

Genetic and Molecular Analyses of *picA*, a Plant-Inducible Locus on the *Agrobacterium tumefaciens* Chromosome

LIJUN RONG, SUSAN J. KARCHER, AND STANTON B. GELVIN*

Department of Biological Sciences, Purdue University, West Lafayette, Indiana 47907

Received 28 February 1991/Accepted 13 June 1991

picA is an *Agrobacterium tumefaciens* chromosomal locus, identified by Mu dI1681 mutagenesis, that is inducible by certain acidic polysaccharides found in carrot root extract. Cloning and genetic analysis of a *picA::lacZ* fusion defined a region of the *picA* promoter that is responsible for the induction of this locus. Furthermore, we identified a possible negative regulator of *picA* expression upstream of the *picA* locus. This sequence, denoted *pgl*, has extensive homology to polygalacturonase genes from several organisms and inhibited the induction of the *picA* promoter when present in multiple copies in *A. tumefaciens*. DNA sequence analysis indicated at least two long open reading frames (ORFs) in the *picA* region. S1 nuclease mapping was used to identify the transcription initiation site of *picA*. Mutation of ORF1, but not ORF2, of the *picA* locus was responsible for an increased aggregation of *A. tumefaciens*, forming "ropes" in the presence of pea root cap cells. In addition, a potato tuber disk virulence assay indicated that a preinduced *picA* mutant was more virulent than was the wild-type control, a further indication that the *picA* locus regulates the surface properties of the bacterium in the presence of plant cells or plant cell extracts.

Agrobacterium tumefaciens, a gram-negative soil bacterium, causes the disease crown gall on many dicotyledonous and some monocotyledonous plant species as a result of the transfer of a segment of DNA, the T (transferred)-DNA, from the Ti (tumor-inducing) plasmid to the plant cell. The interactions of *A. tumefaciens* with its host are mediated by a number of chemical signals. Phenolic compounds such as acetosyringone, secreted by wounded plant cells, play an important role as inducers of transcription of the *vir* (virulence) genes on the Ti plasmid, resulting in the processing of the T-DNA from the Ti plasmid and its subsequent transfer to the plant cell. This induction process is mediated by the Ti plasmid-encoded VirA and VirG signal transduction system, in which the VirA protein functions as the receptor for the phenolic inducing compounds, and the VirG protein serves as a transcriptional activator that induces its own expression and the transcription of other *vir* genes (2, 42, 43, 50). Such two-component sensor-regulator systems have been adapted by many prokaryotes to link the expression of sets of genes with specific environmental stimuli (1). A second group of well-described signal molecules in the *A. tumefaciens*-plant interaction is the opines, a class of low-molecular-weight compounds specifically produced by crown gall tumors. Opines can induce Ti plasmid-encoded genes responsible for opine metabolism, permitting *A. tumefaciens* to use opines as carbon and nitrogen sources to the exclusion of most other soil microorganisms (30, 44). Certain opines, called conjugal opines, can stimulate the conjugal transfer of the Ti plasmid between *Agrobacterium* cells (7, 8, 11, 19, 21, 22, 33, 34). In addition, recent work in our laboratory showed that certain opines can stimulate the induction of *vir* genes by acetosyringone (45) or plant protoplasts (24).

Previously, we reported the discovery of a novel induction system of *A. tumefaciens* genes by plant extracts (37). This system involves the induction of certain *Agrobacterium* chromosomal genes by acidic polysaccharides in a carrot root extract. These polysaccharides are most likely derived

from the pectic portion of the plant cell wall. Comparison of proteins isolated from *A. tumefaciens* A136 and fractionated by one-dimensional sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) indicated that the expression of approximately one dozen proteins was increased in the presence of carrot root extract. To identify plant-inducible genes on the *Agrobacterium* chromosome, we mutagenized *A. tumefaciens* A136 with the promoterless *lacZ*-containing transposon Mu dI1681 and screened for transconjugants that showed greater β -galactosidase activity when grown on medium supplemented with carrot root extract than when grown on unsupplemented medium. One locus that showed a high level of induction was identified and named *picA* (for plant-inducible chromosomal). Although mutation of the *picA* locus had no detectable effect upon bacterial growth or virulence under laboratory conditions, *A. tumefaciens* cells harboring a *picA* mutation aggregated into long "ropes" when incubated with pea root cap cells, suggesting that the *picA* gene product may be involved in determining bacterial surface properties (37).

In an attempt to understand the function and regulation of *picA*, we sequenced the *picA* region. Deletion and S1 nuclease mapping experiments defined the promoter region of this locus. In addition, genetic experiments suggested that the expression of *picA* may be negatively, as well as positively, regulated. DNA sequence analysis of the region responsible for this negative regulation revealed an open reading frame homologous to known genes encoding polygalacturonase.

MATERIALS AND METHODS

Bacterial strains and growth conditions. *A. tumefaciens* and *Escherichia coli* strains as well as plasmids used in this study are listed in Table 1. The growth conditions, media, and antibiotic concentrations used for *A. tumefaciens* and *E. coli* were as described previously (37). The broad-host-range plasmid pCP13/B (4) and its derivatives, as well as other RK2 replicon-derived plasmids, were mobilized from *E. coli* to *A. tumefaciens* by a triparental mating procedure (5) using the mobilizing functions of the plasmid pRK2013 (9). AB

* Corresponding author.

TABLE 1. *E. coli* and *A. tumefaciens* strains and plasmids

Strain or plasmid	Description	Marker	Reference or source
<i>E. coli</i>			
DH5 α	<i>supE44 ΔlacU169(ϕ80lacZΔM15) <i>hsdR17 recA1 endA1 gyrA96 thi-1 relA1</i></i>		13
JM101	<i>supE thi Δ(lac-proAB) F' (traΔ36 proAB⁺ lacI^s lacZΔM15)</i>		27
IT2761	K12 derivative, Δ <i>lacU169 ΔrecA</i>		Irwin Tessman
<i>A. tumefaciens</i>			
A136	Wild-type, C58 chromosomal background, no Ti plasmid	Rif ^r	30
At156	<i>picA::lacZ</i> fusion in the A136 chromosome (contains downstream <i>picA</i> deletion)	Rif ^r Kan ^r	37
At460	<i>Bgl</i> II fragment from At156 chromosome containing <i>picA::lacZ</i> fusion (orientation I in pCP13/B) in A136	Rif ^r Tet ^r	This study
At461	<i>Bgl</i> II fragment from At156 chromosome containing <i>picA::lacZ</i> fusion (orientation II in pCP13/B) in A136	Rif ^r Tet ^r	This study
At484	Reconstructed <i>picA::lacZ</i> fusion in the A136 chromosome, no deletion	Rif ^r Kan ^r	37
At488	Ti plasmid pTiR10 in A136	Rif ^r	This study
At489	Ti plasmid pTiR10 in At484	Rif ^r Kan ^r	This study
At491	pCP13/B in At484	Rif ^r Kan ^r Tet ^r	This study
At492	Cosmid 7 in At484	Rif ^r Kan ^r Tet ^r	This study
At504	<i>Pst</i> I fragment from At156 chromosome containing <i>picA::lacZ</i> fusion (orientation I in pCP13/B) in A136	Rif ^r Tet ^r	This study
At505	<i>Pst</i> I fragment from At156 chromosome containing <i>picA::lacZ</i> fusion (orientation II in pCP13/B)	Rif ^r Tet ^r	This study
At550	ORF2 of <i>picA</i> region in A136 interrupted at <i>Eco</i> RI site by marker exchange with pAM10	Rif ^r Carb ^r	This study
At591	<i>Eco</i> RI fragment 4 of cosmid 7 (pCP13/B) in At484	Rif ^r Kan ^r Tet ^r	This study
At592	<i>Pst</i> I fragment 3 of cosmid 7 (pCP13/B) in At484	Rif ^r Kan ^r Tet ^r	This study
At593	<i>Xho</i> I-deleted cosmid 7 in At484	Rif ^r Kan ^r Tet ^r	This study
At629	<i>Eco</i> RI fragment 4 of cosmid 7 interrupted at <i>Eco</i> RV site by a kanamycin resistance gene (pLAFR1) in At484	Rif ^r Kan ^r Tet ^r	This study
Plasmids			
pUC18		Amp ^r	49
pUC118		Amp ^r	46
pCP13/B		Tet ^r	4
pLAFR1		Tet ^r	4
pVK102		Tet ^r Kan ^r	23
pPH1JI		Gen ^r	18
pCH1	<i>Pst</i> I fragment from At156 containing <i>picA::lacZ</i> fusion in pBR322	Tet ^r	37
pAM10	<i>Pst</i> I fragment 3 of cosmid 7 interrupted at <i>Eco</i> RI site by a T-DNA border region and pUC7 in pVK102	Kan ^r Tet ^r Amp ^r	Walt Ream
pRK2013	Mobilizing plasmid	Kan ^r	5
Cosmid 7	25 kbp of A136 chromosome DNA (pCp13/B) containing <i>picA</i> locus	Tet ^r	37

glucose minimal medium containing crude carrot root extract was prepared as described previously (37).

