The EMBO Journal vol.12 no.11 pp.4445-4452, 1993

Identification of the initiation sequence for viral-strand DNA synthesis of wheat dwarf virus

Françoise Heyraud, Volker Matzeit, Matthias Kammann[2], Sabine Schaefer, Jeff Schell and Bruno Gronenborn[1]

Max-Planck-Institut für Züchtungsforschung, Carl-Von-Linné Weg 10, 5000 Köln 30, Germany and [1]Institut des Sciences Végétales, CNRS, Avenue de la Terrasse, 91198 Gif-sur-Yvette Cedex, France

[2]Present address: Institut für Anthropologie und Humangenetik in der Biologie der Johann Wolfgang Goethe-Universität Frankfurt, Siesmayerstrasse 70, 6000 Frankfurt am Main 11, Germany

Communicated by J.Schell

The intergenic region of the circular single-stranded DNA genome of geminiviruses contains a sequence potentially able to fold into a stem-loop structure. This sequence has been reported to be involved in viral replication by serving as the origin for rolling-circle replication. However, in wheat dwarf virus (WDV) a deletion of 128 bp, removing this sequence, surprisingly does not prevent *de novo* viral DNA synthesis, but instead abrogates the processing of replicative intermediates into monomeric genomes. This deletion mutant permitted us to study the initiation of viral-strand DNA synthesis independently from its termination and also to identify the sequence within which rolling-circle DNA replication of WDV begins. We have mapped the initiation site of replication to a pentanucleotide, TACCC, a sequence that occurs twice in the large intergenic region of WDV: it is found in the right half of the stem-loop sequence and again 170 bases upstream where it is part of a 15 nucleotide sequence highly homologous to the right half of the stem-loop sequence. Here we show that viral-strand DNA synthesis efficiently initiates at both sequences.

Key words: geminivirus/monocotyledonous plants/rolling-circle replication/single-stranded DNA

Introduction

Geminiviruses are plant viruses characterized by twinned icosahedral particles that contain circular single-stranded DNA molecules. They replicate via a double-stranded intermediate in the nuclei of infected cells. Depending on their genome organization, they can be classified into two subgroups: bipartite geminiviruses, which have a divided genome consisting of two DNA components of ~2.7 kb each (designated DNA A and B), and monopartite ones, with a single DNA component of ~2.7 kb (for review see Lazarowitz, 1992). Mutational analysis and expression of viral genes in transgenic plants have shown that only a single viral gene product is required for viral DNA replication, namely the AL1 protein of the bipartite viruses or its counterparts, the ORF III/IV or C1 products, in the monopartite geminiviruses (Rogers *et al.*, 1986; Townsend *et al.*, 1986; Schalk *et al.*, 1989). In addition, the common region (CR) of bipartite geminiviruses and the large intergenic region (LIR) of monopartite geminiviruses contain an ~30 bp sequence which can potentially form a stem-loop structure with an absolutely conserved TAATATTAC motif in the loop sequence. Mutational analyses of the stem-loop sequence have shown that this element is required to produce viral progeny genomes (Revington *et al.*, 1989; Lazarowitz *et al.*, 1992; Schneider *et al.*, 1992).

The detection of double-stranded circular DNA as a replicative intermediate and the observation of sequence similarity between the conserved loop motif and the recognition/cleavage site of gene A protein of bacteriophage ΦX174 have led to the proposition that geminiviruses might replicate via a rolling-circle mechanism, using the stem-loop sequence as part of the origin of replication (Rogers *et al.*, 1986; Elmer *et al.*, 1988). Recently, more direct evidence in favour of this replication model has been provided for a bipartite geminivirus, African cassava mosaic virus (ACMV), as well as for a monopartite geminivirus, beet curly top virus (BCTV) (Saunders *et al.*, 1991; Stenger *et al.*, 1991).

In the case of wheat dwarf virus (WDV), we previously described a mutant (here called WDVΔAB) that lacks part of the LIR, including the stem-loop sequence (Kammann *et al.*, 1991a). This mutant genome was still able to carry out *de novo* viral DNA synthesis in protoplasts since replication products, in the form of high molecular weight DNA, were detected. These data suggest that other sequences in the LIR act as secondary sites for initiation of viral replication and that sequences deleted in WDVΔAB are essential for the generation of unit-length replication products.

Here we report the identification of the secondary site used for initiation of viral-strand DNA synthesis on WDVΔAB and show that the LIR of WDV contains two sites where rolling-circle replication can start. Requirement for a hairpin structure in this initiation process is discussed.

Results

The WDV D3 deletion mutant (Kammann *et al.*, 1991a; here called WDVΔAB, see Figure 1), although lacking the stem-loop sequence, is nonetheless able to replicate in suspension culture cells. In order to identify the sequences used for initiation of viral-strand DNA replication on WDVΔAB, the plasmid pWit (for WDV, initiator, terminator) was constructed. It consists of a WDVΔAB unit cloned via *Sst*I, into which an *Sst*I-*Bcl*I fragment bearing a wild-type LIR was inserted. As a consequence, two different LIRs are present on this construct: one contains the 128 bp *Apa*I-*Bcl*I deletion of WDVΔAB encompassing the stem-loop sequence (ΔAB-LIR, Figure 2B) and the other contains the wild-type counterpart (WT-LIR, Figure 2B). In addition, the 346 bp region between map positions 2362-2707 (*Sst*I-*Apa*I fragment, highlighted by thick arrows in Figure 2B) is duplicated. We hypothesized that initiation of viral-strand

