Proc. Natl. Acad. Sci. USA
Vol. 82, pp. 6522–6526, October 1985
Biochemistry

Nucleotide sequences of the *Pseudomonas savastanoi* indoleacetic acid genes show homology with *Agrobacterium tumefaciens* T-DNA

(plant tumorigenicity/oleander knot/tryptophan monooxygenase/indoleacetamide hydrolase)

TETSUJI YAMADA, CURTIS J. PALM, BOB BROOKS, AND TSUNE KOSUGE

Department of Plant Pathology, University of California, Davis, CA 95616

Communicated by Luis Sequeira, June 17, 1985

ABSTRACT **We report the nucleotide sequences of *iaaM* and *iaaH*, the genetic determinants for, respectively, tryptophan 2-monooxygenase and indoleacetamide hydrolase, the enzymes that catalyze the conversion of L-tryptophan to indoleacetic acid in the tumor-forming bacterium *Pseudomonas syringae* pv. *savastanoi*. The sequence analysis indicates that the *iaaM* locus contains an open reading frame encoding 557 amino acids that would comprise a protein with a molecular weight of 61,783; the *iaaH* locus contains an open reading frame of 455 amino acids that would comprise a protein with a molecular weight of 48,515. Significant amino acid sequence homology was found between the predicted sequence of the tryptophan monooxygenase of *P. savastanoi* and the deduced product of the T-DNA *tms-1* gene of the octopine-type plasmid pTiA6NC from *Agrobacterium tumefaciens*. Strong homology was found in the 25 amino acid sequence in the putative FAD-binding region of tryptophan monooxygenase. Homology was also found in the amino acid sequences representing the central regions of the putative products of *iaaH* and *tms-2* T-DNA. The results suggest a strong similarity in the pathways for indoleacetic acid synthesis encoded by genes in *P. savastanoi* and in *A. tumefaciens* T-DNA.**

The association of the tumor-forming bacterium *Pseudomonas syringae* pv. *savastanoi* (*P. savastanoi*) and its hosts, oleander and olive plants, provides a system for studying the molecular basis of virulence of a bacterium in plants. Tumor formation by these plants is a response to high concentrations of indoleacetic acid (IAA) introduced into infected tissue by the bacterium (1); thus, production of a tumor is used to assess virulence of the bacterium. The bacterium produces IAA from tryptophan, with indoleacetamide as the intermediate. The enzymes involved are tryptophan 2-monooxygenase [L-tryptophan:oxygen 2-oxidoreductase (decarboxylating), EC 1.13.12.3], which catalyzes the conversion of L-tryptophan to indole-3-acetamide, and indoleacetamide hydrolase, which catalyzes the conversion of indoleacetamide to ammonia and IAA (2).

The genes for the two enzymes, *iaaM* and *iaaH*, are part of an operon that is borne on a plasmid, pIAA, in oleander strains of the pathogen; in olive strains these genes are on the chromosome. Mutants cured of pIAA are weakly virulent on oleander; when transformed with pIAA, they are restored to full virulence (3). Moreover, *iaaM* has been cloned and its role in virulence has been demonstrated (3–5). Unlike the crown gall disease caused by *Agrobacterium tumefaciens*, in which transferred DNA (T-DNA) from the tumor-inducing (Ti) plasmid is stably integrated into the nuclear genome (6–12), there appears to be no genetic transformation of host tissue by *P. savastanoi*.

To determine whether there is a common basis for bacteria-induced tumor formation in plants, we compared IAA synthesis in *P. savastanoi* with that in other systems associated with neoplastic growth in plants. Evidence from previous investigations on crown gall tumor tissue suggested that T-DNA, which carries genetic determinants encoding phytohormone synthesis, confers a tumorigenic state when integrated into host tissue (13). Subsequently, Schroeder *et al.* (14) and Thomashow *et al.* (15) demonstrated that the *tms-2* locus of crown gall T-DNA encodes an enzyme possessing indoleacetamide hydrolase activity and further proposed that the *tms-1* locus encodes an enzyme catalyzing the conversion of trytophan to indoleacetamide. However, tryptophan monooxygenase activity has yet to be demonstrated in crown gall tissue or in *A. tumefaciens*. In recent hybridization experiments under low-stringency conditions, homology was demonstrated between DNA sequences bearing the *tms-1* locus of T-DNA and *iaaM* and between *tms-2* and *iaaH* (unpublished results).

