Proc. Natl. Acad. Sci. USA
Vol. 81, pp. 1728–1732, March 1984
Cell Biology

Nucleotide sequence of the *tms* genes of the pTiA6NC octopine Ti plasmid: Two gene products involved in plant tumorigenesis

(*Agrobacterium tumefaciens*)

HARRY KLEE*, ALICE MONTOYA*, FRANK HORODYSKI†, CONRAD LICHTENSTEIN*‡, DAVID GARFINKEL*§, SHERYL FULLER*‡, CARLOS FLORES*, JACQUES PESCHON*, EUGENE NESTER*, AND MILTON GORDON†

*Department of Microbiology and Immunology and †Department of Biochemistry, University of Washington, Seattle, WA 98195

Communicated by Earl W. Davie, December 12, 1983

ABSTRACT **The nucleotide sequence of the tumor morphology locus, *tms*, from pTiA6NC has been determined. The sequence analysis indicates that each of two polyadenylylated transcripts encoded by this locus contains an open reading frame; the predicted transcript 1 gene product has a molecular size of 83,769 daltons, and the predicted transcript 2 gene product, of 49,588 daltons. The precise start and stop positions of the transcript 2 RNA have been mapped with S1 nuclease. Several insertion mutations have been constructed. One of these localizes the transcript 2 promoter within the 72 base pairs 5′ to transcription initiation. Significant homology was observed between the protein encoded by transcript 1 and the adenine binding region of *p*-hydroxybenzoate hydroxylase from *Pseudomonas fluorescens*, suggesting that the transcript 1 protein binds adenine either as substrate or cofactor.**

Crown gall is a neoplastic disease induced by *Agrobacterium tumefaciens* affecting most dicotyledonous plants (reviewed in refs. 1 and 2). Both the ability to transform plant cells and host range of the bacteria reside on a large plasmid, the Ti plasmid. The bacteria transfer a specific portion of the Ti plasmid to the plant cell, where this transferred DNA (T-DNA) is stably integrated into plant nuclear DNA. Axenic cultures of crown gall tissue exhibit autonomous growth in the absence of the phytohormones, auxin, and cytokinin. In addition, transformed tissue directs the synthesis of opines, derivatives of certain amino acids that Ti plasmid-containing strains of *A. tumefaciens* can use as sole sources of carbon and nitrogen.

Tumors induced by the octopine-type Ti plasmid, pTiA6NC, or the closely related pTiB6S3 synthesize eight polyadenylylated transcripts from genes entirely within the T-DNA (3, 4). The DNA sequences of two of these genes, neither essential for tumor formation, have been reported. The product of transcript 3 is responsible for synthesis of octopine by transformed plant tissue (5). The function of the protein encoded by transcript 7 is unknown (6). Five additional transcripts encoded from three distinct genetic loci appear to play a role in tumor morphology (7). Mutations in these loci result in synthesis of shoots (*tms*), synthesis of roots (*tmr*), or enlarged tumors (*tml*). The phenotypes of these different transformed tissues can be correlated with alterations in normal cytokinin/auxin ratios of the tissue (8) and can be corrected by addition of exogenous phytohormones (9). The function of the transcript 2 protein appears to be conversion of indole-3-acetamide to the auxin, indoleacetic acid (J. Schröder, personal communication). It seems likely that the other tumor morphology genes are also involved in phytohormone biosynthesis, although this conclusion is by no means certain. We have determined the nucleotide sequences of two morphology genes, those encoding *tms* transcripts 1 and 2, as well as the sites of transcription initiation and termination for the latter gene. Both of these loci are implicated in conferring auxin autonomy to transformed plant tissue. A mutation in either locus confers the tms phenotype on a tumor. Sequence analysis indicates that transcripts 1 and 2 contain open reading frames (orfs) capable of encoding proteins of 83,769 and 49,588 daltons, respectively.

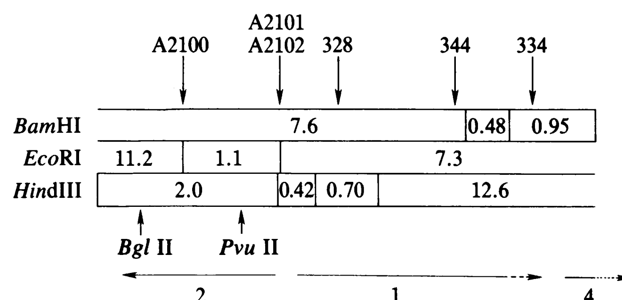

FIG. 1. Restriction endonuclease map of the *tms* region of pTiA6NC. Numbering of fragments is based on sizes in kb. The positions and directions of transcripts 1 and 2 as well as for the *tmr* gene product, transcript 4, are indicated by horizontal arrows. The exact positions of initiation and termination of transcript 1 have not been determined. Positions of insertions resulting in a *tms* morphology (A2100, 328, and 344) or a wild-type morphology (A2101, A2102, and 334) are indicated by vertical arrows.

