Single-Stranded DNA Binding Protein Encoded by the *virE* Locus of *Agrobacterium tumefaciens*

VITALY CITOVSKY, GUIDO DE VOS, PATRICIA ZAMBRYSKI

The transfer process of T (transfer)–DNA of *Agrobacterium tumefaciens* is activated after the induction of the expression of the Ti plasmid virulence (*vir*) loci by plant signal molecules such as acetosyringone. The *vir* gene products then act to generate a free transferable single-stranded copy of the T-DNA, designated the T-strand. Although some *vir* proteins are responsible for the synthesis of the T-strand, others may mediate T-strand transfer to plant cells as part of a DNA-protein complex. Here, a novel 69-kilodalton *vir*-specific single-stranded DNA binding protein is identified in *Agrobacterium* harboring a nopaline-type Ti plasmid. This protein binds single-stranded but not double-stranded DNA regardless of nucleotide sequence composition. The molecular size of the *vir*-specific single-stranded DNA binding protein and its relative abundance in acetosyringone-induced *Agrobacterium* suggested that it might be the product of the *virE* locus; molecular cloning and expression of the *virE* region in *Escherichia coli* confirmed this prediction.

FOR MILLENNIA THE COMMON SOIL phytopathogen *Agrobacterium tumefaciens* has been effectively doing what plant molecular biologists are just learning to do, that is, genetically transforming plant cells by transfer, integration, and subsequent expression of foreign genes. For *Agrobacterium*, this is achieved by transferring a segment of DNA, the transferred DNA (T-DNA), from its large Ti plasmid to the plant nuclear genome. The T-DNA functions solely as a structural element during its transfer; two 25-bp direct repeat sequences (T-DNA borders) at each of its ends provide the sites that are recognized during its mobilization to the plant cell (*1*). The functional products that mediate the recognition of the T-DNA borders, the generation of a transferable T-DNA copy, and the ultimate transfer of the T-DNA molecule to plant cells are encoded primarily in the Ti plasmid virulence (*vir*) region (*2*).

The *vir* region is large, 35 kbp, and includes six loci designated *virA*, *virB*, *virC*, *virD*, *virE*, and *virG*. All these loci constitute a single regulon that is activated by signal molecules from wounded plant cells (*2, 3*). In *Nicotiana tabacum* these factors were identified as the phenolic compounds acetosyringone (AS) and α-hydroxyacetosyringone (*4*). Two *vir* loci, *virA* and *virG*, provide the regulatory system through which *Agrobacterium* responds to plant signal molecules (*3, 5*). The functions of the other *vir* loci are in general less well understood but are presumed to mediate the T-DNA transfer process itself. For example, it has recently been shown that induction of *vir* gene expression results in the production of single-stranded (ss) endonucleolytic cleavages within the T-DNA border sequences (*6–9*). *Vir* gene expression also results in the production of a novel linear ssDNA copy of the bottom strand of the T-DNA, designated the T-strand (*6, 9*); border nicks are presumed to represent initiation and termination sites for T-strand synthesis. The T-strand is proposed to be the T-DNA intermediate molecule that *Agrobacterium* mobilizes to plant cells (*6, 9, 10*). Two polypeptides (VirD1 and VirD2) located in the 5′ half of the *virD* locus specify the T-DNA border endonuclease (*7, 9, 11*). Except for *virD1* and *virD2*, no other *vir*-specific functions have been shown to be directly involved in T-strand production; presumably other functions, such as helicases and polymerases, which mediate later steps in T-strand synthesis, are essential bacterial products encoded by the *A. tumefaciens* chromosome.

We speculated that, once the T-strand is formed, it becomes associated with proteins that facilitate its transfer to the plant cell. Since the T-strand is a linear ssDNA molecule, the most obvious type of T-strand–associated protein to test for would be an ssDNA binding protein. Such a protein might have several possible roles, such as packaging the T-strand into a transferable complex, protecting the T-strand from nucleases, and potentially even participating in active transport and integration of the T-DNA molecule into the plant cell genomic DNA. Indeed, this report shows that the T-DNA transfer process has evolved to specify its own ssDNA binding protein, encoded by the *virE* locus.

