Proc. Natl. Acad. Sci. USA
Vol. 93, pp. 12400–12405, October 1996
Genetics

Intron insertion facilitates amplification of cloned virus cDNA in *Escherichia coli* while biological activity is reestablished after transcription *in vivo*

(pea seedborne mosaic potyvirus/infectious clone)

I. ELISABETH JOHANSEN†

Biotechnology Group, Danish Institute of Plant and Soil Science, Lottenborgvej 2, DK-2800 Lyngby, Denmark

Communicated by George Bruening, University of California, Davis, CA, August 7, 1996 (received for review March 4, 1996)

ABSTRACT **Insertion of introns into cloned cDNA of two isolates of the plant potyvirus pea seedborne mosaic virus facilitated plasmid amplification in *Escherichia coli*. Multiple stop codons in the inserted introns interrupted the open reading frame of the virus cDNA, thereby terminating undesired translation of virus proteins in *E. coli*. Plasmids containing the full-length virus sequences, placed under control of the cauliflower mosaic virus 35S promoter and the nopaline synthase termination signal, were stable and easy to amplify in *E. coli* if one or more introns were inserted into the virus sequence. These plasmids were infectious when inoculated mechanically onto *Pisum sativum* leaves. Examination of the cDNA-derived viruses confirmed that intron splicing of *in vivo* transcribed pre-mRNA had occurred as predicted, reestablishing the virus genome sequences. Symptom development and virus accumulation of the cDNA derived viruses and parental viruses were identical. It is proposed that intron insertion can be used to facilitate manipulation and amplification of cloned DNA fragments that are unstable in, or toxic to, *E. coli*. When transcribed *in vivo* in eukaryotic cells, the introns will be eliminated from the sequence and will not interfere with further analysis of protein expression or virus infection.**

In vitro- or *in vivo*-transcribed infectious RNA derived from full-length cDNA clones is an important tool in the study of RNA viruses. Various aspects of virus pathogenicity have been studied by this technique: symptom development, cell-to-cell and long distance movement, seed and aphid transmission, interaction with resistance genes, replication, proteolysis, and disassembly (1–8). Reporter genes have been inserted into full-length clones to visualize virus spread and accumulation following infection (9, 10). In addition, viruses have been used as vectors for *in vivo* expression of foreign proteins (11, 12).

However, manipulation and amplification of full-length clones may prove difficult due to instability and/or toxicity of some virus sequences in bacteria (13–16). Expression of virus proteins in *Escherichia coli* has been reported to have toxic effects on the host cells (17, 18). It is not known how virus proteins are expressed from full-length clones, since the vector sequences do not contain promoters expected to transcribe the virus RNA in bacterial cells. Fakhfakh *et al.* (14) suggested that virus RNA is transcribed from cryptic promoters and protein synthesis initiated at cryptic ribosomal binding sites present in the virus cDNA sequences. The toxic effects from undesired protein expression can be relieved by cloning in *E. coli* strains that reduce the plasmid copy number (19) or using low copy number cloning vectors (20).

Infectious RNA is obtained from full-length cDNA clones either by *in vitro* transcription of RNA usually from a phage promoter or by *in vivo* transcription from a strong constitutive promoter active in the host cell (13). The latter approach is attractive because it bypasses the need for *in vitro* transcription and incorporation of cap analogue.

Pea seedborne mosaic virus (PSbMV) belongs to the potyvirus group. The potyvirus genome is a single-stranded positive sense RNA of ≈10 kb, encoding a single polyprotein that is processed by three virus proteases (21). PSbMV isolates are divided into three pathotypes—P-1, P-2, and P-4—according to their interaction with four recessive resistance genes in *Pisum sativum* (22). The complete nucleotide sequence of the pathotype P-1 isolate DPD1 and the pathotype P-4 NY isolate have been determined (23, 24)

This paper describes assembly of infectious full-length clones of these two PSbMV isolates in which the open reading frame of the virus sequence is interrupted at one or more positions by intron insertion.

MATERIALS AND METHODS

cDNA, Intron, and Vector Modifications. Synthesis and cloning of cDNA of the two PSbMV isolates P-1 DPD1 (P1) and P-4 NY (P4) were described elsewhere (23, 24). A cistron map of PSbMV is shown in Fig. 1*A*. Locations of natural and engineered restriction sites in the cDNA of P1 and P4 are identified by the first nucleotide of the recognition sequence. cDNA of the 5′ termini of P1 and P4 covering nucleotides 1–*Sph*I 965 and 1–*Sph*I 924, respectively, were amplified by reverse transcription and PCR (25), with primers that added *Dra*I and *Xba*I sites at the 5′ and 3′ ends, respectively. The *Dra*I site allowed precise excision of the virus 5′ termini, and the *Xba*I site facilitated subsequent assembly of full-length clones. A unique *Xba*I site was introduced 3′ to the cDNA covering the virus poly(A) tails by digesting the restriction site of the cloning vector closest to the poly(A) tail, removing nucleotide overhangs by mung bean nuclease and inserting a *Xba*I linker. The cDNA was further modified by site-directed *in vitro* mutagenesis on single-stranded DNA templates (Sculptor; Amersham). Five *Pst*I sites at nucleotides 576, 1044, 3252, 3364, and 3428 were engineered into the sequence of P1. The P4 sequence was modified by introducing a *Bam*HI site at nucleotide 2259 and removing a *Hin*dIII site at nucleotide 7626. All nucleotide substitutions were silent except the substitution at the *Pst*I 3428 site, which resulted in a Gly-to-Cys codon change. A plasmid, pP4(2259-6168)-VΔ*Pst*I, covering nucleotides 2259-6168 of P4, with a 4-nt deletion at nucleotide 3366, was obtained by removal of the 3′ extensions of the *Pst*I site at nucleotide 3366 by T4 DNA polymerase digestion.

