The EMBO Journal vol.3 no.13 pp.3069–3073, 1984

A putative primer for second-strand DNA synthesis of maize streak virus is virion-associated

J. Donson, B.A.M. Morris-Krsinich, P.M. Mullineaux, M.I. Boulton and J.W. Davies

John Innes Institute, Colney Lane, Norwich NR4 7UH, UK

Communicated by D.A. Hopwood

We have isolated, from maize streak virus (MSV) preparations, a population of 'nested' DNA molecules. These molecules have ribonucleotides covalently linked to the DNA species' discrete 5′ deoxyribonucleotide terminus. The major species has a DNA sequence of 80 nucleotides which is complementary to a region 5′ of two hairpin structures on the MSV genome, almost exclusively in an intergenic region. These molecules have been used to prime the synthesis *in vitro* of a complementary strand to virion DNA, initiating this reaction at one site on the genome.

Key words: geminivirus/primer/nucleotide sequence/maize streak virus/replicative form

Introduction

Geminiviruses and a sole representative of the animal viruses – porcine circovirus (PCV; Tischer *et al.*, 1982) are the only eukaryotic viruses having a genome of single-stranded circular DNA. Although little information is available about the replicative cycle of PCV, potential replicative intermediates have been isolated from bean golden mosaic virus (BGMV; Ikegami *et al.*, 1981), tomato golden mosaic virus (TGMV; Hamilton *et al.*, 1982) and cassava latent virus (J. Stanley, personal communication) infected plants. Double-stranded, covalently closed circular DNA (cccDNA) has been isolated from TGMV (Hamilton *et al.*, 1982; Sunter *et al.*, 1984) CLV (J. Stanley, personal communication) and maize streak virus (MSV; unpublished data) infected plants, though in the case of BGMV only an open-circular, double-stranded form has been found (Ikegami *et al.*, 1981). MSV is however distinct from the other geminiviruses characterised at the genome level in apparently having only one species of circular DNA as opposed to two (see preceding paper by Mullineaux *et al.*).

For those autonomous viruses with a single-stranded DNA genome other than the insect densonucleosis viruses, the primary event on invading a host cell is the synthesis of a second DNA strand complementary to the virion DNA. The question of how DNA synthesis on the incoming geminiviral DNA template is primed has not yet been resolved. Here we report the existence of a population of 5′ co-terminal small DNA molecules with covalently linked ribonucleotides, associated with virions. We have also analysed *in vitro* their potential as a primer for second-strand synthesis.

Results

Gel electrophoresis of virion DNA

Gel electrophoresis of MSV DNA under non-denaturing conditions (Figure 1) revealed a single circular species for which the nucleotide sequence has been determined (see preceding paper by Mullineaux *et al.*). Dimethylsuphoxide (DMSO) denaturation and glyoxylation, followed by gel electrophoresis and Southern transfer hybridisation, revealed two other classes of molecules (Figure 1 track E): (i) molecules (band 1), also evident in Southern transfers of non-denatured DNA that migrated slower than the 2687-base circular species (band 2) and are as yet uncharacterised; (ii) molecules migrating faster than species 2 as a single band (band 3). The molecules in band 3 were sensitive to digestion of virion DNA with DNase but apparently resistant to RNase and nuclease S1 digestion (results not shown).

Sequencing of the small DNA molecules

Electrophoresis of 5′ or 3′ end-labelled virion DNA on a 6% denaturing gel (Figure 3, tracks 1–5 and Figure 4, track A) showed that band 3 (Figure 1, track E) observed on Southern transfer hybridisation was in fact a population of molecules of which each member showed an apparent mobility difference from the next of one nucleotide.

Following alkaline treatment the population of DNA molecules was 5′ end-labelled and separated as a band from a strand-separation gel. Total and partial P1 nuclease digestion in conjunction with wandering spot analysis (Figure 2) established that the small DNA molecules had a common 5′

