

Splicing of *Cauliflower mosaic virus* 35S RNA serves to downregulate a toxic gene product

Rémy Froissart,^{1†} Maryline Uzest,¹ Virginia Ruiz-Ferrer,¹ Martin Drucker,¹ Eugénie Hébrard,¹ Thomas Hohn² and Stéphane Blanc¹

Correspondence
Stéphane Blanc
blanc@ensam.inra.fr

¹UMR BGPI, CIRAD-INRA-ENSAM, TA 41/K, Campus International de Baillarguet, 34398 Montpellier Cedex 5, France

²Friedrich Miescher Institut, PO Box 2543, CH-4002 Basel, Switzerland

Alternative splicing usually leads to an increase in the number of gene products that can be derived from a single transcript. Here, a different and novel use of alternative splicing – as a means to control the amount of a potentially toxic gene product in the plant pararetrovirus *Cauliflower mosaic virus* (CaMV) – is reported. About 70% of the CaMV 35S RNA, which serves as a substrate for both reverse transcription and polycistronic mRNA, is spliced into four additional RNA species. Splicing occurs between four donor sites – one in the 5' untranslated region and three within open reading frame (ORF) I – and one unique acceptor site at position 1508 in ORF II. A previous study revealed that the acceptor site is vital for CaMV infectivity and expression of ORFs III and IV from one of the spliced RNA species suggested that splicing may facilitate expression of downstream CaMV ORFs. However, it is shown here that deleting the splice acceptor site and replacing ORF II with a cargo ORF that lacks splice acceptor sites does not interfere with virus proliferation. Furthermore, it is demonstrated that whenever P2 cannot accumulate in infected tissues, the splice acceptor site at position 1508 is no longer vital and has little effect on virus replication. This suggests that the vital role of splicing in CaMV is regulation of P2 expression and that P2 exhibits biological properties that, whilst indispensable for virus–vector interactions, can block *in planta* virus infection if this regulation is abolished.

Received 11 February 2004

Accepted 6 May 2004

INTRODUCTION

Alternative RNA splicing is well-documented in animal DNA viruses and retroviruses as a means of increasing the number of mRNA species that are produced from a primary transcript (Butsch & Boris-Lawrie, 2002; Pongoski *et al.*, 2002; Akusjarvi & Stevenin, 2003). Also, in hepadnaviruses, including *Hepatitis B virus* (HBV; Huang *et al.*, 2000) and *Duck hepatitis B virus* (DHBV; Loeb *et al.*, 2002), several virus proteins have been demonstrated to originate from spliced mRNA. In contrast, RNA splicing events are much less well-documented for plant viruses, regarding both their functional importance and their role in regulating genome expression. In several species of the family *Geminiviridae*, splicing of two major RNA transcripts (C and V transcripts) has been reported and characterized (Wright *et al.*, 1997). Notably, processing of the C transcript results in a fusion between two distinct open reading frames (ORFs), allowing early expression of

the *Rep* protein that is required for virus replication. In the family *Caulimoviridae*, splicing of the polycistronic 35S RNA has been described in the genera 'Rice tungro bacilliform-like viruses' and *Caulimovirus*, where it is believed to enhance expression of downstream ORFs from newly formed mRNAs (Kiss-László & Hohn, 1996; Hull, 2001). In *Rice tungro bacilliform virus* (RTBV), ORF (IV) of the full-length RNA is expressed after removal of a 6.3 kbp intron (Fütterer *et al.*, 1994). However, the biological relevance of the two cases of splicing that have been reported in the genus *Caulimovirus* remains unclear; its role has never been tested for *Figwort mosaic virus* (Scholthof *et al.*, 1991) and, as detailed below, is still undefined for *Cauliflower mosaic virus* (CaMV; Kiss-László *et al.*, 1995).

CaMV is a pararetrovirus, the double-stranded DNA genome of which is replicated by reverse transcription of a pre-genomic 35S RNA. This 35S RNA is also used as a polycistronic mRNA for all virus genes [reviewed by Ryabova *et al.* (2002)]. In earlier work, several additional RNA species that were detected in CaMV-infected plants were suspected to originate from splicing of the 35S RNA (Hirochika *et al.*, 1985; Vaden & Melcher, 1990). That splicing of the 35S RNA is an absolute requirement for

[†]Present address: Equipe 'Ecologie et Evolution des Micro-Organismes', INSERM E 0339, Hôpital X. Bichat, 16 rue Henri Huchard, 75018 Paris, France.

