MOLECULAR PLANT-MICROBE INTERACTIONS
Vol. 4, No. 4, pp. 379-385, 1991

Transcription of the Octopine Catabolism Operon of the *Agrobacterium* Tumor-Inducing Plasmid pTiA6 Is Activated by a LysR-Type Regulatory Protein

Linda F. Habeeb, Lu Wang, and Stephen C. Winans

Section of Microbiology, Cornell University, Ithaca, NY 14853 U.S.A.
Received 8 March 1991. Accepted 12 April 1991.

Agrobacterium tumefaciens **incites crown gall tumors on plant hosts by conjugally transferring a discrete fragment of oncogenic DNA. In addition to oncogenes, the transferred DNA contains genes that direct the synthesis and exudation of opines, which are used as nutrients by the bacteria. The bacterium contains one or more operons of Ti plasmid-encoded genes that are required for the internalization and utilization of opines, and transcription of these catabolic genes is induced by cognate opines. Here we localize the gene required for regulated expression of the octopine degradative operon of the pTiA6 plasmid to a 2-kb fragment of Ti plasmid DNA. The protein encoded by this DNA positively regulates the transcription of the catabolic operon in the presence of octopine. In addition, it negatively regulates its own gene in the presence or absence of octopine. The sequence of this gene was determined and analysis of the inferred protein sequence indicates that the gene encodes a member of the LysR family of prokaryotic transcriptional regulatory proteins.**

Considerable progress has been made elucidating the molecular mechanisms underlying crown gall tumorigenesis by *Agrobacterium tumefaciens* (Smith and Townsend) Conn. These bacteria transfer approximately 20 kilobases (kb) of Ti plasmid-encoded DNA to plant nuclei. This transferred DNA (T-DNA) becomes covalently integrated into the plant genome (Chilton *et al.* 1980) and expresses two types of genes: 1) oncogenes, which upset the balance of phytohormones, resulting in tumors or teratomas (reviewed by Binns and Thomashow 1988; Ream 1989), and 2) opine biosynthetic genes, which direct the synthesis of novel amino acid or sugar conjugates that are then released from plant cells and catabolized by colonizing bacteria (Petit and Tempe 1985). In addition, the transfer process has been investigated and now appears to be a highly specialized form of bacterial conjugation (Zambryski 1988). The two-component signal transduction system that allows *Agrobacterium* to perceive wounded plants susceptible to infection has also been described (Winans *et al.* 1986; Stachel and Zambryski 1986).

By comparison, relatively little is known about molecular aspects of the colonization of crown gall tumors. We would therefore like to learn more about interactions between *Agrobacterium* strains and the plant tumors that they incite. It seems possible (or even probable) that there may be as yet undiscovered genes that are preferentially expressed when the bacterium is living in association with plant tumors. If so, then a signal transduction system must exist that senses compounds associated specifically with transformed plant cells. One class of tumor-specific exudates that could serve as signal molecules is opines.

Indeed, opines are already known to induce specific bacterial genes. First, at least some opines induce the genes that are required for their own transport and catabolism (Montoya *et al.* 1977; Klapwijk *et al.* 1977). Second, certain opines (called conjugal opines) induce genes required for conjugal transfer of the Ti plasmid. These opine catabolism genes and *tra* genes were reported to be controlled by the same negative regulatory system (Ellis *et al.* 1982; Klapwijk *et al.* 1978; Klapwijk and Schilperoort 1979). Third, by screening random Tn*5 gus* fusions for inducibility by the opine octopine (an arginine-pyruvate condensate), we have recently obtained evidence that additional genes of unknown function are inducible by this opine (unpublished observations). Furthermore, opines also potentiate the induction of genes required for T-DNA transfer (Veluthambi *et al.* 1989).

In an effort to understand an octopine-responsive regulatory system, we set out to identify the genes required for the regulated expression of the octopine catabolic operon. This opine was chosen for study primarily because it is a conjugal opine for a variety of Ti plasmids, and therefore is a stronger candidate to regulate other genes than a nonconjugal opine would be. Furthermore, octopine is commercially available and can be readily synthesized enzymatically using radiolabeled precursors, which will be useful in future studies. In apparent contradiction to earlier studies, we report here that octopine catabolic genes are under positive regulation. The single protein required, designated OccR, is a member of the LysR family of regulatory proteins. In addition to inducing the *occ* catabolic operon in the presence of octopine, OccR also negatively regulates its own gene, both in the presence and in the absence of octopine.

Address correspondence to Stephen C. Winans: Section of Microbiology, Wing Hall, Cornell University, Ithaca, NY 14853.

Nucleotide and/or amino acid sequence data submitted to GenBank, EMBL, and DDBJ as accession number JO3691.

© 1991 The American Phytopathological Society

MATERIALS AND METHODS

Strains and plasmids. *Escherichia coli* strain MC4100 (Δ*lac, araD, thiA, rpsL, relA*) was obtained from C. Manoil, Dept. of Genetics, University of Washington. Plasmid pVK261 (Knauf and Nester 1982) was obtained from E. W. Nester, Dept. of Microbiology, University of Washington. Plasmid pSW208 was made by digesting pTZ18R and pACYC184 (Chang and Cohen 1978) with *Hae*II, ligating, and isolating a plasmid containing the α-complementation group of pTZ18R, the *cat* gene of pACYC184, and the origin of replication of pACYC184 (data not shown). The high copy number, Ap^r plasmids pTZ18R and pTZ19R were purchased from U.S. Biochemicals, Cleveland, OH. pMC1403 (Casadaban *et al.* 1980) was obtained from S. Lory, Dept. of Microbiology, University of Washington. *E. coli* strain POII1734 containing Mu dII1734 was obtained from M. Casadaban (Castilho *et al.* 1984) and used according to published procedures.