DNA fragments representing intron IV2 (189 bp) of the ST-LS1 gene from *Solanum tuberosum* (26) and introns 2 (221

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: PSbMV, pea seedborne mosaic virus; CaMV, cauliflower mosaic virus; T-NOS, nopaline synthase termination signal.

†e-mail: e.johansen@dips-lyngby.dk.

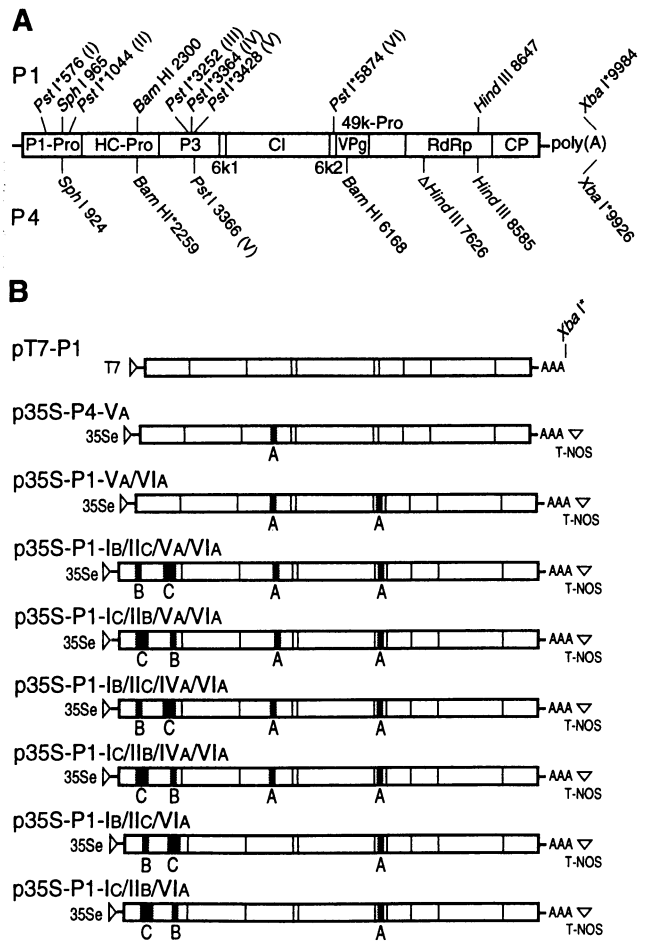

FIG. 1. (*A*) Cistron map of PSbMV showing the noncoding regions (solid line), the open reading frame encoding the potyvirus polyprotein (open box) and putative proteolytic cleavage sites (vertical lines). Restriction sites in the cDNA are shown above (P1) and below (P4) the cistron map. Restrictions sites marked with * were engineered into the sequence, whereas the *Hin*dIII site marked with Δ was removed from the P4 cDNA sequence. The *Pst*I sites marked (I)–(VI) were used for intron insertion. P1-Pro, P1 protease; HC-Pro, helper component-protease; P3, P3 protein; 6k1, 6-kDa protein 1; CI, cylindrical inclusion protein; 6k2: 6-kDa protein 2; 49k-Pro, 49-kDa protease; VPg, genome-linked virus protein; RdRp; putative RNA-dependent RNA polymerase; CP, coat protein. (*B*) Full-length cDNA clones of P1 and P4 from which infectious RNA was derived. T7, T7 RNA polymerase promoter; 35Se, enhanced CaMV 35S promoter; A, intron A; B, intron B; C, intron C.

bp) and 3 (509 bp) of the NiR gene from *Phaseolus vulgaris* (27) were engineered by PCR to have *Pst*I and *Nsi*I sites at the 5′ and 3′ ends, respectively. The introns will be identified as A, B, and C, as summarized in Table 1. The PCR products were digested with *Pst*I and *Nsi*I and inserted into *Pst*I sites in the cDNA of P1 and P4 (*Pst*I and *Nsi*I restriction sites have common 3′ TGCA extensions). The positions of the *Pst*I sites in the cDNA sequence of P1 and P4, which were used for intron insertion are numbered I–VI (Table 2 and Fig. 1*A*).

Table 1. Origin and size of introns inserted in full-length clones of PSbMV

Intron	Origin, intron/gene/species	Size, bp	Ref.
A	IV/ST-LS1/*S. tuberosum*	189	26
B	2/NiR/*Phaseolus vulgaris*	221	27
C	3/NiR/*Phaseolus vulgaris*	509	27

Table 2. Numbering conventions and positions of intron insertion in the genome of PSbMV P1 and P4

Position no.	Nucleotide position in P1	Nucleotide position in P4	Gene product interrupted
I	576		P1 protease
II	1044		P1 protease
III	3254		P3 protein
IV	3364		P3 protein
V	3428	3366	P3 protein
VI	5874		6-kDa protein 2

The nucleotide position identifies the first nucleotide of the *Pst*I site in which the introns were inserted.

