Biologically active transcripts of a large satellite RNA from arabis mosaic nepovirus and the importance of 5' end sequences for its replication

Y. Y. Liu, J. I. Cooper, D. Coates and G. Bauer

1NERC, Institute of Virology and Environmental Microbiology, Mansfield Road, Oxford OX1 3SR, U.K., 2Department of Pure and Applied Biology, University of Leeds, Leeds LS2 9JT, U.K. and 3Max-Planck Institute for Bio-Physical Chemistry, Nikolaufberg, 3400 Goettingen, Germany

Synthetic transcripts of a satellite RNA associated with a lilac isolate of arabis mosaic nepovirus (ArMV) were made from cDNA clones. Transcripts having either six (M1R) or 29 (M3R) extra nucleotides at their 5' ends replicated in the presence of ArMV genomic RNA in manually inoculated Chenopodium quinoa plants, even though M1R also differs from the native sequence at nucleotide position 2. Transcript 12R, which has 11 guanosyl residues and 27 other nucleotides not present in the natural satellite RNA at its 5' end, and also lacks the two 5'-terminal nucleotides (UA), replicated inefficiently, both in transformed tobacco plants and in plants that had been manually inoculated. Transcripts from another construct (M2R) lacking eight 5'-terminal bases of the native sequence did not multiply in plants. Each of these transcripts directed the in vitro synthesis of a protein (Mr 39K) encoded by satellite RNA, although 12R was the least efficient message. Analysis of the 5'-terminal sequences in progeny RNA from M1R showed that the non-native bases were removed and the second nucleotide corrected, suggesting that VPg plus a few initial 5'-terminal bases might serve as a primer for plus-strand synthesis of this satellite RNA. When M1R was inoculated with genomic RNAs from ArMV of ash or ivy, the transcripts replicated and were encapsidated. However, when the same amounts of M1R were inoculated with genomic RNAs of ArMV from hop or sugar-beet, progeny of the transcripts were not detected either in virions or in plants. Less surprisingly, this RNA transcript did not multiply in the presence of dogwood mosaic, strawberry latent ringspot, grapevine fanleaf or cherry leaf roll nepoviruses.

Introduction

Virions of several nepoviruses encapsidate extragenomic RNA species (satellite RNAs) which depend for their replication upon genomic RNA, with which they share few if any nucleotide sequences. Murant & Mayo (1982) and Fritsch & Mayo (1989) have recognized two classes of satellite RNAs associated with nepoviruses: large satellite RNAs of > 1 kb and small satellite RNAs of about 0.3 kb.

The large satellite RNA from a lilac isolate of ArMV (ArMV-L; Liu et al., 1990, 1991) is a linear molecule of 1104 nucleotides which, like other large satellite RNAs of nepoviruses, structurally resembles the genomic species in having 3' polyadenylation and a 5'-linked protein (VPg). Furthermore, the 5'-terminal 11 nucleotides of this satellite RNA (UAUGAAAAAUU) fit the 5' end consensus sequence of nepoviral RNAs noted by Fuchs et al., (1989). ArMV satellite RNA encodes a polypeptide of 360 amino acids with an Mr of 39K (Liu et al., 1990), which is the major translation product in rabbit reticulocyte lysates (Hellen, 1987). The multiplication of ArMV-L satellite RNA has little, if any, effect on the accumulation of the helper virus genome in plants, yet in some plant species it modified symptoms induced by the helper virus (Liu et al., 1991). To study the replication and encapsidation of ArMV-L satellite RNA, the role of the protein it encodes, and to define essential sequences conferring biological properties, we have constructed a variety of clones with deletions or point mutations from which synthetic transcripts have been made. Here, we report the preparation and biological activities of synthetic ArMV-L satellite RNA transcripts with differing 5'-terminal sequences.

Methods

Virus isolates. ArMV-L and a satellite-free derivative of it (ArMV-SF) were the sources of the satellite or viral genomic RNA used in these experiments.

Other ArMV isolates used were from sugar-beet (AB10; Harrison, 1958), ivy, ash (Cooper & Sweet, 1976; Liu et al., 1991), olive (supplied...
by G. Gallitelli, Bari, Italy; Savino et al., 1979) and hop (supplied by M. F. Clark, Horticulture Research International, East Malling, U.K.; Davies & Clark, 1983).

