Proc. Natl. Acad. Sci. USA
Vol. 87, pp. 2680–2684, April 1990
Microbiology

Proposed regulatory pathway encoded by the *nodV* and *nodW* genes, determinants of host specificity in *Bradyrhizobium japonicum*

(symbiosis/two-component regulatory system/signal transduction)

MICHAEL GÖTTFERT, PHILIPP GROB, AND HAUKE HENNECKE

Mikrobiologisches Institut, Eidgenössische Technische Hochschule, ETH-Zentrum, Schmelzbergstrasse 7, CH-8092 Zurich, Switzerland

Communicated by Jeff Schell, January 25, 1990

ABSTRACT ***Bradyrhizobium japonicum* is the root nodule endosymbiont of soybean (*Glycine max*), mung bean (*Vigna radiata*), cowpea (*Vigna unguiculata*), and Siratro (*Macroptilium atropurpureum*). We report the characteristics of a nodulation-gene region of *B. japonicum* that contributes only marginally to the bacterium's ability to nodulate soybean but is essential for the nodulation of the three alternative hosts. This DNA region consists of two open reading frames designated *nodV* and *nodW*. The predicted amino acid sequences of the NodV and NodW proteins suggest that they are members of the family of two-component regulatory systems, which supports the hypothesis that NodV responds to an environmental stimulus and, after signal transduction, NodW may be required to positively regulate the transcription of one or several unknown genes involved in the nodulation process. It seems likely that all host plants produce the necessary signal, whereas host specificity may be brought about by the product(s) of the gene(s) activated by NodW.**

A large number of rhizobial genes are involved in the establishment of a successful root nodule symbiosis with the legume host plants. Mutations within these genes often lead to nodules with a decreased amount of intracellular (endosymbiotic) bacteria. By contrast, mutations in nodulation (*nod*) genes have a strong impact on nodule initiation and nodule number but generally do not influence bacteroid development. The majority of nodulation genes characterized so far share a similar mode of transcriptional regulation in which the *nodD* gene product together with plant-derived flavonoid compounds acts as a transcriptional activator by binding to the *nod* box, a conserved promoter region upstream of *nod* genes (see ref. 1 for a recent review).

In *Bradyrhizobium japonicum* a nodulation gene region (tentatively called "*nod*-1") was identified previously (2) that appeared to be essential for the nodulation of various hosts but had little influence on the nodulation of soybean [*Glycine max* (L.) Merr.]. This region did not hybridize to a *nod*-box probe (3) and might thus represent a different class of nodulation genes. In this paper we present a detailed genetic description of this DNA region and propose a mode of its function that differs from all other nodulation genes identified thus far in rhizobial species.*

MATERIALS AND METHODS

Bacterial Strains and Growth Conditions. *B. japonicum* 110 *spc4* (4), referred to as wild type, and mutant derivatives were grown in PSY medium (4). Strain ΔE1-7d1 is a deletion mutant lacking the *nod*-1 region and adjacent DNA (2). *Escherichia coli* strains were grown in LB medium (5). *E. coli* RR28 (6) was the recipient in plasmid transformations.

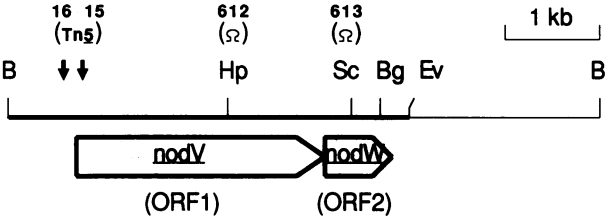

FIG. 1. Physical and genetic map of the *nodVW* region. A 6.2-kb *Bam*HI fragment was cloned in pBR322 to give plasmid pRJ4356. The sequenced 4227-base-pair *Bam*HI–*Eco*RV fragment is emphasized by a bold line. The locations of two Tn*5* insertions within the cosmid pL30-11B are symbolized by arrows (see also Fig. 2); this cosmid contains the entire *Bam*HI fragment plus adjacent DNA (see ref. 2 for further details). The Ω-insertion sites of strains 612 and 613 are also indicated. ORF, open reading frame. Restriction sites: B, *Bam*HI; Bg, *Bgl* II; Ev, *Eco*RV; Hp, *Hpa* I; Sc, *Sac* I.