MATERIALS AND METHODS

Materials. Restriction endonucleases were purchased from New England BioLabs or Bethesda Research Laboratories and used according to the manufacturers' specifications. DNA polymerase I Klenow fragment was obtained from New England BioLabs. S1 nuclease was from Bethesda Research Laboratories. [α-^{32}P]dNTPs were purchased from New England Nuclear, and nonradiolabeled nucleotides were purchased from P-L Biochemicals. Isopropylthio-β-D-galactoside and 5-bromo-4-chloro-indoyl β-D-galactoside were from Bethesda Research Laboratories.

Culture Conditions. LB broth and agar were prepared as described by Miller (10). *Agrobacterium* strains were grown on AB minimal medium (11).

Nucleic Acid Sequence Determination. All plasmid and phage replicative form DNAs were prepared by the method

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

Abbreviations: T-DNA, DNA transferred from the Ti plasmid to a plant cell; orf, open reading frame; bp, base pair(s); kb, kilobase pair(s).

‡Present address: Biotechnology Centre, Imperial College, London SW7 2AZ, England.
§Present address: Biology Department, Massachusetts Institute of Technology, Cambridge, MA 02139.

of Clewell and Helinski (12). The strategies used in the sequence determination are described elsewhere (13). Briefly, 100- to 300-base-pair (bp) fragments of T-DNA were generated by restriction endonuclease digestion or by sonication. These fragments were cloned into M13 mp7, mp8, or mp9 (14) and screened by plaque hybridization (15); their sequences were determined by the dideoxy methods described by Sanger *et al.* (16, 17). The results are shown in Fig. 1.

Isolation of RNA From Crown Gall Tissue. The octopine crown gall line A6S2 (18) was grown as a suspension in hormone-free MS medium (19) and harvested in late logarithmic phase. Total RNA was prepared as described by Taylor and Powell (20) and polyadenylylated RNA was isolated by oligo(dT)-cellulose chromatography (21).

S1 Nuclease Mapping of Transcript 2. Unique single-strand-labeled probes spanning the 5′ and 3′ termini of transcript 2 RNA were prepared as described (13) by *in vitro*-primed synthesis of M13 cloned DNAs. Hybridizations and S1 nuclease digestions were carried out as described (13).

Isolation and Characterization of Mutations. Tn*5* insertion mutations previously isolated in this laboratory (7), which are located in and around the *tms* locus, were accurately mapped (within 10 bp) by restriction endonuclease digestions. Additional mutations in and adjacent to transcript 2 were made in the following way: a clone of *Bam*HI fragment 8 [the 7.6-kilobase-pair (kb) fragment in Fig. 1], in the *Bgl* II site of pRK290 (7), was partially digested with *Eco*RI. A gene encoding kanamycin resistance isolated from pUC4K

