Molecular Microbiology (1996) **19**(3), 443–453

The C-terminal domain of the *Rhizobium leguminosarum* chitin synthase NodC is important for function and determines the orientation of the N-terminal region in the inner membrane

Marie-Anne Barny,[1] Eric Schoonejans,[2] Anastassios Economou,[3] Andrew W. B. Johnston[4] and J. Allan Downie*

John Innes Centre, Colney, Norwich NR4 7UH, UK.

Summary

The *nod*C genes from rhizobia encode an *N*-acetylglucosaminyl transferase (chitin synthase) involved in the formation of lipo-chito-oligosaccharide Nod factors that initiate root nodule morphogenesis in legume plants. NodC proteins have two hydrophobic domains, one of about 21 residues at the N-terminus and a longer one, which could consist of two or three transmembrane spans, near the C-terminus. These two hydrophobic domains flank a large hydrophilic region that shows extensive homology with other β-glycosyl transferases. The topology of NodC in the inner membrane of *Rhizobium leguminosarum* biovar *viciae* was analysed using a series of gene fusions encoding proteins in which NodC was fused to alkaline phosphatase (PhoA) lacking an N-terminal transit sequence or to β-galactosidase (LacZ). Our data support a model in which the N-terminal hydrophobic domain spans the membrane in a N_{out}–C_{in} orientation, with the adjacent large hydrophilic domain being exposed to the cytoplasm. This orientation appears to depend upon the presence of the hydrophobic region near the C-terminus. We propose that this hydrophobic region contains three transmembrane spans, such that the C-terminus of NodC is located in the periplasm. A short region of about 40 amino acids, encompassing the last transmembrane span, is essential for the function of NodC. Our model for NodC topology suggests that most of NodC, including the region showing most similarity to other β-glycosyl transferases, is exposed to the cytoplasm, where it is likely that polymerization of *N*-acetyl glucoasamine occurs. Such a model is incompatible with previous reports suggesting that NodC spans both inner and outer membranes.

Received 5 July, 1995; revised 5 September, 1995; accepted 7 September 1995. Present addresses: [1]Laboratoire de Pathologie Végétale, INRA, 16 rue Claude Bernard, 75231 Paris Cedex 05, France; [2]Institut des Sciences Végétales (CNRS), F91198 Gif-sur-Yvette, France; [3]Department of Biology, University of Crete, GR-71409 Heraclio, Greece; [4]School of Biological Sciences, University of East Anglia, Norwich NR4 7TJ, UK. *For correspondence. E-mail DOWNIE @BBSRC.AC.UK; Tel. 01603 452571; Fax 01603 456844.

Introduction

Recognition between leguminous plants and the rhizobia that nodulate them is mediated via specific signalling molecules exchanged between the two symbiotic partners. The plant roots secrete flavonoid or isoflavonoid molecules that induce the bacterial nodulation (*nod*) genes (Carlson *et al.*, 1994; Downie, 1994). Several of the *nod* gene products are involved in the synthesis of lipo-chito-oligosaccharide Nod factors that induce root nodule morphogenesis (Truchet *et al.*, 1991; Spaink *et al.*, 1991).

The most crucial *nod* gene products required for the synthesis of these lipo-chito-oligosaccharides are NodA, NodB and NodC and the *nodA nodB* and *nodC* genes are highly conserved in all rhizobia studied. In the absence of the other *nod* gene products, they can form a 'core' signal consisting of oligomers of four or five *N*-acetyl glucosamine residues carrying an N-linked acyl group (Spaink *et al.*, 1991). The function of each of the three proteins is known: NodC is an *N*-acetylglucosaminyl transferase (chitin synthase) which makes the chito-oligosaccharide chain (Geremia *et al.*, 1994; Spaink *et al.*, 1994); the *N*-acetyl group from the non-reducing *N*-acetyl glucosamine residue of the chito-oligosaccharide chain is removed by NodB, which acts as a chitin oligosaccharide deacetylase (John *et al.*, 1993; Spaink *et al.*, 1994); NodA is involved in the attachment of the acyl chain to the free amino group generated by the action of NodB (Atkinson *et al.*, 1994; Röhrig *et al.*, 1994).

These reactions probably occur in the cytoplasm. It is evident from their predicted sequences that NodA and NodB (Török *et al.*, 1984; Rossen *et al.*, 1984; Schofield and Watson, 1986; Scott, 1986; Goethals *et al.*, 1989; Krishnan and Pueppke, 1991) are relatively hydrophilic, do not have an N-terminal transit sequence and are, therefore, likely to be cytoplasmically located. NodA was found

to be located predominantly in the cytoplasm of *Rhizobium meliloti* cells in which it was overexpressed (Schmidt *et al*., 1986; Johnson *et al*., 1989), although others found a significant amount of NodA co-purified with the inner membrane, especially when NodA was expressed at lower levels (Egelhoff and Long, 1986; Schlaman, 1992). The NodC proteins have two hydrophobic domains, a short one at the *N*-terminus and a longer one near the C-terminus (Török *et al*., 1984; Rossen *et al*., 1984; Jacobs *et al*., 1985; Goethal *et al*., 1989; Collins-Emerson, 1990; Vázquez *et al*., 1991; Krishan and Pueppke, 1991). It was reported (John *et al.,* 1988; Johnson *et al.,* 1989) that NodC from *R. meliloti* spans both the inner and outer membranes. Hubac *et al*. (1992) found NodC to be present mostly in the inner membrane fraction isolated from *R. meliloti*, but also found significant levels of NodC in the outer membrane fraction. However, using a different membrane fractionation procedure that gave rise to outer membranes that had very low levels of inner membrane contamination, Barny and Downie (1993) reported that NodC from *Rhizobium leguminosarum* biovar *viciae* is exclusively found in the inner membrane.

To analyse the location and membrane topology of NodC, we constructed and characterized fusions of NodC to either alkaline phosphatase or β-galactosidase. The results support a model in which NodC is located in the inner membrane with the bulk of the protein located in the cytoplasm. It is evident that the C-terminal domain of NodC is important in establishing the correct topology of NodC.

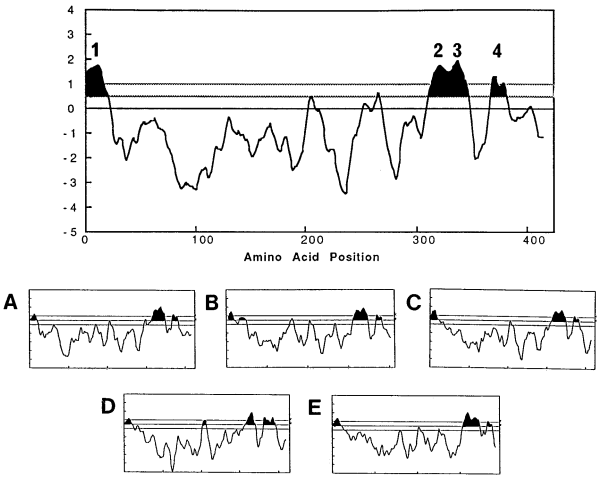

Fig. 1. Hydrophobic and hydrophilic domains in NodC. NodC protein sequences were analysed for hydrophobic regions using the software TopPredII (Claros and von Heijne, 1994). The GES hydrophobicity scale (Engelman *et al*., 1986) was used with the following parameters: full window 21; core window 11; critical length 60. The uppermost graph is the hydropathy plot obtained with the *R. l. viciae* NodC sequence (Rossen *et al*., 1984; SWISSPROT: NODC-RHILV entry, Accession Number P04340). The hydrophobic domains above a cut-off value of 0.5 are shown in black and numbered.
A. *R. meliloti* (Török *et al.,* 1984; Swissprot: NODC-RHIME entry, Accession Number X01649),
B. *R. fredii* (Krishnan and Pueppke, 1991; translation of EMBL entry RFNODABC, Accession Number M73699).
C. *R. loti* (Collins-Emerson *et al*., 1990; Swissprot: NODC-RHILO entry, Accession Number P17862).
D. *R. phaseoli* type I strain (Vázquez *et al*., 1991; SWISSPROT: NODC-RHILP entry, Accession Number P24151).
E. *Azorhizobium caulinodans* (Goethals *et al*., 1989; SWISSPROT: NODC-AZOCA entry, Accession Number Q07755).
Four strongly hydrophobic peaks are conserved in all the NodC proteins.