virus infectivity was later demonstrated by Kiss-László *et al.* (1995). These authors characterized four additional RNA species (C1–C4) that are produced by splicing between four donor sites and one unique acceptor site in the 35S RNA. One of the donor sites is located in the long 5' untranslated region (UTR) of the 35S RNA and the three others are located within ORF I; the unique acceptor site is located in the first half of ORF II at nt 1508 [numbering according to Franck *et al.* (1980)]. Mutations of the consensus sequence of the splice acceptor site, designated ma1 and ma3 (changing AGG to CCG and AGA, respectively), totally abolished splicing and were lethal for CaMV. Rare second-site revertants eventually appeared upon deletion of the central region of ORF II, which allowed the use of a cryptic splice acceptor site that is located further downstream in the ORF II sequence. Several hypotheses were proposed to attribute a biological function to these splicing events, the most commonly accepted of which postulates that splicing might facilitate expression of the downstream genes (Kiss-László & Hohn, 1996; Hull, 2001). Indeed, in RNA species C4, in which the end of the 5' UTR is joined to the acceptor site at position 1508, ORF III is moved to the position of the first translated ORF and efficient translation of ORFs III and IV from C4 RNA was verified in transient expression experiments in protoplasts. Nevertheless, Kiss-László *et al.* (1995) mentioned an alternative (but not mutually exclusive) hypothesis, i.e. that the splicing events observed might also serve to modulate expression of genes I and/or II, which can be translated only from the remaining 30% of unspliced RNA. In this work, we reinvestigate the question of exactly what feature of the splicing events characterized by Kiss-László *et al.* (1995) is vital for CaMV.

ORF II is involved in aphid transmission [for a review, see Blanc *et al.* (2001)] and is otherwise dispensable for host plant infection, as demonstrated by the CaMV isolate CM4-184, where a 421 bp deletion removes almost the entire coding sequence of P2, the ORF II product (Howarth *et al.*, 1981). Surprisingly, and in apparent contradiction, several amino acid changes in P2 are lethal for the virus (S. Blanc, unpublished results), suggesting that, under some circumstances, P2 may be 'toxic' for virus replication and hence may require very tight regulation (Pirone & Blanc, 1996; Blanc *et al.*, 2001). This led to the hypothesis that splicing of 35S RNA is required to downregulate P2 accumulation, and the prediction that the splicing event is vital only where functional P2 accumulates. To test this hypothesis, we produced a number of different CaMV mutants that were debilitated in their capacity to accumulate P2 in infected plants. As predicted, impairing 35S RNA splicing in these mutant viruses was no longer lethal. Moreover, despite disruption of the splice acceptor site at position 1508, symptom appearance and virus accumulation in most of these mutants was not modified, demonstrating that, in the absence of P2, splicing *per se* has no important biological function in CaMV. Hence, we conclude that the vital role of the splicing events analysed is downregulation of P2 expression.

METHODS

Plasmids and mutant virus derivatives. All mutations (Fig. 1) were engineered in the plasmid pCa37 (Franck *et al.*, 1980) that contained the full-length genome of CaMV Cabb-S strain, cloned at the unique *Sall* restriction site in pBR322.

Two partially overlapping primers, Del-S-Forward and Del-S-Reverse, were designed, with the overlap corresponding to the region between nt 1375–1389 and 1811–1825 of Cabb-S. Del-S-Forward contained 16 additional nucleotides downstream of position 1825, corresponding to the end of ORF II and the beginning of ORF III, whereas Del-S-Reverse contained 14 additional nucleotides upstream of position 1375, in the 5' region of ORF II. Del-S-Forward and Del-S-Reverse were used, together with primers RKpn (22 complementary nucleotides overlapping the unique *KpnI* site at position 2039) and FBsiW (22 forward nucleotides overlapping the unique *BsiWI* site at position 208), respectively, to PCR-amplify two DNA fragments, both of which contained a deletion within ORF II (spanning nt 1389–1811). After this first round of PCR amplification, the two DNA fragments obtained were heat-denatured in 0.1 × TE buffer (1 mM Tris/HCl, pH 8; 0.1 mM EDTA) and mixed in a 1:1 ratio before slow cooling. An aliquot (10 ng) of this annealed DNA mixture was used as template for a second round of PCR amplification with primers RKpn and FBsiW. The resulting chimeric PCR fragment was digested by *KpnI* and *BsiWI* and was used to replace the corresponding fragment in pCa37 to generate the clone CaMV Del-S.

The mutant CaMV ΔII-S was obtained similarly, with the partially overlapping primers ΔIIS-Forward (5'-AATAAACTAGTATGGCTA-ATCTTAATC-3') and ΔIIS-Reverse (5'-GCCATACTAGTTTATTCT-CCACAGATTTTC-3') instead of Del-S-Forward and Del-S-Reverse. This construct resulted in a mutant that harboured a total deletion of ORF II, with a unique *SpeI* restriction site separating ORFs I and III.

For mutants CaMV Top-S and Tart-S, similar chimeric PCR fragments were again amplified from the template pCa37. For the first round of PCR, completely overlapping primers were used for both constructs. These primers spanned nt 1343–1387 and 1340–1375 for CaMV Top-S and Tart-S, respectively. For CaMV Top-S, the complementary overlapping primers contained a mutation that altered codon 6 of ORF II from CAA to TAA, thus introducing an early termination codon. For CaMV Tart-S, the complementary overlapping primers contained a mutation that altered the ATG of ORF II to AAG, thus abolishing initiation of translation. The first and second rounds of PCR amplification were performed as above with primers RKpn and FBsiW and the amplified mutated fragments were introduced into pCa37 at the *BsiWI* and *KpnI* restriction sites. To introduce the splicing acceptor mutations ma1 or ma3 into Tart-S, the same procedure was carried out on the plasmid templates pCa37ma1 and pCa37ma3 (Kiss-László *et al.*, 1995).

