Non-viral heterogeneous sequences at the 5' ends of tomato spotted wilt virus mRNAs

Richard Kormelink,* Frank van Poelwijk, Dick Peters and Rob Goldbach

Department of Virology, Binnenhaven 11, 6709 PD Wageningen, The Netherlands

Subgenomic messenger RNAs transcribed from the tomato spotted wilt virus (TSWV) S RNA segment were partially purified from total RNA extracts of TSWV-infected *Nicotiana rustica* and analysed by primer extension analysis. The data obtained show the presence of non-viral sequences, 12 to 20 nucleotides in length, at the 5' ends of the N and NSs mRNAs, indicating a cap-snatching mechanism for the initiation of transcription. This is the first report of a plant virus using such a mechanism for transcription of the viral genome.

Recent studies on the genome of tomato spotted wilt virus (TSWV) have revealed that this virus represents a member of the Bunyaviridae, being unique in its property to infect plants. Mainly based on its distinct host range TSWV has been classified into the newly created genus tospovirus, within this large family of arthropod-borne viruses (Francki et al., 1991). For both the S RNA and L RNA genomic segments, complete nucleotide sequences have become available. The L RNA segment [8897 nucleotides (nt) long] is entirely of negative polarity, encoding the putative (331-5K) viral transcriptase (de Haan et al., 1991). The S RNA segment (2916 nt long), like that of phleboviruses (Ihara et al., 1984; Marriott et al., 1989; Simons et al., 1990; Giorgi et al., 1991), is ambisense and encodes the N protein of Mr 28.8K and a non-structural protein (NSs) of 52-4K (de Haan et al., 1990; Kormelink et al., 1991).

Expression of the L RNA has been shown to occur via the synthesis of a genome-sized mRNA (Kormelink et al., 1992). The two open reading frames (ORFs) in the ambisense S RNA are expressed from two subgenomic mRNAs that are transcribed from opposite strands and terminate at the central, intercistronic region, most probably in a long A-U-rich hairpin (de Haan et al., 1990; Kormelink et al., 1991).

For several animal-infecting members of the Bunyaviridae the process of initiation and termination of transcription has been further studied by characterizing the 5' and 3' ends of viral mRNAs. These studies have demonstrated the presence of short heterogeneous non-viral sequences at the 5' ends of the mRNAs, indicating that the viral transcriptase utilizes RNA primers to initiate transcription (Bishop et al., 1983; Patterson & Kolakofsky, 1984; Eshita et al., 1985; Ihara et al., 1985; Collett, 1986; Gerbaud et al., 1987; Bouloy et al., 1990; Simons & Pettersson, 1991). These primers are generated from capped host mRNA species by a process referred to as 'cap-snatching', i.e. the 5'-terminal sequence of a cellular mRNA is cleaved off by an endonuclease and subsequently used to initiate transcription on the viral genome (Braam et al., 1983; Ulmanen et al., 1981). Less is known about the termination of transcription, but some typical structural features, i.e. palindromic sequences or hairpin structures, have been found close to the sites where termination occurs (Bouloy et al., 1990; Simons & Pettersson, 1991). How these features are involved in transcription termination has so far remained unknown.

