MOLECULAR PLANT-MICROBE INTERACTIONS
Vol. 4, No. 4, pp. 370-378, 1991

Positive Regulators of Opine-Inducible Promoters in the Nopaline and Octopine Catabolism Regions of Ti Plasmids

J. von Lintig, H. Zanker, and J. Schröder

Universität Freiburg, Institut für Biologie II, D-7800 Freiburg, Federal Republic of Germany.
Received 28 December 1991. Accepted 29 March 1991.

The *noc* and *occ* regions of nopaline and octopine Ti plasmids in *Agrobacterium tumefaciens* contain genes for the catabolism of nopaline and octopine, respectively. We investigated the transcriptional organization and regulation of both regions. The *noc* region of pTiC58 contains two nopaline-inducible promoters, and one octopine-inducible promoter was identified in the *occ* region of pTiAch5. All three promoters are positively regulated *in trans* by constitutively expressed genes localized at the right end of the regions. The DNA sequence analysis of these parts revealed genes coding for related proteins (35.6% identity). The two polypeptides share significant similarity with a family of other positive gene regulators, and both contain a protein motif ("LysR" signature) that is characteristic for the DNA binding domain in these polypeptides. These proteins are the only Ti plasmid functions necessary for the activation of the opine-induced promoters. We propose the names *nocR* and *occR* for the regulator genes in the *noc* and in the *occ* regions.

Agrobacterium tumefaciens (Smith and Townsend) Conn induces neoplastic overgrowths on a large variety of plants by transfer of a part of the tumor-inducing (Ti) plasmid DNA into the plant cells (see Ream 1989 for recent review). Genes in the transferred DNA (T-DNA) are responsible for the tumorous growth (reviewed in Weiler and Schröder 1987), and for the synthesis of unusual sugar and amino acid conjugates, collectively called opines (Petit *et al.* 1978). These opines can serve as growth substrates for the tumor-inducing agrobacteria, and it is thought that this benefit was the driving force in the evolution of the interaction with plants (Guyon *et al.* 1980).

The catabolic functions necessary for converting the opines into substances of general metabolism are encoded in the Ti plasmids. They are induced by the opines, and a specific subset (conjugal opines) also activates the conjugative mechanisms that ensure that the majority of the plant-colonizing agrobacteria contain a copy of the appropiate Ti plasmid.

In typical nopaline-type Ti plasmids the catabolic functions for this opine are located directly to the right of the T-region, in a stretch of approximately 17 kbp that is called the *noc* region (Fig. 1A) (Holsters *et al.* 1980; Schardl and Kado 1983; Inzé 1984). The analysis of mutants also suggested that the region is divided into two parts that are separated by several kilobase pairs of DNA of unknown function (Holsters *et al.* 1980; Inzé 1984). The left part contains the genes for nopaline oxidase (EC 1.5.1.19; *nox*, two proteins) (Sans *et al.* 1987), arginase (EC 3.5.3.1; *arc*), and ornithine cyclodeaminase (EC 4.3.1.12; *ocd*). The latter two genes and their proteins have been characterized (Sans *et al.* 1988; Schrell *et al.* 1989). Nopaline is not a conjugal opine for pTiC58; with this plasmid this role is performed by agrocinopine (Ellis *et al.* 1982a, 1982b).

The catabolic functions for octopine (*occ* region; Fig. 2A) have been located to a stretch of approximately 12 kbp (DeGreve *et al.* 1981; Knauf and Nester 1982; Stachel *et al.* 1985). The region maps between coordinates 33 and 45 kbp on the octopine-type Ti plasmid map, which uses as reference a T-region *Sma*I site conserved in nopaline- and octopine-type Ti plasmids (De Vos *et al.* 1981). The genes identified so far code for octopine oxidase (*oox*, two proteins) (H. Zanker and J. Schröder, unpublished results) and ornithine cyclodeaminase (*ocd*) (Farrand and Dessaux 1986; Dessaux *et al.* 1986; Schindler *et al.* 1989). Octopine is a conjugal opine, because it also induces the *tra* functions in octopine-type plasmids.

We investigated opine-inducible promoters and regulatory genes in the catabolic regions of the nopaline-type plasmid pTiC58 and of the octopine-type plasmid pTiAch5. Two nopaline-induced promoters were identified in the *noc* region and one octopine-induced promoter was detected in the *occ* region. Both regions contain a gene that positively regulates the inducible promoters. The regulator proteins are related to each other and to a family of other positive regulators.

Address correspondence to Joachim Schröder: Institut für Biologie II, Universität Freiburg, Schänzlestr. 1, D7800 Freiburg, Federal Republic of Germany.

The nucleotide sequences have been submitted to GenBank with the accession numbers M65107 (nocr) and M65108 (occR).

© 1991 The American Phytopathological Society

MATERIALS AND METHODS

Agrobacterial strains and plasmids. *Agrobacterium* strain APF2, a plasmid-free derivative of LBA275 (= C58Cl) has been described (Hynes *et al.* 1985). Ti plasmid pGV3850 (Zambryski *et al.* 1983) is a derivative of the nopaline plasmid pTiC58. It contains the complete *noc* region, but a part of the T-region is replaced by pBR322. Plasmid pGV2260 (Deblaere *et al.* 1985) was derived from

the octopine-type plasmid pTiB6S3. It harbors the complete *occ* region, but the T-region and some adjoining sequences are replaced by pBR322.

Bacteria were grown in YEB liquid medium (0.5% Bacto beef extract, 0.1% Bacto yeast extract, 0.5% Bacto peptone, 0.5% sucrose, 2 mM $MgSO_4$; pH 7.2) (Van Larebeke *et al.* 1977) at 28° C and with the appropiate antibiotics (rifampicin, 50 μg/ml; carbenicillin, 50 μg/ml; and tetracycline, 2 μg/ml). For long-term storage, cultures grown to the stationary phase were mixed with dimethylsulfoxide (final concentration 3%) and stored at −70° C.

Escherichia coli **strains and plasmids.** The strains JM109 (Yanisch-Perron *et al.* 1985) and S17-1 (Simon *et al.* 1983) were used routinely. *E. coli* S17-1 contains the *trans*-acting mobilization factors of plasmid RP4 integrated in the chromosome. The subcloning of the Ti plasmid fragments was performed in the vectors pINIIA3 (Nakamura and Inouyé 1982) and in the pTZ18R/pTZ19R system, which provides multiple cloning sites in the polylinker (Zagursky and Berman 1984). The molecular techniques have been described (Maniatis *et al.* 1982).