Results

Predicted hydrophobicity of NodC

As shown (Fig. 1), the first 21 amino acid residues of NodC are hydrophobic. This domain differs from classical cleavable leader peptides in that it does not carry a positively charged residue near the N-terminus. The NodC proteins from rhizobial strains can have a negatively charged residue within the first five residues or have no net charge at all. In all of the NodC proteins, the N-terminal hydrophobic domain is followed by a large hydrophilic region that is about 285 residues long.

In NodC from *R. l. viciae*, three strongly hydrophobic peaks are predicted around residues 304–379 (Fig. 1). Hydrophobic peaks 2 and 3 are contiguous while peaks 3 and 4 are separated by a small hydrophilic region (Fig. 1). This domain structure is strongly conserved in NodC proteins from *R. l. viciae*, *R. meliloti*, *Rhizobium fredii*, *Rhizobium loti* and *Azorhizobium caulinodans*. In *Rhizobium etli*, four similar peaks of hydrophobicity are conserved, but the three C-terminal peaks are in slightly different positions (Fig. 1D). The conservation of the amino acid sequence in the hydrophobic region is less pronounced than the conservation of hydrophobicity. Therefore, there are four potential hydrophobic transmembrane domains that could anchor NodC in the inner membrane (Fig. 1).

Isolation and characterization of NodC–PhoA fusions in the C-terminal region

Tn*phoA*, a derivative of Tn*5* carrying *phoA* lacking its own promoter, translational start and export signal, can be used to analyse the topology of inner membrane proteins: highly active *phoA* fusions indicate that Tn*phoA* is inserted in frame within a periplasmic domain of a membrane protein, resulting in translocation of the PhoA moiety to the periplasm where it is most active (Manoil and Beckwith, 1985; 1986). Plasmids pIJ1089 and pIJ1887 were each mutagenized with Tn*phoA*. These plasmids contain, in addition to *nodABC* genes, other *nod* genes including *nodD*, which activate transcription of the *nodABC* operon

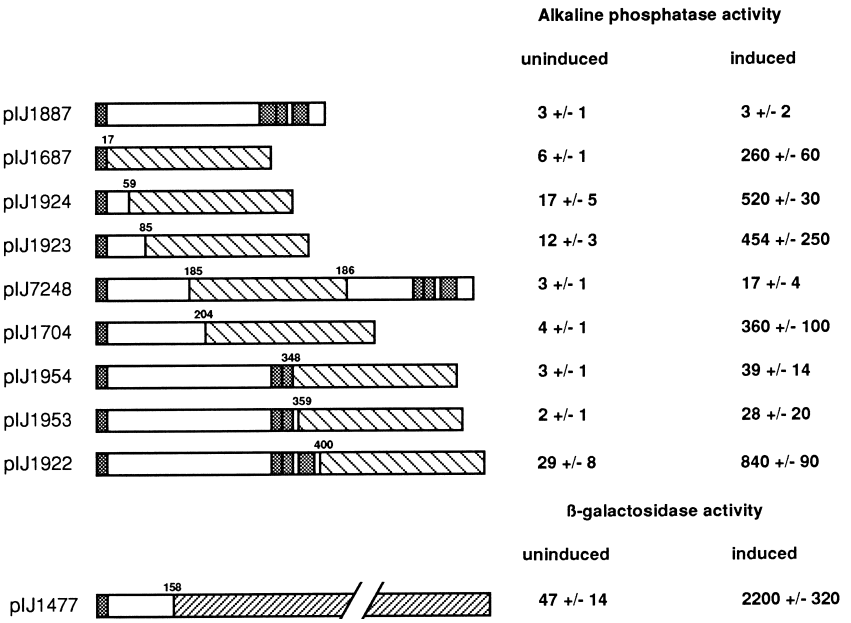

Fig. 2. NodC–PhoA, NodC–PhoA–NodC and NodC–LacZ fusion proteins showing levels of alkaline phosphatase or β-galactosidase activity. The diagram of the native NodC protein (encoded on pIJ1887) shows the four hydrophobic peaks as shaded boxes and the three hydrophilic domains as open boxes. The different NodC–PhoA fusion proteins, encoded by pIJ1687, pIJ1924, pIJ1923, pIJ1704, pIJ1954, pIJ1953 and pIJ1922, are shown along with the NodC′–PhoA–′ NodC sandwich fusion encoded by pIJ7248. The hatched boxes represent the PhoA or LacZ regions and the numbers above the fusions indicate the amino acyl residues of NodC retained
in the fusions. The levels of alkaline phosphatase or β-galactosidase activity obtained with strain 8401 containing each plasmid are shown. The values (± SD) are the average of at least five assays after growth of the bacteria in Y-succinate medium in the presence (induced) or absence (uninduced) of hesperetin.

in the presence of hesperetin. Derivatives of pIJ1089 and pIJ1887 containing Tn*phoA* were transferred to *R. leguminosarum* strain 8401 by conjugation. NodC–PhoA fusions expressing high levels of alkaline phosphatase could be identified in *R. leguminosarum* on plates containing hesperetin to induce *nod* gene expression and the chromogenic indicator 5′-bromo-4′-chloro-3′-indolyl phosphate (XP) to detect alkaline phosphatase activity. Five transconjugants that gave rise to dark blue colonies only when hesperetin was present were studied further. The transposon insertions were mapped with *Eco*RI, *Bam*HI and *Hin*dIII and found to be at different locations within *nodC*. The five plasmids from these strains were called pIJ1687, pIJ1704, pIJ1922, pIJ1923 and pIJ1924 (Fig. 2).

The precise locations of the Tn*phoA* insertions were determined by DNA sequencing of the fusion junctions using a primer from within *phoA*. One of the fusions, carried on pIJ1922, was located near the end of *nodC* such that the protein fusion occurred after residue 400 (NodC400–PhoA), 24 amino acids before the end of NodC (Fig. 2). The strain carrying this fusion was found to be highly induced for alkaline phosphatase activity when grown in the presence of hesperetin (Fig. 2).

A protein extract from 8401 (pIJ1922) (carrying the NodC400–PhoA fusion) was separated by SDS–PAGE, transferred to nitrocellulose and stained with NodC or PhoA antisera. The NodC antiserum recognised a protein at about M_r 95 000, the expected size of the fusion protein (Fig. 3B, lane g). With the PhoA antiserum, the size of the largest stained band was also in good agreement with the predicted size of the NodC400–PhoA fusion protein (Fig. 3A, lane g). However, the strongest signal was seen at a molecular weight smaller than the predicted size of fusion. Since there was no cross-reaction with the strain lacking a NodC–PhoA fusion (Fig. 3A, lane i), it was concluded that there was significant degradation of the fusion protein. Indeed, the major degradation product had a similar size to that of mature alkaline phosphatase. This indicated that the hybrid protein was probably cleaved near the junction site, leading to an accumulation of PhoA in the periplasm. Similar degradation of hybrid proteins containing PhoA fused to a protein with an N-terminal transit signal was observed previously in *R. l. viciae* (Rivilla *et al.*, 1995).