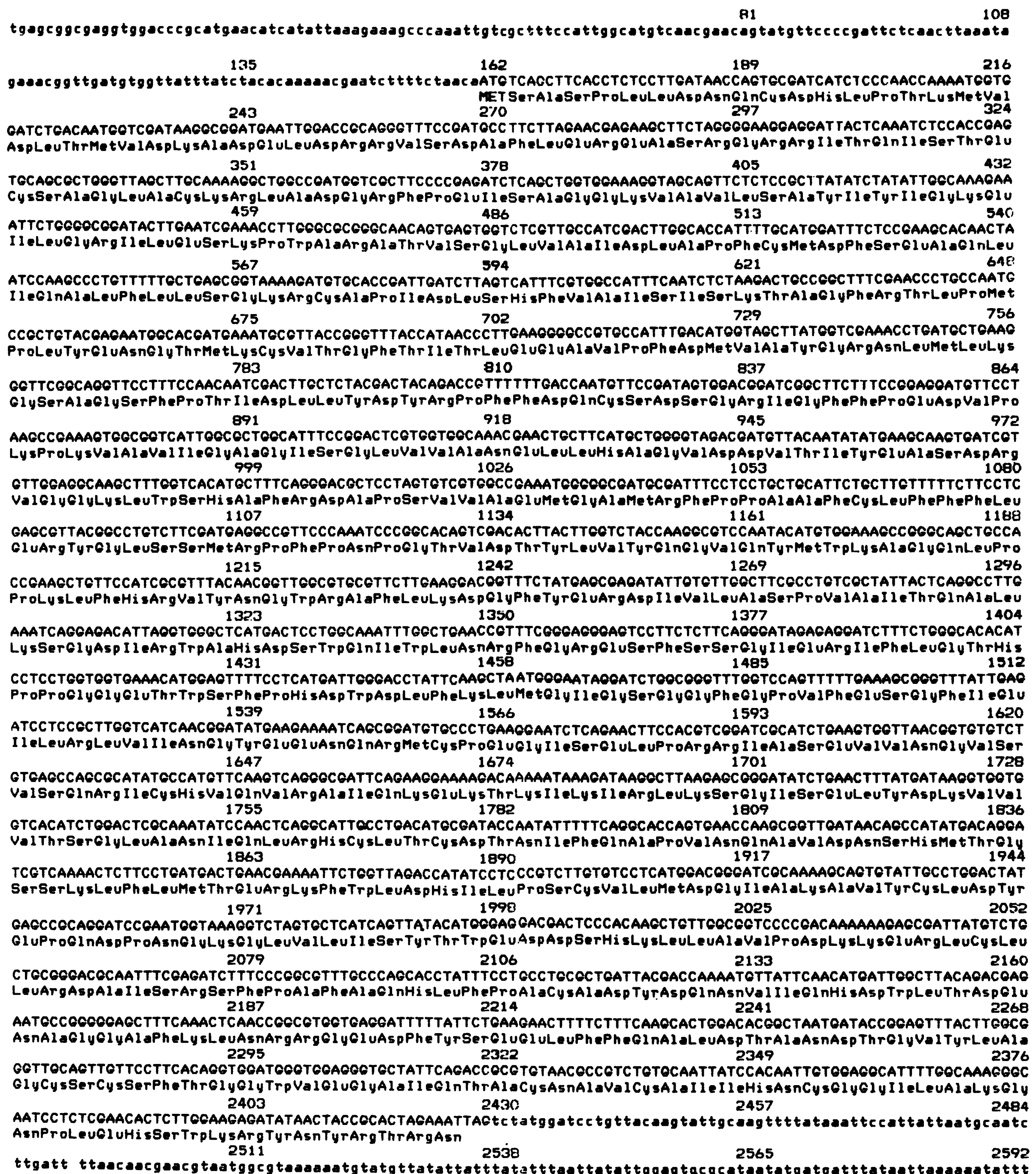

FIG. 2. Nucleotide sequence of transcript 1. The orf is shown in uppercase letters, and the amino acid sequence is shown below the nucleotide sequence.

(22) and containing *Eco*RI sticky ends was ligated into each of the *Eco*RI sites in the Bam 8 clone. After selection on kanamycin-containing medium, clones having the kanamycin-resistance gene inserted into either the left or right end of the 1.1-kb *Eco*RI fragment were identified. These mutations were introduced into pTiA6NC by the marker-rescue technique as described by Garfinkel *et al.* (7). Mutant strains were tested for virulence on *Kalanchoe diagremontiana* and *Nicotiana tabacum* as described (23).

RESULTS

DNA Sequence Determination. Mutational analyses using a variety of techniques have indicated that the *tms* locus in pTiA6NC covers approximately 5 kb (7, 24, 25). RNA blotting of polyadenylylated transcripts has shown that the locus actually encodes two RNAs; transcript 1 (2.7 kb) and transcript 2 (1.75 kb) (26) (Fig. 1). The sequence of the DNA of the region covering these two transcripts has been determined and is shown in Figs. 2 and 3. There are two orfs extending in opposite orientations out from the 0.42-kb *Hin*dIII fragment and separated by 250 bp. The first orf (*tms1*) encodes a protein with a subunit molecular weight of 83,769 and is consistent with the position and direction of transcription of transcript 1. Transcript 2 contains a second orf (*tms2*) encoding a protein with a subunit molecular weight of 49,588.

