UNIVERSIDAD DE SALAMANCA

DEPARTAMENTO DE MICROBIOLOGÍA Y GENÉTICA

CENTRO HISPANOLUSO DE INVESTIGACIONES AGRARIAS

MOLECULAR AND FUNCTIONAL ANALYSIS OF CELULLASE CelC2 IN RHIZOBIA-LEGUMES SYMBIOSIS

MARTA ROBLEDO GARRIDO

DOCTORAL THESIS

2010

CONTENTS

THESIS ABSTRACT

The rhizobia–legume endosymbiosis provides the most efficient source of biologically fixed ammonia fertilizer for agricultural crops. The infection process is tightly regulated in legumes to ensure appropriate microsymbiont penetration into plant cells leading to the establishment of a nitrogen-fixing, intracellular state within the host. It has previously been shown that bacterial Nod Factor and exopolisacarides are involved in plant-bacterial recognition and signal transduction. However, how the process of plant cell wall erosion is balanced to avoid lysis and destruction of the host cell wall is still unknown. In this work we show that cellulase CelC2 from *Rhizobium leguminosarum* bv. trifolii, specifically erodes the noncrystalline apex of the white clover host root hair wall thereby fulfilling an essential role in the canonical primary and secondary infection processes required for development of this nitrogen-fixing symbiosis. Consistent with its involvement in fundamental processes for survival in the environment, the CelC2 encoding gene, *celC*, is widely conserved among species of the genus *Rhizobium*. CelC2 overexpression increases bacterial competitiveness and leads to aberrant infection thread and nodule formation on clover roots as a result of an extensive uncontrolled hydrolysis of the non-crystalline cellulose which distribution is restricted to primary infection sites and the points of release of invading bacteria from the infection threads into nodule cells (secondary infection sites). On the other hand, a *Sinorhizobium meliloti* 1021 derivative over-producing the CelC2 enzyme is delayed in nodulation and nitrogen fixation and displays a similar phenotype in *Medicago sativa* and *M. truncatula* inoculated plants, suggesting that endoglucanase mediated infection is not restricted to clover plants. In addition, we show that application of purified cellulase CelC2 can alter the canonical Nod factor–induced calcium spiking and nodulin gene expression patterns in *M. truncatula*. All this data support a role of the CelC2 cellulase not only in the primary but also in the secondary symbiotic infection of white clover roots and provide experimental evidence for the predicted requirement of a tightly regulation of the production and activity of rhizobial cell wall degrading enzymes during a successful effective nodulation. Furthermore, we show that the CelC2 cellulase is involved in cellulose biosynthesis by modulating the length of the cellulose fibrils, acting in biofilm formation *in planta* and *in vitro*. The data obtained in this experimental work indicate that rhizobia cellulase CelC2 plays a significant role in bacterial colonization and infection of legumes during nitrogen-fixing symbiosis.

THESIS SUMMARY

Mutualistic symbioses between diazotrophic bacteria collectively called rhizobia and legume plants are of critical agronomic and environmental importance, allowing crop production in nitrogen-limited soils without the need for chemical-N fertilizers. Rhizobia grow as free-living organisms but can also induce and colonize root nodules in legume plants thereby establishing a partnership that benefit both organisms. These bacteria can attach and invade their plant hosts through colonization of intercellular epidermal spaces, crack entry at emerging lateral roots or penetration into the root hairs *via* tubular structures called infection threads (1). The later is the best-characterized infection pathway, elicited by a complex molecular dialogue between both symbiotic partners. During early symbiotic stages, flavonoids exuded by legume roots induce the synthesis and secretion of lipochito-oligosaccharide signal molecules or nodulation factors (NFs) in their cognate rhizobia upon the transcriptional activation of the nodulation (*nod*) genes (2). In turn, NFs elicit several responses in the plant leading to nodule organogenesis: calcium spiking in susceptible root hair cells, root hair deformation to facilitate bacteria penetration, dedifferentiation of root cortical cells generating the nodule primordium and transcripcional activation of several early nodulin genes (*Enod*) (3). Concomitantly, the plant cell wall redirects its inward growth, resulting in the formation of the infection thread, which continues to the base of the epidermal root hair towards the nodule primordium where invading bacteria proliferate. In developing nodules, rhizobia are released by endocytosis from intracellular ramifications of the infection threads into membrane-enclosed vesicles within the cytoplasm of some cells of the nodule (infection zone). Released bacteria undergo a process of morphological differentiation to nitrogen-fixing competent bacteroids, which in the mature nodules remain confined to the vesicles (i.e. the symbiosome) surrounded by a plasma membrane of host origin.

Scientists have been working for a number of years to understand the early steps of the symbiosis between legumes and rhizobia to improve the efficiency of the process with the hope that in the future this trait could be transferred to crop plants, the majority of which cannot fix dinitrogen themselves. This work focus on the characterization of the mechanisms governing the *Rhizobium*-legume infection process, an more specifically on the molecular and functional analysis of hydrolytic enzymes produced by *Rhizobium leguminosarum* bv. trifolii in symbiosis with white clover. Microscopy analysis showed that the establishment of rhizobia as nitrogen-fixing endosymbiont within legume root nodules requires the disruption of the host plant cell wall at strategic infection sites in the root hair tip (primary infection) and at points of bacterial release from infection threads within the root cortex (secondary infection). In these two steps of the infection process plant cell wall-degrading enzymes are predicted to participate.

Several previous studies indicated that rhizobia produce enzymes capable of degrading cellulose (4-9), but little is known about their molecular properties, and their specific role in symbiosis is undefined. The relatively low activities of these rhizobial enzymes have hampered research progress in this area. Using improved assays with increased sensitivity reliably detecting these enzyme activities, our research group established that cellulases were produced by several rhizobia wild type strains (6-7). Further studies using *R. leguminosarum* bv. trifolii ANU843 indicated that this model wild-type strain produces constitutively at least two cell-bound cellulase isozymes, C1 and C2 (7). Then, a combination of phase contrast, polarized light microscopy and enzymology indicated that C2 enzyme can completely erode through the white clover root hair wall at a highly localized site on the isotropic, noncrystalline apex of the root hair tip (Hot [Hole on the tip] phenotype) (10). CelC2 enzyme was purified and characterized as a 1,4-β-D-endoglucanase with high substrate specificity for the isotropic non-crystalline cellulose of clover cell walls. It contained a sequence of amino acids which has identity to a sequence encoded by the *celC* gene of *R. leguminosarum* bv. trifolii R200 (11). The *celC*-encoded

protein has sequence homology to several endoglucanases belonging to the glycosyl hydrolase family 8 (12) such as CelC from *Agrobacterium tumefaciens* (13) and CMCax from *Gluconacetobacter xylinum* (14). In all these cellulose-producing bacteria the *celC* genes are located near putative cellulose synthase genes in a region of the chromosome (*celABC*) involved in bacterial cellulose biosynthesis (13).

In this work, we first examined the majority of the currently described official type strains of each rhizobia taxa for cellulase activity, and all were found to be positive to varying degree. Besides this, evaluation of the currently defined genomes of rhizobia and other bacterial plant-symbionts revealed the ubiquitous presence of genes coding for diverse putative cellulases from the glycosyl hydrolase family. With the aim to explore the phylogenetic relationships of *celC* gene, we have sequenced and compared it within several *Rhizobium* type species and biovars isolated from nodules of different hosts and compared its phylogeny with that of classical chromosomal core genes (*rrs, recA* and *atpD*) and other plasmidic symbiotic genes (*nodC*). The results obtained showed that *celC* gene phylogeny correlates to that based on housekeeping genes such as *recA* and *atpD* showing that this gene is widely conserved among species of the genus *Rhizobium,* consistent with its involvement in fundamental processes for survival in the environment. As occurs with housekeeping genes, *celC*-based phylogeny enables the differentiation of species with closely related *rrs* genes thus rendering this *locus* as a novel taxonomic marker for the genus *Rhizobium.* All these results demonstrate that the cellulase genes and the cellulolytic enzymes they encode are widespread and commonly expressed among bacteria establishing plant symbioses (15, Chapter I).

On the other hand, we have studied the symbiotic function of this cell-bound bacterial cellulase (CelC2) enzyme from *Rhizobium leguminosarum* bv. trifolii, the clover-nodulating endosymbiont. Previous CelC2 bioassays on clover roots revealed that the purified enzyme is able to degrade specifically the non-crystalline cellulose located at the apex of clover root-hairs, thereby forming a hole of a precise size that is predicted to be the primary portal of entry of bacteria into clover roots (10). Consistent with this latter assumption, nodulation tests and microscopy analyses revealed that a CelC2 knock-out mutant was unable either to breach the root hair wall at the primary infection site or to form infection threads when inoculated onto white clover seedlings, thus eliciting in its host plant non-fixing nodule structures devoid of bacteria (16, Chapter II). Complementation of the mutant strain with a plasmid expressing the *celC* gene restored the ability to produce the CelC2 cellulase, infect root hairs, and induce efficient nitrogen-fixing nodules on white clover like the parental strain. Thus, CelC2 cellulase fulfils an essential role in the primary infection process required for development of the canonical nitrogen-fixing *R. leguminosarum* bv. trifolii-white clover symbiosis.

Regarding secondary infection, detailed ultrastructural examinations indicate that at the points of bacterial release into the host cells the infection threads are unwalled. However, it remains to be elucidated whether these secondary infection sites are originated either by arrest of the cell wall synthesis during intracellular elongation of the infection threads or by local degradation catalyzed by cell-wall hydrolytic enzymes of bacterial and/or plant origin. In this work, we have used a gain-of-function approach to further explore the impact of CelC2 activity throughout the infection process of white clover roots by its cognate rhizobia. We showed that a *R. leguminosarum* bv. trifolii ANU843 mutant derivative constitutively over-producing the CelC2 enzyme increased the competitiveness of its parental strain, although it elicits aberrant nodule-like structures on clover roots which are inefficient for nitrogen fixation as revealed by shoot length, dry weigh and nitrogen content determinations in inoculated clover plants. Microscopy analyses revealed that this symbiotic phenotype is the outcome of an extensive uncontrolled degradation of the isotropic non-crystalline cellulose in the host cell walls, which distribution is shown to be restricted to the expected primary and secondary infection sites (17, Chapter III). This data support a role of the CelC2 cellulase not only in the primary but also in the secondary symbiotic infection of white clover roots and provide experimental evidence for a predicted requirement of tight regulation in the production and activity of rhizobial cell wall degrading enzymes to establish a compatible effective root nodule symbiosis.

In order to analyze the role of cellulase CelC2 in symbiotic signalling, we have also investigated the nodulation-related responses of the model legume plant *Medicago truncatula* upon inoculation with a *Sinorhizobium meliloti* 1021 derivative overexpressing the *celC* gene from a multicopy plasmid vector. The results show that the aberrant symbiotic phenotypes in root hair deformation, nodule kinetics and development observed in white clover also occur in *Medicago* plants. This uncontrolled disruption of the cell wall polymers are accompanied by a significant alteration in canonical Nod factor-induced responses such as calcium spiking traces and ENOD11 expression after root treatment with purified cellulase CelC2 (19, Chapter IV). Therefore, plant is capable to sense an excess of bacterial cellulase altering the pathways associated with bacterial recognition and nodule formation.

Finally, we examined the role of CelC2 enzyme in cellulose biosynthesis and biofilm formation, both natural processes of biotechnological relevance. Among the many factors involved in development of an effective symbiosis between rhizobia and their host plants, those associated with colonization and adherence of bacteria to the roots and to the root hair tip surface, a key stage for the subsequent entry into the plant, have not been yet characterized widely. It is know that *Rhizobium* synthesize cellulose fibrils, encoded by *celABC* operon, that mediated aggregation to root hairs and anchor to the root surface (11, 20-22). Probably the synthesis of cellulose is one of the most important biochemical processes that is worst known at the molecular level due to its complexity and high regulation. In both bacteria and plants, two proteins directly involved in the biosynthesis of cellulose have been characterized: CesA and Korrigan, in plants, and CelA and CelC in bacteria (23-24). CelA and CesA have homology with glycosyl transferases while CelC and Korrigan have homologies with endoglucanases. The postulated involvement of the *Agrobacterium* CelC cellulase enzyme in bacterial cellulose biosynthesis is to transfer new oligomer primers for chain elongation (25), but this biochemical function has not been definitively established. During the course of our research on the CelC2 protein, we observed all rhizobial strains tested produced cellulose and that cellulose microfibrils produced by the *celC*$^{+}$ overexpression mutant were undetectable by Calcofluor staining. On the other hand, the microfibrils formed the by Δ*celC2* deletion mutant were significantly longer than those seen in wild type. All these data suggest that during bacterial cellulose synthesis, CelC-like enzymes act as endoglucanases that cleaved the cellulose precursor oligomer which is being elongated, following the model proposed in plants for cellulose biosynthesis (29, Chapter V).

Attachment to a surface is the initial step in biofilm formation, followed by establishment of microcolonies by clonal propagation and maturation, and formation of three-dimensional structures that are covered by exopolymers (26-28). The role of cellulose in *Rhizobium* biofilm formation has not been hitherto fully characterized. We gfp-tagged our *celC*$^{-}$ and *celC*$^{+}$ mutant strains to analyze bacterial attachment to clover roots and to several abiotic surfaces such as plastic microtiter plates, tabs and sand. We found that overexpression of the *celC* gene reduced the ability of *R. leguminosarum* bv. trifolii to develop a mature biofilm. Also *celC*$^{-}$ mutant showed altered biofilm structure and formation kinetics (29, Chapter V). Therefore, we conclude that *celC* mutant strains produced altered phenotypes in both bacterial cellulose biosynthesis and biofilm formation.

Together, our results predict a relevant role of this *Rhizobium* cell-bound enzyme in differente stages of the root-nodule symbiosis with legumes, and identifies these enzymes as biotechnological targets to optimize the balance between endophytic and endosymbiotic states in rhizobial biofertilizers.

References

1. Goormachtig, S., W. Capoen, E.K. James, and M. Holsters, Switch from intracellular to intercellular invasion during water stress-tolerant legume nodulation. Proc Natl Acad Sci U S A, 2004. 101(16): p. 6303-8.

2. Gibson, K.E., H. Kobayashi, and G.C. Walker, Molecular determinants of a symbiotic chronic infection. Annu Rev Genet, 2008. 42: p. 413-41.

3. Oldroyd, G.E. and J.A. Downie, Coordinating nodule morphogenesis with rhizobial infection in legumes. Annu Rev Plant Biol, 2008. 59: p. 519-46.

4. Chalifour, F.P. and N. Benhamou, Indirect evidence for cellulase production by *Rhizobium* in pea root nodules during bacteroid differentiation: cytochemical aspects of cellulose breakdown in rhizobial droplets. Can J Microbiol, 1989. 35: p. 821-829.

5. Iannetta, P., G. McMillan, and J.I. Sprent, Plant cell wall-degrading enzymes of *Rhizobium leguminosarum* bv. viciae: their role in avoiding the host-plant defense response. Soil Biol Biochem, 1997. 29: p. 1019-1021.

6. Jiménez-Zurdo, J.I., P.F. Mateos, F.B. Dazzo, and E. Martínez-Molina, Cell-bound cellulase and polygalacturonase production by *Rhizobium* and *Bradyrhizobium* species. Soil Biol Biochem, 1996. 28: p. 917-921.

7. Mateos, P.F., et al., Cell-associated pectinolytic and cellulolytic enzymes in *Rhizobium leguminosarum* biovar trifolii. Appl Environ Microbiol, 1992. 58(6): p. 1816-22.

8. Michaud, P., A. Belaich, B. Courtois, and J. Courtois, Cloning, sequencing and overexpression of a *Sinorhizobium meliloti* M5N1CS carboxymethyl-cellulase gene. Appl Microbiol Biotechnol, 2002. 58(6): p. 767-71.

9. Morales, V., E. Martinez-Molina, and D. Hubbell, Cellulase production by *Rhizobium*. Plant Soil, 1984. 80: p. 407-415.

10. Mateos, P.F., et al., Erosion of root epidermal cell walls by *Rhizobium* polysaccharide-degrading enzymes as related to primary host infection in the *Rhizobium*-legume symbiosis. Can J Microbiol 2001. 47: p. 475-487.

11. Ausmees, N., H. Jonsson, S. Hoglund, H. Ljunggren, and M. Lindberg, Structural and putative regulatory genes involved in cellulose synthesis in *Rhizobium leguminosarum* bv. trifolii. Microbiology, 1999. 145 (Pt 5): p. 1253-62.

12. Coutinho, P.M. and B. Henrissat, The modular structure of cellulases and other carbohydrate-active enzymes: an integrated database approach., in Genetics, Biochemistry and Ecology of Cellulose Degradation, K. Ohmiya, et al., Editors. 1999, Uni Publishers Co: Tokyo. p. 15-23.

13. Matthysse, A.G., S. White, and R. Lightfoot, Genes required for cellulose synthesis in *Agrobacterium tumefaciens*. J Bacteriol, 1995. 177(4): p. 1069-75.

14. Wong, H.C., et al., Genetic organization of the cellulose synthase operon in *Acetobacter xylinum*. Proc Natl Acad Sci USA, 1990. 87(20): p. 8130-4.

15. Robledo, M., *et al.,* The *celC* gene, a new phylogenetic marker useful for taxonomic studies in *Rhizobium*. (Chapter I).

16. Robledo, M., *et al*., *Rhizobium* cellulase CelC2 is essential for primary symbiotic infection of legume host roots. Proc Natl Acad Sci U S A, 2008. 105(19): p. 7064-9. (Chapter II).

17. Jones, K.M., *et al*., How rhizobial symbionts invade plants: the *Sinorhizobium-Medicago* model. Nat Rev Microbiol, 2007. 5(8): p. 619-33.

18. Robledo, M., *et al*., Development of functional symbiotic white clover root nodules requires tightly regulated production of rhizobial cellulase CelC2. (Chapter III).

19. Robledo, M., *et al*., Effect of *Rhizobium* cellulase CelC on Nod factor–induced calcium spiking and nodulin gene expression in *Medicago truncatula*. (Chapter IV).

20. Smit, G., S. Swart, B.J. Lugtenberg, and J.W. Kijne, Molecular mechanisms of attachment of *Rhizobium* bacteria to plant roots. Mol Microbiol, 1992. 6(20): p. 2897-903.

21. Dazzo, F.B., G.L. Truchet, J.E. Sherwood, E.M. Hrabak, M. Abe, and S.H. Pankratz, Specific phases of root hair attachment in the *Rhizobium* trifolii-clover symbiosis. Appl Environ Microbiol, 1984. 48(6): p. 1140-50.

22. Mateos, P.F., et al., Direct *in situ* identification of cellulose microfibrils associated with Rhizobium leguminosarum biovar. trifolii attached to the root epidermis of White clover. Can J Microbiol, 1995. 41: p. 202-207.

23. Doblin, M.S., I. Kurek, D. Jacob-Wilk, and D.P. Delmer, Cellulose biosynthesis in plants: from genes to rosettes. Plant Cell Physiol, 2002. 43(12): p. 1407-20.

24. Ross, P., R. Mayer, and M. Benziman, Cellulose biosynthesis and function in bacteria. Microbiol Rev, 1991. 55(1): p. 35-58.

25. Matthysse, A.G., D.L. Thomas, and A.R. White, Mechanism of cellulose synthesis in *Agrobacterium tumefaciens*. J Bacteriol, 1995. 177(4): p. 1076-81.

26. Davey, M.E. and G.A. O'Toole, Microbial biofilms: from ecology to molecular genetics. Microbiol Mol Biol Rev, 2000. 64(4): p. 847-67.

27. Webb, J.S., M. Givskov, and S. Kjelleberg, Bacterial biofilms: prokaryotic adventures in multicellularity. Curr Opin Microbiol, 2003. 6(6): p. 578-85.

28. Fujishige, N.A., N.N. Kapadia, P.L. De Hoff, and A.M. Hirsch, Investigations of *Rhizobium* biofilm formation. FEMS Microbiol Ecol, 2006. 56(2): p. 195-206.

29. Robledo, M., *et al.*, Role of *Rhizobium* endoglucanase CelC in cellulose biosynthesis and attachment to plant roots and abiotic surfaces. (Chapter V).

CHAPTER I. The *celC* gene, a new phylogenetic marker useful for taxonomic studies in *Rhizobium*

Marta Robledo, Encarna Velázquez*, Martha Helena Ramírez-Bahena, Paula García-Fraile, Ana Pérez-Alonso, Raúl Rivas, Eustoquio Martínez-Molina and Pedro F. Mateos

Departamento de Microbiología y Genética. Universidad de Salamanca. Spain.

The *celC* gene codifies a cellulase that fulfils a very significant role in the infection process of clover by *Rhizobium leguminosarum*. This gene is located in the *celABC* operon present in the chromosome of strains from *R. leguminosarum*, *R. etli* and *R. radiobacter* whose genomes have been completely sequenced. Nevertheless the existence of this gene in other species of genus *Rhizobium* had not been investigated up to date. In this work we analyse for the first time the *celC* gene in several species of this genus isolated from legume nodules and plant tumours in order to compare the *celC* phylogeny to those of other chromosomal and plasmidic genes. The results obtained showed that phylogenies of *celC* and chromosomal genes such as *rrs, recA* and *atpD* are completely congruent, whereas no relation was found with symbiotic or virulence genes. Therefore, the suitability and usefulness of the *celC* gene to differentiate species of genus *Rhizobium* specially those with close related *rrs* genes is pointed out and the taxonomic status of several strains of genus *Rhizobium* whose genomes have been completely sequenced is discussed.

The cellulase genes and the cellulolytic enzymes they encode are widespread and commonly expressed among bacteria establishing root nodule symbiosis [24]. Among them, the *celC* gene, located near putative cellulose synthase genes in a region of the chromosome (*celABC*), is involved in bacterial cellulose biosynthesis [20] and codifies a cellulase essential for symbiotic infection of legume host roots in the genus *Rhizobium* fulfilling a very significant role in infection processes required for development of the nitrogen fixation endosymbiosis [24]. This gene was initially sequenced in some clover-nodulating *Rhizobium* strains and there were some problems with its name since it has been named *celC* when the complete *celABC* operon was sequenced in *R. leguminosarum* bv trifolii R201 [5], *celC8* because it codifies a protein from glycosidases family 8 in *E. coli* [4], *celY* in *Rhizobium leguminosarum* bv viciae 3841 because it codifies an endo-1,4-beta-glucanase Y in *Erwinia chrysanthemi* [10] and finally *celC2* because it codifies for a protein named cellulase C2 [24]. Nevertheless, currently *celC* is the common name used in most of complete genomes of rhizobia sequenced up to date such as *Rhizobium radiobacter* (formerly *Agrobacterium tumefaciens*) C58 whose genome was the first completely sequenced for this group of bacteria (accession number of linear chromosome AE007870). Later, the *celC* gene was localized in several strains from genus *Rhizobium* whose complete sequences have been recently obtained such *R. leguminosarum* bv trifolii WSM1325 and WSM2340 (accession numbers for chromosomes, NC_012850 and NC_011369, respectively), *R. leguminosarum* bv viciae 3841 (accession number for chromosome, NC_008380), *R. etli* CFN42^T and CIAT 652 (accession numbers for chromosomes, NC_007761 and NC_010994, respectively), *R. radiobacter* (currently *R. rhizogenes)* K84 (accession number of chromosome 2, CP000629) and *Rhizobium* sp. NGR234 (accession number of plasmid pNGR234b, CP000874). Moreover, in the annotations of the complete genome of *R. leguminosarum* bv viciae strain 3841 the initial name *celY* has been changed to *celC* (accession number for chromosome, NC_008380). In all these strains the *celC* gene has a chromosomal location, except in the strain *Rhizobium* sp. NGR234 in which the gene is located in a plasmid (accession number of plasmid pNGR234b, CP000874). This gene has not been found in any of the three strains of genus *Bradyrhizobium* whose genomes has been sequenced up to date USDA 110, BTAi1, and ORS 278 (accession numbers for chromosomes, NC_004463, NC_009485, NC_009445, respectively), in *Azorhizobium caulinodans* ORS 571^T (accession number for chromosome, NC_009937), in *Mesorhizobium loti* MAF303099 (accession number for chromosome, NC_002678) or in *Rhizobium* (formerly *Agrobacterium*) *vitis* S4 (accession numbers for chromosomes 1 and 2 CP000633 and CP000634, respectively). In the case of genus *Ensifer* (formerly *Sinorhizobium*), the *celC* gene was not found in the strain *E. meliloti* 1021 (accession number of chromosome, pSymA and pSymB NC_003047.1, NC_003037 and NC_003078, respectively), but in the strain *E. medicae* WSM419 a gene homologous to *celC* was located in a plasmid (accession number for plasmid pSMED02, CP000740).

Considering the importance of *celC* gene in *Rhizobium*-legume symbiosis the aim of this work was to explore the phylogenetic relationships of this gene among species of genus *Rhizobium*. For that, we analysed the sequences of *celC* gene in several strains from different species and biovars of this genus and the results were compared to those obtained on the basis of chromosomal genes (*rrs, recA* and *atpD*) and plasmidic symbiotic or virulence genes (*nodC, virA*). The results of these analyses showed that the phylogeny of *celC* gene is completely congruent with those based on the chromosomal genes, but not with those based on plasmidic genes, in species of genus *Rhizobium,* and thus *celC* represents a new phylogenetic marker useful for taxonomic studies in this genus.

Results and discussion

Comparison of *celC* and chromosomal gene phylogenies.The *celC* gene sequences are currently known in a few number of strains from genus *Rhizobium* nodulating legumes from some of which the sequences of their complete genomes are available. Based on conserved zones of the *celC* gene of these strains we designed several primers pairs to perform a screening of this gene in genus *Rhizobium* (data not shown). The best results were obtained with the primers pair used in this study that yielded a band of about 800 nucleotides. The sequences obtained were identified as those corresponding to the *celC* gene after a comparison with those held in Genbank.

The sequences of *celC* gene obtained in this study were analysed together with those of strains deposited in public databases and the results are shown in fig. 1. The results of this analysis were compared with phylogenies based on the *rrs* gene, because it is the basis of the current classification of rhizobia [17], and with those based on *recA* and *atpD* genes, since they were the first housekeeping genes sequenced in these bacteria [8] and are currently sequenced in almost all species of *Rhizobium*.

The phylogenetic tree corresponding to *celC* gene analysis showed a general arrangement of species completely congruent with that of *rrs* gene (Fig. 2) and *recA*

and *atpD* genes (Figs. S1 and S2, supplementary material) with several subgroups within genus *Rhizobium*. A wide cluster is formed in all cases by the species *R. leguminosarum, R. indigoferae, R. pisi, R. fabae*, *R. etli* and *R. phaseoli*, a second cluster is formed by *R. gallicum* and *R. mongolense*, a third cluster is formed by *R. hainanense, R. tropici* and *R. rhizogenes* and several species formed separated branches in all trees, such as *R. tibeticum*, *R. alamii* and *R. giardinii*. In spite of the coincidence in the arrangement of species, the phylogenetic distances are higher for *celC* gene than for *rrs* gene and even higher than those found for *recA* and *atpD* genes. This fact allows us to make several taxonomic considerations very useful for further taxonomic studies.

The cluster of *R. leguminosarum* contains species, *R. phaseoli, R. etli,* and *R. pisi*, with *rrs* genes highly homologous and whose differentiation should be performed by using protein-coding genes [22]. On the basis of *celC* gene analysis these species formed a cluster together with *R. leguminosarum* supported by a bootstrap of 99% (Fig. 1). The identity values found among species within this group are lower than those found in *recA* and *atpD* genes and therefore *celC* gene allows a better differentiation among close species of *Rhizobium*. Surprisingly, two species from this cluster, *R. indigoferae* and *R. fabae* presented high identity values in their *celC* genes with respect to *R. leguminosarum* and *R. pisi,* respectively. Concretely, the type strain of the species *R. indigoferae* CCBAU 71042^T presented 98.24% identity with respect to *R. leguminosarum* USDA 2370^T. This high identity value suggests that the strain CCBAU 71042^T could belong to the species *R. leguminosarum* in agreement with those found in *rrs* and *atpD* and *recA* genes, 99.7, 99.3 and 99.1%, respectively. Similar results were found between *R. pisi* and *R. fabae* with identity values of 99.9, 98.5 and 97.6%, respectively, in *rrs*, *atpD* and *recA* genes and near to 95% in the *celC* gene. All these high identity values question the affiliation of *R. indigoferae* and *R. fabae* to species different from *R. leguminosarum* and *R. pisi,* respectively, showing that the taxonomic status of *R. indigoferae* and *R. fabae* should be revised.

Several legume nodulating strains whose genomes have been completely sequenced were classified into species *R. leguminosarum*, such as *R. leguminosarum* bv viciae 3841, *R. leguminosarum* bv trifolii WSM1325 and *R. leguminosarum* bv trifolii WSM2304. For all these strains, both housekeeping genes and *celC* gene analysis suggest that they are misclassified since they have identities lower than 91% in *celC* gene and 95% in both *recA* and *atpD* genes with respect to the type strain *R. leguminosarum* USDA 2370^T. Similar values have been previously found between different species of genus *Rhizobium*, such as *R. leguminosarum* and *R. pisi* [22] or *R. rhizogenes* and *R. lusitanum* [31]. Therefore, it is evident that the taxonomic status of the strains 3841, WSM1325 and WSM2304 should be revised according to the current criteria to define bacterial species.

Considering that the *celC* gene was firstly sequenced in strains nodulating clover we obtained the sequences of the *recA* and *atpD* genes of two of these strains, ANU843 and R200, which had also been classified into the biovar trifolii of species *R. leguminosarum*. The identity values of the *celC* gene from these strains were lower than 91% with respect to the type strain of this species. Congruently the identity values of *recA* and *atpD* genes were lower than 96.5 and 93%, respectively. These values are in the limit for species differentiation within genus *Rhizobium* [22, 31] and therefore they could not belong to *R. leguminosarum*.

All these results showed the taxonomic complexity of the *R. leguminosarum* group and the need to revise the criteria used to name the strains it includes because it seems that many of them have been still named on the basis of the host they nodulated even if their genomes have been completely sequenced. The results of the *celC* gene analysis showed that the suitability of this gene to differentiate species of the phylogenetic group of *R. leguminosarum* is higher even than other protein-encoding genes commonly used to this purpose [8].

Phylogentically close to *R. leguminosarum* are located *R. etli* and *R. phaseoli*, two species whose taxonomic history is complex because *R. etli* is the former biovar 1 of the old species *R. phaseoli* [27] and for many years it was thought that this last species had disappeared being a biovar of *R. leguminosarum* [13]. However, we recently showed that *R. phaseoli*, *R. leguminosarum* and *R. etli* are three valid species that can be clearly differentiated on the basis of *recA* and *atpD* genes [22]. This differentiation is even more evident on the basis of the *celC* gene that presented identities lower than 85% among these three species (Fig. 1).

In the case of *R. etli,* as was previously mentioned, the complete genome of two strains has been recently sequenced, the type strain CFN42^T and the strain CIAT652. However, the results of the *recA* and *atpD* analysis showed the nearly complete identity of the last strain with respect to *R. phaseoli* ATCC 14482^T. Although in the *rrs* gene analysis the differences are not clear (Fig. 2), the results on the basis of *recA* and *atpD* genes clearly showed the high phylogenetic distance between strains *R. etli* CFN42^T and CIAT652 (Figs. S1 and S2, supplementary material). These results were corroborated after the analysis of the *celC* gene (Fig. 1), showing that strain CIAT652 is phylogenetically closer to *R. phaseoli* ATCC 14482^T (99% identity) than to *R. etli* CFN42^T (85% identity). Therefore the strain CIAT 652 should be classified into species *R. phaseoli* and not into *R. etli* and thus some conclusions of the work of González *et al*. [9] should be revised, mainly those concerning to the variations of the chromosomal genes within the same species. They found an identity between the sequences of the chromosomes of *R. etli* CFN42^T and CIAT 652 of 90-95% and possibly these values are the limit among different species in rhizobia since strains of *R. phaseoli* and *R. etli* showed values of DNA-DNA relatedness lower than 70% [31]. Nevertheless, it is premature to establish conclusions on the basis of complete genomes comparison and the threshold values of *in silico* matching of core genomes for different species should be estimated after the analysis of several strains of each species previously compared by other methods.

The analysis of the *celC* gene also confirmed the phylogenetic closeness of the group formed by *R. mongolense* and *R. gallicum*, two species whose type strains presented 98.2% and 97.7% identity values in *atpD* and *recA* genes, respectively. Although, as expected, the *celC* gene presented a lower identity value, 95.8%, it is higher than those found among some different species of genus *Rhizobium*. These high identity values evidence that the taxonomic status of these two species should be revised since the DNA-DNA hybridization values are not available in the literature [3, 32].

An interesting group is the one formed by *R. tropici* and *R. hainanense*, two species able to induce nodules, and *R. rhizogenes*, a species that include strains able to cause plant pathogenic symptoms (hairy roots or tumours) formerly classified into *Agrobacterium* and currently into genus *Rhizobium* [36]. In spite of the disagreement of many authors with the reclassification of the complete genus *Agrobacterium* into genus *Rhizobium* [6], it is necessary to point out that the species *R. rhizogenes* really belong to genus *Rhizobium* and not to genus *Agrobacterium* [33]. The type strain of this species, ATCC 11325^T is able to produce hairy roots [14], other strains such as 163C are able to produce plant tumours [34] and K84 is a non-pathogenic strain used in biological control [21]. This strain, whose complete genome has been recently sequenced, was classified as *Agrobacterium* (currently *Rhizobium*) *radiobacter*, however it has *rrs, recA* and *atpD* genes nearly identical to those of *R. rhizogenes* ATCC 11325^T for which we recently reclassified it into this species [33]. Our results on *celC* gene confirmed this reclassification since this gene is nearly identical in K84 and *R. rhizogenes* NCPPB 2991^T (=ATCC 11325^T). Our results also confirmed that the former genus *Agrobacterium*

contained species from divergent phylogenetic origins since the *celC* gene of *R. rhizogenes* strains is highly divergent from that of *R. radiobacter* C58, a tumourigenic strain close to the type strain of *R. radiobacter* NCPPB 2437^T and whose complete genome has also been sequenced. These results, together with those of *rrs* and housekeeping gene analysis, seem to support that *Agrobacterium* is a definable genus within the family *Rhizobiaceae* as proposed Farrand *et al.* [6].

Besides the strains previously cited, another interesting strain is *Rhizobium* sp. NGR234 whose complete genome has also been sequenced. This strain, which has not yet been assigned to a species, has a gene homologous to *celC* located in a plasmid and forms an independent phylogenetic branch (Fig. 1) that is close to *Ensifer medicae* (formerly *Sinorhizobium medicae*) WSM419 whose *celC*-like gene has also been located in a plasmid. The case of strain NGR234 highlights the need of a correct classification and nomenclature of bacterial strains because it has been named *Rhizobium* sp. in Genbank when its correct name is *Ensifer* sp. since Jarvis *et al.* [12] already showed that this strain is close to *Sinorhizobium fredii* (currently *Ensifer fredii*). When considering the correct nomenclature of strain NGR234 we can conclude that all species analysed up to date in the genus *Rhizobium* carry a *celC* gene in the chromosome whereas those of genus *Ensifer* have a *celC*-like gene in a plasmid. It is also remarkable that, considering the low identities of *atpD* and *recA* genes from strain NGR234 with respect to those of the *Ensifer* species (lower than 92 and 95% identity, respectively), this strain probably belongs to a new species of this genus. However, more strains of this group could be necessary to define new species according to the recommendations of the Subcommittee on the taxonomy of *Agrobacterium* and *Rhizobium* [19].

Therefore, it seems that many strains selected for complete genome sequencing are initially missclassified and this represents a real problem more significant than imagined since the characteristics and conclusions obtained after the genome analysis could be linked to a wrong species. It has been already pointed out in the case of *Mesorhizobium loti* MAFF303099 that really belongs to the species *M. huakuii* [30] and *Agrobacterium* (currently *Rhizobium*) *radiobacter* K84 that really belongs to *Rhizobium rhizogenes* [33]. We think that we have currently the necessary tools for proper classification of strains and especially when the complete genomes are available, the correct naming of these strains is unavoidable.

Comparison of *celC* and plasmidic gene phylogenies. All strains of genus *Rhizobium* whose genome have been sequenced up to date have a *celC* gene in the chromosome, whereas a *celC*-like gene has been located in plasmids in the genus *Ensifer*. Therefore, in spite of the complete congruence of the *celC* phylogeny and those of chromosomal core genes in genus *Rhizobium*, we also compared the results of *celC* gene analysis with those of symbiotic and virulence genes located in plasmids in fast-growing rhizobia.

From the symbiotic genes pool, the *nodC* has been widely analysed in many rhizobial strains being related with the host range of rhizobia and the promiscuity degree of the hosts [11, 18, 25]. This gene, located in a symbiotic plasmid (pSym) in the fast-growing rhizobial species, has been used as phylogenetic marker for biovars differentiation within a *Rhizobium* species, such as *R. leguminosarum* biovars viciae, trifolii and phaseoli [7] or *R. gallicum* biovars gallicum and phaseoli [18]. For the comparison of *celC* and *nodC* genes we used strains from different biovars of *R. leguminosarum, R. gallicum* and *R. etli* (Table S1).