CaMV-ΔIIS θ was produced by introducing the *Escherichia coli* thioredoxin gene, PCR-amplified from the plasmid pTrxFus (Invitrogen) and flanked with *SpeI* restriction sites, into the corresponding cloning site in CaMV-ΔIIS.

CaMV mut-S was constructed by amplifying the CaMV Cabb-S region between the unique *XhoI* and *KpnI* sites with primers F-mut-S and Rkpn. The 73 nt F-mut-S primer overlapped the *XhoI* site and the downstream codon that corresponds to aa 116 of P2, which was modified to encode an aspartic acid, instead of the original isoleucine. The PCR fragment obtained was introduced into the *XhoI* and *KpnI* restriction sites of pCa37. To combine mutations ma3 and mut-S, we simply replaced the *XhoI/KpnI* fragment of plasmid pCa37ma3 (Kiss-László *et al.*, 1995) with that of mut-S. All modified plasmids created were verified by sequencing.

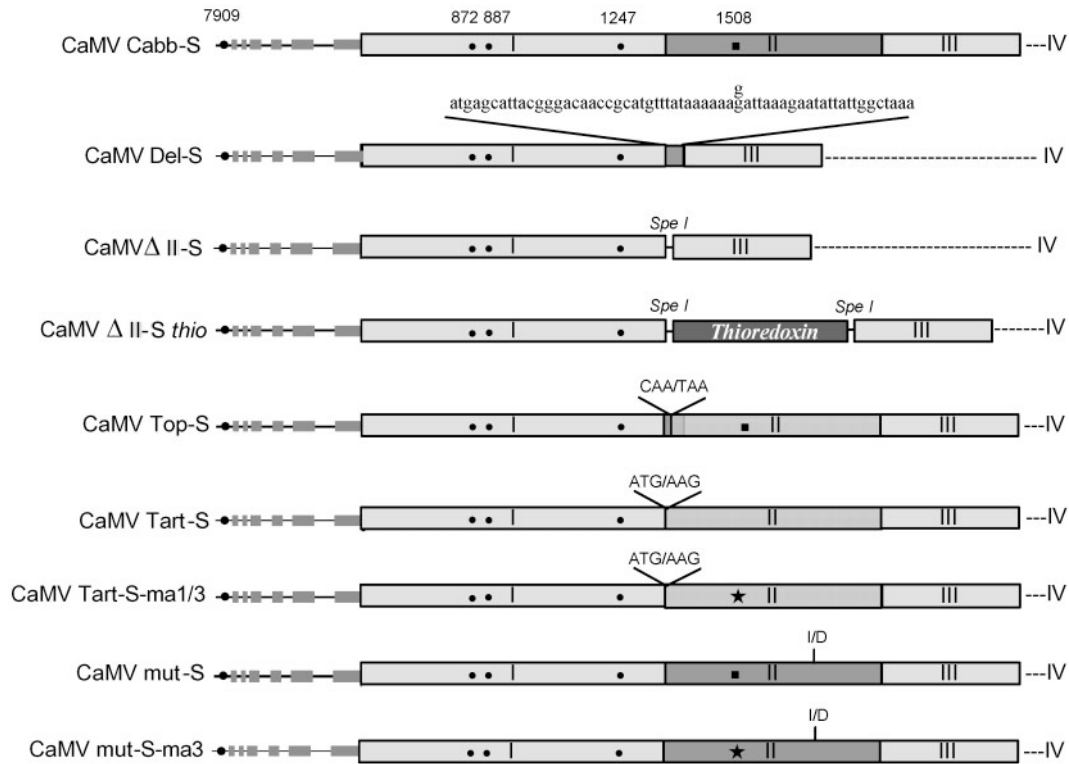


Fig. 1. CaMV genome mutants impaired in the expression of ORF II. First line represents CaMV Cabb-S (wild-type genome) 35S RNA from the 5' extremity to the end of ORF III. Small ORFs in the 5' UTR are shown as grey boxes. The splice donor and acceptor sites described by Kiss-László *et al.* (1995) are indicated by black dots and squares, respectively, and their position in the nucleotide sequence is indicated above [numbering according to Franck *et al.* (1980)]. All mutants used in this study are represented below and their respective names are indicated on the left. The ORF II region remaining in CaMV Del-S is detailed and is similar to that of the naturally occurring mutant CM4-184. Note: 'g' (indicated above the sequence) corresponds to a nucleotide difference that creates a cryptic splice acceptor site in isolate CM4-184 (Kiss-László *et al.*, 1995) that is not present in Del-S. Insertions of *SpeI* cloning sites in CaMV Δ II-S and of a bacterial thioredoxin gene in CaMV Δ II-S*thio* are described in Methods. Mutations ma1 and/or ma3, described previously (Kiss-László *et al.*, 1995), are shown by black stars and the isoleucine (I) to aspartic acid (D) substitution at aa 116 of P2 is indicated in CaMV mut-S and its derivative mutant mut-S-ma3. Dark boxes represent ORF II when translation is abolished; pale boxes represent functional ORFs that can be translated into proteins.

Infectivity of CaMV mutants and stability of mutations.