Identification of an active PhoA fusion near the end of NodC indicates that the C-terminus of NodC is located in the periplasm. This fusion also retains NodC activity in bio-

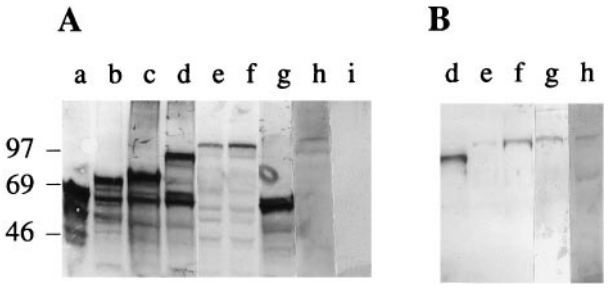

Fig. 3. Identification of NodC–PhoA and NodC–PhoA–NodC fusion proteins. Cell extracts (50 μg protein) from each strain were separated on a SDS–polyacrylamide gel and transferred to nitrocellulose. The lanes are a: 8401 (pIJ1687) (NodC17–PhoA); b: 8401 (pIJ1924) (NodC59–PhoA); c: 8401 (pIJ1923) (NodC85–phoA), d: 8401 (pIJ1704) (NodC204–PhoA); e: 8401 (pIJ1954) (NodC348–PhoA); f: 8401 (pIJ1953) (NodC359–PhoA); g: 8401 (pIJ1922) (NodC400–PhoA); h: 8401 (pIJ7248) (NodC185′–PhoA–′186NodC); and i: 8401 (pIJ1887) (NodC). The different fusions were detected using a PhoA antiserum (A) or a NodC antiserum (B). The sizes of molecular mass markers (kDa) are indicated.

logical tests (see below), indicating that this fusion of PhoA does not interfere with the normal NodC structure. Since the C-terminus is periplasmically located, it follows that the short hydrophilic domain extending from residues 348 to 359 is probably located on the cytoplasmic side of the membrane with hydrophobic peak 4 acting as a transmembrane span. This was confirmed by constructing two fusions, one at each end of the short hydrophilic domain. Plasmids pIJ1953 and pIJ1954 contain *phoA* fusions after amino acids 359 and 348 at either end of the short hydrophilic region. NodC348–PhoA (on pIJ1954) is adjacent to hydrophobic peak 3 and NodC359–PhoA (on pIJ1953) is adjacent to hydrophobic peak 4 (Fig. 2). When strains 8401 (pIJ1953) and 8401 (pIJ1954) were plated on Y medium containing hesperetin and XP, the colonies remained white. Quantitative measurements showed hesperetin-inducible alkaline phosphatase activity at levels less than 10% of that observed with the NodC400–PhoA fusion (Fig. 2). This confirmed that the small hydrophilic loop extending over residues 348–359 is located in the cytoplasm. Protein extracts of strains 8401 (pIJ1953) and 8401 (pIJ1954) were stained with NodC antiserum. As shown in Fig. 3B (lanes e and f), fusion proteins of the expected sizes were detected. Staining with the PhoA antiserum also identified bands of the expected sizes (Fig. 3A, lanes e and f). In contrast to what was seen with the NodC400–PhoA fusion, there was relatively little proteolysis or accumulation of the protein thought to be the free alkaline phosphatase moiety (Fig. 3A, lanes e and f), indicating that cleavage releasing this moiety may only occur when the PhoA region of the hybrid protein is translocated across the membrane. Such specific degradation would be consistent with proteolysis by periplasmically located proteases (possibly including a leader peptidase).

Analysis of fusions in the large hydrophilic region of NodC

A strain carrying pIJ1477 (Rossen *et al*., 1985), in which NodC is fused to LacZ within the large hydrophilic region (at residue 158), has high levels of β-galactosidase activity after growth in the presence of hesperetin (Fig. 2). This indicates that the hydrophilic domain extending from residue 22 to 303 is unlikely to be periplasmic, since low levels of β-galactosidase are expected when LacZ is fused to domains of proteins that are translocated across the membrane (Silhavy and Beckwith, 1985). To check that the *nodC–lacZ* gene fusion constructed on pIJ1477 by Rossen *et al*. (1985) was correct, strain 8401 (pIJ1477) containing *nodD* on pIJ1518 was grown in the presence of hesperetin, and a protein extract was stained with NodC antiserum after SDS–PAGE. A stained band of the expected size for the NodC158–LacZ fusion was detected, confirming that the fusion was as expected (Fig. 4).

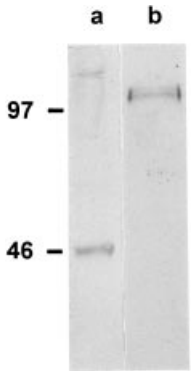

Fig. 4. Identification of the NodC–LacZ fusion protein. Cell extracts (about 50 μg protein) from 8401 (pIJ1887) (NodC) or 8401 (pIJ1518, pIJ1477) (NodC–LacZ) were loaded on a SDS–polyacrylamide gel and transferred to nitrocellulose. The native NodC (lane a) and NodC–LacZ fusions (lane b) were detected using NodC antiserum (Barny and Downie, 1993). The sizes of molecular mass markers (kDa) are indicated.

Four PhoA fusions within the large hydrophilic loop were identified following the Tn*PhoA* mutagenesis of pIJ1089 and pIJ1887. DNA sequencing of the fusion functions in plasmids pIJ1687, pIJ1924, pIJ1923 and pIJ1704 showed that the fusions were in frame and located after amino acids 17 (NodC17–PhoA), 59 (NodC59–PhoA), 85 (NodC85–PhoA) and 204 (NodC204–PhoA), respectively. Strain 8401 carrying each of these plasmids formed dark blue colonies on plates containing XP and hesperetin to induce *nod* gene expression. Quantitative measurements confirmed strong induction by hesperetin of the *nodC–phoA* gene fusions (Fig. 2). NodC antiserum recognized a protein of the appropriate size for the NodC204–PhoA fusions (Fig. 3B, lane d) but, as expected, did not react with the other three fusions (data not shown) because the first 85 amino acids of NodC were absent from the protein used to raise the NodC antiserum (Barny and Downie, 1993). With the PhoA antiserum, the sizes of the largest bands were in good agreement with the predicted sizes of the NodC–PhoA fusion proteins (Fig. 3A, lanes a–d). However, as observed with the NodC400–PhoA fusion, several stained bands were identified and a major product was of similar size to mature alkaline phosphatase.

The high levels of alkaline phosphatase activity observed with these four fusions taken together with the proteolysis (which only appears to occur at high levels when the PhoA region of the fusion proteins is located in the periplasm) indicates that the N-terminal domain of NodC can direct the translocation of these fusions across the inner membranes into the periplasm. However, the observations made with the NodC–LacZ fusion indicates that the large hydrophilic domain is on the cytoplasmic

side of the membrane, rather than the periplasmic side of the membrane as indicated by the results obtained with these four NodC–PhoA fusions.

This apparent contradiction between the conclusions drawn with the NodC–LacZ fusion and the NodC–PhoA fusions suggests that the absence of the C-terminal part of NodC (in the NodC–PhoA fusions) might have caused the proteins to adopt the wrong membrane topology. This would occur if the C-terminal part of NodC were involved in directing NodC folding in the membrane. To test this, a sandwich fusion was constructed in which *phoA* was fused in frame at both ends within that part of *nodC* encoding the large hydrophilic loop, such that the C-terminal part of NodC was conserved.

Plasmid pIJ7248 was constructed *in vitro* to encode a fusion protein containing PhoA spliced into NodC between amino acids 185 and 186 (Fig. 2). A protein extract from strain 8401 (pIJ7248) grown in the presence of hesperetin was stained with NodC antiserum and a protein of the expected size was detected (Fig. 3B, lane h). Colonies of 8401 (pIJ7248) grown on agar medium containing hesperetin and XP were white. In liquid cultures grown with hesperetin, quantitative measurements revealed that 8401 (pIJ7248) does induce alkaline phosphatase activity but the level of activity is low, less than 10% of that observed with the active PhoA fusions (Fig. 2), indicating that this domain of NodC is now cytoplasmic. During the course of making the sandwich fusion, a NodC185–PhoA fusion was also identified (lacking the C-terminal NodC region). Since strain 8401 carrying this fusion formed blue colonies and induced similar high levels of alkaline phosphatase to those found for the four other NodC–PhoA fusions in this domain, it follows that the C-terminal part of NodC influences the folding of the sandwich fusion protein. Staining with the PhoA antiserum (Fig. 3, lane h) identified a protein of the expected size of the sandwich fusion protein and did not detect the extensive protein degradation or appearance of a protein corresponding in size with mature PhoA as had been observed with the highly active PhoA fusions. This is consistent with the hypothesis that the alkaline phosphatase region of the NodC–PhoA–NodC hybrid is not translocated into the periplasm where it would be exposed to proteolysis.