Reagents. Antibiotics, X-Gal, *o*-nitrophenyl β-D-galactopyranoside, and 2-[*N*-morpholino]ethanesulfonic acid were purchased from Sigma Chemical Co., St. Louis, MO. Restriction endonucleases and T4 DNA ligase were purchased from Besthesda Research Laboratories, Gaithersburg, MD. Octopine was purchased from Aldrich Co., Milwaukee, WI. Sequenase DNA sequencing kits were purchased from U. S. Biochemicals. [^{35}S]dATP was purchased from NEN Research Products, Boston, MA.

Induction of *occ-lacZ* gene fusions. Derivatives of *E. coli* strain MC4100 containing the indicated plasmids were grown overnight at 37° C in 2 ml of AB minimal glucose medium (Chilton *et al.* 1974) supplemented with appropriate antibiotics (100 μg of carbenicillin, 50 μg of kanamycin, and/or 50 μg of chloramphenicol per milliliter). The saturated cultures were diluted 100-fold into 10 ml of fresh AB medium without antibiotics. Cultures were grown at 37° C until reaching an $O.D._{600}$ of 0.2. Cultures were then divided into two tubes, and octopine (100 μg/

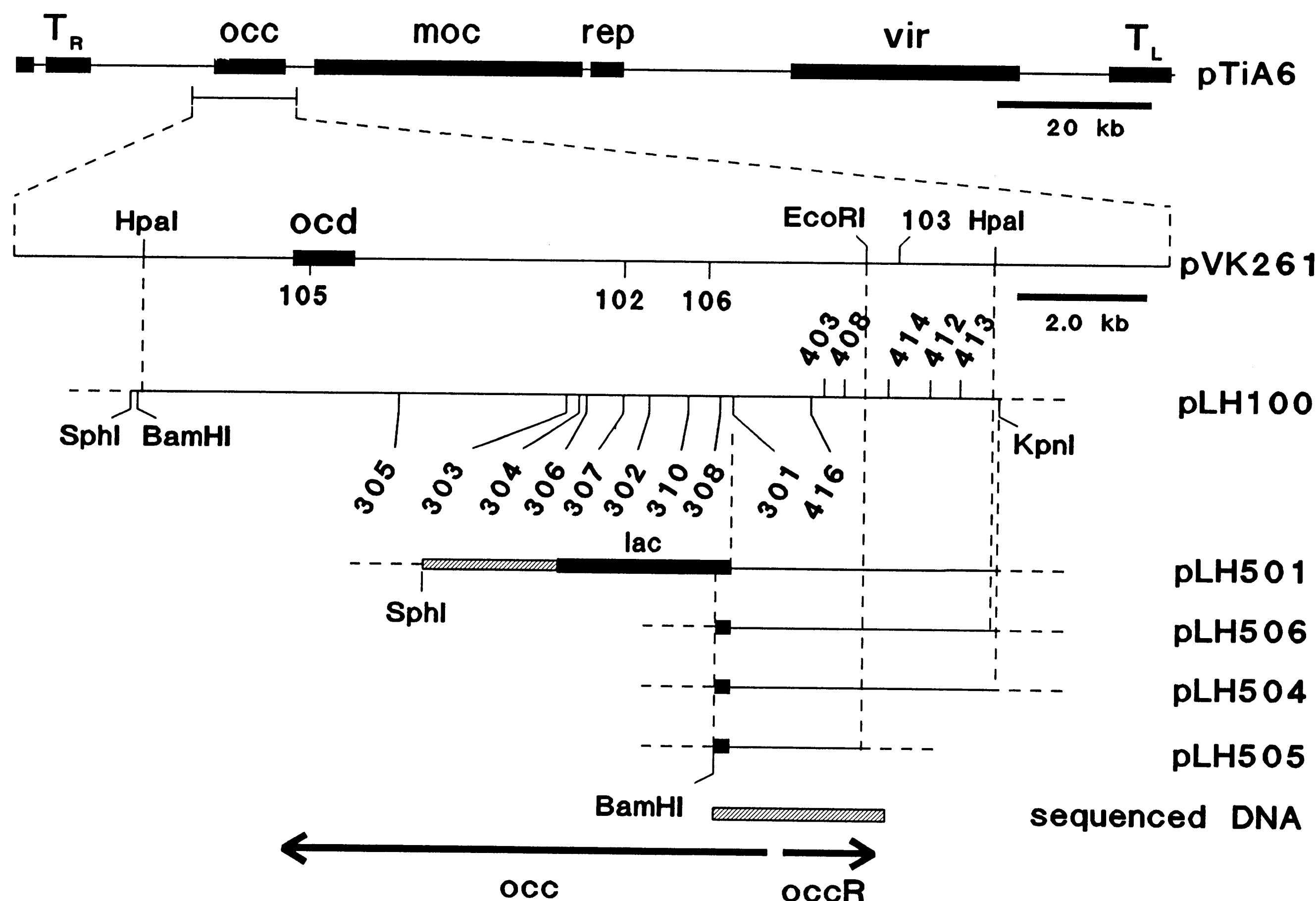

Fig. 1. Physical maps of plasmids used in this study. Insertions Ω102, Ω103, Ω105, and Ω106 in pVK261 were made using Tn*3*HoHo1 (Stachel *et al.* 1985). pLH100 is a construct containing the indicated *Hpa*I fragment cloned into the *Sma*I site of pSW208. Insertions Ω301 to Ω416 were made using Mu dII1734. Insertions shown above the horizontal line are oriented such that *lacZ* is expressed from left to right, while insertions shown below the line have the opposite orientation. Insertions Ω301 to Ω310 create *occ-lacZ* fusions, while insertions Ω403 and Ω408 create *occR-lacZ* fusions, and insertions Ω412, Ω413, Ω414, and Ω416 do not create fusion proteins. Insertin Ω416 disrupts *occR*. pLH501 is a deletion derivative of pLH100 Ω301 made by removing the DNA between the two *Sph*I sites shown. Cross-hatched and filled regions of pLH501 represent Mu dII1734 DNA. pLH506 is a similar deletion of pLH100 Ω301 made using *Bam*HI. pLH504 is a derivative of pTZ19R containing the indicated *Bam*HI-*Kpn*I fragment. pLH505 is a derivative of pMC1403 containing the indicated *Bam*Hi-*Eco*RI fragment. pMC1403 is an Ap^R derivative of pBR322 containing the *lac* operon beginning at the eighth codon of *lacZ*. Horizontal dashed lines indicate vector sequences. Abbreviations: T_R and T_L: "right" and "left" transferred DNAs, respectively (Thomashow *et al.* 1980); *occ:* octopine catabolism (Stachel *et al.* 1985); *moc*: mannityl opine catabolism (Dessaux *et al.* 1987); *rep:* replication origin (Tabata *et al.* 1989); *vir:* virulence genes required for T-DNA transfer (Stachel and Nester 1986); *ocd:* ornithine cyclodeaminase (Sans *et al.* 1988).

ml final concentration) was added to one set. At 30-min intervals, the turbidity of each culture was measured using a Spectronic 20 spectrophotometer (Bausch and Lomb) and 0.5 ml was withdrawn and frozen at −70° C. β-Galactosidase assays (Miller 1972) were performed on each sample at the conclusion of the experiment.

