

Short  
CommunicationIdentification of minimal sequences of the  
*Rhopalosiphum padi* virus 5' untranslated region  
required for internal initiation of protein synthesis in  
mammalian, plant and insect translation systemsElisabetta Groppelli,<sup>1</sup> Graham J. Belsham<sup>2†</sup> and Lisa O. Roberts<sup>1</sup>Correspondence  
Lisa O. Roberts  
l.roberts@surrey.ac.uk<sup>1</sup>School of Biomedical and Molecular Sciences, University of Surrey, Guildford, Surrey GU2 7XH, UK<sup>2</sup>BBSRC Institute for Animal Health, Ash Road, Pirbright, Surrey GU24 0NF, UK

*Rhopalosiphum padi* virus (RhPV) is a member of the family *Dicistroviridae*. The genomes of viruses in this family contain two open reading frames, each preceded by distinct internal ribosome entry site (IRES) elements. The RhPV 5' IRES is functional in mammalian, insect and plant translation systems and can form 48S initiation complexes *in vitro* with just the mammalian initiation factors eIF2, eIF3 and eIF1. Large regions of the 5' untranslated region (UTR) can be deleted without affecting initiation-complex formation. The minimal sequences required for directing internal initiation in mammalian (rabbit reticulocyte lysate), plant (wheatgerm extract) and insect (*Sf21* cells) translation systems have now been defined. A fragment (nt 426–579) from the 3' portion of the 5' UTR can direct translation in each of these translation systems. In addition, a distinct region (nt 300–429) is also active. Thus, unstructured regions within the 5' UTR seem to be critical for IRES function.

Received 3 November 2006  
Accepted 9 January 2007

Initiation of translation on most cellular mRNAs occurs through a cap-dependent mechanism that involves the binding of the cap-binding complex (eukaryotic initiation factor complex 4F) to the cap structure (m<sup>7</sup>GpppN) at the 5' end of the mRNA, followed by recruitment of the 40S ribosomal subunit (reviewed by Hershey & Merrick, 2000). However, a number of viral and cellular mRNAs use a cap-independent mechanism termed internal initiation (reviewed by Belsham & Jackson, 2000; Carter *et al.*, 2000; Hellen & Sarnow, 2001). In this case, initiation of protein synthesis is directed by an internal ribosome entry site (IRES) element within the 5' untranslated region (UTR) of the mRNA. IRES elements have been identified in viral and cellular mRNAs; however, there is much variation in their structures and requirements for initiation factors.

Some of the best-characterized IRES elements are located in the 5' UTRs of picornavirus genomes. The picornavirus IRES elements have generally been grouped into two major classes according to their secondary structure and activity *in vitro*. The cardio- and aphthovirus IRES elements, which function very efficiently in the rabbit reticulocyte lysate

(RRL) system, form one group. The entero- and rhinovirus IRES elements make up the second group and these function poorly in RRL. However, the activity of these IRES elements in RRL is greatly enhanced if supplemented with HeLa cell extracts (Brown & Ehrenfeld, 1979; Dorner *et al.*, 1984). A third type of picornavirus IRES element has recently been described in the porcine teschovirus 1, porcine enterovirus 8 and simian virus 2 genomes; these IRES elements show significant similarity to the IRES from *Hepatitis C virus* (a member of the family *Flaviviridae*) (Pisarev *et al.*, 2004; Chard *et al.*, 2006a, b). It is now clear that picornavirus IRES elements require cellular (*trans-acting*) factors for their function, in addition to the canonical translation initiation factors, and that these factors may play an important role in IRES (and virus) tropism. Examples of these proteins are the La autoantigen, polypyrimidine tract-binding protein (PTB), poly(C)-binding protein and unr (upstream of N-Ras) (reviewed by Belsham & Jackson, 2000).

