



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



Review

Cis-acting RNA elements in human and animal plus-strand RNA viruses

Ying Liu, Eckard Wimmer, Aniko V. Paul*

Department of Molecular Genetics and Microbiology, Stony Brook University, Stony Brook, NY 11790, USA

ARTICLE INFO

Article history:

Received 4 June 2009

Received in revised form 9 September 2009

Accepted 13 September 2009

Available online 23 September 2009

Keywords:

Plus-strand RNA virus

RNA structure in plus-strand RNA virus

IRES element in picornavirus and other

plus-strand RNA virus

Nontranslated region in plus-strand

RNA virus

ABSTRACT

The RNA genomes of plus-strand RNA viruses have the ability to form secondary and higher-order structures that contribute to their stability and to their participation in inter- and intramolecular interactions. Those structures that are functionally important are called *cis*-acting RNA elements because their functions cannot be complemented in *trans*. They can be involved not only in RNA/RNA interactions but also in binding of viral and cellular proteins during the complex processes of translation, RNA replication and encapsidation. Most viral *cis*-acting RNA elements are located in the highly structured 5'- and 3'-nontranslated regions of the genomes but sometimes they also extend into the adjacent coding sequences. In addition, some *cis*-acting RNA elements are embedded within the coding sequences far away from the genomic ends. Although the functional importance of many of these structures has been confirmed by genetic and biochemical analyses, their precise roles are not yet fully understood. In this review we have summarized what is known about *cis*-acting RNA elements in nine families of human and animal plus-strand RNA viruses with an emphasis on the most thoroughly characterized virus families, the *Picornaviridae* and *Flaviviridae*.

© 2009 Elsevier B.V. All rights reserved.

1. Introduction

Most life forms in nature contain dsDNA as their genetic material but viruses frequently possess RNA genomes. Viruses that replicate their genomes entirely through RNA intermediates are classified into plus-strand, minus-strand and double-stranded RNA viruses. Of these three classes, the most abundant are plus-strand RNA viruses, which use their genomic RNA sequence as mRNA to directly encode proteins. Although these viruses vary widely in their virion structures (icosahedral, spherical), mechanisms of RNA translation (cap-dependent, cap-independent), RNA replication (primed or de novo initiated), and encapsidation, they are similar in the basic steps that make up their life cycles. Following binding and entry of the viral RNA into a permissive cell, the incoming mRNA is translated. The products are either the mature viral proteins or a polyprotein, which is subsequently processed into smaller precursor and mature polypeptides. These products include both structural and nonstructural components. The structural proteins are the building blocks of the progeny virion particles. The nonstructural polypeptides, including the RNA-dependent RNA polymerase, are primarily involved in RNA replication, which takes place in replication complexes associated with altered intracellular membranes. First the parental plus-strand RNA is converted into a minus-strand copy, which then serves as the template for the synthesis of progeny plus-strands. Finally, the plus-strand genomes are encapsidated and the newly assembled virion

particles are usually released from the host cells. All of these steps in the viral life cycle are strictly regulated and the RNA structures present in the genomic RNA have important roles in these processes.

The genomes of plus-strand RNA viruses are able to form secondary and higher-order RNA structures that contribute to their stability and to their participation in inter- and intramolecular interactions. The presence of RNA structures in viral genomes can be predicted by computational procedures that combine structure prediction and nucleotide sequence comparison. Subsequently, the biological importance of these predicted structures can be tested by genetic analyses.

Those structures that are functionally important are called *cis*-acting RNA elements because their functions cannot be complemented in *trans*. They can be involved not only in RNA/RNA interactions but also in binding viral and cellular proteins required for the complex processes of the viral life cycle.

Most viral *cis*-acting RNA elements are located in the 5'- and 3'-nontranslated regions (NTRs) of the genomes but now evidence is accumulating for the widespread existence of functional structures embedded within protein-coding sequences. Some of these internal RNA elements comprise individual functional units far away from the genomic ends. Others are simply extensions into the down-stream or up-stream coding sequences of functional RNA elements that are already present in the NTRs. In this review we will attempt to summarize what is known about *cis*-acting RNA elements in nine families of human and animal plus-strand RNA viruses with an emphasis on the most thoroughly characterized virus families, the *Picornaviridae* and *Flaviviridae*. Unfortunately, due to the limited scope of this review article, it is not possible to include a discussion of all of the RNA structures present in all of the viruses that make

* Corresponding author. Tel.: +1 631 632 9777; fax: +1 631 632 8891.

E-mail address: apaul@notes.cc.sunysb.edu (A.V. Paul).

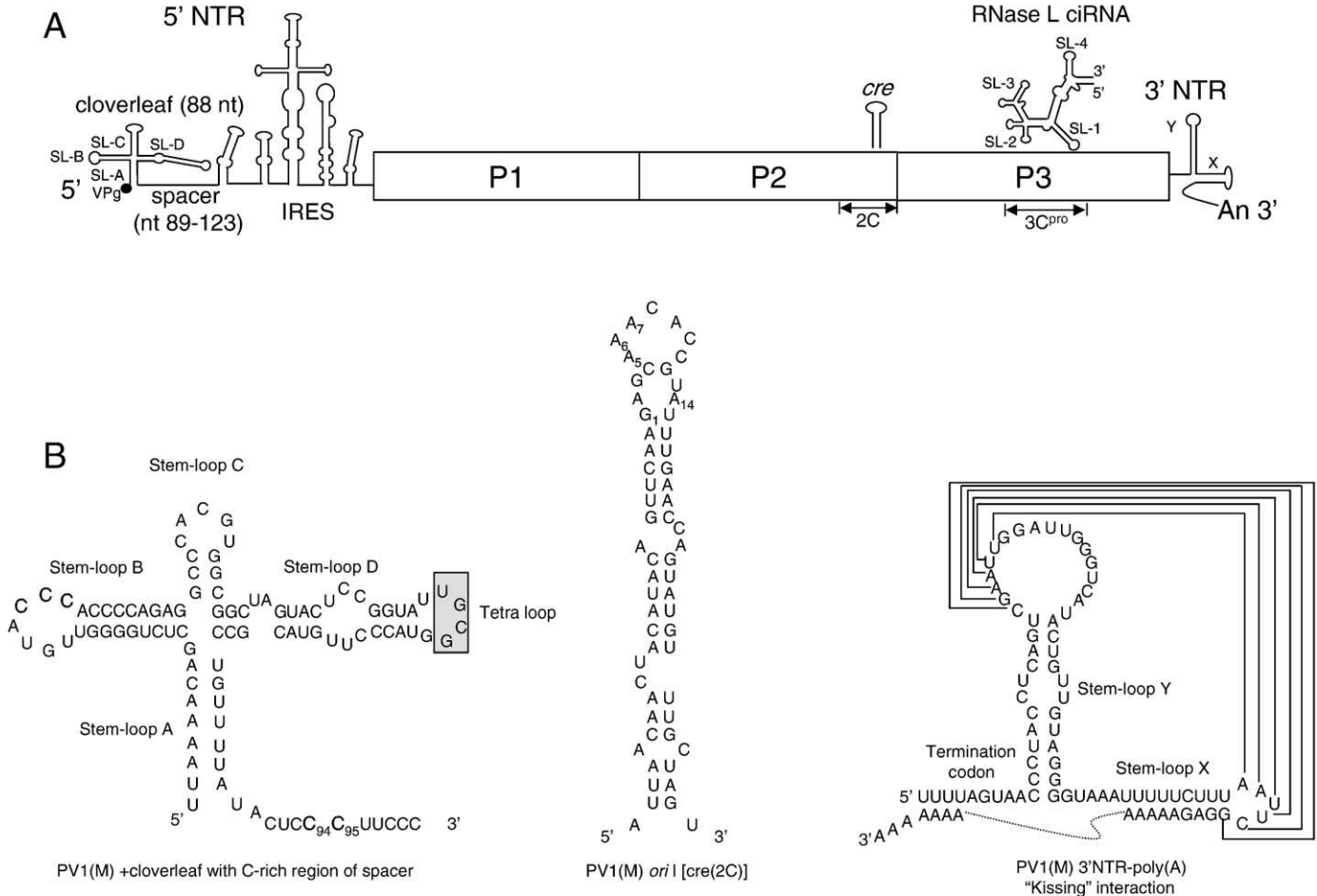
up these virus families. However, we hope that the representative members of each virus family that we have selected for our review will provide the reader with an overview of the nature and functions of the *cis*-acting RNA elements in the life cycle of these viruses.

1.1. Picornaviridae

Members of the *Picornaviridae* virus family include a large number of medically important human and animal pathogens, which cause a wide variety of illnesses. The plus-strand genomes of picornaviruses vary in length from 7209 to 8450 nt and encode a single polyprotein

that consists of one structural (P1) and two nonstructural domains (P2 and P3) (Figs. 1–4). The genomic RNA is covalently linked at the 5'-end to the hydroxyl group of tyrosine in the terminal peptide VPg [1–3] and the 3'-end is polyadenylated [4,5]. There are 12 genera in the *Picornaviridae* family: *Enterovirus*, *Cardiovirus*, *Aphthovirus*, *Hepatovirus*, *Parechovirus*, *Kobuvirus*, *Erbovirus*, *Teschovirus*, *Sapelovirus*, *Senecavirus*, *Tremovirus*, and *Avihepatovirus* ([6], ICTV 2009). It should be noted that recently the genus *Rhinovirus* was incorporated into the *Enterovirus* genus (www.picornavirus.com). For this review we have divided picornaviruses into three major groups based on similarity in their 5' NTRs. In addition, we have briefly discussed the RNA

Enterovirus: Poliovirus (PV)



Rhinovirus: Rhinovirus 14

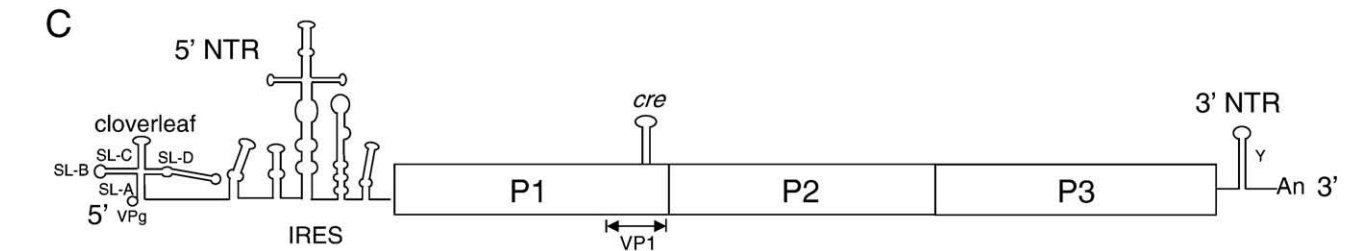


Fig. 1. *Cis*-acting RNA elements in the genomes of PV and of HRV14 (*Enterovirus*, *Picornaviridae*). (A) PV. The 5' NTR consists of the CL and the IRES. The ORF contains three domains, the structural (P1), and the nonstructural domains (P2 and P3). The *cre* is located in the coding sequence of protein 2C^{ATPase}. The ciRNA, inhibitor of RNase L, is located in the 3C^{pro} coding sequence. The 3' NTR contains two stem loops and is followed by a poly(A) tail. (B) Enlarged images of the PV CL, the c-rich region of the spacer (left), the *cre*(2C) hairpin (middle) and the 3' NTR-poly(A) with a "kissing interaction" between stem loops X and Y. The PCBP2 binding site on SL-A of the CL is shown with bold letters and the tetra loop in SL-D is boxed. (C) HRV14. The 5' NTR is similar to the genome of PV but the *cre* RNA is located in VP1 and the 3' NTR has only one stem loop.

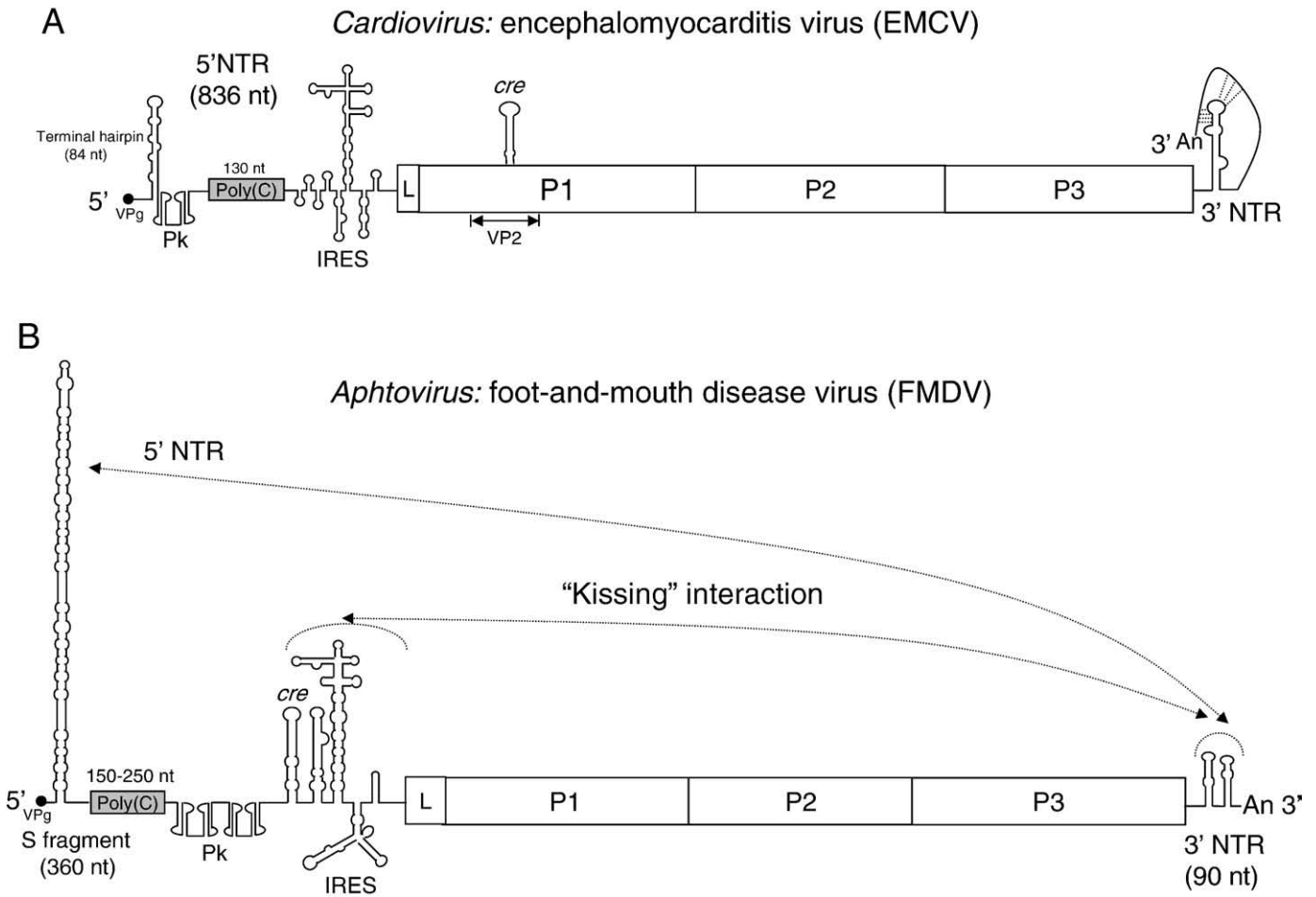


Fig. 2. Cis-acting RNA elements in the genomes of EMCV (*Cardiovirus, Picornaviridae*) and FMDV (*Aphthovirus, Picornaviridae*). (A) EMCV. The 5' NTR contains a large hairpin and two pseudoknots followed by a poly(C) tract and the IRES. The *cre* element is located in the coding region of capsid protein VP2. The ORF contains a leader (L), one structural (P1), and two nonstructural domains (P2 and P3). The one stem loop of the 3' NTR interacts with the poly(A) tail. (B) FMDV. The 5' NTR contains the S fragment, a poly(C) tract, three to four pseudoknots, the *cre* element, and the IRES. The 3' NTR has two stem loops, which interact both with the S fragment and the IRES.

structures present in the genomes of three of the less well-known genera, *Kobuvirus*, *Parechovirus* and *Teschovirus*.

1.1.1. Enteroviruses

Based on RNA folding analyses and sequence comparisons Witwer et al. [7] have predicted the presence of numerous conserved RNA structures in the 5' NTRs and 3' NTRs of enterovirus genomes, including those of human rhinoviruses (HRVs) (Fig. 1A–C). In addition, internal conserved structures were identified in their 3D^{pol} domains, in the 2C^{ATPase} coding sequence of poliovirus (PV) and VP1 or 2A sequences of HRVs.

1.1.1.1. 5' NTR (*oriL*). The 5' noncoding region or *oriL* (origin of replication at the left) of enteroviruses is highly structured and contains two functional domains, which are involved in translation and RNA replication. The first domain in enteroviruses forms a cloverleaf (CL) structure [8] that carries signals to control both translation and RNA replication (Fig. 1A, B) [9]. The PV CL (88 nt) contains four stem loops (SL-A to SL-D), which is followed by a spacer region (nt 89–123) located between the cloverleaf and the IRES (Fig. 1A). The second domain of the enteroviral 5' NTR is the internal ribosomal entry site (IRES) that promotes translation.

Hepatovirus: Hepatitis A virus (HAV)

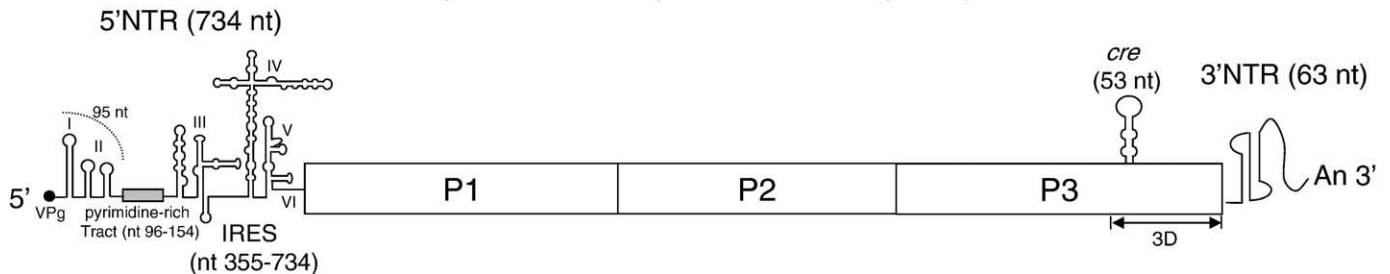


Fig. 3. Cis-acting RNA elements in the genome of HAV (*Hepatovirus, Picornaviridae*). The genome of HAV contains a long 5' NTR, which consists of three stem loops, a pyrimidine-rich tract, and IRES. The ORF contains one structural (P1) and two nonstructural domains (P2, P3). The *cre* element is located in the 3D^{pol}-coding region. The 3' NTR contains a pseudoknot and a poly(A) tail.

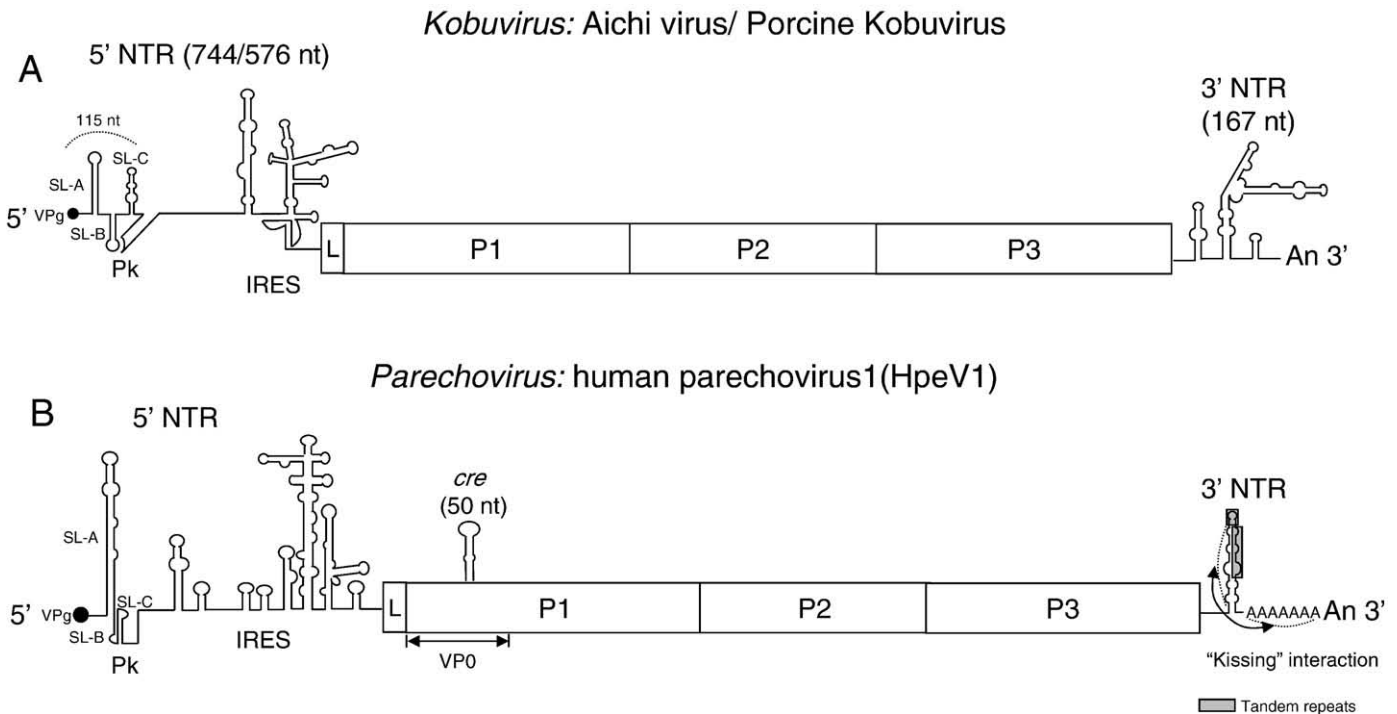


Fig. 4. Cis-acting RNA elements in the genomes of Aichi virus and Porcine kobuvirus (*Kobuvirus*, *Picornaviridae*) and of HpeV1 (*Parechovirus*, *Picornaviridae*). (A) Aichi virus (human kobuvirus) and Porcine kobuvirus. The 5' NTR of Aichi virus consists of three stem loops (SL-A, B, C) and of a pseudoknot (Pk). At the 3'-terminal part of the 5' NTR the putative IRES structure of Porcine kobuvirus is shown. The kobuvirus ORF contains a leader sequence (L), one structural (P1), and two nonstructural domains (P2, P3). The 3' NTR (167 nt) of Porcine kobuvirus is predicted to form three hairpin structures (G. Reuter, unpublished data. Structure is shown with permission of G. Reuter). The genome is terminated with a poly (A) tail. (B) HpeV1. The 5' NTR contains three stem loops (SL-A, B, C) and a pseudoknot, and the IRES. The ORF has a leader sequence, one structural (P1), and 2 nonstructural domains (P2, P3). The *cre* element is located in the VP0 coding sequence. The 3' NTR consists of a stem loop, which is involved in a "kissing interaction" with the poly(A) tail.