DNA sequencing. The *Bam*HI-*Kpn*I fragment shown in Figure 1 was cloned into pTZ18R and pTZ19R. A set of nested deletions of each were created using the Erase-a-base kit (Promega Biotech, Madison, WI) and in some cases by using existing restriction sites. Regions of DNA not covered by these deletions were sequenced by synthesizing oligonucleotides. DNA to be sequenced was isolated in single-stranded form using helper phage M13KO7 (Dente *et al.* 1983), and sequenced using the Sequenase kit (U.S. Biochemicals) according to the recommendations of the manufacturer. DNA sequences were analyzed using the computer programs Gene-Pro (Riverside Scientific Enterprises, Bainbridge Island, WA) and FASTA (Pearson and Lipman 1988).

RESULTS

Localization of the gene required for regulation of octopine catabolism genes. Plasmid pVK261 is an IncP1 cosmid containing 17.9 kb of DNA from the Ti plasmid pTiA6 (Knauf and Nester 1982). When introduced into an *Agrobacterium* host lacking a Ti plasmid, pVK261 confers the ability to catabolize octopine (Knauf and Nester 1982). The catabolic enzymes are inducible by octopine, even in hosts lacking the Ti plasmid (Stachel *et al.* 1985), indicating that all necessary regulatory genes must also be present on this clone (or less likely, on the *Agrobacterium* chromosome). In a previous study, a series of fusions between octopine degradation genes and *lacZ* were generated in pVK261 using Tn*3*HoHo1 (Stachel *et al.* 1985). The locus of insertion of some of these transposons is shown in Figure 1. Insertions Ω102, Ω105, and Ω106 abolished octopine utilization and were inducible by octopine in *Agrobacterium* strains containing or lacking the Ti plasmid (Stachel *et al.* 1985). These fusions were also inducible by octopine in *E. coli*, suggesting even more strongly that this fragment carries all regulatory genes and indicating that the regulatory system is fully functional in this heterologous host.

We digested pVK261 with *Hpa*I and cloned a 12.8-kb *Hpa*I fragment (De Vos *et al.* 1981) into the *Sma*I site of plasmid pSW208. The resulting plasmid, pLH100 (Fig. 1) was mutagenized with Mu dII1734, a transposon that can create translational fusions between *lacZ* and any gene into which it inserts. Two hundred strains of MC4100 containing Mu dII1734 insertion derivatives of pLH100 were screened for octopine-inducible expression of β-galactosidase by stabbing them onto AB agar containing 40 μg of X-Gal and 100 μg of octopine per milliliter (similar plates lacking octopine were used as controls). Nine such fusions were obtained. The induced and uninduced levels of β-galactosidase of insertion Ω301 is shown in Figure 2. The map positions of these inducible insertions are shown in Figure 1 (insertions Ω301 to Ω310 inclusive), and the basal and induced levels of expression after 2 hr of incubation without and with octopine, respectively, are shown in Table 1. The large differences in basal and induced levels could be due to differences in translational efficiency or catalytic activity of these different fusion proteins. Nevertheless, these data indicated that pLH100 contains all genes needed for regulated expression of the catabolic operon. All inducible fusions are oriented such that transcription occurs from right to left as shown in Figure 1. Two insertions, Ω403 and Ω408, shown in Figure 1, were

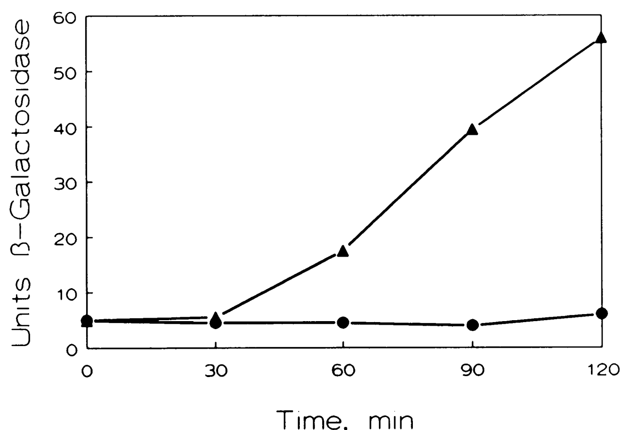

Fig. 2. Kinetics of induction of the *occ-lacZ* fusion of pLH100 Ω301. ▲: 100 μg/ml octopine added; ●: no octopine. Procedures are described in Materials and Methods.