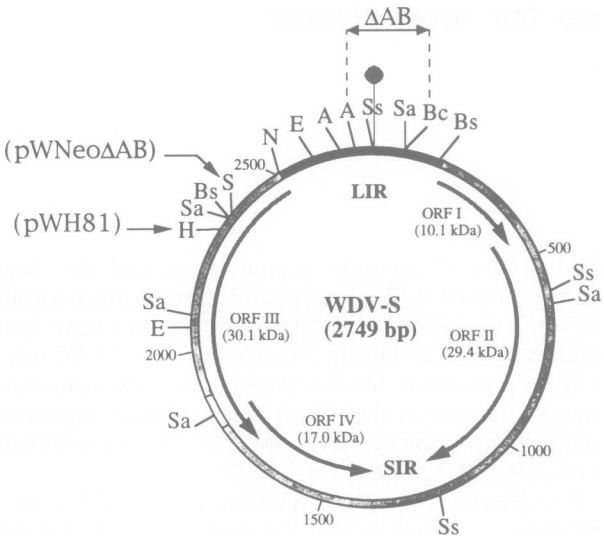

Fig. 1. Organization of the WDV genome (Swedish isolate). The filled circle marks the position of the putative geminivirus stem−loop or hairpin structure. Map position 1 designates the first base of the nonanucleotide motif [1]TAATATTAC contained in the loop sequence of the hairpin (Schalk *et al.*, 1989). Arrows inside the circle represent the four ORFs on the monopartite geminivirus genome; the sizes of the deduced corresponding translation products are given in brackets. The open box at the ORF III/IV junction represents an 86 nucleotide intron; splicing of this intron allows expression of ORFs III and IV as a 41 kDa fusion product. Two intergenic regions are present on the genome of monopartite geminiviruses: a LIR between ORFs I and III, represented by a filled box, and a small intergenic region (SIR) between the ends of ORFs II and IV. Relevant restriction sites are indicated: A, *Apa*I; Bc, *Bcl*I: Bs, *Bst*XI; E, *Eco*RI; H, *Hin*dIII; N, *Nco*I; S, *Sst*I; Sa, *Sau*3AI/*Mbo*I; Ss, *Ssp*I. The positions of the cloning vector in plasmids pWH81 (*Hin*dIII) and pWNeoΔAB (*Sst*I) are marked. The extent of the 128 bp *Apa*I−*Bcl*I deletion in the LIR is indicated by ΔAB.

DNA replication could take place in the ΔAB-LIR and that correct termination, i.e. processing of the replication products into genome-sized molecules, would occur in the WT-LIR, thereby preventing formation of the high molecular weight DNA produced upon replication of WDVΔAB alone. Therefore, replicative viral genomes with a novel LIR that contains composite 'footprints' of the sequences used for initiation and termination of DNA replication might be produced and serve for further analysis.

pWit replicates autonomously and also generates a WDV genome with a hybrid stem−loop

WDV replication in *Triticum monococcum* suspension culture cells was assayed following transfection of protoplasts with cloned DNA, as described previously (Matzeit *et al.*, 1991). Figure 3 shows a typical Southern blot analysis of DNA extracted from protoplast-derived cells that were harvested 1, 2 or 5 days after transfection with supercoiled pWit. Input DNA and *de novo* synthesized DNA were distinguished by their differential cleavability by *Mbo*I, an enzyme sensitive to adenosine methylation in the sequence GATC: *de novo* synthesized DNA in the plant cell is cleavable by *Mbo*I, whereas residual input DNA is not. In contrast, restriction digestion by *Sau*3AI is methylation insensitive. At day 1 post-transfection, only the input DNA is visible (lanes 2 and 3). At days 2 and 5 post-transfection, new restriction fragments appear and these can be attributed to three different replication products.

First, the entire pWit plasmid replicates autonomously in the plant cell, as illustrated in lanes 7 and 8, and 11 and 12. The 1807 bp fragment detected by the WDV probe consists of 883 bp of viral sequence (from the *Sau*3AI/*Mbo*I site located in open reading frame (ORF) III to the *Sst*I cloning site, see Figure 2A) and 924 bp of pUC sequence (from the *Sst*I cloning site to the next *Sau*3AI/*Mbo*I site, see Figure 2A). This particular fragment is produced from the protoplast DNAs prepared 2 and 5 days post-transfection when they are digested with either *Sau*3AI (lanes 8 and 11) or *Mbo*I (lanes 7 and 12). Moreover, a substantial part of pWit DNA becomes susceptible to *Mbo*I digestion (compare lanes 6 and 13 with 7 and 12).

Secondly, a WDVΔAB molecule of 2.62 kb arises, presumably by intramolecular homologous recombination within the 346 bp direct repeat sequences of the plasmid pWit (see Figure 2). The *Hin*dIII digests (lanes 9 and 10) show this as a 2.62 kb fragment which comigrates with the WDVΔAB fragment released from the input plasmid DNA by *Sst*I (lane 5, see also Figure 2). As shown previously (Kammann *et al.*, 1991a), this WDV mutant genome gives rise to large replicative products which cannot be further processed into monomeric units. Without restriction enzyme treatment these products are barely detectable in this experiment; they migrate close to the open circular form (o.c.) of the input DNA (lane 13). However, appearance of the 0.9 kb *Sau*3AI or *Mbo*I fragment (faintly visible in lanes 7 and 8, but more pronounced in lanes 11 and 12) is consistent with the generation of WDVΔAB molecules by recombination since this process would join the 883 bp *Sau*3AI−*Sst*I fragment to the small 20 bp *Sst*I−*Sau*3AI fragment (refer to Figure 2A for the *Sau*3AI restriction map).

Finally, a third replication product, represented by the 2.8 kb *Hin*dIII fragment (lanes 9 and 10), appears. When the DNA is digested by *Sau*3AI or *Mbo*I, a new fragment of 1.05 kb is produced (lanes 7 and 8, and 11 and 12). Unlike WDVΔAB, this 2.8 kb genome cannot arise from intramolecular homologous recombination, but its formation can readily be explained by a replicational release mechanism based on initiation of DNA synthesis in the ΔAB-LIR and processing of the replication products in the WT-LIR. Assuming that the signal for processing is provided by the stem−loop sequence of the WT-LIR, the signal for initiation of viral DNA replication has to be placed in the vicinity of the *Eco*RI site in the ΔAB-LIR. Consequently, a new *Mbo*I or *Sau*3AI fragment of ~1.05 kb, consisting of the 645 bp *Sau*3AI−*Eco*RI and the 405 bp *Ssp*I−*Sau*3AI regions, is expected upon *Mbo*I or *Sau*3AI cleavage (see Figure 2A). In addition, a sequence of ~110 bp (from around the *Eco*RI site to the ΔAB deletion point) has to be duplicated on the 2.8 kb molecule. This new type of replicative WDV molecule is called WDV-Hit (for hybrid initiator/terminator); its open circular (o.c.), linear (l) and supercoiled (s.c.) forms are visible when DNA prepared 5 days after transfection is electrophoresed undigested (bands marked by the arrowhead in lane 13).

Sequencing the LIR of WDV-Hit identifies the region involved in initiation/termination of viral-strand DNA synthesis

In order to characterize WDV-Hit in molecular terms, its LIR was amplified from total DNA prepared from protoplasts transfected with pWit. Using the primers pr.1 and pr.2

