Delineation of the regulatory region sequences of *Agrobacterium tumefaciens* *virB* operon

Anath Das* and Gregory J. Pazour

Department of Biochemistry and Plant Molecular Genetics Institute, University of Minnesota, 1479 Gortner Avenue, St Paul, MN 55108, USA

Received March 29, 1989; Revised and Accepted May 19, 1989

ABSTRACT

A *virB-lacZ* translational fusion was constructed to monitor expression of the *Agrobacterium tumefaciens* *virB* operon. Expression of the fusion gene was dependent on the presence of pTiA6 *virA*, *virG*, and a plant factor acetosyringone. Analysis of deletion mutants, constructed by exonuclease *Bal31* digestion, showed that 68 residues upstream of the *virB* transcription initiation site was necessary for its expression. A TT → CC substitution at positions -62 and -61 led to a 7 fold reduction in *virB* expression. The *virB* upstream region contains a tetradecameric sequence, dPuT/ATDCAATTGHAAPy (D=A, G or T; H=A, C or T), that is conserved in the non-transcribed regions of all *vir* genes. Alteration of the position of this sequence relative to the promoter region sequences had a drastic negative effect on *virB* expression.

INTRODUCTION

Agrobacterium tumefaciens, a phytopathogenic soil bacterium, causes neoplastic crown gall tumor disease on most dicotyledonous plants. During infection at a wound site the bacterium transfers a segment of its plasmid-borne DNA to plant cells. The transferred (T-) DNA is found integrated into the plant nuclear genome (reviewed in ref. 1,2). The T-DNA contains, among others, genes for phytohormone biosynthesis. Expression of these genes in a transformed plant cell leads to the tumorous phenotype associated with crown gall disease. In addition to the T-DNA two other regions, the chromosomal *chv* and the Ti-plasmid virulence (*vir*) regions, are required for tumor formation (3-5). The *vir* region is believed to encode functions necessary for T-DNA transfer to plant cells.

The *vir* region consists of at least six complementation groups, viz. *virA*, *virB*, *virC*, *virD*, *virE* and *virG* (6). Of these only *virA* and *virG* are expressed in free-living bacteria. Expression of other *vir* genes requires *virA*, *virG*, and a plant cell factor(s) (7,8). Two plant cell factors identified were acetosyringone (AS) and α -hydroxyacetosyringone (9). Expression of the two regulatory genes *virA* and *virG* are also controlled by these factors (10,11). DNA sequence analysis of the octopine Ti-plasmid pTiA6 *virA* and *virG* revealed that the encoded polypeptides share considerable homology with several regulatory proteins of *E. coli* and other bacterial species (12,13). *VirA* (the *virA* encoded gene product) is homologous to EnvZ, NtrB, CpxA, PhoB and CheA, and *VirG* is homologous to OmpR, NtrC, SfrA, PhoR, CheB and CheY (12-15). In all cases one component (*VirA* or its homologue) responds to environmental stimuli and modulates the other (*VirG* or its homologue) to serve as a transcriptional activator. Recent studies demonstrate that *VirA* homologues NtrB, EnvZ and CheA, and *VirG* homologues NtrC, CheB and CheY are modified by phosphorylation (16-20).

To identify sequences that interact with the *vir* gene regulators the DNA sequences of

the nontranscribed regions of the *vir* genes were determined (21). Analysis of these sequences identified two conserved hexameric sequences that were postulated to play a role in *vir* gene regulation. To determine if these sequences are required for *vir* gene expression and to identify the *cis*-acting regulatory sequences we constructed a *virB-lacZ* translational fusion. By limited digestion with the exonuclease *Bal31* several deletion mutants that had lost defined regions of *virB* upstream sequences were constructed. Expression of *virB* was monitored by measuring the activity of the reporter gene product β -galactosidase. The present study demonstrates that 68 basepairs (bp) upstream of the *virB* transcription initiation site are sufficient for its expression. Site-specific and insertional mutagenesis procedures were used to identify domains and critical residues required for plant-inducible expression of *virB*. We demonstrate that one of the two conserved hexamers we identified in the previous study (21) constitute part of the upstream regulatory sequences.