In this study, we present the nucleotide sequences of *iaaM* and *iaaH* from *P. savastanoi*; we compare these sequences with the sequences reported by Klee *et al.* (16) and Gielen *et al.* (17) for the *tms-1* and *tms-2* loci in crown gall T-DNA. On the basis of the deduced amino acid sequences, we find significant homology between the *iaaM* and *tms-1* gene products, and lesser, yet significant, homology between the *tms-2* and *iaaH* gene products.

MATERIALS AND METHODS

Bacterial Strains and Plasmids. *Escherichia coli* SK-1592 (pLUC2) (4), *E. coli* HB101 (pCP3) (to be described elsewhere) were used for plasmid isolations. The restriction maps of the cloned fragments of *P. savastanoi* DNA contained in the plasmid pLUC2 and pCP3 are shown in Fig. 1. *E. coli* 7118 was used for nucleotide sequencing experiments (18). Bacteria were grown in LB medium (19).

Plasmid DNA from *E. coli* was isolated by the procedure of Froman as modified by Tait *et al.* (20) and purified by cesium chloride density-gradient centrifugation (21).

Materials. Restriction endonucleases were purchased from New England Biolabs, Bethesda Research Laboratories, or Boehringer Mannheim; DNA polymerase I large fragment, pentadecameric primer, isopropyl β-D-thiogalactoside (IPTG), 5-bromo-4-chloro-3-indolyl β-D-galactoside (X-Gal), and polyacrylamide, from Bethesda Research Laboratories; [α-^{32}P]dATP, [^{35}S]methionine, and coupled transcription/translation system, from Amersham; exonuclease III and nuclease S1, from Boehringer Mannheim; ultrapure urea, from Research Organics, Cleveland, OH; and Kodak x-ray film XAR-5, from Merry X-ray Chemical.

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: IAA, indoleacetic acid; Ti plasmid, tumor-inducing plasmid; T-DNA, DNA transferred from the Ti plasmid to a plant cell.

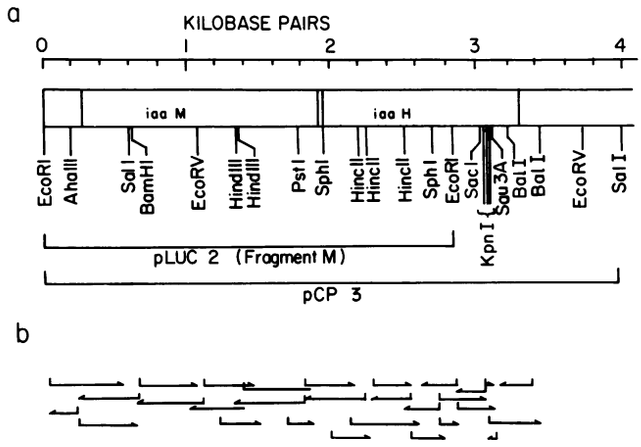

FIG. 1. (*a*) Restriction maps of fragment M (pLUC2) (4) and pCP3. (*b*) Strategy used for nucleotide sequencing of the promoter region, *iaaM*, and *iaaH*. Arrows show the extent and direction of sequence analysis. Both strands have been sequenced.

Nucleic Acid Sequence Determination. The recombinant plasmids pLUC2 and pCP3, constructed from pBR328 and pIAA fragments bearing *iaaM* and *iaaH* (4), provided the DNA to be sequenced (Fig. 1). The plasmids were digested with *Aha* III, *Bal* I, *Bam*HI, *Eco*RI, *Eco*RV, *Hin*cII, *Hin*dIII, *Kpn* I, *Pst* I, *Sac* I, *Sau*3AI, *Sal* I, *Sph* I, and the resulting restriction fragments were cloned in the M13 vectors mp8, mp11, mp18, or mp19 (22). Fragments treated with exonuclease III and nuclease S1 were also cloned in the M13 vectors (23). *E. coli* 7118 transformants (white plaques) were screened by the method of Messing and Vieira (24). Exonuclease III- and nuclease S1-derived clones were screened by plaque-hybridization using fragment M (Fig. 1*a*) as a probe. Phage clones of both strands of the fragments bearing *iaaM* and *iaaH* were isolated for sequencing. A single-stranded phage template was prepared as described by Messing *et al.* (22) and used for the dideoxy sequencing reactions as described by Sanger *et al.* (25). Electrophoresis for nucleotide sequencing was carried out in 8% polyacrylamide (BRL model S0 apparatus; gel 34 × 40 cm, 0.4 mm thick) or 6% polyacrylamide (BRL model S1 apparatus; gel 30 × 84 cm, 0.4 mm thick). Computer analyses of the nucleotide sequences were done using programs kindly provided by R. Larson and J. Messing, University of Minnesota.