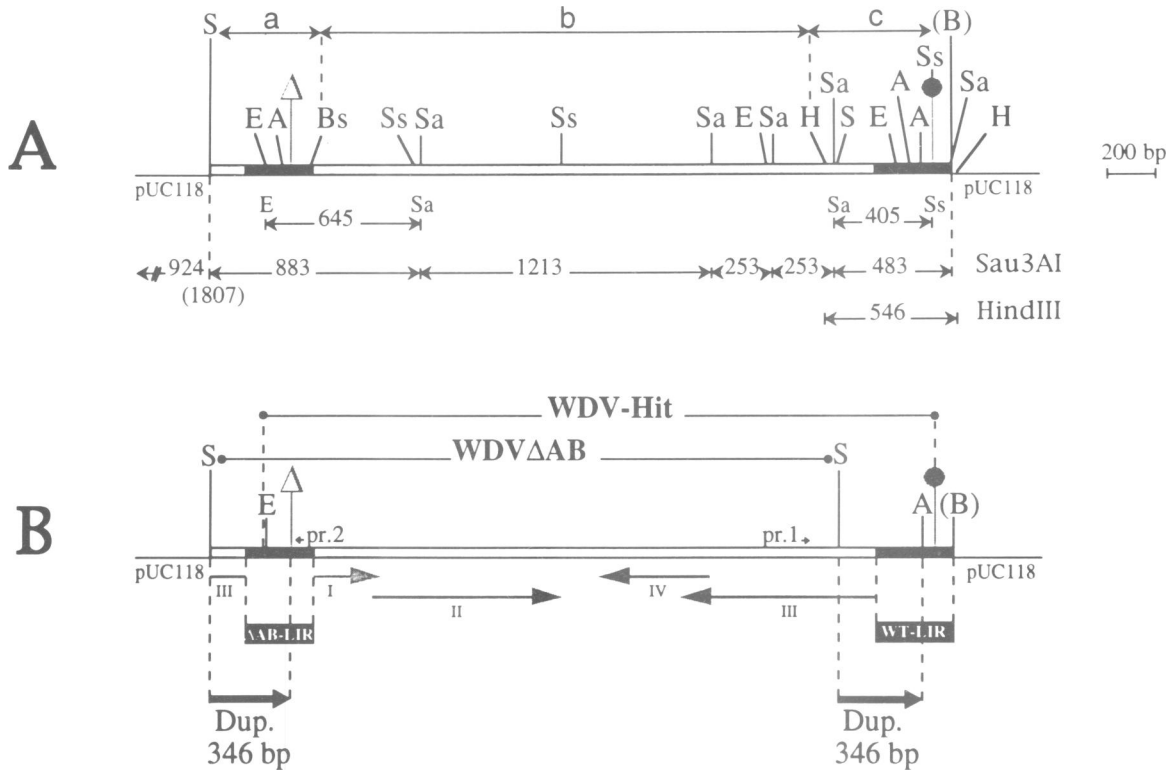

Fig. 2. The pWit plasmid. Only the WDV part of the plasmid is displayed in detail. The thin lines represent pUC118 sequences, while the thick line represents the WDV sequences cloned between the *Sst*I (S) and *Bam*HI (B) sites of pUC118. (**A**) Physical map of pWit. Viral sequences were assembled by ligation of restriction fragments a, b and c, originating from two parent plasmids: pWNeoΔAB (fragment a) and pWH81 (fragments b and c). Relevant restriction sites are indicated: A, *Apa*I; Bs, *Bst*XI; E, *Eco*RI; H, *Hin*dIII: S, *Sst*I; Sa, *Sau*3AI/*Mbo*I; Ss, *Ssp*I. Ligation between the *Bam*HI site of pUC118 and the *Bcl*I site of WDV resulted in the loss of both sites and is marked by (B). Δ indicates the location of the *ApaI*−*Bcl*I deletion and ● the location of the wild-type stem−loop sequence. Below the map restriction fragments relevant for the Southern blot analysis of Figure 3 are shown. The complete *Sau*3AI restriction map of WDV is given. The numbers refer to the size, in bp, of each restriction fragment. (**B**) Genetic map of pWit. Selected restriction sites, defining the borders of different regions, are shown. Δ indicates the location of the *ApaI*−*Bcl*I deletion and ● the location of the wild-type stem−loop sequence. These two positions characterize the two different LIRs, represented by the filled boxes and identified as ΔAB-LIR (for the LIR bearing the *ApaI*−*Bcl*I deletion) and WT-LIR (for the LIR containing the wild-type sequence). The location of the 346 bp repeated sequence is indicated by the thick black arrows. ORFs I, II, III and IV are represented by the grey arrows. The small arrows indicate the location and direction (5′ → 3′) of the primers pr. 1 and pr. 2 used in PCR assays. WDVΔAB refers to the WDV mutant genome bearing on its LIR the 128 bp *ApaI*−*Bcl*I deletion; WDV-Hit refers to the 2.8 kb replicative unit.

shown schematically in Figure 2B, two main PCR products of equal abundance were obtained. First, a 480 bp fragment was amplified from the 2.62 kb WDVΔAB template released from pWit by homologous recombination. The second fragment (640 bp) had the size predicted for the amplification product of the 2.8 kb WDV-Hit genome released from pWit by replication. In a control experiment performed on total DNA extracted from protoplasts transfected with DNA from the WDVΔAB mutant, only the 480 bp fragment was amplified (data not shown).

The 640 bp amplification product was gel purified and cloned in pUC118. Detailed restriction mapping (Figure 4A) and sequence analysis (Figure 4B) of eight recombinant pHit clones showed that they were identical. They all contain a duplication of 124 bp (nucleotides 2584−2707 on the wild-type WDV genome) and an *Ssp*I−*Eco*RI sub-fragment (hatched box above the map) which originates from the two different LIRs of the input plasmid pWit. The potential sites used for initiation and termination of WDV-Hit viral-strand DNA synthesis are therefore located in this sub-fragment. Interestingly, a new stem−loop sequence is found around the *Ssp*I restriction site and it may form a hairpin that has an almost perfect 12 bp stem (one mismatch at the tenth position) and a wild-type WDV loop containing the nonanucleotide TAATATTAC.

Taken together, these results demonstrate that initiation of replication around position 2600 (*Eco*RI site) in the ΔAB-LIR of pWit, and processing of the replication intermediates in the stem−loop sequence of the WT-LIR of pWit, generates WDV-Hit, a replicating 2.8 kb virus genome with a new LIR characterized by a chimeric stem−loop sequence and a 124 bp duplication.

In vitro constructed WDV-Hit replicates efficiently

The newly appearing recombinant WDV-Hit genome was cloned in pUC118 and tested for replication in suspension cultured cells. As shown in Figure 5, replicative s.c. and o.c. forms of WDV-Hit are observed 5 days post-transfection (compare lanes 6 and 5). WDV-Hit replicates equally well in *T.monococcum* cells, whether it is produced from pWit by replicational release or directly introduced into a cell as an individual viral genome (compare lanes 3 and 5). The amount of WDV-Hit replicative forms is nevertheless lower than that of WDV released from pWDV2NB, a plasmid similar to pWit but containing two identical WT-LIRs (compare lanes 3 and 5 with lane 1). It has to be stated,