MATERIALS AND METHODS

Construction of a virB-lacZ Fusion Plasmid

The promoter-regulatory region and part of the first open-reading frame (ORF) of the pTiA6 *virB* operon was isolated as a 329 basepair (bp) *AvaII-HinfI* restriction fragment (21,22). After filling in with DNA polymerase I Klenow enzyme in the presence of deoxynucleoside triphosphates (dNTPs) the fragment was cloned into the *SmaI* restriction site of plasmid pMC1403, a plasmid containing the *E. coli* lactose (*lac*) operon without a promoter sequence (23). The resultant plasmid pAD1041 contained a translational fusion of the 54th codon of the first ORF of *virB* with the 6th codon of *E. coli* β -galactosidase (*lacZ*) gene. The *virB-lacZ* gene fusion in pAD1041 is under the transcriptional control of the *virB* sequences. A 3.6 kb *EcoRI-DraI* restriction fragment containing the *virB-lacZ* fusion gene and the *virB* regulatory sequences was then cloned into a *EcoRI*-filled in *BglIII* site of plasmid pAD1085. Plasmid pAD1085 is a derivative of plasmid pTJS75 that contains a RK2 replicon which allows stable maintenance of plasmids in *Agrobacterium* (24). The resultant plasmid pAD1092 has a unique *EcoRI* restriction site immediately upstream of the *virB* sequences.

Other Plasmids

Plasmid pGP159 is a derivative of plasmid pAD1092. A unique *XbaI* site of pAD1092 was changed to a *KpnI* site by insertion of a linker DNA. This plasmid was linearized with *KpnI* and ligated with a 4.6 kb *KpnI* fragment containing the pTiA6 *virA* gene (24). The resultant plasmid pGP119 was linearized at the *EcoRI* site and a plasmid containing the pTiA6 *virG* (pGP119, a pUC119 derivative (26) that contains the pTiA6 *virG* on a 1.25 kb *BglIII-PstI* fragment (12)) was cloned into this site to yield plasmid pGP159. Plasmid pGP159 contains pTiA6 *virA*, *virG*, and a *virB-lacZ* gene fusion.

Plasmids pAD1199, 1199R and 1200 are derivatives of deletion mutant $\Delta 46$ (discussed below). Plasmid $\Delta 46$ was linearized with *EcoRI*. Before ligating with a synthetic oligonucleotide dAAGCTTCAA (unkinased) linearized $\Delta 46$ DNA was either treated with S1 nuclease (pAD1199, 1199R) or filled-in with DNA polymerase I Klenow enzyme and dNTPs (pAD1200). Plasmids pAD1199 and 1200 contain the insertion sequence dAAGCTTCAA, and pAD1199R has dTTGAAGCTT (complementary sequence) insertion. The insert contains a *HindIII* restriction site that was used to monitor oligodeoxynucleotide insertion.

Plasmids pAD1210, 1212, and 1214 were constructed as follows. Plasmid $\Delta 38$ was digested with *XmnI* (cleaves between -35 and -36 sequences of *virB*). After ligating with a *XbaI* linker DNA (dTCTAGA) the plasmid was cleaved with *XbaI* and *EcoRI*.

Fig. 1. Construction of a *virB-lacZ* Fusion Plasmid: Details of plasmid construction have been described in Materials and Methods. DNA sequences around vector-insert junction regions are shown at the bottom of the figure. Sequences derived from *virB* operon are shown in light type. Lac5, Lac6, and B54 indicate the 5th and 6th codon of *lacZ*, and the 54th codon of *virB1*, respectively. 'Z and *virB-lacZ*' indicates truncated *lacZ* gene and a fusion consisting of partial *virB* and partial *lacZ* coding segments, respectively. Relative positions of restriction endonuclease sites for *Eco*R1 (R), *Bam*H1 (B), *Sma*I (S) and *Dra*I (D) are shown. The *virB* sequences are shown as hatched box and the *virB-lacZ* fusion gene is shown as dark area. Amp^r and Tc^r indicate the ampicillin and tetracycline resistance genes, respectively.

A 42 bp *EcoRI*-*XbaI* fragment containing the -36 to -70 region of *virB* was cloned into *EcoRI*-*XbaI* cleaved pAD1084 (a pGEM1 derivative lacking the SP6 and T7 sequences) to yield plasmid pAD1209. A second plasmid pAD1207 contains *virB* sequences downstream from the -42 region and contains a unique *HindIII* site immediately upstream of the *virB* sequences. Both plasmids were linearized at the respective unique *HindIII* sites and either treated with S1 nuclease and ligated together (plasmid pAD1210) or filled-in with DNA polymerase I Klenow enzyme and dNTPs, and were ligated together (pAD1212).

Table 1. Requirements for *virB-lacZ* gene Expression

Strain	β -gal Activity, units	
	-AS	+AS
A136/pAD1092 (-Ti)	6.3	8.5
A348/pAD1092 (+Ti)	7.2	1210
A136/pGP159 (-Ti)	6.9	1828

Assays were performed in the absence (-) or presence (+) of acetosyringone (AS). -Ti, +Ti indicate strains that lack a Ti plasmid or contain the octopine Ti plasmid pTiA6, respectively.

