodulating the aquatic plant Sesbania, produce NFs possessing an unique feature : an arabinosyl group located on O-3 at the reducing end. In contrast, two biovars of *S. teranga*, one nodulating Acacia, the other nodulating Sesbania, synthesize NFs with completely different features. The former are sulphated, while the later possess an arabinosyl group. Several other exemples show that NFs structures are related to the taxonomy of the plant host and not to that of the bacterium.

When examining the phylogenetic tree of legumes, it cames out that plants of the *Galegae, Trifoliae and Viciae* tribes are nodulated by bacteria producing NFs with fatty acids having conjugated double bonds. Thus, these bacteria do possess the *nodFE* genes.These plant tribes are supposed to derived from a common galegoïd ancestor. No polyunsaturated fatty acids have been found in NFs from rhizobia that nodulate plants belonging to all other legume tribes not deriving from the galegoïd ancestor.

From an evolutionnary point of view, it can be postulated that archaic legumes possess receptors that do not involve a specific recognition of the N-acyl moiety of NFs. Once in the evolution,in the Galegoïd ancestor, appeared a new requirement (a new receptor) for the presence of conjugated double bonds in the acyl moiety of NFs. This "mutation" allowed a new level of selectivity concerning both the presence and the number of double bonds together with the carbon chain length. Coevolution between bacteria and plants was made possible by allelic variations of the *nodFE* genes and also by allelic variations of *nodA*. It is likely that, when one plant genotype with a given type of NFs receptors is introduced in a soil, and if the adequate rhizobia is not present, an horizontal genetic transfer between all soil rhizobia may occur,thus generating new combination of nod genes and the production of various NFs with the expected structures. Only bacteria possessing the appropriate key will enter the legume.

ARE LIPO-CHITOOLIGOSACHARIDES GENERAL PLANT GROWTH REGULATORS?

Several lines of evidence indicated that LCOs are also able to elicit developmental responses in non-legume plants. Arrested embryo development in a carrot mutant cell line can be rescued by the addition of NFs (ref.21). Expression of the *nodAB* genes in transgenic tobacco changes leaf morphology, suggesting that chito-oligomers exist in plants and that their modification influences plant development (ref.22). A more direct proof that LCOs can act as growth regulators in nonlegumes was presented recently (ref.23). It was found that the growth of tobacco protoplasts in the absence of phytohormone could be compensated for by the addition of synthetic LCOs to the medium.Whereas auxin, in the presence of cytokinine, promotes cell division in the micromolar range, some LCOs are able to stimulate auxin-independent growth of protoplasts at femtomolar concentration. Surprisingly, the most active LCOs are chitotetramers

N-acylated on the non reducing end by C18:1 fatty acids with a remote double bond in the *trans* configuration (9E or 11E), whereas the *cis* analogues and the saturated ones are about 10^4 less active. Moreover, LCOs with a trans double bond at 10^{-13} M also stimulates division of tobacco protoplasts in the absence of both auxin and cytokinin . Thus LCOs may be considered as powerful plant growth regulators and may be related to as an yet unidentified family of endogenous plant signals.

CONCLUSION

The results of extensive research over the recent years has clearly established that chitooligomers play major roles in plant-symbiont interactions. The ways by which plants recognize these molecules is one of the most important challenges for the next years. Thresholds for plant responses are as low as 10^{-15}M . One of the first responsive event is a rapid membrane depolarization (ref.24). Plasma membrane depolarization may lead to an enhanced calcium influx through voltage-dependent calcium channels. Are such increases in the cytoplasmic calcium involved in the transduction of the oligosaccharidic signal? The resulting events induced by activation of the cell machinery are complex. It is clear that some of these LCOs can act as powerful plant growth regulators and evidence is increasing that they have yet unidentified endogenous analogues in plant. It is a fascinating task for the future to understand how plants can give appropriate responses to these stimuli.

REFERENCES

1. P. Lerouge, P. Roche, C. Faucher, F. Maillet, G. Truchet, J. C. Promé and J. Denarié, *Nature*, **344**, 781-784 (1990)

2. N. Demont, F. Debellé, H. Aurelle, J. Denarié and J.C. Promé, *J. Biol.Chem.* **268**, 20134-20142 (1993)

3. G. Truchet, P. Roche, P. Lerouge, J. Vaysse, S. Camut, F. de Billy, J.C. Promé and J. Dénarié, *Nature*, **351**, 670-673 (1991).

