Role of N-glycosylation Sites and CXC Motifs in Trafficking of Medicago truncatula Nod Factor Perception Protein to the Plasma Membrane

Benoit Lefebvre1,2, Doerte Klaus-Heisen1,2, Anna Pietraszewska-Bogiel3, Christine Hervé1,2, Sylvie Camut1,2, Marie-Christine Auriae1,2, Virginie Gasciolli1,2, Alessandra Nurisso4, Theodorus W. J. Gadella3, and Julie Cullimore1,2

From 1INRA, Laboratoire des Interactions Plantes-Microorganismes (LIPM), UMR441, F-31326 Castanet-Tolosan, France, 2CNRS, Laboratoire des Interactions Plantes-Microorganismes (LIPM), UMR2594, F-31326 Castanet-Tolosan, France, 3Swammerdam Institute for Life Sciences, Amsterdam 1098 SM, The Netherlands, 4School of Pharmaceutical Sciences, University of Geneva, University of Lausanne, CH-1211, Geneva, Switzerland

Running head: Structure-Function of Medicago truncatula NFP

Background: Nod Factor Perception (NFP) protein is a plant, lysin motif receptor-like kinase.

Results: Disulphide bridges that connect the 3 extracellular lysin motifs, and the intracellular dead-kinase domain are essential for NFP function.

Conclusion: Post-translational modifications are required for NFP folding, trafficking and functioning.

Significance: Structural information will help to determine NFP biochemical function.

ABSTRACT: The lysin motif receptor like kinase, NFP, is a key protein in the legume Medicago truncatula for the perception of lipochitooligosaccharidic Nod Factors, which are secreted bacterial signals essential for establishing the nitrogen-fixing legume-rhizobia symbiosis. Predicted structural and genetic analyses strongly suggest that NFP is at least part of a Nod factor receptor, but few data are available about this protein. Characterization of a variant encoded by the mutant allele nfp-2 revealed the sensitivity of this protein to the endoplasmic reticulum quality control mechanisms, affecting its trafficking to the plasma membrane. Further analysis revealed that the extensive N-glycosylation of the protein is not essential for biological activity. In the NFP extracellular region, two CXC motifs and two other Cys residues were found to be involved in disulphide bridges and these are necessary for correct folding and localization of the protein. Analysis of the intracellular region revealed its importance for biological activity but suggests that it does not rely on kinase activity. This work shows that NFP trafficking to the plasma membrane is highly sensitive to regulation in the endoplasmic reticulum and has identified structural features of the protein, particularly disulphide bridges involving CXC motifs in the extracellular region that are required for its biological function.

The establishment of the nitrogen-fixing symbiosis between legume plants and rhizobia bacteria requires a molecular dialog between the host and the microsymbiont. Secretion of rhizobial lipochitooligosaccharidic signals, called Nod factors (NF), and their perception by the host plant are necessary for activating the plant root nodule organogenesis and infection programmes and these factors are also involved in the partner specificity between the legume species and the bacterial strains. Successful interaction leads to development of infected root nodules (nodulation) in which the bacteria fix dinitrogen.

Genetic analysis has identified genes involved in NF perception and signal transduction. NFP of Medicago truncatula (3), and its orthologues NFR5 of Lotus japonicus and SYM10 of pea (4,5) are required for all NF responses and for nodulation and infection. Another gene, LYK3, is also required for nodulation and infection in M. truncatula, whereas in L. japonicus, a similar gene, NFR1, is required, like NFR5, for all NF responses (5,6).

NFP, LYK3 and their orthologues are
Receptor-Like Kinases (RLKs), belonging to the lysin motif (LysM-RLK) class. LysM-RLKs occur in all plants and phylogenetic analysis has defined two sub-families which have been called the LYT and LYK families (3). In *M. truncatula* there are at least 17 LysM-RLKs and NFP and LYK3 are members of the LYR and LYK sub-families respectively. Like all RLKs (7), the LysM-RLKs are type I membrane proteins with a predicted signal peptide, an extracellular region, followed by a single transmembrane-spanning helix and an intracellular region (IR) exhibiting homology to a protein kinase. The lack of certain highly-conserved kinase features and the lack of *in vitro* kinase activity of the IR of NFP (3) and NFR5 (8) suggest that these proteins and indeed many LYR proteins may be part of the large number of plant RLKs with “dead kinases” (9).

In contrast LYK3 and NFR1 have active kinases, which are necessary for their symbiotic roles (3,8,10).

All the LysM-RLKs are predicted to encode proteins with three LysM domains in their extracellular regions which are separated by characteristic CXC (Cys- any amino acid- Cys) motifs in the inter-domain spacer regions. A similar structure also occurs in related LysM receptor-like proteins (LYM proteins) which lack an IR (11). LysM domains are protein motifs of about 40 amino acids (AA) which were first described in bacterial autolysins but indeed are found in many eukaryota and bacteria proteins, often in association with other domains (12). Only in plants are they associated with kinase domains (12). LysM domains are implicated in the binding of GlcNAc containing molecules. However, although genetic analysis implicates the symbiotic LysM-RLKs of legumes in binding of lipochitooligosaccharide NFs (1,13,14), this has not yet been demonstrated biochemically.

We previously showed that NFP is highly N-glycosylated (15) and of the two mutant alleles of *NFP* that have been identified (3), *nfp-1* appears to be a null allele, whereas *nfp-2* bears a mutation (S67F) located in a putative N-glycosylation site in the first LysM domain. In addition the presence of the highly conserved pairs of CXC motifs between the LysM domains in LysM-RLKs, suggests that they may play a structural role in the protein, perhaps through the formation of disulphide (S-S) bridges.

N-glycosylation and S-S bridges are post-translational modifications (PTMs), which occur in many eukaryotic proteins exposed to the extracellular medium (secreted and surface membrane anchored or transmembrane proteins such as NFP). S-S bridges also occur in prokaryotes, but to a lesser extent than in eukaryotes. N-glycosylation and S-S bridges have been reported to play many roles, including facilitating the folding, trafficking and function of the protein as well as protecting it from an extracellular medium, which, in plants, is acidic and rich in proteases (16). These PTMs have recently been shown to be important for trafficking to the plasma membrane (PM) and/or the functioning of several plant RLKs of the leucine rich-repeat (LRR-RLKs) class (17). Both modifications occur in the endoplasmic reticulum (ER). S-S bridges are formed between two Cys residues, due to the favorable redox condition and with the help of protein disulphide isomerases. N-glycosylation, initially with mannose rich N-glycans, occurs on Asn residues at specific sites which consist of the consensus sequence Asn-X-Ser/Thr, where X is any AA except Pro and is mediated by the oligosaccharyltransferase complex (16). The N-glycans then undergo cycles of trimming and elongation which lead to dissociation / association with the ER luminal lectin chaperones calnexin / calreticulin (CNX/CRX) (18). N-glycans are further modified in the Golgi apparatus to give complex N-glycans, which in plants are often fucosylated on the proximal GlcNAc residue (19). In addition to CNX/CRX, BIP and its partners are other ER non-lectin luminal chaperones that bind hydrophobic exposed domains of unfolded proteins. All together the luminal and the cytosolic chaperones retain the nascent proteins in the ER until they are correctly folded. Long retention of unfolded proteins increases the probability for the protein to be targeted for degradation by the ER associated degradation (ERAD) pathway or eventually by the vacuole (20). This represents the ER quality control (ER-QC).

The aim of this work was to identify the structural features of NFP required for its activity. To this end, we performed a large-scale structure function analysis using a combination of genetic complementation and biochemical approaches. In particular, we examined the role of PTMs in the extracellular region and conserved residues in the intracellular region. The work presented here contributes to our understanding of the structure and regulation of plant LysM-RLKs in signal transduction and to the functioning of NFP in nodulation.
EXPERIMENTAL PROCEDURES

Cloning and Plant Transformation - A cloning vector was built containing an expression cassette consisting of a HindIII site, the cauliflower mosaic virus 35S promoter (Pro35S) or the NFP promoter, -1137 to – 1 bp before ATG (ProNFP), a BgIII site, the coding region of NFP fused in frame to a protein tag, an EcoRI site, the Nos terminator and a Smal site. The protein tags used were the yellow fluorescent protein, sYFP2 (21), monomeric red fluorescent protein, mRFP (22) and 3xFLAG (Sigma). A construction of NFP deleted from its intracellular region and fused to mRFP (NFPΔIR-RFP) contained amino acids 1-283 of NFP. Point mutations were introduced using the quick change mutagenesis kit (Stratagene) using the NFP-RFP cloning vector for the mutations in the NFP extracellular region. The NFP-3xFLAG cloning vector was used for mutations in the NFP IR. The expression cassettes were then transferred to the binary plasmid pBin+ using HindIII and Smal. Plasmids containing PMA4-GFP (23), HDEL-GFP (24), HVR-ROP-NFP extracellular region. The NFP-3xFLAG cloning vector was used for the mutations in the NFP IR. The expression cassettes were then transferred to the binary plasmid pBin+ using HindIII and Smal. Plasmids containing PMA4-GFP (23), HDEL-GFP (24), HVR-ROP-mTurquoise (25,26) are as described.