Numerous studies have shown that the formation of a ribonucleoprotein (RNP) complex, consisting of cloverleaf RNA, cellular protein PCBP and viral protein 3CD^{pro}, is required for viral RNA replication and virus proliferation [10–14]. Protein PCBP2, with its KH1 domain, was shown to bind to stem loop B of the cloverleaf [13,15,16]. The primary PCBP2 binding sites on the CL are located on 3C residues in the loop of SL-B (nt 23–25) and on two Cs (C₉₄ and C₉₅) in the spacer region [15,17]. The other partner in the formation of the CL RNP complex is proteinase and RNA binding protein 3CD^{pro} for PV [12,14,18] or 3C^{pro} for rhinoviruses [19]. PV proteins 3C^{pro}/3CD^{pro} interact with a tetra loop (UGCG) in SL-D of PV (Fig. 1B) but do not bind to the HRV14 CL, which contains only 3 nt (UAU) [20]. Interestingly, the formation of another RNP complex that included CL RNA, and viral proteins 3AB and 3CD^{pro}, was also shown to be important for PV RNA replication [12,14]. Herold and Andino [21] have proposed a model in which the PV genome circularizes prior to minus-strand RNA synthesis by way of an interaction between cellular poly(A) binding protein with the poly(A) tail of PV RNA on the one hand and PCBP2/3CD^{pro}/CL on the other. This model was based on an observation that PABP physically interacts with PCBP2, 3CD^{pro} and the poly(A) tail both *in vivo* and *in vitro*.

1.1.1.2. Minus-strand 3'-CL. The sequence complementary to the plus-strand CL also forms a similar structure. Using an *in vitro* translation/RNA replication system, Sharma et al. [22] have shown that the duplex structure of SL-A in the PV CL (Fig. 1B) was required for minus-strand but not for plus-strand RNA synthesis. On the other hand altering the primary sequence at the 5'-terminal end of SL-A resulted in a striking reduction of plus-strand RNA synthesis. These results suggested that the 3'-terminal end of minus-strands is required for the initiation of plus-strand RNA synthesis. The specific binding of cellular proteins (p36 and p38) and of viral protein 2C^{ATPase} to the 3'-end of minus-strand PV RNA has also been reported [23–25].

1.1.1.3. 3' NTR-poly(A) (*oriR*). The highly structured heteropolymeric regions of picornaviral 3' NTRs are very diverse and their functions are unknown although genetic evidence supports a role in RNA replication [26,27]. The enteroviral and rhinoviral 3' NTRs can be grouped into three types: single stem loop (Y) in rhinoviruses; two stem loops (X and Y) (65 nt) in PVs; 3 stem loops (X, Y and Z) (100 nt) in coxsackieviruses [28]. In PV the poly(A) tail is a part of the overall 3'-terminal structure with poly(A) hybridized to sequences in stem X, just upstream from the termination codon (Fig. 1B). The predicted structures were confirmed by enzymatic and chemical probing [28,29]. It was proposed that the enteroviral 3' NTRs together with the poly(A) tail form the origins of replication *oriRs* for minus-strand RNA synthesis, which serve as specific binding sites for viral and/or cellular proteins.

Mutational analyses of the PV and coxsackie B3 virus (CVB3) 3' NTRs have provided convincing evidence for a "kissing" interaction between 5 nt in the loops of stem loops X and Y, which is functionally important (Fig. 1B) [29,30]. Three-dimensional structures, derived by molecular modeling, displayed the PV and CVB3 *oriRs* as quasi-globular multi-domain structures [29,30].

Unexpectedly, it was observed that the PV 3' NTR (2 stem loops) could be replaced by the 3' NTRs of HRV14 (1 stem loop) or of coxsackievirus B4 (CVB4) (3 stem loops) [23]. Even more surprising was the finding that the PV or HRV14 3' NTRs could be deleted without loss of viability in HeLa cells [31,32]. Subsequently it was shown that the removal of the PV 3' NTR restricted plus-strand RNA synthesis, particularly in cells of neuronal origin [33].

The poly(A) tail of PV [5] and of all other enteroviruses (and of picornaviruses whose genomes have been analyzed) is about 90 nt long and it is genetically encoded [4]. Interestingly, the complementary poly(U) is only 20 nt [34]. These results are in agreement with previous studies showing that a poly(A) tail of 20 nt is sufficient for the replication of PV [35] and for the binding of poly(A) binding

protein to the 3' NTR-poly(A) *in vitro* [21]. The poly(A) tail of CVB3 is markedly shortened in the absence of the OriR, indicating a role for this *cis*-replicating element in poly(A) synthesis [34].

1.1.1.4. Internal RNA elements. Among the best-characterized functional RNA structures are the picornavirus *cre*s (*cis*-replicating RNA elements), which are small hairpins of varying nucleotide sequence located in the viral ORFs [36,37], and in one case so far reported, in the 5' NTR of FMDV [38]. This type of *cre* was first discovered by McKnight and Lemon [39] in the VP1 coding sequence of HRV14 and was shown to be required for RNA replication [40]. Since then similar *cre*s have been identified in the 2C^{ATPase} coding domains of PV (Fig. 1B) [41] and CVB3 [42]. In rhinoviruses the corresponding *cre* elements are found in the 2A ORF of species A rhinoviruses [43], the VP1 ORF of species B rhinoviruses (Fig. 1C) [39], the VP2 ORF of species C rhinoviruses [44]. The precise function of the *cre* (*oril*) is to template the linkage of two UMPs to VPg, to yield VPgpU and VPgpUpU, the primers for both plus and minus-strand RNA synthesis [45–47]. According to current models of PV RNA replication the VPgpUpU made on the *oril* is used for the initiation of both plus and minus-strand RNA synthesis [34,37].

The ability of the enteroviral *cre* elements to support RNA replication is dependent both upon the specific RNA structure, in particular the upper of the part of the stem, and specific sequences within the loop [46,48]. The *cre*s of enteroviruses share a conserved nucleotide sequence motif R₁XXXA₅A₆R₇XXXXXXR₁₄ (R = purine) in the loop of the hairpin (Fig. 1B) [47,48]. Within this motif, the A₅ residue templates the linkage of both UMPs to VPg by a “slide back mechanism” in a reaction catalyzed by RNA polymerase 3D^{pol} and stimulated by viral proteinase 3CD^{pro} [37,45–47,49]. Based on biochemical studies Pathak et al. [50] proposed a model for the assembly of the PV VPg uridylylation RNP complex. According to this model the first step is the binding of a 3CD^{pro} dimer to the upper part of the *oril* stem, which is unwound. In the next step 3D^{pol} associates with the complex via an interaction between the back of the thumb subdomain of 3D^{pol} and a convex surface formed by the top of the subunits of the 3C^{pro} dimer [51].

The function of the enteroviral *cre* is independent of its position within the genome [40,48]. When the endogenous *cre* element of PV is inactivated by mutation, *cre* function can be rescued by the insertion of a second *cre* (PV or HRV14) into the 5' NTR [40,48]. Interestingly, *in vivo* studies indicated that the PV *oril* inhibited PV replication in a *trans*-dominant manner [52].

Recently a second type of internal RNA element was discovered in the coding sequence of protein 3C^{pro} of PV and of several coxsackie A viruses (CAVs) (Fig. 1) [53]. This RNA element potently inhibits the activity of a cellular protein, RNase L, a latent endoribonuclease in an interferon-regulated, dsRNA-activated pathway. The RNA structure consists of 4 stem loops and of these stem loops 1 and 4 is important for function. These are involved in a putative “kissing loop” interaction [54]. Interestingly, RNase L activity was activated late during the course of PV replication in HeLa cells as virus assembly neared completion. This activity did not diminish virus production; rather, it was associated with larger plaques and increased cell-to-cell spread. Han et al. [53] suggested the possibility that RNase L, which is activated late in infection and is proapoptotic, facilitates cythopathic effect and virus release at the end of the PV growth cycle. We have included this RNA structure in our discussion although currently its exact function in the PV life cycle is not known. It is possible, however, that a biological function for it will be discovered in the future but only under certain conditions, such as in different cell lines.

1.1.2. Cardio- and aptoviruses

1.1.2.1. 5' NTR (*oril*). The cardio- and aptoviral 5' NTRs differ from those of enteroviruses in two respects. First, their 5'-terminal RNA elements have a less-defined structure than the corresponding

enteroviral CL [55]. Second, the 5' NTRs contain a long poly(C) tract of about 130 nt in encephalomyocarditis virus (EMCV) (Fig. 2A) and 150–250 nt in foot-and-mouth disease virus (FMDV) (Fig. 2B). EMCV genomic RNA contains a 5'-terminal hairpin (84 nt), linked to two pseudoknots (Pk, Fig. 2A) [56], which are connected to the poly(C) tract, followed by the IRES. In FMDV the 5' NTR (>1300 nt) begins with a very large (360-bp) stem loop structure (S fragment) [7], then the poly(C) tract, three to four tandemly repeated pseudoknots of unknown function [57], the *cre* element [38] and finally the IRES (Fig. 2B). The S fragment is predicted to function in RNA replication although direct evidence for this is lacking. Interestingly, Serrano et al. [58] have identified an interaction *in vitro* between the 3' NTR-poly(A) tail and the S fragment (Fig. 2B).

1.1.2.2. 3' NTR-poly(A) (*oriR*). The 3' NTR of FMDV is 90 nt long and consists of two stem loops, linked to a genetically encoded poly(A) tail (Fig. 2B). In contrast to PV or HRV14 the 3' NTR of FMDV could not be deleted without loss of virus replication [59]. Recent studies indicated that each of the stem loops of the FMDV 3' NTR interacted *in vitro* with the S fragment of the 5' NTR and this interaction was dependent on the presence of the poly(A) tail (Fig. 2B) [58]. Similarly, the 3' NTR was also found to establish a long range RNA/RNA interaction with the IRES.

Duque and Palmenberg [60] have identified three phylogenetically conserved stem loops in mengovirus (cardiovirus) 3' NTR (126 nt). Deletion of stem loop I was found to be dispensable for virus growth while the deletion of stem loop III was lethal. The deletion of stem loop II led to an intermediate growth phenotype. In other studies the 3' NTR of EMCV was predicted to form a single stem loop [61]. Mutational and deletion analyses studies indicated that a U-rich stretch in the loop interacts with the poly(A) tail (Fig. 2A).

1.1.2.3. Internal RNA element (*oril*). In cardioviruses, Theiler's virus and mengovirus, *oril*s were identified in the coding sequence of capsid protein VP2 [62], and these were found to be essential for RNA replication. These stem loop structures had a conserved sequence of 9 nt in mostly unpaired regions of the structure that also contained the AAACA sequence, characteristic of the enterovirus *cre*s. The *cre*s of TMEV and mengovirus could be functionally exchanged but they could not be replaced with the *cre* of HRV14.

Of all the picornaviruses analyzed so far, only FMDV contains a *cre* element outside its protein coding sequence, just upstream of the IRES (Fig. 2B) [38]. Mutations in the conserved AAACA sequence severely reduced RNA replication and yielded quasi-infectious viruses. The structure of the stem was also important for replication. Interestingly, and unlike other picornaviral *cre*s, the FMDV *cre* could be complemented in *trans* in infected cells [63].

1.1.2.4. Genome-scale ordered RNA structures (GORS). Early computational studies by Witwer et al. [7] have predicted the presence of a large number of conserved RNA structures in the genome of FMDV, both in the 5' and 3' NTRs and in the coding sequences. Recently, large-scale thermodynamic prediction methods have detected a genome-scale ordered RNA structure in FMDV [64,65]. Interestingly, in FMDV, like in other mammalian plus-strand RNA viruses, these GORS are associated with a persistent growth phenotype. Although the reason for the association between GORS and persistence is not yet understood, Simmonds et al. [66] suggested the possibility that the formation of an extensive RNA a secondary structure might play a role in the evasion of cell defenses.

1.1.3. Hepatovirus

Hepatitis A virus, a member of the hepatovirus genus, differs from other picornaviruses in the details of the organization and function of its polyprotein. As shown below, however, it also shares many features with other picornaviruses in the overall genome organization and replication strategy (Fig. 3).

1.1.3.1. 5' NTR (*OriL*). Structural predictions and enzymatic probing have shown that the 5' NTR of hepatitis virus A has several motifs common with aphto- and cardioviruses [67]. The 5' NTR contains 6 major structural domains. Domains I and II (1–95 nt) contain a 5'-terminal hairpin and two small stem loops (Fig. 3). This is followed by a pyrimidine-rich tract (96–154 nt). The remainder of

the 5' NTR, domains III to VI (nt 155–734), contains several stem loops and includes the IRES (starting nt 355). The interaction of cellular protein glyceraldehyde-3-dehydrogenase (GAPDH) with the 5' NTR was proposed to modulate protein translation [68]. The binding of viral protein precursor 3ABC in vitro to the 5' NTR was also observed [69].

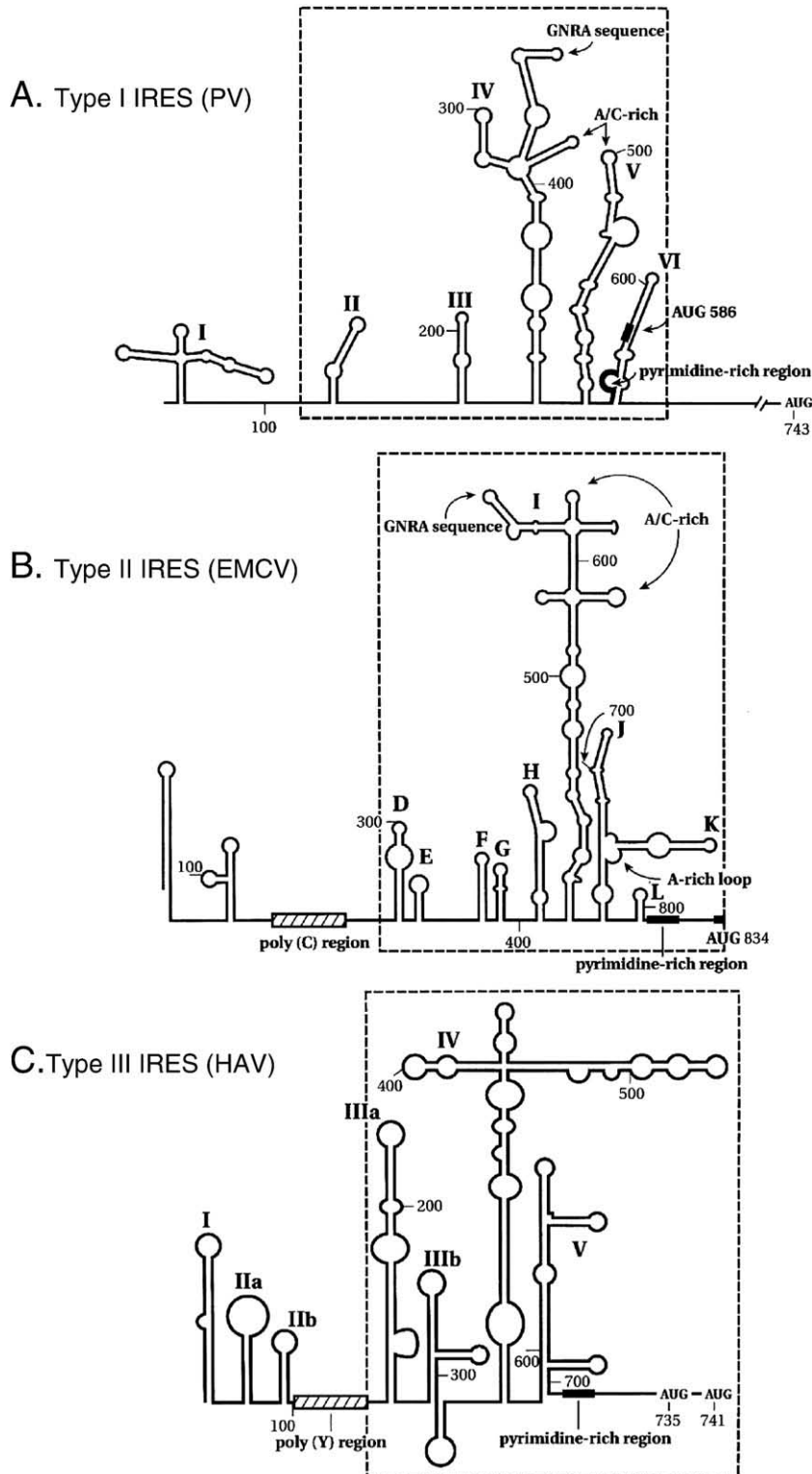


Fig. 5. IRES elements of Picornaviruses. (A–C) The figure shows the type I IRES element of PV, the type II IRES of EMCV, and the type III IRES of HAV. (D–E) Type IV IRESes of Porcine kobuvirus and Porcine teschovirus, respectively. A–C are taken from Ehrenfeld et al., [248], with permission of the publisher. D is taken from Reuter et al., [72], with permission of the publisher. Figure E is taken from Chard et al., [91], with permission of the publisher.

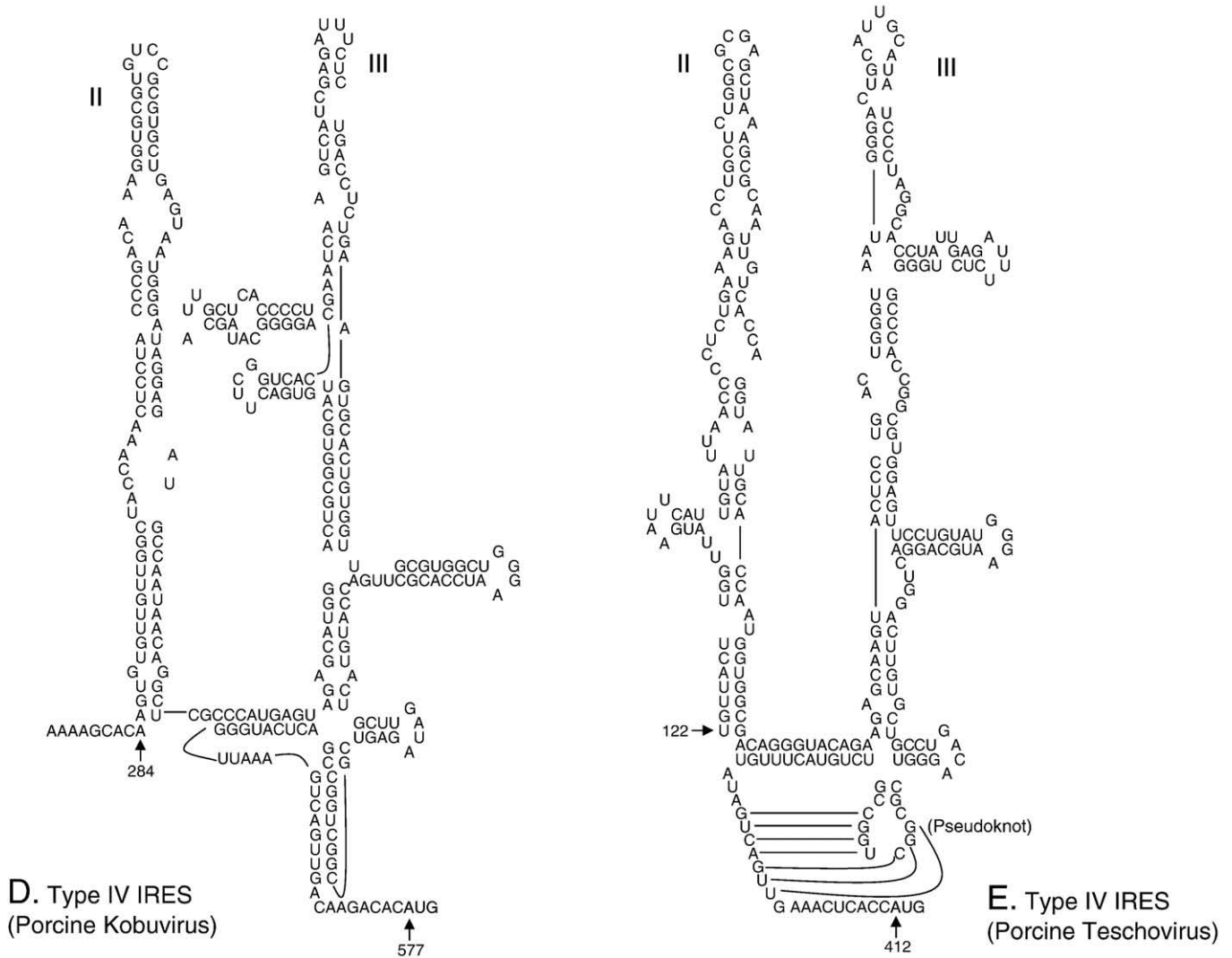


Fig. 5 (continued).

1.1.3.2. 3' NTR (OriR). The 3' NTR of hepatitis A virus is 63 nt long and terminates with a poly(A) tract of 40–80 nt. There are two alternate models for the structure of the 3' NTR. The first consists of two stem loops (X and Y) and a putative “kissing” interaction between the stem loops has been predicted [70]. The second model suggests that the 3' NTR exists in a pseudoknot-like structure and this model was favored by enzymatic probing experiments (Fig. 3). The specific binding of GADPH to the 3' NTR and to a part of the upstream 3D^{pol} coding region was also reported [70].

1.1.3.3. Internal RNA element (OriI). The *cre* element of hepatitis A virus, recently identified, differs from those of other picornaviruses in location, size and stability [71]. It is located near the 5' end of the 3D^{pol} coding sequence and both the stem (35 bp), which is rich in A-U pairs, and the loop (18 nt) are large. Like other picornaviral *cre*s the HAV *cre* also contains an AAACA/G motif and mutations that disrupted the structure of this element ablated the replication of a subgenomic replicon.

1.1.4. Kobu-, Parecho-, and Teschoviruses

Compared to other picornaviruses, relatively little is known about the *Kobovirus* (Aichi virus and Porcine kobovirus), *Parechovirus* (Human parechovirus [HpeV1], Ljungan virus) and *Teschovirus* (Porcine teschovirus) genera of *Picornaviridae*, which have been

recently recognized as important human pathogens. The VPg-linked genomes of Aichi virus and of Porcine kobovirus are 8280 and 8210 nt long, respectively [72]. They encode a single polyprotein that contains a leader protein followed by the capsid proteins and the nonstructural proteins (Fig. 4A). The 3' end of the genomes is terminated with a poly (A) tail.