4. J. Dénarié, F. Debellé and J.C. Promé, *Annu. Rev. Biochem.*, **65** , 503-535 (1996)

5. J.C. Promé and N.Demont, *Plant Microbe Interactions* eds G. Stacey and N.T. Keen, Chapman and Hall, **1**, 272-307 (1996)

6. P. Roche, F. Maillet, C. Plazanet, F. Debellé, M. Ferro, G. Truchet, J.C. Promé and J. Dénarié, *Proc;. Natl. Acad. Sci. USA*, **93**, 15305-15310 (1996)

7. E. Kamst, J. Pilling, L.M. Raamsdonk, B.J.J. Lugtenberg and H.P. Spaink, *J. Bact.* **179**, 2103- 2108 (1997)

8. F. Debellé, C. Plazanet, P. Roche, C. Pujol, A. Savagnac, C. Rosenberg, J.C. Promé and J. Dénarié, *Mol.Microb.* **22**, 303-314 (1996)

9. P.Roche, F.Debellé, F.Maillet, P.Lerouge, C.Faucher, G.Truchet, J. Dénarié and J.C. Promé *Cell*, **67**, 1131-1143 (1991)

10. J. Schwedock, S.R. Long, *Nature*, **348**, 644-647 (1990)

11. J.L.Firmin, K.E.Wilson, R.W.Carlson, A.E.Davies and J.A.Downie , *Mol Microbiol.*, **10** , 351- 360 (1993)

12. S.Jabbouri, R.Fellay, F.Talmont, P.Kamalaprija, U. Burger B.Relic, J.C. Promé and W.J.Broughton, *J. Biol. Chem.*, **270**, 22968-22973 (1995)

13. G.Stacey , S. Luka, J.Sanjuan Z.Banfalvi, A.J.Nieuwkoop, J.Yoon Chun, L.S.Forsberg and R. Carlson, *J.Bact.*, **176**, 620-633 (1994)

14. T.J.W.Stokkermans , S.Ikeshita , J.Cohn, R.W.Carlson, G.Stacey, T. Ogawaand, N.K.Peters , *Plant. Physiol.*, **108** , 1587-1595 (1995)

15. G.V.Bloemberg, E.Kamst, M.Harteveld, K.M.G.M.van der Drift, J.Haverkamp, J.E.Thomas- Oates, B.J.J.Lugtenberg and H.P. Spaink *Mol. Microbiol.*, **16** , 1123-1136 (1995)

16. M.Ardourel, N.Demont, F.Debellé, F.Maillet, F.de Billy, J.C.Promé, J.Dénarié and G.Truchet, *.Plant Cell*, **6**, 1357-1374 (1994)

17. R.Poupot, E.Martinez-Romero, N.Gautier and J.C. Promé, *J. Biol. Chem.*, **270** , 6050-6055 (1995)

18. I.M.Lopez-Lara, J.D.J.van den Berg, J.E. Thomas-Oates, J.Glushka, B.J.J. Lugtenberg and H.P.Spaink *Mol. Microbiol.*, **15**, 627-638 (1995),

19. L.Càrdenas, J.Dominguez , C.Quinto, I.M.Lopez-Lara, B.J.J.Lugtenberg, H.P.Spaink, G.J.Rademaker, J. Haverkamp and J.E.Thomas-Oates , *Plant Mol. Biol.*, **29**, 453-464 (1995)

20. C.Staehelin, M.Schultze, E.Kondorosi, R.B.Mellor, T.Boller and A.Kondorosi , *The Plant J.* , **5**, 319-330 (1994)

21. A.J.de Jong, R.Heidstra, H.P.Spaink, M.V.Hartog, E.A.Meijer, T.Hendriks, F.Lo Schiavo, M. Terzi , T.Bisseling, A.Van Kammen and S.C.De Vries, *Plant Cell,* **5**, 615-620 (1993).

22. J.Schmidt, H.Röhrig, M.John, U.Wieneke, G.Stacey, C. Koncz and J.Schell, *Plant J.*, **4**, 651- 658 (1993)

23. H.Röhrig, J.Schmidt, R.Walden, I.Czaja, E.Miklasecics, U. Wieneke, J.Schell and M.John, *Science* , **269**, 841- 843 (1995)

24. H.H. Felle, E. Kondorosi, A. Kondorosi, M. Schulze, *Plant.J.*, **7**, 939-947 (1995)