Plant Infection Tests. Soybean seeds [*Glycine max* (L.) Merr. cv. Williams] were provided by the Jaques Seed Co. (Prescott, WI). Seeds from cowpea (*Vigna unguiculata* cv. Red Caloona), mung bean (*Vigna radiata*), and Siratro (*Macroptilium atropurpureum*) were kindly provided by W. D. Broughton (University of Geneva, Geneva). The infection tests were done as described (7).

DNA Biochemistry. Standard procedures (5) were used for DNA manipulations. The DNA sequence was established with the help of a DNA sequencer (model 370A, Applied Biosystems), using M13mp18 and M13mp19 (8) as sequencing vectors. For computer-aided DNA and protein sequence analysis the UWGCG (Genetics Computer Group of the University of Wisconsin, Madison, WI) software package was used.

Plasmid Construction and Transfer. In order to mutagenize the *nodV* and *nodW* genes, plasmid pRJ4356 (ref. 2 and Fig. 1) was cut with *Hpa* I or *Sac* I, respectively. The protruding 3′ ends left by *Sac* I were removed by treatment with T4 DNA polymerase. As a selectable marker, the Ω interposon (9) was inserted separately into these sites by using its *Sma* I restriction sites. The mutagenized fragments were cloned into the *Eco*RI site of the mobilizable vector pSUP202 (10). The pSUP202 derivatives were mobilized into *B. japonicum*, and marker exchange events were selected as described (3). The mutations in *nodV* (strain 612) and *nodW* (strain 613) were verified by Southern blot hybridizations. Plasmids pL30-11B::Tn*5*-15 and pL30-11B::Tn*5*-16 are Tn*5*-carrying derivatives of the cosmid pL30-11B; they were used for the complementation of the deletion strain ΔE1-7d1 (see ref. 2 for a more detailed description).

RESULTS AND DISCUSSION

***nodV* and *nodW* Are Essential for the Nodulation Ability of *B. japonicum* on Various Hosts.** A *B. japonicum* mutant

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: ORF, open reading frame.

*The sequence reported in this paper has been deposited in the GenBank data base (accession no. M31765).

(ΔE1-7d1) carrying a large chromosomal deletion was found previously to have a Nod$^-$ phenotype on mung bean, cowpea, and Siratro, whereas it showed only a minor delay of nodulation of soybean (2). A complementing recombinant cosmid has been obtained (pL30-11B) that fully restored nodulation of strain ΔE1-7d1 on all hosts. The new nodulation region (tentatively called *nod*-1) has been delimited further to ≈3.5 kilobases (kb) (2). To analyze the potential coding properties of this region we established its DNA sequence (Figs. 1 and 2). We found two large ORFs spanning >3.3 kb. ORF1 and ORF2 have coding capacities for proteins with 889 and 227 amino acids, respectively, and they were subsequently named *nodV* and *nodW* (Fig. 1). The stop codon of *nodV* (TGA) and the start codon of *nodW* (GTG) overlap (GTGA), suggesting that the two genes are organized in an operon and that they may even be coupled translationally. There are several alternative start codons for *nodV*, two of which are preceded by purine-rich, Shine–Dalgarno-like sequences (underlined in Fig. 2). We favor the ATG at position 648 over the GTG at position 996 as the putative *nodV* start codon; this choice is supported by the following two arguments. (*i*) Two Tn*5* derivatives of the cosmid pL30-11B have been obtained, of which pL30-11B::Tn*5*-16 was able to restore (complement) nodulation of mung bean by mutant