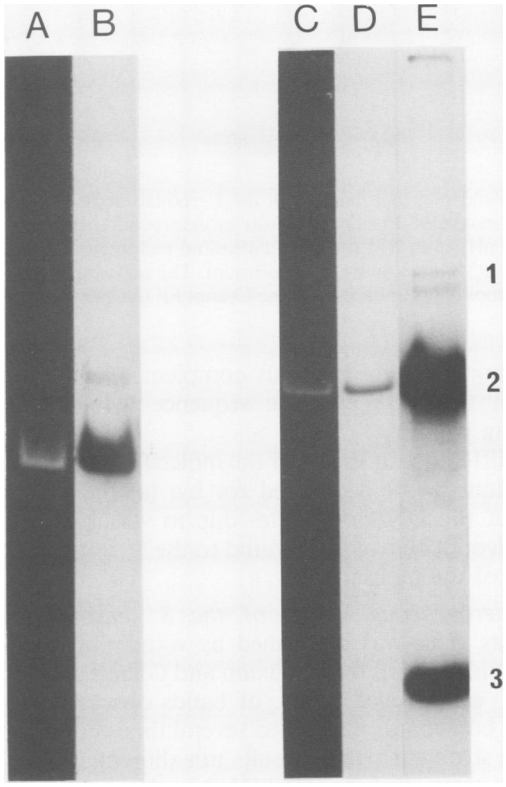

Fig. 1. Electrophoresis on a 1.0% agarose gel **(A)** untreated MSV DNA, **(C)** glyoxal denatured MSV DNA; stained with ethidium bromide. **Tracks B** and **D** represent **tracks A** and **C** transferred to nitrocellulose and hybridised using ^{32}P-labelled MSV cDNA as the probe. **Track E** is an over exposure of **track D**. Bands 1, 2 and 3 are referred to in the text.

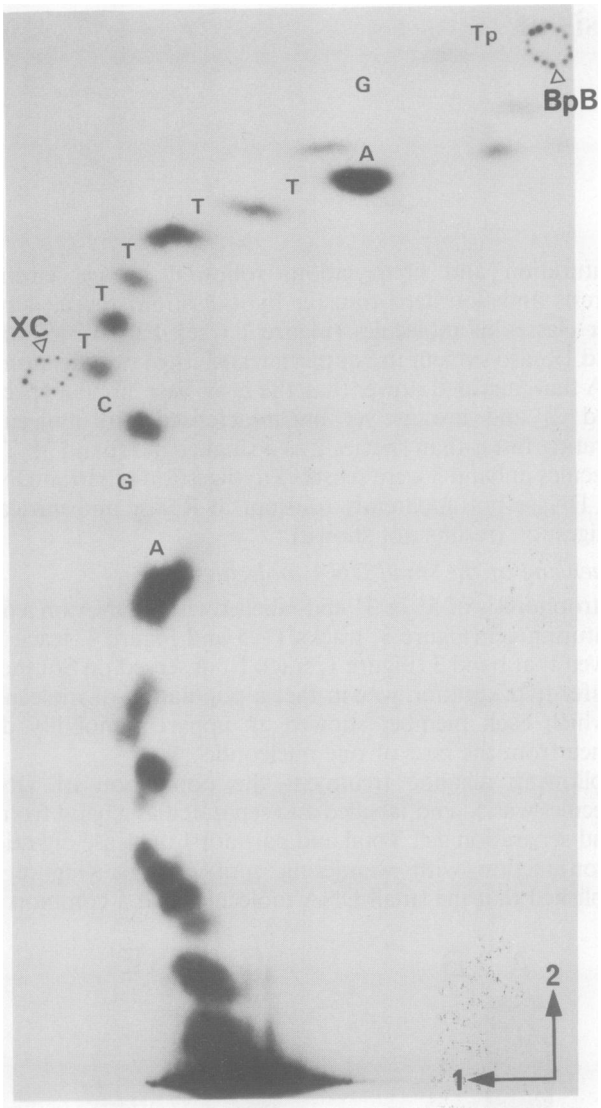

Fig. 2. 'Wandering spot' analysis of the 5′ end-labelled DNA moiety of the small molecules. Fractionation was by electrophoresis on cellulose acetate at pH 3.5 in 7 M urea (1st dimension) and by homochromatography using a 3% homomix (2nd dimension). The positions of the bromophenol blue (BpB) and Xylene Cyanol FF (XC) dye markers are indicated.

deoxyribonucleotide terminus complementary to nucleotide 1202 on the MSV genome sequence (Mullineaux *et al.*, preceding paper).

The difference in length of the molecules represented in the population can be accounted for by: (a) the heterogeneous length of the DNA molecules due to variable 3′ ends; (b) ribonucleotides covalently bound to the 5′ terminus of a proportion of the molecules.