Retardation gel electrophoresis was used as an assay for ssDNA binding activity in extracts of uninduced and AS-induced *Agrobacterium* cells carrying the nopaline Ti plasmid pGV3850 (*12*). Since a general property of ssDNA binding proteins is the ability to bind ssDNA without sequence specificity (*13*), an unrelated (non-Ti) 587-bp Hae III DNA fragment from pUC18 (*14*) was used as a probe for protein-DNA interactions. This fragment was end-labeled and either denatured or not denatured before incubation with bacterial extracts (Fig. 1). When the ss probe was incubated with uninduced extracts, only a slight retardation of the

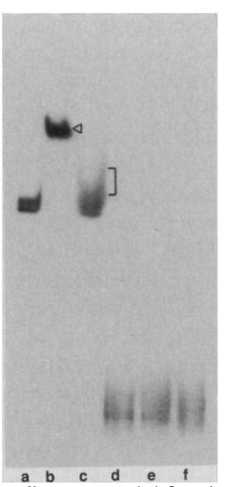

Fig. 1. Single-stranded DNA binding activity in *Agrobacterium* extracts. *Agrobacterium* cells [C58C1 strain carrying Ti plasmid pGV3850 (*12*)] were grown in YEB medium (*32*), pelleted, and resuspended at a density of 0.05 (A_{600}) in M9 minimal medium (*33*), *p*H 5.5, supplemented with 0.2% glucose, and grown for 4 hours at 28°C before the addition of 100 μ*M* acetosyringone (AS); note that we previously used MS (Murashige and Skoog) plant medium at *p*H 5.5 for *vir* induction with AS (*4*). Induction is equally efficient in the M9 medium as long as the *p*H is 5.5, and M9 is easier to prepare. After induction with AS for 14 hours (*4*), 60 ml of cells (A_{600} = 0.1) were harvested (10,000*g* for 5 minutes at 4°C), resuspended in 2 ml of binding buffer (10 m*M* tris-HCl, *p*H 8.0, 200 m*M* NaCl, 20 m*M* EDTA, 0.5 m*M* dithiothreitol, 0.5 m*M* phenylmethylsulfonyl fluoride, and 10% glycerol), and lysed by means of the French press minicell at 20,000 psi. The lysate was passed three times through a 26-gauge needle to shear DNA, and centrifuged (10,000*g* for 5 minutes at 4°C) to remove unlysed cells. The supernatant was collected, heated for 10 minutes at 70°C to inactivate nucleases, and stored at −70°C until use. For retardation experiments, 20 μl of cell extract (5 μg of protein) was incubated for 30 minutes at 4°C with 0.1 ng of probe DNA [the 587-bp Hae III fragment of pUC18 (*14*) blunt end-labeled (*34*) to a specific activity of 5×10^7 cpm/μg]. At the end of the incubation period, the samples were loaded on a 4% native polyacrylamide gel prepared and electrophoresed in low ionic strength buffer (6.7 m*M* tris-HCl, *p*H 7.5, 1 m*M* EDTA, and 3.3 m*M* sodium acetate); gels were run at 4°C with buffer circulation (*35*). After electrophoresis, the gel was dried and exposed to x-ray film with an intensifying screen at −70°C for 12 to 24 hours. (Lane a) Single-stranded denatured probe incubated in binding buffer alone. (Lane b) Single-stranded denatured probe incubated with extract of AS-induced cells. (Lane c) Single-stranded probe incubated with extract of uninduced cells. (Lane d) Double-stranded probe incubated in binding buffer alone. (Lane e) Double-stranded probe incubated with extract of AS-induced cells. (Lane f) Double-stranded probe incubated with extract of the uninduced cells. The arrow indicates strong retardation of the probe, and the bracket indicates slight retardation of the probe.

Division of Molecular Plant Biology, Hilgard Hall, University of California, Berkeley, CA 94720.