Infectivity testing was performed by mechanical inoculation of plasmid DNA into turnip hosts (*Brassica rapa* cv. Just Right) as described by Melcher *et al.* (1986) and symptom appearance was assessed visually daily for over 3 months. When symptoms were fully developed in a host plant, the virus was transferred mechanically to a new series of host plants by sap inoculation, with serial passages being repeated three times. After each passage, viral DNA was purified from systemically infected hosts as described by Gardner & Shepherd (1980). The region that contained the mutation of interest was PCR-amplified and sequenced. Whenever the introduced mutation was maintained and no second-site revertant was detected for at least one passage, the corresponding mutant was declared to be infectious.

Protein analysis. To test for accumulation of virus and P2 (ORF II product) in infected plant cells, we collected leaves with fully developed symptoms and extracted the protein contents. Leaf material (200 mg) was ground in 0.5 ml extraction buffer (200 mM Tris/HCl, pH 7.6; 100 mM EGTA; 50 mM MgCl₂; 0.1% SDS) and centrifuged at 10 000 g for 10 min. The supernatant was then

submitted to 12% SDS-PAGE according to Laemmli (1970) and separated proteins were transferred onto nitrocellulose membranes. Detection of P2 and coat protein was performed with specific antisera that were described by Blanc *et al.* (1993) and Karsies *et al.* (2002), respectively.

RESULTS AND DISCUSSION

The mutants used in this study (Fig. 1) were obtained by mutagenesis of the original pCa37 plasmid (Franck *et al.*, 1980) that contained the full-length genome of CaMV isolate Cabb-S. All tests for infectivity and symptom appearance are summarized in Table 1.

CaMV mutants that harbour a total deletion of ORF II are viable

In isolate CM4-184 (Howarth *et al.*, 1981), the sequence deleted in ORF II includes the splice acceptor site at position

Table 1. Infectivity and timing of symptom appearance of various CaMV Cabb-S derivative mutants

Mutant*	Infection†	P2 accumulation‡	Splice acceptor site§	Symptom delay
CaMV Cabb-S	+	+	+	No delay
CaMV Cabb-S-ma1	–	NA	–	NA
CaMV Cabb-S-ma3	–	NA	–	NA
CaMV Del-S	+	–	–	2–4 weeks
CaMV ΔII-S	+	–	–	2–4 weeks
CaMV ΔII- <i>Sthio</i>	+	–	–	2–4 weeks
CaMV Top-S¶	(–)¶	NA	+	6–8 weeks¶
CaMV Tart-S	+	–	+	2–4 weeks
CaMV Tart-S-ma1	+	–	–	2–4 weeks
CaMV Tart-S-ma3	+	–	–	2–4 weeks
CaMV mut-S	+	–	+	No delay
CaMV mut-S-ma3	+	–	–	1–3 days

*Mutant name and construction are described in Methods, except for Cabb-S-ma1 and ma3, which were described by Kiss-László *et al.* (1995).

†Experiments were repeated four times, with 10 plants tested for each repetition. Mutants were considered to be infectious after determining that test plants became infected with no second-site reversion or genome rearrangements of the mutant CaMV. No. plants infected over the total no. plants tested varied between repetitions but, in all cases of infectious mutants, the infection rate exceeded 30%.

‡In infectious mutants, the presence/absence of P2 was verified experimentally (Fig. 2). This column is not applicable (NA) for non-infectious mutants.

§For the splice acceptor site at nt 1508, a plus sign corresponds to the wild-type sequence and a minus sign to either the absence of the corresponding region or disruption by mutation ma1 or ma3 (Fig. 1).

||Cabb-S is the reference, compared to the other mutants. Experiments were repeated several times and delays in symptom appearance varied, depending on the season and plant conditions, in a range indicated for each mutant.

¶This mutant was infective only after second-site reversion (see text); hence, it is not considered to be infectious here.

1508. The viability of this isolate has been attributed to the use of a cryptic acceptor site, located downstream of the deleted region in the remaining ORF II sequence (Kiss-László *et al.*, 1995). We deleted exactly the same fragment in the background of the wild-type isolate Cabb-S, to generate the mutant Del-S. Due to nucleotide sequence differences between Cabb-S and CM4-184, the deletion in Del-S does not engender a putative cryptic splice acceptor site such as that in CM4-184 (Fig. 1). Nevertheless, the Del-S mutant was infectious (Table 1). To exclude the possibility that an unpredicted sequence was used as a new cryptic splice acceptor site in the remaining ORF II sequence, we engineered the mutant ΔII-S so that the entire ORF II sequence was replaced by a unique *SpeI* restriction site. The CaMV ΔII-S mutant was also infectious, demonstrating that, in the total absence of ORF II and, consequently, of ORF II-based splice acceptor sites, CaMV can still systemically infect its host.

Complete deletion of ORF II resulted in a 35S RNA with ORF III being preceded solely by two ORFs (VII and I), a situation comparable to that described in spliced RNA species C1, C2 and C3 (Kiss-László *et al.*, 1995), where ORF III is preceded by ORF VII and an ORF I–II fusion product. Thus, infectivity of ΔII-S might be enabled simply by the

reduction in the number of ORFs that precede ORF III. To rule out this possibility, we increased the number of ORFs again by including a bacterial thioredoxin gene (330 bp) at the original position of ORF II. This modified 35S RNA is very unlikely to be spliced in a manner comparable to that of wild-type Cabb-S RNA, due to the absence of ORF II and, thus, of the splice acceptor site at position 1508. Nevertheless, CaMV ΔII-*Sthio* proved to be infectious. This result indicates that the presence of three fully constituted ORFs between the 5' UTR and ORF III in a 35S RNA that is impaired in its splicing capacity is not in itself detrimental to CaMV replication. It further suggests that it is specifically the ORF II sequence itself that is lethal in splicing-impaired 35S RNA. Expression of a foreign gene from viable CaMV mutants (DHFR or α interferon at the position of gene II) has been described (Brisson *et al.*, 1984; De Zoeten *et al.*, 1989) but, in those cases, and in contrast to ΔII-*Sthio*, the ORF II sequences had not been totally removed.