Amino Acid Sequence Homology. To obtain some insight into the functions of the *tms* gene products, a search for homology to known proteins was carried out using the protein sequence database of the *Atlas of Protein Sequence and Structure* (27). Tms1 displayed significant homology to *p*-hydroxybenzoate hydroxylase from *Pseudomonas fluorescens*. Specifically, residues 239–263 of tms1 showed a high degree of identity to residues 5–29 of the hydroxylase (Fig. 4). Residues 5–31 of the hydroxylase have been determined by crystallographic analysis to form a pocket that binds the adenine moiety of FAD (28). It thus seems likely that tms1 also binds adenine in some form either as substrate or cofac-

```
                                          -72
                                          EcoRI
tgcctacagtcattcgtatcacacggcgtgattgctgaattcccaccaataatggcgcaagctgggttcaagcttggtatatttatttggtctgaatgggtttgaaat

1
TTCCAACTCAGAGAGATGGTGGCCATTACCTCGTTAGCCCAAAGCCTAGAACACCTGAAACGGAAAGACTACTCCTGCTTAGAACTAGTAGAAACTCTGATAGCGCGT
                 METValAlaIleThrSerLeuAlaGlnSerLeuGluHisLeuLysArgLysAspTyrSerCysLeuGluLeuValGluThrLeuIleAlaArg

109
TGTGAAGCTGCAAAATCATTAAACGCCCTTCTGGCTACAGACTGGGATGGTTTGCGGCGAAGCGCCAAAAAAATTGATCGCCATGGAAACGCCGGAGTAGGTCTTTGC
CysGluAlaAlaLysSerLeuAsnAlaLeuLeuAlaThrAspTrpAspGlyLeuArgArgSerAlaLysLysIleAspArgHisGlyAsnAlaGlyValGlyLeuCys

217
GGCATTCCACTCTGTTTTAAGGCGAACATCGCTACCGGCGTATTTCCCACAAGCGCCGCTACGCCGGCGCTGATAAACCACTTGCCAAAGATACCATCCCGCGTCGCA
GlyIleProLeuCysPheLysAlaAsnIleAlaThrGlyValPheProThrSerAlaAlaThrProAlaLeuIleAsnHisLeuProLysIleProSerArgValAla

325            PvuII
GAAAGACTTTTTTCAGCTGGAGCACTGCCGGGTGCCTCGGGAAATATGCATGAGTTATCGTTTGGAATTACAAGCAACAACTATGCCACCGGGGCGGTGCGAAACCCG
GluArgLeuPheSerAlaGlyAlaLeuProGlyAlaSerGlyAsnMetHisGluLeuSerPheGlyIleThrSerAsnAsnTyrAlaThrGlyAlaValArgAsnPro

433
TGGAATCCAGATCTGATACCAGGGGGCTCAAGCGGTGGTGTGGCTGCTGCGGTAGCAAGCCGATTGATGTTAGGCGGCATAGGCACCGATACCGGTGCATCTGTTCGC
TrpAsnProAspLeuIleProGlyGlySerSerGlyGlyValAlaAlaAlaValAlaSerArgLeuMetLeuGlyGlyIleGlyThrAspThrGlyAlaSerValArg

541
CTACCCGCAGCCCTGTGTGGCGTAGTAGGATTTCGACCGACGCTTGGTAGATATCCGGGAGATCGGATAATACCGGTTAGCCCTACCCGGGACACTCCCGGAATCATA
LeuProAlaAlaLeuCysGlyValValGlyPheArgProThrLeuGlyArgTyrProGlyAspArgIleIleProValSerProThrArgAspThrProGlyIleIle

649
GCGCAGTGCGTAGCCGATGTTGTAATCCTCGACCGGATAATTTCCGGCACACCGGAGAGAATACCACCCGTGCCGCTGAAGGGGCTAAGGATCGGCCTCCCTACAACC
AlaGlnCysValAlaAspValValIleLeuAspArgIleIleSerGlyThrProGluArgIleProProValProLeuLysGlyLeuArgIleGlyLeuProThrThr

757
TACTTTTATGATGACCTTGATGCTGATGTGGCCCTAGCAGCTGAAACAACGATTCGCCTGCTAGCAAACAAAGGCGTAACTTTTGTTGAAGCTAACATTCCCCACCTT
TyrPheTyrAspAspLeuAspAlaAspValAlaLeuAlaAlaGluThrThrIleArgLeuLeuAlaAsnLysGlyValThrPheValGluAlaAsnIleProHisLeu

865
GACGAACTGAATAAAGGGGCCAGCTTCCCAGTTGCACTCTATGAATTTCCACACGCTCTAAAACAGTATCTCGACGACTTTGTAAAAACTGTTTCTTTTTCTGACGTC
AspGluLeuAsnLysGlyAlaSerPheProValAlaLeuTyrGluPheProHisAlaLeuLysGlnTyrLeuAspAspPheValLysThrValSerPheSerAspVal

973    EcoRI
ATCAAAGGAATTCGTAGCCCTGATGTAGCCAACATTGCCAATGCGCAAATTGATGGACATCAAATTTCCAAAGCTGAATATGAACTGGCCCGCCACTCCTTCAGACCA
IleLysGlyIleArgSerProAspValAlaAsnIleAlaAsnAlaGlnIleAspGlyHisGlnIleSerLysAlaGluTyrGluLeuAlaArgHisSerPheArgPro

1081
AGACTTCAAGCCACCTATCGCAACTACTTCAAACTGAATAGATTAGATGCTATTCTCTTCCCAACAGCACCCTTGGTGGCCAGACCCATAGGTCAGGATTCCTCAGTT
ArgLeuGlnAlaThrTyrArgAsnTyrPheLysLeuAsnArgLeuAspAlaIleLeuPheProThrAlaProLeuValAlaArgProIleGlyGlnAspSerSerVal

1189                          BglII
ATCCACAATGGCACGATGCTGGACACATTCAAGATCTACGTGCGAAATGTGGACCCAAGCAGCAACGCAGGCCTACCTGGCTTGAGCATTCCTGTTTGCCTGACACCT
IleHisAsnGlyThrMetLeuAspThrPheLysIleTyrValArgAsnValAspProSerSerAsnAlaGlyLeuProGlyLeuSerIleProValCysLeuThrPro

1297
GATCGCTTGCCTGTTGGAATGGAGATCGATGGATTAGCGGATTCAGACCAACGTCTGTTAGCAATCGGGGGGGCATTGGAAGAAGCCATTGGATTCCGATATTTTGCC
AspArgLeuProValGlyMetGluIleAspGlyLeuAlaAspSerAspGlnArgLeuLeuAlaIleGlyGlyAlaLeuGluGluAlaIleGlyPheArgTyrPheAla

1405
GGTTTACCCAATTAAACTTTCTACCATGTTCGTTTTTTACATTTTTCAGATTGATGACATCAATCCTTGTATTGCGTCTATGAACAACAGTCGCCTTATGTTATAAAT
GlyLeuProAsn

1513       *      *
CGAATAATAACTTgcgatggagatttgaacaaactttaatttatgatttaaccaataaaagtctttgcaataatcatgttcgatcgataaataattttattatcagt
```