The heterogeneous length of the 3′ end of the DNA molecules. This was confirmed by wandering spot analysis and the observation from Maxam and Gilbert sequencing gels using 3′ end-labelled DNA, of bands obscured by 'echos', with the basic pattern repeated several times but out-of-phase with the strongest signal (results not shown).

The 3′ termini of the small DNA molecules were established by (i) excising six of the predominant bands of 3′ end-labelled DNA (corresponding to the bands 2–7 on Figure 4 track A) from a 6% denaturing gel and characterising their 3′ nucleotide by total spleen phosphodiesterase II and DNase II

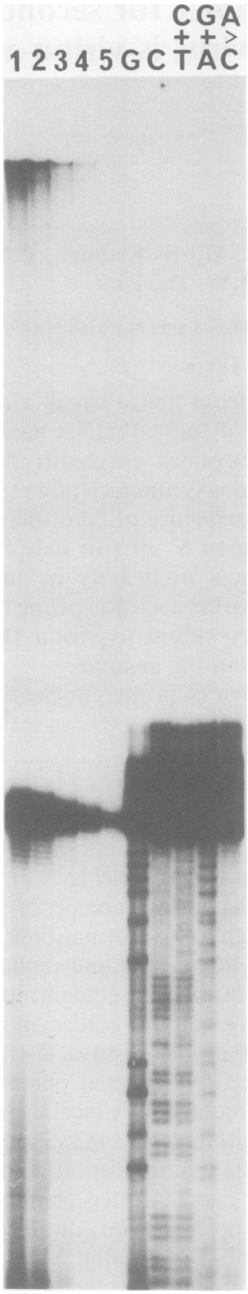

Fig. 3. 6% acrylamide gel: **tracks 1–5** are a dilution series of denatured 5′ end-labelled MSV DNA. **Tracks G, C, C+T, G+A** and **A>C** are Maxam and Gilbert sequencing reactions of the 5′-labelled population of small DNA molecules. The labelled material at the top of **tracks 1–5** is uncharacterised.

digestion, followed by Whatman 3MM paper chromatography in pH 3.5 buffer. All six of the bands 2–7 yielded a labelled thymidine 3′ monophosphate by this nearest neighbour analysis. (ii) By electrophoresis of 5′ end-labelled DNA alongside Maxam and Gilbert sequencing reactions (Figure 3) it was possible to determine the precise length of the members of the small DNA population. This was done by aligning those bands in tracks 1–5 with the Maxam and Gilbert sequencing ladder and by counting back against the sequence determining the length of the more predominant but slower migrating species. By this method the maximal length of the DNA moiety (fragment 5, Figure 4 tracks A and B) was shown to migrate at a position expected of a molecule 81.5

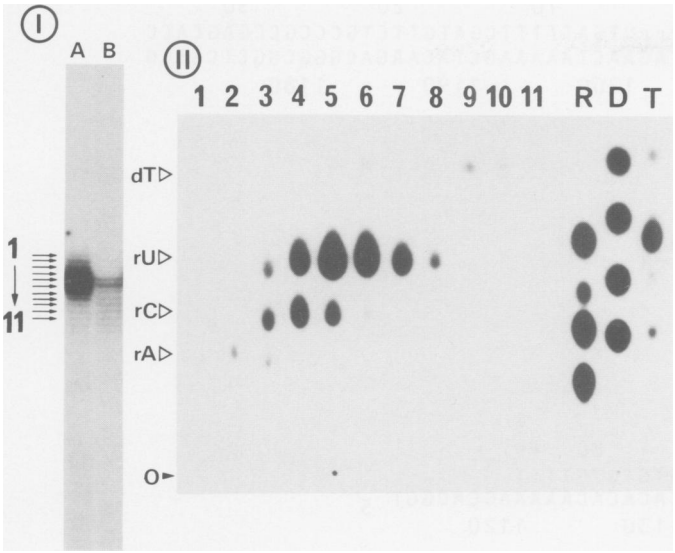

Fig. 4. **(I)** 6% acrylamide gel of 5′-labelled MSV DNA showing the population of small molecules; **(A)** untreated; **(B)** alkaline treated. **(II)** Ascending chromatography on a thin-layer of poly-(ethyleneimine)-cellulose of total P1 nuclease digests of **fragments 1–11 (track A)**. **R** is a total P1 nuclease digestion of tRNA (ribonucleotide 5′ monophosphates G, A, C and U in order of increasing mobility). **D** is a total P1 nuclease digestion of sonicated salmon sperm DNA (deoxyribonucleotide 5′ monophosphates G, A, C and T in order of increasing mobility). **T** is a total P1 nuclease digestion of the 5′-labelled population of small molecules extracted as a single band from a strand sparation gel. O is the origin of the chromatograph. Because of the problem of distinguishing some ribo- and deoxyribo- nucleotides in this manner, the open arrow heads (▷) represent a further experiment in which the base composition of the nucleotide 5′ monophosphates was determined by elution of the spots from the PEI cellulose plate and subsequent paper chromatography (Whatman 3MM) at pH 3.5.