In addition, dogwood mosaic nepovirus (DogMV; supplied by O. W. Barnett, Clemson, S.C., U.S.A.; Barnett et al., 1989), strawberry latent ringspot nepovirus (SLRV-T39, supplied by the Scottish Crop Research Institute, Invergowrie, U.K.), cherry leaf roll nepovirus (CLR-V-I2, from birch; Cooper & Atkinson, 1975) and an isolate of grapevine fanleaf virus (GFLV), from O. W. Barnett, were tested for their ability to support replication and encapsidation of the ArMV satellite RNA.

Virus propagation and preparation of virions or viral RNA. Virus was propagated in Chenopodium quinoa or Nicotiana megalosiphon. Virions were purified from systemically invaded leaves as described by Liu et al. (1990). RNA was extracted from virions using phenol and phenol/chloroform in the presence of 1% (w/v) SDS followed by ethanol precipitation in 0.3 M-sodium acetate (Palukaitis, 1984).

**Construction of cDNA clones.** cDNA representing the satellite RNA was first cloned into the PstI site of the plasmid pBR322 to produce A3/3, which lacked only the two 5'-most nucleotides (UA) of the native satellite RNA. After PstI digestion, the inserted sequence was cloned into the PstI site of pSPT19 (a transcription vector from Pharmacia) to facilitate transcription of the A3/3 sequence by T7 DNA-directed RNA polymerase. Transcripts from this construct (pSPT19-12) had runs of G or C at the 5' and 3' termini respectively and, to eliminate these, a polymerase chain reaction (PCR) was done. Two synthetic oligodeoxyribonucleotide primers (primers 1 and 2; Fig. 1a) were made with an Applied Biosystems 380B synthesizer and, after deprotection at 55 °C overnight, each was purified by electrophoresis in gels containing 20% (w/v) acrylamide/8 M-urea. Each primer comprised 10 or 11 non-complementary nucleotides (Fig. 1a) and 18 residues complementary to the A3/3 sequence. The non-complementary nucleotides introduced an EcoRI site at the 5' end and a BamHI site at the 3' end. The PCR reaction mixture (50 μl) contained 50 ng of A3/3 insert cDNA, 5 μl of 10 × PCR buffer (containing 100 mM-Tris–HCl pH 8.5, 50 mM-KCl, 1 mM-MgCl2, 200 μM-dNTPs, 1.5 U Taq polymerase) and 1 μl of each primer (10 μM).

**Fig. 1.** (a) A schematic view of how M1 A3/3 was constructed. Bold letters indicate non-satellite RNA sequence. pUC18MCS, multiple cloning sites in the pUC18 plasmid. (b) Nucleotide sequences of the 5' (i) and 3' (ii) termini of M1 A3/3.
EcoRI, of 94°C for 1 min, 40°C for 1 min and 70°C for 3 min, and the chloroform treatment. The PCR product was digested with EcoRI and BamHI, cleaned using a GeneClean II kit (Bio 101) and ligated into EcoRl/BamHI-cut pT7T3 18U (Pharmacia), producing clone M1, A3/3 which was under the control of the T7 promoter (Fig. 1a). The nucleotide sequences at both ligation points were checked by direct sequencing of the plasmid DNA (Murphy & Kavanagh, 1988) using the M13 universal sequencing primer and a synthetic primer (5‘ AG-GAAACAGCTATGACCATG 3’) complementary to the sequence containing the A3/3 sequence (as judged by slot blot hybridization to sequence-specific radiolabelled probes) also expressed neomycin phosphotransferase II (NPT II), which was detected by the method of McDonnell et al. (1987).

**Results**

**In vitro transcription and synthesis**

In vitro transcription of BamHI-linearized pSPT19-12, M1 A3/3 and M2 A3/3, or PstI-linearized M1 A3/3 under the control of the T7 promoter yielded transcripts designated 12R, M1R, M2R and M3R, respectively. Each differed from the native sequence in containing nucleotides derived from transcription vectors and from one another in the nucleotide sequences at their 5‘ and 3‘ termini (Fig. 2). Thus, 12R lacks two 5‘-terminal residues (UA) of the native satellite RNA and contains 11 extra G or C residues at the 5‘ and 3‘ termini respectively (Liu et al., 1990); M1R differs from the native sequence in having a C rather than an A in the penultimate position at the 5‘ end. Since M1R was derived from M1R by cleavage at the unique XmnI site (position 8, Fig. 1a), thereby breaking the VpG consensus sequence identified by Fuchs et al. (1989), M2R lacked these eight nucleotides at the 5‘ end. M1R contains the authentic satellite RNA sequence.