The resulting plasmids were introduced into Agrobacterium rhizogenes (ARqua1) or into Agrobacterium tumefaciens (LBA4404) by electroporation. M. truncatula kanamycin resistant roots were produced on nfp-1 plantlets (3), essentially as described (27). After four weeks on agar plates supplemented with 20µg/ml kanamycin, the composite plants were transferred to growth pouches. One week later, the root systems were inoculated with Sinorhizobium meliloti (strain 2011). Nodules were counted at 10 and 14 days post inoculation (dpi) and the root systems were then harvested.

Tunicamycin Treatment - Leaf discs of A. tumefaciens infiltrated N. benthamiana leaves were incubated by flotation for 20h on 10µM tunicamycin (stock 5mM of tunicamycin (Sigma) in DMSO, diluted in water) prior to microscopy analysis. Leaf discs were then stored at -80°C before immunoblotting analysis.

PNGaseF Treatment - PNGaseF treatment was performed on microosomal fractions or on denatured total extracts from transgenic roots. For fractionation, frozen roots were ground for 30s with a 4mm metal bead in 2ml tubes. The powder was diluted in 500µl of 250mM Sorbitol, 50mM Tris-HCl, pH 8.0. 2mM EDTA, 0.6% polyvinylpolypyrrolidone, 5mM DTT, protease inhibitors (1mM phenylmethylsulfonyl fluoride and 1mg/mL each of leupeptin, aprotinin, antipain, chymostatin, and pepstatin, Sigma). 350µl of 0.8mm glass beads were added and samples were re-ground for 90s. Samples were centrifuged for 5min at 10 000g and then the supernatant was re-centrifuged for 30min at 100 000g. For denaturation of total extracts, transgenic roots were solubilized in gel loading buffer (see below) and diluted in water. Proteins were then precipitated by 10% trichloroacetic acid (TCA). The pellets were washed twice with 90% acetone. Microsomal fractions and TCA pellets were resuspended by sonication in 45µl 25mM Tris-HCl pH 7.0. Either 5µl PNGaseF (Roche) or 5µl of 50% glycerol were added. The samples were incubated for 30min at 37°C prior to immunoblotting.

Cysteine Labeling - Microsomal fractions were prepared from 10g of transgenic N. benthamiana leaves as described (29) and resuspended in 2ml denaturing buffer (6M guanidinium HCl, 0.1M Tris-HCl, pH 8.0). All steps were performed under nitrogen. The samples were divided in two; 100µl of freshly-prepared 1M DTT or water were added, before 1h incubation at room temperature. Microsomal fractions were then diluted, pelleted, washed and finally resuspended (450µl) in denaturing buffer. 50µl of Maleimide-PEG2-Biotin (MPBiotin) (Pierce, 10mM in DMSO) was added before 1h incubation at room temperature. Microsomal fractions were then diluted, pelleted and resuspended (200µl) in pull down buffer (150mM NaCl, 25mM Tris-Cl pH 7.5, 10% glycerol). 200µl pull down buffer with 0.5% Dodecyl Maltoside (DDM, Alexis Biochemicals) then cultured in BY2 cell culture medium with only kanamycin. Protoplasts were transformed as described (28).
were added. Samples were incubated for 10 min at 4°C, diluted 2.5-fold with pull down buffer and centrifuged for 15 min at 165 000 g. 15 µl of magnetic RFP trap (Chromotek) were added to the supernatant before 1 h incubation at 4°C. Beads were washed twice in pull down buffer before analysis of the proteins by immunoblotting.

**Immunopurification** - 200 mg of *N. benthamiana* transgenic leaves frozen and crushed were ground for 30 s with two 3 mm metal beads in 2 ml tubes. The powder was diluted in 400 µl of pull down buffer with 0.5% DDM, and shaken for 5 min at 4°C. Samples were centrifuged for 5 min at 5000 g. The supernatant was centrifuged for 30 min at 100 000 g. The supernatant was diluted 5-fold with pull down buffer and 10 µl of magnetic RFP trap (Chromotek) were added before 30 min incubation at 4°C. Beads were washed once in pull down buffer with 0.1% DDM before analysis of the proteins by immunoblotting.

**Immunoblotting** - Protein expression in transgenic roots was followed by analysis of root total extracts (20 mg of roots for 30 µl gel loading buffer). Other samples were prepared as described above. All samples were solubilized in gel loading buffer (2% SDS, 80 mM Tris-Cl pH 6.8, 15% glycerol) heated for 5 min at 95°C and centrifuged for 5 min at 16 000 g. After separation by SDS-PAGE on 8% (w/v) acrylamide gels, proteins were transferred onto nitrocellulose membranes. RFP fusions were detected using the serum from a rabbit immunized with an *Escherichia coli* expressed 6xHIS tagged mCherry protein, at a dilution of 1:5 000. YFP fusions were detected using rabbit anti-GFP (Clontech) at 1:5 000 and BIP was detected using mouse anti-Hsc70 (1D9; Stressgen) at 1:1 000, followed by peroxidase-linked anti-rabbit antibodies (Millipore) or peroxidase-linked anti-mouse antibodies (Pierce). Peroxidase-linked goat anti-biotin antibodies (Sigma) was used at 1:1 000. Bovine serum albumen was used for membrane saturation and incubation of antibiotin antibodies. Peroxidase activity was revealed by chemiluminescence using the Immobilon kit (Millipore). Molecular weights were estimated by comparison to the PageRuler prestained protein ladder (Fermentas).

**Immunocytolabeling** - Root segments transformed with the Pro35S:NFP-RFP construct (or the vector construct) were fixed in 4% formaldehyde, 0.1 M phosphate, 0.1% Triton X-100, buffer pH 7.4 for 30 min at room temperature followed by subsequent fixation for 60 min in fresh fixing medium without triton. After rinsing, the specimens were embedded in low melting point wax (30) and cut into 10 µm thick sections which were deposited on stick-on coated slides and de-waxed prior to the immunolocalization procedure. RFP immunolocalization was performed on sections by treating with anti-RFP antibody diluted 1:20 000 as primary antibody and goat anti rabbit-HRP (Santa Cruz) diluted 1:2 500 as secondary antibody according to the Tyramide Signal Amplification (TSA) Plus Fluorescein System (Perkin Elmer). After staining, sections were treated according to (31).

**Microscopy** - Fluorescence of all cells was imaged using a Leica SP2 confocal microscope except that of the protoplasts, imaged using a Zeiss CLSM510.

**qRT-PCR** - Biological samples analyzed and qRT-PCR analyses were performed as described (32).

**In Silico Analysis** - Alignments were made by ClustalW (http://www.ebi.ac.uk/Tools/msa/clustalw2/) and corrected manually. Transmembrane domains were predicted by TMHMM (http://www.cbs.dtu.dk/services/TMHMM/) and signal peptide by signalP (http://www.cbs.dtu.dk/services/SignalP/).

**Molecular modeling** - Coordinates of the three LysM domains, previously built from the NMR structure of a LysM domain from *E. coli* through homology modeling techniques (15), were herein used to build the whole NFP extracellular region (192 AA, from Ile32 to Pro223). Loops between domains were generated by using the Biopolymer module of Sybyl X (Tripos Inc.). The same module was used for the generation of the mutated NFP S67F model. Hydrogen atoms were added and partial atomic charges were derived from the AMBER force field (33). Titratable groups were considered in their standard protonation state at neutral pH. Minimization cycles were then carried out, by first optimizing the hydrogen atoms and sidechains orientation. The whole structure was then fully optimized. Energy minimization cycles were performed within the AMBER force field (33). The permittivity was set as a distance-dependant function and a Powell-type minimizer was used (34). Connolly surfaces were then computed by employing the MOLCAD program, mapping the lipophilic potential properties (35). Images were created within SybylX graphical...
RESULTS

NFP localizes at the plasma membrane – Putative Nod factor (NF) receptors are expected to localize to the PM where they may encounter NFs being produced by extracellular rhizobia. To study NFP localization and trafficking we made YFP, RFP and 3xFLAG fusions to the C-terminus of NFP and transformed wild type (WT) and the nfp-1 (a predicted null allele) mutant of M. truncatula using the A. rhizogenes transformation system. Using the native promoter (ProNFP), we have not been able to detect specific fluorescence (above the autofluorescence level) from the YFP or RFP tagged NFP proteins in roots and nodules, even though endogenous NFP is expressed in both tissues as shown using a ProNFP:GUS fusion (3) and by NFP mRNA quantification by qRT-PCR (supplemental Fig. S1). We thus tested whether the strong and constitutive promoter (Pro35S) could be used and found that all three tagged NFP fusions expressed from this promoter led to complementation of nfp-1 for nodulation, suggesting that the use of Pro35S and the presence of a C-terminal tag do not compromise the biological function of NFP: indeed comparison of the 3xFLAG constructs showed that the Pro35S allows a stronger nodulation of M. truncatula than the ProNFP (supplemental Fig. S2). However, still no specific fluorescence (above the autofluorescence level) could be observed from the Pro35S:NFP-YFP or -RFP fusions in M. truncatula roots.

