Universiteit Leiden
The Netherlands

Functional analysis of chimeras derived from the Sinorhizobium meliloti and Mesorhizobium loti nodC genes identifies regions controlling chitin oligosaccharide chain length

Kamst, E.; Breek, C.K.D.; Spaink, H.P.

Citation

Kamst, E., Breek, C. K. D., & Spaink, H. P. (2000). Functional analysis of chimeras derived from the Sinorhizobium meliloti and Mesorhizobium loti nodC genes identifies regions controlling chitin oligosaccharide chain length. *Molecular And General Genetics*, *264*(1-2), 75-81. doi:10.1007/s004380000281

Version: Publisher's Version
License: Licensed under Article 25fa Copyright Act/Law (Amendment Taverne)
Downloaded from: https://hdl.handle.net/1887/3677006

Note: To cite this publication please use the final published version (if applicable).

Mol Gen Genet (2000) 264: 75–81
Digital Object Identifier (DOI) 10.1007/s004380000281

ORIGINAL PAPER

E. Kamst · C. K. D. Breek · H. P. Spaink

Functional analysis of chimeras derived from the *Sinorhizobium meliloti* and *Mesorhizobium loti nodC* genes identifies regions controlling chitin oligosaccharide chain length

Received: 4 February 2000 / Accepted: 8 May 2000 / Published online: 10 June 2000
© Springer-Verlag 2000

Abstract The rhizobial nodulation gene *nodC* encodes an *N*-acetylglucosaminyltransferase that is responsible for the synthesis of chitin oligosaccharides. These oligosaccharides are precursors for the synthesis of the lipo-chitin oligosaccharides that induce cell division and differentiation during the development of nitrogen-fixing root nodules in leguminous plants. The NodC proteins of *Mesorhizobium loti* and *Sinorhizobium meliloti* yield chitinpentaose and chitintetraose as their main products, respectively. In order to localize regions in these enzymes that are responsible for this difference in product chain length, a set of six chimeric enzymes, comprising different combinations of regions of the NodC proteins from these two bacteria, was expressed in *Escherichia coli*. The oligosaccharides produced were analyzed using thin-layer chromatography. The major conclusion from this work is that a central 91-amino acid segment does not play any obvious role in determining the difference in the chain length of the major product. Furthermore, the characteristically predominant synthesis of chitintetraose by *S. meliloti* NodC is mainly dependent on a C-terminal region of maximally 164 amino acids; exchange of only this C-terminal region is sufficient to completely convert the *M. loti* chitinpentaose synthase into an *S. meliloti*-like chitintetraose synthase. The N-terminal region of 170 amino acids also plays a role in restricting the length of the major product to a tetramer. However, the role of the C-terminal region is clearly dominant, since exchanging the N-terminal region has no effect on the relative amounts of chitintetraose and -pentaose produced when the C-terminal region of *S. meliloti* NodC is present. The length of a predicted β-strand around residue 300 in the C-terminal region of various NodC proteins is the only structural element that seems to be related to the length of the chitin oligosaccharides produced by these enzymes; the higher the amount of chitintetraose relative to chitinpentaose, the shorter the predicted β-strand. This element may therefore be important for the effect of the C-terminal 164 amino acids on chitin oligosaccharide chain length.

Key words *Sinorhizobium meliloti* · *Mesorhizobium loti* · NodC · Chimeric proteins · Oligosaccharide chain length

Communicated by A. Kondorosi

E. Kamst[1] · C. K. D. Breek · H. P. Spaink (✉)
Clusius Laboratory, Institute of Molecular Plant Sciences,
Leiden University, Wassenaarseweg 64, 2333 AL Leiden,
The Netherlands
E-mail: Spaink@rulbim.leidenuniv.nl
Tel.: +31-71-5275055; Fax: +31-71-5275088

Present address:
[1] DSM anti-infectives, P.O. Box 1, 2600 MA Delft,
The Netherlands

Introduction

Rhizobia are soil bacteria that are able to invade the roots of leguminous plants and induce the formation of a symbiotic organ, the root nodule (Spaink 1996). In the nodule, rhizobia differentiate into bacteroids, which are able to convert atmospheric nitrogen into the biologically utilizable nitrogen source ammonia. Organogenesis in the plant root is elicited by rhizobial lipo-chitin oligosaccharides (LCOs) (Spaink 1996) which are synthesized by the combined action of several nodulation (Nod) proteins (Kamst et al. 1998). The basic structure of the LCOs is an oligosaccharide of β-1,4-linked *N*-acetylglusosamine (GlcNAc) residues, in which the *N*-acetyl group of the non-reducing terminal residue is replaced by a fatty acid. Such a basic LCO is synthesized by the combined activities of the chitin oligosaccharide synthase NodC, the chitin oligosaccharide deacetylase NodB, and the presumed acyltransferase NodA (Kamst et al. 1998). Rhizobia typically produce a mixture of LCOs. The species-specific variations that have been reported are the presence of additional groups on the oligosaccharide moiety and linkage of the oligosaccharide to a unique α,β polyunsaturated fatty acid, instead of a common one. A further variable is the

length of the oligosaccharide, which can vary between two and six GlcNAc residues, although most LCOs have a chitintetraose or chitinpentaose backbone. The majority of LCOs produced by *Sinorhizobium meliloti* have a chitintetraose backbone (Lerouge et al. 1990; Roche et al. 1991; Schultze et al. 1992). In contrast, *Mesorhizobium loti* primarily produces LCOs with a chitinpentaose backbone (López-Lara et al. 1995). It has been shown that NodC determines these differences in chain length (Roche et al. 1996; Kamst et al. 1997).