As expected, the distribution of biovars was coherent with the *nodC* phylogeny and in this way strains nodulating different hosts clustered in separated branches of the trees with independence of the chromosomal species they belong (Fig. S3, supplementary material). For example *R. leguminosarum* nodulating *Vicia*, *Trifolium* and *Phaseolus* strains are located in clearly separated branches whereas they have closely related chromosomal genes. Some of these strains even have nearly identical chromosomal genes as occurs with the strains RVS03, nodulating *Vicia,* ATCC 14480 nodulating *Trifolium* and RPVF18 nodulating *Phaseolus.* In all cases the *celC* gene phylogeny is not related with that of *nodC* gene but with that of chromosomal genes. The same results were found in the case of the two biovars of *R. gallicum*, whose *nodC* occupied different branches (identity lower than 85%) whereas their *celC* genes are identical. We also found the reverse, strains of biovar phaseoli from *Rhizobium gallicum*, *R. etli*, *R. phaseoli* and *R. leguminosarum* have dissimilar *celC* genes since they belong to different species, whereas their *nodC* genes are highly similar (identities higher than 96%) because all of them belong to the same biovar. Also *R. leguminosarum* and *R. pisi*, both isolated from *Pisum* nodules and belonging to the biovar viciae [7], have *nodC* genes nearly identical but, as correspond to different species have divergent *recA, atpD* and *celC* genes between them.

Concerning the virulence genes, the strains able to form tumours or hairy roots are currently included in different species of genus *Rhizobium*. These strains carry virulence genes in plasmids named pTi (Tumour inducing plasmid) or pRi (Root inducing plasmid) some of them located in an operon, which is regulated by the *virA* gene [16]. It has been previously showed the common phylogenetic origin of this gene in strains whose chromosomes are phylogenetically divergent [33]. For example, strains 163C of *R. rhizogenes* and C58 of *R. radiobacter* have nearly identical *virA* gene (Fig. S4, supplementary material) although their *rrs, recA* and *atpD* genes are phylogenetically divergent (Fig. 1 and figs. S1 and S2 from supplementary material). The analysis of *celC* genes of these strains showed that they are highly divergent as occurs in other chromosomal genes. By the contrary, strains ATCC 11325^T and 163C, both belonging to *R. rhizogenes*, have almost identical *celC* genes whereas their *virA* genes are phylogenetically divergent (similarity of 85.6%). All these results showed no relationship between *celC* and symbiotic or virulence genes phylogenies in genus *Rhizobium*.

Potential of *celC* gene for fast-growing rhizobia identification. Considering that the *celC* gene has congruent phylogenies with other chromosomal genes, and it has not relation with plasmidic ones, it could be included in identification processes. To evaluate this possibility we analysed several strains belonging to different species of genus *Rhizobium* that have been previously identified by sequencing of other genes.

In first place we select several strains isolated from the host of the cross-inoculation group of *Vicia*. Some of these strains have been identified as *R. leguminosarum* as occurs for strains RPVF18 isolated from *Phaselous*, RVS03 isolated from *Vicia* and ATCC 14480 isolated from *Trifolium* that present higher similarities between *rrs, recA* and *atpD* genes than *R. leguminosarum* USDA2370^T. The *celC* gene of these strains (Fig. 1) was also very closely related to the type strain with identities higher than 97.6%.

Some strains from *R. leguminosarum* phylogenetic group isolated from *Vicia*, *Phaseolus* and *Trifolium* have identical *rrs* gene with *R. leguminosarum* USDA2370^T but divergent housekeeping gene sequences [2, 7, 22]. Some of these strains presented identities in *recA* and *atpD* genes lower than 94% (RPVR32) or 92% (RTP05), respectively, and probably do not belong to *R. leguminosarum* in spite of the complete identity of their *rrs* genes. In these cases, the *celC* gene analysis allows a better differentiation with identities lower than 87% with respect to the strain USDA 2370^T. One of these strains, RTP05, presents identical housekeeping and *celC* genes as the strain WSM1325, whose genome has been completely sequenced and whose taxonomic status should be revised, as we discussed above.

A strain from the same group of *R. leguminosarum*, CVIII14, was included because it has a *recA* gene highly

similar (97.9%) to that of *R. leguminosarum* USDA2370^T, but an *atpD* gene divergent (92.7%). The *celC* gene has an identity of 89.5% confirming the usefulness of this gene for identification purposes since it has a variability degree only slightly lower than *atpD* gene.

For the group of *R. etli* we analysed the *celC* gene of the strain Mim-2. This strain presented a *rrs* almost identical to that of *R. pisi*, which show *rrs* gene highly homologous with those of *R. etli* and *R. phaseoli.* This strain has been named *R. etli* [35], however, its *recA* and *atpD* genes are closer to *R. phaseoli* ATCC 14482^T than to *R. etli* CFN42^T. Concretely, they are identical to that of other strain, CIAT 652, whose genome has also been sequenced and that was also named *R. etli*. The *celC* genes of these two strains are almost identical showing 99.6% identity with respect to *R. phaseoli* ATCC 14482^T. Therefore, strains Mim-2 and CIAT 652 should be included into *R. phaseoli* instead of *R. etli* and perhaps other strains currently named *R. etli* really belong to *R. phaseoli,* a species that has not been considered valid for many authors for decades. This implies that the taxonomic status of the strains currently named *R. etli* should be revised.

Other strains included in this study able to nodulate *Phaseolus* were PhD12 belonging to *R. gallicum* and Br856 belonging to *R. tropici.* In both cases they presented almost identical *rrs, recA* and *atpD* genes with respect to the type strains of their respective species. As expected, the *celC* genes of the strains PhD12 and Br856 were also identical to *R. gallicum* R602sp^T and *R. tropici* CIAT 899^T, respectively.

An interesting case is that of the strain RPA12 which was isolated from *Prosopis* and presented a *rrs* gene identical to that of *R. giardinii* H152^T but identities lower than 75% in the ITS region [11]. This fact suggested that the strain RPA12 could belongs to a different species. However the analysis of *celC* gene showed 99% identity between strains RPA12 and H152^T suggesting they belong to the same species. Therefore we analysed the *recA* and *atpD* genes of strain RPA12 in this study and their identity values, higher than 99%, are in complete agreement with those found for the *celC* gene showing that the strain RPA12 belongs to the species *R. giardinii*.

Finally, as was previously mentioned, the strain *R. rhizogenes* 163C able to induce legume nodules and plant tumours [34] and the strain *R. rhizogenes* K84 [33], unable to induce plant symptoms, were correctly identified on the basis of their *celC* genes. These genes, as occurs with the *rrs, recA* and *atpD*, were nearly identical between them and with respect to that of the type strain of *R. rhizogenes* ATCC 11325^T, which produce a different symptom, hairy roots. Accordingly no relation was found between phylogenies of *celC* gene and those of virulence genes as occurred in the case of symbiotic genes.

In summary, the results of this work clearly indicate that *celC* gene phylogeny is well correlated with those of chromosomal genes with different evolutionary rate such as *rrs*, *recA* and *atpD* genes in the genus *Rhizobium*. As no relationships were found between *celC* and plasmidic gene phylogenies, we can conclude that the *celC* is a protein-codifying gene useful for taxonomic purposes within genus *Rhizobium* that allows the differentiation of closely related species with high similarity in their *rrs* genes. The phylogenetic analyses of the *celC* and other core genes of *Rhizobium* strains whose genome has been completely sequenced evidence that many of them are misnamed showing that a correct classification and nomenclature of strains should be a requisite for the publication of their complete genomes. Moreover it will be desirable to have the complete genomes of the type strains from each species in order to avoid erroneous conclusions established by comparing misnamed strains.

Aknowledgements. This work was supported by Junta de Castilla y León Grant GR49 and MICINN Grant AGL2008-03360/AGR. M.R. was supported by a PhD fellowship from MICINN and R. R. by a postdoctoral contract of "Ramón y Cajal" Program from Spanish Government. We thank E. Martínez-Romero, P. van Berkum and N. Amarger for providing some of the strains used in this work.

Material and Methods

Bacterial strains. We analysed the type strains of species from genus *Rhizobium* isolated from nodules of different legumes and from tumours formed in different hosts. In addition, several strains representative to biovars viciae, trifolii and phaseoli from *R. leguminosarum*, biovars gallicum and phaseoli from *R. gallicum* and biovars phaseoli and mimosae from *R. etli* were analysed (Table S1). Strains whose complete genome has been sequenced were also included.

Amplification and sequencing of *celC* gene. A 800 bp fragment of the *celC* gene was amplified and sequenced by using the following primers designed in this work from conserved sequences in the *celC* gene available in databases CelC1F (5'-atcagccacagcgaagggca-3') and CelC2R (5'-cacagacactccggatgc-3'). PCR amplification was carried out with an AmpliTaq Gold reagent kit (Applied Biosystems Inc., USA) following the manufacturer's instructions. The conditions were as follows: pre-heating at 95°C for 9 min; 35 cycles of denaturing at 95°C for 1 min; annealing at 54°C for 1 min and extension at 72°C for 1 min (for 16S rRNA gene was 2 min), and a final extension at 72°C for 7 min. The PCR products were electrophoresed in 1% agarose gels and stained with ethidium bromide. The band corresponding to the *celC* gene was purified directly from the gel by room temperature centrifugation using a DNA gel extraction device (Millipore Co., USA) for 10 min at 5000 *g* according to the manufacturer's instructions. The sequence reaction was performed on an ABI377 sequencer (Applied Biosystems Inc., USA) using a BigDye terminator v3.0 cycle sequencing kit as supplied by the manufacturer.

Amplification and sequencing of *rrs, recA* and *atpD* genes. The amplification and sequencing of *rrs* gene was carried out according to Rivas *et al.* [23], that of *recA* and *atpD* genes according to Gaunt *et al.* [8], that of *nodC* gene as described by Laguerre *et al.* [18] and that of *virA* gene as described by Velázquez *et al.* [34]. Gene sequencig was performed with the same primers using for DNA amplification following the same protocol described for *celC* gene. These genes were sequenced in this work for the strains from which their sequences are not available in Genbank.

Phylogenetic analysis. The *celC* sequences obtained were compared with those from the GenBank using the BLASTN program [1]. For phylogenetic analysis sequences were aligned using the Clustal W software [29]. The distances were calculated according to Kimura´s two-parameter model [15]. Phylogenetic trees were inferred using the neighbour-joining method [26] using the MEGA 4.0 [28].

References

[1] Alschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J. (1990) Basic local alignment search tool. J. Mol. Biol. 215, 403-410.

[2] Álvarez-Martínez, E. R., Valverde, A., Ramírez-Bahena, M. H., García-Fraile, P., Tejedor, C., Mateos, P. F., Santillana, N., Zuñiga, D., Peix, A. and Velázquez, E. (2009). The analysis of core and symbiotic genes of rhizobia nodulating *Vicia* from different continents reveals their common phylogenetic origin and suggests the distribution of *Rhizobium leguminosarum* strains together with *Vicia* seeds. *Arch Microbiol* 191, 659-668.

[3] Amarger, N., Macheret, V., Laguerre, G. (1997) *Rhizobium gallicum* sp. nov. and *Rhizobium giardinii* sp. nov., from *Phaseolus vulgaris* nodules. Int. J. Syst. Bacteriol. 47, 996-1006.

[4] An, J.M., Lim, W.J., Hong, S.Y., Shin, E.C., Kim, E.J., Kim, Y.K., Park, S.R., Yun, H.D. (2004) Cloning and characterization of *ce/8A* gene from *Rhizobium leguminosarum* bv. trifolii 1536. Lett. Appl. Microbiol. 38, 296-300.

[5] Ausmees, N., Jonsson, H., Hoglund, S., Ljunggren, H., Lindberg, M. (1999) Structural and putative regulatory genes involved in cellulose synthesis in *Rhizobium leguminosarum* bv. trifolii. Microbiology 145, 1253-1262.

[6] Farrand, S.K., van Berkum, P., Oger, P. (2003). *Agrobacterium* is a definable genus of the family *Rhizobiaceae*. Int. J. Syst. Evol. Microbiol. 53, 1681-1687.

[7] García-Fraile, P., Mulas-García, D., Peix, A., Rivas, R., González-Andres, F., Velázquez, E. (2010) *Phaseolus vulgaris* is

nodulated in northern Spain by *Rhizobium leguminosarum* strains harboring two *nodC* alleles present in American *Rhizobium etli* strains: biogeographical and evolutionary implications. Can. J. Microbiol. 56, 657-666.

[8] Gaunt, M.W., Turner, S.L., Rigottier-Gois, L., Lloyd-Macgilp, S.A., Young, J.P. (2001) Phylogenies of *atpD* and *recA* support the small subunit rRNA-based classification of rhizobia. Int. J. Syst. Evol. Microbiol. 51, 2037-2048.

[9] González, V., Acosta, J.L., Santamaría, R.I., Bustos, P., Fernández, J.L., Hernández-González, I.L., Díaz, R., Flores, M., Palacios, R., Mora, J., Dávila, G. (2010) Conserved symbiotic plasmid DNA sequences in the multireplicon pangenomic structure of *Rhizobium etli*. Appl. Environ. Microbiol. 76, 1604-1614.

[10] Guiseppi, C., Aymeric, J. L., Cami, B., Barras, F., Creuzet, N. (1991) Sequence analysis of the cellulase-encoding celY gene of *Erwinia chrysanthemi*: a possible case of interspecies gene transfer. Gene 106, 109-114.

[11] Iglesias, O., Rivas, R., García-Fraile, P., Abril, A., Mateos, P.F., Martínez-Molina, E, Velázquez, E. (2007) Genetic characterization of fast-growing rhizobia able to nodulate *Prosopis alba* in North Spain. FEMS Microbiol. Lett. 277, 210-216.

[12] Jarvis, B.D., Downer, H.L., Young, J.P. (1002) Phylogeny of fast-growing soybean-nodulating rhizobia support synonymy of *Sinorhizobium* and *Rhizobium* and assignment to *Rhizobium fredii*. Int. J. Syst. Bacteriol. 42, 93-96.

[13] Jordan, D.C. (1984) Family III. *Rhizobiaceae*. In: Krieg, N.R., Holt, J.G. (Eds) Bergey's Manual of Systematic Bacteriology. The Williams and Wilkins Co, Baltimore. pp 234-242.

[14] Kersters, K., de Ley, J. (1984) Genus III. *Agrobacterium* In: Krieg, N.R., Holt, J.G. (Eds.), Bergey's Manual of Systematic Bacteriology, Williams and Wilkins Co., Baltimore, pp. 244-254.

[15] Kimura, M. (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. 16, 111-120.

[16] Krishnamohan, A., Balaji, V., Veluthambi, K. (2001) Efficient *vir* gene induction in *Agrobacterium tumefaciens* requires *virA*, *virG*, and *vir* box from the same Ti plasmid. J. Bacteriol. 183, 4079-4089.

[17] Kuykendall, L.D. (2005) Family I. Rhizobiaceae Conn 1938, 321^{AL}. In: Brenner, D.J., Krieg, N.R., Stanley, J.T (Eds.) Bergey's Manual of Systematic Bacteriology. Vol. 2, part C. Springer, New York. pp 324-361.

[18] Laguerre, G., Nour, S.M., Macheret, V., Sanjuan, J., Drouin, P., Amarger, N. (2001) Classification of rhizobia based on *nodC* and *nifH* gene analysis reveals a close phylogenetic relationship among *Phaseolus vulgaris* symbionts. Microbiology 147, 981-993.

[19] Lindström, K., Young, J.P.W. (2009) International Committee on Systematics of Prokaryotes; Subcommittee on the taxonomy of *Agrobacterium* and *Rhizobium*: Minutes of the meetings, 31 August 2008, Gent, Belgium. Int. J. Syst. Evol. Microbiol. 59, 921-922.

[20] Matthysse, A.G., Thomas, D.L., White, A.R. (1995) Mechanism of cellulose synthesis in *Agrobacterium tumefaciens*. J. Bacteriol. 177, 1076-1081.

[21] Moore, L.W., Warren, G. (1979) *Agrobacterium radiobacter* strain K84 and biological control of crown gall. Annu. Rev. Phytopathol. 17, 163-179.

[22] Ramírez-Bahena, M.H., García-Fraile, P., Peix, A., Valverde, A., Rivas, R., Igual, J.M., Mateos, P.F., Martínez-Molina, E., Velázquez, E. (2008) Revision of the taxonomic status of the species *Rhizobium leguminosarum* (Frank 1879) Frank 1889, *R. phaseoli* Dangeard 1926^{AL} and *R. trifolii* Dangeard 1926^{AL}. *R. trifolii* is a later synonym of *R. leguminosarum.* Reclassification of the strain *Rhizobium leguminosarum* DSM 30132^{T} (=NCIMB 11478^{T}) into the new species *Rhizobium pisi* sp. nov. Int J Syst Evol Microbiol. 58, 2484-2490.

[23] Rivas, R., García-Fraile, P., Mateos, P.F., Martínez-Molina, E., Velázquez, E. (2007) Characterization of xylanolytic bacteria present in the bract phyllosphere of the date palm *Phoenix dactylifera*. Lett. Appl. Microbiol. 44, 181-187.

[24] Robledo, M., Jiménez-Zurdo, J.I., Velázquez, E., Trujillo, M.E., Zurdo-Piñeiro, J.L., Ramírez-Bahena, M.H., Ramos, B., Díaz-Mínguez, J.M., Dazzo, F.B., Martínez-Molina, E., Mateos, P.F. (2008) Rhizobium cellulase CelC2 is essential for primary symbiotic infection of legume host roots. Proc. Natl. Acad. Sci. U S A. 105, 7064-7069.

[25] 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.

[26] Saitou, N., and Nei, M. (1987) A neighbour-joining method: a new method for reconstructing phylogenetics trees. Mol. Biol. Evol. 44, 406-425.

[27] Segovia, L., Young, J.P.W., Martínez-Romero, E. (1993) Reclassification of American *Rhizobium leguminosarum* biovar phaseoli type I strains as *Rhizobium etli* sp. nov. Int. J. Syst. Bacteriol. 43, 374-377.

[28] Tamura, K., Dudley, J., Nei, M., Kumar, S. (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol. Biol. Evol. 24, 1596-1599.

[29] Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G. (1997) The clustalX windows interface: flexible strategies for multiple sequence alignement aided by quality analysis tools. Nucleic Acid Res. 24, 4876-4882.

[30] Turner, S.L., Zhang, X.X., Li, F.D., Young, J.P. (2002) What does a bacterial genome sequence represent? Mis-assignment of MAFF 303099 to the genospecies *Mesorhizobium loti*. Microbiology. 148, 3330-3331.

[31] Valverde, A., Igual, J.M., Peix, A., Cervantes, E., Velázquez, E. (2006) *Rhizobium lusitanum* sp. nov. a bacterium that nodulates *Phaseolus vulgaris*. Int. J. Syst. Evol. Microbiol. 56, 2631-2637.

[32] van Berkum, P., Beyene, D., Bao, G., Campbell, T.A., Eardly, B.D. (1998) *Rhizobium mongolense* sp. nov. is one of three rhizobial genotypes identified which nodulate and form nitrogen-fixing symbioses with *Medicago ruthenica* [(L.) Ledebour]. Int. J. Syst. Bacteriol. 48, 13-22.

[33] Velázquez, E., Palomo, J.L., Rivas, R., Guerra, H., Peix, A., Trujillo, M.E., García-Benavides, P., Mateos, P.F., Wabiko, H., Martínez-Molina, E. (2010) Analysis of core genes supports the reclassification of strains *Agrobacterium radiobacter* K84 and *Agrobacterium tumefaciens* AKE10 into the species *Rhizobium rhizogenes*. Syst. Appl. Microbiol. 33, 247-251.

[34] Velázquez, E., Peix, A., Zurdo-Piñeiro, J.L., Palomo, J.L., Mateos, P.F., Rivas, R., Muñoz-Adelantado, E., Toro, N., García-Benavides, P., Martínez-Molina, E. (2005) The coexistence of symbiosis and pathogenicity-determining genes in *Rhizobium rhizogenes* strains enables them to induce nodules and tumours or hairy roots in plants. Mol. Plant Microbe Interact. 18, 1325-1332.

[35] Wang, E. T., Rogel, M. A., García-de los Santos, A., Martínez-Romero, J., Cevallos, M. A. and Martínez-Romero, E. (1999). *Rhizobium etli* bv. mimosae, a novel biovar isolated from *Mimosa* affinis. *Int J Syst Bacteriol* 49 Pt 4, 1479-1491.

[36] Young, J.M., Kuykendall, L.D., Martínez-Romero, E., Kerr, A., Sawada, H. A (2001) revision of *Rhizobium* Frank 1889, with an emended description of the genus, and the inclusion of all species of *Agrobacterium* Conn 1942 and *Allorhizobium undicola* de Lajudie et al. 1998 as new combinations: *Rhizobium radiobacter*, *R. rhizogenes*, *R. rubi, R. undicola* and *R. vitis*. Int. J. Syst. Evol. Microbiol. 51, 89-103.

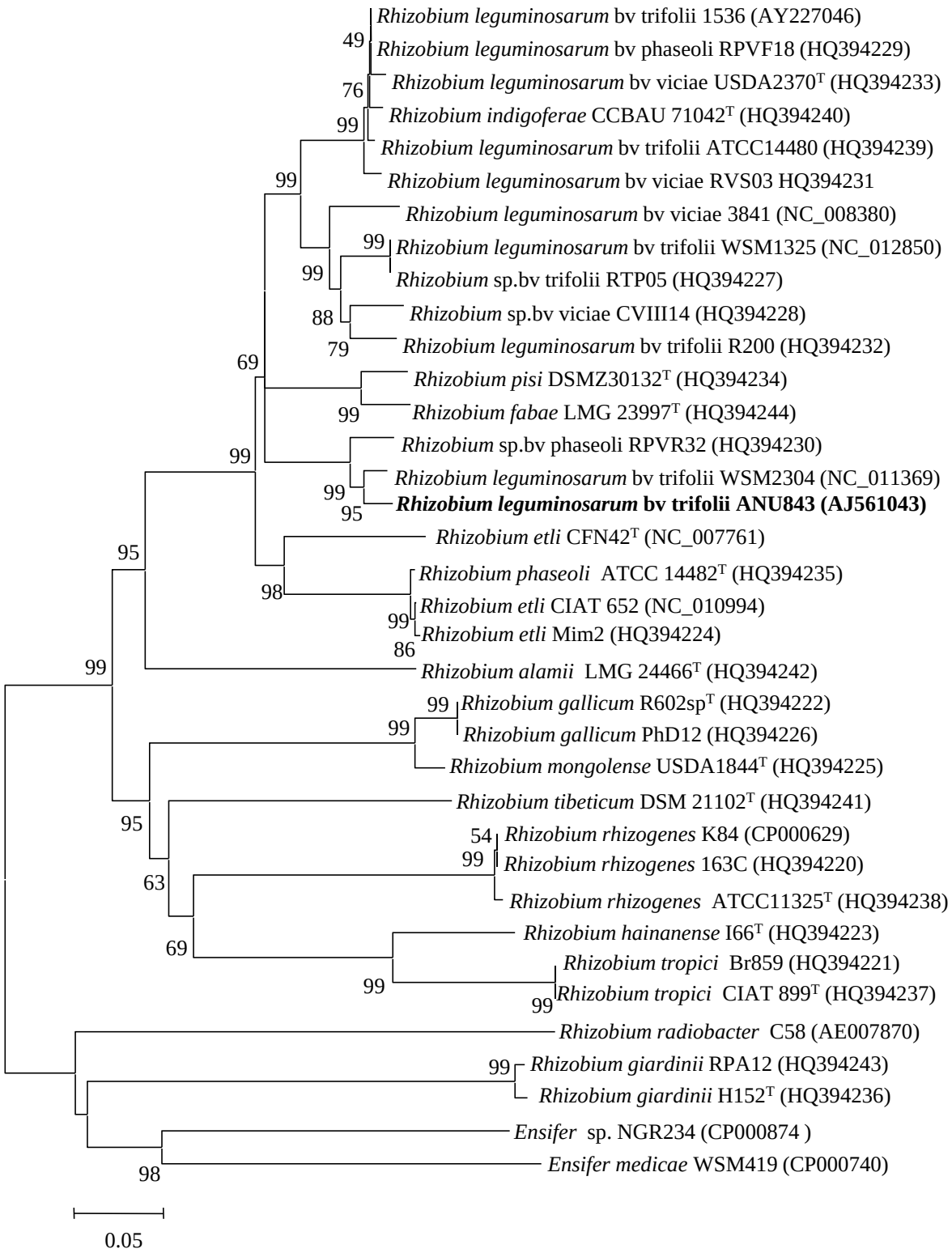

Fig. 1. Neighbour-joining phylogenetic tree based on *celC* gene sequences (800 nt). Bootstrap values calculated for 1000 replications are indicated. Bar, 5 nt substitution per 100 nt. Accession numbers from Genbank are given in brackets.

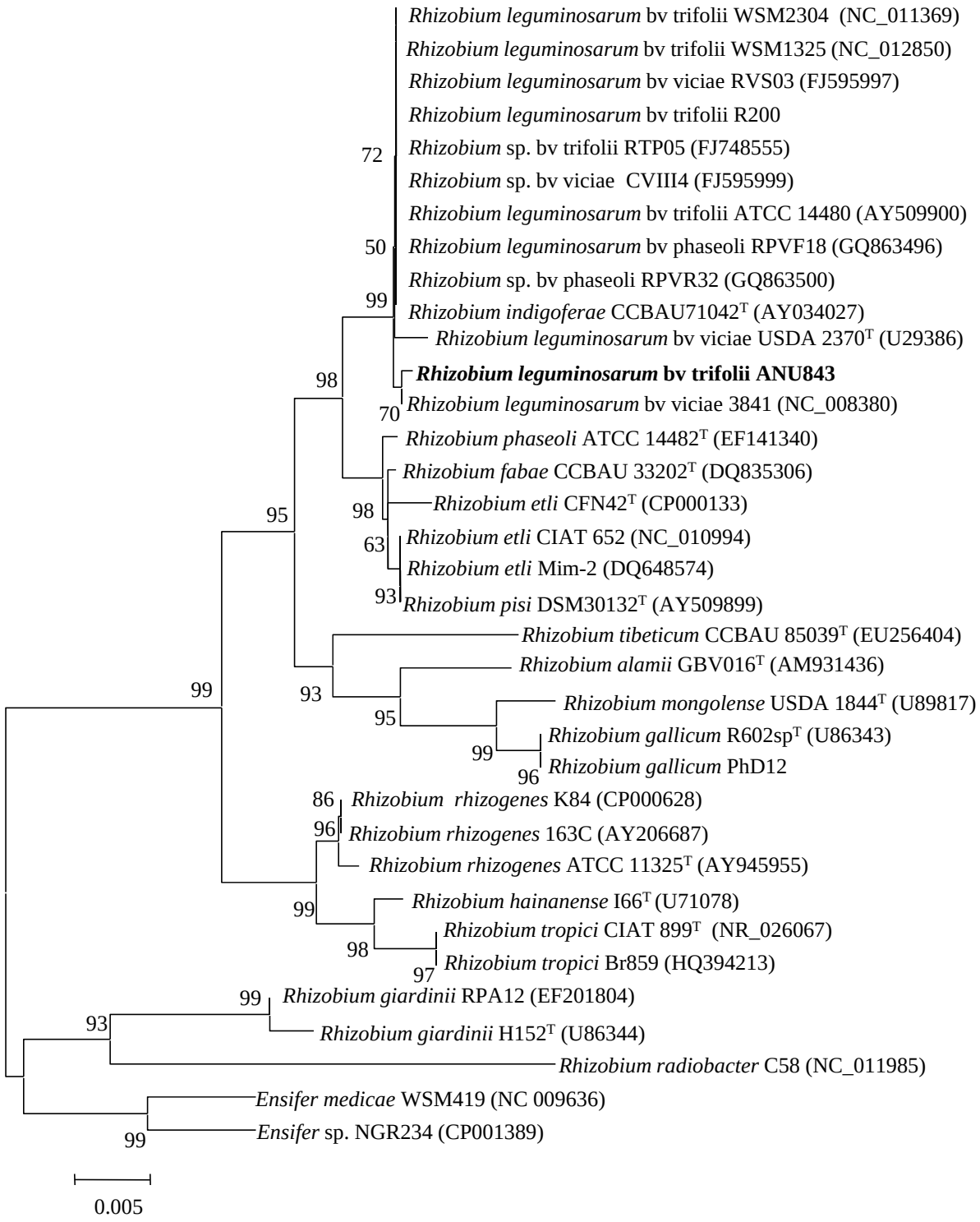

Fig. 2. Neighbour-joining phylogenetic rooted tree based on *rrs* sequences (1475 nt). Bootstrap values calculated for 1000 replications are indicated. Bar, 5 nt substitution per 1000 nt. Accession numbers from Genbank are given in brackets

Supplementary material

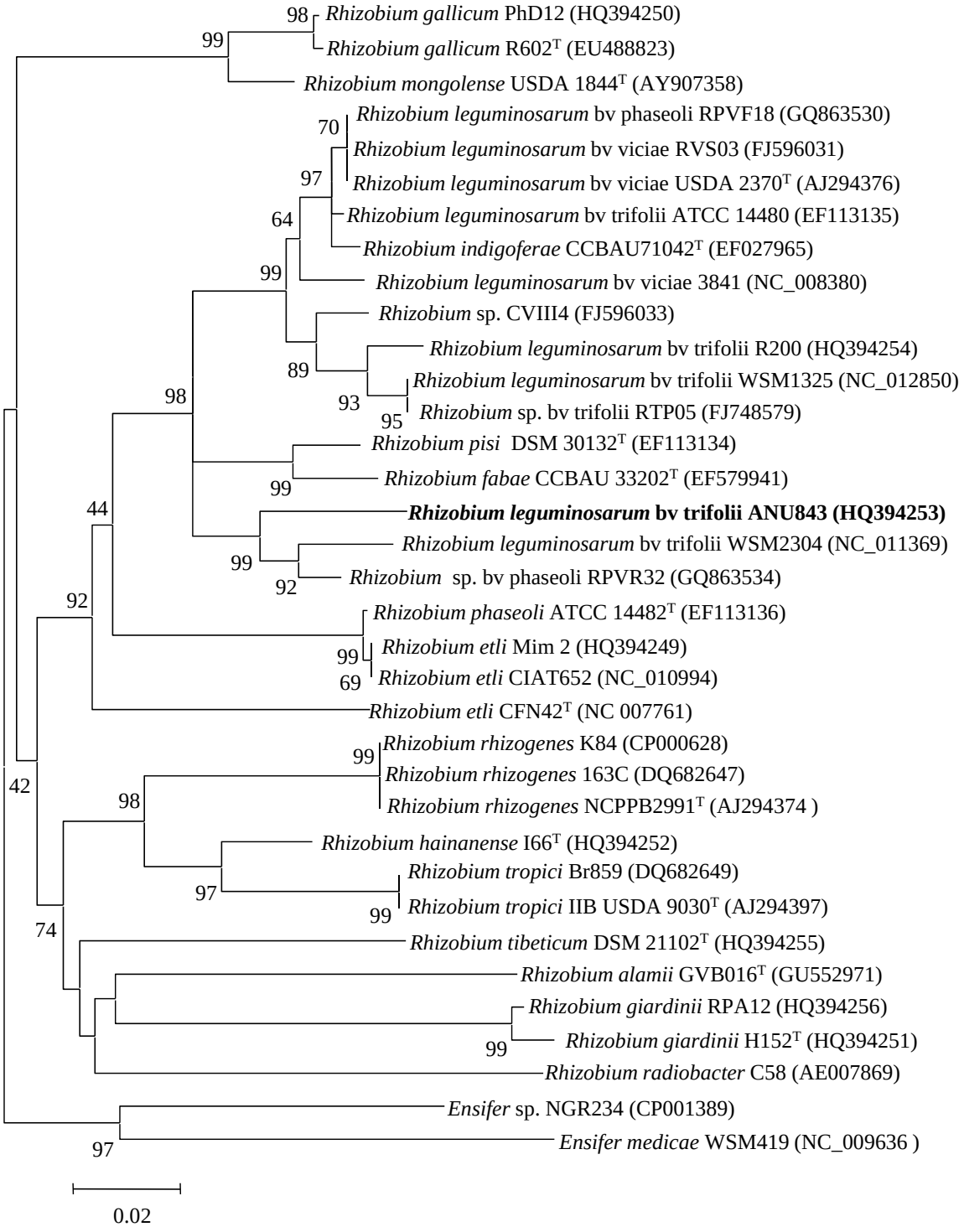

Fig. S1. Neighbour-joining phylogenetic tree based *recA* gene sequences (about 520 nt). Bootstrap values calculated for 1000 replications are indicated. Bar, 1 nt substitution per 100 nt. Accession numbers from Genbank are given in brackets.

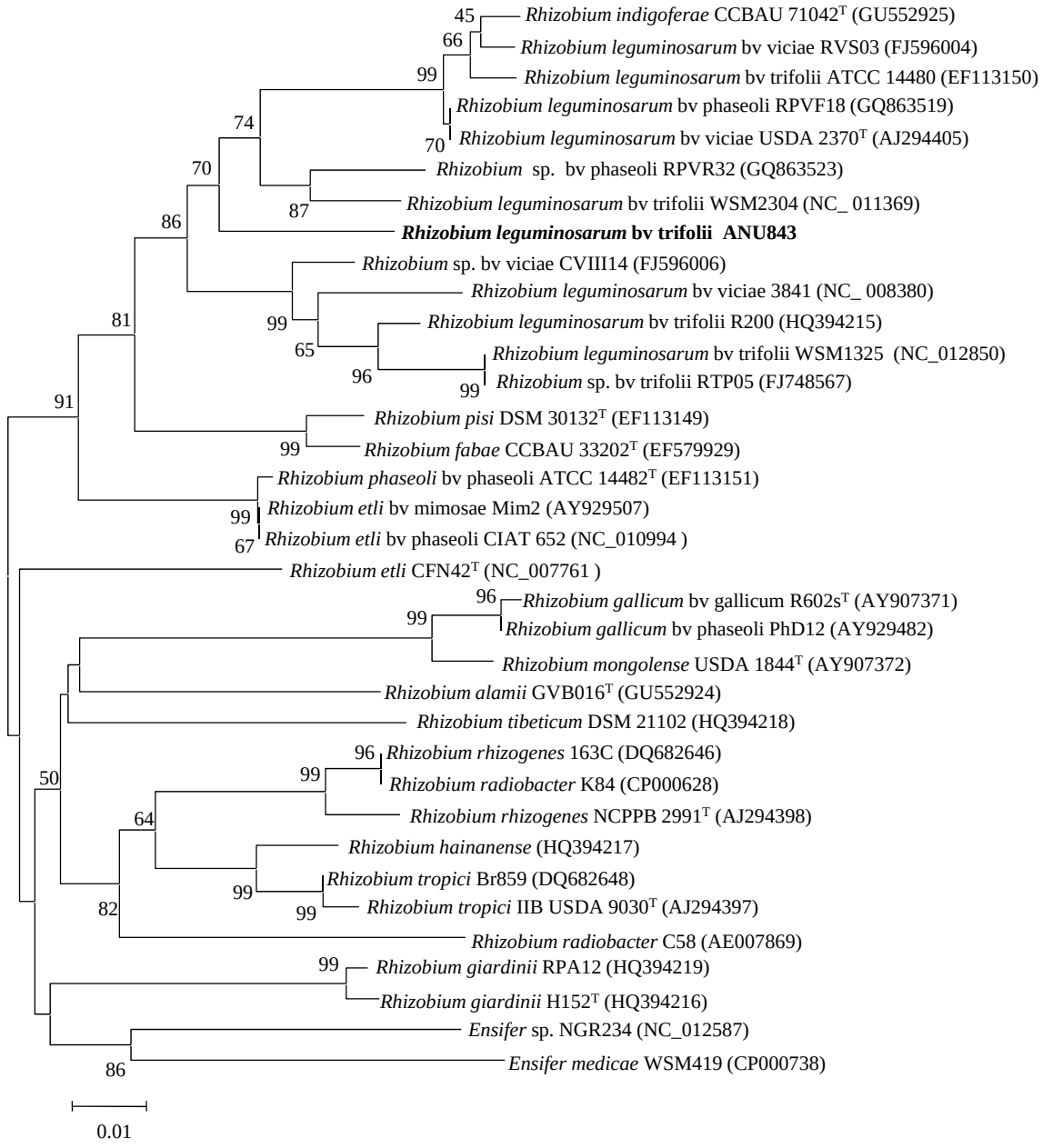

Fig. S2. Neighbour-joining phylogenetic tree based on *atpD* gene sequences (about 500 nt). Bootstrap values calculated for 1000 replications are indicated. Bar, 1 nt substitution per 100 nt. Accession numbers from Genbank are given in brackets.

Table S1. Strains belonging to different biovars included in this study.