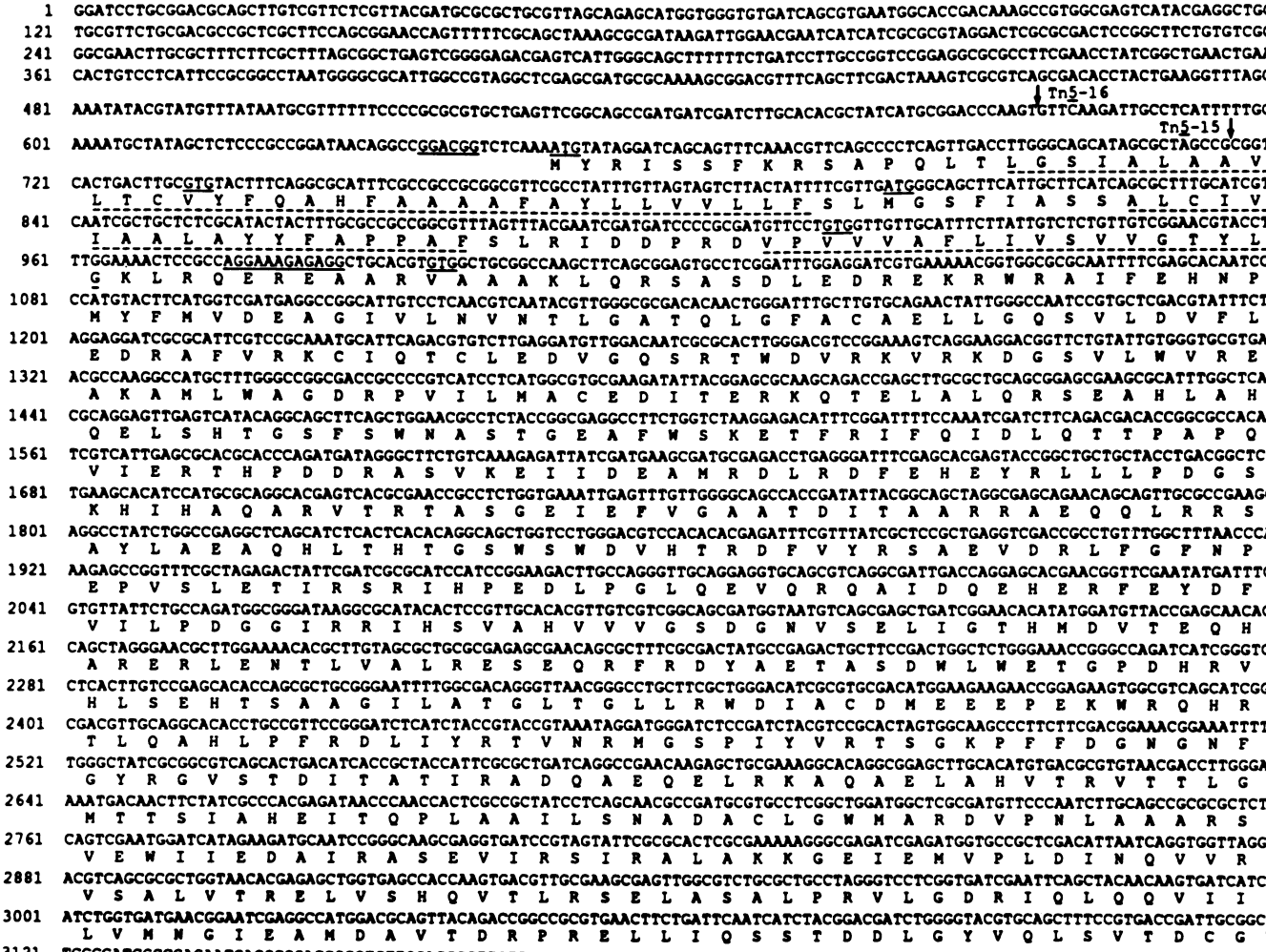

FIG. 2. Nucleotide sequence of *nodVW* and derived amino acid sequences of the NodV and NodW proteins. The exact positions of the two Tn*5* insertions within the cosmid pL30-11B (cf. Fig. 1) are indicated by vertical arrows. As translational start of *nodV*, the first ATG was chosen. Other potential start codons (ATG and GTG only) located further downstream and purine-rich regions preceding them are underlined. Three highly hydrophobic regions within the N-terminal end of the NodV protein are highlighted by broken underlining. The presumptive start codon of *nodW* (GTG) maps at position 3314.