Genetic analysis of the *picA* promoter. To define the plant-inducible promoter of the *picA* locus, a *Bgl*II fragment and a *Pst*I fragment from *A. tumefaciens* At156 were cloned separately into pCP13/B in both orientations. Both of these fragments contain sequences upstream of the *picA::lacZ* fusion derived from the original Mu dI1681 insertion into *picA* in *A. tumefaciens* At156. The plasmids were transformed into *E. coli* DH5 α . Tetracycline-resistant colonies that formed blue colonies when plated on medium containing 5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside (X-Gal) were selected, and DNA minipreparations analyzed by restriction endonuclease digestion. Plasmids harboring the appropriate constructions were mobilized into *A. tumefaciens* A136, and β -galactosidase activity was measured (29) following incubation of the bacteria for 20 h in media either containing or lacking carrot root extract.

To identify a possible negative regulator of the *picA* promoter, different restriction endonuclease fragments (*Eco*RI fragment 4 and *Pst*I fragment 3) of cosmid 7, a cosmid encompassing the *picA* region, were subcloned into pCP13/B. Digestion of cosmid 7 with *Xho*I generated a large deletion with only two small DNA regions remaining from each end of the insert of cosmid 7. To interrupt the *pgl* gene

upstream of the *picA* locus, *Eco*RI fragment 4 was disrupted by insertion of a gene encoding kanamycin resistance (4) into the *Eco*RV site. The disrupted *Eco*RI fragment 4 was cloned into pLAFR1 (4). These plasmids were individually mobilized into *A. tumefaciens* At156 and At484, and β -galactosidase activity was measured (29) following incubation of the bacteria for 20 h in media either containing or lacking carrot root extract.

DNA sequencing and analysis. *Eco*RI fragment 4 of cosmid 7 was cloned into pUC118 (46) in both orientations. A series of overlapping deletions was made, using exonuclease III and S1 nuclease (Promega) (16). Both strands of this fragment were sequenced according to the dideoxy-chain termination method (40), using Sequenase and the -40 forward primer (United States Biochemicals) with either single-stranded or double-stranded DNA templates generated from the overlapping deletions. The single-stranded templates were prepared by using helper phage M13KO7 (46). Occasionally both dGTP and dITP were used to sequence DNA regions with a high G+C composition. Computer sequence analysis of the fragment was performed by using the Genetics Computer Group programs (University of Wisconsin at Madison).

To sequence the Mu dI1681 insertion site in *A. tumefaciens* At156, a double-stranded DNA template was pre-

pared (plasmid pCH1) and the junction region was sequenced by using *Taq* DNA polymerase (Promega) and P21, a primer internal to the end of the Mu dI junction kindly provided by Barry Wanner (Purdue University). The junction sequence was compared to that of *picA* to determine the precise insertion site of Mu dI1681.

S1 nuclease mapping. Total cellular RNA was prepared by a hot phenol method (38) from *A. tumefaciens* A136 grown in either AB glucose minimal medium or AB glucose medium plus carrot root extract. A 214-bp *Sau3A1* fragment spanning the presumed *picA* transcription initiation site was isolated, treated with calf intestinal alkaline phosphatase (Boehringer Mannheim Biochemicals) to remove the 5' phosphate (25), and end labeled with [γ - 32 P]ATP (specific activity, >3,000 Ci/mmol; Amersham Corp.), using polynucleotide kinase (Bethesda Research Laboratories, Inc.). S1 nuclease analysis was performed by a procedure modified from that of Overdier et al. (31). One hundred micrograms of total RNA and the denatured *Sau3A1* probe (about 10,000 cpm) were hybridized at 50°C for 3 h, followed by digestion with 700 U of S1 nuclease (Boehringer Mannheim Biochemicals) for 30 min at 37°C. The S1 nuclease digestion products were ethanol precipitated and subjected to electrophoresis through a 6% acrylamide sequencing gel, and the gel was exposed to Kodak XR-5 X-ray film. A M13 sequencing ladder was used as a molecular weight marker.

Construction of *A. tumefaciens* At550 and incubation with pea root cap cells. To interrupt the second ORF in the *picA* region, plasmid pAM10 (kindly provided by Walt Ream, Oregon State University) was mobilized into *A. tumefaciens* A136. pAM10 contains *PstI* fragment 3 of cosmid 7 interrupted at the *EcoRI* site in ORF2 by a T-DNA border sequence and pUC7, cloned into pVK102 (23). Carbenicillin-resistant colonies were selected. Double-homologous recombinants of this interrupted *A. tumefaciens* region with the wild-type *A. tumefaciens* chromosome were obtained by mobilizing pPH1JI (18) into the bacterium, selecting for carbenicillin and gentamicin resistance, and screening for kanamycin sensitivity (loss of pVK102). The correct construction (At550) was confirmed by Southern blot hybridization after digestion of the bacterial DNA with various restriction endonucleases. *PstI* fragment 3 of cosmid 7 was used as the hybridization probe. Incubation of *A. tumefaciens* A136, At488, and At550 with pea root cap cells was as described previously (14).

Nucleic acid manipulations. Restriction endonucleases were purchased from Bethesda Research Laboratories, Inc., New England BioLabs, Inc., and Boehringer Mannheim Biochemicals and used according to the instructions of the manufacturers. DNA blot hybridizations were performed as previously described (10). Recombinant DNA manipulations were performed as described previously (25) with P1 containment conditions as specified by the National Institutes of Health recombinant DNA guidelines.

Potato tuber disk virulence assay. The virulence of *A. tumefaciens* strains was assayed on red potato tuber disks, using a modification of Rogowsky et al. (35). *A. tumefaciens* At488 or At489 was grown in AB minimal medium or AB medium plus carrot root extract to a Klett reading of 100 (10^9 cells per ml), harvested, and suspended in MS salts (GIBCO-BRL). The potato disks were immersed in various dilutions of bacteria for either 20 min or 2 h. The disks were washed three times in MS salts solution and incubated at 25°C on water agar containing carbenicillin at 500 μ g/ml. The number of tumors per disk was scored after 12 to 14 days.

Maxicells. *EcoRI* fragment 4 and *PstI* fragment 3 of cosmid

7, as well as a *PstI-EcoRI* fragment derived from *PstI* fragment 3, were separately cloned into pUC13 and transformed into *E. coli* IT2761 (Irwin Tessman, Purdue University). A procedure modified from Sancar et al. (39) was used to prepare maxicells. Bacterial cells (10 ml; $A_{600} = 0.5$) were irradiated with a UV dose of 100 J/m². The irradiated bacteria were incubated at 37°C in the dark for 1 h, cycloserine (final concentration of 100 μ g/ml) was added, and the cells were incubated 10 h at 37°C in the dark. The bacteria were pelleted by centrifugation and resuspended in Hershey's sulfate-free medium (48). The bacteria were labeled with [35 S]methionine (5 μ Ci/ml; specific activity, >1,000 Ci/mmol; Amersham) at 37°C for 1 h. Bacterial lysis and protein SDS-PAGE were performed as described previously (37).

Nucleotide sequence accession number. The nucleotide sequence reported in this paper has been submitted to GenBank (accession number M62814).

RESULTS

Determination of the plant-inducible promoter region of the *picA* locus. To understand how the *picA* locus is regulated at the transcriptional level by the inducer in carrot root extract, we felt it important to determine the region of the *picA* locus responsible for this induction. We therefore cloned a *BglII* fragment containing the junction sequence of the *picA::lacZ* fusion from *A. tumefaciens* At156 into pCP13/B in both orientations. This *BglII* fragment contains approximately 10 kbp of DNA upstream of the Mu dI1681 insertion site and the intact promoterless *lacZYA* operon from Mu dI1681. A smaller *PstI* fragment, that contains only about 0.5 kbp of *Agrobacterium* chromosomal DNA upstream of the Mu dI1681 insertion site, also contains an intact promoterless *lacZYA* operon. This fragment was cloned into pCP13/B in both orientations as well. These plasmids were individually mobilized into *A. tumefaciens* A136, creating *A. tumefaciens* At460 and At461 (*BglII* fragment cloned into pCP13/B in two orientations) and At504 and At505 (*PstI* fragment cloned into pCP13/B in two orientations). *A. tumefaciens* At484 (the reconstructed *picA* mutant strain carrying a *picA::lacZ* fusion in the chromosome) containing plasmid pCP13/B (At491) was used as a control.