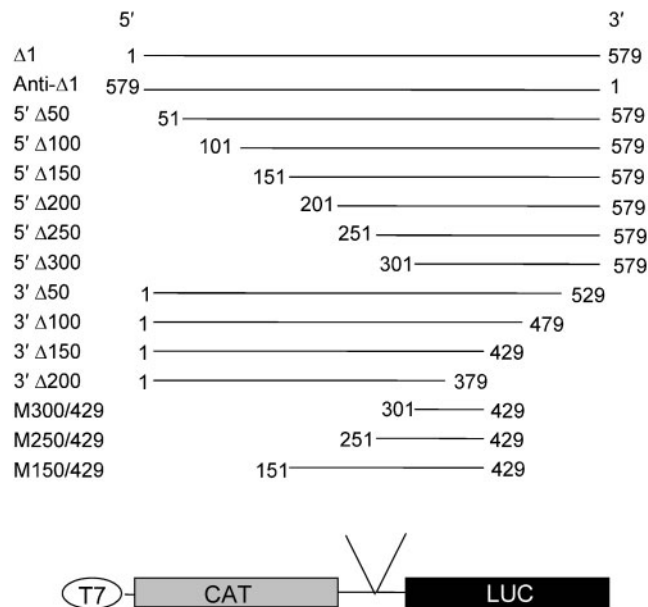
*Rhopalosiphum padi* virus (RhPV) belongs to the family *Dicistroviridae*, which also includes *Cricket paralysis virus* (CrPV), *Plautia stali intestine virus* and *Drosophila C virus*. RhPV has a single-stranded, positive-sense RNA genome that contains two separate open reading frames (ORFs) (Moon *et al.*, 1998). ORF1 encodes the non-structural proteins and ORF2 encodes the structural proteins. The 5' UTRs and the intergenic regions (IGRs) of the dicistrovirus

†Present address: National Veterinary Institute, Technical University of Denmark, Lindholm, 4771 Kalvehave, Denmark.

A supplementary table showing primer sequences is available in JGV Online.

genomes each contain an IRES element (Sasaki & Nakashima, 1999; Domier *et al.*, 2000; Wilson *et al.*, 2000a; Woolaway *et al.*, 2001). The IGR IRES of CrPV does not require any of the initiation factors for the assembly of an initiation complex on the mRNA, which occurs at a non-AUG codon (Wilson *et al.*, 2000b). The properties of the 5' IRES elements are very different and they direct translation initiation from AUG codons (Sasaki & Nakashima, 1999; Domier *et al.*, 2000; Wilson *et al.*, 2000a; Woolaway *et al.*, 2001). We have shown previously that the RhPV 5' IRES functions in mammalian (RRL), insect (*Drosophila* lysate, *Sf21* lysate and *Sf21* cells) and plant (wheatgerm extract; WGE) systems (Woolaway *et al.*, 2001; Kubick *et al.*, 2003; Royall *et al.*, 2004). The ability of the IRES to function in all three systems suggested a mechanism of internal initiation with a less specific requirement for translation initiation factors. Indeed, we have shown recently that the RhPV 5' IRES requires only the mammalian initiation factors eIF2, eIF3 and eIF1 to form 48S complexes *in vitro* (Terenin *et al.*, 2005). The factors eIF1A and eIF4F stimulate the assembly of 48S complexes on the RhPV RNA, but are not essential. We also found that it was possible to delete large regions of the 5' UTR without affecting initiation-complex formation. However, deletion of an unstructured region in the 380 nt proximal to the initiation codon reduced 48S complex formation significantly. As this previous work was carried out using only mammalian factors, we set out to define the minimal RhPV sequences required for directing internal initiation in translation systems from mammals (RRL), plants (WGE) and insects (*Sf21* cells). Here, we report that a small fragment from the 3' end of the RhPV 5' UTR, corresponding to nt 425–579, retains significant IRES activity in all three translation systems, whereas nt 475–579 lack IRES activity. Intriguingly, a fragment corresponding to nt 300–429 is also capable of directing internal initiation in all three systems. The ability of different regions or fragments of the RhPV 5' UTR to display IRES activity is distinct from viral IRES elements described previously, but is reminiscent of some cellular IRES elements.