Symbiotic phenotype of the NodC–PhoA fusions

All the *nodC–phoA* fusions constructed are carried on plasmids derived from either pIJ1089 or pIJ1887, both of which contain the *nodABCDEF* genes, which are sufficient to make a Nod factor that induces root-hair deformation (Rossen *et al*., 1984; Spaink *et al*., 1991). When tested on *Vicia hirsuta,* both strains 8401 (pIJ1089) and 8401 (pIJ1887) induced strong root-hair deformation. To test the effect of the different mutations in *nodC*, derivatives of 8401 carrying pIJ1687, pIJ1704, pIJ1922, pIJ1923, pIJ1924, pIJ1953, pIJ1954 or pIJ7248 were tested for root-hair deformation on *V. hirsuta*. However, of all the strains tested, only 8401 (pIJ1922) (expressing NodC400–PhoA) induced strong root-hair deformation, showing that the last 24 amino acids of NodC are dispensable for NodC function. This result, taken together with the lack of activity observed with strain 8401 (pIJ1954) (encoding the NodC359–PhoA fusion), indicates that the region of NodC between residues 359 and 400, which includes the fourth proposed transmembrane domain, is essential for function. The three Tn*phoA* fusions *nodC17–phoA*, *nodC204–phoA* and *nodC400–phoA*, carried on plasmids pIJ1687, pIJ1704 and pIJ1922, respectively, were recombined onto pRL1JI. The resulting strains, A131, A132 and A290, were tested for nodulation on pea and vetch. As expected, strains A131 and A132, carrying *nodC17–phoA* and *nodC204–phoA*, were unable to nodulate. However, strain A290, carrying *nodC400–phoA*, nodulated both peas and vetch at a level similar to that in the wild-type strain.

Discussion

The rhizobial NodC proteins belong to a family of processive β-glycosyl transferases that interactively transfer sugar residues to an acceptor, thus forming a carbohydrate chain (Saxena *et al*., 1995). Chitin synthases, including NodC proteins and the *Saccharomyces cerevisiae* Chs1 (Bulawa *et al*., 1986) form a subgroup of β-glycosyl transferases. Other proteins in this family (see Saxena *et al*., 1995) include the cellulose synthase AcsAB from *Acetobacter xylinum* (Saxena *et al*., 1990; 1991; 1994), the hyaluronan synthase HasA from *Streptococcus pyogenes* (DeAngelis *et al*., 1993), alginate synthase Alg8 from *Pseudomonas aeruginosa* (Maharaj *et al*., 1993) and the DG42 protein involved in embryo development in *Xenopus laevis* (Rosa *et al*., 1988).

All of these proteins appear to be integral membrane proteins located in the plasma membrane. Work on NodC from *R. meliloti* by John *et al.* (1985; 1988) and Johnson *et al.* (1989) led to the conclusions that NodC is also located in the outer membrane and might form a receptor. However, Barny and Downie (1993) concluded that NodC is an inner but not an outer membrane protein. It seems likely that the biosynthesis of lipo-chito-oligosaccharide Nod factors occurs at the membrane and, indeed, several of the enzymes involved in the biosynthesis appear to be membrane associated. Significant amounts of NodA (which is involved in attachment of the acyl group) and NodE were found to co-purify with the inner membrane (Egelhoff and Long, 1985; Spaink *et al*., 1989; Schlaman, 1992) although neither of these proteins is in our view likely to be an integral (trans)membrane pro-

tein. One possibility is that the biosynthesis of Nod factors occurs at the membrane where several *nod* gene products might form some kind of a loose biosynthetic complex.

Since the Nod proteins (NodA, NodB, NodF, NodE, NodL, NodH, NodP, NodQ, NodX and others, see Downie, 1994; Carlson *et al*., 1994) involved in modifying the chitin precursor are all probably cytoplasmic, it would be expected that the bulk of the NodC chitin synthase would be on the cytoplasmic side of the membrane. Therefore, we were surprised when we identified a series of active NodC–PhoA fusions that indicated that the large hydrophilic domain of NodC (containing the conserved chitin synthase regions) was in the periplasm. However, the observation that a NodC–LacZ protein fusion in this domain expressed high levels of β-galactosidase suggested this region is indeed cytoplasmic. This was confirmed using a NodC–PhoA–NodC sandwich fusion, which had much lower alkaline phosphatase activity than NodC–PhoA fusions in the same region. Such sandwich fusions have been developed by Ehrmann *et al.* (1990), who showed that PhoA retains its enzymic activity when tethered at both the N- and C-termini.

We conclude that the C-terminal part of NodC is essential to establish the correct orientation of the large hydrophilic domain of NodC. The identification of active PhoA fusions in the large hydrophilic region of NodC implies that, in the absence of the C-terminal hydrophobic domain, the hydrophobic N-terminal domain of NodC can act as an N-terminal non-cleavable leader peptide, located in the membrane in an N_{in}–C_{out} orientation. However, it appears unlikely that it plays such a role when the hydrophobic C-terminal region of NodC is present. This implies that the normal insertion of NodC into the membrane may be initiated at the hydrophobic C-terminal region. The presence of negatively charged residues near the extreme N-terminus of NodC proteins from some rhizobia taken together with our analysis of topology suggests that the N-terminal hydrophobic region is transmembranous and normally inserts in the N_{out}–C_{in} orientation, as seen with the N-terminal membrane-spanning region of leader peptidase (Cao and Dalby, 1994; Gafvelin and von Heijne, 1994). Clearly, analysis of the topology of proteins containing such N-terminal membrane-spanning segments using the C-terminal PhoA fusions could lead to incorrect conclusions.

The large hydrophilic loop of NodC has structural similarity to similar regions in the other β-glucosyl transferases: AcsAB, HasA, Alg8, Chs1 and the DG42 protein. This region has been proposed (Saxena *et al*., 1995) to be made up of an A domain (which extends from about residue 45 to 140 in NodC) consisting of alternating β-sheets and α-helices, thought to be essential for

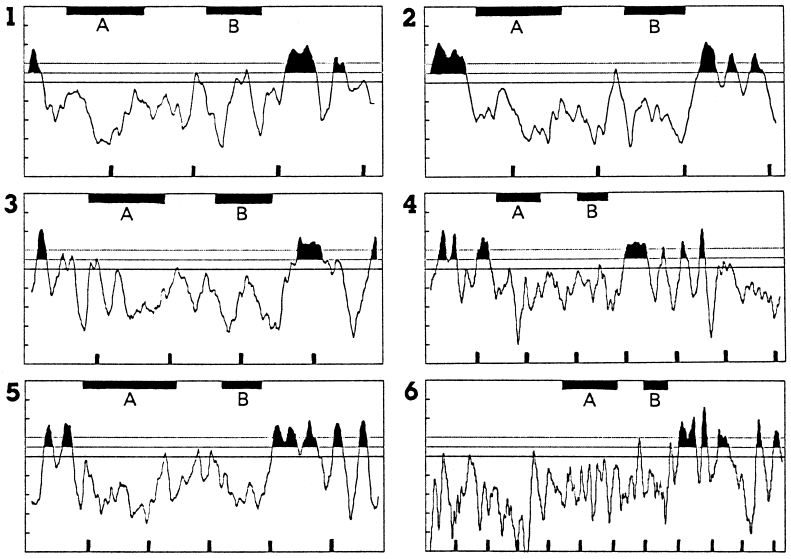

Fig. 5. Comparison of the hydrophobic and hydrophilic domains of NodC with those of processive glycosyl transferase and related proteins. Protein sequences were analysed for hydrophobic regions using the software TopPredII (Claros and von Heijne, 1994). The GES hydrophobicity scale (Engelman *et al*., 1986) was used with the following parameters: full window 21; core window 11; critical length 60. The graphs represent the hydropathy plot obtained with the following protein sequences: 1, *R. l. viciae* NodC (Rossen *et al*., 1984; Swissprot: NODC-RHILV entry, Accession Number PO4340); 2, *S. pyogenes* HasA (De Angelis *et al*., 1993; conceptual translation of EMBL entry SPHASAO, Accession Number L21187); 3. *P. aeruginosa* Alg 8 (Maharaj *et al*., 1993; conceptual translation of EMBL entry PSEALGDB, Accession Number L21611); 4. *A. xylinum* AcsA (Saxena *et al*., 1990; Swissprot: BCSA-ACEXY entry, Accession Number P21877); 5, *X. laevis* Dg42 (Rosa *et al*., 1988; Swissprot: DG42-XENLA entry, Accession Number P13563); 6, *S. cerevisiae* Chs1 (Bulawa *et al*., 1986; Swissprot: CHS1-YEAST entry, Accession Number P08004). The hydrophobic domains above a cut off value of 0.5 are shown in black. The localisation of the A and B domains conserved between the processive glycosyl transferase (Saxena *et al*., 1995) is indicated on the top of each graph. The spacing between the black bars at the bottom of each graph represents 100 amino acids.