nucleotides in length. However the sequencing reactions leave a 3′ phosphate group which would make the oligonucleotide migrate approximately half a nucleotide ahead of the same molecule bearing a 3′ hydroxyl group (Sollner-Webb and Reeder, 1979). In addition, the Maxam and Gilbert sequencing reactions would eliminate the modified nucleoside. Thus the fragments generated by the sequencing reactions will migrate 1.5 nucleotides faster than the corresponding fragments in tracks 1–5, Figure 3. To confirm this observation (Sollner-Webb and Reeder, 1979) the population of small DNA molecules was run alongside a di-deoxy chain termination sequencing ladder generated from a full-length *Sma*I clone of MSV DNA in M13mp9 (data not shown). The result of this was that the maximal length of the DNA moiety (fragment 5, Figure 4 tracks A and B) had a mobility consistent with it having 79.5 nucleotides. However the fragments in the small DNA population have a 5′ phosphate as a result of kinase labelling and as such might be expected to show an increased mobility of 0.5 nucleotide on the fragments in the dideoxy chain termination reaction ladder, which have the 5′-terminal hydroxyl group of the universal primer (Duckworth *et al.*, 1981). Therefore, allowing for these correction factors the maximal length of the DNA moiety was 80 nucleotides.

Presence of ribonucleotides at the 5′ extremity of the single-stranded DNA. Wandering spot analysis of DNA forms that had not been alkaline treated prior to 5′ end-labelling resulted in a complex pattern consistent with the 5′ end of the molecules having some variation in length (results not