Strains	Biovar	Source or reference
R. etli CFN42^T	phaseoli	[7, 27]
R. etli Mim-2	mimosae	[35]
Rhizobium gallicum R602^T	gallicum	[3]
Rhizobium gallicum PhD12	phaseoli	[3]
R. leguminosarum USDA 2370^T	viciae	[7, 13]
R. leguminosarum RPVF18	phaseoli	[7]
R. leguminosarum RVS03	viciae	[7]
R. leguminosarum ATCC 14480	trifolii	[7, 22]
R. pisi DSM 30132^T	viciae	[7]
R. phaseoli ATCC 14482^T	phaseoli	[7]
Rhizobium sp. RTP05	trifolii	[7]
Rhizobium sp. R200	trifolii	[5]
Rhizobium sp. RPVR32	phaseoli	[7]
Rhizobium sp. RPVR06	phaseoli	[7]

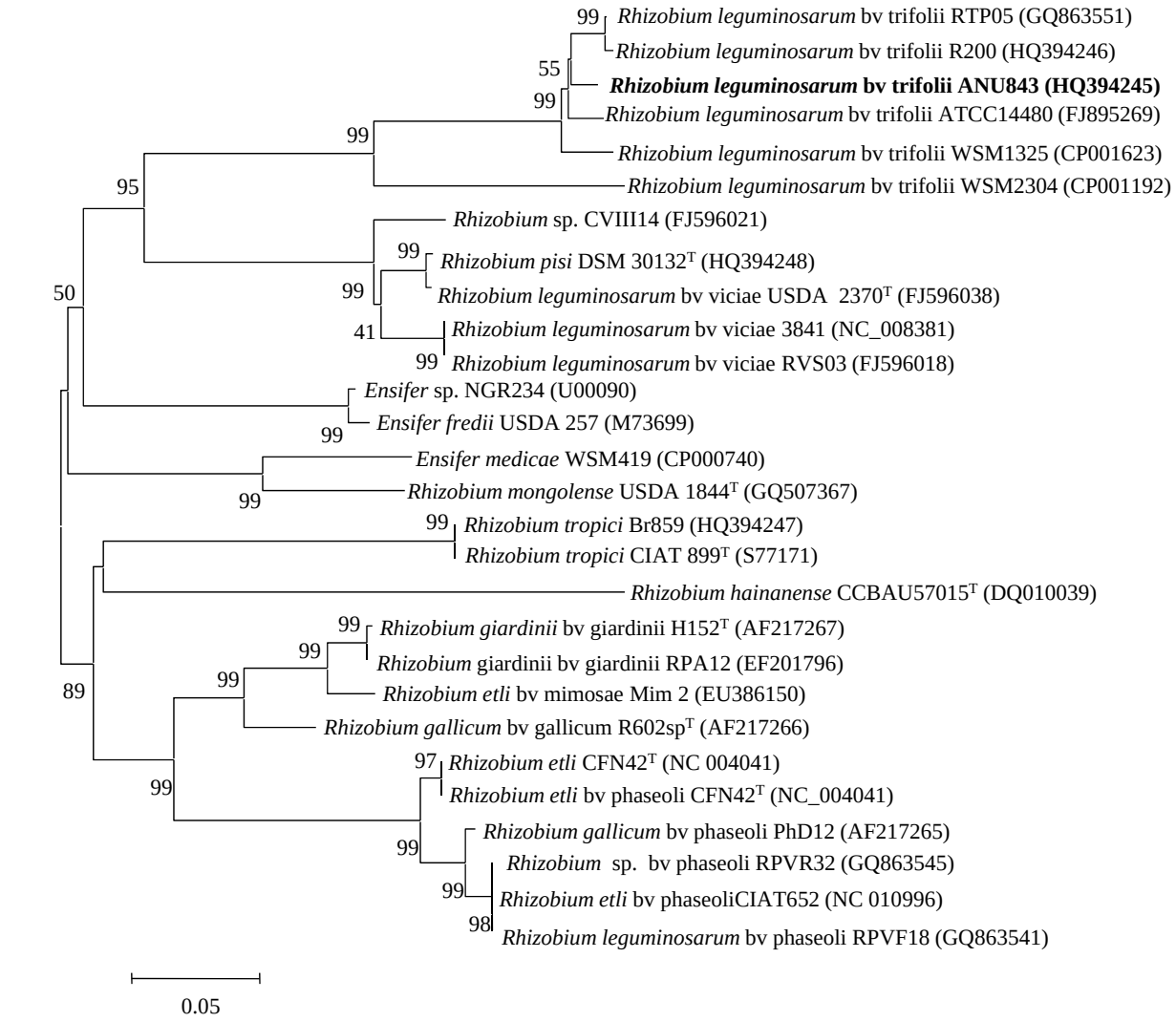

Fig. S3. Neighbour-joining phylogenetic tree based on *nodC* gene sequences (840 nt). Bootstrap values calculated for 1000 replications are indicated. Bar, 5 nt substitution per 100 nt. Accession numbers from Genbank are given in brackets.

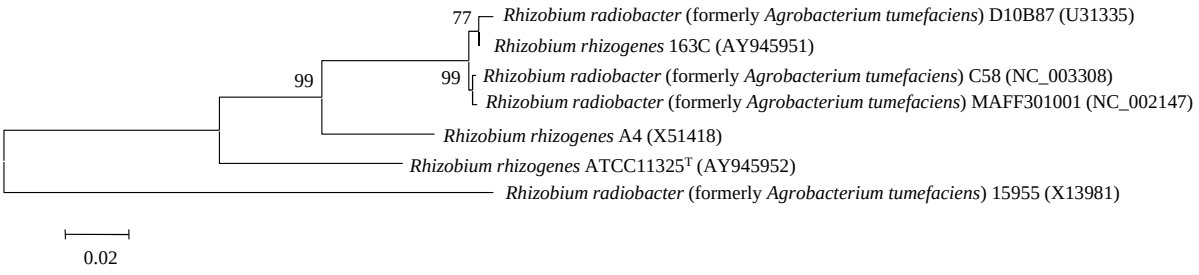

Fig. S4. Neighbour-joining phylogenetic tree based on *virA* gene sequences (950 nt). Bootstrap values calculated for 1000 replications are indicated. Bar, 2 nt substitution per 100 nt. Accession numbers from Genbank are given in brackets.

CHAPTER II. *Rhizobium* cellulase CelC2 is essential for primary symbiotic infection of legume host roots

M. Robledo*, J. I. Jiménez-Zurdo[†], E. Velázquez*, [1] M. E. Trujillo*, J. L. Zurdo-Piñeiro*, M. H. Ramírez-Bahena*, B. Ramos*, J. M. Díaz-Mínguez*, F. Dazzo[‡], E. Martínez-Molina*, and P. F. Mateos*[§]

[1]Departamento de Microbiología y Genética and Centro Hispano Luso de Investigaciones Agrarias, Universidad de Salamanca. Spain; [2]Grupo de Ecología Genética de la Rizosfera. Estación Experimental del Zaidín, CSIC, Granada, Spain; [3]Department of Microbiology and Molecular Genetics, Michigan State University, East Lansing, MI 48824. USA.

The rhizobia-legume, root-nodule symbiosis provides the most efficient source of biologically fixed ammonia fertilizer for agricultural crops. Its development involves pathways of specificity, infectivity and effectivity resulting from expressed traits of the bacterium and host plant. A key event of the infection process required for development of this root-nodule symbiosis is a highly localized, complete erosion of the plant cell wall through which the bacterial symbiont penetrates to establish a nitrogen-fixing, intracellular endosymbiotic state within the host. This process of wall degradation must be delicately balanced in order to avoiod lysis and destruction of the host cell. Here we describe the purification, biochemical characterization, molecular genetic analysis, biological activity, and symbiotic function of a cell-bound bacterial cellulase (CelC2) enzyme from *Rhizobium leguminosarum* bv. trifolii, the clover-nodulating endosymbiont. The purified enzyme can erode the non-crystalline tip of the white clover host root hair wall, making a localized hole of sufficient size to allow wild-type microsymbiont penetration. This CelC2 enzyme is not active on root hairs of the non-host legume alfalfa. Microscopy analysis of the symbiotic phenotypes of the ANU843 wild-type and CelC2 knockout mutant derivative revealed that this enzyme fulfils an essential role in the primary infection process required for development of the canonical nitrogen-fixing *R. leguminosarum* bv. trifolii-white clover symbiosis.

A central event in development of the *Rhizobium*-legume root-nodule symbiosis is the localized erosion of a cellulosic plant wall through which the bacterial symbiont passes to establish a nitrogen-fixing, intracellular endosymbiotic state within its legume host. Plant cell wall-degrading enzymes are predicted to participate in two steps of this infection process: during primary infection of host root hairs leading to infection thread formation (Inf), and later during bacterial release (Bar) from infection threads within host nodule cells. This process of plant cell wall degradation must be delicately balanced in order to allow the localized penetration of the bacterial symbiont into the host cell without its overt lysis and destruction.

Several studies indicate that rhizobia produce enzymes capable of degrading plant cell wall polymers (1-14), but little is known about their molecular properties, none have previously been purified to homogeneity, and their specific role (if any) in symbiosis is undefined. The relatively low activities of these rhizobial enzymes have hampered research progress in this area. Using improved assays with increased sensitivity that reliably detect these enzyme activities, we established that cellulases are produced by wild type strains of *Rhizobium leguminosarum* (biovars trifolii, phaseoli, and viciae), *Bradyrhizobium japonicum, Mesorhizobium loti, and Sinorhizobium meliloti* (7, 9). Further studies using *R. leguminosarum* bv. trifolii ANU843 indicated that this model wild-type strain of clover-nodulating rhizobia produces at least two cellulase isozymes, C1 and C2 (9). Several lines of evidence support the interpretation that both enzymes are cell-bound: i) their activity is detected in extracts following lysozyme-EDTA treatment or brief sonication of washed and pelleted cells, but not detected in the extracellular culture medium (7, 9), ii) both scanning and transmission electron microscopy reveal eroded concave pits that replicate the *in situ* contour of individual bacterial cells attached to root epidermal cell walls of white clover (10), and iii) data reported here show that the amino acid sequence encoded by the *celC2* gene in ANU843 contains a leader signal peptide with several distinctive features used in the general Sec-dependent export pathway of cell envelope proteins. Both C1 and C2 isozymes are made constitutively by ANU843 when grown in defined B-INOS medium, and their approximate molecular masses are 41.5 and 33.2 KDa, respectively (9). Studies using the symbiotic plasmid (pSym)-cured and *nod*-recombinant derivatives of ANU843 indicated that gene(s) required for production of cellulase C1 are located in pSym but outside its 14-Kb *Hin*dIII *nod* region, whereas the locus for the cellulase C2 is not pSym-borne (6).

Recently, a combination of phase contrast/polarized light microscopy and enzymology indicated that these cell-bound enzymes from *R. leguminosarum* bv. trifolii can completely erode through the white clover root hair wall at a highly localized site on the isotropic, noncrystalline apex of its root hair tip (expression of the Hot [Hole on the tip] phenotype), and can more extensively degrade isolated clover root hair walls as a substrate if they were grown in the presence (rather than the absence) of purified chitolipooligosaccharide Nod factors from clover rhizobia (10). Other studies showed that these Nod factors from clover rhizobia evoked a localized disruption in the normal crystallization of the host cell wall architecture in growing root hairs (15). Considered collectively, the results implicate a complimentary role of *Rhizobium* cellulase and Nod factors in promoting root hair infectibility at strategic sites during primary host infection. In the present study, we have purified one of the *R. leguminosarum* bv. trifolii cell-bound cellulase isozymes (CelC2) to homogeneity, and analyzed its symbiotic function by a combination of biochemical, reverse genetics and plant microscopy approaches. The results provide compelling evidence that this enzyme fulfils a very significant role in the primary infection process required for development of the canonical nitrogen-fixing, *R. leguminosarum* bv. trifolii-white clover symbiosis.

Results and Discussion

Cellulases in rhizobia. Currently, the known diversity of bacteria capable of forming a nitrogen-fixing root-nodule symbiosis with legumes includes 10 genera and more than 50 species. All of the official type strains of each of these taxa were examined for cellulase activity, and all were found to be positive to varying degree (Table 1 Supporting Information). Evaluation of the currently defined genomes of rhizobia indicated genes coding for putative cellulases represented by a diversity of glycosyl hydrolase families (Table 2 Supporting Information). Interestingly, these endoglucanase-coding genes found in rhizobial genomes are located near putative cellulose synthase genes.

Purification and biochemical characterization of cellulase CelC2 from wild type *R. leguminosarum* bv. trifolii ANU843. The two cellulase isozymes in the crude extract of ANU843 sonicated cells were resolved by activity gel electrophoresis and were named CelC1 and CelC2 (Fig. S1, lane 6). Anionic exchange chromatography using DEAE Sepharose CL6B separated these two cellulase isozymes, with CelC1 eluting in the peak fraction at 0.1 M NaCl and CelC2 in the peak fraction at 0.2 M NaCl (57.0 U; 0,7 U/mg). Gel filtration chromatography of the second DEAE peak fraction through Sephacryl S-100 HR resolved one peak of cellulolytic activity (24.0 U; 4.1 U/mg) with an apparent molecular mass of 33.2 kDa, consistent with the results of the activity gel electrophoresis step indicating that the second DEAE peak contained only one cellulase isozyme (CelC2). Ion exchange FPLC of this S-100 fraction resolved one peak of cellulolytic activity (6.7 U; 41.9 U/mg) that eluted at 0.2 M NaCl. SDS-PAGE and activity gel electrophoresis of this Mono Q fraction revealed four different proteins, one of which had cellulolytic activity with the same mobility as a silver-stained protein band with an apparent molecular mass of 33.2 kDa (Fig. S1, lane 4). Dialysis of this Mono Q fraction followed by Hydrophobic interaction FPLC using a Phenyl Superose column separated four protein peaks, one of which eluted at 0.26 M ammonium sulphate and contained all the cellulolytic activity (5.0 U; 152 U/mg) of the applied sample (Fig. S1, lane 7).

Final purification of the CelC2 cellulase isozyme to homogeneity was indicated by a single peak of protein with constant cellulase specific activity using gel filtration through Sephacryl S-200, and a single band of silver-staining protein in SDS-PAGE that matched the electrophoretic mobility of cellulase CelC2 detected by activity gel electrophoresis of the final purified enzyme (Fig. S1, lanes 5 and 7). These results indicate that active, purified CelC2 cellulase from *R. leguminosarum* bv. trifolii ANU843 consists of a single peptide with molecular mass of 33.2 kDa. Its hydrolytic enzyme activity using carboxymethyl cellulose (CMC) as substrate has an optimal pH of 5.0, an optimal temperature of 40 ºC, a specific activity of 152 U/mg protein and an apparent Km of 84.4 mg/ml.

Purified CelC2 cellulase hydrolyzed CMC with the release of reducing sugar equivalents, but under the same conditions, this purified enzyme did not release reducing sugars from Avicel microcrystalline cellulose, cellobiose, polygalacturonate, locust bean gum, xylan, *R. leguminosarum* bv. trifolii acidic heteropolysaccharide EPS types I, II or III, or *S. meliloti* succinoglycan exopolysaccharide. These data indicate that purified CelC2 cellulase is a 1,4-ß-D-endoglucanase with high substrate specificity for non-crystalline cellulose, and unlike *R. leguminosarum* PlyA and PlyB glycanases (3, 16), CelC2 does not degrade any of the three different chemotypes of extracellular acidic heteropolysaccharides made by *R. leguminosarum* bv. trifolii. Also, the substrate specificity of ANU843 CelC2 does not extend to the ß 1,4 heptaglucose backbone in the repeated octasaccharide of *S. meliloti* succinoglycan.

Molecular analysis of CelC2. An Edman degradation of purified ANU843 CelC2 cellulase revealed that its N-terminus is blocked. An internal peptide isolated by RP-HPLC from a tryptic digestion of purified CelC2 contained 14 amino acids with a sequence of Ala-Glu-Gly-Phe-Asp-Ala-Glu-Phe-Gly-Tyr-Asn-Ala-Ile-Arg, which has 85% identity to a sequence encoded by the *celC* gene of *R. leguminosarum* bv. trifolii R200 (17). The *celC*-encoded protein has sequence homology to endoglucanase CelC from *Agrobacterium tumefaciens* (18), CMCax from *Gluconoacetobacter xylinus* (19), CelY from *Erwinia chrysanthemi* (20), Cuda from *Cellulomonas uda* (21) and BcsZ from *E. coli* (22), all belonging to the glycosyl hydrolase family 8 (23).

Identification and cloning of the CelC2 encoding gene. The primers CelCexF, and C2R were used for PCR amplification of *celC2* using DNA from strain ANU843 as template. The amplified DNA fragment contained a 1,044 bp ORF (GeneBank accession number AJ561043) encoding a predicted protein of 347 amino acids with a Sec-dependent (General Secretory Pathway) signal peptide of 23 amino acids (Fig.S2). The amino acid sequence of the putative *celC2*-encoded protein contains the sequence of the tryptic peptide isolated from the purified enzyme, confirming that the amplified sequence matched that of our target gene (Fig. S2 Supporting Information). This *celC2* gene from *R. leguminosarum* bv. trifolii ANU843 has 88% identity to *cel8A* from *R. leguminosarum* bv. trifolii 1536 (GeneBank AY227046, 24), 86% to *celC* from *R. leguminosarum* bv. trifolii R200 (GeneBank AF121340, 17), 86% to *celY* from *R. leguminosarum* bv. *viciae* 3841 (GeneBank AM236080, 25), 85% to *celC* from *R. etli* CFN42 (GeneBank CP000133, 26) and 69% to *Smed_5210* from *Sinorhizobium medicae* WSM419 (GeneBank CP000740, unpublished).

Linkage of genes encoding cellulase and cellulose biosynthesis functions in rhizobia. Orthologs to the *celC* endoglucanase-coding gene are found in a diversity of eubacteria that make cellulose (*Rhizobium*, *Agrobacterium*, *Gluconoacetobacter*, *E. coli, Salmonella* and others). In all these cellulose-producing eubacteria the *celC* genes are located near putative cellulose synthase genes localized in a region of the chromosome (*celABC* or *bcsABZ*) involved in bacterial cellulose biosynthesis (17, 18, 19). In the currently defined rhizobial genomes where the *celC* gene is present (*R. leguminosarum*, *R. etli* and *S. medicae*), this gene is located in this cellulose synthase operon. The *celC2* from *R. leguminosarum* bv. trifolii ANU843 is also located in this cellulose synthase operon (Fig. 1 Supporting information). The primary sequence and organization of the *celABC* genes are similar in *R. leguminosarum* bv. trifolii R200 (17) and *A. tumefaciens* (18). The postulated involvement of the *Agrobacterium* CelC cellulase enzyme in bacterial cellulose biosynthesis is to transfer new oligomer primers for chain elongation (27), but that biochemical function has not been definitively established. The lack of cellulose biosynthesis by mutation of *celA* or *celB* of *R. leguminosarum* bv. trifolii did not affect its ability to nodulate clover under controlled lab conditions (17). Nevertheless, it is feasible that the benefit to rhizobia gain by their cellulose microfibril-facilitated firm adhesion during colonization of host roots (28, 29) has significant symbiotic importance under natural conditions in rhizosphere soil.

Biological activity of purified cellulase CelC2 from *R. leguminosarum* bv. trifolii ANU843 on white clover seedling roots. Previous studies (10) showed that the DEAE fraction containing the partially purified cellulase CelC2 isozyme from wild-type ANU843 was capable of degrading the cell wall at the apex of the root hair tip (the "Hot" phenotype) *in vivo* when incubated with intact white clover host seedling roots, and this activity was not found in the

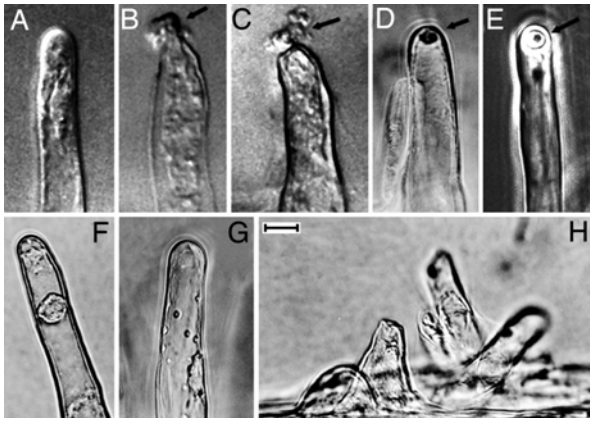

Fig. 1. Biological activity of purified cellulase CelC2 from *R. leguminosarum* bv trifolii ANU843 on clover root hairs. Nomarski interference contrast micrographs (A-C). (A) Typical control root hair without enzyme treatment. (B) Early stage in expression of the Hot phenotype. A portion of the root hair protoplast (arrow) has extruded through the hole made at the tip of the root hair wall by CelC2. (C) Further extrusion of a larger portion of the protoplast (arrow) through the hole made by CelC2 cellulase at a root hair tip. Phase contrast micrographs (D-H). (D, E) The refractile hole (arrow) made by CelC2 cellulase at the tip of a clover root hair after its cytoplasm has already been expelled into the external rooting medium. (F-G) Alfalfa root hairs remain intact in the untreated control (F) and following incubation with CelC2 from ANU843 (G). (H) Extensive degradation of clover root hairs after treatment with a commercial fungal cellulase. Bar scale is 12 μm.

other DEAE fraction containing the cellulase C1 isozyme (10). The symbiotic relevance of this enzyme bioactivity was revealed by studies indicating that (i) its highly localized site of cell wall degradation matched the restricted distribution of the isotropic non-crystalline wall architecture at root hair tips of axenic white clover seedlings that typically undergo marked curling and infection (10), (ii) the discrete hole through the plant wall that is ultimately produced there has a diameter precisely within the 2-3 μm range typically found at primary infection sites in deformed root hairs (10, 30, 31), (iii) it was more extensive in walls of root hairs that had been grown in the presence of wild-type chitolipooligosaccharide Nod factors that increase root hair infectibility (10), and (iv) it was less extensive when clover seedlings were grown with sufficient nitrate supply to inhibit primary infection (10, 32) or with root hair walls isolated from the non-host legume, alfalfa (15).

To test whether CelC2 cellulase exhibits this bioactivity, we performed this direct in situ assay using the fully purified isozyme from wild-type ANU843 on axenic seedling roots of its white clover host. As predicted, the results indicated that the purified CelC2 cellulase exhibits this same infection-related bioactivity as the previously described, partially purified DEAE fraction containing it. The series of images in Fig. 1 show a progression of the morphological features of the Hot phenotype produced in vivo on susceptible root hairs treated with purified CelC2 (Fig. 1, A-E). The enzyme acting at this localized site erodes a hole completely through the root hair wall, resulting in the extrusion of its protoplast / cytoplasm (Fig. 1B, C). The Hot phenotype occurred at an approximate frequency of 14 root hairs per cm of optical median plane on axenic white clover seedlings after 6 hours of enzyme incubation. No Hot activity was found on root hairs of clover seedlings incubated with the isotonic Fähraeus medium without-enzyme as a negative control (Fig. 1A), or when the purified CelC2 enzyme from ANU843 was incubated with axenic seedling roots of the non-host legume, alfalfa (Fig. 1G). As a positive control, incubation of both white clover and alfalfa seedlings with a commercial fungal cellulase from Trichoderma (Sigma) under the same experimental conditions resulted in extensive degradation of

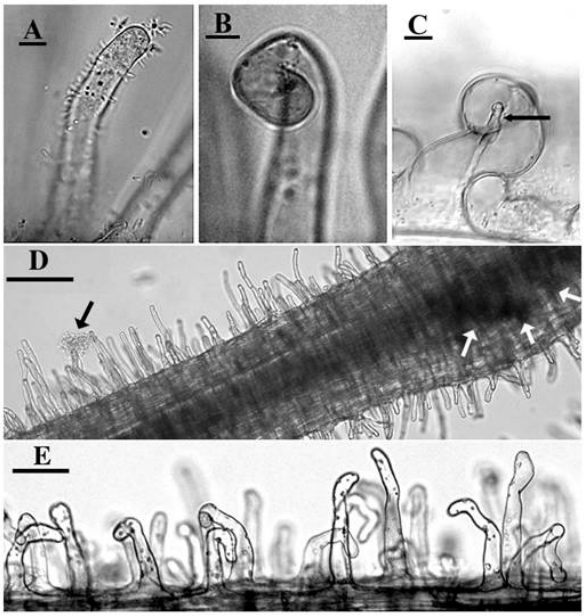

Fig. 2. Primary infection events in white clover inoculated with wild type ANU843 (A-C) and the CelC2 minus mutant Δ*celC2* (D, E). The phase contrast micrographs show the canonical steps in primary host infection of white clover by *R. leguminosarum* bv trifolii; root hair adhesion (A), markedly curled "shepherd crook" (B) and infection thread formation (C arrow). The CelC2 minus mutant *ΔcelC2* exhibits root hair adhesion (D black arrow) and induces root hair deformation (E) and nodule initiation (D white arrows), but no infection threads in white clover root hairs (D, E). Bar scales are 15 μm (A, B, C), 50 μm (E) and 100 μm (D).

root hairs (Fig. 1H, and data not shown). These results provide direct, in situ evidence clearly indicating that purified cellulase CelC2 from ANU843 can degrade the clover host plant cell wall precisely at the strategic location where the bacterial symbiont first penetrates (markedly curled) root hairs, and suggests that the specificity of that enzyme-catalyzed reaction may contribute to the restricted host range of root hair infection expressed in the *Rhizobium*-clover symbiosis.

Symbiotic phenotype of an ANU843 celC2 knockout mutant derivative strain. To further examine if the activities of CelC2 are infection-related, we compared the symbiotic phenotypes of wild type ANU843 to the C2-cellulase knockout mutant derivative ANU843ΔcelC2 obtained by an internal 361 bp BamHI-HindIII fragment deletion that encodes the GH8 glycosyl hydrolase catalytic center motif within the celC2 gene (Fig. S2 Supporting Information). The wild type strain accomplished the canonical steps in the infection process with white clover: root hair adhesion (Fig. 2A), various types of root hair deformation (Had, Fig. 2B), marked root hair curling (Hac, "shepherd crook") and infection thread formation (Inf, Fig. 2C). The ANU843 ΔcelC2 mutant retained the ability to adhere to the host root surface (Fig. 2D) and induced various types of moderate root hair deformation (Fig 3E), but did not evoke formation of complete, marked root hair "shepherd crooks" or complete infection threads when inoculated onto white clover seedlings (Fig. 2 and Table 3 SI).

Detailed examination by phase-contrast microscopy of the seedlings inoculated with the ANU843ΔcelC2 mutant indicated only two rare cases where a localized refractile bright spot developed on deformed root hairs. That morphological feature commonly precedes root hair penetration and infection thread formation (15). As is also the case for root hairs responsive to Nod factors (15), a

subset of root hairs is particularly susceptible to deformation after inoculation with either the wild-type or the CelC2-minus mutant derivative strain (Fig. 2E). The most susceptible root hairs were actively growing in the root hair zones I and II. Root hairs that have finished growing (root hair zone III) were refractory to deformation.

Both ANU843 and ΔcelC2 test strains induced nodule formation on white clover roots (Fig. 2D, Table 3 SI). The first nodules emerged 10 days after inoculation with either strain, and all inoculated plants were nodulated by the third week. The mean number of nodules per plant at 40 days after inoculation was higher with the mutant than with the wild type strain, but the difference in mean values was not statistically significant (Table 3 SI). Increased nodulation is a common feature for ineffective *Rhizobium*-legume symbioses (33).

The nodule organogenesis in the inner cortex (Noi phenotype) was similar in white clover plants inoculated with the wild type ANU843 or the CelC2 minus strain (Table 3SI and Fig. 2D). Emerged nodules induced by the wild type were invaded, typically red, piriform, and indeterminate (Fig. 3A, C). In contrast, emerged nodules induced by the mutant celC2 were small, white, spherical and uninvaded (Fig. 3B, D).

Plants grew well in the nitrogen-free medium when inoculated with the wild type strain ANU843, indicative of an effective nitrogen-fixing symbiosis. Those plants benefited by the symbiosis produced greener leaves and shoots that were significantly longer than plants inoculated with the mutant ΔcelC2 or the uninoculated axenic control plants (Table 3SI and Fig. 3E-G). The latter two groups of plants were stunted in development and became chlorotic once the fixed nitrogen from their cotyledons was exhausted, indicative of nitrogen starvation stress without nitrogen fixation under these N-free growth conditions. Complementation of the mutant strain ANU843ΔcelC2 with a plasmid expressing the celC2 gene restored the ability to produce the CelC2 cellulase isoenzyme, infect root hairs and induce efficient nitrogen-fixing nodules on white clover (Fig. 2 Supporting Information).

To the best of our knowledge, this is the first report on the biochemical characteristics and infection-related biological activity of a purified rhizobial enzyme capable of creating the hollow, localized portal of entry into the root hair cell wall through which the bacterial symbiont can penetrate during primary infection of its legume host. Our purification protocol yields a cell-bound ß-1,4-endoglucanase (EC 3.2.1.4) isozyme from the clover microsymbiont, *R. leguminosarum* bv. trifolii. This enzyme, called CelC2 cellulase, is particularly important for symbiotic development because this rhizobial symbiont requires its activity to breach the host barrier in order to establish an endosymbiotic, nitrogen-fixing association with its specific legume host.

Further studies reported here and elsewhere revealed multiple characteristics of CelC2 cellulase that predictively restrict (thereby tightly control) its symbiotically relevant activity during primary host infection. First, the cell-bound location (7, 9), relatively low quantity produced (rather than excretion of large quantities into the external environment), and high K_m relative to other microbial cellulases are all characteristics that would predictively restrain its degradative action on roots, thereby minimizing indiscriminate and overextended wall hydrolysis inevitably resulting in host cell lysis and death. Elevated levels of root hair wall peroxidase localized at this same site likely play an important complementary role in maintaining the viability of the host cell during incipient microsymbiont penetration (34). Second, these same characteristics would provide further opportunity to restrict its short-range action based on physical positioning of the bacterium at the host wall interface. Third, the substrate specificity of CelC2 cellulase for non-crystalline cellulose significantly restricts its *in vivo*

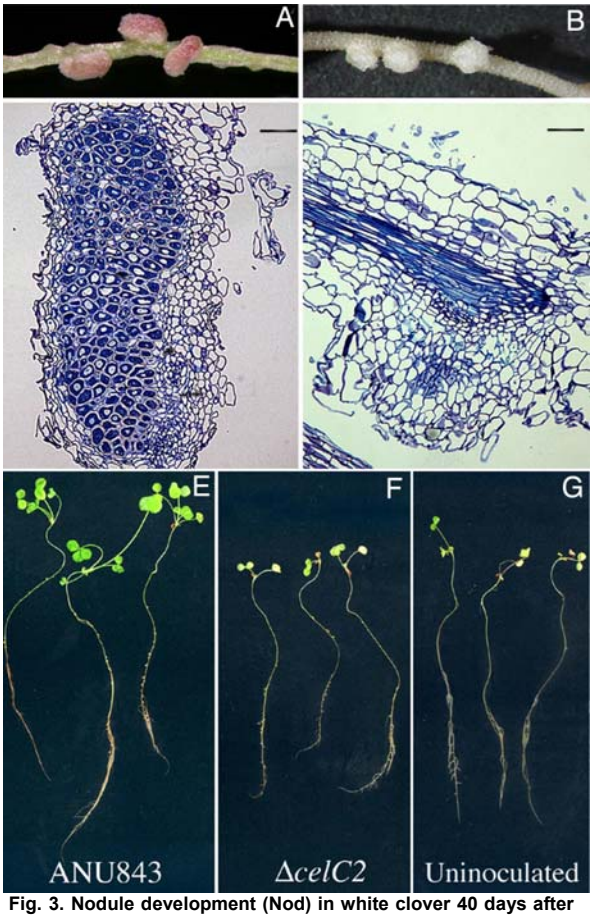

Fig. 3. Nodule development (Nod) in white clover 40 days after inoculation with the wild type ANU843 (A, C, E) or the CelC2 minus mutant Δ*celC2* (B, D, F). Shown are portions of nodulated roots from inoculated plants (A, B), longitudinal sections of the induced nodules (C, D), and whole plants phenotypes (E, F, G). Bar scales are 200 μm.

site of wall-penetrating erosive action to the highly localized root hair infection sites lacking crystalline wall architecture. This finding provides an explanation of the remarkable infection event that rhizobia typically only erode and penetrate one side of the root hair wall despite being sandwiched by both sides when cradled in the overlap of the markedly curled shepherd's crook (10, 15, 29, 30, 35, 36).Development of these localized infection sites is strategically modulated by the action of the rhizobial chitolipooligosaccharide Nod factors that disrupt the crystallization of the root hair wall architecture and activate symbiotic root hair infection (15). Fourth, the diameter of the hole eroded completely through the root hair wall by CelC2 cellulase action is precisely within the same *in situ* size range that wild type *R. leguminosarum* bv trifolii cells use to traverse the clover root hair wall during primary host infection (10, 30, 31). Fifth, the pH 5 optimum for CelC2 cellulase is compatible with the slightly acidic pH measured *in vivo* at the external surface of white clover root hairs (J. Salzwedel and F. Dazzo, unpublished data). All these biochemical and cytological properties of CelC2 cellulase are typical of what one would expect for a *Rhizobium* plant wall-degrading enzyme that plays an important contributing role in creating the unique portal of microsymbiont entry during early steps of primary host infection required for development of the endosymbiotic nitrogen-fixing *Rhizobium*-white clover symbiosis. Finally, CelC2 knockout mutants were unable to breach the host wall at the root hair tip and form infection threads that are necessary to invade nodules allowing them to fix nitrogen, and transfer of the cloned wild-type gene into the CelC2 knockout mutant restored these symbiotic phenotypes, providing further

compelling evidence for the requirement of this enzyme in successful development of the canonical *Rhizobium*-white clover symbiosis.

Materials and Methods

Bacterial strains and growth conditions. Wild-type *Rhizobium leguminosarum* bv. trifolii strain ANU843 and its derivative ANU843Δ*celC2* were grown in defined B-INOS (9) or yeast mannitol agar (YMA) media (37) at 28 ºC. These media were supplemented with kanamycin (50 µg/ml) for growth of the *celC2* complemented strain.

DNA methods and construction of ANU843 derivative strains. DNA was extracted according to Rivas et al. (38). Plasmid DNA was obtained using a High Pure Plasmid Isolation Kit (Boehringer Mannheim) by following manufacturer's instructions. Preparation of plasmid DNA, restriction enzyme digestions, DNA cloning, PCR DNA amplification, and agarose DNA electrophoresis were performed using standard procedures (39).

To generate the Δ*celC2* deletion mutant derivative, CelCexF (TCGCCGCCAACTGGCTGTC) and C2R (CACAGACACTCC-GGATGC) primers were used to amplify *celC2* from ANU843 DNA. The PCR product was cloned into pGEM-T Easy vector (Promega) and a 361 bp *Bam*HI-*Hind*III fragment from *celC2* was removed. The truncated *celC2* gene was cloned as a *Pvu*II fragment into plasmid pK18*mobsacB* (40) and transferred to *R. leguminosarum* bv. trifolii ANU843 by triparental mating using pRK2013 as helper plasmid (41). Allele replacement was selected as described previously (40). PCR analysis with primers C1F (ATCAGCCACAGCGAAGGGCA) and C2R was used to verify the 361 bp deletion in the mutant strain (Fig. 3A Supporting Information). The lack of CelC2 activity in this deletion mutant was confirmed by a zymogram (Fig. 3B Supporting Information).

To generate the complemented strain, an ANU843 genomic region containing the CelC2 coding sequence and 78 nt upstream of the start codon was amplified as a *Sal*I fragment and inserted between the *Xho*I and *Sal*I sites into plasmid pBBRMS-2 (42) to yield pJZC2, which was transferred to *R. leguminosarum* bv. trifolii ANU843Δ*celC2* by conjugation as described.

Enzyme assays. CMCase activities were detected *in vitro* by the double-layer plate enzyme assay using Congo Red staining (9), reducing sugar assays using 2,2'-bicinchonic acid (BCA) (43), and the activity stain overlay technique using SDS-PAGE followed by renaturation (9). One unit of enzyme activity is defined as the amount releasing 1 nmol of reducing sugar equivalent per min at 40 ºC. One unit of specific activity was defined as 1 nmol reducing sugar equivalent released min^{-1} (mg protein)$^{-1}$. The amount of protein was measured by the dye-binding method of Bradford (44) with bovine serum albumin as the standard.