Table 1. Nodulation behavior of *B. japonicum* 110 *spc4* (wild type) and several mutant derivatives

B. japonicum strain	Average nodule number per plant at day 21			
	Glycine max	*Vigna radiata*	*Vigna unguiculata*	*Macroptilium atropurpureum*
Wild type	24	37	42	20
612	24	0	0	0
613	26	0	0	0
ΔE1-7d1/pL30-11B::Tn*5*-15	34	1	NT	NT
ΔE1-7d1/pL30-11B::Tn*5*-16	22	34	NT	NT

Numbers are mean values from 5 plants (*Glycine max*) or 10 plants (*Vigna radiata*, *Vigna unguiculata*, *Macroptilium atropurpureum* cv. Siratro) and were confirmed by at least one repetition of the nodulation experiment. NT, not tested.

ΔE1-7d1, whereas pL30-11B::Tn*5*-15 was not (Table 1). The precise Tn*5* insertion sites were determined by sequencing, which revealed that Tn*5*-16 was located 72 bp upstream of *nodV* whereas Tn*5*-15 mapped within the 23rd codon of the *nodV* reading frame (Fig. 2). (*ii*) The *nodV* gene starting with the first ATG as shown in Fig. 2 would encode a protein with a strongly hydrophobic N terminus, whereas a *nodV* gene starting with GTG at position 996 would not. For reasons given below we believe that these hydrophobic domains are crucial for NodV function.

Having found the *nodV* and *nodW* ORFs, we next wished to construct chromosomal insertion mutations within each of them. For this purpose the Ω interposon was cloned into the *Hpa* I site of *nodV* and into the *Sac* I site of *nodW* (Fig. 1). Marker replacement of these insertions into the *B. japonicum* wild type gave rise to mutants 612 (*nodV*::Ω) and 613 (*nodW*::Ω) (Fig. 1). Both strains were tested for their nodulation ability on soybean, mung bean, cowpea, and Siratro. Table 1 shows that only soybean but none of the other hosts were nodulated. Thus, *nodV* and *nodW* are *B. japonicum* genes involved in host specificity of nodulation. Closer examination of the soybean plants infected with mutant strains 612 and 613 showed that nodules developed further down at the primary root as compared to plants infected with the wild type (data not shown). Such an altered nodule distribution might be indicative of a delay of nodulation. In fact, an experiment aimed at studying the time course of nodule development revealed that soybean nodulation was delayed with both mutant strains by about 2 days (Fig. 3).

NodV and NodW Proteins Belong to the Family of Two-Component Regulatory Systems. When the deduced amino acid sequences of the NodV and NodW proteins were used in a protein data bank search (National Biomedical Research Foundation Protein Identification Resource, release 20.0, March 1989), both proteins were found to share sequence similarity with the prokaryotic signal-transducing regulatory

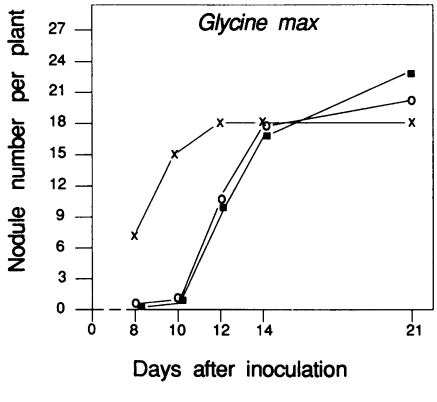

FIG. 3. Nodulation kinetics of *B. japonicum* wild type and *nodVW* mutant strains 612 (*nodV*) and 613 (*nodW*) on soybean. x, Wild type; ■, strain 612; ○, strain 613.

pairs belonging to the large superfamily of two-component regulatory systems. In the proposed model for signal transduction (11), usually a membrane-associated sensor protein responds to an environmental signal by sensing that stimulus near its N-terminal region and transducing this signal to its own C-terminal domain. The activated C-terminal domain may then interact with and modify the conserved N-terminal domain of a regulatory protein. The modified regulator is usually, though not exclusively, an activator of the expression of a specific gene or set of genes.