All three mutants – CaMV Del-S, ΔII-S and ΔII-*Sthio* – were delayed in symptom appearance (Table 1). We verified that the modified sequences were stable and that no additional gene rearrangements or second-site reversions were induced. Other than the absence of splicing, a reason for the delay might be that ORF II includes one of the two

G-rich priming sites that are used for + strand DNA synthesis (Bonneville & Hohn, 1993). The fact that replication is delayed but not abolished is probably due to the presence of an additional priming site for + strand DNA synthesis located within ORF V, which is functionally redundant with that in ORF II.

Preventing translation of P2 renders the splice acceptor site in ORF II non-essential

To test the hypothesis that splicing in wild-type CaMV is required to downregulate P2, we constructed CaMV derivatives in which either the start codon of ORF II was mutated from ATG to AAG (designated CaMV Tart-S) or a premature stop codon was introduced at aa 6 of P2 (designated CaMV Top-S). The aim of this experiment was to disable accumulation of P2 with minimal changes to the CaMV genome sequence. Tart-S was fully infectious, although symptom appearance was slightly delayed (Table 1). In contrast, Top-S was not infectious without acquisition of second-site reversions. In such cases, symptoms appeared after a considerable delay (6–8 weeks) and the stop codon that was introduced had reverted to a new amino acid codon (see one example in Fig. 2). Sequencing and protein analysis controls carried out in parallel (Fig. 2) demonstrated that no reversion had occurred in Tart-S at this time. As wild-type CaMV Tart-S and Top-S theoretically produce the same spliced RNA species, no differences in symptom development would be expected in the mutant,

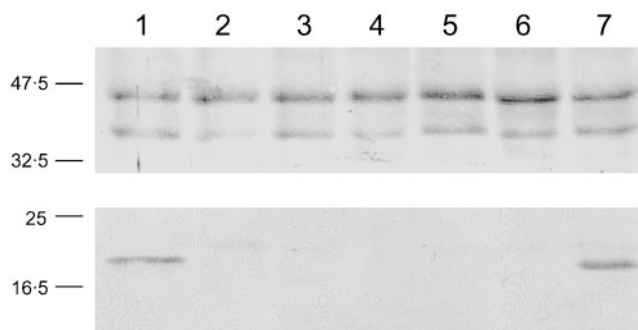


Fig. 2. Accumulation of coat protein and P2 on infection of turnip with CaMV mutants. Total protein extracts of CaMV-infected turnips (60 days post-infection) were prepared as described in Methods. Samples were submitted to 12% SDS-PAGE and transferred to nitrocellulose membranes before specific immunostaining of coat protein (top), as a control of virus accumulation, and P2 protein (below). Virus accumulation in all mutants (lanes 2–7) is comparable to that in wild-type CaMV Cabb-S (lane 1). P2 protein accumulates in the wild-type CaMV Cabb-S (lane 1) and in revertants obtained from the mutant Top-S (one example is shown in lane 7), but is never detected in mutants Tart-S, Tart-S-ma1, Tart-S-ma3, mut-S or mut-S-ma3 (lanes 2–6, respectively). Molecular mass marker positions are indicated in kDa on the left. The very faint band observed slightly above the position of P2 in lanes 2, 3, 6 and 7 is a sporadically appearing, non-specific staining.

compared to wild-type CaMV. The observed delay in symptom appearance in the case of Tart-S is an indication that the unspliced 35S RNA is used, at least partially, for translation of genes downstream of ORF II, with the long non-coding sequence (481 nt) that separates ORFs I and III, and/or the short ORFs that are present in this region in other reading frames, presumably interfering with the polycistronic mode of CaMV translation (Ryabova *et al.*, 2002). In agreement with this latter explanation, we eventually observed a second-site revertant after several successive passages of Tart-S in host plants, with a new initiation codon (ATG) being created at aa 4 of the native ORF II (data not shown). This restores ORF II, thus masking the small ORFs that are present in other reading frames.

The fact that Top-S was not viable is similarly consistent with the hypothesis that at least some translation of ORFs that lie downstream of ORF II occurs from the unspliced RNA. In this sense, the lethality of Top-S might be explained by the presence of several short ORFs, including ORF II truncated after six codons and short ORFs in other reading frames, a situation that is unfavourable for reinitiation of translation of polycistronic RNAs (Ryabova *et al.*, 2002). Again, as the Top-S mutation is located within the introns, it should not have had any effect on CaMV replication if translation of downstream genes was occurring solely from spliced RNA species.

