

The 5' non-translated region of *Varroa destructor* virus 1 (genus *Iflavirus*): structure prediction and IRES activity in *Lymantria dispar* cells

Juliette R. Ongus,¹ Els C. Roode,¹ Cornelis W. A. Pleij,² Just M. Vlask¹ and Monique M. van Oers¹

Correspondence
Monique M. van Oers
monique.vanoers@wur.nl

¹Laboratory of Virology, Wageningen University, Binnenhaven 11, 6709 PD Wageningen, The Netherlands

²Leiden Institute of Chemistry, Leiden University, Einsteinweg 55, 2333 CC Leiden, The Netherlands

Received 10 April 2006
Accepted 11 July 2006

Structure prediction of the 5' non-translated region (NTR) of four iflavirus RNAs revealed two types of potential internal ribosome entry site (IRES), which are discriminated by size and level of complexity, in this group of viruses. In contrast to the intergenic IRES of dicistroviruses, the potential 5' IRES structures of iflaviruses do not have pseudoknots. To test the activity of one of these, a bicistronic construct was made in which the 5' NTR of *Varroa destructor* virus 1 (VDV-1) containing a putative IRES was cloned in between two reporter genes, enhanced green fluorescent protein and firefly luciferase (Fluc). The presence of the 5' NTR of VDV-1 greatly enhanced the expression levels of the second reporter gene (Fluc) in *Lymantria dispar* Ld652Y cells. The 5' NTR was active in a host-specific manner, as it showed lower activity in *Spodoptera frugiperda* Sf21 cells and no activity in *Drosophila melanogaster* S2 cells.

INTRODUCTION

Viruses within the genus *Iflavirus* resemble viruses in the families *Picornaviridae* and *Dicistroviridae* in genome structure and organization. They all contain single-stranded, positive-sense RNA that is active as messenger and is polyadenylated at the 3' terminus. These viruses have a genome-linked viral protein (VPg) on the 5' end of the genomic RNA. Members of the family *Picornaviridae* and iflaviruses have a single, large open reading frame (ORF) encoding a single polyprotein, whilst in the *Dicistroviridae*, two non-overlapping ORFs in the RNA genome are separated by an intergenic region (Christian *et al.*, 2005a, b). In all of these viruses, the ORFs are flanked by 5' and 3' non-translated regions (NTRs), which contain functionally active RNA structures that are utilized at various stages of the infection process, including polyprotein synthesis and virus replication.

Translation of most cellular mRNAs normally starts with the recognition of the 5'-methylated cap, followed by scanning along the molecule to find the AUG initiation codon (Hershey & Merrick, 2000). In this process, a binary complex of eIF-2 and GTP binds Met-tRNA to form a ternary complex that binds to the 40S ribosomal subunit,

forming the 43S pre-initiation complex. This complex associates with the mRNA cap structure through the translation initiation factor eIF-4F, a complex of eIF4A, eIF4B and eIF4G. The eIF4A component exhibits RNA helicase activity to unwind mRNA secondary structure, allowing access of the AUG start codon to the ribosomal subunits. Finally, the 60S ribosomal subunit associates with the pre-initiation complex to form the 80S translation initiation complex, facilitated by the hydrolysis of the GTP bound to eIF-2 (reviewed by Merrick, 2004).

Instead, RNAs of the *Picornaviridae*, *Dicistroviridae* and probably also iflaviruses are translated via cap-independent mechanisms. They have 5' NTRs that form complex secondary structures which constitute, among others, an internal ribosome entry site (IRES) (Hellen & Sarnow, 2001; Witwer *et al.*, 2001). The best-studied IRES structures are those of the family *Picornaviridae*, of which three types have been distinguished: the entero-/rhinovirus type, the aphtho-/cardio-/hepatovirus type (Witwer *et al.*, 2001) and the recently identified porcine teschovirus type, resembling the IRES of *Hepatitis C virus* (family *Flaviviridae*) (Chard *et al.*, 2006; Pisarev *et al.*, 2004). Members of the family *Dicistroviridae* use an IRES element in the 5' NTR to direct the translation of the first ORF (Masoumi *et al.*, 2003; Wilson *et al.*, 2000), whilst a second IRES element located between the two ORFs drives expression of the downstream ORF (Masoumi *et al.*, 2003;

A supplementary table showing PCR primers used for cloning reporter plasmids is available in JGV Online.

Wilson *et al.*, 2000). IRES elements in the intergenic region of dicistroviruses are the ones that have been characterized most (Domier & McCoppin, 2003; Jan & Sarnow, 2002).

The 5' NTRs of a few virus species in the family *Dicistroviridae* have been tested experimentally for IRES activity. *Cricket paralysis virus* (CrPV; Wilson *et al.*, 2000) has an IRES element within the 709 nt long 5' NTR that is active in a variety of dipteran and lepidopteran cell lines, including *Drosophila melanogaster* (DL2 cells) and *Trichoplusia ni* (TN368 cells) (Masoumi *et al.*, 2003). CrPV was also shown to replicate in these two cell lines (Scotti *et al.*, 1996). These results are in line with the host range of CrPV in nature, which includes insects of the orders Diptera, Lepidoptera, Orthoptera and Heteroptera (Christian & Scotti, 1998; Wilson *et al.*, 2000). *Rhopalosiphum padi virus* (RhpV; Moon *et al.*, 1998), an aphid-borne picorna-like virus, has a 5' NTR of 579 nt with an IRES, which also functions in *Drosophila* embryo extracts and a variety of lepidopteran insect cells including *Spodoptera frugiperda* Sf9 and Sf21 cells, in rabbit reticulocyte lysate and in wheatgerm translation systems (Domier & McCoppin, 2003; Pijlman *et al.*, 2006; Royall *et al.*, 2004; Woolaway *et al.*, 2001). *Triatoma virus* (TrV; Czibener *et al.*, 2000), isolated from the trypanosome *Triatoma infestans*, which transmits Chagas' disease in Argentina, has a 5' NTR of 549 nt that contains an IRES element active in *Xenopus* oocytes (Czibener *et al.*, 2005).

Virus species belonging to the genus *Iflavirus* include *Infectious flacherie virus* (IFV; Isawa *et al.*, 1998), *Sacbrood virus* (SBV; Ghosh *et al.*, 1999), *Varroa destructor virus 1* (VDV-1; Ongus *et al.*, 2004), *Deformed wing virus* (DWV; GenBank accession no. AY292384), *Kakugo virus* (KV; Fujiyuki *et al.*, 2004), *Perina nuda virus* (PnV; Wu *et al.*, 2002) and *Ectropis obliqua* picorna-like virus (EoPV; Wang *et al.*, 2004). In this study, computer-assisted analysis was performed to predict the secondary structures of four iflavirus 5' NTRs so as to identify possible IRES motifs. For this analysis, the relatively long 5' NTR sequences of VDV-1 and DWV were chosen and compared with those of PnV and EoPV, which have much shorter 5' NTR sequences (Table 1). IFV and SBV have very short 5' NTR sequences.

Table 1. Nucleotide sequence length of the 5' NTRs of iflaviruses and the G+C content in this region

Iflavirus	Length of 5' NTR (nt)	G+C content (mol%)
PnV	473	36.79
EoPV	390	33.85
IFV	156	39.10
SBV	178	34.27
KV	1156	31.92
DWV	1139	32.31
VDV-1	1117	32.05

Intraspecies variations in VDV-1 isolates were included in the analysis.

The function of the 5' NTR of VDV-1 was tested experimentally in various cultured insect-cell lines. VDV-1 was originally isolated from the *Varroa destructor* mite, an ectoparasite of the honeybee *Apis mellifera*. The virus is able to replicate both in the mite (Ongus *et al.*, 2004) and in the honeybee (Ongus, 2006), but currently no cell-culture systems are available to facilitate the study of this virus.