NodV, the Predicted Sensor. As is characteristic within the sensor class of the two-component regulatory systems, the homology to NodV is restricted to about 240 amino acids at the C-terminal region. This is illustrated in Fig. 4 and further exemplified in Fig. 5*A* by the detailed sequence alignment of the C-terminal parts of the NodV and FixL proteins. Beyond the 240 amino acids the NodV protein shows no homology at all to any of the known sensor proteins identified to date, which makes it likely that NodV does not functionally correspond to any of these and, hence, that it responds to a stimulus different from that to which the other sensors respond. The N-terminal region of NodV is highly hydrophobic (Fig. 2), suggesting that the protein is membrane-bound, as may be true for most other sensor proteins (Fig. 4). A recent study on the membrane topology of the *Agrobacterium tumefaciens* VirA protein showed that the second transmembrane domain following the periplasmic domain is critical in perceiving the stimulus (e.g., acetosyringone) to subsequently modify the VirG protein (19). The Tn*5*-15

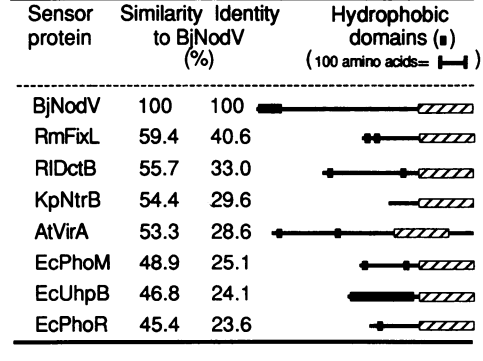

FIG. 4. Relationship between the *B. japonicum* (Bj) NodV protein and other sensor proteins of two-component regulatory systems. The percent similarity/identity values were determined by comparing *only* the C-terminal 240 amino acid residues (GAP program of the UWGCG software package, version 6.0). The lengths of the bars to the right correspond to the sizes of the proteins. Black rectangles within the bars symbolize the presence and location of hydrophobic regions with potential membrane-associated helices. The hatched region marks the conserved stretch of 240 amino acids. The figure is a summary of data on the following selected proteins: *Rhizobium meliloti* FixL (12), *Rhizobium leguminosarum* DctB (13), *Klebsiella pneumoniae* NtrB (14), *Agrobacterium tumefaciens* VirA (15), *E. coli* PhoM (16), *E. coli* UhpB (17), and *E. coli* PhoR (18).