FIG. 3. Nucleotide sequence of transcript 2. The transcribed nucleotides, as determined by S1 nuclease mapping, are indicated by uppercase letters. Numbers are given with respect to the start of transcription. Possible regulatory sequences are underlined with a single line. A sequence capable of forming an 8-bp hairpin structure in the mRNA is indicated by a heavy line. Asterisks indicate the two alternative transcription termination sites.

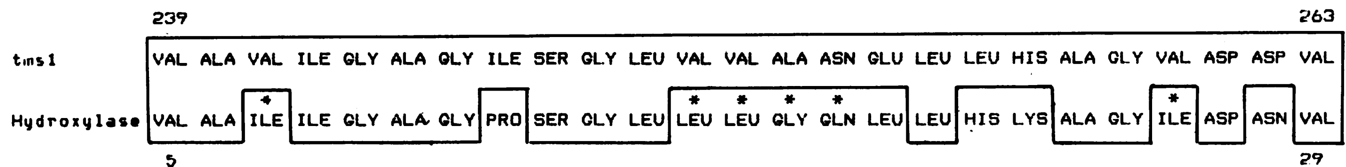

FIG. 4. Amino acid homology between the transcript 1 gene product and *P. fluorescens* p-hydroxybenzoate hydroxylase. Conserved amino acids are enclosed. Functionally similar amino acids are indicated by asterisks.

tor. No significant homology was found for the transcript 2 gene product.

Transcript Characterization. The 5′ end of transcript 2 was precisely mapped by S1 nuclease protection of radiolabeled DNA prepared by *in vitro* fill-in of an M13 clone spanning this region. A 404-bp *Pvu* II/*Eco*RI fragment (Fig. 1) was purified following primed radiolabeling of a single-stranded M13 template containing an insertion of the 1.1-kb *Eco*RI fragment. When the probe complementary to transcript 2 was hybridized to polyadenylylated RNA prepared from plant tumors, a fragment of 338 bp was observed after S1 nuclease digestion (Fig. 5*A*). This fragment determines the 5′ end of transcript 2 to be 16 bp 5′ of the ATG translation initiation codon. When a probe of opposite orientation was used, no S1 nuclease-protected fragments were observed.

The 3′ end of transcript 2 was determined in the same manner except that the labeled probe for protection was a *Bgl* II/*Hin*dIII fragment. In this case, protected fragments present in approximately equal amounts were observed at 313 bp and 323 bp (Fig. 5*B*). Thus, this RNA appears to be heterogeneous, ending 118 and 128 nucleotides beyond the termination codon (positions 1530 and 1540, respectively).