In contrast, the fluorescent NFP fusions were easily detectable when expressed in heterologous systems: cowpea leaf protoplasts, Nicotiana tabacum BY2 cells (supplemental Fig. S3) and Nicotiana benthamiana leaves (Fig. 1A). In all three systems, localization at the cell periphery was observed. In transformed cowpea protoplasts, the fluorescent fusion of NFP showed co-localization with the PM marker (supplemental Fig. S3A). In N. benthamiana leaf epidermal cells, localization of the protein evolved with time following infiltration with A. tumefaciens containing the construction. Early observations, 36hpi revealed NFP fluorescence mainly around the nucleus and in a cortical network resembling the endoplasmic reticulum (ER) (supplemental Fig. S3B). At 72hpi the protein was predominantly localized at the PM, as shown by co-expression with a PM marker (Fig. 1A). Observations at longer times after infiltration revealed fluorescence mainly in the vacuole (supplemental Fig. S3B). Since fluorescence was homogeneous in the vacuole, it likely corresponds to free RFP released following cleavage of the fusion protein in the vacuole. In the case of stably transformed BY2 cells, fluorescence was observed both at the PM and in the vacuole (supplemental Fig. S3C).

To study the localization of NFP in M. truncatula roots, we used immunocytolabeling with anti-RFP antibodies on A. rhizogenes transformed roots expressing NFP-RFP from Pro35S. Using a classical procedure, we were again unable to detect NFP. However, using a highly sensitive amplification method, we were finally able to detect a signal corresponding to NFP-RFP at the cell periphery in root cortical cells while little or no signal was detected in vector-transformed control roots (Fig. 1B). This method also detected a PM localization of the NFP-RFP and NFP-YFP proteins in N. benthamiana leaves similarly to the localization of the fluorescence of these proteins (data not shown).

Together these results show that NFP localizes at the PM in both M. truncatula roots and in heterologous plant expression systems.

The Number of N-glycosylation Sites is Variable between LysM-RLKs and Individual N-Glycans are not Necessary for NFP Activity – Previous work had shown that NFP is highly N-glycosylated and that all the 10-predicted N-glycosylation sites (NXS/T) are most likely occupied (15). The nfp-2 allele contains a point mutation leading to replacement of Ser67 located in a putative N-glycosylation site in the first LysM domain (LysM1) by Phe and, like nfp-1, is a strong allele showing no NF responses (3). In some receptors, N-glycosylation can affect receptor function by affecting ligand binding and signal transduction (36). In NFP eight of the N-glycosylation sites occur in the LysM domains (Fig. 2A, supplemental Fig. S4) and four of them, including the one affected in the nfp-2 allele, are conserved in NFP orthologous proteins of both pea (SYM10) and L. japonicus (NFR5) (Fig. 2B). However the NFP N-glycosylation sites are not well conserved within the Lyr and Lyk proteins of the LysM-RLK family from M. truncatula, although in some cases other sites are found in close vicinity to those of NFP (Fig. 2B, supplemental Fig. S4). The number of N-glycosylation sites in the LysM domains of these proteins is variable and can be as few as two (in...
lyr2, lyr3 and lyr4), whereas nfp, with eight, contains the most (supplemental fig. s4).

In order to test whether N-glycosylation is required for NFP function, we first treated N. benthamiana leaves expressing NFP-YFP with tunicamycin, a drug inhibiting an early step of N-glycosylation. Immunoblot analysis showed that NFP from the tunicamycin-treated sample migrated at the size predicted from its unmodified AA sequence (fig. 2c), suggesting that the tunicamycin treatment had substantially inhibited NFP N-glycosylation. Microscopy analysis revealed that the protein was located predominantly at the cell periphery at 72hpi, probably at the PM (fig. 2da), although in some cells NFP fluorescence was observed in perinuclear membranes (arrow in fig. 2db), as also seen in control leaves treated only with diluted DMSO.

We then individually replaced the Asn of each potential N-glycosylation site present in the LysM domains by Gln and took advantage of our nfp-1 complementation system using Pro35S:NFP-RFP tagged proteins to study the effect of the mutations. All mutated proteins were able to complement the absence of nfp-1 mutant (table 1), showing that each N-glycosylation site, individually, is not essential for NFP function. We combined some of these mutations in order to remove all N-glycosylation sites from LysM1, LysM2 or LysM3. Again, expression of the mutated proteins restored the nodulation in the nfp-1 mutant (table 1), showing that N-glycosylation of each individual LysM domain is not essential for the biological role of the protein and that the loss of function of nfp-2 is thus not due to the absence of an N-glycan.

NFP trafficking to the PM is regulated in the ER – In order to investigate the reason for the lack of function of the nfp-2 allele, a construct coding for NFP-S67F fused to RFP was used to transiently transform N. benthamiana leaves. In contrast to NFP which localized at the PM at 72hpi, NFP-S67F localized to the periphery of the nucleus and in a reticulated network throughout the cell, and showed co-localization with an ER marker, rather than with the PM marker (fig. 1c).

To study the trafficking of these proteins in M. truncatula we made use of the information that the N-glycans of NFP are mostly insensitive to PNGaseF (15), probably because they are fucosylated on the proximal GlcNAc residue in the Golgi apparatus (19). Therefore, the sensitivity of NFP to PNGaseF treatment could be used to test whether the protein has passed through the Golgi, and presumably inserted into the PM. We thus studied the effect of PNGaseF on different variants of NFP, expressed in M. truncatula roots. NFP showed a small reduction in size following PNGaseF treatment (fig. 3a), probably due to one or two N-glycans not being modified to a complex, fucosylated form, as shown previously (15). In contrast, NFP-S67F was completely sensitive to PNGaseF leading to a protein of the size predicted for unmodified NFP (fig. 3a). This result suggests that its N-glycans were not modified in the Golgi apparatus, consistent with an ER localization of this protein in M. truncatula.

In order to investigate in more details the effect of the S67F mutation, we replaced Ser67 by Ala (S67A) or made a conservative substitution by replacing Ser67 by Thr (S67T), which does not destroy the putative N-glycosylation site. Both mutations allowed complementation of the nfp-1 for nodulation (table 1), showing that the presence of a Phe at position 67 was responsible for the loss of function of the protein coded by nfp-2.

Sensitivity to PNGaseF was different for NFP-S67A (similar to NFP: mostly insensitive) compared to NFP-S67T (partially sensitive) (fig. 3a). In both cases, a fraction of these proteins contained PNGaseF-insensitive N-glycans and thus has passed through the Golgi, coherent with this fraction being located at the PM and restoring nodulation. This result shows that only a small amount of correctly localized NFP is enough to perform its function.

In order to understand how NFP-S67F is retained in the ER, NFP and NFP-S67F expressed in N. benthamiana leaves were solubilized from membranes, purified using the RFP tag and submitted to SDS-PAGE. Antibodies raised against the ER luminal BIP chaperones which bind and retain unfolded proteins in the ER, were used to test whether BIPs interact with NFP or NFP-S67F. Indeed, BIPs were co-purified with NFP-S67F to a much larger extent than with NFP (fig. 3b) suggesting that NFP-S67F is not correctly folded. Misfolded proteins occasionally induce an "unfolded protein response" which leads to over-expression of BIP chaperones (20). We compared the level of BIP in roots from nfp-1 and nfp-2 mutants and wild-type plants grown in aeroponic conditions. No differences in the levels of BIP were detected between the plants (fig. 3c) despite the
accumulation of NFP-S67F / BIP complexes in the ER in nfp-2 plants.

Cysteine Pairs are Conserved in all LysM-RLKs and are Essential for NFP Activity - An intriguing feature of all plant LysM-RLKs characterized to date is the presence of CXC motifs between the three LysM domains. In NFP these occur at Cys102/Cys104 between LysM1/LysM2 and at Cys162/Cys164 between LysM2/LysM3 (Fig. 2A and B). Further inspection of the extracellular regions revealed two other Cys residues (Cys39 and Cys47), separated by seven other residues, present in the N-terminal region of NFP before LysM1 and these residues are clearly conserved in NFR5, SYM10 and LYR1 (supplemental Fig. S4, Fig. 2B). Two Cys residues in similar positions are also found before LysM1 in all M. truncatula LysM-RLKs, although their exact positions and separation (between three and six residues) are less well conserved. Some of these proteins also contain Cys residues in the predicted signal peptide and thus would not be present in the mature protein. It is noteworthy that five of the proteins (LYR2 to LYR6) contain an additional Cys residue in the N-terminal region, after the predicted signal peptide, and also a Cys residue in the LysM2. In addition LYR5 and LYR6 contain a third CXC motif after LysM3 and these two proteins and LYR2 contain two additional Cys residues in this extracellular juxtamembrane region (supplemental Fig. S4). In all cases the LysM-RLKs contain an even number of Cys residues in the extracellular regions of the predicted mature proteins which suggests the presence of S-S bridges.