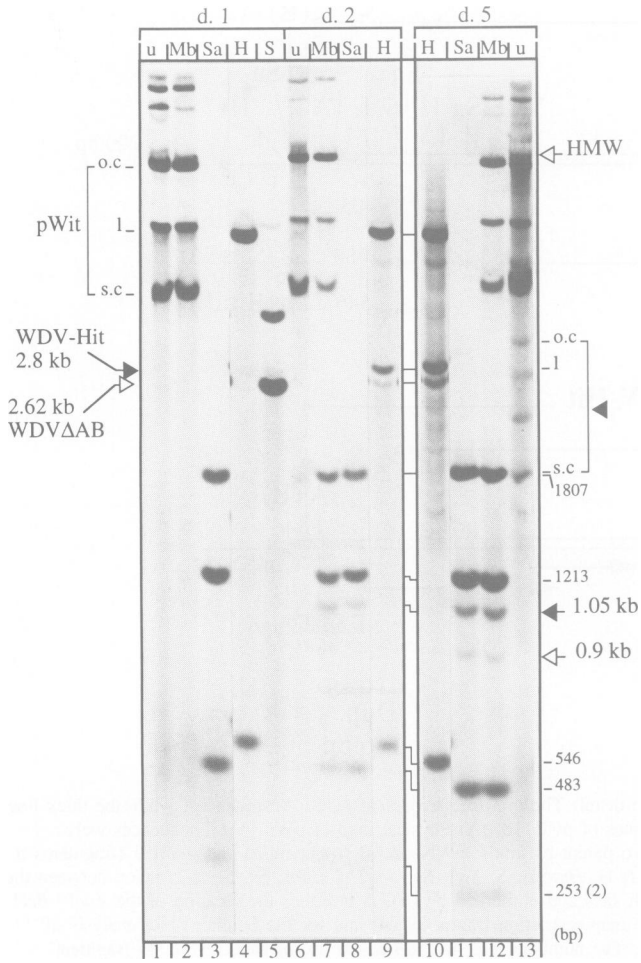

Fig. 3. Southern blot analysis of total DNA prepared from transfected protoplasts. Supercoiled pWit DNA was used to transfect *T.monococcum* protoplasts. Cells were harvested 1 (d. 1), 2 (d. 2) or 5 (d. 5) days after transfection and total DNAs were prepared. These DNAs, either uncut (u; lanes 1, 6 and 13) or after digestion with selected restriction endonucleases: *Mbo*I (Mb; lanes 2, 7 and 12), *Sau*3AI (Sa; lanes 3, 8 and 11), *Hin*dIII (H; lanes 4, 9 and 10) or *Sst*I (S; lane 5), were run on a 0.9% agarose gel. After transfer of the DNAs onto a nylon membrane, the blot was hybridized with a full-length digoxigenin-labelled WDV probe. The filled arrows at 2.8 kb and 1.05 kb indicate the respective unit size and a diagnostic *Sau*3AI–*Mbo*I fragment of the replicating viral genome WDV-Hit produced from pWit by replicational release. The corresponding o.c., 1 and s.c. forms are indicated by the filled arrowhead next to lane 13. The open arrows at 2.62 kb and 0.9 kb indicate the respective monomer size and a diagnostic *Sau*3AI–*Mbo*I fragment of the WDVΔAB product generated from pWit by intramolecular homologous recombination. The high molecular weight DNA products (HMW) associated with replication of WDVΔAB are also indicated. The scale (in bp) on the right refers to the *Hin*dIII (546) or *Sau*3AI fragments of pWit hybridized with the WDV probe.

however, that transfection with pWDV2NB, where the WDV genome is flanked by two wild-type stem–loop sequences, leads to an earlier appearance (day 1, see also lane 2) and a faster accumulation of WDV replicative forms, than transfections with an individual complete genome (F.Heyraud, unpublished). In addition, intramolecular homologous recombination occurs within the 124 bp duplicated region of WDV-Hit, generating WDVΔAB which in turn yields high molecular weight intermediates (see lanes 5 and 11).

Discussion

A WDV mutant with a 128 bp deletion including the conserved geminivirus potential stem–loop structure in the intergenic region replicates when introduced into *T.monococcum* suspension cultured cells. However, it does not yield genome-sized (2.62 kb) molecules of double- or single-stranded DNA, but produces high molecular weight replicative double-stranded intermediates (Kammann *et al.*, 1991a; see also Figure 5). In this mutant, here called WDVΔAB, the geminivirus DNA replication *per se* is uncoupled from termination which occurs when the replicative intermediates are processed into genome-length units. This allowed the analysis of the sequence requirements for initiation as opposed to termination of WDV DNA replication.

The LIR of WDV contains two sites for initiation of rolling-circle replication

A wild-type large intergenic region (WT-LIR) was added to WDVΔAB, generating pWit, a plasmid with two different LIRs (Figure 2B). Upon protoplast transfection, recombinant viral molecules (WDV-Hit) were released from pWit (Figure 3). They possess a hybrid LIR (Figure 6A) that consists of a left part originating from the WT-LIR (up to within the pentanucleotide TACCC of the stem–loop sequence) and a right part derived from the ΔAB-LIR (starting from within the TACCC pentanucleotide located upstream of the *Eco*RI site). Hence, the stem–loop sequence of the WT-LIR functions as a processing signal to release viral genomes having initiated rolling-circle replication at the TACCC pentanucleotide in the ΔAB-LIR. Interestingly, this initiator sequence is part of a 15 nucleotide sequence (nucleotides 2584–2598 on the viral-strand of WDV) which shares 13 nucleotides with the right half of the WDV stem–loop sequence (Figure 6B). As a consequence, the LIR of WDV-Hit contains a hybrid stem–loop sequence with a wild-type WDV loop and a new 12 bp stem (Figure 6A).

We propose that these two highly homologous regions of the LIR of WDV are two separate sites where rolling-circle replication can initiate. The stem–loop sequence contains the primary site for initiation of replication; this site coincides with the site of resolution of the replicative intermediates into genome-sized molecules. The sequence located between nucleotides 2584 and 2598 (i.e. 170 nucleotides upstream from the stem–loop sequence) constitutes the secondary site for initiation of replication. In contrast to the primary initiation site, this site cannot be used for processing of the replication products into genome-sized molecules, as demonstrated by the behaviour of the WDVΔAB mutant. Generation of different replication products from pWit demonstrates that both sites are used (see Figure 2); if rolling-circle replication starts at the primary site in the WT-LIR of pWit, the replication complex has to proceed through the bacterial plasmid sequence (pUC) and then through the whole WDV sequence before a resolution signal is reached (i.e. the stem–loop sequence in the WT-LIR). In this way, a complete pWit plasmid, synthesized *de novo* in the plant cell, is produced. If, on the other hand, rolling-circle replication starts at the secondary site for initiation of DNA synthesis in the ΔAB-LIR, the replication fork proceeds through the WDV genome and reaches the wild-type stem–loop sequence which allows resolution into genome-sized units; a 2.8 kb WDV-Hit molecule is released from the input