Constructions with the broad host range vectors pCB303 and pUCD1002. Plasmid pCB303 has been described (Schneider and Beck 1987). It contains a small polylinker (*Sma*I, *Xba*I, *Pst*I, *Sal*I, *Bam*HI, and *Eco*RI), which is flanked to the right by the reporter gene *lacZ* and to the left by *phoA*. The Ti plasmid fragments were inserted into the unique cloning sites (*Xba*I, *Pst*I, and *Bam*HI). Suitable restriction sites were usually obtained via intermediate clones, in which the fragments had been inserted into the pTZ18R/pTZ19R polylinker. If this was not feasible, the fragments were excised and inserted blunt-ended after fill-in of the overhangs. The orientation of the inserts was determined with asymmetric sites in the cloned fragments and with help of the flanking restriction sites in the vector. Plasmid pUCD1002 (Gallie *et al.* 1985) has been described. Fragments were cloned into the Asp718 site of the kanamycin resistance gene.

Conjugative crosses. Plasmids were conjugated from *E. coli* S17-1 into *Agrobacterium* APF2, C58C1(pGV3850), or C58C1(pGV2260); the procedures have been described (Alt-Mörbe *et al.* 1989). The presence of the plasmids was

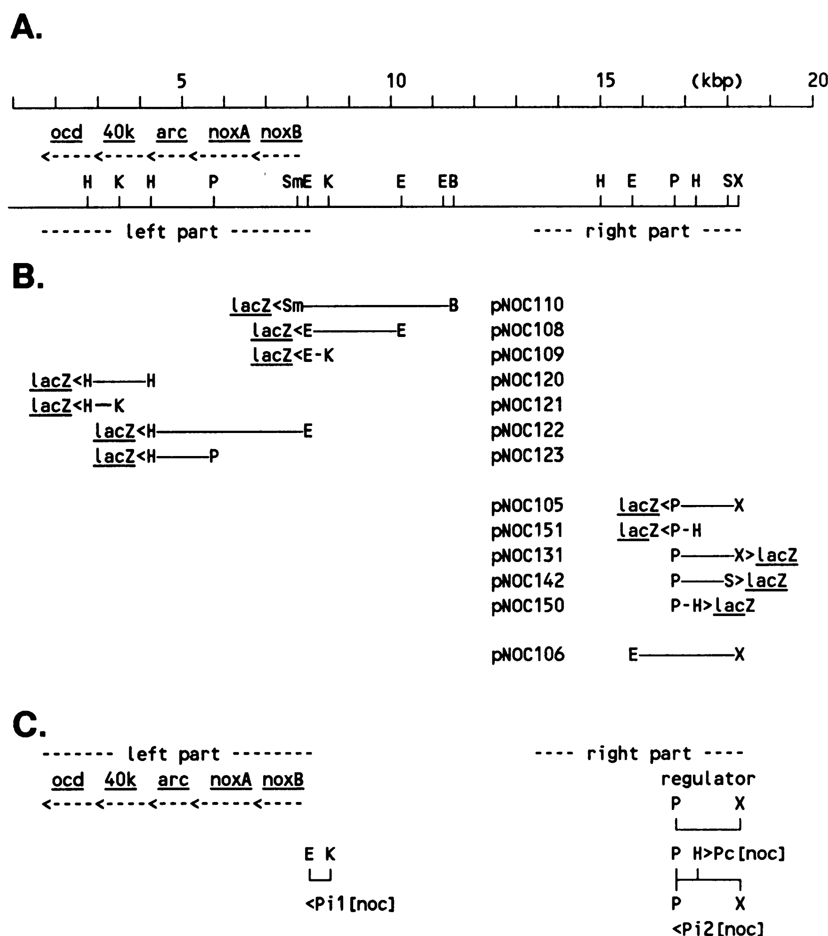


Fig. 1. The *noc* region of nopaline plasmid pTiC58. **A,** Identified catabolic genes and restriction sites used for construction of the subclones. The map coordinates refer to the *Sma*I site in the T-region that is conserved in nopaline and octopine Ti plasmids (De Vos *et al.* 1981). ocd = ornithine cyclodeaminase; 40k = 40 kDa polypeptide of unknown function; arc = arginase; *noxA* and *noxB* = two proteins necessary for oxidative cleavage of nopaline (Sans *et al.* 1987); B = *Bam*HI; E = *Eco*RI; H = *Hin*dIII; K = *Kpn*I; P = *Pst*I; S = *Sal*I; Sm = *Sma*I; X = *Xba*I. **B,** Fragments subcloned in broad host range vector pCB303. The arrowheads (< or >) indicate the location of the reporter gene *lacZ*. Plasmid pNOC106 contains an *Eco*RI/*Xba*I fragment in vector pUCD1002 (no reporter gene). **C,** Positions of the identified promoters with the direction of transcription and of the regulator gene.

routinely confirmed by miniscale plasmid isolation from the *Agrobacterium* strains, retransformation of *E. coli*, and restriction analysis. The promoter analysis was routinely performed with several independent *Agrobacterium* colonies.

Measurements of opine-inducible promoter activities. Liquid cultures were grown in YEB medium at 28° C to $O.D._{600} = 0.4$–0.8 and harvested by centrifugation. After resuspension to $O.D._{600} = 0.4$ in induction medium (2% sucrose, 20 mM sodium citrate, pH 5.5) without or with the opines (0.4 mM) the cells were incubated on a rotary shaker (200 rpm) at 28° C. A pH of 5.5 was chosen because it approaches the conditions in the close environment of plant cells. In the studies performed so far, however, no significant differences were noted when pH 7.0 was used in the experiments. Detailed kinetics of the time course of induction showed detectable increases of β-galactosidase activity after a lag phase of 25–30 min, and the activity increased for at least 6 hr. Standard experiments were therefore terminated by harvesting the cells after 6 hr. The preparation of the extracts and the determination of β-galactoside activity have been described (Stachel *et al.* 1985). The units are defined as: $1{,}000 \times O.D._{410}/\text{time (min)} \times O.D._{600}$ of the culture (Schneider and Beck 1987).