DNA was observed (Fig. 1, lane c); this result may reflect the detection of nonabundant constitutively expressed ssDNA binding proteins. In contrast, incubation of the ss probe with extract from cells induced for *vir* gene expression resulted in a very strong retardation of the DNA (Fig. 1, lane b). No retardation of the undenatured, double-stranded (ds) DNA probe occurred upon incubation with either extract (Fig. 1, lanes d to f). Exactly the same results, namely, strong *vir*-specific ssDNA binding activity,

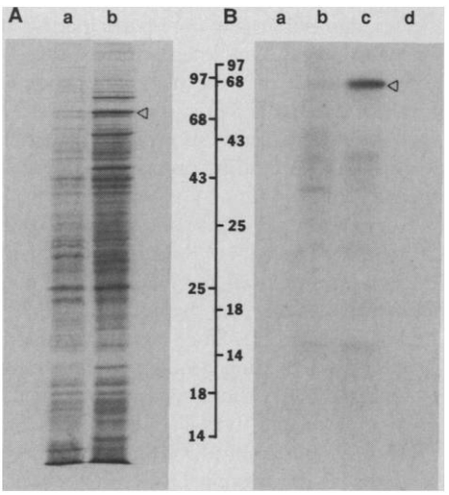

Fig. 2. SDS–polyacrylamide gel electrophoresis (PAGE) analysis of the AS-induced ssDNA binding activity. (**A**) AS-induced protein synthesis in *Agrobacterium* cells. Uninduced (lane a) and AS-induced (lane b) cells were prepared as described in Fig. 1. Cells were labeled for 30 minutes with 6 μCi of [^{35}S]methionine (1129 Ci/mmol), harvested, and separated by electrophoresis on a 12.5% SDS-polyacrylamide gel as previously described (*23*). The arrow indicates the position of the 69-kD AS-induced protein. (**B**) SDS-PAGE analysis of *Agrobacterium* ssDNA binding proteins. *Agrobacterium* cell extracts were obtained and assayed on two 4% polyacrylamide retardation gels; gel 1 assayed extract from uninduced cells, and gel 2 assayed extract from AS-induced cells. Electrophoresis conditions were as described in Fig. 1, except that preparative (600 μl) combs were used to load the samples (0.03 μg of probe DNA incubated with 125 μg of total protein derived from 10^9 cells). The gels were autoradiographed wet in a sealed bag, and the following regions of the gels were cut out for analysis: (lane a) region of gel 1 corresponding to the position of the retarded probe in gel 2; (lane b) region of gel 1 containing the slightly retarded (including the nonretarded DNA) DNA band (equivalent to lane c, Fig. 1); (lane c) region of gel 2 containing the retarded probe (equivalent to lane b, Fig. 1); and (lane d) region of gel 2 corresponding to the slightly retarded and nonretarded band in gel 1. Proteins from these gel regions were electroeluted in 50 m*M* sodium borate buffer (*p*H 8.5), concentrated twofold by lyophilization and radioiodinated with diiodo [^{125}I]Bolton-Hunter reagent (5 μCi per 100 ng of protein for 1 hour at 4°C) (*17*). After iodination, the protein samples were extensively dialyzed against binding buffer and then analyzed on a 12.5% SDS-polyacrylamide gel.

were obtained with another nopaline Ti plasmid, pMP90 (*15*), which contains a deletion of the T-DNA region and consequently cannot produce T-strand molecules. These last results suggest that we are detecting, under the experimental conditions used, either mostly free excess ssDNA binding protein or that any T-strand protein complex dissociates during preparation of the extracts from the T-DNA (pGV3850)–containing cells. In addition, *vir*-induced extracts were able to retard a variety of ssDNA probes with equal efficiency (*16*), confirming that the *vir*-specific ssDNA binding activity binds DNA independent of sequence composition.

Four samples were electroeluted from different parts of the retardation gel to purify and further characterize the *Agrobacterium vir*-specific ssDNA binding protein: (i) the strongly retarded (low mobility) radioactive ssDNA region resulting from incubation with extract from AS-induced cells (as in Fig. 1, lane b), (ii) the corresponding (unlabeled) region of the gel resulting from incubation with extract from uninduced cells, (iii) the partially retarded (high mobility) radioactive ssDNA region resulting from incubation with extract from uninduced cells (as in Fig. 1, lane c), and (iv) the corresponding unlabeled region of the gel resulting from incubation with extract from induced cells. The protein composition of these four samples was determined after radioiodination with Bolton-Hunter reagent (*17*) and electrophoresis on an SDS-polyacrylamide gel (*18*). The protein pattern (Fig. 2B, lane c) clearly indicates the presence of a dominant band of molecular size 69 kD; this intense band is only seen in the sample derived from the gel-retarded ssDNA probe resulting from incubation with extract from AS-induced cells. This sample also shows the presence of additional less abundant ssDNA binding proteins covering a wide range of molecular weights; a similar distribution of proteins is found in the sample from the ssDNA retarded probe obtained with extracts from uninduced cells (Fig. 2B, lane b). Presumably the latter proteins represent constitutively expressed *Agrobacterium*-specific ssDNA binding proteins. No iodinated bands were detected in the two control samples that did not comigrate with the radioactive DNA probe (Fig. 2B, lanes a and d). The method we used, namely, gel retardation followed by electroelution and radioiodination with Bolton-Hunter reagent, may be generally useful to isolate either sequence-specific or nonspecific DNA binding proteins. This method efficiently recovers DNA binding proteins; the yield obtained (Fig. 2B, lane b) correlates well with that expected from the intensity of this protein in total extracts from AS-induced cells [see below and (*19*)].