Although the enzymatic activity (Geremia et al. 1994; Spaink et al. 1994; Mergaert et al. 1995; Kamst et al. 1997) and topology (Barny et al. 1996) of NodC have been described, the function of the different domains has so far been a subject of speculation based upon protein sequence alignments and secondary-structure predictions (Saxena et al. 1995; Kamst and Spaink 1999). In the present paper we describe the construction and functional analysis of *nodC* chimeras in which a range of analogous regions from the *nodC* genes of the two species have been exchanged. The results show that the central part of NodC does not play a major in chain length determination and that the C-terminal region of the *S. meliloti* NodC is sufficient to convert *M. loti* NodC into an *S. meliloti*-like chitintetraose synthase.

Materials and methods

Bacterial strains and plasmids

The bacterial strains and plasmids used in this study, together with their relevant characteristics are listed in Table 1. *E. coli* strains derived from XL1-Blue (Stratagene) were routinely grown at 37 °C in LC medium to which kanamycin (50 μg/ml), carbenicillin (analogue of ampicillin, 100 μg/ml), or tetracycline (250 μg/ml) was added when appropriate.

Cloning procedures

An overview of the plasmids used in this study is given in Table 1. Only two unique restriction sites (*Nhe*I and *Bgl*I) are conserved in the *nodC* genes of *M. loti* and *S. meliloti* (Fig. 1). In order to use these sites for the construction of chimeras, we introduced unique restriction sites at the 5′- and 3′-termini of these genes using PCR. In addition, the cloning vector pOK12 (Vieira and Messing 1991) was adapted by removing the *Bgl*I fragment, filling in the resulting 5′ overhangs using the Klenow fragment of DNA polymerase, and recircularizing the resulting fragment. This procedure yielded plasmid pMP3520, in which the restriction sites required to construct the *nodC* chimeras remain unique after insertion of the parental *nodC* genes. Amplification of *nodC* sequences was performed as described previously (Kamst et al. 1997). The oligonucleotides oMP151 [5′-TTTAGATCTCTCGAGCTATTGCTGTTCGCTGTAAG-3′] and oMP152[5′-TTTGGATCCCTCGAGTCACTCGCCGCTGCAAGTGC-3′] contain a *Bgl*II and *Bam*HI restriction site, respectively, and both contain an *Xho*I restriction site to allow cloning of the amplification products into the expression vector pET9a, and subsequent transfer of the *Nde*I-*Xho*I fragment into the cloning vector pMP3520 (Fig. 1). These primers were purchased from Isogen Bioscience (Maarsen, Netherlands). The other oligonucleotides have been described before (Kamst et al. 1997). Chimeric *nodC* genes were constructed, using pMP3531 and pMP3532 (Fig. 1) as parental plasmids, by exchange of the *Nhe*I-*Xho*I fragments, resulting in the formation of chimeras A and D (Fig. 2), and exchange of the *Nde*I-*Bgl*I fragment resulting in chimeras C and F. Chimera E was obtained by ligating the *Nhe*I-*Bgl*I fragment of pMP3532 in the corresponding sites of pMP3531. Introduction of the *Nde*I-*Bgl*I fragment of chimera F into the parental *M. loti nodC* gene in plasmid pMP3532 yielded chimera B. Restriction analysis using *Pst*I, *Sca*I and *Hin*dIII was used to confirm the identity of the chimeras. Both the parental genes and the chimeras were then transferred to the *E. coli* expression vector pET3a as *Nde*I-*Xho*I fragments.