***In Vitro* Protein Synthesis and Enzyme Activities.** The proteins encoded by pCP3 were labeled with [^{35}S]methionine by the procedures in the coupled transcription/translation system of Chen and Zubay (26). Proteins were subjected to $NaDodSO_4$/PAGE in a modified Laemmli system (27). Tryptophan monooxygenase and indoleacetamide hydrolase activities were assayed as described (3).

RESULTS

DNA Sequence Determination. The nucleotide sequences of *iaaM* and *iaaH* are shown in Fig. 2. The sequence (5′ to 3′) from the *Eco*RI site of fragment M is presented for only one strand. The open reading frame of *iaaM* encodes a 557 amino acid protein with a molecular weight of 61,783, which corresponds closely to the molecular weight of 62,000 determined for the purified tryptophan monooxygenase monomer by $NaDodSO_4$/PAGE (29). The first 10 amino acids of the coding region correspond exactly to the amino-terminal amino acid sequence determined for the purified protein (29). The nucleotide sequence of the *tms-1* region of the octopine-type plasmid pTiA6NC from *A. tumefaciens* shows an open reading frame that encodes a 755-residue protein of molecular weight 83,769 (16). As shown in Fig. 3*a*, the deduced amino acid sequences of the *iaaM* and *tms-1* products show strong homology throughout the entire length of the region coding for tryptophan monooxygenase, provided that tyrosine at position 90 and proline, methionine, and threonine at positions 494–496 in the *iaaM* sequence are skipped. Overall, the deduced amino acid sequences of tryptophan monooxygenase and the *tms-1*-encoded protein show 50% perfect matches.

The open reading frame of *iaaH* encodes a protein of 455 amino acids, with a molecular weight of 48,515. As shown in Fig. 3*b*, the deduced amino acid sequences of the *iaaH* and *tms-2* products show lesser homology (27% for perfect matches, if threonine at position 251 in the *tms-2* sequence is skipped) than the *iaaM* and *tms-1* products. Strongest homology occurs in the sequences of the core of each protein. Comparison of the nucleotide sequences shows there is 54% homology between *iaaM* and *tms-1* and 38% homology between *iaaH* and *tms-2*.

Predicted Sequence of the FAD-Binding Region. Previous studies showed that tryptophan monooxygenase possesses FAD as a cofactor (29). Further, Klee *et al.* (16) detected homology between amino acids 239–263 of the predicted *tms-1* product and amino acids 5–29 in the FAD-linked hydroxybenzoate hydroxylase from *P. fluorescens*: the latter sequence is rich in hydrophobic amino acids and has been shown by x-ray crystallography to comprise the pocket in the hydroxylase protein that binds the adenine moiety of FAD (30). The same deduced sequence of the *tms-1* product shows strong homology with amino acid residues 42–66 of tryptophan monooxygenase (Fig. 4). We suggest that residues 42–66 comprise the FAD-binding site in tryptophan monooxygenase.

***In Vitro* Protein Synthesis and Enzyme Activities.** To verify that the cloned sequences encode intact *iaaM* and *iaaH* products, we determined protein synthesis in a DNA-directed transcription/translation system. pCP3 encoded two proteins, of M_r 62,000 and 47,000 (Fig. 5). In the control reaction mixture with pCP3ΔR1, which has the 2.8-kilobase-pair fragment M deleted (Fig. 1), no radioactive protein of M_r > 30,000 was found. Fragment M in pLUC2 encodes proteins of M_r 62,000 (tryptophan monooxygenase) and 39,000 (which appears to be a truncated indoleacetamide hydrolase protein). The two proteins encoded by pCP3 correspond in size to those predicted by the open reading frames in the nucleotide sequence; they were identified as tryptophan monooxygenase and indoleacetamide hydrolase, since activities of both enzymes were detected in cell-free preparations of *E. coli* HB101 transformed with pCP3. IAA accumulated to levels of 58 μg/mg (dry weight of cells) in culture filtrates in which *E. coli* HB101 transformed with pCP3 had been grown overnight. Neither the above enzyme activities nor IAA were detected in cell-free preparations of *E. coli* HB101 cells transformed by the vector plasmid pBR328 alone.