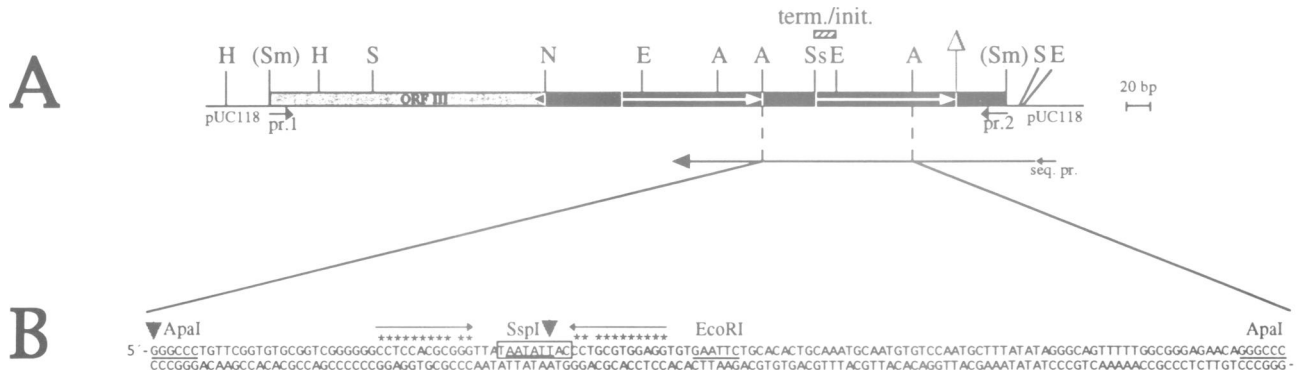

Fig. 4. Organization of the LIR of the replicating 2.8 kb viral genome (WDV-Hit). (**A**) Restriction map of the cloned PCR products. Using primers pr. 1 (viral strand polarity) and pr. 2 (complementary-strand polarity), a 640 bp fragment was amplified from total DNA prepared from protoplasts transfected with pWit. (Sm) refers to the former *Sma*I cloning site of pUC118 which was destroyed during the cloning procedure. Other relevant restriction sites are indicated: A, *Apa*I; E, *Eco*RI; H, *Hin*dIII; N, *Nco*I; S, *Sst*I; Ss, *Ssp*I. Δ indicates the location of the 128 bp *Apa*I−*Bcl*I deletion. The shaded box represents the 5′ end of the ORF III coding sequence (the black triangle at the *Nco*I site represents the AUG codon) and the filled boxes denote the LIR. The white arrows in the LIR indicate the location of the 124 bp duplication. The hatched box above the restriction map refers to the *Ssp*I−*Eco*RI sub-fragment containing initiation and termination signals for viral-strand DNA replication of WDV-Hit. The extent of the LIR for which the nucleotide sequence was determined (long thin arrow below part A), as well as the location of the sequencing primer (seq. pr.) are indicated. (**B**) Nucleotide sequence of the 140 bp internal *Apa*I fragment of pHit. Grey arrowheads indicate the end-points of the 124 bp duplicated sequence. Arrows represent the stem of a potential stem−loop structure. Asterisks denote nucleotides which are complementary. The nonanucleotide motif TAATATTAC of the loop sequence is framed.

plasmid. Efficient replication of the cloned WDV-Hit genome is also observed when it is transfected as an individual molecule into protoplasts (Figure 5). Concomitantly, homologous recombination in the redundant 124 bp sequence leads to a loss of the composite LIR and generates WDVΔAB which, upon further replication, will yield only long intermediates of multiple genome size.

The secondary site for initiation is used in nature

Interestingly, the same stem−loop sequence as the one contained on WDV-Hit had been observed previously in a Czechoslovakian isolate of WDV (Commandeur, 1987). Due to a duplication of 172 bp in the LIR (from nucleotide position 2584 to position 6; see Figure 6B), this clone had two potential hairpins, one with the wild-type sequence and the other with the same sequence as that present on the LIR of WDV-Hit. Most probably, the same mechanism as the one proposed above for the generation of the 2.8 kb WDV-Hit led to the generation of this natural WDV mutant. The replication complex started viral-strand DNA synthesis within the TACCC sequence of the secondary initiation site and proceeded through the LIR, passing along the stem−loop sequence without efficient resolution (or producing only a 172 bp minicircle that has so far gone undetected). After a complete round of replication, the canonical stem−loop sequence was reached for the second time and a resolution structure was formed, allowing the release of a monomeric genome unit, yet bearing a 172 bp duplication in the LIR. The fact that one of three WDV clones isolated from naturally infected Czechoslovakian wheat contained this duplication suggests that the secondary initiation site may be used rather frequently *in planta*.

Sequence requirements for rolling-circle replication of the WDV genome

The fact that the secondary site for initiation of replication cannot be used as a resolution signal for the release of genome-sized molecules indicates that the constraints for a correct termination are more stringent than those for

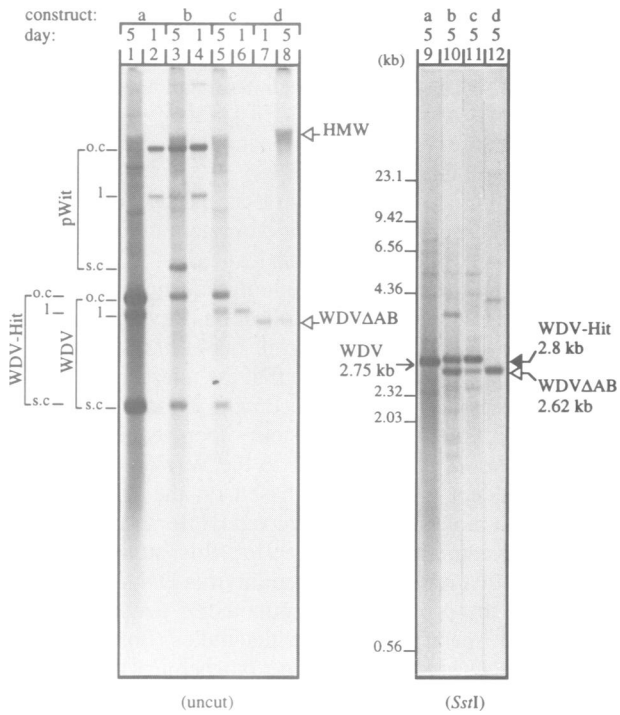

Fig. 5. Replication of WDV-Hit compared with that of WDV. This Southern blot analysis was performed as indicated in Figure 3. The following constructs were used: a, pWDV2NB; b, pWit; c, WDV-Hit; d, WDVΔAB. pWDV2NB and pWit were transfected as supercoiled plasmid DNAs whereas WDV-Hit and WDVΔAB were released from pWDV-Hit and pWDVΔAB respectively, by *Sst*I digestion prior to transfection. Total DNAs prepared 1 or 5 days post-transfection were run either uncut (lanes 1−8) or after digestion with *Sst*I (lanes 9−12). Each lane contains 7 μg of total DNA. o.c., l and s.c. forms of replicating pWit, WDV-Hit and WDV are indicated, as well as the unit sizes of replicating WDVΔAB, WDV and WDV-Hit genomes. HMW refers to the high molecular weight DNA products associated with replication of WDVΔAB. The scale (in kb) next to the *Sst*I blot corresponds to the restriction fragments generated by *Hin*dIII digestion of bacteriophage λ.