In order to investigate the presence of such putative S-S bridges, we expressed, in N. benthamiana leaves, an RFP fusion of NFP truncated from its intracellular region in order to eliminate unwanted Cys residues in the IR. This truncated form (NFPΔIR-RFP) contains the six Cys residues from the NFP extracellular region, potentially involved in S-S bridges and one Cys residue in the RFP tag which faces the cytosol and cannot be involved in a S-S bridge. NFPΔIR-RFP localized at the PM similarly to the full length NFP (supplemental Fig. S5A). Membrane fractions were prepared from leaves and free Cys residues were labeled with MPBiotin before and after treatment with DTT, which reduces the S-S bridges. The truncated NFP proteins were then purified using the RFP tag. An immunoblot using anti-biotin antibodies (Fig. 4A) revealed a substantial increase in Cys residue labeling of NFP after the reduction of the S-S bridges: the increase corresponds to the expected difference (1/7) for all the six Cys residues of NFP extracellular region being involved in S-S bridges.

In order to determine the requirement of these S-S bridges in NFP activity, we replaced each Cys residue, individually or in combination by Ala and analyzed the proteins for their ability to complement the nfp-1 for nodulation (Table 1). Among the six Cys residues in the extracellular region, Cys47 and Cys166 were found to be essential for NFP nodulation activity (Table 1). Analysis of combinations of Cys residue mutations revealed that either Cys102 or Cys104 must be present for nodulation activity and that Cys164 is required if Cys104 is mutated (Table 1). The mutated proteins that were unable to complement nfp-1 for nodulation were then tested for their sensitivity to PNGase F and all of them were found to be sensitive (Fig. 4B), suggesting that these proteins contain non-modified N-glycans and that these proteins did not reach the Golgi apparatus. Even the mutated proteins complementing the absence of nodulation in nfp-1 were detected as sensitive to PNGase F (supplemental Fig. S5C), suggesting that these proteins were severely affected in trafficking and that only a small amount of functional protein reached the PM to allow complementation. The ER localization of proteins mutated in the Cys residue is supported by the finding that when the proteins were expressed in N. benthamiana, for example C102/104/164/166A, they localized to regions, similar to the ER marker (supplemental Fig. 5B, Table 1). Further analysis of the protein encoded by NFP mutated in the two CXC motifs showed that it co-purified with BIP chaperones from N. benthamiana leaves (Fig. 4C).

Finally, we deduced from the mutation analysis the positions of the NFP S-S bridges (see discussion for details) and used this information to build a model of the entire NFP extracellular region (Fig. 5A). Stability of the model supports S-S bridges linking Cys39 to Cys104, Cys47 to Cys166 and Cys102 to Cys164. Such S-S bridges pack together the three LysM domains and the N-ter (Fig. 5A).

Together, these data suggest that S-S bridges in the NFP extracellular region are essential for correct folding of NFP and trafficking to the PM and that in their absence the protein is retained in the ER by binding to BIP chaperones.
**NFP Intracellular Region is Necessary for Nodulation** - We recently showed that homology with the kinase domain of human interleukin-1 receptor-associated kinase 4 (HsIRAK4), (of which the structure is known) can be used to compare the structures of the predicted kinase domains of plant RLKs, due to the common evolutionary origins of these plant and animal kinases (10). Alignment of the NFP IR with the kinase domains of HsIRAK-4, LYK3 and an *Arabidopsis thaliana* flagellin receptor (FLS2) confirmed that the NFP IR lacks some essential features which are conserved in active kinases (3), including part of the Gly-rich loop and most of the activation segment (supplemental Fig. S6). However, other residues conserved in kinases, such as the “VAIK” (phosphotransfer) motif and the HRD motif in the catalytic loop, which are involved in ATP hydrolysis, are still present in NFP. Thus, we investigated whether the intracellular region of NFP and the conserved kinase residues are important for its biological function.

We first tested whether the NFP IR is essential for nodulation and found that a construct lacking the IR (NFPΔIR) did not complement *nfp-1* for nodulation. We then tested Ala substitutions in conserved residues in the “VAIK” motif (Lys339) and in the HRD motif (Asp435) and found that both mutant proteins complemented *nfp-1* for nodulation (Table 1). We then tested Ala substitutions in residues that are predicted to be phosphorylated (NetPhos programme) in the juxtamembrane region (T281, S282 and S283), in α-helix F (T482), in the incomplete activation segment (T459, S460 and T461) and near the C-terminus of the protein (T578 and S579). All the constructs coding NFP mutated in these residues complemented *nfp-1* for nodulation (Table 1). Finally we tested the role of Gly474, which is a highly conserved kinase residue in predicted α-helix F and which has been shown to be essential for the activity of *A. thaliana* FLS2 (37). An NFP-G474E mutant protein did not allow complementation of *nfp-1* for nodulation, whereas a protein with a more conservative substitution (G474A) complemented. A comparison of the sensitivity to PNGaseF of the NFP-G474E and NFP-G474A proteins from *M. truncatula* roots suggests that at least a fraction of both proteins has passed the Golgi apparatus and should be located at the PM (supplemental Fig. 5D). This result suggests that the G474E mutant is affected, not only in trafficking, but also in biological function.

**DISCUSSION**

Genetic studies have shown that NFP plays an essential role in lipochitooligosaccharidic NF perception and signaling. In this article, using a structure function approach, we have identified key elements of the protein which are required for folding, localization to the PM and biological activity of the protein.

We used FP fusions to the C-terminus of the protein to follow NFP localization. Studies in three plant heterologous expression systems suggest that NFP follows a typical cycle for a cell surface receptor: synthesis at the ER, transit through the Golgi apparatus, localization at the PM and later internalization and degradation in the vacuole (Fig. 1, supplemental Fig. S3). Although the fluorescent of the fusion proteins was well detected by microscopy when expressed in heterologous systems under the control of a strong promoter (Pro35S), it was not detected in *M. truncatula* roots either when expressed under the control of its own promoter (ProNFP) or the Pro35S. In contrast, RLKs of the lectin family, expressed under the control of the Pro35S, were detected in *M. truncatula* roots using similar microscopy (38). Plasma membrane localization was confirmed in *M. truncatula* roots by immunocytolabeling, but only using a highly-sensitive amplification method. Together these data suggest that the level of NFP is tightly controlled and kept at a relatively low level in *M. truncatula*, by post-transcriptional or post-translational regulation mechanism(s). Nevertheless, NFP fusion proteins expressed from the Pro35S localized to the PM and complemented the *nfp-1* mutant for nodulation, suggesting that NFP fusion proteins are biologically active and thus can be used to monitor the activity of NFP. Moreover the fusion proteins were detected by immunoblotting and we used the resistance to PNGaseF of Golgi-modified NFP N-glycans to follow NFP trafficking in *M. truncatula*. We thus used a combination of microscopy and protein biochemistry to examine whether particular NFP structures play a role in its biological function.

Concerning the intracellular region (IR) of NFP, our data show that it is essential for biological activity as its deletion leads to a protein that is not functional in nodulation but which is still capable of locating to the PM and thus is probably defective in signal transduction. Previous studies had shown that the IR of NFP...
lacks kinase activity (3) and this is reinforced by our observation that mutants in two highly conserved “catalytic” residues (Lys339 in the VAIK motif and Asp435 in the HRD motif) are still able to complement for nodulation (Table 1). In the animal epidermal growth factor receptor kinase family, the kinase-dead ErbB3 which binds the ligand, activates signal transduction through oligomerization and phosphorylation by the kinase-active ErbB2 which does not bind the ligand (39). In our studies of M. truncatula NFP, we have not been able to show transphosphorylation by its potential partner LYK3 (10), even though the LYK3 kinase domain is capable of transphosphorylating another symbiotic protein, PUB1 (29). In L. japonicus, the kinase domain of the LYK3 ortholog, NFR1, has been shown in vitro to transphosphorylate the intracellular region of the NFP ortholog, NFR5 on Ser282 in the juxtamembrane region (8). However, this Ser residue is not essential for nodulation activity of NFR5 in L. japonicus, nor, as shown here of NFP in M. truncatula. Unlike in L. japonicus, where NFR5 and NFR1 are both implicated in the initial steps of nodulation, in M. truncatula LYK3 is only required for later stages of nodulation and infection. Thus for these early steps, NFP may interact with a yet unidentified RLK. Alternatively its dead-kinase domain could be activated by ligand-dependent conformational changes leading to activation of a different signal transduction protein.