Ribosome Binding Sites of *iaaM*. The nucleotide sequences upstream from the initiation codons (ATG) of the *iaaM* and *iaaH* coding regions are shown in Fig. 2. There are two possible ribosome binding sites of *iaaM*, AAGAG and AGAG, which are similar to sequences near the 3′ end of the 16S rRNA of *P. aeruginosa* (31). The putative ribosome binding site of *iaaH*, AAGAG, is also shown (Fig. 2). All three of these sites show a high degree of homology with ribosome binding sites of other genes characterized from Pseudomonads (32, 33). However, we find no consensus sequence characteristic of the Pribnow box or a −35 region (34).

DISCUSSION

Production of IAA confers virulence in *P. savastanoi* for its hosts, oleander and olive. Knowledge of the mechanisms regulating production of IAA therefore is important for the understanding of mechanisms controlling expression of vir-

FIG. 2. Nucleotide sequences of the promoter region, *iaaM*, and *iaaH* and amino acid sequences predicted by the *iaaM* (*Left*) and *iaaH* (*Right*) open reading frames. Numbers above each line refer to amino acid positions in the predicted *iaaM* and *iaaH* products; numbers below each line refer to nucleotide positions starting from the *Eco*RI site of fragment M at the 5′ end. Sequence is shown from 5′ to 3′. The boxed sequences are the proposed Shine–Dalgarno (ribosome-binding) regions (28).

ulence. Previous investigations showed that insertions in *iaaM* had polar effects on *iaaH* expression. This suggested that expression of *iaaH* depends upon the promoter for *iaaM* and that the two determinants occur in an operon. To further elucidate their organization in an operon, we determined the fine structure of the two genes and confirmed their expression in *E. coli*.