METHODS

Prediction of the secondary structure of the iflavirus RNA 5' NTR. Predictions for RNA secondary structures were obtained by using two different algorithms. The Mfold program predicts the overall minimum free energy of the selected RNA molecule (Zuker, 2003), whilst in the genetic algorithm of Gultyaev *et al.* (1995), the RNA is folded sequentially from the 5' to 3' end, thereby simulating the folding process during replication. In the latter program, stems are added to the growing structure by stepwise selection from a list containing only those stems that are compatible with those already incorporated. The selection of the stem to be added depends, among others, on the free energy of the stem. As the folding progresses, less stable structures are removed in favour of more stable motifs, also ensuring that short-range interactions prevail over long-range pairings (Gultyaev *et al.*, 1995). An advantage of the genetic algorithm is the possibility to predict RNA pseudoknot structures, which influence translation initiation by frame-shifting or ribosome read-through (Giedroc *et al.*, 2000). The secondary structures predicted in this way are analysed by comparison with homologous sequences, in order to find support for the proposed secondary structures from co-variations or compensated base changes in stem regions. The preferred sequence requirement for an optimal comparison is an identity ranging from approximately 60 to 80% between the compared homologous RNA molecules (Gardner *et al.*, 2005) to allow sufficient compensatory co-variations that support the existence of a motif. The homologous pairs VDV-1/DWV and PnV/EoPV were selected because they show high conservation in the respective pairwise alignments (see also Table 2).

Comparison of the 5' NTR of various VDV-1 isolates. VDV-1 isolates were obtained from *V. destructor* mites collected from beehives in different locations in Europe: Wageningen in the Netherlands, Avignon in France, Oberürsel and Kirchhain in Germany and Grugliasco and Udine in Italy. A portion of the 5' NTR of the VDV-1 sequences, including the 5' end of the coding region, was amplified by RT-PCR using the primers 5'-CATTACTAGTCCATCTACTCC-3' (1875–1895) to make cDNA from the VDV-1 genome and, for PCR, 5'-GAAGTCGAATACTTGTGTATAGT-3' (397–419) and 5'-ATTACTGATTGAAATGGGGACA-3' (1429–1408), which amplified a product of 1033 bp of VDV-1 sequence. The nucleotide location for each primer is indicated in parentheses and corresponds to the sequence with GenBank accession no. AY251269. The PCR-amplified products were sequenced. The sequences were aligned by using the CLUSTAL_X program (Thompson *et al.*, 1997) and viewed in GeneDoc (Nicholas *et al.*, 1997).

Cloning of reporter plasmids. To test for IRES function in the 5' NTR of VDV-1, a series of reporter plasmids were constructed. The backbone of these reporter plasmids was the pIZ/V5-His plasmid (Invitrogen). The ORFs of enhanced green fluorescent protein (EGFP) and firefly luciferase (Fluc) were cloned into this plasmid

Table 2. Comparison of the nucleotide composition of the 5' NTRs of VDV-1 and DWV

Composition	5' NTR of VDV-1		5' NTR of DWV	
	1–810 (structured)	811–1117	1–818 (structured)	819–1139
A + U	65.80 mol%	73.62 mol%	66.01 mol%	71.96 mol%
G + C	34.20 mol%	26.38 mol%	33.99 mol%	28.04 mol%

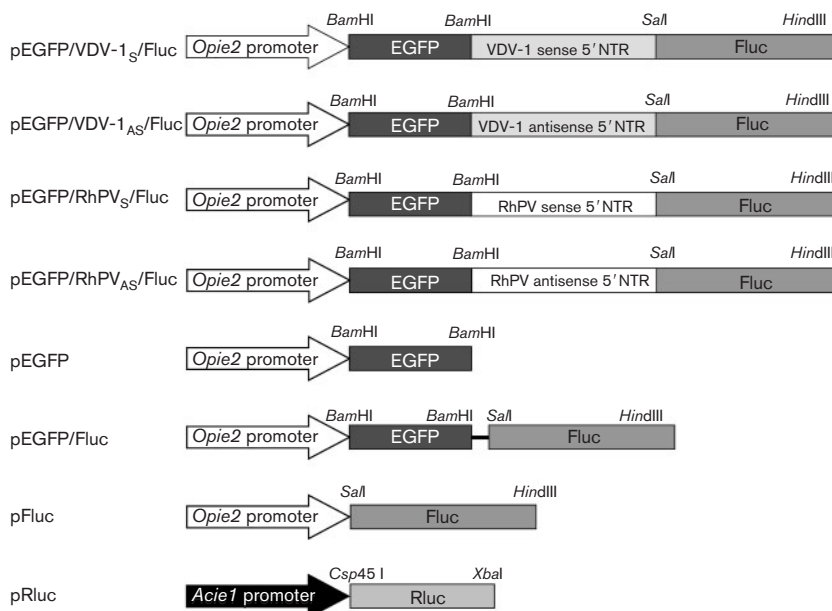
downstream of the *Orgyia pseudotsugata* multiple capsid nucleopolyhedrovirus (OpMNPV) immediate-early 2 (*Opie2*) promoter (Theilmann & Stewart, 1992) (Fig. 1).

In order to make these constructs, PCR products were amplified and first cloned into pGEM-T Easy (Promega). The primers used contained extensions to introduce the appropriate restriction sites (see Supplementary Table S1, available in JGV Online). The sequences of the PCR products were verified by automated sequence analysis (Greenomics, Wageningen, The Netherlands). Fluc was amplified from the plasmid pGL3-control (Promega) and cloned via pGEM-T Easy into pET 28a(+) (Novagen) as a *SalI*–*HindIII* fragment. The Fluc fragment was then recloned into the pIZ/V5 plasmid as a *Bam*HI–*NotI* fragment, generating pFluc. EGFP with an added eukaryotic Kozak consensus ribosome-binding site (GCCACC) (Kozak, 1986) before the translation start codon was amplified by PCR from the plasmid pEGFP (Clontech Laboratories) and cloned as a *Bam*HI fragment into pFluc to give pEGFP/Fluc. The entire 5' NTR of VDV-1, extended with the first six codons of the polyprotein, was amplified from a plasmid clone containing a 5' segment of the virus genome (GenBank accession no. AY251269), again by using primers described in Supplementary Table S1 (available in JGV Online). It was cloned as a *Bam*HI–*SalI* insert into pFluc either in the sense orientation and in frame, giving a fusion with Fluc, or in the antisense orientation. EGFP was cloned into these two constructs as a *Bam*HI fragment to generate the plasmids pEGFP/VDV-1_S/Fluc and pEGFP/VDV-1_{AS}/Fluc, respectively.

Three control plasmids were used in these assays: one with EGFP only (pEGFP), another with EGFP and Fluc not fused and without the

5' NTR (pEGFP/Fluc) and a third with Fluc directly downstream of the *Opie2* promoter (pFluc). As a positive control, the RhPV 5' NTR with the first six codons of the flanking polyprotein was amplified from the pGEM-CAT/RhPVΔ1/LUC plasmid (Woolaway *et al.*, 2001) and cloned as a *Bam*HI–*SalI* fragment between EGFP and Fluc in the sense or antisense orientation (pEGFP/RhPV_S/Fluc and pEGFP/RhPV_{AS}/Fluc). An internal transfection-control plasmid containing *Renilla* luciferase (pRluc) under control of the *Autographa californica* MNPV immediate-early 1 (*Acie1*) promoter (Jarvis *et al.*, 1996) was used to correct for variations in transfection efficiency (Nalçacıoğlu *et al.*, 2003).