Table 1. β-Galactosidase synthesis of *occ*::*lac* gene fusions in response to octopine[a]

Plasmid	β-Galactosidase activity[b]	
	− Octopine	+ Octopine
pLH301[c]	5.3	60.0
pLH302	3.4	36.4
pLH303	1.6	13.5
pLH304	0.7	3.9
pLH305	0.3	11.9
pLH306	0.9	16.4
pLH307	0.4	3.6
pLH308	1.0	16.0
pLH310	0.7	17.4

[a] Strains of MC4100 containing the indicated plasmids were assayed for β-galactosidase activity 2 h after addition of octopine, as described in Materials and Methods.
[b] Miller units (Miller 1972).
[c] Derivatives of pLH100 containing Mu dII1734. Sites of insertions are shown in Figure 1.

Table 2. Negative autoregulation of *occR*[a]

Plasmids	β-Galactosidase[b]	
	− Octopine	+ Octopine
pTZ19R, pLH403[c]	8.9	8.6
pLH504, pLH403	0.6	0.6
pTZ19R, pLH408	7.4	7.6
pLH504, pLH408	1.8	1.2

[a] Assay conditions are the same as in Figure 1.
[b] Miller units (Miller 1972).
[c] pLH403 and pLH408 contain *occR-lacZ* gene fusions. pLH504 expresses *occR*, while pTZ19R is a vector control not expressing *occR*.

isolated, which constitutively synthesized β-galactosidase (see Table 2). Insertions Ω412, Ω413, Ω414, and Ω416, also shown in Figure 1, do not create gene fusions, and were isolated by determining the map positions of a larger number of transposon derivatives of pLH100.

We reasoned that, regardless how many octopine-inducible promoters might exist on this plasmid, at least one must be located to the right of the right-most inducible insertion. Because the genes for regulatory proteins are often located close to the promoters they regulate, it was a reasonable hypothesis that the regulatory gene(s) might lie at the right end of the *Hpa*I fragment. Therefore, we digested pLH100 Ω301 with *Sph*I, which cuts at the left end of Mu dII1734 (as drawn) and in the extreme left end of the parent plasmid, but not within the vector or to the right of the insertion. These digestion products were ligated upon themselves, and introduced into *E. coli* strain MC4100. The resulting plasmid, pLH501 (Figure 1), showed a restriction pattern consistent with a deletion of the DNA between the *Sph*I sites, and as expected, had lost the kanamycin-resistance gene of Mu dII1734. pLH501 conferred upon its host octopine-inducible expression of β-galactosidase, indicating that all necessary regulatory genes must lie to the right of the Mu dII1734 insertion in pLH100 Ω301.

To localize the regulatory gene more precisely, we wanted to create a plasmid that contained the inducible promoter but not the regulatory gene. We subcloned the fragment of pLH501 containing the inducible promoter as an *Eco*RI-*Bam*HI fragment into plasmid pMC1403, creating pLH505 (Fig. 1). pMC1403 contains the *E. coli lac* operon beginning at the eighth codon of *lacZ* and contains cleavage sites for *Eco*RI, *Sma*I, and *Bam*HI directly upstream. The *Bam*HI site of Mu dII1734 and of pMC1403 are in the same reading frame. MC4100(pSW208)(pLH505) expressed β-galactosidase at low levels and was not inducible by octopine (Table 3). This result suggested that we had succeeded in separating the regulated promoter from the regulatory gene, and that the regulatory protein present in pLH100 must have been an activator rather than a repressor of transcription.

Table 3. Transcriptional activation of the *occ* catabolic promoter of pLH505 in merodiploid strains expressing OccR[a]

	β-Galactosidase activity[b]	
Plasmids	**− Octopine**	**+ Octopine**
pSW208, pLH505[c]	5.2	6
pLH100, pLH505	3.7	39.3
pLH506, pLH505	3.4	31.5
pLH412	0.7	0.8
pLH412, pLH505	4.9	27.9
pLH413	0.5	0.5
pLH413, pLH505	2.6	22.1
pLH414	1.0	0.9
pLH414, pLH505	4.1	28.3
pLH416	0.4	0.4
pLH416, pLH505	2.7	2.6

[a] Assay conditions are the same as in Table 1.
[b] Miller units (Miller 1972).
[c] pLH505 contains an *occ-lacZ* gene fusion. The other plasmids shown on lines 3–13 are insertion or deletion derivatives of pLH100.