Our work has also shown an important role of Gly474 in NFP activity. This residue is important for the activity of another plant RLK (37) and is located in the predicted α-helix F just after the activation segment, which is the region that is most highly conserved between NFP, other plant RLKs and a related human kinase, IRAK-4 (supplementary Fig. S6). This helix is postulated to relay conformational changes leading to substrate-binding. Flexibility due to the Gly residue could be essential for these conformational changes (40).

Concerning the extracellular region of NFP, through studies on the protein encoded by the strong nfp-2 allele, NFP-S67F, we have shown that NFP is exquisitely sensitive to ER-QC. In N. benthamiana the NFP-S67F fusion protein accumulated in the ER and associated with the chaperones BIP (Fig. 3), which recognize misfolded proteins (20). In M. truncatula, the PNGaseF sensitivity of the N-glycans of NFP-S67F (Fig. 3) suggests that the mutant protein also accumulates in the ER in the homologous system.

Modeling the effect of the S67F mutation, suggests that the Phe residue should not perturb the global 3D structure of LysM1 (Fig. 5B), but causes changes in the local hydrophobic properties of LysM1 (Fig. 5C and D). However it is more likely that steric hindrance, rather than increased hydrophobicity is responsible for the ER localization of proteins mutated at position 67 since an Ala residue is more hydrophobic and less bulky than a Thr residue and S67A does not affect NFP, whereas the S67F mutation led to most of the protein accumulating in the ER (Fig. 3A). In the case of a mutation by a Phe residue, a change in the surface hydrophobicity (Fig. 5C and D) could participate, in addition to the changes in the LysM1 structure, to ER-QC detection. Thus it is not surprising to find that Ser67, which is located in an α-helix of the LysM1 (15), is conserved in most members of the M. truncatula LysM-RLK family or is substituted by Ala (supplemental Fig. S4). These results suggest that the phenotype of the strong nfp-2 allele is due to misfolding of the protein and that NFP trafficking to the PM requires correct folding of its extracellular region.

Although Ser67 is part of an N-glycosylation site, the lack of N-glycosylation is not the reason for the ER-QC of NFP-S67F as NFP-S67A, locates to the PM (Fig. 3). Previous work (15) and this study (Fig. 2) have shown that all of the ten predicted N-glycosylation sites in NFP are probably occupied and that eight of them occur on the surface of the LysM domains. NFP indeed contains more N-glycosylation sites than other M. truncatula LysM-RLKs (Fig. 2). In N. benthamiana, following inhibition of N-glycosylation by tunicamycin treatment, NFP was still able to reach the PM, suggesting that the N-glycans are not essential for NFP folding and trafficking to the PM, although it does not exclude that they facilitate it. Moreover, by mutating each or combinations of these N-glycosylation sites, we have shown that the proteins are still capable of restoring nodulation of the nfp-1 mutant (Table 1), suggesting that N-glycosylation of the LysM domains is not essential for NFP activity and NF perception. Mutations of the N-glycosylation site in the LysM1, N65Q and of combinations of N-glycosylation sites in the LysM domains, led to partial sensitivity to PNGaseF (data not shown).  

In the case of N65Q, the N-glycan itself is not responsible for this partial sensitivity since the
mutation S67A destroying the same N-glycosylation site is as insensitive to PNGaseF as NFP. Thus the nature of the residues modifying a predicted N-glycosylation site is important as even conservative substitutions in NFP (N65Q and S67T) affect trafficking to the PM.

Nevertheless, although N-glycans might facilitate NFP folding and trafficking to the PM, they are not essential for its function. In contrast, several other RLKs show N-glycan requirement for their activity. This property is protein specific as differences in sensitivity to the ER-QC and in N-glycan requirement for activity have been reported for the closely related LRR-RLKs of Arabidopsis thaliana FLS2 and EFR (36,41-44).

In contrast to N-glycosylation, S-S bridge formation appears to be essential for NFP activity. By studying a plant expressed NFP fusion, we have shown that all the six Cys residues in the NFP extracellular region are likely to be involved in S-S bridges (Fig. 4). NFP variants mutated in these Cys residues have PNGaseF-sensitive N-glycans (Fig. 4) suggesting that the proteins are retained in the ER. This localization has been demonstrated for the C102/104/164/166A protein in N. benthamiana and moreover this protein interacts with BIP chaperones (Fig. 4). Protein mutated in the Cys residues might thus be not correctly folded and/or might interact with ER-QC proteins through either the N-glycan, hydrophobic exposed region or mixed S-S bridges. These results suggest that the S-S bridges are necessary for NFP folding and trafficking to the PM. Cys residues are present in the extracellular region of many plant RLKs and receptor-like proteins (RLPs) and have been shown to be important for the function of some of these proteins, including Cf9 (36,45), BR11 (46) but not CLV2 (47). For BR11, the bri1-5 allele contains a point mutation in Cys69 which leads to retention of the protein in the ER and interaction with the ER chaperones, calnexin and BIP (46).

LysM-RLKs have a very characteristic arrangement of Cys residues with the presence of a CXC motif in each of the two spacer regions between the three LysM domains (Fig. 2). This arrangement occurs also in related LYM proteins (11). In addition, some LysM-RLKs, including M. truncatula LYR5 and LYR6 contain a third CXC motif before the transmembrane spanning segment (supplemental Fig. S4). CXC motifs are found in several protein families including some plant proteins (e.g. prolamins and lipid transfer proteins containing an 8-Cys motif – (48)), some scorpion toxins and related natural peptides (49) and the animal CXC-type chemokines (50). In some of these proteins the CXC motif has been shown to be involved in S-S bridges. In most of these cases each Cys residue of the CXC motif forms a S-S bridge with another Cys residue in the protein and in only a few very rare examples does the CXC motif form an intra-motif S-S bridge leading to a thermodynamically disfavored 11-member ring structure (49,51).

Examination of the extracellular region of NFP and other LysM-RLKs revealed two other Cys residues before LysM1, separated by 3 to 7 residues (Fig. S4). Mutation analysis revealed that the only individual Cys residues with an essential role in NFP activity are Cys47 in the N-ter and Cys166, located between LysM2 and LysM3. Thus a S-S bridge of primordial importance might occur between these two residues. Two other S-S bridges which are individually less essential might occur between the two CXC motifs and between the first CXC motif and Cys39 in the N-ter. Double mutations of NFP-C102A/C104A and NFP-C104A/C164A that would disrupt these two S-S bridges lead to non-functional NFP proteins and suggest that the S-S bridges link Cys39 to Cys104 and Cys102 to Cys164. An arrangement involving each of the Cys residues before LysM1 interacting with one of the CXC motifs, although speculative, is supported by the observation that some plant LysM-RLKs (for example LYR2-LYR6 of M. truncatula) contain a Cys residue in the middle of LysM2 and an additional one in the N-ter. In these proteins a fourth S-S bridge could form between these residues, forming a link between the N-ter region and LysM2. Taking into account this prediction, we modeled the entire extracellular region of NFP (Fig. 5A) and the stability of the model supports the S-S bridge arrangement we propose. However, our model is an interpretation of the data and we cannot exclude other arrangements.

Protein disulphide isomerase, thioredoxin and some prokaryotic periplasmic enzymes involved in protein oxidation (e.g. DsbA) contain CXXC motifs which have disulphide isomerase activity. Deletion of one AA to form a CXC motif in E. coli thioredoxin, greatly increases the disulphide isomerase activity and even a synthetic CXC tripeptide possesses such activity (52). In addition, the presence of a CXC motif at the C-terminus of two enzymes from hyperthermophilic Archaea was shown to be necessary for folding of these enzymes (53) and...
in their absence, addition of synthetic peptides, containing the CXC region, were able to restore the folding. These peptides were also shown to have disulphide isomerase activities (53). Thus it is possible that the CXC motifs in NFP and other LysM-RLKs have disulphide isomerase activities and that these proteins have the ability to catalyze the formation of their own S-S bridges. Because mutations in the Cys residues inhibit NFP trafficking, it is difficult to assess whether the S-S bridges in the extracellular region are also important for ligand binding. In the LysM domains of the chitinase of the fern Pteris ryukyuensis four Cys residues are involved in intra-domain S-S bridges and stabilization of the LysM-domain structure (54). For NFP and other plant LysM-RLKs it is likely that the S-S bridges are important for stabilizing the 3-dimensional geometry of the three LysM domains relative to one another (Fig. 5A). Indeed, biochemical analysis showed that the three LysM domains of CERK1 are required for ligand binding (55). Homology modeling and docking suggests that each LysM domain of NFP may bind a NF ligand (15) and genetic analysis, suggests differences in importance of the three domains, with LysM2 of NFP (13) and NFR5 (14) playing the most important role in rhizobial selectivity. The configuration defined by the proposed S-S bridges (Fig. 5A) in which the N and C termini of LysM1 and LysM2 and the N-terminus of LysM3 are physically very close, may be important for the perception of complex or multiple ligands and/or for cooperativity of ligand binding. The demonstration in NFP of the importance of the Cys residues, and their involvement in S-S bridges, will aid strategies to produce functional plant LysM-RLK and LYM proteins for ligand binding studies.