The plasmid vectors pT7E19(+) (28) and pAGUS1 (29) were used to assemble the full-length clones. pT7E19(+) has a *Sac*I site and pAGUS1 has a *Bam*HI site, which allow cloning directly at the transcription initiation site of the bacteriophage T7 promoter and the enhanced cauliflower mosaic virus (CaMV) 35S promoter, respectively. pT7E19(+) was modified to remove the *Hin*dIII site from the polylinker, and the *Sac*I site of pAGUS1 was replaced with *Xba*I. The restriction sites in the cDNA sequence of P1 and P4, which were used to assemble the full-length clones, are shown on the cistron map (Fig. 1*A*).

Assembly of P1 cDNA in pT7E19(+). pT7E19(+) was digested with *Sac*I, T4 DNA polymerase, and *Xba*I, and the 5′ terminus of P1 was inserted as a *Dra*I–*Xba*I fragment behind the T7 promoter. This plasmid was digested with *Sph*I and *Xba*I, and cloned cDNA covering *Sph*I 965 to *Bam*HI 2300 was inserted as a *Sph*I–*Xba*I fragment, the *Xba*I site originating from the polylinker of the cloning vector, generating pT7-P1(1-2300). Cloned cDNA covering *Hin*dIII 8647 to the poly(A) tail was inserted 3′ to *Bam*HI 2300 of pT7-P1(1-2300) as a *Bam*HI–*Xba*I fragment, the *Bam*HI site originating from the polylinker of the cloning vector, resulting in pT7-P1(1-2300/8647-A). Full-length pT7-P1 was assembled by inserting a cDNA fragment covering *Bam*HI 2300 to *Hin*dIII 8647 into *Bam*HI–*Hin*dIII-digested pT7-P1(1-2300/8647-A).

Assembly of P4 cDNA in pAGUS1. pAGUS1 was digested with *Bam*HI, mung bean nuclease, and *Xba*I and the 5′ terminus of P4 was inserted as a *Dra*I–*Xba*I fragment behind the enhanced CaMV 35S promoter. This plasmid was digested with *Sph*I and *Xba*I, and cloned cDNA covering *Sph*I 924 to *Bam*HI 2259 was inserted as a *Sph*I–*Xba*I fragment, the *Xba*I site originating from the polylinker of the cloning vector, generating p35S-P4(1-2259). cDNA of nucleotide 6168 to the poly(A) tail was inserted as a *Bam*HI–*Xba*I fragment into *Bam*HI–*Xba*I-digested p35S-P4(1-2259) creating p35S-P4(1-2259/6168-A). Intron A was inserted at position V of pP4(3366-6168) generating pP4(3366-6168)-VA. pP4(2259-6168)-VA was subsequently assembled from pP4(2259-3366) and pP4(3366-6168)-VA. The full-length clone p35S-P4-VA was assembled by inserting the *Bam*HI fragment of pP4(2259-6168)-VA into *Bam*HI-digested p35S-P4(1-2259/6168-A).

Assembly of P1 cDNA in pAGUS1. The initial steps of P1 assembly were identical to P4 creating p35S-P1(1-2300) analogous to p35S-P4(1-2259). Cloned cDNA of nucleotide 8647 to the poly(A) tail was inserted into p35S-P1(1-2300) as a *Bam*HI–*Xba*I fragment, resulting in p35S-P1(1-2300/8647-A). Introns B and C were inserted at position I and/or II of p35S-P1(1-2300/8647-A) in different combinations, creating p35S-P1(1-2300/8647-A)-IB, p35S-P1(1-2300/8647-A)-IIB, p35S-P1(1-2300/8647-A)-IB/IIC, and p35S-P1(1-2300/8647-A)-IC/IIB. Intron A was inserted at position III, IV, V, and/or VI of P1(2300-8647) to create pP1(2300-8647)-IIIA, pP1(2300-8647)-IVA, pP1(2300-8647)-VA, pP1(2300-8647)-VIA, pP1(2300-8647)-IVA/VIA, and pP1(2300-8647)-VA/

VIA. Full-length p35S-P1 clones with different intron combinations (listed in Table 3) were assembled by insertion of the *Bam*HI 2300 to *Hin*dIII 8647 fragments of pP1(2300-8652)-"XN" clones into *Bam*HI-*Hin*dIII-digested p35S-P1(1-2300/8647-A)-"XN" clones.

Amplification and Inoculation of Full-Length Clones. Plasmids were propagated in *E. coli* strain SURE (Stratagene). Clones of pT7-P1 and p35S-P1-VA/VIA were grown in small volumes (3.5 ml), whereas all other clones were grown in 50-ml cultures. Plasmid yield was estimated by comparing the intensity of linear plasmid to *Hin*dIII-digested λ marker (Boehringer Mannheim) after electrophoresis in ethidium bromide-stained agarose gels (30). Capped transcripts were synthesized from pT7-P1 using 2 μg of cesium chloride-purified, *Xba*I-digested plasmid DNA as template following the mMESSAGE mMACHINE T7 kit protocol recommended by the manufacturer (Ambion, Austin, TX). Transcripts were diluted to 100 ng/μl in 3 mM sodium phosphate, 0.5 mg of bentonite per ml, and 0.5 mg of carborundum per ml (31), and 2 μg of transcript was inoculated manually to two leaves of each healthy *P. sativum* plant. p35S full-length clones were diluted in distilled H_2O, and 500, 250, 100, 50, 25, or 10 ng of plasmid was inoculated to healthy *P. sativum* plants with the aid of carborundum.