In order to define the minimal sequence that can direct internal initiation of protein synthesis in different translation systems, we created truncated versions of the cDNA corresponding to the RhPV 5' UTR and inserted the fragments between two reporter sequences. Fragments lacking sequences from either the 5' or 3' end were obtained by PCR using specific primers (see Supplementary Table S1, available in JGV Online). All primers contained *Bam*HI restriction-enzyme sites to facilitate cloning into the pGEM-CAT/LUC vector between the two ORFs, as described previously (van der Velden *et al.*, 1995). The plasmid pGEM-CAT/RhPV $\Delta$ 1/LUC, which contains the entire 5' UTR (nt 1–579; Woolaway *et al.*, 2001), was used as template for each PCR. Deletions of 50, 100, 150, 200, 250 and 300 nt from the 5' end were created (Fig. 1). Similarly, deletions of 50, 100, 150 and 200 nt from the 3' end of the RhPV IRES were also generated. Further

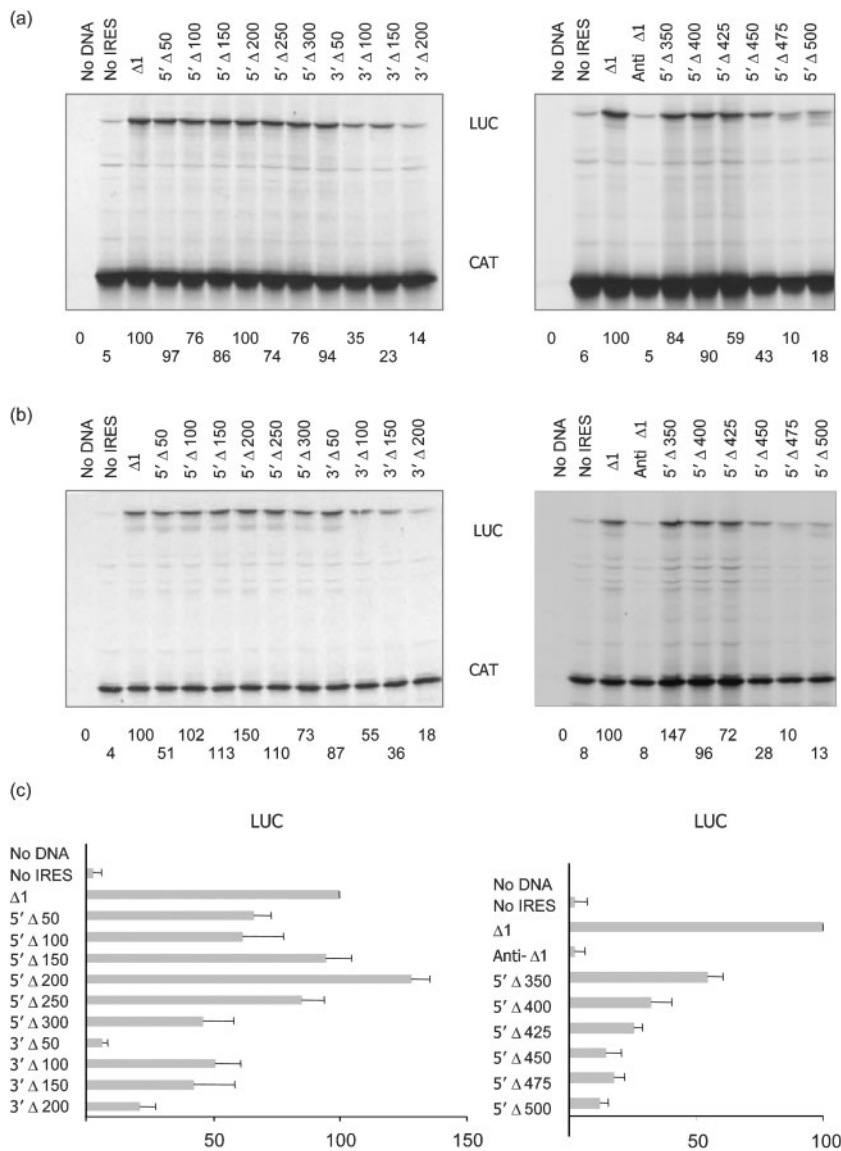


**Fig. 1.** Schematic representation of the truncated versions of the RhPV 5' IRES. The cDNA fragments were obtained by PCR and inserted in the intercistronic space of the dicistronic vector pGEM-CAT/LUC. The first cistron, CAT, encodes chloramphenicol acetyltransferase and is translated by a cap-dependent mechanism, whilst the second cistron encodes firefly luciferase (LUC) and is translated by an IRES-dependent mechanism.

deletions at the 3' end were not made, as we have shown previously that a 200 nt deletion from the 3' end reduced IRES activity significantly (Woolaway *et al.*, 2001). All resulting constructs were verified by sequencing.