Transfection of reporter vectors into insect cells. The insect cells used in this study were *D. melanogaster* S2 cells (Schneider, 1972), *S. frugiperda* Sf21 cells (Vaughn *et al.*, 1977) and *Lymantria dispar* Ld652Y cells (Goodwin *et al.*, 1978). Sf21 and Ld652Y cells were grown in Grace's supplemented insect medium (Invitrogen) with 10% fetal bovine serum (FBS) and the S2 cells in Schneider's *Drosophila* medium (Invitrogen) supplemented with 20% FBS. One million cells were seeded into 35 mm Petri dishes and incubated overnight at 27 °C. Following overnight incubation, the medium over Sf21 and Ld652Y cells was replaced with serum-free Grace's insect medium (SFM) (Sigma) and S2 cells with serum-free Schneider's *Drosophila* medium 2 h before transfection. The amount of Rluc internal plasmid was kept constant in all transfections by making mastermixes. Reporter plasmid (5 µg) and 5 µg of the pRluc internal control plasmid were added to sterile deionized water to a final volume of 25 µl. Ten microlitres of Cellfectin reagent (Invitrogen) was diluted with 15 µl sterile water and added to the

**Fig. 1.** Reporter plasmids used to study IRES activity of the VDV-1 5' NTR. Restriction sites used in cloning are indicated at the top of each bar. The names of the plasmids are indicated on the left.

DNA mix. After 15 min incubation, 500 µl SFM was added to the DNA transfection mix, which was then used to cover the cells. After 1 h incubation at 27 °C with gentle rocking, 500 µl SFM was added to the cells, which were further incubated for 3 h at 27 °C. After that, the medium was replaced by 2 ml Grace's insect medium supplemented with 10% FBS to the Sf21 and Ld652Y cells, and medium with 20% FBS to S2 cells. Incubation was continued at 27 °C for 48 h. The transfections were performed in duplicate.

Luciferase assay. Cells were examined under a fluorescence microscope for EGFP expression and harvested 48 h post-transfection, washed once with PBS (pH 7.4) and suspended in 250 µl 'passive lysis buffer' (Promega's Dual Luciferase reporter assay system). Lysis was allowed to proceed at room temperature for 15 min, followed by two freeze-thaw cycles and vortexing. The lysates were clarified by centrifugation at 14 000 r.p.m. for 30 s in an Eppendorf centrifuge 5417C. Fluc and Rluc expression were measured sequentially by using reagents and protocol from the Dual Luciferase reporter assay system and the readings were made by using a TD-20/20 single-tube luminometer (Turner BioSystems), designed to run Promega's genetic reporter assays. The Fluc values were calibrated by dividing them by the Rluc values to correct for slight variations in transfection efficiency. The means of the Fluc:Rluc ratio for the duplicate experiments for each reporter plasmid were plotted, as well as the standard deviation from the mean.

RNA analysis. RNA was isolated at 40 h post-transfection from 3×10^6 Ld652Y cells that were either mock-transfected or transfected with 10 µg plasmid DNA of pEGFP/VDV-1_S/FLuc or pFLuc. RNA was isolated with TRIzol (Invitrogen). Samples of 5 µg total RNA were glyoxylated and analysed in 1.5% agarose in 15 mM sodium phosphate buffer (pH 6.5). RNA was transferred to Hybond N (Amersham Biosciences) by capillary transfer in 25 mM sodium phosphate buffer (pH 6.5) and hybridized to a 250 bp luciferase probe labelled with [α -³²P]dATP in a PCR amplification with pFLuc as template and with the primers FlucF2 (5'-AAAGAATCCGATCCTGTTCATTCATCACGGTT-3') and FlucR2 (5'-AAAGCGGCCGCTCTAGAGAGAGGGGAGCGCCACCAG-3'). A glyoxylated RNA marker (Promega) was used as a reference.

RESULTS

Iflavirus RNA 5' NTR structures

VDV-1 and DWV have 81% nucleotide identity in the 5' NTR region. The 5' NTR of VDV-1 is 1117 nt long, whereas that of DWV is 22 nt longer as a result of nucleotide insertions (Table 1). PnV and EoPV have much shorter 5' NTRs of 473 and 390 nt, respectively. A pairwise alignment of the entire 5' NTR of PnV and EoPV shows that they have 64% overall nucleotide identity in this region. The PnV sequence has an extension of 83 nt at the 5' end. If the extra 5' sequence is not considered in the alignment, then the identity level rises to 78%. The entire 5' NTR of these four viruses was used for structure prediction.

In the VDV-1/DWV structure (Fig. 2), seven dominant and conserved structural elements are predicted, five of which are hairpins and two (I and V) that formed branched, cloverleaf-like structures. It is interesting to note that the branched structure labelled V has two identical palindromic domains in the VDV-1 sequence at the bulges, highlighted by the rectangles, and parallel sequences that are highlighted by the arrows. In hairpin structure III, the VDV-1 and DWV

sequences were quite different, but the overall configuration was maintained in both. No conserved secondary structure could be proposed so far for the regions in between structures II and III, III and IV and the region from structure VII until just before the underlined AUG start codon, because no co-variation support for the predicted structures could be obtained from the homologous sequence. It may well be that the secondary structure for these regions can be determined more easily when other homologous iflavirus RNA sequences can be included in the analysis.

The conserved structural features predicted for VDV-1 and DWV were located approximately within the first 800 nt. These have a slightly higher G+C content than the remaining portion before the AUG start codon (Table 2), for which no structure could be predicted with high probability.

The PnV/EoPV 5' NTRs had eight dominant structural features (Fig. 3), most of which were hairpins of which two (V and VII) were branched. The hairpins labelled I–IV had a conserved stem and tetraloop at their tips (highlighted in Fig. 3 by using rectangles). The predicted structures were spread out almost evenly along the entire 5' NTR.

Compensatory mutations are the basis of co-variations that support the existence of stems in the two sequences being compared (Witwer *et al.*, 2001). Elements I, III and V of the VDV-1/DWV type structure and element V of the PnV/EoPV type showed good examples of such support. Many other variations between the sequences did not alter the structure predicted, for example, when a G–U pair was changed to a G–C pair and vice versa or, similarly, upon an A–U to G–U pair change. At least two types of conserved structural feature in the 5' NTR of iflaviruses were predicted; the longer VDV-1/DWV type that was based on the VDV-1 sequence (Fig. 2), which is probably also conserved in the closely related KV (Fujiyuki *et al.*, 2004), and the shorter PnV/EoPV type based on the PnV sequence (Fig. 3). Extension of the secondary structures proposed here may well be possible if other iflavirus RNA sequences become available in the near future.

Effect of genetic variation of the predicted structure of the 5' NTR of VDV-1

VDV-1 sequences obtained from virus isolated from *V. destructor* mite specimens collected from various European countries showed 96–99% mutual identity (or 1–4% single-nucleotide polymorphisms). At certain nucleotide positions in the amplified genome region (indicated with an asterisk in Fig. 4), the same nucleotide change was observed in more than one sequence. Real hot spots with more than two nucleotide possibilities at a particular position were not observed; nor were longer stretches of variable nucleotides. The virus sampled from Avignon (France) was related most closely (99% identity) to the VDV-1 sequence from Wageningen (the Netherlands). The variations in the nucleotide sequence do not affect the