To confirm these results and to localize the regulatory genes required for induction, pLH505 was introduced by transformation into derivatives of MC4100 containing pLH100 or its various insertion derivatives. Inducible expression was restored by pLH100 (Table 3), indicating that the inducible promoter is controlled by a transcriptional activator. We also tested whether pLH506 (a derivative of pLH100 Ω301 made by deleting across the two *Bam*HI sites shown in Fig. 1) could also restore induction. As shown in Table 3, pLH506 does fully restore inducible expression, indicating that it must contain all necessary regulatory genes. To localize the genes required for regulation, we introduced pLH505 into various Mu dII1734 insertion derivatives of pLH100. We chose only plasmids containing insertions (Ω412, Ω413, Ω414, and Ω416) that had insertions that did not result in β-galactosidase fusions. Each strain was screened for octopine-inducible β-galactosidase

```
CAT AGCGTTCCCTGCGCGCGACCGTTTCGCCGCACGTCTGATCCCGGTTGGCCCGGAAG   60
                                 -35                -10
AGTTGTGATTTATAGGGTGAGCTTTACATTCGATATGCATTCGGTCAAATTCATAATGAC  120

CGGGCAAGAATAAGCAGATGTTATGGTGCGTTTATGAATCTCAGGCAGGTCGAGGCGTTC  180
                                 M  N  L  R  Q  V  E  A  F

CGGGCAGTCATGCTGACGGGGCAAATGACGGCGGCGGCTGAACTAATGCTGGTGACTCAG  240
R  A  V  M  L  T  G  Q  M  T  A  A  A  E  L  M  L  V  T  Q

CCGGCCATCAGTCGCCTAATCAAGGACTTTGAACAGGCGACAAAACTGCAGCTCTTCGAG  300
P  A  I  S  R  L  I  K  D  F  E  Q  A  T  K  L  Q  L  F  E

AGGCGTGGGAACCATATTATCCCGACACAGGAGGCAAAGACGCTGTGGAAAGAGGTCGAT  360
R  R  G  N  H  I  I  P  T  Q  E  A  K  T  L  W  K  E  V  D

CGGGCGTTCGTCGGGCTTAATCATATAGGCAACCTGGCTGCCGACATCGGCAGGCAGGCA  420
R  A  F  V  G  L  N  H  I  G  N  L  A  A  D  I  G  R  Q  A

GCGGGGACGCTCCGCATTGCTGCAATGCCTGCTCTGGCAAACGGCCTCTTGCCGCGGTTT  480
A  G  T  L  R  I  A  A  M  P  A  L  A  N  G  L  L  P  R  F

CTTGCTCAGTTCATCCGTGACAGACCAAATCTCCAGGTCTCCCTAATGGGACTGCCCTCA  540
L  A  Q  F  I  R  D  R  P  N  L  Q  V  S  L  M  G  L  P  S

AGCATGGTCATGGAAGCCGTTGCGTCCGGCAGGGCCGACATCGGTTATGCCGATGGCCCA  600
S  M  V  M  E  A  V  A  S  G  R  A  D  I  G  Y  A  D  G  P

CAGGAGCGCCAAGGTTTTCTAATCGAAACCCGGTCGCTTCCCGCTGTTGTCGCTGTCCCG  660
Q  E  R  Q  G  F  L  I  E  T  R  S  L  P  A  V  V  A  V  P

ATGGGACATCGACTTGCTGGCCTTGACCGTGTCACGCCACAGGACCTTGCCGGTGAGCGT  720
M  G  H  R  L  A  G  L  D  R  V  T  P  Q  D  L  A  G  E  R

ATTATAAAACAGGAGACTGGCACTCTCTTCGCCATGCGGGTAGAGGTGGCGATTGGTGGT  780
I  I  K  Q  E  T  G  T  L  F  A  M  R  V  E  V  A  I  G  G

ATTCAACGCCGGCCGTCAATTGAAGTGAGCCTGTCGCATACTGCGCTAAGTCTCGTCCGC  840
I  Q  R  R  P  S  I  E  V  S  L  S  H  T  A  L  S  L  V  R

GAAGGCGCCGGGATCGCAATTATCGATCCAGCCGCGGCGATCGAGTTCACGGACAGGATC  900
E  G  A  G  I  A  I  I  D  P  A  A  A  I  E  F  T  D  R  I

GTACTGCGACCGTTCTCGATCTTCATTGACGCCGGATTCCTCGAAGTCCGGTCAGCAATT  960
V  L  R  P  F  S  I  F  I  D  A  G  F  L  E  V  R  S  A  I

GGCGCTCCCTCAACCATCGTCGATCGTTTCACAACCGAATTCTGGAGGTTTCATGATGAC 1020
G  A  P  S  T  I  V  D  R  F  T  T  E  F  W  R  F  H  D  D

TTGATGAAGCAGAACGGCCTAATGGAGTAATGCATATTCGGCATCTTGTGTAGGTGGGTC 1080
L  M  K  Q  N  G  L  M  E  *

GATTAGACGTCACAATGAGAAAATGCCGTCGGCAAGCGAGGCCCGCACCGATTTCGAGTT 1140

AATAGTTCCGCGCGG                                              1155
```

Fig. 3. Nucleotide and inferred amino acid sequence of the *occR* regulatory gene. The DNA sequence shown includes 153 bp of DNA upstream of the putative ATG initiation codon. The first three bases of the sequence (CAT) correspond to the putative initiation codon of a divergently transcribed gene which is regulated by octopine (data not shown). Possible -10, -35, and Shine and Dalgano sequences are underlined.

activity, and the results are shown in Table 3. These results indicated that one or more regulatory genes required for induction were located between insertions Ω414 and Ω301. It remains a formal possibility, however, that additional regulatory genes could exist to the right of Ω414 that are not disrupted by insertions Ω412, Ω413, or Ω414. Insertion Ω416 appears to disrupt at least one gene required for regulation.

Nucleotide sequence of the regulatory gene. We then determined the nucleotide sequence of a 2.2-kb region of DNA to the right of insertion Ω301. The DNA sequence and the inferred amino acid sequence of a portion of this region is shown on Figure 3. We find one open reading frame reading from left to right (as drawn in Fig. 1), which could encode a hydrophilic protein of 298 amino acids, having a molecular weight of 32,666, and having a net charge of +0.5 at pH 7. We find a putative ribosome binding site upstream of the putative ATG initiation codon. This open reading frame lies in the region disrupted by insertion Ω416 but not by insertions Ω412, Ω413, or Ω414, and was truncated in pLH505. It is therefore highly probable that this open reading frame corresponds to the regulatory locus. We designate this gene *occR* (for *oc*topine *c*atabolism *R*egulator). The FASTA program of Lipman and Pearson was used to compare the amino acid sequence of this protein to that of all proteins in the Protein Information Resource database (release 26) of the National Biomedical Research Foundation. We found a strong similarity between this protein and many members of the LysR family of prokaryotic transcriptional regulatory proteins (Henikoff *et al.* 1988). The greatest similarity was found with LysR itself (Fig. 4), although OccR was strongly similar to all members of the family (data not shown). The similarity

