assist in protecting plants against fungal diseases.

REFERENCES AND NOTES

1. P. E. Kolattukudy, *Annu. Rev. Phytopathol.* **23**, 223 (1985).
2. ———, in *The Biochemistry of Plants*, P. K. Stumpf and E. E. Conn, Eds. (Academic Press, New York, 1980), vol. 4, p. 571; *Annu. Rev. Plant Physiol.* **32**, 539 (1981); P. J. Holloway, in *The Plant Cuticle*, D. F. Cutler, K. L. Alvin, C. E. Price, Eds. (Academic Press, London, 1982), p. 45.
3. C. P. Woloshuk and P. E. Kolattukudy, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 1704 (1986).
4. C. X. Huang *et al.*, unpublished results.
5. P. E. Kolattukudy *et al.*, *Arch. Biochem. Biophys.* **256**, 446 (1987).
6. M. L. Webb, K. A. Maguire, S. T. Jacob, *Nucleic Acids Res.* **15**, 8547 (1987).
7. M. L. Webb and S. T. Jacob, *J. Biol. Chem.* **263**, 4745 (1988).
8. K. R. Yamamoto, *Annu. Rev. Genet.* **19**, 209 (1985).
9. B. Corthésy, R. Hipskind, I. Theulaz, W. Wahli, *Science* **239**, 1137 (1988).
10. T. S. Lin and P. E. Kolattukudy, *J. Bacteriol.* **133**, 942 (1978).
11. Nuclei from *F. solani pisi* were isolated as described [W. E. Timberlake, in *Isolation of Stage- and Cell-Specific Genes from Fungi* (North Atlantic Treaty Organization Advanced Study Institutes Series), J. Bailey, Ed. (Springer-Verlag, Heidelberg, 1986), vol. H1, pp. 343–357] with some modifications. Briefly, mycelia collected by filtration were washed several times with distilled water. For each gram of tissue, 2 ml of SSE.5 nuclei isolation buffer [100 m*M* KCl, 10 m*M* tris-Cl (*p*H 7.0), 10 m*M* EDTA, 4 m*M* spermidine, 1 m*M* spermine, 1 m*M* phenylmethylsulfonyl fluoride, 0.1% β-mercaptoethanol, and 17% sucrose (w/v)] was added, allowed to soak, and ground in liquid nitrogen; then it was ground in an Omni-Mixer (Sorvall) to a fine powder. Upon thawing, the powder was resuspended by short bursts in a Polytron (Brinkmann). The homogenate was filtered gently through two layers of sterile Mira-Cloth (Calbiochem) and centrifuged at 4000*g* for 20 min at 4°C. The pellet was collected, resuspended in SSE.5 with a Dounce homogenizer, centrifuged, and suspended in 3 ml of SSE.5. The suspension was layered over 7 ml of SSE 2.1, which is similar to SSE.5 except that the sucrose concentration is increased to 72% (w/v). After centrifugation for 1 hour at 55,000*g* in a Beckman Ti 50 rotor at 4°C, the pellet containing nuclei was resuspended in storage buffer [40% glycerol, 50 m*M* tris-Cl (*p*H 8.3), 5 m*M* $MgCl_2$, and 0.1 m*M* EDTA]. The purity of nuclei was checked by microscopy.
12. G. S. McKnight and R. D. Palmitter, *J. Biol. Chem.* **254**, 9050 (1979).
13. C. L. Soliday, W. H. Flurkey, T. W. Okita, P. E. Kolattukudy, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 3939 (1984).
14. R. M. Fourney, J. Miyakoshi, R. S. Day III, M. C. Paterson, *Focus (Rochester, NY)* **10** (no. 1), 4 (1988).