Enzyme purification. All purification steps were performed at 4 ºC and protein profiles in chromatographic fractions were analysed by SDS-PAGE followed by silver staining (45). Broth cultures (12 l) of wild-type ANU843 in the early stationary phase (9 x 10^8 cells per ml) were centrifuged at 20,000 x *g* for 15 min. Pelleted cells were washed by centrifugation with water, suspended in 30 ml of 10 mM Tris/HCl buffer (pH 8.0) on ice, sonicated in five cycles of 1-min bursts at 32 W with a microprobe, and centrifuged at 12,000 x *g* for 20 min. This sonicated "crude extract" was applied to a DEAE-Sepharose CL-6B column (18 x 3 cm) previously equilibrated with 10 mM Tris/HCl buffer. Unbound protein was washed through the column with 10 mM Tris/HCl buffer. Bound protein was eluted in a linear gradient of 0 to 0.4 M NaCl in 10 mM Tris/HCl buffer at a flow rate of 1.75 ml/min. Fractions (2.5 ml) containing cellulase activity were pooled and concentrated by ultrafiltration using an Amicon YM 30 membrane, which does not bind these cellulases. This ultrafiltration-concentrated DEAE fraction was chromatographed through a Sephacryl S-100 HR column (95 x 2 cm) using 100 mM Tris/HCl buffer (pH 8.0, 0.25 ml/min) to avoid non-specific binding of rhizobial cellulases. The S-100 fractions (2.5 ml) containing cellulase activity were pooled and applied to a Pharmacia Mono Q HR 5/5 column. Unbound protein was eluted with 10 mM Tris/HCl buffer (pH 8). The bound cellulase activity was eluted by a linear gradient of 0 to 0.4 M NaCl in 10 mM Tris/HCl buffer at 0.25 ml/min. Mono Q-fractions containing cellulase activity were pooled and mixed with an equal volume of 3.4 M ammonium sulfate in PCA buffer (pH 5.0) to yield a final concentration of 1.7 M ammonium sulfate. This mixture was applied to a Pharmacia Phenyl-Superose HR 5/5 column equilibrated with 1.7 M ammonium sulfate in 10 mM PCA buffer (pH 5.0, 0.5 ml/min, 0.5 ml fractions). Bound proteins ($A_{280\ nm}$) were eluted with a linear gradient of 1.7 M to 0 M ammonium sulfate in 10 mM PCA buffer. Collected fractions (0.5 ml) containing cellulase activity were concentrated using Centricon YM-10 ultrafilters and stored at 4 ºC.

Biochemical characterization of purified CelC2 Cellulase. 1,4-ß-D-endoglucanase activity assays were run under standard conditions (9) to determine the optimum pH over the range of 4.0 to 7.0 (PCA 200mM) and 8.0 to 9.0 (Tris-HCl 200 mM), and the influence of temperature on enzyme activity in the range of 4 ºC to 70 ºC. The molecular mass of the purified native enzyme was analyzed by gel filtration chromatography using a Sephacryl S-200 column (100 x 2.5 cm, 2.5 ml fractions) in 100 mM Tris/HCl buffer. The protein content of collected fractions were estimated by $A_{280\ nm}$ and assayed for enzyme activity (9). The molecular mass was also estimated by activity gel electrophoresis (9). To examine substrate specificity, potential substrates of Avicel crystalline cellulose, cellobiose, Na polygalacturonate, locust bean gum, xylan, *R. leguminosarum* bv. trifolii EPS types I, II and III (46) and *S. meliloti* succinoglycan (each 1 mg/ml in PCA buffer) were compared to CMC under standard assay conditions (9). When necessary, the remaining insoluble substrates were removed by centrifugation (10,000 x *g*, 15 min) at 4 ºC before analyzing the supernatant for depolymerized products. To determine the K_m for cellulase CelC2, its rate of CMC hydrolysis was measured at substrate concentrations ranging from 0.1 to 1 % (w/v) at pH 5 and 40 ºC, followed by graphing its kinetics in a plot of V vs. S and computing its S at $0.5V_{max}$. For sequencing analysis, tryptic peptides of the purified enzyme were isolated using a RP-HPLC column. The automated Applied Biosystems model 477A pulse-liquid sequenator equipment used to perform the Edman degradation was connected on-line to an RP-HPLC unit for identification of the step-wise released PTH-aminoacids.

Biological activity assay. Purified CelC2 cellulase (40 µl, 0.2 U) from strain ANU843 was assayed *in vivo* for its ability to evoke the Hot (Hole on the tip) phenotype (10) on root hairs of intact axenic seedlings of white clover (*Trifolium repens* L. var. Dutch) and alfalfa (*Medicago sativa* var. Aragón). Control treatments included nitrogen-free Fahraeus medium (44) without enzyme, and commercial fungal cellulase (0.2 U) from *Trichoderma* (Sigma) in the same medium. Periodically, root hairs were examined individually for localized cell wall degradation by phase-contrast and Nomarski interference light microscopy at 400 X using a 546-nm narrow band-pass interference contrast filter to enhance optical refractility by minimizing chromatic aberration.

Analysis of symbiotic phenotypes. Surface-sterilized white clover seeds were germinated for 2 days in humid air on inverted plates of water agar incubated at room temperature in the dark. Five seedlings with straight roots ca. 10 mm in length were transferred to the surface of 2% agar plates containing nitrogen-free Fähraeus medium (35) and overlaided with a 9 cm diameter disk of Whatman No. 1 sterile filter paper. One ml of a suspension (OD_{600nm} of 0.5) of 5 day-old *Rhizobium* culture that had been grown on YMA was deposited on the overlying filter paper. Inoculated seedlings were incubated at room temperature in the dark for 5 days and then examined by light microscopy to assess root hair deformation and infection.

For nodulation phenotypes, axenic seedlings were transferred to a Whatman No. 1 filter paper support (15 x 1.5 cm) placed inside test tubes (one plant per tube) containing 20 ml of nitrogen-free Fähraeus medium. Plants were grown at room temperature in the dark for 5 days, then inoculated (20 per test strain) with 1 ml of a suspension (OD_{600nm} of 0.5) of a 5 day-old *Rhizobium* culture grown on YMA, and transferred to a growth chamber with mixed incandescent and fluorescent illumination (400 microeinsteins m^{-2} s^{-1}; 400 to 700 nm) programmed for a 16 h photoperiod, 24 ºC: 20 ºC day-night cycle, and 50 to 60% relative humidity. Roots were examined by stereomicroscopy every 48 h over 40 consecutive days. Then roots and nodules were excised and processed for brightfield light microscopy (47).

Acknowledgements. This work was supported by Ministerio de Educación y Ciencia grant AGL 2005-07796 from Spain and by the Michigan Agricultural Experiment Station, USA. M.R. was supported by a fellowship from Ministerio de Educación y Ciencia, Spain. We dedicate this paper in memory of Professor Antonio Palomares, former president of the Spanish Society for Nitrogen Fixation (Sociedad Española de Fijación de Nitrógeno, SEFIN).

References

1. Angle J-S (1986) *Soil Biol Biochem* 18:115-116.
2. Chalifour F-P, Benhamou, N (1989) *Can J Microbiol* 35:821-829.
3. Finnie C, Zorregueita A, Hartley N-M, Downie J-A (1998) *J Bacteriol* 180:1691-1699.
4. Hubbell D-H, Morales V, Umali-Garcia M (1978) *Appl Environ Microbiol* 35:210-224.
5. Iannetta P, McMillan G, Sprent J-I (1997) *Soil Biol Biochem* 29:1019-1021.
6. Jiménez-Zurdo J-I, Mateos P-F, Dazzo F, Martínez-Molina E (1996) *Soil Biol Biochem* 28:131-133.
7. Jiménez-Zurdo J-I, Mateos P-F, Dazzo F, Martínez-Molina E (1996) *Soil Biol Biochem* 28:917-921.
8. Martínez-Molina E, Morales V, Hubbell D (1979) *Appl Environ Microbiol* 38:1186-1188.
9. Mateos P-F, Jiménez-Zurdo J-I, Chen J, Squartini A, Haack S, Martínez-Molina E, Hubbell D-H, Dazzo F (1992) *Appl Environ Microbiol* 58:1816-1822.
10. Mateos P-F, Baker D, Petersen M, Velázquez E, Jiménez-Zurdo J-I, Martínez-Molina E, Squartini A, Orgambide G, Hubbell D, Dazzo F-B (2001) *Can J Microbiol* 47:475-487.
11. Michaud P, Belaich A, Courtois B, Courtois J (2002) *Appl Microbiol Biotechnol* 58:767-771.
12. Morales V, Martínez-Molina E, Hubbell D (1984) *Plant Soil* 80:407-415.
13. Plazinski J, Rolfe B-G (1985) *J Plant Physiol* 120:181-187.
14. Verma D-P, Zogbi V, Bal A-K (1978) *Plant Science Lett* 13:137-142.
15. Dazzo F-B, Orgambide G-G, Philip-Hollingsworth S, Hollingsworth R-I, Ninke K, Salzwedel J (1996) *J Bacteriol* 178:3621-3627.
16. Zorreguieta A, Finnie C, Downie, JA (2000) J Bacteriol 182:1304-1312.
17. Ausmees N, Jonsson H, Höglund S, Ljunggren H, Lindberg M (1999) *Microbiology* 145:1253-1262.
18. Matthyssee A-G, White S, Lightfoot R (1995) *J Bacteriol* 177:1069-1075.
19. Wong H-C, Fear A-L, Calhoon R-D, Eichinger G-H, Mayer R, Amikan D, Benziman M, Gelfand D-H, Meade J-H, Emerick A-W, *et al*. (1990) *Proc Natl Acad Sci USA* 87:8130-8134.
20. Guiseppi A, Aymeric J-L, Cami B, Barras F, Creuzet N (1991) *Gene* 106:109-114.
21. Nakamura K, Misawa N, Kitamura K (1986) *J Biotechnol* 4:247-254.
22. Park Y-W, Yun H-D (1999) *Mol Gen Genet* 261:236-241.
23. Coutinho P-M, Henrissat B (1999) In *Genetics, Biochemistry and Ecology of Cellulose Degradation*, eds Ohmiya K, Hayashi K, Sakka K, Kobayashi Y, Karita S, Kimura T (Uni Publishers Co., Tokyo), pp 15-23.
24. An J-M, Lim W-J, Hong S-Y, Shin E-C, Kim E-J, Kim Y-K, Park S-R, Yun H-D (2004) *Lett Appl Microbiol* 38:296-300.
25. Young J-W, Crossman L-C, Johnston A-W, Thomson N-R, Ghazoui Z-F, Hull K-H, Wexler M, Curson A-R, Todd J-D, Poole P-S *et al*. (2006) *Genome Biol* 7:R4-R4.
26. González V, Santamaría R-I, Bustos P, Hernández-González I, Medrano-Soto A, Moreno-Hagelsieb G, Janga S-C, Ramírez M-A, Jiménez-Jacinto V, Collado-Vides J, Davila G (2006) *Proc Natl Acad Sci USA* 103:3834-3839.
27. Matthyssee A-G, Thomas D-L, White A-R (1995) *J Bacteriol* 177:1076-1081.
28. Dazzo F-B, Truchet G, Sherwood J, Hrabak E, Abe M, Pankratz H-S (1984) *Appl Environ Microbiol* 48:1140-1150.
29. Mateos P-F, Baker D-L, Philip-Hollingsworth S, Squartini A, Peruffo D-B, Nuti M-P, Dazzo F-B (1995) *Can J Microbiol* 41:202-207.
30. Callaham D-A, Torrey J-G (1981) *Can J Bot* 59:1647-1664.
31. Napoli C, Hubbell D-H (1975) *Appl Microbiol* 30:1003-1009.
32. Dazzo F-B, Brill W J (1978) Plant Physiol 62(1): 18-21.
33. Fred E-B, Baldwin I, McCoy E (1932) The root nodule bacteria, Univ. of Wisconsin Press, Madison, WI USA.
34. Salzwedel J, Dazzo F (1993) *Molec. Plant-Microbe Interact* 6: 127-134.
35. Fähraeus G (1957) *J Gen Microbiol* 14:374-381.
36. Turgeon B, Bauer W (1985) *Planta* 163: 328-349.
37. Somasegaran P, Hoben H-J (1994) *Handbook for rhizobia: Methods in legume-Rhizobium technology* (Springer-Verlag, Heidelberg).
38. Rivas R, Velázquez E, Valverde A, Mateos P, Martínez-Molina E (2001) *Electrophoresis* 22:1086-1089.
39. Sambrook J, Fritsch E-F, Maniatis T (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor).
40. Schaefer A, Tauch A, Jaeger W, Kalinowski J, Thierbach G, Pühler A (1994) *Gene* 145:69-73.
41. Ditta G, Stanfield S, Corbin D, Helinski D-R (1980) *Proc Natl Acad Sci USA* 77:7347-7351.
42. Kovach, M. E., Phillips, R. W., Elzer, P. H., Roop, R. M., 2nd & Peterson, K. M. (1994) *Biotechniques* 16, 800-802.
43. Waffenschmidt S, Jaenicke L (1987) *Anal Biochem* 165:337-340.
44. Bradford M-M (1976) *Anal Biochem* 72:248-254.
45. Morrisey J (1981) *Anal Biochem* 117:307-310.
46. Philip-Hollingsworth S, Hollingsworth R, Dazzo F-B (1989) *J Biol Chem* 264:542-545.
47. Velázquez E, Mateos P, Pedrero P, Dazzo F, Martínez-Molina E (1995) *Appl Environ Microbiol* 61: 2033-2036.

Suplemental Information

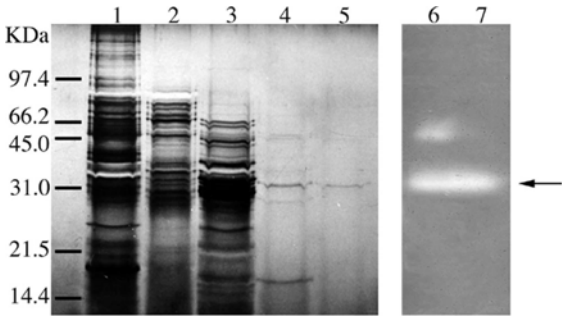

Fig. 1SI. SDS-PAGE of fractions at various steps in the purification of CelC2 cellulase made by *R. leguminosarum* bv. trifolii ANU843. Silver staining (1-5) and zymogram (6, 7). Lanes: 1 and 6, sonicated extract; 2, DEAE Sepharose; 3, Sephacryl S-100; 4, Mono Q; 5 and 7, Phenyl Superose. The molecular masses of protein standards are indicated on the left. The position corresponding to CelC2 cellulase is indicated with an arrow.

TNILSYASLLVVFVLLIGVTTSTMLKKLGRSMRRWRALLLAASVAVAPGLPATA QQAMINADAWSAYKAKFLDPSGRIVDNGNGNISHSEGQGYGLLLAYLSASPAD FEQIWYFTRTELLLRDDGLAVWKWDPNVKPHVADTNNATDGDMLIAYALALA GTAWKREDYILAASRMAQALLAETVGSSQGRTLLMPGTEGFTGSDRDDGPVV NPSYWIYEAIPVMAALAPSDAWKKLSDDGVELLKTMQFGPRKLPAEWVSLHD KPRPAEGFDAEFGYNAIRIPLYLARGGITDKALLVRLQKGMSQDGVPATIDLTT GRPKTVLSDPGYRIVNDVVACVVDGTRLPSSALQFAPALYYPSTLQLLGLAYIG EKHPECL

Fig. 2SI. *In silico* amino acid sequence analysis of translated product obtained by DNA amplification of *Rhizobium leguminosarum* bv. trifolii ANU843 using CelCexF and C2R primers. Brown: CelB C-terminus; Blue: CelC2 Sec leader signal peptide; Red: GH8 glycosyl hydrolase catalytic center; Green: BamHI - HindIII deletion fragment; Underlined: the internal sequence that is identical to the isolated internal tryptic peptide obtained from the purified enzyme.More comments on the CelC2 signal peptide: this 23 aa peptide has the three distinct regions of the signal peptide used in the general secretory (Sec) pathway in Gram negative bacteria. These include a positively charged N-terminal region (n-region), a hydrophobic α helical region (h-region) and a c-domain that contains the site of cleavage by the signal peptidase. The signal peptides for the Tat secretory pathway have a similar tripartite organization, but they also have a conserved sequence motif (S/T)-R-R-x-F-L-K containing the invariantly consecutive arginine residues (R) not found in CelC2.

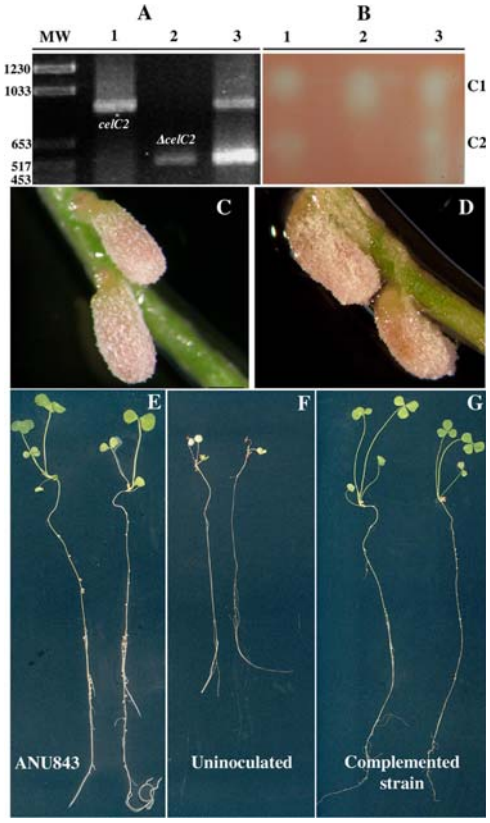

Fig. 3SI. PCR amplification of DNA using primers C1F/C2R (A) and zymogram of CM-cellulase isozymes in sonicated cell extracts (B) from *Rhizobium legumuminosarum* bv trifolii ANU843 wt (lanes 1), ANU843Δ*celC2* (lanes 2) and *celC2* complemented strain (lanes 3). PCR products were analyzed by electrophoresis on a 1% agarose gel. MW: DNA molecular weight marker, fragment lengths in bp are indicated. Typical nodules and whole phenotypes of white clover plants 40 days after inoculation with the wild type ANU843 (C, E) or the *celC2* complemented strain (D, G). Uninoculated plants (F).

Table 1SI. Cellulase production by representative type strains of root-nodule legume symbionts.

GENUS	TYPE STRAIN	NODULATED	CELLULASE
Rhizobium	*R. etli* CFN 42^T	*Phaseolus*	+
	R. galegae ATCC 43677^T	*Galega*	+
	R. gallicum R602sp^T	*Phaseolus*	+
	R. giardinii H152^T	*Phaseolus*	+
	R. hainanense I66^T	*Desmodium*	+
	R. huautlense SO2^T	*Sesbania*	+
	R. indigoferae CCBAU 71042^T	*Indigofera*	+
	R. leguminosarum ATCC10004^T	*Pisum*	+
	R. loessense CCBAU 7190B^T	*Astragalus*	+
	R. lusitanum P1-7^T	*Phaseolus*	+
	R. mongolense USDA 1844^T	*Medicago*	+
	R. sullae IS123^T	*Hedysarum*	+
	R. tropici CIAT 899^T	*Phaseolus*	+
	R. yanglingense CCBAU 71623	*Amphicarpaea*	+
	R . cellulosilyticum ALA10B2^T	*Medicago*	+
	R. undicola LMG 11875^T	*Neptunia*	+
Sinorhizobium	*S. arboris* LMG 14919^T	*Acacia*	+
	S. fredii LMG 6217^T	*Glycine*	+
	S. kostiense LMG 19227^T	*Acacia*	+
	S. meliloti ATCC 9930^T	*Medicago*	+
	S. medicae LMG1037^T	*Medicago*	+
	S. saheli LMG7837^T	*Acacia*	+
	S. terangae LMG6463^T	*Acacia*	+
	S. xinjiangense LMG17930^T	*Glycine*	+
Mesorhizobium	*M. amorphae* ACCC 19665^T	*Amorpha*	+
	M. chacoense Pr5^T	*Prosopis*	w
	M. ciceri USDA 3383^T	*Cicer*	+
	M. huakuii USDA 4779^T	*Astragalus*	+
	M. loti ATCC 33669^T	*Lotus*	+
	M. mediterraneum USDA 3392^T	*Cicer*	+
	M. plurifarium LMG 7836^T	*Acacia*	+
	M. septentrionale HAMBI 2582^T	*Astragalus*	+
	M. tianshanense LMG 18976^T	*Sophora*	+
	M. temperatum HAMBI 2583^T	*Astragalus*	+
Phyllobacterium	*P. trifolii* pETPO2^T	*Trifolium*	+
Bradyrhizobium	*B. canariense* BTA1^T	*Chamaecytisus*	w
	B. elkanii LMG 6134^T	*Glycine*	+
	B. japonicum LMG 6138^T	*Glycine*	+
	B. liaoningense LMG 18230^T	*Glycine*	w
	B. yuanmingense LMG 21827^T	*Lespedeza*	w
Azorhizobium	*A. caulinodans* ORS 571^T	*Sesbania*	+
Devosia	*D. neptuniae* J1^T	*Neptunia*	w

* Detected by the double-layer plate assay. +: postive w: weakly positive

Table 2SI. Putative rhizobial cellulase genes

Rhizobia	Locus tag*	Glycosyl Hydrolase Family
Rhizobium leguminosarum 3841	RL1648 RL0081	GH 8 GH 26
Rhizobium etli CFN42^{T}	RHE_CH01544 RHE_CH00072	GH 8 GH 26
Sinorhizobium medicae WSM419	Smed_5210 Smed_3669	GH 8 GH 26
Sinorhizobium meliloti 1021	SMb20462	GH 26
Mesorhizobium loti MAFF33099	mll7872 mlr2086	GH 26 GH 5
Bradyrhizobium japonicum USDA110	blr3367	GH 5

*Data from GenBank

Table3SI. Symbiotic phenotypes in white clover after inoculation with wild type ANU843 or the CelC2 cellulase minus mutant derivative Δ*celC2*. The results reported are the mean ± standard deviation of at least 4 (Had, Hac, Inf) or 20 (Noi+Nod, Shoot length) replicate samples per treatment. Phenotype designations are Noi (nodule primordia), Nod (emerged nodules), Had (moderate root hair deformations), Hac (marked curling of root hairs, the so-called "shepherd's crook"), and Inf (infection thread formation within root hairs). Values followed by the same letter are not significantly different from each other at *P=0.01* according to Fisher´s Protected LSD (Least Significant Differences test statistic).

Inoculant Strain	Symbiotic Phenotypes On White Clover Seedlings				
	Noi+Nod plant^{-1}	Had	Hac per plant^{-1}	Inf per plant^{-1}	Shoot length per plant^{-1} (cm)
ANU843wt	10.8^{a} ± 5.2	+	16.8 ± 1.7	13.8 ± 1.7	5.2^{b} ± 3.1
Δ*celC2*	12.0^{a} ± 5.8	+	0 ± 0	0 ± 0	3.0^{a} ± 1.4
Uninoculated	0 ± 0	-	0 ± 0	0 ± 0	3.0^{a} ± 0.6

CHAPTER III. Development of functional symbiotic white clover root hairs and nodules requires tightly regulated production of rhizobial cellulase CelC2

Robledo M[1], Jiménez-Zurdo JI[2], Soto MJ[2], Velázquez E[1], Dazzo F[3], Martínez-Molina E[1] and Mateos PF[1]

[1]Departamento de Microbiología y Genética, Universidad de Salamanca, SPAIN, pfmg@usal.es [2]Estación Experimental del Zaidín, CSIC, Granada, SPAIN. [3]Department of Microbiology and Molecular Genetics, Michigan State University, USA.

The establishment of rhizobia as nitrogen-fixing endosymbionts within legume root nodules requires the disruption of the plant cell wall to breach the host barrier at strategic infection sites in the root hair tip (primary infection), and at points of bacterial release from infection threads within the root cortex (secondary infection). Using biochemical and loss-of-function genetic approaches, we previously found that *Rhizobium leguminosarum* bv. trifolii uses its cell-bound CelC2 cellulase to erode the highly localized non-crystalline wall at the apex of root hairs thereby creating the primary portal of its entry into white clover roots. Here, we show that a recombinant derivative of *R. leguminosarum* bv. trifolii ANU843 that constitutively over-produces the CelC2 enzyme has increased competitiveness in occupancy of aberrant nodule-like structures on clover roots that are inefficient in nitrogen fixation. Microscopy revealed that this aberrant symbiotic phenotype involves an extensive uncontrolled degradation of the host cell walls restricted to the expected infection sites. Furthermore, signs of Reactive Oxygen Species (ROS) production in root tissues were more evident during infection by the recombinant strain than its wild-type parent. Our data supports a role of CelC2 cellulase in both primary and secondary symbiotic infection of white clover roots, and also provides experimental evidence for a predicted requirement of tight regulation in the production and activity of rhizobial cell wall degrading enzymes to establish a compatible effective root nodule symbiosis.

Mutualistic symbioses between nitrogen-fixing rhizobia and their legume plant hosts are of critical agronomic and environmental importance, allowing successful crop production in nitrogen-limited soils without the need for chemical-N fertilizers. Rhizobia invade their plant hosts through colonization of intercellular epidermal spaces, crack entry at emerging lateral roots or penetration into the root hairs *via* tubular structures called infection threads (1). The later route is the best characterized infection pathway, elicited by a complex molecular dialogue between both symbiotic partners. During early symbiotic stages, flavonoids exuded from legume roots induce the synthesis and secretion of lipochito-oligosaccharide signal molecules or nodulation factors (NFs) in their cognate rhizobia upon the transcriptional activation of the nodulation (*nod*) genes (2). In turn, NFs trigger specific morphogenetic responses in the plant such as root hair deformation and dedifferentiation of the cortical cells which reactivate their mitotic activity to divide, thereby generating the nodule primordium as the precursor of the nodular meristem (3). Primary penetration of the microsymbiont into their legume hosts involves the degradation of the root hair cell wall restricted to a precise location. This hydrolysis been shown to be mediated by cell-bound cellulolytic enzymes produced by the symbiotic bacteria sandwiched in the typical "shepherd's crook" structures of curled root hairs (4, 5). Concomitantly, the root hair cell wall redirects its inward growth, resulting in the formation of the infection thread that continues its growth to the base of the epidermal cell and then towards the nodule primordium where invading bacteria proliferate. Within developing nodules, rhizobia are released by endocytosis from intracellular ramifications of the infection threads into membrane-enclosed vesicles within the cytoplasm of some nodule cells (infection zone). Released bacteria undergo a process of morphological differentiation into nitrogen-fixing competent bacteroids, which in the mature nodules remain confined to the vesicles (*i.e.* the symbiosome) defined by a plasma membrane of host origin. Detailed ultrastructural examinations indicate that at the points of bacterial release into the host cells the infection threads are unwalled (6, 7). However, it remains to be elucidated whether these secondary infection sites arise by arrest of the cell wall synthesis during intracellular elongation of the infection threads or by local degradation of the wall surrounding infection threads catalyzed by cell-wall hydrolytic enzymes of bacterial and/or plant origin.

Nonetheless, the process of plant cell wall degradation during the primary and secondary infection events described above is predicted to be delicately balanced in order to preserve the integrity of the host cells throughout the mutually symbiotic interaction and to maintain compatibility by avoiding the plant defense response. It is known that in response to early stages of rhizobial infection, legumes show transient, localized defence-like responses, suggesting that rhizobia, like phytopathogenic bacteria, are perceived as intruders. Amongst plant defence-like responses detected in the *Rhizobium*-legume symbiosis are the production of root hair peroxidase, salicylic acid (SA) and Reactive Oxygen Species (ROS). The elevation in peroxidase activity, that requires pSym *nod* gene expression, is of short duration and confined to the localized site of incipient penetration through the small portal of entry at the primary infection site in the compatible combinations, whereas it is more intense and of longer duration in incompatible combinations (8). Although there is evidence that the controlled production of ROS in response to rhizobial infection might be necessary for an effective establishment of the symbiosis (9-11), an intense oxidative burst like the one detected in aborted infection threads in alfalfa plants (12) or the accumulation of H_2O_2 detected in response to incompatible rhizobia (8, 13) could interfere with rhizobial infection.

We have previously reported on a chromosomically-encoded cell-bound cellulase (CelC2) from *Rhizobium leguminosarum* bv. trifolii ANU843 that fulfils an essential role in the primary infection of white clover roots (5, 14). This enzyme was purified and characterized as a 1,4-β-D-endoglucanase with high substrate specificity for the non-crystalline cellulose as exists in clover root cell walls. These features restrict its early symbiotic activity to the *transmuro* erosion of the isotropic apex of growing root hair precisely at the point of the microsymbiont penetration into the host roots. That enzyme-catalyzed reaction is enhanced by the ability of Nod factors from *R. leguminosarum* bv. trifolii to disrupt the crystallization architecture of the growing clover root hair cell wall (15). Consequently, an ANU843 *celC2* deletion mutant was found to be unable to breach the root hair wall at the primary infection site or to form infection threads when inoculated onto white clover seedlings, thus eliciting non-fixing nodule structures devoid of bacteria in its host plant (5). In work presented here, we have used a gain-of-function approach to further explore the impact of the CelC2 activity on the infection process during development of the white clover-rhizobia symbiosis. Our findings suggest that a deregulation in the production and localized activity of this rhizobial CelC2 cellulase disrupts the balance between the plasticity of degradation and assembly of the cell wall architecture required for development of a canonical effective root-nodule symbiosis.

Results

Overexpression of the *celC2* gene in *Rhizobium leguminosarum* bv. trifolii ANU843. The wild-type ANU843 strain and its derivative ANU843C2$^+$ overexpressing the *celC2* gene were compared for their ability to hydrolyze carboxymethylcellulose (CMC) as revealed by different sensitive methods previously reported to reliably detect this enzymatic activity (Fig. 1). The double-layer plate assay using culture supernatants and sonicated cell extracts as the source of enzymes indicated that there were no differences in CMC activity between ANU843 and ANU843EV strains and that *trans*-overexpression of *celC2* resulted in an increase of the wild-type CM-cellulase activity as inferred from the intensity and diameter of the clearing halos resulting from substrate hydrolysis (Fig. 1A). Despite this overproduction, most of the activity detected in ANU843C2$^+$ was in the sonicated cell extracts rather than in the culture supernatants fluid confirming the cell-bound nature of CMC-degrading enzymes in *R. leguminosarum* bv. trifolii. The quantitative BCA assay with whole cell extracts further confirmed this observation, revealing a 3-fold increase of the total CMCase activity in ANU843C2$^+$ as compared to its parent strain (Fig. 1B). Finally, SDS-PAGE separation (Fig. 1C) of proteins in the same cell extracts followed by their activity assay resolved in gels (Fig. 1D) unequivocally demonstrated that the observed increase of CMCase activity in ANU843C2$^+$ was due to the overproduction of the 31-KDa CelC2 enzyme.

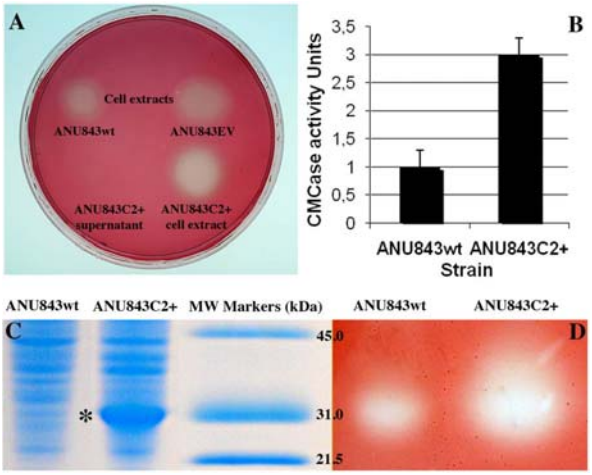

Fig. 1. Carboxymethylcelullase activity of *R. leguminosarum* bv. trifolii strains detected by different methods: (A) double-layer plate assay of 20 μL samples, (B) quantitative enzyme BCA assay, (C) SDS-PAGE, (D) activity stain overlay. 1, ANU843; 2, ANU843C2$^+$.

Irregular symbiotic performance of the ANU843C2$^+$ recombinant strain. Fig. 2 shows typical 3-month old white clover plants that were either uninoculated, inoculated with wild type ANU843 or with the *celC2*$^+$-overexpressing recombinant strain. ANU843-infected plants displayed a typical Fix$^+$ symbiotic phenotype (Fig. 2a). In contrast, most plants inoculated with the ANU843C2$^+$ recombinant strain exhibited a Fix$^-$ phenotype (Fig. 2b) characterized by stunted growth, shorter shoots, less dry weight and less N-content than plants inoculated with the wild-type strain ANU843. Development of clover or wounded tomato and tobacco plants grown in nitrogen-containing media and inoculated with the ANU843C2$^+$ strain was uncompromised (not shown). Together these findings revealed a negative influence of the overproduction of the CelC2 cellulase in the outcome of the *R. leguminosarum* symbiosis with its cognate legume host white clover under N-limited conditions. However, this effect is specific to the symbiotic root-nodule interaction since it was

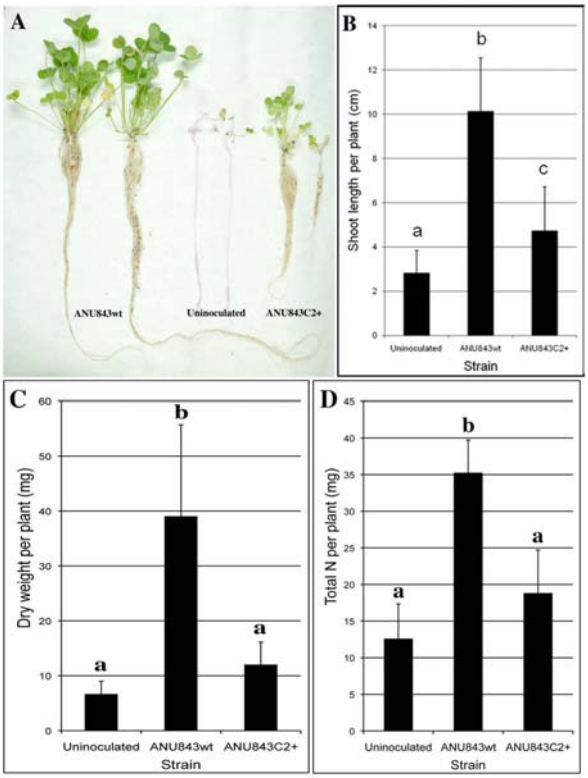

Fig. 2. Symbiotic phenotype of *R. leguminosarum* bv. trifolii strains. Nodule occupancy, shoot length, dry weight and Nitrogen content per plant reported in graphs are the mean of 24 replicate samples per treatment. Three replicates per treatment were analyzed. Values followed by the same letter are not significantly different from each other at P <0.01 according to the Fisher´s Protected LSD (least-significant differences) test statistic.

not observed either under conditions that inhibit nodulation or in non-legume plants.

***celC2* overexpression increased the competitiveness of *Rhizobium leguminosarum* bv trifolii strain ANU843 on clover roots.** The frequencies of nodule occupancy by ANU843pGUS3 in coinoculation assays with ANU843EV and ANU843C2$^+$ in greenhouse experiments are indicated in Fig. 2. Two months after inoculation, 76% of the nodules analyzed were occupied by the strain containing the pJZC2 plasmid, which overexpresses the CelC2 cellulase. This proportion was significantly higher than that obtained with the wild type strain under the same conditions (48,5%). The results of this competitive assay indicated that an increase in cellulase CelC2 production allowed ANU843 to invade and occupy a higher proportion of nodules than the control strain.

The ANU843C2$^+$ strain induces aberrant nodules in white clover roots. To better define the effect of CelC2 overproduction on symbiotic development, we examined the histology of the ANU843C2$^+$-induced nodules in more detail. The wild-type ANU843 strain induced elongated, pink, N_2-fixing nodules on clover roots (Fig. 3a). Brightfield microscopy of longitudinal nodule sections revealed the typical histology of indeterminate nodules with plant cells homogeneously occupied by differentiated bacteroids (Fig. 3b-c). Tracking of the nodulation process with an ANU843 GFP-tagged strain further confirmed the typical Fix$^+$ pattern of nodule colonization by wild-type endosymbiotic rhizobia (Fig. 3D). In contrast, the ANU843C2$^+$ strain induced predominantly nodule-like structures in white clover roots with abnormal histology including a marked dark brown necrotic zone containing undifferentiated bacteria in the distal region of the

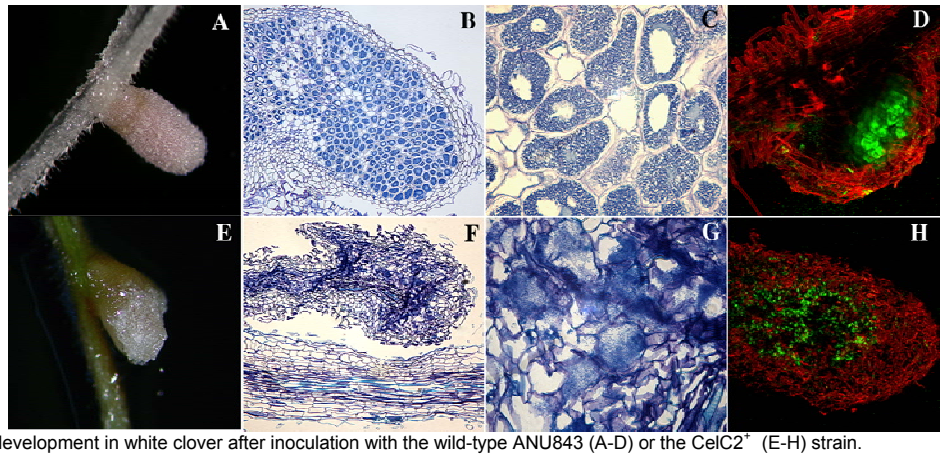

Fig. 3. Nodule development in white clover after inoculation with the wild-type ANU843 (A-D) or the CelC2^{+} (E-H) strain.

infection zone (Fig. 3 e-g). This disorganized histology is confined only to the nodule whereas the other root tissues had a normal histology (Fig. 3 f). Confocal microscopy of nodules induced by a GFP-tagged derivative of ANU843C2^{+} revealed that this recombinant was able to colonize the nodule tissues.