To examine the activity of the RhPV sequences in RRL, the plasmids were used to program *in vitro* coupled transcription/translation (TNT) reactions (Promega). Reactions contained [<sup>35</sup>S]methionine and the products were analysed by using SDS-PAGE and autoradiography. Firefly luciferase (LUC) expression was also measured by luciferase assay (Promega) according to the manufacturer's instructions. Plasmids pGEM-CAT/LUC, which lacks any IRES sequence, and pGEM-CAT/ $\Delta$ 1/LUC, containing the wild-type (*wt*) RhPV IRES, were used as negative and positive controls, respectively. All of the constructs expressed similar levels of the upstream cistron, chloramphenicol acetyltransferase (CAT), as expected (Fig. 2a). All constructs containing regions of the RhPV 5' UTR were also able to direct expression of LUC, although with different efficiencies. Deletion of up to 300 nt from the 5' end had little effect on IRES activity, with between 74 and 100% of *wt* activity being observed.

As removal of 300 nt from the 5' end of the RhPV UTR had little effect on IRES activity, further deletions of 350, 400, 425, 450, 475 and 500 nt were made (primer sequences are given in Supplementary Table S1, available in JGV



**Fig. 2.** Delimitation of the 5' and 3' borders of the RhPV 5' UTR sequence required for IRES activity in mammalian, plant and insect systems. Plasmids that express mRNAs of the form CAT/IRES/LUC were analysed in RRL (a), WGE (b) and *Sf21* cells (c) as described in the text. For RRL and WGE assays, the samples were analysed by SDS-PAGE and autoradiography. LUC assays were also performed and values relative to the *wt*  $\Delta 1$  construct are shown. LUC values were normalized against CAT expression. Results are representative of three separate experiments. For the *Sf21* cells, CAT and LUC expression was measured as described in the text and LUC values were normalized against CAT expression. Results are the means (+SEM) from three separate experiments.

Online). The fragments lacking up to 450 nt from the 5' end retained significant IRES activity in RRL (Fig. 2a, right panel), but further deletion to nt 475 reduced activity to near background level.

We then explored the 3' boundary of the IRES. Deletion of 50 nt from the 3' end of the RhPV 5' UTR (construct 3'  $\Delta 50$ ) did not have any major effect on IRES activity, but a further deletion of 50 nt (construct 3'  $\Delta 100$ ) reduced activity to about 35% of *wt*. Deletion of 150 nt (construct 3'  $\Delta 150$ ) reduced activity further, and deletion of the 3' 200 nt resulted in a complete loss of activity (c.f. no IRES control). Most of these results are consistent with previous experiments (Terenin *et al.*, 2005). However, in the previous study, we observed that deletions from the 3' end of the IRES affected IRES activity in RRL only marginally. There is a key difference between the two studies: the earlier experiments were performed by using monocistronic transcripts with a hairpin located at the 5' end,

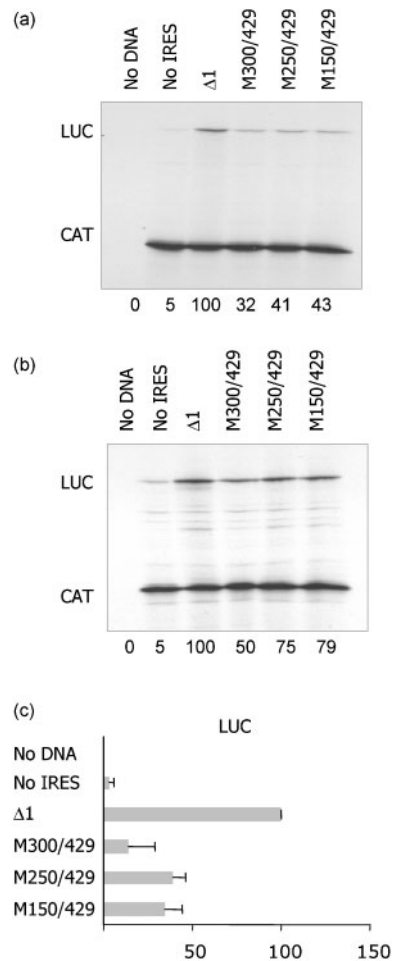
rather than by using dicistronic constructs (as used in the current study). The latter appear more stringent (Robertson *et al.*, 1999).