15. Supported in part by grant DMB-8306835 from the National Science Foundation.

20 June 1988; accepted 16 September 1988

A Chemically Synthesized *Antennapedia* Homeo Domain Binds to a Specific DNA Sequence

HISAKAZU MIHARA* AND EMIL THOMAS KAISER

A peptide 60 residues in length that corresponds to the homeo domain of *Antennapedia* (*Antp*), a protein governing development in *Drosophila*, was synthesized by segment condensation with protected peptide segments prepared on an oxime resin. A footprinting assay showed that the homeo domain binds specifically to a TAA repeat DNA sequence in the *Antp* gene. Thus the *Antp* homeo domain has a sequence-specific DNA binding property. The circular dichroism spectra of the homeo domain peptide showed the presence of a significant amount of α-helical structure in aqueous solution and in 50 percent trifluoroethanol. The α helicity measured in water appears to depend on the peptide concentration, which suggests that the peptide aggregates. These results support the hypothesis that the homeo domain binds to DNA through a helix-turn-helix motif.

THE DEVELOPMENT OF *DROSOPHILA* is governed by the segmentation and homeotic genes, of which many contain a highly conserved sequence, the "homeo box" (*1–3*). The homeo box encodes the 60–amino-acid sequence of the homeo domain in the products of many developmental genes. Genetic studies indicate that the developmental genes encode regulatory molecules that could affect the expression of other developmental genes through binding to specific DNA sequences (*4–10*). Sequence comparison with prokaryotic gene regulatory proteins and yeast mating-type proteins suggests that the homeo domain adopts a helix-turn-helix motif in binding to DNA (*1, 4, 11–13*).

We report the preparation of a 60–amino-acid peptide **1** corresponding to the *Antp* homeo domain by a segment-synthesis–condensation procedure that made use of an oxime resin (Fig. 1) (*14*). Because the protected peptide segments we used were purified before segment coupling, and because coupled products are readily separated from uncoupled segments, this method produces higher purity proteins through less purification than the classical solid-phase method (*15*). Another advantage of the segment-synthesis–condensation approach is that mutants can be readily prepared by the replacement of segments in a cassette mode. The purified form of the homeotic proteins is needed to examine correct interactions

Laboratory of Bioorganic Chemistry and Biochemistry, Rockefeller University, New York, NY 10021.

*Present address: Department of Industrial Chemistry, Kyushu Institute of Technology, Kitakyushu, Japan.

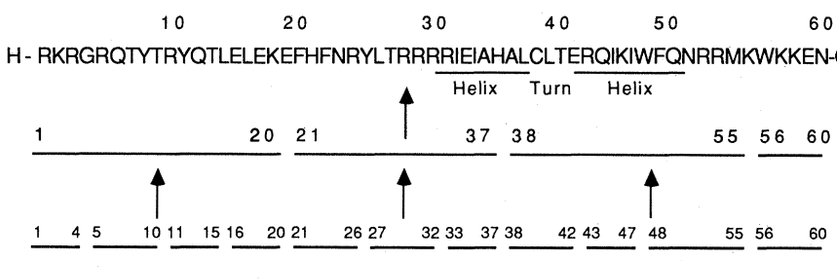

Fig. 1. The amino acid sequence of the *Antp* homeo domain and the strategy of the synthesis of peptide **1** by segment condensation are shown. The putative helix-turn-helix portion is indicated (*11*). The side chain–protected amino acids used were Asp(OBzl), Glu(OBzl), Arg(Tos), Cys(MeBzl), His(Bom), Lys(Cl-Z), Ser(Bzl), Thr(Bzl), Trp(HCO), and Tyr(Cl_2Bzl) (Bom, π-benzyloxymethyl; Bzl, benzyl; OBzl, benzyl-protected COOH; Cl-Z, 2-chlorobenzyloxycarbonyl; Cl_2Bzl, 2,6-dichlorobenzyl; HCO, formyl; MeBzl, 4-methylbenzyl; and Tos, toluenesulfonyl). Eleven small fragments (4 to 7 amino acids) were prepared stepwise on the oxime resin (*14*) and then purified with RP-HPLC. The larger sized (17 to 20 amino acids) protected peptides were synthesized by coupling of the small segments with dicyclohexylcarbodiimide-hydroxybenzotriazole (HOBt) on the oxime resin. The peptides were removed from the resin and then purified by passage through a column of Sephadex LH 60 (dimethylformamide). The large segments were coupled starting from the COOH-terminal pentapeptide benzylester to each other with (*N*-ethyl-*N*′-dimethylaminopropyl) carbodiimide-HOBt in solution and were purified with Sephadex LH 60 to give the protected homeo domain. The protected protein was deprotected by the low- and high-HF method (*23*) and purified by gel-filtration and RP-HPLC (TSK Phenyl 5PW · RP, Toyo Soda) to give a single peak of protein as observed on the same RP-HPLC column and a single band corresponding 8000 daltons on SDS-PAGE. All intermediates were characterized by amino acid analysis and mass spectroscopy.