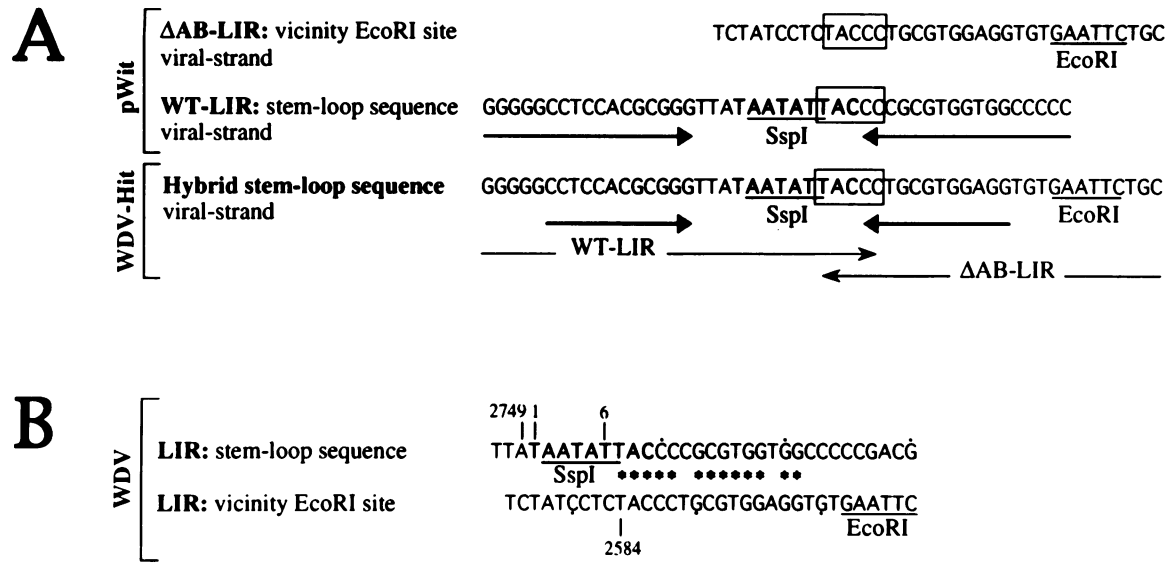

Fig. 6. Sequences potentially involved in initiation and termination of WDV rolling-circle replication. (**A**) Generation of the LIR of WDV-Hit from the two different LIRs of pWit. The upper sequence shows the region of ΔAB-LIR which was used for initiation of WDV-Hit replication on pWit. The middle sequence shows the region of WT-LIR which was used for processing of the corresponding replication products on pWit. The lower sequence shows the hybrid stem–loop sequence found in the LIR of WDV-Hit: the left half originates from WT-LIR and the right half from ΔAB-LIR. The TACCC motif present on both WT- and ΔAB-LIR, and constituting the junction between the two halves of the new stem–loop sequence, is framed. The loop sequence motif, TAATATTAC, is shown in bold; the stem inverted repeat sequences are underlined with arrows. *Eco*RI and *Ssp*I restriction sites are indicated. (**B**) WDV LIR contains two initiation sites for viral-strand DNA synthesis. The right half of the WDV stem–loop sequence (upper sequence) and the region of the LIR immediately upstream from the *Eco*RI restriction site (lower sequence) are compared at the nucleotide level. Sequence homology is denoted by asterisks: 13 positions, over a stretch of 15 nucleotides, are conserved. The loop sequence motif TAATATTAC is shown in bold, *Ssp*I and *Eco*RI restriction sites are marked and relevant nucleotide positions are indicated.

initiation. Presumably, the latter requires the formation of a particular secondary structure.

Secondary initiation sites for viral-strand DNA synthesis have not been described so far for other geminiviruses. For tomato golden mosaic virus, squash leaf curl virus and maize streak virus, the respective stem–loop sequences have been shown to be essential for DNA replication (Revington *et al.*, 1989; Lazarowitz *et al.*, 1992; Schneider *et al.*, 1992). This correlates quite well with the fact that the intergenic regions of these viruses do not seem to contain sequences sharing homology with their respective stem–loop sequences. For WDV, a deletion mutant similar to the WDVΔAB mutant was reported not to replicate, based on the fact that no genome-sized molecules were detected (Hofer *et al.*, 1992). Our data (Kammann *et al.*, 1991a; this study) clearly demonstrate that replication of geminivirus DNA cannot be restricted to the production of monomeric genome units.

Apart from specific DNA binding studies (Fontes *et al.*, 1992; Thömmes *et al.*, 1993) direct biochemical data on the action of the *rep* proteins of geminiviruses (i.e. the ORF III/IV, C1 or AL1 proteins) on their DNA target(s) are so far lacking. From the data reported here, we predict that the specific nick that initiates viral-strand DNA synthesis of WDV is introduced within the pentanucleotide TACCC.

Rolling-circle replication of single-stranded DNA phages and single-stranded plasmids: models for geminivirus DNA replication?

Rolling-circle replication has been described in great detail for the single-stranded DNA phages of *Escherichia coli*, exemplified by ΦX174 or fd (Kornberg and Baker, 1992, and references therein). Also single-stranded plasmids of Gram-positive bacteria, like pT181, pC194 or pUB110 (Koepsel *et al.*, 1985; Gros *et al.*, 1987, 1989; Koepsel and Khan, 1987; Gruss and Ehrlich, 1989; Murray *et al.*, 1989), or of Gram-negative bacteria, like pKYM (Yasukawa *et al.*, 1991), multiply by rolling-circle replication. Sequence comparisons between the *rep* proteins of these plasmids and the ORF III/IV, C1 or AL1 gene products of geminiviruses revealed conserved motifs, therefore suggesting a similar mechanism of action (Koonin and Ilyina, 1992). The similarities can even be extended to proteins involved in the initial nicking of the T-strand of the T-DNA of the *Agrobacterium tumefaciens* Ti plasmid prior to its transfer into the plant cell (Ilyina and Koonin, 1992).

A detailed model for the interaction between the *rep* protein of pT181 and the plasmid origin region has been described recently (Wang *et al.*, 1992, 1993). Molecular and biochemical analyses of the geminivirus *rep* protein function(s) are the subject of current studies. They will help to elucidate how general the mechanisms of rolling-circle initiation are among prokaryotic and eukaryotic single-stranded DNA replicons.