Figure 1 shows the results of β -galactosidase assays of these strains grown in either AB glucose minimal medium or AB glucose minimal medium plus carrot root extract. The uninduced level of expression of the *lacZ* gene in *A. tumefaciens* strains carrying either the *BglII* fragment or the *PstI* fragment in pCP13/B was about three- to fivefold higher than that of the control strain (At491). We believe that this increased basal activity was due to the high copy number of the *picA::lacZ* fusion in these strains, because pCP13/B normally exists in *A. tumefaciens* with a copy number of 5 to 10 (data not shown). Nevertheless, these strains still showed a 4- to 12-fold induction of β -galactosidase activity when grown in AB minimal medium plus carrot root extract relative to AB minimal medium. We therefore conclude that a region of *Agrobacterium* DNA defined by this *PstI* fragment is sufficient for the induction of the *picA* promoter by carrot root extract. As discussed below, this *PstI* fragment contains only 120 bp of DNA upstream of the transcription initiation site of the *picA* promoter. Because *A. tumefaciens* At460 and At461 displayed a higher level of induced β -galactosidase activity than did *A. tumefaciens* At504 and At505, it is possible that additional elements required for maximal *picA* induction exist upstream of the *PstI* site.

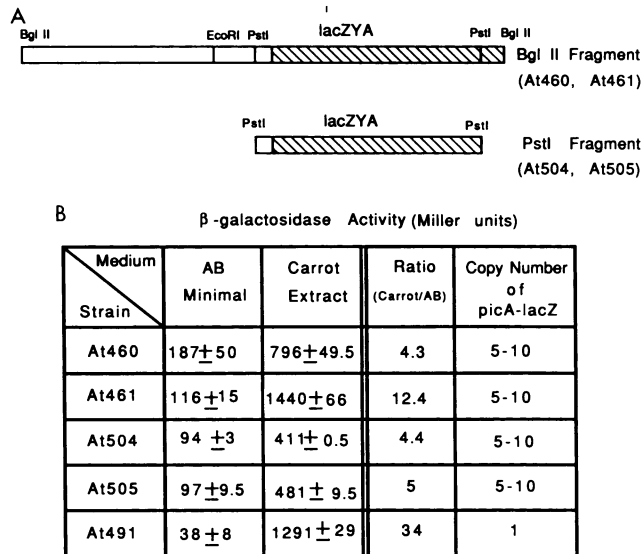


FIG. 1. (A) Restriction endonuclease map of the *Bgl*II fragment and the *Pst*I fragment from *A. tumefaciens* At156 DNA containing the *picA::lacZ* fusion. Hatched region is from the transposon Mu dl1681; open region is from the *Agrobacterium* chromosome upstream of the Mu dl1681 insertion site. Only restriction endonuclease sites of interest are indicated. The *Bgl*II fragment cloned into pCP13/B in both orientations (I and II) was mobilized into *A. tumefaciens* A136, creating At460 (I) and At461 (II). The *Pst*I fragment cloned into pCP13/B in both orientations (I and II) was mobilized into *A. tumefaciens* A136, creating At504 (I) and At505 (II). pCP13/B was mobilized into *A. tumefaciens* At484 to create At491 as a control. (B) β-Galactosidase activity of these *A. tumefaciens* strains grown in AB glucose minimal medium (AB minimal) or AB minimal plus carrot root extract (carrot extract). The induction level was calculated as the ratio of β-galactosidase activity from strains grown in carrot root extract to strains grown in AB minimal medium.

Identification of a possible negative regulator of *picA* expression. We initially observed that when cosmid 7 or *Eco*RI fragment 4 of cosmid 7 (cloned into pCP13/B) was mobilized into *A. tumefaciens* At156, we could not induce the *picA::lacZ* fusion gene by carrot root extract (data not shown). As described previously, *A. tumefaciens* At156 contains both a Mu dl1681 insertion and a deletion of the *Agrobacterium* chromosome downstream from the site of Mu dl1681 insertion (37). To avoid possible complications in interpreting results due to this deletion, we used *A. tumefaciens* At484 for further genetic analyses.

To determine which region of cosmid 7 could inhibit the induction of the *picA* promoter by carrot root extract, various plasmids were mobilized into *A. tumefaciens* At484, and β-galactosidase activity was measured following incubation of the bacteria in different media. As can be seen in Fig. 2, cosmid 7 (At492) and *Eco*RI fragment 4 (At591) inhibited the induction of the *picA* promoter of *A. tumefaciens* At484. *Xho*I-deleted cosmid 7 (At593) had a similar effect. Disruption of *Eco*RI fragment 4 by insertion of a gene encoding kanamycin resistance into the *Eco*RV site (At629) restored the induction of *picA* by carrot root extract. However, *Pst*I fragment 3 (At592) did not have any effect on the induction of the *picA::lacZ* fusion. The vector pCP13/B (At491) had no effect on induction. These results indicate that when present on a multiple-copy plasmid, a DNA fragment delimited by

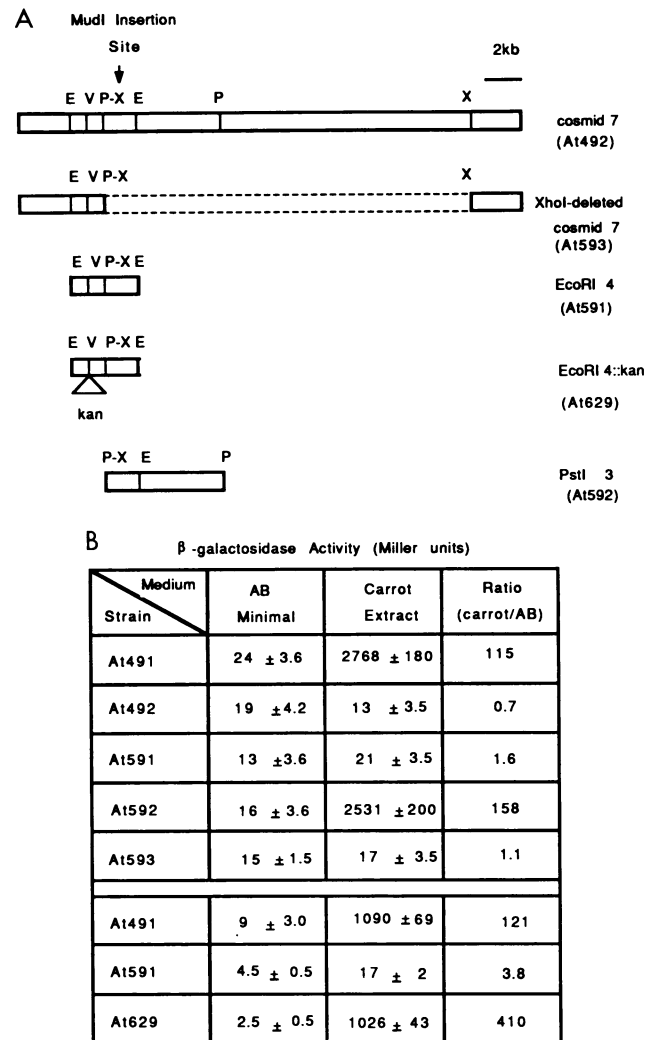


FIG. 2. (A) Restriction endonuclease fragments of cosmid 7 used to identify a putative negative regulatory element of *picA*. All of the fragments were cloned into pCP13/B or pLAFR1 and mobilized into *A. tumefaciens* At484. pCP13/B was mobilized into strain At484, creating strain At491 as a control. Restriction endonuclease sites shown are only those of importance for this experiment and are not unique. Transposon Mu dl1681 insertion site is indicated. E, *Eco*RI; P, *Pst*I; X, *Xho*I; V, *Eco*RV. (B) β-Galactosidase activity of *A. tumefaciens* strains grown in AB glucose minimal medium (AB minimal) or AB minimal medium plus carrot root extract (carrot extract).

the *Eco*RI-to-*Pst*I sites upstream of *picA* can inhibit *picA* induction.