Mutational Analysis of the *tms* Region. Several Tn*5* insertion mutations were previously mapped to the *tms* region (7). Since the DNA sequences of both the *tms* region and Tn*5* (29) have been determined, it was possible to map the sites of

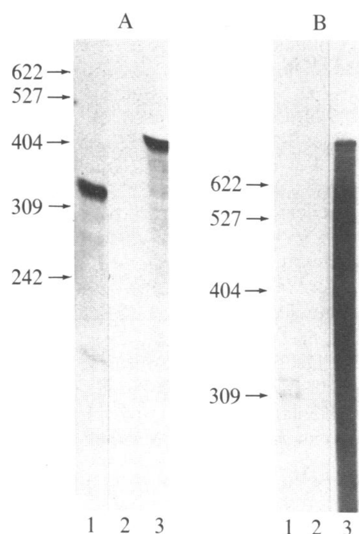

FIG. 5. S1 nuclease mapping of the initiation and termination sites for transcript 2. (*A*) 5′-End mapping. The S1 nuclease-resistant fragment resulting from hybridization of the *Pvu* II/*Eco*RI fragment (Fig. 1) ^{32}P-labeled in the coding strand with poly(A)$^+$ A6S2 RNA was sized on a 6% polyacrylamide urea gel (lane 1). As a control, the probe was hybridized to yeast tRNA and the hybrid was digested with nuclease S1 (lane 2). The undigested *Pvu* II/*Eco*RI probe is also shown (lane 3). (*B*) 3′-End mapping. The *Hin*dIII/*Bgl* II fragment was labeled in the coding strand (Fig. 1). The lanes are arranged as in *A*: hybridization with tumor RNA (lane 1) and tRNA (lane 2) followed by S1 digestion, undigested *Hin*dIII/*Bgl* II probe (lane 3). The positions of *Hpa* II-digested pBR322 size markers are shown to the left of the gels.

insertion by restriction endonuclease digestion. Three mutations, 328, 334, and 344, were mapped in this way. Both 328 and 344 were localized within the orf of transcript 1 (positions 337 and 1934, respectively) and are *tms* mutants. Mutation 334 mapped 85 bp downstream from the *tms1* orf (position 2531) and is wild type.

Because no transposon insertions in the 0.42-kb *Hin*dIII fragment were available and this restriction fragment likely contains the promoters for both *tms* transcripts, we made additional mutations by inserting an *Eco*RI fragment encoding kanamycin resistance into each of the two *Eco*RI sites located in the 7.6-kb *Bam*HI fragment (Fig. 1). The leftward insertion, A2100, is within the orf of transcript 2 (position 982) and results in a tms phenotype. The rightward insertion is located 72 bp 5′ to the start of transcript 2 (position −72). Insertion of the kanamycin-resistance gene in either orientation (A2101 and A2102) results in a wild-type phenotype. These insertions thus appear to separate the two *tms* genes as well as localize the entire promoter for transcript 2 within the 72 bp between the *Eco*RI site and the start of transcription.

Analysis of Nontranslated Regions. We have observed a conserved 9-nucleotide sequence, T-T-T-C-A-A-G-G-A, located 100–200 nucleotides upstream from the start of transcription in a number of T-DNA genes including transcripts 3 (octopine synthase), 4, and 7 from pTiA6NC as well as nopaline synthase from the nopaline Ti plasmid, pTiT37 (13). Although the exact position of the start of transcription for transcript 1 is not known, a sequence closely resembling the consensus, T-G-T-C-A-A-A-G-A, appears 85 nucleotides upstream from the start of translation. In contrast, transcript 2 does not have anything resembling this conserved sequence in the proper orientation in the 5′ nontranscribed region. The upstream region of transcript 2 does contain a sequence homologous to the "TATA" box observed in most eukaryotic genes (30) as well as a sequence resembling the "CCAAT" box, usually located 5′ to the TATA box (31).

Both of the *tms* transcripts are polyadenylylated in plant cells. A search of the sequences 3′ to the orfs of these two transcripts reveals sequences resembling the canonical polyadenylylation sequence, A-A-T-A-A-A. Sequences with a five out of six base match appear in several places 3′ to the orf in both transcripts. Two of these, T-A-T-A-A-A and A-A-T-A-A-T, occur immediately preceding the sites of termination in transcript 2. There is also a sequence located midway between translational and transcriptional termination (position 1459) that is a perfect 8-bp palindrome separated by one nucleotide. Such hairpin structures have been implicated in prokaryotic transcription termination but do not appear to be involved in eukaryotic termination (reviewed in ref. 32).

