Supporting Information

Robledo *et al.* 10.1073/pnas.0802547105

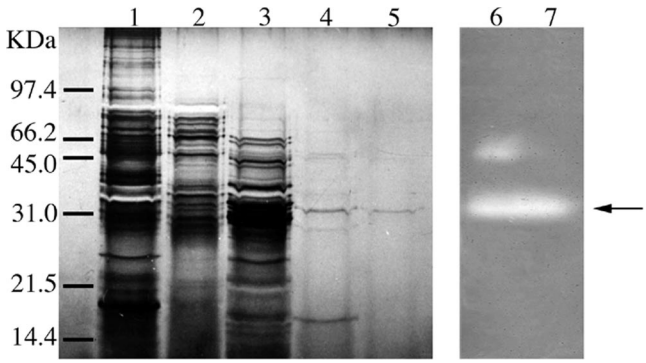

Fig. S1. SDS/PAGE of fractions at various steps in the purification of CelC2 cellulase made by *R. leguminosarum* bv. trifolii ANU843. Silver staining (lanes 1–5) and zymogram (lanes 6 and 7). Lanes: 1 and 6, sonicated extract; lane 2, DEAE Sepharose; lane 3, Sephacryl S-100; lane 4, Mono Q; lanes 5 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.

TNILSYASLLVVFVLLIGVTTSTMLKKLGRSMRRWRALLLAASVAVAPGLPA
TAQQAMINADAWSAYKAKFLDPSGRIVDNGNGNISHSEGQGYGLLLAYLSA
SPADFEQIWYFTRTELLLRDDGLAVWKWDPNVKPHVADTNNATDGDMLIA
YALALAGTAWKREDYILAASRMAQALLAETVGSSQGRTLLMPGTEGFTGS
DRDDGPVVNPSYWIYEAIPVMAALAPSDAWKKLSDDGVELLKTMQFGPRK
LPAEWVSLHDKPRPAEGFDAEFGYNAIRIPLYLARGGITDKALLVRLQKGM
SQDGVPATIDLTTGRPKTVLSDPGYRIVNDVVACVVDGTRLPSSALQFAPAL
YYPSTLQLLGLAYIGEKHPECL

Fig. S2. *In silico* amino acid sequence analysis of translated product obtained by DNA amplification of *R. 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.

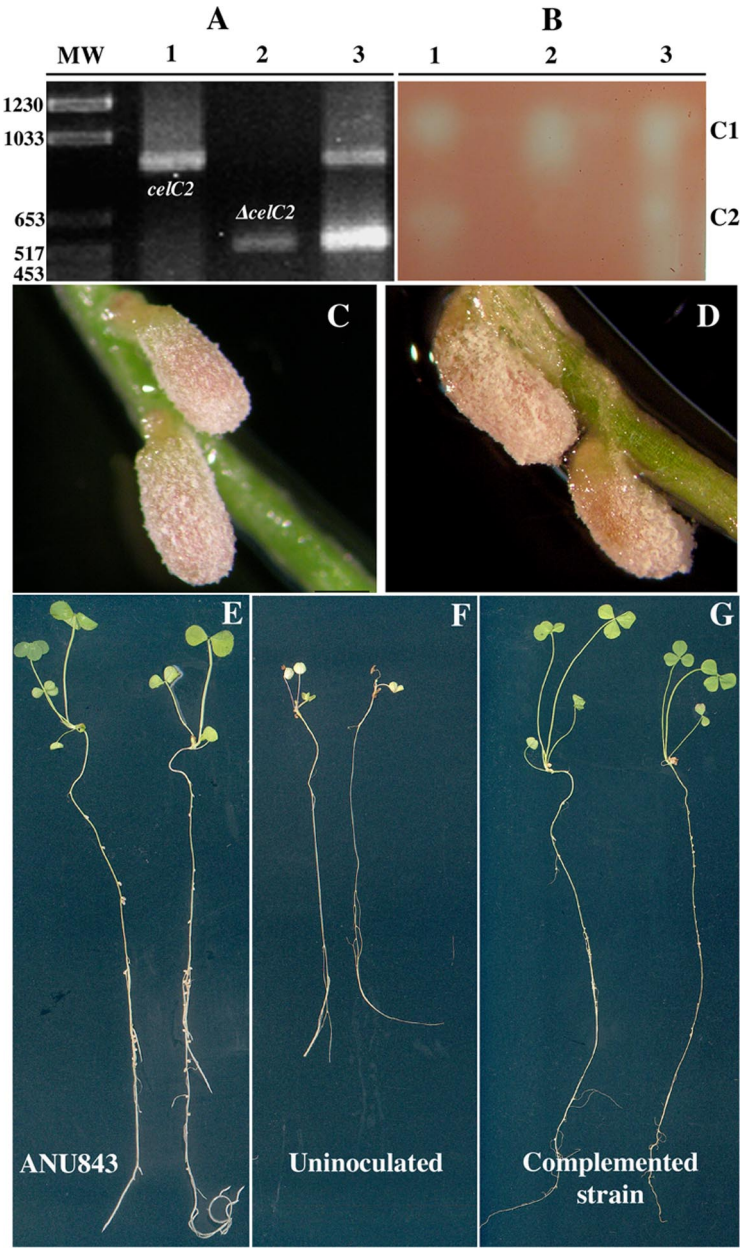

Fig. S3. Analysis of the *celC2*-complemented strain. (*A* and *B*) PCR amplification of DNA using primers C1F/C2R (*A*) and zymogram of CM-cellulase isozymes in sonicated cell extracts (*B*) from *R. legumuminosarum* bv trifolii ANU843 wt (lanes 1), ANU843Δ*celC2* (lanes 2), and *celC2* complemented strain (lanes 3). PCR products were analyzed by electrophoresis on a 1% agarose gel. MW, DNA molecular weight marker; fragment lengths in base pairs are indicated. (*C*–*G*) Typical nodules and whole phenotypes of white clover plants 40 days after inoculation with the wild-type ANU843 (*C* and *E*) or the *celC2*-complemented strain (*D* and *G*) and uninoculated plants (*F*).

Table S1. Cellulase production by representative type strains of root-nodule legume symbionts

Genus	Type strain	Nodulated legume host	Cellulase production*
Rhizobium	*R. etli* CFN 42^T	*Phaseolus*	+
	R. galegae ATCC 43677^T	*Galega*	+
	R. gallicum R602 sp^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

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

Table S2. Putative rhizobial cellulase genes

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

*Data from GenBank.

Table S3. Symbiotic phenotypes in white clover after inoculation with wild-type ANU843 or the CelC2 cellulase$^-$ mutant derivative Δ*celC2*

	Symbiotic phenotypes on white clover seedlings				
Inoculant strain	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

The results reported are the mean ± SD 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 (a or b) are not significantly different from each other at $P = 0.01$ according to Fisher´s Protected LSD (least-significant differences) test statistic.