To determine whether TSWV (differing from all other bunyaviruses in being completely adapted to multiplication in plant cells) also utilizes cap-snatching to initiate transcription, the 5' ends of the N and NSs mRNAs were analysed. To allow unequivocal analysis of the 5' ends of the N and NSs subgenomic mRNAs were analysed. Recent studies on viral RNA synthesis in TSWV-infected *Nicotiana rustica* revealed the presence of low amounts of the S-specific mRNAs and viral complementary (vc) S RNA strands, relative to the full-length viral (v) S RNA (Kormelink et al., 1992). To allow unequivocal analysis of the 5' ends of the N and NSs mRNAs, and of the 5' end of the S vcRNA, it was necessary to purify these RNA species to a certain extent. To this end, total RNA from TSWV-infected *N. rustica* was extracted 8 days post-inoculation (p.i.) according to De Vries et al. (1982), and resolved on 15 to 22.5% sucrose gradients (Ulmanen et al., 1981b; Bishop et al., 1983). The fractions collected were analysed for their absorbance at 254 nm (A254) (Fig. 1a), and for their...
Fig. 1. Sedimentation analysis and separation of RNA species from TSWV-infected N. rustica plants, 8 days p.i. Five-hundred μg of total RNA from infected tissues was layered on a 15 to 22.5% sucrose gradient. (a) Absorbance profile of RNA fractions collected from the gradient. Sedimentation was from left to right. Two μg RNA from each fraction was resolved on a 1% agarose gel, transferred to a Genescreen membrane and hybridized to riboprobe S2-v, specific for the N gene. (b) RNA pellets obtained after centrifugation through a CsCl cushion (as described in the text), enriched for the N mRNA (lane 4), and for the S vcRNA (lane 3), analysed on a Northern blot using riboprobes S1-vc (lanes 1 and 2) and S2-v (lanes 3 and 4). Lane 1 contains purified nucleocapsid RNA which was used for primer extension analysis of the 5' end of the S vcRNA strands. The riboprobes were prepared according to Kormelink et al. (1992).

RNA content. For the latter, RNA samples were resolved on a 1% agarose gel (Bailey & Davidson, 1976), transferred to a Genescreen membrane and hybridized to strand-specific probes corresponding to the N- (Fig. 1a) and NS\textsubscript{S} (data not shown) coding regions in S RNA. The relevant fractions, enriched for the N and NS\textsubscript{S} mRNA (fractions 3 and 4), and enriched for the S vcRNA strands (fractions 7 and 8), were pooled and the RNA was precipitated after the addition of 0.1 volume 3 M-sodium acetate and 2.5 volumes of ethanol. The RNA was resuspended in GIT buffer (4 M-guanidinium isothiocyanate, 25 mM-sodium acetate pH 6, 0-12 M-2-mercaptoethanol) and subsequently pelleted through a 5.7 M-CsCl cushion according to Davis et al. (1986). The RNA pellets obtained were analysed on Northern blots for the contents of full-length S vcRNA, and the subgenomic mRNAs (Fig. 1b). Ten μg of the selected RNA samples, and 2 μg of purified nucleocapsid RNA were used for primer extension experiments. For this purpose, two oligonucleotides, one complementary to the vc sense RNA at positions 2834 to 2852 (primer pN) and another complementary to the v sense RNA at positions 32 to 51 (primer pNS\textsubscript{S}) of the S RNA (Fig. 2a), were synthesized. Both primers were labelled at their 5' end using [γ-\textsuperscript{32}P]ATP and T4 polynucleotide kinase and subsequently purified from an 8% sequencing gel. The primers were mixed with the RNA in 10 μl annealing buffer (250 mM-KCl, 10 mM-Tris-HCl pH 8.3), heated at 90 °C for 2 min and subsequently incubated at 37 °C for 5 min. Ten μl of the reverse transcriptase mix (100 mM-Tris-HCl pH 8.3, 10 mM-MgCl\textsubscript{2}, 10 mM-DTT, 50 units of Moloney murine leukaemia virus reverse transcriptase, 20 units of RNase, 1 μM-dNTP) was added and the reaction incubated at 37 °C for 30 min to synthesize run-off copies of the upstream sequences in the S-specific RNAs.