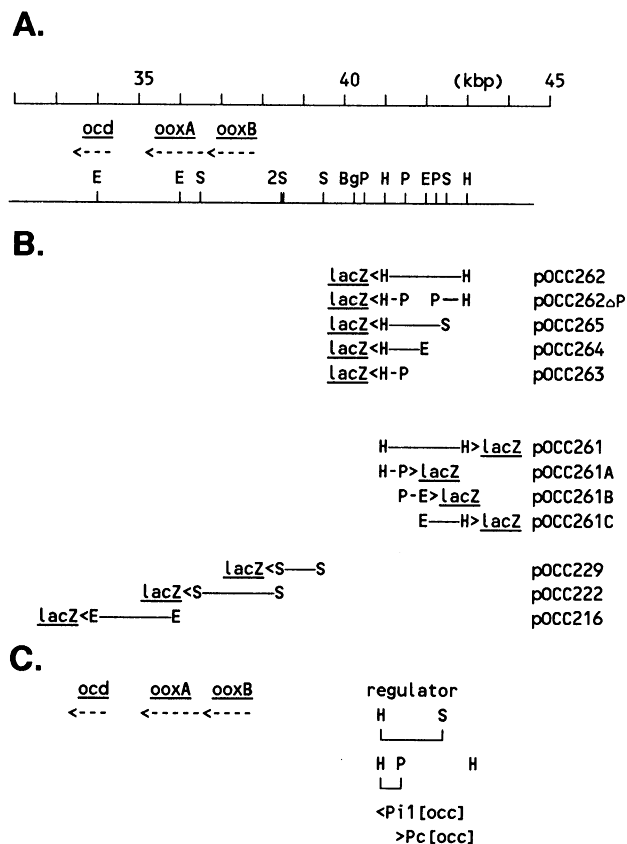


Fig. 2. The *occ* region of octopine plasmid pTiAch5. **A,** Identified catabolic functions and restriction sites used for establishment of the subclones. The map coordinates refer to a *Sma*I site in the T region that is conserved in nopaline and octopine Ti plasmids (De Vos *et al.* 1981). ocd = ornithine cyclodeaminase; *ooxA* and *ooxB* = two proteins necessary for oxidative cleavage of octopine. E = *Eco*RI; Bg = *Bgl*II; H = *Hin*dIII; P = *Pst*I; S = *Sal*I; 2S: two closely spaced *Sal*I sites. **B,** Fragments subcloned in pCB303. The arrowheads (< or >) indicate the location of the reporter gene *lacZ*. **C,** Position of the regulator element and of the identified promoters with the direction of transcription (< or >).

DNA sequence analysis. The DNA sequence of the desired fragments was determined by the dideoxy nucleotide chain termination technique (Sanger *et al.* 1977; Sanger and Coulson 1978). The pTZ18R and pTZ19R system (Zagursky and Berman 1984), helper phage M13K07 (Vieira and Messing 1987), *E. coli* strain JM109 (Yanisch-Perron *et al.* 1985) (all from Pharmacia, Molecular Biology Division, Freiburg, Germany), and the reverse-sequencing primer (Boehringer, Mannheim, Germany) were used routinely. Suitable plasmids were obtained by subcloning of DNA fragments; standard molecular techniques were employed in this work (Maniatis *et al.* 1982). DNA polymerization reactions were performed with [^{35}S]dATP (>37 TBq/mmol, Amersham Buchler, Braunschweig, Germany) and modified T7-DNA-polymerase (Sequenase, United States Biochemical Corporation, Cleveland, OH). Both strands were fully sequenced with overlapping subclones or with help of synthetic oligonucleotides.

Computer analysis. Standard sequence compilation and evaluation was performed with published programs (Kroeger and Kroeger-Block 1984). The alignment of the sequences was carried out with the ALIGN2 program of M. Trippel (Infomed GmbH, Freiburg, Germany), which is based on the algorithm of Needleman and Wunsch (1970) and the PCGENE software package (IntelliGenetics Inc., Mountain View, CA). The FSTPSCN program of PCGENE was employed for homology searches in the SWISSPROT database (release 14).

Table 1. Promoter activities of *noc* region fragments[a]

Agrobacterium strain	Resident plasmid	*noc* Region subclones	*lacZ* Activity − (nopaline)	+
Left part				
C58C1	pGV3850	pNOC110	10	383
APF2	. . .	pNOC110	12	10
C58C1	pGV3850	pNOC108	8	297
APF2	. . .	pNOC108	5	7
C58C1	pGV3850	pNOC109	5	410
APF2	. . .	pNOC109	17	<1
C58C1	pGV3850	pNOC120	14	23
APF2	. . .	pNOC120	22	20
C58C1	pGV3850	pNOC121	10	17
APF2	. . .	pNOC121	7	7
C58C1	pGV3850	pNOC122	10	8
APF2	. . .	pNOC122	7	5
C58C1	pGV3850	pNOC123	7	9
APF2	. . .	pNOC123	8	4
Right part				
C58C1	pGV3850	pNOC105	18	657
APF2	. . .	pNOC105	18	818
C58C1	pGV3850	pNOC151	10	9
APF2	. . .	pNOC151	14	13
C58C1	pGV3850	pNOC131	10	12
APF2	. . .	pNOC131	3	6
C58C1	pGV3850	pNOC142	39	47
APF2	. . .	pNOC142	3	3
C58C1	pGV3850	pNOC150	40	37
APF2	. . .	pNOC150	30	33
APF2	. . .	pNOC109+pNOC106	26	480
Controls				
C58C1	pGV3850	pCB303	9	8
APF2	. . .	pCB303	7	16

[a] The subclones are described in Figure 1B. Induction: 6 hr with 0.4 mM nopaline.

RESULTS

Promoter analysis of the *noc* region. All of these experiments were performed in *Agrobacterium* with fragments from the *noc* region of pTiC58. The fragments were cloned in broad host range vector pCB303, which contains the reporter gene *lacZ*, and the activity of β-galactosidase was taken as measure for promoter activities reading out of the Ti plasmid fragments. The position of the fragments in the *noc* region and the location of the reporter gene (either to the right or the left of the fragments) are shown in Figure 1B. The results of the β-galactosidase measurements are summarized in Table 1.

The data obtained from analysis of the left part of the region with plasmids pNOC110, pNOC108, and pNOC109 in presence of Ti plasmid pGV3850 indicated that the identified catabolic genes are preceded by a promoter (Pi1[noc]) which is induced by nopaline. The experiment with pNOC109 localized it between the *Eco*RI and the *Kpn*I site. The subclones were also tested in absence of the Ti plasmid, i.e., in strain APF2. Significant activity was not detected with any of the clones, regardless of whether nopaline was present or not. This indicated that the activity of Pi1[noc] requires additional function(s) and that these are encoded on the Ti plasmid. The low activity in absence of the opine suggested positive control of the promoter. The experiments with pNOC120, 121, 122, and 123 showed that the left part of the *noc* region contained no other promoter with significant induction by nopaline.

The analysis of the right part of the *noc* region revealed a second nopaline-induced promoter that also reads leftwards (Pi2[noc]). It was detected with a *Pst*I/*Xba*I fragment (pNOC105) in the presence of Ti plasmid pGV3850. Experiments with the same clone in APF2 revealed that the promoter was also activated by nopaline in the Ti plasmid-free background (Table 1). This indicated that pNOC105 not only contains the promoter, but also all of the Ti plasmid functions that are necessary for the induction in *Agrobacterium*. These functions could include transport of the opine into the cells and also regulatory elements. The sequence analysis of the fragment in pNOC105 (see below) identified a regulatory gene and a part of a gene for a polypeptide that may be involved in opine transport. Other experiments (data not shown) showed that Ti plasmid-encoded nopaline transport requires the expression of at least four different polypeptides, and that the part contained in pNOC105 is not sufficient. It must therefore be postulated that nopaline is taken up by other transport systems in quantities that are sufficient for induction of the promoter. This notion is supported by uptake studies with radioactive opines (nopaline and octopine; unpublished results) and by a recent report (Krishnan *et al.* 1991).