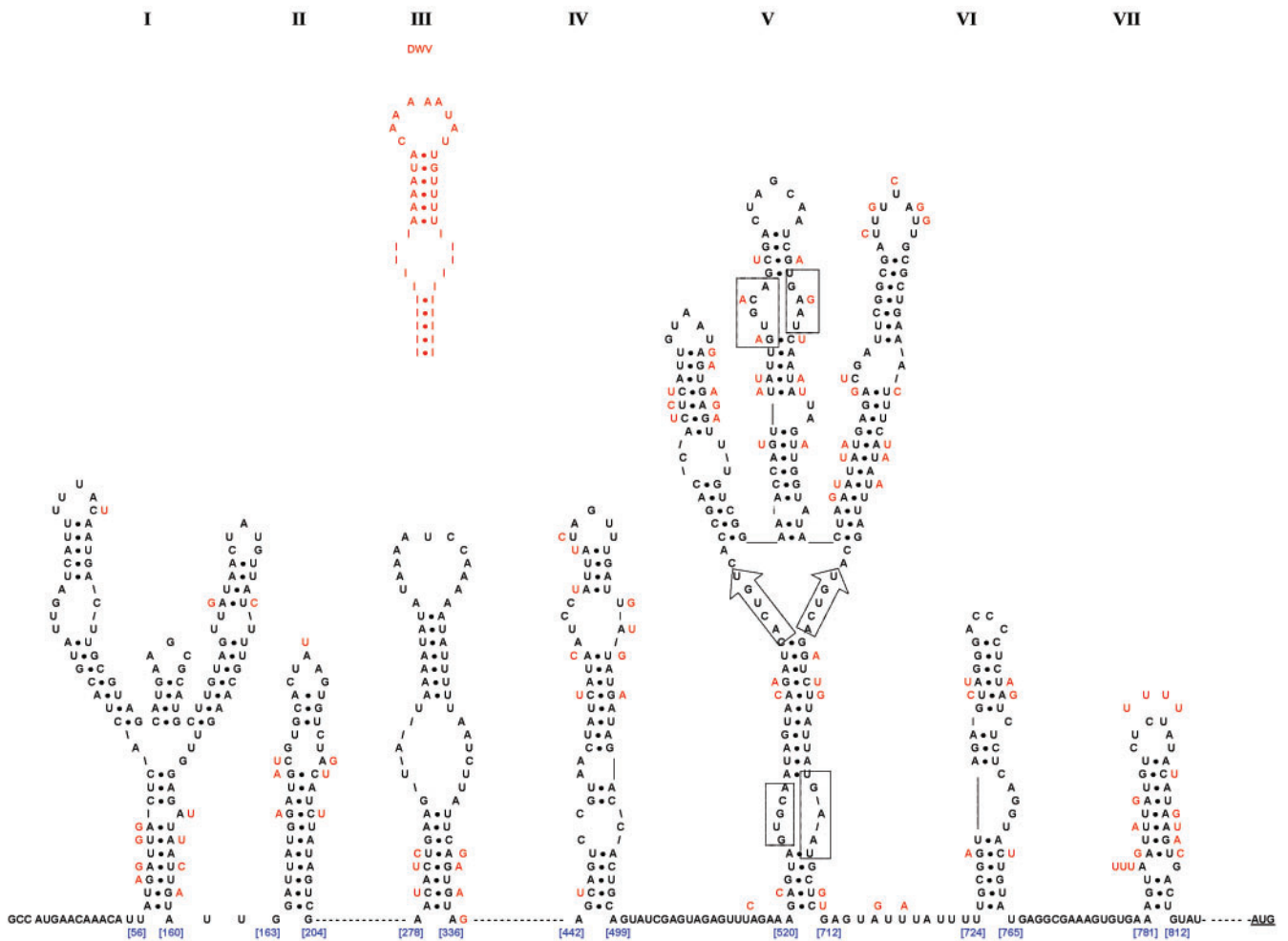


Fig. 2. Predicted conserved secondary structure of VDV-1 and DWV 5' NTRs based on the VDV-1 sequence. Red letters at the periphery indicate the differences in the DWV sequence and thereby show the occurrence of the co-variations in the stem regions. The sequence of DWV in hairpin III is different from that of VDV-1, but the structure is maintained. Numbers at the bottom indicate the nucleotide sequence position in the VDV-1 genome from the 5' end. The lines in between the sequence have been added to give continuity to the two-dimensional structure, except for the dotted breaks at the bottom between structures II and III, between III and IV, and just before the underlined AUG start codon, which indicate the omission of unsupported structures (Table 2).

structure of the 5' NTR. Six nucleotide changes (Fig. 4) were located in cloverleaf structure V (Fig. 2) (nucleotide position 520–712) and did not interfere with this structural element, as they either occurred in the non-paired regions or were substitutions that would not melt the structure.

Fluc expression due to IRES activity in the 5' NTR of VDV-1

The conserved complex secondary structures predicted in the 5' NTR of the iflavirus RNAs point to the presence of putative IRES elements in these regions. Therefore, the ability of the 5' NTR of VDV-1 to function as an IRES in cell culture was examined. An IRES element functions by binding to cellular translation initiation factors, leading to

the recruitment of ribosomes for the translation of a downstream ORF. For this analysis, bicistronic reporter plasmids were designed (Fig. 1) and tested. The translation of the first ORF (EGFP) located directly downstream of the promoter will proceed via a 5'-methylated cap-dependent mechanism. The stop signal at the end of the EGFP ORF prevents the ribosomes from reading through to the second ORF (Fluc). Expression of the second ORF (Fluc) will only occur if there is an intervening IRES element, which will recruit ribosomes, allowing translation of the latter ORF to proceed (Royall *et al.*, 2004).

Following transfection of the bicistronic constructs into *Drosophila* S2 or lepidopteran Sf21 and Ld652Y cells, EGFP expression was verified by UV microscopy. The transfection

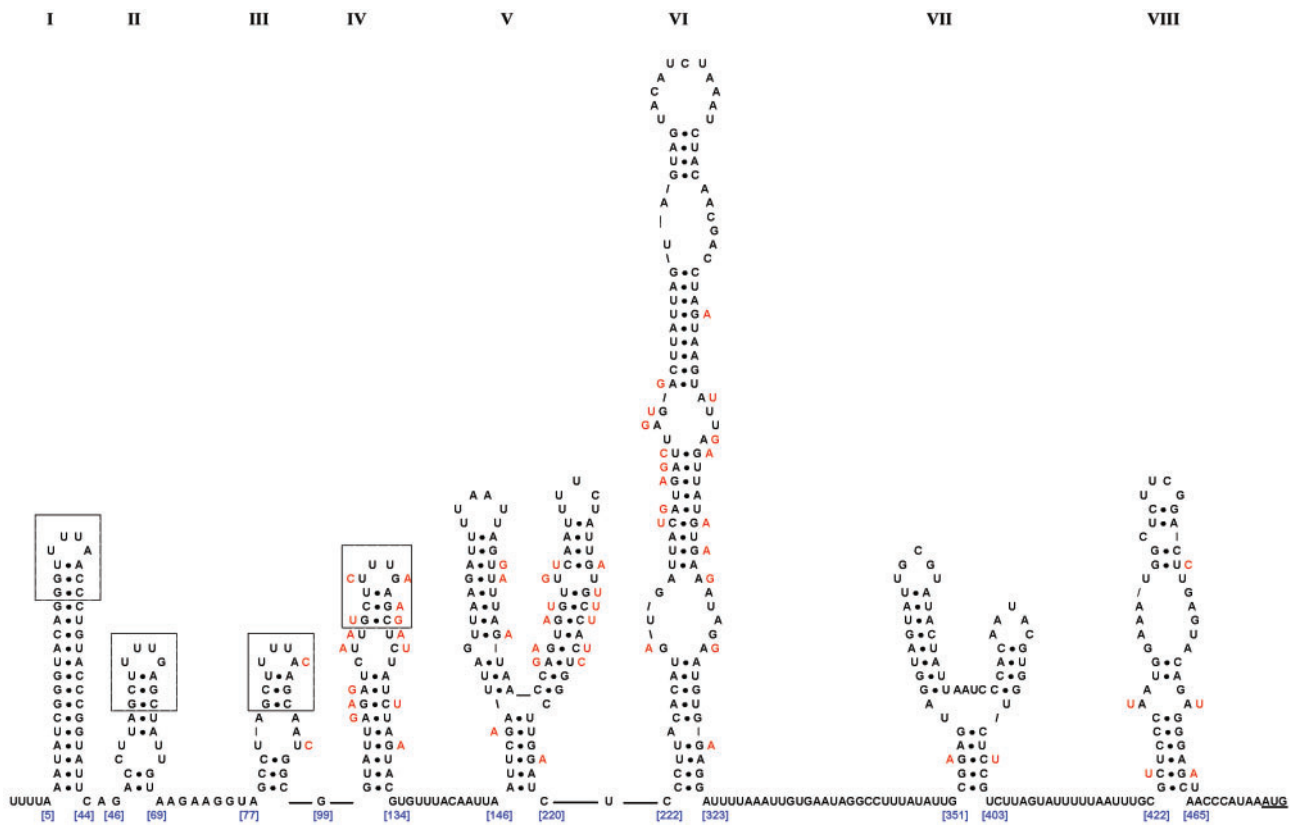


Fig. 3. Predicted conserved secondary structure of PnV and EoPV 5' NTRs based on the PnV sequence. Red letters at the periphery indicate the differences in the EoPV sequence. The numbers at the bottom indicate the nucleotide sequence position in the PnV genome from the 5' end. The lines in between the sequence have been added to prevent overlap in a two-dimensional drawing. The rectangles indicate conserved stem and tetraloop motifs at the tips of hairpins I–IV.