DISCUSSION

A. tumefaciens containing mutations in the *tms* locus induce tumors that have a proliferation of shoots on many host plants. These tumors have greatly reduced levels of auxins compared with tumors induced by the wild-type strain (8). At least some of these mutants contain elevated cytokinin levels as well. Addition of exogenous auxins to tumor tissue results in a return to the wild-type, unorganized phenotype

(9). Thus there is a suggestion that the *tms* locus codes for some product(s) that synthesizes or stabilizes auxins or activates the plant's own auxin biosynthetic pathway. Indeed, recent evidence suggests that the function of the transcript 2 protein in tumors is to convert indole-3-acetamide to the phytohormone indoleacetic acid (J. Schröder, personal communication).

RNA blot analyses suggest that the *tms* loci of pTiA6NC and closely related octopine Ti plasmids code for two α-amanatin-sensitive polyadenylylated transcripts (3, 4). We have shown that each of these transcripts contains an orf; transcript 1 can encode a protein of 83,769 daltons and transcript 2 can encode a protein of 49,588 daltons. The predicted size of the transcript 2 gene product agrees with the data of Schroder *et al.* (33), who observe proteins of 74,000, 49,000, and 28,000 daltons in *Escherichia coli* minicells containing clones covering the *tms* region. The predicted size of tms1 is somewhat larger than observed; however, this may simply reflect an aberrant electrophoretic mobility on $NaDodSO_4$/polyacrylamide gels. Alternatively, the 74,000-dalton protein may represent a processed or degraded form of the tms1 protein. We observe no gene in the *tms* or the adjacent *tmr* region corresponding to a 28,000-dalton protein.

A significant homology between tms1 and the adenine-binding domain of *P. fluorescens p*-hydroxybenzoate hydroxylase suggests that tms1 uses adenine as either a substrate or a cofactor. Although no significant homology to other regions of the hydroxylase was found, it is interesting to note that the product of the hydroxylase, a diphenolic acid, enhances the growth effect of indoleacetic acid by blocking its decarboxylation (34).

The DNA sequences of these and other T-DNA genes provide few clues about the origin of these genes. That is, are they prokaryotic or eukaryotic? Codon usage for these two genes reveals neither a prokaryotic nor a eukaryotic bias (data not shown). The observation that the genes are expressed at low levels in both plant cells and *E. coli* minicells suggests that these genes are somewhat adaptable but are not entirely suited for either situation. They are, however, capable of expressing at high enough levels to bring about transformation of the plant cell. Quantitation of the levels of transcripts 1 and 2 by dot blot hybridization indicates that each of these mRNAs represents $<0.001\%$ of polyadenylylated tumor mRNA (unpublished observations).

The fact that a sequence of 9 bp is conserved upstream of transcription initiation in four octopine T-DNA genes and one nopaline T-DNA gene suggests that this sequence is or once was important for expression. This sequence is not, however, found upstream of transcript 2. Examination of all T-DNA promoters for which sequences have been determined reveals no other striking similarity in either primary or secondary structure. It is possible that the conserved sequence was originally important for expression in *Agrobacterium* and is not necessary for expression in plant cells. Alternatively, the consensus sequence upstream of the transcript 1 orf may also stimulate transcript 2 expression.

Insertions A2101 and A2102, both wild type, localize the entire transcript 2 promoter within the 72 bp between the *Eco*RI site and the start of transcription. Quantitation of transcript 2 in tumor lines containing these insertions may, however, show decreased expression of this gene. Transcript 2 closely resembles other eukaryotic genes in that it contains sequences homologous to the TATA and CCAAT boxes. It should be noted, however, in light of the probable prokaryotic origin of the T-DNA genes that the TATA sequence, centered at -28, also bears a remarkable resemblance to the Pribnow box, found prior to the start of transcription in *E. coli* (35).

We would like to give special thanks to Tom Hollon for his help with the computer programming and to Dr. Kenneth Walsh and Paul Karplus for help with the protein homology searches. This research was supported by a grant from Standard Oil (Indiana). H.K. was supported by National Cancer Institute Fellowship CA 07249-01. C.L. was supported by a Science and Engineering Research Council–North Atlantic Treaty Organization fellowship. F.H. was supported by American Cancer Society Fellowship PF 2150.