When the region contained in pNOC105 was shortened to a *Pst*I/*Hin*dIII subfragment (pNOC151), the inducible activity was lost in both *Agrobacterium* strains, indicating that Pi2[noc] extended beyond the *Hin*dIII site.

The same *Pst*I/*Xba*I fragment as in pNOC105, but with the reporter gene at the *Xba*I site (= pNOC131) possessed no significant promoter activity. A weak constitutive activity (Pc[noc]) was detected when the fragment was shortened to an *Pst*I/*Sal*I subfragment (pNOC142). The data in Table 1 also suggested that this activity requires the presence of Ti plasmid pGV3850. A further reduction of the fragment to the small piece between the *Pst*I and *Hin*dIII sites (pNOC150) revealed a constitutive activity in presence and in absence of the Ti plasmid. The lack of activity with pNOC142 in absence of pGV3850 remains unexplained. The DNA sequence analysis of this region (see below) shows that the *Sal*I site is in the coding region of the regulator protein.

We next investigated whether the regulatory functions located in the right part of the *noc* region are also sufficient to induce the promoter in the left part (Pi1[noc] in pNOC109). For this purpose, pNOC109 was introduced into *Agrobacterium* strain APF2, which already contained the *Eco*RI/*Xba*I fragment from the right part of the *noc* region in vector pUCD1002 (pNOC106, Fig. 1B). The result showed that Pi1[noc] was now inducible in the presence of nopaline (Table 1).

Promoter analysis of the *occ* region. The strategy for the analysis of the promoter activities was the same as described above for the *noc* region. All of the fragments used in these experiments were from the *occ* region of pTiAch5. The position of the fragments and the relative position of the reporter gene *lacZ* are shown in Figure 2B. The results are summarized in Table 2.

The experiments with pOCC262, pOCC262ΔP, pOCC265, pOCC264, and pOCC263 in presence of octopine plasmid pGV2260 showed that the right end of the *occ* region contains a promoter (Pi1[occ]), which reads leftwards and is induced in presence of octopine. Plasmid pOCC263 defined the position of the promoter in the sequences between a *Hin*dIII and a *Pst*I site. When the same subclones

Table 2. Promoter activities of *occ* region fragments[a]

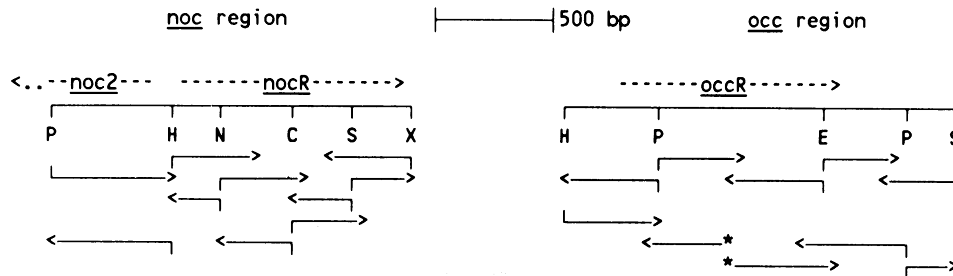
Agrobacterium strain	Resident plasmid	*occ* Region subclones	*lacZ* Activity − (octopine)	+
C58C1	pGV2260	pOCC262	73	2,840
APF2	...	pOCC262	114	2,283
C58C1	pGV2260	pOCC262ΔP	40	2,287
APF2	...	pOCC262ΔP	64	57
C58C1	pGV2260	pOCC265	32	1,870
APF2	...	pOCC265	67	1,861
C58C1	pGV2260	pOCC264	42	2,580
APF2	...	pOCC264	35	56
C58C1	pGV2260	pOCC263	41	3,200
APF2	...	pOCC263	28	27
C58C1	pGV2260	pOCC261	40	50
APF2	...	pOCC261	68	44
C58C1	pGV2260	pOCC261A	111	92
APF2	...	pOCC261A	136	123
C58C1	pGV2260	pOCC261B	11	10
APF2	...	pOCC261B	12	11
C58C1	pGV2260	pOCC261C	45	53
APF2	...	pOCC261C	48	44
C58C1	pGV2260	pOCC229	11	15
APF2	...	pOCC229	12	13
C58C1	pGV2260	pOCC222	12	14
APF2	...	pOCC222	11	12
C58C1	pGV2260	pOCC216	23	25
APF2	...	pOCC216	27	22
Controls				
C58C1	pGV2260	pCB303	8	6
APF2	...	pCB303	7	16

[a] The plasmids are described in Figure 2B. Induction: 6 hr with 0.4 mM octopine.

were assayed in the Ti plasmid-free strain APF2, only pOCC262 and pOCC265 remained inducible by the opine. This indicated 1) that the *Hin*dIII/*Sal*I subfragment in pOCC265 not only contained the inducible promoter at its left side, but also all of the Ti plasmid functions that are necessary for its activation, and 2) that a further shortening of the fragment (pOCC264 and pOCC263) or the *Pst*I deletion (pOCC262ΔP) destroyed the regulatory element. The low activity of these subclones in APF2 in absence of octopine also strongly suggests that the changes inactivated an activator element, not a repressor.

We also investigated whether the region containing Pil[occ] and the regulatory functions contained promoter activities reading rightwards (Table 2, part 2). The results

A.