between the proteins and DNA. Synthesis of peptide **1** has allowed us to establish that binding occurs to the upstream region of the *Antp* gene, and to provide information on the conformation of the domain. The synthetic *Antp* homeo domain peptide was characterized by reversed-phase high-performance liquid chromatography (RP-HPLC), sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE), amino acid analysis, automated sequencing by the Edman procedure, and ^{252}Cf fission-fragment mass spectroscopy (*16*).

The specific DNA binding ability of the *Antp* homeo domain was determined by the assay of its protection of the 5′ upstream region (472 bp in length and −6 kbp upstream from the *Antp* P_1 transcription start site) of the *Antp* gene (*17*) from digestion by deoxyribonuclease I (DNase I) (*18*) (Fig. 2). Peptide **1** protected the DNA from DNase I digestion at a TAA repeat sequence (Fig. 2A) and at the consecutive TAA-rich region (Fig. 2B). This result demonstrates that the homeo domain itself can bind to a specific DNA sequence and that it recognizes a TAA repeat sequence. The *Ultrabithorax* (*Ubx*) protein and the fusion protein containing the *fushi tarazu* (*ftz*) homeo domain bind to the same TAA repeat region (*17, 19*). This coincidence is quite reasonable, because the *Antp*, *ftz*, and *Ubx* homeo domains have the same amino acid sequence in the putative recognition helix (residues 42 to 50) according to the helix-turn-helix model for DNA binding (*11*), which suggests that this model is correct. The fusion protein containing the *engrailed* (*en*) homeo domain binds to the same TAA tandem repeat sequence (*19*). The *en* homeo domain varies at one position in the recognition helix compared with *Antp* and *ftz*. These observations are consistent with the proposal that the homeo domain–containing proteins recognize and bind to similar DNA sequences because the amino acid sequences of their homeo domains are closely related; that is, the competition of the homeo domain–containing proteins for binding to similar DNA sequences may play a role in the mechanism for cross-regulatory interactions among the developmental genes (*1, 4, 7, 19, 20*). However, the *en* and *even-skipped* (*eve*) proteins and the *ftz* fusion protein also recognize the different consensus sequence TCAATTAAAT (*19, 20*), suggesting that the homeo domains can recognize one or more types of DNA sequence and that the regulatory mechanism may be more complex than the simple competition model.