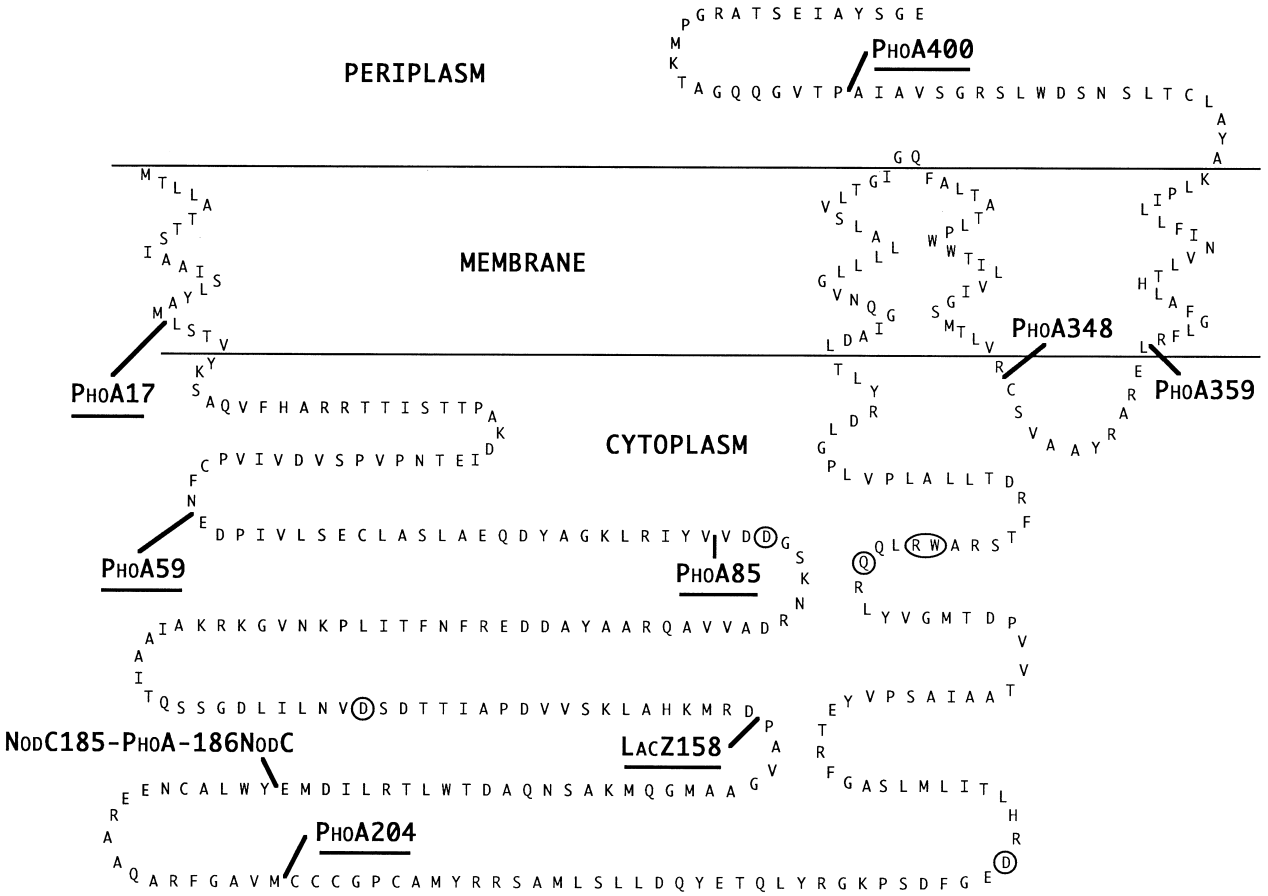

Fig. 6. Model for the topology of the NodC protein in the inner membrane. The sequence shown is of NodC from *R. l. viciae*. The positions of the PhoA and LacZ fusions described in the text are shown along the amino acid sequence. The fusions giving rise to high β-galactosidase and high alkaline phosphatase activity are underlined. The residues conserved among processive β-glycosyl transferases are circled. The transmembrane segments were predicted using the TopPredII software (Claros and von Heijne, 1994).

β-glucosyl transferase activity, and a B domain (corresponding to residues 215–280 of NodC) thought to be responsible for processivity (Fig. 5). In the A domain, two aspartate residues are conserved among all the sequences (residues 88 and 139 of NodC, see Fig. 6); in the B domains one aspartate residue and the motif QXXRW (residue 240 and 276–280 of NodC, see Fig. 6) are also conserved and thought to be crucial for catalytic activity.

Interestingly, all of the processive β-glycosyl transferases share with NodC the presence of a large hydrophobic region immediately following the B domain (Fig. 5). This hydrophobic region consists of multiple hydrophobic peaks made of one large hydrophobic domain followed by one or several smaller hydrophobic peaks. This suggests that all these proteins might have in common with NodC a related membrane topology. Our observation that the PhoA fusions at amino acid positions 348 and 359 block the biological activity of the *R. l. viciae* NodC, whereas PhoA insertions at residue 400 does not, demonstrates that the region of about 40 residues encompassing the proposed fourth transmembrane span is essential for the function of the protein.

Saxena *et al.* (1994) used Tn*phoA* fusions to study the transmembrane nature of the AcsAB cellulose synthase from *A. xylinum*. Several active fusions were identified in the hydrophilic region that they (Saxena *et al.*, 1995) concluded was structurally similar to the large hydrophilic domain of NodC. In our view, it is unlikely that this domain would be on the periplasmic side of the membrane in *A. xylinum* and the cytoplasmic side of the membrane in *R. l. viciae*. In the light of the results presented here, the interpretation of the data with simple AcsAB–PhoA fusions may be more complex than anticipated.

The model for the membrane topology of NodC (Fig. 6) is consistent with NodC being located in the inner membrane, in accord with our previous work (Barny and Downie, 1993). It is evident that the use of PhoA fusions has limitations in the analysis of NodC topology and that the final topology of NodC is dependent upon the protein remaining intact. Based on the results presented here, we propose that a major part of NodC is exposed to the

Table 1. Bacterial strains and plasmids.

Strain/ Plasmid	Relevant characteristics	Source/Reference
Rhizobium		
8401	*R. leguminosarum*, lacking symbiotic plasmid (Str^R)	Lamb *et al.* (1982)
8401/pRL1JI	Derivative of 8401 carrying the bv. *phaseoli* plasmid pRL1JI (Str^R)	Downie *et al.* (1983)
A131	8401 (pRL1JI*nodC17*::Tn*phoA*)	This work
A132	8401 (pRL1JI*nodC204*::Tn*phoA*)	This work
A290	8401 (pRL1JI*nodC400*::Tn*phoA*)	This work
Plasmid		
pIJ1089	pLAFR1 derivative carrying *nodABCDEFIJLMNTO* (Tet^R)	Downie *et al.* (1983)
pIJ1477	*nodC–lacZ* derivative of pIJ1363	Rossen *et al.* (1985)
pIJ1687	Derivative of pIJ1089 carrying *nodC17*::Tn*phoA*	This work
IJ1704	Derivative of pIJ1089 carrying *nodC204*::Tn*phoA* mutation (Tet^R, Kan^R)	This work
pIJ1887	Derivative of pRK290 carrying *nodABCDEF* (Tet^R)	C. Marie
pIJ1922	Derivative of pIJ1887 carrying the *nodC400*::Tn*phoA*	This work
pIJ1923	Derivative of pIJ1887 carrying *nodC85*::Tn*phoA*	This work
pIJ1924	Derivative of pIJ1887 carrying *nodC59*::Tn*phoA*	This work
pIJ1953	Derivative of pIJ1887 carrying *nodC359–phoA*	This work
pIJ1954	Derivative of pIJ1887 carrying *nodC348–phoA*	This work
pPHO7	Plasmid carrying a *phoA* cassette	Gutierrez and Devedjian (1989)
pSWFII	Derivative of pPHO7 carrying a 'sandwich' *phoA* cassette	Ehrmann *et al.* (1990)

cytoplasm, where it could be involved in the biosynthesis of N-acetyl glucosamine oligomers and potentially interact with other *nod* gene products. This model is not consistent with other proposals (John *et al.*, 1985; 1988; Johnson *et al.*, 1989) suggesting that NodC spans the inner and outer membranes. Lipo-chito-oligosaccharides synthesized at the cytoplasmic side of the membrane must be secreted and it has been proposed that a specific membrane translocation complex encoded by the *nodI*, *nodJ* and *nodT* genes may secrete Nod factors across both membranes into the legume rhizosphere (Downie, 1994; McKay and Djordjevic, 1993; Vázquez *et al.*, 1993; Rivilla *et al.*, 1995) where they can initiate the development of nitrogen-fixing nodules.