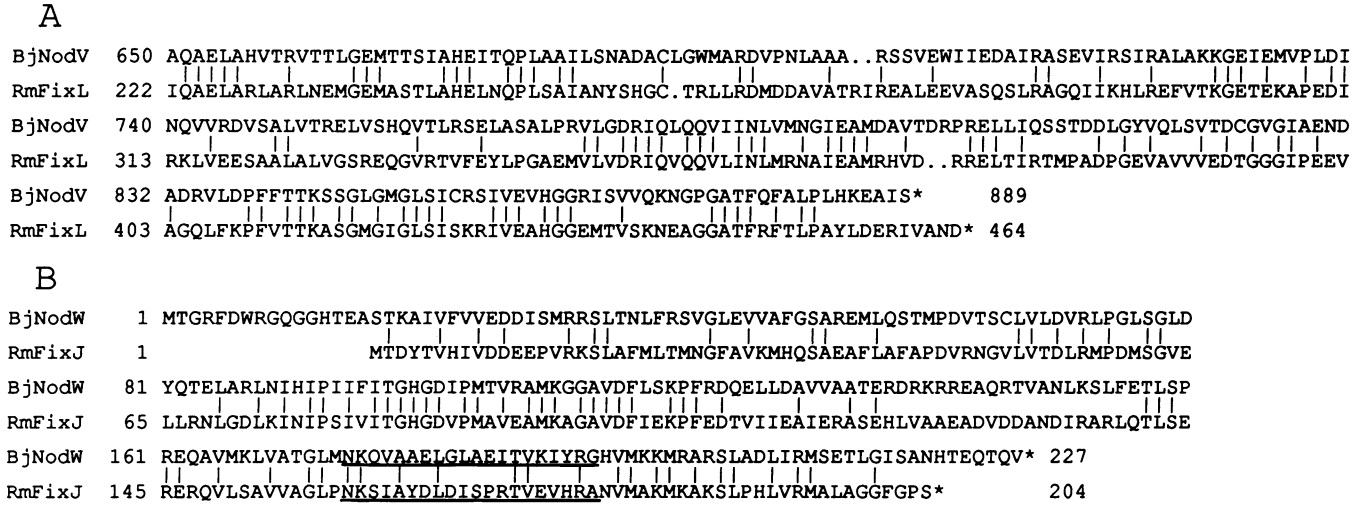

FIG. 5. Amino acid sequence alignment of the *B. japonicum* (Bj) NodV and NodW proteins with the *R. meliloti* (Rm) FixL and FixJ proteins. (*A*) NodV/FixL comparison, showing only the conserved C-terminal parts of the proteins. Gaps (.) were introduced to allow an optimal alignment. (*B*) NodW/FixJ comparison over the whole lengths of the two proteins. A potential helix–turn–helix motif is underlined.

mutation in *nodV* (see above), which produces a nodulation-defective phenotype, is located in the DNA region coding for the N-terminal, hydrophobic NodV domain. This is consistent with the idea that the *nodV* ORF starts with the first ATG as shown in Fig. 2.

NodW, the Predicted Regulator. In general, the similarity between members of the regulator class is restricted to approximately 120 amino acids located at the N terminus. The same holds true for the NodW protein (data not shown). However, NodW also exhibits more extended homology to a subclass of the regulatory proteins to which the FixJ, UhpA, and NarL proteins belong (12, 17, 20). As an example, the amino acid sequence alignment of the NodW and FixJ proteins is presented in Fig. 5*B*. One notable feature in the homology of almost all regulator proteins is the conservation of two aspartic residues (in the NodW protein at positions 27 and 70). One of these may be a target for phosphorylation, analogous to the situation with NtrC in which NtrB-mediated phosphorylation has been shown to occur within the first 12.5 kDa of the protein (21, 22). Most regulators identified to date are believed to be transcriptional activators. This may apply to NodW as well. Based on the screening method of Dodd and Egan (23), a potential DNA-binding domain (helix–turn–helix motif) was identified close to the C-terminal end of the protein (underlined in Fig. 5*B*). The important traits of such a DNA-binding domain are conserved within a stretch of 20 amino acids of NodW: (*i*) a glycine at the 9th position, (*ii*) the nonpolar residues alanine and valine at the 5th and 15th positions, and (*iii*) the space-filling residue tyrosine at the 18th position.

A Genetic Regulatory Circuit Involved in Host-Specific Nodulation. Among the sensor/regulator pairs homologous to NodV and NodW, the most homologous one was the *Rhizobium meliloti* FixL/FixJ pair (Figs. 4 and 5). This might suggest certain similarities in the signal-transduction pathway and/or the interactions between the regulators and their target DNAs. It is clear, however, that the *B. japonicum* NodV and NodW proteins are not functional homologs of the *R. meliloti* FixL and FixJ proteins. First, the phenotype of *fixLJ* mutations [impaired regulation of N_2 fixation (12)] and that of *nodVW* mutations (impaired host-specific nodulation) are completely different, and second, the *fixL*- and *fixJ*-homologous genes of *B. japonicum* have been identified just recently in our laboratory (D. Anthamatten, unpublished results). Concerning the possible function of the NodV and NodW proteins, the data reported here can be taken to formulate a working hypothesis according to which a signal-transduction event proceeds via a membrane-bound NodV. The model further predicts that the activated NodW protein acts as a transcriptional activator of an unknown gene (or several genes). It is the absence of the corresponding gene product(s) that leads to the observed nodulation defect of *nodVW* mutants. Evidently, the presence of these gene products is a strong requirement for the nodulation of mung bean, cowpea, and Siratro but much less so for nodulation of the soybean variety we used. Therefore, one goal will be to identify the NodW-regulated gene(s). One potential candidate for such a target gene may be the *B. japonicum hsn* locus described by Nieuwkoop *et al.* (24). Mutations in this region prevented the nodulation of Siratro but not that of soybean (24). Moreover, this *hsn* locus is not preceded by a *nod*-box sequence and, consequently, is probably not controlled by *nodD*. The identification and characterization of the NodW-regulated genes may help to resolve the details of this novel regulatory circuit involved in host-specific nodulation and may also lead to an approach to track down the nature of the environmental stimulus needed to induce this process.