B.

```
1               HindIII
nocR         taaAAGCTTGCCTTCCATATCGCAAGGGCAGTAATAGGCATGCGCCCATGTGTATTTGTTATGGAGAAAGCAATGATTCAATCGCGTCAG
occR  tagGGTGAGCTTTACATTCGATATGCATTCGGTCAAATTCATAATGACCGGGCAAGAATAAGCAGATGTTATGG----TGCGTTTATGAATCTCAGGCAG

101                                                                                      NaeI
nocR  TTAGAAGCGTTCCGGGCAGTCATGCTGACAGGAGGTATGACGTCAGCAGCGAATCTGGTGAGGATCACGCAGCCCGCGATCAGCCGGCTGATCAGGGATC
occR  GTCGAGGCGTTCCGGGCAGTCATGCTGACGGGGCAAATGACGGCGGCGGCTGAACTAATGCTGGTGACTCAGCCGGCCATCAGTCGCCTAATCAAGGACT

201
nocR  TCGAAGAGGAAATTGGGATCAGCCTCTTCGAAAGAACGGGCAACCGGTTACGTCCTACGCGGGAGGCCGGTATTCTGTTCAAGGAAGTGTCGCGACATTT
occR  TTGAACAGGCGACAAAACTGCAGCTCTTCGAGAGGCGTGGGAACCATATTATCCCGACACAGGAGGCAAAGACGCTGTGGAAAGAGGTCGATCGGGCGTT
                       PstI
301
nocR  CAACGGGATTCAGCACATCGACAAAGTCGCGGCTGAACTGAAGAAGTCTCATATGGGGTCCCTAAGGGTCGCCTGTTATACAGCGCCAGCTCTGAGTTTT
occR  CGTCGGGCTTAATCATATAGGCAACCTGGCTGCCGACATCGGCAGGCAGGCAGCGGGGACGCTCCGCATTGCTGCAATGCCTGCTCTGGCAAACGGCCTC

401
nocR  ATGTCCGGCGTCATTCAGACGTTCATCGCCGATCGGCCCGACGTGTCGGTCTACCTCGACACAGTTCCTTCCCAGACGGTCCTCGAATTGGTCTCGCTCC
occR  TTGCCGCGGTTTCTTGCTCAGTTCATCCGTGACAGACCAAATCTCCAGGTCTCCCTAATGGGACTGCCCTCAAGCATGGTCATGGAAGCCGTTGCGTCCG

501                ClaI
nocR  AGCACTACGATCTCGGAATATCGATATTGGCTGGCGACTATCCTGGTCTCACCACCGAACCTGTCCCTTCCTTTCGTGCGGTCTGCCTGCTGCCGCCGGG
occR  GCAGGGCCGACATCGGTTATGCCGATGGCCCACAGGAGCGCCAAGGTTTTCTAATCGAAAC---CCGGTCGCTTCCCGCTGTTGTCGCTGTCCCGATGGG

601
nocR  GCATCGTCTCGAAGACAAGGAAACTGTTCATGCGACGGACCTTGAAGGAGAGTCATTGATTTGCCTCTCTCCAGTGAGCCTTCTACGGATGCAAACGGAC
occR  ACATCGACTTGCTGGCCTTGACCGTGTCACGCCACAGGACCTTGCCGGTGAGCGTATTATAAAACAGGAGACTGGCACTCTCTTCGCCATGCGGGTAGAG

701
nocR  GCCGCACTGGACAGCTGCGGCGTCCAC---TGTAATCGCAGGATAGAAAGTAGTCTGGCGCTGAATCTCTGCGATCTGGTAAGCAGGGGAATGGGGGTTG
occR  GTGGCGATTGGTGGTATTCAACGCCGGCCGTCAATTGAAGTGAGCCTGTCGCATACTGCGCTAAGTCTCGTCCGCGAAGGCGCCGGGATCGCAATTATCG

801    SalI
nocR  GTATCGTCGACCCCTTCACTGCCGACTACTACAGTGCAAATCCGGTTATTCAGCGCTCCTTTGATCCGGTTGTCCCCTACCATTTTGCTATAGTTCTTCC
occR  ATCCAGCCGCGGCGATCGAGTTCACGGACAGGATCGTACTGCGACCGTTCTCGATCTTCATTGACGCCGGATTCCTCGAAGTCCGGTCAGCAATTGGCGC

901
nocR  GACCGACAGCCCACCGCCGCGCTTGGTTAGCGAGTTCCGGGCAGCGTTGCTTGATGCTTTGAAAGCCTTGCCCTATGAAACCATTtgaTCGTCAGGATCG
occR  TCCCTCAACCATCGTCGATCGTTTCACAACCGAATTCTGGAGGTTTCATGATGACTTGATGAAGCAGAACGGCCTAATGGAGtaaTGCATATTTCGGCAT
                                     EcoRI
```

Fig. 3. Molecular analysis of the regulator gene regions. **A,** Sequence strategy. Abbreviations: C = *Cla*I; H = *Hin*dIII; P = *Pst*I; E = *Eco*RI; S = *Sal*I; X = *Xba*I; N = *Nae*I. * = Sequences obtained with synthesized oligonucleotides as primers. **B,** Alignment of the open reading frames in the regulator gene regions in the *noc* region (*nocR*) and in the *occ* region (*occR*). Indicated are: the stop codons defining the start and end of the open reading frames (lowercase letters) potential start methionines (▬) and some of the restriction sites used in the sequence analysis. - - - = Gaps introduced for maximizing the alignment.

revealed no inducible promoter, but significant constitutive activities with pOCC261A, and these were independent of Ti plasmid pGV2260. This indicates that the small *Hin*dIII/*Pst*I contained not only Pil[occ] (reading leftwards) but also a constitutive promoter reading rightwards (Pc[occ]).

The plasmids pOCC229, pOCC222, and pOCC216 were used to investigate whether the left part of the *occ* region down to the end of the *ocd* gene contained additional octopine-induced promoters in the orientation predicted by the catabolic genes. Table 2 (part 3) shows that in all cases the activities were very low, regardless of the absence/presence of pGV2260 or of the opine. However, the fragments used in these experiments were large and did not cover all of the region, and therefore the presence of other promoters cannot be excluded at present.

Sequence analysis of the regulator gene regions. The functional analysis identified regulatory elements in the 1.5-kbp *Pst*I/*Xba*I fragment (see Fig. 1C) at the right end of the *noc* region, and in the 1.6-kbp *Hin*dIII/*Sal*I fragment (see Fig. 2C) at the right end of the *occ* region. Both fragments were sequenced, and Figure 3 summarizes the sequence strategy and the results.

The *noc* region fragment contained two large open reading frames (Fig. 3A). *Noc2* is located in the left part of the *Pst*I/*Xba*I fragment. It reads from right to left (with respect to the map in Fig. 3A) and the open reading frame continues into the *Pst*I fragment to the left (not shown). The analysis of the predicted polypeptide indicated that it is related to proteins involved in transport functions in other bacteria (data not shown).

The DNA sequence of the other open reading frame was designated *nocR* (Fig. 3B). The open reading frame proceeds from left to right; it begins close to the *Hin*dIII site and ends a few base pairs before the *Xba*I site (not contained in the sequence of *nocR* in Fig. 3B). The predicted amino acid sequence is shown in Figure 4 (NocR). A search for related sequences in the SWISSPROT library revealed that the polypeptide shares significant similarity with several regulatory proteins, and the polypeptide with the highest values (LysR) is shown in Figure 4. All of these polypeptides contain the LysR signature, an amino acid sequence motif, with some allowed variations, which is typical for a specific family of positive regulators (Henikoff *et al.* 1988). NocR corresponds to the signature in 10 out of 11 positions (Table 3). The precise start amino acid of NocR was not determined, but a comparison with the other regulatory proteins (not shown) and with the corresponding protein from the *occ* region suggests that the second or third methionine are likely candidates.