Table 1 Plasmids used in this study

Plasmids	Characteristics	Reference
Vectors		
pOK12	Cloning vector containing *lacZ* promoter, and multiple cloning site in α-peptide coding region. Km^R	(Vieira and Messing 1991)
pMP3520	pOK12, Δ*Bgl*I fragment (containing the α-peptide coding region). Km^R	This work
pET3a	T7 promoter-based expression vector. Amp^R	(Studier et al. 1990)
pET9a	T7 promoter-based expression vector. Km^R	(Studier et al. 1990)
Constructs		
pMP3521	*S. meliloti* strain 1021 *nodC* in pET9a	This work
pMP3522	*M. loti* strain NZP2037 *nodC* in pET9a	This work
pMP3531	*S. meliloti* strain 1021 *nodC* in pMP3520	This work
pMP3532	*M. loti* strain NZP2037 *nodC* in pMP3520	This work
pMP3551	*S. meliloti* strain 1021 *nodC* in pET3a	This work
pMP3552	*M. loti* strain NZP2037 *nodC* in pET3a	This work
pMP3534	*nodC* chimera A in pMP3520	This work
pMP3536	*nodC* chimera B in pMP3520	This work
pMP3533	*nodC* chimera C in pMP3520	This work
pMP3537	*nodC* chimera D in pMP3520	This work
pMP3535	*nodC* chimera E in pMP3520	This work
pMP3538	*nodC* chimera F in pMP3520	This work
pMP3554	*nodC* chimera A in pET3a	This work
pMP3556	*nodC* chimera B in pET3a	This work
pMP3553	*nodC* chimera C in pET3a	This work
pMP3557	*nodC* chimera D in pET3a	This work
pMP3555	*nodC* chimera E in pET3a	This work
pMP3558	*nodC* chimera F in pET3a	This work

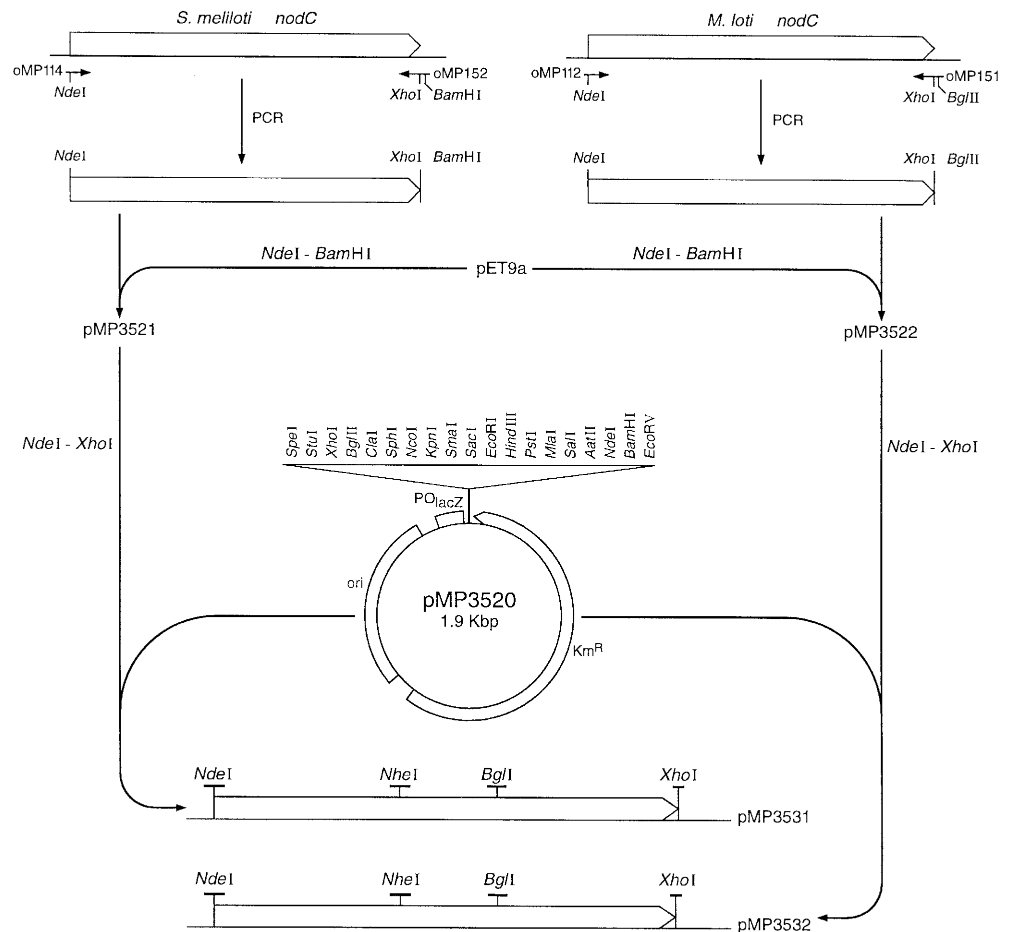

Fig. 1 Subcloning of *S. meliloti* and *M. loti nodC* genes. *Nde*I and *Xho*I restriction sites were introduced at the beginning and the end of the *nodC* ORFs using PCR. The resulting fragments were cloned in the expression vector pET9a, and transferred to the cloning vector pMP3520. This cloning vector was obtained by deleting the *Bgl*I fragment, containing the 3′ portion of the *lacZ* α-peptide coding region, from the vector pOK12 (Vieira and Messing 1991). The *underlined* restriction sites in the resulting *nodC* contructs pMP3521 and pMP3522 are unique to these plasmids. Chimeric *nodC* genes, and the parental *nodC* genes from plasmids pMP3521 and 3522 were transferred as *Nde*I-*Xho*I fragments to the *E. coli* expression vector pET3a (Studier et al. 1990), resulting in expression of the parental NodCs from *M. loti* and *S. meliloti*, and the chimeric NodC proteins depicted in Fig. 2A