The 5' NTRs of Aichi virus (744 nt) and of HPeV resemble that of cardio- and aptoviruses (Fig. 4A, B). In Aichi virus the first 115 nt of the 5' NTR are predicted to form three stem loops (SL-A to SL-C) and a pseudoknot [73]. Disruption of these RNA elements impaired minus-strand RNA synthesis while the stem of SL-A was found to be important for plus-strand RNA synthesis. Nagashima et al. [74] have observed a specific binding of viral protein 3ABC to the 5' NTR of Aichi virus and this binding was required for minus-strand RNA synthesis. Recent studies by Sasaki and Taniguchi [75] indicated that a 7-nt long segment (UCCCACU and its complementary sequence) in the stem of SL-A of Aichi virus has an important role in encapsidation. The 5' NTR of Porcine kobovirus (S-1-HUN) is 576-nt long and is predicted to form seven stem loops [72]. The first three of these stem loops are similar to those of Aichi virus. A polypyrimidine tract is located between nt 459 and 465, within a predicted IRES element (nt 284–576), which has a strong sequence similarity to that of Porcine teschovirus and belongs to the type IV hepaci/pestivirus IRES group (Figs. 4A and 5E). The 3' NTR of Porcine kobovirus is 167 nucleotides

long and is predicted to form three stem loops (Fig. 4A) [72]. Using bioinformatics analysis methods Simmonds et al. [65] have predicted the presence of genome-scale ordered RNA structures in the genomes of kobuviruses.

In HpeV1 the 5' NTR also contains the same functional stem loops (SL-A to SL-C) as in Aichi virus (Fig. 4B) [76]. The 3' NTR of HpeV1 is made up of two highly conserved tandem repeats, which is predicted to form a single stem loop structure with extensive base-pairing between the poly(A) tail and a U-rich region of the 3' NTR [76,77]. The biological function of this 3'-terminal structure is not yet known. The *oril* of HpeV1 has been identified in the coding sequence of VP0 as a hairpin structure of about 50 nt containing a large loop (23 nt). The loop contains the CAAAC motif, typical of picornaviruses [77]. In this same study the *cre* element of Ljungan virus was predicted to exist in the VPg coding region.

Computer analysis has predicted five highly conserved RNA structures in the 5' NTR and six such structures in the noncoding regions of Teschoviruses [7]. The 5' NTR of Porcine teschovirus is 412 nt long and contains an HCV-like IRES element (nt 122–411) (Fig. 5D) [78].

1.1.5. Picornavirus IRES elements

The discovery of IRESes in picornaviruses has fundamentally changed the perception of gene expression by translation in eukaryotic cells. Since until 1988 a single mechanism of cap-dependent initiation of translation was accepted almost as dogma, the experiments illustrating IRES-mediated translation (translation of artificially engineered dicistronic mRNAs) [79–82] was received with skepticism. Two experiments can be cited that lent unambiguous support for IRES function. The first provided proof *in vivo* through the construction and propagation of a dicistronic poliovirus whose translation of the segmented polyprotein was controlled by two different IRES elements and whose infection of cells followed single hit kinetics [83]. The second experiment described the construction of a circular mRNA containing a picornavirus IRES element. This circular mRNA was efficiently translated *in vitro* [84]. In the following section, the structural and functional hallmarks of picornavirus IRESes will be summarized from several recent general reviews [85–87].

1.1.5.1. Location of IRES elements in picornavirus genomes. Picornavirus IRESes that all reside in the 5' NTR, are 290–450 nt long. They are preceded by RNA structures of varying length, the nature of the structures depending upon the picornavirus genus (Fig. 5). In the case of the cardio- and aphthoviruses, IRESes are preceded by long poly(C) tracts. In the case of the aphthoviruses the critical *cre* element maps between the poly(C) and the IRES (Fig. 2A, B). In hepatitis A virus, the polypyrimidine tract consists of C and U residues (Fig. 3).

1.1.5.2. Picornavirus IRESes types I–III. Based on the primary sequence as determinants of structures, the IRESes of enteroviruses, cardio- and aphthoviruses, and hepatitis A virus have been classified as type I, type II, and type III IRESes, respectively [88] (Fig. 5A–C). The nucleotide sequence homology between the three IRES types is barely 50%. In spite of the general sequence differences, all three IRES types fold roughly into domains dominated by a large domain in the center (Fig. 5). In addition, representatives of type I and type II IRESes carry tetra loops (GNRA or GNAA; N, any nucleotide; R, purine) in the loops of their domains (Fig. 5A, B). The integrity of the tetra loops is required for IRES function.

1.1.5.3. A novel, unexpected picornavirus IRES, type IV. Sequence analyses of Porcine teschoviruses, which are members of the new picornavirus genus *Teschovirus*, have revealed that their highly conserved 5' NTRs exhibited no apparent similarity with picornavirus 5' NTRs known at the time [89]. Closer analyses of the 5' NTR RNA of Porcine teschovirus-1 (PTV-1) led to the unexpected discovery that in

structure and function the IRES of this virus is more related to the IRES of hepatitis C virus (HCV) than to the types I–III IRESes discussed above (Fig. 5E) [90]. As can be seen, the PTV-1 IRES (Fig. 5E) consists only of two major stem loop domains (domains II and III), forming a pseudoknot at the base of domain III, and it is only 290-nt long (between the base of domain II and the initiating AUG at nt 412) [91]. The small size, the architecture and genetic properties resemble the IRES of HCV. Indeed, it now appears that HCV-related IRES elements are found in genomes of many genera of picornaviruses [78]. The HCV-like IRES elements of picornaviruses are designated as type IV IRESes. For a more detailed discussion of the HCV IRES, see the section on *Flaviviridae*. A recent report by Reuter et al. [72] shows the putative IRES element (nt 284–576) of Porcine kobuvirus (Fig. 5D), which has a 74% sequence similarity to the IRES of Porcine teschovirus (Fig. 5E). The predicted structure of this IRES element (nt 122–411) is also that of type IV IRESes, similar to those of hepac- and pestiviruses.

1.1.5.4. The YnXmAUG motif. Types I–III IRES elements contain a highly conserved motif consisting of an oligo pyrimidine tract (Yn), followed by a tract of an unspecified sequence of 15–20 nucleotides (Xm), followed by an AUG [92]. The AUG codon in type II IRES elements is the initiating codon (in aphthoviruses, a downstream codon may also be used) whereas in type I IRESes, the motif is upstream of the initiating codon and its AUG codon is, therefore, cryptic. Genetic analyses have shown that the integrity of the motif is essential for IRES function. Eliminating Yn or shortening or enlarging Xm, etc., have resulted in severe deficiencies in translation of PV RNA; the same is true for the EMCV type II IRES [93–96]. It has been speculated that the YnXmAUG motif may constitute the “landing pad” for the small ribosomal subunit but direct evidence for this hypothesis is still lacking. In contrast to enteroviruses and EMCV, Pilipenko et al [97] reported that the YnXmAUG motif in Theiler's virus, another species of the genus *Cardiovirus*, is dispensable for translation *in vitro* and for replication of the virus in tissue culture [97]. These authors found, however, that elimination of the motif strongly influenced the pathogenesis of Theiler's virus in mice. Meanwhile, available evidence has clearly indicated that IRES elements do not function by the same mechanism in recruiting the ribosomal subunits, followed by selecting the initiating AUG codon. Indeed, the HCV-related type IV IRESes of picornaviruses, discussed above, lack the YnXmAUG motif altogether.

1.1.5.5. Exchange of IRES elements amongst different picornaviruses. As has been pointed out before, the nt sequences and apparent higher-order structures of the four picornavirus IRES types vary widely. On the other hand, the overall function of IRES elements – internal initiation of translation – is the same, regardless of the underlying mechanism. Would an exchange of these different IRES elements between different picornaviruses yield viable viruses? The answer is yes. This was surprising since it was originally expected that these genetic elements that are huge in relation to the small viral genomes would contribute to functions other than translational control. Many IRES chimeras were analyzed using PV as the backbone. They included the exchange of the PV IRES with other type I IRESes such as CVB3 [98], HRV2 and HRV14 [99], with type II IRESes such as EMCV [100], and even with the IRES of HCV [101], a virus belonging to a different family. In HeLa cells at 37 °C, these chimeric viruses replicated with wt kinetics, an observation suggesting that surrounding sequences did not influence the function of the IRES elements. Chimeras, however, may express interesting phenotypes, for example host range phenotypes, when assayed under different conditions. For example, PV carrying the HRV2 IRES [“PV(RIPO)”] is highly attenuated in cells of neuronal origin [99,102].

These studies allow two conclusions. First, the IRESes of poliovirus, HRV, CVB3, EMCV, or HCV carry no *essential* signals necessary for PV genome replication. Second, IRESes are defined solely on the basis of

function rather than on the basis of a specific structure. This notwithstanding, Le and Maizel [103] have suggested that picornavirus IRESes types I–III have evolved from a common structural core.

1.2. Flaviviridae

The family *Flaviviridae* includes the genera of *Hepacivirus* [HCV genotypes 1–7], *Pestivirus* [bovine viral diarrhoea virus (BVDV) and classical swine fever virus (CSFV)] and *Flavivirus* [dengue virus (DENV), Japanese Encephalitis virus (JEV), Kunjin virus (KUNV), tick-borne encephalitis virus (TBEV), yellow fever virus (YFV), West Nile virus (WNV)]. The members of this family are enveloped virions composed of a lipid bilayer with two or more species of envelope glycoproteins surrounding a nucleocapsid. The nucleocapsid contains the plus-strand RNA genome complexed with multiple copies of a small capsid protein. The RNA genomes (9–11 kb) of *Flavivirus* are capped at the 5'-end, that of the *Hepacivirus* and *Pestivirus* genera are terminated with p(pp)pGCC.... Translation of members of the *Hepacivirus* and *Pestivirus* is controlled by IRES elements; the first of these was discovered for HCV by Tsukiyama-Kohara et al. [104]. All genomes of the *Flaviviridae* encode an open reading frame (ORF) specifying a polyprotein with structural (SP) and nonstructural (NSP) domains (Fig. 6A). The polyprotein is processed by cellular and viral proteinases. The HCV genome contains, in addition, a small extra-frame ORF in the 5' end of the coding region. The viral RNA lacks a 3'-terminal poly(A).

1.2.1. Hepacivirus

Hepatitis C virus (HCV) (genotypes 1–7) is the only member of the *Hepacivirus* genus. A comprehensive computational survey of conserved structure motifs has predicted a large number of structural elements both in the nontranslated regions and the ORF of the HCV genome [105].

1.2.1.1. 5'-Terminal elements including the HCV IRES. The 5' NTR of HCV folds into a complex structure containing four distinct domains (domains I–IV) [106] of which domains II–IV belong to the IRES (Fig. 5A), altogether 341 nt long. Using chimeric subgenomic replicons of HCV, Kim et al. [106] have demonstrated that the first 40 nt of HCV RNA (including domain I) as well as domain II of the IRES are essential for HCV genome replication whereas the other two IRES domains III and IV aid in, but are not essential for, efficient RNA replication. In other studies Friebe et al. [107] demonstrated that although the first 125 nt of the 5' NTR are sufficient for RNA replication the efficiency could be greatly increased by the presence of the complete HCV 5' NTR. A short and highly conserved sequence mapping next to domain I binds the liver-specific micro-RNA, miR122, an interaction that appears to be required for efficient HCV RNA replication [108].

Numerous mutational studies of the HCV IRES [109,110] and analyses by NMR and X-ray crystallography [111–113] have led to its apparent secondary structure (Fig. 6A) and subsequently to higher-order structures in the absence or presence of the ribosomal subunit [114].

Domain II is organized into basal domain IIa and apical domain IIb of which IIa appears to provide the proper configuration for the apical domain to interact with the ribosomal subunit. Domains III and IV, in turn, provide the platform for binding the very limited menu of canonical translational factors plus the ribosomal subunit. In contrast to the picornavirus type I and II IRESes, none of the canonical translation factors eIF4A, eIF4B and eIF4G are required for HCV function [85,114].

Note that the secondary structure of domain IIa of the IRES was solved by genetic analysis in 2001 [109] and it was subsequently confirmed in studies of the entire domain II by NMR [112]. Surprisingly, nearly all research publications up to the time of writing this review, and also most review articles, reproduce domain IIa with the

wrong, pre-2001 secondary structure. The downstream boundary of the HCV IRES extends unexpectedly into the core-coding region [101,115–117]. In the HCV core-coding region, adjacent to the 5' NTR, there is evidence for both phylogenetically conserved RNA structures and an overlapping reading frame (ARF) in the 1+ frame, the latter encoding a putative polypeptide of approximately 124 aa [118]. Antibodies to polypeptides from this reading frame have been found in HCV infected patients, but the significance of its expression remains unclear [119–122]. Two RNA stem loop structures SL-V and SL-VI have been predicted to exist in the region containing the ARF and were confirmed by enzymatic structure probing (Fig. 6A) [120]. This region was found to be dispensable for the replication of the subgenomic HCV replicon in cell culture [122]. It was proposed that these stem loops might be involved in the stimulation of HCV IRES function by reducing inhibitory interactions between the 5' NTR and the core region [116,123]. Using a chimpanzee model McMullan et al. [122] have also demonstrated that the ARF protein was not essential for HCV genotype 1a H77 RNA replication. However, the ARF region contains a functionally important RNA element (SL-VI) whose role remains unknown.

A recent study by Vassilaki et al. [121] has examined the roles of 4 predicted RNA stem loops in the core-encoding region (SL47, SL87, SL248 and SL 443) and also evaluated the function of the core+1 ORF. In agreement with the results of McMullan et al. [122], the expression of the core + 1 ORF exhibited no role in the replication of HCV JFH1 isolate either in tissue culture or in xenografted mice. Using a mutational analysis it was observed that SL47 and SL87 (corresponding to SL-V and S-VI, respectively) [122] are important RNA elements contributing to HCV genome translation and replication both in cell culture and *in vivo* but SL248 and SL443 have no function. In addition, an interaction between a sequence at the 5' end of SL87 (nt 428–442) and nt 24–38 of the 5' NTR was predicted to be detrimental to IRES-dependent translation [124].

1.2.1.2. 3'-Terminal elements. The 3' NTR of HCV contains three domains, namely, the variable region (VR), the poly(U/UC) tract, and the highly conserved 3' X tail, which contains three putative stem loop structures SL1, SL2, and SL3 [125] (Fig. 5A). Using deletion mutants of the HCV-N subgenomic replicon RNA, Yi and Lemon [125] demonstrated that the 3'-most 150 nt [the 3'X tail and the 3' 52 nt of the poly(U/UC) tract] contained essential signals for RNA replication. The remaining segment of the 3' NTR enhanced replication but was not absolutely required. Similarly, in another study the 3'X tail and the poly(U/UC) tract were found to be required for infectivity of genome length RNA inoculated into chimpanzees [126].

Several independent groups have predicted the existence of extensive RNA secondary structure within the C-terminal NS5B encoding region of HCV, just upstream from the 3' NTR, suggesting the possibility that these have a function *in vivo* [120,127–130]. The C-terminal coding region (nt 9126–9374) of NS5B contains four highly conserved and stable stem loops (SL-IV to SL-VII) (Fig. 6A) [131]. A mutational analysis of these stem loops, in the context of the subgenomic HCV replicon, revealed that SL-V (nt 9262–9311) and SL-VI (nt 9215–9260) are essential for RNA replication in Huh-7 cells while the integrity of SL-IV (nt 9318–9355) and SL-VII (9129–9189) is less important. *In vitro* gel shift and filter-binding assays have shown that purified RNA polymerase NS5B specifically binds to SL-V. A similar study by You et al. [132] has also identified SL-V, designated as 5BSL3.2, as an essential *cis*-acting RNA for the replication of the HCV replicon in tissue culture. This hairpin is about 50 nt in length and consists of an 8-bp lower helix, a 6-bp upper helix, a 12-nt terminal loop and an 8-nt long internal loop. Primary sequences in the loops and RNA structure of the upper helix were found to be important for function. Interestingly, this RNA stem loop could be functionally moved to the 3' NTR, albeit with a reduction in replication efficiency. This same hairpin 5BSL3.2 was found to be involved in a “kissing”

interaction between the upper loop and the loop of another hairpin SL2, located in the 3' NTR (Fig. 6A) [133]. The importance of this tertiary RNA structure was confirmed by the rescue of RNA replication by compensating changes in both stem loops. Recent studies with the HCV cell culture system (HCVcc), using fully infectious HCV, demonstrated that both the “kissing” loop interaction and the length and composition of the poly(U/UC) tract were critical important for the replication in the genotype 2a HCVcc context [134]. A model was proposed in which one or more *trans* acting factors interact with the highly conserved X tail to aid the formation of the “kissing” loop interaction. Using various bioinformatics methods to detect phylogenetically conserved RNA secondary structures, Diviney et al. [135] have predicted an additional long range interaction between the bulge loop of 5SBL3.2 (also known as SL9266) and an unpaired sequence located about 200 nucleotides upstream (around nt 9100) in the NS5B coding sequence. Mutational analyses of the two interacting sequences provided genetic support for this interaction. Fig. 6 illustrates the proposed structure of this pseudoknot with long-range interactions of SL9266 with 3' and 5' sequences. Despite extensive studies in several laboratories the precise role of this stem loop structure (SL9266) in HCV RNA replication remains unknown.

1.2.1.3. Genome-scale ordered structures

Recently, large-scale ordered RNA structures (GORS) were detected in the genomes of HCV by advanced computational methods [64,65]. The presence of these large RNA structures was confirmed by hybridization accessibility assays. In addition, using atomic force microscopy HCV was visualized as tightly compacted spheroids. HCV belongs into a group of plus-strand RNA viruses where the presence of GORS is associated with viral persistence in the host.

1.2.2. Pestiviruses

Pestiviruses such as BVDV and CSFV are the causative agents of important livestock diseases.

1.2.2.1. 5'-Terminal elements. The 5' NTRs of pestiviruses and hepaciviruses have similar structural and functional organization. Just as the HCV genome, the pestivirus genome is not capped. It starts with the tetra nucleotide pGUAU whose integrity is required for efficient genome replication [136]. The 5' NTR (nt 385) of BVDV (NADL strain) contains stem loops designated as Ia, Ib, IIa, IIb, and IIIa-IIIc (Fig. 6B). The IRES resides within domain II (near nt 75) through nt ~310; that is, the 3'-end of the IRES extends into the ORF just as in the case of HCV [137]. Domain Ia is a bifunctional structure element in that it modulates both translation and RNA replication [138,139]. Mutations within the highly conserved 5'-terminal 4 nt (GUAU) of the genomic BVDV RNA resulted in severely impaired replication phenotypes [140].

1.2.2.2. 3' NTR. The 3' NTRs of pestiviruses are about 190- to 270-nt long. According to computer analyses the 3'-most 70 nt represent a highly conserved element of which the last 56-60 nt form a stable stem loop structure (SLI) [139,141]. The remaining region, with variable sequences, is predicted to form two less-stable stem loop structures (SLII, SLIII), designated as 3'V [142] (Fig. 6B). Studies with subgenomic BVDV replicon RNAs indicated that the 3'-terminal SLI

and part of the single-stranded region between SLI and SLII were essential for RNA replication [141]. The same results were obtained in the context of an infectious full-length BVDV cDNA clone [143]. The deletion of both SLII and SLIII were also found to result in a lethal growth phenotype [143]. However, deletions of either SLII or SLIII did not have a significant effect on viral replication. Interestingly, using a replicon system, Isken et al. [142] observed that the proper conformation of 3'V is required for accurate termination of translation at the stop codon of the viral ORF and this is essential for efficient RNA replication. However, this observation could not be confirmed in the context of the full length BVDV genome [143]. The binding of the so-called NFAR proteins both to the 3'V domain and to the 5' NTR was observed suggesting the possibility of an interaction between the 5'- and 3'-termini of the viral RNA [144].

1.2.3. Flaviviruses

The *Flavivirus* genus includes the medically important arthropod-borne viruses such as the mosquito transmitted DENV, YFV, KUNV and WNV or the tick-borne encephalitis virus (TBEV). Just like HCV or pestiviruses, the flaviviruses are enveloped plus-strand RNA (11 kb) viruses whose genome encodes a single ORF and lacks a 3'-terminal poly(A). In contrast to the two other genera, however, the flavivirus genome lacks an IRES. Instead it controls its translation via a type 1 cap at the 5'-end. It remains a puzzle why within the family of *Flaviviridae*, the member viruses of the genus *Flavivirus* use a cap-dependent mechanism of initiation of translation whereas member viruses of the *Hepacivirus* and *Pestivirus* use IRES elements that are structurally related to each other and to the above-mentioned picornavirus type IV IRESes. It has been speculated that “...since all picornaviruses contain IRES elements but only certain flaviviruses contain such elements then it is more likely that the flaviviruses acquired IRES elements (by recombination?) from picornaviruses” [145].