efficiency was in the same order of magnitude (50–60%) for all cell lines, as was the intensity of EGFP expression. Fluc activity was quantified to measure the translation activity of the second ORF. The construct in which the 5' NTR of VDV-1 was present in the sense orientation (pEGFP/VDV-1_S/Fluc) gave significantly higher luciferase activity, about ten times more, than did the antisense construct in Ld652Y cells (Fig. 5a). In Sf21 cells (Fig. 5b), there was an approximately fourfold increase in the luciferase activity when the 5' NTR was placed in a sense orientation between the reporter genes compared with the antisense construct, but the activity was approximately threefold lower than that in Ld652Y cells. There was limited luciferase expression in these two cell lines in the absence of an intervening 5' NTR sequence between EGFP and Fluc, in the same order of magnitude as that measured for the 'antisense-orientation' constructs. Mock-transfected cells and the monocistronic EGFP construct showed no luciferase activity, as expected. Minimal luciferase activity was detected in S2 cells transfected with the various test constructs (Fig. 5c) and no difference was observed in luminescence between the sense and antisense orientations of the 5' NTR of VDV-1. The control clone with Fluc directly behind the *Opie2*

promoter, as expected, gave very high levels of Fluc activity in all cell lines including S2 and had to be diluted 100 times to be in the same range as the samples obtained from the reporter plasmids. For this reason, it was not included in Fig. 5.

The 5' IRES of the dicistrovirus RhPV has been reported as functional in Sf21 cells (Royall *et al.*, 2004) and *Drosophila* embryo extracts (Woolaway *et al.*, 2001), and was included here as a positive control and to estimate the relative strength of the VDV-1 IRES. The 5' NTR of VDV-1 was about twice as strong as that of RhPV in Ld652Y cells, but in Sf21 cells, the strength of these two sequences was similar. In general, Ld652Y cells gave the best results for both VDV-1 and RhPV. The performance of both 5' NTR sequences was poor in *Drosophila* S2 cells, despite the fact that EGFP expression in the *Drosophila* S2 cells was similar to that observed for the other two cell lines.

As this is the first time that a 5' NTR has been analysed for an iflavirus, RNA analysis was also performed on the bicistronic construct in Ld652Y cells to be able to exclude the possibility of monocistronic Fluc transcripts being translated via cap-dependent mechanisms. A complicating

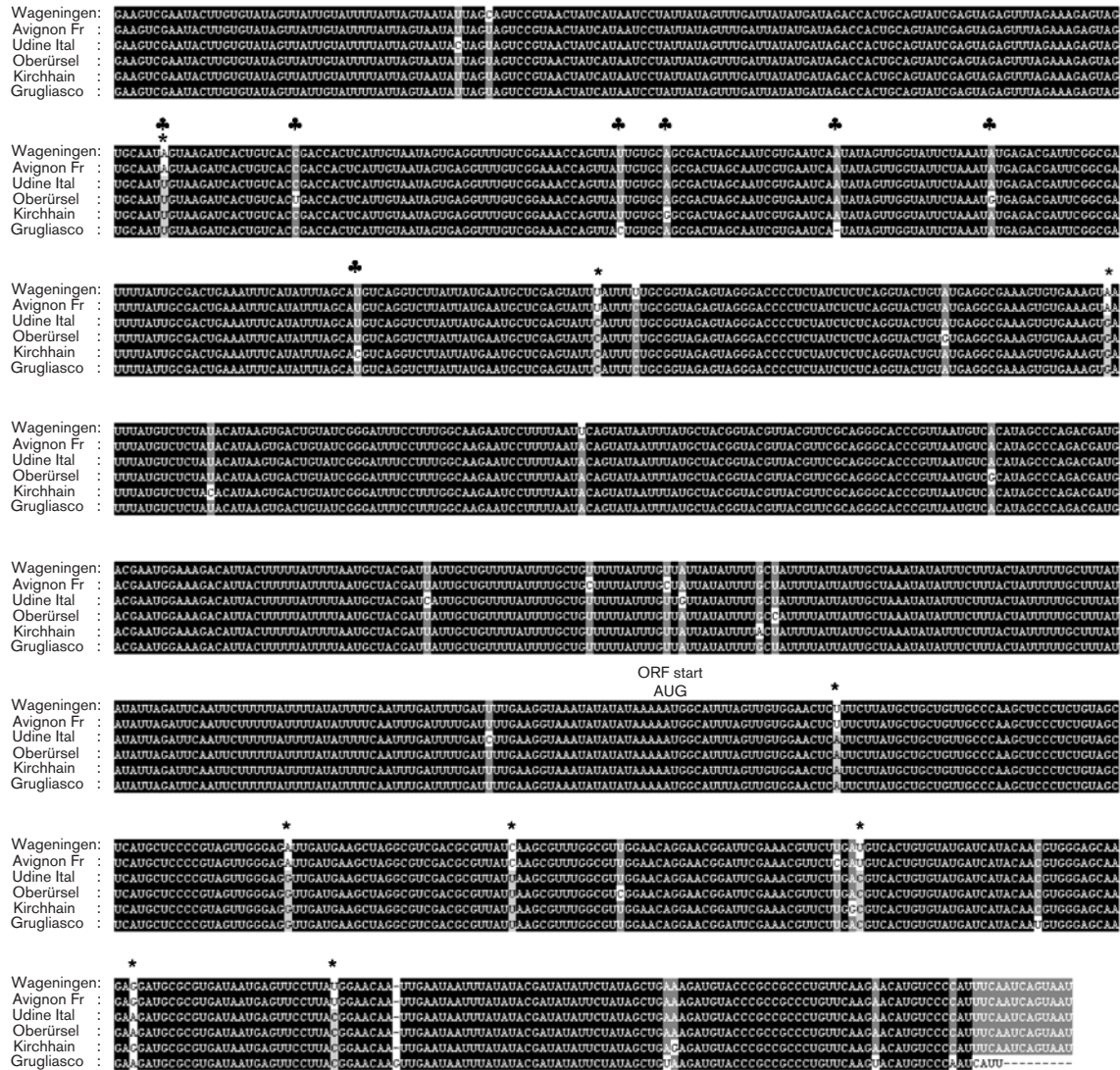


Fig. 4. Alignment of European VDV-1 sequences mapping to the nucleotide position 397–1429 in the virus genome (GenBank accession no. AY251269). An asterisk (*) indicates the location where the same nucleotide change was observed in more than one sequence. The cloverleaf symbol (♣) indicates variations in the cloverleaf structure V (Fig. 2).

factor in these studies is the presence of a sequence motif in the VDV-1 5' NTR resembling baculovirus early gene motifs: a CAGT motif (position 329) is preceded in this case by two TATA boxes (TATAAAT, at positions 289 and 298) that may be active in insect cells and give rise to transcripts initiating within the IRES element.

When the monocistronic pFluc control was used for transfection of Ld652Y cells, a transcript was found of approximately 2000 nt, which is within the expectation range for a transcript of a 1655 nt ORF when polyadenylated (Fig. 6a). Upon transfection with pEGFP/VDV-1_S/Fluc, two transcripts of approximately 3800 and 2800 nt were detected with an Fluc probe. A transcript containing the EGFP, IRES and Fluc sequences should have a size well over 3500 nt, especially when polyadenylated, which corresponds

well with the 3800 nt transcript. The size of the 2800 nt transcript appears to correspond to a putative mRNA start site within the IRES, starting about 800 nt before the Fluc ORF, a position corresponding with the baculovirus early gene motif located at position 329.