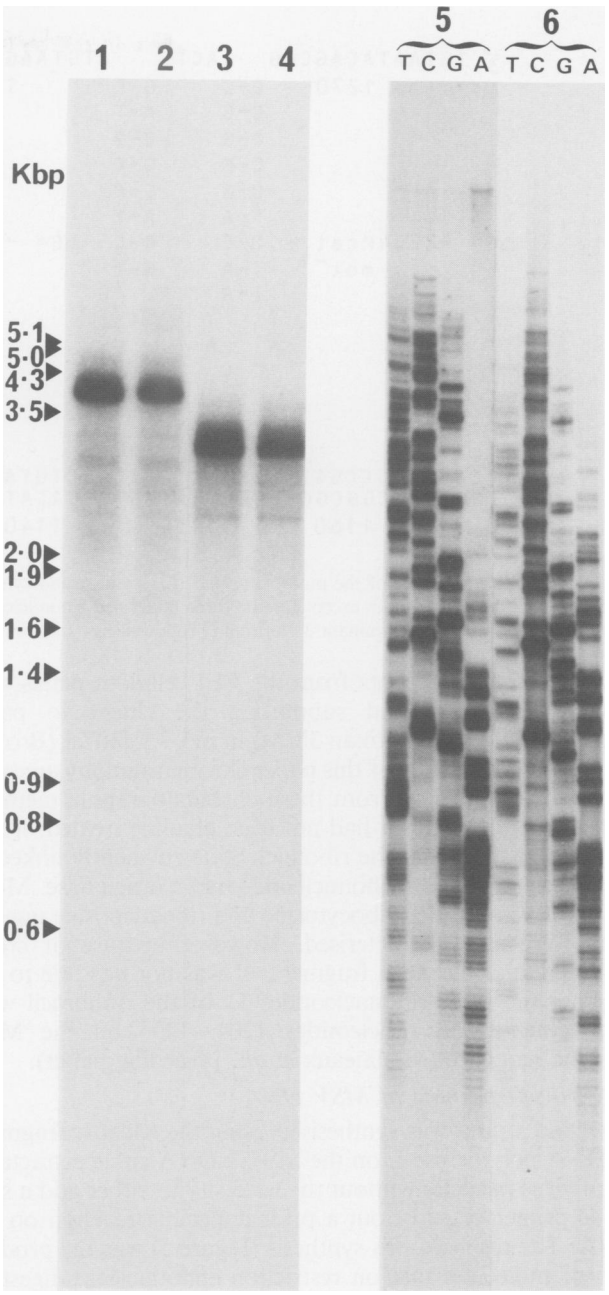

Fig. 5. Self priming of MSV DNA electrophoresed on a 1.4% agarose gel: **(1)** and **(3)** with, **(2)** and **(4)** without an annealing reaction; **(1)** and **(2)** uncut; **(3)** and **(4)** cut with *Xho*I. Numbers indicated (▶) to the left of the gel are phage λ *Hind*III-*Eco*RI size markers in kilobase pairs (kbp). **(5)** and **(6)**: 8% acrylamide gel of self-primed dideoxy chain termination reactions on MSV DNA **(5)** untreated and **(6)** piperidine treated.

shown). Indeed alkaline treatment of 5′ end-labelled MSV DNA resulted in a loss of label from bands 1–4 (Figure 4B) in the population of small molecules. Bands 1–11 (Figure 4A) from non-alkaline treated 5′-labelled DNA were excised from a 6% denaturing gel, digested to completion with P1 nuclease and separated out on PEI cellulose (Figure 4 part II). Figure 4 part II in fact represents a summary of a broader experiment in which PEI thin layer plates were developed with 5′ end-labelled sonicated salmon sperm DNA and tRNA as internal markers. However because of the problem of distinguishing some ribo- and deoxyribo- nucleotides in this manner, nucleotide base composition was also confirmed by

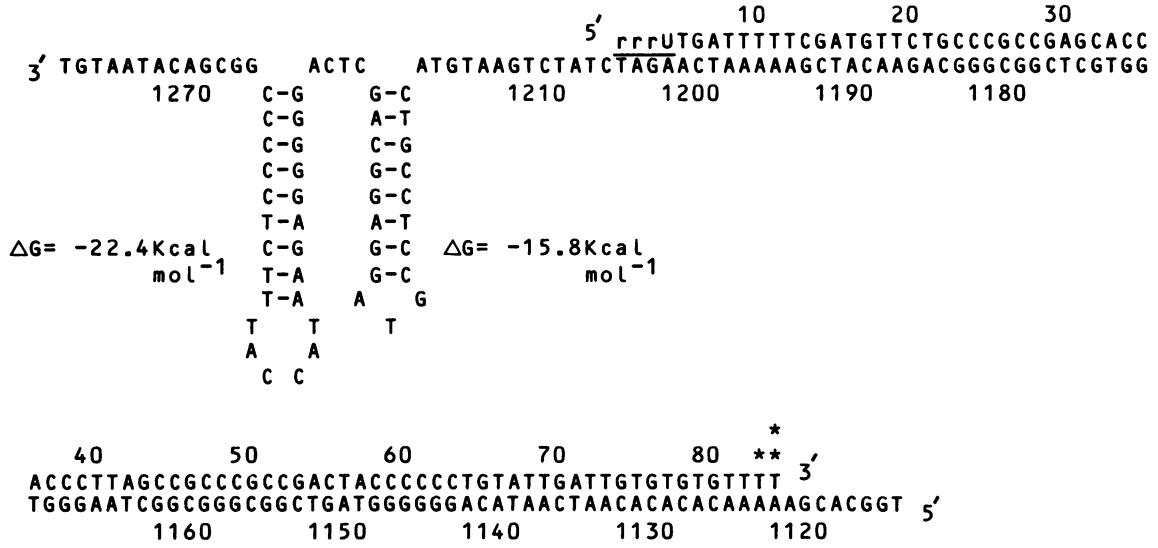

Fig. 6. Nucleotide sequence of the putative primer of second-strand synthesis. The sequence is shown aligned with the viral strand (+) sequence. All the bases are deoxyribonucleotides except for the three undefined ribonucleotides (r) and the ribouracil representing the first four nucleotides of primer sequence. The two most predominant 3′ termini (: followed by *) are indicated.

eluting each of the spots from the PEI cellulose plates with 0.5 M $(NH_4)_2CO_3$ and submitting the eluent to paper chromatography (Whatman 3MM) in pH 3.5 buffer (Brownlee, 1972). The results of this paper chromatography are summarised in Figure 4. From these chromatographic methods on samples that had or had not been alkaline treated, it was possible to show that the ribonucleotide covalently linked to the 5′-terminal deoxyribonucleotide had a uracil base. Molecules with 5′-labelled ribocytosine and riboadenosine nucleotides were also characterised. However because of cross-contamination between fragments it was not possible to say whether the three ribonucleotides 5′ of the ribouracil were complementary to nucleotides 1204–1206 on the MSV genome sequence (Mullineaux *et al.*, preceding paper).