1.2.3.1. 5'-Terminal elements. The 5' NTRs of flaviviruses, which are about 100 nt in length, exhibit a high degree of sequence conservation among different strains of the same virus but less conservation among the different members of the genus [146]. The predicted secondary structures all contain a large stem loop (SLA) (~70 nt), which for DENV was proposed to function as the promoter for negative strand RNA synthesis (Fig. 6C) [147,148]. Two helical regions were identified in DENV SLA, a side stem loop, a top loop and an U bulge as critical for RNA replication [147]. A conserved oligo(U) track present downstream of SLA was found to modulate RNA synthesis in transfected cells [147]. A smaller hairpin (SLB) is located at or near the translation initiation codon. This RNA element contains a 16 nt long stretch, called the 5' upstream AUG region (5'UAR), which is complementary to a sequence in the 3' NTR (3'UAR). This pair of complementary sequences was named cyclization sequence 5'-3' UAR and the importance of this long-range interaction was shown to be essential for RNA replication of both DENV and WNV [149–151]. Additional complementary sequences have been identified between the coding sequence of protein C (5'CS) and a sequence just upstream of the 3'SL in the 3' NTR (3'CS). The requirement for complementarity between 5'CS and 3'CS for RNA replication was demonstrated for KUNV, DENV and WNV [150,152,153]. Long range interactions between the ends of

Fig. 6. Cis-acting RNA elements in the genomes of HCV (*Hepacivirus, Flaviviridae*), BVDV (*Pestivirus, Flaviviridae*), and DENV (*Flavivirus, Flaviviridae*). (A) HCV. The 5' NTR contains one stem loop and the IRES, which extends into the core coding sequence. Stem loops II and III of the IRES are shown enlarged. (Figure of IRES is taken from Lukowsky et al. [114] with permission of the publisher). The ORF contains a structural (SP) and a nonstructural (NSP) domain. The 3' NTR contains the 3'X tail, a polyU/C tract, and a variable (VR) domain. There are four stem loops (3'SL-IV to 3'SL-VII) in the NS5B C-terminal coding sequence, one of which (3'SL-V) is involved in “kissing interactions” with stem loop 2 of the 3' NTR and an unpaired sequence upstream in the NS5B coding sequence. (B) BVDV. The 5' NTR contains 2 stem loops (1a, 1b) and the IRES (stem loops IIa, IIb, IIIa-IIIc). The ORF contains a structural (SP) and a nonstructural (NSP) domain. The 3' NTR contains three stem loops (SLI to SLIII). (C) DENV. The 5' NTR contains two stem loops (SLA, B) with an additional small stem loop (cHP) in the capsid-coding region. SL-B contains a short sequence, UAR (upstream AUG region), which is complementary to a sequence in the 3' NTR. Additional complementary sequences (CS) are located near the 5'- and 3'-terminus. The ORF contains a structural (SP) and a nonstructural (NSP) domain. The 3' NTR consists of three domains with a total of five stem loops. The interaction between the 5'UAR/3'UAR and the 5'CS/3'CS are shown below. Fig. 6C is modified from two figures published in a review by Iglesias et al., [148], with permission of the publisher.

flavivirus RNA were confirmed by visualization of individual molecules using atomic force microscopy [150].

In many mosquito-borne flaviviruses the start codon of the polyprotein is in a poor Kozak initiation context suggesting that this start codon would be utilized inefficiently for the initiation of translation. A small RNA hairpin (cHP) located 14 nt into the capsid coding region of DENV and WNV was found to be involved both in the recognition of the 5' C protein start codon and also in RNA replication [154,155] (Fig. 6C). Among different flaviviruses cHP sequences are poorly conserved. Interestingly, no sequence-specific element has been identified in cHP but an intact structure is required for function.

1.2.3.2. 3' NTR. The 3' NTRs of flaviviruses are relatively large, and exhibit extensive heterogeneity in size and sequence. The 3' NTR of DENV can be divided into three domains. Domain I is located immediately downstream from the stop codon and is variable in sequence (Fig. 6C) [156]. Domain II contains moderately conserved sequences and two stem loops (A2 and A3), including a dumbbell structure with a conserved sequence motif (CS2) [105,157]. Deletions of A2 or A3 resulted in reduced RNA replication [150]. The most conserved domain (III) of the 3' NTR includes a stable stem loop at the terminus (3'SL), whose structure was confirmed by biochemical probing and which is absolutely required for RNA replication [150,158]. Upstream of the 3'SL there is the conserved cyclization sequence (CS1 or 3'CS) (Fig. 6C) [150].

The 3' NTR of tick-borne flaviviruses, exemplified by TBE (tick-borne encephalitis virus), is subdivided into a variable domain, just

downstream from the stop codon, and a 3'-terminal core element [159–161]. The variable region ranges in size from 50 nt to 400 nt and in some cases includes an internal poly(A) stretch. The core element, which is about 350 nt in length, is mostly conserved in sequence and structure. The very 3'-terminal 90–100 nt forms a highly conserved structural element, characteristic of all members of the *Flaviviridae*. Recently cyclization sequences (5'-CS-A and 3'-CS-A) have been identified for TBEV that are unrelated to the mosquito-borne flavivirus CS elements and are located at different genomic positions [162]. The 5'-CS-A is located in the 5' NTR upstream of the AUG start codon rather than down stream. The 3'-CS-A is situated in the stem of the 3'-SL structure. In contrast to mosquito-borne flaviviruses, the C protein-coding region is not required for RNA replication.

1.3. Dicistroviridae

The family of *Picornaviridae* contains vertebrate viruses, while plus-strand RNA viruses of invertebrates that have properties similar to those of picornaviruses have been termed as “picorna-like” viruses [163]. Several of these have genomes that are organized differently than those of picornaviruses. They have two large, nonoverlapping ORFs, and have been classified into a new family *Dicistroviridae*. The 5'-proximal ORF1 encodes the nonstructural proteins and the 3'-proximal ORF2 encodes the capsid precursor. ORF1 contains motifs for “picorna-like” replicases (helicase, protease, polymerase). Some dicistroviruses contain 2A-like sequences but others do not. Interestingly, they contain repeated VPg coding sequences, just like FMDV.

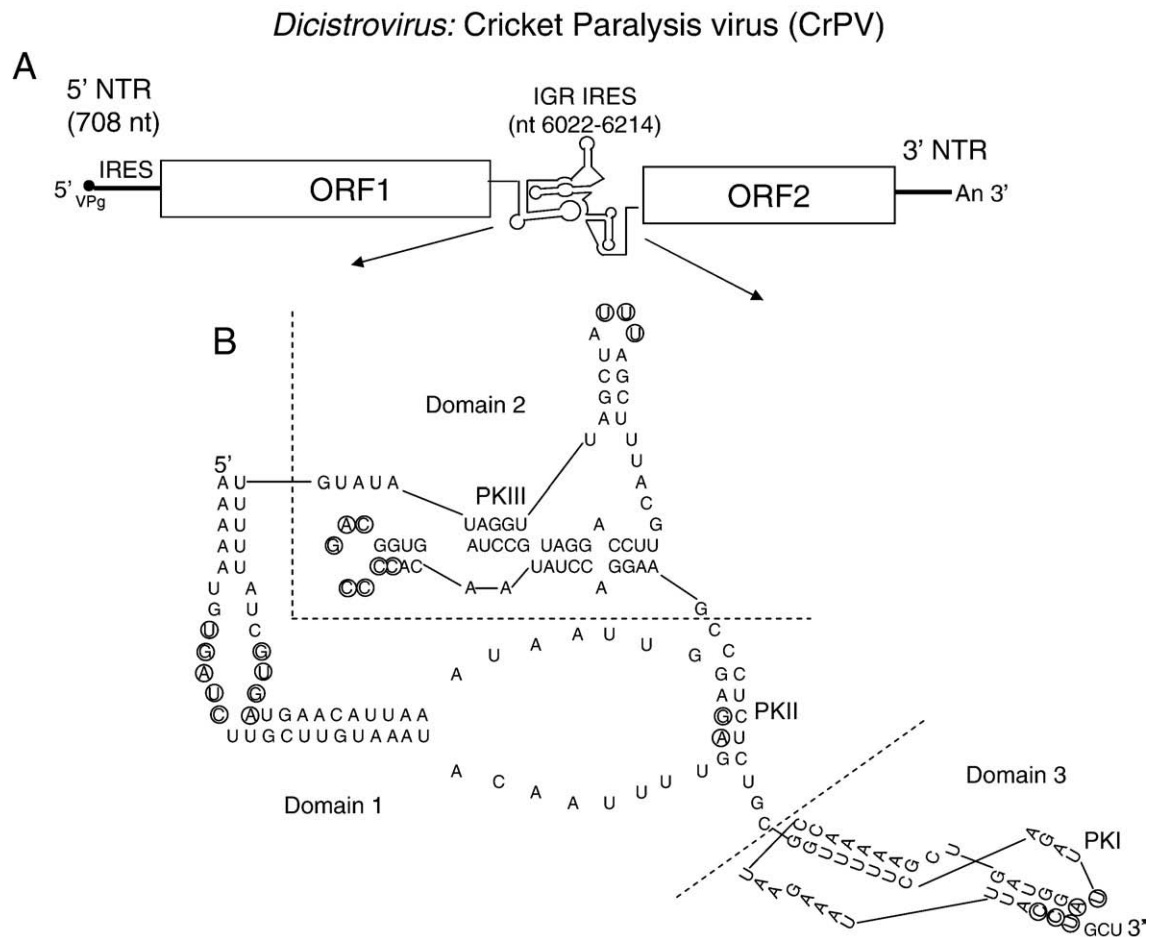


Fig. 7. Cis-acting RNA elements in the genome of CrPV (insect picorna-like virus, *Dicistroviridae*). (A) The genome structure of Cricket paralysis virus. The single-stranded RNA genome is linked to VPg at the 5'-end and contains two ORFs. The upstream ORF1 encodes the nonstructural proteins and the downstream ORF2 the capsid proteins. The two ORFs are separated by the intergenomic (IGR) IRES. Another IRES is located in the 5' NTR. The 3' NTR is polyadenylated. (B) Structure of the CrPV IGR-IRES element. Dotted lines indicate separation of domains. IRES nucleotides that are likely to interact with the ribosome are circled. Figure of IRES is taken from Schuler et al. [249], with permission of the publisher.

The second ORF is composed of coding sequences for capsid proteins VP2, VP4, VP3 and VP1 [164]. VP4 is cleaved from the VP0 precursor (VP4-VP3).

The best-characterized member of the family is Cricket paralysis virus (CrPV) belonging to the *Cripavirus* genus. Other well-known members of the genus include *Drosophila C virus* (DCV) [165], and *Plautia stali intestine virus* (PSIV) [166]. The genome of Cricket paralysis virus is about 8900-nt long, and contains two nonoverlapping ORFs translated from two different IRES elements (Fig. 7A). The 5' NTR (708 nt) contains the first IRES element [167] and an inter-genomic region (IGR) IRES is located between nt 6022 and 6214 (Fig. 7B). The length of the 3' NTR in dicistroviruses varies from 150 to 500 nt [163].

What is unusual about the IGR-IRESes of *Dicistroviridae* is that they do not require any of the canonical translation initiation factors to recruit the ribosome to the viral RNA, initiation is from a non-AUG start codon and from the ribosome's A-site rather than P-site [167–170]. The cryo-EM structure of the CrPV IRES, bound to ribosome, was determined by Schuler et al. (Fig. 7B). The structure indicates the presence of three pseudoknots (PKI, PKII and PKIII). Domain 1 of the IRES interacts with the 60S ribosomal subunit in the E- and P-site regions, domain 2 interacts with the 40S subunit in the E-site region, and domain 3 places the alanine start codon into the A-site.

1.4. Togaviridae

The *Togaviridae* family of animal viruses contains two genera, the *Alphavirus* and *Rubivirus*. These enveloped viruses contain an icosahedral nucleocapsid consisting of the viral RNA and multiple

copies of a single capsid protein (C). The RNA has a capped 5'-end and a poly(A) tail at the 3' end. The genome of these viruses is 9–12 kb long and it contains two ORFs. The 5'-terminal two thirds of the genome is translated into nonstructural proteins. The 3'-proximal ORF encodes the structural proteins, which is translated from a sub-genomic RNA (SG) (Fig. 8). The synthesis of both genomic (G) RNA and SG RNA is templated by the minus-strand RNA. However, the mechanism is different because the synthesis of G RNA begins at or near the 3' end while synthesis of SG RNA begins with internal initiation on the minus-strand template [171,172].

1.4.1. Alphavirus

The alphavirus genus contains about 30 members with Sindbis virus (SIN) as the prototype. The majority of alphaviruses are transmitted by mosquitoes to birds and mammals that serve as hosts. In cultured mosquito cells they establish persistent infections but infection of vertebrate cells leads to progressive cythopathic effect and cell death.

1.4.1.1. 5'-Terminal elements. The 5'-terminal 200 nt of SIN RNA contains two *cis*-acting elements, one located in the 5' NTR and the other in the ORF, called conserved sequence element, CSE (Fig. 8A). The 5' NTR (nt 1–59) contains a small (SL1) and a large stem loop structure (SL2). The 51-nt long CSE is located in the nsP1 coding sequence (nt 155–205), that is predicted to form two smaller stem loop structures (SL3 and SL4) [173]. Mutations in SL1 (nt 1–44) were either lethal or resulted in poor growth phenotypes [174] and were essential for efficient minus-strand RNA synthesis [173,175,176]. The effect of the mutations differed in mosquito cells and in mammalian

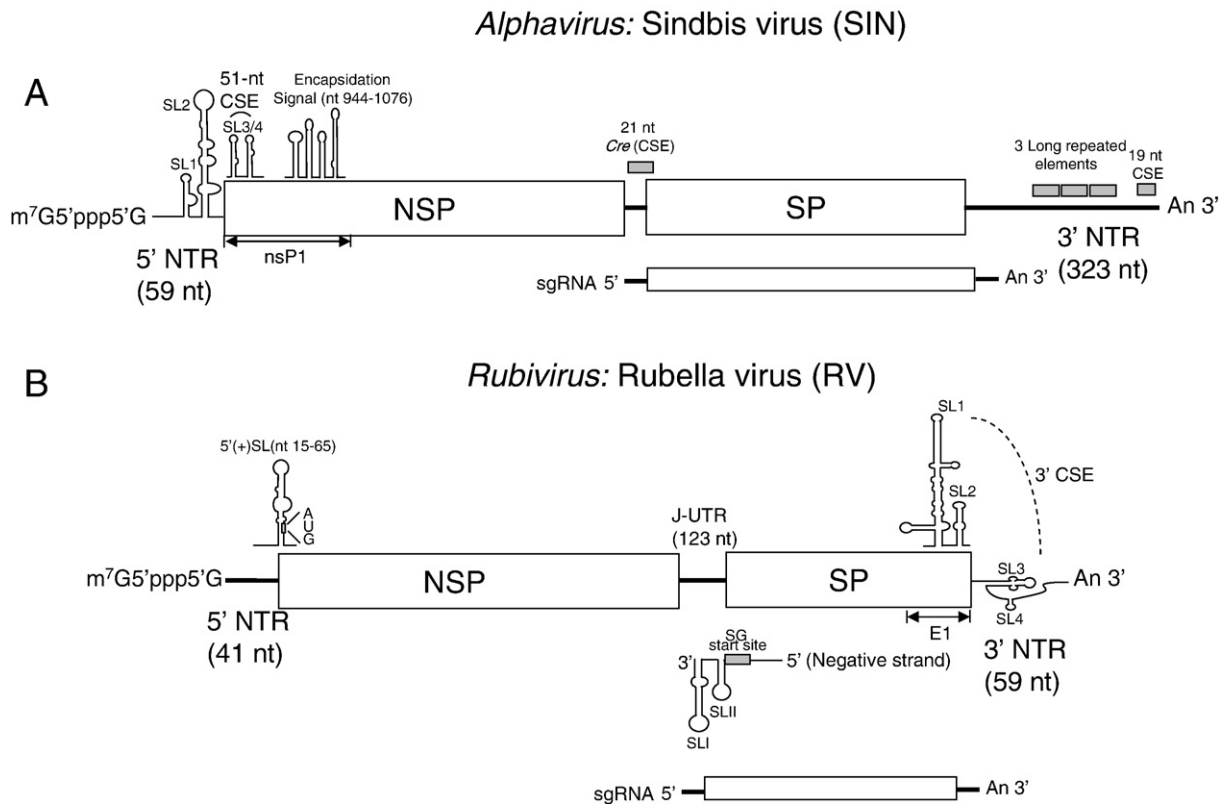


Fig. 8. *Cis*-acting RNA elements in the genomes of SIN (*Alphavirus, Togaviridae*) and of RV (*Rubivirus, Togaviridae*). (A) SIN. The 5' NTR contains two stem loops (SL1 and SL2). Just downstream of SL2, in the adjacent nsP1 coding sequence, there are two stem loops comprising the CSE. Four additional stem loops, located further downstream, form an encapsidation signal. The ORF contains a nonstructural domain (NSP) and a structural domain (SP), which is translated from a subgenomic RNA (sg RNA). There is a short connecting segment (CSE) between the NSP and SP domains. The 3' NTR contains repeated sequence elements, a short CSE segment and a poly(A) tail. (B) RV. The short 5' NTR and the adjacent NSP coding sequence contain a single stem loop. The ORF consists of a nonstructural (NSP) and a structural domain (SP). A connecting region, J-UTR, contains two stem loops in the minus-strand (SLI and SLII). SLII is located within the subgenomic promoter. The 3' NTR contains two stem loops (SL3 and SL4) followed by a poly(A) tail with two additional stem loops in the C-terminus of the E1 coding region (SL1 and SL2).

cells. SL2–SL4 comprises a replication enhancer but it is not essential for replication. The 51 nt CSE also enhances RNA replication and its integrity is more important for SIN replication in mosquito cells than in mammalian cells [173,174,177]. The complement of the 5'-terminal 200–250 nt was found to be required for the synthesis of viral genome RNAs from minus-strand intermediates [173,178].

It is interesting to note that some DI particles of Sindbis virus, obtained after serial high multiplicity passaging on cultured cells, were found to contain a cellular tRNA^{ASP} sequence (nt 10–75) at their 5'-ends [179,180]. This RNA sequence could be folded into a cloverleaf like structure but missing the 5'-terminal stem. In the two DI RNAs examined, the 3'-end of the cellular tRNA^{ASP} was covalently attached to viral sequences and therefore did not possess amino acid accepting activity.

1.4.1.2. 3' NTR. The 3' NTR of alphaviruses varies from 121 to 524 nt in length and that of SIN is 323 nt long (Fig. 8A). Immediately preceding the poly(A) tail a highly conserved 19 nt long CSE, with no discernable secondary structure, was identified. This sequence was postulated to play a key role in the initiation of minus-strand RNA synthesis from a plus-strand genomic template [181]. Upstream of the 3'-terminal CSE the 3' NTR contains an AU rich segment and repeated sequence elements 25–72 nt in length. In the 3' NTR of SIN there are 3 copies of a 40 nt long repeated element. Deletion of the repeated elements resulted in a viable virus but the growth of the virus was impaired [181]. It is not yet known whether the 3' and 5' ends of the genomic RNA interact to facilitate minus-strand RNA synthesis. Panhandle structures have been visualized by electron microscopy [182] but complementary sequences near the 5' and 3' ends are not evident.

1.4.1.3. Internal RNA elements. In SIN the initiation of nucleocapsid assembly begins with the recognition and binding of the encapsidation signal by capsid protein. The encapsidation signal of SIN genomic RNA was localized to a 132-nt long segment of the nsp1 coding sequence, spanning nt 944–1076, which is predicted to form four stem loop structures connected by single-stranded RNA sequences (Fig. 8A) [183–185]. The capsid protein of SIN is 264 aa long and of these amino acids 81–112 are involved in the recognition of the encapsidation signal [185]. Two purine-rich loops were found to be essential for the binding of the capsid protein to the RNA.

A second internal *cis*-acting RNA element in the SIN genome is located in the “junction region,” which includes nucleotide sequences of the genomic RNA preceding and including the beginning of the 26S SG RNA. This region contains the promoter for the synthesis of the subgenomic RNA (26S), in the context of the minus-strand. This is a highly conserved sequence in alphaviruses that encompasses 19 nt upstream and 2 nt downstream from the initiation site (Fig. 8A) [186]. Subsequent studies with DI particles indicated that the minimal subgenomic promoter for SIN was 18–19 nt upstream and 5 nt downstream from the start of the subgenomic RNA [172].

1.4.2. Rubivirus

Rubella virus (RV) is the only member of the Rubivirus genus. An interesting feature of rubella virus genotypes is their genomic uniformity [187]. A comparison of eight genotypes showed that 78% of nucleotides in the genomes were conserved. In addition, the viruses in all eight genotypes had the same number of nucleotides in each of the two ORFs and NTRs.

1.4.2.1. 5'-Terminal elements. The 5'-end of RV genomic RNA contains a 14-nt single-stranded leader followed by a stem loop structure [5'(+)SL] (nt 15–65), which contains the AUG (nt 41) for initiation of translation of the nonstructural proteins (Fig. 8B). Mutations of an AA dinucleotide at positions nt 2 or 3 were lethal

but no other mutation within the leader or the 5'(+)SL was lethal [188]. Some mutations in the leader or the stem of 5'(+)SL resulted in viruses that grew to a low titer. Interestingly, mutations in the 5'(+)SL resulted in a significant reduction in the synthesis of nonstructural proteins indicating that this structure is important for translation and not for plus-strand RNA synthesis.

The complementary negative strand equivalent of the 5' SL structure of RV [3'(–)SL] was found to bind specifically three cellular proteins (p56, p79, p97). Altering the SL structure in either one of the two predicted loops has abolished the binding interaction [189].

1.4.2.2. 3'-Terminal elements. Four stem loop structures were thermodynamically predicted to exist within the 3'-terminal 305 nt (3'CSE) of the RV RNA (SL1 to SL4) (Fig. 8B) [190]. SL1 and SL2 are located in the E1 coding region while SL3 and SL4 are within the 59 nt long 3' NTR. Mutational analyses indicated that most of the 3' NTR is required for RNA replication except for the 3'-terminal 5 nt and the poly(A) tail. Maintenance of the SL3 structure, an 11-nt long single-stranded sequence between SL3 and SL4, and the sequences forming SL4 were all important for virus viability [190]. A cellular protein, calreticulin (CAL), was found to bind to SL2 but this was independent of virus viability. Subsequent studies by Chen et al. [191] showed that the part of the 3'CSE, overlapping the E1 coding region, affected plus-strand but not minus-strand RNA synthesis.

1.4.2.3. Internal RNA elements. The two ORFs of RV are separated by an untranslated region known as the J-UTR (junction UTR), which contains on the negative strand the promoter for subgenomic (SG) RNA synthesis. The secondary structure predicted for the junction region contains two stem loop structures (SLI and SLII) (Fig. 8B) [192] of which SLII is within the subgenomic promoter (SGP). The minimal SGP starts from nt –26 through the SG start site and extends to nt +6, although a larger region is required for the generation of virus with wt phenotype.

1.5. Coronaviridae

The family *Coronaviridae*, in the order of *Nidovirales*, includes the largest plus-strand RNA viruses with 5'-capped and 3'-polyadenylated genomes ranging from 27.3 kb [human coronavirus (hCoV)] to 31.3 kb [mouse hepatitis virus (MHV)] [193]. At the 5'-end there is a 60–80 nt long leader sequence followed by a 200–500 nt long non-translated region. Coronavirus genomes contain multiple ORFs with the replicase gene occupying the 5'-most two thirds of the genome followed by the nonstructural proteins (Fig. 9). Additional accessory genes are interspersed among the structural proteins. There are 5–7 overlapping, 3'-coterminal subgenomic mRNAs, each of which is capped, polyadenylated, and contain a common leader sequence with the genomic RNA.

Most studies of coronavirus *cis*-acting RNA elements were conducted with DI RNAs because of the large size of the genomic coronavirus RNAs. DI RNAs, which contain extensive deletions, are not viable but they can propagate by using the RNA synthetic machinery provided by a homologous helper virus.