Materials and methods

Recombinant DNA techniques

Restriction endonucleases and other DNA-modifying enzymes were used as recommended by the suppliers (Boehringer Mannheim, New England Biolabs and Pharmacia LKB). Standard techniques in molecular biology were applied as described by Sambrook *et al.* (1989).

Construction of plasmids

A restriction map of the Swedish isolate of WDV is shown in Figure 1.

pWit was derived from the plasmids pWNeoΔAB and pWH81. pWNeoΔAB (D3 mutant of Kammann *et al.*, 1991a) consists of a viral unit, bearing the neomycin phosphotransferase (NPTII) reporter gene in place of the capsid protein gene (ORF II), cloned in the *Sst*I site of pUCBY (a pUC18 derivative with a modified polylinker sequence; B.Gronenborn, unpublished). In addition, it contains a 128 bp deletion between the *Apa*I

and *Bcl*I restriction sites in the LIR [map positions 2708 to 86, as defined in Schalk *et al.* (1989)]. pWH81 (Kammann *et al.*, 1991b) carries a wild-type WDV genome (Swedish isolate) cloned in the *Hin*dIII site of pUC8 (Yanisch-Perron *et al.*, 1985). From these two plasmids three restriction fragments (a, b and c, Figure 2A) were gel purified and ligated to an *Sst*I- and *Bam*HI-digested pUC118 vector (Vieira and Messing, 1987). Fragment a is a 433 bp *Sst*I−*Bst*XI fragment bearing the *Apa*I−*Bcl*I deletion in the LIR and isolated from pWNeoΔAB. Fragment b is a 2137 bp *Bst*XI−*Hin*dIII fragment, containing ORFs I, II and IV and part of ORF III, isolated from pWH81. Fragment c is a 516 bp *Hin*dIII−*Bcl*I fragment containing the N-terminal part of ORF III and most of the LIR (wild-type sequence) and isolated from pWH81. The resulting plasmid, pWit, is shown in Figure 2 in a linear organization.

pWDV-Hit was derived from pWDVΔAB and pHit. pWDVΔAB consists of a WDV mutant genome bearing the 128 bp ΔAB deletion in the LIR, cloned in the *Sst*I site of pUC118. pHit contains the 640 bp PCR product derived from replicating WDV-Hit genomes, cloned in the *Sma*I site of pUC118. To generate pWDV-Hit, the 159 bp *Nco*I−*Apa*I fragment located in the LIR of pWDVΔAB (see Figure 1) was replaced with the 331 bp *Nco*I−*Apa*I fragment of pHit, which contains the hybrid stem−loop sequence (see Figure 4).

pWDV2NB was obtained by ligating two fragments derived from pWH81 to an *Nco*I- and *Bam*HI-digested pUC120 vector (kindly provided by J.Vieira and J.Messing): the 2549 bp *Nco*I−*Hin*dIII fragment and the 516 bp *Hin*dIII−*Bcl*I fragment, both encompassing the LIR (see Figure 1). pWDV2NB therefore contains 1.1 copies of the WDV genome.

Tissue culture, protoplast transfection and preparation of DNA

The constructs to be tested were introduced into *T.monococcum* protoplasts derived from a cell suspension culture, via a polyethylene glycol-mediated DNA uptake, as described previously (Matzeit *et al.*, 1991). Routinely, 15 μg of purified plasmid DNA [prepared using disposable Qiagen columns (Diagen)] were used to transfect 1−3 × 10^6 protoplasts per assay. Total DNA was prepared from protoplast-derived cells using standard phenol−chloroform extractions.

Analysis of protoplast DNA

Viral DNA replication was monitored by Southern blot analysis (Southern, 1975) of total DNA extracted from protoplast-derived cells at various time points after transfection. Typically, 5−10 μg of DNA, digested with restriction enzymes or untreated, were separated on 0.9% agarose gels and blotted onto nylon membranes (Hybond-N, Amersham). The membranes were probed with a full-length WDV fragment labelled by incorporation of a nucleotide analogue (digoxigenin-11-dUTP) into DNA by the random primed labelling technique (non-radioactive DNA labelling kit, Boehringer Mannheim). All treatments of the membranes were carried out as recommended by the manufacturer. Visualization of the hybridized fragments was performed using either a chemiluminescent substrate [AMPPD: 3-(2′-spiroadamantane)-4-methoxy-4-(3″-phosphoryloxy)-phenyl-1,2-dioxetane] or a coloured substrate (BCIP: 5-bromo-4-chloro-3-indolyl phosphate + NBT: 4-nitroblue tetrazolium chloride) for alkaline phosphatase conjugated to an anti-digoxigenin antibody.

Distinction between input DNA and *de novo* synthesized DNA was accomplished by differential digestion with *Mbo*I and *Sau*3AI. These two enzymes recognize the same target sequence (GATC) but, whereas *Mbo*I only cleaves when the A residue is not methylated, *Sau*3AI activity is not affected by the methylation state of the GATC sequences. Consequently, plasmid DNA isolated from *dam*$^+$ *E.coli* strains is completely cut by *Sau*3AI but resistant to *Mbo*I digestion, as in these strains the GATC sequences are methylated on the A residues. In contrast, *de novo* synthesized DNA in the plant cell contains unmethylated GATC sequences, and hence is cleavable by both *Mbo*I and *Sau*3AI.

Amplification of viral-specific sequences by polymerase chain reaction

Using a thermostable DNA polymerase (Saiki *et al.*, 1988), PCR was performed on protoplast DNA in order to amplify LIR sequences of replicating viral genomes. The following oligonucleotides were used as primers: pr. 1, 5′-TCCATACGGAGGTTTAAGGC-3′ [map positions 2270−2289, as defined in Schalk *et al.* (1989)] and pr. 2, 5′-ATTTCC-CCGTCAAAAGTGTG-3′ (complementary to map positions 130−111). pr. 1 is part of the ORF III coding sequence and has viral-strand polarity; pr. 2 is located in the LIR upstream of ORF I and has complementary-strand polarity.

The oligonucleotides were synthesized on an Applied Biosystems DNA synthesizer (model 380B) and used after purification on NAP G-25 columns (Pharmacia).

PCR assays on total DNA isolated from protoplast-derived cells were performed in a final volume of 50 μl consisting of 2.5 μg total DNA, 5 μl 10 × reaction buffer (100 mM Tris−HCl, pH 8.3, 500 mM KCl, 15 mM $MgCl_2$, 0.01% gelatin) and 100 pmol of each primer (final concentration: 2 μM). The DNA was denatured for 5 min at 94°C and subsequently the following reaction components were added: 10 nmol of each dNTP (final concentration: 200 μM each) and 2 U of *Taq* DNA polymerase ('Native Taq DNA Polymerase', Perkin Elmer−Cetus). Finally, 50 μl of mineral oil were added to prevent evaporation.