DISCUSSION

The nucleotide sequence of single-stranded RNA plays a significant role in how the molecule will be folded. The secondary structure is composed of short-range interactions involving the base pairing of neighbouring nucleotides to form hairpin and loop structures. The structural elements predicted in this study for the 5' NTRs of four iflavirus species are mainly simple hairpins, in which the VDV-1/

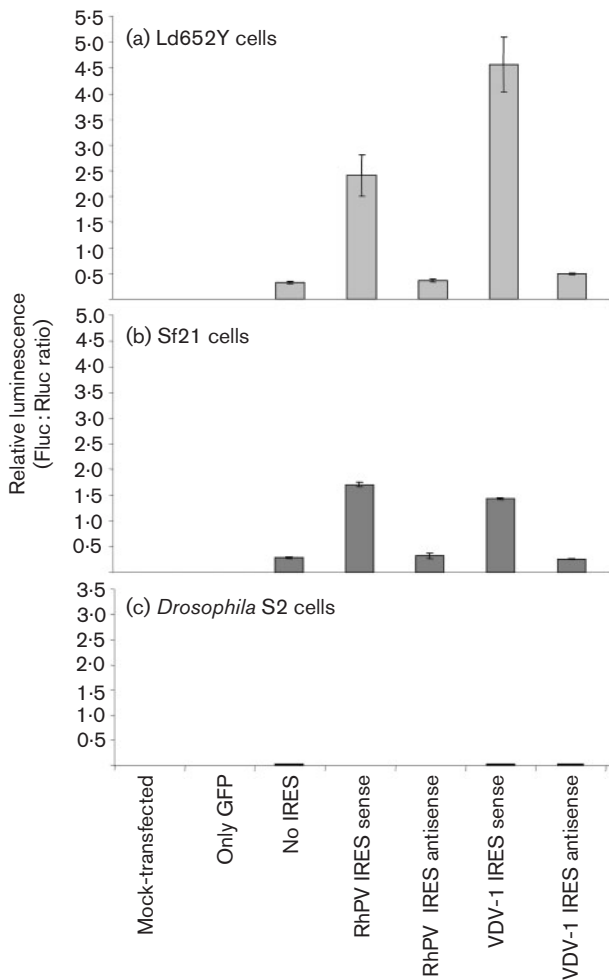


Fig. 5. VDV-1 IRES activity in transfected cells. Luciferase activity was measured in (a) Ld652Y cells (*L. dispar*), (b) Sf21 cells (*S. frugiperda*) and (c) S2 cells (*D. melanogaster*) transfected with various sense and antisense RhPV and VDV-1 IRES constructs. Cells were co-transfected with an Rluc construct and Rluc readings were used to correct for differences in transfection efficiency by plotting the ratio between Fluc and Rluc luminescence. Each bar represents the mean of an experiment done in duplicate for each reporter plasmid and the variation lines show the standard deviation from the mean.

DWV type structure shows more complexity than that of PnV/EoPV. This might be related to the difference in size (about 600 nt) of the 5' NTR. So far, no pseudoknot structures have been detected in either iflavivirus structure type. IRES structures in the intergenic region of the genomes of members of the family *Dicistroviridae* have been shown to have pseudoknot formations that influence translation initiation, for example, the translation of the second ORF of CrPV (Jan & Sarnow, 2002) and RhPV (Domier & McCoppin, 2003) starts from the codon CCU (encoding the amino acid proline).

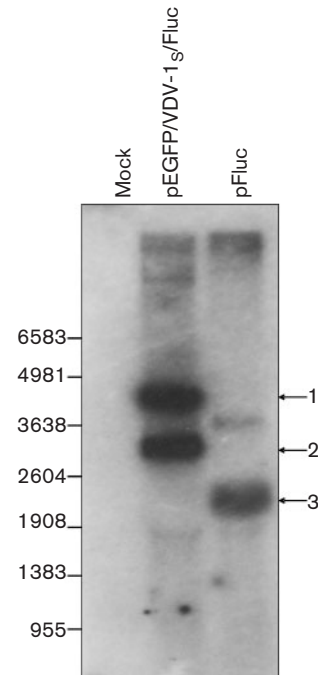


Fig. 6. RNA analysis of transfected *L. dispar* cells. Total RNA was isolated from Ld652Y cells that were either mock-transfected or transfected with the bicistronic plasmid pEGFP/VDV-1_S/Fluc or with the monocistronic pFluc transcript. Northern blot analysis was performed by using an Fluc-specific probe. Arrow 1 indicates the full-length bicistronic transcript, arrow 2 points at a 2800 nt transcript generated from the bicistronic plasmid and arrow 3 indicates the monocistronic Fluc transcript. Size markers (nt) are indicated.

Among the *Picornaviridae*, 5' NTR structures are divided into three groups based on structural similarities: the entero-/rhinovirus group, the aphtho-/cardio-/hepatovirus group and the porcine teschovirus type. However, the IRES structure can vary even within a genus (Witwer *et al.*, 2001). The VDV-1/DWV type structure showed some resemblance to the entero-/rhinovirus type, especially that of *Poliovirus*, which also starts with a three-branched or cloverleaf-like structure. In contrast, the PnV/EoPV structure started with a very stable hairpin on the 5' end, which had eight supporting G-C pairs. The EoPV sequence is 83 nt shorter on the 5' end, where the stable hairpin was located in the PnV sequence. The hairpin appears to be stable ($\Delta G = -28.5 \text{ kcal mol}^{-1}$) such that it was assumed to exist even without supporting substitutions from the EoPV sequence. PnV and EoPV have shorter 5' NTRs with simpler structures that are unique and do not resemble any 5' NTR structures among picornaviruses.

The conserved structural features in the 5' NTR of VDV-1 were mainly confined to the first 810 nt, indicating that the essential elements for the putative IRES are probably located in this region. The three-branched element V (Fig. 2) could

have a vital role for the IRES. This element was the most prominent and contains parallel and palindromic features that could be important recognition signals for translation initiation and binding of translation factors. There was not sufficient evidence, based on co-variations or stable structure formations between structures II and III, III and IV and in the remaining A+U-rich region (the approx. 300 nt before the AUG start codon) of the 5' NTR, to support the prediction of conserved structures between VDV-1 and DWV. These regions could remain mostly single-stranded or could form an elaborate fold at the 3' end of the 5' NTR, such as that seen in enteroviruses (Witwer *et al.*, 2001).

Conserved features were also present in the PnV/EoPV structure (Fig. 3), such as the repeated stem and loop at the tip of hairpins I–IV, which could form important recognition points for translation initiation factors. López de Quinto & Martínez-Salas (1997) compared picornavirus IRES secondary structures and revealed the existence of conserved motifs in the loops. They found that a GNRA motif in the loop of stem 2 and RAAA on stem 4 are required for IRES activity. Similar conserved motifs were identified in the intergenic IRES of some members of the *Dicistroviridae* (Domier *et al.*, 2000). The PnV/EoPV motif in loops I–IV is YUUV (Y stands for C or U in that position and V is for A, C or G). VDV-1 and DWV did not have such conserved loop motifs.

Changes in the 5' NTR can affect virus replication and translation. For example, a C-to-U mutation at base 472 in the poliovirus IRES was demonstrated to alter its tissue tropism and lead to attenuation (Kauder & Racaniello, 2004). It is not known whether the variations observed in the different VDV-1 isolates affect the activity of the 5' NTR in initiating polyprotein translation and perhaps the infectivity of the virus strain, but this may be less likely, as they do not alter the predicted 5' NTR structure.

Apart from the complex secondary structure, the presence of 26 upstream ATGs and three upstream ORFs (positions 44–184, 858–1054 and 915–1038) in the 5' NTR of VDV-1 form another strong indication for IRES-mediated translation. This was further analysed by using bicistronic constructs. Due to the absence of mite and bee cell-culture systems and the fact that the virus is able to replicate in very different cells, alternative cell systems were used to assay the 5' NTR for IRES activity. Ultimately, this may lead to a cell system supporting viral genome replication, as IRES activity enables the translation of viral proteins needed for replication.