1.5.1. 5'-Terminal elements

The 5'-terminal leader sequence of Bovine coronavirus (BCoV) contains a stem loop structure that is involved in DI particle replication [194]. In groups 1 and 2 coronaviruses the 5' NTR ranges in size from 210 nt [MHV, hCoV and BCoV] to 314 nt [transmissible gastroenteritis virus (TGEV)]. The minimal 5' *cis*-replicating element for RNA replication usually extends into the replicase ORF. In the case of group 3 coronaviruses such as IBV the 5' NTR (528 nt) appears to contain the entire *cis*-acting RNA element. The 5'-genomic region of BCoV contains at least 6 stem loop structures, numbered SL-I to SL-VI in the 5' to 3' direction, whose structures were confirmed by

Group 2 Coronavirus: Bovine coronavirus (BCoV), mouse hepatitis virus (MHV)

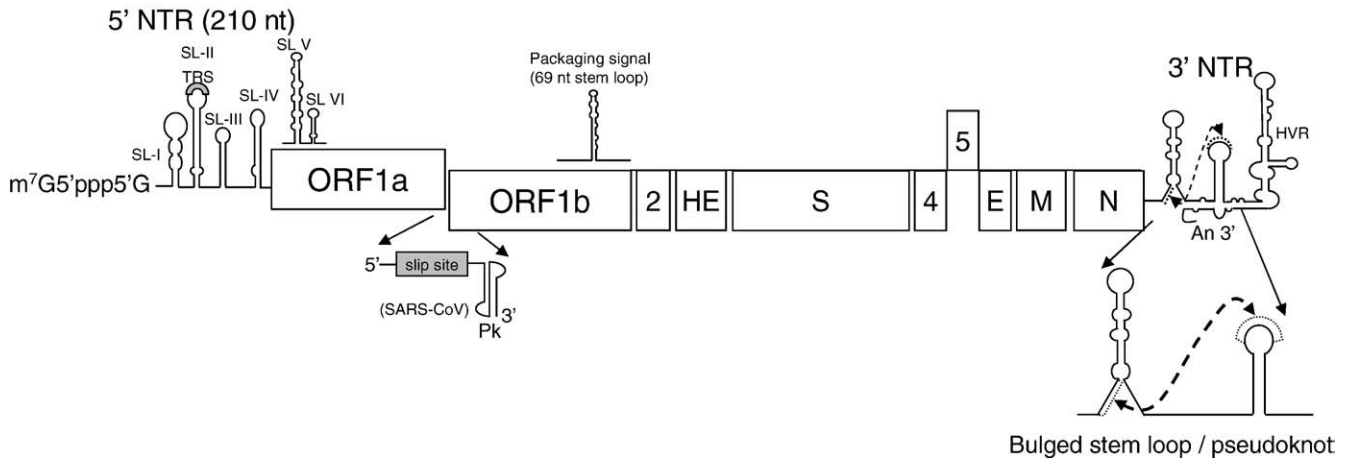


Fig. 9. *Cis*-acting RNA elements in the genomes of BCoV (*Coronaviridae*) and of mouse MHV (*Coronaviridae*). In BCoV the 5' NTR contains 4 stem loops (SL-I to SL-IV) with two additional stem loops in the adjacent ORF1a. A packaging signal is located in ORF1b. The genome contains multiple ORFs. The 3' NTRs of BCoV and MHV both contain a pseudoknot, a bulged stem loop, and a variable region (VAR). There is an interaction between the pseudoknot and the 3' end of the MHV genome. The genome is terminated with a poly(A) tail. In SARS-CoV there is a slippery site and a pseudoknot (Pk) in the overlapping sequence of ORF1a and 1b.

enzymatic probing (Fig. 9) [195–199]. Stem loops I and II were found to be essential for RNA replication. SL II contains a “transcription regulating sequence” (TRS), which is involved in the transcription of SG RNAs. SL-III and SL-IV were identified as *cis*-acting elements for DI RNA replication. In these stem loops the structure of the stem rather than the primary sequences are important for RNA replication. SL-IV contains the start codon for the replicase ORF in the downstream arm of its stem. The binding of purified nsp1, the N-terminal ORF 1a protein, to SLIII and its flanking sequences was also demonstrated [195]. A recent study suggests that the base of SL-I mediates a long range interaction between the 5' NTR and the 3' NTR that is an essential step in the transcription of SG RNAs but is not required for genomic minus-strand synthesis [200].

Studies with BCoV DI particles have determined that a 186 nt long sequence from the nsp1 coding region of the DI RNA ORF is required for RNA replication [201]. This sequence contains two stem loops (SL V and SL VI) whose structures were confirmed by RNase probing. SL VI maps within the nsp1 coding region (ORF 1a) found in all naturally occurring DI particles (Fig. 9) [196]. Mutational analyses have indicated that SLVI is required for DI RNA replication and is likely to be also involved in the replication of the full-length viral genome [201]. Recently, SLV was also identified as a *cis*-acting RNA element [195]. In SLV the overall higher-order structure and the integrity of the lower stem contributed to optimal RNA replication. Two cellular proteins were found to bind SLV and SLVI *in vitro* but the biological relevance of these interactions is not yet proven [195].

1.5.2. 3'-Terminal elements

The 3' NTR of coronaviruses ranges in size from 273 nt in TGEV to about 505 nt in infectious bronchitis virus (IBV) and is followed by a poly(A) tail. For MHV the domain required for RNA replication not only encompassed the 3' NTR but also ran into the adjacent N ORF [202]. In contrast the required *cis*-acting signals for coronaviruses groups 1 and 2 did not contain any of the N coding sequences [203]. The 3' NTR of MHV and BCoV upstream end contains a bulged stem loop [204] and a pseudoknot (Fig. 9) [205]. The bulged stem loop and the pseudoknot overlap and they cannot form simultaneously. This has led to a proposal that these structures are part of a molecular switch that regulates different steps of replication. The pseudoknot is followed by a variable region containing nt 46–156 from the 3'-end of MHV RNA (HVR), within which a highly conserved octanucleotide (GGAAGAGC) is located. Interestingly, the MHV HVR is required

for DI RNA replication but not for the replication of full length genomic RNA [206,207]. However, the HVR has a significant role in pathogenesis [207]. The very 3'-end of the genome (~55 nt) is required for MHV minus-strand RNA synthesis, as determined by DI RNA analysis, in conjunction with a poly(A) tail of 5–10 nt [205,208,209]. Recently, a direct interaction was demonstrated between loop 1 of the pseudoknot and the extreme 3' end of the MHV genome (Fig. 9) [210]. In the same study an interaction between the pseudoknot and gene 1a replicase products nsp8 and nsp9 was also reported.

1.5.3. Internal RNA elements

An internal *cis*-acting RNA element is located in the ORF1b of group 2 coronavirus MHV. A deletion analysis of MHV DI RNAs has identified a 190 nt long sequence within which a 69-nt bulged stem loop was required for packaging of DI RNAs into particles (Fig. 9) [211]. RNA structures similar to that found in MHV are also found in the ORF 1b regions of BCoV [212] and SCoV but their role as encapsidation signals remains to be demonstrated [213]. It should be noted that in other coronaviruses, such as IBV and TGEV, packaging signals are located at the genome ends [214].

The two replicase proteins of SARS-CoV, which are coterminal with their N-terminal ends, are produced by ribosomal frameshifting [215,216]. This process is dependent upon a slippery heptanucleotide sequence (UUUAAAC) and a closely spaced pseudoknot, located in the replicase ORF (Fig. 9).

1.6. Arteriviridae

Members of the *Arteriviridae* family are also called “small nidoviruses” because they have smaller genomes (13–16 kb) than the other virus families in the *Nidovirales* order. They are enveloped animal viruses that cause asymptomatic, persistent or acute infections. The nidovirus genomes contain a 5' cap structure, 5' and 3' NTRs, and a poly(A) tail. Their RNA genomes encode multiple (4–14) ORFs (Fig. 8). There are two large replicase ORFs and a set of downstream ORFs, which encode structural proteins from SG RNAs [217]. Not much is known about RNA structures in *Arteriviridae* because DI particles have not been observed for any arteriviruses. Recently, a reverse genetic system was used to study RNA elements in Porcine reproductive and respiratory syndrome virus (PRRSV), which uses an infectious cDNA as a bacterial artificial chromosome [218].

1.6.1. 5'-Terminal elements

The 5' NTRs of arterivirus genomic and subgenomic RNAs contain a leader sequence (nt 1–211) [217]. For Equine Arteritis Virus (EAV) the 5'-terminal 313 nt were shown to be sufficient for genome translation, RNA replication and SG RNA synthesis [219]. RNA folding of this region, containing the leader and part of ORF1a, predicted five stem loops (SL A–E), a pyrimidine rich segment followed by an additional five stem loops (SL F–J) (Fig. 10A) [220]. In stem loop G, contained in the leader sequence, there is a conserved transcription regulating sequence (TRS), which base pairs with downstream complementary TRS sequences ("body" TRSs). This interaction is essential for SG mRNA synthesis [220–222].

1.6.2. 3'-Terminal elements

Both full-length and SG minus-strand RNA synthesis are initiated at the 3'-end of the EAV genomic RNA. Using computer aided analysis and chemical and enzymatic probing of the 3'-terminal region (200 nt) of EAV RNA two domains were identified that are required for RNA synthesis [223]. The first domain, directly upstream of the 3' NTR (nt 12610–12654), contains one small stem loop (SL IV) and a single-stranded region (Fig. 10A). The second domain is located within the 3' NTR (nt 12661–12690) and is predicted to fold into a prominent stem loop (SL V) with a large loop. Two cellular proteins (PTB and aldolase) were found to specifically interact with the 3' NTR

but the biological significance of these interactions has not been established [224].

ORF7 of PRRSV, which encodes the nucleocapsid protein, is located just upstream of the 3' NTR. It contains a 34-nt long sequence that is required for minus-strand RNA synthesis [225]. This sequence is highly conserved among PRRSV isolates and is predicted to form a hairpin (Fig. 10B). A 7-nt long sequence within the loop of this structure was predicted to form a "kissing interaction" with the loop of a hairpin in the 3' NTR. Mutational analyses have confirmed that the "kissing interaction" was required for replication but the ability of the two loops to base pair and not their sequence was functionally relevant.

1.6.3. Internal RNA elements

The arterivirus genomes contain two large replicase ORFs (ORF1a and ORF1b). ORF1b translation requires a ribosomal frameshift just before ORF1a translation is terminated. The ORF1a/1b overlap region contains two signals that are believed to promote this function. The first is a "slippery" sequence (GUUAAC) that is the actual frame shift site and a downstream pseudoknot structure (Fig. 10A) [217,226].

1.7. Caliciviridae

The family *Caliciviridae* contains small plus-strand RNA viruses, which are members of four genera: *Norovirus*, *Sapovirus*, *Vesivirus*, and

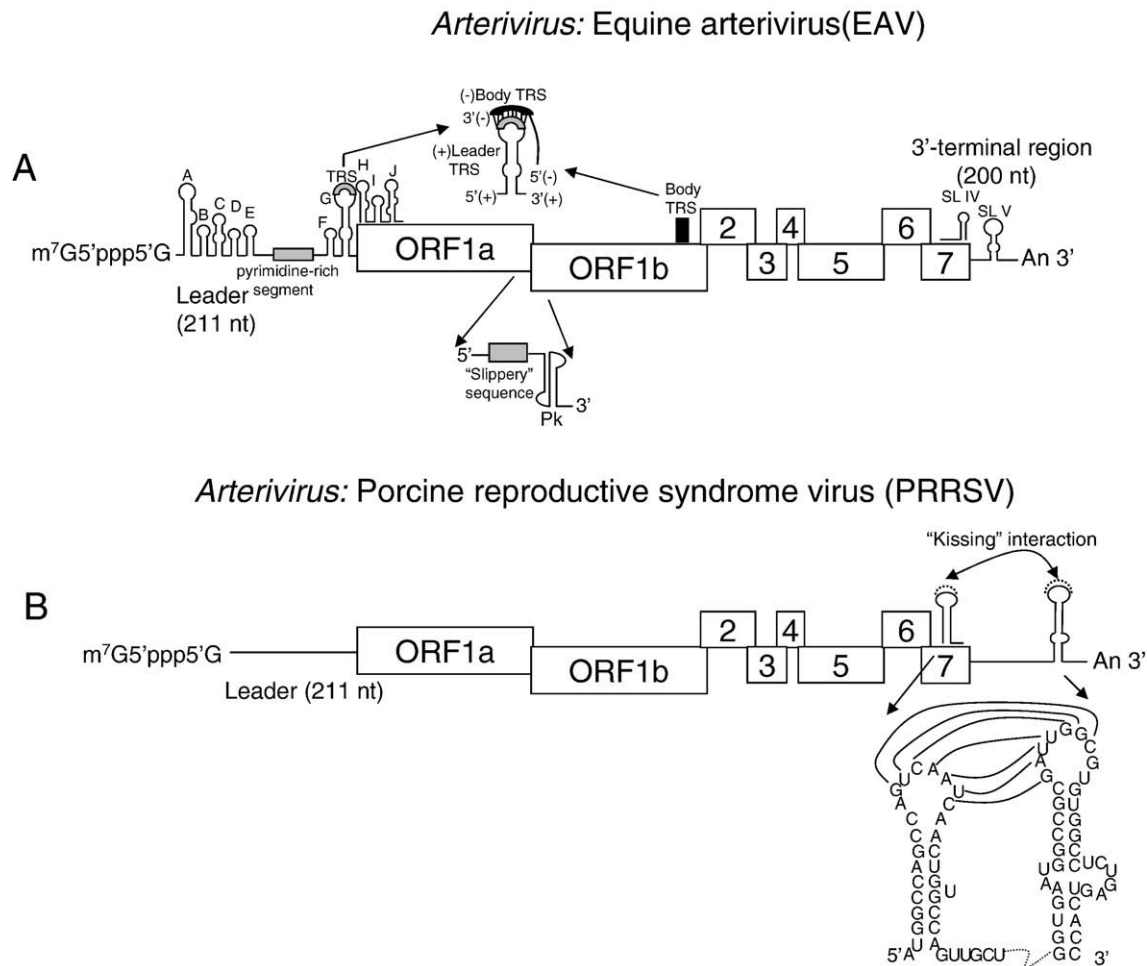


Fig. 10. Cis-acting RNA elements in the genomes of EAV (*Arteriviridae*) and of PRRSV (*Arteriviridae*). (A) EAV. The 5' NTR contains a leader sequence with 5 stem loops (SL-A to SL-E) and an additional five stem loops (SL-F–SL-J) of which the last two is located in the ORF1a coding sequence. The genome contains multiple ORFs. The 3' NTR contains one stem loop (SL-V) and another stem loop (SL-IV) is located just upstream in ORF 7. The genome is terminated with a poly(A) tail. (B) PRRSV. In PRRSV there is a leader sequence in the 5' NTR. The genome contains multiple ORFs. The 3' NTR contains one stem loop, which interacts with another stem loop located in the coding sequence of ORF7. The genome is terminated with a poly(A) tail.

Lagovirus. Human noroviruses and sapoviruses are important medical pathogens of humans, which cause gastroenteritis. Murine norovirus (MNV), the most thoroughly studied calicivirus, causes diarrhea and lethality in mice deficient in components of the innate immune system. Vesiviruses such as feline calicivirus (FCV) cause respiratory disease in cats and lagoviruses such as rabbit hemorrhagic disease virus (RHDV) causes hemorrhagic disease in rabbits. The genome of caliciviruses is 7.3–8.5 kb in length and is organized into two or three ORFs (Fig. 11). The nonstructural proteins are encoded near the 5'-end and the structural proteins near the 3'-end of the genome. The viral RNA is covalently linked to a terminal protein (VPg) at the 5'-end and the 3'-end is polyadenylated. So far very limited amount of information is available about *cis*-acting RNA structures in caliciviruses. Using a variety of bioinformatics methods Simmonds et al. [66] have recently searched for conserved RNA structures in the genomes of different human and animal caliciviruses.

1.7.1. 5'-Terminal elements

Caliciviruses have short 5' NTRs of 5–20 nt preceding their coding regions. Simmonds et al. [66] have predicted the presence of RNA structures in the first 150 nt of calicivirus genomes but there was no similarity in shape or position between groups or even within a genus. Nucleotide substitutions introduced into the 5'-terminal two stem loops starting at nt 8 and 29 of the murine norovirus (MNV) RNA genome resulted in a 20-fold reduction in virus titer and no significant inhibition of translation (Fig. 11) [66]. In the 5' NTR of Norwalk virus (NV) (norovirus genus) the existence of a double stem loop was predicted (nt 20–98) but its biological function was not tested [227]. Interestingly, caliciviruses use a novel protein-directed translation initiation, which relies on the interaction of eIF4E and VPg-linked to the RNA [228]. The 5'-terminal RNA structures in MNV RNA do not appear to have a significant role during the translation process [66].

1.7.2. 3'-Terminal elements

The 3' NTRs of caliciviruses vary in length from about 45–100 nt. Conserved secondary structures were reported in the 3' NTRs of MNV, human noroviruses, vesiviruses and sapoviruses [66,229]. Two large stem loops were identified in MNV sequences, a terminal hairpin from nt 7330 to the end of the genome and another large stem loop (164 nt) centering around position 7239 (Fig. 11) [66]. Mutational analysis of the terminal stem loop confirmed its essential role in RNA replication. The presence of 3'-terminal RNA stem loop structures was also reported in FCV genomes, both in the 3' NTR and in the adjacent coding sequences of ORF3 [230]. Binding of cellular proteins La, PTB, and PAB to the 3' NTR-poly(A) of Norwalk virus was reported but the biological significance of this observation is not yet established [231].

1.7.3. Internal RNA elements

A highly structured region was observed at the junction of the nonstructural and structural proteins in human norovirus, vesivirus, lagovirus and sapovirus [66]. A stem loop of varying sizes was located in the minus-strand just 6 nt upstream from the transcription site of the SG RNA in all of the caliciviruses tested indicating putative roles as promoters for SG RNA synthesis (Fig. 11). Nucleotide substitutions and repair mutations in the stem of the MNV stem loop resulted in a loss of viral viability and confirmed the importance of the RNA structure but not of the sequence.

1.8. Astroviridae

The family *Astroviridae* includes nonenveloped human and animal plus-strand RNA viruses, which usually cause gastroenteritis. The plus-strand RNA genome of astroviruses (6.4–7.3 kb) is linked to a terminal protein (VPg) at the 5'-end [232] and has a poly(A) tail at the 3' end (Fig. 12). The viral genome contains three overlapping ORFs. ORF1 (1a and 1b) encodes the nonstructural proteins while ORF 2, at the 3' end of both genomic and SG RNAs, encodes the structural proteins.

1.8.1. 5'- and 3'-terminal elements

A short 5' NTR (11–85 nt) precedes ORF1, which has not yet been characterized. The size of the 3' NTR varies from 80 nt in human astroviruses up to 305 nt in avian astroviruses. The genome is terminated in a poly(A) tail of about 30 nt in human astrovirus [233]. About 100 nt (80 from 3' NTR and 19 from ORF2) adjacent to the poly(A) tail of human astroviruses are very well conserved and are predicted to fold into two RNA stem loops (Fig. 12) [234].

1.8.2. Internal RNA elements

A ribosomal frame shift site was identified in the overlap region (70 nt) of ORF1a and 1b, consisting of a heptanucleotide and an RNA stem loop structure, which may form a pseudoknot with a downstream sequence (Fig. 12) [233]. At the junction of ORF1b and ORF2 a 120-nt long sequence, highly conserved among astroviruses, was predicted to form three hairpins (Fig. 12) [235]. The structure, which contains the start codon for the capsid gene as well as the stop codon for the polymerase gene, was proposed to be part of the promoter for sgRNA synthesis.

1.8.3. Genome-scale ordered structures

Recent bioinformatics analyses by Davis et al. [64] have predicted the presence of genome-scale ordered RNA structures in the genome of MNV, which was confirmed by probe hybridization accessibility assays. MNV is another example of mammalian viruses where the presence of GORS is correlated with the persistence of the virus in its host.

Calicivirus: Murine norovirus (MNV)

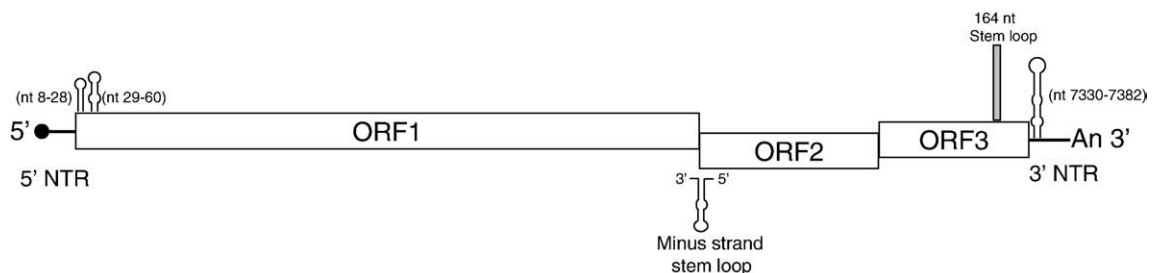


Fig. 11. *Cis*-acting RNA elements in the genome of MNV (*Norovirus, Caliciviridae*). The 5' NTR contains one stem loop and an adjacent stem loop in the ORF1 coding sequence. The genome contains three ORFs. The 3' NTR contains one large stem loop with an adjacent stem loop in the ORF3 coding sequence. The genome is terminated with a poly(A) tail. A minus-strand stem loop is located at the ORF1 and ORF2 junctions.