Sequence of *Eco*RI fragment 4. The nucleotide sequence of *Eco*RI fragment 4, including the *picA* region (3,155 nucleotides in length), is presented in Fig. 3. Three major ORFs are indicated. One ORF (designated *pgl*; see below) starts at nucleotide 390 and extends to nucleotide 1328 and can encode a protein of approximate molecular weight 34,300. The first ORF of *picA* (ORF1) begins at nucleotide 1728 and extends 699 bp (stop codon at nucleotide 2429; TGA). ORF1 can encode a polypeptide with a predicted molecular weight of 25,500. ORF1 is preceded by a consensus -10 promoter region of *E. coli* (TATAAT from positions 1542 to 1547) but not by a consensus Shine-Dalgarno sequence. ORF2 initiates

```

EcoRI 10          30          50          70
GAATTCAGAACCGCGCCCTGTGCCGACTCGTCGAAGCCACAGCGTTTTCCCTACGCCCGGATATTGCGCTTGACGATGA
          90          110          130          150
AGAAGGTGGCGCGCCAATGGCACGGCGTCGAAAACGCTGTCCCGCGGTGCCGTGGCGGCACCCCTGCGTCTGGGGCCCCG
          170          190          210          230
GTGCTGGACGGCCTTTCGGTGGCGCTGAAAAGCGACATGACGCTGCACCTTGGCGAAGGTCGCGTGTGCGCGCCATC
          250          Sau3AI 270          290SphI  PvuII  310
CGTCCGTAACAGGTGGCCGATCCTGCCCGCGCGATGAGGCCGGCGCATGCTTGGCAGCTGGGAGGGACTACCGGATG
          330          350          Sau3AI370          390
CCTGTTTCGCCCGCCTGTTTCATGCGATCGGGCGGATAATCTCGTGATCGAGGGCAGGGGTATCCTGGATGGCTCTGGC
          410          430          450          (PGL)  M A L A
GACAAGGGCGACTGGTGGAGCTGGCCGAAGGAAACCCGTGAGGGCGCGATGCGCACGGCCCTGCATCTCGTGCTCTGCC
          470
T R A T G G A G R R K P V R A R C A R G L H L V S C H
          490          Sau3AI510          530Sau3AI          550 Sau3AI
ACAAAACACAGCTTTTGGGGTTCACGATCCGCAATGCCGCCTCGTGGACGATCCATCCGCAGGGTGGCAGGATCTGACG
          570          590          610          630
K T Q L L G F T I R N A A S W T I H P Q G C E D L T
GCCGCGCGTCCACCATCATCGCGCCGATGACAGTCCCAATACCGATGGTTTCAACCCCTGAAAGCTGCCGTAACGTGAT
          650          670          690          710
A A A S T I I A P H D S P N T D G F N P E S C R N V M
EcoRV 650          670          690          710
GATATCAGGCGTGCCTTTCCGTGGGTGATGACTGCATCGCGGTGAAGCGGGGAAACGCGGGCTGATGGTGGAGACG
          730          750          770          NcoI  790
I S G V R F S V G D D C I A V K A G K R G P D G E D D
Sau3AI 730          750          770          NcoI  790
ATCATTGGCGGAGACACGGGTATCACGGTGCGCCATTGCCTGATGCAGCCGGCCATGGCGGGCTGGTCATCGGTTCCG
          810          830          Sau3AI850          Sau3AI 870
H L A E T R G I T V R H C L M Q P G H G G L V I G S
GAAATGCCGGGGTCCATGATGTGACGGTGAAGATTGCGACATGATCGGCACGGATCGCGGCTGCGTCTCAAGC
          890          910          930          950
E M S G G V H D V T V E D C D M I G T D R G L R L K T
CGGCGCGGTTCCGGCGGGTATGGTCCGGAACATCACCATGCGCGGGTGTGCTGGACGGTGTGCAGACTGCACTTT
          970          990          1010          1030
G A R S G G G M V G N I T M R R V L L D G V Q T A L S
CCGCCAACGCCCATATTATTGATGCCGATGGGATGATGACTGGGTGCGAGTCGCGGAAACCCGGCGCCGTCAATGAC
          1050          SaI  1070          1090          1110
A N A H Y H C D A D G H D D W V Q S R N P A P V N D
GGCAGCGGTTTCGACGGCATCCCGTGAAGATGTCGAAATCCGCAATCTGCCCATGCGGCAGGTGTCTTCTCGG
          1130          1150          1170          1190
G T P F V D G I T V E D V E I R N L A H A A G V F L G

```

FIG. 3. Nucleotide sequence of *EcoRI* fragment 4. The deduced amino acid sequences for the three major ORFs (*pgl*, ORF1, and ORF2) are shown. The consensus -10 region near the *picA* transcription initiation site is underlined (---). The putative transcription initiation site (*) and the Mu dI1681 insertion site are also indicated.

157 bp downstream of the termination site of ORF1 and continues beyond the *EcoRI* site into a region that we have yet to sequence. ORF2 can encode a protein with a molecular weight of at least 21,000. ORF2 is not preceded by a consensus Shine-Dalgarno sequence. Computer analysis suggests that ORF2 may be membrane localized because the sequenced region shows three putative membrane-spanning domains. In addition, a smaller ORF starts at nucleotide 1830, which can encode a polypeptide of 146 amino acids in length with a predicted molecular weight of 16,000 (not shown).

We searched for sequence homology of the ORFs within *EcoRI* fragment 4 with those sequences in the GenBank and EMBL libraries at both the nucleotide sequence and amino acid sequence levels, using the FASTA and TFASTA programs (32) provided by the Genetics Computer Group program package (University of Wisconsin at Madison). We were unable to match at a significant level ORF1 and ORF2 with any of these sequences. Neither did we identify homol-

ogy of ORF1 and ORF2 with a number of consensus active site motifs. These results suggest that the *picA* locus encodes novel gene products.

Computer analysis of the ORF encoded by nucleotides 390 to 1,328 (*pgl*) revealed significant homology with genes encoding polygalacturonase from a number of organisms. Figure 4 shows that over a 201-amino-acid sequence, *pgl* from *A. tumefaciens* has 28.9% identity, and an additional 44% conserved amino acid replacement homology, with polygalacturonase from *Erwinia carotovora* (*peh*; 17). The calculated amino acid homology with other polygalacturonase proteins is as follows: *Pseudomonas solanacearum* (*pglA*; 20), 67.4% (over a 151-amino-acid sequence); and tomato (*pg*; 12), 72.2% (over a 176-amino-acid sequence). In addition, *pgl* is very homologous to polygalacturonosidase from *Erwinia chrysanthemi* (*pehX*; 15), with 71% homology over a 227-amino-acid sequence. This high degree of amino acid homology between *pgl* and known polygalacturonase and polygalacturonosidase genes from other organisms sug-

```

1130          1150          1170          Sau3AI 1190
TCTGCCGGACGTCCCATCCGCAACATCGTTGTCGCCAACCTCACCCATCGTCTCGCATGATCCTTCGGCTGTTGGCAGCG
L P D V P S A T S L S A T S P I V S H D P S A V A T P
Sau3AI1210          1230          1250          1270 Sau3AI
CGCCGATCATGGCCGACCGCGTGCGCCCATGCGCATGAGGCTCGTCTTCGAGCAGGCGGATGTCGTCGCGATGATCCG
P I M A D R V R P M R M R L V F E Q A D V V C D D P
1290          1310          1330          1350 BclI
GCGCTTCTGAATGACGCCCCCGTTTCCATTTTCGTCATATTTTCGATTGAGAAAAGCCATGAAAGCCACTGAATATTTTAT
A L L N D A P V S I S S Y F D
Sau3AI1370          1390 PvuII          1410          1430 PstI
CAATCTCTCGCGGATACAAACATTACAAAGGCGGCAGCTGGTGTATGAGGATGGTTGTGTCTATCGCGGTCTGCAGCA
PvuII 1450XhoI          1470          1490 Sau3AI          1510
GCTGCTCGAGGCGACGGGCGAGGCTGGAATGACCATTGACCGCTCTGCCGATCCCCAGATTGGTGCAGGATGGAACGCT
1530Sau3AI          1550          1570          1590
GGCCGGTTATGATCCGAGAATAATAATCGACCATATTCTTGCCTGCGGACGATTCCTTTCCCTGTCGCGCAAACCGG
1610          1630          1650          1670
GGATGCACGCTATCTGGCGGACGGGCGCATCTGGCAGGCCAACTTCGAAGCCATCCACGCAATCGCGGCAATATTAT
1690          1710          1730          Sau3AI1750
GGCACAGAAGCGTTACCCGCATCAGGTCTGGCTCGATGGCCTCTATATGGGGCTGCCATTCCAGATCGAATATGGTCA
(picA) (ORF1) M G L P F Q I E Y G Q
1770          Sau3AI1790          1810          1830
ACGACGGGCGCCCGGAGCTGATCGAGGATGCGTTGCGTCAGTTTTTCAGCGGCACTTGCCTGACGGCGGATGCTGGTGG
T T G R P E L I E D A L R Q F S A A L A L T A D A G G
1850          1870          1890          1910
TCTATACGTTACGGTTATGACGAGAGCCGCAACCAGCGCTGGGCTAATCCTGCGAGCGGCAATCACCGGCCATCTGGG
L Y V H G Y D E S R N Q R W A N P A S G K S P A I W A
1930          NcoI MudI Insertion Site 1970          1990
CGCGGGCGTGGGCTGGCTTGCATGGCGCTGGTGGATGCGCTGGTCTATCTTGCCTGACGACAGTGCAGCGCCGAGCTT
R A V G W L A M A L V D A L V I L P D D S A T A E L
2010          2030          2050          2070
CGCGAGAGGACGCGACGCTTCTGGCTGGTATCATTGCCCGGACGCGAGGCGGCTCTATGGATGCAGGTGCTCGACAA
R E R T R R L L A G I I A R Q T Q A G L W M Q V L D N
2090          2110          2130          2150
TCAGGGCCTTGCCTGCAATTATGCGGAAACATCCGCCCTCCGCCATGTTTCGCTTATGCCCTGCTGCGTGCAGCCCGGCTGG
Q G L A G N Y A E T S A S A M F A Y A L L R A A R L G
2170          2190          2210          2230
GGCTTTCGCGGGTGAAGAGGCGAAGGCTGCCCTTTCTGCTGGTGCAGGCGCTTCCGCACTTCTGGAACCGCCCTC
L L R G E E A K A A L S A G R Q A L A A L L E T R L