For pAD1214 construction, pAD1207 was treated with *Hind*III followed by DNA polymerase 1 Klenow enzyme and dNTPs, and pAD1209 was treated with *Xba*I followed by S1 nuclease. The two DNAs were then ligated together. In all cases the orientation that places the *virB* upstream sequences of pAD1209 in front of the *virB* promoter of pAD1207 was selected. In the three plasmids pAD1210, 1212, and 1214 the upstream sequences are separated from the promoter sequences by 34, 38, and 10 bp, respectively. Also the -36 to -42 region is present as a tandem duplication in these plasmids.

Plasmid pAD1223 was constructed by a site-specific mutagenesis procedure. The *virB* sequences of pAD1092 were cloned as a 0.33 kb *Bam*HI fragment into the *Bam*HI site of pUC118. The resultant plasmid was introduced into *dut*⁻ *ung*⁻ *E. coli* CJ236 and single-stranded DNA containing uracil was prepared. The desired mutation was introduced using the site-specific mutagenesis procedure of Kunkel (27) using pdTGAAATCATCTAGAAGCAAT as a mutator. The mutated *virB* fragment was recloned into pAD1092 to yield pAD1223. The mutant was screened by scoring for the newly introduced *Xba*I restriction site.

Plasmids pAD1224, 1225 and 1226 are derivatives of pAD1223. pAD1223 DNA was first linearized by digestion with *Xba*I. The linear DNA was either treated with S1 nuclease and recircularized (pAD1224), or was treated with dNTPs and DNA polymerase 1 Klenow enzyme and recircularized (pAD1225), or was ligated with a synthetic *Xba*I-KpnI adaptor DNA dCTAGGTACCGGTAC (pAD1226). Plasmids pAD1224, 1225, and 1226 have altered spacings of -4, +4, and +10 bp between the -48 and upstream sequences.

All plasmids were confirmed by DNA sequence analysis of the relevant region by dideoxy chain termination method (28,29).

Construction of Deletion Mutations

Plasmid pAD1092 was linearized at its unique *Eco*RI site and treated with exonuclease *Bal*31 for limited period of time (30). The reaction was terminated by addition of ethylene glycol tetraacetic acid (25 mM, final concentration). The ends of *Bal*31 treated DNA were repaired by treatment with T4 DNA polymerase for 15 min at 37°C in the presence of dNTPs. The DNA was extracted with phenol and chloroform, precipitated by addition of ethyl alcohol, dried, and resuspended in TE (10 mM Tris-HCl, 1 mM EDTA, pH 8.0). 0.1–0.2 μ g of this DNA was ligated overnight with *Eco*RI linkers, pdGGAATTCC, in the presence of T4 DNA ligase. The DNA was digested with *Eco*RI, extracted with phenol and chloroform, religated to circularize, and used to transform *E. coli* DH5 α (F⁻, *endA*1, *hsdR*17, (*r*_k-*m*_k), *supE*44, *thi*-1, λ ⁻, *recA*1, *gyrA*96, ϕ 80*dlacZ*M15, *relA*1) to tetracycline resistance. Plasmid DNA was isolated from more than one hundred independent transformants (31). The extent of the deletion was determined, to a first approximation, by size determination of restriction fragments produced by double digestion of plasmid

Plasmid	<i>virB</i> sequences deleted, bp	DNA sequences Around Vector- <i>virB</i> Junction Region	β -gal Activity	
			-AS	+AS
		-108 -80 -70 -60		
WT	0	CGAATTCCC...27 bp...CGAAGGTTTTCGCTTCAATT GAAAT	7.2	1210
$\Delta 70$	32	CGGAATTCC-----GGTTTTCGCTTCAATT GAAAT	8.1	1537
$\Delta 38$	38	CGGAATTCC-----CGCTTCAATT GAAAT	15.3	1252
$\Delta 83$	41	CGGAATTCC-----TTCAATT GAAAT	60.0	3301
$\Delta 100$	46	CGGAATTCC-----TT GAAAT	5.6	7.0
$\Delta 46$	48	CGGAATTCC-----GAAAT	12.1	12.7
pAD1199	*	CGGAA-----GCTTCAACCGAAAT	6.7	207
pAD1199R	*	CGGTT-----GAAGCTTCCGAAAT	6.7	8.3
pAD1200	*	CGGAATTAA-----GCTTCAAAATTCGAAAT	3.7	4.6

Fig. 2. Expression of *virB-lacZ* Gene Fusion of Deletion Plasmids: DNA sequences around the vector-*virB* junction regions are shown. The deleted sequences are shown as a dotted line and the EcoRI linker DNA is shown in bold. Enzyme assays were performed in the absence (–) or presence (+) of acetosyringone (AS). Data shown are that obtained with *Agrobacterium* A348 containing the indicated plasmids. In the *Agrobacterium* A136 host all plasmids showed a basal level of expression (data not shown). Numbers at top indicate basepair position relative to the *virB* transcription initiation site (+1, ref. 21). An asterisk indicates that the plasmid is a derivative of $\Delta 46$.