To show IRES activity in cell-culture systems, bicistronic reporter plasmids were used. In these constructs, both ORFs are located on the same transcript and the translation of the second ORF will occur only when preceded by an active IRES element. The very low Fluc expression (Fig. 5) with the pEGFP/Fluc construct, where the IRES element was absent, indicated that there was hardly any read-through by

ribosomes once they reached the stop codon of the EGFP ORF. A tenfold enhancement of the Fluc signal was observed in Ld652Y cells when the 5' NTR of VDV-1 was incorporated between the two ORFs. A complication in the analysis is that, apart from full-length transcripts, RNA molecules were detected that appeared to initiate within the IRES sequence, probably due to a cryptic baculovirus-like early promoter motif (Fig. 6). Cap-dependent translation of the Fluc ORF from this truncated transcript is very unlikely, as such a transcript still contains 14 upstream ATG codons (of the original 26) and two overlapping upstream ORFs. Whether such a truncated sequence still has IRES activity is unknown. To further analyse the IRES activity of the VDV-1 IRES, the cryptic promoter may be inactivated by mutating the CAGT motif. As this motif is part of a conserved stem-loop (number III; Fig. 2), this requires additional mutations on the other side of the stem to preserve secondary structure. The CAGT sequence itself is probably not crucial for IRES activity, as it shows co-variation in the DWV 5' NTR (Fig. 2).

The 5' RhPV IRES has been shown to be stronger if the extra codons of the virus polyprotein are omitted from the construct when cloning the 5' NTR (Woolaway *et al.*, 2001). This may also enhance the expression of Fluc in the reporter plasmid construct described in this paper. Even though 5' RhPV IRES activity has been demonstrated in *Drosophila* embryo extracts (Woolaway *et al.*, 2001), no activity was observed in S2 cells in this study.

It is known that picornaviruses exhibit variation in host range, as well as differences in the tissues that they infect (Pilipenko *et al.*, 2000, 2001). Tissue specificity depends on the ability of a picornavirus to bind and enter particular cells and on host-specific factors that interact with the viral genome and that support, for instance, the function of the IRES element (Rueckert, 1996). The 5' and intergenic IRES elements of the dicistrovirus CrPV also displayed different levels of activity in different insect cells (Masoumi *et al.*, 2003). CrPV 5' IRES activity in a cell line is a prerequisite to support genome replication, but does not guarantee that the cell line can support replication. The strength of the 5' IRES activity was not correlated directly to the ability to support CrPV replication, as moderate IRES activity could also support CrPV replication (Masoumi *et al.*, 2003). Host specificity also appears to exist for the VDV-1 IRES (Fig. 5).

The availability of a predicted secondary structure for the 5' NTR of VDV-1 and the demonstration of its activity in Ld652Y cells will facilitate targeted mutation studies to identify motifs crucial for 5' IRES activity. To this aim, the CAGT motif at position 329 should be eliminated first to avoid similar complications in future studies. Disruptions caused by mutations that alter key elements in the structure can then be applied to map the boundaries and most important elements for IRES activity. Mutations may be in the form of deletions that will melt a prominent structural element, for example the branched cloverleaf structure

labelled V in the VDV-1 5' NTR structure (Fig. 2), or nucleotide substitutions that will either undo a structural element or disrupt a conserved motif, for example the conserved palindrome and parallel domains in the structure labelled V in the VDV-1 structure, which are thought to be important recognition sites for translation initiation factors and could be targeted for mutation. By extrapolation from the poliovirus 5' NTR structure and mapping of the IRES (Wimmer *et al.*, 1993), the hypothesis is that the VDV-1 IRES includes the prominent cloverleaf element V and the two hairpins immediately flanking it on either side.

ACKNOWLEDGEMENTS

Dr Lisa Roberts (University of Surrey, UK), is acknowledged for sharing the pGEM-CAT/RhPVΔ1/LUC plasmid. We thank Dr J. Calis and Dr W.-J. Boot of Wageningen University apiary for making available untreated beehives used as the source of Wageningen mites. We are thankful to Dr F. Nazzi (Università di Udine, Udine, Italy), Dr C. Garrido (Bee Institute, Kirchhain, Germany), Dr M. Ollion (INRA/UAPV Écologie des Invertébrés, Avignon, France), Dr R. Ferrero (Università di Torino, Grugliasco, Italy) and Dr J. Kralj (Institut für Bienenkunde, Oberürsel, Germany). J. R. O. was supported financially by the 1221/97 EC programme in co-operation with CNRS, Orléans, France (contract 081.0870.00).

REFERENCES

- Chard, L. S., Kaku, Y., Jones, B., Nayak, A. & Belsham, G. J. (2006). Functional analyses of RNA structures shared between the internal ribosome entry sites of hepatitis C virus and the picornavirus porcine teschovirus 1 Talfan. *J Virol* **80**, 1271–1279.
- Christian, P. D. & Scotti, P. D. (1998). The picorna-like viruses of insects. In *The Insect Viruses*, pp. 301–336. Edited by L. K. Miller & L. A. Ball. New York: Plenum.
- Christian, P., Carstens, E., Domier, L., Johnson, J., Johnson, K., Nakashima, N., Scotti, P. & van der Wilk, F. (2005a). Genus *Iflavirus*. In *Virus Taxonomy: Eighth Report of the International Committee on Taxonomy of Viruses*, pp. 779–782. Edited by C. M. Fauquet, M. A. Mayo, J. Maniloff, U. Desselberger & L. A. Ball. San Diego: Elsevier Academic Press.
- Christian, P., Carstens, E., Domier, L., Johnson, J., Johnson, K., Nakashima, N., Scotti, P. & van der Wilk, F. (2005b). Family *Dicistroviridae*. In *Virus Taxonomy: Eighth Report of the International Committee on Taxonomy of Viruses*, pp. 783–788. Edited by C. M. Fauquet, M. A. Mayo, J. Maniloff, U. Desselberger & L. A. Ball. San Diego: Elsevier Academic Press.
- Czibener, C., La Torre, J. L., Muscio, O. A., Ugalde, R. A. & Scodeller, E. A. (2000). Nucleotide sequence analysis of *Triatoma* virus shows that it is a member of a novel group of insect RNA viruses. *J Gen Virol* **81**, 1149–1154.
- Czibener, C., Alvarez, D., Scodeller, E. & Gamarnik, A. V. (2005). Characterization of internal ribosomal entry sites of *Triatoma* virus. *J Gen Virol* **86**, 2275–2280.
- Domier, L. L. & McCoppin, N. K. (2003). In vivo activity of *Rhopalosiphum padi* virus internal ribosome entry sites. *J Gen Virol* **84**, 415–419.
- Domier, L. L., McCoppin, N. K. & D'Arcy, C. J. (2000). Sequence requirements for translation initiation of *Rhopalosiphum padi* virus ORF2. *Virology* **268**, 264–271.
- Fujiyuki, T., Takeuchi, H., Ono, M., Ohka, S., Sasaki, T., Nomoto, A. & Kubo, T. (2004). Novel insect picorna-like virus identified in the brains of aggressive worker honeybees. *J Virol* **78**, 1093–1100.
- Gardner, P. P., Wilm, A. & Washietl, S. (2005). A benchmark of multiple sequence alignment programs upon structural RNAs. *Nucleic Acids Res* **33**, 2433–2439.
- Ghosh, R. C., Ball, B. V., Willcocks, M. M. & Carter, M. J. (1999). The nucleotide sequence of sacbrood virus of the honey bee: an insect picorna-like virus. *J Gen Virol* **80**, 1541–1549.
- Giedroc, D. P., Theimer, C. A. & Nixon, P. L. (2000). Structure, stability and function of RNA pseudoknots involved in stimulating ribosomal frameshifting. *J Mol Biol* **298**, 167–185.
- Goodwin, R. H., Tompkins, G. J. & McCawley, P. (1978). Gypsy moth cell lines divergent in viral susceptibility. I. Culture and identification. *In Vitro* **14**, 485–494.
- Gultyaev, A. P., van Batenburg, F. H. D. & Pleij, C. W. A. (1995). The computer simulation of RNA folding pathways using a genetic algorithm. *J Mol Biol* **250**, 37–51.
- Hellen, C. U. T. & Sarnow, P. (2001). Internal ribosome entry sites in eukaryotic mRNA molecules. *Genes Dev* **15**, 1593–1612.
- Hershey, J. W. B. & Merrick, W. C. (2000). The pathway and mechanism of initiation of protein synthesis. In *Translational Control of Gene Expression*, pp. 33–88. Edited by N. Sonenberg, J. W. B. Hershey & M. B. Mathews. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- Isawa, H., Asano, S., Sahara, K., Iizuka, T. & Bando, H. (1998). Analysis of genetic information of an insect picorna-like virus, infectious flacherie virus of silkworm: evidence for evolutionary relationships among insect, mammalian and plant picorna(-like) viruses. *Arch Virol* **143**, 127–143.
- Jan, E. & Sarnow, P. (2002). Factorless ribosome assembly on the internal ribosome entry site of cricket paralysis virus. *J Mol Biol* **324**, 889–902.
- Jarvis, D. L., Weinkauff, C. & Guarino, L. A. (1996). Immediate-early baculovirus vectors for foreign gene expression in transformed or infected insect cells. *Protein Expr Purif* **8**, 191–203.
- Kauder, S. E. & Racaniello, V. R. (2004). Poliovirus tropism and attenuation are determined after internal ribosome entry. *J Clin Invest* **113**, 1743–1753.
- Kozak, M. (1986). Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell* **44**, 283–292.
- López de Quinto, S. & Martínez-Salas, E. (1997). Conserved structural motifs located in distal loops of aphthovirus internal ribosome entry site domain 3 are required for internal initiation of translation. *J Virol* **71**, 4171–4175.
- Masoumi, A., Hanzlik, T. N. & Christian, P. D. (2003). Functionality of the 5'- and intergenic IRES elements of *cricket paralysis virus* in a range of insect cell lines, and its relationship with viral activities. *Virus Res* **94**, 113–120.
- Merrick, W. C. (2004). Cap-dependent and cap-independent translation in eukaryotic systems. *Gene* **332**, 1–11.
- Moon, J. S., Domier, L. L., McCoppin, N. K., D'Arcy, C. J. & Jin, H. (1998). Nucleotide sequence analysis shows that *Rhopalosiphum padi* virus is a member of a novel group of insect-infecting RNA virus. *Virology* **243**, 54–65.
- Nałcacioğlu, R., Marks, H., Vlak, J. M., Demirbağ, Z. & van Oers, M. M. (2003). Promoter analysis of the *Chilo* iridescent virus