Since the abundant 69-kD ssDNA binding protein is expressed only in AS-induced cells, it could be encoded by one of the *vir* loci. Thus, we looked for a correlation between its molecular size and the relative mobility of the known AS-induced proteins of *Agrobacterium*. Figure 2A displays the patterns of AS-induced and uninduced proteins from *Agrobacterium* carrying the pGV3850 nopaline-derived Ti plasmid. There is a strong protein band of 69 kD in the AS-induced pattern. The relative abundance and mobility of this protein suggest that it is one of the products of the *virE* locus. The nucleotide sequence of the nopaline *virE* locus has been determined and predicts a 64-kD polypeptide in the largest open reading frame, and this polypeptide has a relative mobility of 69 kD on SDS-polyacrylamide gels (*20, 21*). Similarly, the nucleotide sequence of the 2.0-kbp octopine *virE* region reveals two open reading frames specifying 7.0- and 60.5-kD polypeptides (*22*); the larger octopine *virE* polypeptide migrates with a relative mobility corresponding to 65 kD and it is the most abundant protein detected in AS-induced cell extracts (*23*).

To test directly the possibility that the AS-induced ssDNA binding protein is the product of the *virE* locus, we cloned the *virE* region of the nopaline Ti plasmid, expressed it in *Escherichia coli*, and subsequently assayed it for ssDNA binding activity. A restriction fragment carrying the *virE* structural gene (*21*) was cloned in two orientations into a plasmid carrying a bactriophage T7–specific promoter (*24*). When *E. coli* carrying the plasmid with *virE* in the correct orientation was infected with lambda phage carrying the T7-specific RNA polymerase (*25, 26*), a high level of the 69-kD *virE* product was detected (Fig. 3A, lane b); no *virE* product was detected from the plasmid carrying *virE* in the reverse orientation (Fig. 3A, lane a). Cell-free extracts were prepared and used in the gel retardation assay. Extracts derived from cells expressing *virE* reveal a strong binding activity for ssDNA but not for dsDNA (Fig. 3B, lanes d and c, respectively). This dramatic shift in mobility of the ssDNA probe directly mimics the results obtained with extracts from AS-induced *Agrobacterium* (Fig. 1). No DNA binding activity was observed with extracts from cells carrying *virE* in the wrong orientation for expression (Fig. 3B, lanes e and f), or from nonexpressing (that is, not infected with lambda-carrying T7 RNA polymerase) cells carrying the proper orientation of *virE* sequences (Fig. 3B, lanes g and h). Thus, our results show the function of the 69-kD