In order to define the minimal 5' UTR sequence required for IRES activity in plant systems, the plasmids were also assayed in the WGE TNT system (Promega). Reactions were analysed by SDS-PAGE/autoradiography and LUC assay as described above (Fig. 2b). All of the constructs expressed similar levels of CAT, as expected. As in the RRL system, the 5'-end deletions were tolerated much more than deletions from the 3' end. However, in the WGE, five of the six 5'-end mutants directed LUC expression more efficiently than the full-length 5' UTR in the  $\Delta 1$  construct. The 5'  $\Delta 200$  mutant displayed IRES activity of 150% compared with the full-length 5' UTR. This suggests that the 5' sequences may actually be inhibitory in WGE and that their removal creates a more efficient IRES element. As seen in RRL, the 3' deletions had a major effect on activity, and

removal of the 3' 200 nt resulted in very little internal initiation.

The same reporter plasmids were also assayed in an insect-cell system by using a transient expression system in *Sf21* cells, as described previously (Royall *et al.*, 2004). Cells (in 60 mm dishes) were infected with the recombinant baculovirus AcT7N, which expresses T7 RNA polymerase (a gift from Monique van Oers, University of Wageningen, The Netherlands; van Poelwijk *et al.*, 1995) and subsequently transfected with plasmid DNA by using Lipofectin (Invitrogen). Lysates (prepared 48 h later) were analysed for CAT expression by using a CAT ELISA (Roche) and for LUC expression as described above. LUC activities were normalized against CAT expression to take account of any differences in transfection efficiency between the plasmids (Fig. 2c). Consistent with the observations made in RRL and WGE, the 5'-end deletion mutants displayed higher IRES activity than the 3'-end mutants. As observed in the WGE system, four of the six 5'-end mutants directed higher LUC expression than the full-length 5' UTR; the 5'  $\Delta 200$  construct exhibited about twice this activity. Interestingly, removal of the 3' 50 nt resulted in a dramatic decrease in LUC expression (to about 10% of the  $\Delta 1$  value), which contrasts with the activity of this sequence in the RRL and WGE systems. However, a further deletion of 50 nt (construct 3'  $\Delta 100$ ) restored the activity. This suggests that the 3'-proximal structures may play a more critical role in translation initiation in the insect-cell system. It should be noted that, in all three systems, the presence of the 5' UTR sequences in the antisense orientation resulted in no significant IRES activity (Fig. 2).

From these results, it was apparent that deletion of up to 300 nt from the 5' end and 150 nt from the 3' end of the RhPV 5' UTR could be tolerated without losing IRES activity completely in any of these three assay systems. We therefore created three further mutants, with deletions at both ends, to define the minimal sequence that can display IRES activity. Removal of the 5' 150 nt and 3' 150 nt resulted in construct M150/429 and was obtained by PCR using the primers 5'  $\Delta 150$  F and 3'  $\Delta 150$  R. Deletion of the 5' 250 nt and 3' 150 nt created construct M250/429 and was generated with primers 5'  $\Delta 250$  F and 3'  $\Delta 150$  R. Loss of the 5' 300 nt and 3' 150 nt created construct M300/429 and this was generated by using primers 5'  $\Delta 300$  F and 3'  $\Delta 150$  R. All plasmids were assayed in RRL, WGE and *Sf21* cells. Remarkably, the combined deletion mutants still retained at least 30–40% of the activity of the full-length  $\Delta 1$  sequence in all three systems tested, and up to about 80% of the *wt* IRES activity in WGE (Fig. 3). The smallest combined deletion mutant (M300/429) retained about half of the activity of  $\Delta 1$ , indicating that the region from nt 300 to 429 can function as an IRES. The same results were obtained in RRL programmed with *in vitro*-derived RNA transcripts (data not shown). Since, as shown above, nt 426–579 in 5'  $\Delta 425$  also functions, these data suggest that more than one region of the RhPV 5' UTR can function to direct internal initiation.