In summary, we found that overproduction of the CelC2 cellulase impairs normal nodule organogenesis in clover roots leading to the formation of abnormal nodule structures which nonetheless are invaded by undifferentiated CelC2-overproducing bacteria.

Cellulose architecture of white clover roots. We previously reported that the most intense activity of the purified rhizobial CelC2 enzyme on white clover roots is restricted to the localized hydrolysis of the non-crystalline cellulosic wall at the isotropic apex of root hairs (Hot phenotype), thereby creating the primary portal of entry of the microsymbiont into the plant (5, 14). To identify new potential sites for CelC2 activity, we have further analyzed the cellulose architecture of white clover cell walls in developing nodules upon infection with the wild-type strain ANU843. Polarized light microscopy of epidermal and vascular cells of white clover roots confirmed the uniform distribution of crystalline cellulose polymers throughout the cell walls except at the reported primary infection sites at the tips of the growing root hairs (Fig. 4a, and enlarged insert). Polarized light microscopy of nodule infection zone sections showed that the infection thread wall synthesized *de novo* upon ANU843 infection also exhibits a crystalline architecture similar in appearance to root hair cell walls (Fig. 4b-c, red arrows) except at the localized points of as isotropic non-crystalline cellulose where the bacteria are released into the plant cells of the nodule primordial (Fig. 4b-c, red circles).

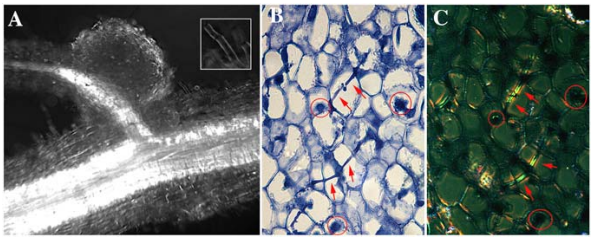

Fig. 4: Polarized (A, C) and brightfield (B) light microscopy of *Trifolium repens* roots (A) and nodules (B, C) induced by wild type ANU843. (A) Overview of crystalline architecture of the root epidermis and vascular cell walls (white retractile zones). Enlarged insert shows growing root hairs with non-crystalline structure at the tip. (B, C) Histological sections of the nodule infection zone. Walls of infection threads are also crystalline (red arrows) except at sites where bacteria are released (red circles).

Because of the restricted activity of cellulase CelC2 we predicted that only noncrystalline regions of polar-growth of plant walls can be hydrolyzed, therefore plant cell wall degradation by CelC2 would be restricted to root hair tips and infection thread branches and tips leading to bacterial point erosion instead of massive plant cell wall destruction.

Effect of CelC2 overproduction on primary infection. Root hair deformation and infection thread formation assays were performed on white clover plants inoculated with the wild-type strain of *R. leguminosarum* bv. trifolii or the recombinant strain that overproduces CelC2 (Fig. 5).

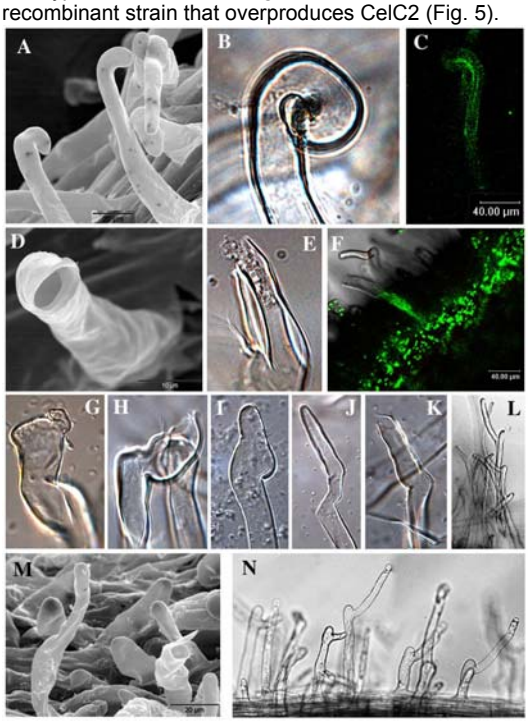

Fig. 5: Scanning electron (A, D, M), Nomarski-contrast (B, E, G-K), brightfield (L, N) and confocal epifluorescence (C, F) microscopy of primary infection events in white clover inoculated with the ANU843 wild-type (A-C), CelC2 overproducing recombinant strain ANU843C2^{+} (D-M) and their gfp-tagged derivatives (C, F), or treated with purified CelC2 (N). Root hairs inoculated with ANU843 display typical "curling" deformations (A, B) and infection thread formation (C). However, instead of hydrolysis only at the apex of the hair tip, the overexpressed ANU843C2^{+} recombinant strain produced extensive degradation throughout the distal end of root hairs (D-F). This exaggerated degradation of the root-hair wall resulted in the extrusion of the root hair protoplast/cytoplasm (E) and an enlarged portal of entry of the bacteria, resulting in their uncontrolled, massive penetration into the host plant through these dead cells (F). Deposition of new cell wall material before *transmuro* erosion of the original root hair tip, followed by redirection of new tip growth resulting in development of a branch in the root hair that is again susceptible to hydrolysis at its new growing tip (G-M). Biological activity of purified cellulase CelC2 from *R. leguminosarum* bv trifolii ANU843 on clover root hairs after 12 h of enzyme incubation (N).

As predicted, ANU843C2$^+$ induced various alterations at the root hair tip that included: (i) canonical primary infection stages as occurs in wild type with curling deformations (Fig. 5a), shepherd's crook (Fig. 5b) and infection thread formation (Fig. 5c); (ii) Hot phenotype as occurs in response to purified CelC2 enzyme (10, 16) where wall at the tip of some growing root hairs is degraded to produce a hole at this location (Fig. 5d) resulting in the extrusion of the root hair protoplast/cytoplasm (Fig. 5e and table 1SI).

Confocal microscopy of GFP- tagged bacteria showed that the dead root hairs with extensive wall degradation at their tip are invaded by ANU843C2$^+$ but did not form normal infection threads (Fig. 5F); (iii) new cell wall material deposited before *transmuro* erosion of the original root hair apex, followed by redirection of new tip growth displaced at a vertex angle of 119 ± 11 degrees from the longitudinal cell axis and development of a branch in the root hair that is again susceptible to hydrolysis at its new growing tip (Fig. 5G-M). This redirections are also observed after 12 hours of incubation of clover roots with purified cellulase CelC2 (Fig. 5N). These features show that the CelC2 cellulase-overproducing strain evokes severe aberrant phenotypes in root hair development, disrupting the balance between biosynthesis and degradation of its host´s cell wall.

Effect of CelC2 overproduction on secondary infection. To examine if CelC2 performs a symbiotic function during the secondary infection process, we monitored both ANU843wt and CelC2-overexpressing strains behaviour within nodules by TEM. ANU843wt strain followed the infection process steps: bacteria penetrate the plant root through the localized disruption of the root hair tip leading to infection thread formation where invading bacteria proliferate and reach the nodule (Fig. 6A). In developing nodules, rhizobia are released by a previous localized hydrolysis of infection thread cell wall (Fig. 6B) and subsequent endocytosis surrounded by a plasma membrane. This feature leads to the "infection droplet" formation (Fig. 6C) and differentiation into nitrogen-fixing competent bacteroids, forming the nitrogen-fixing symbiosomes (Fig. 6D).

Bacterial progression through the infection thread is similar in both ANU843wt (Fig. 6A) and the recombinant strain ANU843C2$^+$ (Fig. 6E). Microscopy analysis of the ANU843C2$^+$ strain in the nodule invasion zone revealed a strong hydrolysis of the infection thread cell wall noncrystalline zone from which the bacteria are released (Fig. 6F). Following the loss of stiffness, infection threads literally explode releasing CelC2 overproducing bacteria into the cytoplasm without being surrounded by the citoplasmic membrane (Fig. 6 G). Thus, bacteria invade the nodule cells by an uncontrolled way (Fig. 6H) originating plant cell disorganization and absence of symbiosome formation. These bacteria do not differentiate into bacteroids, and therefore do not fix nitrogen. As a result of the absence of peribacteroid membrane, bacteria are in direct contact with the nodule cells causing a complete disruption of the plant cell wall (Fig. 6H). These observations indicate that the CelC2 cellulase intervenes decisively in bacteria liberation from infection threads and that its concentration is critical. Thus an excess in CelC2 production leads to naked bacteria release without peribacteroid membrane that induce a nodule premature senescence and eventually the symbiotic nitrogen-fixing process abortion.

CelC2 overproducing strain of *R. leguminosarum* bv. trifolii induces higher accumulation of ROS than does the wild type strain in clover plants. Sensing of cell wall integrity may be one of the mechanisms used by plants to detect pathogen attack and in response, activate a variety of defenses like the production of ROS or SA aimed at limiting pathogen ingress (16). Our findings described above suggest that the overproduction of the CelC2 cellulase by *R. leguminosarum* bv trifolii would cause a loss in plant cell wall integrity that is severe enough to activate plant defenses and in turn, could explain the observed impairment in the establishment of an effective symbiosis. To test this hypothesis, SA and ROS accumulation were measured at different time intervals during the *Trifolium repens- Rhizobium leguminosarum* bv. trifolii ANU843EV and ANU843C2$^+$ interactions. No significant differences in SA accumulation were detected in clover plants inoculated with the CelC2 overexpressing strain compared to those inoculated with the control strain harboring the empty vector. However, fluorescence microscopy using the ROS-sensitive dye 2', 7'-dichlorodihydrofluorescein-diacetate (H_2DCFDA) detected a higher percentage of root hairs showing fluorescence of this defense response marker in plants inoculated with the CelC2 overexpressing-strain than in those treated with ANU843EV at different time points and locations (Fig. 7H).

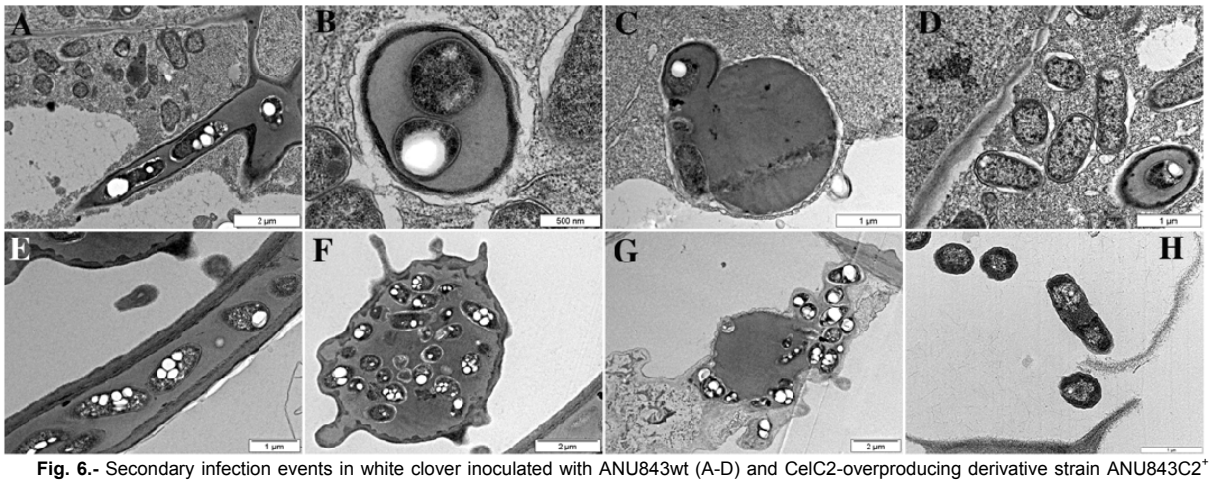

Fig. 6.- Secondary infection events in white clover inoculated with ANU843wt (A-D) and CelC2-overproducing derivative strain ANU843C2$^+$ (E-H). A) and E) are longitudinal sections of nodule infection threads containing vegetative rhizobia within their lumen B) and F) Shows a cross-section of an infection thread containing vegetative rhizobia. Micrograph C) shows a wild-type localized erosion of the infection thread wall to form the infection droplet and D) is an infected nodule cell with a diagonal section through a walled (electron-dense) infection thread (lower right corner) plus vegetative rhizobia that have been released from the infection thread into the intact nodule cell and are surrounded by peribacteroid membranes. And F to H TEM micrographs show the abnormal and destructive features of the overexpressing recombinant within nodule cells which are lysed without cytoplasm or intracellular organelles. Note that in F) and in G) there is a significantly widened infection threads with extensively eroded walls and unstable "exploding" infection droplets and micrograph H) shows naked vegetative rhizobia released within a lysed nodule cell whose wall has been significantly disrupted.

We observed that ANU843wt can elicit soft signs of oxidative burst on clover at the early and late stages of symbiotic development (24 hours and 25 days after inoculation respectively). The early time course is similar to the peroxidise activity elicitation found before in clover after inoculation with strain ANU843 (8). Moreover, this biphasic oxidative burst matches the already described response in the compatible *S. meliloti- Medicago truncatula* interactions where ROS transiently and weakly accumulates during the initial steps of the interaction between those symbionts (17). Therefore ROS is thought to contribute to the infection process progression (8), linked to the NFs signaling transduction pathway (11). However, the time course of ANU843C2$^+$ ROS production increase to a higher level than after inoculation with ANU843EV or the uninoculated control. This increase in ROS activity has been also reported in incompatible symbiosis (8).

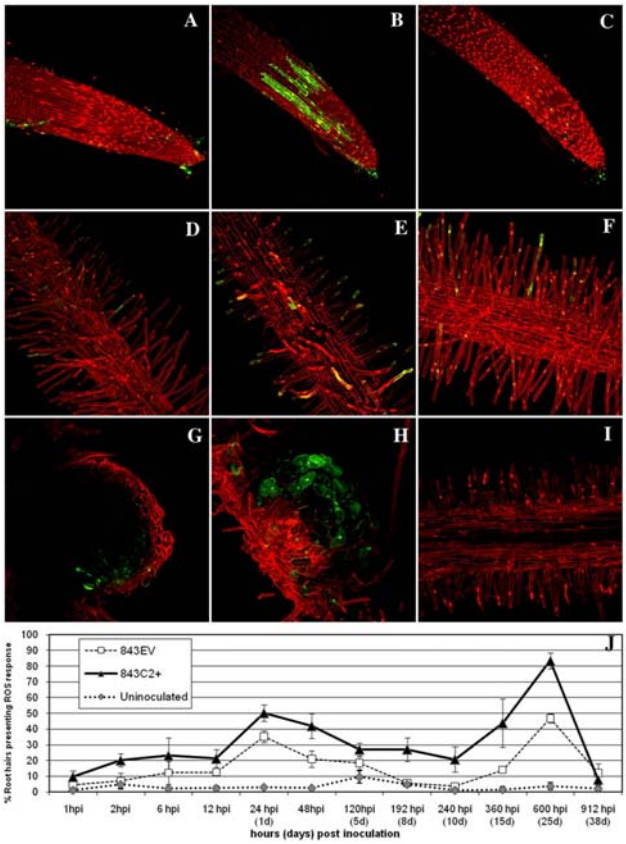

Fig. 7: Confocal microscopy showing ROS production (green fluorescence) in *Trifolium repens*. Images showing root meristem and root hairs of plants 8 days after inoculation with strain ANU843EV (A, D), ANU843C2$^+$ (B, E) or uninoculated (C, F); and nodules form by ANU843EV (G) and ANU843C2$^+$ (H) in plants 15 days after inoculation. (I) Involvement of NADPH oxidase in ROS production of seedlings inoculated with ANU843C2$^+$ was indicated by decreased fluorescence brightness following addition of diphenylene iodonium (DPI) inhibitor. A representative experiment from three with similar results is shown. (J) Percentage of root hairs expressing the ROS response at different time intervals after inoculation.

Clover plants inoculated with ANU843C2$^+$ also displayed a stronger and more prolonged emission of fluorescence in its cortical, meristematic and root hair cells (Fig. 7B, E) than ANU843EV (Fig. 7A, D). One week after inoculation, the oxidative burst was especially localized at the root hari tip matching the distribution of cellulase CelC2 substrate, suggesting that the localized action of this rhizobial enzyme may possibly contribute to the elicitation of the oxidative burst by the on the host cell (Fig. 7E).

It has been reported that abundant H_2O_2 associated with degrading bacteroids is localized in the senescent zone (IV), reflecting the oxidative stress ensued during nodule senescence (18). We have observed signs of weak oxidative burst in senescent zones of 2 week old nodules on clover plants inoculated with the wild type strain (Fig. 7G), however, it seems that there is a massive production of ROS in the invasion zone of nodules occupied by the ANU843C2$^+$ recombinant (Fig. 7H), showing that the uncontrolled hydrolysis of the noncrystalline zone of the infection thread cell wall where the bacteria are released is concomitant with ROS production and may prevent symbiosome formation from occurring.

In both treatments, the *Rhizobium*-induced fluorescence was inhibited by DPI (Fig. 7I), suggesting that the detected fluorescence is the consequence of ROS accumulation after activation of a plant NADPH oxidase (19) and that ANU843C2$^+$ induces higher accumulation of ROS as a defense response than does ANU843EV.

Taking into account these elevated ROS responses in plants inoculated with the ANU843C2$^+$ overexpression recombinant strain, it seems probable that this strain is recognize as a strongly invasive pathogen by clover plants at certain stages of the nodulation process.

Discussion

Root hairs and pollen tubes are the best-studied examples of plant cells that elongate through the process of tip growth. The polar growth during the elongation phase occurs through the loosening of cell wall components in specific zones of the growing root hair, i.e., just behind the root hair apex. In this way, root hair growth can be limited to only the tip, with the majority of the cell wall at the flanks of the root hair reinforced to prevent expansion (20). Infection threads develop from growing root hairs and are also thought to be tip-growing structures, and therefore are most likely to elongate by using at least some of the machinery involved in root hair growth before infection (4). The tip of the developing infection thread is a site of new wall and membrane synthesis, and has been proposed to involve inversion of the tip growth that is normally exhibited by the root hair, and to be the result of the reorganization of cellular polarity (4). Intracellular infection of legume host cells begins when rhizobia become entrapped between adjacent root hair cell walls within the shepherd crook structure. The entrapped microcolony of rhizobial cells apparently generates a high local concentration of cellulase, which may be required in order to reach a threshold concentration for localized degradation of the cell wall generating a point of intrusion. At the same time, the continued growth and division of entrapped bacteria generates an inward pressure counteracting the turgor pressure of the host plant cell and avoiding the extrusion of its protoplast. Elevation of peroxidase activity at that same location of incipient penetration likely cross-links some residual wall polymers, thereby aiding to avoid protoplast lysis (32). Entrapped live bacterial cells are apparently a requirement for the initiation of an inwardly growing IT at the center of a curled root hair cell (21). It is still not clear how an IT is initiated, but several elements in the process can be identified as a result of genetic and cytological analysis (22). Our results show how one of these elements, cell wall rupture, must be tightly regulated in order to generate an intracellular infection via the infection thread. Both the absence or an excess of CelC2 enzyme activity block infection threads initiation. This tightly regulation seems crucial not only in primary infection within root hairs but also in secondary infection within the nodule. When the bacteria reach the nodular infection zone, they must be internalized by a cortical cell and they must establish a niche within that cell. In a target cell each bacterial cell is endocytosed in an individual, unwalled membrane compartment that originates from the infection thread. Once the released bacteria have

been engulfed within host cell membranes, they must survive within the symbiosome compartment and differentiate into the N_2-fixing bacteroid form. An overproduction of CelC2 leads to a loosening of cell wall at the tip of infection thread. This weak cell wall cannot retain the pressure of the infection thread lumen and literally explodes. As a consequence, naked bacteria are released into the interior of nodule cells. Without surrounding peribacteroid membranes, bacteria cannot complete the bacteroid differentiation program that includes expression of the enzymes of the nitrogenase complex and thus cannot begin to fix N_2. Furthermore, these naked bacteria induce a premature senescence of the nodule.

There are two host plant physical defense barriers: the cell wall (in direct contact with external bacteria) and the cytoplasmic membrane (in contact with the inward plant cell). An effective symbiosis needs to maintain at least cell membrane integrity through infection threads. However, overt cell wall hydrolysis could lead to lost of turgescence and destruction of the cell membrane. Without the presence of the overlapping shepherd crook, the hole made by CelC2 cellulase at a root-hair tip leads to extrusion of a larger portion of the protoplast which ultimately lyses, killing the root hair cell. However, during bacterial released from infection threads (secondary infection), overt-hydrolysis of the noncrystalline cell wall lead to bacterial “intrusion” in the nodule cells. Both effects abort the infection process: the first one blocks normal development of functional infection threads and the second one induces a premature nodule senescence.

Our results document the crucial role of this *Rhizobium* cell-bound enzyme in the root-nodule symbiosis with legumes, and identifies these enzymes as biotechnological targets to optimize the balance between endophytic and endosymbiotic states in rhizobial biofertilizers. Confocal microscopy of roots showed that 843C2 invaded the plant but unlike the wild type strain, sometimes it did not form infection threads in clover root hairs and also failed to nodulate properly. Some of the nodules induced in clover by this CelC2 overproducing strain are white and deformed but, unlike the *celC2* minus mutant (5), these nodules are infected by bacteria. Moreover, the overproducing strain is more infectious that the wild type to create more portals of entry into the plant. TEM of the nodule invasion zone revealed that the non crystalline zone of the infection thread cell wall is highly hydrolyzed by these CelC2 overproducing bacteria. The CelC2 overproducing bacteria invade the nodule cells by an uncontrolled way. This behaviour breaks the canonical symbiotic process leading to invaded but ineffective nodules.

The results reported here show that the N_2-fixing *Rhizobium*- legume symbiosis is a tightly regulated process, which needs a molecular balance between both symbionts. The cell-bound rhizobial cellulase CelC2 enzyme is a clear example of how imbalances resulting in either loss or excess of its activity ultimately lead to abortion of the infection process. Although the mechanisms differ, the consequences are the same: absence of biological nitrogen fixation in this root-nodule symbiosis.

Matherials and methods

Bacterial strains, plasmids and growth conditions. Bacterial strains and plasmids used in this study are listed in Table 2 (SI). *R. leguminosarum* bv. trifolii strains were routinely grown in TY (23) or yeast mannitol agar (YMA) media (24) at 28°C. *Escherichia coli* was grown in Luria Bertani (LB) broth at 37ºC. Antibiotics were added when required to the media at the following concentrations: kanamycin (50 µg/ml) or/and tetracycline (10 µg/ml).

DNA methods and construction of ANU843 derivative strains. Plasmid DNA preparation and recombinant DNA techniques were performed according to standard procedures (25). To generate the recombinant ANU843C2$^+$, a DNA fragment of parental strain ANU843 containing the cellulase CelC2 coding sequence along with 78 nt upstream of the start codon was inserted between the *Xho*I and *Sal*I sites of pBBR1MCS-2, thus yielding plasmid pJZC2 which was mobilized into *R. leguminosarum* bv. trifolii ANU843 by a triparental mating using pRK2013 as a helper. GFP- or GUS-tagged bacteria were similarly constructed by conjugation of plasmids pHC60 or pGUS3, respectively, into the corresponding recipient strains (ANU843 or ANU843C2$^+$).

Enzyme Assays. Cellulase activity in sonicated bacterial cultures standardized to cell density was examined by a double-layer plate enzyme assay using Congo red staining, a reducing sugar assay using 2, 2-bicinchonic acid (BCA), and SDS/PAGE followed by activity stain overlay (26,27). One unit of enzyme activity (U) is defined as the amount releasing 1 nmol of reducing sugar equivalent per min at 40°C.

Plant growth conditions. White clover (*Trifolium repens* L.) seeds were surface-sterilized and germinated as previously described (5). For long-term greenhouse experiments, surface-sterilized white clover seeds were germinated for 2 days in humid air on inverted plates and incubated at room temperature in the dark. Seedlings with straight roots were transferred to pots containing sterile vermiculite as support. After 48 h, each plant was inoculated with 1 ml of a bacterial suspension (OD_{600nm} 0.5; $4x10^8$ bacterial cells) in nitrogen-free Fähraeus medium (28). Pots were watered as needed and fertilized once every week with either nitrogen-free Fähraeus nutrient solution or half-strength Hoagland's no. 2 (Sigma) complete plant growth medium. Plants were harvested 3 months after inoculation to determine the shoot length, dry weight and nitrogen content of the aerial parts according to the A. O. A. C. Method (29). Data were analyzed by one-way analysis of variance, with the mean values compared using the Fisher´s Protected LSD (P= 0.01). For competition assays, germinated clover plants were inoculated with bacterial suspensions containing 1:1 mixtures of ANU843EV:ANU843pGUS3 or ANU843C2$^+$: ANU843pGUS3. Two months after inoculation, plant roots from each of the three replicates per treatment were pooled and stained to detect GUS activity in nodules after incubation of roots in X-Gluc solution [1 mM X-Gluc, 1% (w/v) SDS, 50 mM sodium phosphate buffer (pH 7.5)] at 37°C in the dark (30). Nodule occupancy was assessed by blue (occupied) and white (unoccupied) nodule counting.

For the analysis of ROS responses, seedlings with straight roots were transferred to the surface of 1.5% agar plates containing nitrogen-free Fähraeus medium and overlaid with Whatman No. 1 sterile filter paper wetted with four milliliters of sterile water. The plates were incubated vertically at room temperature in the dark for 48 h and then transferred to a plant growth chamber programmed for a 16:8 photoperiod. After 5 days, plants were inoculated with a suspension ($OD_{600nm}$of 0.5, 80 µl/ plant) of 5 day-old *Rhizobium* culture grown on YMA.

Microscopy. To assess root hair deformation and primary infection events, root hairs were examined by phase-contrast and Nomarski interference-contrast light microscopy (5). Confocal epifluorescence microscopy was carried out with a Leica confocal microscope equipped with a krypton–argon laser using a blue excitation filter (excitation maximum 488 nm; 530-nm long-pass filter). Roots and root hairs were stained with 10 µM propidium iodide (Sigma). Projections were made from adjusted individual channels in the image stacks using Leica software.

ROS was detected by treating clover roots at different times after *Rhizobium* inoculation with 10 µM 2',7'-diclorodihidrofluorescein-diacetate (H_2DCFDA, Sigma), which converts to a fluorescent product when oxidized by H_2O_2 (54). After 15 min, the plants were carefully washed and examined by confocal microscopy as described above using a different excitation filter (excitation maximum 514 nm). The percentage of fluorescent root hairs in the infection zone II (4), root meristem and nodules was determined in at least three different plants for each time point and treatment.

Roots and nodules were excised from 40-days-old plants and processed for light microscopy using brightfield, cross-polarized and phase contrast optics, and transmission and scanning electron microscopy as previously described (14, 15, 31). In the latter case, samples were coated with osmium instead of gold.

***In situ* detection of ROS production in clover roots.** The ROS-sensitive dye 2', 7'-dichlorodihydrofluorescein diacetate (H_2DCFDA, Sigma) was used to detect *in situ* ROS production (32). Uninoculated and inoculated clover roots were harvested at different times and submerged in the dye solution prepared at a final concentration of 10 µM in DMSO. After 15 min, the plants were carefully washed with sterile water and root hairs and nodules were examined for fluorescence by confocal microscopy. The percentage of root hairs showing fluorescence present in zone II, was determined. To corroborate that fluorescence was specific for ROS production, a treatment with the NADPH oxidase inhibitor diphenylene iodonium (DPI) was performed (19). Briefly, DPI (20 µM) was added at the onset of *Rhizobium* inoculation and again 30 min before observations to test for inhibition of *Rhizobium*-induced ROS production. At least three different plants were analyzed for each time point and treatment.

Supplemental information

Table 1. Symbiotic phenotypes in white clover after inoculation with wild-type ANU843 or the CelC2 cellulase-overproducing strain

	Symbiotic phenotypes on white clover seedling		
Inoculant strain	Hac per cm plant^{-1}	Hot per cm plant^{-1}	Inf per cm plant^{-1}
ANU843EV	3,15 ± 0,69	0,18 ± 0,08	1,11 ± 0,54
ANU843C2^{+}	2,18 ± 0,29	3,82 ± 0,66	0,15 ± 0,04
UNINOCULATED	0 ± 0	0,12 ± 0,09	0 ± 0

Table 2.Bacterial strains and plasmids used in this study

Strains and plasmids	Description	Source or reference
***Rhizobium leguminosarum* bv. trifolii**		
ANU843	Wild-type Nod+ Fix+	(57)
ANU843EV	Derivative of wild type ANU843 containing pBBR1MCS-2; Km^{r}	This work
ANU843C2^{+}	ANU843 containing plasmid pJZC2 expressing*celC2*; Km^{r}	This work
ANU843pGUS3	ANU843 containing plasmid pGUS3; Km^{r}	This work
ANU843GFP	ANU843 containing pHC60 ; Tc^{r}	This work
ANU843C2^{+}GFP	ANU843C2^{+} containing pHC60; Km^{r}, Tc^{r}	This work
Escherichia coli		
DH5α	Host for cloning	(50)
S17.1	*thi pro hsd*R– *hsd*M+ *recA* RP4 2-Tc::Mu-Km::Tn*7* (Sp^{r}/Sm^{r})	(58)
Plasmids		
pBBR1MCS-2	Km^{r}, *lacPOZ mob*,cloning vector	(59)
pJZC2	CelC2 encoding fragment cloned into the *XhoI/SalI* site of pBBR1MCS-2	This work
pGUS3	*pnfeD-gusA* translational fusion in pB1101: Km^{r}	(60)
pHC60	pHC41 derivative containing the *gfp* gene constitutively expressed; Tc^{r}	(61)

Acknowledgements

We thank Asunción García for help in confocal microscopy and Luis Rey for providing pHC60. This work was supported by Junta de Castilla y León Grant GR49 and Ministerio de Ciencia e Innovación Grants AGL 2005-07796 and AGL 2008-03360. M. R. was supported by a PhD fellowship from Spanish Government.

Author contributions: M.R., E.M.-M., and P.F.M. designed research; M.R., J.I.J.-Z., M.J.S., E.V., F.D., and P.F.M. performed research; M.R., J.I.J.-Z., E.V., M.J.S., F.D., E.M.-M., and P.F.M. analyzed data; and M.R., J.I.J.-Z., M.J.S., F.D., and P.F.M. wrote the paper.

REFERENCES

1. Goormachtig, S., W. Capoen, E.K. James, and M. Holsters, Switch from intracellular to intercellular invasion during water stress-tolerant legume nodulation. Proc Natl Acad Sci U S A, 2004. 101(16): p. 6303-8.
2. Gibson, K.E., H. Kobayashi, and G.C. Walker, Molecular determinants of a symbiotic chronic infection. Annu Rev Genet, 2008. 42: p. 413-41.
3. Oldroyd, G.E. and J.A. Downie, Coordinating nodule morphogenesis with rhizobial infection in legumes. Annu Rev Plant Biol, 2008. 59: p. 519-46.
4. Gage, D.J., Infection and invasion of roots by symbiotic, nitrogen-fixing rhizobia during nodulation of temperate legumes. Microbiol Mol Biol Rev, 2004. 68(2): p. 280-300.
5. Robledo, M., et al., *Rhizobium* cellulase CelC2 is essential for primary symbiotic infection of legume host roots. Proc Natl Acad Sci U S A, 2008. 105(19): p. 7064-9.
6. Jones, K.M., H. Kobayashi, B.W. Davies, M.E. Taga, and G.C. Walker, How rhizobial symbionts invade plants: the *Sinorhizobium-Medicago* model. Nat Rev Microbiol, 2007. 5(8): p. 619-33.
7. Subba-Rao, N.S., P.F. Mateos, D. Baker, H.S. Pankratz, J. Palma, and F.B. Dazzo, The unique root-nodule symbiosis between *Rhizobium* and the aquatic legume, *Neptunia natans* (L. f.) druce. Planta, 1995. 196: p. 311-320.
8. Salzwedel, J.L. and F.B. Dazzo, pSym *nod* gene influence on elicitation of peroxidase activity from white clover and pea roots by rhizobia and their cell-free supernatants. Mol Plant Microbe Interact, 1993. 6(1): p. 127-34.
9. Cárdenas, L., A. Martínez, F. Sánchez, and C. Quinto, Fast, transient and specific intracellular ROS changes in living root hair cells responding to Nod factors (NFs). Plant J, 2008. 56(5): p. 802-13.
10. Jamet, A., K. Mandon, A. Puppo, and D. Herouart, H_2O_2 is required for optimal establishment of the *Medicago sativa/ Sinorhizobium meliloti* symbiosis. J Bacteriol, 2007. 189(23): p. 8741-5.
11. Ramu, S.K., H.M. Peng, and D.R. Cook, Nod factor induction of reactive oxygen species production is correlated with expression of the early nodulin gene rip1 in *Medicago truncatula*. Mol Plant Microbe Interact, 2002. 15(6): p. 522-8.
12. Vasse, J., F. Debilly, and G. Truchet, Abortion of infection during the *Rhizobium-meliloti*-alfalfa symbiotic interaction is accompanied by a hypersensitive reaction. Plant Journal, 1993. 4: p. 555-66.
13. Bueno, P., M.J. Soto, M.P. Rodríguez-Rosales, J. Sanjuan, J. Olivares, and J. Donaire, Time course of lipoxygenase, antioxidant enzyme acitivities and H_2O_2 accumulation during the early stages of *Rhizobium*-legume symbiosis. New Phytol., 2001. 152: p. 91-96.
14. Mateos, P.F., et al., Erosion of root epidermal cell walls by *Rhizobium* polysaccharide-degrading enzymes as related to primary host infection in the *Rhizobium*-legume symbiosis. Can J Microbiol 2001. 47: p. 475-487.
15. Dazzo, F.B., G.G. Orgambide, S. Philip-Hollingsworth, R.I. Hollingsworth, K.O. Ninke, and J.L. Salzwedel, Modulation of development, growth dynamics, wall crystallinity, and infection sites in white clover root hairs by membrane chitolipooligosaccharides from *Rhizobium leguminosarum* biovar trifolii. J Bacteriol, 1996. 178(12): p. 3621-7.
16. Hematy, K., C. Cherk, and S. Somerville, Host-pathogen warfare at the plant cell wall. Curr Opin Plant Biol, 2009. 12(4): p. 406-13.
17. Santos, R., D. Herouart, S. Sigaud, D. Touati, and A. Puppo, Oxidative burst in alfalfa-*Sinorhizobium meliloti* symbiotic interaction. Mol Plant Microbe Interact, 2001. 14(1): p. 86-9.
18. Rubio, M.C., et al., Localization of superoxide dismutases and hydrogen peroxide in legume root nodules. Mol Plant Microbe Interact, 2004. 17(12): p. 1294-305.
19. Mazel, A., Y. Leshem, B.S. Tiwari, and A. Levine, Induction of salt and osmotic stress tolerance by overexpression of an intracellular vesicle trafficking protein AtRab7 (AtRabG3e). Plant Physiol, 2004. 134(1): p. 118-28.
20. Knight, M.R., New ideas on root hair growth appear from the flanks. Proc Natl Acad Sci U S A, 2007. 104(52): p. 20649-50.
21. Gage, D.J. and W. Margolin, Hanging by a thread: invasion of legume plants by rhizobia. Curr Opin Microbiol, 2000. 3(6): p. 613-7.
22. Brewin, N.J., Plant cell wall remodelling in the *Rhizobium*–Legume symbiosis. . Critical Reviews in Plant Sciences, 2004. 23: p. 293–316.
23. Beringer, J.E., R factor transfer in *Rhizobium leguminosarum*. J Gen Microbiol, 1974. 84(1): p. 188-98.
24. Somasegaran, P. and H.J. Hoben, Handbook for rhizobia: methods in Legume-*Rhizobium* technology. 1994, New York: Springer-Verlag. xvi, 450 p.
25. Sambrook, J., E.F. Fritsch, and T. Maniatis, Molecular cloning: a laboratory manual. 2nd ed. ed. 1989, Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory.
26. Jiménez-Zurdo, J.I., P.F. Mateos, F.B. Dazzo, and E. Martínez-Molina, Cell-bound cellulase and polygalacturonase production by *Rhizobium* and *Bradyrhizobium* species. Soil Biol Biochem, 1996. 28: p. 917-921.
27. Mateos, P.F., et al., Cell-associated pectinolytic and cellulolytic enzymes in *Rhizobium leguminosarum* biovar trifolii. Appl Environ Microbiol, 1992. 58(6): p. 1816-22.
28. Fähraeus, G., The infection of clover root hairs by nodule bacteria studied by a simple glass slide technique. J Gen Microbiol, 1957. 16(2): p. 374-81.
29. Jonhson, F., Detection method of nitrogen (total) in fertilizers, in Methods of analysis of the Association of Official Analytical Chemists., K. Elrich, Editor. 1990, Association of Official Analytical Chemists, USA. p. 17–19.
30. van Dillewijn, P., M.J. Soto, P.J. Villadas, and N. Toro, Construction and environmental release of a *Sinorhizobium meliloti* strain genetically modified to be more competitive for alfalfa nodulation. Appl Environ Microbiol, 2001. 67(9): p. 3860-5.
31. Velázquez, E., P.F. Mateos, P. Pedrero, F.B. Dazzo, and E. Martínez-Molina, Attenuation of Symbiotic Effectiveness by *Rhizobium meliloti* SAF22 Related to the Presence of a Cryptic Plasmid. Appl Environ Microbiol, 1995. 61(5): p. 2033-2036.
32. Ubezio, P. and F. Civoli, Flow cytometric detection of hydrogen peroxide production induced by doxorubicin in cancer cells. Free Rad ic Biol Med, 1994. 16(4): p. 509-16.