The yield of M1R, M2R or M3R transcripts was in the range of 10 to 40 μg/μg template DNA. However, under the same transcription conditions the production of 12R was 2 to 10 μg/μg template. The lower efficiency of transcription of pSPT19-12 may be due to the Y-terminal poly(A) of the native satellite RNA and contains 11 extra G or C residues at the 5‘ and 3‘ termini respectively (Liu et al., 1990); M1R differs from the native sequence in having a C rather than an A in the penultimate position at the 5‘ end. Since M1R was derived from M1R by cleavage at the unique XmnI site (position 8, Fig. 1a), thereby breaking the VpG consensus sequence identified by Fuchs et al. (1989), M2R lacked these eight nucleotides at the 5‘ end. M1R contains the authentic satellite RNA sequence.

In vitro transcription of pBR322 and pUC32 linearized with BamHI, HindIII and EcoRI, respectively. Each differed from the native sequence in containing nucleotides derived from transcription vectors and from one another in the nucleotide sequences at their 5‘ and 3‘ termini (Fig. 2). Thus, 12R lacks two 5‘-terminal residues (UA) of the native satellite RNA and contains 11 extra G or C residues at the 5‘ and 3‘ termini respectively (Liu et al., 1990); M1R differs from the native sequence in having a C rather than an A in the penultimate position at the 5‘ end. Since M1R was derived from M1R by cleavage at the unique XmnI site (position 8, Fig. 1a), thereby breaking the VpG consensus sequence identified by Fuchs et al. (1989), M2R lacked these eight nucleotides at the 5‘ end. M1R contains the authentic satellite RNA sequence.

The yield of M1R, M2R or M3R transcripts was in the range of 10 to 40 μg/μg template DNA. However, under the same transcription conditions the production of 12R was 2 to 10 μg/μg template. The lower efficiency of transcription of pSPT19-12 may be due to the 5‘-terminal G residues. Analysis of the transcripts by non-denaturing agarose gel electrophoresis showed product heterogeneity, but the transcripts were predominantly of two sizes (Fig. 3). The RNA products may have included dimers because the transcripts migrated as a single species after they had been heated at 65°C for 5 min.

**Nucleotide sequencing of 5‘ terminal of progeny RNA.** This was done with fluorescent dideoxynucleotide terminators and a DuPont Genesis 2000 DNA sequencer (Bauer, 1990). In this work, a synthetic oligomer of 17 nucleotides (5‘ TATAATTCGCGCTTATG 3’) complementary to residues 25 to 41 of the satellite RNA was used with virion-derived RNA (RNA-1, RNA-2 and satellite RNA).

**Transformation of tobacco leaf discs.** The A3/3 insert excised from A3/3 in pBR322 using PstI was first cloned into the PstI site in pUC4K (Pharmacia) and released therefrom by cleavage with BamHI. Thereafter, the fragment was ligated into the unique BamHI site in pRok1, a derivative of pBin19 (kindly supplied by D. Baulcombe, Sainsbury Laboratory, Norwich, U.K.; Bevan, 1984). Standard methods (Horsch et al., 1985) were used to transform leaf discs of N. tabacum cv. xanthi-nc with Agrobacterium tumefaciens (LBA 4404) containing the pRok1 construct. Putative transformants, identified by their ability to grow and form roots in the presence of kanamycin (100 μg/ml), were maintained. A total of 15 independent tobacco clones containing the A3/3 sequence (as judged by slot blot hybridization to sequence-specific radiolabelled probes) also expressed neomycin phosphotransferase II (NPT II), which was detected by the method of McDonnell et al. (1987).
Translation of the transcripts

In rabbit reticulocyte lysates, each transcript directed the synthesis of a 39K protein indistinguishable in Mr from that encoded by the native satellite. When an equivalent amount of each transcript was translated, 12R produced a very much smaller quantity of the 39K protein than either M1R or M2R (Fig. 4).