From the sequence data of the S RNA segment (de Haan et al., 1990) it was predicted that primer pNS\textsubscript{S} should be extended for 31 nt on the v sense S RNA template, resulting in a run-off product of 51 nt. Primer extension synthesis on nucleocapsid RNA using pNS\textsubscript{S}, and subsequent analysis of the run-off products on an 8% sequencing gel, indeed showed the presence of the expected band (Fig. 2b, lane 1). Similarly, when primer pN was extended on vc sense S RNA, a run-off product with an expected size of 85 nt was found (Fig. 2b, lane 3). Analyses of NS\textsubscript{S} and N mRNA-enriched fractions revealed an additional ladder of products 12 to 20 nt larger in size than the 51 and 85 nt products, respectively (Fig. 2b, lanes 2 and 4). This indicates the presence of extra sequences, heterogeneous in length, at the 5' ends of the TSWV-specific mRNAs. As these additional terminal sequences do not appear at the 5' end of the nucleocapsid S RNA, except for a distinct extension product of 63 nt, it is unlikely that the extra sequences in both the N and NS\textsubscript{S} mRNA are templated by the viral RNA. They probably originate from host mRNAs, and are utilized by the viral RNA transcription machinery to
Fig. 2. (a) Genetic organization and expression of the TSWV S RNA segment. Nucleotides are numbered from the 5' end of the v strand, and ORFs are presented as open bars. The locations of the primers (pNSs and pN) used for primer extension analysis are indicated. (b) Primer extension analyses on unfractionated nucleocapsid RNA (lane 1), NSs mRNA-enriched RNA fraction (lane 2), S vcRNA-enriched RNA fraction (lane 3) and the N mRNA-enriched RNA fraction (lane 4) using primers pNSs or pN are as described in the text. A sequence ladder is included, as a standard to determine sizes of the various primer extension products.

initiate transcription. An additional single extension product of 63 nt was found in minor amounts when pNSs was extended on nucleocapsid RNA (Fig. 2b, lane 1). The origin of this band has not been investigated further, but could be caused by spurious amounts of viral mRNA copurified (or even co-encapsidated) with the genomic RNA.

The presence of non-viral sequences in mRNAs has previously been found for other members of the Bunyaviridae (Bishop et al., 1983; Patterson & Kolakofsky, 1984; Eshita et al., 1985; Ihara et al., 1985; Collett, 1986; Gerbaud et al., 1987; Bouloy et al., 1990; Simons & Pettersson, 1991), and also for members of the Orthomyxoviridae (Caton & Robertson, 1980; Dhar et al., 1980) and Arenaviridae (Garcin & Kolakofsky, 1990; Raju et al., 1990). In most cases, the added sequences range in size from 15 to 18 nt; for the arenavirus Tacaribe however, only 1 to 4 extra nucleotides have been reported (Garcin & Kolakofsky, 1990; Raju et al., 1990). Here it is demonstrated that TSWV, a bunyavirus with a very distinct host range, also uses cap-snatching to initiate transcription. This indicates that all members of the Bunyaviridae, irrespective of whether they infect animal or plant cells, use the same mechanism to initiate transcription of their genome. The process of cap-snatching, in particular the involvement of a host- or virus-encoded endonuclease activity, needs to be investigated further. For influenza virus (Plotch et al., 1981), La Crosse virus (Patterson et al., 1984) and Germiston virus (Vialat & Bouloy, 1992), the responsible endonuclease activity has already been demonstrated to be virus-encoded. During influenza virus replication three different viral proteins are involved in genome transcription and replication, i.e. proteins PA, PB1 and PB2, of which PB1 represents the core RNA polymerase and PB2 probably the endonuclease activity required for cap-snatching (Plotch et al., 1981; Ulmanen et al., 1981 a). In view of the size of the L protein of TSWV (331.5K),
which exceeds the sum of the sizes of the three replication proteins of influenza virus, it is tempting to assume that this L RNA-encoded product encompasses, replication proteins of influenza virus, it is tempting to which exceeds the sum of the sizes of the three

The authors wish to thank Jeroen Corver and Arie Smolders for their technical assistance. This work was supported in part by the Netherlands Foundation for Chemical Research (S.O.N.) with financial aid from the Netherlands Organization for Scientific Research (N.W.O.).

References