CHAPTER IV. Bacterial cellulase CelC2 affects Nod factor–induced calcium spiking, nodulin gene expression and infection during symbiosis establishment in *Medicago*

Marta Robledo[1], Esther Menéndez[1], Jose Ignacio Jiménez-Zurdo[2], Encarna Velázquez[1], Eustoquio Martínez-Molina[1], Giles Oldroyd[3], and Pedro F. Mateos[1]

[1]Departamento de Microbiología y Genética, Centro Hispano- Luso de Investigaciones Agrarias, Universidad de Salamanca, Spain. [2]Estación Experimental del Zaidín, CSIC, Granada, Spain. [3]Department of Disease and Stress Biology, John Innes Centre, Norwich, UK.

The rhizobia–legume endosymbiosis provides the most efficient source of biologically fixed ammonia fertilizer for agricultural crops. The infection process is tightly regulated in legumes to ensure appropriate bacterial symbiont penetration to establish a nitrogen-fixing, intracellular state within the host. It has previously been shown that Nod Factor and exopolisacarides are involved in plant-bacterial recognition and signal transduction (1,2). However, how the process of plant cell wall erosion, required for development of root-nodule symbiosis is balanced to avoid lysis and destruction of the host cell wall is still unknown. Previous studies have shown that cellulase CelC2 from *Rhizobium leguminosarum* bv. trifolii can erode the noncrystalline apex of the white clover host root hair wall, fulfils an essential role in the canonical primary infection process required for development of this nitrogen-fixing symbiosis. CelC2 enzyme overexpression elicits signs of plant defense response and lead to aberrant infection thread and nodule formation on clover roots as a result of an extensive uncontrolled hydrolysis of the non-crystalline cellulose plant cell walls located just in primary and secondary infection sites. Here, we study the symbiotic phenotypes of a *Sinorhizobium meliloti* 1021 derivative strain constitutively expresing the CelC2 enzyme in plants of *Trifolium repens* and *Medicago*. This strain can reproduce the Hot phenotype in clover and is retarded in nodulation and nitrogen fixation in *Medicago sativa and M. truncatula* inoculated plants, suggesting that endoglucanase mediated infection is not restricted to clover plants. In addition, we show that application of purified cellulase CelC2 can interfere the nature Nod factor–induced calcium spiking and nodulin gene expression in *Medicago truncatula*. Therefore, plant is capable to sense an excess of bacterial cellulase altering the pathways associated with bacterial recognition and nodule formation.

The fixation of nitrogen by some microorganism is critical to maintaining the fertility of soil as it converts the inert atmospheric form of nitrogen into compounds usable by plants. Bacteria known as rhizobia, as used in this study, are an important group of microorganism as they have the ability to fix nitrogen – which they owe to a symbiotic relationship with plants in root nodules. Legumes are often used as a rotation crop to naturally enhance the nitrogen content of soils. Scientists have been working for a number of years to understand the first steps of the symbiosis between legumes and rhizobial bacteria, with the hope that one day they can transfer this trait to crop plants, the majority of which cannot fix nitrogen in association with rhizobia.

The infection process during the establishment of a rhizobia-legume symbiotic relationship is a complex molecular dialogue between the soil bacteria and the plant that follows a series of steps guided by reciprocal signal exchanges. The process is activated upon flavonoid perception by bacterial Nod factors (NFs) (3). In response to them, root hairs form a curl in order to let the microorganism penetrate into the cell and develop an infection thread that grows until the plant nodule, a new organ where nitrogen fixation takes place. Purified NFs elicit several responses in susceptible root hair cells, such as rhythmic Ca^{2+} oscillations termed calcium spiking (1), that are involved in signal transduction leading to transcripcional induction of several early nodulin genes (ENODs). A central event of this process is the erosion of the cell wall through which the bacterial symbiont penetrates into the plant because this degradation must be delicately balanced in order to occur without overdestruction of the host cell.

Our previous studies about the role of hydrolytic enzymes in the primary host infection by rhizobia (4) have shown that *Rhizobium leguminosarum* biovar trifolii strain ANU843 produces a cellulase, called CelC2, which can completely erode the highly localized noncrystalline tip of the host root hair, making a complete hole whose geometry and location match the entry point of primary host infection into white clover. CelC2 cellulase is a 1,4-ß-D-endoglucanase belonging to the glycosyl hydrolase family 8 with high substrate specificity for non-crystalline cellulose. ANU843 *celC* null mutants were unable to invade the plant at the entry point of the endosymbiont into clover. On the other hand, microscopy analysis of the overexpressed ANU843CelC2$^+$ strain in the clover root hair tips and the nodule invasion zone revealed an exaggerated tip degradation of host root hairs and a highly hydrolysis of the noncrystalline zone of the infection thread cell wall where the bacteria are released. Furthermore, signs of ROS production in root tissues were more evident during CelC2 overproducing derivative strain than wild-type infection (5).

All these results document the crucial physiological role of this cell-bound enzyme in the root-nodule symbiosis with clover, but nothing is known about its role neither in symbiotic signalling nor in other plant species. In this work we have assessed the infection and nodulation responses ot the clover, alfalfa and the model legume plant *Medicago truncatula* upon inoculation with *S. meliloti* 1021 derivative overexpresing the *celC* gene from a multicopy plasmid vector.

Our results show that, this strain is able to induce the hot phenotype in clover, invading the root hairs without infection thread development. Furthermore, we have shown aberrant symbiotic phenotypes in root hair deformation and nodule development in *Medicago* plants inoculated with the *S. meliloti* 1021C2$^+$ strain. This was the result of the extensive degradation of the isotropic non-crystalline cellulose which has been found exclusively located at the root hairs tip and the point of release of bacteria from the infection threads. These uncontrolled alterations of the cell wall polymers are accompanied by a significant alteration in canonical Nod-Factor induced responses such as calcium spiking traces and ENOD11 expression after treatment with purified cellulase CelC2 in the root hairs as compared to the Nod Factor. Our data support a role of the CelC2 cellulase not only in the primary but also in the secondary symbiotic infection of white clover and *Medicago* roots and provide experimental evidence for the predicted requirement of a tightly regulation of the production and activity of rhizobial cell wall degrading enzymes during a successful effective nodulation .

Results and discussion

Heterologous expresion of *celC* gene in *S. meliloti* 1021 resulted in an increase of CM-cellulase activity. In order to analyze the possible distribution of the effect of this endoglucanase coding gene in symbiosis with other plants, we have introduced the *celC* gene into *Sinorhizobium meliloti* 1021, which is a bacterium capable of infect and nodulate *Medicago* but not *Trifolium*. This genome has been sequenced and contains no *celC* homologs.

The wild-type *S. meliloti* 1021 strain and its derivative 1021C2$^+$ overexpressing the *celC* gene from a multicopy plasmid were compared for their ability to hydrolyze carboxymetylcellulose (CMC) using the double-layer plate assay, indicating that *trans*-overexpression of *celC* resulted in an increase of the cell-bound CM-cellulase activity (data not show). BCA assay with whole cell extracts not detect CMC-ase activity in 1021 wild-type strain but revealed a 6-fold increase of the total CMCase activity in 1021C2$^+$ as compared to *R. leguminosarum* bv. trifolii ANU843 parental strain. Furthermore, SDS-PAGE (Fig. 1A) analysis of the same cell extracts and the activity assay of the proteins (Fig. 1B) resolved on gels unequivocally demonstrated that the CMCase activity in 1021C2$^+$ was due to the production of the 31-KDa CelC2 enzyme. These assays confirm that cellulase CelC2 is expressed, processed and localized in 1021C2$^+$ strain as in *R. leguminosarum* bv. trifolii ANU843. We can also confirm that the *Sinorhizobium meliloti* 1021wild type strain does not produce this enzyme.

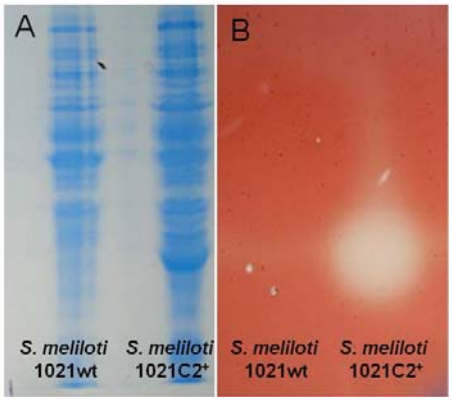

Fig 1. SDS-PAGE (A) and activity stain overlay (B) showing carboxymethylcelullase activity of the strains under study.

Effect of *S. meliloti* CelC2 heterologous expression on primary infection in *Trifolim repens* and *Medicago sativa*. Previous studies showed that the purified cellulase CelC2 isozyme from wild-type ANU843 was capable of degrading the cell wall at the apex of the root hair tip *in vivo* (“Hot” phenotype) when incubated with intact white clover host-seedling roots (4). Here, sterile plants of *Trifolium repens* var. HUIA and *Medicago sativa* var. Aragon were inoculated with *S. meliloti* 1021 or 1021C2$^+$ derivative strain labelled with green fluorescence protein.

Both *S. meliloti* strains attach to the root and to the root hair surface (Fig. 2 A, B, D) and induced typical non-specific deformations in clover root hairs (Fig. 2 A, B green arrows). *Medicago sativa* plants inoculated with *S. meliloti* 1021 display typical root hair curling and infection thread formation (Fig. 2C), but this strain is unable to induce these phenotypes on its non-host *Trifolium repens.*

On the other hand, *S. meliloti* 1021C2$^+$ derivative strain constitutively expressing CelC2 displayed several relevant differences in root hair deformation and invasion in *Trifolium repens* (Fig. 2E) and *Medicago sativa* (Fig. 2F) plants as a result of an extensive degradation of the non-crystalline cellulose located at root hair tips (shown as * in fig. 2E). Some of the tip-hydrolysed root hairs were able to continue its growth forming polar redirections (Fig 2D, black arrows). However, the majority of roots showed a large proportion of degraded root hair tips that even lead to massive root invasion without infection threads formation in both plants (Fig. 2 E, F). This extensive Hot phenotype is produced similarly in clover after treatment with *R. leguminosarum* bv. trifolii ANU843 derivative strain overproducing CelC2 cellulase (5) and has been also observed in *M. truncatula* plants upon inoculation with 1021C2$^+$ strain (data not shown).

Our results show that CelC2 highly localized site of cell wall degradation in *Medicago* plants may match the restricted distribution of the isotropic non-crystalline wall architecture at root hair tips found in *Trifolium repens*; and also suggest that heterologous endoglucanase expression may mediate invasion of non-host legumes, encouraging and spreading CelC2 enzyme role in root hair invasion of legumes.

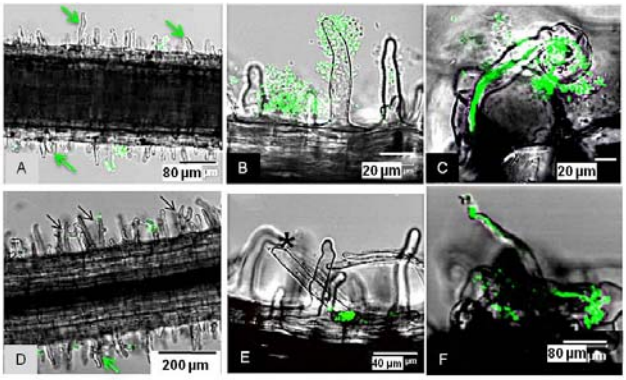

Fig. 2. Confocal micrographies of root hair deformation phenotypes of *Trifolium repens* (A,B,D,E) and *Medicago sativa* (C,F) inoculated with gfp-tagged *S. meliloti* 1021 (A-C) and 1021C2$^+$ derivative strain over expressing cellulase CelC2 (D-F). Green arrows show root hair common redirections and black arrow show redirections that apparently are produced after root hair tip degradation. (*) Root hair deformations in the form of tip-degraded and decapitated root hairs and showing an exaggerated Hot (Hole on the tip) phenotype.

Symbiotic performance of the 1021C2$^+$ recombinant strain in *Medicago* plants. Cells of *Sinorhizobium meliloti* 1021, both wild type and the strain which harbours CelC2 from *Rhizobium leguminosarum* bv. trifolii ANU843, called *S. meliloti* 1021C2$^+$, were taken for further symbiotic analysis involving *Medicago sativa* and *M. truncatula* plants, such as nodulation and greenhouse assays.

In plants inoculated with *Sinorhizobium meliloti* 1021 appeared red characteristic indeterminate nodules (Fig 3 A, B), but in those inoculated with 1021C2$^+$ the 26% show an aberrant morphology (Fig. 3 E, F) specially in the meristemática and infection nodular zones I and II. Histological sections of these aberrant nodules show disorganized cellular structures (Fig. 3 G, H) that differ to those containing nitrogen fixing symbiosomes observed in the wild type (Fig. 3 C, D). These abnormal nodules induced by the 1021C2$^+$ strain, contain bacteria that are not surrounded by the peribacteriodeal membrane, leading to non-fixing structures as a result of the CelC over-expression (Fig. 3 H).

Comparing the plants during the *in vitro* axenic assay, we observed that the mean number of nodules per plant was significantly higher after inoculation with the wild-type strain than with the 1021C2$^+$ derivative (Fig. 3I). Beside this, alfalfa plants inoculated with the wild type strain at different inoculum size *in vitro* can grow normally in a nitrogen-free medium indicative of an effective nitrogen-fixing symbiosis, whereas plants inoculated with the cellulase over-expressing recombinant displayed aberrant development that is more evident as the inocula size is increased. The percentage of nitrogen fixing-plants dropped from 50% (1X, DO_{600nm}: 0.5) to 40% (2X) and finally at 3X inoculum size almost no plants show a Fix$^+$ phenotype (15%) and the percentage of these abnormal nodules was 96%.

As a result of this abnormal development, the greenhouse assays with *Medicago sativa* plants inoculated with *S. meliloti* 1021 or 1021C2$^+$ derivative strain showed significant differences when grown in nitrogen free medium (Fig. 3I) as consequence of a reduction in nitrogen fixation. This significant reduction in shoot lenght, dry weight and nitrogen content of 1021C2$^+$-inoculated

plants compare with the wild-type strain (Fig. 3 K, L, M) may be related with the degradation of the infection thread noncrystalline

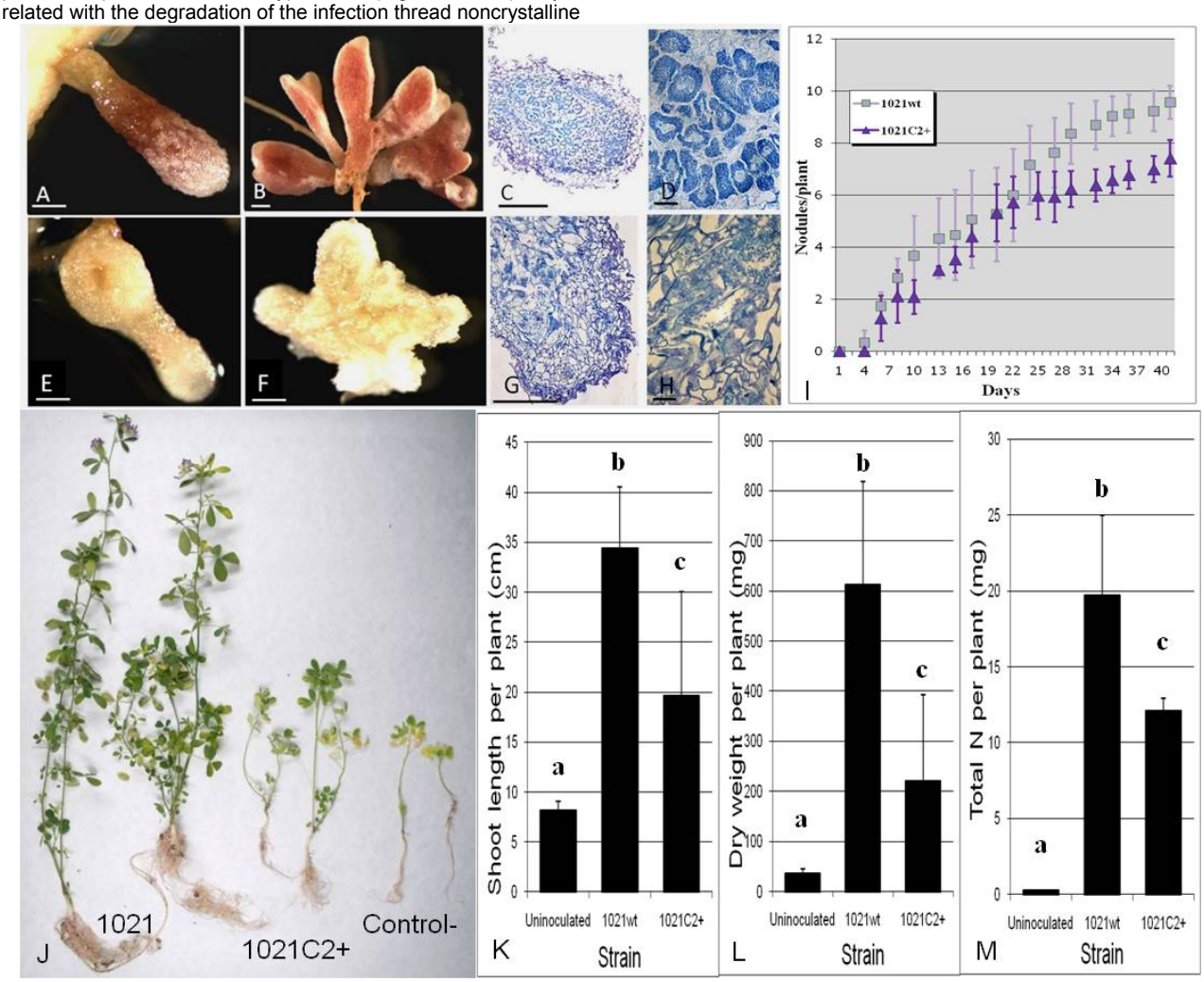

Figure 3. Symbiotic Phenotype of *M. sativa*-1021wt vs. 1021C2^{+}, CelC2-overproducer. Nodule development (Nod) in plants 33 days after inoculation with the wt 1021 (Fix^{+} and elongated) (A-D) or the CelC2^{+} derivative strain (Fix$^{+/-}$ senescent or more spherical) (E-H). Bar 1 mm (A-C); 500 nm (D-H). All samples were collected from the same experiment. (I) Time course of *S. meliloti*–induced nodule production in wild-type *M. truncatula* versus the CelC2^{+} derivative strain. Whole phenotype (J), shoot length (K), dry weight (L) and nitrogen content (M) per plant reported in graphs are the mean of 24 replicate samples per treatment. Three biological repetitions are represented. Error bars SD. Values followed by the same letter are not significantly different from each other at P <0.01 according to Fisher´s Protected LSD (least-significant differences) test statistic.

zone of the cell wall where bacteria are released, leading to strain 1021C2^{+} nodule invasion by an uncontrolled way, without being surrounded by the citoplasmatic membrane compartment, as we have previously seen in *Trifolium repens*.

We carried out infection and nodulation assays in *M. truncatula* the model legume plant in which the appropriate contructions to carry out assays regarding the signalling perception pathway are available. The results we obtained were similar to those described above in *M. sativa* (Fig. 4). All these features confirm that this strain invade the nodule cells by an uncontrolled way forming infective but ineffective nodules.

Ca^{2+} spiking in *M. truncatula* root hairs is influenced by CelC2 cellulase. Ca^{2+} spiking is a common feature in root nodule symbioses which can be induced by very low concentrations of Nod factor (10^{-11}–10^{-12} M) (6). Although the mechanism of induction of Ca^{2+} spiking is not yet well understood, it has been suggested that Nod factors may cause Ca^{2+} spiking by transiently opening a Ca^{2+} channel leading to a rapid increase in Ca^{2+}, followed by a slower phase of pumping Ca^{2+} out of the cytoplasm back into the Ca^{2+} store. The Ca^{2+} spiking signal is linked to the induction of early nodulation genes such as *ENOD11* and this requires the calcium-calmodulin-dependent kinase (CCaMK), which is thought to integrate the sequential changes in Ca^{2+} (7, 8).

To determine whether the infection and nodulation deficiency of CelC2^{+} derivative strain resulted also from decreased sensitivity to Nod factors, we assessed the effect of purified CelC2 cellulase on the Nod Factor-induced Ca^{2+} spiking responses in root hairs. We use changes in root hairs using *M. truncatula* expressing the YC2.1 cameleon calcium reporter, which can be used to analyse changes in Ca^{2+} in plant cells, binds of Ca^{2+} causing a conformational change in the chimeric fluorescent cameleon protein that increases the efficiency of fluorescent resonance energy transfer (FRET) from the cyan fluorescent protein (CFP) to the yellow fluorescent protein (YFP); thus as the Ca^{2+} concentration increases, fluorescence from CFP decreases and fluorescence from YFP increases (9).

Quantification of the distribution of spiking signatures showed that Ca^{2+} spiking was totally abolished in 43% of analyze root hairs by 5 pg/µl (1U) of purified CelC2 (Fig. 5A). In 40%, the signature was strongly affected (the average period is reduced without inhibition of Ca^{2+} spiking) (Fig. 5C) and only 17% of root hairs remain spiking normally after the treatment. There were no signs of root hair tip degradation during the time course of the calcium spiking assay.

Hence, Ca^{2+} spiking was suppressed or slowed down by CelC2, nevertheless, whether the addition of this enzyme has a direct inhibitory effect because it is causing a damage on the root hair tip, where the main nod factor receptors are located, or a degradation of the nod factor molecules due to their similar structure to its natural substrate is still unknown. Other possible

hypothesis to this NF-induced Ca^{2+} spiking inhibition is that the oligomers produced after CelC2-mediated plant cell wall hidrolysis elicit plant defence responses, as it has been previously reported in ANU843C2^{+} - *Trifolium repens* symbiosis (5).

Considering the importance of Nod factor signaling for bacterial infection and nodule initiation, the inhibition or retardation of the Ca^{2+} spiking frequency by CelC2 will leads to a decrease in the number of cells responsive to Nod factor signaling, which is sufficient to explain the nodulation retardation shown in plants inoculated with *S. meliloti* overproducing CelC2.

Effect of CelC2^{+} overexpression in Nod Factor signal transduction pathway. After response to Nod Factor induction, calcium signal is transduced via two nodulation signalling proteins, NSP1 and NSP2, which belong to the required for ENOD11 induction (10-12). To assess the effects of CelC2 cellulase on Nod factor–induced gene expression, we used *M.* (13)*meliloti* wild-type and CelC2^{+} derivative strain or purified Nod Factor and CelC2 enzyme. At different times after treatment, GUS activity was tested either by plant root staining or quantificacion of by the fluorimetric assay as described (14).

After GUS staining, we found typical expression following Nod-factor induction in a zone of epidermal cells starting about 0.6 cm from the root tip in plants treated with either *S. meliloti* 1021 wild type or purified Nod Factor (Fig. 4A, B and D). However, we found that inoculation with S. meliloti 1021C2^{+} reduced induction of ENOD11, especially 24 hours after treatment. CelC2 overexpression also modifies the timing of ENOD11–GUS induction, as treatment with 1021C2^{+} led to a decrease in activity in several time points, especially 36 h post treatment (Fig. 4C). Beside this, co-treatment with 5 pg/µl CelC2 (1U) almost abolished Nod factor induction of ENOD11 (Fig. 4 D, E). However, it seems that the differences observed upon the *in vitro* assays using purified NF and CelC2 protein are less noticeable than those obtained in the *in vivo* assays with the bacteria strains.

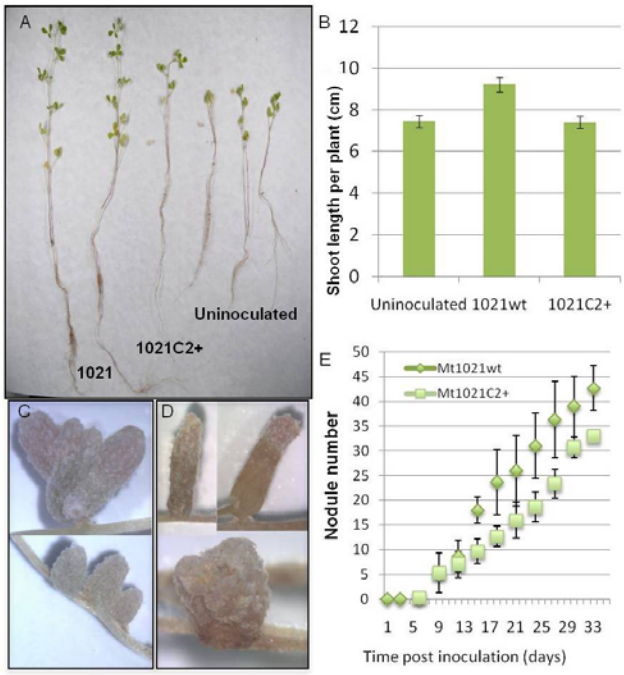

Figure 4. Symbiotic Phenotype of *M. truncatula*-1021wt vs. 1021C2^{+}, CelC2-overproducer. (A) Whole-plant phenotypes inoculated with *S. meliloti* 1021wt (left), the CelC2 overproducing derivative strain (center) and uninoculated (right). (B) Plant shoot length. Nodule development (Nod) in plants 33 days after inoculation with the wt 1021 (C) (Fix+ and elongated) or the CelC2^{+} derivative strain (D) (Fix+/- senescent or more spherical). All samples were collected from the same experiment. (E) Time course of *S. meliloti*–induced nodule production in wild-type *M. truncatula* versus the CelC2^{+} derivative strain. Three biological repetitions are represented. Error bars SE.

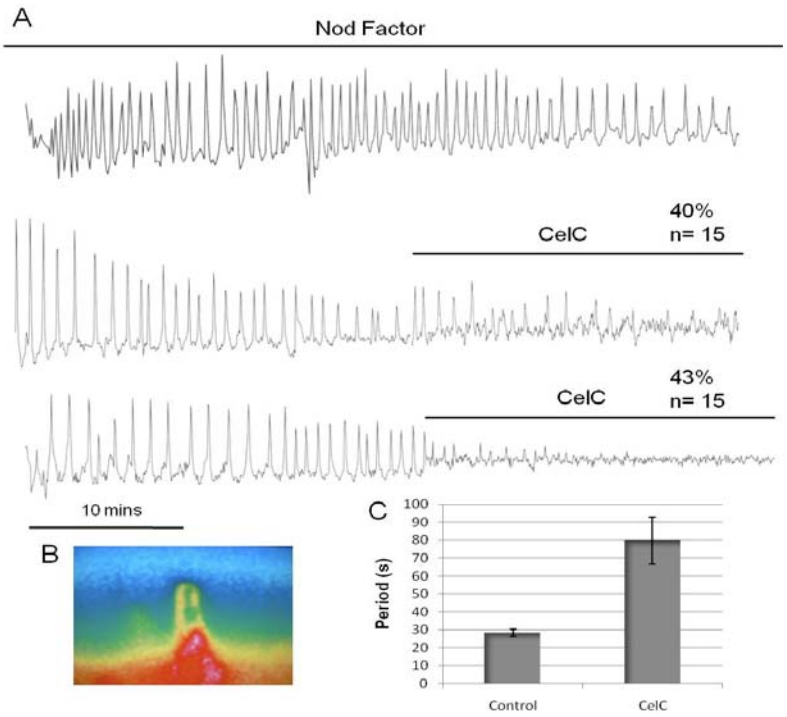

Figure 5. CelC Modulates Nod Factor–Induced Calcium Spiking.(A) Representative calcium traces of *M. truncatula* root hair cells preinduced with 10 pM Nod factor and secondarily treated with 5 pg/µl CelC (1U). The y axes represent the ratio of cyan fluorescent protein: yellow fluorescent protein (CFP:YFP) in arbitrary units. The top trace is a reference, showing NF-induced Ca^{2+} spiking in roots. Treatment with CelC2 causes a lengthening of the interval between individual calcium or inhibits calcium spiking. The percentage of visualized cells that follow each behaviour and the number of cells tested is shown. (B) Image of a *M. truncatula* root hair nuclear localization of the calcium responses changes in the root hair tip during a typical experiment. Calcium measurements were generated from cameleon transformed plants. (C) Effect of treatment with 5 pg/µl CelC2 on the average period of Ca^{2+} spiking. To analyze the periodicity of the spikes, the time between Ca^{2+} spike maxima was measured and averaged. Error bars indicate SE.

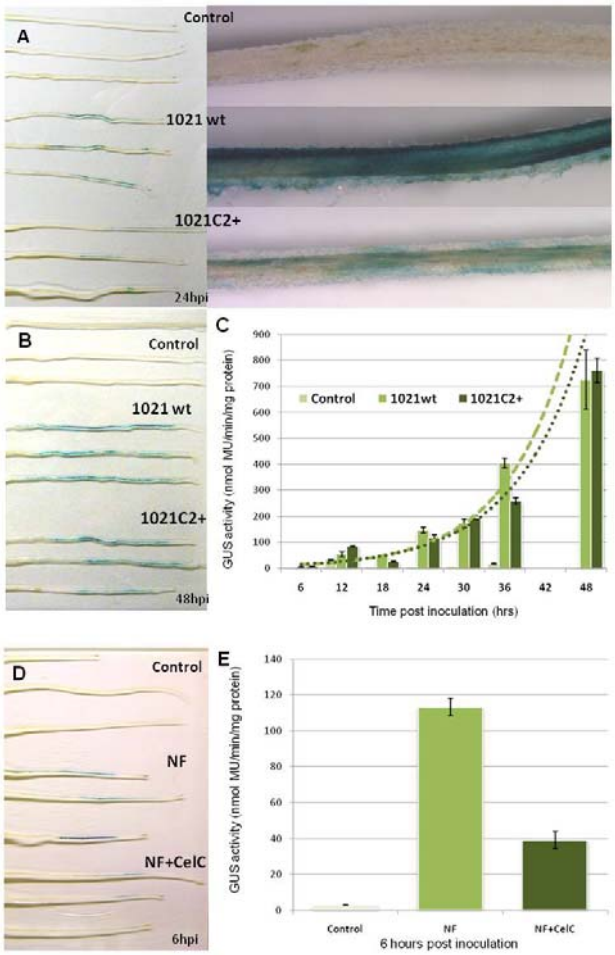

Figure 6. Histochemical Localization and Quantificacion of GUS Activity in Roots of Transgenic Plants Carrying a Fusion between the Promoter of *MtENOD11* and the *GUS* Reporter Gene after Treatment with *S. meliloti* or Nod Factors.
(A) ENOD11:GUS induction by *S. meliloti* strains 24 hours post inoculation and (B) 48 hpi. Plants without Nod factor treatment show GUS activity only in root caps, while bacteria activate ENOD11 in epidermal cells within a specific region of the root. *S. meliloti* 1021 wt shows slightly higher levels of ENOD11 induction than CelC2$^+$ overproducing derivative strain, especially at 24 hpi. (C) Quantification of GUS in wild-type plants carrying ENOD11:GUS after treatment with *S. meliloti* 1021 and 1021C2$^+$ at 6, 12, 18, 24, 30, 36, 42 and 48 hours shows slightly higher levels of ENOD11 induction tendency after addition of wild type strain. (D, E) ENOD11:GUS induction by 1 pM Nod factor at 6 hr is almost suppressed in the presence of purified CelC2.

All this results show that heterologous expresion of the CelC2 cellulase from *R. leguminosarum* bv. trifolli in *Sinorhizobium meliloti* mediates bacteria root hairs invasion on clover, and that a compatible canonical effective nodulation requires tightly regulated production of cellulases by the invading bacteria in *Medicago* plants, as assessed by early Nod factor signal transduction events, symbiosome development and nodule organogenesis.

Taken together, these results correlate with the Calcium signaling observations and indicate that CelC2 can negatively regulate early gene expression induced by rhizobial Nod factor and implies that CelC2 inhibition of nodulation may be through the regulation of the Nod factor signal transduction pathway. This data are in agreement with those obtained previously showing that cell wall-degrading enzymes interfere with Nod factor signalling in *Medicago sativa* root hairs (15) during an study that was designed to understand the loss Nod factor response after protoplasts preparation from plant cells using cell wall-digesting enzymes including cellulases.

Matherials and methods

Bacterial Strains and Plant Material.

Wild-type strain *Sinorhizobium meliloti* 1021, its derivative strain 1021C2$^+$ constitutively overexpressing the *celC* gene (this work) were routinely grown in tryptone yeast or yeast manitol agar medium at 28°C. *Escherichia coli* DH5α used for cloning and maintenance of plasmid constructs was grown in Luria Bertani (LB) broth at 37ºC. Antibiotics were added when required to the media at the following concentrations: Kanamycin (50 µg/ml for *E. coli* and 200 µg/ml for rhizobia) and Tetracycline (10 µg/ml). *M. sativa* and *T. repens* infection and nodulation assays and CelC2 cellulase extracts obtainment and quantification were performed as described before (4).

Medicago truncatula seeds of cv Jemalong line A17 and R108-1 were used as the wild type of the L416 line, transformed with pMtENOD11-GUS (16) and YC2.1, transformed with the yellow cameleon YC2.1 on p35S-yc2.1-kan (9). They were scarified with sand paper, surface-sterilized in 100% bleach, imbibed in sterile water for 5 hours, and plated on 1% deionized water agar plates. Seeds were incubated for 24 to 48 h at 4ºC and germinated on inverted agar plates at room temperature for 30 h. Seedlings were then transferred to Fähraeus medium containing 0.1 mM 2-aminoethoxyvinyl-glycine (Sigma-Aldrich) and placed in a 24°C growth chamber with a photoperiod of 16 h light, 8 h dark and a light intensity of 60 photons per m^2 s^{-1}. The bottom half of the plates was covered with foil to shade the roots until they were used for the different assays. Nod factor was prepared and quantified as described previously (17).

DNA methods and construction of 1021 derivative strains. Plasmid DNA preparation and recombinant DNA techniques were performed according to standard procedures (18). To generate the recombinant 1021C2$^+$, a DNA fragment containing the cellulase CelC2 coding sequence along with 78 bp upstream of the start codon was inserted between the *Xho*I and *Sal*I sites of pBBRMS-2, thus yielding plasmid pJZC2 which was mobilized into *S. meliloti* 1021 by a triparental mating using pRK2013 as helper. Transformants were analyzed by PCR and enzyme assays (4) to confirm the presence of the transgene of interest. GFP-bacteria were similarly constructed by conjugation of plasmid pHC60, into the corresponding recipient strains (1021 or 1021C2$^+$).

Microscopy. To assess root hair deformation and primary infection events, root hairs were examined by phase-contrast and Nomarski interference-contrast light microscopy (4). Confocal epifluorescence microscopy was carried out with a Leica confocal microscope equipped with a krypton–argon laser using a blue excitation filter (excitation maximum 488 nm; 530-nm long-pass filter). Roots and nodules were excised from 40-days-old plants and processed for brightfield and phase contrast light microscopy as previously described (Velázquez et al., 1995; Dazzo et al., 1996; Mateos et al., 2001).

Ca^{2+} Spiking Analysis. Spiking changes in root hairs using *M. truncatula* YC2.1 lines expressing the calcium reporter cameleon (9) was tested 1 to 2 d after germination. Seedlings were mounted on slides and Nod Factors were added to a final concentration of 10 pM as a positive control for the root hairs' ability to initiate calcium spiking. Root hairs were imaged for a minimum of 30 min after addition of either buffer or a solution containing 5 pg/µl CelC (1U) to determine whether calcium spiking was triggered. Treatments were added to the edge of the liquid in the incubation chamber and allowed to diffuse to the root to avoid disturbing those root hairs which were being imaged under the microscope.

Analyses of changes in fluorescence due to changes in calcium were done using epifluorescence microscopy as described (19) using a Nikon TE2000U inverted microscope (Nikon Instruments Inc., Melville, NY, USA) coupled to a Hamamatsu Photonics digital charge device camera (Hamamatsu Photonics UK Ltd., Welwyn Gardern City, UK).

The cian fluorescence protein (CFP) component of the cameleon was excited at a wavelength of 437 nm with an 11 nm bandpass using an Optoscan Monochrometer (Cairn Research, Faversham, Kent, UK). Emitted fluorescence was separated by an image splitter with a dichroic mirror and then passed through an emission filter of 485 nm for CFP fluorescence and 535 (15 nm) (Cairn Research) for the yellow protein (YFP) fluorescence. Images were collected every 5 sec with a 200–900 msec exposure and analysed using MetaFluor software (Universal Imaging Corp., Downington, PA, USA).

Regions of interest were outlined on the image and the average pixel intensity was averaged for each image in the data set. The ratio changes were calculated by dividing YFP intensities using CFP intensities by exporting the values into Microsoft Excel, converting them to a ratio and plotting them against time as described (20).