The *occ* region fragment contained a large open reading frame; it was designated *occR* (Fig. 3A). It starts about 200 bp after the *Hin*dIII site and terminates after the *Eco*RI site. The DNA sequence has similarities with *nocR* (Fig. 3B). The predicted amino acid sequence (Fig. 4, OccR) shares significant similarities with NocR (35.6%) and with the other proteins listed in Table 3. OccR corresponds in nine out of 11 positions to the amino acids conserved in the LysR signature (Table 3). A complete alignment with NocR and LysR is shown in Figure 4.

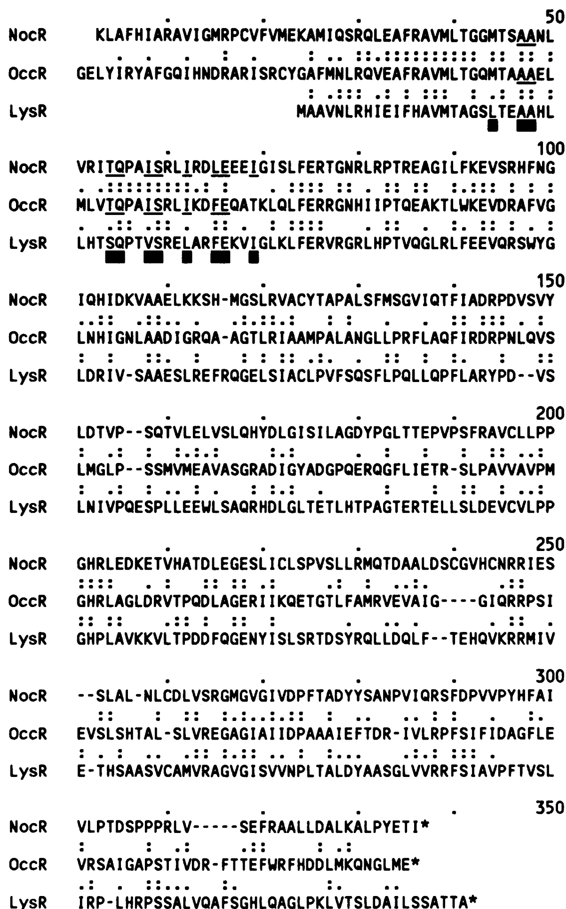


Fig. 4. Alignment of NocR with OccR, and of OccR with LysR. Double dots = sequence identity; dots = conservative exchanges; ■ = amino acids corresponding to the LysR signature (underlined in NocR and OccR).

Functional exchange of elements from the *noc* and the *occ* region. The sequence analysis indicated that the regulator proteins from the two catabolic regions are related. We therefore investigated whether NocR (OccR) can acti-

Table 3. Overall similarity (% identity) of NocR and OccR with other regulatory proteins in bacteria, and comparison with their LysR signatures[a]

	NocR	OccR	LysR signature	reference
LysR	32.2	28.0	LTEAAHLLHTSQPTVSRELARFEKVI	Stragier and Patte (1983)
TrpI	22.9	23.2	ISLAAEELHVTHGAVSRQVRLLEEDL	Chang et al. (1989)
AmpR	23.2	21.5	FTKAAIELRVTQAAVSHQVARLEDLL	Campbell et al. (1989)
OxyR	22.3	22.3	FRRAADSCHVSQPTLSGQIRKLEDEL	Bölker and Kahmann (1989)
CysB	22.6	20.4	SSTAEGLYTSQPGISKQVRMLEDELG	Ostrowski et al. (1987)
NodD	21.8	22.8	LTAAARSINLSQPAMSAAISRLRDYF	Shearman et al. (1986)
IlvY	19.9	19.2	FGRSARAMHVSPSTLSRQIQRLEEDL	Wek and Hatfield (1986)
NocR			MTSAANLVRITQPAISRLIRDLEEEI	
OccR			MTAAAELMLVTQPAISRLIKDFEQAT	

[a] Note that the signature does not represent a strict consensus; it allows variations of the motif without loss of function.

vate an inducible promoter from the *occ* region (*noc* region). The results are summarized in Table 4. They showed that OccR (or more precisely, OccR or other protein[s] from the octopine plasmid pGV2260) did not induce Pi1[noc] from the *noc* region in presence of octopine, nopaline, or both. The combination of *noc* region regulator and inducible promoter from the *occ* region gave a more complex picture. These experiments were performed with Ti plasmid pGV3850 or pNOC106 as sources of the regulatory protein, and with two different plasmids containing the inducible promoter from the *occ* region. NocR could activate Pi1[occ], but only with nopaline, not with octopine, and octopine did not reduce the induction significantly when supplied simultaneously with nopaline. The values were low when compared with OccR in presence of octopine (Table 2), but significantly higher than the controls.

DISCUSSION

Opine inducible promoters. The results indicate two nopaline-induced promoters in the *noc* region (Fig. 1C). The first is located in the left part and governs the expression of the identified catabolic enzymes. The second was discovered in the right part. It controls the expression of several polypeptides, and sequence analyses suggest that at least some of them are involved in transport functions (J. von Lintig and J. Schröder, unpublished results; D. W. R. White, M. Pritchard, and F. Marincs. Genetic and sequence analysis of *Agrobacterium* opine catabolism genes. Fifth International Symposium on the Molecular Genetics of Plant-Microbe Interactions, poster no. 39). Only one octopine-induced promoter was identified in the *occ* region (Fig. 2C). There is a gap between the promoter and the genes with identified functions, and sequence data suggest that this region contains the coding sequences for four polypeptides that resemble transport proteins in other gram-negative bacteria (H. Zanker and J. Schröder, unpublished results; S. Winans, personal communication). Additional promoters are not rigorously excluded at present.

The sum of the results suggests that the main regulation of the catabolic functions is controlled by these inducible promoters. This, however, does not exclude more complex mechanisms of fine tuning. More detailed experiments analyzing the regulated expression of catabolic enzymes from the *noc* region suggest that other factors can modulate the expression (J. von Lintig and J. Schröder, unpublished results). Interestingly, the basal and induced promoter activities are in general much higher in the *occ* than in the *noc* region (compare Tables 1 and 2). The reasons are not known. The differences were also observed in APF2, and therefore they cannot reflect the effects of other genes in nopaline (pGV3850) or octopine (pGV2260) Ti plasmids.

Regulator proteins. A regulatory gene in the *noc* region has been postulated before (Inzé 1984). Our results show that the *noc* and the *occ* regions both code for a protein that regulates the inducible promoters. The available evidence indicates that these are the only Ti plasmid proteins that are necessary, and that the regulation is by activation, not by repression; this follows from the result that the promoters are silent in the absence of the proteins. The results could be explained by a model in which a protein/opine complex interacts directly with the promoters; this would not exclude additional contributions by chromosomal functions.