Analysis of chitin oligosaccharide production by NodC in *E. coli*

E. coli strains carrying either a *nodC*-containing plasmid or the vector pET3a without an insert were grown overnight at 37 °C in LC medium in the presence of kanamycin (50 µg/ml), diluted 1:100 in fresh medium and grown to an OD_{620} of between 0.1 and 0.2. To 1 ml of these cultures, 0.2 µCi of [1-^{14}C]$GlcNH_2$ was added, and *nodC* expression was induced by infection with phage mGP1-2, encoding the T7 RNA polymerase as described previously (Kamst et al. 1997). After incubation at 28 °C for 2 h, bacteria were collected by centrifugation and extracted with 200 µl of a 1:1 (vol:vol) mixture of chloroform and water. The aqueous phase was dried and the material was dissolved in 10 µl of water. A volume of 1 µl of this sample was analyzed by thin-layer chromatography (TLC). Chitin oligosaccharides were analyzed using NH_2-TLC plates (Merck, Darmstadt, Germany), using acetonitrile:water (7:3 vol/vol) as the mobile phase. Results were visualized using a PhosphorImager system in combination with the ImageQuant software (Molecular Dynamics, Sunnyvale, Calif.).

Results and discussion

Construction and functional analysis of chimeric NodC proteins

The NodC proteins of *S. meliloti* and *M. loti* produce chitin oligosaccharides having a chain length of up to five GlcNAc residues. The main product of *S. meliloti* is chitintetraose (synthesized in about tenfold excess over the pentaose), whereas *M. loti* NodC predominantly synthesizes chitinpentaose (in about tenfold excess over the tetraose) (Kamst et al. 1997). In order to localize regions that determine this difference between the enzymes, we made chimeras by exchanging three corresponding regions from the genes from the two bacteria (Fig. 1, Table 1). The synthesis of chitin oligosaccharides by the parental and chimeric NodC proteins was studied by expressing the genes in the *E. coli* strain XL1-blue (Stratagene). Gene expression was induced in the presence of radiolabeled glucosamine, resulting in the metabolic labeling of chitin oligosaccharides synthesized in vivo, as reported previously (Kamst et al. 1997). Radiolabelled oligosaccharides were analyzed using thin-layer chromatography (TLC, Fig. 2). The NodC enzymes encoded by the subcloned parental genes showed similar levels of chitin oligosaccharide formation, and produced the same distribution of oligosaccharide chain lengths as reported previously for similar constructs (Kamst et al. 1997), indicating that mutations affecting enzyme activity or oligosaccharide chain length control had not been introduced during the subcloning procedure. All chimeras were shown to be enzymatically active. The total amount of chitin oligosaccharides produced by strains expressing chimeras C and D

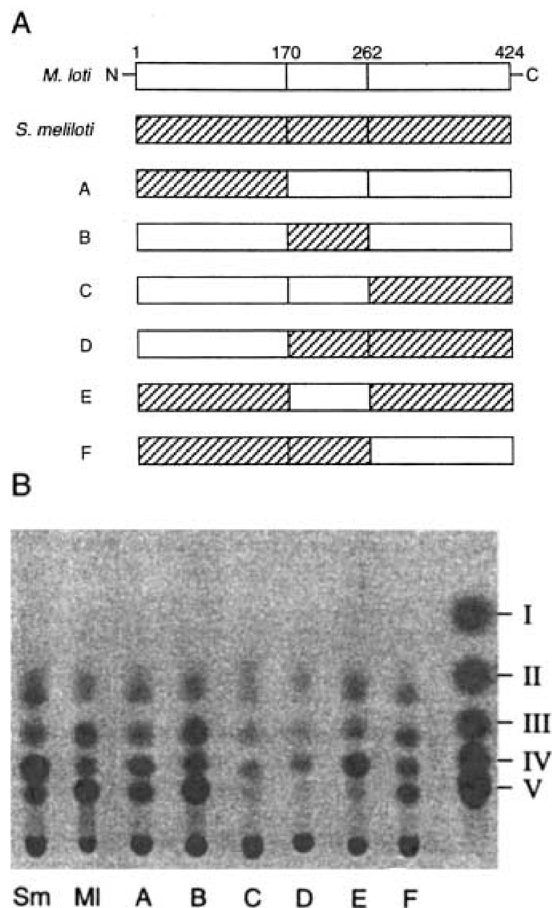

Fig. 2A, B TLC analysis of chitin oligosaccharides produced by NodC proteins in *E. coli*. **A** Schematic representation of wild-type *M. loti* (Ml) NodC, *S. meliloti* (Sm) NodC, and the constructed chimeras. **B** Chitin oligosaccharides were extracted from *E. coli* cells after induction of NodC-expresssion in the presence of ^{14}C-labeled glucosamine in the medium. The extracts were analyzed on thin-layer chromatography (TLC) plates and the result was visualized using a PhosphoImager in combination with the ImageQuant software. The *Roman numerals* indicate the position of the radiolabelled *N*-acetylglucosamine (I) and the chitin oligosaccharides (II–V), prepared using NodC in vitro (Kamst et al. 1997). Lanes A–F show the reaction products of the corresponding NodC chimeras. In a control experiment with the empty vector the amounts of radioactivity measured at the tetramer and pentamer positions were negligible (data not shown)