Mutations ma1 or ma3, which suppressed splicing totally by disrupting the splice acceptor site at nucleotide position 1508 and were lethal for wild-type CaMV Cabb-S (Kiss-László *et al.*, 1995), were introduced in the Tart-S mutant background to generate CaMV Tart-S-ma1 and Tart-S-ma3. Interestingly, both proved to be infectious, with no detectable delay in symptom appearance compared to Tart-S (Table 1). These results confirm that, when P2 cannot be translated and thus does not accumulate in infected tissues (Fig. 2), disruption of the splice acceptor site at position 1508 no longer interferes with the CaMV infection cycle and has little or no effect on the timing of infection. We checked the virus progeny carefully by sequencing the whole ORF II region and observed neither direct reversion of ma1 or ma3, nor second-site reversions of any sort.

CaMV mutants expressing unstable P2 are not affected by their splicing capability

Amino acid changes in the C-terminal α -helix 1 of P2 resulted in the complete absence of P2 accumulation in CaMV-infected plants by serendipity (Hébrard *et al.*, 2001), which was explained by induced conformational changes leading to protein instability. Accordingly, we assumed that the mutation I116D in P2 (CaMV mut-S) would also prevent P2 accumulation, as position 116 falls into the hydrophobic interface that is responsible for the formation of coiled-coil structures that lead to P2 self-interaction (Hébrard *et al.*, 2001). This additional construct is of particular interest. It was created in order to yield a mutant as comparable as possible to wild-type CaMV, where all

genuine ORFs are present and can be translated normally, but where the ORF II gene product P2 would nevertheless be absent. As mut-S has no modification that is predicted to affect genome translation or replication, in contrast to all other mutants described above, we expected this mutant to be infectious without any noticeable delay in symptom appearance when compared with wild-type Cabb-S.

We first verified that mut-S was viable and displayed no delays in symptom appearance (Table 1), then verified the absence of P2 in infected tissues (Fig. 2). Mutation ma3 was then introduced in mut-S to yield CaMV mut-S-ma3, which was assessed for infectivity. Table 1 shows that it is indeed infectious, with very short delays (if any) in symptom appearance, further confirming that the splice acceptor site at position 1508 of CaMV Cabb-S is not essential in the absence of P2 accumulation in infected plants. Moreover, the very short delays that were recorded in disease development on infection of plants with CaMV mut-S-ma3, together with normal virus accumulation in infected tissues (Fig. 2), suggest that splicing of the 35S RNA of CaMV (Kiss-László *et al.*, 1995) has no major biological function other than regulating expression of P2. However, as very short delays did sometimes occur, we could not totally exclude a secondary, minor, dispensable effect that was caused either by enhancement of translation of downstream ORFs or modulation of P1 expression, or both.

Three of the four spliced RNA species C1–C4 contain in-frame fusions between ORFs I and II, the putative function of which was questioned by Kiss-László *et al.* (1995). As demonstrated here, the presence or absence of such P1–P2 fusions (mutants with an unmodified acceptor site at position 1508, and any mutant ma1 or ma3 where the acceptor site is disrupted, respectively) is not critical for CaMV viability. We therefore believe that these putative fusion proteins have no function, at least in the absence of P2, and are rather unstable, underlined by the fact that they were barely detectable in infected tissues (Kiss-László *et al.*, 1995). We propose that the P1–P2 fusions in RNA species C2, C3 and C4 are in-frame in order to prevent the deleterious appearance of many small ORFs that would result from out-of-frame fusions. This idea is supported by the lack of infectivity of our mutant Top-S, demonstrating that a series of small ORFs located between coding ORFs is deleterious for the CaMV life cycle.

Spliced RNA species are extremely difficult to detect in CaMV-infected plants, due to the presence of many partially digested virus RNA molecules that are released during reverse transcription (Kiss-László *et al.*, 1995). For this reason, we did not analyse putative alternative spliced RNA species in our various mutants, but designed the latter to allow solid conclusions from simple testing of infectivity, timing of symptom appearance and virus and P2 accumulation. The logic underlying the design of the mutants used is discussed above, together with the results in each case, but we draw attention particularly to the comparison of mutants mut-S and mut-S-ma3. The change in mut-S is

restricted to a single amino acid change in the last third of the protein. Whilst it results in the total disappearance of P2, it affects neither virus fitness, as judged by timing of infection (Table 1), nor virus accumulation (Fig. 2). The ma3 disruption of the splice acceptor site at position 1508 has virtually no effect on mut-S (Table 1; Fig. 2), whilst it kills wild-type Cabb-S. The only difference between the two is the presence/absence of P2; thus, the most logical conclusion is that P2 aborts CaMV infection in the absence of splicing. Indeed, other hypotheses that invoke putative alternative splicing events that could compensate for the ma3 mutation in mut-S are very unlikely, as these alternative splicing events obviously do not occur in the wild-type Cabb-S.