The regulator proteins are related, and therefore it was an interesting question whether they could be exchanged. This was unsuccessful with the *occ* regulatory gene, in the presence of either octopine or nopaline. The *noc* region regulator in combination with nopaline but not with octopine, however, led to a low, but significant induction of the *occ* region promoter (Table 4). This was observed not only in the presence of pGV3850 (which may contain other genes influencing the regulation), but also with

Table 4. Exchange of elements from the *noc* and the *occ* region[a]

Agrobacterium tumafaciens strain	Source of regulator	Promoter tested	Induction	*lacZ* Activity
C58C1	OccR (pGV2260)	Pi1[noc] (pNOC109)	Octopine	13
			Nopaline	11
			Both	13
C58C1	NocR (pGV3850)	Pi1[occ] (pOCC262ΔP)	None	51
			Octopine	60
			Nopaline	167
			Both	166
C58C1	NocR (pGV3850)	Pi1[occ] (pOCC264)	None	75
			Octopine	80
			Nopaline	207
			Both	172
APF2	NocR (pNOC106)	Pi1[occ] (pOCC262ΔP)	None	81
			Octopine	75
			Nopaline	287
			Both	290
APF2	NocR (pNOC106)	Pi1[occ] (pOCC264)	None	80
			Otopine	60
			Nopaline	161
			Both	154
APF2	NocR (pNOC106)	None	Octopine	13
			Nopaline	15

[a] The plasmids are described in Figures 1B and 2B. Induction with the opines: 6 hr with 0.4 mM.

pNOC106, which contains *nocR* as the only regulatory gene. Simultaneous addition of octopine did not block the induction with nopaline. In the framework of the model proposed above this would suggest that NocR does not bind octopine.

The regulation of the *occ* functions by activation is in contrast to previous reports that proposed from the analysis of octopine Ti plasmid mutants that *occ* and *tra* are negatively controlled by a common repressor (Klapwijk *et al.* 1978; Klapwijk and Schilperoort 1982). These authors, however, did not exclude a contribution of positive elements. The regulation of the *tra* functions has not been investigated on the molecular level, and more detailed studies may resolve the apparent contradiction.

NocR and OccR are related to each other and to a family of other gene activators that share a common DNA binding motif (Henikoff *et al.* 1988). This motif is also present in the Ti plasmid proteins, and the similarity to the other polypeptides is most pronounced in the aminoterminal part of the proteins where the signature is located. Its presence strongly supports the conclusion from the functional assays that NocR and OccR are positive gene regulators. We therefore propose the names *nocR* and *occR* for the regulator genes from the *noc* and the *occ* regions.

ACKNOWLEDGMENTS

We thank G. Lurz for her competent help, Ch. Beck for vector pCB303, and C. I. Kado for vector pUCD1002. This work was supported by Deutsche Forschungsgemeinschaft (SFB206) and Fonds der Chemischen Industrie.

LITERATURE CITED

Alt-Mörbe, J., Kühlmann, H., and Schröder, J. 1989. Differences in induction of Ti plasmid virulence genes *virG* and *virD*, and continued control of *virD* expression by four external factors. Mol. Plant-Microbe Interact. 2:301-308.

Bolker, M., and Kahmann, R. 1989. The *Escherichia coli* regulatory protein *OxyR* discriminates between methylated and unmethylated states of the phage Mu *mom* promoter. EMBO J. 8:2403-2410.

Campbell, J. I. A., Scahill, S., Gibson, T., and Ambler, R. P. 1989. The phototrophic bacterium *Rhodopseudomonas capsulata* sp108 encodes an indigenous class A β-lactamase. Biochem. J. 260:803-812.

Chang, M., Hadero, A., and Crawford, I. P. 1989. Sequence of the *Pseudomonas aeruginosa trpI* activator gene and relatedness of *trpI* to other procaryotic regulatory genes. J. Bacteriol. 171:172-183.

De Greve, H., Decraemer, H., Seurinck, J., Van Montagu, M., and Schell, J. 1981. The functional organization of the octopine *Agrobacterium tumefaciens* plasmid pTiB6S3. Plasmid 6:235-248.

De Vos, G., De Beuckeleer, M., Van Montagu, M., and Schell, J. 1981. Restriction endonuclease mapping of the octopine tumor-inducing plasmid pTiAch5 of *Agrobacterium tumefaciens*. Plasmid 6:249-253.

Deblaere, R., Bytebier, B., De Greve, H., Deboeck, F., Schell, J., Van Montagu, M., and Leemans, J. 1985. Efficient octopine Ti plasmid-derived vectors for *Agrobacterium*-mediated gene transfer to plants. Nucleic Acids Res. 13:4777-4788.

Dessaux, Y., Petit, A., Tempé, J., Demarez, M., Legrain, Ch., and Wiame, J.-M. 1986. Arginine catabolism in *Agrobacterium* strains: Role of the Ti-plasmid. J. Bacteriol. 166:44-50.

Ellis, J. G., Murphy, P. J., and Kerr, A. 1982b. Isolation and properties of transfer regulatory mutants of the nopaline Ti-plasmid pTiC58. Mol. Gen. Genet. 186:275-281.

Ellis, J. G., Kerr, A., Petit, A., and Tempé, J. 1982a. Conjugal transfer of nopaline and agropine Ti-plasmids—The role of agrocinopines. Mol. Gen. Genet. 186:269-274.

Farrand, S. K., and Dessaux, Y. 1986. Proline biosynthesis encoded by the *noc* and *occ* loci of *Agrobacterium* Ti plasmid. J. Bacteriol. 167:732-734.

Gallie, D. R., Novak, S., and Kado, C. I. 1985. Novel high- and low-copy stable cosmids for use in *Agrobacterium* and *Rhizobium*. Plasmid 14:171-175.

Guyon, P., Chilton, M. D., Petit, A., and Tempe, J. 1980. Agropine in "null-type" crown gall tumors: Evidence for generality of the opine concept. Proc. Natl. Acad. Sci. USA 77:2693-2697.

Henikoff, S., Haughn, G. W., Calvo, J. M., and Wallace, J. C. 1988. A large family of bacterial activator proteins. Proc. Natl. Acad. Sci. USA 85:6602-6606.

Holsters, M., Silva, B., Van Vliet, F., Genetello, C., De Block, M., Dhaese, P., Depicker, A., Inzé, D., Engler, G., Villarroel, R., Van Montagu, M., and Schell, J. 1980. The functional organization of nopaline *Agrobacterium tumefaciens* plasmid pTiC58. Plasmid 3:212-230.

Hynes, M. F., Simon, R., and Pühler, A. 1985. The development of plasmid-free strains of *Agrobacterium tumefaciens* by using incompatibility with a *Rhizobium meliloti* plasmid to eliminate pAtC58. Plasmid 13:99-105.