DNAs with enzymes *EcoRI* and *BamHI*. Deletion endpoints of selected mutants were determined by DNA sequence analysis on double-stranded DNA template by dideoxy chain termination method (28,29). A synthetic oligodeoxynucleotide dTGTGACGATTATCGGGA (sequence complementary to the *virB* –17 to +1 region) was used as a sequencing primer.

Bacterial Conjugation, Induction, Enzyme Assay, and Others

All plasmids were mobilized from *E. coli* to *Agrobacterium* strains A136 or A348 by triparental mating procedures (32). *Agrobacterium* A136 is avirulent and lacks a Ti-plasmid, and *Agrobacterium* A348 contains the virulent octopine Ti plasmid pTiA6. Growth of agrobacteria and measurement of β -galactosidase enzyme activity were according to published procedures (11). All enzyme activities presented here are an average of three independent determinations. Sequencing primer, and other deoxyoligonucleotides were purchased from Microchemical Facilities, University of Minnesota and Genetic Designs, Inc., Houston, TX, respectively. Restriction and modifying enzymes, and EcoRI linker were purchased from New England Biolabs. Radioisotopes were purchased from Amersham Corporation.

Table 2. Effect of Deletion/Insertion Within the 5' Nontranscribed Region of *VirB* on *VirB-LacZ* Expression

Plasmid	Deletion/ Insertion, bp	β -gal Activity, units	
		–AS	+AS
pAD1092	0	5.1	1333
pAD1223	0	16.8	5816
pAD1224	–4	5.4	17
pAD1225	4	10.3	11
pAD1226	10	4.2	13
pAD1210	34	9.0	12
pAD1212	38	7.0	7
pAD1214	10	13.4	40

RESULTS

To monitor expression of the *Agrobacterium tumefaciens* octopine Ti plasmid pTiA6 *virB* operon we constructed a plasmid pAD1092 that contains a *virB-lacZ* translational fusion (Fig. 1). The fusion gene contains 108 bp upstream of the first *virB* transcription initiation site (21) and has the 54th codon of the first open-reading frame of *virB* fused to the 6th codon of *E. coli* β -galactosidase (*lacZ*) gene. A Ti-plasmidless *Agrobacterium* strain A136 containing the fusion plasmid pAD1092 showed low, basal level of β -galactosidase (β -gal) activity both in the absence and in the presence of the chemical inducer acetosyringone (Table 1, line 1). When the same plasmid was introduced into the Ti plasmid containing *Agrobacterium* A348 low level of *virB-lacZ* expression was observed in the absence of acetosyringone (line 2). In the presence of acetosyringone a 168 fold increase in β -gal activity was observed (line 2). This is similar to that observed with *virB-lacZ* fusions on a Ti-plasmid (9). This indicated that the 108 bp upstream sequences that are present in pAD1092 are sufficient for inducible expression of *virB*. To determine if *virA* and *virG* are the only Ti-plasmid encoded genes required for *virB-lacZ* expression we constructed plasmid pGP159. This plasmid contains the pTiA6 *virA*, *virG*, and the *virB-lacZ* gene fusion. In the presence of acetosyringone *Agrobacterium* A136/pGP159 showed a high level of β -gal activity (line 3) indicating that the two regulatory genes of pTiA6 were sufficient for *virB* expression. Similar results were obtained when *Nicotiana tabacum* culture cells were used as inducer (data not shown).