REFERENCES
1. Gough, C., and Cullimore, J. (2011) Mol Plant Microbe Interact 24, 867-878
2. Oldroyd, G. E., and Downie, J. A. (2008) Annu Rev Plant Biol 59, 519-546
3. Arrighi, J. F., Barre, A., Ben Amor, B., Bersoult, A., Soriano, L. C., Mirabella, R., de Carvalho-Niebel, F., Journet, E. P., Gherardi, M., Huguet, T., Geurts, R., Denarie, J., Rouge, P., and Gough, C. (2006) Plant Physiol 142, 265-279
4. Madsen, E. B., Madsen, L. H., Radutoiu, S., Olbyt, M., Rakwalska, M., Szczyglowski, K., Sato, S., Kaneko, T., Tabata, S., Sandal, N., and Stougaard, J. (2003) Nature 425, 637-640
5. Radutoiu, S., Madsen, L. H., Madsen, E. B., Felle, H. H., Umehara, Y., Gronlund, M., Sato, S., Nakamura, Y., Tabata, S., Sandal, N., and Stougaard, J. (2003) Nature 425, 585-592
6. Smit, P., Limpons, E., Geurts, R., Fedorova, E., Dolgikh, E., Gough, C., and Bisseling, T. (2007) Plant Physiol 145, 183-191
7. Shiu, S. H., and Bleecker, A. B. (2001) Proc Natl Acad Sci U S A 98, 10763-10768
8. Madsen, E. B., Antolin-Llovera, M., Grossmann, C., Ye, J., Vieweg, S., Broghammer, A., Krussell, L., Radutoiu, S., Jensen, O. N., Stougaard, J., and Parniske, M. (2011) Plant Journal 65, 404-417
9. Castells, E., and Casacuberta, J. M. (2007) J Exp Bot 58, 3503-3511
10. Klaus-Heisen, D., Nurisso, A., Pietraszewska-Bogtj, A., Mbengue, M., Camut, S., Timmers, T., Pichereaux, C., Rossignol, M., Gadella, T. W., Imberty, A., Lefebvre, B., and Cullimore, J. V. (2011) J Biol Chem 286, 11202-11210
11. Fliegmann, J., Uhlenbroich, S., Shinya, T., Martinez, Y., Lefebvre, B., Shibuya, N., and Bono, J. J. (2011) Plant Physio Biochem 49, 709-720
12. Buist, G., Steen, A., Kok, J., and Kuipers, O. P. (2008) Mol Microbiol 68, 838-847
13. Bensmihen, S., de Billy, F., and Gough, C. (2011) PLoS One 6, e26114
14. Radutoiu, S., Madsen, L. H., Madsen, E. B., Jurkiewicz, A., Fukai, E., Quistgaard, E. M., Albrectsen, A. S., James, E. K., Thirup, S., and Stougaard, J. (2007) Embo J 26, 3923-3935
15. Mulder, L., Lefebvre, B., Cullimore, J., and Imberty, A. (2006) Glycobiology 16, 801-809
16. Leroque, P., Cabanes-Macheteau, M., Rayon, C., Fischette-Laine, A. C., Gomord, V., and Faye, L. (1998) Plant Mol Biol 38, 31-48
17. Saijo, Y. (2010) Cell Microbiol 12, 716-724
18. Pattison, R. J., and Amtmann, A. (2009) Trends Plant Sci 14, 92-99
19. Gomord, V., Fitchette, A. C., Menu-Bouaouiche, L., Saint-Jore-Dupas, C., Plasson, C., Michaud, D., and Faye, L. (2010) Plant Biotechnol J 8, 564-587
Structure-Function of *Medicago trunculata* NFP

20. Vitale, A., and Boston, R. S. (2008) *Traffic* **9**, 1581-1588
21. Kremers, G. J., Goedhart, J., van Munster, E. B., and Gadella, T. W., Jr. (2006) *Biochemistry* **45**, 6570-6580
22. Shaner, N. C., Campbell, R. E., Steinbach, P. A., Giepmans, B. N., Palmer, A. E., and Tsien, R. Y. (2004) *Nat Biotechnol* **22**, 1567-1572
23. Lefebvre, B., Batoko, H., Duby, G., and Boutry, M. (2004) *Plant Cell* **16**, 1772-1789
24. Haseloff, J., Siemering, K. R., Prasher, D. C., and Hodge, S. (1997) *Proc Natl Acad Sci U S A* **94**, 2122-2127
25. Ivanchenko, M., Vejlupkova, Z., Quatrano, R. S., and Fowler, J. E. (2000) *Plant J* **24**, 79-90
26. Lefebvre, B., Batoko, H., Duby, G., and Boutry, M. (2004) *Plant Cell* **16**, 1772-1789
27. Haseloff, J., Siemering, K. R., Prasher, D. C., and Hodge, S. (1997) *Proc Natl Acad Sci U S A* **94**, 2122-2127
28. van Bokhoven, H., Verver, J., Wellink, J., and van Kammen, A. (1993) *J Gen Virol* **74**, 2233-2241
29. Mbengue, M., Camut, S., de Carvalho-Niebel, F., Deslandes, L., Froidure, S., Klaus-Heisen, D., Moreau, S., Rivas, S., Timmers, T., Herve, C., Cullimore, J., and Lefebvre, B. (2010) *Plant Cell* **22**, 3474-3488
30. Vitha, S., Baluska, F., Mews, M., and Volkmann, D. (1997) *J Histochem Cytochem* **45**, 89-95
31. Timmers, A. C., Auric, M. C., and Truchet, G. (1999) *Development* **126**, 3617-3628
32. Moreau, S., Verdenaud, M., Ott, T., Letort, S., de Billy, F., Niebel, A., Gouzy, J., de Carvalho-Niebel, F., and Gamas, P. (2011) *PLoS One* **6**, e16463
33. Pearlman, D. A., Case, D. A., Caldwell, J. W., Ross, W. S., and Cheatham, T. E. (1995) *Comp Phys Comm* **91**, 1-41
34. Powell, M. J. D. (1964) *Comp J* **7**, 55–162
35. Waldherr-Teschner, M., Goetze, T., Heiden, W., Knoblauch, M., Vollhardt, H., and Brickmann, J. (1992) *Advances in Scientific Visualization. Springer, Heidelberg*, 58-67
36. Haweker, H., Rips, S., Koiwa, H., Salomon, S., Saijo, Y., Chinchilla, D., Robatzek, S., and von Schaewen, A. (2010) *J Biol Chem* **285**, 4629-4636
37. Gomez-Gomez, L., Bauer, Z., and Boller, T. (2001) *Plant Cell* **13**, 1155-1163
38. Navarro-Gochicoa, M. T., Camut, S., Timmers, A. C., Niebel, A., Herve, C., Boutet, E., Bono, J. J., Imberty, A., and Cullimore, J. V. (2003) *Plant Physiol* **133**, 1893-1910
39. Boudeau, J., Miranda-Saavedra, D., Barton, G. J., and Alessi, D. R. (2006) *Trends Cell Biol* **16**, 443-452
40. Cannan, N., and Neuwald, A. F. (2005) *J Mol Biol* **351**, 956-972
41. Li, J., Zhao-Hui, C., Batoux, M., Nekrasov, V., Roux, M., Chinchilla, D., Zipfel, C., and Jones, J. D. (2009) *Proc Natl Acad Sci U S A* **106**, 15973-15978
42. Lu, X., Tintor, N., Mentzel, T., Kombrink, E., Boller, T., Robatzek, S., Schulze-Lefert, P., and Saijo, Y. (2009) *Proc Natl Acad Sci U S A* **106**, 22522-22527
43. Nekrasov, V., Li, J., Batoux, M., Roux, M., Chu, Z. H., Lacombe, S., Rougon, A., Bittel, P., Kiss-Papp, M., Chinchilla, D., van Esse, H. P., Jorda, L., Schwessinger, B., Nicaise, V., Thomma, B. P., Molina, A., Jones, J. D., and Zipfel, C. (2009) *Embo J* **28**, 3428-3438
44. Saijo, Y., Tintor, N., Lu, X., Rauf, P., Pajerowska-Mukhtar, K., Haweker, H., Dong, X., Robatzek, S., and Schulze-Lefert, P. (2009) *Embo J* **28**, 3439-3449
45. van der Hoorn, R. A., Wulff, B. B., Rivas, S., Durrant, M. C., van der Ploeg, A., de Wit, P. J., and Jones, J. D. (2005) *Plant Cell* **17**, 1000-1015
46. Hong, Z., Jin, H., Tzfira, T., and Li, J. (2008) *Plant Cell* **20**, 3418-3429
47. Song, X., Guo, P., Li, C., and Liu, C. M. (2010) *J Integr Plant Biol* **52**, 774-781
48. Jose-Estanyol, M., Gomis-Ruth, F. X., and Puigdomenech, P. (2004) *Plant Physiol Biochem* **42**, 355-365
49. Zhu, Q., Liang, S., Martin, L., Gasparini, S., Menez, A., and Vita, C. (2002) *Biochemistry* **41**, 11488-11494
50. Joseph, P. R., Sarmiento, J. M., Mishra, A. K., Das, S. T., Garofalo, R. P., Navarro, J., and Rajarathnam, K. (2010) *J Biol Chem* **285**, 29262-29269
51. Derewenda, U., Boczek, T., Gorres, K. L., Yu, M., Hung, L. W., Cooper, D., Joachimiak, A., Raines, R. T., and Derewenda, Z. S. (2009) *Biochemistry* **48**, 8664-8671
52. Woycechowsky, K. J., and Raines, R. T. (2003) *Biochemistry* **42**, 5387-5394
53. Cacciapuoti, G., Peluso, I., Fuccio, F., and Porcelli, M. (2009) *Fes J* **276**, 5799-5805
54. Ohnuma, T., Onaga, S., Murata, K., Taira, T., and Katoh, E. (2008) *J Biol Chem* **283**, 5178-5187
55. Petutschnig, E. K., Jones, A. M., Serazetdinova, L., Lipka, U., and Lipka, V. (2010) *J Biol Chem* **285**, 28902-28911