In vitro self-priming of MSV DNA

A second strand was synthesised using the Klenow fragment of DNA polymerase I, on the MSV ssDNA circle extractable from virus particles without the necessity to either add a synthetic primer or carry out a prior annealing reaction on the DNA. The result of this synthesis (Figure 5) was the production of molecules that on restriction endonuclease digestion and gel electrophoresis migrated in a manner expected of linearised full length cccDNA extracted from MSV-infected tissue (Figure 5 tracks 3 and 4).

To establish that this priming reaction occurred at a single position on the genome, chain termination sequencing reactions were undertaken using no added synthetic primer. The result of these reactions and subsequent piperidine treatment was a ladder of readable sequence >100 bases long (Figure 5 sample 6). This sequence was directly complementary to the MSV genome from nucleotides 1102–1003 (Mullineaux *et al.*, preceding paper). Lack of piperidine treatment (Figure 5 sample 5) resulted in bands obscured by 'echos' presumably due to the variable lengths of ribonucleotides on the 5′ terminus of a proportion of molecules.

Discussion

From the variety of techniques used, the sequence of the small DNA molecules, isolated along with the genomic ssDNA circle (Mullineaux *et al.*, preceding paper) from virus particles, is given in Figure 6. Alignment with the MSV plus-strand (Mullineaux *et al.*, preceding paper) established the sequence to be directly complementary to nucleotides 1123–1206, lying almost exclusively in one of the two non-coding regions of the genome (nucleotides 1047–1202).

The common intergenic regions of CLV (Stanley and Gay, 1983) and TGMV (Hamilton *et al.*, 1984) have been implicated in the priming of second-strand synthesis on the basis that a process common to both DNA components might require similar if not identical recognition signals. Interestingly the most stable loop structures in CLV (Stanley and Gay, 1983) and TGMV (Hamilton *et al.*, 1984) are found within this region common to the two circles of each viral genome. MSV DNA however shows no detectable sequence homology with CLV or TGMV (Mullineaux *et al.*, preceding paper). Even so, applying the procedures of Tinoco *et al.* (1973), the sequence shows that MSV contains a number of stable hairpin structures. Two of the nine inverted repeat sequences which have a potential of forming hairpin structures of $\Delta G \geq -14$ kcal/mol (Mullineaux *et al.*, preceding paper) on the MSV genome are positioned 3′ of the small DNA molecules in the viral + sense (Figure 6).

By analogy with bacteriophage ϕX174 single-stranded DNA, stable hairpin secondary structure has been implicated in the host-directed synthesis of double-stranded replicative intermediates (Arai and Kornberg, 1981). More specifically bacteriophage G4 negative strand is initiated at a unique intercistronic site, by the transcription of a hairpin region in the viral DNA (Sims and Dressler, 1978). The indication from the results presented in this paper is that second-strand synthesis of MSV DNA also involves RNA-primed DNA synthesis at a unique site close to stable hairpin secondary structure. This paper records the presence of four ribonucleotides at the 5′ terminus of a DNA molecule with a maximal, though variable, length of 80 nucleotides. It is possible however that these four ribonucleotides observed represent only vestiges of a larger primer that might extend into the region of hairpin secondary structure. Such a primer might have been degraded prior to encapsidation, or as a consequence of experimental procedures.

Although the DNA moiety represents essentially a replica-

tive intermediate one would predict that it would itself serve as a primer for further second-strand synthesis on subsequent re-infection with virus particles, just as it has been shown to do *in vitro* (Figure 5). It has already been established that 75% of the complementary (−) strand of MSV could be involved in protein coding (Mullineaux *et al.*, preceding paper). If early transcription of the MSV minus strand is a prerequisite for successful infection, priming involving this small DNA molecule may have some selective advantage. The reason for the DNA molecules having a maximal length may be a consequence of packaging constraints, but might equally be sequence-specific termination of a DNA polymerase. Even so the variable length of a proportion of the DNA moiety might suggest a level of non-specific interruption of DNA polymerase activity by encapsidation. Further work on the replication of this newly sequenced virus will hopefully reveal the reason for this intermediate in second-strand synthesis being virion-associated.