Pirone & Blanc (1996) first suggested the hypothetical requirement for tight regulation of P2 expression and accumulation. They speculated that the properties of P2 that are required for aphid transmission might, in turn, be 'toxic' for the virus replication cycle or cell-to-cell movement. Thus, precise mechanisms of regulation may exist to harmonize the different steps of the CaMV life cycle: plant invasion versus aphid transmission. Here, we provide the first indirect experimental evidence for such toxic properties by showing that, whilst P2 can be removed totally without abolishing host plant infection, its accumulation can be lethal in the absence of splicing of the 35S RNA. At this point, we can speculate that some properties of P2 may potentially interfere with CaMV infection. The affinity of P2 for the microtubule network of the host cell (Blanc *et al.*, 1996) or its association with the virion–P3 complex for aphid transmission (Schmidt *et al.*, 1994; Leh *et al.*, 1999, 2001) appear to be good candidates. Both properties might interfere with replication or cell-to-cell movement and unregulated binding to the P3–virion complex could, for instance, mask the nuclear localization signal motif of the coat protein and prevent nuclear targeting of the virions. It was shown that P3 is also required for cell-to-cell movement (Kobayashi *et al.*, 2002) and that P3 virions interact with the movement protein (L. Stavalone, personal communication). P2 could inhibit such an interaction and, consequently, cell-to-cell movement. Surprisingly, it was demonstrated recently that, in infected plant cells, P2 is not massively bound either to virion–P3 complexes or to the microtubule network. Instead, it is stored entirely within specific, electron-lucent inclusion bodies (ELIBs) as P2–P3 aggregates (Drucker *et al.*, 2002). The mechanisms that sequester P2 in these ELIBs remain to be discovered, but P3 may play a pivotal role. In this sense, when P3 is targeted to the ELIB together with P2, one could postulate that the P3/P2 ratio is important and might be regulated by the splicing events that are discussed here. Whilst P3 is probably expressed from both spliced and unspliced viral RNAs, P2 can be derived only from unspliced 35S species.

More generally, alternative RNA splicing is commonly interpreted as producing new RNA species that allow (or enhance) the expression of a protein with a specific

biological function (Huang *et al.*, 2000; Kondrashov & Koonin, 2003; Kriventseva *et al.*, 2003; Lu & Cullen, 2003). To the best of our knowledge, the work presented here is the first example of a splicing event of which the primary purpose appears to be prevention of excess accumulation of a specific gene product that would otherwise be produced from the unspliced RNA.

ACKNOWLEDGEMENTS

We are very grateful to Helen Rothnie and Lyubov Ryabova for critical reading of the manuscript. Turnip cv. Just Right seeds were kindly provided by Takii Ltd. We thank Pascal Espérandieu for technical assistance in virus DNA inoculation.