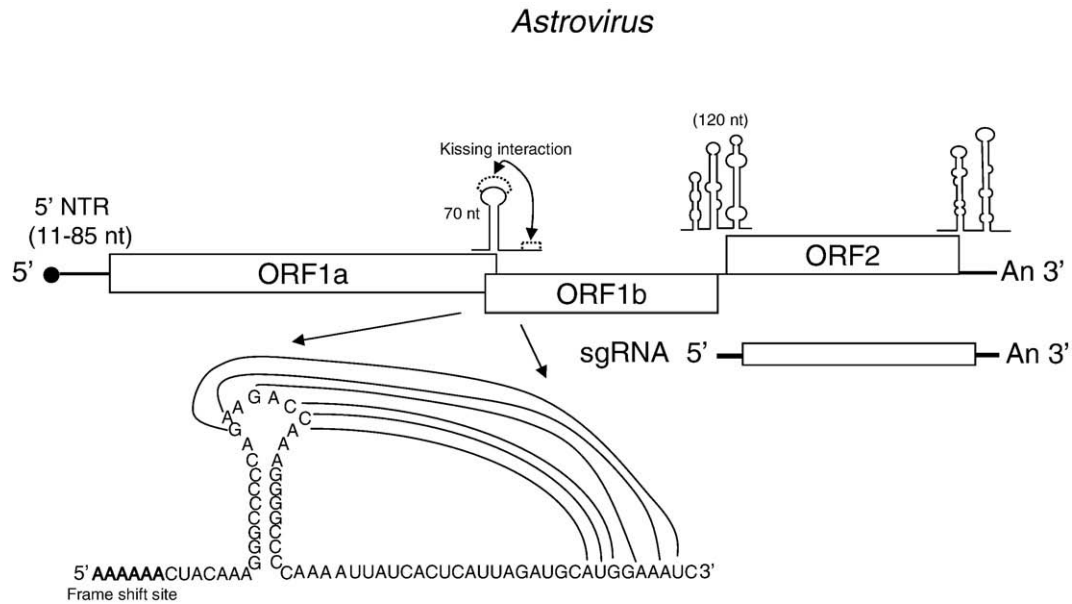


Fig. 12. Cis-acting RNA elements in the genome of Astrovirus (*Astroviridae*). The genome contains 3 ORFs (ORF1a, 1b and 2). The short 5' NTR is not yet characterized. A ribosomal frame shift site is located at the junction of ORFs 1a and 1b, which is predicted to interact with a downstream sequence. A 120 nt long segment, containing three stem loops, overlaps the junction of ORF1b and ORF2. The 3'-terminal region (110 nt) contains two highly conserved stem loops.

1.9. Nodaviridae

Nodaviruses are nonenveloped, icosahedral viruses. They are divided into two genera, α and β , that infect insects and fish, respectively. The most thoroughly studied insect viruses in this family are flock house virus (FHV) and blackbeetle virus (BBV). In tissue culture these viruses can replicate in cells derived from a variety of organisms such as yeast, mammals and plants.

The nodavirus genome consists of two plus-stranded RNAs, RNA1 (3.1 kb) and RNA2 (1.4 kb), that are capped at the 5'-end and do not contain a poly(A) tail at the 3'-end (Fig. 13). The combined length of RNA1 and RNA2 is 4507 nt, one of the smallest known viral genomes for animal viruses. Both RNAs are required for infectivity and are packaged within the same virion. RNA1 encodes the multifunctional replication protein A (112 kDa), which serves as the RNA polymerase for the synthesis of both genomic RNAs and of a subgenomic RNA, RNA3 (387 nt). Protein A also contains signals for targeting and insertion into outer mitochondrial membranes where RNA replication takes place [236]. RNA3, which corresponds to the 3'-end of RNA1, is not packaged into virions. It encodes two ORFs for the expression of proteins B1, whose function is unknown, and B2 (12 kDa), which is a suppressor of host-mediated silencing. RNA2 encodes the capsid

protein precursor α (43 kDa) [237], whose subunits assemble into a precursor viral particle, the provirion. Interestingly, RNA3 also acts as a transactivator of RNA2 replication [238]. At the onset of RNA2 replication the synthesis of RNA3 is suppressed [239].

Various RNA dimers can be observed during FHV replication, which appear to have functional roles in RNA replication [240]. These include dimeric RNAs covalently linked, head to tail monomers of RNAs 1, 2, and 3, as well as heterodimers of RNA2 and RNA3.

1.9.1. 5'- and 3'-terminal elements

FHV RNA replication is regulated by *cis*-acting RNA elements both at the 5'- and 3'-end of the genome [241]. The 5' NTR of FHV is 39 nt in length and the 3' NTR is 71 nt long [242]. Minus-strand RNA synthesis depends on the 3'-proximal 108 and 50 nt of RNA1 and RNA2, respectively. The 3'-terminal RNA element of RNA2 is predicted to form two stem loop structures [242]. In contrast, plus-strand RNA synthesis requires only about 3–14 nt at the 3'-terminus of minus-strands [243].

1.9.2. Internal RNA elements

There are also internal *cis*-acting RNA elements on the plus-strand of both RNA1 (intRE; nt 2322–2501) and RNA2 (nt 520–720) [244]. A

Nodavirus: Flock House Virus (FHV)

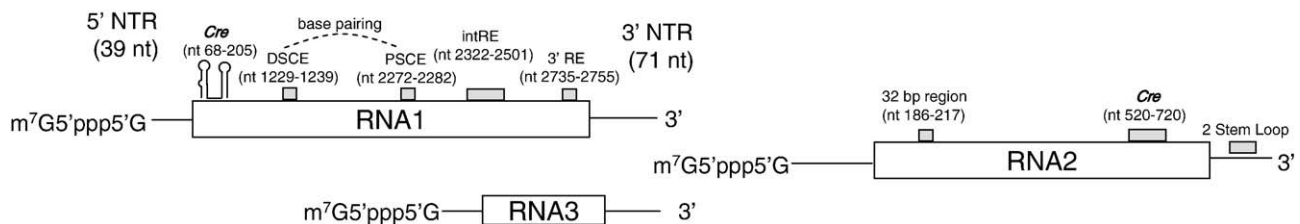


Fig. 13. Cis-acting RNA elements in the genome of Flock House virus (*Nodaviridae*). The bipartite genome of FHV consists of RNA1 and RNA2, each of which encode a single ORF. The subgenomic RNA (RNA3) encodes two ORFs. minus-strand RNA synthesis requires the 3-proximal 108 and 50 nt of RNA 1 and RNA2, respectively. Plus-strand RNA synthesis requires only 3–14 nt at the 3'-end of minus-strands. The synthesis of RNA3 is governed by a long-range interaction between two *cis*-elements on RNA1 (DSCE and PSCE). A packaging signal (32 nt) is located in RNA2. Additional RNA elements are located on RNA1 and RNA2, as shown.

second region with a 5' boundary between nt 2735–2755 (3'RE) is also important for RNA1 replication. Studies with a DI particle, isolated from persistently infected *Drosophila* cells, suggested that a 32-bp region of RNA2 (nt 186–217) is required for packaging of RNA2 into virions [245].

A recent report described the identification of a *cis*-acting element in FHV RNA1 (nt 68–205), which is predicted to form two stem loops with nearly identical loop sequences [246]. Mutational analyses of the stem loops showed that both are required for the recruitment of RNA1 to the outer mitochondrial membranes, and for minus and plus-strand RNA synthesis. The 5' NTR of FHV, comprising the 5'-terminal 39 nt, had no effect on RNA1 recruitment to the membranes.

RNA3 synthesis requires two *cis*-acting RNA elements on RNA1, which take part in a long distance base-pairing, involving about 6 base pairs (ACCGGU/UGGCCA) [244]. The proximal subgenomic control element (PSCE) (nt 2272–2282) is located just upstream of the RNA3 start site. A short distal subgenomic control element (DSCE) (nt 1229–1239) is 1.5 kb upstream from the RNA3 start site. Disruption of the DSCE-PSCE interaction abolished both negative and positive strand RNA3 replication. RNA3 can also replicate independently of RNA1 and this process requires the 3'-terminal 540 nt of plus-strand RNA3 [247].

2. Comparison of RNA structures in the different virus families

From the nine different virus families discussed in this review, those that have been studied in detail, all possess *cis*-acting RNA elements in the 5'- and 3'-terminal regions of their genomes. These RNA structures are involved in the complex process of RNA replication and in some cases also in promoting the initiation of translation. The virus families differ from each other greatly in the total number of functional RNA structures in their genomes and also in the lengths of their 5' and 3' NTRs. For example some viral RNAs contain large IRES elements in their 5' NTRs that are required for the translation of the polyprotein while other viruses with shorter 5' NTRs use cap-dependent translation. In general, where the 5' or 3' NTRs are too short to include all of the required functional elements, the RNA structures extend into the downstream or upstream coding sequences, respectively. In some viral genomes functional structural elements are also located in the coding sequences far away from the genomic ends that are not an extension of the 5' and 3' NTRs. These "internal" *cis*-acting RNA structures are involved in a variety of processes such as ribosomal frameshifting, RNA replication or encapsidation. Although the functional importance of many of these structures have been confirmed by genetic and biochemical studies their precise roles are not yet fully understood.

Recently, a less well-characterized, but much more extensive set of RNA secondary structures, designated as genome-scale ordered structure (GORS), have also been identified in the coding sequences of *Picornaviridae* (FMDV, kobuvirus, cardiovirus), *Caliciviridae* (OG caliciviruses), and *Flaviviridae* (HCV, HGV/GBV-C) genomes [64,65]. Interestingly, there was remarkable variability between genera of each family that exhibited this characteristic. GORS were not associated with translation or replication strategies of the various genera but they were strikingly correlated with the ability of each genus to persist in their natural host. The mechanism that underlines the association between GORS and host persistence is not yet understood. Davis et al. [64] suggested the possibility that the role of GORS is in the modulation of innate intracellular defense mechanisms, and the acquired immune system triggered by dsRNA. GORS may delay or prevent interferon induction early in viral infection or replication.

3. Concluding remarks

The accumulation of knowledge about RNA structures in the genomes of plus-strand RNA viruses during the last decade not only

enhances our understanding of the proliferation of individual viruses but also reveal certain unifying principles that link many aspects of translation, RNA replication and encapsidation. Future studies in this field will surely extend and refine our knowledge of how these structures function at the various stages of the viral life cycle and will have an impact on our ability to limit the toll of viral diseases.

Acknowledgements

We thank SK Jang for helpful discussions. This work was supported by grant AI15122 from the National Institute of Allergy and Infectious Diseases.

References

- [1] P.G. Rothberg, T.J. Harris, A. Nomoto, E. Wimmer, O4-(5'-uridylyl)tyrosine is the bond between the genome-linked protein and the RNA of poliovirus, *Proc. Natl. Acad. Sci. U. S. A.* 75 (1978) 4868–4872.
- [2] V. Ambros, D. Baltimore, Protein is linked to the 5' end of poliovirus RNA by a phosphodiester linkage to tyrosine, *J. Biol. Chem.* 253 (1978) 5263–5266.
- [3] Y.F. Lee, A. Nomoto, B.M. Detjen, E. Wimmer, A protein covalently linked to poliovirus genome RNA, *Proc. Natl. Acad. Sci. U. S. A.* 74 (1977) 59–63.
- [4] K. Dorsch-Hasler, Y. Yogo, E. Wimmer, Replication of picornaviruses: I. Evidence from *in vitro* RNA synthesis that poly(A) of the poliovirus genome is genetically coded, *J. Virol.* 16 (1975) 1512–1517.
- [5] Y. Yogo, E. Wimmer, Polyadenylic acid at the 3'-terminus of poliovirus RNA, *Proc. Natl. Acad. Sci. U. S. A.* 69 (1972) 1877–1882.
- [6] T.H.G. Stanway, N.J. Knowles, T. Hovi, Molecular and biological basis of picornavirus taxonomy, in: B.L. Semler, E. Wimmer (Eds.), *Molecular Biology of Picornaviruses*, ASM Press, Washington DC, 2002, pp. 17–24.
- [7] C. Witwer, S. Rauscher, I.L. Hofacker, P.F. Stadler, Conserved RNA secondary structures in Picornaviridae genomes, *Nucleic Acids Res.* 29 (2001) 5079–5089.
- [8] V.M. Rivera, J.D. Welsh, J.V. Maizel Jr., Comparative sequence analysis of the 5' noncoding region of the enteroviruses and rhinoviruses, *Virology* 165 (1988) 42–50.
- [9] A.V. Gamarnik, R. Andino, Switch from translation to RNA replication in a positive-stranded RNA virus, *Genes Dev.* 12 (1998) 2293–2304.
- [10] R. Andino, G.E. Rieckhof, P.L. Achacoso, D. Baltimore, Poliovirus RNA synthesis utilizes an RNP complex formed around the 5'-end of viral RNA, *EMBO J.* 12 (1993) 3587–3598.
- [11] R. Andino, G.E. Rieckhof, D. Baltimore, A functional ribonucleoprotein complex forms around the 5' end of poliovirus RNA, *Cell* 63 (1990) 369–380.
- [12] K.S. Harris, W. Xiang, L. Alexander, W.S. Lane, A.V. Paul, E. Wimmer, Interaction of poliovirus polypeptide 3CDpro with the 5' and 3' termini of the poliovirus genome. Identification of viral and cellular cofactors needed for efficient binding, *J. Biol. Chem.* 269 (1994) 27004–27014.
- [13] T.B. Parsley, J.S. Towner, L.B. Blyn, E. Ehrenfeld, B.L. Semler, Poly (rC) binding protein 2 forms a ternary complex with the 5'-terminal sequences of poliovirus RNA and the viral 3CD proteinase, *RNA* 3 (1997) 1124–1134.
- [14] W. Xiang, K.S. Harris, L. Alexander, E. Wimmer, Interaction between the 5'-terminal cloverleaf and 3AB/3CDpro of poliovirus is essential for RNA replication, *J. Virol.* 69 (1995) 3658–3667.
- [15] A.V. Gamarnik, R. Andino, Two functional complexes formed by KH domain containing proteins with the 5' noncoding region of poliovirus RNA, *RNA* 3 (1997) 882–892.
- [16] B.L. Walter, T.B. Parsley, E. Ehrenfeld, B.L. Semler, Distinct poly(rC) binding protein KH domain determinants for poliovirus translation initiation and viral RNA replication, *J. Virol.* 76 (2002) 12008–12022.
- [17] H. Toyoda, D. Franco, K. Fujita, A.V. Paul, E. Wimmer, Replication of poliovirus requires binding of the poly(rC) binding protein to the cloverleaf as well as to the adjacent C-rich spacer sequence between the cloverleaf and the internal ribosomal entry site, *J. Virol.* 81 (2007) 10017–10028.
- [18] A.V. Gamarnik, R. Andino, Interactions of viral protein 3CD and poly(rC) binding protein with the 5' untranslated region of the poliovirus genome, *J. Virol.* 74 (2000) 2219–2226.
- [19] L.E. Leong, P.A. Walker, A.G. Porter, Human rhinovirus-14 protease 3C (3Cpro) binds specifically to the 5'-noncoding region of the viral RNA. Evidence that 3Cpro has different domains for the RNA binding and proteolytic activities, *J. Biol. Chem.* 268 (1993) 25735–25739.
- [20] E. Rieder, W. Xiang, A. Paul, E. Wimmer, Analysis of the cloverleaf element in a human rhinovirus type 14/poliovirus chimera: correlation of subdomain D structure, ternary protein complex formation and virus replication, *J. Gen. Virol.* 84 (2003) 2203–2216.
- [21] J. Herold, R. Andino, Poliovirus RNA replication requires genome circularization through a protein-protein bridge, *Mol. Cell* 7 (2001) 581–591.
- [22] N. Sharma, B.J. O'Donnell, J.B. Flanagan, 3'-Terminal sequence in poliovirus negative-strand templates is the primary *cis*-acting element required for VPgUpU-primed positive-strand initiation, *J. Virol.* 79 (2005) 3565–3577.
- [23] H.H. Roehl, B.L. Semler, Poliovirus infection enhances the formation of two ribonucleoprotein complexes at the 3' end of viral negative-strand RNA, *J. Virol.* 69 (1995) 2954–2961.

- [24] R. Banerjee, A. Echeverri, A. Dasgupta, Poliovirus-encoded 2C polypeptide specifically binds to the 3'-terminal sequences of viral negative-strand RNA, *J. Virol.* 71 (1997) 9570–9578.
- [25] H.H. Roehl, T.B. Parsley, T.V. Ho, B.L. Semler, Processing of a cellular polypeptide by 3CD proteinase is required for poliovirus ribonucleoprotein complex formation, *J. Virol.* 71 (1997) 578–585.
- [26] V.I. Agol, A.V. Paul, E. Wimmer, Paradoxes of the replication of picornavirus genomes, *Virus Res.* 62 (1999) 129–147.
- [27] J. Zoll, H.A. Heus, F.J. van Kuppeveld, W.J. Melchers, The structure–function relationship of the enterovirus 3'-UTR, *Virus Res.* 139 (2009) 209–216.
- [28] E.V. Piliipenko, S.V. Maslova, A.N. Sinyakov, V.I. Agol, Towards identification of cis-acting elements involved in the replication of enterovirus and rhinovirus RNAs: a proposal for the existence of tRNA-like terminal structures, *Nucleic Acids Res.* 20 (1992) 1739–1745.
- [29] W.J. Melchers, J.G. Hoenderop, H.J. Bruins Slot, et al., Kissing of the two predominant hairpin loops in the coxsackie B virus 3' untranslated region is the essential structural feature of the origin of replication required for negative-strand RNA synthesis, *J. Virol.* 71 (1997) 686–696.
- [30] E.V. Piliipenko, K.V. Poperechny, S.V. Maslova, W.J. Melchers, H.J. Slot, V.I. Agol, Cis-element, oriR, involved in the initiation of (-) strand poliovirus RNA: a quasi-globular multi-domain RNA structure maintained by tertiary ('kissing') interactions, *EMBO J.* 15 (1996) 5428–5436.
- [31] S. Todd, J.H. Nguyen, B.L. Semler, RNA-protein interactions directed by the 3' end of human rhinovirus genomic RNA, *J. Virol.* 69 (1995) 3605–3614.
- [32] S. Todd, J.S. Townner, D.M. Brown, B.L. Semler, Replication-competent picornaviruses with complete genomic RNA 3' noncoding region deletions, *J. Virol.* 71 (1997) 8868–8874.
- [33] D.M. Brown, S.E. Kauder, C.T. Cornell, G.M. Jang, V.R. Racaniello, B.L. Semler, Cell-dependent role for the poliovirus 3' noncoding region in positive-strand RNA synthesis, *J. Virol.* 78 (2004) 1344–1351.
- [34] M.J. van Ooij, C. Polacek, D.H. Gludemans, et al., Polyadenylation of genomic RNA and initiation of antigenomic RNA in a positive-strand RNA virus are controlled by the same cis-element, *Nucleic Acids Res.* 34 (2006) 2953–2965.
- [35] J. Herold, R. Andino, Poliovirus requires a precise 5' end for efficient positive-strand RNA synthesis, *J. Virol.* 74 (2000) 6394–6400.
- [36] A.V. Paul, Possible unifying mechanism of picornavirus genome replication, in: B. L. Semler, E. Wimmer (Eds.), *Molecular Biology of Picornaviruses*, ASM Press, Washington DC, 2002, pp. 227–246.
- [37] B.P. Steil, D.J. Barton, Poliovirus cis-acting replication element-dependent VPg uridylylation lowers the Km of the initiating nucleoside triphosphate for viral RNA replication, *J. Virol.* 82 (2008) 9400–9408.
- [38] P.W. Mason, S.V. Bezborodova, T.M. Henry, Identification and characterization of a cis-acting replication element (cre) adjacent to the internal ribosome entry site of foot-and-mouth disease virus, *J. Virol.* 76 (2002) 9686–9694.
- [39] K.L. McKnight, S.M. Lemon, Capsid coding sequence is required for efficient replication of human rhinovirus 14 RNA, *J. Virol.* 70 (1996) 1941–1952.
- [40] K.L. McKnight, S.M. Lemon, The rhinovirus type 14 genome contains an internally located RNA structure that is required for viral replication, *Rna* 4 (1998) 1569–1584.
- [41] I. Goodfellow, Y. Chaudhry, A. Richardson, et al., Identification of a cis-acting replication element within the poliovirus coding region, *J. Virol.* 74 (2000) 4590–4600.
- [42] M.J. van Ooij, D.A. Vogt, A. Paul, et al., Structural and functional characterization of the coxsackievirus B3 CRE(2C): role of CRE(2C) in negative- and positive-strand RNA synthesis, *J. Gen. Virol.* 87 (2006) 103–113.
- [43] K. Gerber, E. Wimmer, A.V. Paul, Biochemical and genetic studies of the initiation of human rhinovirus 2 RNA replication: identification of a cis-replicating element in the coding sequence of 2A(pro), *J. Virol.* 75 (2001) 10979–10990.
- [44] S. Cordey, D. Gerlach, T. Junier, E.M. Zdobnov, L. Kaiser, C. Tapparel, The cis-acting replication elements define human enterovirus and rhinovirus species, *RNA* 14 (2008) 1568–1578.
- [45] A.V. Paul, E. Rieder, D.W. Kim, J.H. van Boom, E. Wimmer, Identification of an RNA hairpin in poliovirus RNA that serves as the primary template in the *in vitro* uridylylation of VPg, *J. Virol.* 74 (2000) 10359–10370.
- [46] E. Rieder, A.V. Paul, D.W. Kim, J.H. van Boom, E. Wimmer, Genetic and biochemical studies of poliovirus cis-acting replication element cre in relation to VPg uridylylation, *J. Virol.* 74 (2000) 10371–10380.
- [47] Y. Yang, R. Rijnbrand, K.L. McKnight, et al., Sequence requirements for viral RNA replication and VPg uridylylation directed by the internal cis-acting replication element (cre) of human rhinovirus type 14, *J. Virol.* 76 (2002) 7485–7494.
- [48] J. Yin, A.V. Paul, E. Wimmer, E. Rieder, Functional dissection of a poliovirus cis-acting replication element [PV-cre(2C)]: analysis of single- and dual-cre viral genomes and proteins that bind specifically to PV-cre RNA, *J. Virol.* 77 (2003) 5152–5166.
- [49] A.V. Paul, J. Yin, J. Mugavero, E. Rieder, Y. Liu, E. Wimmer, A "slide-back" mechanism for the initiation of protein-primed RNA synthesis by the RNA polymerase of poliovirus, *J. Biol. Chem.* 278 (2003) 43951–43960.
- [50] H.B. Pathak, J.J. Arnold, P.N. Wiegand, M.R. Hargittai, C.E. Cameron, Picornavirus genome replication: assembly and organization of the VPg uridylylation ribonucleoprotein (initiation) complex, *J. Biol. Chem.* 282 (2007) 16202–16213.
- [51] M. Shen, Z.J. Reitman, Y. Zhao, et al., Picornavirus genome replication. Identification of the surface of the poliovirus (PV) 3C dimer that interacts with PV 3Dpol during VPg uridylylation and construction of a structural model for the PV 3C2-3Dpol complex, *J. Biol. Chem.* 283 (2008) 875–888.
- [52] S. Crowder, K. Kirkegaard, Trans-dominant inhibition of RNA viral replication can slow growth of drug-resistant viruses, *Nat. Genet.* 37 (2005) 701–709.
- [53] J.Q. Han, H.L. Townsend, B.K. Jha, J.M. Paranjape, R.H. Silverman, D.J. Barton, A phylogenetically conserved RNA structure in the poliovirus open reading frame inhibits the antiviral endoribonuclease RNase L, *J. Virol.* 81 (2007) 5561–5572.
- [54] H.L. Townsend, B.K. Jha, R.H. Silverman, D.J. Barton, A putative loop E motif and an H-H kissing loop interaction are conserved and functional features in a group C enterovirus RNA that inhibits ribonuclease L, *RNA Biol.* 5 (2008) 263–272.
- [55] S.Y. Le, J.H. Chen, N. Sonenberg, J.V. Maizel Jr., Conserved tertiary structural elements in the 5' nontranslated region of cardiovirus, aphthovirus and hepatitis A virus RNAs, *Nucleic Acids Res.* 21 (1993) 2445–2451.
- [56] G.M. Duke, M.A. Hoffman, A.C. Palmenberg, Sequence and structural elements that contribute to efficient encephalomyocarditis virus RNA translation, *J. Virol.* 66 (1992) 1602–1609.
- [57] B.E. Clarke, A.L. Brown, K.M. Currey, S.E. Newton, D.J. Rowlands, A.R. Carroll, Potential secondary and tertiary structure in the genomic RNA of foot and mouth disease virus, *Nucleic Acids Res.* 15 (1987) 7067–7079.
- [58] P. Serrano, M.R. Pulido, M. Saiz, E. Martinez-Salas, The 3' end of the foot-and-mouth disease virus genome establishes two distinct long-range RNA–RNA interactions with the 5' end region, *J. Gen. Virol.* 87 (2006) 3013–3022.
- [59] M. Saiz, S. Gomez, E. Martinez-Salas, F. Sobrino, Deletion or substitution of the aphthovirus 3' NCR abrogates infectivity and virus replication, *J. Gen. Virol.* 82 (2001) 93–101.
- [60] H. Duque, A.C. Palmenberg, Phenotypic characterization of three phylogenetically conserved stem loop motifs in the mengovirus 3' untranslated region, *J. Virol.* 75 (2001) 3111–3120.
- [61] T. Cui, A.G. Porter, Localization of binding site for encephalomyocarditis virus RNA polymerase in the 3'-noncoding region of the viral RNA, *Nucleic Acids Res.* 23 (1995) 377–382.
- [62] P.E. Lobert, N. Escriou, J. Ruelle, T. Michiels, A coding RNA sequence acts as a replication signal in cardioviruses, *Proc. Natl. Acad. Sci. U. S. A.* 96 (1999) 11560–11565.
- [63] L. Tiley, A.M. King, G.J. Belsham, The foot-and-mouth disease virus cis-acting replication element (cre) can be complemented in trans within infected cells, *J. Virol.* 77 (2003) 2243–2246.
- [64] M. Davis, S.M. Sagan, J.P. Pezacki, D.J. Evans, P. Simmonds, Bioinformatic and physical characterizations of genome-scale ordered RNA structure in mammalian RNA viruses, *J. Virol.* 82 (2008) 11824–11836.
- [65] P. Simmonds, A. Tuplin, D.J. Evans, Detection of genome-scale ordered RNA structure (GORS) in genomes of positive-stranded RNA viruses: implications for virus evolution and host persistence, *Rna* 10 (2004) 1337–1351.
- [66] P. Simmonds, I. Karakasiliotis, D. Bailey, Y. Chaudhry, D.J. Evans, I.G. Goodfellow, Bioinformatic and functional analysis of RNA secondary structure elements among different genera of human and animal calciviruses, *Nucleic Acids Res.* 36 (2008) 2530–2546.
- [67] E.A. Brown, S.P. Day, R.W. Jansen, S.M. Lemon, The 5' nontranslated region of hepatitis A virus RNA: secondary structure and elements required for translation *in vitro*, *J. Virol.* 65 (1991) 5828–5838.
- [68] D.E. Schultz, C.C. Hardin, S.M. Lemon, Specific interaction of glyceraldehyde 3-phosphate dehydrogenase with the 5'-nontranslated RNA of hepatitis A virus, *J. Biol. Chem.* 271 (1996) 14134–14142.
- [69] Y.Y. Kusov, G. Morace, C. Probst, V. Gauss-Muller, Interaction of hepatitis A virus (HAV) precursor proteins 3AB and 3ABC with the 5' and 3' termini of the HAV RNA, *Virus Res.* 51 (1997) 151–157.
- [70] G. Dollenmaier, M. Weitz, Interaction of glyceraldehyde-3-phosphate dehydrogenase with secondary and tertiary RNA structural elements of the hepatitis A virus 3' translated and non-translated regions, *J. Gen. Virol.* 84 (2003) 403–414.
- [71] Y. Yang, M. Yi, D.J. Evans, P. Simmonds, S.M. Lemon, Identification of a conserved RNA replication element (cre) within the 3Dpol-coding sequence of hepatoviruses, *J. Virol.* 82 (2008) 10118–10128.
- [72] G. Reuter, A. Boldizar, P. Pankovics, Complete nucleotide and amino acid sequences and genetic organization of porcine kobuvirus, a member of a new species in the genus Kobuvirus, family Picornaviridae, *Arch. Virol.* 154 (2009) 101–108.
- [73] S. Nagashima, J. Sasaki, K. Taniguchi, The 5'-terminal region of the Aichi virus genome encodes cis-acting replication elements required for positive- and negative-strand RNA synthesis, *J. Virol.* 79 (2005) 6918–6931.
- [74] S. Nagashima, J. Sasaki, K. Taniguchi, Interaction between polypeptide 3ABC and the 5'-terminal structural elements of the genome of Aichi virus: implication for negative-strand RNA synthesis, *J. Virol.* 82 (2008) 6161–6171.
- [75] J. Sasaki, K. Taniguchi, The 5'-end sequence of the genome of Aichi virus, a picornavirus, contains an element critical for viral RNA encapsidation, *J. Virol.* 77 (2003) 3542–3548.
- [76] A.S. Nateri, P.J. Hughes, G. Stanway, Terminal RNA replication elements in human parechovirus 1, *J. Virol.* 76 (2002) 13116–13122.
- [77] M. Al-Sunaidi, C.H. Williams, P.J. Hughes, D.P. Schnurr, G. Stanway, Analysis of a new human parechovirus allows the definition of parechovirus types and the identification of RNA structural domains, *J. Virol.* 81 (2007) 1013–1021.
- [78] L.S. Chard, M.E. Bordeleau, J. Pelletier, J. Tanaka, G.J. Belsham, Hepatitis C virus-related internal ribosome entry sites are found in multiple genera of the family Picornaviridae, *J. Gen. Virol.* 87 (2006) 927–936.
- [79] S.K. Jang, H.G. Krausslich, M.J. Nicklin, G.M. Duke, A.C. Palmenberg, E. Wimmer, A segment of the 5' nontranslated region of encephalomyocarditis virus RNA directs internal entry of ribosomes during *in vitro* translation, *J. Virol.* 62 (1988) 2636–2643.
- [80] S.K. Jang, M.V. Davies, R.J. Kaufman, E. Wimmer, Initiation of protein synthesis by internal entry of ribosomes into the 5' nontranslated region of encephalomyocarditis virus RNA *in vivo*, *J. Virol.* 63 (1989) 1651–1660.