The *virB* sequences in plasmid pAD1092 are preceded by a unique *EcoR*I restriction site. To define the minimum sequences necessary for *virB* expression plasmid pAD1092 was linearized with restriction endonuclease *EcoR*I and treated briefly with the *Bal*31 exonuclease. After repairing the ends with T4 DNA polymerase and deoxynucleoside triphosphates the DNA was recircularized with T4 DNA ligase in the presence of an *EcoR*I linker DNA (dGGAATTCC). A large number of independent deletion derivatives were introduced into *Agrobacterium* A136 and A348, and the β -gal activities of the new strains were determined in the absence and presence of acetosyringone. The deletion endpoints of the mutants were determined by DNA sequence analysis (28,29). Results obtained with five selected mutants are presented in Fig. 2 (lines 2–6). These mutants have lost between 32 and 48 bp of *virB* upstream sequences. Deletions that had lost sequences up to residue –68 (plasmids Δ 70, Δ 38, and Δ 83) showed high level of inducible β -gal activity (Fig. 2, lines 2–4). The induced levels were comparable to that observed with strains containing the parent plasmid pAD1092 (line 1). One mutant, Δ 83, consistently exhibited a higher level of expression both in the absence and in the presence of acetosyringone (line 4), and in a *virA-virG* dependent manner (data not shown). Whether the increased expression is due to vector DNA sequences that were placed adjacent to the –68 region as a consequence of deletion mutagenesis is currently under investigation. Deletions that had lost sequences up to residue –63 and beyond (Δ 100, Δ 46) showed no inducible β -gal expression (lines 5,6). Therefore, sequences required for inducible expression of *virB* must lie within the –68 region. As promoter sequences in bacteria lie within the –35 region these sequences must constitute the upstream regulatory sequences (33). That the difference in β -gal. activity in the mutant plasmids were not due to difference in plasmid copy number was confirmed by plasmid copy number determination (data not shown).

Comparison of DNA sequences around vector-*virB* junction regions of the various deletion mutants indicate that the 'dCTTCAA' sequence at positions –68 to –63 is required for *virB* expression. To determine if these sequences are sufficient for *virB* expression we

Sequence #	Locus	Relevant DNA Sequence																														
		-15	-10	-5	.	5	10	15																								
1	<u>virA</u>	T	C	T	T	C	T	G	T	T	C	A	C	T	T	G	A	A	A	C	A	A	C	T	G	A	A	-59				
2	<u>virB</u>	A	G	G	T	T	T	C	G	C	T	T	C	A	A	T	T	G	A	A	A	T	C	A	T	A	A	G	A	-47		
3		A	T	C	A	T	A	A	G	A	A	G	C	A	A	T	T	G	A	A	A	T	T	T	T	C	G	A	G	-28		
4	<u>virC</u>	A	A	A	A	A	C	C	G	A	A	T	A	T	A	T	T	G	C	T	A	C	A	A	T	T	G	C	A	A	-48	
5		A	T	A	T	A	A	T	T	G	C	T	A	C	A	A	T	T	G	C	A	A	T	T	T	A	A	T	T	A	T	-39
6		T	T	G	C	T	A	C	A	A	T	T	G	C	A	A	T	T	T	A	A	T	T	A	T	A	A	C	G	A	A	-33
7		A	C	A	A	T	T	G	C	A	A	T	T	T	A	A	T	T	A	T	A	A	C	G	A	A	T	T	T	G	A	-28
8	<u>virD</u>	T	T	C	G	T	T	A	T	A	A	T	T	A	A	T	T	G	C	A	A	T	T	G	T	A	G	C	A	A	-67	
9		A	T	A	A	T	T	A	A	T	T	G	C	A	A	T	T	T	A	G	C	A	A	T	T	A	T	A	T	A	T	-61
10		T	T	G	C	A	A	T	T	G	T	A	G	C	A	A	T	T	A	T	T	C	G	G	T	T	T	T	T	T	-52	
11	<u>virE</u>	G	C	C	T	G	C	T	C	A	T	T	G	C	A	G	T	T	G	A	A	C	G	C	G	A	T	A	T	C	-47	
12	<u>virG</u>	C	T	C	G	A	A	G	C	G	T	T	T	C	A	C	T	T	G	T	A	A	C	A	C	G	A	T	T	G	-43	
13		C	A	C	T	T	G	T	A	A	C	A	A	C	G	A	T	T	G	A	G	A	A	T	T	T	T	G	T	C	-31	
14		A	T	T	T	T	G	T	C	A	T	A	A	A	A	T	T	G	A	A	A	T	A	C	T	T	G	G	T	T	-10	
	A	7	2	4	4	4	5	3	4	7	5	3	4	2	13	11	0	0	2	7	12	11	2	6	6	4	5	3	2	7	6	
	C	2	3	5	2	1	2	3	4	1	3	0	0	10	2	0	0	0	3	0	0	6	2	2	1	1	2	2	0	2		
	G	1	1	3	2	1	1	3	1	6	0	0	5	0	1	1	0	0	11	0	1	1	0	2	2	2	1	3	5	2	2	
	T	4	8	2	6	8	6	5	5	0	6	11	5	2	0	0	14	14	1	4	1	2	6	4	4	7	7	6	5	5	4	
	Consensus	A	T	C	T	T	T																	A	A	T	T	G	A	A		
		-----N	HPu																													
		T	A	A	A	A	A																									