Experimental procedures

Strains and microbiological techniques

Bacterial strains and plasmids are listed in Table 1 or in the text. Unless otherwise stated, the *Escherichia coli* strain used for transformation was DH5α (Sambrook *et al.*, 1989). General growth conditions for *R. leguminosarum* were as described by Beringer (1974) and appropriate antibiotics were added: streptomycin (200 $\mu g\,ml^{-1}$), tetracycline (1 $\mu g\,ml^{-1}$ in minimal and 5 $\mu g\,ml^{-1}$ in complete medium), kanamycin (20 $\mu g\,ml^{-1}$) and gentamicin (10 $\mu g\,ml^{-1}$). The flavone hesperetin was used at a final concentration of 2 μM. To detect alkaline phosphatase active fusions on plates, XP was added to solid Y medium (Sherwood, 1970) at a final concentration of 40 $\mu g\,ml^{-1}$. Plasmids were mobilized in triparental conjugational matings from *E. coli* to *R. leguminosarum* strains using pRK2013 (Figurski and Helinski, 1979) as helper plasmid. Mutations in the *nodC* gene were transferred to pRLIJI by homologous recombination, as described by Downie *et al.* (1985). Nodulation tests were carried out as described by Knight *et al.* (1986). Tests of root hair curling on *V. hirsuta* were made as described by Bhuvaneswari and Solheim (1985). Alkaline phosphatase assays were performed as described by Cubo *et al.* (1992).

DNA manipulations

Routine DNA manipulations were carried out as described by Sambrook *et al.* (1989). Plasmid pIJ1887 was constructed by subcloning a 6.4 kb *Eco*RI fragment carrying the *nodABCDEF* genes into the *Eco*RI site of pRK290 (Ditta *et al.*, 1980). Polymerase chain reactions (PCR) amplifications were carried out as described by Young *et al.* (1991).

*Isolation and characterization of Tn*phoA *insertions*

Tn*phoA* Manoil and Beckwith (1985) mutagenesis of pIJ1089 and pIJ1887 was carried out as described by Davis and Johnston (1990). To sequence junctions of the *nodC–phoA* gene fusions, DNA from the different plasmids carrying the fusions were used as templates to generate PCR products that span the sites of the fusions. Two primers were used for this purpose: primer 2 binds to the end of *nodB* with its 3′ end toward *nodC* (Fig. 7) and primer 1 binds in *phoA* with its 3′ end toward the edge of Tn*phoA* (Fig. 5). The PCR products generated were purified and sequenced with primer 1, as described by Young *et al.* (1991).

Construction of NodC–PhoA–NodC and NodC–PhoA fusions in vitro

To construct a NodC–PhoA–NodC sandwich fusion, plasmid pSWFII containing the *phoA* cassette (lacking the sequence encoding the N-terminal leader peptide) with polylinkers at both ends was digested with *Sma*I and *Hin*dIII. The 1.3 kb fragment carrying the truncated *phoA* gene was isolated and the 5′ end protruding from the *Hin*dIII site was filled in with the Klenow fragment of DNA polymerase. This fragment was then cloned into pIJ1887 linearized with *Sca*I (Fig. 7).

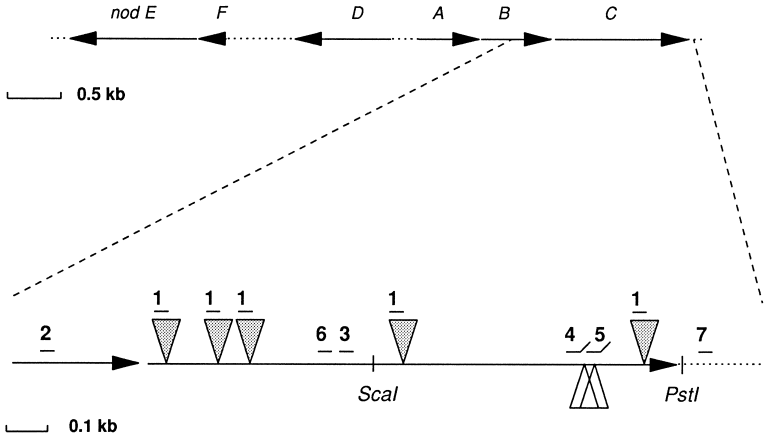

Fig. 7. Description of the construction of NodC–PhoA gene fusions. The top part shows the *nod* genes present on plasmid pIJ1887 and, in the bottom part, the *nodC* region is enlarged. The triangles indicate the location of the different NodC–PhoA fusions that have been constructed The shaded triangles represent the fusions giving rise to a high alkaline phosphatase activity while the open triangles represent the PhoA fusions with low activity. The *Sca*I site and the *Pst*I site that were used in cloning are also indicated. The numbers indicate the different primers that have been used in the PCR and sequencing reactions: primer 1 (5′-CGCCCTGAGCAGCCCGGTTTTCCAG-3′); primer 2 (5′-GAGCCGTTCACTGGTCGGCCGACCC-3′; nucleotide positions 1023 to 1046; Rossen *et al*., 1984); primer 3 (5′-CGGACACCTGGC-TAACTCGC-3′; nucleotide positions 1904–1923; Rossen *et al*., 1984); primer 4 (5′-CACGCTGCAGCGTACAAGAGTCATGGATC-3′); primer 5 (5′-GCACTGCAGCTAAGTTCGCGGGCGCGATAG-3′); primer 6 (5′-CGATGGGCCAAATGAAAGCC-3′; nucleotides 1874–1893; Rossen *et al*., 1984); and primer 7 (5′-CCGATAGGATGAGATGATCC-3′; nucleotides 110–92; Evans and Downie, 1986).

A plasmid containing the insert in the correct orientation was named pIJ7248. The fact that the fusion was in frame at both ends was confirmed by DNA sequencing and by the size of the protein detected when a Western blot was stained with NodC antiserum.

NodC348–PhoA and NodC359–PhoA fusions carried on plasmids pIJ1954 and plasmid pIJ1953, respectively, were generated by a two-step cloning procedure. The first step involved PCR using the following primers: primer 3 binds to the *nodC* sequence immediately upstream of the *ScaI* site (Fig. 7); primers 4 and 5 were designed such that they would create a *Pst*I site in the part of NodC encoding amino acids 348 and 359, respectively (Fig. 7). After PCR amplification using pIJ1887 as template and either primers 3 and 4 or primers 3 and 5, two different PCR products were obtained. These were digested with *Pst*I and *Sca*I and cloned into pIJ1887 digested with the same enzymes (Fig. 7). This generated pIJ1952 and pIJ1951, which both contain a derivative of *nodC* containing a novel *Pst*I site. To check that no mutations were introduced as a result of the PCR reactions, the cloned fragments were sequenced. This was done by generating PCR products with primers 6 and 7 using pIJ1952 and pIJ1951 as template and direct sequencing of these PCR products using primers 6 and 7 (Fig. 7). To generate *nodC–phoA* gene fusions using pIJ1952 and pIJ1951, a 2.6 kb *Pst*I fragment containing a *phoA* cassette from pPHO7 (Gutierrez and Devedjian, 1989) was cloned into the unique *Pst*I sites forming pIJ1954 and pIJ1953, respectively.