Assessment of ENOD11–GUS expression. GUS staining and activity measurement were performed as described (21). *M. truncatula* line 416K plants carrying the ENOD11 promoter driving the expression of GUS were grown on Fähraeus medium for 3 d. Plants were inoculated with 100 µl of a suspension of *S. meliloti* 1021 or *S. meliloti* 1021C2^{+} (OD_{600} 0.5) or 1 pM NF, with or without 0.5 ng CelC2 protein (1U) at different time points and all treatments left until the final time point. GUS staining was performed 2h at 37ºC with 1 mM 5-bromo-4-chloro-3-indolyl-ß-glucuronic acid, 5 mM EDTA, 20% Methanol, and 0.1 M potassium phosphate, pH 7.0. Subsequently, roots were observed under a stereomicroscope.

The experiments were repeated three times (n=18) with similar results. GUS activity was measured by the fluorimetric assay as described (14). Six plant roots were ground in liquid nitrogen and homogenized in GUS extraction buffer (50 mM sodium phosphate buffer, pH 7,5; 10 mM 2-mercaptoethanol; 10 mM EDTA; 0.1% Triton X100; and 0.1% sodium lauryl-sarcosine) for total protein extraction. Enzymatic reactions were performed during 1h at 37ºC in GUS extraction buffer using 1 µg of total protein extract with 1mM 4-methylumbelliferyl-ß-D-glucuronide as substrate (Sigma-Aldrich) and terminated with 1M Na_2CO_3. GUS activities were measured using a microtiter fluorimeter (FL600, Bio-Tek, Highland Park, VT, U.S.A.). Standard curves were prepared with a range of increasing concentrations of 4-methylumbelliferone (Sigma-Aldrich).

Acknowledgements. We thank Ward Capoen and the rest of the people in Oldroyd´s lab for help during my stay. This work was supported by grant AGL2008-03360. M.R. was supported by a FEMS Research Fellowship 2009-2.

REFERENCES

1. Oldroyd, G.E. and J.A. Downie, Coordinating nodule morphogenesis with rhizobial infection in legumes. Annu Rev Plant Biol, 2008. 59: p. 519-46.
2. Cheng, H.P. and G.C. Walker, Succinoglycan is required for initiation and elongation of infection threads during nodulation of alfalfa by *Rhizobium meliloti*. J Bacteriol, 1998. 180(19): p. 5183-91.
3. Jones, K.M., H. Kobayashi, B.W. Davies, M.E. Taga, and G.C. Walker, How rhizobial symbionts invade plants: the *Sinorhizobium-Medicago* model. Nat Rev Microbiol, 2007. 5(8): p. 619-33.
4. Robledo, M., *et al*.., *Rhizobium* cellulase CelC2 is essential for primary symbiotic infection of legume host roots. Proc Natl Acad Sci U S A, 2008. 105(19): p. 7064-9.
5. Robledo, M., *et al*.., Development of functional symbiotic white clover root nodules requires tightly regulated production of rhizobial cellulase CelC2. In Preparation.
6. Shaw, S.L. and S.R. Long, Nod factor elicits two separable calcium responses in *Medicago truncatula* root hair cells. Plant Physiol, 2003. 131(3): p. 976-84.
7. Levy, J., *et al*.., A putative Ca^{2+} and calmodulin-dependent protein kinase required for bacterial and fungal symbioses. Science, 2004. 303(5662): p. 1361-4.
8. Mitra, R.M., *et al*.., A Ca^{2+}/calmodulin-dependent protein kinase required for symbiotic nodule development: Gene identification by transcript-based cloning. Proc Natl Acad Sci U S A, 2004. 101(13): p. 4701-5.
9. Miwa, H., J. Sun, G.E. Oldroyd, and J.A. Downie, Analysis of calcium spiking using a cameleon calcium sensor reveals that nodulation gene expression is regulated by calcium spike number and the developmental status of the cell. Plant J, 2006. 48(6): p. 883-94.
10. Gleason, C., S. Chaudhuri, T. Yang, A. Munoz, B.W. Poovaiah, and G.E. Oldroyd, Nodulation independent of rhizobia induced by a calcium-activated kinase lacking autoinhibition. Nature, 2006. 441(7097): p. 1149-52.
11. Kalo, P., *et al*.., Nodulation signaling in legumes requires NSP2, a member of the GRAS family of transcriptional regulators. Science, 2005. 308(5729): p. 1786-9.
12. Smit, G., J.W. Kijne, and B.J. Lugtenberg, Involvement of both cellulose fibrils and a Ca^{2+}-dependent adhesin in the attachment of *Rhizobium leguminosarum* to pea root hair tips. J Bacteriol, 1987. 169(9): p. 4294-301.
13. Journet, E.P., *et al*.., *Medicago truncatula* ENOD11: a novel RPRP-encoding early nodulin gene expressed during mycorrhization in arbuscule-containing cells. Mol Plant Microbe Interact, 2001. 14(6): p. 737-48.
14. Jefferson, R.A., T.A. Kavanagh, and M.W. Bevan, GUS fusions: beta-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. Embo J, 1987. 6(13): p. 3901-7.
15. Carden, D.E. and H.H. Felle, The mode of action of cell wall-degrading enzymes and their interference with Nod factor signalling in *Medicago sativa* root hairs. Planta, 2003. 216(6): p. 993-1002.
16. Charron, D., J.L. Pingret, M. Chabaud, E.P. Journet, and D.G. Barker, Pharmacological evidence that multiple phospholipid signaling pathways link *Rhizobium* nodulation factor perception in *Medicago truncatula* root hairs to intracellular responses, including Ca^{2+} spiking and specific ENOD gene expression. Plant Physiol, 2004. 136(3): p. 3582-93.
17. Ehrhardt, D.W., R. Wais, and S.R. Long, Calcium spiking in plant root hairs responding to *Rhizobium* nodulation signals. Cell, 1996. 85(5): p. 673-81.
18. Sambrook, J., E.F. Fritsch, and T. Maniatis, Molecular cloning: a laboratory manual. 2nd ed. ed. 1989, Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory.
19. Wais, R.J., D.H. Keating, and S.R. Long, Structure-function analysis of nod factor-induced root hair calcium spiking in *Rhizobium*-legume symbiosis. Plant Physiol, 2002. 129(1): p. 211-24.
20. Capoen, W., *et al*.., Calcium spiking patterns and the role of the calcium/calmodulin-dependent kinase CCaMK in lateral root base nodulation of *Sesbania rostrata*. Plant Cell, 2009. 21(5): p. 1526-40.
21. Ding, Y., *et al*.., Abscisic acid coordinates nod factor and cytokinin signaling during the regulation of nodulation in *Medicago truncatula*. Plant Cell, 2008. 20(10): p. 2681-95.

CHAPTER V. Role of *Rhizobium* endoglucanase CelC2 in cellulose biosynthesis and biofilm formation on plant roots and abiotic surfaces

Robledo, M[1], Rivera, L[1], Jiménez-Zurdo, JI[2], Dazzo, F[3], Velázquez, E[1], Martínez-Molina, E[1], Hirsch, AM[4] and Mateos, PF[1]

[1]Departamento de Microbiología y Genética and CIALE, Universidad de Salamanca, SPAIN. [2]Estación Experimental del Zaidín, CSIC, Granada, SPAIN. [3]Department of Microbiology and Molecular Genetics, Michigan State University, USA. [4]Department of Molecular, Cell and Developmental Biology, University of California-Los Angeles, USA.

The synthesis of cellulose is probably one of the most important, but least known at the molecular level, biochemical processes, especially for bacteria, due to its complexity and high degree of regulation. In this article, we report that cellulose is produced by all tested members of the *Rhizobiaceae*, and that the *Rhizobium celABC* operon is predicted to be involved in bacterial cellulose biosynthesis. The *celC* gene encodes an endoglucanase that is required for canonical root hair infection in *Rhizobium leguminosarum* bv. trifolii. The encoded protein, CelC2 has a high substrate specificity for carboxymethyl cellulose (CMC) and does not degrade other polysaccharides. CelC2 mutants lacking (*celC*$^-$) or overproducing (*celC*$^+$) cellulase in *R. leguminosarum* bv. trifolii strain ANU843 produced greatly increased or reduced amounts of cellulose, respectively, as revealed by Congo red staining. They also developed a different colony morphology and thickness compared to the wild-type strain. Calcofluor-staining *celC*$^-$ bacteria formed cellulose microfibrils that were considerably longer than the wild-type strain. In addition, the *celC*$^-$ phenotype was correlated with a significant increase in culture flocculation. In contrast, neither microfibrils nor flocculation were detectable in *celC*$^+$ cells. Because these two mutants constitute an excellent model system to clarify the role of cellulose synthesis in aggregation and attachment in *Rhizobium*, we analyzed their ability to produce biofilms on plant roots and abiotic surfaces. Our results indicate that CelC plays a role in cellulose biosynthesis by modulating the length of the cellulose fibrils that mediate attachment among *Rhizobium* bacteria and that cellulose-deficient mutants as well as cellulose overproducers exhibit reduced biofilm phenotypes *in planta* and *in vitro*.

Diazotrophic bacteria–plant symbioses are of critical agronomic and environmental importance, making crop production possible in nitrogen-limited soils without added fertilizer supply. Rhizobia grow as free-living organisms, but can also induce and colonize root nodules in legume plants thereby establishing a partnership that benefits both organisms. This process begins when flavonoids secreted by the plant induce *Rhizobium nod* genes, which are involved in the synthesis and secretion of lipo-chitooligosaccharide signal molecules, known as Nod factors. In response, plant root hairs deform and exhibit a typical marked curling to facilitate bacteria penetration. The interaction continues with the initiation of the nodule, where bacterial cells are released into the host cells, differentiate into bacteroids, and afterwards fix atmospheric nitrogen into ammonia.

Among the many factors involved in development of an effective symbiosis between rhizobia and their host plants, those associated with colonization and adherence of bacteria to the roots and to the root hair tip surface, a key stage for the subsequent entry into the plant, have not been fully characterized. Root attachment and colonization by rhizobia follow the two-step system previously described for bacteria in general. For example, several bacterial proteins, such as rhicadhesin, lectins, and flagellar proteins (1-3), have been proposed to be important factors for the early reversible, specific binding events, whereas cell-surface polysaccharides are the main components involved in the later, irreversible attachment stages. Some of these cell surface components include exopolysaccharides, lipopolysaccharides, and cyclic β-1,2-glucans (4), but especially cellulose fibrils, which mediate attachment to root hairs and anchor bacteria to the root surface (5-7).

Attachment to a surface is the initial step in biofilm formation, which is followed by the establishment of microcolonies by clonal propagation, and finally maturation, the formation of three-dimensional structures that are covered by exopolymer and other matrix materials (3, 4, 8). Initially, the proposed function for cellulose in bacteria was not linked to biofilm formation. However, recent studies have revealed that other species of the family *Enterobacteriaceae*, such as *Citrobacter* spp., *Enterobacter* spp., and *Klebsiella* spp., produce cellulose as a crucial component of the bacterial extracellular matrix (reviewed in Lasa, 2006).

Cellulose biosynthesis, a characteristic feature of plants, has also been described in some bacteria and even in one group of animals, the urochordates (9). Cellulose biosynthesis in bacteria has been widely studied in *Gluconacetobacter xylinum* (formerly called *Acetobacter xylinum*) (10), but it is widespread in gram negative bacteria, some species of cyanobacteria, and the gram-positive species *Sarcina ventriculi* (11-13).

We have found orthologs to the putative cellulose synthase genes *celABC* in a region of the chromosome that is involved in cellulose biosynthesis (14-16) among a variety of organisms that synthesize cellulose. The first gene in the operon is *bcsA* (bacterial cellulose synthesis), also named *acsA* or *celA*. It encodes a cellulose synthase and harbors a β-glycosyltransferase 2 domain. It is the longer and the most conserved gene of the *bcs* operon among diverse species, and binds the substrate UDP-glucose (17). The second gene, *bcsB* (synonyms: *acsB*, *celB*), encodes a di-GMP binding protein and is less conserved. The last gene, *bcsZ* (also called *celC* in *A. tumefaciens* and *R. leguminosarum* bv. trifolii) has been shown to encode a cellulase (family 8 glycosyl hydrolase), which is present in all cellulose-producing species, but is located outside of, but adjacent to the cellulose biosynthesis operon in several *Gluconacetobacter xylinum* strains (18). Other genes such as *bcsC* and *bcsD* are required *in vivo*, but not *in vitro* for cellulose biosynthesis (19).

In both bacteria and plants, two proteins directly involved in the biosynthesis of cellulose have been characterized: CesA and Korrigan in plants, and CelA and CelC in bacteria (12, 20). CelA and CesA have homology with glycosyl transferases while CelC and Korrigan are homologous with endoglucanases. The postulated involvement of the *Agrobacterium* CelC cellulase enzyme in bacterial cellulose biosynthesis is to transfer new oligomer primers for chain elongation (21), but this biochemical function has not been definitively established.

Previous studies have shown that cellulose is involved in anchoring the pathogen *Agrobacterium tumefaciens* to plant tissue, thereby affecting virulence (15). In culture, *Rhizobium leguminosarum* bv. trifolii produces cellulose microfibrils by the polymerization of β-(1-4) glucose residues that are anchored to the roots of the host plant *Trifolium repens* (7). The lack of cellulose in *celA* or *celB* mutants affected neither *R. leguminosarum* bv. trifolii's ability to nodulate clover under controlled laboratory conditions (14) nor *celC* cellulase activity

(data not shown). In wild-type rhizobia, this polysaccharide is induced upon contact with the roots of the host plant and may be needed for optimal infection and biofilm formation on long root hairs (14, 22, 23). However, the need for cellulose in *Rhizobium* biofilm formation has not been fully characterized.

By obtaining knockout (*celC*$^{-}$) and a CelC2 overproducing (*celC*$^{+}$) mutants of *Rhizobium leguminosarum* bv. trifolii ANU843, we determined that cellulase is essential for both primary and secondary infection of *Trifolium repens*, most likely because it is responsible for the localized erosion of the root hair tip through which the bacteria enter and initiate the symbiotic process and the infection thread zone in which the bacteria are released (24, 25). Moreover, the *celC* gene is widely conserved among species of the genus *Rhizobium* (25), consistent with its involvement in fundamental processes for survival in the environment. In this work, we analyze the role of CelC in cellulose biosynthesis and biofilm formation. We found that *celC* mutant strains produced altered phenotypes in bacterial cellulose biosynthesis and biofilm formation.

Results

Cellulose production is widespread among Rhizobia. The tertiary structure of cellulose microfibrils makes them water-insoluble, which is the likely explanation for why cellulose production by bacteria is normally not detected using methodologies employed for the analysis of exopolysaccharide occurrence. However, different screens for qualitatively assaying cellulose-producing organisms, based on the characteristic binding behavior of dyes, such as Congo red or Calcofluor (14), to cellulose are frequently used. These assays may be followed by genetic and chemical analyses to confirm the results.

Our research group has previously detected cellulose microfibrils produced by *Rhizobium leguminosarum* bv. trifolii ANU843 that were associated with the roots of its host plant, clover (7). This feature has been described in other strains belonging to the same biovar (14) and in *Rhizobium leguminosarum* bv. *viciae* 3841 (23). However, the ability to synthesize cellulose is not one of the phenotypic characters that are routinely studied when describing this bacterial group. Thus, it is not known whether cellulose production is conserved among the different genera of currently known species, and therefore we decided to study this in the different rhizobia capable of establishing symbiotic nitrogen-fixing legume nodules. Currently, the known diversity of these bacteria is found within the order *Hyphomicrobiales*, which includes about 83 species distributed in different families and genera. Most of official type strains of each of these taxa were examined for cellulose production by Congo Red staining, and all were found to be positive to varying degrees (Figure 1 and Table 1).

The relative amount of cellulose production varies significantly among species from the different genera (fig. 1), and also between species of the same genus (Table 2). The organisms with a greater capacity to produce cellulose were generally found within the genera *Rhizobium* and *Mesorhizobium*, whereas representatives of *Sinorhizobium* and *Bradyrhizobium* showed lower Congo Red affinity.

A search of the databases revealed that all the currently defined Rhizobiaceae genomes possess genes coding for cellulose synthase that belong to Glycosyl Transferase family 2. Interestingly, these GT2 coding genes are located near endoglucanase *celC* homologs (belonging to Glycosyl Hydrolase family 8) in the *celABC* operon or near cellulase genes from Glycosyl Hydrolase family 26, forming a potential operon that contains a cellulose synthase associated with a cellulase and a hypothetical protein of unknown function. We have named this putative cellulose operon *celIJK*, and it has homologs in all finished and unfinished bacterial genomes belonging to *Rhizobiaceae* (Table 2). There are several organisms that share both cellulose production operons, and in *Agrobacterium tumefaciens* C58, *celABC* and *celIJK* are very close together in the genome.

We have sequenced the *celABC* genes from *R. leguminosarum* bv. trifolii ANU843. The nucleotide primary sequence and organization are similar and highly conserved among various Rhizobiaceae, with ca. 100% identity to *Rleg1292/Rleg1293/Rleg1294* and *Rleg21204/Rleg21205/ Rleg21206* from *R. leguminosarum* bv. trifolii WSM2304 and WSM1325 respectively, 89% to *celABC* from *R. leguminosarum* bv. trifolii R200 (14), 88% to *bcsa/celB/celY* from *R. leguminosarum* bv. *viciae* 3841 (26), 86% to *celABC* from *R. etli* CFN42 (27), 77% to *celABC* from *Rhizobium* sp. NGR234 (28), 91% to *celABC* from *A. tumefaciens* C58 (29), and 73% to Smed_5208/Smed_5209/Smed_5210 from *Sinorhizobium medicae* WSM419 (Table 3). Orthologs to this operon are also found in other bacteria that can nodulate legumes such as *Methylobacterium nodulans* ORS 2060.

Detection of a polymer that binds specifically to Congo Red in diverse species of the family Rhizobiaceae and the fact that cellulose synthases are found in the sequenced genomes strongly suggests that the ability to synthesize cellulose is fairly common in this group. Moreover, the fact that there is a cellulose synthase-encoding gene near one that codes for cellulase in all species with accessible data in GeneBank, together with the high degree of conservation of the *celABC* operon within *Rhizobiaceae* as well as the existence of at least *celC* among all tested species of the genus *Rhizobium* (25), further suggest that the involvement of a cellulase in cellulose production pathway is well conserved among these species.

CelC2 cellulase is involved in cellulose microfibrils formation. During the course of our research on the CelC2 protein; we observed that ANU843ΔC2 (*celC*$^{-}$) and ANU843C2^{+} (*celC*$^{+}$) show some significant differences in colony shape and in the degree of gumminess when they were grown in YMA solid media (Fig. 2A-C). Beside this, ANU843C2^{-} mutant, strain defective in the synthesis of cellulase CelC2, floculate heavily and form clumps in YMB liquid culture, that were undetectable in the *celC* complemented strain (data not shown) and in ANU843C2^{+} cellulase CelC2 overproduce derivative strain (Fig. 2B). The autoaggregation of each strain assay, assessed at 24 h determined that there was a reduction of 67% in the OD_{600} of the upper part of the standing culture tubes in comparision with the wild type. These clumps were completely dissolved after treatment with commercial cellulase from Sigma (data not shown).

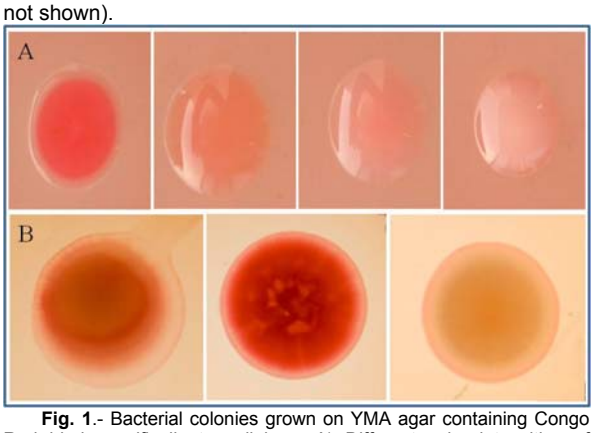

Fig. 1.- Bacterial colonies grown on YMA agar containing Congo Red bind specifically to cellulose. A) Different color intensities of cellulose-producing colonies, from left to right: *R. cellulosilyticum* ALA10B2 ^{T} (+++), *R. leguminosarum* bv. trifolii ANU843 (++), *R. hainanense* I66^{T} (+), and *S. kostiense* LMG 19227 ^{T} (w). B) Wild-type strain of *Rhizobium leguminosarum* bv. trifolii ANU843 (left, cellulose ++) and its derivatives ANU843ΔC2 (cellulose +++) and ANU843C2+ (cellulose-minus).

Aggregation of bacteria, usually mediated by cellulose microfibrils, and elimination of cell clumps by a cellulase are enough hints that indicate that an excess of cellulose was being produced (31). Nevertheless, we decided to observe Congo Red staining of *celC*$^{-}$ and *celC*$^{+}$ mutants, detecting

greatly increased and reduced amounts of cellulose, respectively, compared to the wild type (Fig. 1B). We also performed Calcofluor staining because this fluorescent molecule binds to cellulose showing the presence or absence of cellulosic microfibrils, showing that the microfibrils in the *celC*$^-$ mutant were significantly longer than the average found in wild-type and that they were absent in the *celC*$^+$ mutant (Fig. 2 C-E).

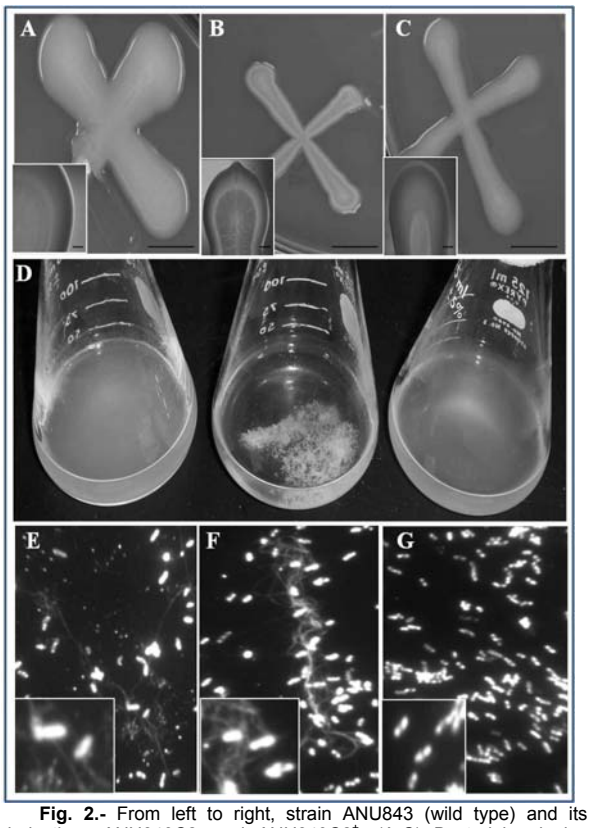

Fig. 2.- From left to right, strain ANU843 (wild type) and its derivatives ANU843C2$^-$ and ANU843C2$^+$. (A-C) Bacterial colonies grown on YMA (bar 1cm). At the down left side there is a magnification of the images (bar 1 mm). (D) Growth in stationary phase in liquid YMB shaken at 180 rpm. (E-G) Calcofluor staining showed the presence of microfibrils.

Ausmees *et al.* (1999) cloned *Rhizobium leguminosarum* bv. trifolii R200 strain genes involved in cellulose biosynthesis and found that they have homology and the same organization as the corresponding genes in *A. tumefaciens*. These authors found these genes by Tn*5* mutagenesis followed by screening with Calcofluor staining for mutants showed less ability to synthesize cellulose. They obtained *celA*, *celB,* and *celE* mutants but not *celC*. Their results are consistent with ours in that the *celC* mutant is a cellulose overproducer,so their screening method will discard this strain.

Therefore, we conclude that *celC* gene is involved in the formation and elongation of cellulose microfibrils. It is likely that cellulose oligomers synthesized by *celA* gene elongate indefinitely in the absence of an endoglucanase, causing very long microfibrils that lead to aggregate formation among *celC*$^-$ mutant bacterial cells and flocculation in liquid medium (fig. 2B). By contrast, in the *celC*$^+$ mutant cultures, there is an excess of activity corresponding to the encoding endoglucanase, which leads to an uncontrolled degradation of cellulose oligomers synthesized by CelA preventing their transport and subsequent formation of microfibrils that extend outside the cell (fig. 2G). It seems that these degraded oligomers without the lipid carrier accumulate inside the cell, more precisely in the areas of growth (poles and septum).

Cellulose microfibrils are involved in biofilm formation *in vitro.* Among the molecules involved in biofilm development are different proteins and some exopolysaccharides, such as cellulose, which is the major component of the biofilm matrix of several bacterial species (18). Previous studies have shown that cellulose is involved in anchoring the pathogen *Agrobacterium tumefaciens* to plant tissues, and also in biofilm formation, which affects virulence (31). In this work, we analyzed the biofilm forming ability of these cellulose mutants because similar mutants have often been shown to be impaired in biofilm development in other bacterial species (31, 33). By testing biofilm formation *in vitro*, using different abiotic substrates, and examining attachment *in planta* with the host plant clover, we qualitatively and quantitatively assessed biofilm formation by ANU843ΔC2 and ANU843C2$^+$ mutants in comparison to the wild type strain ANU843, based on methods previously described (3). Different patterns of biofilm formation on both abiotic and biotic substrates were assessed for the different strains.

Biofilm rings at the liquid-air interface in ANU843ΔC2 appeared to be thicker and more compact, but were released from the glass surface very easily compared to the wild type strain. In ANU843C2$^+$, the rings were much thinner, suggesting a role for CelC2 in biofilm formation (Fig 3A).

In an examination of mutant strain growth using the microtiter plate assay method, we found that YMA was not the appropriate medium to observe biofilm formation because the growth of ANU843ΔC2 bacteria was slightly retarded in comparison with wild-type due to flocculation. The use of either TY or minimal media containing mannitol resulted in no differences in growth among the strains (data not shown), but did show statistically significantly reduced biofilm formation for the mutants compared to the wild-type (fig. 3B-D). ANU843ΔC2 mutants exhibited a 30–50% decrease in biofilm formation in the microtiter plate assay (Fig.3 B-D) in both media. ANU843C2$^+$ exhibited ca. 30–60% reduced biofilm formation compared to the wild-type strain 24 h and 48 h after the start of the experiment. However, the differences were less pronounced at 72 h.

The most striking change was found between the *celC*$^-$ and wild-type strains because prior to washing, the biofilms of the microfibril-overproducing mutant appeared to be robust and compact. However, *celC*$^-$ cells were easily removed with each successive washing step, whereas almost all of the wild-type cells remained attached.

We also used a sand assay for monitoring biofilm development because sand is a natural environment for many bacteria and also because it is less complex than soil. This enabled us to test potential biofilm mutants obtained in the plate assay in a defined, but natural substrate. PVC tabs were also used to examine the three-dimensional structure of biofilms formed at the edges of inert surfaces. Biofilm growth on sand (fig. 4, A-G) and on plastic tabs (fig. 4, H-J) in static cultures was monitored by fluorescence microscopy.

A week after inoculation, we took aliquot equivalent of grains of sand to determine their ability to form biofilms by observing the particles under fluorescence microscopy. These experiments showed significant differences, which agree with the observations in figure 3, regarding the adhesion capacities of the different strains under study (fig. 4 A-C). On sand, wild-type ANU843 developed microcolonies, which progress into a characteristic three-dimensional biofilm (fig. 4A). In contrast, ANU843C2$^-$ established unstable aggregates on surfaces that were easily removed (fig. 4B), explaining why a significantly fewer number of bacteria were attached in both the microtiter plate wells (fig. 3) and in the sand test. Interestingly, ANU843C2$^+$ did not form any structures on the sand surface.

Mutants were tested for biofilming ability on PVC tabs by confocal imaging of GFP-labelled rhizobia. Wild-type mature biofilms consisted of distinct towers and ridges by the termination of the experiment (Fig. 4H). This biofilm morphology sharply contrasted with that produced by the *celC*$^-$ (fig. 4I) mutant in which only thick, tightly appressed

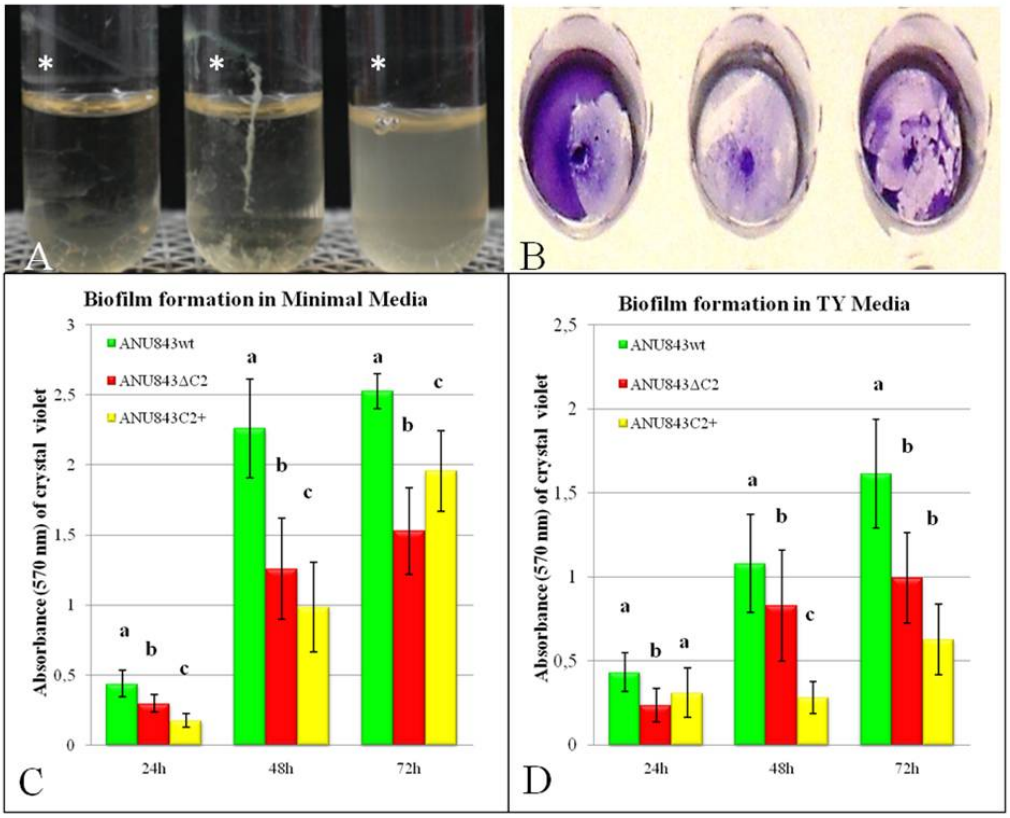

Fig. 3: Biofilm-forming capacity on glass (A) and polystyrene plates (B-D) by ANU843 (wild-type strain) and its derivatives ANU843ΔC2 and ANU843C2^{+} (from left to right). A) Formation of a biofilm ring (*) in shaken glass tubes. B) Microtiter assay test wells after staining the bacteria forming a film on the bottom of the wells with crystal violet (cv). C-D) The data show absorbance values obtained in minimal media and TY at different times after inoculation. The culture medium of the static culture was removed and differences in biofilm matrix development were observed by cv staining. Before that, the OD_{600} was measured in a Microtiter Plate reader to verify that no differences in growth rate among the wells had occurred. Each data point is the average of at least 20 wells. Error bars indicate the standard deviation from the mean. Each experiment was repeated at least three times. For each experiment and time, values followed by the same letter do not differ significantly according to Fisher protected LSD test at P = 0.01. In both media, we found significant differences in biofilm formation among the strains under study, although in complex medium, these differences were less noticeable. A significant reduction in biofilm-forming capacity by both mutant strains ANU843ΔC2 and ANU843C2^{+} was observed.

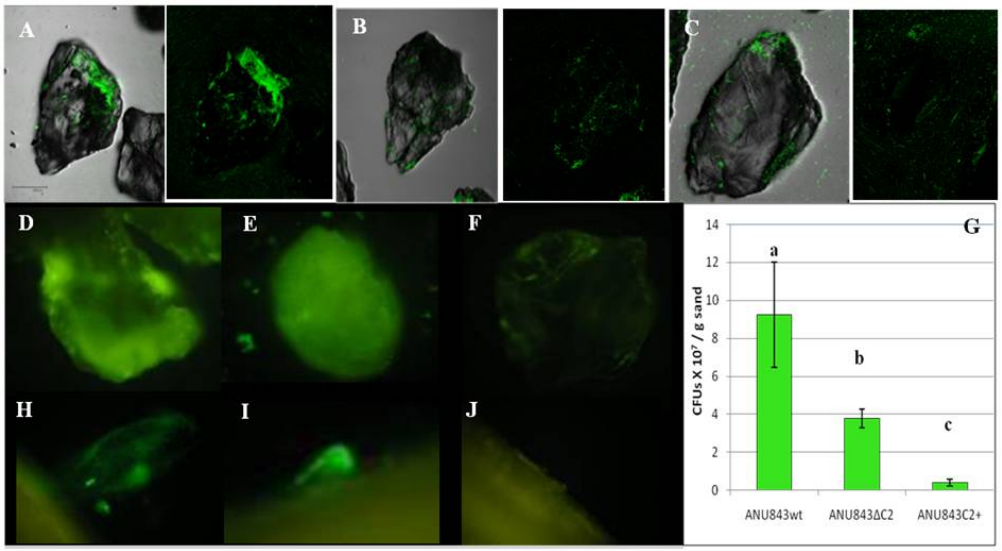

Fig. 4: Test for adhesion of the studied strains marked with GFP to sand (A- G) and biofilm formation on PVC tabs (H- J). The wild-type strain forms well-defined three-dimensional structures (A, D, H) whereas ANU843ΔC2 establishes a thin, weak layer that coats the sand grains (B, E, I), and ANU843C2^{+} barely adheres (C, F, J). G) We also evaluated attachment quantitatively, by counting cfus of bacteria attached to sand grains. Error bars indicate the standard deviation from the mean. Each experiment was repeated three times. Values followed by the same letter do not differ significantly according to the Fisher protected LSD test at P = 0.01.

mounds of cells that did not bind to the surface were apparent. On the other hand, *celC*$^{+}$ mutant bacteria did not establish a mature biofilm, just a monolayer with few bacteria attached to one another (fig. 4J).

This experiment confirms the role of *celC* gene in cellulose biosynthesis and the modulation of the biofilm structure by cellulose. It further suggests that a *Rhizobium leguminosarum* bv. trifolii biofilm does not develop normally if cellulose microfibrils are not cleaved into shorter lengths by *celC*.

Cellulose microfibrils are involved in biofilm formation *in planta*. Rhizobia are capable of colonizing the plant roots and generally they attach to the roots of legume hosts better than to those of non-legume roots (34). We examined attachment and biofilm formation of *Rhizobium leguminosarum* bv. trifolii and the derivative strains on roots and root hair tips of the host plant *Trifolium repens* in Fähraeus medium to extend the results found in the abiotic surface tests. We found that all strains colonized the root, especially root hair zone II (35), and remained firmly attached to the root surface even after thorough washing (fig. 5). Even the cellulose mutant (ANU843C2^{+}) attached to plant roots, demonstrating that cellulose is not required for plant root attachment.

In ANU843, the majority of root hairs in the growing root hair zone were covered with *R. leguminosarum* bv. trifolii cells growing in a three-dimensional biofilm (fig. 5A, D) whereas ANU843C2^{-} formed large aggregates that covered the root in an irregular fashion (fig. 5B,E). In contrast, very few individual cells of the ANU843C2^{+} mutant (cellulose-) were attached to root hairs (fig. 5C,F).

As observed previously (24), there was attachment of individual GFP-labeled *R. leguminosarum* bv. trifolii ANU843 wild-type and strain ANU843C2^{-} bacterial cells, followed by "cap" formation, in which additional bacterial cells attach to bacteria that directly bind to root hairs (fig. 5A, D,H and B,E,I). However, ANU843C2^{-} cells more often clustered in discrete clumps on the root surface (fig. 5B,E) and were also closely associated with root hairs (fig. 5I). In contrast, ANU843C2^{+} mutant biofilm caps were absent from most of the root hairs (fig. 5). These results confirm that cellulose is not required for lectin-mediated attachment to root hairs, although its presence is important for three-dimensional biofilm formation (23). The differences in root and root hair attachment seen with the cellulose mutants implies that the structure of the root hair "cap" biofilm is a three-dimensional structure in which cellulose fibrils play a significant role.