- [81] J. Pelletier, N. Sonenberg, Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA, *Nature* 334 (1988) 320–325.
- [82] J. Pelletier, N. Sonenberg, Internal binding of eucaryotic ribosomes on poliovirus RNA: translation in HeLa cell extracts, *J. Virol.* 63 (1989) 441–444.
- [83] A. Molla, S.K. Jang, A.V. Paul, Q. Reuer, E. Wimmer, Cardiovascular internal ribosomal entry site is functional in a genetically engineered dicistronic poliovirus, *Nature* 356 (1992) 255–257.
- [84] C.Y. Chen, P. Sarnow, Initiation of protein synthesis by the eukaryotic translational apparatus on circular RNAs, *Science* 268 (1995) 415–417.
- [85] S.K. Jang, Internal initiation: IRES elements of picornaviruses and hepatitis C virus, *Virus Res.* 119 (2006) 2–15.
- [86] E. Martinez-Salas, The impact of RNA structure on picornavirus IRES activity, *Trends Microbiol.* 16 (2008) 230–237.
- [87] M. Niepmann, Internal translation initiation of picornaviruses and hepatitis C virus, *Biochim. Biophys. Acta* 1789 (2009) 529–541.
- [88] E. Wimmer, C.U. Hellen, X. Cao, Genetics of poliovirus, *Annu. Rev. Genet.* 27 (1993) 353–436.
- [89] R. Zell, M. Dauber, A. Krumbholz, et al., Porcine teschoviruses comprise at least eleven distinct serotypes: molecular and evolutionary aspects, *J. Virol.* 75 (2001) 1620–1631.
- [90] A.V. Pisarev, L.S. Chard, Y. Kaku, H.L. Johns, I.N. Shatsky, G.J. Belsham, Functional and structural similarities between the internal ribosome entry sites of hepatitis C virus and porcine teschovirus, a picornavirus, *J. Virol.* 78 (2004) 4487–4497.
- [91] L.S. Chard, Y. Kaku, B. Jones, A. Nayak, G.J. Belsham, 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 (2006) 1271–1279.
- [92] S.K. Jang, T.V. Pestova, C.U. Hellen, G.W. Withereill, E. Wimmer, Cap-independent translation of picornavirus RNAs: structure and function of the internal ribosomal entry site, *Enzyme* 44 (1990) 292–309.
- [93] S.K. Jang, E. Wimmer, Cap-independent translation of encephalomyocarditis virus RNA: structural elements of the internal ribosomal entry site and involvement of a cellular 57-kD RNA-binding protein, *Genes Dev.* 4 (1990) 1560–1572.
- [94] T.V. Pestova, C.U. Hellen, E. Wimmer, Translation of poliovirus RNA: role of an essential cis-acting oligopyrimidine element within the 5' nontranslated region and involvement of a cellular 57-kilodalton protein, *J. Virol.* 65 (1991) 6194–6204.
- [95] E.V. Pilipenko, A.P. Gmyl, S.V. Maslova, Y.V. Svitkin, A.N. Sinyakov, V.I. Agol, Prokaryotic-like cis elements in the cap-independent internal initiation of translation on picornavirus RNA, *Cell* 68 (1992) 119–131.
- [96] A.P. Gmyl, E.V. Pilipenko, S.V. Maslova, G.A. Belov, V.I. Agol, Functional and genetic plasticities of the poliovirus genome: quasi-infectious RNAs modified in the 5'-untranslated region yield a variety of pseudorevertants, *J. Virol.* 67 (1993) 6309–6316.
- [97] E.V. Pilipenko, A.P. Gmyl, S.V. Maslova, et al., Starting window, a distinct element in the cap-independent internal initiation of translation on picornavirus RNA, *J. Mol. Biol.* 241 (1994) 398–414.
- [98] V.H. Johnson, B.L. Semler, Defined recombinants of poliovirus and coxsackievirus: sequence-specific deletions and functional substitutions in the 5'-noncoding regions of viral RNAs, *Virology* 162 (1988) 47–57.
- [99] M. Gromeier, L. Alexander, E. Wimmer, Internal ribosomal entry site substitution eliminates neurovirulence in intergeneric poliovirus recombinants, *Proc. Natl. Acad. Sci. U. S. A.* 93 (1996) 2370–2375.
- [100] L. Alexander, H.H. Lu, E. Wimmer, Polioviruses containing picornavirus type 1 and/or type 2 internal ribosomal entry site elements: genetic hybrids and the expression of a foreign gene, *Proc. Natl. Acad. Sci. U. S. A.* 91 (1994) 1406–1410.
- [101] H.H. Lu, E. Wimmer, Poliovirus chimeras replicating under the translational control of genetic elements of hepatitis C virus reveal unusual properties of the internal ribosomal entry site of hepatitis C virus, *Proc. Natl. Acad. Sci. U. S. A.* 93 (1996) 1412–1417.
- [102] M. Gromeier, S. Lachmann, M.R. Rosenfeld, P.H. Gutin, E. Wimmer, Intergenic poliovirus recombinants for the treatment of malignant glioma, *Proc. Natl. Acad. Sci. U. S. A.* 97 (2000) 6803–6808.
- [103] S.Y. Le, J.V. Maizel Jr, Evolution of a common structural core in the internal ribosome entry sites of picornavirus, *Virus Genes* 16 (1998) 25–38.
- [104] K. Tsukiyama-Kohara, N. Iizuka, M. Kohara, A. Nomoto, Internal ribosome entry site within hepatitis C virus RNA, *J. Virol.* 66 (1992) 1476–1483.
- [105] C. Thurner, C. Witwer, I.L. Hofacker, P.F. Stadler, Conserved RNA secondary structures in Flaviviridae genomes, *J. Gen. Virol.* 85 (2004) 1113–1124.
- [106] Y.K. Kim, C.S. Kim, S.H. Lee, S.K. Jang, Domains I and II in the 5' nontranslated region of the HCV genome are required for RNA replication, *Biochem. Biophys. Res. Commun.* 290 (2002) 105–112.
- [107] P. Friebe, V. Lohmann, N. Krieger, R. Bartenschlager, Sequences in the 5' nontranslated region of hepatitis C virus required for RNA replication, *J. Virol.* 75 (2001) 12047–12057.
- [108] C.L. Jopling, M. Yi, A.M. Lancaster, S.M. Lemon, P. Sarnow, Modulation of hepatitis C virus RNA abundance by a liver-specific MicroRNA, *Science* 309 (2005) 1577–1581.
- [109] W.D. Zhao, E. Wimmer, Genetic analysis of a poliovirus/hepatitis C virus chimera: new structure for domain II of the internal ribosomal entry site of hepatitis C virus, *J. Virol.* 75 (2001) 3719–3730.
- [110] J.S. Kieft, K. Zhou, R. Jubin, J.A. Doudna, Mechanism of ribosome recruitment by hepatitis C IRES RNA, *Rna* 7 (2001) 194–206.
- [111] P.J. Lukavsky, G.A. Otto, A.M. Lancaster, P. Sarnow, J.D. Puglisi, Structures of two RNA domains essential for hepatitis C virus internal ribosome entry site function, *Nat. Struct. Biol.* 7 (2000) 1105–1110.
- [112] P.J. Lukavsky, I. Kim, G.A. Otto, J.D. Puglisi, Structure of HCV IRES domain II determined by NMR, *Nat. Struct. Biol.* 10 (2003) 1033–1038.
- [113] J.S. Kieft, K. Zhou, A. Grech, R. Jubin, J.A. Doudna, Crystal structure of an RNA tertiary domain essential to HCV IRES-mediated translation initiation, *Nat. Struct. Biol.* 9 (2002) 370–374.
- [114] P.J. Lukavsky, Structure and function of HCV IRES domains, *Virus Res.* 139 (2009) 166–171.
- [115] S. Boni, J.P. Laverne, S. Boulant, A. Cahour, Hepatitis C virus core protein acts as a trans-modulating factor on internal translation initiation of the viral RNA, *J. Biol. Chem.* 280 (2005) 17737–17748.
- [116] T.H. Wang, R.C. Rijnbrand, S.M. Lemon, Core protein-coding sequence, but not core protein, modulates the efficiency of cap-independent translation directed by the internal ribosome entry site of hepatitis C virus, *J. Virol.* 74 (2000) 11347–11358.
- [117] J.E. Reynolds, A. Kaminski, H.J. Kettinen, et al., Unique features of internal initiation of hepatitis C virus RNA translation, *EMBO J.* 14 (1995) 6010–6020.
- [118] J.L. Walewski, T.R. Keller, D.D. Stump, A.D. Branch, Evidence for a new hepatitis C virus antigen encoded in an overlapping reading frame, *Rna* 7 (2001) 710–721.
- [119] Y. Morice, M. Ratniner, A. Miladi, et al., Seroconversion to hepatitis C virus alternate reading frame protein during acute infection, *Hepatology* 49 (2009) 1449–1459.
- [120] A. Tuplin, D.J. Evans, P. Simmonds, Detailed mapping of RNA secondary structures in core and NS5B-encoding region sequences of hepatitis C virus by RNase cleavage and novel bioinformatic prediction methods, *J. Gen. Virol.* 85 (2004) 3037–3047.
- [121] N. Vassilaki, P. Friebe, P. Meuleman, et al., Role of the hepatitis C virus core+1 open reading frame and core cis-acting RNA elements in viral RNA translation and replication, *J. Virol.* 82 (2008) 11503–11515.
- [122] L.K. McMullan, A. Grakoui, M.J. Evans, et al., Evidence for a functional RNA element in the hepatitis C virus core gene, *Proc. Natl. Acad. Sci. U. S. A.* 104 (2007) 2879–2884.
- [123] M. Honda, M.R. Beard, L.H. Ping, S.M. Lemon, A phylogenetically conserved stem loop structure at the 5' border of the internal ribosome entry site of hepatitis C virus is required for cap-independent viral translation, *J. Virol.* 73 (1999) 1165–1174.
- [124] Y.K. Kim, S.H. Lee, C.S. Kim, S.K. Seol, S.K. Jang, Long-range RNA–RNA interaction between the 5' nontranslated region and the core-coding sequences of hepatitis C virus modulates the IRES-dependent translation, *Rna* 9 (2003) 599–606.
- [125] M. Yi, S.M. Lemon, Structure-function analysis of the 3' stem loop of hepatitis C virus genomic RNA and its role in viral RNA replication, *Rna* 9 (2003) 331–345.
- [126] A.A. Kolykhalov, K. Mihalik, S.M. Feinstone, C.M. Rice, Hepatitis C virus-encoded enzymatic activities and conserved RNA elements in the 3' nontranslated region are essential for virus replication in vivo, *J. Virol.* 74 (2000) 2046–2051.
- [127] A. Tuplin, J. Wood, D.J. Evans, A.H. Patel, P. Simmonds, Thermodynamic and phylogenetic prediction of RNA secondary structures in the coding region of hepatitis C virus, *Rna* 8 (2002) 824–841.
- [128] J.H. Han, M. Houghton, Group specific sequences and conserved secondary structures at the 3' end of HCV genome and its implication for viral replication, *Nucleic Acids Res.* 20 (1992) 3520.
- [129] D.B. Smith, P. Simmonds, Characteristics of nucleotide substitution in the hepatitis C virus genome: constraints on sequence change in coding regions at both ends of the genome, *J. Mol. Evol.* 45 (1997) 238–246.
- [130] I.L. Hofacker, M. Fekete, C. Flamm, et al., Automatic detection of conserved RNA structure elements in complete RNA virus genomes, *Nucleic Acids Res.* 26 (1998) 3825–3836.
- [131] H. Lee, H. Shin, E. Wimmer, A.V. Paul, cis-acting RNA signals in the NS5B C-terminal coding sequence of the hepatitis C virus genome, *J. Virol.* 78 (2004) 10865–10877.
- [132] S. You, D.D. Stump, A.D. Branch, C.M. Rice, A cis-acting replication element in the sequence encoding the NS5B RNA-dependent RNA polymerase is required for hepatitis C virus RNA replication, *J. Virol.* 78 (2004) 1352–1366.
- [133] P. Friebe, J. Boudet, J.P. Simorre, R. Bartenschlager, Kissing-loop interaction in the 3' end of the hepatitis C virus genome essential for RNA replication, *J. Virol.* 79 (2005) 380–392.
- [134] S. You, C.M. Rice, 3' RNA elements in hepatitis C virus replication: kissing partners and long poly(U), *J. Virol.* 82 (2008) 184–195.
- [135] S. Diviney, A. Tuplin, M. Struthers, et al., A hepatitis C virus cis-acting replication element forms a long-range RNA–RNA interaction with upstream RNA sequences in NS5B, *J. Virol.* 82 (2008) 9008–9022.
- [136] I. Frolov, M.S. McBride, C.M. Rice, cis-acting RNA elements required for replication of bovine viral diarrhoea virus-hepatitis C virus 5' nontranslated region chimeras, *Rna* 4 (1998) 1418–1435.
- [137] S.K. Chon, D.R. Perez, R.O. Donis, Genetic analysis of the internal ribosome entry sequence of bovine viral diarrhoea virus, *Virology* 251 (1998) 370–382.
- [138] H. Yu, O. Isken, C.W. Grassmann, S.E. Behrens, A stem loop motif formed by the immediate 5' terminus of the bovine viral diarrhoea virus genome modulates translation as well as replication of the viral RNA, *J. Virol.* 74 (2000) 5825–5835.
- [139] R. Deng, K.V. Brock, 5' and 3' untranslated regions of pestivirus genome: primary and secondary structure analyses, *Nucleic Acids Res.* 21 (1993) 1949–1957.
- [140] P. Becher, M. Orlich, H.J. Thiel, Mutations in the 5' nontranslated region of bovine viral diarrhoea virus result in altered growth characteristics, *J. Virol.* 74 (2000) 7884–7894.
- [141] H. Yu, C.W. Grassmann, S.E. Behrens, Sequence and structural elements at the 3' terminus of bovine viral diarrhoea virus genomic RNA: functional role during RNA replication, *J. Virol.* 73 (1999) 3638–3648.
- [142] O. Isken, C.W. Grassmann, H. Yu, S.E. Behrens, Complex signals in the genomic 3' nontranslated region of bovine viral diarrhoea virus coordinate translation and replication of the viral RNA, *Rna* 10 (2004) 1637–1652.