Biological activity of *in vitro*- and *in vivo*-transcribed RNA in *P. sativum* was determined 2 weeks after inoculation by assaying plants for the presence of PSbMV coat protein by ELISA using a PSbMV-specific polyclonal antiserum (25). Symptom induction and virus accumulation of cDNA-derived PSbMV P1 and P4 were compared with native PSbMV P1 and P4 in *P. sativum* cultivars Dark Skinned Perfection, Fjord, and Vedette and in *Chenopodium quinoa*. Virus accumulation was quantitated by ELISA in six plants of each host–virus combination 3 weeks after inoculation on leaf extracts of systemically infected leaves of *P. sativum* and inoculated leaves of *C. quinoa*.

Table 3. Plasmid amplification of selected subclones and of full-length clones of PSbMV P1 and P4

Plasmid	DNA, μg per ml of culture
Subclones	
pP4(2259-3366)	10–15
pP4(3366-6168)	25–35
pP4(2259-6168)	NA
pP4(2259-6168)-VΔPstI	10–15
pP4(2259-6168)-VA	15–20
pP1(494-965)	2.5–7.5
Full-length clones	
pT7-P1*	0.2–0.3
p35S-P4-VA*	12.5–17.5
p35S-P1	Unstable
p35S-P1-IIIA	NA
p35S-P1-IVA	NA
p35S-P1-VA	NA
p35S-P1-VIA	NA
p35S-P1-VA/VIA*	0.1–0.2
p35S-P1-IB/IIC	NA
p35S-P1-IB/VA/VIA	NA
p35S-P1-IIB/VA/VIA	NA
p35S-P1-IB/IIC/VA/VIA*	20–25
p35S-P1-IC/IIB/VA/VIA*	15–20
p35S-P1-IB/IIC/IVA/VIA*	17.5–22.5
p35S-P1-IC/IIB/IVA/VIA*	15–20
p35S-P1-IB/IIC/VIA*	15–20
p35S-P1-IC/IIB/VIA*	15–20
p35S-P1-IB/IIC/IIIA	Unstable
p35S-P1-IB/IIC/IVA	Unstable

Plasmid amplification was quantitated by the intensity of restriction enzyme-digested plasmid on ethidium bromide-stained agarose gels. Plasmids were purified from *E. coli* strain SURE and yield was determined on four cultures of each plasmid. NA, not amplified.
*Transcripts derived from these clones were infectious on *P. sativum*.

Virus Purification and Analysis of Intron Splicing. Virus particles were purified from infected *P. sativum* plants inoculated with p35S-P4-VA, p35S-P1-IB/IIC/VA/VIA, or p35S-P1-IC/IIB/VA/VIA (22), and virus RNA was extracted from the purified virus particles (32). The RNA was used as template for reverse transcription and PCR amplification of cDNA fragments covering nucleotides 2259-4744 of p35S-P4-VA-derived P4; and nucleotides 1-2760 p35S-P1-IB/IIC/VA/VIA, and p35S-P1-IC/IIB/VA/VIA-derived P1. The cDNA sequences surrounding the intron splice sites were determined using Sequenase 2.1 (United States Biochemical) after subcloning *Sph*I 3274 to *Eco*RI 4054 of P4, *Eco*RI 494 to *Sph*I 965, and *Sph*I 965 to *Eco*RI 1600 of P1 in pUC18 or pUC19.

RESULTS

Amplification of Cloned PSbMV cDNA in *E. coli*. During the sequence analysis of PSbMV P1 and P4 cDNA (23, 24), it was observed that the plasmid yield after amplification in *E. coli* depended on the particular PSbMV cDNA sequence carried by the cloning vector (Table 3). Some plasmids were difficult to amplify and colonies produced by *E. coli* carrying these plasmids were small. It was observed that plasmid yield was often higher if the open reading frame of the virus sequence was oriented opposite the open reading frame of the *lac*Z gene of the cloning vector. Also, the *E. coli* strain SURE was generally more efficient for amplification of difficult plasmids compared with the strains TG-1, DH5α, and XL1-blue (data not shown). One plasmid, pP4(2259-6168), covering nucleotides 2259-6168 of P4 in pUC19, appeared to be lethal to the host cells at high copy numbers, since no colonies were observed after transformation of the *E. coli* strain SURE with pP4(2259-6168). Reduction of plasmid copy number should reduce toxicity to the host cell. Therefore, the *E. coli* strain ABLE K, which reduces the plasmid copy number 10-fold (19), was transformed with pP4(2259-6168), and nucleotides 2259-6168 were cloned in runaway replication vector pOU61, which is only present in one copy per cell at restrictive temperatures (33). Both approaches resulted in formation of normal-sized colonies, but plasmids containing nucleotides 2259-6168 of P4 could not be recovered from liquid cultures of these clones (data not shown). The two plasmids pP4(2259-3366) and pP4(3366-6168), which were used to assemble pP4(2259-6168), were easy to amplify in *E. coli* (Table 3). To determine if the apparent toxicity of the pP4(2259-6168) could be caused by translation of a virus protein product, the virus open reading frame was interrupted by making a 4-nt deletion at position V. In the resulting plasmid pP4(2259-6168)-VΔ*Pst*I, the 4-nt deletion changes the open reading frame of the P4 sequence, and, in consequence, a stop codon was reached five codons downstream position V in the new frame. Transformants containing pP4(2259-6168)-VΔ*Pst*I formed normal-sized colonies, and the plasmid was easy to amplify (Table 3), suggesting that a protein product translated from pP4(2259-6168) was actually toxic to the *E. coli* cells.