Protein electrophoresis and Western blotting

R. leguminosarum strains were grown at 28°C in 100 ml Y medium (containing 10 mM succinate) to an OD_{600} of 0.2. Hesperetin was then added, and the cells grown for a further 10 h. The cells were harvested, resuspended in 0.5 to 1 ml of water, sonicated using an MSE soniprep (5 pulses of 20 s each with 20 s intervals) and unbroken cells were pelleted by centrifugation. The supernatant was mixed with loading buffer and heated at 100°C for 5 min. Samples of 2–10 μl were loaded onto acrylamide gels and proteins were separated by electrophoresis. The gels were then stained with Coomassie brilliant blue or transferred onto nitrocellulose and immunostained with NodC antiserum (Barny and Downie, 1993) or PhoA antiserum, as described by Bradley *et al.* (1988).

Acknowledgements

We thank C. Marie for the gift of pIJ1887, D.A. Hopwood for critically reading the manuscript, and J.M. Sutton for advice and critical comments. We are grateful to P. Götz and W. Boos for the gift of PhoA antiserum. The work was supported by a PMB grant from the BBSRC and a BBSRC–INRA fellowship to E.S., and an EU sectoral training grant to A.E.

References

Atkinson, E.M., Palcic, M.M., Hindsgaul, O., and Long, S.R. (1994) Biosynthesis of *Rhizobium meliloti* lipo-oligosaccharide Nod factors: NodA is required for an *N*-acyl transferase activity. *Proc Natl Acad Sci USA* **91:** 8418–8422.

Barny, M.A., and Downie, J.A. (1993) Identification of the NodC protein in the inner but not the outer membrane of *Rhizobium leguminosarum*. *Mol Plant–Microbe Interact* **6:** 669–672.

Beringer, J.E. (1974) R factor transfer in *Rhizobium leguminosarum*. *J Gen Microbiol* **84:** 188–198.

Bhuvaneswari, T.V., and Solheim, B. (1985) Root hair deformation in the white clover/*Rhizobium trifolii* symbiosis. *Physiol Plant* **63:** 25–34.

Bradley, D.J., Wood, E.A., Larkins, A.P., Galfrè, G., Butcher, G.W., and Brewin, N.J. (1988) Isolation of monoclonal antibodies reacting with peribacteroid membranes and other components of pea root nodules containing *Rhizobium leguminosarum*. *Planta* **173:** 149–160.

Bulawa, C.E., Slater, M., Cabib, E., Au-Young, J., Sburlati, A., Adair, W.L., and Robbins, P.W. (1986) The *S. cerevisiae* structural gene for chitin synthase is not required for chitin synthesis *in vivo*. *Cell* **46:** 213–225.

Cao, G., and Dalbey, R.E. (1994) Translocation of *N*-terminal tails across the plasma membrane. *EMBO J* **19:** 4662–4669.

Carlson, R.W., Price, N.P.J., and Stacey, G. (1994) The biosynthesis of rhizobial lipo-oligosaccharide nodulation signal molecules. *Mol Plant–Microbe Interact* **7:** 684–695.

Claros, M.G., and von Heijne, G. (1994) TopPred II: an improved software for membrane protein structure predictions. *CABIOS Application Notes* **10:** 685–686.

Collins-Emerson, J.M., Terzaghi, E.A., and Scott, D.B. (1990) Nucleotide sequence of *Rhizobium loti nodC*. *Nucl Acids Res* **18:** 6990.

Cubo, T., Economou, A., Murphy, G., Johnston, A.W.B., and Downie, J.A. (1992) Molecular characterization and regulation of the rhizosphere-expressed genes *rhiABCR* that can influence nodulation by *Rhizobium leguminosarum* biovar *viciae*. *J Bacteriol* **174:** 4026–4035.

Davis, E.O., and Johnston, A.W.B. (1990) Analysis of three *nodD* genes in *Rhizobium leguminosarum* bv. *phaseoli*; $nodD_1$ is preceded by *nolE*, a gene whose product is secreted from the cytoplasm. *Mol Microbiol* **4:** 921–923.

DeAngelis, P.L., Papaconstantinou, J., and Weigel, P.H. (1993) Molecular cloning, identification, and sequence of the hyaluronan synthase gene from Group A *Streptococcus pyogenes*. *J Biol Chem* **268:** 19181–19184.

Ditta, G.S., Stanfield, S.W., Corby, H.D.L., and Helinski, D.R. (1980) Broad host range DNA cloning system from Gram-negative bacteria: construction of a gene bank of *Rhizobium meliloti*. *Proc Natl Acad Sci USA* **77:** 7347–7351.

Downie, J.A. (1994) Signalling strategies for nodulation of legumes by rhizobia. *Trends Microbiol* **2:** 318–324.

Downie, J.A., Hombrecher, G., Ma, Q.-S., Knight, C.D., Wells, B., and Johnston, A.W.B. (1983) Cloned nodulation genes of *Rhizobium leguminosarum* determine host-range specificity. *Mol Gen Genet* **190:** 359–365.

Downie, J.A., Knight, C.D., and Johnston, A.W.B. (1985) Identification of genes and gene products involved in nodulation of peas by *Rhizobium leguminosarum*. *Mol Gen Genet* **198:** 255–262.

Egelhoff, T.T., and Long, S.R. (1986) *Rhizobium meliloti* nodulation genes: identification of *nodDABC* gene products, purification of Nod protein, and expression of *nodA* in *Rhizobium meliloti*. *J Bacteriol.* **164:** 591–599.

Ehrmann, M., Boyd, D., and Beckwith, J. (1990) Genetic analysis of membrane protein topology by a sandwich gene fusion approach. *Proc Natl Acad Sci USA* **87:** 7574–7578.

Engelman, D.M., Steitz, T.A., and Goldman, A. (1986) Identifying nonpolar transbilayer helices in amino acid sequences of membrane proteins. *Annu Rev Biophys Chem* **15:** 321–353.

Evans, I.J., and Downie, J.A. (1986) The *NodI* gene product of *Rhizobium leguminosarum* is closely related to ATP-binding bacterial transport proteins: nucleotide sequence analysis of the *nodI* and *nodJ* genes. *Gene* **43:** 95–101.

Figurski, D.H., and Helinski, D.R. (1979) Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided *in trans*. *Proc Natl Acad Sci USA* **76:** 1648–1652.

Gafvelin, G., and von Heijne, G. (1994) Topological 'Frustration' in multispanning E. coli inner membrane proteins. *Cell* **77:** 401–412.

Geremia, R.A., Mergaert, P., Geelen, D., Van Montagu, M., and Holsters, M. (1994) The NodC protein of *Azorhizobium caulinodans* is an *N*-acetylglucosaminyltransferase. *Proc Natl Acad Sci USA* **91:** 2669–2673.

Goethals, K., Gao, M., Tomekpe, K., Van Montagu, M., and Holsters, M. (1989) Common *nodABC* genes in Nod locus 1 of *Azorhizobium caulinodans*: nucleotide sequence and plant-inducible expression. *Mol Gen Genet* **219:** 289–298.

Gutierrez, C., and Devedjian, J.C. (1989) A plasmid facilitating *in vitro* construction of *phoA* gene fusions in *Escherichia coli*. *Nucl Acids Res* **17:** 3999.

Hubac, H.C., Guerrier, D., Ferran, J., Tremoliers, A., and Kondorosi, A. (1992) Lipid and protein composition of outer and inner membranes in wild-type strains and *nod* mutants of *Rhizobium meliloti*. *J Gen Microbiol* **138:** 1973–1983.

Jacobs, T.W., Egelhoff, T.T., and Long, S.R. (1985) Physical and genetic map of a *Rhizobium meliloti* nodulation gene region and nucleotide sequence of *nodC*. *J Bacteriol* **162:** 469–476.

John, M., Schmidt, J., Wieneke, U., Kondorosi, E., Kondorosi, A., and Schell, J. (1985) Expression of the nodulation gene nodC of *Rhizobium meliloti* in *Escherichia coli*: role of the *nodC* gene product in nodulation. *EMBO J* **4:** 2425–2430.

John, M., Schmidt, J., Wieneke, U., Krussmann, H.-D., and Schell, J. (1988) Transmembrane orientation and receptor-like structure of the *Rhizobium meliloti* common nodulation protein NodC. *EMBO J* **7:** 583–588.