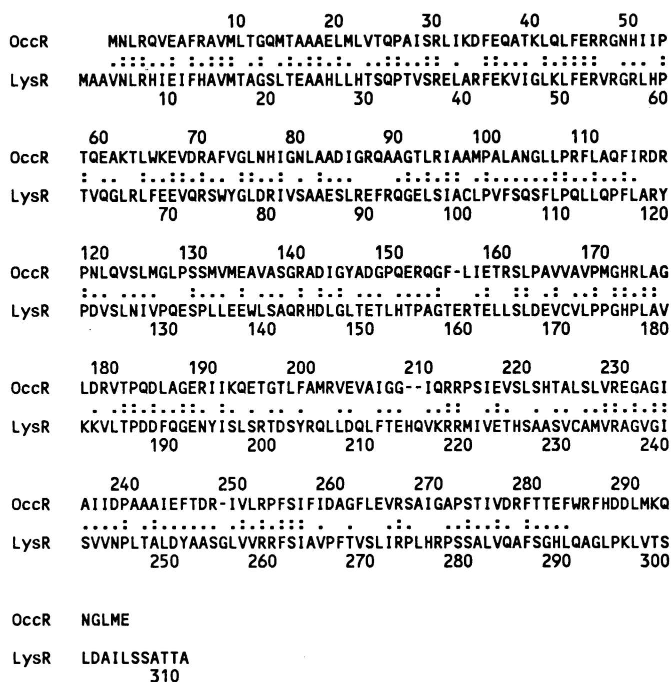

Fig. 4. Alignment of OccR and LysR as obtained from the FASTA program. Colons indicate identical residues, while periods indicate conservative substitutions.

appears strongest close to the amino termini of the proteins, which contain putative DNA binding sites (Henikoff *et al.* 1988), and considerably weaker in the carboxyl terminal half of the two proteins, which contain sites for the binding of coinducers. It remains formally possible that DNA to the left of *occR* may also contain genes required for regulation. However, in this region, we found an open reading frame transcribed divergently from *occR* that is homologous to a family of amino acid permeases (data not shown), suggesting that it is not required for transcriptional regulation.

Transcriptional regulation of *occR*. We sought to determine whether the *occR* gene was itself expressed in a regulated fashion. Two insertions of Mu dII1734 showing constitutive expression of β-galactosidase (insertions Ω403 and Ω408) were described above. Our sequence indicates that these insertions are within the *occR* structural gene, and the observed β-galactosidase expression must be due to fusions to this gene. However, because these insertions would also have disrupted the gene, no functional OccR protein was made in those strains. We supplied the wild-type allele of *occR* in a merodiploid strain by using pLH504 (a derivative of pTZ19R containing the indicated region cloned as a *Bam*HI-*Kpn*I fragment as shown in Fig. 1). We introduced pLH504 into strains of *E. coli* containing pLH100 Ω403 or pLH100 Ω408. Strains containing a functional *occR* gene expressed an *occR*::*lacZ* fusion at about 10-fold lower levels than strains lacking *occR* (Table 2). Repression was not influenced by the presence or absence of octopine. These data indicate that OccR protein is required for negative regulation of the *occR* gene, and that this autoregulation does not require octopine. These data do not preclude the possibility that other pLH100-encoded genes may also be required to regulate *occR*.

DISCUSSION

We have identified a region of *Agrobacterium* Ti plasmid DNA sufficient to confer octopine-inducible expression of the octopine catabolism pathway. This region lies to the right of insertion Ω301 and most likely to the left of insertion Ω414. In addition, this regulatory locus is disrupted by insertion Ω416 and is truncated in pLH505. Sequencing this DNA revealed one gene, designated *occR*, which fits all these criteria. OccR is homologous to the LysR family of transcriptional regulatory proteins. This protein positively regulates the octopine catabolic operon in the presence of octopine, and represses its own gene in the presence or absence of octopine.

Earlier studies have shown that octopine degradation and conjugal transfer genes are coregulated (Klapwijk *et al.* 1977; Klapwijk and Schilperoort 1979). A number of regulatory mutations that cause an Occ^c, Tra^c phenotype were recessive to the wild-type genes, suggesting that *occ* and *tra* genes are under negative regulation (Klapwijk and Schilperoort 1979). However, a regulatory mutation causing an Occ$^-$, Tra$^-$ phenotype was also recessive to wild type (Klapwijk and Schilperoort 1979), suggesting positive regulation. These two sets of data are somewhat difficult to reconcile with each other and the second set are far

easier to reconcile with the results presented in this study. There is now evidence that at least two proteins homologous to OccR have a tetrameric quaternary structure (Chang and Crawford 1990; Schell *et al.* 1990), suggesting a similar structure for OccR. If so, perhaps the Occ^c, Tra^c strains described above actually contained mutant activators that are constitutively active, but that lose this property when mixed with wild-type subunits. Because the plasmid supplying wild-type gene products contained an IncP origin (which has a higher copy number than the Ti plasmid) these merodiploid strains probably contained considerably more wild-type protein than mutant protein. This could explain the recessive nature of the Occ^c, Tra^c regulatory mutations.