**FOOTNOTES**

* This work was supported in part by the European Community Marie Curie Research Training Network Programme Contract MRTN-CT-2006-035546 “NODPERCEPTION,” French National Research Agency Contracts “NodBindsLysM”, “LCOinNONLEGUMES” and “SYMPASIGNAL.” This work has been done in the LIPM, part of the ”Laboratoire d'Excellence (LABEX) entitled TULIP (ANR-10-LABX-41).

* The online version of this article (available at http://www.jbc.org) contains supplemental Figs S1 - S6.

& The abbreviations used are: RLK, receptor-like-kinase; AA, amino acid; h/dpi, hours/days post-inoculation; FP, fluorescent protein; IR, intracellular region; S-S, disulphide; LysM, lysin motif; NF, Nod factor; PM, plasma membrane; PTM, post-translational modification; ER, endoplasmic reticulum; ERAD, ER associated degradation; PNGaseF, peptide-N-glycosidase F; qRT-PCR, quantitative reverse transcriptase polymerase chain reaction.

Acknowledgments - We thank Sophie Valière for help with NFP purification; Sandra Bensmihen, Jean-Jacques Bono and Clare Gough (LIPM) for critically reading the manuscript; Prof. Pierre-Alain Carrupt (School of Pharmaceutical Sciences, University of Geneva, University of Lausanne) for technical support in molecular modeling tools.
FIGURE LEGENDS

FIGURE 1. NFP localizes to the PM and NFP-S67F localizes in the ER in N. benthamiana. A, NFP localizes to the PM in N. benthamiana. NFP-RFP was co-expressed in N. benthamiana leaves with a PM marker (PMA4-GFP) and the leaves were analyzed by confocal microscopy at 3dpi. Superposition of the fluorescence images shows clear co-localization of NFP-RFP with the PM marker at the cell boundary. B, NFP localizes to the PM in M. truncatula roots. Roots of M. truncatula nfp-1 were transformed with the Pro35S:NFP-RFP construct or a control vector. Immunocytolabeling was performed using anti-RFP as primary antibodies and Tyramide Signal Amplification method (green) on root longitudinal sections corresponding to cortical cells. Proteins are stained with Evans blue (red) and the nucleus with DAPI (blue). C, NFP-S67F is retained in the ER in N. benthamiana. The protein encoded by the nfp-2 allele (NFP-S67F) was fused to RFP and co-expressed with the PM (PMA4-GFP) or ER (HDEL-GFP) markers and the leaves were analyzed by confocal microscopy at 3dpi. Co-localization with the ER marker in a reticulated network in the cell and around the nucleus suggests that the protein is retained in the ER. Bars represent 20µm.

FIGURE 2. N-glycosylation sites are not conserved between NFP and various LysM-RLKs and trafficking of NFP to the PM is not blocked by inhibition of N-glycosylation. A, Scheme of the extracellular region of NFP showing positions of the Asn (N) and Cys (C) residues (putatively being involved in post-translational modifications) in relation to the LysM domains. SP = signal peptide, TM = transmembrane domain, numbering is from the start of the SP. B, Conservation of the NFP Cys and N-glycosylation site residues in the orthologues from L. japonicus (NFR5) and P. sativum (SYM10) and the LYR and LYK proteins from M. truncatula. + indicates a site found at the same position, +/- indicates a site found at a close position (see alignment in supplemental Fig. S4). C, Tunicamycin effectively inhibits N-glycosylation of NFP. N. benthamiana leaves expressing NFP-YFP at 2 and 3dpi were treated with 10µM tunicamycin or solvent only (diluted DMSO) for 20h prior to analysis by immunoblotting. Tunicamycin treatment led to NFP-YFP detected at the size predicted from the unmodified AA sequence (93 kDa) whereas the protein in the untreated leaves contains about 20 kDa of N-glycans. D, NFP can reach the PM in absence of N-glycosylation. The same material described in C (leaves at 3dpi treated with 10µM tunicamycin or solvent only (diluted DMSO) for 20h prior to analysis by immunoblotting) was observed by confocal microscopy: two images are shown. The arrow points to labeling of perinuclear ER, whereas most of the NFP protein localizes to the PM. Bars represent 20µm.

FIGURE 3. Trafficking of NFP-S67F to the PM is blocked in M. truncatula and NFP-S67F protein interacts with the ER-located BIP chaperones in N. benthamiana. A, NFP-S67F lacks Golgi matured N-glycans in M. truncatula. Protein extracts from nfp-1 roots expressing NFP-RFP or indicated mutated variants were treated with PNGaseF and analyzed by immunoblotting. Lower size band after treatment indicates a sensitivity of NFP N-glycans to PNGaseF and reflects an absence of maturation of NFP N-glycans in the Golgi apparatus and hence retention of the protein in the ER. The higher band size reflects transport through the Golgi to the PM. B, NFP-S67F interacts with the BIP chaperones in N. benthamiana. NFP-RFP or NFP-S67F-RFP expressed in N. benthamiana leaves were solubilized and purified using anti-RFP antibodies. Immunoblotting reveals that NFP-S67F, but not NFP, interacts strongly with BIP chaperones. NT: non transformed leaves. C, BIP chaperones are not over-expressed in roots of nfp-2 mutant plants. An immunoblot of protein extracts from roots of WT, nfp-1, nfp-2 plants reveals similar quantities of BIP chaperones, suggesting that the accumulation of NFP-S67F protein in the ER does not lead to an “unfolded protein response”. The ponceau red staining shows the protein loading.

FIGURE 4. NFP contains three disulphide bridges and mutants in the CXC motifs are retained in the ER. A, NFP bears three S-S bridges. A truncated version of NFP, without its intracellular region, fused to RFP (NFPΔIR-RFP) was expressed in N. benthamiana leaves. The microsomal
fraction derived from these leaves was first treated or not with DTT and then free Cys residues were labeled with MPBiotin. NFP was then solubilized and purified using anti-RFP antibodies. Proteins were separated by SDS-PAGE and analyzed by immunoblotting: fusion proteins were detected using anti-RFP antibodies and labeled Cys residues were detected using anti-biotin antibodies. The result suggests that the six Cys residues of NFP extracellular region are involved in S-S bridges and only the single Cys residue of the RFP is labeled in the absence of DTT. B, NFP mutated in the Cys pairs localizes to the ER in *M. truncatula*. Protein extracts from *nfp*-1 roots expressing the indicated NFP-RFP mutated variants were treated as described in Figure 3A. The sensitivity of the proteins to PNGaseF suggests lack of N-glycan modification and hence retention in the ER. C, NFP mutated in the CXC motifs interacts with BIP chaperones. NFP and the quadruple-Cys to Ala mutant proteins were expressed and purified from *N. benthamiana* leaves and complexes with BIP were analyzed by SDS-PAGE and immunoblotting. The mutant NFP protein co-purifies BIP chaperones, suggesting that it is retained in the ER by binding to BIP chaperones. NT: non transformed leaves.