- [143] A. Pankraz, H.J. Thiel, P. Becher, Essential and nonessential elements in the 3' nontranslated region of Bovine viral diarrhoea virus, *J. Virol.* 79 (2005) 9119–9127.
- [144] O. Isken, C.W. Grassmann, R.T. Sarisky, et al., Members of the NF90/NFAR protein group are involved in the life cycle of a positive-strand RNA virus, *EMBO J.* 22 (2003) 5655–5665.
- [145] G.J. Belsham, Divergent picornavirus IRES elements, *Virus Res.* 139 (2009) 183–192.
- [146] M.A. Brinton, J.H. Dispoto, Sequence and secondary structure analysis of the 5'-terminal region of flavivirus genome RNA, *Virology* 162 (1988) 290–299.
- [147] M.F. Lodeiro, C.V. Filomatori, A.V. Gamarnik, Structural and functional studies of the promoter element for dengue virus RNA replication, *J. Virol.* 83 (2009) 993–1008.
- [148] N.G. Iglesias, C.V. Filomatori, D.E. Alvarez, A.V. Gamarnik, Flaviviruses, in: C.E. Cameron, M. Gotte, K.D. Raney (Eds.), *Viral genome replication*, Springer, 2009, pp. 41–60.
- [149] D.E. Alvarez, C.V. Filomatori, A.V. Gamarnik, Functional analysis of dengue virus cyclization sequences located at the 5' and 3'UTRs, *Virology* 375 (2008) 223–235.
- [150] D.E. Alvarez, M.F. Lodeiro, S.J. Luduena, L.I. Pietrasanta, A.V. Gamarnik, Long-range RNA–RNA interactions circularize the dengue virus genome, *J. Virol.* 79 (2005) 6631–6643.
- [151] B. Zhang, H. Dong, D.A. Stein, P.L. Iversen, P.Y. Shi, West Nile virus genome cyclization and RNA replication require two pairs of long-distance RNA interactions, *Virology* 373 (2008) 1–13.
- [152] A.A. Khromykh, H. Meka, K.J. Guyatt, E.G. Westaway, Essential role of cyclization sequences in flavivirus RNA replication, *J. Virol.* 75 (2001) 6719–6728.
- [153] M.K. Lo, M. Tilgner, K.A. Bernard, P.Y. Shi, Functional analysis of mosquito-borne flavivirus conserved sequence elements within 3' untranslated region of West Nile virus by use of a reporting replicon that differentiates between viral translation and RNA replication, *J. Virol.* 77 (2003) 10004–10014.
- [154] K. Clyde, J. Barrera, E. Harris, The capsid-coding region hairpin element (cHP) is a critical determinant of dengue virus and West Nile virus RNA synthesis, *Virology* 379 (2008) 314–323.
- [155] K. Clyde, E. Harris, RNA secondary structure in the coding region of dengue virus type 2 directs translation start codon selection and is required for viral replication, *J. Virol.* 80 (2006) 2170–2182.
- [156] J.E. Bryant, P.F. Vasconcelos, R.C. Rijnbrand, J.P. Mutebi, S. Higgs, A.D. Barrett, Size heterogeneity in the 3' noncoding region of South American isolates of yellow fever virus, *J. Virol.* 79 (2005) 3807–3821.
- [157] T.A. Romero, E. Tumban, J. Jun, W.B. Lott, K.A. Hanley, Secondary structure of dengue virus type 4 3' untranslated region: impact of deletion and substitution mutations, *J. Gen. Virol.* 87 (2006) 3291–3296.
- [158] L. Zeng, B. Falgout, L. Markoff, Identification of specific nucleotide sequences within the conserved 3'-SL in the dengue type 2 virus genome required for replication, *J. Virol.* 72 (1998) 7510–7522.
- [159] G. Wallner, C.W. Mandl, C. Kunz, F.X. Heinz, The flavivirus 3'-noncoding region: extensive size heterogeneity independent of evolutionary relationships among strains of tick-borne encephalitis virus, *Virology* 213 (1995) 169–178.
- [160] S. Rauscher, C. Flamm, C.W. Mandl, F.X. Heinz, P.F. Stadler, Secondary structure of the 3'-noncoding region of flavivirus genomes: comparative analysis of base pairing probabilities, *Rna* 3 (1997) 779–791.
- [161] T.S. Gritsun, K. Venugopal, P.M. Zanotto, et al., Complete sequence of two tick-borne flaviviruses isolated from Siberia and the UK: analysis and significance of the 5' and 3'-UTRs, *Virus Res* 49 (1997) 27–39.
- [162] R.M. Kofler, V.M. Hoenninger, C. Thurner, C.W. Mandl, Functional analysis of the tick-borne encephalitis virus cyclization elements indicates major differences between mosquito-borne and tick-borne flaviviruses, *J. Virol.* 80 (2006) 4099–4113.
- [163] N. Nakashima, T. Uchiyama, Functional analysis of structural motifs in dicistroviruses, *Virus Res.* 139 (2009) 137–147.
- [164] J. Sasaki, N. Nakashima, H. Saito, H. Noda, An insect picorna-like virus, *Plautia stali* intestine virus, has genes of capsid proteins in the 3' part of the genome, *Virology* 244 (1998) 50–58.
- [165] K.N. Johnson, P.D. Christian, The novel genome organization of the insect picorna-like virus *Drosophila C* virus suggests this virus belongs to a previously undescribed virus family, *J. Gen. Virol.* 79 (Pt. 1) (1998) 191–203.
- [166] Y. Kanamori, N. Nakashima, A tertiary structure model of the internal ribosome entry site (IRES) for methionine-independent initiation of translation, *Rna* 7 (2001) 266–274.
- [167] J.E. Wilson, M.J. Powell, S.E. Hoover, P. Sarnow, Naturally occurring dicistronic cricket paralysis virus RNA is regulated by two internal ribosome entry sites, *Mol. Cell Biol.* 20 (2000) 4990–4999.
- [168] J. Sasaki, N. Nakashima, 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 (1999) 1219–1226.
- [169] N. Deniz, E.M. Lenarcic, D.M. Landry, S.R. Thompson, Translation initiation factors are not required for Dicistroviridae IRES function in vivo, *Rna* 15 (2009) 932–946.
- [170] J.S. Kieft, Comparing the three-dimensional structures of Dicistroviridae IGR IRES RNAs with other viral RNA structures, *Virus Res.* 139 (2009) 148–156.
- [171] M.L. Li, T.Y. Kwan, H.A. Simmonds, V. Stollar, Synthesis of genomic and subgenomic RNA in mosquito cells infected with two Sindbis virus nsp4 mutants: influence of intracellular nucleoside triphosphate concentrations, *J. Virol.* 82 (2008) 6880–6888.
- [172] R. Levis, S. Schlesinger, H.V. Huang, Promoter for Sindbis virus RNA-dependent subgenomic RNA transcription, *J. Virol.* 64 (1990) 1726–1733.
- [173] I. Frolov, R. Hardy, C.M. Rice, Cis-acting RNA elements at the 5' end of Sindbis virus genome RNA regulate minus- and plus-strand RNA synthesis, *Rna* 7 (2001) 1638–1651.
- [174] H.G. Niesters, J.H. Strauss, Mutagenesis of the conserved 51-nucleotide region of Sindbis virus, *J. Virol.* 64 (1990) 1639–1647.
- [175] D.G. Nickens, R.W. Hardy, Structural and functional analyses of stem loop 1 of the Sindbis virus genome, *Virology* 370 (2008) 158–172.
- [176] R. Gorchakov, R. Hardy, C.M. Rice, I. Frolov, Selection of functional 5' cis-acting elements promoting efficient sindbis virus genome replication, *J. Virol.* 78 (2004) 61–75.
- [177] R. Fayzulin, I. Frolov, Changes of the secondary structure of the 5' end of the Sindbis virus genome inhibit virus growth in mosquito cells and lead to accumulation of adaptive mutations, *J. Virol.* 78 (2004) 4953–4964.
- [178] J.H. Ou, E.G. Strauss, J.H. Strauss, The 5'-terminal sequences of the genomic RNAs of several alphaviruses, *J. Mol. Biol.* 168 (1983) 1–15.
- [179] S.S. Monroe, S. Schlesinger, RNAs from two independently isolated defective interfering particles of Sindbis virus contain a cellular tRNA sequence at their 5' ends, *Proc. Natl. Acad. Sci. U. S. A.* 80 (1983) 3279–3283.
- [180] S.S. Monroe, S. Schlesinger, Common and distinct regions of defective-interfering RNAs of Sindbis virus, *J. Virol.* 49 (1984) 865–872.
- [181] R.J. Kuhn, Z. Hong, J.H. Strauss, Mutagenesis of the 3' nontranslated region of Sindbis virus RNA, *J. Virol.* 64 (1990) 1465–1476.
- [182] T.K. Frey, D.L. Gard, J.H. Strauss, Biophysical studies on circle formation by Sindbis virus 49 S RNA, *J. Mol. Biol.* 132 (1979) 1–18.
- [183] B. Weiss, U. Geigenmüller-Girke, S. Schlesinger, Interactions between Sindbis virus RNAs and a 68 amino acid derivative of the viral capsid protein further defines the capsid binding site, *Nucleic Acids Res.* 22 (1994) 780–786.
- [184] E. Frolova, I. Frolov, S. Schlesinger, Packaging signals in alphaviruses, *J. Virol.* 71 (1997) 248–258.
- [185] B.R. Linger, L. Kunovska, R.J. Kuhn, B.L. Golden, Sindbis virus nucleocapsid assembly: RNA folding promotes capsid protein dimerization, *Rna* 10 (2004) 128–138.
- [186] J.H. Ou, C.M. Rice, L. Dalgarno, E.G. Strauss, J.H. Strauss, Sequence studies of several alphavirus genomic RNAs in the region containing the start of the subgenomic RNA, *Proc. Natl. Acad. Sci. U. S. A.* 79 (1982) 5235–5239.
- [187] Y. Zhou, H. Ushijima, T.K. Frey, Genomic analysis of diverse rubella virus genotypes, *J. Gen. Virol.* 88 (2007) 932–941.
- [188] K.V. Pugachev, T.K. Frey, Effects of defined mutations in the 5' nontranslated region of rubella virus genomic RNA on virus viability and macromolecule synthesis, *J. Virol.* 72 (1998) 641–650.
- [189] H.L. Nakhasi, X.Q. Cao, T.A. Rouault, T.Y. Liu, Specific binding of host cell proteins to the 3'-terminal stem loop structure of rubella virus negative-strand RNA, *J. Virol.* 65 (1991) 5961–5967.
- [190] M.H. Chen, T.K. Frey, Mutagenic analysis of the 3' cis-acting elements of the rubella virus genome, *J. Virol.* 73 (1999) 3386–3403.
- [191] M.H. Chen, I. Frolov, J. Icenogle, T.K. Frey, Analysis of the 3' cis-acting elements of rubella virus by using replicons expressing a puromycin resistance gene, *J. Virol.* 78 (2004) 2553–2561.
- [192] W.P. Tzeng, T.K. Frey, Mapping the rubella virus subgenomic promoter, *J. Virol.* 76 (2002) 3189–3201.
- [193] P.S. Masters, The molecular biology of coronaviruses, *Adv. Virus Res.* 66 (2006) 193–292.
- [194] R.Y. Chang, M.A. Hofmann, P.B. Sethna, D.A. Brian, A cis-acting function for the coronavirus leader in defective interfering RNA replication, *J. Virol.* 68 (1994) 8223–8231.
- [195] K.M. Gustin, B.J. Guan, A. Dziduszko, D.A. Brian, Bovine coronavirus nonstructural protein 1 (p28) is an RNA binding protein that binds terminal genomic cis-replication elements, *J. Virol.* 83 (2009) 6087–6097.
- [196] D.A. Brian, R.S. Baric, Coronavirus genome structure and replication, *Curr. Top Microbiol. Immunol.* 287 (2005) 1–30.
- [197] R.Y. Chang, R. Krishnan, D.A. Brian, The UCUAAC promoter motif is not required for high-frequency leader recombination in bovine coronavirus defective interfering RNA, *J. Virol.* 70 (1996) 2720–2729.
- [198] S. Raman, P. Bouma, G.D. Williams, D.A. Brian, Stem loop III in the 5' untranslated region is a cis-acting element in bovine coronavirus defective interfering RNA replication, *J. Virol.* 77 (2003) 6720–6730.
- [199] S. Raman, D.A. Brian, Stem loop IV in the 5' untranslated region is a cis-acting element in bovine coronavirus defective interfering RNA replication, *J. Virol.* 79 (2005) 12434–12446.
- [200] L. Li, H. Kang, P. Liu, et al., Structural lability in stem loop 1 drives a 5' UTR–3' UTR interaction in coronavirus replication, *J. Mol. Biol.* 377 (2008) 790–803.
- [201] C.G. Brown, K.S. Nixon, S.D. Senanayake, D.A. Brian, An RNA stem loop within the bovine coronavirus nsp1 coding region is a cis-acting element in defective interfering RNA replication, *J. Virol.* 81 (2007) 7716–7724.
- [202] Y.N. Kim, Y.S. Jeong, S. Makino, Analysis of cis-acting sequences essential for coronavirus defective interfering RNA replication, *Virology* 197 (1993) 53–63.
- [203] W. Luytjes, H. Gerritsma, W.J. Spaan, Replication of synthetic defective interfering RNAs derived from coronavirus mouse hepatitis virus-A59, *Virology* 216 (1996) 174–183.
- [204] B. Hsue, P.S. Masters, A bulged stem loop structure in the 3' untranslated region of the genome of the coronavirus mouse hepatitis virus is essential for replication, *J. Virol.* 71 (1997) 7567–7578.
- [205] G.D. Williams, R.Y. Chang, D.A. Brian, A phylogenetically conserved hairpin-type 3' untranslated region pseudoknot functions in coronavirus RNA replication, *J. Virol.* 73 (1999) 8349–8355.

- [206] P. Huang, M.M. Lai, Heterogeneous nuclear ribonucleoprotein a1 binds to the 3'-untranslated region and mediates potential 5'-3'-end cross talks of mouse hepatitis virus RNA, *J. Virol.* 75 (2001) 5009–5017.
- [207] S.J. Goebel, T.B. Miller, C.J. Bennett, K.A. Bernard, P.S. Masters, A hypervariable region within the 3' cis-acting element of the murine coronavirus genome is nonessential for RNA synthesis but affects pathogenesis, *J. Virol.* 81 (2007) 1274–1287.
- [208] Y.J. Lin, C.L. Liao, M.M. Lai, Identification of the cis-acting signal for minus-strand RNA synthesis of a murine coronavirus: implications for the role of minus-strand RNA in RNA replication and transcription, *J. Virol.* 68 (1994) 8131–8140.
- [209] J.F. Spagnolo, B.G. Hogue, Host protein interactions with the 3' end of bovine coronavirus RNA and the requirement of the poly(A) tail for coronavirus defective genome replication, *J. Virol.* 74 (2000) 5053–5065.
- [210] R. Züst, T.B. Miller, S.J. Goebel, V. Thiel, P.S. Masters, Genetic interactions between an essential 3' cis-acting RNA pseudoknot, replicase gene products, and the extreme 3' end of the mouse coronavirus genome, *J. Virol.* 82 (2008) 1214–1228.
- [211] J.A. Fosmire, K. Hwang, S. Makino, Identification and characterization of a coronavirus packaging signal, *J. Virol.* 66 (1992) 3522–3530.
- [212] R. Cologna, B.G. Hogue, Identification of a bovine coronavirus packaging signal, *J. Virol.* 74 (2000) 580–583.
- [213] P.K. Hsieh, S.C. Chang, C.C. Huang, et al., Assembly of severe acute respiratory syndrome coronavirus RNA packaging signal into virus-like particles is nucleocapsid dependent, *J. Virol.* 79 (2005) 13848–13855.
- [214] S.M.K. Narayanan, Coronavirus genome packaging, in: V. Thiel (Ed.), *Coronaviruses Molecular and Cellular Biology*, Caister Academic Press, Norfolk, UK, 2007, pp. 131–142.
- [215] I. Brierley, M.E. Bournsnel, M.M. Binns, et al., An efficient ribosomal frame-shifting signal in the polymerase-encoding region of the coronavirus IBV, *EMBO J.* 6 (1987) 3779–3785.
- [216] F. Dos Ramos, M. Carrasco, T. Doyle, I. Brierley, Programmed-1 ribosomal frame-shifting in the SARS coronavirus, *Biochem. Soc. Trans.* 32 (2004) 1081–1083.
- [217] E.J. Snijder, J.J. Meulenbergh, The molecular biology of arteriviruses, *J. Gen. Virol.* 79 (Pt. 5) (1998) 961–979.
- [218] Y.J. Choi, S.I. Yun, S.Y. Kang, Y.M. Lee, Identification of 5' and 3' cis-acting elements of the porcine reproductive and respiratory syndrome virus: acquisition of novel 5' AU-rich sequences restored replication of a 5'-proximal 7-nucleotide deletion mutant, *J. Virol.* 80 (2006) 723–736.
- [219] M.A. Tijms, L.C. van Dinten, A.E. Gorbalenya, E.J. Snijder, A zinc finger-containing papain-like protease couples subgenomic mRNA synthesis to genome translation in a positive-stranded RNA virus, *Proc. Natl. Acad. Sci. U. S. A.* 98 (2001) 1889–1894.
- [220] E. Van Den Born, A.P. Gultyaev, E.J. Snijder, Secondary structure and function of the 5'-proximal region of the equine arteritis virus RNA genome, *Rna* 10 (2004) 424–437.
- [221] E. van den Born, C.C. Posthuma, A.P. Gultyaev, E.J. Snijder, Discontinuous subgenomic RNA synthesis in arteriviruses is guided by an RNA hairpin structure located in the genomic leader region, *J. Virol.* 79 (2005) 6312–6324.
- [222] A.O. Pasternak, E. van den Born, W.J. Spaan, E.J. Snijder, The stability of the duplex between sense and antisense transcription-regulating sequences is a crucial factor in arterivirus subgenomic mRNA synthesis, *J. Virol.* 77 (2003) 1175–1183.
- [223] N. Beerens, E.J. Snijder, RNA signals in the 3' terminus of the genome of Equine arteritis virus are required for viral RNA synthesis, *J. Gen. Virol.* 87 (2006) 1977–1983.
- [224] T.R. Maines, M. Young, N.N. Dinh, M.A. Brinton, Two cellular proteins that interact with a stem loop in the simian hemorrhagic fever virus 3'(+)NCR RNA, *Virus Res.* 109 (2005) 109–124.
- [225] M.H. Verheije, R.C. Olsthoorn, M.V. Kroese, P.J. Rottier, J.J. Meulenbergh, Kissing interaction between 3' noncoding and coding sequences is essential for porcine arterivirus RNA replication, *J. Virol.* 76 (2002) 1521–1526.
- [226] D.P. Giedroc, P.V. Cornish, Frameshifting RNA pseudoknots: structure and mechanism, *Virus Res.* 139 (2009) 193–208.
- [227] X. Jiang, M. Wang, K. Wang, M.K. Estes, Sequence and genomic organization of Norwalk virus, *Virology* 195 (1993) 51–61.
- [228] I. Goodfellow, Y. Chaudhry, I. Gioldasi, et al., Calcivirus translation initiation requires an interaction between VPg and eIF 4 E, *EMBO Rep.* 6 (2005) 968–972.
- [229] B.S. Seal, J.D. Neill, J.F. Ridpath, Predicted stem loop structures and variation in nucleotide sequence of 3' noncoding regions among animal calcivirus genomes, *Virus Genes* 8 (1994) 243–247.
- [230] S.V. Sosnovtsev, G. Belliot, K.O. Chang, O. Onwudiwe, K.Y. Green, Feline calcivirus VP2 is essential for the production of infectious virions, *J. Virol.* 79 (2005) 4012–4024.
- [231] A.L. Gutierrez-Escolano, M. Vazquez-Ochoa, J. Escobar-Herrera, J. Hernandez-Acosta, La, PTB, and PAB proteins bind to the 3' untranslated region of Norwalk virus genomic RNA, *Biochem. Biophys. Res. Commun.* 311 (2003) 759–766.
- [232] B. Al-Mutairy, J.E. Walter, A. Pothen, D.K. Mitchell, Genome Prediction of Putative Genome-Linked Viral Protein (VPg) of Astroviruses, *Virus Genes* 31 (2005) 21–30.
- [233] B. Jiang, S.S. Monroe, E.V. Koonin, S.E. Stine, R.I. Glass, RNA sequence of astrovirus: distinctive genomic organization and a putative retrovirus-like ribosomal frameshifting signal that directs the viral replicase synthesis, *Proc. Natl. Acad. Sci. U. S. A.* 90 (1993) 10539–10543.
- [234] C. Monceyron, B. Grinde, T.O. Jonassen, Molecular characterisation of the 3'-end of the astrovirus genome, *Arch. Virol.* 142 (1997) 699–706.
- [235] J.E. Walter, J. Briggs, M.L. Guerrero, et al., Molecular characterization of a novel recombinant strain of human astrovirus associated with gastroenteritis in children, *Arch. Virol.* 146 (2001) 2357–2367.
- [236] D.J. Miller, M.D. Schwartz, P. Ahlquist, Flock house virus RNA replicates on outer mitochondrial membranes in *Drosophila* cells, *J. Virol.* 75 (2001) 11664–11676.
- [237] P.D. Friesen, R.R. Rueckert, Synthesis of black beetle virus proteins in cultured *Drosophila* cells: differential expression of RNAs 1 and 2, *J. Virol.* 37 (1981) 876–886.
- [238] L.D. Eckerle, L.A. Ball, Replication of the RNA segments of a bipartite viral genome is coordinated by a transactivating subgenomic RNA, *Virology* 296 (2002) 165–176.
- [239] W. Zhong, R.R. Rueckert, Flock house virus: down-regulation of subgenomic RNA3 synthesis does not involve coat protein and is targeted to synthesis of its positive strand, *J. Virol.* 67 (1993) 2716–2722.
- [240] C.G. Albarino, B.D. Price, L.D. Eckerle, L.A. Ball, Characterization and template properties of RNA dimers generated during flock house virus RNA replication, *Virology* 289 (2001) 269–282.
- [241] P.A. Venter, A. Schneemann, Recent insights into the biology and biomedical applications of Flock House virus, *Cell Mol. Life Sci.* 65 (2008) 2675–2687.
- [242] K.N. Johnson, K.L. Johnson, R. Dasgupta, T. Gratsch, L.A. Ball, Comparisons among the larger genome segments of six nodaviruses and their encoded RNA replicases, *J. Gen. Virol.* 82 (2001) 1855–1866.
- [243] L.A. Ball, Replication of the genomic RNA of a positive-strand RNA animal virus from negative-sense transcripts, *Proc. Natl. Acad. Sci. U. S. A.* 91 (1994) 12443–12447.
- [244] B.D. Lindenbach, J.Y. Sgro, P. Ahlquist, Long-distance base pairing in flock house virus RNA1 regulates subgenomic RNA3 synthesis and RNA2 replication, *J. Virol.* 76 (2002) 3905–3919.
- [245] W. Zhong, R. Dasgupta, R. Rueckert, Evidence that the packaging signal for nodaviral RNA2 is a bulged stem loop, *Proc. Natl. Acad. Sci. U. S. A.* 89 (1992) 11146–11150.
- [246] P.M. Van Wynsberghe, P. Ahlquist, 5' cis elements direct nodavirus RNA1 recruitment to mitochondrial sites of replication complex formation, *J. Virol.* 83 (2009) 2976–2988.
- [247] C.G. Albarino, L.D. Eckerle, L.A. Ball, The cis-acting replication signal at the 3' end of Flock House virus RNA2 is RNA3-dependent, *Virology* 311 (2003) 181–191.
- [248] G. Ehrenfeld, N.L. Teterina, Initiation of Translation of Picornavirus RNAs: Structure and Function of the Internal Ribosome Entry Site, in: B.L. Semler, E. Wimmer (Eds.), *Molecular Biology of Picornaviruses*, ASM Press, Washington DC, 2002.
- [249] M. Schuler, S.R. Connell, A. Lescoute, et al., Structure of the ribosome-bound cricket paralysis virus IRES RNA, *Nat. Struct. Mol. Biol.* 13 (2006) 1092–1096.