An internal ribosome entry site located upstream of the crucifer-infecting tobamovirus coat protein (CP) gene can be used for CP synthesis in vivo

Yu. L. Dorokhov,1 P. A. Ivanov,2 T. V. Komarova,1 M. V. Skulachev2 and J. G. Atabekov2

A. N. Belozersky Institute of Physico-Chemical Biology1 and Department of Virology2, Moscow State University, Leninsky Gory 1, Laboratory Building A, Moscow 119992, Russia

It was previously shown that, unlike the type member of the genus *Tobamovirus* (TMV U1), a crucifer-infecting tobamovirus (crTMV) contains a 148 nt internal ribosome entry site (IRES)CR upstream of the coat protein (CP) gene. Here, viral vectors with substitutions in the stem–loop (SL) region of CP subgenomic promoters (TMV U1–CP–GFP/SL-mut and crTMV–CP–GFP/SL-mut) were constructed and the levels of CP synthesis in agroinoculation experiments were compared. No CP–GFP (green fluorescent protein) synthesis was detected in *Nicotiana benthamiana* leaves inoculated with TMV U1–CP–GFP/SL-mut, whereas a small amount of CP–GFP synthesis was obtained in crTMV–CP–GFP/SL-mut-injected leaves. Northern blots proved that both promoters were inactive. It could be hypothesized that IRES-mediated early production of the CP by crTMV is needed for realization of its crucifer-infecting capacity.

Tobacco mosaic virus (TMV) RNA encodes four major proteins. The 126 and 183 kDa replicase proteins are expressed via individual 3′-co-terminal subgenomic (sg) RNAs (Deom et al., 1991; Lewandowski & Dawson, 2000; Mishi et al., 1987). The 30 kDa movement protein (MP) and 17 kDa 4 kDa coat protein (CP) are expressed via individual 3′-co-terminal subgenomic (sg) RNAs (Deom et al., 1987; Hunter et al., 1976; Mishi et al., 1987). The dicistronic intermediate-length RNA-2 called sgRNA I2 is translated to produce the subgenomic (sg) RNAs (Deom et al., 1991; Lewandowski & Dawson, 2000; Mishi et al., 1987). The 30 kDa movement protein (MP) and 17-4 kDa coat protein (CP) are expressed via individual 3′-co-terminal subgenomic (sg) RNAs (Deom et al., 1991; Lewandowski & Dawson, 2000; Mishi et al., 1987). The dicistronic intermediate-length RNA-2 called sgRNA I2 is translated to produce the MP (Bruening et al., 1976; Higgins et al., 1976; Beachy & Zaitlin, 1977; Goelet & Karn, 1982), whereas the 3′-proximal CP gene of I2 RNA is translationally silent. This gene is expressed from a monocistronic sgRNA called low molecular component, LMC (Beachy & Zaitlin, 1977). The 75 nt leader sequence of TMV U1 sgRNA I2, called internal ribosome entry site (IRES)MP,75CR, promotes translation of the downstream ORF in dicistronic reporter constructs (Skulachev et al., 1999). A similar element called IRESMP,75CR has been detected in the RNA of crucifer-infecting tobamovirus (crTMV) (Dorokhov et al., 1993, 1994; Skulachev et al., 1999). The intracellular transport of a movement-deficient TMV U1–KK6 mutant (Lehto et al., 1990) lacking IRESMP,75CR was largely restored by the insertion of IRESMP,75CR, which was apparently due to the translation-enhancing ability of IRESMP,75CR (Zvereva et al., 2004). Furthermore, the 148 nt IRES (IRESCP,148CR) is located upstream of the crTMV-CP gene (Ivanov et al., 1997).

Previously, we suggested that, concurrently with conventional TMV CP gene expression via cap-dependent translation of LMC, the IRESCP,148CR allows cap-independent translation of the CP gene from full-length genomic RNA and/or from RNA I2 of crTMV via an internal ribosome entry mechanism (Ivanov et al., 1997; Dorokhov et al., 2002). It has been shown that cap-independent translation activity mediated by IRESCP,148CR in cell-free systems from plant, animal and yeast cells was higher than that of a widely used IRES from encephalomyocarditis virus RNA (Dorokhov et al., 2002). Analysis of IRESCP,148CR sequence and structure revealed a bulged stem–loop (SL) structure flanked by two polyuridine (A)-rich sequences (PARS), crucial for IRES activity (Dorokhov et al., 2002; Ivanov et al., 1997). Remarkably, the equivalent 148 nt sequence from TMV U1 RNA (U1CP,148CR) was incapable of mediating internal initiation of *in vitro* translation (Dorokhov et al., 2002; Ivanov et al., 1997).