Amplification of P1 Full-Length Clones pT7-P1 and p35S-P1. A full-length clone of PSbMV P1, pT7-P1, was assembled behind the T7 promoter of pT7E19(+) (28). This clone produced very small colonies but was stable in *E. coli* long enough to amplify the plasmid. Cultures were always inoculated with colonies from agar plates within 24 hr after transformation. If inoculum was taken from plates which had been stored at 4°C or from glycerol stocks kept at −80°C, no plasmid was recovered from overnight cultures (no detectable plasmid band on ethidium bromide-stained agarose gel). Also it was observed that plasmid yield from cultures grown in small volumes (3.5 ml) was higher than from cultures grown in larger

volumes (50 ml). However, despite these precautions, plasmid yield was low: 0.2–0.3 μg of DNA per ml of culture (Table 3).

To avoid the *in vitro* transcription reaction, the P1 cDNA sequence was placed under control of the enhanced CaMV 35S promoter and the nopaline synthase termination signal (T-NOS) of pAGUS1 so that plasmid could be inoculated directly and transcribed *in vivo*. This plasmid, p35S-P1, was not stable during amplification in *E. coli*, and analysis of restriction enzyme digestions of recovered plasmid showed strong bands of rearranged plasmids on a background of weaker bands representing the full-length p35S-P1 (data not shown).

Cloning of Intron Containing Full-Length Clones. The low plasmid yield of pT7-P1 and the instability of p35S-P1 suggest that these plasmids, like pP4(2259-6168), have deleterious effects on *E. coli*. Since interruption of the virus open frame appeared to relieve this problem, it was decided to insert introns containing multiple stop codons into the virus sequence and thereby prevent expression of virus proteins in *E. coli*. *In vivo* transcription of intron containing full-length clones from the cauliflower 35S promoter should assure splicing of the introns in the nucleus of the plant cells, thus restoring the virus open reading frame and the infectivity of the plasmid-derived RNA. The introns were inserted into the virus sequence to generate 5′ and 3′ intron splice sites matching the consensus of natural plant introns (34). This was accomplished by engineering intron A, B, and C from natural plant introns with a *Pst*I site 5′ to the intron sequence and a *Nsi*I site within the last six nucleotides of the intron. The introns were inserted as *Pst*I–*Nsi*I cassettes into native or engineered *Pst*I sites of P1 and P4. In natural plant genes the first nucleotide 3′ to the intron splice site is usually a guanine (34), and, accordingly, the nucleotide 3′ to the *Pst*I sites in the virus sequences were, if necessary, changed to a guanine. Fig. 2 shows design and insertion of the introns.

Since interruption of the open reading frame of P4 at position V had allowed amplification of pP4(2259-6168)-VΔ*Pst*I, intron A was inserted in position V of the P4 sequence. The sequence of intron A contains multiple stop codons and, when inserted in position V, the first in-frame stop codon was reached two codons downstream of nucleotide 3366. Like pP4(2259-6168)-VΔ*Pst*I, in which a stop codon is also present just 3′ to nucleotide 3366, pP4(2259-6168)-VA was easy to amplify (Table 3). pP4(2259-6168)-VA was subsequently used to assemble the full-length clone p35S-P4-VA. *E. coli* transformed with p35S-P4-VA contained the desired plasmid and though the colonies were small the plasmid was amplified to yield 12.5–17.5 μg DNA per ml culture.

Amplification of the intron containing P4 full-length clone, p35S-P4-VA, yielded more than 50 times more DNA per milliliter of *E. coli* culture than pT7-P1. It therefore seemed feasible to clone the full-length P1 sequence behind the cauliflower 35S promoter and improve plasmid yield by inserting introns into the P1 sequence. Construction of a P1 full-length clone analogous to p35S-P4-VA, required engineering of a *Pst*I site at position V in the P1 sequence, resulting in a Gly-to-Cys codon change. To avoid any codon changes, intron A was first inserted at two engineered *Pst*I sites at positions III and IV. However, all attempts to clone full-length P1 clones p35S-P1-IIIA and p35S-P1-IVA failed. The failure to assemble p35S-P1-IIIA and p35S-P1-IVA suggested that possibly insertion of the intron in position V was required to eliminate the toxicity of the virus sequence to *E. coli*. Consequently full-length clones were assembled with intron A inserted in position V and also at a downstream native *Pst*I site at position VI. Of the three clones p35S-P1-VA, p35S-P1-VIA, and p35S-P1-VA/VIA, only p35S-P1-VA/VIA was successfully cloned. Colonies of *E. coli* transformed with p35S-P1-VA/VIA were very small and plasmid recovery was only 0.1–0.2 μg of DNA per ml of culture. To improve amplification of the P1 full-length clone, introns B and C were inserted into