Thirty amplification cycles were carried out using a Biometra Trio-Thermoblock. Each cycle included a heat denaturation step at 92°C for 1 min, followed by annealing of the primers to the DNA at 42°C for 30 s and DNA chain extension with *Taq* polymerase at 72°C for 1 min. After cycle completion, one-fifth (10 μl) of each reaction was loaded on a 1.5% agarose gel and subjected to electrophoresis in order to control the amplification. Subsequently, the remaining 80% of the reactions was similarly electrophoresed and the fragments of interest were recovered.

Cloning and sequencing of PCR products

PCR products were gel purified using the QIAEX agarose gel extraction kit (Diagen). To facilitate cloning, the ends of the fragments were made flush by treatment with the Klenow fragment of DNA polymerase I (30 min at room temperature) prior to ligation to a *Sma*I-cleaved pUC118 vector. Double-stranded DNA of selected recombinant pHit clones was prepared for sequencing and the sequencing reactions were performed by the dideoxy chain termination method (Sanger *et al.*, 1977) using a modified T7 DNA polymerase (Tabor and Richardson, 1989) and a fluorescent primer (reverse sequence primer, Pharmacia LKB). Products were run and analysed on an automated DNA sequencer (ALF DNA Sequencer, Pharmacia LKB).

Acknowledgements

We thank S.Rechmann for performing the DNA sequencing. We are indebted to K.Moore and R.Walden for critical comments on the manuscript, and to J.Leung and I.Moore for stimulating discussions. This work was supported by the Bundesministerium für Forschung und Technologie; F.H. was in receipt of a MPG/CNRS fellowship.

References

Commandeur,U. (1987) Diploma Thesis. University of Köln, Germany.
Elmer,J.S., Sunter,G., Gardiner,W.E., Brand,L., Browning,C.K., Bisaro,D.M. and Rogers,S.G. (1988) *Plant. Mol. Biol.*, **10**, 225−234.
Fontes,E.P.B., Lucknow,V.A. and Hanley-Bowdoin,L. (1992) *Plant Cell*, **4**, 597−608.
Gros,M.-F., te Riele,H. and Ehrlich,S.D. (1987) *EMBO J.*, **6**, 3863−3869.
Gros,M.-F., te Riele,H. and Ehrlich,S.D. (1989) *EMBO J.*, **8**, 2711−2716.
Gruss,A. and Ehrlich,S.D. (1989) *Microbiol. Rev.*, **53**, 231−241.
Hofer,J.M.I., Dekke,E.L., Reynolds,H.V., Woolston,C.J., Cox,B.S. and Mullineaux,P.M. (1992) *Plant Cell*, **4**, 213−223.
Ilyina,T.V. and Koonin,E.V. (1992) *Nucleic Acids Res.*, **20**, 3279−3285.
Kammann,M., Schalk,H.-J., Matzeit,V., Schaefer,S., Schell,J. and Gronenborn,B. (1991a) *Virology*, **184**, 786−790.
Kammann,M., Matzeit,V., Schmidt,B., Schell,J., Walden,R. and Gronenborn,B. (1991b) *Gene*, **104**, 247−252.
Koepsel,R.R. and Khan,S.A. (1987) *Nucleic Acids Res.*, **15**, 4085−4097.
Koepsel,R.R., Murray,R.W., Rosenblum,W.D. and Khan,S.A. (1985) *Proc. Natl Acad. Sci. USA*, **82**, 6845−6849.
Koonin,E.V. and Ilyina,T.V. (1992) *J. Gen. Virol.*, **73**, 2763−2766.
Kornberg,A. and Baker,T.A. (1992) *DNA Replication*. W.H.Freeman & Co., New York.
Lazarowitz,S.G. (1992) *CRC Crit. Rev. Plant Sci.*, **11**, 327−349.
Lazarowitz,S.G., Wu,L.C., Rogers,S.G. and Elmer,J.S. (1992) *Plant Cell*, **4**, 799−809.
Matzeit,V., Schaefer,S., Kammann,M., Schalk,H.-J., Schell,J. and Gronenborn,B. (1991) *Plant Cell*, **3**, 247−258.
Murray,R.W., Koepsel,R.R. and Khan,S.A. (1989) *J. Biol. Chem.*, **264**, 1051−1057.
Revington,G.N., Sunter,G. and Bisaro,D.M. (1989) *Plant Cell*, **1**, 985−992.
Rogers,S.G., Bisaro,D.M., Horsch,R.B., Fraley,R.T., Hoffmann,N.L., Brand,L., Elmer,J.S. and Lloyd,A.M. (1986) *Cell*, **45**, 593−600.
Saiki,R.K., Gelfand,D.H., Stoffel,S., Scharf,S.J, Higuchi,R., Horn,G.T., Mullis,K.B. and Erlich,H.A. (1988) *Science*, **239**, 487−491.
Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A*

Laboratory Manual. 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl Acad. Sci. USA*, **74**, 5463–5467.
Saunders,K., Lucy,A. and Stanley,J. (1991) *Nucleic Acids Res.*, **19**, 2325–2330.
Schalk,H.-J., Matzeit,V., Schiller,B., Schell,J. and Gronenborn,B. (1989) *EMBO J.*, **8**, 359–364.
Schneider,M., Jarchow,E. and Hohn,B. (1992) *Plant. Mol. Biol.*, **19**, 601–610.
Southern,E.M. (1975) *J. Mol. Biol.*, **98**, 503–517.
Stenger,D.C., Revington,G.N., Stevenson,M.C. and Bisaro,D.M. (1991) *Proc. Natl Acad. Sci. USA*, **88**, 8029–8033.
Tabor,S. and Richardson,C.C. (1989) *Proc. Natl Acad. Sci. USA*, **86**, 4076–4080.
Thömmes,P., Osman,T.A.M., Hayes,R.J. and Buck,K.W. (1993) *FEBS Lett.*, **319**, 95–99.
Townsend,R., Watts,J. and Stanley,J. (1986) *Nucleic Acids Res.*, **14**, 1253–1265.
Vieira,J. and Messing,J. (1987) *Methods Enzymol.*, **153**, 3–11.
Wang,P.-H., Projan,S.J., Henriquez,V. and Novick,R.P. (1992) *J. Mol. Biol.*, **223**, 145–158.
Wang,P.-H., Projan,S.J., Henriquez,V. and Novick,R.P. (1993) *EMBO J.*, **12**, 45–52.
Yanisch-Perron,C., Vieira,J. and Messing,J. (1985) *Gene*, **33**, 103–119.
Yasukawa,H., Hase,T., Sakai,A. and Masamune,Y. (1991) *Proc. Natl Acad. Sci. USA*, **88**, 10282–10286.

Received on February 1, 1993; revised on July 23, 1993