```

FIG. 3.—Continued.

gests that *pgl* encodes a polygalacturonase or related enzymatic activity in *A. tumefaciens*.

Mu dI1681 insertion site in the *picA* locus. Using restriction endonuclease mapping, we previously determined that the transposon Mu dI1681 inserted into the *A. tumefaciens* At156 chromosome at a site about 0.5 kbp downstream from the *PstI* site (37). To localize the insertion site more precisely, we used P21, a primer internal to the Mu dI1681 junction, to sequence the junction region at the insertion site. Because an inverted repeat at the end of the Mu dI1681 could form a hairpin secondary structure that would make sequencing with Sequenase difficult (28), we used *Taq* DNA polymerase to sequence this junction region. Our results indicated that transposon Mu dI1681 inserted after nucleotide 1,950 (a few nucleotides from the *NcoI* site) (Fig. 3). The transposon, therefore, interrupted ORF1 of *picA*. The distance between the *PstI* site and the Mu dI1681 insertion site is 520 bp.

Transcription initiation site of the *picA* locus. To determine the transcription initiation site of the *picA* locus, we isolated total RNA from *A. tumefaciens* A136 grown in AB glucose

minimal medium and AB glucose minimal medium plus carrot root extract. A *Sau3AI* fragment (214 bp, from positions 1532 to 1745; Fig. 3) was end labeled and used as a S1 nuclease protection probe. The S1 nuclease mapping results are shown in Fig. 5. The major band protected from S1 nuclease digestion by RNA from induced bacteria was a fragment of about 187 bp in length (Fig. 5, lane 3). This band was not detectable in RNA from uninduced bacteria (Fig. 5, lane 2), although an equal amount of total RNA (100 µg) from each bacterial culture was used. Because the *picA* promoter is relatively inactive in AB minimal medium but is induced by the addition of carrot root extract, we conclude that the band protected in lane 3 represents the major, if not the only, transcription initiation site of the *picA* locus. This site corresponds to nucleotide 1557 of the *picA* sequence presented in Fig. 3. It is interesting that this site is immediately downstream of the consensus *E. coli* -10 sequence, suggesting that this -10 region is used as a signal for transcription initiation of the *picA* locus.

Maxicell expression of the ORFs. We attempted to express the sequenced ORFs in *E. coli* maxicells. Figure 6 shows

```

      2250              2270              2290              2310
GAGCTGGATGAGCAGGGCGTCGCGCGGCTAACCGGCATCGTGCATGTCGCCGGGCTCGGGGTTTCGACGGTAATTATCG
E L D E Q G V A R L T G I V H V A G L G G F D G N Y R
      2330              2350              2370              2390
AGATGGAAACGCCGGACTATTATCTGACGGAGCCGGTCGTATCCGACGATGCAAAGGGTGTCCGGCCGCTGATGATGGCCT
D G T P D Y Y L T E P V V S D D A K G V G P L M M A Y
      2410              2430              2450              2470
ATGCGAAAGCCTGCTTCTGGCCCGCTGAAGCCGGATACTTGCCGGAAATGGTGATATAAACCGGGCCGATAGAATCC
A E S L L L A R
      2490              2510              2530              2550
GGGTGCGCTGAACGGTTTGTGACTGGAAGAAATGAACCGCTCGGGATAAACCTCGTTTTTGCATATAACCGGAATCC
      2570              2590              2610              2630
AGTGTCAACGAGCCCAATGCCATGAGCCAGTCACCCCGGAACGTTTCATCCTGCTGGACGGCATAAGGGGTGTGGCTG
      (ORF2) M S Q S P P E R F I L L D G I R G V A A
      2650              2670              2690              2710
CGCTTTTTATCGTTACCCGCCATGCCGAACAGTTTTTTCGACGGGACCCGGCCTCGAGCTATCTTGCCTGGACCTGTTT
L F I V H R H A E Q F F G R D P A S S Y L A V D L F
      2730              2750              2770              2790
TTTGCCTCAGCGGCTTCTGCTGGCCCATGCCTATGGCAAGAACTTATGAGGGGACAATCACGCCGGGTTTTTCTCT
F A L S G F V L A H A Y G K K L Y E G T I T P G F F L
      2810              2830              2850              2870
GAAGGCTCGCTTTGCCGCTTTATCCACTCTATGTGCTGGCGCTGGCGCTGATGGCGGCTATTTCACTGCCTTTACG
K A R F A R L Y P L Y V L A L A L M A A Y F I C L Y V
      2890              Sau3aIPvuI2910Sau3AI              Sau3AI2930              2950
TGCTGGGCTCGCCGACCGCGATCGATGATCTGCATCGTCTCATTGATCCCGGAGCTGGCTTTTGCCTTGTGTACAGGA
L G L P T P I D D L H R L I D P G E L A F A L V T G
      2970              2990              3010              3030
CTTCTGTTCTGCCCGCCCTTCACTCACCTGAACGGCGCGCTGTTTTCGGTTCAGCCCGCATGGTCCGCTGTTCAA
L L F L P A P F T L T L N G A L F L V S P A W S L F N
      3050              3070              3090              3110
CGAACTGGTGGTGAATGCGGTTTATGCCCGCTGGGGTGC GCGCGCAATGAAACAGACGGTTTGGTGTCTCGCCGTC
E L V V N A V Y A R W G A R A T M K Q T V L V L A V S
      3130              3150EcoRI
GCGCCGTCGTGCTGATGGTGGCGCGCGGAATTC
A V V L M V A A A E F

```

FIG. 3.—Continued.

that although we could detect the product of the β -lactamase gene (30 kDa) carried by pUC13, we could not detect any expression of ORF1 or ORF2 (lanes 1 to 3). This failure most likely results from the lack of a consensus Shine-Dalgarno sequence preceding these ORFs. The 15-kDa protein detected in lane 1 must be encoded by an ORF beyond *EcoRI* fragment 4. Lane 3 shows an approximately 34-kDa protein encoded by the region that contains the *pgl* locus. The size of the expressed protein corresponds well to the predicted size of the protein encoded by *pgl*, suggesting that *pgl* can be expressed in *E. coli*.