1. Nester, E. & Kosuge, T. (1981) *Annu. Rev. Microbiol.* **35**, 531–565.
2. Bevan, M. & Chilton, M.-D. (1982) *Annu. Rev. Genet.* **16**, 357–384.
3. Gelvin, S., Thomashow, M., McPherson, J., Gordon, M. & Nester, E. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 76–80.
4. Willmitzer, L., Simons, G. & Schell, J. (1982) *EMBO J.* **1**, 139–146.
5. De Greve, H., Dhaese, P., Seurinck, J., Lemmers, M., Van Montagu, M. & Schell, J. (1983) *J. Mol. Appl. Genet.* **1**, 499–511.
6. Dhaese, P., De Greve, H., Gielen, J., Seurinck, J., Van Montagu, M. & Schell, J. (1983) *EMBO J.* **2**, 419–426.
7. Garfinkel, D., Simpson, R., Ream, L., White, F., Gordon, M. & Nester, E. (1981) *Cell* **27**, 143–153.
8. Akiyoshi, D., Morris, R., Hinz, R., Mischike, B., Kosuge, T., Garfinkel, D., Gordon, M. & Nester, E. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 407–411.
9. Ooms, G., Hooykaas, P., Moolenaar, G. & Schilperoort, R. (1981) *Gene* **14**, 33–50.
10. Miller, J. H. (1972) *Experiments in Molecular Genetics* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
11. Chilton, M.-D., Currier, T., Farrand, S., Bendich, A., Gordon, M. & Nester, E. (1974) *Proc. Natl. Acad. Sci. USA* **71**, 3672–3676.
12. Clewell, D. & Helinski, D. (1972) *J. Bacteriol.* **110**, 667–676.
13. Lichtenstein, C., Klee, H., Montoya, A., Garfinkel, D., Fuller, S., Flores, C., Nester, E. & Gordon, M. (1984) *J. Mol. Appl. Genet.*, in press.
14. Messing, J. & Vieira, J. (1982) *Gene* **19**, 269–276.
15. Grunstein, M. & Hogness, D. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 3961–3965.
16. Sanger, F., Nicklen, S. & Coulson, A. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
17. Sanger, F., Coulson, A., Smith, A. & Roe, B. (1980) *J. Mol. Biol.* **143**, 161–178.
18. Thomashow, M., Nutter, R., Montoya, A., Gordon, M. & Nester, E. (1980) *Cell* **19**, 729–739.
19. Murashige, T. & Skoog, F. (1962) *Physiol. Plant.* **15**, 473–497.
20. Taylor, B. & Powell, A. (1982) *Bethesda Res. Lab. Focus* **4**, 4–6.
21. Aviv, H. & Leder, P. (1972) *Proc. Natl. Acad. Sci. USA* **69**, 1408–1412.
22. Vieira, J. & Messing, J. (1982) *Gene* **19**, 259–268.
23. Garfinkel, D. & Nester, E. (1980) *J. Bacteriol.* **144**, 732–743.
24. Ooms, H., Klapwijk, P., Poulsi, J. & Schilperoort, R. (1980) *J. Bacteriol.* **144**, 82–91.
25. De Greve, H., Decraemer, H., Seurinck, J., Van Montagu, M. & Schell, J. (1981) *Plasmid* **6**, 235–248.
26. Leemans, J., Deblaere, R., Willmitzer, L., De Greve, H., Hernalsteens, J., Van Montagu, M. & Schell, J. (1982) *EMBO J.* **1**, 147–152.
27. Dayhoff, M. O., ed. (1983) *Atlas of Protein Sequence and Structure* (Natl. Biomed. Res. Found., Washington, DC), 6th Ed.
28. Wierenga, R., De Jong, R., Kalk, K., Hol, W. & Drenth, J. (1979) *J. Mol. Biol.* **131**, 55–73.
29. Auerswald, E., Ludwig, G. & Schaller, H. (1980) *Cold Spring Harbor Symp. Quant. Biol.* **45**, 107–114.
30. Breathnach, R. & Chambon, P. (1981) *Annu. Rev. Biochem.* **50**, 349–383.
31. Shenk, T. (1981) *Curr. Top. Microbiol. Immunol.* **93**, 25–46.
32. Kozak, M. (1983) *Microbiol. Rev.* **47**, 1–45.
33. Schroder, G., Klipp, W., Hillebrand, A., Ehring, R., Koncz, C. & Schroder, J. (1983) *EMBO J.* **2**, 403–409.
34. Tomaszewski, M. & Thimann, K. (1966) *Plant Physiol.* **41**, 1443–1454.
35. Rosenberg, M. & Court, D. (1979) *Annu. Rev. Genet.* **13**, 319–353.