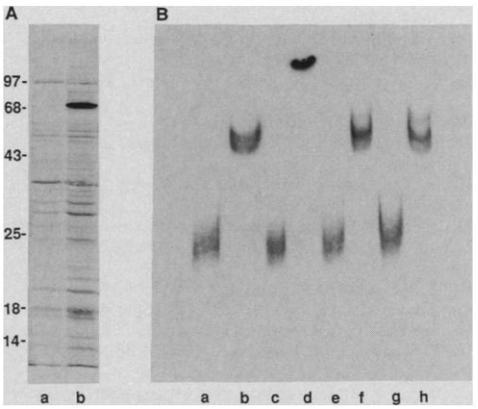

Fig. 3. Single-stranded DNA binding activity in extracts of *E. coli* expressing the *virE* product of *Agrobacterium*. *Escherichia coli* cells containing plasmids carrying the correct and reverse orientations of the nopaline *virE* gene (*24*) were grown in M9 medium supplemented with 0.4% maltose, 2 m*M* $MgCl_2$, and 20 μg of essential amino acids per milliliter. At a cell density of A_{600} = 0.3 (2×10^8 to 3×10^8 viable cells per milliliter), CE6 phage (*25*), which express the T7-specific RNA polymerase, were added at a multiplicity of 7. Cells were harvested 3 hours after infection. For labeling experiments in (**A**), 6 μCi of [^{35}S]methionine was added 2.5 hours after infection to cells carrying the reverse (lane a) or correct (lane b) orientation of *virE* sequences behind the T7 promoter. The labeled cell extracts were prepared and analyzed by SDS-PAGE as in Fig. 2A. For the retardation assay (**B**), ssDNA (lanes b, d, f, and h) or dsDNA (lanes a, c, e, and g) were incubated alone (lanes a and b) or with extracts of: CE6-infected cells carrying the correct orientation of *virE* (lanes c and d); CE6-infected cells carrying the reverse orientation of *virE* (lanes e and f); or uninfected cells carrying the correct orientation of *virE* (lanes g and h). The overproduced VirE protein in *E. coli* occurs mainly in an insoluble fraction after cell lysis. This fraction was dissolved in binding buffer containing 4*M* urea, centrifuged (25,000*g* for 15 minutes at 4°C), and the supernatant dialyzed against binding buffer. This treatment solubilizes about 50% of the overproduced VirE protein. All other experimental conditions were as in Fig. 1 for *Agrobacterium* extracts.

protein product of the *virE* locus is binding of ssDNA.

Single-stranded DNA binding proteins can be classified into two major groups: (i) enzymes and cofactors: for example, RecA protein, DNA polymerase, lactate dehydrogenase, RNA polymerase, and C_1C_2 polymerase alpha cofactors; and (ii) true SSB (ssDNA binding protein) (*13*). The latter class designation implies similar modes of action to *E. coli* SSB and T4 bacteriophage gene 32 proteins, (i) ssDNA binding, (ii) requirement in stoichiometric rather than catalytic amounts, (iii) cooperative DNA binding, (iv) stimulation of DNA polymerase, and (v) absence of intrinsic enzymatic activity. The VirE protein fulfills the first and likely the second criterion for an SSB. Whether it can bind ssDNA in a cooperative fashion, its involvement in genomic replication, or its possible enzymatic activities, such as DNA-dependent adenosine triphosphatase or helicase, can be directly tested. Nevertheless, that the VirE protein is one of the most abundantly produced *vir* proteins supports the idea that its role is structural rather than enzymatic (Fig. 2) (*23*). For example, to completely protect a 20-kb T-strand molecule would require 350 to 700 protein molecules, assuming a 60- to 30-nucleotide substrate [by analogy to the size of the binding sites for the smooth contoured or beaded structure of the *E. coli* SSB tetramer (*27*)]; the amount of *virE* product observed in AS-induced *Agrobacterium* is consistent with this estimate (*19*).

That *Agrobacterium* has evolved an ssDNA binding protein specifically under *vir* control strengthens the hypothesis that an ssDNA, presumably the T-strand, is indeed the molecule that is transferred by *Agrobacterium* to plant cells (*6, 9, 10*). Presumably the VirE protein is part of protein-DNA complex that forms the actual transfer structure. Since it has been proposed that the T-strand is transferred to plant cells by a mechanism similar to bacterial conjugation (*6, 10*), the VirE protein may act to protect the T-strand during transfer. However, it should be noted that the role of ssDNA binding proteins in conjugative DNA transfer is still obsure. For example, mutants defective in F-factor specific SSB are still capable of conjugative DNA transfer; this result has been explained by assuming that *E. coli* SSB can substitute for F-factor SSB (*28*).

Nevertheless, the analogy of VirE protein to F-factor SSB may help to explain some of the unusual phenotypes of *virE* mutants. For example, *virE* mutants are avirulent only on some plant hosts (*2, 20, 29*). Since ssDNA binding proteins have high affinity for ssDNA without regard to sequence composition (*13*), *Agrobacterium* chromosomally encoded ssDNA binding proteins may substitute for *vir* ssDNA binding activity in *virE* mutants; these less abundant SSBs may be able to protect the T-strand from degradation in some, but not all, plant hosts. Furthermore, *virE* mutants can be complemented to wild-type virulence if coinoculated on plants with *Agrobacterium* carrying a wild-type *vir* region (*30*). These data have led to the suggestion that the VirE protein is secreted and functions extracellularly, potentially as part of a transferable T-DNA–protein complex; that *virE* specifies a DNA binding protein supports these earlier hypotheses.