Interruption of ORF2 did not result in "rope" formation by *A. tumefaciens*. We showed previously that mutation of the *picA* locus by Mu dI1681 (*A. tumefaciens* At156 and At484) caused bacteria to aggregate, forming ropes in the presence of pea root cap cells (37). DNA sequence analysis showed that the insertion of Mu dI1681 into the *picA* locus interrupted ORF1 of the *picA* region. Because the *picA* locus may encode two (or more) polypeptides, we were concerned that the phenotype of bacterial aggregation may be specified by ORF2. This could occur if ORF1 and ORF2 constitute an operon, and Mu dI1681 insertion into ORF1 had downstream polar effects upon ORF2. We therefore constructed an *A. tumefaciens* strain, At550, that contained an insertion into ORF2 (at the *EcoRI* site) but contained an uninterrupted

ORF1 (see Materials and Methods). Following incubation of *A. tumefaciens* A136, At484, and At550 with pea root cap cells for 16 h, *A. tumefaciens* At484 aggregated extensively to form ropes. However, *A. tumefaciens* At550 (interrupted ORF2) behaved as did the wild-type parental strain *A. tumefaciens* A136 (data not shown). We conclude that the phenotype of bacterial aggregation in the presence of pea cells is conferred by mutation of ORF1 of the *picA* locus.

Potato tuber disk virulence assay. We previously showed that mutation of the *picA* locus did not noticeably affect the virulence of *Agrobacterium* strains on tobacco leaf disks, carrot root disks, or sunflower stems (37). Because mutation of *picA* may alter the surface properties of agrobacteria, we performed a more quantitative virulence assay using potato tuber disks (35). *A. tumefaciens* At488 (wild type) and At489 (mutant *picA*) were grown to a Klett reading of 100 (approximately 10^9 cells per ml) in the presence or absence of carrot root extract, diluted into a MS salts solution (10^4 to 10^6 bacteria per ml), and incubated with potato disks for 20 min or 2 h. The disks were washed in a MS salts solution three times, and the number of tumors was scored 12 to 14 days later. Table 2 shows that there was no significant difference between the number of tumors incited by the uninduced mutant and wild-type bacteria when incubated with potato disks for either 20 min or 2 h. Similarly, there was no

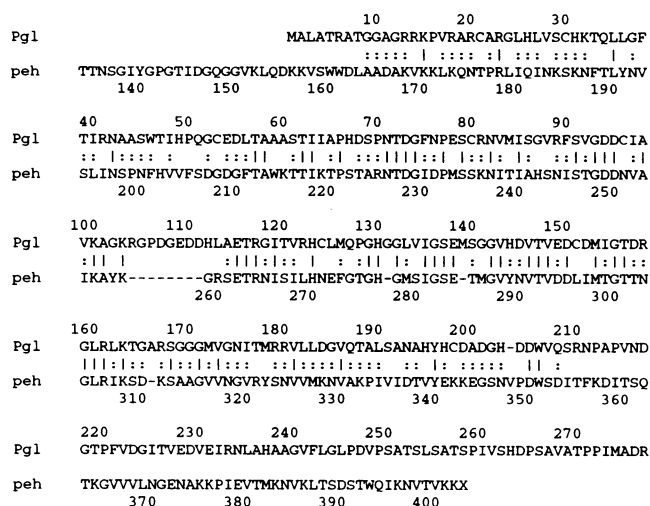


FIG. 4. Homology between *pgl* from *A. tumefaciens* and *peh* from *Erwinia carotovora*, determined by using TFASTA program. Amino acid identity (|) and conserved amino acid replacement homology (:) are indicated.

significant difference between the induced mutant and wild-type strains upon incubation with potato disks for 20 min. When the induced mutant and wild-type strains were incubated with disks for 2 h, however, there was a significant difference ($P < 0.01$) between the strains at two of three concentrations: the mutant *A. tumefaciens* At489 at 10^4 and 10^5 cells per ml incited 109 to 111% more tumors than did the wild-type *A. tumefaciens* At488. Although the number of tumors incited by bacterial strains at 10^5 cells per ml was not significantly different (as calculated by using Student's *t* test), the mutant strain nevertheless incited 39% more tumors than did the wild-type strain. A similar large variation in the number of tumors per disk has been reported by others (47). A repetition of this experiment, using induced bacteria, yielded qualitatively similar results (data not shown). It therefore appears that with this potato disk virulence assay, induced bacteria mutant in *picA* are more virulent than are induced wild-type bacteria.

DISCUSSION

picA is an *A. tumefaciens* chromosomal locus, identified by transposon mutagenesis, that is inducible by certain acidic polysaccharides such as polygalaturonic acid found in carrot root extract. We are interested in how the *picA* locus is regulated at the molecular level by plant signal molecules and what role this locus plays in *A. tumefaciens*-plant interactions. Cloning and genetic analysis of the *picA::lacZ* fusion presented in this report defined a region of the *picA* promoter that is responsible for the induction of this locus. Furthermore, a DNA sequence that inhibited the induction of the *picA* promoter by carrot root extract when present in multiple copies in *A. tumefaciens* cells was identified upstream of the *picA* locus. DNA sequence analysis of the *picA* region revealed two major ORFs. Mutation of ORF1, but not ORF2, was responsible for the increased aggregation of *A. tumefaciens*, resulting in the formation of ropes in the presence of pea root cap cells.

Although we previously did not detect a difference in virulence between wild-type *A. tumefaciens* cells and bacteria mutant in *picA* when assayed on tobacco, carrot, or sunflower (37), further analysis indicated that strains mutant in *picA* were more virulent on potato disks than were wild-type cells. This difference was expressed only when the bacteria were induced by carrot root extract prior to inoculation on the potato disks, and only when the bacteria were incubated with the disks for 2 h prior to washing the disks. We interpret these data to indicate that preinduced mutant bacteria are more adherent and bind more efficiently to each other and perhaps to plant cells than do preinduced wild-type agrobacteria. These results coincide well with previous experimental data indicating that *picA* induction by carrot root extract requires 6 to 8 h (37) and that the saturatable binding of *A. tumefaciens* to plant cells requires 30 to 60 min (6, 26). Taken together with our previous data showing that a *picA* mutant self-aggregates in the presence of pea root cap cells to a greater extent than do wild-type bacterial cells (37), these virulence data further indicate that the *picA* locus influences the surface properties of the bacterium in the presence of plant cells or plant cell extracts.

The details as to how the *picA* promoter is regulated are not yet clear. We initially favored a model in which *picA* was regulated by a repressor encoded by the *EcoRI*-to-*PstI* fragment upstream of *picA*. When present in multiple copies,

TABLE 2. Potato tuber disk virulence assays

Bacteria/ml	Uninduced		Induced	
	At488	At489	At488	At489
	Avg tumors/disk, total tumors	Avg tumors/disk, total tumors	Avg tumors/disk, total tumors	Avg tumors/disk, total tumors
20 min				
10 ⁴	6.8 ± 8.2 (29) ^a , 198	5.7 ± 5.8 (30), 171	9.2 ± 12.9 (22), 203	9.4 ± 13.0 (29), 274
10 ⁵	21.6 ± 15.7 (24), 518	19.8 ± 13.1 (27), 536	16.8 ± 20.4 (25), 421	21.7 ± 19.2 (27), 585
10 ⁶	38.6 ± 20.8 (23), 887	19.9 ± 19.3 (20), 398	44.3 ± 29.7 (21), 930	41.4 ± 37.8 (28), 1,159
2 h				
10 ⁴	10.2 ± 18.0 (19), 193	10.6 ± 9.4 (21), 222	5.6 ± 5.0 (16), 90	11.8 ± 6.8 (21), 248 ^b
10 ⁵	24.0 ± 29.6 (20), 481	40.9 ± 47.5 (17), 696	32.4 ± 30.2 (21), 682	45.0 ± 31.0 (21), 946 ^c
10 ⁶	36.2 ± 17.2 (18), 562	39.3 ± 32.3 (19), 746	32.3 ± 20.0 (24), 776	67.4 ± 50.1 (20), 1349 ^d

^a Values in parentheses are the numbers of disks.

^b *t* = 3.06, $P < 0.01$.

^c *t* = 0.9, not significant.

^d *t* = 3.18, $P < 0.01$.