(Fig. 2) was clearly reduced as compared to the total amounts produced by the parental NodC proteins. Although we cannot exclude the possibility that this is due to differences in expression levels, it is unlikely that only these two chimeras are expressed at a lower level than the parental NodC proteins. Since chimeras C and D are the only proteins in which the N-terminal region of *M. loti* NodC is combined with the C-terminal region of *S. meliloti* NodC, the reduction in chitin oligosaccharide production rather suggests that a wild-type level of enzyme activity requires a specific interaction between these regions.

The relative amounts of chitintetraose and chitinpentaose produced by the parental NodC proteins and the chimeras was quantified and statistically analyzed (Table 2). The chimeric NodC proteins that are expected are outlined in Fig. 3. They are arranged according the origin of the C-terminal region, which derives either from *S. meliloti* NodC (left panel) or *M. loti* NodC (right panel). The data obtained with these proteins clearly show that the presence of this C-terminal region from *S. meliloti* NodC always results in an increase in the relative amount of chitintetraose produced relative to the amounts produced when the proteins contain the corresponding region from *M. loti* NodC. Moreover, the replacement of only the C-terminal region of *M. loti* NodC by that of *S. meliloti* NodC (chimera C) is sufficient to convert completely the pattern of chitintetraose and chitinpentaose produced from the excess of chitinpentaose normally made by *M. loti* NodC to the excess of chitintetraose characteristic of the parental NodC from *S. meliloti*. This C-terminal region is therefore likely to be the most important factor determining the characteristic excess of chitintetraose over chitinpentaose produced by *S. meliloti* NodC.

Although the replacement of only the C-terminal region of *M. loti* NodC by that of *S. meliloti* NodC is sufficient to convert the *M. loti* chitinpentaose synthase into an *S. meliloti*-like chitintetraose synthase, the reverse was not observed: Chimera F has the C-terminal region of *M. loti* NodC, while the remaining regions originate from *S. meliloti* NodC. Although this chimera does produce more chitinpentaose than chitintetraose,

Table 2 Analysis of chitintetraose and -pentaose production by wild-type and chimeric NodC enzymes

NodC	Number of experiments	Amount of chitintetraose relative to chitinpentaose (%)[a]		
		Mean (SD)	95% confidence interval	Significance groups[b]
Wild-type *M. loti*	5	11 (2)	9–14	a
Wild-type *S. meliloti*	5	89 (3)	86–93	b, c
Chimera A	5	65 (6)	57–72	–
Chimera B	4	16 (3)	12–20	a
Chimera C	5	85 (8)	75–95	b
Chimera D	4	89 (5)	82–97	b, c
Chimera E	5	95 (2)	93–98	c
Chimera F	4	41 (5)	32–50	–

[a] The combined amounts of chitintetraose and -pentaose were set to 100%. SD, standard deviation
[b] Data were analyzed for significant differences using the Least Significant Difference (LSD) test. This analysis was performed using SPSS software. Groups of values whose members are not statistically different from each other are indicated by a, b or c