Biological activity of the transcripts

Transcripts (12R, M1R, M2R or M3R) plus RNA-1 and RNA-2 from ArMV-SF were inoculated into N. megalosiphon or C. quinoa plants and 7 to 10 days later total RNA was extracted from systemically invaded foliage and used for slot-blot hybridization (Fig. 5). A nick-translated plasmid (32P-labelled) containing the complete satellite sequence was used as a probe. Crude sap from tip leaves of the inoculated plants was also used as inoculum for other C. quinoa plants. After each such passage cycle, virions were used to provide RNA which was analysed by agarose gel electrophoresis (Fig. 6). In five of six trials with M1R, satellite RNA was apparent after only one passage cycle (Fig. 5, lane 2) and it seemed to be encapsidated when tested after the second replication cycle (Fig. 6, lane 1). In the sixth trial, progeny of M1R were detected only after the third passage cycle. Even though M2R was inoculated with the genomic RNA at the same concentration as M1R, progeny of M2R were not detected even after four passage cycles in C. quinoa plants. M3R was replicated as efficiently as M1R; in three experiments satellite RNA was detected in systemically invaded foliage of the M3R-inoculated C. quinoa plants and in virions purified after a second passage cycle.

The transcript designated 12R was inefficiently produced but, in four of seven trials in which 5 μg of linearized pSPT19-12 was used as template, satellite RNA was detected in virions purified after the second or
ArMV satellite RNA transcripts

Fig. 6. Northern transfer hybridization of RNA from virions purified from C. quinoa foliage containing second passage derivatives of the following: RNAs 1 and 2 from ArMV-SF plus M₁₆R (lane 1) or 12R (lane 2). Lane 3 contained derivatives of M₁₆R plus RNA-1 and -2 from the isolate of ArMV from ash. Lane 4 RNA derived from M₁₆R plus the isolate of ArMV from ivy.

third passage cycle in C. quinoa plants (Fig. 6, lane 2). In the other three tests, no evidence of satellite RNA replication was detected after four cycles of replication in C. quinoa plants. When RNA-1 and RNA-2 were inoculated with M₁₆A₃/₃ DNA template, only the genomic species were encapsidated.

Tests for support of M₁₆R replication by nepoviruses other than ArMV-L

The hop isolate of ArMV used contains a 300 nucleotide RNA satellite (Davies & Clark, 1983; Kaper et al., 1988), but the ArMV isolates from ivy, ash or sugar-beet do not contain extragenomic RNAs (Liu et al., 1991; unpublished data). When M₁₆R transcribed from 2 μg of linearized template DNA was inoculated with RNA-1 and RNA-2 from the ivy or ash isolates of ArMV, satellite RNA was detected (by slot blot hybridization) in extracts from the tip leaves of C. quinoa plants and was also found to be encapsidated in tests done after the second passage cycle (Fig. 6, lanes 3 and 4). However, when the same amount of this transcript was inoculated with the genomic RNAs from the hop or sugar-beet isolates of ArMV, or from DogMV, SLRV, GFLV or CLRV, no evidence of satellite replication was detected after four passage cycles. This was consistent with previous observations (unpublished data) in which the ash and ivy isolates of ArMV supported multiplication of virion-derived satellite RNA (from ArMV-L), whereas the sugar-beet isolate of ArMV, SLRV and CLRV did not.

Terminal sequence of progeny

Analysis of the 5' terminal sequence of the progeny of M₁₆R (Fig. 7) showed that the non-satellite sequence at the 5' terminus had been lost and the second base (C) had reverted to A (T in Fig. 7). There was sequence heterogeneity at position 20 (in the coding region) but this did not alter the product (GAU and GAC encode Asp). The sequence heterogeneity in the non-coding region at positions 3 to 5 may be attributable to the existence of a minor satellite RNA population with four

GGGAATCCATAGAAATTTCATA
( G C A T )

Fig. 7. Raw data from a 12% polyacrylamide gel showing sequences at the 5' end of cDNA representing progeny of M₁₆R replication. Sequence heterogeneities are shown in parentheses.
A residues in positions 5 to 8 (because they were detected in the native satellite RNA). Alternatively or additionally, the apparent heterogeneity may be a result of reverse transcriptase ‘slipping’ at the A residue cluster.

**Satellite RNA transcription from transgenic tobacco**

Northern transfer hybridization with total nucleic acid extracts from kanamycin-resistant tobacco plants in tissue culture did not reveal the presence of satellite RNA nor was it detected when these plants were transferred into potting compost and inoculated with RNA-1 plus RNA-2 from ArMV-SF. However, the satellite RNA was detected after virus from these plants had been passaged in *C. quinoa* plants. This was shown by agarose gel electrophoresis and Northern analysis of RNA from virions purified from *C. quinoa* plants after the second passage. The presence of terminal guanosines or cytosines in the transcript may explain the inefficient expression of the satellite RNA in these transgenic tobacco plants. The construct we used for plant transformation would be expected to produce a transcript closely resembling the biologically inefficient 12R. When untransformed tobacco plants or another transgenic line (expressing GUS) were inoculated with RNA-1 and RNA-2, only these genomic species were detected, even after passage in *C. quinoa* plants. Presumably these data indicate that the 35S promoter was expressing the satellite RNA only inefficiently, but that the few initial molecules of transcript were amplified when the genomic RNAs of the helper virus replicated in the transgenic tissue.

