

Sequence of the *Rhizobium leguminosarum* biovar *phaseoli* *syrM* gene

Jan Michiels, Petra De Wilde and Jos Vanderleyden*

F.A.Janssens Laboratory of Genetics, Catholic University of Leuven, Willem de Croylaan 42, B-3001 Heverlee, Belgium

Received June 24, 1993; Accepted June 26, 1993

EMBL accession no. Z23013

Rhizobium forms nitrogen-fixing nodules in symbiosis with its leguminous host plant. In *R. meliloti*, the symbiont of alfalfa (*Medicago sativa*), the symbiotic regulatory gene *syrM* is involved in the coordinated expression of genes controlling nodulation (*nod*) and exopolysaccharide (*exo*) synthesis (1). *SyrM* activates the expression of the regulatory *nod* gene *nodD3* (1). In turn, *NodD3* stimulates transcription of *syrM* (2). *NodD* and *SyrM* proteins belong to the LysR family of transcriptional regulatory proteins (3). When carried on a multicopy plasmid, *nodD3* and *syrM* allow high levels of *nodABC* gene expression in the absence of plant inducer (1). In addition *syrM* stimulates the expression of *exoA*, *exoB* and *exoF* (1). *syrM*-dependent *exo* gene expression is mediated by *syrA* (1). So far, a *syrM* gene has been identified only in *R. meliloti* (2, 4).

Rhizobium leguminosarum biovar *phaseoli*, the symbiont of common beans (*Phaseolus vulgaris*), possesses three *nodD* genes (5). While transcription of *nodD2* and *nodD3* is constitutive, *nodD1* is inducible in the presence of bean exudates (6). Here we report the sequence of the *R. l. bv. phaseoli* *syrM* gene.

Using a *R. l. bv. viciae* *nifAB* probe (7), we have previously identified a *nifB*-like gene in *R. l. bv. phaseoli* CNPAF512 (unpublished results). DNA sequence analysis of a 1.8 kb *SalI*–*HindIII* fragment, located downstream of this *nifB*-like gene, indicated the presence of two open reading frames (ORF) coding for polypeptides of 171 and 336 amino acids respectively. The polypeptide encoded by the largest ORF, designated as *R. l. bv. phaseoli* *SyrM*, has a calculated molecular mass of 37,800 Da and displays homology to *R. meliloti* *SyrM* and to various *NodD* proteins (Figure 1). *R. l. bv. phaseoli* and *R. meliloti* *SyrM* proteins display 55% identity and 68% similarity (Figure 1). Amino acid similarity between *R. l. bv. phaseoli* *SyrM* and different *NodD* proteins is between 33% (*Azorhizobium caulinodans* *NodD*) and 41% (*R. l. bv. phaseoli* *NodD3*). The conserved amino acids between *NodD* and *SyrM* proteins are found predominantly in their amino terminus (underlined in Figure 1). This region contains a helix-turn-helix motif, characteristic of DNA-binding proteins (3).

The ORF encoding the 171 amino acids protein (19,187 Da) is located upstream from *syrM* and reads in the same direction. This ORF displays 81% similarity to the product of a partially sequenced unidentified ORF downstream of the *Bradyrhizobium japonicum* *nodIJ* genes (8). The function of this protein is unknown. No significant homology to any protein sequence in the Swiss-Prot data base was detected.

ACKNOWLEDGEMENTS

J.M. is a recipient of a fellowship from the Instituut voor Wetenschappelijk Onderzoek in de Nijverheid en de Landbouw. Part of this work was supported by an EEC-STDII program [(TS2-0199-C(GDF))].

REFERENCES

- Mulligan, J.T. and Long, S.R. (1989) *Genetics* **122**, 7–18.
- Kondorosi, E., Buiré, M., Cren, M., Iyer, N., Hoffmann, B. and Kondorosi, A. (1991) *Mol. Microbiol.* **5**, 3035–3048.
- Henikoff, S., Haughn, G.W., Calvo, J.M. and Wallace, J.C. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 6602–6606.
- Barnett, M.J. and Long, S.R. (1990) *J. Bacteriol.* **172**, 3695–3700.
- Davis, E.O. and Johnston, A.W.B. (1990) *Mol. Microbiol.* **4**, 924–932.
- Davis, E.O. and Johnston, A.W.B. (1990) *Mol. Microbiol.* **4**, 933–941.
- Schetgens, T.M.P., Hontelez, J.G.J., Van den Bos, R.C. and Van Kammen, A. (1985) *Mol. Gen. Genet.* **200**, 368–374.
- Göttfert, M., Hitz, S. and Hennecke, H. (1990) *Mol. Plant-Microbe Interact.* **3**, 308–316.

Rp	MNFMKRNVKEVTDQLAYRRKMLQDWGRERLTGHQI---NLASIDLNLLV	47
Rm	MDQPTWKRPHRAKFAVGSDAQQRQMPNLASIDLNLLV	38
	* * * * *	
Rp	ALEALLEYRNVTAGQHIGRSQPAMSRALGRLRGLFNDDLLVRSSTGLIP	97
Rm	DLEALLQYRHRITQAAQHVGRSQPAMSRALRGLRGLNDDLLVAGSRGLVL	88
	* * * * *	
Rp	TPOGEHLAQRLPALRTIREMVTSSRSVSKEMGRGATLAIPOHQALAVLP	147
Rm	TPLAECLTQMLPSVLDLAIQRM-NMLSLAPAQRRKVTMAMPDQAVVLLP	137
	* * * * *	
Rp	RLLPWLRERAPHLDTLACLPFRVAVRGLQGGIDDLAVGHIDVQLPGYFRR	197
Rm	HLLPRLHERAPHLDIVTDPDLGGALGLLEQGEIDVVVQMGAAAPLGYLRR	187
	* * * * *	
Rp	SLYTDRFACLLRHGHAPALAQEWITDNFATLRHAAISTDSDPHFGPIYDHL	247
Rm	RLYADSFVCLRHHPALAQEWITFAALRHVAIASEPDELFCQIYDRL	237
	* * * * *	
Rp	PNLRADRS-PILFSSVLTAAVVASATDVLVLPVRRVATQVSAMPLRVVD	296
Rm	TKLGLQRGDPHVVTVLTAAVLIATDSSLVLPVSRVATRVAAMLSIAVIP	287
	* * * * *	
Rp	PPLEPAPYKVMILNHERCHRDPEHRMLRGEITAAAST-AG	336
Rm	PPVELRPFYVALIWHERCHRDPEHRMLRGEITAAAST-AG	326
	* * * * *	

Figure 1. Amino acid sequence alignment of the *SyrM* proteins of *R. l. bv. phaseoli* (Rp) and *R. meliloti* (Rm, ref. 3). Identical amino acid residues are marked by asterisks and positions with conservative substitutions (S-T-A; L-V-I-M; K-R; D-E; Q-N; F-Y-W) are indicated with dots. Identical amino acids conserved in the *NodD* proteins of *R. meliloti* (*NodD1*, *NodD2*, *NodD3*), *R. l. bv. phaseoli* (*NodD1*, *NodD2*, *NodD3*), *R. l. bv. viciae*, *R. l. bv. trifolii*, *B. japonicum* (*NodD1*, *NodD2*), *A. caulinodans* and *B. japonicum* sp. ANU289 and in both *SyrM* proteins are underlined.

* To whom correspondence should be addressed