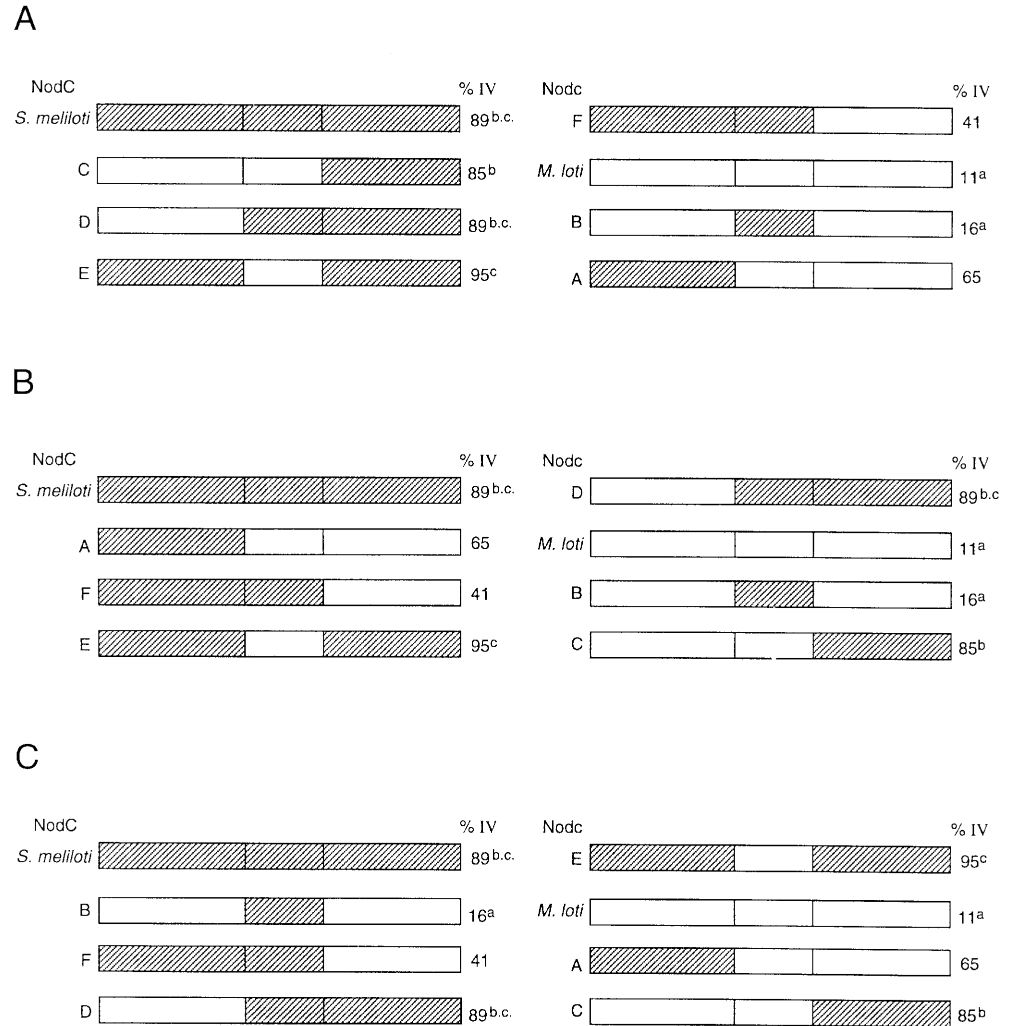

Fig. 3A–C Comparison of the relative amounts of chitintetraose produced by wild-type and comparable pairs of chimeric NodC proteins. Chitin oligosaccharide formation was quantified and statistically analyzed as described in Fig. 2 and Table 2. The relative amounts of chitintetraose (%IV) produced by the NodC proteins depicted were determined after setting the combined amounts of chitintetraose and -pentaose to 100%. The superscripts (a, b and c) indicate groups of values whose members do not differ statistically from each other (see Table 2). Comparisons of NodC chimeras in which either the C-terminal 164 amino acids (**A**), the N-terminal 170 amino acids (**B**), or the central 91 amino acids are derived from either *S. meliloti* NodC (*left panels*) or *M. loti* NodC (*right panels*) are shown

the relative amount of chitintetraose (41%) is substantially higher than that produced by the wild-type *M. loti* NodC (11%) (Fig. 3A). This indicates that, in addition to the C-terminal region, some other region in *S. meliloti* NodC also plays a role in the production of excess chitintetraose. Figure 3B shows the NodC proteins grouped with respect to the origin of the N-terminal region of 170 amino acids. The presence of this N-terminal region from *S. meliloti* NodC generally increases the relative amount of chitintetraose produced, but replacement of this region of the *M. loti* NodC by that of *S. meliloti* NodC (chimera A) does not lead to a complete reversion to an *S. meliloti* wild-type excess of chitintetraose, as was observed with the C-terminal region. This shows that, in addition to the C-terminus, the N-terminal region of *S. meliloti* also plays a role in determining the length of the main product, chitintetraose. However, the influence of the N-terminal region is minor or negligible when the C-terminal region of *S. meliloti* NodC is present, confirming our earlier conclusion that the major factor in restricting the length of the main product to a tetramer is the C-terminal region.

When the parental and chimeric NodC proteins are grouped according to the origin of the central 91-amino acid region (Fig. 3C), it appears that there is only one case in which exchange of this central region leads to a statistically significant change in the relative amounts of chitintetraose and -pentaose. Surprisingly, the presence of only the central region from the chitintetraose-producing *S. meliloti* NodC in chimera F results in a 24% decrease in the relative amount of chitintetraose, compared to chimera A. In addition we have observed that although the absolute values for the relative amounts of chitintetraose produced by the respective parental NodC proteins and chimeras B and E, in which only the central region was exchanged, are not significantly different statistically; there is a difference of approximately 5% in the relative amounts of chitintetraose produced by these proteins in every experiment that we performed. In the case of *M. loti* NodC and chimera B, the central region induces a slight increase in the relative amount of chitintetraose, whereas a slight decrease is observed in *S. meliloti* NodC with respect to chimera E. These variations do not correlate with a particular combination of regions in these chimeras. Taken together, these

data indicate that the central region has no significant function in determining the difference in the relative amounts of chitintetraose and -pentaose formed by *M. loti* and *S. meliloti* NodC.