**Discussion**

Our observation that, in the presence of an appropriate helper virus, synthetic transcripts of a satellite RNA replicate is, in itself, not exceptional; analogous data have been recorded for the satellite RNAs of tomato black ring nepovirus (Greif et al., 1990), tobacco ringspot nepovirus (Gerlach et al., 1986), turnip crinkle carnovirus (Simon & Howell, 1987) and cucumber mosaic cucumovirus (Collmer & Kaper, 1986). However, in most recorded instances the biological activity of synthetic transcripts has drastically diminished with increasing numbers of ‘extra’, non-viral nucleotides at the 5’ end of the molecule (Dawson et al., 1986; Janda et al., 1987; Ziegler-Graff et al., 1988; Masuta et al., 1988; Heaton et al., 1989; Eggen et al., 1989). In our experiments by contrast, there was no noticeable difference in the biological activity of M1R and M2R, which have either six or 29 extra nucleotides (derived from the cloning vectors) at the 5’ terminus. However, these constructs were less active than native (virion-derived) satellite RNA (Fig. 5).

Our experience when translating the synthetic satellite RNA 12R was broadly in line with that of Gough et al. (1985) or Jobling & Gehlke (1987). The presence of 11 G residues 5’ to and separated by a further 12 nucleotides from the first AUG sequence was probably responsible for the markedly decreased translation efficiency of this construct. Perhaps the inefficient transcription in vitro and poor replication in vivo of this construct were other manifestations of the same phenomenon. The ability of these replicating RNA sequences to edit-out 5’ defects (e.g. the deletion of the six extra nucleotides in M1R) is noteworthy even though others have recorded analogous changes. Analysis of the 5’-terminal sequence of progeny RNA derived from M1R showed heterogeneity in the region which encodes the 39K protein, but did not change the amino acid encoded. Perhaps this indicates the importance of this product in satellite RNA replication. With regard to the satellite RNA from ArMV-L, it seems that the first two nucleotides are not essential for replication because both 12R (lacking these two) and M1R (having an incorrect residue at the second position) are biologically active. Interestingly, both recombinant coxsackievirus B3 (CVB3) cDNA and transcripts of CVB3 lacking the two 5’-most uridine residues present in authentic virion RNA, are infectious (Klump et al., 1990), and during replication in transfected cells the authentic 5’-terminal sequence is produced. The finding of 5’-terminal self-correction with CVB is consistent with the suggestion that VPg-pUpU is a component of the initiation complex for poliovirus RNA (Crawford & Baltimore, 1983; Takegami et al., 1983; Takeda et al., 1986; Wimmer et al., 1987). Our experience with satellite RNA, though not identical, is perhaps analogous and may imply that VPg plus a few nucleotides at the 5’ terminus (which possibly arise as a result of the replication of the ArMV genomic RNAs) serve as a primer for plus-strand synthesis, independent of the complementary terminal residues on the minus-strand template. If true, this might explain why M2R did not seem to replicate in planta. It is plausible that initiation of plus-strand synthesis could not occur because the deletion of eight nucleotides in the 5’ end consensus sequence prevented the putative primer from interacting with the 5’ terminus of the minus-strand template RNA.

In common with other satellite RNAs (Schneider, 1977; Kaper et al., 1978; Murant & Mayo, 1982; Ponz et al., 1987), the satellite RNA from ArMV-L is somewhat helper virus-specific. Thus, whereas the helper activity of the ArMV isolates from ivy or ash seemed to be no worse than that of the homologous helper (ArMV-SF), other ArMV isolates having few if any different antigenic
determinants (e.g. a sugar-beet isolate, AB10) were incompetent. This suggests that the genetic determinants coding region of RNA-2. The availability of biologically active transcripts will facilitate the search for those parts of the helper virus sequences that are essential for satellite RNA replication/encapsulation.

References


(Received 25 June 1991; Accepted 20 August 1991)