The *Agrobacterium virE* ssDNA binding protein is one of the largest prokaryotic ssDNA binding proteins identified to date (*13*); potentially the VirE protein has different functional domains that act at different steps of the T-DNA transfer process. In fact, the ssDNA binding activity may not be its major function. Although the existing data are tentative, it has been suggested that the VirE protein also plays a role in the integration of the T-DNA into the plant cell genome (*31*). Or, VirE protein may facilitate passage of the T-DNA transfer complex through the bacterial and plant cell and nuclear membranes; the latter function is consistent with the above-mentioned observation that *virE* mutants can be complemented extracellularly (*30*). Further experiments are needed to elucidate the biological role of the *Agrobacterium vir*-specific ssDNA binding protein.

REFERENCES AND NOTES

1. P. Zambryski, A. Depicker, K. Kruger, H. Goodman, *J. Mol. Appl. Genet.* **1**, 361 (1982); N. S. Yadav, J. Vanderleyden, D. R. Bennett, W. M. Barnes, M.-D. Chilton, *Proc. Natl. Acad. Sci. U.S.A.* **79**, 6322 (1982); R. F. Barker *et al.*, *Plant Mol. Biol.* **2**, 335 (1983); J. L. Slightom *et al.*, *EMBO J.* **4**, 3069 (1985); K. Wang, L. Herrera-Estrella, M. Van Montagu, P. Zambryski, *Cell* **38**, 35 (1984); E. G. Peralta and L. W. Ream, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 5112 (1985); K. Wang, C. Genetello, M. Van Montagu, P. Zambryski, *Mol. Gen. Genet.* **210**, 338 (1987).
2. S. Stachel and E. W. Nester, *EMBO J.* **5**, 1445 (1986).
3. S. Stachel and P. Zambryski, *Cell* **46**, 325 (1986).
4. S. Stachel, E. Messens, M. Van Montagu, P. Zambryski, *Nature (London)* **318**, 624 (1985).
5. S. C. Winans, P. R. Ebert, S. E. Stachel, M. P. Gordon, E. W. Nester, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 8278 (1986); B. Leroux *et al.*, *EMBO J.* **6**, 1020 (1987).
6. S. Stachel, B. Timmerman, P. Zambryski, *Nature (London)* **322**, 706 (1986).
7. M. Yanofsky *et al.*, *Cell* **47**, 471 (1986).
8. K. Wang, S. E. Stachel, B. Timmerman, M. Van Montagu, P. C. Zambryski, *Science* **235**, 587 (1987); L. M. Albright *et al.*, *J. Bacteriol.* **169**, 1046 (1987).
9. S. Stachel, B. Timmerman, P. Zambryski, *EMBO J.* **6**, 857 (1987).
10. S. Stachel and P. Zambryski, *Cell* **47**, 155 (1986).
11. K. Veluthambi, R. K. Jayaswal, S. B. Gelvin, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 1881 (1987); R. K. Jayaswal, K. Veluthambi, S. B. Gelvin, J. L. Slightom, *J. Bacteriol.* **169**, 5035 (1987).
12. P. Zambryski *et al.*, *EMBO J.* **2**, 2143 (1983).
13. J. W. Chase and K. R. Williams, *Annu. Rev. Biochem.* **55**, 103 (1986).
14. C. Yanisch-Perron, J. Vieira, J. Messing, *Gene* **33**, 103 (1985).
15. C. Koncz and J. Schell, *Mol. Gen. Genet.* **204**, 383 (1986).
16. Hae III DNA fragment (267 bp) of pUC18 (*14*); 750-bp Cla I–Hind III DNA fragment of pAGS111 [P. Van Den Elsen *et al.*, *Plant Mol. Biol.* **5**, 149 (1985)]; and 230-bp Pvu II–Bam HI DNA fragment of pE8 containing a portion of an ethylene-induced complementary DNA clone from *Lycopersicon esculentum* [J. E. Lincoln, S. Cordes, E. Read, R. L. Fischer, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 2793 (1987)].
17. A. E. Bolton and W. M. Hunter, *Biochem. J.* **133**, 529 (1973).
18. U. K. Laemmli, *Nature (London)* **227**, 680 (1970).
19. VirE represents 0.05 to 0.1% of the total protein from AS-induced cells (Fig. 2A), and we recover