Here, we have examined the contribution of IRESCP,148CR to CP production under conditions when the functional activity of the CP gene sg promoter (SGP) was abolished. Binary vectors containing a tobamovirus genome were delivered to plant cells by the agroinjection technique (Dorokhov et al., 2004), which is known to infect at least 94 % of the cells of injected leaves (Marillonnet et al., 2005). cDNA copies of TMV U1 and crTMV containing GFP fused with the N-terminal part of their CP genes (TMV U1–CP–GFP and crTMV–CP–GFP) were constructed (Fig. 1a and b, respectively). TMV U1–CP–GFP and crTMV–CP–GFP vectors contain viral cDNA that is...
Fig. 1. Agroinjection system for testing of IRESCP,148CR activity. (a and b) Genetic maps of viral vectors based on cDNA copies of TMV U1 and crTMV used for agroinjections of N. benthamiana leaves. Schematic diagram of vector virus genomes: (a) TMV U1-CP–GFP, (b) crTMV-CP–GFP, and the CP-SGP mutants, TMV U1-CP–GFP/SL-mut (a) and crTMV-CP–GFP/SL-mut (b). MET, Methyltransferase; HEL, helicase; POL, polymerase domains of TMV RNA-dependent RNA polymerase; MP, movement protein; CP, coat protein. The positions of MP-SGP including IRESMP,75CR (SGP/RES) and crTMV-CP-SGP including IRESCP,148CR (CP-SGP/RES) are indicated. Schematic drawing of putative stem–loop (SL) structure of wild-type CP-SGP sequence in the minus-copy of genomic RNA is shown. Nucleotide substitutions of SL mutants are indicated. Black arrows mark the intron in the GFP gene and location of the CP mRNA start. Nucleotides are numbered from the CP mRNA start. (c) Image of leaf spots expressing GFP 3 days after agroinjection (UV illumination, 380 nm).
fused to the transcription start site of the actin 2 promoter of *Arabidopsis thaliana* and the nos transcription terminator. Most of the CP gene was substituted by GFP, using the additional *Bam*HI, *Apa*I and *Xba*I sites introduced into the CP sequence and in front of the 3'-non-translated region, respectively. The whole cassette was inserted into the binary vector pBin19 between *Kpn*I and *Sal*I sites (U1) or *Hind*III sites. To allow comparisons, the size of the remaining CP gene sequence (25 codons) was similar for both viral vectors. It should be noted that: (i) an enhancer element is located between nt +25 and +55 with respect to the TMV U1-CP translation start site (Man & Epel, 2004) and (ii) the crTMV-CP gene overlaps the MP gene by 75 nt (Dorokhov et al., 1994). In order to prevent CP–GFP synthesis in agrobacteria, we inserted a small synthetic intron into the GFP ORF.

Fully active TMV U1-CP-SGP was mapped between nt −157 and +54 of the CP ORF and it can be folded into one long SL structure (Grdzelishvili et al., 2000). Deletions of 39 or 59 nt caused unfolding of the stem in this putative structure. Activity of the sgRNA promoter decreased significantly when the length of base-paired sequence in the stem was shortened (Grdzelishvili et al., 2000). In order to inactivate TMV U1-CP–GFP-SGP, we inserted, by overlapping PCRs, 3 nt substitutions into the SL as indicated in the resulting construct TMV U1-CP–GFP/CP-mut (Fig. 1a). Although there is low similarity (18%) between the sequences of CP sgRNA promoters of different tobamoviruses, most of them could be folded into similar SL structures (Grdzelishvili et al., 2000). Our computer-predicted folding of the putative crTMV-CP-SGP also predicted an SL structure. We substituted 4 nt in the stem and, in addition, changed a putative start nucleotide of crTMV-CP–GFP sgRNA to obtain crTMV-CP–GFP/SL-mut (Fig. 1b).