**Fig. 3.** Internal fragments of the RhPV 5' IRES display IRES activity. Indicated plasmids of the form CAT/IRES/LUC were analysed in RRL (a), WGE (b) and *Sf21* cells (c) as described in the text and legend to Fig. 2. The names of the plasmids refer to the nucleotides from the RhPV 5' UTR contained within them.

We have demonstrated previously that the RhPV 5' IRES functions efficiently in mammalian, plant and insect translation systems (Woolaway *et al.*, 2001; Royall *et al.*, 2004). The unusual characteristics of this IRES element were studied further by analysing the factors required for initiation-complex formation and the secondary structure of the IRES. We showed that large deletions could be made in the RhPV 5' UTR sequence without losing IRES function by using mammalian initiation factors and ribosomes (Terenin *et al.*, 2005). In this study, we now demonstrate that such large deletions (up to 425 nt from the 5' end) can also be tolerated without affecting activity significantly in WGE and *Sf21* cells. In fact, the 3'-proximal sequences functioned at least as well as the complete 5' UTR in the WGE and *Sf21* systems, and sometimes more efficiently. Furthermore, we have shown that small, internal fragments of the RhPV 5' UTR (130 nt in the case of the smallest fragment, nt 300–429) are able to

direct internal initiation in all three systems to about 40% of the efficiency of the *wt*  $\Delta$ 1 sequence. From previous work (Terenin *et al.*, 2005), we know that these regions are largely unstructured. Other viral IRES elements, for example those from the picornaviruses, cannot tolerate such large deletions. However, certain IRES elements from cellular mRNAs have this characteristic. For example, fragments of the immunoglobulin heavy chain-binding protein (BiP) and *c-myc* IRES elements retain function; however, they do not work in RRL (Yang & Sarnow, 1997; Stoneley *et al.*, 2000).

We have proposed previously (Terenin *et al.*, 2005) that the RhPV 5' IRES probably consists of several regions that are able to bind ribosomes and factors, and that the absence of structure facilitates this process. The studies presented here support this idea and show that minimal sequences of approximately 130 nt are important for IRES activity in mammalian, plant and insect systems. It is worth noting, however, that this 'cross-kingdom' activity has not been demonstrated for other 5' IRES elements from the family *Dicistroviridae*. For example, the CrPV 5' UTR does not function in WGE (Wilson *et al.*, 2000a). It may be that the ability of the RhPV 5' UTR to function in all three systems allows (or has historically allowed) the virus mRNA to be translated within cells of different origin, and has therefore aided survival and competition.

Finally, we have previously reported the utility of the RhPV 5' IRES element in the baculovirus expression system (Pijlman *et al.*, 2006) and *in vitro* protein expression systems (Royall *et al.*, 2004). Use of smaller fragments of the RhPV IRES in such systems, rather than the full-length version, may be beneficial in the construction of such vectors.

## Acknowledgements

This work was supported by the BBSRC. E. G. gratefully acknowledges receipt of a UniS School of Biomedical and Molecular Sciences studentship award. We also thank Marion Chadd for excellent technical assistance and Elizabeth Royall for critical reading of the manuscript.

## References

Belsham, G. J. & Jackson, R. J. (2000). Translation initiation on picornavirus RNA. In *Translational Control of Gene Expression*, 2nd edn, pp. 869–900. Edited by N. Sonenberg, J. W. B. Hershey & M. B. Mathews. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Brown, B. A. & Ehrenfeld, E. (1979). Translation of poliovirus RNA in vitro: changes in cleavage pattern and initiation sites by ribosomal salt wash. *Virology* **97**, 396–405.

Carter, M. S., Kuhn, K. M. & Sarnow, P. (2000). Cellular internal ribosome entry site elements and the use of cDNA microarrays in their investigation. In *Translational Control of Gene Expression*, 2nd edn, pp. 615–634. Edited by N. Sonenberg, J. W. B. Hershey & M. B. Mathews. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Chard, L. S., Kaku, Y., Jones, B., Nayak, A. & Belsham, G. J. (2006a). 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.

Chard, L. S., Bordeleau, M. E., Pelletier, J. & Belsham, G. J. (2006b). Hepatitis C virus-related internal ribosome entry sites are found in multiple genera of the family *Picornaviridae*. *J Gen Virol* **87**, 927–936.