approximately 0.05 μg of VirE protein from the preparative gel using extracts containing 125 μg of total protein, that is, 0.04% of the total which corresponds to roughly 40 to 80% recovery.
20. T. Hirooka and C. I. Kado, *J. Bacteriol.* **168**, 237 (1986).
21. T. Hirooka, P. M. Rogowsky, C. I. Kado, *ibid.* **169**, 1529 (1987).
22. S. C. Winans *et al.*, *Nucleic Acids Res.* **15**, 825 (1987).
23. P. Engstrom, P. Zambryski, M. Van Montagu, S. Stachel, *J. Mol. Biol.* **197**, 635 (1987).
24. The Bam HI restriction fragment 15 of pTiC58, which contains sequences directing translation of all of the 69-kD VirE protein except for the first amino acid (*21*), was cloned into the T7 RNA polymerase expression vector pET3b (*26*). The fragment was isolated from pGV0361 [A. Depicker *et al.*, *Plasmid* **3**, 193 (1980)] and cloned in both orientations into the vector by standard procedures (*34*). In the correctly orientated clone, the first methionine of the VirE protein is replaced by 12 amino acids encoded by the vector, generating a hybrid protein in the proper reading frame.
25. F. W. Studier and B. A. Moffatt, *J. Mol. Biol.* **189**, 113 (1986).
26. A. H. Rosenberg *et al.*, *Gene* **56**, 125 (1987).
27. T. Lohman and L. B. Overman, *J. Biol. Chem.* **260**, 3594 (1985).
28. K. A. Ippen-Ihler and E. G. Minkley, *Annu. Rev. Genet.* **20**, 593 (1986).
29. M. Yanofsky *et al.*, *Mol. Gen. Genet.* **201**, 237 (1985).
30. L. Otten *et al.*, *ibid.* **195**, 159 (1984).
31. R. C. Gardner and V. C. Knauf, *Science* **231**, 725 (1986).
32. N. Van Larebeke *et al.*, *Mol. Gen. Genet.* **152**, 119 (1977).
33. J. H. Miller, Ed., *Experiments in Molecular Genetics* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1972).
34. T. Maniatis, E. F. Fritsch, J. Sambrook, *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1982).
35. F. Strauss and A. Varshavsky, *Cell* **37**, 889 (1984).
36. P.Z. and G.D.V. thank S. Stachel for suggesting that the product of the *virE* locus might be an ssDNA binding protein. We are grateful to W. Studier for T7 expression plasmids, J. Deikman for providing 230-bp fragment of pE8, E. Howard and B. Winsor for critical reading of the manuscript. V.C. is supported by a Chaim Weizmann Postdoctoral Fellowship. The research was supported by NSF grant DMB-8617772 to P.Z.

28 December 1987; accepted 7 March 1988

Cleaving DNA at Any Predetermined Site with Adapter-Primers and Class-IIS Restriction Enzymes

SUN C. KIM, ANNA J. PODHAJSKA,* WACLAW SZYBALSKI†

A four-component system has been designed that makes it possible to prepare a double-stranded (ds) DNA fragment; one fragment end is predesigned (by the use of a class-IIS restriction enzyme and adapter-primer), and the other end corresponds to any normal restriction cut. The system is composed of the phage M13mp7 single-stranded (ss) target DNA; the Fok I restriction enzyme; an oligodeoxynucleotide adapter-primer, which permits one to introduce Fok I cuts at any specified site in the target DNA; and DNA polymerase, which converts the ss target into a ds form ready for cloning. In this system, the oligodeoxynucleotide adapter-primer serves several purposes. The 5′ hairpin ds domain of the adapter-primer contains the Fok I recognition site. Its 3′ ss domain selects a complementary site on the target ss DNA, hybridizes with it to form the ds cleavage site, and serves as a primer to convert the ss M13mp7 target to ds DNA.