TMV U1-CP–GFP, crTMV-CP–GFP and their CP-SGP mutants were agroinjected into four locations of the same *Nicotiana benthamiana* leaf (Fig. 1c). Efficient fluorescence could be detected under UV illumination in sites injected with TMV U1-CP–GFP and crTMV-CP–GFP. TMV U1-CP–GFP/SL-mut did not produce any fluorescence, whereas TMV U1-CP–GFP, crTMV-CP–GFP and their CP-SGP vectors and the CP gene led to a significant increase of RNA degradation in the absence of the CP. In separate experiments, we showed that joint agroinjection of viral vectors and the CP gene led to a significant increase of genomic RNA accumulation (data not shown). This is consistent with the conclusion (Asurmendi et al., 2004) that the replication of genomic RNA is much more efficient in the presence of the CP.

It is known that agroinjection-mediated transient gene expression is accompanied by induction of gene silencing (Voinnet et al., 2003), suppressing genomic RNA accumulation. In our experiments with another CP-lacking crTMV-based vector, crTMV–GFP (Dorokhov et al., 2004), accumulation of genomic and sgRNA was low (Fig. 2, lane 10), but co-expression of potato virus Y Hc-Pro suppressor of RNA silencing drastically stimulated genomic RNA accumulation (Fig. 2, lane 11).

Importantly, the nucleotide substitutions in CP-SGP abolished the synthesis of CP–GFP sgRNA by crTMV–CP–GFP/SL-mut (Fig. 2, lanes 2 and 4) and TMV U1-CP–GFP/SL-mut (Fig. 2, lanes 6 and 8), indicating that SGPs were completely inactivated. In accordance with this observation, no CP–GFP production by TMV U1-CP–GFP/SL-mut vector could be detected by Western blot analysis (Fig. 3, lanes 4 and 6). It was particularly noteworthy that under the same conditions the CP–GFP
IRESCP, constructed upstream of the CP gene. The location of ORF6, which is not present in the corresponding region of crTMV (Canto et al., 2004).

**Acknowledgements**

This work was partly supported by the Russian Foundation for Basic Research (grants 04-04-48490, 05-04-48674 and 05-04-08002) and Russian Federal Agency of Science and Innovations (contract 02.435.11.3012). The authors would like to thank A. D. Klyushin for assistance in this work.

**References**


Dorokhov, Y. L., Skurat, E. V., Frolova, O. Yu. & 7 other authors (2004). Reciprocal dependence between pectinmethylesterase gene

fusion protein was detected in leaves agroinjected with crTMV-CP–GFP/SL-mut that contained IRES<sub>CP,148</sub><sup>CR</sup> upstream of the CP gene. Fig. 3 shows that protein samples without dilution (lane 3) or after fivefold dilution (lane 5) displayed the CP–GFP band. The protein samples with different dilutions were loaded into the lanes: 1 and 2 (30-fold dilution), 3 and 4 (undiluted), 5 and 6 (fivefold dilution), GFP (5 ng, lane 7) sample was used as a positive control. The membrane was used as a loading control after blotting and staining with amido black (bottom panel).

**Fig. 3.** GFP production by IRES<sub>CP,148</sub><sup>CR</sup>-containing vector virus after blocking the synthesis of CP–GFP sgRNA. Western blot analysis of GFP in samples from leaf material agroinjected with crTMV-CP–GFP (lane 1), TMV U1-CP–GFP (lane 2), crTMV-CP–GFP/SL-mut (lanes 3 and 5) and TMV U1-CP–GFP/SL-mut (lanes 4 and 6). The protein samples with different dilutions were loaded into the lanes: 1 and 2 (30-fold dilution), 3 and 4 (undiluted), 5 and 6 (fivefold dilution), GFP (5 ng, lane 7) sample was used as a positive control. The membrane was used as a loading control after blotting and staining with amido black (bottom panel).

Experiments on in vitro translation proved that nucleotide substitutions in crTMV-CP-SGP did not affect the activity of IRES<sub>CP,148</sub><sup>CR</sup> in expression of the 3’-proximal gene of the bicistronic transcript (data not shown).

Our results indicate that IRES-mediated translation is less efficient than cap-dependent translation. However, IRES-mediated expression of the 3’-proximal gene of polycistronic viral RNA might provide an advantage in virus genome expression. It is possible that IRES<sub>CP,148</sub><sup>CR</sup> provides early CP synthesis that enhances systemic movement of virus, replication and formation of an efficient viral replicative complex (‘virus factory’) (Asurmendi et al., 2004). It could be speculated that early in the viral replicative cycle production of the CP is needed for crucifier-infecting capacity of crTMV. The different strategies of crTMV and TMV U1 in the expression of CP gene might be explained by the location of ORF6, which is not present in the corresponding region of crTMV (Canto et al., 2004).