Materials and methods

Virus and viral DNA purification

MSV (Nigerian isolate) was propagated in *Zea mays* L. var. Golden Cross Bantam, following transmission by *Cicadulina mbila* (Naudé). *Z. mays* L. plants were harvested 20–30 days post-inoculation. Virus particles and viral DNA were isolated essentially by the methods of Harrison *et al.* (1977).

Southern transfer hybridisation analysis

Undenatured and glyoxal-treated DNA (McMaster and Carmichael, 1977) was electrophoresed in 1.0% agarose in 40 mM Tris acetate pH 7.9, 1 mM EDTA and transferred to nitrocellulose as described by Southern (1975). The transfer was probed with ^{32}P-labelled random-primed complementary DNA synthesized to the virus DNA as described by Bisaro and Siegel (1980).

5′ and 3′ end-labelling of the DNA, strand separation and sequence analysis

5′ ^{32}P end-labelling with calf intestinal phosphatase (Boehringer), polynucleotide kinase (PL Biochemicals) and [γ-^{32}P]ATP (Amersham International or New England Nuclear), was performed as described by Maniatis *et al.* (1982). 3′ ^{32}P end-labelling using terminal transferase (BRL) and [α-^{32}P]ddATP (Amersham International) followed the procedure described by Yousaf *et al.* (1984).

To isolate the population of small molecules the end-labelled virion DNA was denatured by boiling in 80% (v/v) deionised formamide, 20 mM EDTA, followed by quick-chilling and loading on a 5% strand-separation gel (Maxam and Gilbert, 1980). To isolate molecules with single nucleotide differences in length, the end-labelled virion DNA was electrophoresed on a 6% (w/v) denaturing polyacrylamide gel (Sanger and Coulson, 1978). DNA fragments were recovered from the crushed gel bands by the method of Maxam and Gilbert (1980). The population of small DNA molecules was sequenced by the partial chemical cleavage protocol of Maxam and Gilbert (1980). All sequencing gels were fixed, dried and subjected to autoradiography as described in Biggin *et al.* (1983).

Characterisation of 5′-terminal structure

The 5′-terminal nucleotides labelled with or without prior alkaline treatment (0.2 N NaOH, 16 h, 37°C), were sequenced by total and partial P1 nuclease (Sigma) digestion. The nucleotide 5′ monophosphates resulting from total P1 nuclease digestion were characterised by chromatography on poly-(ethyleneimine)-cellulose (Randerath and Randerath, 1964) developed in 2% boric acid, 2 M LiCl (2:1). Total P1 digestion of 5′-labelled partial alkaline hydrolysed tRNA and sonicated salmon sperm DNA were used as markers. Nucleotide base composition was confirmed by eluting spots in 0.5 M ammonium carbonate followed by paper chromatography (Whatman 3MM) in pH 3.5 buffer (Brownlee, 1972).

Partial P1 nuclease digestions were characterised by the 'wandering spot' method (Brownlee, 1972) – high voltage electrophoresis on cellulose acetate at pH 3.5 in 7 M urea (first dimension) and by homochromatography on DEAE-cellulose thin layers using a 3% homomix (second dimension; Brownlee, 1972). Sequences were deduced by mobility shift analysis (Siberklang *et al.*, 1977).

Characterisation of the 3′ terminus

Nucleotide 3′ monophosphates were generated by total spleen phosphodiesterase II (Worthington) and DNase II (Sigma) digestion of [α-^{32}P]ddATP-labelled DNA. These were characterised by Whatman 3MM paper chromatography in pH 3.5 buffer (Brownlee, 1972).

Partial P1 nuclease (Sigma) digests of 3′-labelled DNA were characterised by the 'wandering spot' method described above (Brownlee, 1972; Silberklang *et al.*, 1977).

Length distributions of the 3′ termini were also established by electrophoresis of the 5′-labelled MSV DNA alongside Maxam and Gilbert sequencing reactions (Maxam and Gilbert, 1980) on a 6% (w/v) denaturing polyacrylamide gel (Sanger and Coulson, 1978) which was then treated as described above.