Fig. 3. Sequence Homology in the Upstream Regions of *Vir* Genes: Partial sequences of the *vir* gene nontranscribed regions are according to Das et al. (21). Numbers on the right indicates the position of the rightmost residue relative to the transcription initiation site (+1) defined by Das et al. (21). Numbers at top indicate nucleotide position from the center of symmetry (●). The numbers at lines 15–18 indicate frequency of bases at a particular position. Pu: Purine, Py: Pyrimidine, D: A, G or T, H: A, C or T.

used oligonucleotide directed mutagenesis to alter two basepairs downstream from this hexameric sequence. A plasmid pAD1199, that contains the dCTTCAA sequence at the native positions and a TT → CC substitution at positions –62 and –61, was constructed by modifying the unique *Eco*R1 site of plasmid Δ46 (see methods). Analysis of β-gal activity showed that reintroduction of the dCTTCAA sequence into Δ46 restored inducible β-gal expression of *virB-lacZ* but the induced level was 7 fold lower than that observed with the wild type plasmid (Fig. 2, line 7). Since sequences upstream of –68 region did not significantly reduce *virB* induction the low level of induction in pAD1199 must be due to the TT → CC substitutions at positions –62 and –61. Therefore, the right boundary of the upstream regulatory sequences must extend at least up to residue –61. Insertion of the dCTTCAA sequence in the opposite orientation (pAD1199R) at positions –71 to –66 or in the native orientation at positions –72 to –67 (pAD1200) did not restore inducibility of the *virB* promoter (lines 8,9).

The sequences dCAATTGAAA at positions –65 to –57 ('box A' sequence) are also present at position –45 to –37 ('box B' sequence). The two regions are separated by 11 bp which will constitute approximately a single helical turn of the DNA double helix. To determine if the spacing between the box A and box B sequences could be altered we first constructed plasmid pAD1223 which contains AA → CT substitutions at positions –52 and –51. These substitutions create a unique *Xba*I restriction site that could be used to construct mutants with altered spacing between box A and box B sequences. Analysis of *virB-lacZ* expression showed that in the presence of acetosyringone plasmid pAD1223 exhibited a 4.4 fold higher induced level of β-gal activity compared to plasmid pAD1092 (Table 2, line 2). A 3 fold increase in *lacZ* expression was also observed in the absence of acetosyringone. To alter spacing between the box A and box B sequences the *Xba*I

site of pAD1223 was modified to produce three derivatives pAD1224, pAD1225 and pAD1226. Plasmids pAD1224, pAD1225, or pAD1226 contains a deletion of 4 bp, an insertion of 4 bp, or an insertion of 10 bp, respectively, between the box A and box B sequences. When *Agrobacterium* A348 containing these plasmids were assayed for *virB-lacZ* expression none of them showed any inducible β -gal synthesis (Table 2, lines 3–5). This indicates that neither the position of the box A sequence relative to the promoter region nor the spacing between the box A and box B sequences (if the box B sequence is required for *virB* expression) could be altered without deleterious effect.

We then sought to determine whether both box A and box B sequences together could be moved further upstream from the promoter (–35 and –10) sequences. We constructed three mutants, plasmids pAD1210, pAD1212, and pAD1214, that had insertions of 34, 38, and 10 bp, respectively, between the –70 to –36, and the –42 to downstream sequences (see methods for details). In these plasmids the –36 to –42 sequences are duplicated immediately upstream of the native sequences. While two of the mutants, pAD1210 and pAD1212, showed no inducible β -gal synthesis the third one, pAD1214, showed a low level (3 fold) of induction in the presence of acetosyringone (Table 2, lines 6–8). Since both box A and box B sequences are present in all these mutants and the sequences up to –42 are identical to the wild type sequence these results indicate that for maximal expression of *virB* the upstream regulatory sequences must lie at a specific location relative to the promoter region.

DISCUSSION

The *virA* and *virG* loci of the *Agrobacterium* Ti plasmid in conjunction with the plant cell factor acetosyringone positively control *vir* gene expression (7–9). According to a model proposed by Winans et al. (12) interaction with acetosyringone modulates VirA to activate VirG. The activated VirG positively controls *vir* gene expression.