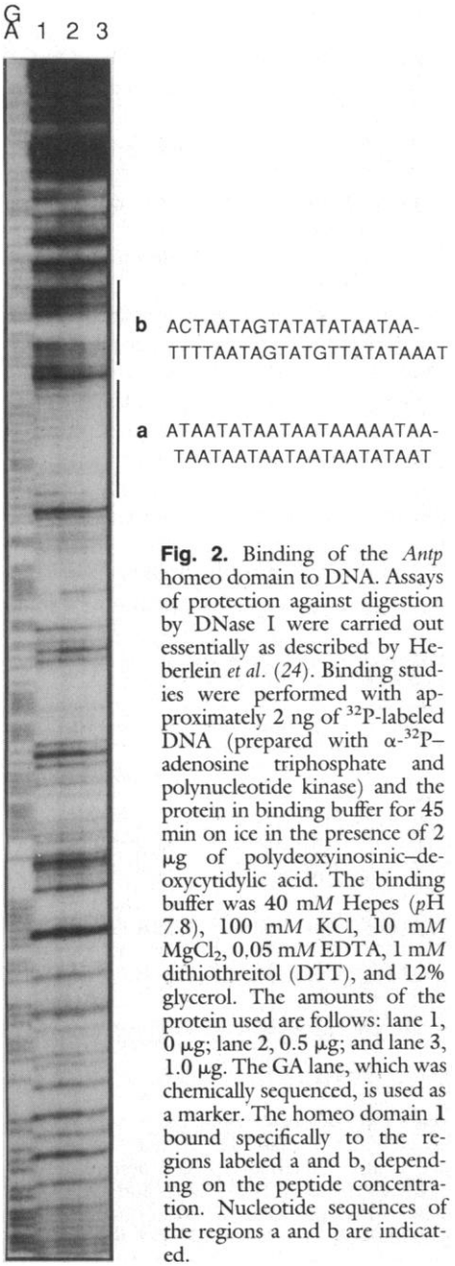

Fig. 2. Binding of the *Antp* homeo domain to DNA. Assays of protection against digestion by DNase I were carried out essentially as described by Heberlein *et al.* (*24*). Binding studies were performed with approximately 2 ng of ^{32}P-labeled DNA (prepared with α-^{32}P–adenosine triphosphate and polynucleotide kinase) and the protein in binding buffer for 45 min on ice in the presence of 2 μg of polydeoxyinosinic–deoxycytidylic acid. The binding buffer was 40 m*M* Hepes (*p*H 7.8), 100 m*M* KCl, 10 m*M* $MgCl_2$, 0.05 m*M* EDTA, 1 m*M* dithiothreitol (DTT), and 12% glycerol. The amounts of the protein used are follows: lane 1, 0 μg; lane 2, 0.5 μg; and lane 3, 1.0 μg. The GA lane, which was chemically sequenced, is used as a marker. The homeo domain **1** bound specifically to the regions labeled a and b, depending on the peptide concentration. Nucleotide sequences of the regions a and b are indicated.

The circular dichroism (CD) spectra of the *Antp* homeo domain were measured in aqueous buffer at two different concentrations of **1** and in 50% trifluoroethanol (TFE) (Fig. 3). The *Antp* homeo domain showed typical α-helical CD patterns in both aqueous buffer and 50% TFE, indicating that a significant amount of α-helical structure was adopted by **1**. The amounts of secondary structure calculated from the CD data are α helix 36%, β sheet 18%, β turn 14%, and a remainder of 32% in aqueous buffer (peptide **1** concentration = $4.7 \times 10^{-6}M$) and α helix 78%, β sheet 0%, β turn 7%, and a remainder of 15% in 50% TFE (*21*). The α helicity in aqueous buffer is comparable to the expected value (30%) for two helices in the putative helix-turn structure in peptide **1**. Also, the estimates of the β turn content in buffer and 50% TFE are consistent with the expected values for two (14%) and one β turns, respectively, in peptide **1**. The amount of α helix increases from 36 to 78% from aqueous buffer to TFE, whereas the amount of β-sheet structure and β turn decreases. The curve at a lower concentration of **1** ($1.3 \times 10^{-6}M$) in aqueous medium gives a lower estimate of the helical content (15%). The concentration dependence of the α helicity indicates that the domain aggregates in aqueous solution. These results are consistent with the proposal that the homeo domain can adopt the helix-turn-helix motif and that it may bind to DNA as an aggregate (*11, 12, 22*). The CD data for **1** provide direct information about the conformation of the homeo domain.