In conclusion, our results show that the characteristic synthesis of chitintetraose as the major product by *S. meliloti* NodC is mainly dependent on the C-terminal region of maximally 164 amino acids, whereas the N-terminal segment of 170 amino acids also plays a role in restricting the length of the major product to a tetramer, possibly through a specific interaction with the C-terminal region.

Potential chain-length element

This C-terminal region of NodC should contain a structural element that is important for the oligosaccharides chain length. The predicted structures of the C-terminal regions of NodC proteins whose metabolic products have been determined were compared with each other (Fig. 4). The length of the predicted β-strand around position 300 is the only feature that appears to be related to the chain-length of the major product of NodC; the higher the relative amount of chitintetraose over chitinpentaose, the shorter the predicted β-strand. This element may therefore be responsible for the effect of the C-terminal 164 amino acids on chitin oligosaccharide chain length that we describe. Analysis of the complete NodC protein did not result in the recognition of any other structure that was correlated with the chitin oligo chain length (data not shown).

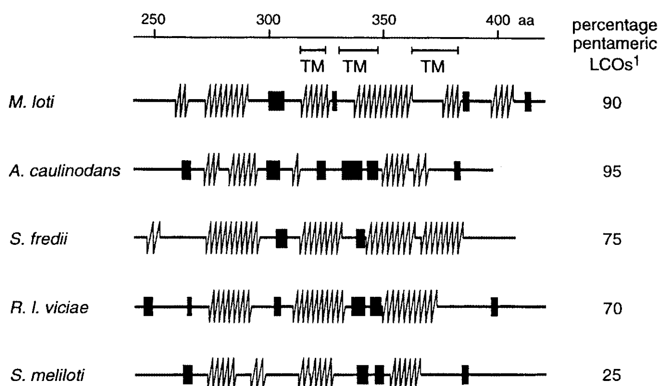

Fig. 4 Secondary structure predictions for the C-termini of NodC proteins. A consensus in secondary-structure prediction was obtained using the methods of Deleage and colleagues (Levin et al. 1986; Deleage et al. 1987; Gibrat et al. 1987; Kneller et al. 1990; Rost et al. 1994), using the software available on the servers http://www.ibcp.fr/serv_pred.html (Institute of Biology and Chemistry of Proteins CNRS-UPR, Lyon, France) and http://kiwi.imgen.bem.tmc.edu:8088/search-launcher/launcher.html (Baylor College of Medicine, Houston, Tex.). Predicted α-helices (*zig-zag pattern*) and β-strands (*black boxes*) are indicated. Predictions of transmembrane (TM) regions were obtained using the TopPred 2 program and the DAS transmembrane prediction server at http://www.biokemi.su.se/~server (Stockholm University, Sweden), the TMpred program (Swiss Institute for Experimental Cancer Research, Epalinges, Switzerland), and the PhdTopology program (EMBL, Heidelberg, Germany) via the BCM server mentioned above. Shown are the C-termini of NodC proteins from *M. loti* (only chitinpentaose-containing LCOs detected; Collins-Emerson et al. 1990; López-Lara et al. 1995), *Azorhizobium caulinodans* (95% of LCOs contain chitinpentaose; Goethals et al. 1989; Mergaert et al. 1993), *S. fredii* (75% of LCOs contain chitinpentaose; Krishnan et al. 1991; Bec-Ferté et al. 1994), *Rhizobium leguminosarum* bv. *viciae* (70% of LCOs contain chitinpentaose; Rossen et al. 1984; Spaink et al. 1991) and *S. meliloti* (5% of LCOs contain chitinpentaose Jacobs et al. 1985; Schultze et al. 1992)

Acknowledgements The authors thank H. Ceelie and J. Pilling for their help in constructing some of the chimeric *nodC* genes, and Dr. J. E. Thomas-Oates for critically reading the manuscript. This work was funded by the Netherlands Organization for Scientific Research (NWO-PIONIER grant to H.P.S.).

References

Barny MA, Schoonejans E, Economou A, Johnston AW, Downie JA (1996) 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. Mol Microbiol 19: 443–453

Bec-Ferté MP, Krishnan HB, Promé D, Savagnac A, Pueppke SG, Promé J-C (1994) Structures of nodulation factors from the nitrogen-fixing soybean symbiont *Rhizobium fredii* USDA257. Biochemistry 33: 11782–11788

Collins-Emerson JM, Terzaghi EA, Scott DB (1990) Nucleotide sequence of *Rhizobium loti nodC*. Nucleic Acids Res 18: 6690

Deleage G, Roux B (1987) An algorithm for protein secondary structure prediction based on class prediction. Protein Eng 1: 289–294

Geremia RA, Mergaert P, Geelen D, Vanmontagu M, Holsters M (1994) The NodC Protein of *Azorhizobium caulinodans* is an *N*-acetylglucosaminyltransferase. Proc Natl Acad Sci USA 91: 2669–2673