The deletion mutagenesis studies presented here demonstrate that sequences within the –68 region of *virB* are sufficient for its expression (Fig. 2). Within this region lies a nonameric sequence (box A and box B sequences) that is present in two copies in the *virB* upstream region. This region is a subset of a tetradecameric sequence, dPuT/ATDCAATTGHAAPy (D=A,G or T; H=A, C or T), Vir Box, that is conserved in all *vir* promoters (Fig. 3 and ref. 33). This region also contains a dyad-symmetry and is flanked by highly AT rich sequences. The central core region of the tetradecamer is the most conserved region and contains the hexamer dGCAATT we postulated in an earlier study (21) to play a role in *vir* gene regulation.

Analysis of mutant $\Delta 46$ show that deletion of the left-half of the *vir* box completely abolished *virB* expression. Reinstating these sequences (pAD1199) restored *virB* expression (Fig. 2). The mutant plasmid pAD1199, however, exhibited a 7 fold lower inducible expression which appears to be due to a TT $\rightarrow$ CC transition at the most conserved positions 8 and 9 of the *vir* box sequence. As these transitions affect the dyad-symmetry of the *vir* box it is likely that the mutant contains a poor site for interaction with the transcription activator VirG protein. We have recently shown that *in vitro* VirG binds specifically at the *vir* box sequence of *virB* (Pazour and Das, submitted). One mutant, pAD1200, contains a 4 bp insertion within the *vir* box. The insertion sequences (dAATT), however, do not affect the dyad symmetry, yet this mutant has a null phenotype. Therefore, both primary sequence and the dyad symmetry of the *vir* box must be important for *vir* gene expression. The presence of the dyad symmetry suggests that the regulatory sequences are composed

of two half-sites each of which interacts with a protein molecule. Consequently a dimer or multimer of VirG is expected to be the active form of the protein. In *E. coli* the cyclic-AMP dependent activator protein interacts with a conserved sequence having a dyad symmetry (35).

Insertion and deletion mutagenesis studies indicate that the *vir* box must lie at a specific position relative to the promoter region for maximal expression of *virB* (Tables 3,4). Two sets of these sequences are present in the *virB* upstream region. Whether both copies are required for expression is not clear at present. At least two *vir* genes, *virA* and *virE*, contain a single copy of these sequences (Fig. 3). Both these genes are inducible and *virE* is induced to a very high level (9). Multiple copies of the *vir* box are present in *virC*, *virD* and *virG*. In these cases which copy(copies) is biologically active is not known.

The presence of the *vir* box at different locations apparently contradicts our observations that in *virB* the sequences must lie at a specific location relative to the promoter sequences. An analysis of the conserved sequences indicates that this may not be true. In four of the six *vir* genes, *virB*, *virC*, *virD*, and *virE*, there is one copy of the sequence present in the same relative position (sequence no. 2, 4, 10, and 11 in Fig. 3). In one case, *virA*, this sequence is located a single helical turn away from the position of the other *vir* box sequences while in *virG* the sequences are closer to the promoter sequence. Coincidentally, *virA* and *virG* are induced to a low level. *VirA* is induced 9 fold (11) compared to a 15, and 30–90 fold induction observed with *virG*, and *virB*–*E*, respectively (8). The low level of *virA* induction may be due to the distal positioning of the *vir* box sequence. One of our mutants supports this hypothesis. Mutant plasmid pAD1214, where the *vir* box was placed 10 bp (approx. 1 helical turn) away from its native position, showed a markedly reduced (3 fold vs. 135 fold for wild type) level of induction. For *virG* the converse may lead to its low level of induction. The requirement that the *vir* box be present at a particular location suggests that the regulatory protein bound at the *vir* box must interact with a RNA polymerase molecule for activation of *vir* gene transcription. Future experiments will test this hypothesis.

ACKNOWLEDGEMENTS

We thank Mark Mersereau for excellent technical assistance, Bonnie Allen for typing of the manuscript, Malcolm Casadaban for a gift of pMC1403, and John Larkin for a critical reading of the manuscript. This work was supported by a grant from the National Institute of Health (GM 37555). AD is the recipient of a Junior Faculty Research Award of the American Cancer Society (JFRA-170) and a McKnight Land-Grant Professorship of the University of Minnesota.