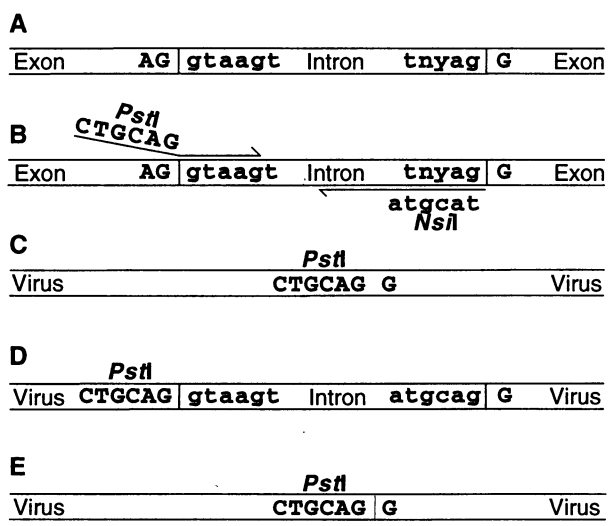

FIG. 2. Design and insertion of introns into the virus sequence. Intron sequences are shown in lowercase letters, and exon and virus sequences are shown in uppercase letters. (*A*) The consensus of plant introns 5′ and 3′ splice sites (34). (*B*) The PCR primers (arrows) used to engineer a *Pst*I site 5′ to the introns and a *Nsi*I site within the last six nucleotides of the introns. Only the 5′ sequence of the primers containing the restriction sites are shown; the 3′ sequence of the primers were designed to anneal to the sequence of introns A, B, and C, respectively, as indicated in Table 1. (*C*) The native or engineered *Pst*I site followed by a guanine in the virus cDNA sequence before intron insertion. (*D*) The virus cDNA sequence after insertion of the *Pst*I–*Nsi*I intron cassette. (*E*) The sequence at the intron splice site in viruses derived from intron containing full-length clones. The vertical line shows the position of the predicted splice site.

two *Pst*I sites introduced at positions I and II in the P1-protease coding region of the P1 genome. This region was chosen for intron insertion because subclones of the 5′ region of P1 were difficult to amplify yielding only 2.5–7.5 μg of plasmid per ml culture, despite the small size of the cDNA inserts. The introns B and C were used for insertion at position I and II to avoid the risk of homologous recombination in the plasmids caused by the presence of multiple identical sequences. Additionally, 11 full-length clones of P1 with various combinations of introns A, B, and C were assembled and, of these, six were successfully cloned (Fig. 1*B*). The colonies of *E. coli* transformed with these six plasmids were small but plasmids yield was 15–25 μg of DNA per ml culture (Table 3).

Biological Activity of *in Vitro*- and *in Vivo*-Transcribed RNA. The ability to initiate a virus infection on *P. sativum* was used to assay biological activity of capped *in vitro* synthesized transcripts and transcripts synthesized *in vivo* from p35S full-length plasmids. Healthy *P. sativum* were inoculated mechanically by rubbing transcript or plasmid on the leaflets in the presence of carborundum, and 2 weeks later the presence of virus in uninoculated upper leaves was assayed by ELISA. Inoculation with 2 μg of capped transcripts of pT7-P1 typically resulted in infection of 50–100% of inoculated plants. All *P. sativum* plants became infected when inoculated with 500 ng of the p35S full-length plasmids p35S-P4-VA, p35S-P1-VA/VIA, p35S-P1-IB/IIC/VA/VIA, p35S-P1-IC/IIB/VA/VIA, p35S-P1-IB/IIC/IVA/VIA, p35S-P1-IC/IIB/IVA/VIA, p35S-P1-IB/IIC/VIA, and p35S-P1-IC/IIB/VIA. The dilution end point of infectivity of the *in vivo*-transcribed plasmids was estimated by inoculation with various amounts of p35S-P4-VA, p35S-P1-IB/IIC/VA/VIA, and p35S-P1-IC/IIB/VA/VIA. Inoculation with ≥50 ng of plasmid always resulted in virus infection, whereas some plants escaped infection when less DNA was used. Symptom induction and virus accumulation of cDNA-derived P1 and P4 were identical to native P1 and P4

on three cultivars of *P. sativum* and on the indicator host *C. quinoa* (data not shown).

Fidelity of Intron Splicing. The nucleotide sequences of virus cDNA obtained from *P. sativum* plants inoculated with p35S-P4-VA, p35S-P1-IB/IIC/VA/VIA, and p35S-P1-IC/IIB/VA/VIA were determined to confirm that the introns A, B, and C were spliced correctly from the *in vivo*-transcribed pre-mRNA. The cDNA sequence surrounding intron A insertion site in P4 at position V and introns B and C insertion sites in P1 at both position I and II confirmed that the three introns were spliced as predicted (Fig. 2). Intron splicing of intron A in P4 restored the native P4 sequence whereas splicing of introns B and C in P1 restored the *Pst*I sites which had been engineered into the P1 sequence to insert the introns (data not shown).

DISCUSSION

This study has shown that insertion of introns into cloned cDNA of two isolates of the plant potyvirus PSbMV can facilitate amplification of full-length infectious clones in *E. coli*. The intron containing PSbMV P1 and P4 sequences were placed under control of the enhanced CaMV 35S promoter and T-NOS. The full-length plasmids were inoculated directly onto pea leaves, and, after *in vivo* transcription and splicing, the virus infections were established.