FIGURE 5. *Molecular modeling of the NFP extracellular region*. A, Molecular modeling showing the suggested positions of the S-S bridges. LysM domains are colored in blue (LysM1), red (LysM2) and green (LysM3) ribbons whereas the preceding loops are represented in pale blue, pale red and pale green, respectively. Cys residues involved in S-S bridges are shown as yellow sticks. The S-S bridge positions deduced from mutation and complementation analysis (Table 1) were used for modeling. B, Superposed molecular modeling of NFP and NFP S67F. NFP model before (violet ribbons) and after mutation (pink ribbons): local structural changes are highlighted at the LysM1 level. The arrow points the position of the residue 67. C and D, Solvent accessible surfaces colored according to the lipophilic potential, from brown (hydrophobic) to blue (polar) calculated for the NFP model before (C) and after the S67F mutation (D). The mutation zone is circled.
**TABLE 1. Extracellular Cys residues and an intracellular kinase residue, but not N-glycosylation sites are essential for NFP biological activity.** Constructs coding for tagged NFP or the indicated mutant variants of NFP were transformed into *M. truncatula nfp-*1. Transgenic roots were inoculated with *S. melliloti* and complementation for nodulation was analyzed two weeks later. When indicated, proteins from transgenic roots were analyzed for sensitivity to PNGaseF as shown in Fig. 3 (an indicator of ER localization) and localization in *N. benthamiana* (as shown in Fig. 1 and supplemental Figs. S3 and S5). Abbreviations for the intracellular region features are JM = juxtamembrane, A-segment = Activation segment, C-tail = C-terminal tail. AA numbering starts from the first Met of the predicted signal peptide.

| Construct / NFP mutation | Domain mutated | Nodulation | PNGaseF sensitivity | Localization in *N. benthamiana* |
|--------------------------|----------------|------------|---------------------|---------------------------------|
| Empty vector             |                |            |                     |                                 |
| **NFP and NFP-S67F**     |                |            |                     |                                 |
| NFP                      |                | +          | -                   |                                 |
| S67F                     | LysM1          | -          | +                   | PM                              |
| S67T                     | LysM1          | +          | -                   | ER                              |
| S67A                     | LysM1          | +          | -                   |                                 |
| **N-glycosylation sites**|                |            |                     |                                 |
| N65Q                     | LysM1          | +          | -                   |                                 |
| N71Q                     | LysM1          | -          | -                   |                                 |
| N65/71Q                  | LysM1          | +          | -                   |                                 |
| N112Q                    | LysM2          | +          | -                   |                                 |
| N133Q                    | LysM2          | +          | -                   |                                 |
| N146Q                    | LysM2          | +          | -                   |                                 |
| N112/133/146Q            | LysM2          | +          | -                   |                                 |
| N189Q                    | LysM3          | +          | -                   |                                 |
| N211Q                    | LysM3          | +          | -                   |                                 |
| N219Q                    | LysM3          | +          | -                   |                                 |
| N189/211/219Q            | LysM3          | +          | -                   |                                 |
| **Cys residues**         |                |            |                     |                                 |
| C39A                     | <LysM1         | +          | -                   |                                 |
| C47A                     | <LysM1         | -          | -                   |                                 |
| C39/47A                  | <LysM1         | -          | +                   |                                 |
| C102A                    | LysM1<>LysM2   | +          | +                   |                                 |
| C104A                    | LysM1<>LysM2   | +          | +                   |                                 |
| C102/104A                | LysM1<>LysM2   | -          | +                   |                                 |
| C164A                    | LysM2<>LysM3   | +          | +                   |                                 |
| C166A                    | LysM2<>LysM3   | -          | +                   |                                 |
| C164/166A                | LysM2<>LysM3   | -          | +                   |                                 |
| C102/164A                | +              | +          | +                   |                                 |
| C102/166A                | -              | -          | +                   |                                 |
| C104/164A                | -              | -          | +                   |                                 |
| C104/166A                | -              | +          | -                   |                                 |
| C102/104/164/166A        | -              | +          | -                   |                                 |
| **Intracellular region** |                |            |                     |                                 |
| T281/S282/S283A          | JM             | +          | -                   |                                 |
| K339A                    | VAiK           | +          | -                   |                                 |
| D435A                    | HRD            | +          | -                   |                                 |
| T459/S460/T461A          | A-segment      | +          | -                   |                                 |
| G474A                    | α-helix F      | +          | -                   |                                 |
| G474E                    | α-helix F      | -          | +                   |                                 |
| T482A                    | α-helix F      | +          | -                   |                                 |
| T578/S579A               | C-tail         | +          | -                   |                                 |
Figure 1

A  NFP-RFP  PM marker  Merged

B  Anti-RFP  Evans blue  DAPI
    NFP-RFP
    Control

C  NFP-S67F-RFP  PM marker  Merged
    NFP-S67F- RFP  ER marker  Merged
Figure 2

A

[Diagram showing amino acid sequence with LysM1, LysM2, and LysM3 domains.]  

B

| NFP     | C39 | C47 | N65 | N71 | C102 | C104 | N112 | N133 | N146 | C164 | N189 | N211 | N219 |
|---------|-----|-----|-----|-----|------|------|------|------|------|------|------|------|------|
| NFR5    | +   | +   | +   | +   | +    | +    | +    | +    | +    | +    | +    | +    | +    |
| SYM10   | +   | +   | +   | +   | +    | +    | +    | +    | +    | +    | +    | +    | +    |
| LYR1    | +   | +   | +   | +   | +    | +    | +    | +    | +    | +    | +    | +    | +    |
| LYR2    | +/- | +   | +   | +   | +    | +/-  | +    | +    | +    | +    | +    | +    | +    |
| LYR3    | +/- | +   | +   | +   | +    | +/-  | +    | +    | +    | +    | +    | +    | +    |
| LYR4    | +/- | +   | +   | +   | +    | +/-  | +    | +    | +    | +    | +    | +    | +    |
| LYR5    | +/- | +   | +   | +   | +    | +/-  | +    | +    | +    | +    | +    | +    | +/-  |
| LYR6    | +/- | +   | +   | +   | +    | +    | +    | +    | +    | +    | +    | +    | +/-  |
| LYK1    | +/- | +/- | +/- | +   | +    | +    | +    | +    | +    | +    | +    | +    | +    |
| LYK2    | +/- | +/- | +/- | +   | +    | +    | +    | +    | +    | +    | +    | +    | +/-  |
| LYK3    | +/- | +/- | +/- | +   | +    | +    | +    | +    | +    | +    | +    | +    | +    |
| LYK4    | +/- | +/- | +/- | +   | +    | +    | +    | +    | +    | +    | +    | +    | +    |
| LYK5    | +/- | +/- | +/- | +   | +    | +    | +    | +    | +    | +    | +    | +    | +    |
| LYK6    | +/- | +/- | +/- | +   | +    | +    | +    | +    | +    | +    | +    | +    | +    |
| LYK7    | +/- | +/- | +/- | +   | +    | +    | +    | +    | +    | +    | +    | +    | +    |
| LYK8    | +/- | +/- | +/- | +   | +    | +    | +    | +    | +    | +    | +    | +    | +    |
| LYK9    | +/- | +/- | +/- | +   | +    | +    | +    | +    | +    | +    | +    | +    | +    |
| LYK10   | +/- | +/- | +/- | +   | +    | +    | +    | +    | +    | +    | +    | +    | +    |

C

| Tunicamycin | 2dpi | 3dpi |
|-------------|------|------|
| +           | -    | +    |
| -           | +    | -    |

D

NFP-YFP 3dpi + Tunicamycin

[Images of Western Blot and Immunofluorescence Microscopy with arrows indicating changes.]
Figure 4

A

| MPBiotin | + | + | MPBiotin | + | + |
|----------|---|---|----------|---|---|
| DTT      | + | - | kDa      |
| NFPΔIR-RFP | 72 |
| WB: Anti biotin |

B

| PNGaseF | - | + | - | + | - | + | - | + | kDa |
|---------|---|---|---|---|---|---|---|---|---|
| NFP-RFP PM |
| NFP-RFP ER |
| WB: Anti RFP |

C

| WT | C102/104/164A/166A |
|----|--------------------|
| NFP-RFP |
| WB: Anti RFP |
| BIP |
| WB: Anti BIP |
Role of N-glycosylation sites and CXC motifs in trafficking of Medicago trunculata Nod Factor Perception protein to the plasma membrane
Benoit Lefebvre, Doerte Klaus-Heisen, Anna Pietraszewska-Bogiel, Christine Hervé, Sylvie Camut, Marie-Christine Auriac, Virginie Gasciolli, Alessandra Nurisso, Benoit Lefebvre, Virginie Gasciolli, Alessandra Nurisso, Theodorus W. J. Gadella and Julie Cullimore

J. Biol. Chem. published online February 9, 2012

Access the most updated version of this article at doi: 10.1074/jbc.M111.281634

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

Supplemental material:
http://www.jbc.org/content/suppl/2012/02/09/M111.281634.DC1