- DNA polymerase and major capsid protein genes. *Virology* **317**, 321–329.
- Nicholas, K. B., Nicholas, H. B., Jr & Deerfield, D. W., II (1997).** GeneDoc: analysis and visualization of genetic variation. *EMBNEW News* **4**, 14.
- Ongus, J. R. (2006).** *Varroa destructor* virus 1: a new picorna-like virus in *Varroa* mites as well as honey bees. PhD thesis, Wageningen University, The Netherlands.
- Ongus, J. R., Peters, D., Bonmatin, J.-M., Bengsch, E., Vlak, J. M. & van Oers, M. M. (2004).** Complete sequence of a picorna-like virus of the genus *Iflavirus* replicating in the mite *Varroa destructor*. *J Gen Virol* **85**, 3747–3755.
- Pijlman, G. P., Roode, E. C., Fan, X., Roberts, L. O., Belsham, G. J., Vlak, J. M. & van Oers, M. M. (2006).** Stabilized baculovirus vector expressing a heterologous gene and GP64 from a single bicistronic transcript. *J Biotechnol* **123**, 13–21.
- Pilipenko, E. V., Pestova, T. V., Kolupaeva, V. G., Khitrina, E. V., Poperechnaya, A. N., Agol, V. I. & Hellen, C. U. T. (2000).** A cell cycle-dependent protein serves as a template-specific translation initiation factor. *Genes Dev* **14**, 2028–2045.
- Pilipenko, E. V., Viktorova, E. G., Guest, S. T., Agol, V. I. & Roos, R. P. (2001).** Cell-specific proteins regulate viral RNA translation and virus-induced disease. *EMBO J* **20**, 6899–6908.
- Pisarev, A. V., Chard, L. S., Kaku, Y., Johns, H. L., Shatsky, I. N. & Belsham, G. J. (2004).** Functional and structural similarities between the internal ribosome entry sites of hepatitis C virus and porcine teschovirus, a picornavirus. *J Virol* **78**, 4487–4497.
- Royall, E., Woolaway, K. E., Schacherl, J., Kubick, S., Belsham, G. J. & Roberts, L. O. (2004).** The *Rhopalosiphum padi* virus 5' internal ribosome entry site is functional in *Spodoptera frugiperda* 21 cells and in their cell-free lysates: implications for the baculovirus expression system. *J Gen Virol* **85**, 1565–1569.
- Rueckert, R. R. (1996).** *Picornaviridae: the viruses and their replication*. In *Fields Virology*, pp. 609–654. Edited by B. N. Fields, D. M. Knipe & P. M. Howley. Philadelphia, PA: Lippincott-Raven.
- Schneider, I. (1972).** Cell lines derived from late embryonic stages of *Drosophila melanogaster*. *J Embryol Exp Morphol* **27**, 353–365.
- Scotti, P. D., Hoefacker, P. & Dearing, S. (1996).** The production of cricket paralysis virus in suspension cultures of insect cell lines. *J Invertebr Pathol* **68**, 109–112.
- Theilmann, D. A. & Stewart, S. (1992).** Molecular analysis of the *trans*-activating IE-2 gene of *Orgyia pseudotsugata* multicapsid nuclear polyhedrosis virus. *Virology* **187**, 84–96.
- Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. (1997).** The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* **25**, 4876–4882.
- Vaughn, J. L., Goodwin, R. H., Tompkins, G. J. & McCawley, P. (1977).** The establishment of two cell lines from the insect *Spodoptera frugiperda* (Lepidoptera; Noctuidae). *In Vitro* **13**, 213–217.
- Wang, X., Zhang, J., Lu, J., Yi, F., Liu, C. & Hu, Y. (2004).** Sequence analysis and genomic organization of a new insect picorna-like virus, *Ectropis obliqua* picorna-like virus, isolated from *Ectropis obliqua*. *J Gen Virol* **85**, 1145–1151.
- Wilson, J. E., Powell, M. J., Hoover, S. E. & Sarnow, P. (2000).** Naturally occurring dicistronic cricket paralysis virus RNA is regulated by two internal ribosome entry sites. *Mol Cell Biol* **20**, 4990–4999.
- Wimmer, E., Hellen, C. U. T. & Cao, X. (1993).** Genetics of poliovirus. *Annu Rev Genet* **27**, 353–436.
- Witwer, C., Rauscher, S., Hofacker, I. L. & Stadler, P. F. (2001).** Conserved RNA secondary structures in *Picornaviridae* genomes. *Nucleic Acids Res* **29**, 5079–5089.
- Woolaway, K. E., Lazaridis, K., Belsham, G. J., Carter, M. J. & Roberts, L. O. (2001).** The 5' untranslated region of *Rhopalosiphum padi* virus contains an internal ribosome entry site which functions efficiently in mammalian, plant, and insect translation systems. *J Virol* **75**, 10244–10249.
- Wu, C.-Y., Lo, C.-F., Huang, C.-J., Yu, H.-T. & Wang, C.-H. (2002).** The complete genome sequence of *Perina nuda* picorna-like virus, an insect-infecting RNA virus with a genome organization similar to that of the mammalian picornaviruses. *Virology* **294**, 312–323.
- Zuker, M. (2003).** Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* **31**, 3406–3415.