The sequence analysis of *iaaM* shows an open reading

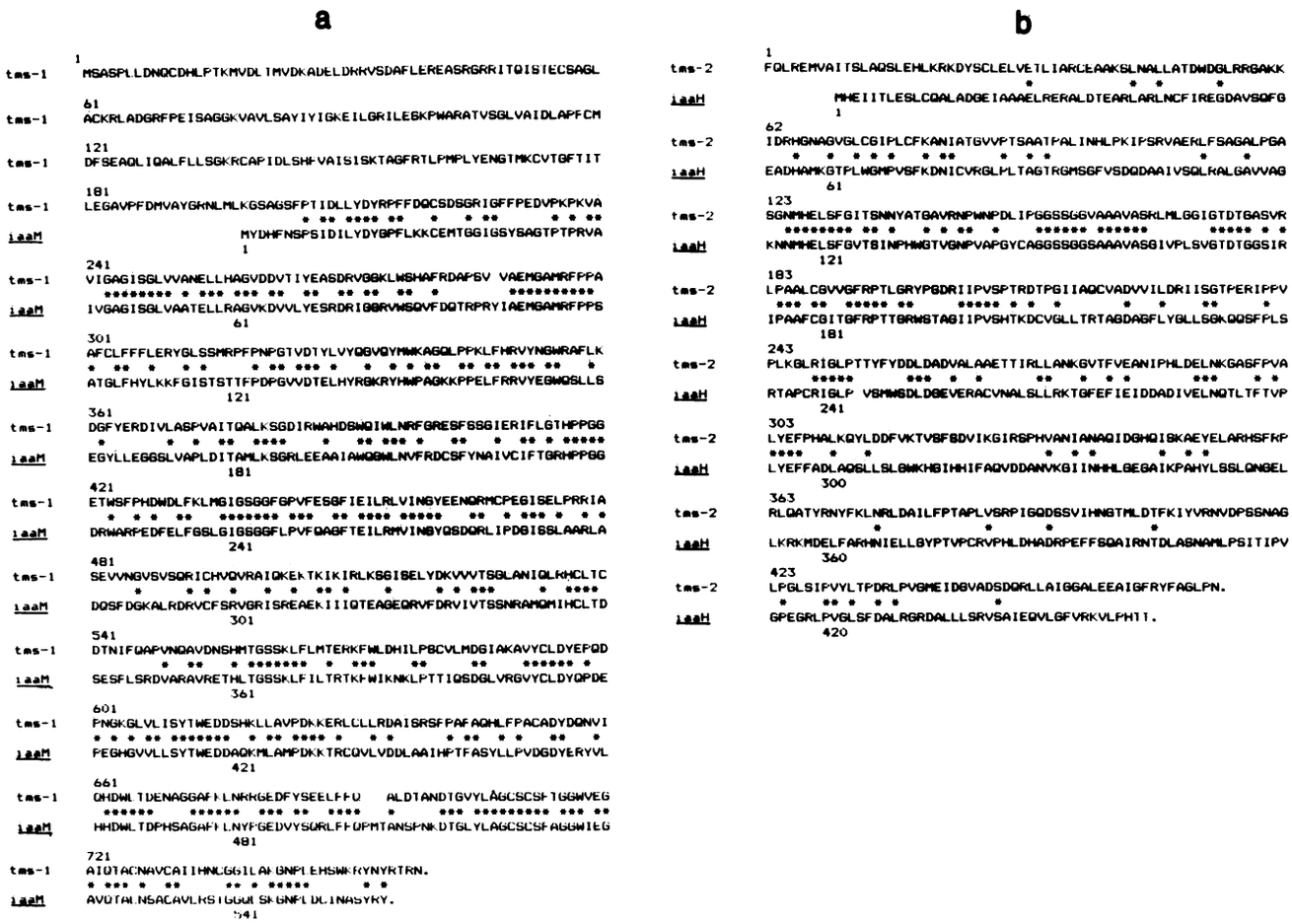

FIG. 3. Deduced amino acid sequences of the *iaaM* and *tms-1* products (*a*) and of the *iaaH* and *tms-2* products (*b*). One-letter amino acid abbreviations are used. Asterisks indicate matching amino acids. (*a*) Numbers above each line refer to the position of amino acids in the *tms-1* sequence; those below refer to *iaaM*. To obtain maximum homology, gaps have been inserted in the *tms-1*-encoded sequence at positions corresponding to positions 90 and 494–496 in the *iaaM* product. (*b*) Numbers above each line refer to the position of amino acids in the *tms-2* product; those below refer to *iaaH*. A gap has been inserted in the *iaaH*-predicted sequence (corresponding to position 251 in the *tms-2*-predicted sequence). The *tms-1* and *tms-2* sequences shown are those determined by Klee *et al.* (16) for the octopine-type plasmid pTiA6NC from *A. tumefaciens*; identical sequences were reported by Gielen *et al.* (17) for the comparable regions (transcript 1 and transcript 2) of the T-DNA of the octopine-type plasmid pTiAch5 except that N, R, and P are at positions 718, 719, and 721 in the transcript 1-encoded sequence.

frame sufficient to encode a protein of M_r 61,783, which is consistent with the apparent M_r of 62,000 estimated for the monomer of purified tryptophan monooxygenase (29). Similarly, the predicted amino-terminal amino acid sequence (Met-Tyr-Asp-His-Phe-Asn-Ser-Pro-Ser-Ile-Asp-) is in perfect agreement with that determined by chemical analysis of the purified protein. As shown by the coupled *in vitro* transcription/translation system, determinants borne on the DNA fragments sequenced in this study encoded two proteins of molecular weights (62,000 and 47,000) that correspond closely to those determined for the products of *iaaM* and *iaaH* from the deduced amino acid sequences. That the encoded proteins are tryptophan monooxygenase and indoleacetamide hydrolase was demonstrated by detection of their activities in extracts of *E. coli* transformed with pCP3.

Similarities between the *P. savastanoi* and crown gall systems for IAA synthesis are further evident in the nucleotide sequence and in the deduced amino acid sequences of the *iaa* operon and the *tms* locus of *A. tumefaciens* T-DNA. The strong homology observed in the apparent FAD-binding domains suggests that the *tms-1* product is functionally very similar to the tryptophan monooxygenase from *P. savastanoi*. No similarities are seen in the possible regulatory sequences of the genes from the two sources. Moreover, *iaaM* and *iaaH* are organized in an operon in *P. savastanoi* (35), whereas the comparable genes in T-DNA are monocistronic (16, 17), as might be predicted, since they function in the plant cell.