THERE IS A LIMITED NUMBER OF restriction enzymes, with which it is possible to cleave DNA only at preexisting recognition sites (*1*). It would be of great advantage, however, to be able to cut DNA at any predetermined site. To attain this goal it is necessary to use a combination of two moieties, one with the DNA site specificity and another with the DNA cleaving function. We have used a specially designed oligodeoxynucleotide adapter and a class-IIS restriction enzyme, which permitted a precise cut between any predetermined two nucleotides of the single-stranded (ss) DNA target (*2–4*). Other investigators have used either an oligodeoxynucleotide (*5, 6*) or site-specific DNA binding protein (*7*) to guide the cutting complex to the specific site on DNA; the target DNA was cut either chemically, by EDTA·Fe complex (*5, 7*), or by an otherwise nonspecific nuclease (*6*). In these three methods (*5, 7*), the DNA cuts were imprecise and statistically distributed around the target sequence. Our present system produces double-stranded (ds) DNA fragments with one predetermined and precisely cut end.

Most of the class-II restriction endonucleases (*1*) recognize a specific sequence on ds DNA and cleave it at specific points within this sequence. Subclass IIS (*2*) cuts DNA, not within the recognition sequence, but at a precise distance from it on the same molecule. This physical separation between the recognition and cleavage sites permitted us to design a system (*2, 3*) where the Fok I recognition site and one of the cuts are on an adapter oligodeoxynucleotide, whereas the second cut site is on M13mp7 ss target DNA.

In the present study we have tried to redesign our system (*3*) so as to obtain defined ds DNA fragments, ready for cloning, instead of just introducing cuts in the ss target DNA at the predetermined sites. The general design of the new four-component system, as adapted for the Fok I enzyme, is shown in Fig. 1. The 34-mer adapter has a 5′-$\overset{GGATG}{CCTAC}$ Fok I recognition site within its ds hairpin domain and a 14-nucleotide (nt) ss domain; the latter is complementary to nt 1339–1353 of the M13mp7 (*4*) ss target DNA. This ss domain plays two roles in our system: (i) when paired with M13mp7 ss DNA, it provides a ds cut site for Fok I (indicated by vertical arrows in Fig. 1A) and (ii) it serves as a primer for PolIk-catalyzed DNA synthesis starting at its 3′-OH terminus, using M13mp7 ss target DNA as a template, and thus creating the ds M13mp7 molecule with a novel predesigned restriction site for the Fok I enzyme (Fig. 1A). The products of this novel Fok I cut are shown in Fig. 1B. The cohesive ends are filled in by PolIk, as shown by the italicized letters in Fig. 1B.

The sizes of the expected Fok I digestion products are summarized in Fig. 2. The replicative form (RF) ds DNA of M13mp7 is cut into 218-, 910-, 2775-, and 3335-bp fragments by Fok I (see the outer ring of Fig. 2 and lanes 6 and 7 of Fig. 3). The 34-mer adapter would induce a new cut between nt 1341 and 1342 of the M13mp7 (+) strand, and thus a new 1116-bp fragment should be produced (Fig. 2) shortly after the intitiation of the adapter-primed, PolIk–catalyzed DNA synthesis. Indeed, as shown in Fig. 3, a new 1116-bp fragment appeared as a major digestion product within 10 minutes after the start of DNA synthesis (lane 2); this fragment was totally absent in lanes 6 and 7, which represent the Fok I–cleaved RF ds DNA of M13mp7. As DNA synthesis proceeded, the regular products (analogous to those in lanes 6 and 7) of Fok I cleavage appeared, first the shortest (218-bp) fragment and later the 910-bp, 2775-bp, and finally the 3335-bp products (lanes 2 through 5), in the same order as they are primed by the 34-mer adapter-primer oligodeoxynucleotide (Fig. 2). Actually, the restriction patterns in Fig. 3 are somewhat more complex, because the 34-mer has a secondary binding site on M13mp7 DNA, including the second cut next to nt 6938 (Fig. 2) (*3*). This cut would generate a 603-bp fragment (Fig. 2), which is present in the form of an early appearing 603-bp minor

McArdle Laboratory for Cancer Research, University of Wisconsin, Madison, WI 53706.

*Permanent address: Department of Microbiology, University of Gdańsk, ul. Kładki 24, 80-822 Gdańsk, Poland.
†To whom correspondence should be addressed.