Inzé, D. 1984. Een genetische studie van het pTiC58 plasmide van *Agrobacterium tumefaciens*: De constructie en karakterisatie van T-DNA mutanten en de identificatie van functies voor ornithine catabolisme. Ph.D. thesis. Rijksuniversiteit Gent, Belgium.

Klapwijk, P. M., and Schilperoort, R. A. 1982. Genetic determination of octopine degradation. Pages 475-495 in: Molecular Biology of Plant Tumors. G. Kahl and J. Schell, eds. Academic Press, New York.

Klapwijk, P. M., Scheulderman, T., and Schilperoort, R. A. 1978. Coordinated regulation of octopine degradation and conjugative transfer of Ti-plasmid in *Agrobacterium tumefaciens*: Evidence for a common regulatory gene and separate operons. J. Bacteriol. 136:775-785.

Knauf, V. C., and Nester, E. W. 1982. Wide host range cloning vectors: A cosmid clone bank of an *Agrobacterium* Ti plasmid. Plasmid 8:45-54.

Krishnan, M., Burgner, J. W., Chilton, W. S., and Gelvin, S. B. 1991. Transport of nonmetabolizable opines by *Agrobacterium tumefaciens*. J. Bacteriol. 173:903-905.

Kroeger, M., and Kroeger-Block, A. 1984. Simplified computer programs for search of homology within nucleotide sequences. Nucleic Acids Res. 12:193-201.

Maniatis, T., Fritsch, E. F., and Sambrook, J. 1982. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Nakamura, K., and Inouyé, M. 1982. Construction of versatile expression cloning vehicles using the lipoprotein gene of *E. coli*. EMBO J. 1:771-775.

Needleman, S. B., and Wunsch, C. D. 1970. A general method applicable to the search of similarities in the amino acid sequences of two proteins. J. Mol. Biol. 48:443-453.

Ostrowski, J., Jagura-Burdzy, G., and Kredich, N. M. 1987. DNA sequences of the *cysB* regions of *Salmonella typhimurium* and *Escherichia coli*. J. Biol. Chem. 262:5999-6005.

Petit, A., Dessaux, Y., and Tempé, J. 1978. The biological significance of opines. I. A study of opine catabolism by *Agrobacterium tumefaciens*. Pages 143-152 in: Proc. Int. Conf. Plant Pathol. Bacteriol., 4th. M. Ride, ed. Institut National de Recherche Agronomique, Angers, France.

Ream, W. 1989. *Agrobacterium tumefaciens* and interkingdom genetic exchange. Annu. Rev. Phytopathol. 27:583-618.

Sanger, F., and Coulson, A. R. 1978. The use of thin acrylamide gels for DNA sequencing. FEBS Lett. 87:107-110.

Sanger, F., Nicklen, S., and Coulson, A. R. 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74:5463-5467.

Sans, N., Schindler, U., and Schröder, J. 1988. Ornithine cyclodeaminase from Ti plasmid C58: DNA sequence, enzyme properties, and regulation of activity by arginine. Eur. J. Biochem. 173:123-130.

Sans, N., Schröder, G., and Schröder, J. 1987. The *noc* region of Ti plasmid C58 codes for arginase and ornithine cyclodeaminase. Eur. J. Biochem. 167:81-87.

Schardl, C. L., and Kado, C. I. 1983. A functional map of the nopaline catabolism genes on the Ti plasmid of *Agrobacterium tumefaciens* C58. Mol. Gen. Genet. 191:10-16.

Schindler, U., Sans, N., and Schröder, J. 1989. Ornithine cyclodeaminase from octopine Ti plasmid Ach5: Identification, DNA sequence, enzyme properties and comparison with gene and enzyme from nopaline Ti plasmid C58. J. Bacteriol. 171:847-854.

Schneider, K., and Beck, C. F. 1987. New expression vectors for identifying and testing signal structures for initiation and termination of transcription. Methods Enzymol. 153:452-461.

Schrell, A., Alt-Mörbe, J., Lanz, T., and Schröder, J. 1989. Arginase of *Agrobacterium* Ti plasmid C58—DNA sequence, properties, and comparison with eucaryotic enzymes. Eur. J. Biochem. 184:635-641.

Shearman, C. A., Rossen, L., Johnston, A. W. B., and Downie, J. A. 1986. The *Rhizobium leguminosarum* nodulation gene *nodF* encodes a polypeptide similar to acyl-carrier protein and is regulated by *nodD* plus a factor in pea root exudate. EMBO J. 5:647-652.

Simon, R., Priefer, U., and Pühler, A. 1983. Vector plasmids for in vivo manipulations of gram-negative bacteria. Pages 98-106 in: Molecular Genetics of the Bacteria-Plant Interaction. A. Pühler, ed. Springer-Verlag, Heidelberg, New York.

Stachel, S. E., An, G., Flores, C., and Nester, E. W. 1985. A Tn3 *lacZ* transposon for the random generation of β-galactosidase gene fusions: Application to the analysis of gene expression in Agrobacterium. EMBO J. 4:891-898.

Stragier, P., and Patte, J.-C. 1983. Regulation of diaminopimelate decarboxylase synthesis in *Escherichia coli*. III. Nucleotide sequence and regulation of the *lysR* gene. J. Mol. Biol. 168:333-350.

Van Larebeke, N., Genetello, C., Hernalsteens, J. P., Depicker, A., Zaenen, I., Messens, E., Van Montagu, M., and Schell, J. 1977. Transfer of Ti plasmids between *Agrobacterium* strains by mobilization with the conjugative plasmid RP4. Mol. Gen. Genet. 152:119-124.

Vieira, J., and Messing, J. 1987. Production of single-stranded plasmid DNA. Methods Enzymol. 153:3-11.

Weiler, E. W., and Schröder, J. 1987. Hormone genes and crown gall disease. Trends Biochem. Sci. 12:271-275.

Wek, R. C., and Hatfield, G. W. 1986. Nucleotide sequence and *in vivo* expression of the *ilvY* and *ilvC* genes in *Escherichia coli* K12. J. Biol. Chem. 261:2441-2450.

Yanisch-Perron, C., Vieira, J., and Messing, J. 1985. Improved M13 phage cloning vectors and host strains: Nucleotide sequence of the M13mp18 and pUC19 vectors. Gene 33:103-119.

Zagursky, R. J., and Berman, M. L. 1984. Cloning vectors that yield high levels of single-stranded DNA for rapid DNA sequencing. Gene 27:183-191.

Zambryski, P., Joos, H., Genetello, C., Leemans, J., Van Montagu, M., and Schell, J. 1983. Ti plasmid vector for the introduction of DNA into plant cells without alteration of their normal regeneration capacity. EMBO J. 2:2143-2150.