We thank S. Hitz for expert technical assistance and W. J. Broughton for seeds. M.G. acknowledges the receipt of a fellowship from the Deutsche Forschungsgemeinschaft. This work was supported by a grant from the Swiss Federal Institute of Technology, Zurich.

1. Long, S. R. (1989) *Cell* **56**, 203–214.
2. Hahn, M. & Hennecke, H. (1988) *Appl. Environ. Microbiol.* **54**, 55–61.
3. Göttfert, M., Lamb, J. W., Gasser, R., Semenza, J. & Hennecke, H. (1989) *Mol. Gen. Genet.* **215**, 407–415.
4. Regensburger, B. & Hennecke, H. (1983) *Arch. Microbiol.* **135**, 103–109.
5. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
6. Hennecke, H., Günther, I. & Binder, F. (1982) *Gene* **19**, 231–234.
7. Hahn, M. & Studer, D. (1986) *FEMS Microbiol. Lett.* **33**, 143–148.
8. Norrander, J., Kempe, T. & Messing, J. (1983) *Gene* **26**, 101–106.
9. Prentky, P. & Krisch, H. M. (1984) *Gene* **29**, 303–313.
10. Simon, R., Priefer, U. & Pühler, A. (1983) in *Molecular Genetics of the Bacteria-Plant Interaction*, ed. Pühler, A. (Springer, Berlin), pp. 98–106.
11. Nixon, B. T., Ronson, C. W. & Ausubel, F. M. (1986) *Proc.*

Natl. Acad. Sci. USA **83,** 7850–7854.
12. David, M., Daveron, M. L., Batut, J., Dedieu, A., Domergue, O., Ghai, J., Hertig, C., Boistard, P. & Kahn, D. (1988) *Cell* **54,** 671–683.
13. Ronson, C. W., Astwood, P. M., Nixon, B. T. & Ausubel, F. M. (1987) *Nucleic Acids Res.* **15,** 7921–7934.
14. MacFarlane, S. A. & Merrick, M. (1985) *Nucleic Acids Res.* **13,** 7591–7606.
15. Melchers, L. S., Thompson, D. V., Idler, K. B., Neuteboom, S. T. C., de Maagd, R. A., Schilperoort, R. A. & Hooykaas, P. J. J. (1987) *Plant Mol. Biol.* **11,** 227–237.
16. Amemura, M., Makino, K., Shinagawa, H. & Nakata, A. (1986) *J. Bacteriol.* **168,** 294–302.
17. Friedrich, M. J. & Kadner, R. J. (1987) *J. Bacteriol.* **179,** 3556–3563.
18. Makino, K., Shinagawa, H., Amemura, M. & Nakata, A. (1986) *J. Mol. Biol.* **192,** 549–556.
19. Melchers, L. S., Regensburg-Tuink, J. G., Bourret, R. B., Sedee, N. J. A., Schilperoort, R. A. & Hooykaas, P. J. J. (1989) *EMBO J.* **8,** 1919–1925.
20. Stewart, V., Parales, J. & Merkel, S. M. (1989) *J. Bacteriol.* **171,** 2229–2234.
21. Keener, J. & Kustu, S. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 4976–4980.
22. Weiss, V. & Magasanik, B. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 8919–8923.
23. Dodd, I. B. & Egan, J. B. (1987) *J. Mol. Biol.* **194,** 557–564.
24. Nieuwkoop, A. J., Banfalvi, Z., Deshmane, N., Gerhold, D., Schell, M. G., Sirotkin, K. M. & Stacey, G. (1987) *J. Bacteriol.* **169,** 2631–2638.