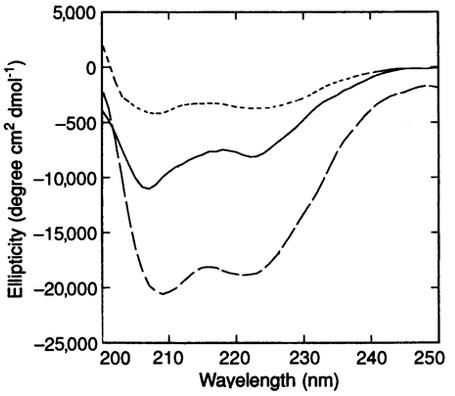

Fig. 3. CD spectra of the *Antp* homeo domain; $4.7 \times 10^{-6}M$ **1** in 100 m*M* KCl–20 m*M* phosphate, *p*H 7.5 (———); $1.3 \times 10^{-6}M$ in the same buffer (--------); $5.7 \times 10^{-6}M$ in 50% TFE (– – –). CD measurements (mean residue molar ellipticity) were taken on an Aviv (Philadelphia) modification of a Cary model 60ds spectropolarimeter.

Synthesis of the *Antp* homeo domain has provided evidence in support of the proposal that the homeo domain is a sequence-specific DNA binding region in the homeo domain–containing protein. A careful comparison of the DNA binding specificity between the *Antp* homeo domain and the purified, intact *Antp* protein needs to be performed. The *Antp* homeo domain (**1**) recognized a TAA repeat sequence in the 5′ region upstream from the P_1 transcription start site in the *Antp* gene itself. This result is consistent with the genetic evidence that expression of the *Antp* gene is regulated by its own gene product (*10*).

The present work demonstrates the usefulness of the oxime resin method (*14*) for the preparation of pure proteins. The advantages of this method are that it avoids the microheterogeneity of the proteins synthesized by the stepwise solid-phase method (*15*) and that it should simplify the synthesis of mutants of proteins, including ones in which there are unnatural amino acids and nonpeptidic regions. A chemical approach for synthesizing small proteins and domains

with modified structures would be a useful tool for clarifying protein-substrate interactions. Furthermore, scaling up the synthesis of small proteins by our procedure is feasible because the protected segments needed can be rapidly prepared on the oxime resin.

REFERENCES AND NOTES

1. W. J. Ghering, *Science* **236**, 1245 (1987); _____ and Y. Hiromi, *Annu. Rev. Genet.* **20**, 147 (1986).
2. W. McGinnis, M. S. Levine, E. Hafen, A. Kuroiwa, W. J. Gehring, *Nature* **308**, 428 (1984).
3. M. P. Scott and A. J. Weiner, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 4115 (1984).
4. G. Desplan, J. Theis, P. H. O'Farrell, *Nature* **318**, 630 (1985).
5. R. A. H. White and M. Wilcox, *Cell* **39**, 163 (1984).
6. E. Hafen, A. Kuroiwa, W. J. Gehring, *ibid.* **37**, 833 (1984).
7. K. Harding, C. Wedeen, W. McGinnis, M. Levine, *Science* **229**, 1236 (1985).
8. S. B. Carroll and M. P. Scott, *Cell* **43**, 47 (1985).
9. P. W. Ingham and A. Martinez-Arias, *Nature* **324**, 592 (1986).
10. S. B. Carroll *et al.*, *Cell* **47**, 113 (1986).
11. A. Laughon and M. P. Scott, *Nature* **310**, 25 (1984).
12. J. C. W. Shepherd, W. McGinnis, A. E. Carrasco, E. M. DeRobertis, W. J. Gehring, *ibid.*, p. 70.
13. C. O. Pabo and R. T. Sauer, *Annu. Rev. Biochem.* **53**, 293 (1984); M. Ptashne, in *A Genetic Switch: Gene Control and Phage Lambda* (Cell and Blackwell, Cambridge, 1987), chap. 2.
14. W. F. DeGrado and E. T. Kaiser, *J. Org. Chem.* **47**, 3258 (1982); S. H. Nakagawa and E. T. Kaiser, *ibid.* **48**, 678 (1984); S. H. Nakagawa, H. S. H. Lau, F. J. Kézdy, E. T. Kaiser, *J. Am. Chem. Soc.* **107**, 7087 (1985).
15. R. B. Merrifield, *J. Am. Chem. Soc.* **85**, 2149 (1963); G. Barany and R. B. Merrifield, in *The Peptides*, E. Gross and J. Meinhoper, Eds. (Academic Press, New York, 1980), vol. 2, pp. 1–284.
16. B. T. Chait, W. C. Agosta, F. H. Field, *Int. J. Mass Spectrom. Ion Phys.* **39**, 339 (1981).
17. P. A. Beachy, unpublished data.
18. D. Galas and A. Schmitz, *Nucleic Acids Res.* **5**, 3157 (1978).
19. C. Desplan, J. Theis, P. O'Farrell, *Cell*, in press.
20. T. Hoey and M. Levine, *Nature* **332**, 858 (1988).
21. S. W. Provencher and J. Glöckner, *Biochemistry* **20**, 33 (1981).
22. A. M. Miller, V. L. Mackay, K. A. Nasmyth, *Nature* **314**, 598 (1985).
23. J. P. Tam, W. F. Heath, R. B. Merrifield, *J. Am. Chem. Soc.* **105**, 6442 (1983).
24. U. Heberlein, B. England, R. Tjian, *Cell* **41**, 965 (1985).