*To whom correspondence should be addressed

REFERENCES

1. Hooykaas, P. and Schilperoort, R. (1984) *Adv. Genetics* 22:210–283.
2. Nester, E., Gordon, M., Amasino, R. and Yanofsky, M. (1984) *Ann. Rev. Plant Physiol.* 35:387–413.
3. Garfinkel, D. and Nester, E. (1980) *J. Bacteriol.* 144:732–743.
4. Ooms, G., Kalpwijk, P., Poulis, J. and Schilperoort, R. (1980) *J. Bacteriol.* 144:82–91.
5. Douglas, R., Staneloni, R., Rubin, R. and Nester, E. (1985) *J. Bacteriol.* 161:850–860.
6. Stachel, S., Nester, E. and Zambryski, P. (1986) *EMBO J.* 5:1445–1454.
7. Stachel, S., Nester, E. and Zambryski, P. (1986) *Proc. Natl. Acad. Sci. USA* 83:379–383.
8. Stachel, S. and Zambryski, P. (1986) *Cell* 46:325–333.
9. Stachel, S., Messens, E., van Montagu, M. and Zambryski, P. (1985) *Nature (London)* 318:624–629.

10. Rogowsky, P., Close, T., Chimera, J., Shaw, J. and Kado, C. (1987) *J. Bacteriol.* **169**:5101–5112.
11. Winans, S., Kerstetter, R. and Nester, E. (1988) *J. Bacteriol.* **170**:4047–4054.
12. Winans, S., Ebert, P., Stachel, S., Gordon, M. and Nester, E. (1986) *Proc. Natl. Acad. Sci. USA* **83**:8278–8282.
13. Leroux, B., Yanofsky, M., Winans, S., Ward, J., Ziegler, S. and Nester, E. (1987) *EMBO J.* **6**:849–856.
14. Nixon, B., Ronson, C. and Ausubel, F. (1986) *Proc. Natl. Acad. Sci. USA* **83**:7850–7854.
15. Ronson, C., Nixon, B. and Ausubel, F. (1987) *Cell* **49**:579–581.
16. Ninfa, A. and Magasanik, B. (1986) *Proc. Natl. Acad. Sci. USA* **83**:5909–5913.
17. Hess, J.F., Oosawa, K., Matsumura, P. and Simon, M. (1987) *Proc. Natl. Acad. Sci. USA* **84**:7609–7613.
18. Ninfa, A., Ninfa, E., Lupas, A., Stock, A., Magasanik, B. and Stock, J. (1988) *Proc. Natl. Acad. Sci. USA* **85**:5492–5496.
19. Hess, J.F., Oosawa, K., Kaplan, N. and Simon, M. (1988) *Cell* **53**:79–87.
20. Igo, M. and Silhavy, T. (1988) *J. Bacteriol.* **170**:5971–5973.
21. Das, A., Stachel, S., Ebert, P., Allenza, P., Montoya, A. and Nester, E. (1986) *Nucleic Acids Res.* **14**:1355–1364.
22. Ward, J., Akiyoshi, D., Regier, D., Datta, A., Gordon, M. and Nester, E. (1988) *J. Biol. Chem.* **263**:5804–5814.
23. Casadaban, M., Chou, J and Cohen, S. (1980) *J. Bacteriol.* **143**:971–980.
24. Schmidhauser, T and Helsinki, D. (1985) *J. Bacteriol.* **164**:446–455.
25. Engler, G., Depicker, A., Maenhaut, R., ViMarroel-Mandiola, R., van Montagu, M. and Schell, J. (1981) *J. Mol. Biol.* **152**:183–208.
26. Vieira, J. and Messing, J. (1987) *Methods Enzymol.* **15**:3–11.
27. Kunkel, T. (1985) *Proc. Natl. Acad. Sci. USA* **82**:488–492.
28. Sanger, F., Nicklen, S. and Coulson, A. (1977) *Proc. Natl. Acad. Sci. USA* **74**:5463–5467.
29. Garber, R. and Gaiser, C.J. (1986) *Promega Notes, Promega Biotec. Madison, WI.* #4, pp. 3–5.
30. Gray, H., Ostrander, D., Hodnett, J., Legerski, R. and Robertson, D. (1975) *Nucleic Acids Res.* **2**:1459–1492.
31. Birnboim, H. and Doly, J. (1979) *Nucleic Acids Res.* **7**:1513–1523.
32. Ditta, G., Stanfield, S., Corbin, D. and Helinski, D. (1980) *Proc. Natl. Acad. Sci. USA* **77**: 7347–7351.
33. McClure, W. (1985) *Ann. Rev. Biochem.* **54**:171–204.
34. Winans, S., Jin, S., Komari, T., Johnson, K. and Nester, E. (1987) in *Plant Molecular Biology*, von Wettstein, D. and Chua, N.-H. , eds. Plenum Press, NY, pp. 573–582.
35. deCrombrughe, B., Busby, S. and Buc, H (1984) *Science* **224**:831–837.

**This article, submitted on disc, has been automatically
converted into this typeset format by the publisher.**