REFERENCES

- Akusjarvi, G. & Stevenin, J. (2003). Remodelling of the host cell RNA splicing machinery during an adenovirus infection. *Curr Top Microbiol Immunol* **272**, 253–286.
- Blanc, S., Cerutti, M., Usmany, M., Vlak, J. M. & Hull, R. (1993). Biological activity of cauliflower mosaic virus aphid transmission factor expressed in a heterologous system. *Virology* **192**, 643–650.
- Blanc, S., Schmidt, I., Vantard, M., Scholthof, H. B., Khul, G., Espérandieu, P., Cerutti, M. & Louis, C. (1996). The aphid transmission factor of cauliflower mosaic virus forms a stable complex with microtubules in both insect and plant cells. *Proc Natl Acad Sci U S A* **93**, 15158–15163.
- Blanc, S., Hébrard, E., Drucker, M. & Froissart, R. (2001). Molecular basis of vector transmission: Caulimoviruses. In *Virus–Insect–Plant Interactions*, pp. 143–166. Edited by K. Harris, O. P. Smith & J. E. Duffus. San Diego, CA: Academic Press.
- Bonneville, J.-M. & Hohn, T. (1993). A reverse transcriptase for cauliflower mosaic virus: state of the art, 1991. In *Reverse Transcriptase*, pp. 357–390. Edited by A. M. Skalka & S. P. Goff. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.
- Brisson, N., Paszkowski, J., Penswick, J. R., Gronenborn, B., Potrykus, I. & Hohn, T. (1984). Expression of a bacterial gene in plants by using a viral vector. *Nature* **310**, 511–514.
- Butsch, M. & Boris-Lawrie, K. (2002). Destiny of unspliced retroviral RNA: ribosome and/or virion? *J Virol* **76**, 3089–3094.
- De Zoeten, G. A., Penswick, J. R., Horisberger, M. A., Ahl, P., Schultze, M. & Hohn, T. (1989). The expression, localization, and effect of human interferon in plants. *Virology* **172**, 213–222.
- Drucker, M., Froissart, R., Hébrard, E. & 8 other authors (2002). Intracellular distribution of viral gene products regulates a complex mechanism of cauliflower mosaic virus acquisition by its aphid vector. *Proc Natl Acad Sci U S A* **99**, 2422–2427.
- Franck, A., Guilley, H., Jonard, G., Richards, K. & Hirth, L. (1980). Nucleotide sequence of cauliflower mosaic virus DNA. *Cell* **21**, 285–294.
- Fütterer, J., Potrykus, I., Valles Brau, M. P., Dasgupta, I., Hull, R. & Hohn, T. (1994). Splicing in a plant pararetrovirus. *Virology* **198**, 663–670.
- Gardner, R. C. & Shepherd, R. J. (1980). A procedure for rapid isolation and analysis of cauliflower mosaic virus DNA. *Virology* **106**, 159–161.
- Hébrard, E., Drucker, M., Leclerc, D. & 9 other authors (2001). Biochemical characterization of the helper component of *Cauliflower mosaic virus*. *J Virol* **75**, 8538–8546.
- Hirochika, H., Takatsuji, H., Ubasawa, A. & Ikeda, J. (1985). Site-specific deletion in cauliflower mosaic virus DNA: possible involvement of RNA splicing and reverse transcription. *EMBO J* **4**, 1673–1680.
- Howarth, A. J., Gardner, R. C., Messing, J. & Shepherd, R. J. (1981). Nucleotide sequence of naturally occurring deletion mutants of cauliflower mosaic virus. *Virology* **112**, 678–685.
- Huang, H.-L., Jeng, K.-S., Hu, C.-P., Tsai, C.-H., Lo, S. J. & Chang, C. (2000). Identification and characterization of a structural protein of hepatitis B virus: a polymerase and surface fusion protein encoded by a spliced RNA. *Virology* **275**, 398–410.
- Hull, R. (2001). *Matthews' Plant Virology*, 4th edn. San Diego, CA: Academic Press.
- Karsies, A., Merkle, T., Szurek, B., Bonas, U., Hohn, T. & Leclerc, D. (2002). Regulated nuclear targeting of cauliflower mosaic virus. *J Gen Virol* **83**, 1783–1790.
- Kiss-László, Z. & Hohn, T. (1996). Pararetro- and retrovirus RNA: splicing and the control of nuclear export. *Trends Microbiol* **4**, 480–485.
- Kiss-László, Z., Blanc, S. & Hohn, T. (1995). Splicing of cauliflower mosaic virus 35S RNA is essential for viral infectivity. *EMBO J* **14**, 3552–3562.
- Kobayashi, K., Tsuge, S., Stavalone, L. & Hohn, T. (2002). The cauliflower mosaic virus virion-associated protein is dispensable for viral replication in single cells. *J Virol* **76**, 9457–9464.
- Kondrashov, F. A. & Koonin, E. V. (2003). Evolution of alternative splicing: deletions, insertions and origin of functional parts of proteins from intron sequences. *Trends Genet* **19**, 115–119.
- Kriventseva, E. V., Koch, I., Apweiler, R., Vingron, M., Bork, P., Gelfand, M. S. & Sunyaev, S. (2003). Increase of functional diversity by alternative splicing. *Trends Genet* **19**, 124–128.
- Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685.
- Leh, V., Jacquot, E., Geldreich, A., Hermann, T., Leclerc, D., Cerutti, M., Yot, P., Keller, M. & Blanc, S. (1999). Aphid transmission of cauliflower mosaic virus requires the viral PIII protein. *EMBO J* **18**, 7077–7085.
- Leh, V., Jacquot, E., Geldreich, A., Haas, M., Blanc, S., Keller, M. & Yot, P. (2001). Interaction between the open reading frame III product and the coat protein is required for transmission of cauliflower mosaic virus by aphids. *J Virol* **75**, 100–106.
- Loeb, D. D., Mack, A. A. & Tian, R. (2002). A secondary structure that contains the 5' and 3' splice sites suppresses splicing of duck hepatitis B virus pregenomic RNA. *J Virol* **76**, 10195–10202.
- Lu, S. & Cullen, B. R. (2003). Analysis of the stimulatory effect of splicing on mRNA production and utilization in mammalian cells. *RNA* **9**, 618–630.
- Melcher, U., Steffens, D. L., Lyttle, D. J., Lebourier, G., Lin, H., Choe, I. S. & Essenberg, R. C. (1986). Infectious and non-infectious mutants of cauliflower mosaic virus DNA. *J Gen Virol* **67**, 1491–1498.
- Pirone, T. P. & Blanc, S. (1996). Helper-dependent vector transmission of plant viruses. *Annu Rev Phytopathol* **34**, 227–247.
- Pongoski, J., Asai, K. & Cochrane, A. (2002). Positive and negative modulation of human immunodeficiency virus type 1 Rev function by *cis* and *trans* regulators of viral RNA splicing. *J Virol* **76**, 5108–5120.
- Ryabova, L. A., Pooggin, M. M. & Hohn, T. (2002). Viral strategies of translation initiation: ribosomal shunt and reinitiation. *Prog Nucleic Acid Res Mol Biol* **72**, 1–39.
- Schmidt, I., Blanc, S., Espérandieu, P., Kuhl, G., Devauchelle, G., Louis, C. & Cerutti, M. (1994). Interaction between the aphid

transmission factor and virus particles is a part of the molecular mechanism of cauliflower mosaic virus aphid transmission. *Proc Natl Acad Sci U S A* **91**, 8885–8889.

Scholthof, H. B., Wu, F. C., Richins, R. D. & Shepherd, R. J. (1991). A naturally occurring deletion mutant of figwort mosaic virus (caulimovirus) is generated by RNA splicing. *Virology* **184**, 290–298.

Vaden, V. R. & Melcher, U. (1990). Recombination sites in cauliflower mosaic virus DNAs: implications for mechanisms of recombination. *Virology* **177**, 717–726.

Wright, E. A., Heckel, T., Groenendijk, J., Davies, J. W. & Boulton, M. (1997). Splicing features in maize streak virus virion- and complementary-sense gene expression. *Plant J* **12**, 1285–1297.