25. We thank G. A. Laforet for her suggestions which aided us in the initiation of this work, P. Freimuth, D. A. Kendall, T. Sasaki, J. Taylor, and C. Desplan for helpful discussions, T. Hoey and M. S. Levine for teaching us the footprinting technique, and P. A. Beachy for providing us with the plasmid containing the 5′ upstream region of the *Antp* gene. All mass spectra were done by B. Chait at the Rockefeller University Biotechnology Mass Spectrometric Research Resource, which is supported by the Division of Research Resources, NIH grant RR 862. Peptide sequencing analysis was performed at the Yale University School of Medicine Sequencing Facility. Supported in part by a grant from Hoffmann–La Roche.

18 May 1988; accepted 7 September 1988

VirD2 Protein of *Agrobacterium tumefaciens* Very Tightly Linked to the 5′ End of T-Strand DNA

ERIC R. WARD* AND WAYNE M. BARNES†

The T-strand, a probable intermediate of *Agrobacterium* plant transformation, is bound by a nondenaturable linkage to a protein moiety at its 5′ end. The protein is shown to be the polypeptide VirD2, previously identified as a component of the T-DNA border endonuclease that initiates T-strand production. T-strands from an *Agrobacterium* strain expressing a *virD2-lacZ* fusion are bound to a protein of larger size than the wild-type protein and are immunoprecipitable by antibody to β-galactosidase.

Agrobacterium tumefaciens CAUSES THE neoplastic disease crown gall by transforming the nuclear genome of higher plants (*1*). The segment of DNA transferred (the T-DNA) resides on the large tumor-inducing (Ti) plasmid and is flanked by short direct repeats called border sequences (*2*). Six virulence (*vir*) loci, also located on the Ti plasmid, act on the T-DNA in trans to promote its transfer to plant cells (*3*). The *vir* genes are inducible by crude plant exudates and by specific phenolic compounds found therein (*4*). *Agrobacterium tumefaciens* cells grown under *vir*-inducing conditions produce single-stranded copies of the lower strand of the T-DNA called T-strands (*5–7*), which are thought to be intermediates of plant transformation. The first two open reading frames (ORFs) of *virD* (*virD1* and *virD2*) encode a site-specific endonuclease that nicks the lower strand of the border sequences (*8–10*), presumably initiating T-strand production (*5, 6, 10*).