Of the many proteins homologous to LysR, a subset has been characterized more thoroughly than the rest. These include OxyR from *E. coli*, which regulates the oxidative stress response (Storz *et al.* 1990), NodD from several *Rhizobium* species, which regulate those organisms' nodulation genes (Long 1989), NahR from the *Pseudomonas aeruginosa* (Schroeter) Migula Tol plasmid, which regulates genes required for degradation of aromatic hydrocarbons (Schell and Poser 1989), and TrpI from *P. aeruginosa*, which regulates the *trpAB* biosynthetic operon (Chang and Crawford 1990). The OxyR protein is activated by an oxidized intracellular environment, while NodD is activated by flavones and isoflavones released from the roots of leguminous plants. The NahR protein is activated by aromatic hydrocarbons and finally, TrpI is activated by indole acetamide. Each of these proteins has been shown to bind to the promoters that they regulate, and binding occurs both in the presence or absence of coinducers. In the case of OxyR and TrpI, however, inducing conditions result in alterations in the binding of the protein to DNA (as measured in a footprint assay), while activators of NodD or NahR do not detectably alter the binding of these proteins. LysR, OxyR, TrpI, and some but not all NodD proteins also repress their own gene (Chang and Crawford 1990; Stragier and Patte, 1983; Christman *et al.* 1989; Rossen *et al.* 1985; Schell and Faris 1987). In most cases the regulatory gene is transcribed divergently from the inducible operon, and regulation of both promoters may involve a single binding site.

By analogy, we predict that OccR binds to the promoters that it regulates, and that it also binds to octopine and related opines. We predict further that binding to promoters should occur in the presence or absence of octopine and related opines, but that these opines may cause alterations in the contacts between OccR and DNA. The fact that OccR autoregulates both in the presence and absence of octopine indicates that if it binds to DNA, it must be able to do so independent of octopine availability. The sequence shown in Figure 3 starts with the sequence "CAT," which corresponds to the probable ATG initiation codon of the divergent catabolic operon (unpublished data). Therefore, the sequence shown may contain all DNA required for both activation of the catabolic operon and for negative autoregulation. Given the limited amount of DNA between these two open reading frames, if OccR does bind to these promoters, then this protein could both activate the catabolic operon and repress its own gene by binding to a single site somewhere in this region.

The member of the LysR family most similar to OccR is LysR itself (data not shown). This could in part be due to the fact that OccR is responsive to compounds that are somewhat similar to those recognized by LysR. *occ* and *tra* genes can be induced by octopine, octopinic acid (an ornithine-pyruvate conjugate), and lysopine (a lysine-pyruvate conjugate), suggesting that OccR is activated by all three of these compounds (Klapwijk *et al.* 1978, Klapwijk and Schilperoort 1979). LysR is activated by diaminopimelic acid, which is an intermediate in lysine biosynthesis (Stragier and Patte 1983).

A related Ti plasmid directs plant tumors to synthesize nopaline (an arginine-α-ketoglutarate conjugate) rather than octopine. The genes required for nopaline catabolism are controlled by a transcriptional activator, NocR, which has an amino acid sequence 37% identical to OccR (J. Schroder, personal communication; D. White, personal communication). In contrast, the catabolism of a third opine, agrocinopine A/B (made by conjugation of sucrose or fructose with arabinose), is regulated by a repressor protein (Ellis *et al.* 1982) that is homologous to the family of regulatory proteins including the *E. coli* FucR protein (S. Farrand, personal communication). It is striking that this protein should be evolutionarily related, not to OccR or NocR, but rather to a different family of prokaryotic regulatory genes. These findings indicate that at least some genes required for opine catabolism have been recruited from genes having other roles on two or more occasions.

ACKNOWLEDGMENTS

We thank N. Mantis, C.-Y. Chen, and C.-H. Chang for helpful discussions and critical appraisal of the manuscript. We also thank S. Farrand, J. Schroder, and D. White for openly sharing unpublished data. This work was supported by National Institutes of Health grant 1 R29 GM2893-01. This work was also supported by a grant from the Cornell Biotechnology Program, which is sponsored by the New York State Science and Technology Foundation, a consortium of industries, the U.S. Army Research Office, and the National Science Foundation.

LITERATURE CITED

Binns, A. N., and Thomashow, M. F. 1988. Cell biology of *Agrobacterium* infection and transformation of plants. Annu. Rev. Microbiol. 42:575-606.

Casadaban, M. J., Chou, J., and Cohen, S. N. 1980. In vitro gene fusions that join an enzymatically active β-galactosidase segment to amino terminal fragments of exogenous proteins: *E. coli* plasmid vectors for the detection and cloning of translational initiation signals. J. Bacteriol. 143:971-980.

Castilho, B. A., Olfson, P., and Casadaban, M. J. 1984. Plasmid insertion mutagenesis and *lac* gene fusion with mini-Mu bacteriophage transposons. J. Bacteriol. 158:488-495.

Chang, A. C. Y., and Cohen, S. N. 1978. Construction and characterization of amplifiable DNA cloning vehicles derived from the P15A cryptic miniplasmid. J. Bacteriol. 134:1141-1156.

Chang, M., and Crawford, I. P. 1990. The roles of indoleglycerol phosphate and the TrpI protein in the expression of *trpBA* from *Pseudomonas aeruginosa.* Nucleic Acids Res. 18:979-988.

Chilton, M.-D., Currier, T. C., Farrand, S. K., Bendich, A. J., Gordon, M. P., and Nester, E. W. 1974. Agrobacterium DNA and PS8 bacteriophage DNA not detected in crown gall tumors. Proc. Natl. Acad. Sci. USA 71:3672-3676.

Chilton, M.-D., Saiki, R. K.,Yadav, N., Gordon, M. P., and Quetier, F. 1980. T-DNA from *Agrobacterium* Ti plasmid is in the nuclear DNA fraction of crown gall tumor cells. Proc. Natl. Acad. Sci. USA 77:4060-4064.