Self-priming of the MSV genome

Second-strand synthesis was directed using the Klenow fragment of DNA polymerase I essentially as described by Hong (1981), save that no synthetic primer was added and [α-^{32}P]dATP (Amersham) was incorporated into the reaction mix. In half the reactions the annealing reaction was also omitted. After incubation the products of the reactions were digested with restriction enzymes and electrophoresed on a 1.4% agarose gel in 40 mM Tris acetate pH 7.9, 1 mM EDTA. Marker tracks (phage λ digested with *Eco*RI and *Hind*III) were ethidium bromide stained and photographed. The rest of the gel was dried down and autoradiographed.

MSV DNA was incorporated into the di-deoxy chain termination sequencing method of Sanger *et al.* (1977). However no synthetic primer was added or any annealing reactions undertaken. Half the samples from the sequencing reactions were then treated with 1.0 M piperidine for 30 min at 90°C. Piperidine was eliminated by three cycles of lyophilisation. The products of the sequencing reactions were then subjected to electrophoresis on 8% (w/v) polyacrylamide gels (Sanger and Coulson, 1978) and treated as described above.

Acknowledgements

The authors wish to thank Drs. George Lomonossoff and John Stanley for both technical advice and interesting discussion throughout this work; Drs. Rod Townsend, Simon Covey and Colwyn Thomas for useful discussion. This work was funded by Agrigenetics Research Corporation, and carried out under MAFF licence number 49/152.

References

Arai,K. and Kornberg,A. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 69-73.
Biggin,M.D., Gibson,T.J. and Hong,G.F. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 3963-3965.
Bisaro,D.M. and Siegel,A. (1980) *Virology*, **107**, 194-201.
Brownlee,G.G. (1972) *Determination of Sequences in RNA*, published by North Holland, Amsterdam.
Duckworth,M.L., Gait,M.J., Goelet,P., Hong,G.F., Singh,M. and Titmas, R.C. (1981) *Nucleic Acids Res.*, **9**, 1691-1706.
Hamilton,W.D.O., Bisaro,D.M. and Buck,K.W. (1982) *Nucleic Acids Res.*, **10**, 4901-4912.
Hamilton,W.D.O., Stein,V.E., Coutts,R.H.A. and Buck,K.W. (1984) *EMBO J.*, **3**, 2197-2205.
Harrison,B.D., Barker,H., Bock,K.R., Guthrie,E.J., Meredith,G. and Atkinson,M. (1977) *Nature*, **270**, 760-762.
Hong,G.F. (1981) *Biosci. Rep.*, **1**, 243-252.
Ikegami,M., Haber,S. and Goodman,R.M. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 4102-4106.
Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning, A Laboratory Manual*, published by Cold Spring Harbor Laboratory Press, NY.
Maxam,A.M. and Gilbert,W. (1980) *Methods Enzymol.*, **65**, 499-560.
McMaster,G.K. and Carmichael,G.C. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 4835-4838.
Randerath,K. and Randerath,E. (1964) *J. Chromatogr.*, **16**, 111-125.
Sanger,F. and Coulson,A.R. (1978) *FEBS Lett.*, **87**, 107-110.
Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463-5467.
Silberklang,M., Gillum,A.M. and RajBhandary,U.L. (1977) *Nucleic Acids Res.*, **4**, 4091-4108.
Sims,J. and Dressler,D. (1978) *Proc. Natl. Acad. Sci. USA*, **75**, 3094-3098.
Sollner-Webb,B. and Reeder,R.H. (1979) *Cell*, **18**, 485-499.
Southern,E.M. (1975) *J. Mol. Biol.*, **98**, 503-517.
Stanley,J. and Gay,M.R. (1983) *Nature*, **301**, 260-262.
Sunter,G., Coutts,R.H.A. and Buck,K.W. (1984) *Biochem. Biophys. Res. Commun.*, **118**, 747-752.
Tinoco,I., Borer,P.N., Dengler,B., Levine,M.D., Uhlenbeck,O.C., Crothers, D.M. and Gralla,J. (1973) *Nature, New Biol.*, **246**, 40-41.
Tischer,I., Gelderblom,H., Vettermann,W. and Koch,M.A. (1982) *Nature*, **295**, 64-66.
Yousaf,S.I., Carroll,A.R. and Clarke,B.E. (1984) *Gene*, **27**, 309-313.

Received on 19 September; revised on 19 October 1984