Although *tms-2* and *iaaH* are similar in size, the open reading frame of *tms-1* is substantially larger (594 base pairs)

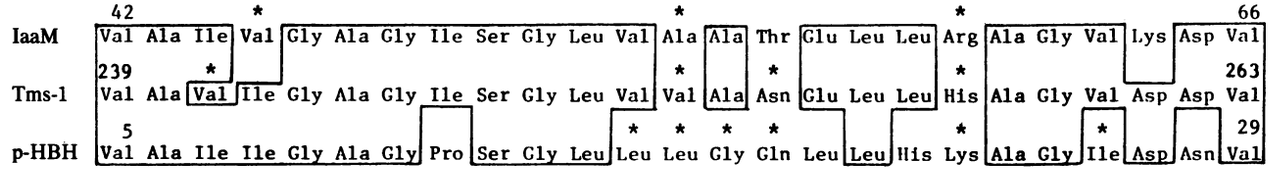

FIG. 4. Amino acid sequence homology between the predicted *iaaM* product (IaaM), the predicted *tms-1* product (Tms-1), and *p*-hydroxybenzoate hydroxylase (p-HBH) at the nucleotide-binding site. The conserved amino acids are boxed. Functionally similar amino acids are indicated by asterisks.

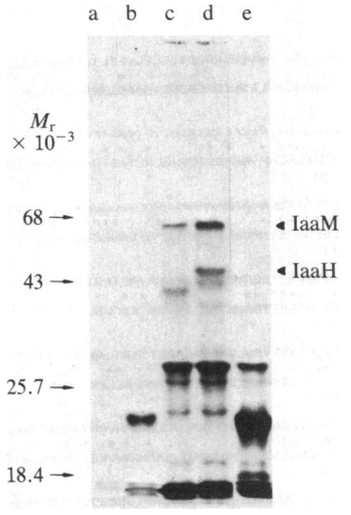

FIG. 5. Autoradiograph obtained after $NaDodSO_4$/PAGE of [^{35}S]methionine-labeled proteins encoded by various plasmids using the coupled transcription/translation system. Lanes: a, no DNA; b, pBR328; c, pLUC 2; d, pCP3; e, pCP3ΔR1 (pCP3 with fragment M deleted). Positions of *iaaM* and *iaaH* gene products (IaaM and IaaH) are indicated at right. Positions and molecular weights of concurrently electrophoresed standards (bovine serum albumin, ovalbumin, α-chymotrypsinogen, and β-lactoglobulin) are at left.

than the open reading frame of *iaaM*. The significance of the size differences is unknown, but the additional polypeptide in the *tms-1* product may reflect the structural requirements for its function in the plant cell.

Since tryptophan monooxygenase exhibits broad substrate specificity for methylated and halogenated tryptophan derivatives (4, 29), IAA^+ strains of *P. savastanoi* are resistant to tryptophan analogues such as 5-methyltryptophan. IAA^- mutants, which lack tryptophan monooxygenase, are sensitive to 5-methyltryptophan (1). The same phenotype is exhibited by T-DNA-transformed cultured tobacco cells, which are more resistant to 5-methyltryptophan than are nontransformed cells (36). These results suggest that T-DNA-transformed cells contain tryptophan monooxygenase, which provides detoxifying activity toward 5-methyltryptophan. This further argues for the similarity between *P. savastanoi* and *A. tumefaciens* in IAA synthesis and suggests there is a common origin for the genes encoding the synthetic pathway for IAA in the two systems.

We thank Gene Nester, Frank White, and Martin Yanofsky (Univ. of Washington) for valuable discussions during the course of this work. This material is based upon work supported by Grants PCM-8011794 and DMB-831872 from the National Science Foundation.