Christman, M. F., Storz, G., and Ames, B. N. 1989. OxyR, a positive regulator of hydrogen peroxide-inducible genes in *Escherichia coli* and *Salmonella typhimurium*, is homologous to a family of bacterial regulatory proteins. Proc. Natl. Acad. Sci. USA 86:3484-3488.

Dessaux, Y., Tempe, J., and Farrand, S. K. 1987. Genetic analysis of mannityl opine catabolism in octopine-type *Agrobacterium tumefaciens* strain 15955. Mol. Gen. Genet. 208:301-308.

Dente, L., Cesareni, G., and Cortese, R. 1983. pEMBL: A new family of single stranded plasmids. Nucleic Acids Res. 11:1645-1655.

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.

Ellis, J. G., Kerr. A., Petit, A., and Tempe, J. 1982. Conjugal transfer of nopaline and agropine Ti-plasmids—The role of agrocinopines. Mol. Gen. Genet. 186:269-273.

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.

Klapwijk, P. M., Oudshoorn, M., and Schilperoort. R. A. 1977. Inducible permease involved in the uptake of octopine, lysopine and octopinic acid by *Agrobacterium tumefaciens* strains carrying virulence-associated plasmids. J. Gen. Microbiol. 102:1-11.

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

Klapwijk, P. M., and Schilperoort, R. A. 1979. Negative control of octopine degradation and transfer genes of octopine Ti plasmids in *Agrobacterium tumefaciens*. J. Bacteriol. 141:424-431.

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.

Long, S. R. 1989. *Rhizobium*-legume nodulation: Life together in the underground. Cell 56:203-214.

Miller, J. H. 1972. Experiments in Molecular Genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Montoya, A., Chilton, M. D., Gordon, M. P., Sciaky, D., and Nester, E. W. 1977. Octopine and nopaline metabolism in *Agrobacterium tumefaciens* and crown gall tumor cells: Role of plasmid genes. J. Bacteriol. 129:101-107.

Pearson, W. R., and Lipman, D. J. 1988. Improved tools for biological sequence comparison. Proc. Natl. Acad. Sci. USA. 85:2444-2448.

Petit, A., and Tempe, J. 1978. Isolation of *Agrobacterium* Ti-plasmid regulatory mutants. Mol. Gen. Genet. 167:147-155.

Petit, A., and Tempe, J. 1985. The function of T-DNA in nature. Pages 625-636 in: Molecular Form and Function of the Plant Genome. L. van Vloten-Doting, G. Groot, and T. Hall, eds. Plenum, New York.

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

Rossen, L., Shearman, C. A., Johnston, A. W. B., and Downie, J. A. 1985. The *nodD* gene of *R. leguminosarum* is autoregulatory and in the presence of plant exudate induces the *nodABC* genes. EMBO J. 4:3369-3374.

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

Schell, M. A., Brown, P. H., and Raju, S. 1990. Use of saturation mutagenesis to localize probable functional domains in the NahR protein, a LysR-type transcription activator. J. Biol. Chem. 265:3844-3850.

Schell, M. A., and Faris, E. 1987. Transcriptional regulation of the *nah* and *sal* naphthaline degradation operons of plasmid NAH7 of *Pseudomonas putida*. Pages 455-458 in: Proc. Steenbock Symposium, 16th. Madison, WI. W. Reznikoff, R. Burgess, C. Gross, and M. Record, ed. Elsevier Science Publishing, Inc., New York.

Schell, M. A., and Poser, E. F. 1989. Demonstration, characterization, and mutational analysis of NahR protein binding to *nah* and *sal* promoters. J. Bacteriol. 171:837-846.

Stachel, S. E., An, G., Flores, C., and Nester, E. W. 1985. A Tn*3* *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.

Stachel, S. E., and Nester, E. W. 1986. The genetic and transcriptional organization of the *vir* region of the A6 Ti plasmid of *Agrobacterium tumefaciens*. EMBO J. 5:1445-1454.

Stachel, S. E., and Zambryski, P. C. 1986. *Agrobacterium tumefaciens* and the susceptible plant cell: A novel adaptation of extracellular recognition and DNA conjugation. Cell 47:155-157.

Storz, G., Tartaglia, L. A., and Ames, B. A. 1990. Transcriptional regulator genes of oxidative stress-inducible genes: Direct activation by oxidation. Science 248:189-194.

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.

Tabata, S., Hooykaas, P. J., and Oka, A. 1989. Sequence determination and characterization of the replicator region in the tumor-inducing plasmid pTiB6S3. J. Bacteriol. 171:1665-1672.

Thomashow, M. F., Nutter, R., Montoya, A. L., Gordon, M. P., and Nester, E. W. 1980. Integration and organization of Ti plasmid sequences in crown gall tumors. Cell 19:729-739.

Veluthambi, K., Krishman, M., Gould, J. H., Smith, R. H., and Gelvin, S. B. 1989. Opines stimulate induction of the *vir* genes of *Agrobacterium tumefaciens* Ti plasmid. J. Bacteriol. 171:3696-3703.

White, D. W. R., Pritchard, M., and Marincs, F. 1990. Genetic and sequence analysis of *Agrobacterium* opine catabolism genes. Page 53 in: Int. Symp. Mol. Genetics Plant-Microbe Interact., 5th. M. Gottfert, H. Hennecke, and H. Paul, eds. Eidgenossische Technische Hochschule, Zurich.

Winans, S. C., Ebert, P. R., Stachel, S. E., Gordon, M. P., and Nester, E. W. 1986. A gene essential for *Agrobacterium* virulence is homologous to a family of positive regulatory loci. Proc. Natl. Acad. Sci. USA 83:8278-8282.

Zambryski, P. 1988. Basic processes underlying *Agrobacterium*-mediated DNA transfer to plant cells. Ann. Rev. Genet. 22:1-30.