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.

Dorner, A. J., Semler, B. L., Jackson, R. J., Hanecak, R., Duprey, E. & Wimmer, E. (1984). *In vitro* translation of poliovirus RNA: utilization of internal initiation sites in reticulocyte lysate. *J Virol* **50**, 507–514.

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). Pathway and mechanism of initiation of protein synthesis. In *Translational Control of Gene Expression*, 2nd edn, pp. 33–88. Edited by N. Sonenberg, J. W. B. Hershey & M. B. Mathews. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Kubick, S., Schacherl, J., Fleischer-Notter, H., Royall, E., Roberts, L. O. & Stiege, W. (2003). *In vitro* translation in an insect-based cell-free system. In *Cell-Free Protein Expression*, pp. 209–217. Edited by J. R. Swartz. Berlin: Springer.

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 viruses. *Virology* **243**, 54–65.

Pijlman, G. P., Roode, E. C., Fan, X., Roberts, L. O., Belsham, G. J., Viak, 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.

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.

Robertson, M. E. M., Seamons, R. A. & Belsham, G. J. (1999). A selection system for functional internal ribosome entry site (IRES) elements: analysis of the requirement for a conserved GNRA tetraloop in the encephalomyocarditis virus IRES. *RNA* **5**, 1167–1179.

Royall, E., Woolaway, K. E., Schacherl, J., Kubick, S., Belsham, G. J. & Roberts, L. O. (2004). The *Rhopalosiphum padi* virus 5' IRES is functional in *Spodoptera frugiperda* cells and in their cell-free lysates: implications for the baculovirus expression system. *J Gen Virol* **85**, 1565–1569.

Sasaki, J. & Nakashima, N. (1999). Translation initiation at the CUU codon is mediated by the internal ribosome entry site of an insect picorna-like virus in vitro. *J Virol* **73**, 1219–1226.

Stoneley, M., Subkhanulova, T., LeQuesne, J. P. C., Coldwell, M., Jopling, C., Belsham, G. J. & Willis, A. E. (2000). Analysis of the *c-myc* IRES; a potential role for the cell-type specific trans-acting factors and the nuclear compartment. *Nucleic Acids Res* **28**, 687–694.

Terenin, I. M., Dmitriev, S. E., Andreev, D. E., Royall, E., Belsham, G. J., Roberts, L. O. & Shatsky, I. N. (2005). A 'cross-kingdom' IRES reveals a simplified mode of internal ribosome entry. *Mol Cell Biol* **25**, 7879–7888.

van der Velden, A., Kaminski, A., Jackson, R. J. & Belsham, G. J. (1995). Defective point mutants of the encephalomyocarditis virus internal ribosome entry site can be complemented *in trans*. *Virology* **214**, 82–90.

van Poelwijk, F., Broer, R., Belsham, G. J., Oudshoorn, P., Viak, J. M. & Goldbach, R. W. (1995). A hybrid baculovirus-bacteriophage T7 transient expression system. *Biotechnology (N Y)* **13**, 261–264.

**Wilson, J. E., Powell, M. J., Hoover, S. E. & Sarnow, P. (2000a).** Naturally occurring dicistronic cricket paralysis virus RNA is regulated by two internal ribosome entry sites. *Mol Cell Biol* **20**, 4990–4999.

**Wilson, J. E., Pestova, T. V., Hellen, C. U. T. & Sarnow, P. (2000b).** Initiation from the A site of the ribosome. *Cell* **102**, 511–520.

**Woolaway, K. E., Lazaridis, K., Belsham, G. J., Carter, M. J. & Roberts, L. O. (2001).** The 5' UTR of *Rhopalosiphum padi* virus (RhPV)

contains an internal ribosome entry site (IRES) which functions efficiently in mammalian, insect and plant translation systems. *J Virol* **75**, 10244–10249.

**Yang, Q. & Sarnow, P. (1997).** Location of the internal ribosome entry site in the 5' non-coding region of the immunoglobulin heavy-chain binding protein (Bip) mRNA: evidence for specific RNA-protein interactions. *Nucleic Acids Res* **25**, 2800–2807.