John, M., Röhrig, H., Schmidt, J., Wieneke, U., and Schell, J. (1993) *Rhizobium* NodB protein involved in nodulation signal synthesis is a chitooligosaccharide deacetylase. *Proc Natl Acad Sci USA* **90:** 625–629.

Johnson, D., Roth, L.E., and Stacey, G. (1989) Immunogold localization of the NodC and NodA proteins of *Rhizobium meliloti*. *J Bacteriol* **171:** 4583–4588.

Knight, C.D., Rossen, L., Robertson, J.G., Wells, B., and Downie, J.A. (1986) Nodulation inhibition by *Rhizobium leguminosarum* multicopy *nodABC* genes and analysis of early stages of plant infection. *J Bacteriol* **166:** 552–558.

Krishnan, H.B., and Pueppke, S.G. (1991) Sequence and analysis of the *nodABC* region of *Rhizobium fredii* USDA257, a nitrogen-fixing symbiont of soybean and other legumes. *Mol Plant–Microbe Interact* **4:** 512–520.

Lamb, J.W., Hombrecher, G., and Johnston, A.W.B. (1982) Plasmid-determined nodulation and nitrogen-fixation abilities in *Rhizobium phaseoli*. *Mol Gen Genet* **186:** 449–452.

McKay, I.A., and Djordjevic, M.A. (1993) Production and secretion of Nod metabolites by *Rhizobium leguminosarum* biovar *trifolii* are disrupted by the same environmental factors that reduce nodulation in the field. *Appl Environ Microbiol* **59:** 3385–3392.

Maharaj, C., May, T.B., Wang, S.-K., and Chakrabarty, A.M. (1993) Sequence of the *alg8* and *alg44* genes involved in

the synthesis of alginate by *Pseudomonas aeruginosa*. *Gene* **136:** 267–269.
Manoil, C., and Beckwith, J. (1985) Tn*phoA*: a transposon probe for protein export signals. *Proc Natl Acad Sci USA* **82:** 8129–8133.
Manoil, C., and Beckwith, J. (1986) A genetic approach to analysing membrane protein topology. *Science* **233:** 1403.
Rivilla, R., Sutton, J.M., and Downie, J.A. (1995) *Rhizobium leguminosarum* NodT is related to a family of outer-membrane transport proteins that includes TolC, PrtF, CyaE and AprF. *Gene* **161:** 27–31.
Röhrig, H., Schmidt, J., Wieneke, U., Kondorosi, E., Barlier, I., Schell, J., and John, M. (1994) Biosynthesis of lipooligosaccharide nodulation factors: *Rhizobium* NodA protein is involved in *N*-acylation of the chitooligosaccharide backbone. *Proc Natl Acad Sci USA* **91:** 3122–3126.
Rosa, F., Sargent, T.D., Rebbert, M.L., Michaels, G.S., Jamrich, M., Grunz, H., Jonas, E., Winkles, J.A., and David, I.B. (1988) Accumulation and decay of DG42 gene products follow a gradient pattern *Xenopus* embryogenesis. *Devel Biol* **129:** 114–123.
Rossen, L., Johnston, A.W.B., and Downie, J.A. (1984) DNA sequence of the *Rhizobium leguminosarum* nodulation genes *nodA, B* and *C* required for root hair curling. *Nucl Acids Res* **12:** 9497–9508.
Rossen, L., Shearman, C.A., Johnston, A.W.B., and Downie, J.A. (1985) The *nodD* gene of *Rhizobium leguminosarum* is autoregulatory and in the presence of plant exudate induces the *nodABC* genes. *EMBO J* **4:** 3369–3373.
Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989) *Molecular Cloning. A Laboratory Manual*, 2nd edn. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.
Saxena, I.M., Chyr Lin, F., and Brown, Jr, R.M. (1990) Cloning and sequencing of the cellulose synthase catalytic subunit gene of *Acetobacter xylinum*. *Plant Mol Biol* **15:** 673–683.
Saxena, I.M., Chyr Lin, F., and Brown, Jr, R.M. (1991) Identification of a new gene in an operon for cellulose biosynthesis in *Acetobacter xylinum*. *Plant Mol Biol* **16:** 947–954.
Saxena, I.M., Kudlicka, K., Okuda, K., and Brown, Jr, R.M. (1994) Characterization of genes in the cellulose-synthesizing operon (*acs* operon) of *Acetobacter xylinum*: implications for cellulose crystallization. *J Bacteriol* **176:** 5735–5752.
Saxena, I.M., Brown, Jr, R.M., Febre, M., Geremia, R.A., and Henrissat, B. (1995) Multidomain architecture of β-glycosyl transferases: implications for mechanism of action. *J Bacteriol* **177:** 1419–1424.
Schlaman, W.R.M. (1992). Regulation of nodulation gene expression in *Rhizobium leguminosarum* biovar *viceae*, PhD Thesis, University of Leiden.
Schmidt, J., John, M., Wieneke, U., Krussmann, H.-D., and Schell, J. (1986) Expression of the nodulation gene *nodA* in *R. meliloti* and localization of the gene product in the cytosol. *Proc Natl Acad Sci USA* **93:** 9581–9585.
Schofield, P.R., and Watson, J.M. (1986) DNA sequence of *Rhizobium trifolii* nodulation genes reveals a reiterated and potentially regulatory sequence preceding *nodABC* and *nodFE*. *Nucl Acids Res* **14:** 2891–2903.
Scott, K.F. (1986) Conserved nodulation genes from the non-legume symbiont *Bradyrhizobium* sp. (Parasponia). *Nucl Acids Res* **14:** 2905–2919.
Sherwood, M.T. (1970) Improved synthetic medium for the growth of *Rhizobium*. *J Appl Bacteriol* **33:** 708–713.
Silhavy, T.J., and Beckwith, J.R. (1985) Use of *lac* fusions for the study of biological problems. *Microbiol Rev* **49:** 398–418.
Spaink, H.P., Weinman, J., Djordjevic, M.A., Wijffelman, C.A., Okker, R.J.H., and Lugtenberg, B.J.J. (1989) Genetic analysis and cellular localization of the *Rhizobium* host specificity-determining NodE protein. *EMBO J* **8:** 2811–2818.
Spaink, H.P., Sheeley, D.M., Van Brussel, A.A.N., Glushka, J.N., York, W.S., Tak, T., Geiger, O., Kennedy, E.P., Reinhold, V.N., and Lugtenberg, B.J.J. (1991) A novel, highly unsaturated, fatty acid moeity of lipo-oligosaccharide signals determines host specificity of *Rhizobium leguminosarum*. *Nature* **354:** 125–130.
Spaink, H.P., Wijfjes, A.H.M., van der Drift, K.M.G.M., Haverkamp, J., Thomas-Oates, J.E., and Lugtenberg, B.J.J. (1994) Structural identification of metabolites produced by the NodB and NodC proteins of *Rhizobium leguminosarum*. *Mol Microbiol* **13:** 821–831.
Török, I., Kondorosi, E., Stepkowski, T., Posfai, J., and Kondorosi, A. (1984) Nucleotide sequence of *Rhizobium meliloti* nodulation genes. *Nucl Acids Res* **12:** 9509–9524.
Truchet, G., Roche, P., Lerouge, P., Vasse, J., Camut, S., de Billy, F., Prome, J.-C., and Dénarié, J. (1991) Sulphated lipo-oligosaccharide signals of *Rhizobium meliloti* elicit root nodule organogenesis in alfalfa. *Nature* **351:** 670–673.
Vázquez, M., Dávalos, A., de las Peñas, A., Sánchez, F., and Quinto, C. (1991) Novel organization of the common nodulation genes in *Rhizobium leguminosarum* bv. *phaseoli* strains. *J Bacteriol* **173:** 1250–1258.
Vázquez, M., Santana, O., and Quinto, C. (1993) The NodI and NodJ proteins from *Rhizobium* and *Bradyrhizobium* strains are similar to capsular polysaccharide secretion proteins from Gram-negative bacteria. *Mol Microbiol* **8:** 369–377.
Young, J.P.W., Downer, H.L., and Eardly, B.D. (1991) Phylogeny of the phototrophic *Rhizobium* strain BTAi1 by polymerase chain reaction-based sequencing of a 16S rRNA gene segment. *J Bacteriol* **173:** 2271–2277.