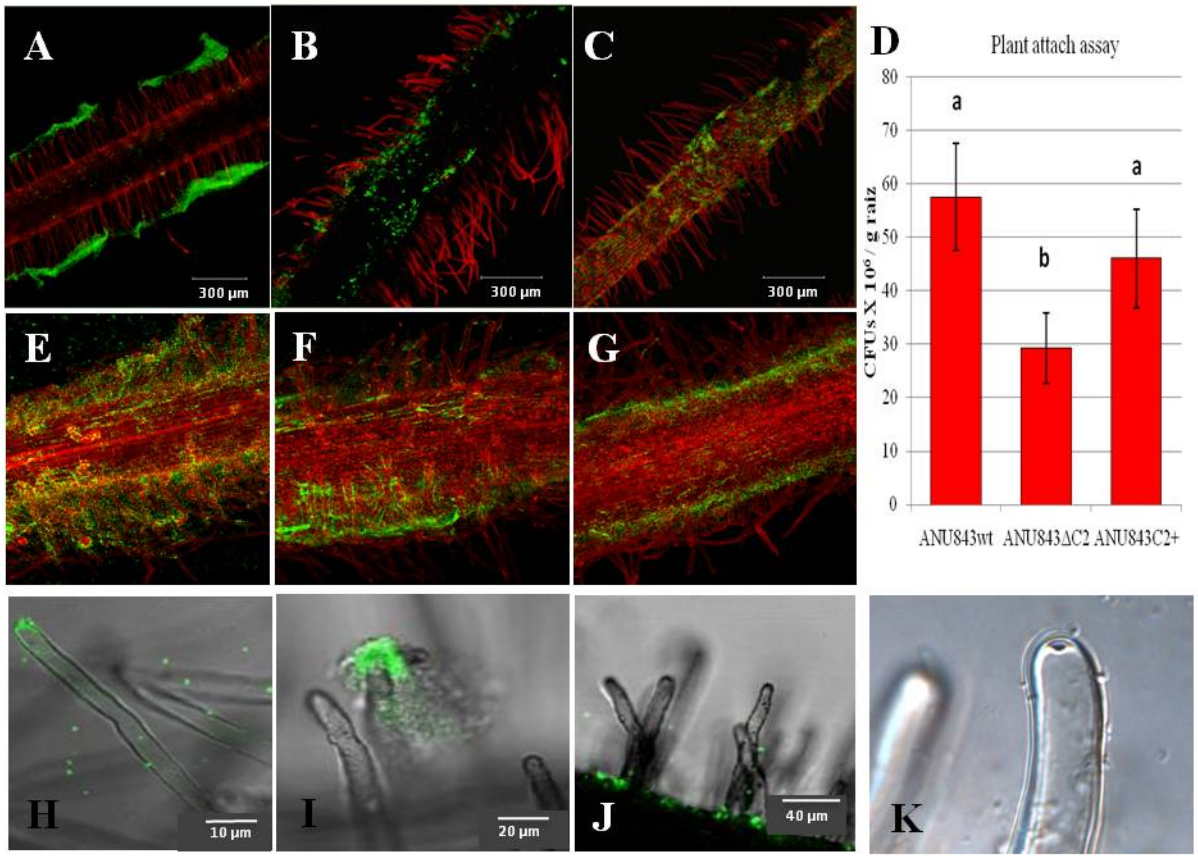

Fig. 5. . Plant attachment assays using *Trifolium repens* to study the ability of the strains under study to form biofilms. (A-C, E-G) Confocal laser scanning microscopy of propidium iodide-stained roots inoculated with gfp-tagged ANU843 and its derivatives showing biofilm formation along the root surface at 10X (A-C) or 20X (E-G). D) Number of colony-forming units (cfu) per gram of root tissue after washing and then sonicating the roots. Each data point is the average of at least 9 determinations. Error bars indicate the standard error from the mean. Root biofilms were harvested 72 h post-inoculation with either wild-type or *celC* mutant bacteria. Fluorescence (H-J) and optical (K) microcopy shows root hair colonization in detail. The wild-type strain forms three-dimensional biofilms (A) that cover both root surface and root hairs (E) forming distinct "caps" on the tip (H). In contrast, ANU843C2^{-} establishes aggregates that cover the root irregularly (B, F) and forms a thicker cap on the root hairs (I) whereas ANU843C2^{+} appears to coat the root surface (C, G, J) and cap formation does not occur. Nevertheless, sufficient adhesion of individual bacteria occurs on the tip to produce the hot (hole on the tip) phenotype (K).

Clover roots were also harvested and sonicated for colony counts as previously described. For strain ANU843, an average of 5.8 x 10^7 rhizobia cells per mg root fresh weight were attached to the root. By contrast, only 3 x 10^7 ANU843ΔC2 cells per mg root fresh weight attached to clover roots, showing that, as occurred in the abiotic surface experiments, the unstable aggregates on surfaces are easily removed (fig. 5G). Regarding ANU843C2$^+$, 4.6 x 10^7 cells per mg of root fresh weight were attached, showing that the inability to form three-dimensional structures led to loose bacterial attachment. However, no significant differences with respect to ANU843 were detected in terms of CFU/mg. This could be related to the involvement of the plant in bacterial adhesion to the root surface but not to the root hair tips.

To test this, we analyzed by scanning electron microscopy the effect of overproduction of cellulase CelC2 in adhesion of these mutants to the root surface. Figure 6 shows that most of the cells corresponding to the wild type are attached to the root surface and surrounded by cellulose microfibrils (82%) (fig. 6 A ,C). However, samples inoculated with the CelC2 overproducing strain had fewer bacteria coated with cellulose microfibrils (50%) (fig. 6 BD). It remains to be determined if this difference is due to cellulose production by the plant or to the regulation of the bacterial cellulose by the plant, perhaps by inhibiting the enzyme (CelC2) or by repressing expression of the *celC* gene.

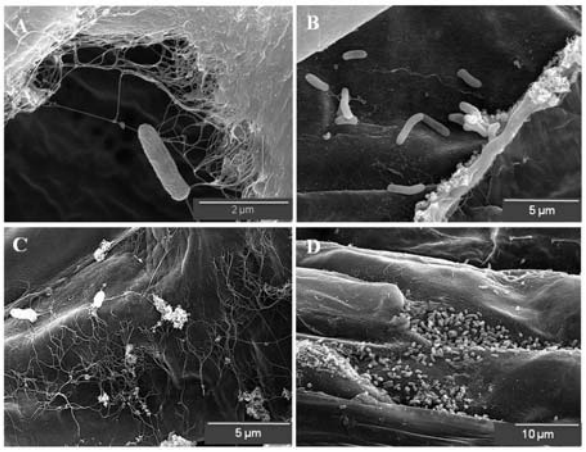

Fig. 6.- Scanning Electron Microscopy (SEM) of clover roots inoculated with the strains under study: (A,C) ANU843 wild-type strain and (B, D) ANU843C2$^+$.

Discussion

Cellulose production in bacteria is of potential economical interest because of its impact in medical settings and in the paper and food industry (18). In addition, cellulose production in a wide variety of bacteria has increased the possibility for the elucidation of regulation and molecular mechanisms of cellulose biosynthesis, which still remains largely unknown, and for the role of cellulose in attachment and biofilm formation during the interaction of the bacteria with the environment and a potential host.

Although Congo Red has been described as being able to also bind outer membrane proteins in some animal pathogens (36), several authors have noted that among *Rhizobium* spp., staining with this dye can be related to the content of cellulose in the bacterial cultures (14). The results obtained in our work, which analyzed mutants impaired in cellulose production (fig. 1B), seem to confirm that *Rhizobium* does not produces other substances that bind Congo Red.

The ability of almost all tested microorganisms capable of forming nodules to synthesize cellulose seems to support the importance of this polymer in the survival of these bacteria in the environment. Furthermore, although the inactivation of cellulose biosynthesis in R. leguminosarum bv. trifolii does not affect the ability to nodulate clover in controlled laboratory conditions (14, 25), the ability of rhizobia to adhere firmly to a substrate using cellulose microfibrils facilitates host root colonization (6-7). Moreover, it is very possible that cellulose microfibril-mediated firm adhesion during host root colonization is important under natural conditions in the rhizosphere soil, where bacteria have to compete for survival to gain access to plant carbon sources.

During the course of our research on the CelC2 protein, we observed that cellulose microfibrils produced by the *celC*$^+$ overexpression mutant were undetectable by Calcofluor staining, and that the microfibrils of the *celC*$^-$ deletion mutant were significantly longer than those seen in wild type, suggesting that *R. leguminosarum* bv. trifolii cellulose is cleaved by CelC2.

In the model proposed for the synthesis of cellulose in plants (37), CesA protein catalyzes ß-1,4-glucan chain elongation by transferring UDP-glucose moieties to the sitosterol-ß-sitosterol-glucoside intermediary, forming celodextrines. The other protein that has been proposed to participate in this process is Korrigan (Kor) cellulase, which may act to release the sitosterol-ß-glucoside from the cellulose polymer chain formed (fig. 7). How the cellulose chain is finally formed and which molecule(s) intervene in this last step remains unknown. However, in the model for cellulose synthesis in bacteria (21), glucose monomers are added to the intermediary lipid-glucose resulting from CelA activity to form lipid-glucose$_x$ (x = 2-4 glucose). This protein is homolog to CesA, whereas Korrigan cellulase has homology to the cellulase encoded by the *celC* gene, which, according to Mathysse, may act in this case as a transferase, by transferring glucose moieties from the lipooligosacharide to the cellulose polymer chain being formed, instead of as an endoglucanase.

According to our findings, the role of CelC cellulase in cellulose biosynthesis is to catalyze the hydrolysis reaction, in which 3-glucose moieties are released from the lipid-4-glucose intermediary to form the cellulose microfibril, and which may be mediated by CelA acting again as a glycosyl transferase. Thus, in the absence of CelC protein, the chain grows longer whereas under conditions of cellulase overproduction, no fibrils are formed.

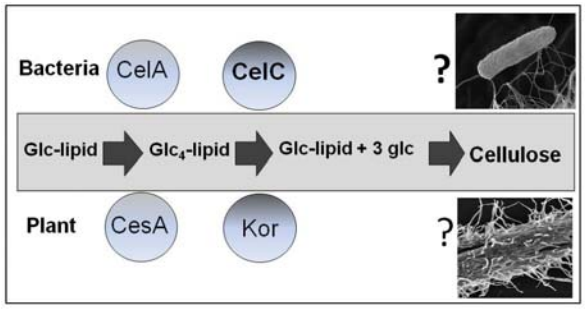

Fig. 7: Model of cellulose biosynthesis in Bacteria and Plant.

This function for the *celC* gene is distinctly different from its established role as an endoglucanase involved in the infection process (24). We propose that cellulose production may reflect an earlier evolutionary development because this property is encoded by an operon common to all nodulating rhizobia, including *Burkholderia* and *Cupriavidus* strains, the so-called β-rhizobia (38), whereas the cellulase *celC* gene function for the infection developed later. The fact that core *celC* has both colonization and infection functions, i.e. in cellulose biosynthesis and independently as a molecule leading to *Rhizobium leguminosarum* bv. trifolii penetration into the root hair and liberation in the symbiosomes, implies the likely existence of two different sets of control mechanisms.

The role of cellulose in *Rhizobium* biofilm establishment has been unrecognized not only in part because of the lack of mutants that are impaired in cellulose microfibril synthesis as

well as cellulose overproducing mutants, but also methods that precisely measure biofilming ability (23, 30). To our knowledge, this is the first time that *Rhizobium* mutants either lacking and overexpressing cellulose have been analysed with respect to attachment to abiotic surfaces, root colonization, and biofilm formation.

Our results show that these three parameters were markedly reduced in *celC*$^-$ and *celC*$^+$ mutants in that the mutants exhibited significantly reduced (25–50%) in biofilm formation, not only on abiotic surfaces (PVC microtiter plate assays (fig. 3) and on sand (fig. 4), but also on plant roots (fig. 5). Microtiter well and sand attachment assays of gfp-tagged bacterial cultures confirmed the role of cellulase in affecting biofilm structure. We propose that cellulase CelC is associated with cellulose cleavage and processing and that the quantity of cellulose is essential for the three-dimensional structure of a mature biofilm of *Rhizobium leguminosarum* bv. trifolii.

Matherials and methods

Bacterial strains, plasmids and growth conditions. Bacterial strains and plasmids used are described in Table 3. R. leguminosarum strains were grown at 28°C in TY (40), YMA (41) or minimal medium containing 1% mannitol (42) (MM) or 0.5% inositol (BINOS) (43) as the carbon source. Escherichia coli was grown at 37°C in LB medium (44). These media were supplemented with kanamycin (50 µg/ml) or tetracycline (10 µg/ml) as required. Plasmid pHC60 [45] was introduced via biparental mating to yield GFP-tagged bacteria.

Cellulose detection assays. Cellulose production was qualitatively assayed by growing the strains under study in YMA plates containing 25 mg/l Congo red for 5 days to stain for cellulose. For visualization of cellulose microfibrils, cells grown on plates were suspended in 0.025% Calcofluor (Sigma), placed onto teflon printed slides and observed under a fluorescence microscope. Aggregation assays were performed basically as described in (39). Overnight cultures were adjusted to an OD_{600} of 0.6 by dilution with YMB medium. Three-ml aliquots of cultures were incubated in 10 ml standing tubes at room temperature, and the OD_{600} of the upper parts of the cultures was measured after 24 h. These tests were performed in triplicate.

Determination of biofilm formation. Rings of biofilm at the air-liquid interface were qualitatively scored after 3 to 5 days of bacterial growth in 100 ml of TY medium in a 250-ml conical flask shaken at 300 rpm in an orbital shaker. *In vitro* biofilms were established as described earlier (3). Basically, to prepare the initial inocula, gfp-tagged bacteria were grown in the corresponding liquid medium for 2 days (optical density at 600 nm [OD_{600}], to approximately 2.0). They were then washed and diluted in the same medium to OD_{600} = 0.2 (ca. $1x10^7$ cells/ml) for all the subsequent assays. Bacterial attachment to PVC plates was assayed by pipeting 100 µl of this culture into individual PVC wells in a 96-well plate (Falcon 3911, Becton Dickinson, Franklin Lakes, NY). The plates were sealed with parafilm and incubated at 28ºC. At defined times, the growth media and unbound cells were removed, and attached bacteria were quantified by staining them with 0.3% (wt/vol) Crystal Violet (CV) (Sigma) for 10 min. After the excess dye was washed away and the wells allowed to dry, the remaining dye was solubilized with 80% ethanol-20% acetone and quantified by measuring the absorbance at 570 nm in an ASYS (Biochrom, UK) Microtiter Plate reader (Model No. UVM340).

Biofilm formation on abiotic surfaces was assayed by transferring 0.5 ml of each strain's culture to individual wells of a 24-well Costar PS (polystyrene) plate containing either a PVC tab or sterile sand. After 1 week at 28ºC, either the tabs or 100 µl of sand were examined under a fluorescence microscope to see the three-dimensional architecture of the biofilms. Sand particles were also washed and attached cells were dissociated in an Elmasonic (Singen, Germany) sonication bath and quantified by counting cfus (normalized to the amount of sand weighed).

Root biofilms were prepared by dipping one-week-old seedling roots into a *Rhizobium* suspension of OD_{600} = 0.02 (ca. 1 x 10^6 cells/ml) for 5 min. The seedlings were then grown on a Whatman paper support placed inside glass tubes containing 20 ml Fähraeus medium without nitrogen. The attachment of the bacteria to the root and the root hairs was evaluated using confocal microscopy at different time points. To count the number of attached rhizobia three days after inoculation, the roots were washed three times with sterile water under vigorous shaking to remove loosely associated cells, then dried on sterile Whatman paper, and weighed. The dried roots were immersed in 1 ml of sterile TE buffer and sonicated to release attached cells, which were quantified by colony counting (normalized to the roots weigh). For each strain tested, we counted the attachment to at least four roots. All binding experiments were repeated at least three times.

Microscopy. Calcofluor-stained bacteria were observed with a Zeiss fluorescence microscope with a 366 nm UV light source. Scanning electron microscopy (SEM) was performed as previously described (40), except that the samples were coated with osmium instead of gold.

GFP-tagged rhizobia growing on sand particles or plastic tabs were aseptically removed from the wells, washed, and placed into the wells of depression slides, topped with a coverslip, and examined under epifluorescence using a Zeiss Axiophot microscope in conjunction with fluorescence microscopy.

Bacterial attachment to the root hairs was observed by confocal scanning laser microscopy with a Leica microscope using 488-nm argon laser excitation and a 500-nm long-pass emission filter, which allowed observation of GFP-labeled bacteria, and phase optics, which showed root hairs. Images were processed using Leica confocal software. GFP-labelled biofilms formed on roots were examined on a Zeiss LSM510 confocal microscope. Roots were stained with 10 µl of propidium iodide (Sigma). Projections were made from adjusted individual channels in the image stacks using Zeiss LSM510 imaging software.

Acknowledgements. We thank Nancy Fujishige and Peter De Hoff for help in biofilm assays and Luis Rey for providing pHC60. This work was supported by grants AGL 2005-07796 and AGL 2008-03360 and M. R. by a PhD fellowship both from Spanish government.

REFERENCES

1. Ausmees, N., K. Jacobsson, and M. Lindberg, A unipolarly located, cell-surface-associated agglutinin, RapA, belongs to a family of *Rhizobium*-adhering proteins (Rap) in *Rhizobium leguminosarum* bv. trifolii. Microbiology, 2001. 147(Pt 3): p. 549-59.
2. De Hoff, P.L., L.M. Brill, and A.M. Hirsch, Plant lectins: the ties that bind in root symbiosis and plant defense. Mol Genet Genomics, 2009. 282(1): p. 1-15.
3. Fujishige, N.A., N.N. Kapadia, P.L. De Hoff, and A.M. Hirsch, Investigations of *Rhizobium* biofilm formation. FEMS Microbiol Ecol, 2006. 56(2): p. 195-206.
4. Davey, M.E. and G.A. O'Toole, Microbial biofilms: from ecology to molecular genetics. Microbiol Mol Biol Rev, 2000. 64(4): p. 847-67.
5. Smit, G., S. Swart, B.J. Lugtenberg, and J.W. Kijne, Molecular mechanisms of attachment of *Rhizobium* bacteria to plant roots. Mol Microbiol, 1992. 6(20): p. 2897-903.
6. Dazzo, F.B., G.L. Truchet, J.E. Sherwood, E.M. Hrabak, M. Abe, and S.H. Pankratz, Specific phases of root hair attachment in the *Rhizobium trifolii*-clover symbiosis. Appl Environ Microbiol, 1984. 48(6): p. 1140-50.
7. Mateos, P.F., et al., Direct *in situ* identification of cellulose microfibrils associated with *Rhizobium leguminosarum* biovar. trifolii attached to the root epidermis of White clover. Can J Microbiol, 1995. 41: p. 202-207.
8. Webb, J.S., M. Givskov, and S. Kjelleberg, Bacterial biofilms: prokaryotic adventures in multicellularity. Curr Opin Microbiol, 2003. 6(6): p. 578-85.
9. Matthysse, A.G., K. Deschet, M. Williams, M. Marry, A.R. White, and W.C. Smith, A functional cellulose synthase from ascidian epidermis. Proc Natl Acad Sci U S A, 2004. 101(4): p. 986-91.

10. Ross, P., et al., Regulation of cellulose synthesis in Acetobacter xylinum by cyclic diguanylic acid. Nature, 1987. 325(6101): p. 279-81.
11. Nobles, D.R., D.K. Romanovicz, and R.M. Brown, Jr., Cellulose in cyanobacteria. Origin of vascular plant cellulose synthase? Plant Physiol, 2001. 127(2): p. 529-42.
12. Ross, P., R. Mayer, and M. Benziman, Cellulose biosynthesis and function in bacteria. Microbiol Rev, 1991. 55(1): p. 35-58.
13. Lasa, I., Towards the identification of the common features of bacterial biofilm development. Int Microbiol, 2006. 9(1): p. 21-8.
14. Ausmees, N., H. Jonsson, S. Hoglund, H. Ljunggren, and M. Lindberg, Structural and putative regulatory genes involved in cellulose synthesis in *Rhizobium leguminosarum* bv. trifolii. Microbiology, 1999. 145 (Pt 5): p. 1253-62.
15. Matthysse, A.G., S. White, and R. Lightfoot, Genes required for cellulose synthesis in *Agrobacterium tumefaciens*. J Bacteriol, 1995. 177(4): p. 1069-75.
16. Wong, H.C., et al., Genetic organization of the cellulose synthase operon in *Acetobacter xylinum*. Proc Natl Acad Sci USA, 1990. 87(20): p. 8130-4.
17. Lin, F.C., R.M. Brown, Jr., R.R. Drake, Jr., and B.E. Haley, Identification of the uridine 5'-diphosphoglucose (UDP-Glc) binding subunit of cellulose synthase in *Acetobacter xylinum* using the photoaffinity probe 5-azido-UDP-Glc. J Biol Chem, 1990. 265(9): p. 4782-4.
18. Romling, U., Molecular biology of cellulose production in bacteria. Res Microbiol, 2002. 153(4): p. 205-12.
19. Saxena, I.M., K. Kudlicka, K. Okuda, and R.M. Brown, Jr., Characterization of genes in the cellulose-synthesizing operon (acs operon) of *Acetobacter xylinum*: implications for cellulose crystallization. J Bacteriol, 1994. 176(18): p. 5735-52.
20. Doblin, M.S., I. Kurek, D. Jacob-Wilk, and D.P. Delmer, Cellulose biosynthesis in plants: from genes to rosettes. Plant Cell Physiol, 2002. 43(12): p. 1407-20.
21. Matthysse, A.G., D.L. Thomas, and A.R. White, Mechanism of cellulose synthesis in *Agrobacterium tumefaciens*. J Bacteriol, 1995. 177(4): p. 1076-81.
22. Laus, M.C., A.A. van Brussel, and J.W. Kijne, Role of cellulose fibrils and exopolysaccharides of *Rhizobium leguminosarum* in attachment to and infection of *Vicia sativa* root hairs. Mol Plant Microbe Interact, 2005. 18(6): p. 533-8.
23. Williams, A., A. Wilkinson, M. Krehenbrink, D.M. Russo, A. Zorreguieta, and J.A. Downie, Glucomannan-mediated attachment of Rhizobium leguminosarum to pea root hairs is required for competitive nodule infection. J Bacteriol, 2008. 190(13): p. 4706-15.
24. Robledo, M., *et al., Rhizobium* cellulase CelC2 is essential for primary symbiotic infection of legume host roots. Proc Natl Acad Sci U S A, 2008. 105(19): p. 7064-9.
25. Robledo, M., *et al.,* Development of functional symbiotic white clover root nodules requires tightly regulated production of rhizobial cellulase CelC2. In preparation.
26. Young, J.P., *et al.*, The genome of *Rhizobium* l*eguminosarum* has recognizable core and accessory components. Genome Biol, 2006. 7(4): p. R34.
27. González, V., et al., The partitioned *Rhizobium etli* genome: genetic and metabolic redundancy in seven interacting replicons. Proc Natl Acad Sci USA, 2006. 103(10): p. 3834-9.
28. Schmeisser, C., *et al., Rhizobium* sp. strain NGR234 possesses a remarkable number of secretion systems. Appl Environ Microbiol, 2009. 75(12): p. 4035-45.
29. Goodner, B., et al., Genome sequence of the plant pathogen and biotechnology agent *Agrobacterium tumefaciens* C58. Science, 2001. 294(5550): p. 2323-8.
30. Deinema, M.H. and L.P. Zevenhuizen, Formation of cellulose fibrils by gram-negative bacteria and their role in bacterial flocculation. Arch Mikrobiol, 1971. 78(1): p. 42-51.
31. Matthysse, A.G., *et al.,* The effect of cellulose overproduction on binding and biofilm formation on roots by *Agrobacterium tumefaciens*. Mol Plant Microbe Interact, 2005. 18(9): p. 1002-10.
32. Zogaj, X., M. Nimtz, M. Rohde, W. Bokranz, and U. Romling, The multicellular morphotypes of *Salmonella typhimurium* and *Escherichia coli* produce cellulose as the second component of the extracellular matrix. Mol Microbiol, 2001. 39(6): p. 1452-63.
33. Saldana, Z., *et al*., Synergistic role of curli and cellulose in cell adherence and biofilm formation of attaching and effacing *Escherichia coli* and identification of Fis as a negative regulator of curli. Environ Microbiol, 2009. 11(4): p. 992-1006.
34. Solano, C., et al., Genetic analysis of *Salmonella enteritidis* biofilm formation: critical role of cellulose. Mol Microbiol, 2002. 43(3): p. 793-808.
35. Fujishige, N.A., N.N. Kapadia, and A.M. Hirsch, A feeling for the micro-organism: structure on a small scale. Biofilms on plant roots. Botanical Journal of the Linnean Society, 2006. 150: p. 79-88.
36. Gage, D.J., Infection and invasion of roots by symbiotic, nitrogen-fixing rhizobia during nodulation of temperate legumes. Microbiol Mol Biol Rev, 2004. 68(2): p. 280-300.
37. Theisen, M., C.R. Rioux, and A.A. Potter, Molecular cloning, nucleotide sequence, and characterization of *lppB,* encoding an antigenic 40-kilodalton lipoprotein of *Haemophilus somnus*. Infect Immun, 1993. 61(5): p. 1793-8.
38. Peng, L., Y. Kawagoe, P. Hogan, and D. Delmer, Sitosterol-beta-glucoside as primer for cellulose synthesis in plants. Science, 2002. 295(5552): p. 147-50.
39. Moulin, L., A. Munive, B. Dreyfus, and C. Boivin-Masson, Nodulation of legumes by members of the beta-subclass of Proteobacteria. Nature, 2001. 411(6840): p. 948-50.
40. Beringer, J.E., R factor transfer in *Rhizobium leguminosarum*. J Gen Microbiol, 1974. 84(1): p. 188-98.
41. Somasegaran, P. and H.J. Hoben, Handbook for rhizobia: methods in Legume-*Rhizobium* technology. 1994, New York: Springer-Verlag. xvi, 450 p.
42. O'Gara, F. and K.T. Shanmugam, Regulation of nitrogen fixation by Rhizobia. Export of fixed N_2 as NH^{+4}. Biochim Biophys Acta, 1976. 437(2): p. 313-21.
43. Mateos, P.F., *et al.*, Cell-associated pectinolytic and cellulolytic enzymes in *Rhizobium leguminosarum* biovar trifolii. Appl Environ Microbiol, 1992. 58(6): p. 1816-22.
44. Sambrook, J., E.F. Fritsch, and T. Maniatis, Molecular cloning: a laboratory manual. 2nd ed. ed. 1989, Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory.
45. Cheng, H.P. and G.C. Walker, Succinoglycan is required for initiation and elongation of infection threads during nodulation of alfalfa by *Rhizobium meliloti*. J Bacteriol, 1998. 180(19): p. 5183-91.
46. Rolfe, B.G., P.M. Gresshoff, and J. Shine, Rapid screening method for symbiotic mutants of *Rhizobium leguminosarum* biovar trifolii on white clover plants. Plant Sci. Lett., 1980. 19: p. 277-284.
47. Robledo, M., *et al*., Development of functional symbiotic white clover root nodules requires tightly regulated production of rhizobial cellulase CelC2. In Proccess.
48. Da Re, S., B. Le Quere, J.M. Ghigo, and C. Beloin, Tight modulation of *Escherichia coli* bacterial biofilm formation through controlled expression of adhesion factors. Appl Environ Microbiol, 2007. 73(10): p. 3391-403.
49. Chao, Y., T. Ishida, Y. Sugano, and M. Shoda, Bacterial cellulose production by *Acetobacter xylinum* in a 50-L internal-loop airlift reactor. Biotechnol Bioeng, 2000. 68(3): p. 345-52.
50. Mateos, P.F., *et al.,* Erosion of root epidermal cell walls by *Rhizobium* polysaccharide-degrading enzymes as related to primary host infection in the *Rhizobium*-legume symbiosis. Can J Microbiol 2001. 47: p. 475-487.

SUPLEMENTAL INFORMATION

Table 1. Cellulose production by representative type strains of root-nodule legume symbionts and related detected by Congo Red staining. +++/++/+: positive to different degree w: weakly positive. "Classical rhizobia" species are in bold, the rest are non-rhizobia species that are able to nodule. * Species whose nodulation capacity has not been described, including in genera traditionally considered as "rhizobia".

GENUS	TYPE STRAIN	LEGUME HOST	CELLULOSE
Rhizobium	***R . cellulosilyticum* ALA10B2 ^{T}**	*Medicago*	+++
	***R. etli* CFN 42^{T}**	*Phaseolus*	+++
	***R. galegae* ATCC 43677^{T}**	*Galega*	+++
	***R. gallicum* R602sp^{T}**	*Phaseolus*	+
	***R. giardinii* H152^{T}**	*Phaseolus*	+++
	***R. hainanense* I66^{T}**	*Desmodium*	+
	***R. huautlense* SO2^{T}**	*Sesbania*	+
	***R. indigoferae* CCBAU 71042^{T}**	*Indigofera*	+
	***R. leguminosarum* ATCC10004^{T}**	*Pisum*	+
	***R. leguminosarum* bv. trifolii**	*Trifolium*	++
	***R. loessense* CCBAU 7190B^{T}**	*Astragalus*	++
	***R. lusitanum* P1-7^{T}**	*Phaseolus*	+
	***R. mongolense* USDA 1844^{T}**	*Medicago*	+++
	***R. radiobacter* ATCC 19358^{T}**	*	+
	***R. rhizogenes* ATCC 11325^{T}**	*Phaseolus*	++
	***R. sullae* IS123 ^{T}**	*Hedysarum*	+
	***R. rubi* ATCC 13335^{T}**	*	++
	***R. phaseoli* ATCC 14482^{T}**	*Phaseolus*	+
	***R. pisi* DSM 30132^{T}**	*Pisum*	++
	***R. tropici* CIAT 899 ^{T}**	*Phaseolus*	+
	***R. yanglingense* CCBAU 71623**	*Amphicarpaea*	+
Sinorhizobium	***S. fredii* LMG 6217^{T}**	*Glycine*	+
	***S. kostiense* LMG 19227 ^{T}**	*Acacia*	w
	***S. meliloti* ATCC 9930 ^{T}**	*Medicago*	+
	***S. meliloti* 1021**	*Medicago*	w
	***S. medicae* LMG1037 ^{T}**	*Medicago*	+
	***S. medicae* WSM419**	*Medicago*	+
	***S. morelense* Lc04 ^{T}**	*	++
	***S. saheli* LMG7837 ^{T}**	*Acacia*	++
	***S. terangae* LMG6463 ^{T}**	*Acacia*	+
	***S. xinjiangense* LMG17930 ^{T}**	*Glycine*	+
Mesorhizobium	***M. amorphae* ACCC 19665 ^{T}**	*Amorpha*	++
	***M. chacoense* Pr5 ^{T}**	*Prosopis*	w
	***M. ciceri* USDA 3383 ^{T}**	*Cicer*	+
	***M. huakuii* USDA 4779 ^{T}**	*Astragalus*	+
	***M. loti* ATCC 33669 ^{T}**	*Lotus*	+
	***M. mediterraneum* USDA 3392^{T}**	*Cicer*	++
	***M. plurifarium* LMG 7836^{T}**	*Acacia*	+
	***M. septentrionale* HAMBI 2582 ^{T}**	*Astragalus*	+
	***M. tianshanense* LMG 18976 ^{T}**	*Sophora*	++
	***M. temperatum* HAMBI 2583 ^{T}**	*Astragalus*	w
Phyllobacterium	*P. trifolii* pETPO2 ^{T}	*Trifolium*	+
Bradyrhizobium	***B. betae* pl7Hg^{T}**	*	+
	***B. canariense* BTA1^{T}**	*Chamaecytisus*	+++
	***B. elkanii* LMG 6134^{T}**	*Glycine*	+
	***B. japonicum* LMG 6138^{T}**	*Glycine*	+
	***B. liaoningense* LMG 18230^{T}**	*Glycine*	+
	***B. yuanmingense* LMG 21827^{T}**	*Lespedeza*	+
Azorhizobium	***A. caulinodans* ORS 571^{T}**	*Sesbania*	+
Devosia	*D. neptuniae* J1^{T}	*Neptunia*	++
Ochrobactrum	*O. lupini* LUP21^{T}	*Lupinus*	+
	O. triciti LMG18957^{T}	*Cytisus*	+

Table 2. Putative rhizobial operons involved in cellulose biosynthesis located in the sequenced genome of rhizobia.

Organism	**Cellulose**	**Locus Tag***
Rhizobium leguminosarum bv. trifolii WSM1325	*celABC*	Rleg1293, Rleg1294, Rleg1295
	cellJK	Rleg43342, Rleg43343, Rleg4334
Rhizobium leguminosarum bv. trifolii WSM2304	*celABC*	Rleg21204, Rleg21205, Rleg21206
	cellJK	Rleg24003, Rleg24004, Rleg24005
Rhizobium leguminosarum bv. viciae 3841	*celABC*	RL1646, RL1647, RL1648
	cellJK	RL0079, RL0080, RL0081
Rhizobium etli CFN42	*celABC*	RHECH01542, RHECH01543, RHECH01544
	cellJK	RHECH00070,RHECH00071, RHECH00072
Rhizobium etli CIAT652	*celABC*	RHECIATCH0001610, RHECIATCH0001611,
	cellJK	RHECIATCH0000074,RHECIATCH0000075,RHECIAT
Rhizobium sp. NGR234	*celABC*	NGRb15360, NGRb15370, NGRb15380
	cellJK	NGRb17870NGRb17890NGRb17880
Agrobacterium tumefaciens C58	*celABC*	ATU3309, ATU08, ATU3307
	cellJK	ATU3314, ATU3313, ATU3312
Agrobacterium radiobacter K84	*celABC*	Arad9974, Arad9975, Arad9976
	cellJK	Arad7615,Arad7617,Arad7618
Sinorhizobium meliloti 1021	*cellJK*	SMb20460SMb20461 SMb20462
Sinorhizobium medicae WSM419	*celABC*	Smed5208, Smed5209, Smed5210
	cellJK	Smed3671, Smed3670, Smed3669
Mesorhizobium loti MAFF303099	*cellKJ*	mll7871, mll7872, mll7873

*Data from GenBank**.**

Table 3. Strains used in this study

Strains and plasmids	Description	Source
Rhizobium leguminosarum bv. trifolii		
ANU843	Wild-type Nod^+ Fix^+	(46)
ANU843ΔC2 (*celC*$^-$)	Wild-type containing a 361-bp deletion in *celC* gene	(24)
ANU843C2$^+$ (*celC*$^+$)	Wild-type containing plasmid pJZC2, carries *celC2*; Km^r	(47)
ANU843GFP	Wild type containing pHC60 for gfp expression; Tc^r	(47)
ANU843GFPΔC2	ANU843ΔC2 containing pHC60 for gfp expression; Tc^r	This work
ANU843GFPC2$^+$	ANU843C2$^+$ containing pHC60 for gfp expression; Km^r, Tc^r	(47)

CONCLUSIONS

1. The production of cell-bound enzymes with cellulase activity is a feature common to all nitrogen-fixing root-nodule legume endosymbionts. In general these rhizobial enzyme activities are relatively low as compared to pathogenic bacteria; exceptions are some representatives of the genus *Rhizobium*, *Azorhizobium* and *Phyllobacterium* which exhibited enhanced cellulase production.

2. In *Rhizobium leguminosarum* bv trifolii ANU843 the cellulase CelC2 is encoded by the chromosomal *celC* gene which is present as single copy in the genome and well conserved in rhizobia. The analysis of its inferred amino acid sequence revealed that this protein belongs to the glycosyl hydrolase family 8 and predicts that it is targeted to the periplasmic space to bind the outer bacterial membrane.

3. *celC*-based phylogeny of *Rhizobium* species is in good agreement with that determined by core housekeeping chromosomal genes but not with accessory genes phylogenies. This gene even allows better differentiation than currently used housekeeping genes among phylogenetically close related species. These findings render *celC* as a novel reliable taxonomic marker for the genus *Rhizobium*.

4. The CelC2 cellulase fulfils an essential role both in primary and secondary infection events required for the development of the canonical nitrogen-fixing *R. leguminosarum* bv. trifolii-white clover symbiosis. This enzyme specifically degrades the isotropic non-crystalline cellulose polymers, which distribution in white clover roots is restricted to the tips of both growing root hairs and infection threads, thereby creating the portal of entry of the microsymbiont into the plant root and facilitating the release of bacteria into the cells of developing nodules. However, CelC2 is dispensable for plant invasion through intercellular spaces or “cracks”.

5. Overexpression of the CelC2 cellulase either in its natural host *R. leguminosarum* bv. trifolli or in its non-host *Sinorhizobium meliloti* enhances rhizobial competitiveness and mediates bacteria root hairs invasion on clover, respectively.

6. A compatible canonical effective nodulation as assessed by early Nod factor signal transduction events, elicitation of plant defense responses, symbiosome development and nodule organogenesis requires tightly regulated production of CelC2 enzyme by the invading bacteria.

7. Cellulose biosynthesis is a trait common to all studied legume-nodulating bacteria. The organisms with a greater capacity to produce cellulose were generally found within the genera *Rhizobium* and *Mesorhizobium*, whereas reference strains from genus *Sinorhizobium* and *Bradyrhizobium* showed lower Congo Red affinity

8. The *celA* gene homologs coding for cellulose synthase of the glycosil transferase family localization upstream of *celC* gene ortologs is also a common feature in family *Rhizobiaceae* members and thus are predicted to form the *celABC* operon specifying cellulose biosynthesis, which indeed is transcribed as a single polycistronic mRNA in *Rhizobium leguminosarum* bv trifolii ANU843.

9. CelC2 plays also a role in cellulose biosynthesis by modulating the length of the cellulose fibrils which likely support *Rhizobium* attachment and biofilm formation to plant roots and abiotic surfaces.