Because mating bacterial cells also produce single strands of the DNA to be transferred (*11*), it has been suggested that *Agrobacterium*-mediated plant transformation may be similar to conjugation. Recent genetic evidence, showing that the mobilization (*mob*) genes and origin of transfer (*oriT*) of a mobilizable plasmid can substitute for T-DNA border sequences (*12*), has strengthened the case. In conjugation, one of the *mob*-encoded proteins involved in *oriT* nicking is thought to bind covalently to the 5′ end of the nick and remain associated with the conjugal plasmid strand (*11, 13*). We show that the *Agrobacterium* T-strand is bound tightly at or very near its 5′ end to a component of the border endonuclease.

We developed a method to construct Ti plasmids that bear tandem arrays of many

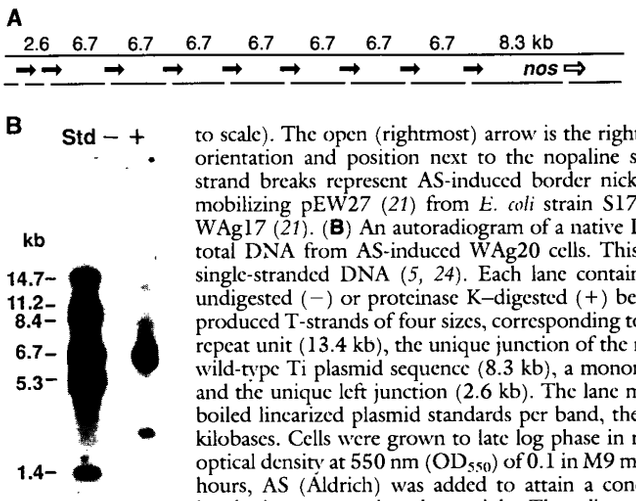

Fig. 1. (**A**) A schematic of the tandem T-DNA array carried by WAg20. Left border sequences are designated by filled arrows (not to scale). The open (rightmost) arrow is the right border of pTiT37 in its native orientation and position next to the nopaline synthase gene (*nos*). The lower strand breaks represent AS-induced border nicks. WAg20 was constructed by mobilizing pEW27 (*21*) from *E. coli* strain S17-1 (*22*) to *A. tumefaciens* strain WAg17 (*21*). (**B**) An autoradiogram of a native DNA blot transfer (*5, 23, 24*) of total DNA from AS-induced WAg20 cells. This type of transfer is specific for single-stranded DNA (*5, 24*). Each lane contains 40 ng of total DNA, either undigested (−) or proteinase K–digested (+) before phenol extraction. WAg20 produced T-strands of four sizes, corresponding to a dimer of the pEW27-derived repeat unit (13.4 kb), the unique junction of the right end of the tandem array to wild-type Ti plasmid sequence (8.3 kb), a monomer of the repeat unit (6.7 kb), and the unique left junction (2.6 kb). The lane marked "std" contains 200 pg of boiled linearized plasmid standards per band, the sizes of which are indicated in kilobases. Cells were grown to late log phase in rich medium, then diluted to an optical density at 550 nm (OD_{550}) of 0.1 in M9 medium, *p*H 5.4 (*25*). After 2 to 3 hours, AS (Aldrich) was added to attain a concentration of 100 μ*M* and the incubation was continued overnight. The cells were lysed in 50 m*M* tris-HCl, *p*H 7.9, 20 m*M* EDTA, and 1% sarkosyl (Sigma) and digested (or not) with proteinase K (Boehringer) at 200 μg/ml, followed by phenol and chloroform extraction and ethanol precipitation. The DNA was resuspended and separated by electrophoresis through 1% agarose in a tris-acetate buffer, then blotted as described (*5*). Hybridization, with nick-translated (*26*) pEW27 DNA (10 to 20 ng/ml) and washing were as described (*27*).

Department of Biological Chemistry, Washington University School of Medicine, St. Louis, MO 63110.

*Present address: CIBA-GEIGY Biotechnology Research, Box 12257, Research Triangle Park, NC, 27709.
†To whom correspondence should be addressed.