1. Smidt, M. & Kosuge, T. (1978) *Physiol. Plant Pathol.* **13,** 203–214.
2. Kosuge, T., Heskett, M. G. & Wilson, E. E. (1966) *J. Biol. Chem.* **241,** 3738–3744.
3. Comai, L. & Kosuge, T. (1980) *J. Bacteriol.* **143,** 950–957.
4. Comai, L. & Kosuge, T. (1982) *J. Bacteriol.* **149,** 40–46.
5. Kosuge, T. & Comai, L. (1982) *Plant Infection: The Physiological and Biochemical Basis*, ed. Asada, Y. (Springer, Berlin), pp. 175–186.
6. Chilton, M.-D., Drummond, M. H., Merlo, D. J., Sciaky, D., Montoya, A. L., Gordon, M. P. & Nester, E. W. (1977) *Cell* **11,** 263–271.
7. Chilton, M.-D., Saiki, R. K., Yadav, N., Gordon, M. P. & Quetier, F. (1980) *Proc. Natl. Acad. Sci. USA* **77,** 4060–4064.
8. Lemmers, M., De Beuckeleer, M., Holsters, M., Zambryski, P., Hernalsteens, J. P., Van Montagu, M. & Schell, J. (1980) *J. Mol. Biol.* **144,** 353–376.
9. Thomashow, M. F., Nutter, R., Montoya, A. L., Gordon, M. P. & Nester, E. W. (1980) *Cell* **19,** 729–739.
10. Thomashow, M. F., Nutter, R., Postle, K., Chilton, M.-D., Blattner, F. R., Powell, A., Gordon, M. P. & Nester, E. W. (1980) *Proc. Natl. Acad. Sci. USA* **77,** 6448–6452.
11. Willmitzer, L., De Beuckeleer, M., Lemmers, M., Van Montagu, M. & Schell, J. (1980) *Nature (London)* **287,** 359–361.
12. Yadav, N. S., Postle, K., Saiki, R. K., Thomashow, M. F. & Chilton, M.-D. (1980) *Nature (London)* **287,** 458–461.
13. Akiyoshi, D. E., Morris, R. O., Hinz, R., Mischke, B. S., Kosuge, T., Garfinkel, D. J., Gordon, M. P. & Nester, E. W. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 407–411.
14. Schroeder, G., Waffenschmidt, S., Weiler, E. W. & Schroeder, J. (1984) *Eur. J. Biochem.* **138,** 387–391.
15. Thomashow, L. S., Reeves, S. & Thomashow, M. F. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 5071–5075.
16. Klee, H., Montoya, A., Horodyski, F., Lichtenstein, C., Garfinkel, D., Fuller, S., Flores, C., Peschon, J., Nester, E. & Gordon, M. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 1728–1732.
17. Gielen, J., De Beuckeleer, M., Seurinck, J., Beboeck, F., De Greve, H., Lemmers, M., Van Montagu, M. & Schell, J. (1984) *EMBO J.* **3,** 835–846.
18. Vieira, J. & Messing, J. (1982) *Gene* **19,** 259–268.
19. Miller, J. H. (1972) *Experiments in Molecular Genetics* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), pp. 431–435.
20. Tait, R. C., Lundquist, R. C. & Kado, C. I. (1982) *Mol. Gen. Genet.* **186,** 10–15.
21. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) in *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), p. 93.
22. Messing, J. (1983) *Methods Enzymol.* **101,** 20–78.
23. Henikoff, S. (1984) *Gene* **28,** 351–359.
24. Messing, J. & Vieira, J. (1982) *Gene* **19,** 269–276.
25. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
26. Chen, H. & Zubay, G. (1983) *Methods Enzymol.* **101,** 674–690.
27. Laemmli, U. K. (1970) *Nature (London)* **227,** 680–685.
28. Shine, J. & Dalgarno, L. (1975) *Nature (London)* **254,** 34–38.
29. Hutcheson, S. & Kosuge, T. (1985) *J. Biol. Chem.* **260,** 6281–6285.
30. Wierenga, R. K., Jong, R. J. D., Kalk, K. H., Hol, W. G. J. & Drenth, J. (1979) *J. Mol. Biol.* **131,** 55–73.
31. Bassel, B. A. (1979) *Nucleic Acids Res.* **6,** 2003–2016.
32. Gray, G. L., Smith, D. H., Baldridge, J. S., Harkins, R. N., Vasil, M. L., Chen, E. Y. & Heyheker, H. L. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 2645–2649.
33. Inouye, S., Elina, Y., Nakazawa, A. & Nakazawa, T. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 1688–1691.
34. Rosenberg, M. & Court, D. (1979) *Annu. Rev. Genet.* **13,** 319–353.
35. Comai, L. & Kosuge, T. (1983) *J. Bacteriol.* **154,** 1162–1167.
36. Sanger, M. & Kosuge, T. (1984) *Plant Physiol.* Suppl. **75,** 42.