Gibrat JF, Garnier J, Robson B (1987) Further developments of protein secondary structure prediction using information theory. New parameters and consideration of residue pairs. J Mol Biol 198: 425–443

Goethals K, Gao M, Tomekpe K, Van Montagu M, 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

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

Kamst E, Spaink HP (1999) Functional domains in the chitin oligosaccharide synthase NodC and related β-polysaccharide synthases. Trends Glycosci Glycotechnol 11: 187–200

Kamst E, Pilling J, Raamsdonk LM, Lugtenberg BJJ, Spaink HP (1997) *Rhizobium* nodulation protein NodC is an important determinant of chitin oligosaccharide chain length in Nod factor biosynthesis. J Bacteriol 179: 2103–2108

Kamst E, Spaink HP, Kafetzopoulos D (1998) Biosynthesis and secretion of rhizobial lipochitin-oligosaccharide signal molecules. In: Biswas BB, Das HK (eds) Plant-microbe interactions. Plenum Press, New York, pp 29–72

Kneller DG, Cohen FE, Langridge R (1990) Improvements in protein secondary structure prediction by an enhanced neural network. J Mol Biol 214: 171–182

Krishnan HB, Pueppke SG (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

Lerouge P, Roche P, Faucher C, Maillet F, Truchet G, Promé J-C, Dénarié J (1990) Symbiotic host-specificity of *Rhizobium*

meliloti is determined by a sulphated and acylated glucosamine oligosaccharide signal. Nature 344: 781–784

Levin JM, Robson B, Garnier J (1986) An algorithm for secondary structure determination in proteins based on sequence similarity. FEBS Lett 205: 303–308

López-Lara IM, van den Berg JDJ, Thomas-Oates JE, Glushka J, Lugtenberg BJJ, Spaink HP (1995) Structural identification of the lipo-chitin oligosaccharide nodulation signals of *Rhizobium loti*. Mol Microbiol 15: 627–638

Mergaert P, Van Montagu M, Promé J-C, Holsters M (1993) Three unusual modifications, a D-arabinosyl, a N-methyl, and a carbamoyl group, are present on the Nod factors of *Azorhizobium caulinodans* strain ORS571. Proc Natl Acad Sci USA 90: 1551–1555

Mergaert P, D'Haeze W, Geelen D, Promé D, Van Montagu M, Geremia R, Promé J-C, Holsters M (1995) Biosynthesis of *Azorhizobium caulinodans* Nod factors: study of the activity of the NodABC proteins by expression of the genes in *Escherichia coli*. J Biol Chem 270: 29217–29223

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

Roche P, Maillet F, Plazanet C, Debellé F, Ferro M, Truchet G, Promé J-C, Dénarié J (1996) The common *nodABC* genes of *Rhizobium meliloti* are host-range determinants. Proc Natl Acad Sci USA 93: 15305–15310

Rossen L, Johnston AWB, Downie JA (1984) DNA sequence of the *Rhizobium leguminosarum* nodulation genes *nodAB* and *C* required for root hair induction. Nucleic Acids Res 12: 9497–9508

Rost B, Sander C (1994) Combining evolutionary information and neural networks to predict protein secondary structure. Proteins 19: 55–72

Saxena IM, Brown RM, Fevre M, Geremia RA, Henrissat B (1995) Multidomain architecture of beta-glycosyl transferases: implications for mechanism of action. J Bacteriol 177: 1419–1424

Schultze M, Quiclet-Sire B, Kondorosi E, Virelizier H, Glushka JN, Endre G, Géro SD, Kondorosi A (1992) *Rhizobium meliloti* produces a family of sulphated lipo-oligosaccharides exhibiting different degrees of plant host specificity. Proc Natl Acad Sci USA 89: 192–196

Spaink HP (1996) Regulation of plant morphogenesis by lipo-chitin oligosaccharides. Crit Rev Plant Sci 15: 559–582

Spaink HP, Sheeley DM, van Brussel AAN, Glushka J, York WS, Tak T, Geiger O, Kennedy EP, Reinhold VN, Lugtenberg BJJ (1991) A novel highly unsaturated fatty acid moiety of lipo- oligosaccharide signals determines host specificity of *Rhizobium*. Nature 354: 125–130

Spaink HP, Wijfjes AHM, van der Drift KMGM, Haverkamp J, Thomas-Oates JE, Lugtenberg BJJ (1994) Structural identification of metabolites produced by the NodB and NodC proteins of *Rhizobium leguminosarum*. Mol Microbiol 13: 821–831

Studier FW, Rosenberg AH, Dunn JJ, Dubendorf JW (1990) Use of T7 RNA polymerase to direct expression of cloned genes. Methods Enzymol 185: 60–89

Vieira J, Messing J (1991) New pUC-derived cloning vectors with different selectable markers and DNA replication origins. Gene 100: 189–194