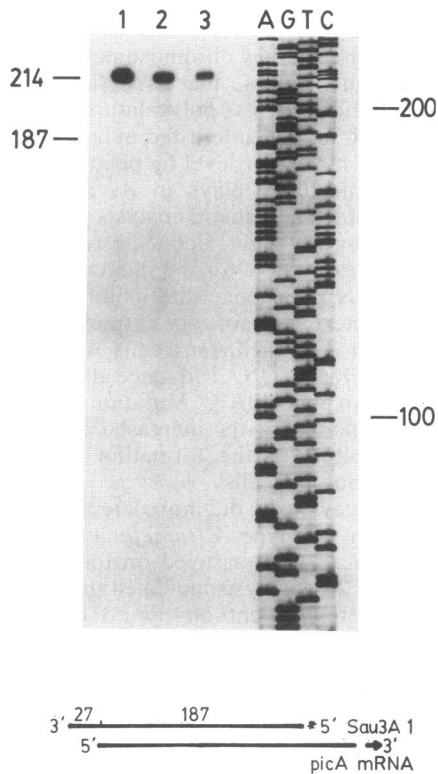


FIG. 5. S1 nuclease mapping of the *picA* transcription initiation site. Lanes: 1, *Sau3A1* probe and yeast tRNA digested with 700 U of S1 nuclease; 2, *Sau3A1* probe and 100 μ g of total RNA isolated from *A. tumefaciens* A136 (grown in AB minimal medium) digested with 700 U of S1 nuclease; 3, *Sau3A1* probe and 100 μ g of total RNA isolated from *A. tumefaciens* A136 (grown in AB minimal medium plus carrot root extract) digested with 700 U of S1 nuclease. A M13 sequencing ladder (lanes A, G, T, and C) was used as a molecular weight marker. Numbers beside the gel refer to length in nucleotides. The map below the gel indicates the *Sau3A1* S1 nuclease probe (labeled at the 5' end; *) and the 5' end of the *picA* mRNA.

this fragment would overexpress the putative repressor, resulting in the inhibition of *picA* induction by carrot root extract. Analysis of the DNA sequence of this region unexpectedly revealed an ORF (*pgl*) with significant homology to known polygalacturonase protein genes. Disruption of *pgl* carried by a plasmid prevented the inhibition of *picA* induction, suggesting that *pgl* is expressed in agrobacteria.

How can the expression of *pgl* regulate *picA* induction? Biotype I strains of *A. tumefaciens* such as the strains used in this study do not have secreted polygalacturonase activity (3). In addition, the strains used in this study cannot utilize polygalacturonic acid as the sole carbon source (36). However, the presence of an intracellular polygalacturonase activity in biotype I *Agrobacterium* strains has not been examined. Although we have not yet proven that *pgl* encodes a polygalacturonase, the nature of the inducing compound of *picA* may provide some insight regarding the mechanism of regulation of *picA* by *pgl*. Chemical analysis of the inducing substance in carrot root extract indicated that the inducer is related to polygalacturonic acid. Indeed, pure polygalacturonic acid can induce *picA* (37). However, we have found that only a certain size range of polygalacturonic acid (degree of polymerization 6 to 16) can effectively induce *picA*. Galacturonic acid polymers of lesser or greater chain

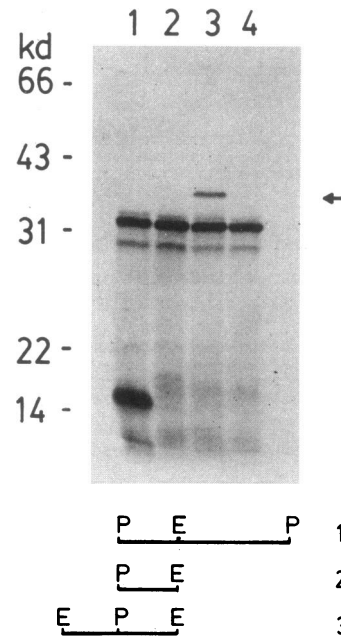


FIG. 6. Expression of the *pgl* gene of *A. tumefaciens* in *E. coli* maxicells. The three fragments derived from cosmid 7 used in this experiment are shown at the bottom; the number at the right of each fragment corresponds to the lane number of SDS-PAGE. These fragments were cloned into plasmid pUC13 and transformed into *E. coli* IT2761. The maxicell expression procedure was as described in Materials and Methods. Lanes: 1, *Pst*I fragment 3 (4.9 kbp) of cosmid 7; 2, *Pst*I-*Eco*RI fragment (1.7 kbp) derived from *Eco*RI fragment 4 of cosmid 7; 3, *Eco*RI fragment 4 (3.2 kbp) of cosmid 7; 4, pUC13. The arrow on the right indicates the protein expressed in lane 3, which is the predicted size of *pgl* from *A. tumefaciens* according to DNA sequence analysis. The 31-kDa protein detected in all four lanes is the β -lactamase protein encoded by plasmid pUC13.

length are relatively poor inducers (36). If *pgl* encodes a polygalacturonase, the overexpression of this gene may result in the rapid degradation of galacturonic acid polymers to a size too small to induce *picA*. We realize that this model of *picA* regulation by *pgl* depends, in part, on proof that *pgl* actually encodes a polygalacturonase. We are currently investigating this possibility. Alternatively, our initial model, in which *pgl* encodes a repressor that acts directly upon the *picA* promoter, cannot yet be discounted. In addition, there may be other mechanisms by which *pgl* regulates *picA*.

DNA sequence analysis of the *picA* locus did not provide information regarding the functions of the ORFs of this locus. In fact, because there are 155 nucleotides between *picA* ORF1 and ORF2, we do not know whether ORF2 is controlled by the same plant-inducible promoter that regulates ORF1 expression. We noted, however, that no consensus prokaryotic promoterlike element exists between these two ORFs. We were not able to detect an inducible RNA species from the *picA* region by Northern (RNA) blot analysis, probably because the RNA was not stable or abundant enough to yield a defined band under our experimental conditions (data not shown). Neither were we able to express ORF1 or ORF2 in *E. coli* maxicells. This failure most likely resulted from the lack of consensus Shine-Dalgarno sequences preceding these ORFs. We are currently using transcriptional *lacZ* fusion transposon mutagen-

esis and protein fusion approaches to define the length of the transcription unit of the *picA* locus. Both of these approaches have been highly successful in other systems (41).

Studies of bacterium-plant interactions are important for the understanding of pathogenicity and symbiosis of bacteria. Signal molecules from plants play major roles in regulating specific sets of genes of many bacteria that allow them to adapt rapidly to a new environment. *A. tumefaciens* senses phenolic compounds and opines and responds to new plant environments by changing the expression of sets of genes required for virulence and for the metabolism of carbon and nitrogen sources, respectively (19, 30, 50). Although the role of the *picA* locus in *A. tumefaciens*-plant interactions is not yet clear, it is interesting that acidic polysaccharides from plant cell walls act as signal molecules to induce *picA* and probably other unidentified loci on the *A. tumefaciens* chromosome (37). Studies of this novel induction system should reveal new aspects of *A. tumefaciens*-plant interactions.

ACKNOWLEDGMENTS

We thank Laszlo Csonka, and Todd Steck and other members of the Gelvin laboratory for useful discussions, David Overdier for help with S1 nuclease mapping, Gaochao Zhou and Jianke Zhang for help in DNA sequencing, Eureka Wang for help with Northern blotting, Barry Wanner for primer P21, Walt Ream for plasmid pAM10, Irwin Tessman for *E. coli* IT2761, and Rick Westerman for computer assistance. We are grateful to Lillian Habeck for technical assistance, Wilma Foust for assistance in preparation of the manuscript, and Laszlo Csonka and Todd Steck for critical reading of the manuscript.

Computer analysis involving the University of Wisconsin Genetics Computer Group programs was supported by NIH grant 27713. This work was supported by National Science Foundation Presidential Young Investigator Award DMB-835112 to S.B.G. and by U.S. Department of Agriculture grant 89-37262-4643.