Amplification of full-length clones of P1 and P4 without introns was difficult or impossible. Initially a full-length clone of P1, pT7-P1, was assembled from which infectious transcripts could be synthesized *in vitro*. However, the plasmid pT7-P1 was very difficult to amplify in *E. coli*, and a full-length clone of P4 for *in vitro* transcription could not be assembled because the P4 cDNA sequence covering nucleotides 2259-6168 was toxic to the *E. coli* host cells. The full-length sequence of P1 was also placed under control of the enhanced CaMV 35S promoter and T-NOS, but this plasmid was not stable during amplification in *E. coli*. Toxicity and instability of cloned virus cDNA in *E. coli* are not uncommon. Potato virus Y potyvirus full-length cDNA and cDNA of beet necrotic yellow vein furovirus RNA 2 was found to be refractory to cloning in *E. coli* (14, 16), and, in *E. coli* cultures harboring plasmids containing cDNA of RNA 1 from cowpea severe mosaic comovirus, <2% of the cells retained the plasmid after growing overnight (15). Infectious cDNA of potato virus Y was instead obtained by synthesizing a full-length cDNA copy by reverse transcription-PCR using double-stranded megaprimers, including the CaMV 35S promoter and T-NOS (14). Full-length *in vitro*-synthesized transcripts of beet necrotic yellow vein virus RNA 2 were produced by transcription of cDNA ligation products without amplification in bacteria (16), and Chen and Bruening (15) managed to reduce plasmid loss by washing the cells twice during growth and resuspending them in fresh antibiotic-containing medium.

The toxicity and instability of cloned virus cDNA sequences has been suggested to be caused by expression of virus proteins in *E. coli* (14). The A+T-rich virus cDNA potentially contains cryptic promoter-like sequences (14), directing transcription of the virus sequences in *E. coli* irrespective of the cloning vector used and the orientation of the virus cDNA sequence in the vector. Subsequent translation could be initiated at cryptic ribosomal binding sites in the transcribed virus sequences. This view was supported by the observation that plasmid containing the P4 cDNA fragment covering nucleotides 2259-6168 could be cloned if the virus open reading frame was interrupted by a 4-nt deletion or insertion of intron A in position V. Nucleotides 2259-6168 of P4 covers the region encoding the potyvirus P3 protein, which has been reported to have a toxic effect on the host cells when expressed in *E. coli* (18). Position V lies within the region encoding the C terminus of the P3 protein, and it is therefore possible that insertion of intron A in position V of P4 facilitated plasmid amplification by preventing expression of this part of the P3 protein in the *E. coli* cells.

Insertion of intron A in position V was also sufficient to allow cloning of the full-length P4 cDNA sequence under control of enhanced CaMV 35S promoter and T-NOS. Plasmid yield of p35S-P4-VA was >50 times higher than of pT7-P1. However, insertion of intron A in the same region of the P1 genome at position III, IV, or V did not result in amplification of the expected plasmids p35S-P1-IIIA, p35S-P1-IVA, or p35S-P1-VA. Introns were subsequently inserted at three more positions, I, II, and VI. Intron insertion at both position V and VI resulted in amplification of the plasmid p35S-P1-VA/VIA, which was infectious, but plasmid yield was even less than that in pT7-P1. However, the plasmid p35S-P1-VA/VIA demonstrated that it was possible to obtain infectious *in vivo*-synthesized RNA from plasmids containing more than one intron. Subsequent insertion of introns B and C in position I and II greatly improved plasmid yield. p35S-P1-IB/IIC/VA/VIA and p35S-P1-IB/IIC/VA/VIA amplified to the same level as p35S-P4-VA. Since the *Pst*I site engineered into the sequence of P1 in position V results in a Gly-to-Cys codon change the plasmids p35S-P1-IB/IIC/IVA/VIA and p35S-P1-IC/IIB/IVA/VIA were constructed in which intron A was moved from position V to position IV, and finally two clones, which do not have an intron inserted in the P3 protein coding region, p35S-P1-IB/IIC/VIA and p35S-P1-IC/IIB/VIA, were constructed. These four clones were all amplified to the same level as p35S-P4-VA, demonstrating that intron insertion in the P3 coding region of P1 was not necessary to obtain good amplification, whereas intron insertion at positions I, II, and VI were all needed to obtain plasmids that were stable and easy to amplify. These results demonstrated that, although PSbMV P1 and P4 are closely related viruses, it was apparently not the same sequences that were toxic to *E. coli*. The identity of the intron inserted in a particular positions did not appear to affect plasmid stability and amplification, since plasmids in which introns B and C were exchanged between positions I and II were amplified to approximately the same level. Furthermore, there was no evidence that the presence of two introns with the same sequence decreased stability or yield of the plasmids.

Fifty nanograms of DNA of the intron containing PSbMV cDNA clones was sufficient to assure 100% infection by mechanical inoculation, whereas 500 ng to 10 μg of DNA has been used to obtain infection from other *in vivo*-transcribed virus full-length clones (ref. 13 and references therein). For example, inoculation with 10 μg of DNA of an *in vivo*-transcribed full-length clone of plum pox potyvirus was not sufficient to assure 100% infection (35), and inoculation with 5 μg of *in vivo*-transcribed cDNA of each of the three cucumber mosaic virus genomic RNAs gave infection in only 30% of the plants, although infectivity could be improved to almost 100% by excision of the viral expression cassette (36). These results suggest that even in those situations where infectious cDNA clones have been obtained, intron insertion may be of great value not only by improving the efficiency of plasmid production but also by increasing infectivity.

This strategy could potentially be used to facilitate manipulation of any DNA sequence from which toxic products are expressed in *E. coli* from cryptic promoters and initiation codons. By selection of suitable promoter, terminator, and intron sequences, the biological function of cloned eukaryotic sequences can subsequently be analyzed *in vivo*, as demonstrated here by the establishment of virus infections.