REFERENCES

- Albright, L. M., E. Huala, and F. M. Ausubel. 1989. Prokaryotic signal transduction mediated by sensor and regulator protein pairs. *Annu. Rev. Genet.* **23**:311-336.
- Bolton, G. W., E. W. Nester, and M. P. Gordon. 1986. Plant phenolic compounds induce expression of the *Agrobacterium tumefaciens* loci needed for virulence. *Science* **232**:983-985.
- Burr, Thomas. 1991. Personal communication.
- Darzins, A., and A. M. Chakrabarty. 1984. Cloning of genes controlling alginate biosynthesis from a mucoid cystic fibrosis isolate of *Pseudomonas aeruginosa*. *J. Bacteriol.* **159**:9-18.
- Ditta, G., S. Stanfield, D. Corbin, and D. R. Helinski. 1980. Broad host range DNA cloning system for Gram-negative bacteria: construction of a gene bank of *Rhizobium meliloti*. *Proc. Natl. Acad. Sci. USA* **77**:7347-7351.
- Douglas, C. J., W. Halperin, and E. W. Nester. 1982. *Agrobacterium tumefaciens* mutants affected in attachment to plant cells. *J. Bacteriol.* **152**:1265-1275.
- Ellis, J. G., A. Kerr, A. Petit, and J. Tempe. 1982. Conjugal transfer of nopaline and agropine Ti-plasmids—the role of agrocinnopines. *Mol. Gen. Genet.* **186**:275-281.
- Ellis, J. G., P. J. Murphy, and A. Kerr. 1982. Isolation and properties of transfer regulatory mutants of the nopaline Ti-plasmid pTiC58. *Mol. Gen. Genet.* **186**:275-281.
- Figurski, D. H., and D. R. Helinski. 1979. Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided in *trans*. *Proc. Natl. Acad. Sci. USA* **76**:1648-1652.
- Gelvin, S. B., S. J. Karcher, and V. J. DiRita. 1983. Methylation of the T-DNA in *Agrobacterium tumefaciens* and in several crown gall tumors. *Nucleic Acids Res.* **11**:159-174.
- Genetello, C., N. Van Larebeke, M. Holsters, A. DePicker, M. Van Montagu, and J. Schell. 1977. Ti-plasmid of *Agrobacterium tumefaciens* as conjugative plasmids. *Nature (London)* **265**:561-563.
- Grierson, D., G. A. Tucker, J. Keen, J. Ray, C. R. Bird, and W. Schuch. 1986. Sequencing and identification of a cDNA clone for tomato polygalacturonase. *Nucleic Acids Res.* **14**:8595-8603.
- Hanahan, D. 1983. Studies on transformation of *Escherichia coli* with plasmids. *J. Mol. Biol.* **166**:557-580.
- Hawes, M. C., and S. G. Pueppke. 1987. Correlation between binding of *Agrobacterium tumefaciens* by root cap cells and susceptibility of plants to crown gall. *Plant Cell Rep.* **6**:289-290.
- He, S. Y., and A. Collmer. 1990. Molecular cloning, nucleotide sequence, and marker exchange mutagenesis of the exo-poly- α -D-galacturonosidase-encoding *pehX* gene of *Erwinia chrysanthemi* EC16. *J. Bacteriol.* **172**:4988-4995.
- Henikoff, S. 1980. Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. *Gene* **28**:351-359.
- Hinton, J. C. D., D. R. Gill, D. Lalo, G. S. Plastow, and G. P. C. Salmond. 1990. Sequence of the *peh* gene of *Erwinia carotovora*: homology between *Erwinia* and plant enzymes. *Mol. Microbiol.* **4**:1029-1036.
- Hirsch, P. R., and J. E. Beringer. 1984. A physical map of pPH1J1 and pJB4J1. *Plasmid* **12**:139-141.
- Hooykaas, P. J. J., C. Roobol, and R. A. Schilperoort. 1979. Regulation of the transfer of Ti-plasmids of *Agrobacterium tumefaciens*. *J. Gen. Microbiol.* **110**:99-109.
- Huang, J., and M. A. Schell. 1990. DNA sequence analysis of *pgla* and mechanism of export of its polygalacturonase product from *Pseudomonas solanacearum*. *J. Bacteriol.* **172**:3879-3887.
- Kerr, A., P. Manigault, and J. Tempe. 1977. Transfer of virulence *in vivo* and *in vitro* in *Agrobacterium*. *Nature (London)* **265**:560-561.
- Klapwijk, P. M., T. Scheuldermon, and R. A. Schilperoort. 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.
- Knauf, V. C., and E. W. Nester. 1982. Wide host range cloning vectors: a cosmid clone bank of an *Agrobacterium* Ti plasmid. *Plasmid* **8**:45-54.
- Krishnan, M., and S. B. Gelvin. Unpublished data.
- Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Matthyse, A. 1987. Effect of plasmid pSa and of auxin on attachment of *Agrobacterium tumefaciens* to carrot cells. *Appl. Environ. Microbiol.* **53**:2574-2582.
- Messing, J., R. Crea, and P. H. Seeburg. 1981. A system for shotgun DNA sequencing. *Nucleic Acids Res.* **9**:309-321.
- Metcalf, W. W., P. M. Steed, and B. L. Wanner. 1990. Identification of phosphate starvation-inducible genes in *Escherichia coli* K-12 by DNA sequence analysis of *psi::lacZ* (Mu dl) transcription fusions. *J. Bacteriol.* **172**:3191-3200.
- Miller, J. M. 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Montoya, A. L., L. W. Moore, M. P. Gordon, and E. W. Nester. 1978. Multiple genes coding for octopine-degrading enzymes in *Agrobacterium*. *J. Bacteriol.* **136**:909-915.
- Overdier, D. G., E. R. Olson, B. D. Erickson, M. M. Ederer, and L. N. Csonka. 1989. Nucleotide sequence of the transcriptional control region of the osmotically regulated *proU* operon of *Salmonella typhimurium* and identification of the 5' endpoint of the *proU* mRNA. *J. Bacteriol.* **171**:4694-4706.
- Pearson, W. R., and D. J. Lipman. 1988. Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci. USA* **85**:2444-2448.
- Petit, A., and J. Tempe. 1978. Isolation of *Agrobacterium* Ti-plasmid regulatory mutants. *Mol. Gen. Genet.* **167**:147-155.
- Petit, A., J. Tempe, A. Kerr, M. Holsters, M. Van Montagu, and J. Schell. 1978. Substrate induction of conjugal activity of *Agrobacterium tumefaciens* Ti-plasmids. *Nature (London)* **271**:570-572.

35. Rogowsky, P. M., B. S. Powell, K. Shirasu, T.-S. Lin, P. Morel, E. M. Zyprian, T. R. Steck, and C. I. Kado. 1990. Molecular characterization of the *vir* regulon of *Agrobacterium tumefaciens*: complete nucleotide sequence and gene organization of the 28.63-kbp regulon cloned as a single unit. *Plasmid* **23**:85–106.
36. Rong, L., and S. B. Gelvin. Unpublished data.
37. Rong, L., S. J. Karcher, K. O'Neal, M. C. Hawes, C. D. Yerkes, R. K. Jayaswal, C. A. Hallberg, and S. B. Gelvin. 1990. *picA*, a novel plant-inducible locus on the *Agrobacterium tumefaciens* chromosome. *J. Bacteriol.* **172**:5828–5836.
38. Salser, W., R. F. Gesteland, and A. Bolle. 1967. *In vitro* synthesis of bacteriophage lysozyme. *Nature (London)* **215**:588–591.
39. Sancar, A., A. M. Hack, and W. D. Rupp. 1979. Simple method for identification of plasmid-encoded proteins. *J. Bacteriol.* **137**:692–693.
40. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**:5463–5467.
41. Silhavy, T., and J. R. Beckwith. 1985. Uses of *lac* fusions for the study of biological problems. *Microbiol. Rev.* **49**:398–418.
42. Stachel, S. E., E. Messens, M. Van Montagu, and P. Zambryski. 1985. Identification of the signal molecules produced by wounded plant cells that activate T-DNA transfer in *Agrobacterium tumefaciens vir* gene expression. *Nature (London)* **318**:624–629.
43. Stachel, S. E., E. W. Nester, and P. Zambryski. 1986. A plant factor induces *Agrobacterium tumefaciens vir* gene expression. *Proc. Natl. Acad. Sci. USA* **83**:379–383.
44. Tempe, J., and A. Petit. 1982. Opine utilization by *Agrobacterium*, p. 451–459. In G. Kahl and J. S. Schell (ed.), *Molecular biology of plant tumors*. Academic Press, Inc., Orlando, Fla.
45. Veluthambi, K., M. Krishnan, J. H. Gould, R. H. Smith, and S. B. Gelvin. 1989. Opines stimulate induction of the *vir* genes of the *Agrobacterium tumefaciens* Ti plasmid. *J. Bacteriol.* **171**:3696–3703.
46. Vieira, J., and J. Messing. 1987. Production of single stranded DNA. *Methods Enzymol.* **153**:3–11.
47. Wang, K., A. Herrera-Estrella, and M. Van Montagu. 1990. Overexpression of *virD1* and *virD2* genes in *Agrobacterium tumefaciens* enhances T-complex formation and plant transformation. *J. Bacteriol.* **172**:4432–4440.
48. Worcel, A., and E. Burgi. 1974. Properties of a membrane-attached form of the folded chromosome of *Escherichia coli*. *J. Mol. Biol.* **82**:91–105.
49. Yanisch-Perron, C., J. Vieira, and J. Messing. 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**:103–119.
50. Zambryski, P. 1988. Basic processes underlying *Agrobacterium*-mediated DNA transfer to plant cells. *Annu. Rev. Genet.* **22**:1–3.