I would like to thank Dr. John Weiland for directing my attention to the potential use of introns; Dr. Bernhard Borkhardt for valuable discussions on the experimental work; Dr. I. T. D. Petty for pT7E19(+); Dr. J. M. Skuzeski for pAGUS1; Lilly Sander for the nitrite reductase clone; Dr. Molin for pOU61 and advice on amplification of runaway plasmids; and Dr. Merete Albrechtsen for com-

ments on the manuscript. This work was supported in part by the Danish Veterinary and Agricultural Research Council Grant 13-4837-1.

1. Zhang, L., Hanada, K. & Palukaitis, P. (1994) *J. Gen. Virol.* **75,** 3185–3191.
2. De Jong, W. & Ahlquist, P. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 6808–6812.
3. Edwards, M. C. (1995) *Mol. Plant–Microbe Interact.* **8,** 906–915.
4. Atreya, C. D., Atreya, P. L., Thornbury, D. W. & Pirone, T. P. (1992) *Virology* **191,** 106–111.
5. Kavanagh, T., Goulden, M., Santa Cruz, S., Chapman, S., Barker, I. & Baulcombe, D. (1992) *Virology* **189,** 609–617.
6. Tsai, C.-H. & Dreher, T. W. (1991) *J. Virol.* **65,** 3060–3067.
7. Carrington, J. C., Haldeman, R., Dolja, V. V. & Restrepo-Hartvig, M. A. (1993) *J. Virol.* **67,** 6995–7000.
8. Culver, J. N., Dawson, W. O., Plonk, K. & Stubbs, G. (1995) *Virology* **206,** 724–730.
9. Dolja, V. V., McBride, H. J. & Carrington, J. C. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 10208–10212.
10. Baulcombe, D. C., Chapman, S. & Santa Cruz, S. (1995) *Plant J.* **7,** 1045–1053.
11. Joshi, R. L., Joshi, V. & Ow, D. W. (1990) *EMBO J.* **9,** 2663–2669.
12. Mori, M., Zhang, G.-H., Kaido, M., Okuno, T. & Furusawa, I. (1993) *J. Gen. Virol.* **74,** 1255–1260.
13. Boyer, J.-C. & Haenni, A.-L. (1994) *Virology* **198,** 415–426.
14. Fakhfakh, H., Vilaine, F., Makni, M. & Robaglia, C. (1996) *J. Gen. Virol.* **77,** 519–523.
15. Chen, X. & Bruening, G. (1992) *Virology* **191,** 607–618.
16. Quillet, L., Guilley, H., Jonard, G. & Richards, K. (1989) *Virology* **172,** 293–301.
17. Lama, J. & Carrasco, L. (1992) *J. Biol. Chem.* **267,** 15932–15937.
18. Rodriguez-Cerezo, E. & Shaw, J. G. (1991) *Virology* **185,** 572–579.
19. Greener, A. (1993) *Strategies Mol. Biol.* **6,** 7–9.
20. Uhlin, B. E., Molin, S., Gustafsson, P. & Nordstrom, K. (1979) *Gene* **6,** 91–106.
21. Riechmann, J. L., Lain, S. & Garcia, J. A. (1992) *J. Gen. Virol.* **73,** 1–16.
22. Alconero, R., Providenti, R. & Gonsalves, D. (1986) *Plant Dis.* **70,** 783–786.
23. Johansen, E., Rasmussen, O. F., Heide, M. & Borkhardt, B. (1991) *J. Gen. Virol.* **72,** 2625–2632.
24. Johansen, I. E., Keller, K. E., Dougherty, W. G. & Hampton, R. O. (1996) *J. Gen. Virol.* **77,** 1329–1333.
25. Kohnen, P. D., Dougherty, W. G. & Hampton, R. O. (1992) *J. Virol. Methods* **37,** 253–258.
26. Eckes, P., Rosahl, S., Schell, J. & Willmitzer, L. (1986) *Mol. Gen. Genet.* **199,** 216–224.
27. Sander, L., Jensen, P. E., Back, L. F., Stuman, B. M. & Henningsen, K. W. (1995) *Plant Mol. Biol.* **27,** 165–177.
28. Petty, I. T. D. (1988) *Nucleic Acids Res.* **16,** 8738.
29. Skuzeski, J. M., Nichols, L. M. & Gesteland, R. F. (1990) *Plant Mol. Biol.* **15,** 65–79.
30. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY), 2nd Ed., pp. E.5–E.7.
31. Singer, B. & Fraenkel-Conrat, H. (1961) *Virology* **14,** 59–65.
32. Spiegel, S. & Martin, R. R. (1993) *Ann. Appl. Biol.* **122,** 493–500.
33. Løve Larsen, F. E, Gerdes, K., Light, F. & Molin, S. (1984) *Gene* **28,** 45–54.
34. Wiebauer, K., Herrero, J.-J. & Filipowicz, W. (1988) *Mol. Cell. Biol.* **8,** 2042–2051.
35. Maiss, E., Timpe, U., Brisske-Rode, A., Lesemann, D.-E. & Casper, R. (1992) *J. Gen. Virol.* **73,** 709–713.
36. Ding, S.-W., Rathjen, J. P., Li, W.-X., Swanson, R., Healy, H. & Symons, R. H. (1995) *J. Gen. Virol.* **76,** 459–464.